



User Guide

AWS Clean Rooms



AWS Clean Rooms: User Guide

Copyright © 2026 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Die Handelsmarken und Handelsaufmachung von Amazon dürfen nicht in einer Weise in Verbindung mit nicht von Amazon stammenden Produkten oder Services verwendet werden, durch die Kunden irregeführt werden könnten oder Amazon in schlechtem Licht dargestellt oder diskreditiert werden könnte. Alle anderen Handelsmarken, die nicht Eigentum von Amazon sind, gehören den jeweiligen Besitzern, die möglicherweise zu Amazon gehören oder nicht, mit Amazon verbunden sind oder von Amazon gesponsert werden.

Table of Contents

Was ist AWS Clean Rooms?	1
Sind Sie ein AWS Clean Rooms Erstbenutzer?	2
Wie funktioniert AWS Clean Rooms	2
Zugehörige Services	2
AWS Dienstleistungen	2
Dienste von Drittanbietern	4
Zugreifen AWS Clean Rooms	5
Preisgestaltung für AWS Clean Rooms	5
Abrechnung für AWS Clean Rooms	5
Regeln für die Analyse	6
Typen von Analyseregeln	7
Regel für die Aggregationsanalyse	10
Regel für die Listenanalyse	31
Benutzerdefinierte Analyseregeln	40
Regel für die Analyse der ID-Zuordnungstabelle	47
AWS Clean Rooms Differenzierter Datenschutz	58
Differentieller Datenschutz	59
So funktioniert Differential Privacy AWS Clean Rooms	59
Differenzielle Datenschutzrichtlinie	60
SQL-Funktionen	62
Tipps und Beispiele für SQL-Abfragen	102
Einschränkungen	103
AWS-Reinräume ML	105
ML-Terminologie von AWS Clean Rooms	105
So funktioniert AWS Clean Rooms ML mit AWS Modellen	107
So funktioniert AWS Clean Rooms ML mit benutzerdefinierten Modellen	108
AWS Modelle in Reinräumen (ML)	110
Benutzerdefinierte Modelle in Clean Rooms ML	120
Kryptografisches Rechnen	128
Überlegungen	130
Unterstützte Datei- und Datentypen	133
Spaltennamen	138
Spaltentypen	139
Parameters	141

Optionale Flags	147
Abfragen mit C3R	150
Richtlinien	151
Einrichten AWS Clean Rooms	178
Melden Sie sich an für AWS	178
Richten Sie Servicerollen ein für AWS Clean Rooms	178
Erstellen Sie einen Administratorbenutzer	179
Erstellen Sie eine IAM-Rolle für ein Kollaborationsmitglied	180
Erstellen Sie eine Servicerolle zum Lesen von Daten aus Amazon S3	181
Erstellen Sie eine Servicerolle zum Lesen von Daten aus Amazon Athena	186
Erstellen Sie eine Servicerolle, um Daten aus Snowflake zu lesen	189
Erstellen Sie eine Servicerolle, um Code aus einem S3-Bucket zu lesen (PySpark Analysevorlagenrolle)	193
Erstellen Sie eine Servicerolle, um die Ergebnisse eines PySpark Jobs zu schreiben	197
Erstellen Sie eine Servicerolle, um Ergebnisse zu erhalten	200
Servicerollen für AWS Clean Rooms ML einrichten	204
Richten Sie Servicerollen für die Lookalike-Modellierung ein	204
Richten Sie Servicerollen für die benutzerdefinierte Modellierung ein	220
Kooperationen und Mitgliedschaften	236
Eine Zusammenarbeit erstellen	237
Eine Zusammenarbeit für Abfragen erstellen	237
Eine Kollaboration für Anfragen und Jobs erstellen	249
Eine Zusammenarbeit für ML-Modellierung erstellen	261
Eine Mitgliedschaft erstellen und einer Kollaboration beitreten	272
Kollaborationen bearbeiten	279
Bearbeiten Sie den Namen und die Beschreibung der Zusammenarbeit	279
Aktualisieren Sie die Analyse-Engine für die Zusammenarbeit	280
Schalten Sie den Protokollspeicher aus	280
Bearbeiten Sie die Einstellungen für Kollaborationsprotokolle	281
Tags für die Zusammenarbeit bearbeiten	282
Bearbeiten Sie Mitgliedskennungen	283
Ein neues Mitglied hinzufügen	284
Bearbeiten Sie die Fähigkeiten vorhandener Mitglieder	284
Einstellungen für die automatische Genehmigung der Zusammenarbeit bearbeiten	284
Bearbeiten Sie die zugehörigen Tabellen-Tags	284
Bearbeiten Sie die Tags der Analysevorlage	285

Bearbeiten Sie unterschiedliche Datenschutzrichtlinientags	285
Anfragen ändern	286
Fügen Sie einer Kollaboration ein neues Mitglied hinzu	286
Aktualisiere die Fähigkeiten vorhandener Mitglieder	287
Einstellungen für die automatische Genehmigung der Zusammenarbeit bearbeiten	289
Kollaborationen löschen	290
Kollaborationen anzeigen	290
Mitglieder einladen	291
Mitglieder überwachen	291
Mitglieder hinzufügen	292
Mitglieder entfernen	293
Verlassen einer Zusammenarbeit	294
Datentabellen	296
Datenformate	297
Unterstützte Datenformate für Jobs PySpark	297
Unterstützte Datenformate für SQL-Abfragen	297
Unterstützte Datentypen	298
Arten der Dateikomprimierung für AWS Clean Rooms	299
Serverseitige Verschlüsselung für AWS Clean Rooms	299
Apache Iceberg Tabellen	300
Unterstützte Datentypen für Iceberg-Tabellen	301
Vorbereiten von Datentabellen	302
Vorbereiten von Datentabellen in Amazon S3	302
Vorbereiten von Datentabellen in Amazon Athena	305
Datentabellen in Snowflake vorbereiten	308
Vorbereiten verschlüsselter Datentabellen	310
Schritt 1: Erfüllen der Voraussetzungen	311
Schritt 2: Laden Sie den C3R-Verschlüsselungsclient herunter	312
Schritt 3: (Optional) Verfügbare Befehle im C3R-Verschlüsselungsclient anzeigen	312
Schritt 4: Generieren Sie ein Verschlüsselungsschema für eine tabellarische Datei	313
Schritt 5: Erstellen Sie einen gemeinsamen geheimen Schlüssel	321
Schritt 6: Speichern Sie den gemeinsamen geheimen Schlüssel in einer Umgebungsvariablen	322
Schritt 7: Daten verschlüsseln	324
Schritt 8: Überprüfen Sie die Datenverschlüsselung	325
(Optional) Erstellen Sie ein Schema (fortgeschrittene Benutzer)	326

Datentabellen entschlüsseln	336
Konfigurierte Tabellen	338
Eine konfigurierte Tabelle erstellen	339
Amazon-S3-Datenquelle	340
Amazon Athena Athena-Datenquelle	343
Snowflake-Datenquelle	347
Hinzufügen einer Analyseregeln zu einer konfigurierten Tabelle	351
Hinzufügen einer Aggregationsanalyseregeln zu einer Tabelle (geführter Ablauf)	352
Hinzufügen einer Listenanalyseregeln zu einer Tabelle (geführter Ablauf)	357
Hinzufügen einer benutzerdefinierten Analyseregeln zu einer Tabelle (geführter Ablauf)	359
Analyseregeln zu einer Tabelle hinzufügen (JSON-Editor)	364
Nächste Schritte	365
Eine konfigurierte Tabelle einer Kooperation zuordnen	366
Budget für den Datenzugriff	366
Ordnen Sie eine konfigurierte Tabelle zu	367
Ein Datenzugriffsbudget konfigurieren	374
Ein Budget für den Datenzugriff anzeigen	374
Hinzufügen eines Datenzugriffsbudgets zu einer vorhandenen verknüpften Tabelle	375
Ein Datenzugriffsbudget bearbeiten	377
Löschen eines Datenzugriffsbudgets	378
Eine Regel für die Kollaborationsanalyse zu einer konfigurierten Tabelle hinzufügen	379
Konfiguration der differenzierten Datenschutzrichtlinie (optional)	381
Differenzielle Nutzungsprotokolle zum Datenschutz anzeigen	382
Eine differenzierte Datenschutzrichtlinie bearbeiten	382
Löschen einer differenzierten Datenschutzrichtlinie	384
Anzeige der berechneten differenziellen Datenschutzparameter	384
Tabellen und Analyseregeln anzeigen	386
Eine konfigurierte Tabelle bearbeiten	386
Bearbeiten konfigurierter Tabellen-Tags	387
Bearbeitung der konfigurierten Tabellenanalyseregeln	387
Löschen der konfigurierten Tabellenanalyseregeln	388
In der konfigurierten Tabelle sind keine Spalten zulässig	389
Bearbeiten konfigurierter Tabellenzuordnungen	392
Aufheben der Zuordnung konfigurierter Tabellen	393
AWS Entity Resolution in AWS Clean Rooms	394
ID-Namespaces	395

Einen neuen ID-Namespace erstellen und zuordnen	395
Zuordnen eines vorhandenen ID-Namespace	398
ID-Namespace-Zuordnungen bearbeiten	401
Zuordnung von ID-Namespace-Zuordnungen aufheben	402
Tabellen mit ID-Zuordnungen	403
Eine neue ID-Zuordnungstabelle erstellen und auffüllen	404
Auffüllen einer vorhandenen ID-Zuordnungstabelle	412
Bearbeiten einer ID-Zuordnungstabelle	413
Löschen einer ID-Zuordnungstabelle	413
Analysevorlagen	415
SQL-Analysevorlagen	415
Erstellen einer SQL-Analysevorlage	416
Überprüfen einer SQL-Analysevorlage	419
PySpark Analysevorlagen	421
Sicherheit	421
Einschränkungen	422
Best Practices	423
Ein Benutzerskript erstellen	424
Arbeiten mit Parametern in PySpark Analysevorlagen	428
Erstellen einer virtuellen Umgebung (optional)	432
Speichern eines Benutzerskripts und einer virtuellen Umgebung in S3	433
Eine PySpark Analysevorlage erstellen	435
Überprüfen einer PySpark Analysevorlage	439
PySpark Analysevorlagen zur Fehlerbehebung	443
Problembehandlung bei Ihrem Code	443
Der Job mit der Analysevorlage wird nicht gestartet	444
Der Job für die Analysevorlage wird gestartet, schlägt aber bei der Verarbeitung fehl	445
Die Einrichtung der virtuellen Umgebung schlägt fehl	447
Analyse	449
Ausführen von SQL-Abfragen	449
Voraussetzungen	450
Konfiguration der Spark-Eigenschaften für SQL-Abfragen	451
Konfigurierte Tabellen abfragen	452
Abfragen von ID-Zuordnungstabellen	471
Abfragen konfigurierter Tabellen mithilfe einer Analysevorlage	474
Abfragen mit dem Analysis Builder	490

Überblick über die Auswirkungen des unterschiedlichen Datenschutzes	496
Anzeige kürzlicher Abfragen	497
Anzeigen von Abfragedetails	498
PySpark Jobs werden ausgeführt	498
Einen Job mithilfe einer Analysevorlage ausführen	499
Aktuelle Jobs anzeigen	502
Anzeigen von Auftragsdetails	503
Analyseergebnisse	504
Abfrageergebnisse werden empfangen	505
Empfangen von Auftragsergebnissen	506
Standardwerte für Einstellungen für Abfrageergebnisse bearbeiten	507
Bearbeiten der Standardwerte für die Einstellungen für die Auftragsergebnisse	509
Verwenden der Abfrageausgabe in anderen AWS-Services	510
Benutzerdefiniertes Modellieren	511
Generierung synthetischer Datensätze mit verbessertem Datenschutz	512
Überlegungen zur Generierung synthetischer Daten	513
Gründung und Beitritt zur Zusammenarbeit	516
Aufbau einer Zusammenarbeit für maschinelles Lernen	516
Einer Kollaboration beitreten	521
Bereitstellung von Trainingsdaten	522
Konfiguration eines Modellalgorithmus	529
Den konfigurierten Modellalgorithmus zuordnen	532
Einen ML-Eingangskanal erstellen	537
Ein trainiertes Modell erstellen	557
Inkrementelles Training verwenden	561
Nutzung verteilter Schulungen	565
Exportieren von Modellartefakten	569
Inferenz auf einem trainierten Modell ausführen	571
ML-Modellierung für Anbieter von Trainingsdaten	574
Trainingsdaten importieren	575
Ein Lookalike-Modell erstellen	576
Konfiguration eines Lookalike-Modells	577
Zuordnen eines konfigurierten Lookalike-Modells	579
Aktualisierung eines konfigurierten Lookalike-Modells	579
ML-Modellierung für Seed-Datenanbieter	581
Ein Lookalike-Segment erstellen	581

Exportieren eines Lookalike-Segments	597
Fehlerbehebung	599
Auf eine oder mehrere Tabellen, auf die in der Abfrage verwiesen wird, kann über die zugehörige Dienstrolle nicht zugegriffen werden. Der table/role Besitzer muss der Servicerolle Zugriff auf die Tabelle gewähren.	599
Einer der zugrunde liegenden Datensätze hat ein nicht unterstütztes Dateiformat.	599
Die Abfrageergebnisse entsprechen nicht den Erwartungen, wenn Sie Cryptographic Computing for Clean Rooms verwenden.	600
AWS Clean Rooms Spark SQL: Fehlende Partitionsdaten	600
Sicherheit	601
Datenschutz	602
Verschlüsselung im Ruhezustand	603
Verschlüsselung während der Übertragung	604
Verschlüsselung der zugrunde liegenden Daten	604
Schlüsselrichtlinie	604
Verwenden von servicegebundenen Rollen	611
Berechtigungen für dienstverknüpfte Rollen AWS Clean Rooms	611
Erstellen einer dienstbezogenen Rolle für AWS Clean Rooms	612
Bearbeiten einer dienstbezogenen Rolle für AWS Clean Rooms	612
Löschen einer serviceverknüpften Rolle für AWS Clean Rooms	612
Unterstützte Regionen für serviceverknüpfte Rollen AWS Clean Rooms	613
Datenaufbewahrung	613
Best Practices	614
Bewährte Methoden mit AWS Clean Rooms	615
Bewährte Methoden für die Verwendung von Analyseregeln in AWS Clean Rooms	615
Identitäts- und Zugriffsverwaltung	617
Zielgruppe	618
Authentifizierung mit Identitäten	618
Verwalten des Zugriffs mit Richtlinien	620
Wie AWS Clean Rooms funktioniert mit IAM	623
Beispiele für identitätsbasierte Richtlinien	628
AWS verwaltete Richtlinien	632
Fehlerbehebung	640
Serviceübergreifende Confused-Deputy-Prävention	642
IAM-Verhalten für AWS Clean Rooms ML	644
IAM-Verhalten für benutzerdefinierte Clean Rooms ML-Modelle	648

Compliance-Validierung	650
Ausfallsicherheit	651
Sicherheit der Infrastruktur	651
Netzwerksicherheit	651
AWS PrivateLink	652
Überlegungen	652
Erstellen eines Schnittstellenendpunkts	653
Überwachen	654
Analyse Einloggen AWS Clean Rooms	655
Empfangen von Abfrage- und Jobprotokollen	656
Empfohlene Aktionen für Abfrage- und Jobprotokolle	657
Detailliertes Monitoring mit CloudWatch in AWS Clean Rooms	658
Arten von Metriken	658
Metrikdimensionen	660
Wer kann auf Metriken zugreifen	660
Preisgestaltung	660
CloudTrail protokolliert	661
AWS Clean Rooms Informationen in CloudTrail	661
AWS Clean Rooms Logdateieinträge verstehen	662
Beispiele für AWS Clean Rooms CloudTrail Ereignisse	662
Integration in AWS Clean Rooms EDAs	666
AWS Clean Rooms Ereignisse	667
AWS Clean Rooms Ereignisse weiterleiten	672
Detailreferenz zu Ereignissen	673
Kostenzuordnungs-Tags	682
Taggbare Ressourcen für die Kostenzuweisung	683
AWS CloudFormation Ressourcen	684
AWS Clean Rooms und CloudFormation Vorlagen	684
Erfahren Sie mehr über CloudFormation	684
Kontingente	686
AWS Clean Rooms Kontingente	686
AWS Clean Rooms Grenzwerte für Ressourcenparameter	701
ML-Kontingente für AWS Clean Rooms	702
Drosselungsquoten für die ML-API von Clean Rooms	737
Dokumentverlauf	745
Glossar	762

Regel für die Aggregationsanalyse	762
Analyseregeln	762
Analysevorlage	762
C3R-Verschlüsselungsclient	763
Spalte mit klarem Text	763
Zusammenarbeit	763
Ersteller der Kollaboration	763
Konfigurierte Tabelle	764
Benutzerdefinierte Analyseregeln	764
Entschlüsselung	764
Differenzielle Privatsphäre	765
Verschlüsselung	765
Spalte „Fingerabdruck“	765
Workflow-Methode für die ID-Zuordnung	765
Tabelle mit ID-Zuordnung	765
Regel zur Analyse von ID-Zuordnungstabellen	766
Arbeitsablauf bei der ID-Zuordnung	766
ID-Namespaces	766
Zuordnung des ID-Namespaces	766
Aufgabe	766
Analyseregeln auflisten	767
Lookalike-Modell	767
Ähnliches Segment	767
Mitglied	767
Mitglied, das Abfragen durchführen kann	767
Mitglied, das Abfragen und Jobs ausführen kann	768
Mitglied, das Ergebnisse erhalten kann	768
Das Mitglied zahlt die Kosten für die Berechnung von Abfragen	768
Das Mitglied zahlt für die Kosten der Abfrage- und Auftragsverarbeitung	769
Mitgliedschaften	769
Versiegelte Spalte	769
Ausgangsdaten	769
Spark-Analyse-Engine	770
Query	770
.....	dcclxxi

Was ist AWS Clean Rooms?

AWS Clean Rooms hilft Ihnen und Ihren Partnern, Ihre kollektiven Datensätze zu analysieren und gemeinsam daran zu arbeiten, um neue Erkenntnisse zu gewinnen, ohne sich gegenseitig die zugrunde liegenden Daten preiszugeben. AWS Clean Rooms ist ein sicherer Arbeitsbereich für die Zusammenarbeit, in dem Sie in wenigen Minuten Ihre eigenen Reindräume einrichten und Ihre kollektiven Datensätze mit nur wenigen Schritten analysieren können. Sie wählen die Partner aus, mit denen Sie zusammenarbeiten möchten, wählen deren Datensätze aus und konfigurieren Kontrollen zur Verbesserung der Privatsphäre für diese Partner.

Mit können Sie mit Tausenden von Unternehmen zusammenarbeiten AWS Clean Rooms, die dies bereits nutzen. AWS Für die Zusammenarbeit ist es nicht erforderlich, Daten aus einem anderen Cloud-Dienstanbieter zu verschieben AWS oder in diesen zu laden. Wenn Sie Abfragen oder Jobs ausführen, AWS Clean Rooms liest es Daten vom ursprünglichen Speicherort dieser Daten und wendet integrierte Analyseregeln an, damit Sie die Kontrolle über diese Daten behalten.

AWS Clean Rooms bietet integrierte Datenzugriffskontrollen und Kontrollfunktionen zur Unterstützung von Audits, die Sie konfigurieren können. Zu diesen Kontrollen gehören:

- [Analyseregeln](#) zur Einschränkung von SQL-Abfragen und zur Bereitstellung von Ausgabebeschränkungen.
- [Kryptografisches Rechnen Clean Rooms zur](#) Verschlüsselung von Daten, auch wenn Abfragen verarbeitet werden, um strenge Datenverarbeitungsrichtlinien einzuhalten.
- [Analyseprotokolle](#) zur Überprüfung von Abfragen und Aufträgen und zur Unterstützung von Audits. AWS Clean Rooms
- [Differenzierter Datenschutz](#) zum Schutz vor Versuchen zur Benutzeridentifikation. AWS Clean Rooms Differential Privacy ist eine vollständig verwaltete Funktion, die die Privatsphäre Ihrer Benutzer mit mathematisch gestützten Techniken und intuitiven Steuerelementen schützt, die Sie in wenigen Schritten anwenden können.
- [AWS Clean Rooms ML](#) ermöglicht es zwei Parteien, ähnliche Benutzer in ihren Daten zu identifizieren, ohne ihre Daten miteinander teilen zu müssen. Die erste Partei erstellt und konfiguriert anhand ihrer Trainingsdaten ein Lookalike-Modell. Anschließend werden Ausgangsdaten in die Zusammenarbeit eingebracht, um ein Lookalike-Segment zu erstellen, das den Trainingsdaten ähnelt.

Das folgende Video erklärt mehr über AWS Clean Rooms.

[AWS Clean Rooms](#)

Sind Sie ein AWS Clean Rooms Erstbenutzer?

Wenn Sie zum ersten Mal Benutzer von sind AWS Clean Rooms, empfehlen wir Ihnen, zunächst die folgenden Abschnitte zu lesen:

- [Wie funktioniert AWS Clean Rooms](#)
- [Zugreifen AWS Clean Rooms](#)
- [Einrichten AWS Clean Rooms](#)
- [AWS Clean Rooms Glossar](#)

Wie funktioniert AWS Clean Rooms

In AWS Clean Rooms erstellen Sie eine Kollaboration und fügen die hinzu AWS-Konten , die Sie einladen möchten, oder erstellen eine Mitgliedschaft, um einer Kollaboration beizutreten, zu der Sie eingeladen wurden. Anschließend verknüpfen Sie die für Ihren Anwendungsfall benötigten Datenressourcen: konfigurierte Tabellen für Ereignisdaten, konfigurierte Modelle für die ML-Modellierung oder ID-Namespaces für die Entitätsauflösung. Sie haben die Möglichkeit, Analysevorlagen zu erstellen oder zu genehmigen, um sich im Voraus auf die genauen Abfragen und Jobs zu einigen, die Sie in einer Zusammenarbeit zulassen möchten. Schließlich analysieren Sie die gemeinsamen Daten, indem Sie SQL-Abfragen oder PySpark Jobs für die konfigurierten Tabellen ausführen, Entitätsauflösung in ID-Zuordnungstabellen durchführen oder ML-Modellierung verwenden, um ähnliche Zielgruppensegmente zu generieren.

Das folgende Diagramm zeigt, wie das AWS Clean Rooms funktioniert.

Zugehörige Services

AWS Dienstleistungen

Folgendes bezieht AWS-Services sich auf AWS Clean Rooms:

- Amazon Athena

Mitglieder der Kollaboration können Daten, die sie einbringen, AWS Clean Rooms als AWS Glue Data Catalog Ansichten in Amazon Athena speichern. Weitere Informationen finden Sie unter den folgenden Themen:

Weitere Informationen finden Sie unter den folgenden Themen:

[Vorbereiten von Datentabellen für Abfragen in AWS Clean Rooms](#)

[Eine konfigurierte Tabelle erstellen — Amazon Athena Athena-Datenquelle](#)

[Was ist Amazon Athena?](#) im Amazon Athena Athena-Benutzerhandbuch

- CloudFormation

Erstellen Sie die folgenden Ressourcen in CloudFormation: Kollaborationen, konfigurierte Tabellen, konfigurierte Tabellenzuordnungen und Mitgliedschaften

Weitere Informationen finden Sie unter [AWS Clean Rooms Ressourcen erstellen mit AWS CloudFormation](#).

- AWS CloudTrail

Verwenden Sie es AWS Clean Rooms zusammen mit CloudTrail Protokollen, um Ihre Aktivitätsanalyse zu verbessern. AWS-Service

Weitere Informationen finden Sie unter [Protokollieren von AWS Clean Rooms API-Aufrufen mit AWS CloudTrail](#).

- AWS Entity Resolution

Verwenden Sie AWS Clean Rooms with AWS Entity Resolution , um eine Entitätsauflösung durchzuführen.

Weitere Informationen finden Sie unter [AWS Entity Resolution in AWS Clean Rooms](#).

- AWS Glue

Mitglieder der Kollaboration können aus ihren Daten in Amazon S3 AWS Glue Tabellen zur Verwendung in erstellen AWS Clean Rooms.

Weitere Informationen finden Sie unter den folgenden Themen:

[Vorbereiten von Datentabellen für Abfragen in AWS Clean Rooms](#)

[Was ist AWS Glue?](#) im Entwicklerhandbuch für AWS Glue

- Amazon Simple Storage Service (Amazon-S3)

Mitglieder der Kollaboration können Daten, die sie einbringen, AWS Clean Rooms in Amazon S3 speichern.

Weitere Informationen finden Sie unter den folgenden Themen:

[Vorbereiten von Datentabellen für Abfragen in AWS Clean Rooms](#)

[Eine konfigurierte Tabelle erstellen — Amazon S3 S3-Datenquelle](#)

[Was ist Amazon S3?](#) im Entwicklerhandbuch für Amazon Simple Storage Service

- AWS Secrets Manager

Mitglieder der Kollaboration können Geheimnisse erstellen, um auf in Snowflake gespeicherte Daten zuzugreifen und diese zu lesen.

Weitere Informationen finden Sie unter den folgenden Themen:

[Erstellen Sie eine Servicerolle, um Daten aus Snowflake zu lesen](#)

[Vorbereiten von Datentabellen für Abfragen in AWS Clean Rooms](#)

[Was ist AWS Secrets Manager?](#) im AWS Secrets Manager Benutzerhandbuch

Dienste von Drittanbietern

Der folgende Drittanbieter-Service bezieht sich auf AWS Clean Rooms:

- Snowflake

Mitglieder der Kollaboration können Daten, die sie einbringen, AWS Clean Rooms in einem Snowflake-Warehouse speichern.

Weitere Informationen finden Sie unter den folgenden Themen:

[Vorbereiten von Datentabellen für Abfragen in AWS Clean Rooms](#)

[Eine konfigurierte Tabelle erstellen — Snowflake-Datenquelle](#)

Zugreifen AWS Clean Rooms

Sie können AWS Clean Rooms über die folgenden Optionen darauf zugreifen:

- Direkt über die AWS Clean Rooms Konsole unter <https://console.aws.amazon.com/cleanrooms/>.
- Programmgesteuert über die AWS Clean Rooms API. Weitere Informationen finden Sie in der [AWS Clean Rooms -API-Referenz](#).

Preisgestaltung für AWS Clean Rooms

Preisinformationen finden Sie unter [AWS Clean Rooms – Preise](#).

Note

Für Mitglieder von Collaboration, die Daten verknüpft haben, die in Snowflake gespeichert sind, wird Ihnen jedes Mal, wenn eine Abfrage ausgeführt wird, die Daten verwendet, die an diesen Speicherorten gespeichert sind, von ihrem jeweiligen Data Warehouse- oder Cloud-Anbieter sowohl für ausgehende Daten als auch für die Berechnung in Rechnung gestellt.

Abrechnung für AWS Clean Rooms

AWS Clean Rooms gibt dem Kollaborationsersteller die Möglichkeit, festzulegen, welches Mitglied die Kosten für die Abfrage- oder Auftragsverarbeitung in der Kollaboration bezahlt.

In den meisten Fällen sind das [Mitglied, das Abfragen durchführen kann](#), und das [Mitglied, das die Kosten für die Query Compute bezahlt](#), identisch. Wenn jedoch das Mitglied, das Abfragen durchführen kann, und das Mitglied, das die Kosten für die Query Compute bezahlt, unterschiedlich sind, wird, wenn das Mitglied, das Abfragen durchführen kann, Abfragen für seine eigene Mitgliedschaftsressource ausführt, der Mitgliedschaftsressource des Mitglieds, das die Kosten für die Abfrage-Compute bezahlt, in Rechnung gestellt.

Das Mitglied, das die Kosten für die Abfragerechnung bezahlt, sieht in seinem Ereignisverlauf kein CloudTrail Ereignis für ausgeführte Abfragen, da der Zahler weder derjenige ist, der die Abfragen ausführt, noch der Besitzer der Ressource ist, für die die Abfragen ausgeführt werden. Der Zahler sieht jedoch die Gebühren, die auf seiner Mitgliedschaftsressource für alle Abfragen anfallen, die von dem Mitglied ausgeführt werden, das Abfragen in der Kollaboration ausführen kann.

Weitere Informationen zum Erstellen einer Kollaboration und zur Konfiguration des Mitglieds, das die Kosten für die Berechnung von Abfragen bezahlt, finden Sie unter [Eine Zusammenarbeit erstellen](#).

Analyseregeln in AWS Clean Rooms

Im Rahmen der Aktivierung einer Tabelle AWS Clean Rooms für die Kollaborationsanalyse muss das Kollaborationsmitglied eine Analyseregel konfigurieren.

Bei einer Analyseregel handelt es sich um eine Kontrolle zur Verbesserung des Datenschutzes, die jeder Datenbesitzer in einer konfigurierten Tabelle einrichtet. Eine Analyseregel bestimmt, wie die konfigurierte Tabelle analysiert werden kann.

Bei der Analyseregel handelt es sich um eine Kontrolle auf Kontoebene für die konfigurierte Tabelle (eine Ressource auf Kontoebene). Sie wird in jeder Zusammenarbeit durchgesetzt, der die konfigurierte Tabelle zugeordnet ist. Wenn keine Analyseregel konfiguriert ist, kann die konfigurierte Tabelle Kollaborationen zugeordnet, aber nicht abgefragt werden. Abfragen können nur auf konfigurierte Tabellen mit demselben Analyseregeltyp verweisen.

Um eine Analyseregel zu konfigurieren, wählen Sie zuerst einen Analysetyp aus und geben dann die Analyseregel an. Bei beiden Schritten sollten Sie berücksichtigen, welchen Anwendungsfall Sie aktivieren möchten und wie Sie Ihre zugrunde liegenden Daten schützen möchten.

AWS Clean Rooms erzwingt die restriktiveren Kontrollen für alle konfigurierten Tabellen, auf die in einer Abfrage verwiesen wird.

Die folgenden Beispiele veranschaulichen die restriktiven Kontrollen.

Example Restriktive Kontrolle: Ausgabebeschränkung

- Mitarbeiter A hat eine Ausgabebeschränkung für die Kennungsspalte 100.
- Mitarbeiter B hat eine Ausgabebeschränkung für die Identifikatorspalte von 150.

Eine Aggregationsabfrage, die auf beide konfigurierten Tabellen verweist, benötigt mindestens 150 unterschiedliche Bezeichnerwerte innerhalb einer Ausgabezeile, damit sie in der Abfrageausgabe angezeigt werden kann. Die Abfrageausgabe gibt nicht an, dass Ergebnisse aufgrund der Ausgabebeschränkung entfernt wurden.

Example Restriktive Kontrolle: Analysevorlage nicht genehmigt

- Mitarbeiter A hat eine Analysevorlage mit einer Abfrage zugelassen, die in ihrer benutzerdefinierten Analyseregeln auf konfigurierte Tabellen von Collaborator A und Collaborator B verweist.
- Mitarbeiter B hat die Analysevorlage nicht zugelassen.

Da Mitarbeiter B die Analysevorlage nicht zugelassen hat, kann das Mitglied, das Abfragen durchführen kann, diese Analysevorlage nicht ausführen.

Typen von Analyseregeln

Es gibt drei Arten von Analyseregeln: [Aggregationsregeln](#), [Listenregeln](#) und [benutzerdefinierte Regeln](#). In den folgenden Tabellen werden die Analyseregeltypen verglichen. Jeder Typ hat einen eigenen Abschnitt, in dem die Angabe der Analyseregeln beschrieben wird.

Note

Es gibt einen Analyseregeltyp, der als Analyseregel für ID-Zuordnungstabellen bezeichnet wird. Diese Analyseregel wird jedoch von verwaltet AWS Clean Rooms und kann nicht geändert werden. Weitere Informationen finden Sie unter [Regel zur Analyse von ID-Zuordnungstabellen](#).

In den folgenden Abschnitten werden unterstützte Anwendungsfälle und Kontrollen für jeden Analyseregeltyp beschrieben.

Unterstützte Anwendungsfälle

Die folgenden Tabellen enthalten eine Vergleichszusammenfassung der unterstützten Anwendungsfälle für jeden Analyseregeltyp.

Anwendungsfall	Aggregation	Liste	Custom (Benutzer definiert)
Unterstützte Analysen	Abfragen, die Statistik mithilfe der Funktionen COUNT,	Abfragen, die Listen mit Überschneidungen zwischen	Jede benutzerdefinierte Analyse, sofern die Analysevo

Anwendungsfall	Aggregation	Liste	Custom (Benutzer definiert)
	SUM und AVG anhand optionaler Dimensionen aggregieren	mehreren Tabellen auf Zeilenebene ausgeben	Ursprache oder der Analyseersteller überprüft und zugelassen wurden
Allgemeine Anwendungsfälle	Segmentanalyse, Messung, Zuordnung	Bereicherung, Segmentbildung	Zuordnung auf Anfrage, inkrementelle Analysen, Zielgruppenfindung
SQL-Konstrukte	• JOIN-Anweisungen: INNER JOIN • Aggregatfunktionen: COUNT/COUNT DISTINCT, SUM/SUM DISTINCT und AVG • Skalarfunktionen: Eingeschränkte Teilmenge	• JOIN-Anweisungen: INNER JOIN • Skalarfunktionen: Keine	Die meisten SQL-Funktionen und SQL-Konstrukte sind mit dem SELECT-Befehl verfügbar
Unterabfragen und allgemeine Tabellenausdrücke () CTEs	Nein	Nein	Ja
Vorlagen für Analysen	Nein	Nein	Ja

Unterstützte Steuerelemente

Die folgenden Tabellen zeigen eine vergleichende Zusammenfassung darüber, wie die einzelnen Analyseregeltypen Ihre zugrunde liegenden Daten schützen.

Kontrolle	Aggregation	Liste	Custom (Benutzer definiert)
Kontrollmechanismus	Steuern Sie, wie Daten in der Tabelle in einer Abfrage verwendet werden können (Lassen Sie beispielsweise COUNT und SUM der Spalte hashed_email zu.)	Steuern Sie, wie Daten in der Tabelle in einer Abfrage verwendet werden können (Erlauben Sie beispielsweise die Verwendung der Spalte hashed_email nur für den Beitritt.)	Steuern Sie, welche Abfragen in der Tabelle ausgeführt werden dürfen (Lassen Sie beispielsweise nur Abfragen zu, die in den Analysevorlagen „Benutzerdefinierte Abfrage 1“ definiert sind.)
Integrierte Techniken zur Verbesserung der Privatsphäre	• Blindes Spiel • Aggregation erforderlich • Minimaler Aggregationsschwellenwert $\geq$ • 2 Vordefinierte Abfragestruktur	• Blindes Spiel • Überlappung erforderlich • Vordefinierte Abfragestruktur • Zusätzliche Analysen sind zulässig	• Differenzierter Datenschutz • Unzulässige Ausgabespalten
Überprüfen Sie die Abfrage, bevor sie ausgeführt werden kann	Nein	Nein	Ja, mithilfe von Analysevorlagen

Weitere Informationen zu den Analyseregeln, die in verfügbar sind AWS Clean Rooms, finden Sie in den folgenden Themen.

- [Regel für die Aggregationsanalyse](#)
- [Regel für die Listenanalyse](#)
- [Benutzerdefinierte Analyseregeln in AWS Clean Rooms](#)

Regel für die Aggregationsanalyse

In AWS Clean Rooms generiert eine Aggregationsanalyseregeln aggregierte Statistiken mithilfe der Funktionen COUNT, SUM und and/or AVG anhand optionaler Dimensionen. Wenn die Aggregationsanalyseregeln zu einer konfigurierten Tabelle hinzugefügt wird, ermöglicht sie dem Mitglied, das Abfragen durchführen kann, Abfragen in der konfigurierten Tabelle auszuführen.

Die Aggregationsanalyseregeln unterstützt Anwendungsfälle wie Kampagnenplanung, Medienreichweite, Frequenzmessung und Zuordnung.

Die unterstützte Abfragestruktur und Syntax sind in definiert. [Struktur und Syntax von Aggregationsabfragen](#)

Zu den Parametern der Analyseregeln, die in definiert sind [Regel für die Aggregationsanalyse — Steuerelemente abfragen](#), gehören Abfragesteuerelemente und Steuerelemente für Abfrageergebnisse. Zu den Abfragesteuerelementen gehört die Möglichkeit, zu verlangen, dass eine konfigurierte Tabelle mit mindestens einer konfigurierten Tabelle verknüpft wird, deren Eigentümer das Mitglied ist, das Abfragen entweder direkt oder transitiv durchführen kann. Mit dieser Anforderung können Sie sicherstellen, dass die Abfrage an der Schnittstelle (INNERJOIN) zwischen Ihrer und ihrer Tabelle ausgeführt wird.

Struktur und Syntax von Aggregationsabfragen

Abfragen in Tabellen, für die eine Aggregationsanalyseregeln gilt, müssen der folgenden Syntax entsprechen.

```
--select_aggregate_function_expression
SELECT
aggregation_function(column_name) [[AS] column_alias ] [, ...]

--select_grouping_column_expression
[, {column_name|scalar_function(arguments)} [[AS] column_alias ]][, ...]

--table_expression
FROM table_name [[AS] table_alias ]
    [[INNER] JOIN table_name [[AS] table_alias] ON join_condition] [...]

--where_expression
[WHERE where_condition]

--group_by_expression
[GROUP BY {column_name|scalar_function(arguments)}, ...]
```

```
--having_expression
[HAVING having_condition]

--order_by_expression
[ORDER BY {column_name|scalar_function(arguments)} [{ASC|DESC}]] [,...]]
```

In der folgenden Tabelle werden alle in der vorherigen Syntax aufgeführten Ausdrücke erklärt.

Expression	Definition	Beispiele
<i>select_aggregate_function_expression</i>	Eine durch Kommas getrennte Liste mit den folgenden Ausdrücken: - select_aggregation_function_expression - select_aggregate_expression Note Es muss mindestens einen select_aggregation_function_expression in der geben. select_aggregate_expression	SELECT SUM(PRICE), user_segment
<i>select_aggregation_function_expression</i>	Eine oder mehrere unterstützte Aggregationsfunktionen, die auf eine oder mehrere Spalten angewende	AVG(PRICE) COUNT(DISTINCT user_id)

Expression	Definition	Beispiele
	t werden. Nur Spalten sind als Argumente von Aggregationsfunktionen zulässig.  Note Es muss mindestens einen <code>select_aggregation_function_expression</code> in der <code>select_aggregate_expression</code> geben.	

Expression	Definition	Beispiele
<i><code>select_grouping_column_expression</code></i>	Ein Ausdruck, der einen beliebigen Ausdruck enthalten kann, wobei Folgendes verwendet wird: • Tabellenspaltennamen • Unterstützte Skalarfunktionen • Zeichenkettenlitterale • Numerische Litterale	<code>TRUNC(timestampColumn)</code> <code>UPPER(campaignName)</code>

 **Note**

`select_aggregate_expression` kann Spalten mit oder ohne den AS Parameter als Alias kennzeichnen. Weitere Informationen finden Sie in der [AWS Clean Rooms SQL-Referenz](#).

Expression	Definition	Beispiele
<i>table_expression</i>	Eine Tabelle oder eine Verknüpfung von Tabellen, mit der bedingte Join-Ausdrücke miteinander verbunden werden. <code>join_condition</code> gibt einen booleschen Wert zurück. Die <code>table_expression</code> Stützen: • Ein bestimmter JOIN Typ (INNERJOIN) • Die Gleichheitsvergleichsbedingung innerhalb von <code>join_condition</code> (=) • Logische Operatoren (AND,OR).	<pre>FROM consumer_table INNER JOIN provider_ table ON consumer_table.id entifier1 = provider_ table.identifier1 AND consumer_table .identifier2 = provider_table.id entifier2</pre>

Expression	Definition	Beispiele
<i>where_expression</i>	Ein bedingter Ausdruck, der einen booleschen Wert zurückgibt. Er kann aus Folgendem bestehen: • Tabellenspaltennamen • Unterstützte Skalarfunktionen • Mathematische Operatoren • Zeichenkettenlitterale • Numerische Litterale Unterstützte Vergleichsbedingungen sind (=, >, <, <=, >=, <>, !=, NOT, IN, NOT IN, LIKE, IS NULL, IS NOT NULL). Unterstützte logische Operatoren sind (AND, OR). Die where_expression ist optional.	<pre>WHERE where_condition</pre> <pre>WHERE price > 100</pre> <pre>WHERE TRUNC(timestampColumn) = '1/1/2022'</pre> <pre>WHERE timestampColumn = timestampColumn2 - 14</pre>
<i>group_by_expression</i>	Eine durch Kommas getrennte Liste von Ausdrücken, die den Anforderungen für entsprechenden. select_grouping_column_expression	<pre>GROUP BY TRUNC(timestampColumn), UPPER(campaignName), segment</pre>

Expression	Definition	Beispiele
<i>having_expression</i>	Ein bedingter Ausdruck, der einen booleschen Wert zurückgibt. Sie verfügen über eine unterstützte Aggregationsfunktion, die auf eine einzelne Spalte angewendet wird (z. B. <code>SUM(price)</code>), und sie werden mit einem numerischen Literal verglichen. Unterstützte Bedingungen sind <code>()=</code>, <code>></code>, <code><</code>, <code><=</code>, <code>>=</code>, <code><></code>, <code>!=</code>. Unterstützte logische Operatoren sind <code>AND</code>, <code>OR</code>. Die <code>having_expression</code> ist optional.	<code>HAVING SUM(SALES) > 500</code>

Expression	Definition	Beispiele
<i>order_by_expression</i>	Eine durch Kommas getrennte Liste von Ausdrücken, die mit den gleichen Anforderungen kompatibel ist, die zuvor definiert <code>select_aggregate_expression</code> wurden. Die <code>order_by_expression</code> ist optional.  Note <code>order_by_expression</code> Genehmigungen <code>ASC</code> und Parameter <code>. DESC</code> Weitere Informationen finden Sie unter <code>ASC DESC-Parameter</code> in der AWS Clean Rooms SQL-Referenz.	<pre>ORDER BY SUM(SALES), UPPER(campaignName)</pre>

Beachten Sie bei der Struktur und Syntax von Aggregationsabfragen Folgendes:

- Andere SQL-Befehle als `SELECT` werden nicht unterstützt.
- Unterabfragen und allgemeine Tabellenausdrücke (z. B. `WITH`) werden nicht unterstützt.
- Operatoren, die mehrere Abfragen kombinieren (z. B. `UNION`), werden nicht unterstützt.
- `TOPLIMIT`, und `OFFSET` Parameter werden nicht unterstützt.

Regel für die Aggregationsanalyse — Steuerelemente abfragen

Mit Steuerelementen für Aggregationsabfragen können Sie steuern, wie die Spalten in Ihrer Tabelle zur Abfrage der Tabelle verwendet werden. Sie können beispielsweise steuern, welche Spalte für die Verknüpfung verwendet wird, welche Spalte gezählt werden kann oder welche Spalte in WHERE Anweisungen verwendet werden kann.

In den folgenden Abschnitten werden die einzelnen Steuerelemente erläutert.

Themen

- [Steuerelemente für die Aggregation](#)
- [Steuerelemente verbinden](#)
- [Steuerelemente für Dimensionen](#)
- [Skalarfunktionen](#)

Steuerelemente für die Aggregation

Mithilfe von Aggregationssteuerelementen können Sie definieren, welche Aggregationsfunktionen zulässig sind und auf welche Spalten sie angewendet werden müssen. Aggregationsfunktionen können in den Ausdrücken `SELECTHAVING`, und verwendet werden. `ORDER BY`

Steuerung	Definition	Usage
<code>aggregateColumns</code>	Spalten konfigurierter Tabellenspalten, die Sie für die Verwendung innerhalb von Aggregationsfunktionen zulassen.	<code>aggregateColumns</code> kann innerhalb einer Aggregationsfunktion in den Ausdrücken <code>SELECTHAVING</code>, und <code>ORDER BY</code> verwendet werden. Einige <code>aggregateColumns</code> können auch als <code>joinColumn</code> (später definiert) kategorisiert werden. Given <code>aggregateColumn</code> kann nicht auch als

Steuerung	Definition	Usage
		<code>dimensionColumn</code> (später definiert) kategorisiert werden.
<code>function</code>	Die Funktionen COUNT, SUM und AVG, die Sie zusätzlich zu verwenden zulassen <code>aggregate Columns</code> .	<code>function</code> kann auf eine angewendet werden <code>aggregate Columns</code> , die damit verknüpft ist.

Steuerelemente verbinden

Eine JOIN Klausel wird verwendet, um Zeilen aus zwei oder mehr Tabellen auf der Grundlage einer zugehörigen Spalte miteinander zu kombinieren.

Mithilfe von Join-Steuerelementen können Sie steuern, wie Ihre Tabelle mit anderen Tabellen in der verknüpft werden kann `table_expression`. AWS Clean Rooms unterstützt nur INNERJOIN. INNERJOINAnweisungen können nur Spalten verwenden, die `joinColumn` in Ihrer Analyseregeln explizit als `a` kategorisiert wurden, und zwar vorbehaltlich der von Ihnen definierten Kontrollen.

Sie INNER JOIN müssen mit einer Tabelle `joinColumn` aus Ihrer konfigurierten Tabelle und mit einer Tabelle `joinColumn` aus einer anderen konfigurierten Tabelle in der Kollaboration arbeiten. Sie entscheiden, als welche Spalten aus Ihrer Tabelle verwendet werden können `joinColumn`.

Jede Übereinstimmungsbedingung innerhalb der ON Klausel muss die Gleichheitsvergleichsbedingung (=) zwischen zwei Spalten verwenden.

Es können mehrere Übereinstimmungsbedingungen innerhalb einer ON Klausel sein:

- Kombiniert mit dem AND logischen Operator
- Mit dem OR logischen Operator getrennt

Note

Alle JOIN Spielbedingungen müssen einer Zeile auf jeder Seite von entsprechen JOIN. Alle Bedingungen, die durch einen OR oder einen AND logischen Operator miteinander verbunden sind, müssen dieser Anforderung ebenfalls entsprechen.

Das Folgende ist ein Beispiel für eine Abfrage mit einem AND logischen Operator.

```
SELECT some_col, other_col
FROM table1
JOIN table2
ON table1.id = table2.id AND table1.name = table2.name
```

Das Folgende ist ein Beispiel für eine Abfrage mit einem OR logischen Operator.

```
SELECT some_col, other_col
FROM table1
JOIN table2
ON table1.id = table2.id OR table1.name = table2.name
```

Steuerung	Definition	Usage
<code>joinColumns</code>	Die Spalten (falls vorhanden), die Sie dem Mitglied, das Abfragen durchführen kann, in der INNER JOIN Anweisung verwenden dürfen.	Ein bestimmtes <code>joinColumn</code> Objekt kann auch als <code>aggregateColumn</code> (siehe Steuerelemente für die Aggregation) kategorisiert werden. Dieselbe Spalte kann nicht gleichzeitig als <code>joinColumn</code> und verwendet werden <code>dimensionColumns</code> (siehe später). Sofern sie nicht auch als <code>a</code> kategorisiert wurde <code>aggregateColumn</code>, <code>joinColumn</code> kann sie in keinem anderen Teil der Abfrage als dem verwendet werden INNERJOIN.
<code>joinRequired</code>	Steuern Sie von dem Mitglied, das Abfragen durchführen	Wenn Sie diesen Parameter aktivieren, INNER JOIN ist ein

Steuerung	Definition	Usage
	kann, ob Sie eine INNER JOIN mit einer konfigurierten Tabelle benötigen.	erforderlich. Wenn Sie diesen Parameter nicht aktivieren, INNER JOIN ist optional. Angenommen, Sie aktivieren diesen Parameter, muss das Mitglied, das Abfragen durchführen kann, eine Tabelle, deren Eigentümer es ist, in die aufnehmen INNERJOIN. Sie müssen JOIN Ihre Tabelle mit ihrer Tabelle entweder direkt oder transitiv verbinden (d. h. sie verknüpfen ihre Tabelle mit einer anderen Tabelle, die wiederum mit Ihrer Tabelle verknüpft ist).

Im Folgenden finden Sie ein Beispiel für Transitivität.

```
ON
my_table.identifer = third_party_table.identifier
....
ON
third_party_table.identifier = member_who_can_query_table.id
```

Note

Das Mitglied, das Abfragen durchführen kann, kann den `joinRequired` Parameter auch verwenden. In diesem Fall muss die Abfrage ihre Tabelle mit mindestens einer anderen Tabelle verknüpfen.

Steuerelemente für Dimensionen

Dimensionssteuerelemente steuern die Spalte, anhand derer die Aggregationsspalten gefiltert, gruppiert oder aggregiert werden können.

Steuerung	Definition	Usage
<code>dimensionColumns</code>	Die Spalten (falls vorhanden), die Sie dem Mitglied, das Abfragen durchführen kann SELECT, in, WHERE GROUPBY, und verwenden dürfen. ORDER BY	A <code>dimensionColumn</code> kann in SELECT (<code>select_grouping_column_expression</code>), WHERE GROUPBY, und verwendet werden ORDERBY. Dieselbe Spalte kann nicht gleichzeitig a <code>dimensionColumn</code> <code>joinColumn</code>, a und and/or an sein <code>aggregateColumn</code>.

Skalarfunktionen

Skalarfunktionen steuern, welche Skalarfunktionen für Dimensionsspalten verwendet werden können.

Steuerung	Definition	Usage
<code>scalarFunctions</code>	Die Skalarfunktionen, die <code>dimensionColumns</code> in der Abfrage verwendet werden können.	Gibt die Skalarfunktionen (falls vorhanden) an, auf die Sie (z. B. CAST) zulassen. <code>dimensionColumns</code> Skalarfunktionen können nicht zusätzlich zu anderen Funktionen oder innerhalb anderer Funktionen verwendet werden. Argumente von Skalarfunktionen können

Steuerung	Definition	Usage
		Spalten, Zeichenkettenlitterale oder numerische Litterale sein.

Die folgenden Skalarfunktionen werden unterstützt:

- Mathematische Funktionen — ABS, CEILING, FLOOR, LOG, LN, ROUND, SQRT
- Funktionen zur Formatierung von Datentypen — CAST, CONVERT, TO_CHAR, TO_DATE, TO_NUMBER, TO_TIMESTAMP
- Zeichenkettenfunktionen — LOWER, UPPER, TRIM, RTRIM, SUBSTRING
 - Für RTRIM sind benutzerdefinierte Zeichensätze zum Kürzen nicht zulässig.
- Bedingte Ausdrücke — COALESCE
- Datumsfunktionen — EXTRACT, GETDATE, CURRENT_DATE, DATEADD
- Andere Funktionen — TRUNC

Weitere Informationen finden Sie in der [AWS Clean Rooms SQL-Referenz](#).

Regel für die Aggregationsanalyse — Steuerelemente für Abfrageergebnisse

Mit den Steuerelementen für Aggregationsabfrageergebnisse können Sie steuern, welche Ergebnisse zurückgegeben werden, indem Sie eine oder mehrere Bedingungen angeben, die jede Ausgabezeile erfüllen muss, damit sie zurückgegeben wird. AWS Clean Rooms unterstützt Aggregationseinschränkungen in der Form von. `COUNT (DISTINCT column) >= X` Dieses Formular erfordert, dass jede Zeile mindestens X verschiedene Werte einer Auswahl aus Ihrer konfigurierten Tabelle aggregiert (z. B. eine Mindestanzahl von unterschiedlichen `user_id` Werten). Dieser Mindestschwellenwert wird automatisch durchgesetzt, auch wenn die übermittelte Abfrage selbst die angegebene Spalte nicht verwendet. Sie werden gemeinsam für jede konfigurierte Tabelle in der Abfrage anhand der konfigurierten Tabellen aller Mitglieder der Kollaboration durchgesetzt.

Jede konfigurierte Tabelle muss mindestens eine Aggregationsbeschränkung in ihrer Analyseregeln enthalten. Besitzer konfiguierter Tabellen können mehrere `columnName` und zugeordnete Tabellen hinzufügen, `minimum` und sie werden gemeinsam durchgesetzt.

Einschränkungen bei der Aggregation

Aggregationseinschränkungen steuern, welche Zeilen in den Abfrageergebnissen zurückgegeben werden. Um zurückgegeben zu werden, muss eine Zeile die angegebene Mindestanzahl an unterschiedlichen Werten in jeder Spalte erfüllen, die in der Aggregationsbeschränkung angegeben ist. Diese Anforderung gilt auch dann, wenn die Spalte in der Abfrage oder in anderen Teilen der Analyseregeln nicht ausdrücklich erwähnt wird.

Steuerung	Definition	Usage
<code>columnName</code>	Die <code>aggregateColumn</code> , die in der Bedingung verwendet wird, die jede Ausgabezeile erfüllen muss.	Es kann sich um eine beliebige Spalte in der konfigurierten Tabelle handeln.
<code>minimum</code>	Die Mindestanzahl an eindeutigen Werten für die Verknüpfung <code>aggregateColumn</code> , die die Ausgabezeile haben muss (z. B. <code>COUNT DISTINCT</code>), damit sie in den Abfrageergebnissen zurückgegeben wird.	Das <code>minimum</code> muss mindestens den Wert 2 haben.

Struktur der Regeln für die Aggregationsanalyse

Das folgende Beispiel zeigt eine vordefinierte Struktur für eine Aggregationsanalyseregeln.

MyTable Bezieht sich im folgenden Beispiel auf Ihre Datentabelle. Sie können jede Information *user input placeholder* durch Ihre eigenen Informationen ersetzen.

```
{
  "aggregateColumns": [
    {
      "columnNames": [MyTable column names], "function": [Allowed Agg Functions]
    },
  ],
  "joinRequired": ["QUERY_RUNNER"],
```

```
"joinColumns": [MyTable column names],
"dimensionColumns": [MyTable column names],
"scalarFunctions": [Allowed Scalar functions],
"outputConstraints": [
  {
    "columnName": [MyTable column names], "minimum": [Numeric value]
  },
]
}
```

Regel für die Aggregationsanalyse — Beispiel

Das folgende Beispiel zeigt, wie zwei Unternehmen AWS Clean Rooms mithilfe der Aggregationsanalyse zusammenarbeiten können.

Unternehmen A verfügt über Kunden- und Vertriebsdaten. Unternehmen A ist daran interessiert, die Aktivitäten zur Produktrückgabe zu verstehen. Unternehmen B ist einer der Einzelhändler von Unternehmen A und verfügt über Rückgabedaten. Unternehmen B verfügt auch über Segmentattribute für Kunden, die für Unternehmen A nützlich sind (z. B. ähnliche Produkte gekauft, den Kundendienst des Einzelhändlers in Anspruch genommen). Unternehmen B möchte keine Kundenrückgabedaten und Attributinformationen auf Zeilenebene bereitstellen. Unternehmen B möchte nur eine Reihe von Abfragen für Unternehmen A aktivieren, um aggregierte Statistiken über sich überschneidende Kunden bei einem Mindestaggregationsschwellenwert zu erhalten.

Unternehmen A und Unternehmen B beschließen, zusammenzuarbeiten, damit Unternehmen A die Produktrückgabeaktivitäten nachvollziehen und bessere Produkte für Unternehmen B und andere Vertriebskanäle liefern kann.

Um die Zusammenarbeit aufzubauen und eine Aggregationsanalyse durchzuführen, gehen die Unternehmen wie folgt vor:

1. Unternehmen A erstellt eine Kollaboration und erstellt eine Mitgliedschaft. Die Kollaboration hat Firma B als weiteres Mitglied der Kollaboration. Unternehmen A aktiviert die Abfrageprotokollierung in der Kollaboration und aktiviert die Abfrageprotokollierung in ihrem Konto.
2. Unternehmen B erstellt eine Mitgliedschaft in der Kollaboration. Es aktiviert die Abfrageprotokollierung in seinem Konto.
3. Firma A erstellt eine für den Vertrieb konfigurierte Tabelle.
4. Unternehmen A fügt der konfigurierten Tabelle für Verkäufe die folgende Aggregationsanalyseregel hinzu.

```
{
  "aggregateColumns": [
    {
      "columnNames": [
        "identifier"
      ],
      "function": "COUNT_DISTINCT"
    },
    {
      "columnNames": [
        "purchases"
      ],
      "function": "AVG"
    },
    {
      "columnNames": [
        "purchases"
      ],
      "function": "SUM"
    }
  ],
  "joinColumns": [
    "hashedemail"
  ],
  "dimensionColumns": [
    "demoseg",
    "purchasedate",
    "productline"
  ],
  "scalarFunctions": [
    "CAST",
    "COALESCE",
    "TRUNC"
  ],
  "outputConstraints": [
    {
      "columnName": "hashedemail",
      "minimum": 2,
      "type": "COUNT_DISTINCT"
    }
  ]
}
```

aggregateColumns— Unternehmen A möchte die Anzahl der einzelnen Kunden in der Überschneidung zwischen Verkaufsdaten und Retourendaten zählen. Unternehmen A möchte auch die Anzahl der purchases hergestellten Produkte summieren, um sie mit der Anzahl von zu vergleichenreturns.

joinColumns— Unternehmen A **identifizier** möchte damit Kunden aus Verkaufsdaten mit Kunden aus Retourendaten abgleichen. Dies hilft Unternehmen A dabei, Retouren den richtigen Käufen zuzuordnen. Es hilft Unternehmen A auch dabei, Kunden zu segmentieren, die sich überschneiden.

dimensionColumns— Unternehmen A filtert **dimensionColumns** nach einem bestimmten Produkt, vergleicht Käufe und Rücksendungen über einen bestimmten Zeitraum, stellt sicher, dass das Rückgabedatum nach dem Produktdatum liegt, und hilft dabei, sich überschneidende Kunden zu segmentieren.

scalarFunctions— Unternehmen A wählt die CAST Skalarfunktion aus, um Datentypformate bei Bedarf auf der Grundlage der konfigurierten Tabelle, die Unternehmen A der Zusammenarbeit zuordnet, zu aktualisieren. Außerdem werden Skalarfunktionen hinzugefügt, um bei Bedarf die Formatierung von Spalten zu erleichtern.

outputConstraints— Unternehmen A legt Mindestbeschränkungen für die Produktion fest. Die Ergebnisse müssen nicht eingeschränkt werden, da der Analyst Daten auf Zeilenebene aus seiner Verkaufstabelle einsehen kann

 Note

Unternehmen A nimmt in der Analyseregeln nichts **joinRequired** auf. Es bietet ihren Analysten die Flexibilität, nur die Verkaufstabelle abzufragen.

5. Firma B erstellt eine konfigurierte Tabelle für Renditen.
6. Unternehmen B fügt der konfigurierten Tabelle für Rücksendungen die folgende Aggregationsanalyseregeln hinzu.

```
{
  "aggregateColumns": [
    {
      "columnNames": [
        "identifizier"
```

```
    ],
    "function": "COUNT_DISTINCT"
  },
  {
    "columnNames": [
      "returns"
    ],
    "function": "AVG"
  },
  {
    "columnNames": [
      "returns"
    ],
    "function": "SUM"
  }
],
"joinColumns": [
  "hashedemail"
],
"joinRequired": [
  "QUERY_RUNNER"
],
"dimensionColumns": [
  "state",
  "popularpurchases",
  "customerserviceuser",
  "productline",
  "returndate"
],
"scalarFunctions": [
  "CAST",
  "LOWER",
  "UPPER",
  "TRUNC"
],
"outputConstraints": [
  {
    "columnName": "hashedemail",
    "minimum": 100,
    "type": "COUNT_DISTINCT"
  },
  {
    "columnName": "producttype",
    "minimum": 2,
```

```
    "type": "COUNT_DISTINCT"  
  }  
]  
}
```

aggregateColumns— Unternehmen B ermöglicht es Unternehmen A, eine Summe zu erstellen `returns`, um sie mit der Anzahl der Käufe zu vergleichen. Sie haben mindestens eine Aggregatspalte, da sie eine Aggregatabfrage ermöglichen.

joinColumns— Unternehmen B ermöglicht es Unternehmen A, sich zusammenzutun `unidentifiziert`, um Kunden anhand von Rückgabedaten mit Kunden aus Verkaufsdaten abzugleichen. `identifiziert` Daten sind besonders sensibel, und wenn sie als A verwendet werden, wird `joinColumn` sichergestellt, dass die Daten niemals in einer Abfrage ausgegeben werden.

joinRequired— Unternehmen B verlangt, dass sich Abfragen zu den Rückgabedaten mit den Verkaufsdaten überschneiden. Sie möchten es Unternehmen A nicht ermöglichen, alle Personen in ihrem Datensatz abzufragen. Sie haben sich auch in ihrer Kooperationsvereinbarung auf diese Einschränkung geeinigt.

dimensionColumns— Unternehmen B ermöglicht es Unternehmen A, nach `statepopularpurchases`, und eindeutigen Attributen zu filtern und zu gruppieren, `customerserviceuser` die bei der Analyse für Unternehmen A hilfreich sein könnten. Unternehmen B ermöglicht es Unternehmen A, die Ausgabe `returndate` danach `returndate` zu filtern `purchasedate`. Mit dieser Filterung ist die Ausgabe genauer, was die Bewertung der Auswirkungen der Produktänderung ermöglicht.

scalarFunctions— Unternehmen B ermöglicht Folgendes:

- `TRUNC` für Daten
- `LOWER` und `UPPER`, falls `producttype` die in ihren Daten in einem anderen Format eingegeben wurden
- `CAST` wenn Unternehmen A die Datentypen im Vertrieb so konvertieren muss, dass sie den Datentypen in Rücksendungen entsprechen

Unternehmen A aktiviert keine anderen Skalarfunktionen, da sie nicht der Meinung sind, dass sie für Abfragen erforderlich sind.

outputConstraints— Unternehmen B legt Mindestbeschränkungen für die Produktion fest hashedemail, um die Möglichkeit zu verringern, Kunden neu zu identifizieren. Außerdem werden Mindestbeschränkungen für die Produktion eingeführt producttype, um die Möglichkeit zu verringern, bestimmte Produkte, die zurückgegeben wurden, erneut zu identifizieren. Bestimmte Produkttypen könnten aufgrund der Größe der Produktion dominanter sein (z. B. state). Ihre Produktionsbeschränkungen werden immer durchgesetzt, unabhängig von den Produktionsbeschränkungen, die Unternehmen A ihren Daten hinzugefügt hat.

7. Firma A erstellt eine Verkaufstabelle, die der Zusammenarbeit zugeordnet ist.
8. Firma B erstellt eine Verknüpfung der Tabelle mit Rücksendungen zur Zusammenarbeit.
9. Unternehmen A führt Abfragen wie das folgende Beispiel durch, um besser zu verstehen, wie viele Retouren in Unternehmen B im Vergleich zu den Gesamteinkäufen pro Standort im Jahr 2022 getätigt wurden.

```
SELECT
    companyB.state,
    SUM(companyB.returns),
    COUNT(DISTINCT companyA.hashemail)
FROM
    sales companyA
    INNER JOIN returns companyB ON companyA.identifier = companyB.identifier
WHERE
    companyA.purchasedate BETWEEN '2022-01-01' AND '2022-12-31' AND
    TRUNC(companyB.returndate) > companyA.purchasedate
GROUP BY
    companyB.state;
```

10. Unternehmen A und Unternehmen B überprüfen die Abfrageprotokolle. Unternehmen B überprüft, ob die Anfrage mit dem übereinstimmt, was in der Kooperationsvereinbarung vereinbart wurde.

Behebung von Problemen mit Regeln für die Aggregationsanalyse

Verwenden Sie die Informationen hier, um häufig auftretende Probleme bei der Arbeit mit Aggregationsanalyseregeln zu diagnostizieren und zu beheben.

Problembereiche

- [Meine Abfrage lieferte keine Ergebnisse](#)

Meine Abfrage lieferte keine Ergebnisse

Dies kann passieren, wenn es keine passenden Ergebnisse gibt oder wenn die übereinstimmenden Ergebnisse einen oder mehrere Mindestaggregationsschwellenwerte nicht erreichen.

Weitere Informationen zu Mindestschwellenwerten für die Aggregation finden Sie unter. [Regel für die Aggregationsanalyse — Beispiel](#)

Regel für die Listenanalyse

In AWS Clean Rooms gibt eine Listenanalyseregel Listen mit Überschneidungen zwischen der konfigurierten Tabelle, der sie hinzugefügt wurde, und den konfigurierten Tabellen des Mitglieds, das Abfragen durchführen kann, auf Zeilenebene aus. Das Mitglied, das Abfragen durchführen kann, führt Abfragen aus, die eine Listenanalyseregel enthalten.

Der Regeltyp „Listenanalyse“ unterstützt Anwendungsfälle wie Anreicherung und Zielgruppenbildung.

Weitere Informationen zur vordefinierten Abfragestruktur und Syntax für diese Analyseregel finden Sie unter [Vordefinierte Struktur der Listenanalyseregel](#).

Die Parameter der in [Regel für die Listenanalyse — Steuerelemente abfragen](#) definierten Listenanalyseregel verfügen über Abfragesteuerelemente. Zu den Abfragesteuerelementen gehört die Möglichkeit, die Spalten auszuwählen, die in der Ausgabe aufgeführt werden können. Für die Abfrage ist mindestens eine Verknüpfung mit einer konfigurierten Tabelle des Mitglieds erforderlich, das Abfragen entweder direkt oder transitiv durchführen kann.

Es gibt keine Steuerelemente für Abfrageergebnisse wie bei der [Aggregationsanalyseregel](#).

Listenabfragen können nur mathematische Operatoren verwenden. Sie können keine anderen Funktionen (wie Aggregation oder Skalar) verwenden.

Topics

- [Struktur und Syntax von Abfragen auflisten](#)
- [Regel für die Listenanalyse — Steuerelemente abfragen](#)
- [Vordefinierte Struktur der Listenanalyseregel](#)
- [Regel zur Listenanalyse — Beispiel](#)

Struktur und Syntax von Abfragen auflisten

Abfragen in Tabellen, für die eine Listenanalyseregel gilt, müssen der folgenden Syntax entsprechen.

```

--select_list_expression
SELECT DISTINCT column_name [[AS] column_alias ] [, ...]

--table_expression
FROM table_name [[AS] table_alias ]
    [[INNER] JOIN table_name [[AS] table_alias] ON join_condition] [...]

--where_expression
[WHERE where_condition]

--limit_expression
[LIMIT number]

```

In der folgenden Tabelle werden alle in der vorherigen Syntax aufgeführten Ausdrücke erklärt.

Expression	Definition	Beispiele
<i>select_list_expression</i>	Eine durch Kommas getrennte Liste, die mindestens einen Tabellenspaltennamen enthält. Ein DISTINCT Parameter ist erforderlich.  Note select_list_expression Sie können Spalten mit oder ohne den AS Parameter als Alias verwenden. Weitere Informationen finden Sie in der AWS Clean Rooms SQL-Referenz.	SELECT DISTINCT segment

Expression	Definition	Beispiele
<i>table_expression</i>	Eine Tabelle oder eine Verknüpfung von Tabellen, mit der eine <code>join_condition</code> Verbindung hergestellt werden soll. <code>join_condition</code> gibt einen booleschen Wert zurück. Die <code>table_expression</code> Stützen: • Ein bestimmter JOIN-Typ (INNERJOIN) • Die Bedingungen für den Gleichheitsvergleich in a <code>join_condition</code> (=) • Logische Operatoren (AND,OR).	<pre>FROM consumer_table INNER JOIN provider_ table ON consumer_table.ide ntifier1 = provider_ table.identifrier1 AND consumer_table .identifrier2 = provider_table.ide ntifier2</pre>

Expression	Definition	Beispiele
<i>where_expression</i>	Ein bedingter Ausdruck, der einen booleschen Wert zurückgibt. Er kann aus folgenden Elementen bestehen: • Tabellenspaltennamen • Mathematische Operatoren • Zeichenkettenlitterale • Numerische Literale Unterstützte Vergleichsbedingungen sind (=, >, <, <=, >=, <>, !=, NOT, IN, NOT IN, LIKE, IS NULL, IS NOT NULL). Unterstützte logische Operatoren sind (AND, OR). Die <code>where_expression</code> ist optional.	<pre>WHERE state + '_' + city = 'NY_NYC'</pre> <pre>WHERE timestampColumn = timestampColumn2 - 14</pre>
<i>limit_expression</i>	Dieser Ausdruck muss eine positive Ganzzahl enthalten. Die <code>limit_expression</code> ist optional.	<pre>LIMIT 100</pre>

Beachten Sie bei der Struktur und Syntax von Listenabfragen Folgendes:

- Andere SQL-Befehle als SELECT werden nicht unterstützt.
- Unterabfragen und allgemeine Tabellenausdrücke (z. B. WITH) werden nicht unterstützt
- Die BY Klauseln GROUP BY HAVING, und ORDER werden nicht unterstützt

- Der OFFSET-Parameter wird nicht unterstützt

Regel für die Listenanalyse — Steuerelemente abfragen

Mit Steuerelementen für Listenabfragen können Sie steuern, wie die Spalten in Ihrer Tabelle zum Abfragen der Tabelle verwendet werden. Sie können beispielsweise steuern, welche Spalte für die Verknüpfung verwendet wird oder welche Spalte in der SELECT-Anweisung und WHERE -Klausel verwendet werden kann.

In den folgenden Abschnitten werden die einzelnen Steuerelemente erläutert.

Themen

- [Steuerelemente verbinden](#)
- [Steuerelemente auflisten](#)

Steuerelemente verbinden

Mit Join-Steuerelementen können Sie steuern, wie Ihre Tabelle mit anderen Tabellen in `table_expression` verknüpft werden kann. AWS Clean Rooms unterstützt INNER nur JOIN. In der Regel für die Listenanalyse ist mindestens ein INNER JOIN erforderlich, und das Mitglied, das Abfragen durchführen kann, muss eine Tabelle, deren Eigentümer es ist, in den INNER JOIN aufnehmen. Das bedeutet, dass sie Ihre Tabelle entweder direkt oder transitiv mit ihrer Tabelle verknüpfen müssen.

Im Folgenden finden Sie ein Beispiel für Transitivität.

```
ON
my_table.identifer = third_party_table.identifier
....
ON
third_party_table.identifer = member_who_can_query_table.id
```

INNERJOIN-Anweisungen können nur Spalten verwenden, die `joinColumn` in Ihrer Analyseregeln explizit als `a` kategorisiert wurden.

Der INNER JOIN-Befehl muss für eine Tabelle `joinColumn` aus Ihrer konfigurierten Tabelle und für eine Tabelle `joinColumn` aus einer anderen konfigurierten Tabelle in der Kollaboration ausgeführt werden. Sie entscheiden, als welche Spalten aus Ihrer Tabelle verwendet werden können `joinColumn`.

Jede Übereinstimmungsbedingung innerhalb der ON Klausel muss die Gleichheitsvergleichsbedingung (=) zwischen zwei Spalten verwenden.

Es können mehrere Übereinstimmungsbedingungen innerhalb einer ON Klausel sein:

- Kombiniert mit dem AND logischen Operator
- Mit dem OR logischen Operator getrennt

 Note

Alle JOIN Spielbedingungen müssen einer Zeile auf jeder Seite von entsprechenJOIN. Alle Bedingungen, die durch einen OR oder einen AND logischen Operator miteinander verbunden sind, müssen dieser Anforderung ebenfalls entsprechen.

Das Folgende ist ein Beispiel für eine Abfrage mit einem AND logischen Operator.

```
SELECT some_col, other_col
FROM table1
  JOIN table2
    ON table1.id = table2.id AND table1.name = table2.name
```

Im Folgenden finden Sie ein Beispiel für eine Abfrage mit einem OR logischen Operator.

```
SELECT some_col, other_col
FROM table1
  JOIN table2
    ON table1.id = table2.id OR table1.name = table2.name
```

Steuerung	Definition	Usage
<code>joinColumns</code>	Die Spalten, die Sie dem Mitglied, das Abfragen durchführen kann, in der INNER JOIN-Anweisung verwenden dürfen.	Dieselbe Spalte kann nicht sowohl als <code>joinColumn</code> n auch kategorisiert werden <code>listColumn</code> (siehe Steuerelemente auflisten).

Steuerung	Definition	Usage
		<code>joinColumn</code> kann in keinem anderen Teil der Abfrage als INNER JOIN verwendet werden.

Steuerelemente auflisten

Listensteuerelemente steuern die Spalten, die in der Abfrageausgabe aufgeführt (d. h. in der SELECT-Anweisung verwendet) oder zum Filtern von Ergebnissen (d. h. in der WHERE Anweisung verwendet) verwendet werden können.

Steuerung	Definition	Usage
<code>listColumns</code>	Die Spalten, die Sie dem Mitglied, das Abfragen abfragen kann, in SELECT verwenden dürfen und WHERE	A <code>listColumn</code> kann in SELECT und verwendet werden WHERE. Dieselbe Spalte kann nicht gleichzeitig als ein <code>listColumn</code> und verwendet werden <code>joinColumn</code> .

Vordefinierte Struktur der Listenanalyseregel

Das folgende Beispiel enthält eine vordefinierte Struktur, die zeigt, wie Sie eine Listenanalyseregel abschließen.

MyTable Bezieht sich im folgenden Beispiel auf Ihre Datentabelle. Sie können jede Information *user input placeholder* durch Ihre eigenen Informationen ersetzen.

```
{
  "joinColumns": [MyTable column name(s)],
  "listColumns": [MyTable column name(s)],
}
```


Regel zur Listenanalyse — Beispiel

Das folgende Beispiel zeigt, wie zwei Unternehmen AWS Clean Rooms mithilfe der Listenanalyse zusammenarbeiten können.

Unternehmen A verfügt über Daten zum Kundenbeziehungsmanagement (CRM). Unternehmen A möchte zusätzliche Segmentdaten über seine Kunden erhalten, um mehr über ihre Kunden zu erfahren und möglicherweise Attribute als Input für andere Analysen zu verwenden. Unternehmen B verfügt über Segmentdaten, die aus eindeutigen Segmentattributen bestehen, die das Unternehmen auf der Grundlage seiner eigenen Daten erstellt hat. Unternehmen B möchte Unternehmen A die eindeutigen Segmentattribute nur für Kunden zur Verfügung stellen, deren Daten sich mit den Daten von Unternehmen A überschneiden.

Die Unternehmen beschließen, zusammenzuarbeiten, damit Unternehmen A die sich überschneidenden Daten anreichern kann. Unternehmen A ist das Mitglied, das Abfragen durchführen kann, und Unternehmen B ist der Mitwirkende.

Um eine Zusammenarbeit zu erstellen und gemeinsam eine Listenanalyse durchzuführen, gehen die Unternehmen wie folgt vor:

1. Unternehmen A erstellt eine Kollaboration und erstellt eine Mitgliedschaft. Die Kollaboration hat Firma B als weiteres Mitglied der Kollaboration. Unternehmen A aktiviert die Abfrageprotokollierung in der Kollaboration und sie aktiviert die Abfrageprotokollierung in ihrem Konto.
2. Unternehmen B erstellt eine Mitgliedschaft in der Kollaboration. Es aktiviert die Abfrageprotokollierung in seinem Konto.
3. Firma A erstellt eine für CRM konfigurierte Tabelle
4. Unternehmen A fügt die Analyseregeln der vom Kunden konfigurierten Tabelle hinzu, wie im folgenden Beispiel gezeigt.

```
{
  "joinColumns": [
    "identifizier1",
    "identifizier2"
  ],
  "listColumns": [
    "internalid",
    "segment1",
    "segment2",
    "customercategory"
  ]
}
```

```
]
}
```

joinColumns— Unternehmen A möchte `hashedemail` and/or `thirdpartyid` (von einem Identitätsanbieter bezogen) verwenden, um Kunden aus CRM-Daten mit Kunden aus Segmentdaten abzugleichen. Auf diese Weise kann sichergestellt werden, dass Unternehmen A angereicherte Daten den richtigen Kunden zuordnet. Sie verfügen über zwei `JoinColumns`, um die Trefferquote der Analyse potenziell zu verbessern.

listColumns— Unternehmen A verwendet `listColumns`, um zusätzlich angereicherte Spalten zu erhalten und `internalid` sie in ihren eigenen Systemen zu verwenden. Sie fügen `segment1` hinzu `segment2` und `customercategory` beschränken die Anreicherung möglicherweise auf bestimmte Segmente, indem sie sie in Filtern verwenden.

5. Firma B erstellt eine segmentkonfigurierte Tabelle.

6. Firma B fügt die Analyseregeln zur segmentkonfigurierten Tabelle hinzu.

```
{
  "joinColumns": [
    "identifizier2"
  ],
  "listColumns": [
    "segment3",
    "segment4"
  ]
}
```

joinColumns— Unternehmen B ermöglicht es Unternehmen A, gemeinsam Kunden `identifizier2` anhand von Segmentdaten mit CRM-Daten abzugleichen. Unternehmen A und Unternehmen B arbeiteten mit dem Identitätsanbieter zusammen, um herauszufinden, `identifizier2` welcher Anbieter für diese Zusammenarbeit geeignet wäre. Andere wurden nicht hinzugefügt, `joinColumns` da sie der Meinung waren, dass sie die höchste und genaueste Trefferquote `identifizier2` bieten und andere Identifikatoren für die Abfragen nicht erforderlich sind.

listColumns— Unternehmen B ermöglicht es Unternehmen A, seine Daten mit `segment3` `segment4` Attributen anzureichern. Dabei handelt es sich um einzigartige Attribute, die das Unternehmen im Rahmen der Datenanreicherung erstellt, gesammelt und (mit Kunde A) abgestimmt hat. Sie möchten, dass Unternehmen A diese Segmente für die Überschneidung

auf Zeilenebene erhält, da es sich um eine Zusammenarbeit im Bereich der Datenanreicherung handelt.

7. Unternehmen A erstellt eine CRM-Tabellenzuordnung zur Kollaboration.
8. Unternehmen B erstellt eine Segmenttabellenzuordnung zur Kollaboration.
9. Unternehmen A führt Abfragen wie die folgende aus, um sich überschneidende Kundendaten anzureichern.

```
SELECT companyA.internalid, companyB.segment3, companyB.segment4
INNER JOIN returns companyB
  ON companyA.identifizier2 = companyB.identifizier2
WHERE companyA.customercategory > 'xxx'
```

10. Unternehmen A und Unternehmen B überprüfen die Abfrageprotokolle. Unternehmen B überprüft, ob die Anfrage mit dem übereinstimmt, was in der Kooperationsvereinbarung vereinbart wurde.

Benutzerdefinierte Analyseregeln in AWS Clean Rooms

In AWS Clean Rooms ist eine benutzerdefinierte Analyseregeln ein neuer Typ von Analyseregeln, mit dem benutzerdefinierte Abfragen für die konfigurierte Tabelle ausgeführt werden können. Benutzerdefinierte SQL-Abfragen sind weiterhin darauf beschränkt, nur den SELECT Befehl zu verwenden, können aber mehr SQL-Konstrukte als [Aggregations](#) - und [Listenabfragen](#) verwenden (z. B. Fensterfunktionen, OUTER JOIN oder Unterabfragen; eine vollständige Liste finden Sie in der [AWS Clean Rooms SQL-Referenz](#)). CTEs [Benutzerdefinierte SQL-Abfragen müssen nicht wie Aggregations - und Listenabfragen einer Abfragestruktur folgen](#).

Die benutzerdefinierte Analyseregeln unterstützt komplexere Anwendungsfälle als solche, die von der Aggregations- und Listenanalyseregeln unterstützt werden können, z. B. benutzerdefinierte Attributionsanalysen, Benchmarking, Inkrementalitätsanalysen und Zielgruppenerkennung. Dies gilt zusätzlich zu einer Vielzahl von Anwendungsfällen, die von der Aggregations- und Listenanalyseregeln unterstützt werden.

Die benutzerdefinierte Analyseregeln unterstützt auch den differenziellen Datenschutz. Differential Privacy ist ein mathematisch strenges Rahmenwerk für den Datenschutz. Weitere Informationen finden Sie unter [AWS Clean Rooms Differenzierter Datenschutz](#). Wenn Sie eine Analysevorlage erstellen, überprüft AWS Clean Rooms Differential Privacy die Vorlage, um festzustellen, ob sie mit der allgemeinen Abfragestruktur für Differential Privacy kompatibel ist. AWS Clean Rooms Durch diese Überprüfung wird sichergestellt, dass Sie keine Analysevorlage erstellen, die in einer durch Differential Privacy geschützten Tabelle nicht zulässig ist.

Um die benutzerdefinierte Analyseregeln zu konfigurieren, können Datenbesitzer festlegen, dass bestimmte benutzerdefinierte Abfragen, die in [Analysevorlagen gespeichert sind, für](#) ihre konfigurierten Tabellen ausgeführt werden. Datenbesitzer überprüfen Analysevorlagen, bevor sie sie der zulässigen Analysesteuerung in der benutzerdefinierten Analyseregeln hinzufügen. Analysevorlagen sind nur in der Kollaboration verfügbar und sichtbar, in der sie erstellt wurden (auch wenn die Tabelle mit anderen Kollaborationen verknüpft ist). Sie können nur von dem Mitglied ausgeführt werden, das in dieser Kollaboration Abfragen durchführen kann.

Alternativ können sich Mitglieder dafür entscheiden, anderen Mitgliedern (Abfrageanbietern) zu gestatten, Abfragen ohne Überprüfung zu erstellen. Mitglieder fügen in der benutzerdefinierten Analyseregeln Konten von Abfrageanbietern hinzu, die über die zulässigen Abfrageanbieter verfügen. Wenn der Abfrageanbieter das Mitglied ist, das Abfragen durchführen kann, könnten sie jede Abfrage direkt in der konfigurierten Tabelle ausführen. Abfrageanbieter könnten Abfragen auch erstellen, indem sie [Analysevorlagen erstellen](#). Alle Abfragen, die von den Abfrageanbietern erstellt wurden, dürfen automatisch für die Tabelle in allen Kollaborationen ausgeführt werden, in denen die AWS-Konto vorhanden und die Tabelle verknüpft ist.

Datenbesitzer können nur Analysevorlagen oder Konten erlauben, Abfragen zu erstellen, nicht beides. Wenn der Datenbesitzer dieses Feld leer lässt, kann das Mitglied, das Abfragen durchführen kann, keine Abfragen für die konfigurierte Tabelle ausführen.

Topics

- [Benutzerdefinierte Analyseregeln, vordefinierte Struktur](#)
- [Beispiel für eine benutzerdefinierte Analyseregeln](#)
- [Benutzerdefinierte Analyseregeln mit differenziellem Datenschutz](#)

Benutzerdefinierte Analyseregeln, vordefinierte Struktur

Das folgende Beispiel enthält eine vordefinierte Struktur, die zeigt, wie Sie eine benutzerdefinierte Analyseregeln mit aktiviertem differenziellen Datenschutz abschließen. Der `userIdentifizier` Wert ist die Spalte, die Ihre Benutzer eindeutig identifiziert, z. B. `user_id`. Wenn Sie in einer Kollaboration zwei oder mehr Tabellen mit aktiviertem differenziellen Datenschutz haben, AWS Clean Rooms müssen Sie in beiden Analyseregeln dieselbe Spalte wie die Benutzer-ID-Spalte konfigurieren, um eine konsistente Definition der Benutzer in allen Tabellen aufrechtzuerhalten.

```
{
  "allowedAnalyses": ["ANY_QUERY"] | string[],
  "allowedAnalysisProviders": [],
```

```
"differentialPrivacy": {
  "columns": [
    {
      "name": "userIdentifier"
    }
  ]
}
```

Führen Sie dazu einen der folgenden Schritte aus:

- Fügen Sie der Steuerung „Zulässige Analysen“ eine Analysevorlage ARNs hinzu. In diesem Fall ist das `allowedAnalysisProviders` Steuerelement nicht enthalten.

```
{
  allowedAnalyses: string[]
}
```

- Fügt AWS-Konto IDs dem `allowedAnalysisProviders` Steuerelement ein Mitglied hinzu. In diesem Fall fügen Sie dem `allowedAnalyses` Steuerelement etwas `ANY_QUERY` hinzu.

```
{
  allowedAnalyses: ["ANY_QUERY"],
  allowedAnalysisProviders: string[]
}
```

Beispiel für eine benutzerdefinierte Analyseregeln

Das folgende Beispiel zeigt, wie zwei Unternehmen AWS Clean Rooms mithilfe der benutzerdefinierten Analyseregeln zusammenarbeiten können.

Unternehmen A verfügt über Kunden- und Vertriebsdaten. Unternehmen A ist daran interessiert, die Umsatzsteigerung einer Werbekampagne auf der Website von Unternehmen B zu verstehen. Unternehmen B verfügt über Zuschauerdaten und Segmentattribute, die für Unternehmen nützlich sind (z. B. das Gerät, mit dem sie sich die Werbung angesehen haben).

Unternehmen A hat eine spezielle Inkrementalitätsabfrage, die im Rahmen der Zusammenarbeit ausgeführt werden soll.

Um eine Zusammenarbeit zu erstellen und gemeinsam eine benutzerdefinierte Analyse durchzuführen, gehen die Unternehmen wie folgt vor:

1. Unternehmen A erstellt eine Kollaboration und erstellt eine Mitgliedschaft. Die Kollaboration hat Firma B als weiteres Mitglied der Kollaboration. Unternehmen A aktiviert die Abfrageprotokollierung in der Kollaboration und sie aktiviert die Abfrageprotokollierung in ihrem Konto.
2. Unternehmen B erstellt eine Mitgliedschaft in der Kollaboration. Es aktiviert die Abfrageprotokollierung in seinem Konto.
3. Firma A erstellt eine für CRM konfigurierte Tabelle
4. Unternehmen A fügt der für den Vertrieb konfigurierten Tabelle eine leere benutzerdefinierte Analyseregeln hinzu.
5. Firma A ordnet der Kollaboration eine für den Vertrieb konfigurierte Tabelle zu.
6. Unternehmen B erstellt eine für die Zuschauerzahl konfigurierte Tabelle.
7. Unternehmen B fügt der für die Zuschauerzahl konfigurierten Tabelle eine leere benutzerdefinierte Analyseregeln hinzu.
8. Unternehmen B ordnet der Kollaboration eine für die Zuschauerzahl konfigurierte Tabelle zu.
9. Unternehmen A zeigt die der Kollaboration zugeordnete Verkaufstabelle und die Tabelle mit den Zuschauerzahlen an und erstellt eine Analysevorlage, in der die Inkrementalitätsabfrage und der Parameter für den Kampagnenmonat hinzugefügt werden.

```
{
  "analysisParameters": [
    {
      "defaultValue": ""
      "type": "DATE"
      "name": "campaign_month"
    }
  ],
  "description": "Monthly incrementality query using sales and viewership data"
  "format": "SQL"
  "name": "Incrementality analysis"
  "source":
    "WITH labeleddata AS
    (
      SELECT hashedemail, deviceid, purchases, unitprice, purchasedate,
      CASE
        WHEN testvalue IN ('value1', 'value2', 'value3') THEN 0
        ELSE 1
      END AS testgroup
    FROM viewershipdata
```

```

    )
    SELECT labeleddata.purchases, provider.impressions
    FROM labeleddata
    INNER JOIN salesdata
        ON labeleddata.hashedemail = provider.hashedemail
    WHERE MONTH(labeleddata.purchasedate) > :campaignmonth
    AND testgroup = :group
    "
}

```

10. Unternehmen A fügt sein Konto (z. B. 444455556666) zur Steuerung des zulässigen Analyseanbieters in der benutzerdefinierten Analyseregeln hinzu. Sie verwenden das Steuerelement für zugelassene Analyseanbieter, weil sie zulassen möchten, dass alle von ihnen erstellten Abfragen in ihrer für den Vertrieb konfigurierten Tabelle ausgeführt werden.

```

{
  "allowedAnalyses": [
    "ANY_QUERY"
  ],
  "allowedAnalysisProviders": [
    "444455556666"
  ]
}

```

11. Unternehmen B sieht die erstellte Analysevorlage in der Kollaboration und überprüft ihren Inhalt, einschließlich der Abfragezeichenfolge und des Parameters.
12. Unternehmen B stellt fest, dass die Analysevorlage den Anwendungsfall Inkrementalität erfüllt und die Datenschutzanforderungen hinsichtlich der Art und Weise, wie die für die Zuschauerzahl konfigurierte Tabelle abgefragt werden kann, erfüllt.
13. Unternehmen B fügt den ARN der Analysevorlage zur zulässigen Analysesteuerung in der benutzerdefinierten Analyseregeln der Zuschauerschaftstabelle hinzu. Sie verwenden das zulässige Analysesteuerelement, weil sie nur zulassen möchten, dass die Inkrementalitätsabfrage für ihre für die Zuschauerzahl konfigurierte Tabelle ausgeführt wird.

```

{
  "allowedAnalyses": [
    "arn:aws:cleanrooms:us-east-1:111122223333:membership/41327cc4-bbf0-43f1-b70c-a160dddceb08/analysistemplate/1ff1bf9d-781c-418d-a6ac-2b80c09d6292"
  ]
}

```

14. Unternehmen A führt die Analysevorlage aus und verwendet den Parameterwert. 05-01-2023

Benutzerdefinierte Analyseregeln mit differenziellem Datenschutz

In AWS Clean Rooms, die benutzerdefinierte Analyseregeln unterstützen differenziellen Datenschutz. Differential Privacy ist ein mathematisch strenger Rahmen für den Datenschutz, der Ihnen hilft, Ihre Daten vor Reidentifikationsversuchen zu schützen.

Differential Privacy unterstützt aggregierte Analysen wie die Planung und post-ad-campaign Messung von Werbekampagnen, Benchmarking in einem Konsortium von Finanzinstituten und A/B-Tests für die Gesundheitsforschung.

Die unterstützte Abfragestruktur und Syntax sind in definiert. [Struktur und Syntax der Abfrage](#)

Beispiel für eine benutzerdefinierte Analyseregeln mit differenziertem Datenschutz

Note

AWS Clean Rooms Differential Privacy ist nur für Kollaborationen verfügbar, bei denen die Daten in Amazon S3 gespeichert sind.

Sehen Sie sich das [Beispiel für eine benutzerdefinierte Analyseregeln](#) an, das im vorherigen Abschnitt vorgestellt wurde. Dieses Beispiel zeigt, wie Sie Differential Privacy nutzen können, um Ihre Daten vor Reidentifikationsversuchen zu schützen und gleichzeitig Ihrem Partner die Möglichkeit zu geben, geschäftskritische Erkenntnisse aus Ihren Daten zu gewinnen. Gehen Sie davon aus, dass Unternehmen B, das über die Zuschauerdaten verfügt, seine Daten mithilfe von Differential Privacy schützen möchte. Um die Einrichtung des differenzierten Datenschutzes abzuschließen, führt Unternehmen B die folgenden Schritte durch:

1. Unternehmen B aktiviert den differenziellen Datenschutz und fügt gleichzeitig eine benutzerdefinierte Analyseregeln zur konfigurierten Tabelle für die Zuschauerzahl hinzu. Unternehmen B wählt diese `viewershipdata.hashemail` Spalte als Benutzer-ID aus.
2. Unternehmen B [fügt der Zusammenarbeit eine differenzierte Datenschutzrichtlinie](#) hinzu, um die Tabelle mit den Zuschauerzahlen für Abfragen verfügbar zu machen. Unternehmen B wählt die Standardrichtlinie aus, um die Einrichtung schnell abzuschließen.

Unternehmen A, das die Umsatzsteigerung einer Werbekampagne auf der Website von Unternehmen B verstehen möchte, führt die Analysevorlage aus. Da die Abfrage mit der allgemeinen [Abfragestruktur von AWS Clean Rooms Differential Privacy kompatibel ist, wird die Abfrage](#) erfolgreich ausgeführt.

Struktur und Syntax der Abfrage

Abfragen, die mindestens eine Tabelle enthalten, für die Differential Privacy aktiviert ist, müssen der folgenden Syntax entsprechen.

```
query_statement:
    [cte, ...] final_select

cte:
    WITH sub_query AS (
        inner_select
        [ UNION | INTERSECT | UNION_ALL | EXCEPT/MINUS ]
        [ inner_select ]
    )

inner_select:
    SELECT [user_id_column, ] expression [, ...]
    FROM table_reference [, ...]
    [ WHERE condition ]
    [ GROUP BY user_id_column[, expression] [, ...] ]
    [ HAVING condition ]

final_select:
    SELECT [expression, ...] | COUNT | COUNT_DISTINCT | SUM | AVG | STDDEV
    FROM table_reference [, ...]
    [ WHERE condition ]
    [ GROUP BY expression [, ...] ]
    [ HAVING COUNT | COUNT_DISTINCT | SUM | AVG | STDDEV | condition ]
    [ ORDER BY column_list ASC | DESC ]
    [ OFFSET literal ]
    [ LIMIT literal ]

expression:
    column_name [, ...] | expression AS alias | aggregation_functions |
    window_functions_on_user_id | scalar_function | CASE | column_name math_expression [,
    expression]

window_functions_on_user_id:
```

```
function () OVER (PARTITION BY user_id_column, [column_name] [ORDER BY column_list  
ASC|DESC])
```

Note

Beachten Sie bei der Struktur und Syntax von Differential Privacy Abfragen Folgendes:

- Unterabfragen werden nicht unterstützt.
- Common Table Expressions (CTEs) sollte die Benutzer-ID-Spalte ausgeben, wenn eine Tabelle oder ein CTE Daten enthält, die durch Differential Privacy geschützt sind. Filter, Gruppierungen und Aggregationen sollten auf Benutzerebene vorgenommen werden.
- Final_Select ermöglicht die Aggregatfunktionen COUNT DISTINCT, COUNT, SUM, AVG und STDDEV.

Weitere Informationen darüber, welche SQL-Schlüsselwörter für Differential Privacy unterstützt werden, finden Sie unter [SQL-Funktionen von AWS Clean Rooms Differential Privacy](#)

Regel zur Analyse von ID-Zuordnungstabellen

In AWS Clean Rooms ist eine Analyseregeln für ID-Zuordnungstabellen keine eigenständige Analyseregeln. Diese Art von Analyseregeln wird von unterschiedlichen Identitätsdaten verwaltet AWS Clean Rooms und verwendet, um das Abfragen zu erleichtern. Sie wird automatisch zu ID-Zuordnungstabellen hinzugefügt und kann nicht bearbeitet werden. Es erbt das Verhalten der anderen Analyseregeln in der Zusammenarbeit — sofern diese Analyseregeln homogen sind.

Die Analyseregeln für die ID-Zuordnungstabelle erzwingt die Sicherheit einer ID-Zuordnungstabelle. Sie verhindert, dass ein Mitglied der Kollaboration anhand der ID-Zuordnungstabelle die Population, die sich nicht überschneidet, zwischen den Datensätzen der beiden Mitglieder direkt auswählt oder überprüft. Die Analyseregeln für die ID-Zuordnungstabelle wird verwendet, um die sensiblen Daten in der ID-Zuordnungstabelle zu schützen, wenn sie in Abfragen mit impliziten anderen Analyseregeln verwendet wird.

AWS Clean Rooms Erzwingt mit der Analyseregeln für ID-Zuordnungstabellen in erweitertem SQL eine Überlappung auf beiden Seiten der ID-Zuordnungstabelle. Auf diese Weise können Sie die folgenden Aufgaben ausführen:

- Verwenden Sie die Überlappung der ID-Zuordnungstabelle in JOIN Anweisungen.

AWS Clean Rooms erlaubt einen INNERLEFT, oder RIGHT -Join in der ID-Zuordnungstabelle, sofern dabei die Überlappung berücksichtigt wird. Um vertrauliche Zuordnungsinformationen zu schützen, muss sich die ID-Zuordnungstabelle bei jeder JOIN Operation immer auf der Seite inner "" befinden. Beispielsweise sind die folgenden JOIN Operationen gültig:

- table LEFT JOIN id_mapping_table
- id_mapping_table RIGHT JOIN table
- table INNER JOIN id_mapping_table

Die folgenden JOIN Operationen sind ungültig:

- id_mapping_table LEFT JOIN table
- table RIGHT JOIN id_mapping_table

Dadurch wird verhindert, dass Kartendatensätze offengelegt werden, für die Ihr Datensatz keine entsprechenden Treffer enthält. Wenn Sie solche Operationen zulassen, könnten möglicherweise vertrauliche Informationen über die Datenzuordnungen anderer Kollaborationsmitglieder preisgegeben werden.

- Verwenden Sie die Spalten der Zuordnungstabelle in Anweisungen. JOIN

Sie können die Spalten der Zuordnungstabelle nicht in den folgenden Anweisungen verwenden: SELECT, WHEREHAVING, oder ORDER BY (es sei denn GROUP BY, die Schutzmaßnahmen für die Quell-ID-Namespaces-Zuordnung oder die Ziel-ID-Namespaces-Zuordnung wurden geändert).

- Unterstützt in erweitertem SQL AWS Clean Rooms auch implizit OUTER JOIN und CROSSJOIN. JOIN Diese Verknüpfungen können die Anforderungen an Überschneidungen nicht erfüllen. Wird stattdessen AWS Clean Rooms verwendet, requireOverlap um anzugeben, für welche Spalten eine Verbindung hergestellt werden muss.

Die unterstützte Abfragestruktur und Syntax sind in definiert. [Struktur und Syntax der Abfrage in der ID-Zuordnungstabelle](#)

Zu den Parametern der Analyseregeln, die unter definiert sind [Steuerelemente für die Abfrage von ID-Zuordnungstabellen für Analysen](#), gehören Abfragesteuerelemente und Steuerelemente für Abfrageergebnisse. Zu den Abfragesteuerelementen gehört die Möglichkeit, eine Überlappung der ID-Zuordnungstabelle in JOIN Anweisungen vorzuschreiben (d. h. requireOverlap).

Topics

- [Struktur und Syntax der Abfrage in der ID-Zuordnungstabelle](#)
- [Steuerelemente für die Abfrage von ID-Zuordnungstabellen für Analysen](#)
- [Vordefinierte Struktur der Regel für die Analyse der ID-Zuordnungstabelle](#)
- [Analyseregel für ID-Zuordnungstabellen — Beispiel](#)

Struktur und Syntax der Abfrage in der ID-Zuordnungstabelle

Abfragen für Tabellen, für die eine Analyseregel für ID-Zuordnungstabellen gilt, müssen der folgenden Syntax entsprechen.

```
--select_list_expression
SELECT
provider.data_col, consumer.data_col

--table_expression
FROM provider

JOIN idMappingTable idmt ON provider.id = idmt.sourceId

JOIN consumer ON consumer.id = idmt.targetId
```

Tabellen für die Zusammenarbeit

Die folgenden Tabellen stellen konfigurierte Tabellen dar, die in einer AWS Clean Rooms Kollaboration existieren. Die ID-Spalte in den Tabellen `cr_drivers_license` und `cr_insurance` stellt eine Spalte dar, die mit der ID-Zuordnungstabelle übereinstimmt.

`cr_drivers_license`

id	Treibername	Staat der Registrierung
1	Eduard	TX
2	Dana	MA
3	Gweneth	IL

Autoversicherung

id	E-Mail des Versicherungsnehmers	Policy_number
a	eduardo@internal.company.com	17f9d04e-f5be-4426-bdc4-250ed59c6529
b	gwen@internal.company.com	3f0092db-2316-48a8-8d44-09cf8f6e6c64
c	rosa@internal.company.com	d7692e84-3d3c-47b8-b46d-a0d5345f0601

ID-Zuordnungstabelle

Die folgende Tabelle stellt eine bestehende ID-Zuordnungstabelle dar, die mit den Tabellen `cr_drivers_license` und `cr_insurance` übereinstimmt. Nicht alle Einträge sind für beide Kollaborationstabellen gültig. IDs

cr_drivers_license_id	cr_insurance_id
1	a
2	Null
3	b
Null	c

Die Analyseregeln für die ID-Zuordnungstabelle lässt nur zu, dass Abfragen für den Satz sich überschneidender Daten ausgeführt werden, was wie folgt aussehen würde:

cr_driver_license_id	cr_insurance_id	Treibernamen	Staat der Registrierung	E-Mail des Versicherungsnehmers	Policy_number

1	a	Eduard	TX	eduardo@internal.com	17f9d04e-f5be-4426-bdc4-250ed59c6529
3	b	Gweneth	IL	gwen@internal.com	3f0092db-2316-48a8-8d44-09cf8f6e6c64

Beispielabfragen

Die folgenden Beispiele zeigen gültige Speicherorte für die Verknüpfungen der ID-Zuordnungstabellen:

```
-- Single ID mapping table
SELECT
    [ select_items ]FROM
    cr_drivers_license cr_dl
    [ INNER | LEFT ] JOIN cr_identity_mapping_table idmt ON idmt.cr_drivers_license_id
= cr_dl.id
    [ INNER | RIGHT ] JOIN cr_insurance cr_in          ON idmt.cr_insurance_id
= cr_in.id
;
-- Single ID mapping table (Subquery)
SELECT
    [ select_items ]FROM (
    SELECT
        [ select_items ]
    FROM
        cr_drivers_license cr_dl
        [ INNER | LEFT ] JOIN cr_identity_mapping_table idmt ON
idmt.cr_drivers_license_id = cr_dl.id
        [ INNER | RIGHT ] JOIN cr_insurance cr_in          ON idmt.cr_insurance_id
= cr_in.id
    )
;
-- Single ID mapping table (CTE)
WITH
    matched_ids AS (
        SELECT
```

```

        [ select_items ]
    FROM
        cr_drivers_license cr_dl
        [ INNER | LEFT ] JOIN cr_identity_mapping_table idmt ON
idmt.cr_drivers_license_id = cr_dl.id
        [ INNER | RIGHT ] JOIN cr_insurance cr_in          ON
idmt.cr_insurance_id      = cr_in.id
    )SELECT
    [ select_items ]FROM
    matched_ids
;

```

Überlegungen

Beachten Sie bei der Struktur und Syntax der Abfrage von ID-Zuordnungstabellen Folgendes:

- Sie können es nicht bearbeiten.
- Es wird standardmäßig auf die ID-Zuordnungstabelle angewendet.
- Es verwendet eine Quell- und Ziel-ID-Namespaces-Zuordnung innerhalb der Kollaboration.
- Die ID-Zuordnungstabelle ist standardmäßig so konfiguriert, dass sie Standardschutz für die Spalte bietet, die aus dem ID-Namespaces stammt. Sie können diese Konfiguration so ändern, dass die Spalte, die aus dem ID-Namespaces stammt (entweder `sourceID` oder `targetID`), an beliebiger Stelle in der Abfrage zulässig ist. Weitere Informationen finden Sie unter [ID-Namespaces in AWS Clean Rooms](#).
- Die Analyseregeln für die ID-Zuordnungstabelle erbt die SQL-Einschränkungen der anderen Analyseregeln in der Kollaboration.

Steuerelemente für die Abfrage von ID-Zuordnungstabellen für Analysen

Steuert mit Abfragesteuerelementen für ID-Zuordnungstabellen, AWS Clean Rooms wie die Spalten in Ihrer Tabelle zur Abfrage der Tabelle verwendet werden. Sie steuert beispielsweise, welche Spalten für die Verknüpfung verwendet werden und welche Spalten sich überlappen müssen. Die Analyseregeln für ID-Zuordnungstabellen umfasst auch Funktionen, mit denen Sie die Projektion von `sourceID`/`targetID`, dem oder beiden zulassen können, ohne dass ein JOIN erforderlich ist.

In der folgenden Tabelle werden die einzelnen Steuerelemente erläutert.

Steuerung	Definition	Usage
<code>joinColumns</code>	Die Spalten, die das Mitglied, das Abfragen durchführen kann, in der INNER JOIN-Anweisung verwenden kann.	Sie können sie <code>joinColumns</code> in keinem anderen Teil der Abfrage als INNER JOIN verwenden. Weitere Informationen finden Sie unter Steuerelemente verbinden.
<code>dimensionColumns</code>	Die Spalten (falls vorhanden), die das Mitglied, das Abfragen durchführen kann, in SELECT- und GROUP BY-Anweisungen verwenden kann.	A <code>dimensionColumn</code> kann in SELECT und verwendet werden GROUPBY. A <code>dimensionColumn</code> kann erscheinen als <code>joinKeys</code>. Sie können es <code>dimensionColumns</code> in der JOIN-Klausel nur verwenden, wenn Sie es in Klammern angeben.
<code>queryConstraints:RequireOverlap</code>	Die Spalten in der ID-Zuordnungstabelle, die verknüpft werden müssen, damit die Abfrage ausgeführt werden kann.	Diese Spalten müssen verwendet werden, um die ID-Zuordnungstabelle und eine Kollaborationstabelle miteinander zu verknüpfen.

Vordefinierte Struktur der Regel für die Analyse der ID-Zuordnungstabelle

Die vordefinierte Struktur für eine Analyseregeln für ID-Zuordnungstabellen enthält Standardschutzmaßnahmen, die auf das und angewendet werden. `sourceID targetID` Das bedeutet, dass die Spalte mit den angewendeten Schutzmaßnahmen in Abfragen verwendet werden muss.

Sie können die Analyseregeln für die ID-Zuordnungstabelle auf folgende Weise konfigurieren:

- Beides sourceID und targetID geschützt

In dieser Konfiguration targetID können sourceID sowohl das als auch das andere projiziert werden. Das sourceID und targetID muss in einem JOIN verwendet werden, wenn auf die ID-Zuordnungstabelle verwiesen wird.

- Nur targetID geschützt

In dieser Konfiguration targetID kann das nicht projiziert werden. Das targetID muss in einem JOIN verwendet werden, wenn auf die ID-Zuordnungstabelle verwiesen wird. Das sourceID kann in einer Abfrage verwendet werden.

- Nur sourceID geschützt

In dieser Konfiguration sourceID kann das nicht projiziert werden. Das sourceID muss in einem JOIN verwendet werden, wenn auf die ID-Zuordnungstabelle verwiesen wird. Das targetID kann in einer Abfrage verwendet werden.

- Weder sourceID noch targetID geschützt

In dieser Konfiguration unterliegt die ID-Zuordnungstabelle keiner bestimmten Erzwingung, die in Abfragen verwendet werden kann.

Das folgende Beispiel zeigt eine vordefinierte Struktur für eine Analyseregeln für ID-Zuordnungstabellen, bei der die Standardschutzmaßnahmen auf und angewendet werden.

sourceID targetID In diesem Beispiel erlaubt die Analyseregeln für die ID-Zuordnungstabelle nur einen INNER JOIN sowohl für die Spalte als auch für die sourceID Spalte. targetID

```
{
  "joinColumns": [
    "source_id",
    "target_id"
  ],
  "queryConstraints": [
    {
      "requireOverlap": {
        "columns": [
          "source_id",
          "target_id"
        ]
      }
    }
  ]
}
```

```

],
"dimensionColumns": [] // columns that can be used in SELECT and JOIN
}

```

Das folgende Beispiel zeigt eine vordefinierte Struktur für eine Analyseregeln für eine ID-Zuordnungstabelle, auf die Schutzmaßnahmen angewendet wurden. `targetID` In diesem Beispiel erlaubt die Analyseregeln für ID-Zuordnungstabellen nur einen INNER JOIN für die Spalte. `sourceID`

```

{
  "joinColumns": [
    "source_id",
    "target_id"
  ],
  "queryConstraints": [
    {
      "requireOverlap": {
        "columns": [
          "target_id"
        ]
      }
    }
  ],
  "dimensionColumns": [
    "source_id"
  ]
}

```

Das folgende Beispiel zeigt eine vordefinierte Struktur für eine Analyseregeln für ID-Zuordnungstabellen mit Schutzmaßnahmen für. `sourceID` In diesem Beispiel erlaubt die Analyseregeln für ID-Zuordnungstabellen nur einen INNER JOIN für die Spalte. `targetID`

```

{
  "joinColumns": [
    "source_id",
    "target_id"
  ],
  "queryConstraints": [
    {
      "requireOverlap": {
        "columns": [
          "source_id"
        ]
      }
    }
  ]
}

```

```

    }
  }
],
"dimensionColumns": [
  "target_id"
]
}

```

Das folgende Beispiel zeigt eine vordefinierte Struktur für eine Analyseregeln für ID-Zuordnungstabellen ohne Schutzmaßnahmen für das Oder. `sourceID` `targetID` In diesem Beispiel ermöglicht die Analyseregeln für die ID-Zuordnungstabelle einen INNER JOIN sowohl für die Spalte als auch für die `sourceID` Spalte. `targetID`

```

{
  "joinColumns": [
    "source_id",
    "target_id"
  ],
  "queryConstraints": [
    {
      "requireOverlap": {
        "columns": []
      }
    }
  ],
  "dimensionColumns": [
    "source_id",
    "target_id"
  ]
}

```

Analyseregeln für ID-Zuordnungstabellen — Beispiel

Anstatt eine lange Wasserfall-Anweisung zu verfassen, die beispielsweise auf personenbezogene Daten (PII) verweist, können Unternehmen die Analyseregeln für die ID-Zuordnungstabelle verwenden, um die Transcodierung mehrerer Parteien LiveRamp zu verwenden. Das folgende Beispiel zeigt, wie Sie gemeinsam die Analyseregeln für ID-Zuordnungstabellen AWS Clean Rooms verwenden können.

Unternehmen A ist ein Werbetreibender, der über Kunden- und Verkaufsdaten verfügt, die als Quelle verwendet werden. Unternehmen A führt auch die Transcodierung im Namen der an der Zusammenarbeit beteiligten Parteien durch und bringt die LiveRamp Anmeldeinformationen mit.

Firma B ist ein Herausgeber, der über Veranstaltungsdaten verfügt, die als Ziel verwendet werden.

 Note

Entweder Unternehmen A oder Unternehmen B können Anmeldeinformationen für die LiveRamp Transcodierung bereitstellen und die Transcodierung durchführen.

Um eine Zusammenarbeit aufzubauen, die die Analyse der ID-Zuordnungstabellen in Zusammenarbeit ermöglicht, gehen die Unternehmen wie folgt vor:

1. Unternehmen A erstellt eine Kollaboration und erstellt eine Mitgliedschaft. Es fügt Unternehmen B hinzu, das auch eine Mitgliedschaft in der Kollaboration einrichtet.
2. Firma A ordnet entweder eine vorhandene ID-Namespace-Quelle zu oder erstellt AWS Entity Resolution mithilfe der AWS Clean Rooms Konsole eine neue.

Firma A erstellt eine konfigurierte Tabelle mit ihren Verkaufsdaten und einer Spalte mit dem Schlüssel `sourceId` in der ID-Zuordnungstabelle.

Die ID-Namespace-Quelle stellt Daten für die Transcodierung bereit.

3. Firma B ordnet entweder ein vorhandenes ID-Namespace-Ziel zu oder erstellt AWS Entity Resolution mithilfe der Konsole ein neues. AWS Clean Rooms

Firma B erstellt eine konfigurierte Tabelle mit ihren Ereignisdaten und einer Spalte mit dem Schlüssel `targetId` in der ID-Zuordnungstabelle.

Das ID-Namespace-Ziel stellt keine Daten für die Transcodierung bereit, sondern nur Metadaten rund um die Konfiguration. LiveRamp

4. Unternehmen A erkennt die beiden ID-Namespace, die der Kollaboration zugeordnet sind, und erstellt eine ID-Zuordnungstabelle und füllt sie auf.
5. Unternehmen A führt eine Abfrage über die beiden Datensätze durch, indem es die ID-Zuordnungstabelle verknüpft.

```
--- this would be valid for Custom or List
SELECT provider.data_col, consumer.data_col
FROM provider
  JOIN idMappingTable-123123123123-myMappingWFName idmt
    ON provider.id = idmt.sourceId
  JOIN consumer
```

```
ON consumer.id = idmt.targetId
```

AWS Clean Rooms Differenzierter Datenschutz

AWS Clean Rooms Differential Privacy hilft Ihnen dabei, die Privatsphäre Ihrer Benutzer mit einer mathematisch gestützten Technik zu schützen, die mit intuitiven Steuerungen mit wenigen Klicks implementiert wird. Da es sich um eine vollständig verwaltete Funktion handelt, sind keine vorherigen Erfahrungen mit differenziertem Datenschutz erforderlich, um die erneute Identifizierung Ihrer Benutzer zu verhindern. AWS Clean Rooms fügt den Abfrageergebnissen zur Laufzeit automatisch eine sorgfältig kalibrierte Menge an Rauschen hinzu, um Ihre Daten auf individueller Ebene zu schützen.

AWS Clean Rooms Differential Privacy unterstützt eine Vielzahl analytischer Abfragen und eignet sich für eine Vielzahl von Anwendungsfällen, bei denen ein geringfügiger Fehler in den Abfrageergebnissen die Nützlichkeit Ihrer Analyse nicht beeinträchtigt. Damit können Ihre Partner geschäftskritische Erkenntnisse über Werbekampagnen, Investitionsentscheidungen, klinische Studien und mehr gewinnen, ohne dass Ihre Partner zusätzliche Einstellungen vornehmen müssen.

AWS Clean Rooms Differential Privacy schützt vor Überlauffehlern oder ungültigen Umwandlungsfehlern, bei denen Skalarfunktionen oder mathematische Operatorsymbole auf böswillige Weise verwendet werden.

Weitere Informationen zu AWS Clean Rooms Differential Privacy finden Sie in den folgenden Themen.

Themen

- [Differentieller Datenschutz](#)
- [So funktioniert Differential Privacy AWS Clean Rooms](#)
- [Differenzielle Datenschutzrichtlinie](#)
- [SQL-Funktionen von AWS Clean Rooms Differential Privacy](#)
- [Tipps und Beispiele für Differential Privacy-Abfragen](#)
- [Einschränkungen von AWS Clean Rooms Differential Privacy](#)

Differentieller Datenschutz

Differenzierter Datenschutz ermöglicht nur aggregierte Erkenntnisse und verschleiert den Beitrag der Daten einer Person zu diesen Erkenntnissen. Differentieller Datenschutz schützt die Kooperationsdaten vor dem Mitglied, das Ergebnisse erhalten kann, wenn es mehr über eine bestimmte Person erfährt. Ohne Differential Privacy kann das Mitglied, das Ergebnisse erhalten kann, versuchen, auf individuelle Benutzerdaten zu schließen, indem es Datensätze über eine Person hinzufügt oder entfernt und den Unterschied zwischen den Abfrageergebnissen beobachtet.

Wenn die Option „Differenzierter Datenschutz“ aktiviert ist, wird den Abfrageergebnissen eine bestimmte Menge an Rauschen hinzugefügt, um den Beitrag einzelner Benutzer zu verschleiern. Wenn das Mitglied, das Ergebnisse erhalten kann, versucht, den Unterschied in den Abfrageergebnissen zu beobachten, nachdem es Datensätze über eine Person aus seinem Datensatz entfernt hat, verhindert die Variabilität des Abfrageergebnisses, dass die Daten der Person identifiziert werden können. AWS Clean Rooms Differential Privacy verwendet den [SampCert](#) Sampler, eine bewährte Correct-Sampler-Implementierung, die von entwickelt wurde. AWS

So funktioniert Differential Privacy AWS Clean Rooms

Der Workflow zur Aktivierung des differenziellen AWS Clean Rooms Datenschutzes in erfordert die folgenden zusätzlichen Schritte, wenn [der Workflow abgeschlossen wird für AWS Clean Rooms](#):

1. Sie aktivieren den differenziellen Datenschutz, wenn Sie eine [benutzerdefinierte Analyseregel](#) hinzufügen.
2. [Sie konfigurieren die differenzielle Datenschutzrichtlinie für die Zusammenarbeit](#), um Ihre Datentabellen, die mit differentiellem Datenschutz geschützt sind, für Abfragen verfügbar zu machen.

Nachdem Sie diese Schritte abgeschlossen haben, kann das Mitglied, das Abfragen durchführen kann, Abfragen für durch Differential Privacy geschützte Daten ausführen. AWS Clean Rooms gibt Ergebnisse zurück, die der differenziellen Datenschutzrichtlinie entsprechen. AWS Clean Rooms Differential Privacy verfolgt die geschätzte Anzahl der verbleibenden Abfragen, die Sie ausführen können, ähnlich der Tankanzeige in einem Auto, die Ihnen den aktuellen Kraftstoffstand des Fahrzeugs anzeigt. Die Anzahl der Abfragen, die ein Mitglied, das Abfragen durchführen kann, ist durch das Datenschutzbudget und die in der festgelegten Parameter für die Anzahl der pro Abfrage hinzugefügten Störungen begrenzt [Differenzielle Datenschutzrichtlinie](#).

Überlegungen

Beachten Sie bei der Verwendung von Differential Privacy in AWS Clean Rooms Folgendes:

- Das Mitglied, das Ergebnisse erhalten kann, kann Differential Privacy nicht verwenden. Sie konfigurieren eine benutzerdefinierte Analyseregeln mit deaktiviertem Differential Privacy für ihre konfigurierten Tabellen.
- Das Mitglied, das Abfragen durchführen kann, kann keine Tabellen von zwei oder mehr Datenanbietern verknüpfen, wenn für beide der differenzielle Datenschutz aktiviert ist.

Differenzielle Datenschutzrichtlinie

Die differenzielle Datenschutzrichtlinie legt fest, wie viele Aggregationsfunktionen das Mitglied, das Abfragen durchführen kann, in einer Kollaboration ausführen darf. Das Datenschutzbudget definiert eine gemeinsame, begrenzte Ressource, die auf alle Tabellen in einer Kollaboration angewendet wird. Das pro Abfrage hinzugefügte Rauschen bestimmt, wie schnell das Datenschutzbudget aufgebraucht wird.

Eine differenzielle Datenschutzrichtlinie ist erforderlich, um Ihre durch Differentialdatenschutz geschützten Tabellen für Abfragen verfügbar zu machen. Dies ist ein einmaliger Schritt in einer Zusammenarbeit und umfasst zwei Eingaben:

- **Datenschutzbudget** — In Epsilon ausgedrückt, bestimmt das Datenschutzbudget das Datenschutzniveau. Es handelt sich um eine gemeinsame, begrenzte Ressource, die für all Ihre Tabellen verwendet wird, die in der Zusammenarbeit mit unterschiedlichem Datenschutz geschützt sind, da das Ziel darin besteht, die Privatsphäre Ihrer Benutzer zu schützen, deren Informationen in mehreren Tabellen vorhanden sein können.

Das Datenschutzbudget wird jedes Mal aufgebraucht, wenn eine Abfrage an Ihren Tabellen ausgeführt wird. Wenn das Datenschutzbudget vollständig aufgebraucht ist, kann das Collaboration-Mitglied, das Abfragen durchführen kann, keine weiteren Abfragen ausführen, bis es erhöht oder aktualisiert wird. Durch die Festlegung eines höheren Datenschutzbudgets kann das Mitglied, das Ergebnisse erhalten kann, seine Unsicherheit über die einzelnen Personen in den Daten verringern. Wählen Sie nach Rücksprache mit den Entscheidungsträgern im Unternehmen ein Datenschutzbudget, das Ihre Anforderungen an die Zusammenarbeit mit Ihren Datenschutzanforderungen in Einklang bringt.

Sie können die Option Datenschutzbudget monatlich aktualisieren auswählen, um jeden Kalendermonat automatisch ein neues Datenschutzbudget zu erstellen, wenn Sie planen, regelmäßig neue Daten in die Zusammenarbeit einzubeziehen. Wenn Sie diese Option wählen, können beliebig viele Informationen über Datenzeilen angezeigt werden, wenn diese bei Aktualisierungen wiederholt abgefragt werden. Vermeiden Sie diese Option, wenn dieselben Zeilen zwischen Aktualisierungen des Datenschutzbudgets wiederholt abgefragt werden.

- Das pro Anfrage hinzugefügte Rauschen wird anhand der Anzahl der Nutzer gemessen, deren Beiträge Sie unkenntlich machen möchten. Dieser Wert bestimmt, wie schnell das Datenschutzbudget aufgebraucht wird. Ein höherer Rauschwert verringert die Geschwindigkeit, mit der das Datenschutzbudget aufgebraucht wird, und ermöglicht somit, dass mehr Abfragen mit Ihren Daten ausgeführt werden können. Dies sollte jedoch gegen die Veröffentlichung weniger genauer Dateneinblicke abgewogen werden. Berücksichtigen Sie bei der Festlegung dieses Werts die gewünschte Genauigkeit für Erkenntnisse aus der Zusammenarbeit.

Sie können die standardmäßige differenzielle Datenschutzrichtlinie verwenden, um die Einrichtung schnell abzuschließen, oder Ihre differenzielle Datenschutzrichtlinie an Ihren Anwendungsfall anpassen. AWS Clean Rooms Differential Privacy bietet intuitive Steuerelemente zur Konfiguration der Richtlinie. AWS Clean Rooms Mit Differential Privacy können Sie eine Vorschau des Dienstprogramms im Hinblick auf die Anzahl der möglichen Aggregationen für alle Abfragen Ihrer Daten anzeigen und abschätzen, wie viele Abfragen in einer Datenzusammenarbeit ausgeführt werden können.

Anhand der interaktiven Beispiele können Sie sich ein Bild davon machen, wie sich unterschiedliche Werte für Privacy Budget und Noise, die pro Abfrage hinzugefügt werden, auf die Ergebnisse verschiedener Typen von SQL-Abfragen auswirken würden. Im Allgemeinen müssen Sie Ihre Datenschutzanforderungen mit der Anzahl der Abfragen, die Sie zulassen möchten, und der Genauigkeit dieser Abfragen abwägen. Ein kleineres Datenschutzbudget oder mehr Rauschen pro Anfrage können die Privatsphäre der Nutzer besser schützen, bieten Ihren Kooperationspartnern aber weniger aussagekräftige Erkenntnisse.

Wenn Sie das Datenschutzbudget erhöhen und gleichzeitig den Parameter „Pro Abfrage hinzugefügtes Rauschen“ beibehalten, kann das Mitglied, das Abfragen durchführen kann, mehr Aggregationen für Ihre Tabellen in der Kollaboration ausführen. Sie können das Datenschutzbudget jederzeit während der Zusammenarbeit erhöhen. Wenn Sie das Datenschutzbudget verringern und gleichzeitig den Parameter „Pro Abfrage hinzugefügtes Rauschen“ beibehalten, kann das Mitglied, das Abfragen durchführen kann, weniger Aggregationen ausführen. Sie können das

Datenschutzbudget nicht verringern, nachdem das Mitglied, das Abfragen durchführen kann, mit der Analyse Ihrer Daten begonnen hat.

Wenn Sie die Anzahl der pro Abfrage hinzugefügten Störungen erhöhen und gleichzeitig die Eingabe für das Datenschutzbudget beibehalten, kann das Mitglied, das Abfragen durchführen kann, mehr Aggregationen für Ihre Tabellen in der Kollaboration ausführen. Wenn Sie die Anzahl der pro Abfrage hinzugefügten Störungen verringern und gleichzeitig die Eingabe für das Datenschutzbudget beibehalten, kann das Mitglied, das Abfragen durchführen kann, weniger Aggregationen ausführen. Sie können das pro Abfrage hinzugefügte Rauschen jederzeit während der Zusammenarbeit erhöhen oder verringern.

Die differenzierte Datenschutzrichtlinie wird durch die API-Aktionen für die Vorlage „Datenschutzbudget“ verwaltet.

SQL-Funktionen von AWS Clean Rooms Differential Privacy

AWS Clean Rooms Differential Privacy verwendet eine allgemeine Abfragestruktur zur Unterstützung komplexer SQL-Abfragen. Benutzerdefinierte Analysevorlagen werden anhand dieser Struktur validiert, um sicherzustellen, dass sie auf Tabellen ausgeführt werden können, die durch Differential Privacy geschützt sind. Die folgende Tabelle zeigt, welche Funktionen unterstützt werden. Weitere Informationen finden Sie unter [Struktur und Syntax der Abfrage](#).

Abfrage

SELECT-

Tabella

Ansicht

Analytics-

OTES

gine

unterstüt

zte

SQL-

Konst

rukte

Aggregation

ANY_VALUE

Aggregati

Mittelwert

SELECT-

Tabellen

Ansicht

Metrics-

ETEs

gine

unterstüt

zte

SQL-

Konst

rukte

Bedingung

Aggregation

APPROXIMA

COUNT

PERCENTIL

STDDEV

Verwendun

SUM

AVG

Tabellen,

• Die

die Funktione

durch

Different

ial und

Privacy

geschützt

• Die

zu

Funktion

Daten

MAX

mit

• Die

Datensatz

en

au

Mitgeben
 SELECT-
 Tabellena
 Ansätze
 Metrics-
 DTEs
 gine
 unterstüt
 zte
 SQL-
 Konst
 runkte
 Bereizere
 be Funktion
 für MEN
 muss.
 Die
 Sie Funktion
 sollten
 PERCENTIL
 den CONT
 SELECT-
 • Die
 Au Funktion
 sdruck
 in
 STDDEV_SA
 denen
 MP
 schreiben
 und
 ' STDDEV_PO
 die
 `SELECT
 • Funktion
 userIdent
 n.
 ifierColu
 SUM,
 mn...'
 und
 das
 SUM
 Format
 DISTINCT
 CTEs
 • Die
 Funktion

Abfragen

SELECT-

Tabellen

Ansichten

Analytics-

ETEs

Engine

unterstüt

zte

SQL-

Konstruk

te

verwenden

. VAR_SAMP

und

VAR_POP

Mitgliedern

SELECT-

Tabellen

Ansichten

Analytics-

CTEs

gine

unterstüt

zte

SQL-

Konstrukte

rukte

WITH-

Kluster

sel,

Bedingung

Klusterstut

zel,

Abfrage

alge

CTEs

Verwendun

g

von

Tabellen,

die

durch

Differential

ial

Privacy

geschützt

sind,

zu

Daten

mit

Wählen Sie
SELECT-
Tabellen
Ansicht
Metrics-
DTEs
Engine
unterstüt
zte
SQL-
Konst
rukte

Datensatz
en
auf
Benutzere
bene
führen
muss.
Sie
sollten
den
SELECT-
Au
sdruck
in
denen
schreiben
,
die
`SELECT
userIdent
ifierColu
mn...`
das
Format

Abbildung
SELECT-
Tabellen
Ansicht
Analytics-
CTEs
Engine
unterstüt
zte
SQL-
Konst
rukte
CTEs
verwenden
.

Multi-tenancy

SELECT-

Tabellen

Ansichten

Metrics-

ETEs

gine

unterstüt

zte

SQL-

Konst

rukte

INSERT

abgeben

SHAVING

Konstrukt

• JOIN

en

• JOIN-

können

Bedi

Sie

ngung

jede

FROM

Unterabir

WHERE

verwenden

,

die

nicht

auf

unterschi

edliche

Datenschu

tzbeziehu

ngen

verweist.

Sie

können

jede

Mitgliedern
SELECT-
Tabellen
Ansicht
Metrics-
DTEs
gine
unterstüt
zte
SQL-
Konst
rukte

Unterabfr
age
verwenden
,
die
auf
unterschi
edliche
Datenschu
tzbeziehu
ngen
verweist,
nur
in
einer
FROM-
und
JOIN-
Klau
sel.

Multi-join

SELECT-

Tabellena

Ansicht

lytics-

ETEs

gine

unterstüt

zte

SQL-

Konst

rukte

JOINNER

JOIN

LEFT

Bedingung

unterstüt

• LINKS

zt, HALB

dass VERBINDEN

nur

• LINKS

JOIN-

GEGEN

Funk

JOIN

tionen

• RIGHT

unterstüt

zt JOIN

• FULL

STÄND

bed

GER

den

BEITRITT

es

BEITRETE

sich

un

ODER

Gleichver

Operator

knüpfunge

• CROSS

n JOIN

für

Wählen Sie
SELECT-
Tabellen
Ansicht
Metrics-
DTEs
Engine
unterstüt
zte
SQL-
Konst
rukte

Benutzer-
ID-
Spalte
n
handelt.
Diese
sind
erforderl
ich,
wenn
zwei
oder
mehr
Tabellen
mit
aktiviert
em
Differential
ial
Privacy
abgefragt
werden.
Stellen
Sie

Wird ein
SELECT-
Tabella
Ansicht
Metrics-
DTEs
gine
unterstüt
zte
SQL-
Konst
rukte
sicher,
dass
die
obligator
ischen
Equi-
Join
-
Bedingun
gen
korrekt
sind.
Vergewiss
ern
Sie
sich,
dass
der
Tabellenb
esitzer
in
allen
Tabellen
dieselbe

Wird ein

SELECT-

Tabellena

Ansicht

lytics-

ETEs

gine

unterstüt

zte

SQL-

Konst

rukte

Benutzer-

ID-

Spalte

konfiguri

ert

hat,

sodass

die

Definitio

n

eines

Benutzers

tabellenü

bergreife

nd

konsisten

t

bleibt.

CROSS

JOIN-

Funk

tionen

werden

Mitgliedern
SELECT-
Tabellen
Ansichten
Metrics-
DTEs
Engine
unterstüt
zte
SQL-
Konst
rukte
nicht
unterstüt
zt,
wenn
zwei
oder
mehr
Beziehung
en
mit
aktiviert
em
Different
ial
Privacy
kombinier
t
werden.

Aggregation

SELECT-

Tabellena

Ansicht

Metrics-

ETEs

gine

unterstüt

zte

SQL-

Konst

rukte

UNION,ra

UNIONüt

ALL,

INTERSECT

”

EXCEPT

||

MINUS

(das

sind

Synonyme))

Aggregation

SELECT-

Tabellena

Ansicht

Metrics-

ETEs

gine

unterstüt

zte

SQL-

Konst

rukte

Aggregation

unterstüt

unter

der

• Die

Bedingung

Fensterfu

unterstüt

z, AVG

dass

• Die

die

Fensterfu

Benutzer

IDaktion

Spalte

in CUME_DIST

der

Partition

funktion

der

• Die

Fensterfu

Fensterfu

nktion

erforderl

DENSE_RAN

ich

ist

• Die

wenn

Fensterfu

• **Aggregieren**

• **SELECT-**

• **Tabellena**

• **Ansicht**

• **Metrics-**

• **ETEs**

• **gine**

• **unterstüt**

• **zte**

• **SQL-**

• **Konst**

• **rukte**

• **Siektion**

• **FIRST_VAL**

• **Beziehung**

• **Die**

• **aktiviert**

• **Fensterfu**

• **nktion**

• **Differenz**

• **LAG**

• **ial**

• **Die**

• **Privacy**

• **Fensterfu**

• **abfragen.**

• **nktion**

• **LAST_VALU**

• **E**

• **Die**

• **Fensterfu**

• **nktion**

• **LEAD**

• **MAX-**

• **Fenst**

• **erfunktio**

• **nen**

• **Funktion**

• **n**

• **des**

• **Aggregation**

• **SELECT-**

• **Tabellen**

• **Ansichten**

• **Analytics-**

• **ETEs**

• **Engine**

• **unterstüt**

• **zte**

• **SQL-**

• **Konst**

• **rukte**

• **MEDIAN-**

• **Fe**

• **nsters**

• **Funktion**

• **n**

• **im**

• **MIN-**

• **Fenst**

• **er**

• **Die**

• **Fensterfu**

• **nktion**

• **NTH_VALUE**

• **Fensterfu**

• **nktionen**

• **STDDEV_SA**

• **MP**

• **und**

• **STDDEV_PO**

• **P**

• **(STDDEV_S**

• **AMP**

• **und**

Mathematische

SELECT-

Tabellen

Ansichten

Analytics-

ETEs

Engine

unterstüt

zte

SQL-

Konst

rukte

STDDEV

sind

Synonyme)

- SUM-

Fenster

erfunktio

nen

- Fensterfu

nktionen

VAR_SAMP

und

VAR_POP

(VAR_SAMP

und

VARIANCE

sind

Synonyme)

Rangfestl

egungsfun

ktionen

Aggregation

SELECT-

Tabellena

Ansichten

Analytics-

ETEs

gine

unterstüt

zte

SQL-

Konst

rukte

- Die
Fensterfu
nktion
DENSE_RAN
K
- Die
Fensterfu
nktion
NTILE
- Die
Fensterfu
nktion
PERCENT_R
ANK
- Die
Fensterfu
nktion
RANK
- Die
Fensterfu
nktion
ROW_NUMBE
R

Unterstützung

• SELECT-

Tabellen

Ansichten

Analytics-

DETs

Engine

unterstüt

zte

SQL-

Konstrukte

unterstüt

• ANALYSE-

Ausdrücke

unterstüt

zdruck

- COALESCE-
Ausdruck

- Funktionen
n
GREATEST
und
LEAST

- NVL-
und
COALESCE-
Funktionen
n

- NVL2
Funktion

- NULLIF-
Fu
nktion

Werte in

SELECT-

Tabellen

Ansicht

lytics-

ETEs

gine

unterstüt

zte

SQL-

Konst

rukte

Werte in

Tabellen

körpers

nicht

Logische

verwendet

Bedingung

werden,

da

• Patternma

sie

tching-

Unterabfr

Be

agen

dingungen

erfordern

• Bedingung

.

en

Alle

zwischen

anderen

den

werden

Reichweit

unterstüt

en

zt.

• „Null“-

Be

dingung

Werte in

SELECT-

Tabellena

Ansicht

lytics-

ETEs

gine

unterstüt

zte

SQL-

Konst

rukte

Attributions-

werden

unterstüt

Datumen

unich

Uhrzeit

onen

- Verkettun
gsoperato
r

- ADD_MONTH
S-
Funktio
nen

- Funktion
CONVERT_T
IMEZONE

- Funktion
CURRENT_D
ATE

- Funktion
DATEADD

Mitgliedern

SELECT-

Tabellen

Ansicht

lytics-

DEs

gine

unterstüt

zte

SQL-

Konst

rukte

- Funktion
DATEDIFF
- DATE_PART
-
Funktion
en
- Funktion
DATE_TRUNC
- Funktion
EXTRACT
- Funktion
TO_TIMESTAMP
- Datumstei
le
für
Datums-
oder
Zeitstemp
elfunktio
nen

Mittelwert

SELECT-

Tabellen

Ansicht

lytics-

ETEs

gine

unterstüt

zte

SQL-

Konst

rukte

Zellenfo

ly (Vektor)

unterstüt

zOperator

- Die Funktion BTRIM
- Die Funktion CHAR_LENGTH
- Die Funktion CHARACTER_LENGTH
- Funktion CONCAT
- Die Funktion LEFT

Mitgliedern

SELECT-

Tabellen

Ansichten

Analytics-

ETEs

gine

unterstüt

zte

SQL-

Konstr

rukte

und

RIGHT

- Die Funktion LEN
- Die Funktion LENGTH
- Die Funktion LOWER
- Die Funktionen LPAD und RPAD
- Die Funktion LTRIM
- POSITION-Funktionen

Multi-tenancy

SELECT-

Tabellen

Ansichten

Analytics-

ETEs

Engine

unterstüt

zte

SQL-

Konst

rukte

- Die
Funktion
REGEXP_CO
UNT
- Die
Funktion
REGEXP_IN
STR
- Die
Funktion
REGEXP_RE
PLACE
- Die
Funktion
REGEXP_SU
BSTR
- Die
Funktion
REPEAT
- Die
Funktion
REPLACE

Werte in

SELECT-

Tabellen

Ansicht

lytics-

ETEs

gine

unterstüt

zte

SQL-

Konst

rukte

- Die
Funktion
REVERSE
- Die
Funktion
RTRIM
- Die
Funktion
SPLIT_PAR
T
- Die
Funktion
SUBSTRING
- Die
Funktion
TRANSLATE
- TRIM-
Funk
tionen
- Die
Funktion
UPPER

Abbildung

SELECT-

Tabellen

Ansicht

lytics-

ETEs

gine

unterstüt

zte

SQL-

Konst

rukte

ALTER

werden

unterstüt

TO_CHAR

Datentyp

• **TO_DATE-**

ormatieru

ng

• **TO_NUMBER**

• **Datum-/**

Uh

rzeit-

For

matzeiche

nfolgen

• **Numerisch**

e

Formatzei

chenfolge

n

Mitgliedern

SELECT-

Tabellen

Ansichten

Metrics-

ETEs

gine

unterstüt

zte

SQL-

Konst

rukte

AES_ENCRY

Funktion

AES_ENTSC

HLÜSSELN

- ENCODE

- DECODE

- MD5

Funktion

- SHA1

Funktion

- SHA2

Funktion

- XX_

HASH64

Symbole

würden

unterstüt

the

Operatore

@

~~Alle Funktionen~~

~~SELECT-~~

~~Tabellen~~

~~Ansichten~~

~~Analytics-~~

~~ETEs~~

~~Engine~~

~~unterstüt~~

~~zte~~

~~SQL-~~

~~Konst~~

~~rukte~~

~~Alle Funktionen~~

~~werden~~

~~Funktion~~

~~Die~~

~~ACOS~~

- Die Funktion ASIN

- Die Funktion ATAN

- ATAN2 Funktion

- Die Funktion CBRT

- Die Funktion CEILING (oder CEIL)

Mitgliedern

SELECT-

Tabellen

Ansichten

Analytics-

ETEs

gine

unterstüt

zte

SQL-

Konst

rukte

- Die
Funktion
COS
- Die
Funktion
COT
- Die
Funktion
DEGREES
- Die
Funktion
LTRIM
- Die
Funktion
EXP
- Die
Funktion
FLOOR
- Die
Funktion
LN

Mittelwert

SELECT-

Tabellen

Ansicht

Analytics-

ETEs

gine

unterstüt

zte

SQL-

Konst

rukte

- Die Funktion LOG
- Die Funktion MOD
- Die Funktion PI
- Die Funktion POWER
- Die Funktion RADIANS
- Die Funktion RANDOM
- Die Funktion ROUND

Mitgliedern

SELECT-

Tabellen

Ansichten

Analytics-

ETEs

gine

unterstüt

zte

SQL-

Konst

rukte

- Die
Funktion
SIGN
- Die
Funktion
SIN
- SQRT-
Funk
tionen
- Die
Funktion
TRUNC

Aggregieren

SELECT-

Tabellena

Ansicht

Analytics-

ETEs

gine

unterstüt

zte

SQL-

Konst

rukte

MIN BY EX-

werden

UNBASE64

umtänstlich

• **SECHSKANT**

zt

• **HLL_SKETC**

H_AGG,

• **HLL_SKETC**

H_ESTIMAT

E

• **HLL_UNION**

• **HLL_UNION**

_AGG

AND JSON

werden

• **GET_JSON_**

unterstüt

OBJECT

zt

~~Multiplan~~

~~SELECT-~~

~~Tabellena~~

~~Ansicht~~

~~lytics-~~

~~ETEs~~

~~gine~~

~~unterstüt~~

~~zte~~

~~SQL-~~

~~Konst~~

~~rukte~~

~~ARRAY_CON~~

~~FUNCTION~~

~~Unterstüt~~

~~zungen~~

~~ARRAY_DIS~~

~~TINCT~~

~~• ARRAY_EXC~~

~~EPT~~

~~• ARRAY_INT~~

~~ERSECT~~

~~• ARRAY_JOI~~

~~N~~

~~• ARRAY_ENT~~

~~FERNEN~~

~~• ARRAY_SOR~~

~~T~~

~~• ARRAY_UNI~~

~~ON~~

~~NUMBERU~~

~~NUMERU~~

~~NUMERU~~

~~NUMERU~~

~~NUMERU~~

Wird ein

SELECT-

Tabellena

Ansicht

lytics-

ETEs

gine

unterstüt

zte

SQL-

Konst

rukte

Wird ein

Filteren

der

Bedingung

unterstüt

zt,

dass

eine

ORDER

BY-

Klause

I

nur

in

der

Partition

sklausel

einer

Fensterfu

nktion

unterstüt

zt

wird,

Wird ein
SELECT-
Tabella
Ansicht
Metrics-
DTEs
gine
unterstüt
zte
SQL-
Konst
rukte
wenn
Tabellen
mit
aktiviert
em
Differential
ial
Privacy
abgefragt
werden.

Aggregation

SELECT-

Tabellena

Ansicht

Metrics-

CTEs

gine

unterstüt

zte

SQL-

Konst

rukte

WITH beg

OFFSET

unterstüt

CTEs

Verwendun

g

von

differenz

uell

datenschu

tzgeschüt

zten

Tabellen

nicht

unterstüt

zt

Unterstüt

zt

Tabellen

und

Spalten

Kategorie	SQL-Konstrukt	Alternative
Symbole für mathematische Operatoren	• \$column 2 • \$Spalte / 2 • \$Spalte ^ 2	• CBRT • SQRT • MACHT (\$Spalte, 2)
Skalarfunktionen	• SYSDATE • \$column: :Ganzzahl • konvertieren (Typ, \$Spalte)	• CURRENT_DATE • CAST \$column ALS Ganzzahl • CAST \$column AS-Typ
Literale	INTERVALL '1 SEKUNDE'	INTERVALL '1' SEKUNDE
Zeilenbegrenzung	TOP n	GRENZE n
Join	• USING • NATURAL	Die ON-Klausel sollte explizit ein Join-Kriterium enthalten.

Tipps und Beispiele für Differential Privacy-Abfragen

AWS Clean Rooms Differential Privacy verwendet eine [allgemeine Abfragestruktur](#), um eine Vielzahl von SQL-Konstrukten wie Common Table Expressions (CTEs) für die Datenaufbereitung und häufig verwendete Aggregatfunktionen wie, oder zu unterstützen. COUNT SUM Um den Beitrag jedes möglichen Benutzers in Ihren Daten zu verschleiern, indem den aggregierten Abfrageergebnissen zur Laufzeit Rauschen hinzugefügt wird, erfordert AWS Clean Rooms Differential Privacy, dass Aggregatfunktionen in der Endversion auf Daten auf Benutzerebene ausgeführt werden. SELECT statement

Im folgenden Beispiel werden zwei Tabellen mit dem Namen `socialco_impressions` und `socialco_users` von einem Medienverlag verwendet, der Daten mithilfe von Differential Privacy schützen möchte, während er mit einer Sportmarke mit Daten zusammenarbeitet. `athletic_brand_sales` Der Medienherausgeber hat die `user_id` Spalte als Benutzer-ID-Spalte konfiguriert und gleichzeitig den differenziellen Datenschutz in AWS Clean Rooms aktiviert. Der Werbetreibende benötigt keinen differenzierten Datenschutz und möchte eine Abfrage CTEs anhand kombinierter Daten ausführen. Da sein CTE differenzielle datenschutzgeschützte Tabellen verwendet, nimmt der Werbetreibende die Benutzer-ID-Spalte aus diesen geschützten Tabellen in die Liste der CTE-Spalten auf und fügt die geschützten Tabellen in der Benutzer-ID-Spalte zusammen.

```

WITH matches_table AS(
    SELECT si.user_id, si.campaign_id, s.sale_id, s.sale_price
    FROM socialco_impressions si
    JOIN socialco_users su
        ON su.user_id = si.user_id
    JOIN athletic_brand_sales s
        ON s.emailsha256 = su.emailsha256
    WHERE s.timestamp > si.timestamp

UNION ALL

    SELECT si.user_id, si.campaign_id, s.sale_id, s.sale_price
    FROM socialco_impressions si
    JOIN socialco_users su
        ON su.user_id = si.user_id
    JOIN athletic_brand_sales s
        ON s.phonesha256 = su.phonesha256
    WHERE s.timestamp > si.timestamp
)

SELECT COUNT (DISTINCT user_id) as unique_users
FROM matches_table
GROUP BY campaign_id
ORDER BY COUNT (DISTINCT user_id) DESC
LIMIT 5

```

Ebenso müssen Sie, wenn Sie Fensterfunktionen für Tabellen mit differenziertem Datenschutz ausführen möchten, die Spalte mit der Benutzerkennung in die Klausel aufnehmen. `PARTITION BY`

```

ROW_NUMBER() OVER (PARTITION BY conversion_id, user_id ORDER BY match_type, match_age)
AS row

```

Einschränkungen von AWS Clean Rooms Differential Privacy

AWS Clean Rooms Differential Privacy befasst sich nicht mit den folgenden Situationen:

1. AWS Clean Rooms Differential Privacy unterstützt nur Abfragen mit Amazon AWS Glue S3-gestützten Tabellen. Abfragen mit Snowflake- oder Amazon Athena Athena-Tabellen werden nicht unterstützt.
2. AWS Clean Rooms Differential Privacy befasst sich nicht mit Timing-Angriffen. Diese Angriffe sind beispielsweise in Szenarien möglich, in denen ein einzelner Benutzer eine große Anzahl von

Zeilen beisteuert und das Hinzufügen oder Entfernen dieses Benutzers die Berechnungszeit für Abfragen erheblich verändert.

3. AWS Clean Rooms Differential Privacy garantiert keinen differenziellen Datenschutz, wenn eine SQL-Abfrage aufgrund der Verwendung bestimmter SQL-Konstrukte zur Laufzeit zu einem Überlauf oder zu ungültigen Umwandlungsfehlern führen kann.

Die folgende Tabelle enthält eine Liste einiger, aber nicht aller SQL-Konstrukte, die zu Laufzeitfehlern führen können und die in Analysevorlagen überprüft werden sollten. Es wird empfohlen, Analysevorlagen zu genehmigen, die die Wahrscheinlichkeit solcher Laufzeitfehler minimieren, und die Abfrageprotokolle regelmäßig zu überprüfen, um festzustellen, ob die Abfragen mit der Kooperationsvereinbarung übereinstimmen.

Die folgenden SQL-Konstrukte sind anfällig für Überlauffehler:

Kategorie	Für Überlauffehler anfällige SQL-Konstrukte in der Spark-SQL-Analyse-Engine
Aggregationsfunktionen	• AVG • SUM/SUM_DISTINCT
Funktionen für die Datentypformatierung	• TO_TIMESTAMP • TO_DATE
Datums- und Zeitfunktionen	• ADD_MONTHS • DATEADD • DATEDIFF
Mathematische Funktionen	• +, -, *, / • POWER
Zeichenfolgenfunktionen	• • CONCAT • REPEAT
Fensterfunktionen	• AVG • SUM

4. Die Formatierungsfunktion für CAST-Datentypen ist anfällig für ungültige Umwandlungsfehler.

Sie können so konfigurieren [CloudWatch](#) , dass ein [Metrikfilter für eine Protokollgruppe](#) und anschließend [ein CloudWatch Alarm für diesen Metrikfilter erstellt](#) wird, um Benachrichtigungen zu erhalten, wenn ein potenzieller Überlauf- oder Umwandlungsfehler aufgetreten ist.

Insbesondere sollten Sie auf die Fehlercodes `CastError`, `OverflowError`, `ConversionError` achten. Das Vorhandensein dieser Fehlercodes deutet auf einen möglichen Seitenkanalangriff hin, kann aber auch auf eine fehlerhafte SQL-Abfrage hinweisen.

Weitere Informationen finden Sie unter [Analyse Einloggen AWS Clean Rooms](#).

AWS-Reinräume ML

Mit AWS Clean Rooms ML können zwei oder mehr Parteien Modelle für maschinelles Lernen auf ihren Daten ausführen, ohne ihre Daten miteinander teilen zu müssen. Der Service bietet Kontrollen zur Verbesserung des Datenschutzes, mit denen Dateneigentümer ihre Daten und ihre Modell-IP schützen können. Sie können von Hand AWS erstellte Modelle verwenden oder Ihr eigenes benutzerdefiniertes Modell mitbringen.

Eine detailliertere Erklärung, wie das funktioniert, finden Sie unter. [Kontoübergreifende Jobs](#)

Weitere Informationen zu den Funktionen von Clean Rooms ML-Modellen finden Sie in den folgenden Themen.

Topics

- [ML-Terminologie von AWS Clean Rooms](#)
- [So funktioniert AWS Clean Rooms ML mit AWS Modellen](#)
- [So funktioniert AWS Clean Rooms ML mit benutzerdefinierten Modellen](#)
- [AWS Modelle in Reinräumen \(ML\)](#)
- [Benutzerdefinierte Modelle in Clean Rooms ML](#)

ML-Terminologie von AWS Clean Rooms

Bei der Verwendung von Clean Rooms ML ist es wichtig, die folgende Terminologie zu verstehen:

- **Anbieter von Trainingsdaten** — Die Partei, die die Trainingsdaten bereitstellt, ein Lookalike-Modell erstellt und konfiguriert und dieses Lookalike-Modell dann einer Zusammenarbeit zuordnet.
- **Seed-Datenanbieter** — Die Partei, die die Ausgangsdaten bereitstellt, generiert ein Lookalike-Segment und exportiert ihr Lookalike-Segment.
- **Trainingsdaten** — Die Daten des Trainingsdatenanbieters, die zur Generierung eines Lookalike-Modells verwendet werden. Die Trainingsdaten werden verwendet, um die Ähnlichkeit des Benutzerverhaltens zu messen.

Die Trainingsdaten müssen eine Benutzer-ID, eine Element-ID und eine Zeitstempelspalte enthalten. Optional können die Trainingsdaten auch andere Interaktionen als numerische oder kategoriale Merkmale enthalten. Beispiele für Interaktionen sind eine Liste von angesehenen Videos, gekauften Artikeln oder gelesenen Artikeln.

- **Seed-Daten** — Die Daten des Seed-Datenanbieters, die zur Erstellung eines Lookalike-Segments verwendet werden. Die Ausgangsdaten können direkt bereitgestellt werden oder aus den Ergebnissen einer AWS Clean Rooms Abfrage stammen. Bei der Ausgabe des Lookalike-Segments handelt es sich um eine Gruppe von Benutzern aus den Trainingsdaten, die den Ausgangsbenutzern am ähnlichsten sind.
- **Lookalike-Modell** — Ein maschinelles Lernmodell der Trainingsdaten, das verwendet wird, um ähnliche Benutzer in anderen Datensätzen zu finden.

Bei der Verwendung der API wird der Begriff Zielgruppenmodell gleichwertig mit dem Lookalike-Modell verwendet. Beispielsweise verwenden Sie die [CreateAudienceModel](#)API, um ein Lookalike-Modell zu erstellen.

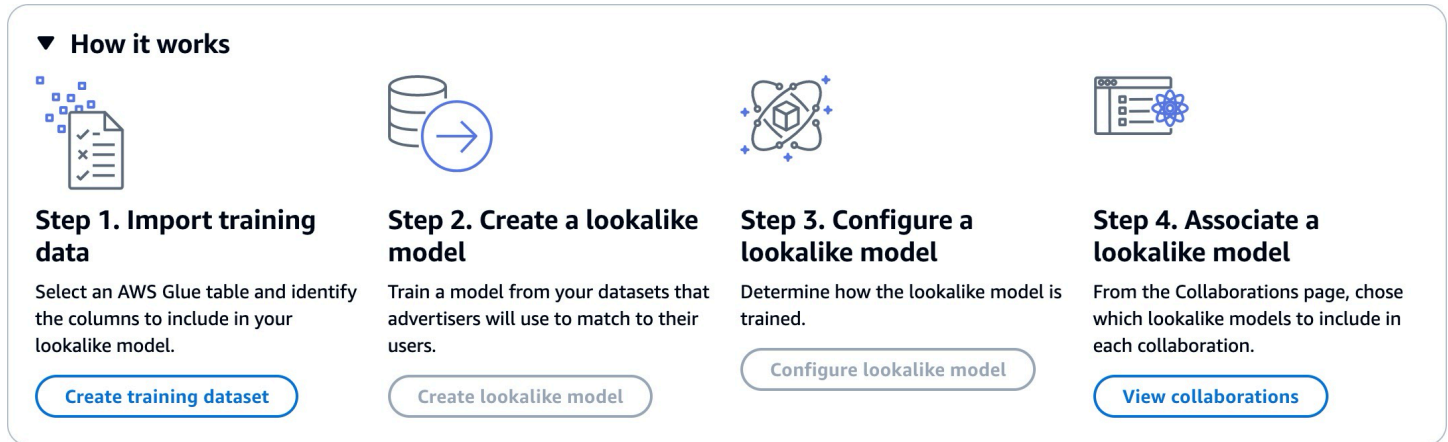
- **Lookalike-Segment** — Eine Teilmenge der Trainingsdaten, die den Ausgangsdaten am ähnlichsten ist.

Wenn Sie die API verwenden, erstellen Sie mit der API ein Lookalike-Segment.

[StartAudienceGenerationJob](#)

Die Daten des Trainingsdatenanbieters werden niemals mit dem Startdatenanbieter geteilt, und die Daten des Ausgangsdatenanbieters werden niemals mit dem Trainingsdatenanbieter geteilt. Die Ausgabe des Lookalike-Segments wird mit dem Trainingsdatenanbieter geteilt, aber niemals mit dem Seed-Datenanbieter.

So funktioniert AWS Clean Rooms ML mit AWS Modellen



Die Arbeit mit Lookalike-Modellen erfordert, dass zwei Parteien, ein Anbieter von Trainingsdaten und ein Anbieter von Startdaten, nacheinander zusammenarbeiten, AWS Clean Rooms um ihre Daten in eine Zusammenarbeit einzubringen. Dies ist der Workflow, den der Trainingsdatenanbieter zuerst abschließen muss:

1. Die Daten des Trainingsdatenanbieters müssen in einer AWS Glue Datenkatalogtabelle mit Interaktionen zwischen Benutzern und Elementen gespeichert werden. Die Trainingsdaten müssen mindestens eine Benutzer-ID-Spalte, eine Interaktions-ID-Spalte und eine Zeitstempelspalte enthalten.
2. Der Trainingsdatenanbieter registriert die Trainingsdaten bei AWS Clean Rooms.
3. Der Trainingsdatenanbieter erstellt ein Lookalike-Modell, das mit mehreren Startdatenanbietern gemeinsam genutzt werden kann. Das Lookalike-Modell ist ein tiefes neuronales Netzwerk, dessen Training bis zu 24 Stunden dauern kann. Es wird nicht automatisch neu trainiert und wir empfehlen, dass Sie das Modell wöchentlich neu trainieren.
4. Der Anbieter von Trainingsdaten konfiguriert das Lookalike-Modell, einschließlich der Frage, ob Relevanzkennzahlen und der Amazon S3 S3-Speicherort der Ausgabesegmente geteilt werden sollen. Der Anbieter von Trainingsdaten kann mehrere konfigurierte Lookalike-Modelle aus einem einzigen Lookalike-Modell erstellen.
5. Der Anbieter von Trainingsdaten ordnet das konfigurierte Zielgruppenmodell einer Zusammenarbeit zu, die mit einem Startdatenanbieter geteilt wird.

Dies ist der Workflow, den der Seed-Datenanbieter als Nächstes abschließen muss:

1. Die Daten des Seed-Datenanbieters können in einem Amazon S3 S3-Bucket gespeichert werden oder aus den Ergebnissen einer Abfrage stammen.
2. Der Seed-Datenanbieter eröffnet die Zusammenarbeit, die er mit dem Trainingsdatenanbieter teilt.
3. Der Seed-Datenanbieter erstellt auf der Registerkarte Clean Rooms ML der Kollaborationsseite ein ähnliches Segment.
4. Der Seed-Datenanbieter kann die Relevanzkennzahlen auswerten, sofern sie geteilt wurden, und das Lookalike-Segment zur externen Verwendung exportieren. AWS Clean Rooms

So funktioniert AWS Clean Rooms ML mit benutzerdefinierten Modellen

Mit Clean Rooms ML können Mitglieder einer Kollaboration einen angedockten benutzerdefinierten Modellalgorithmus verwenden, der in Amazon ECR gespeichert ist, um ihre Daten gemeinsam zu analysieren. Dazu muss der Modellanbieter ein Bild erstellen und es in Amazon ECR speichern. Folgen Sie den Schritten im [Amazon Elastic Container Registry User Guide](#), um ein privates Repository zu erstellen, das das benutzerdefinierte ML-Modell enthalten wird.

Jedes Mitglied einer Kollaboration kann der Modellanbieter sein, vorausgesetzt, es verfügt über die richtigen Berechtigungen. Alle Mitglieder einer Kollaboration können Trainingsdaten, Inferenzdaten oder beides zum Modell beitragen. Für die Zwecke dieses Leitfadens werden Mitglieder, die Daten beisteuern, als Datenanbieter bezeichnet. Das Mitglied, das die Kollaboration erstellt, ist der Ersteller der Kollaboration, und dieses Mitglied kann entweder der Modellanbieter, einer der Datenanbieter oder beides sein.

Auf der höchsten Ebene sind die folgenden Schritte aufgeführt, die ausgeführt werden müssen, um eine benutzerdefinierte ML-Modellierung durchzuführen:

1. Der Kollaborationsersteller erstellt eine Kollaboration und weist jedem Mitglied die richtigen Mitgliederfähigkeiten und Zahlungskonfigurationen zu. Der Kollaborationsersteller muss in diesem Schritt dem jeweiligen Mitglied die Fähigkeit zuweisen, entweder Modellausgaben zu empfangen oder Inferenzergebnisse zu empfangen, da diese Fähigkeit nach der Erstellung der Kollaboration nicht mehr aktualisiert werden kann. Weitere Informationen finden Sie unter [Aufbau und Beitritt zur Zusammenarbeit in AWS Clean Rooms ML](#).
2. Der Modellanbieter konfiguriert und ordnet sein containerisiertes ML-Modell der Kollaboration zu und stellt sicher, dass Datenschutzbeschränkungen für exportierte Daten festgelegt werden. Weitere Informationen finden Sie unter [Konfiguration eines Modellalgorithmus in AWS Clean Rooms ML](#).

3. Die Datenanbieter tragen ihre Daten zur Zusammenarbeit bei und stellen sicher, dass ihre Datenschutzerfordernisse spezifiziert werden. Datenanbieter müssen dem Modell den Zugriff auf ihre Daten ermöglichen. Weitere Informationen erhalten Sie unter [Bereitstellung von Trainingsdaten in AWS Clean Rooms ML](#) und [Zuordnen des konfigurierten Modellalgorithmus in AWS Clean Rooms ML](#).
4. Ein Mitglied der Kollaboration erstellt die ML-Konfiguration, die definiert, wohin die Modellartefakte oder Inferenzergebnisse exportiert werden.
5. Ein Kollaborationsmitglied erstellt einen ML-Eingabekanal, der Eingaben für den Trainings- oder Inferenzcontainer bereitstellt. Der ML-Eingabekanal ist eine Abfrage, die die Daten definiert, die im Kontext des Modellalgorithmus verwendet werden sollen.
6. Ein Kollaborationsmitglied ruft das Modelltraining mithilfe des ML-Eingabekanals und des konfigurierten Modellalgorithmus auf. Weitere Informationen finden Sie unter [Erstellen eines trainierten Modells in AWS Clean Rooms ML](#).
7. (Optional) Der Modeltrainer ruft den Modellexportjob auf und die Modellartefakte werden an den Empfänger der Modellergebnisse gesendet. Nur Mitglieder mit einer gültigen ML-Konfiguration und der Fähigkeit, Modellausgaben zu empfangen, können Modellartefakte empfangen. Weitere Informationen finden Sie unter [Exportieren von Modellartefakten aus AWS Clean Rooms ML](#).
8. (Optional) Ein Kollaborationsmitglied ruft die Modellinferenz mithilfe des ML-Eingangskanals, des trainierten Modell-ARN und des mit Inferenz konfigurierten Modellalgorithmus auf. Die Inferenzergebnisse werden an den Empfänger für die Inferenzausgabe gesendet. Nur Mitglieder mit einer gültigen ML-Konfiguration und der Fähigkeit eines Mitglieds, Inferenzausgaben zu empfangen, können Inferenzergebnisse empfangen.

Hier sind die Schritte, die vom Modellanbieter ausgeführt werden müssen:

1. Erstellen Sie ein SageMaker KI-kompatibles Amazon ECR-Docker-Image. Clean Rooms ML unterstützt nur SageMaker KI-kompatible Docker-Images.
2. Nachdem Sie ein SageMaker KI-kompatibles Docker-Image erstellt haben, übertragen Sie das Image an Amazon ECR. Folgen Sie den Anweisungen im [Amazon Elastic Container Registry User Guide](#), um ein Container-Training-Image zu erstellen.
3. Konfigurieren Sie den Modellalgorithmus für die Verwendung in Clean Rooms ML.
 - a. Geben Sie den Amazon ECR-Repository-Link und alle Argumente an, die für die Konfiguration des Modellalgorithmus erforderlich sind.
 - b. Stellen Sie eine Servicezugriffsrolle bereit, die Clean Rooms ML den Zugriff auf das Amazon ECR-Repository ermöglicht.

- c. Ordnen Sie den konfigurierten Modellalgorithmus der Kollaboration zu. Dazu gehört die Bereitstellung einer Datenschutzrichtlinie, die Kontrollen für Container-Logs, Fehlerprotokolle und CloudWatch Metriken sowie Beschränkungen dafür definiert, wie viele Daten aus den Container-Ergebnissen exportiert werden können.

Die folgenden Schritte müssen vom Datenanbieter ausgeführt werden, um mit einem benutzerdefinierten ML-Modell zusammenzuarbeiten:

1. Konfigurieren Sie eine vorhandene AWS Glue Tabelle mit einer benutzerdefinierten Analyseregeln. Auf diese Weise können bestimmte vorab genehmigte Abfragen oder vorab genehmigte Konten Ihre Daten verwenden.
2. Ordnen Sie Ihre konfigurierte Tabelle einer Kollaboration zu und stellen Sie eine Dienstzugriffsrolle bereit, die auf Ihre AWS Glue Tabellen zugreifen kann.
3. [Fügen Sie der Tabelle eine Regel zur Kollaborationsanalyse](#) hinzu, die es der konfigurierten Modellalgorithmus-Assoziation ermöglicht, auf die konfigurierte Tabelle zuzugreifen.
4. Nachdem das Modell und die Daten in Clean Rooms ML verknüpft und konfiguriert wurden, stellt das Mitglied, das Abfragen ausführen kann, eine SQL-Abfrage bereit und wählt den zu verwendenden Modellalgorithmus aus.

Nach Abschluss des Modelltrainings initiiert dieses Mitglied den Export von Modelltrainingsartefakten oder Inferenzergebnissen. Diese Artefakte oder Ergebnisse werden an das Mitglied gesendet, das die Ausgabe des trainierten Modells empfangen kann. Der Empfänger der Ergebnisse muss sie konfigurieren, `MachineLearningConfiguration` bevor er die Modellausgabe empfangen kann.

AWS Modelle in Reinräumen (ML)

AWS Clean Rooms ML bietet eine Methode zur Wahrung der Privatsphäre, mit der zwei Parteien ähnliche Benutzer in ihren Daten identifizieren können, ohne ihre Daten miteinander teilen zu müssen. Die erste Partei stellt die Trainingsdaten zur Verfügung, AWS Clean Rooms sodass sie ein Lookalike-Modell erstellen und konfigurieren und es mit einer Zusammenarbeit verknüpfen kann. Anschließend werden Ausgangsdaten in die Zusammenarbeit übernommen, um ein ähnliches Segment zu erstellen, das den Trainingsdaten ähnelt.

Eine detailliertere Erklärung, wie das funktioniert, finden Sie unter [Kontoübergreifende Jobs](#).

Die folgenden Themen enthalten Informationen zum Erstellen und Konfigurieren von AWS Modellen in Clean Rooms ML.

Topics

- [Datenschutz von AWS Clean Rooms ML](#)
- [Anforderungen an Schulungsdaten für Clean Rooms ML](#)
- [Anforderungen an Saatgutdaten für Clean Rooms ML](#)
- [Bewertungsmetriken für das ML-Modell von AWS Clean Rooms](#)

Datenschutz von AWS Clean Rooms ML

Clean Rooms ML wurde entwickelt, um das Risiko von Inferenzangriffen auf Mitglieder zu verringern, bei denen der Anbieter der Trainingsdaten herausfinden kann, wer in den Startdaten enthalten ist, und der Anbieter der Startdaten kann herausfinden, wer in den Trainingsdaten enthalten ist. Es wurden mehrere Schritte unternommen, um diesen Angriff zu verhindern.

Erstens beobachten Anbieter von Saatgutdaten die Ergebnisse von Clean Rooms ML nicht direkt, und Anbieter von Trainingsdaten können die Saatgutdaten niemals beobachten. Anbieter von Startdaten können sich dafür entscheiden, die Ausgangsdaten in das Output-Segment aufzunehmen.

Als Nächstes wird das Lookalike-Modell aus einer Zufallsstichprobe der Trainingsdaten erstellt. Diese Stichprobe umfasst eine beträchtliche Anzahl von Benutzern, die nicht der Stammzielgruppe entsprechen. Durch diesen Prozess ist es schwieriger festzustellen, ob ein Benutzer nicht in den Daten enthalten war. Dies ist eine weitere Möglichkeit, Rückschlüsse auf die Mitgliedschaft zu ziehen.

Außerdem können mehrere Startkunden für jeden Parameter des samenspezifischen Lookalike-Modell-Trainings verwendet werden. Dadurch wird begrenzt, wie stark das Modell übermäßig angepasst werden kann und wie viel Rückschlüsse auf einen Benutzer gezogen werden können. Daher empfehlen wir, dass die Mindestgröße der Ausgangsdaten 500 Benutzer beträgt.

Schließlich werden den Anbietern von Trainingsdaten niemals Kennzahlen auf Benutzerebene zur Verfügung gestellt, wodurch eine weitere Möglichkeit für einen Angriff auf Mitgliedschaftsabschlüsse ausgeschlossen wird.

Anforderungen an Schulungsdaten für Clean Rooms ML

Um erfolgreich ein Lookalike-Modell zu erstellen, müssen Ihre Trainingsdaten die folgenden Anforderungen erfüllen:

- Die Trainingsdaten müssen im Parquet-, CSV- oder JSON-Format vorliegen.

Note

Zstandard (ZSTD) komprimierte Parquet-Daten werden nicht unterstützt.

- Ihre Trainingsdaten müssen katalogisiert sein. AWS Glue Weitere Informationen finden Sie unter [Erste Schritte mit dem AWS Glue Glue-Datenkatalog](#) im AWS Glue Entwicklerhandbuch. Wir empfehlen die Verwendung von AWS Glue Crawlern zur Erstellung Ihrer Tabellen, da das Schema automatisch abgeleitet wird.
- Der Amazon S3 S3-Bucket, der die Trainings- und Startdaten enthält, befindet sich in derselben AWS Region wie Ihre anderen Clean Rooms ML-Ressourcen.
- Die Trainingsdaten müssen mindestens 100.000 eindeutige Benutzer IDs mit jeweils mindestens zwei Artikelinteraktionen enthalten.
- Die Trainingsdaten müssen mindestens 1 Million Datensätze enthalten.
- Das in der [CreateTrainingDataset](#)Aktion angegebene Schema muss mit dem Schema übereinstimmen, das bei der Erstellung der AWS Glue Tabelle definiert wurde.
- Die erforderlichen Felder, wie sie in der bereitgestellten Tabelle definiert sind, sind in der [CreateTrainingDataset](#)Aktion definiert.

Feldtyp	Unterstützte Datentypen	Erforderlich	Description
USER_ID	Zeichenfolge, Ganzzahl, Ganzzahl	Ja	Eine eindeutige Kennung für jeden Benutzer im Datensatz. Es sollte sich um einen Wert für nicht persönlich identifizierbare

Feldtyp	Unterstützte Datentypen	Erforderlich	Description
			Informationen (PII) handeln. Dabei kann es sich um eine Hash-ID oder eine Kunden-ID handeln.
ITEM_ID	Zeichenfolge, Ganzzahl, Ganzzahl	Ja	Eine eindeutige Kennung für jedes Objekt, mit dem ein Benutzer interagiert.

Feldtyp	Unterstützte Datentypen	Erforderlich	Description
TIMESTAMP (ZEITSTEMPEL)	bigint, int, timestamp	Ja	Die Zeit, zu der ein Benutzer mit dem Objekt interagiert hat. Die Werte müssen im Format Unix-Epochenzeit in Sekunden angegeben werden.

Feldtyp	Unterstützte Datentypen	Erforderlich	Description
KATEGORISCHES MERKMAL	string, int, float, bigint, double, boolean, array	Nein	Erfasst kategorische Daten, die sich auf den Benutzer oder das Objekt beziehen. Dies kann Dinge wie Ereignistyp (wie Klick oder Kauf), Nutzerdemografie (Altersgruppe, Geschlecht — anonymisiert), Nutzerstandort (Stadt, Land — anonymisiert), Artikelkategorie (z. B. Kleidung oder

Feldtyp	Unterstützte Datentypen	Erforderlich	Description
			Elektronik) oder Artikelmarke beinhalten.

Feldtyp	Unterstützte Datentypen	Erforderlich	Description
NUMERISCHES_MERKMAL	doppelt, Gleitkomm a, Ganzzahl, große Ganzzahl	Nein	Erfasst numerische Daten, die sich auf den Benutzer oder das Objekt beziehen. Dazu können Dinge wie die Kaufhistorie von Nutzern (Gesamtbetrag der Ausgaben), der Artikelpreis, die Anzahl der Besuche eines Artikels oder Benutzerbewertungen für Artikel gehören.

- Optional können Sie insgesamt bis zu 10 kategoriale oder numerische Merkmale angeben.

Hier ist ein Beispiel für einen gültigen Trainingsdatensatz im CSV-Format

```
USER_ID,ITEM_ID,TIMESTAMP,EVENT_TYPE(CATEGORICAL FEATURE),EVENT_VALUE (NUMERICAL FEATURE)
196,242,881250949,click,15
186,302,891717742,click,13
22,377,878887116,click,10
244,51,880606923,click,20
166,346,886397596,click,10
```

Anforderungen an Saatgutdaten für Clean Rooms ML

Die Ausgangsdaten für ein Lookalike-Modell können entweder direkt aus einem Amazon S3 S3-Bucket oder aus den Ergebnissen einer SQL-Abfrage stammen.

Direkt bereitgestellte Ausgangsdaten müssen die folgenden Anforderungen erfüllen:

- Die Ausgangsdaten müssen im JSON-Zeilenformat mit einer Benutzerliste vorliegen IDs.
- Die Ausgangsgröße sollte zwischen 25 und 500.000 einzelnen Benutzern IDs liegen.
- Die Mindestanzahl von Startbenutzern muss dem Mindestwert für die passende Ausgangsgröße entsprechen, der bei der Erstellung des konfigurierten Zielgruppenmodells angegeben wurde.

Im Folgenden finden Sie ein Beispiel für einen gültigen Trainingsdatensatz im CSV-Format

```
{"user_id": "abc"}
{"user_id": "def"}
{"user_id": "ghijkl"}
{"user_id": "123"}
{"user_id": "456"}
{"user_id": "7890"}
```

Bewertungsmetriken für das ML-Modell von AWS Clean Rooms

Clean Rooms ML berechnet den Erinnerungs - und den Relevanzwert, um festzustellen, wie gut Ihr Modell abschneidet. Recall vergleicht die Ähnlichkeit zwischen den Lookalike-Daten und den Trainingsdaten. Der Relevanzwert wird verwendet, um zu entscheiden, wie groß die Zielgruppe sein sollte, und nicht, ob das Modell gut abschneidet.

Der Recall ist ein unvoreingenommenes Maß dafür, wie ähnlich das Lookalike-Segment den Trainingsdaten ist. Die Rückrufaktion ist der Prozentsatz der Nutzer, die sich am ähnlichsten sind (standardmäßig die ähnlichsten 20%) aus einer Stichprobe von Trainingsdaten, die in der Startzielgruppe nach dem Job zur Zielgruppengenerierung enthalten sind. Die Werte liegen zwischen 0 und 1, größere Werte deuten auf eine bessere Zielgruppe hin. Ein Erinnerungswert, der in etwa dem maximalen Prozentsatz entspricht, gibt an, dass das Zielgruppenmodell einer zufälligen Auswahl entspricht.

Wir halten dies für eine bessere Bewertungsmetrik als Genauigkeit, Präzision und F1-Werte, da Clean Rooms ML bei der Erstellung seines Modells nicht genau als negativ eingestufte Nutzer eingestuft hat.

Der Relevanzwert auf Segmentebene ist ein Maß für die Ähnlichkeit mit Werten im Bereich von -1 (am wenigsten ähnlich) bis 1 (am ähnlichsten). Clean Rooms ML berechnet eine Reihe von Relevanzwerten für verschiedene Segmentgrößen, damit Sie die beste Segmentgröße für Ihre Daten ermitteln können. Die Relevanzwerte nehmen mit zunehmender Segmentgröße monoton ab, sodass sie mit zunehmender Segmentgröße den Ausgangsdaten weniger ähnlich sein können. Wenn der Relevanzwert auf Segmentebene 0 erreicht, prognostiziert das Modell, dass alle Benutzer im Lookalike-Segment aus derselben Verteilung stammen wie die Ausgangsdaten. Durch eine Erhöhung der Ausgabegröße werden wahrscheinlich auch Benutzer im Lookalike-Segment berücksichtigt, die nicht aus derselben Verteilung wie die Ausgangsdaten stammen.

Die Relevanzwerte werden innerhalb einer einzelnen Kampagne normalisiert und sollten nicht für Vergleiche zwischen Kampagnen verwendet werden. Relevanzwerte sollten nicht als Einzelnachweis für Geschäftsergebnisse verwendet werden, da diese neben der Relevanz auch von mehreren komplexen Faktoren beeinflusst werden, wie z. B. Bestandsqualität, Inventarart, Zeitpunkt der Werbung usw.

Relevanzwerte sollten nicht dazu verwendet werden, die Qualität des Saatguts zu beurteilen, sondern eher, ob sie erhöht oder verringert werden kann. Betrachten Sie die folgenden Beispiele:

- Durchweg positive Werte — Dies deutet darauf hin, dass es mehr Output-Nutzer gibt, die als ähnlich prognostiziert werden, als dass sie im Lookalike-Segment enthalten sind. Dies ist häufig bei Saatgutdaten der Fall, die Teil eines großen Marktes sind, z. B. bei allen, die im letzten Monat Zahnpasta gekauft haben. Wir empfehlen, sich kleinere Samendaten anzusehen, z. B. alle, die im letzten Monat mehr als einmal Zahnpasta gekauft haben.
- Alle Werte sind negativ oder negativ für Ihre gewünschte Lookalike-Segmentgröße — Dies deutet darauf hin, dass Clean Rooms ML davon ausgeht, dass es in der gewünschten Lookalike-

Segmentgröße nicht genügend ähnliche Benutzer gibt. Das kann daran liegen, dass die Startdaten zu spezifisch sind oder der Markt zu klein ist. Wir empfehlen, entweder weniger Filter auf die Saatgutdaten anzuwenden oder den Markt zu erweitern. Wenn es sich bei den ursprünglichen Ausgangsdaten beispielsweise um Kunden handelte, die einen Kinderwagen und einen Kindersitz gekauft haben, könnten Sie den Markt auf Kunden ausdehnen, die mehrere Babyartikel gekauft haben.

Die Anbieter von Schulungsdaten bestimmen, ob die Relevanzwerte veröffentlicht werden und welche Felder für die Berechnung der Relevanzwerte verwendet werden.

Benutzerdefinierte Modelle in Clean Rooms ML

Mit Clean Rooms ML können Mitglieder einer Kollaboration einen angedockten benutzerdefinierten Modellalgorithmus verwenden, der in Amazon ECR gespeichert ist, um ihre Daten gemeinsam zu analysieren. Dazu muss der Modellanbieter ein Bild erstellen und es in Amazon ECR speichern. Folgen Sie den Schritten im [Amazon Elastic Container Registry User Guide](#), um ein privates Repository zu erstellen, das das benutzerdefinierte ML-Modell enthalten wird.

Jedes Mitglied einer Kollaboration kann der Modellanbieter sein, vorausgesetzt, es verfügt über die richtigen Berechtigungen. Alle Mitglieder einer Kollaboration können Daten zum Modell beitragen. Für die Zwecke dieses Leitfadens werden Mitglieder, die Daten beisteuern, als Datenanbieter bezeichnet. Das Mitglied, das die Kollaboration erstellt, ist der Ersteller der Kollaboration, und dieses Mitglied kann entweder der Modellanbieter, einer der Datenanbieter oder beides sein.

In den folgenden Themen werden die Informationen beschrieben, die zum Erstellen eines benutzerdefinierten ML-Modells erforderlich sind

Topics

- [Voraussetzungen für die benutzerdefinierte ML-Modellierung](#)
- [Richtlinien für die Modellerstellung für den Trainingscontainer](#)
- [Richtlinien für die Modellerstellung für den Inferenzcontainer](#)
- [Empfangen von Modellprotokollen und Metriken](#)

Voraussetzungen für die benutzerdefinierte ML-Modellierung

Bevor Sie eine benutzerdefinierte ML-Modellierung durchführen können, sollten Sie Folgendes berücksichtigen:

- Stellen Sie fest, ob im Rahmen der Zusammenarbeit sowohl das Modelltraining als auch die Inferenz am trainierten Modell durchgeführt werden sollen.
- Legen Sie fest, welche Rolle jedes Kollaborationsmitglied übernehmen soll, und weisen Sie ihm die entsprechenden Fähigkeiten zu.
 - Weisen Sie die `CAN_QUERY` Fähigkeit dem Mitglied zu, das das Modell trainiert und Inferenzen anhand des trainierten Modells durchführt.
 - Weisen Sie `CAN_RECEIVE_RESULTS` die mindestens einem Mitglied der Kollaboration zu.
 - Weisen Sie `CAN_RECEIVE_MODEL_OUTPUT` dem Mitglied, das trainierte Modellexporte bzw. Inferenzergebnisse erhalten soll, `CAN_RECEIVE_INFERENCE_OUTPUT` Fähigkeiten zu. Sie können sich dafür entscheiden, beide Fähigkeiten zu verwenden, wenn sie für Ihren Anwendungsfall erforderlich sind.
- Bestimmen Sie die maximale Größe der trainierten Modellartefakte oder Inferenzergebnisse, die exportiert werden dürfen.
- Wir empfehlen, dass alle Benutzer die `CleanroomsMLFullAccess` Richtlinien `CleanroomsFullAccess` und ihrer Rolle zuordnen. Für die Verwendung benutzerdefinierter ML-Modelle müssen AWS Clean Rooms sowohl das ML als auch AWS Clean Rooms ML verwendet SDKs werden.
- Beachten Sie die folgenden Informationen zu IAM-Rollen.
 - Alle Datenanbieter müssen über eine Servicezugriffsrolle verfügen, die es AWS Clean Rooms ermöglicht, Daten aus ihren AWS Glue Katalogen und Tabellen sowie den zugrunde liegenden Amazon S3 S3-Standorten zu lesen. Diese Rollen ähneln denen, die für SQL-Abfragen erforderlich sind. Dadurch können Sie die `CreateConfiguredTableAssociation` Aktion verwenden. Weitere Informationen finden Sie unter [Erstellen Sie eine Servicerolle, um eine konfigurierte Tabellenzuordnung zu erstellen](#).
 - Alle Mitglieder, die Metriken erhalten möchten, müssen über eine Dienstzugriffsrolle verfügen, die es ihnen ermöglicht, CloudWatch Metriken und Protokolle zu schreiben. Diese Rolle wird von Clean Rooms ML verwendet, um AWS-Konto während des Modelltrainings und der Inferenz alle Modellmetriken und -protokolle in die Mitglieder zu schreiben. Wir bieten auch Datenschutzkontrollen an, um zu bestimmen, welche Mitglieder Zugriff auf die Metriken und Protokolle haben. Auf diese Weise können Sie die `CreateMLConfiguration` Aktion verwenden. Weitere Informationen finden Sie unter [Erstellen Sie eine Servicerolle für die benutzerdefinierte ML-Modellierung — ML-Konfiguration](#).

Das Mitglied, das Ergebnisse erhält, muss eine Servicezugriffsrolle mit Schreibberechtigungen für seinen Amazon S3 S3-Bucket bereitstellen. Diese Rolle ermöglicht es Clean Rooms ML,

Ergebnisse (trainierte Modellartefakte oder Inferenzergebnisse) in einen Amazon S3 S3-Bucket zu exportieren. Auf diese Weise können Sie die `CreateMLConfiguration` Aktion verwenden. Weitere Informationen finden Sie unter [Erstellen Sie eine Servicerolle für die benutzerdefinierte ML-Modellierung — ML-Konfiguration](#).

- Der Modellanbieter muss eine Servicezugriffsrolle mit Berechtigungen zum Lesen seines Amazon ECR-Repositorys und Images bereitstellen. Auf diese Weise können Sie die `CreateConfigureModelAlgorithm` Aktion verwenden. Weitere Informationen finden Sie unter [Erstellen Sie eine Servicerolle, um ein benutzerdefiniertes ML-Modell bereitzustellen](#).
- Das Mitglied, das die erstellt, `MLInputChannel` um Datensätze für Training oder Inferenz zu generieren, muss eine Dienstzugriffsrolle bereitstellen, in der Clean Rooms ML eine SQL-Abfrage ausführen kann. AWS Clean Rooms Auf diese Weise können Sie die Aktionen `CreateTrainedModel` und `StartTrainedModelInferenceJob` verwenden. Weitere Informationen finden Sie unter [Erstellen Sie eine Servicerolle, um einen Datensatz abzufragen](#).
- Modellautoren sollten sich an die Regeln [Richtlinien für die Modellerstellung für den Trainingscontainer](#) und halten [Richtlinien für die Modellerstellung für den Inferenzcontainer](#), um sicherzustellen, dass die Modelleingaben und -ausgaben wie erwartet von konfiguriert sind AWS Clean Rooms.

Richtlinien für die Modellerstellung für den Trainingscontainer

In diesem Abschnitt werden die Richtlinien beschrieben, die Modellanbieter bei der Erstellung eines benutzerdefinierten ML-Modellalgorithmus für Clean Rooms ML beachten sollten.

- Verwenden Sie das entsprechende, von SageMaker KI-Schulungen unterstützte Container-Basis-Image, wie im [SageMaker AI Developer Guide](#) beschrieben. Mit dem folgenden Code können Sie die unterstützten Container-Basis-Images von öffentlichen SageMaker KI-Endpunkten abrufen.

```
ecr_registry_endpoint='763104351884.dkr.ecr.$REGION.amazonaws.com'  
base_image='pytorch-training:2.3.0-cpu-py311-ubuntu20.04-sagemaker'  
aws ecr get-login-password --region $REGION | docker login --username AWS --password-  
stdin $ecr_registry_endpoint  
docker pull $ecr_registry_endpoint/$base_image
```

- Achten Sie bei der lokalen Erstellung des Modells auf Folgendes, damit Sie Ihr Modell lokal, auf einer Entwicklungsinstanz, auf SageMaker AI Training in Ihrem AWS-Konto und auf Clean Rooms ML testen können.

- Wir empfehlen, ein Trainingsskript zu schreiben, das über verschiedene Umgebungsvariablen auf nützliche Eigenschaften der Trainingsumgebung zugreift. Clean Rooms ML verwendet die folgenden Argumente, um das Training für Ihren Modellcode aufzurufen: `SM_MODEL_DIR`, `SM_OUTPUT_DIR`, `SM_CHANNEL_TRAIN`, und `FILE_FORMAT`. Diese Standardwerte werden von Clean Rooms ML verwendet, um Ihr ML-Modell in einer eigenen Ausführungsumgebung mit den Daten aller Parteien zu trainieren.
- Clean Rooms ML stellt Ihre Trainingseingabekanäle über die `/opt/ml/input/data/channel-name` Verzeichnisse im Docker-Container zur Verfügung. Jeder ML-Eingangskanal wird auf der Grundlage seines in der Anfrage `channel_name` angegebenen entsprechenden Kanals zugeordnet. `CreateTrainedModel`

```
parser = argparse.ArgumentParser()# Data, model, and output directories

parser.add_argument('--model_dir', type=str, default=os.environ.get('SM_MODEL_DIR',
"/opt/ml/model"))
parser.add_argument('--output_dir', type=str,
default=os.environ.get('SM_OUTPUT_DIR', "/opt/ml/output/data"))
parser.add_argument('--train_dir', type=str,
default=os.environ.get('SM_CHANNEL_TRAIN', "/opt/ml/input/data/train"))
parser.add_argument('--train_file_format', type=str,
default=os.environ.get('FILE_FORMAT', "csv"))
```

- Stellen Sie sicher, dass Sie in der Lage sind, einen synthetischen Datensatz oder einen Testdatensatz auf der Grundlage des Schemas der Mitarbeiter zu generieren, das in Ihrem Modellcode verwendet wird.
- Stellen Sie sicher, dass Sie einen SageMaker KI-Trainingsjob selbst ausführen können, AWS-Konto bevor Sie den Modellalgorithmus einer AWS Clean Rooms Kollaboration zuordnen.

Der folgende Code enthält eine Docker-Beispieldatei, die mit lokalen Tests, Tests der SageMaker KI-Trainingsumgebung und Clean Rooms ML kompatibel ist

```
FROM 763104351884.dkr.ecr.us-west-2.amazonaws.com/pytorch-training:2.3.0-cpu-
py311-ubuntu20.04-sagemaker
MAINTAINER $author_name

ENV PYTHONDONTWRITEBYTECODE=1 \
    PYTHONUNBUFFERED=1 \
    LD_LIBRARY_PATH="${LD_LIBRARY_PATH}:/usr/local/lib"

ENV PATH="/opt/ml/code:${PATH}"
```

```
# this environment variable is used by the SageMaker PyTorch container to determine
our user code directory
ENV SAGEMAKER_SUBMIT_DIRECTORY /opt/ml/code

# copy the training script inside the container
COPY train.py /opt/ml/code/train.py
# define train.py as the script entry point
ENV SAGEMAKER_PROGRAM train.py
ENTRYPOINT ["python", "/opt/ml/code/train.py"]
```

- Um Container-Fehler bestmöglich überwachen zu können, empfehlen wir, Protokolle zu exportieren und aus Fehlergründen zu debuggen. In einer `GetTrainedModel` Antwort gibt Clean Rooms ML die ersten 1024 Zeichen aus dieser Datei unter `StatusDetails` zurück.
- Nachdem Sie alle Modelländerungen vorgenommen haben und bereit sind, es in der SageMaker KI-Umgebung zu testen, führen Sie die folgenden Befehle in der angegebenen Reihenfolge aus.

```
export ACCOUNT_ID=xxx
export REPO_NAME=xxx
export REPO_TAG=xxx
export REGION=xxx

docker build -t $ACCOUNT_ID.dkr.ecr.us-west-2.amazonaws.com/$REPO_NAME:$REPO_TAG

# Sign into AWS $ACCOUNT_ID/ Run aws configure
# Check the account and make sure it is the correct role/credentials
aws sts get-caller-identity
aws ecr create-repository --repository-name $REPO_NAME --region $REGION
aws ecr describe-repositories --repository-name $REPO_NAME --region $REGION

# Authenticate Docker
aws ecr get-login-password --region $REGION | docker login --username AWS --password-
stdin $ACCOUNT_ID.dkr.ecr.$REGION.amazonaws.com

# Push To ECR Images
docker push $ACCOUNT_ID.dkr.ecr.$REGION.amazonaws.com$REPO_NAME:$REPO_TAG

# Create Sagemaker Training job
# Configure the training_job.json with
# 1. TrainingImage
# 2. Input DataConfig
# 3. Output DataConfig
```

```
aws sagemaker create-training-job --cli-input-json file://training_job.json --region $REGION
```

Nachdem der SageMaker KI-Job abgeschlossen ist und Sie mit Ihrem Modellalgorithmus zufrieden sind, können Sie das Amazon ECR Registry bei AWS Clean Rooms ML registrieren. Verwenden Sie die `CreateConfiguredModelAlgorithm` Aktion, um den Modellalgorithmus `CreateConfiguredModelAlgorithmAssociation` zu registrieren und ihn anschließend einer Kollaboration zuzuordnen.

Richtlinien für die Modellerstellung für den Inferenzcontainer

In diesem Abschnitt werden die Richtlinien beschrieben, die Modellanbieter bei der Erstellung eines Inferenzalgorithmus für Clean Rooms ML beachten sollten.

- Verwenden Sie das entsprechende, von SageMaker KI-Inferenzen unterstützte Container-Basis-Image, wie im [SageMaker AI Developer Guide](#) beschrieben. Mit dem folgenden Code können Sie die unterstützten Container-Basis-Images von öffentlichen SageMaker KI-Endpunkten abrufen.

```
ecr_registry_endpoint='763104351884.dkr.ecr.$REGION.amazonaws.com'  
base_image='pytorch-inference:2.3.0-cpu-py311-ubuntu20.04-sagemaker'  
aws ecr get-login-password --region $REGION | docker login --username AWS --password-stdin $ecr_registry_endpoint  
docker pull $ecr_registry_endpoint/$base_image
```

- Wenn Sie das Modell lokal erstellen, stellen Sie Folgendes sicher, damit Sie Ihr Modell lokal, auf einer Entwicklungsinstanz, auf SageMaker AI Batch Transform in Ihrem AWS-Konto und auf Clean Rooms ML testen können.
 - Clean Rooms ML stellt Ihre Modellartefakte aus der Inferenz über das `/opt/ml/model` Verzeichnis im Docker-Container zur Verwendung durch Ihren Inferenzcode zur Verfügung.
 - Clean Rooms ML teilt die Eingabe zeilenweise auf, verwendet eine `MultiRecordBatch` Strategie und fügt am Ende jedes transformierten Datensatzes ein Zeilenumbruchzeichen hinzu.
 - Stellen Sie sicher, dass Sie auf der Grundlage des Schemas der Mitarbeiter, die in Ihrem Modellcode verwendet werden, einen synthetischen Datensatz oder einen Test-Inferenzdatensatz generieren können.
 - Stellen Sie sicher, dass Sie einen SageMaker AI-Batch-Transformationsjob selbst ausführen können, AWS-Konto bevor Sie den Modellalgorithmus einer AWS Clean Rooms Kollaboration zuordnen.

Der folgende Code enthält eine Docker-Beispieldatei, die mit lokalen Tests, Tests von SageMaker KI-Transformationsumgebungen und Clean Rooms ML kompatibel ist

```
FROM 763104351884.dkr.ecr.us-east-1.amazonaws.com/pytorch-inference:1.12.1-cpu-
py38-ubuntu20.04-sagemaker

ENV PYTHONUNBUFFERED=1

COPY serve.py /opt/ml/code/serve.py
COPY inference_handler.py /opt/ml/code/inference_handler.py
COPY handler_service.py /opt/ml/code/handler_service.py
COPY model.py /opt/ml/code/model.py

RUN chmod +x /opt/ml/code/serve.py

ENTRYPOINT ["/opt/ml/code/serve.py"]
```

- Nachdem Sie alle Modelländerungen vorgenommen haben und bereit sind, sie in der SageMaker KI-Umgebung zu testen, führen Sie die folgenden Befehle in der angegebenen Reihenfolge aus.

```
export ACCOUNT_ID=xxx
export REPO_NAME=xxx
export REPO_TAG=xxx
export REGION=xxx

docker build -t $ACCOUNT_ID.dkr.ecr.us-west-2.amazonaws.com/$REPO_NAME:$REPO_TAG

# Sign into AWS $ACCOUNT_ID/ Run aws configure
# Check the account and make sure it is the correct role/credentials
aws sts get-caller-identity
aws ecr create-repository --repository-name $REPO_NAME --region $REGION
aws ecr describe-repositories --repository-name $REPO_NAME --region $REGION

# Authenticate Docker
aws ecr get-login-password --region $REGION | docker login --username AWS --password-
stdin $ACCOUNT_ID.dkr.ecr.$REGION.amazonaws.com

# Push To ECR Repository
docker push $ACCOUNT_ID.dkr.ecr.$REGION.amazonaws.com$REPO_NAME:$REPO_TAG

# Create Sagemaker Model
# Configure the create_model.json with
```

```
# 1. Primary container -
# a. ModelDataUrl - S3 Uri of the model.tar from your training job
aws sagemaker create-model --cli-input-json file://create_model.json --region $REGION

# Create Sagemaker Transform Job
# Configure the transform_job.json with
# 1. Model created in the step above
# 2. MultiRecord batch strategy
# 3. Line SplitType for TransformInput
# 4. AssembleWith Line for TransformOutput
aws sagemaker create-transform-job --cli-input-json file://transform_job.json --
region $REGION
```

Nachdem der SageMaker KI-Job abgeschlossen ist und Sie mit Ihrer Batch-Transformation zufrieden sind, können Sie das Amazon ECR Registry bei AWS Clean Rooms ML registrieren. Verwenden Sie die `CreateConfiguredModelAlgorithm` Aktion, um den Modellalgorithmus `CreateConfiguredModelAlgorithmAssociation` zu registrieren und ihn anschließend einer Kollaboration zuzuordnen.

Empfangen von Modellprotokollen und Metriken

Um Protokolle und Metriken aus Schulungen oder Inferenzen mit benutzerdefinierten Modellen zu erhalten, müssen die Mitglieder [eine ML-Konfiguration mit einer gültigen Rolle erstellt](#) haben, die die erforderlichen CloudWatch Berechtigungen bereitstellt (siehe [Erstellen einer Servicerolle für benutzerdefinierte ML-Modellierung — ML-Konfiguration](#)).

System-Metrik

Systemmetriken für Training und Inferenz, wie CPU- und Speicherauslastung, werden allen Mitgliedern der Zusammenarbeit mit gültigen ML-Konfigurationen zur Verfügung gestellt. Diese Metriken können im Verlauf des Jobs über CloudWatch Metriken in den jeweiligen `/aws/cleanroomsm1/TrainedModelInferenceJobs` Namespaces `/aws/cleanroomsm1/TrainedModels` eingesehen werden.

Modellieren Sie Logs

Der Zugriff auf die Modellprotokolle erfolgt über die Datenschutzrichtlinien der einzelnen konfigurierten Modellalgorithmen. Der Modellautor legt die Datenschutzrichtlinie fest, wenn er einen konfigurierten Modellalgorithmus (entweder über die Konsole oder die `CreateConfiguredModelAlgorithmAssociation` API) einer Kollaboration zuordnet. Durch die

Festlegung der Datenschutzrichtlinie wird gesteuert, welche Mitglieder die Modellprotokolle erhalten können.

Darüber hinaus kann der Modellautor in der Datenschutzrichtlinie ein Filtermuster festlegen, um Protokollereignisse zu filtern. Alle Logs, die ein Modellcontainer an `stdout` oder `stderr` sendet und die dem Filtermuster entsprechen (falls gesetzt), werden an Amazon CloudWatch Logs gesendet. Modellprotokolle sind jeweils in CloudWatch Protokollgruppen `/aws/cleanroomsm1/TrainedModels` verfügbar. `/aws/cleanroomsm1/TrainedModelInferenceJobs`

Benutzerdefinierte Metriken

Wenn Sie einen Modellalgorithmus konfigurieren (entweder über die Konsole oder die `CreateConfiguredModelAlgorithm` API), kann der Modellautor bestimmte Metrikenamen und Regex-Anweisungen angeben, nach denen in den Ausgabeprotokollen gesucht werden soll. Diese können im Verlauf des Jobs über CloudWatch Metrics im Namespace eingesehen werden. `/aws/cleanroomsm1/TrainedModels` Bei der Zuordnung eines konfigurierten Modellalgorithmus kann der Modellautor in der Datenschutzkonfiguration der Metriken einen optionalen Geräuschpegel festlegen, um die Ausgabe von Rohdaten zu vermeiden und dennoch Einblick in benutzerdefinierte Metriktrends zu erhalten. Wenn ein Geräuschpegel festgelegt ist, werden die Messwerte am Ende des Jobs veröffentlicht und nicht in Echtzeit.

Kryptografisches Rechnen für Clean Rooms

Kryptografisches Rechnen für Clean Rooms (C3R) ist eine Funktion AWS Clean Rooms, die zusätzlich zu [Analyiseregeln](#) verwendet werden kann. Mit C3R können Unternehmen sensible Daten zusammenführen, um neue Erkenntnisse aus der Datenanalyse zu gewinnen, und gleichzeitig kryptografisch einschränken, was von jeder Partei im Prozess gelernt werden kann. C3R kann von zwei oder mehr Parteien verwendet werden, die mit ihren sensiblen Daten zusammenarbeiten möchten, aber nur verschlüsselte Daten in der Cloud verwenden müssen.

Der C3R-Verschlüsselungscient ist ein clientseitiges Verschlüsselungstool, mit dem Sie Ihre Daten für die Verwendung mit [verschlüsseln](#) können. AWS Clean Rooms Wenn Sie den C3R-Verschlüsselungscient verwenden, bleiben Daten während der Verwendung in einer Zusammenarbeit kryptografisch geschützt. AWS Clean Rooms Wie bei einer normalen AWS Clean Rooms Zusammenarbeit handelt es sich bei den Eingabedaten um relationale Datenbanktabellen, und die Berechnung wird als SQL-Abfrage ausgedrückt. C3R unterstützt jedoch nur eine begrenzte Teilmenge von SQL-Abfragen für verschlüsselte Daten.

Insbesondere unterstützt C3R SQL JOIN and SELECT Aussagen zu kryptografisch geschützten Daten. Jede Spalte in der Eingabetabelle kann in genau einem der folgenden SQL-Anweisungstypen verwendet werden:

- Spalten, die kryptografisch geschützt sind für die Verwendung in JOIN Anweisungen werden aufgerufen fingerprint Spalten.
- Spalten, die kryptografisch geschützt sind für die Verwendung in SELECT Anweisungen werden aufgerufen sealed Spalten.
- Spalten, die nicht kryptografisch geschützt sind, für die Verwendung in JOIN or SELECT Anweisungen werden aufgerufen cleartext Spalten.

In einigen Fällen GROUP BY Aussagen werden gestützt auf fingerprint Spalten. Weitere Informationen finden Sie unter [FingerprintSpalten](#). Derzeit unterstützt C3R nicht die Verwendung anderer SQL-Konstrukte für verschlüsselte Daten, wie z. B. WHERE Klauseln oder Aggregatfunktionen wie SUM and AVERAGE, auch wenn sie sonst nach den entsprechenden Analyseregeln zulässig wären.

C3R wurde entwickelt, um Daten in einzelnen Zellen einer Tabelle zu schützen. Bei Verwendung der Standardkonfiguration für C3R bleiben die zugrunde liegenden Daten, die ein Kunde im Rahmen einer Zusammenarbeit Dritten zur Verfügung stellt, verschlüsselt, während der Inhalt darin verwendet wird. AWS Clean Rooms C3R verwendet die branchenübliche AES-GCM-Verschlüsselung für alle sealed Spalten und eine dem Industriestandard entsprechende Pseudozufallsfunktion, bekannt als Hash-Based Message Authentication Code (HMAC), zum Schutz fingerprint Spalten.

Obwohl C3R die Daten in Ihren Tabellen verschlüsselt, können die folgenden Informationen möglicherweise dennoch abgeleitet werden:

- Informationen zu den Tabellen selbst, einschließlich der Anzahl der Spalten, der Spaltennamen und der Anzahl der Zeilen in Ihrer Tabelle.
- Wie bei den meisten Standardverschlüsselungsformen versucht C3R nicht, die Länge der verschlüsselten Werte zu verbergen. C3R bietet die Möglichkeit, verschlüsselte Werte aufzufüllen, um die genaue Länge von Klartexten zu verbergen. Eine Obergrenze für die Länge der Klartexte in jeder Spalte könnte jedoch immer noch einer anderen Partei offengelegt werden.
- Informationen auf Protokollebene, z. B. wann eine bestimmte Zeile zu einer verschlüsselten C3R-Tabelle hinzugefügt wurde.

Weitere Informationen zu C3R finden Sie in den folgenden Themen.

Themen

- [Überlegungen zur Verwendung von Cryptographic Computing für Clean Rooms](#)
- [Unterstützte Datei- und Datentypen in Cryptographic Computing für Clean Rooms](#)
- [Spaltennamen in Cryptographic Computing für Clean Rooms](#)
- [Spaltentypen in Cryptographic Computing für Clean Rooms](#)
- [Kryptografische Rechenparameter](#)
- [Optionale Flags in Cryptographic Computing für Clean Rooms](#)
- [Abfragen mit Cryptographic Computing für Clean Rooms](#)
- [Richtlinien für den C3R-Verschlüsselungsclient](#)

Überlegungen zur Verwendung von Cryptographic Computing für Clean Rooms

Cryptographic Computing for Clean Rooms (C3R) zielt darauf ab, den Datenschutz zu maximieren. Einige Anwendungsfälle könnten jedoch von einem geringeren Datenschutzniveau im Austausch für zusätzliche Funktionen profitieren. Sie können diese spezifischen Kompromisse eingehen, indem Sie C3R von der sichersten Konfiguration aus ändern. Als Kunde sollten Sie sich dieser Kompromisse bewusst sein und entscheiden, ob sie für Ihren Anwendungsfall geeignet sind. Zu den Kompromissen, die es zu berücksichtigen gilt, gehören:

Themen

- [Zulassen gemischter cleartext und verschlüsselter Daten in Ihren Tabellen](#)
- [Wiederholte Werte in fingerprint Spalten zulassen](#)
- [Lockerung der Beschränkungen für die Benennung von fingerprint Spalten](#)
- [Bestimmen, wie NULL Werte dargestellt werden](#)

Weitere Informationen zum Einstellen von Parametern für diese Szenarien finden Sie unter.

[Kryptografische Rechenparameter](#)

Zulassen gemischter cleartext und verschlüsselter Daten in Ihren Tabellen

Die clientseitige Verschlüsselung aller Daten bietet maximalen Datenschutz. Dadurch werden jedoch bestimmte Arten von Abfragen eingeschränkt (z. B. die SUM Aggregatfunktion). Das Risiko, cleartext Daten zuzulassen, besteht darin, dass es möglich ist, dass jeder, der Zugriff auf die verschlüsselten

Tabellen hat, Informationen über verschlüsselte Werte ableiten kann. Dies könnte durch eine statistische Analyse der Daten cleartext und der zugehörigen Daten geschehen.

Stellen Sie sich zum Beispiel vor, Sie hätten die Spalten City und State. Die City Spalte ist cleartext und die State Spalte ist verschlüsselt. Wenn Sie den Wert Chicago in der City Spalte sehen, können Sie mit hoher Wahrscheinlichkeit feststellen, State dass der Illinois. Im Gegensatz dazu, wenn eine Spalte City und die andere Spalte ist EmailAddress, ist es unwahrscheinlich, dass a cleartext City etwas über eine verschlüsselte Spalte aussagt EmailAddress.

Weitere Informationen zu dem Parameter für dieses Szenario finden Sie unter [Parameter „Spalten zulassencleartext“](#).

Wiederholte Werte in fingerprint Spalten zulassen

Für den sichersten Ansatz gehen wir davon aus, dass jede fingerprint Spalte genau eine Instanz einer Variablen enthält. Kein Element kann in einer fingerprint Spalte wiederholt werden. Der C3R-Verschlüsselungscient ordnet diese cleartext Werte eindeutigen Werten zu, die nicht von Zufallswerten zu unterscheiden sind. Daher ist es unmöglich, aus diesen Zufallswerten Informationen über die abzuleiten. cleartext

Das Risiko wiederholter Werte in einer fingerprint Spalte besteht darin, dass wiederholte Werte zu wiederholten zufällig aussehenden Werten führen. Somit könnte theoretisch jeder, der Zugriff auf die verschlüsselten Tabellen hat, eine statistische Analyse der fingerprint Spalten durchführen, die Informationen über cleartext Werte liefern könnte.

Nehmen wir erneut an State, die fingerprint Spalte entspricht einem US-Haushalt und jede Zeile der Tabelle entspricht. Durch eine Frequenzanalyse könnte man ableiten, um welchen Bundesstaat es sich handelt California und welcher Wyoming mit hoher Wahrscheinlichkeit. Diese Schlussfolgerung ist möglich, weil es viel mehr Einwohner California hat als. Wyoming Nehmen wir dagegen an, die fingerprint Spalte bezieht sich auf eine Haushalts-ID und jeder Haushalt tauchte in der Datenbank ein- bis viermal in einer Datenbank mit Millionen von Einträgen auf. Es ist unwahrscheinlich, dass eine Frequenzanalyse nützliche Informationen liefern würde.

Weitere Informationen zu den Parametern für dieses Szenario finden Sie unter [Parameter „Duplikate zulassen“](#).

Lockerung der Beschränkungen für die Benennung von fingerprint Spalten

Standardmäßig gehen wir davon aus, dass, wenn zwei Tabellen mithilfe verschlüsselter fingerprint Spalten verknüpft werden, diese Spalten in jeder Tabelle denselben Namen haben. Der technische Grund für dieses Ergebnis ist, dass wir standardmäßig einen anderen kryptografischen Schlüssel für die Verschlüsselung jeder Spalte ableiten. Dieser Schlüssel wird aus einer Kombination aus dem gemeinsamen geheimen Schlüssel für die Zusammenarbeit und dem Spaltennamen abgeleitet. Wenn wir versuchen, zwei Spalten mit unterschiedlichen Spaltennamen zu verbinden, leiten wir unterschiedliche Schlüssel ab und können keinen gültigen Join berechnen.

Um dieses Problem zu beheben, können Sie die Funktion deaktivieren, die Schlüssel aus jedem Spaltennamen ableitet. Anschließend verwendet der C3R-Verschlüsselungsclient einen einzigen abgeleiteten Schlüssel für alle fingerprint Spalten. Das Risiko besteht darin, dass eine andere Art der Frequenzanalyse durchgeführt werden kann, die Informationen preisgeben könnte.

Lassen Sie uns das State Beispiel City und noch einmal verwenden. Wenn wir für jede fingerprint Spalte dieselben Zufallswerte ableiten (indem wir den Spaltennamen nicht einbeziehen). New York hat den gleichen Zufallswert in den Spalten City und State. New York ist eine der wenigen Städte in den USA, in denen der City Name mit dem State Namen identisch ist. Wenn Ihr Datensatz dagegen in jeder Spalte völlig unterschiedliche Werte enthält, werden keine Informationen durchgesickert.

Weitere Informationen zum Parameter für dieses Szenario finden Sie unter [Parameter „Zulassen JOIN von Spalten mit unterschiedlichen Namen“](#).

Bestimmen, wie NULL Werte dargestellt werden

Sie haben die Wahl, ob Sie Werte wie alle anderen NULL Werte kryptografisch (verschlüsseln und HMAC) verarbeiten möchten. Wenn Sie Werte nicht wie alle anderen NULL Werte verarbeiten, können Informationen preisgegeben werden.

Nehmen wir zum Beispiel an, dass NULL in der Middle Name Spalte in der Personen ohne zweiten Vornamen cleartext angegeben werden. Wenn Sie diese Werte nicht verschlüsseln, können Sie durchsickern lassen, welche Zeilen in der verschlüsselten Tabelle für Personen ohne zweiten Vornamen verwendet werden. Diese Informationen könnten für einige Menschen in bestimmten Bevölkerungsgruppen ein Identifikationssignal sein. Wenn Sie NULL Werte jedoch kryptografisch verarbeiten, verhalten sich bestimmte SQL-Abfragen anders. Beispielsweise gruppieren GROUP BY Klauseln fingerprint NULL Werte in fingerprint Spalten nicht zusammen.

Weitere Informationen zum Parameter für dieses Szenario finden Sie unter [Parameter „NULLWerte beibehalten“](#).

Unterstützte Datei- und Datentypen in Cryptographic Computing für Clean Rooms

Der C3R-Verschlüsselungscient erkennt die folgenden Dateitypen:

- CSV-Dateien
- ParquetDateien

Sie können das `--fileFormat` Flag im C3R-Verschlüsselungscient verwenden, um ein Dateiformat explizit anzugeben. Wenn das Dateiformat explizit angegeben wird, wird es nicht durch die Dateierweiterung bestimmt.

Themen

- [CSV-Dateien](#)
- [ParquetDateien](#)
- [Verschlüsseln von Werten, die keine Zeichenfolge sind](#)

CSV-Dateien

Es wird davon ausgegangen, dass eine Datei mit der Erweiterung.csv im CSV-Format ist und UTF-8-codierten Text enthält. Der C3R-Verschlüsselungscient behandelt alle Werte als Zeichenketten.

Unterstützte Eigenschaften in CSV-Dateien

Der C3R-Verschlüsselungscient erfordert, dass CSV-Dateien die folgenden Eigenschaften haben:

- Kann eine erste Kopfzeile enthalten, die jede Spalte eindeutig benennt, oder auch nicht.
- Durch Kommas getrennt. (Derzeit werden benutzerdefinierte Trennzeichen nicht unterstützt.)
- UTF-8-codierter Text.

Löschen von Leerzeichen aus CSV-Einträgen

Sowohl führende als auch nachfolgende Leerzeichen werden aus CSV-Einträgen entfernt.

Benutzerdefinierte NULL Kodierung für eine CSV-Datei

Eine CSV-Datei kann eine benutzerdefinierte NULL Kodierung verwenden.

Mit dem C3R-Verschlüsselungsclient können Sie mithilfe des Flags benutzerdefinierte Kodierungen für NULL Einträge in den Eingabedaten angeben. `--csvInputNULLValue=<csv-input-null>`
Der C3R-Verschlüsselungsclient kann mithilfe des Flags benutzerdefinierte Kodierungen in der generierten Ausgabedatei für NULL-Einträge verwenden. `--csvOutputNULLValue=<csv-output-null>`

 **Note**

Es wird NULL davon ausgegangen, dass es einem Eintrag an Inhalt mangelt, insbesondere im Zusammenhang mit einem umfangreicheren Tabellenformat wie einer SQL-Tabelle. Obwohl .csv diese Charakterisierung aus historischen Gründen nicht ausdrücklich unterstützt, ist es üblich, einen leeren Eintrag, der nur Leerraum enthält, als solche zu betrachten. NULL Daher ist dies das Standardverhalten des C3R-Verschlüsselungsclients und kann nach Bedarf angepasst werden.

Wie werden CSV-Einträge von C3R interpretiert

Die folgende Tabelle enthält Beispiele dafür, wie .csv-Einträge auf der Grundlage der Werte (falls vorhanden)cleartext, die cleartext für die Flags und angegeben wurden, zusammengefasst werden (aus Gründen der Übersichtlichkeit). `--csvInputNULLValue=<csv-input-null>`
`--csvOutputNULLValue=<csv-output-null>` Leerzeichen am Anfang und am Ende von Anführungszeichen werden gekürzt, bevor C3R die Bedeutung eines Werts interpretiert.

<code><csv-input-null></code>	<code><csv-output-null></code>	Eingabeeintrag	Ausgangseintrag
Keine	Keine	,AnyProduct,	,AnyProduct,
Keine	Keine	, AnyProduct ,	,AnyProduct,
Keine	Keine	,"AnyProduct",	,AnyProduct,
Keine	Keine	, "AnyProdu ct" ,	,AnyProduct,

<csv-input-null>	<csv-output-null>	Eingabeeintrag	Ausgangseintrag
Keine	Keine	,,	,,
Keine	Keine	, ,	,,
Keine	Keine	, "",	,,
Keine	Keine	, " ",	, " ",
Keine	Keine	, " " ,	, " ",
"AnyProduct"	"NULL"	,AnyProduct,	,NULL,
"AnyProduct"	"NULL"	, AnyProduct ,	,NULL,
"AnyProduct"	"NULL"	, "AnyProduct",	,NULL,
"AnyProduct"	"NULL"	, "AnyProdu ct" ,	,NULL,
Keine	"NULL"	,,	,NULL,
Keine	"NULL"	, ,	,NULL,
Keine	"NULL"	, "",	,NULL,
Keine	"NULL"	, " ",	, " ",
Keine	"NULL"	, " " ,	, " ",
""	"NULL"	,,	,NULL,
""	"NULL"	, ,	,NULL,
""	"NULL"	, "",	, "",
""	"NULL"	, " ",	, " ",
""	"NULL"	, " " ,	, " ",

<csv-input-null>	<csv-output-null>	Eingabeeintrag	Ausgangseintrag
"\""	"NULL"	,,	,,
"\""	"NULL"	, ,	,,
"\""	"NULL"	, "",	, NULL,
"\""	"NULL"	, " ",	, " ",
"\""	"NULL"	, " " ,	, " ",

CSV-Datei ohne Header

Die CSV-Quelldatei muss keine Kopfzeilen in der ersten Zeile haben, die jede Spalte eindeutig benennen. Für eine CSV-Datei ohne Kopfzeile ist jedoch ein positionsbezogenes Verschlüsselungsschema erforderlich. Das Positionsverschlüsselungsschema ist anstelle des typischen Mapping-Schemas erforderlich, das sowohl für CSV-Dateien mit einer Kopfzeile als auch für Dateien verwendet wird. Parquet

Ein positionsabhängiges Verschlüsselungsschema spezifiziert Ausgabespalten nach Position statt nach Namen. Ein zugeordnetes Verschlüsselungsschema ordnet Quellspaltennamen Zielspaltennamen zu. Weitere Informationen, einschließlich einer ausführlichen Erläuterung und Beispielen für beide Schemaformate, finden Sie unter [Schemas für zugeordnete und positionierte Tabellen](#).

ParquetDateien

Es wird davon ausgegangen, dass eine Datei mit einer .parquet Erweiterung das Apache Parquet Format hat.

Unterstützte Parquet Datentypen

Der C3R-Verschlüsselungsclient kann alle nicht komplexen (d. h. primitiven Datentypen) Daten in einer Parquet Datei verarbeiten, die einen Datentyp darstellt, der von unterstützt wird. AWS Clean Rooms

Für Spalten können jedoch nur Zeichenkettenspalten verwendet werden. sealed

Die folgenden Parquet-Datentypen werden unterstützt:

- Binaryprimitiver Typ mit den folgenden logischen Anmerkungen:
 - Keine, wenn der gesetzte `--parquetBinaryAsString` ist (STRINGDatentyp)
 - `Decimal(scale, precision)`(DECIMALDatentyp)
 - `String`(STRINGDatentyp)
- Booleanprimitiver Datentyp ohne logische Anmerkung (BOOLEANDatentyp)
- Doubleprimitiver Datentyp ohne logische Anmerkung (DOUBLEDatentyp)
- Fixed_Len_Binary_Arrayprimitiver Typ mit der `Decimal(scale, precision)` logischen Anmerkung (DECIMALDatentyp)
- Floatprimitiver Datentyp ohne logische Anmerkung (FLOATDatentyp)
- Int32primitiver Typ mit den folgenden logischen Anmerkungen:
 - Keiner (INTDatentyp)
 - `Date`(DATEDatentyp)
 - `Decimal(scale, precision)`(DECIMALDatentyp)
 - `Int(16, true)`(SMALLINTDatentyp)
 - `Int(32, true)`(INTDatentyp)
- Int64primitiver Datentyp mit den folgenden logischen Anmerkungen:
 - Keiner (BIGINTDatentyp)
 - `Decimal(scale, precision)`(DECIMALDatentyp)
 - `Int(64, true)`(BIGINTDatentyp)
 - `Timestamp(isUTCAdjusted, TimeUnit.MILLIS)`(TIMESTAMPDatentyp)
 - `Timestamp(isUTCAdjusted, TimeUnit.MICROS)`(TIMESTAMPDatentyp)
 - `Timestamp(isUTCAdjusted, TimeUnit.NANOS)`(TIMESTAMPDatentyp)

Verschlüsseln von Werten, die keine Zeichenfolge sind

Derzeit werden nur Zeichenkettenwerte für sealed Spalten unterstützt.

Bei CSV-Dateien behandelt der C3R-Verschlüsselungsclient alle Werte als UTF-8-codierten Text und versucht nicht, sie vor der Verschlüsselung unterschiedlich zu interpretieren.

Bei Fingerabdruckspalten werden die Typen in Äquivalenzklassen eingeteilt. Eine Äquivalenzklasse ist ein Satz von Datentypen, deren Gleichheit anhand eines repräsentativen Datentyps eindeutig verglichen werden kann.

Äquivalenzklassen ermöglichen es, identische Fingerabdrücke demselben semantischen Wert zuzuweisen, unabhängig von der ursprünglichen Darstellung. Derselbe Wert in zwei Äquivalenzklassen führt jedoch nicht zu derselben Fingerabdruckspalte.

Beispielsweise 42 wird dem INTEGRAL Wert derselbe Fingerabdruck zugewiesen, unabhängig davon, ob es sich ursprünglich um ein SMALLINT, oder BIGINT handelte. Außerdem 0 wird der INTEGRAL Wert niemals mit dem BOOLEAN Wert FALSE (der durch den Wert repräsentiert wird) übereinstimmen.

Die folgenden Äquivalenzklassen und die entsprechenden AWS Clean Rooms Datentypen werden von Fingerabdruckspalten unterstützt:

Äquivalenzklasse	Unterstützter AWS Clean Rooms Datentyp
BOOLEAN	BOOLEAN
DATE	DATE
INTEGRAL	BIGINT, INT, SMALLINT
STRING	CHAR, STRING, VARCHAR

Spaltennamen in Cryptographic Computing für Clean Rooms

Standardmäßig sind die Namen von Spalten in Cryptographic Computing für Clean Rooms

Wenn der Wert des Parameters Spalten mit unterschiedlichen Namen zulassen JOIN auf False gesetzt ist, werden bei der Verschlüsselung von fingerprint Spalten Spaltennamen verwendet. Aus diesem Grund müssen sich Mitarbeiter standardmäßig im Voraus abstimmen und dieselben Zielspaltennamen für Daten verwenden, für die JOIN Anweisungen in Abfragen verwendet werden. Standardmäßig können Spalten, für die unterschiedliche Namen verschlüsselt wurden, JOIN JOIN bei keinem Wert erfolgreich verwendet werden.

Wenn der Wert des Parameters Spalten mit unterschiedlichen Namen zulassen JOIN wahr ist, sind JOIN Anweisungen für alle als Spalten verschlüsselten fingerprint Spalten erfolgreich. Das Verschlüsseln von Daten mit diesem Parameter ermöglicht möglicherweise einige Rückschlüsse auf die cleartext Werte. Wenn eine Zeile beispielsweise denselben HMAC-Wert (Hash-Based Message

Authentication Code) sowohl in der Spalte als auch in der City Spalte hat, könnte State der Wert lauten. New York

Normalisierung der Namen der Spaltenüberschriften

Die Namen der Spaltenüberschriften werden vom C3R-Verschlüsselungsclient normalisiert. Alle Leerzeichen am Anfang und Ende werden entfernt, und der Spaltenname wird für die transformierte Ausgabe in Kleinbuchstaben geschrieben.

Die Normalisierung wird vor allen anderen Berechnungen, Berechnungen oder anderen Operationen angewendet, die möglicherweise durch Spaltennamen beeinflusst werden könnten. Die ausgegebene Ausgabedatei enthält nur die normalisierten Namen.

Spaltentypen in Cryptographic Computing für Clean Rooms

Dieses Thema enthält Informationen zu Spaltentypen in Cryptographic Computing for. Clean Rooms

Themen

- [FingerprintSpalten](#)
- [Versiegelte Spalten](#)
- [CleartextSpalten](#)

FingerprintSpalten

FingerprintSpalten sind Spalten, die kryptografisch für die Verwendung in JOIN Anweisungen geschützt sind.

Daten aus fingerprint Spalten können nicht entschlüsselt werden. Nur Daten aus versiegelten Spalten können entschlüsselt werden.

FingerprintSpalten dürfen nur in den folgenden SQL-Klauseln und Funktionen verwendet werden:

- JOIN (INNER, OUTER, LEFT, RIGHT, or FULL) gegen andere fingerprint Spalten:
 - Wenn der Wert des `allowJoinsOnColumnsWithDifferentNames` Parameters auf `gesetzt` ist `false`, JOIN müssen beide fingerprint Spalten von ebenfalls denselben Namen haben.
- SELECT COUNT()
- SELECT COUNT(DISTINCT)

- GROUP BY (Nur verwenden, wenn die Kollaboration den Wert des `preserveNulls` Parameters auf festgelegt hat `true`.)

Abfragen, die gegen diese Einschränkungen verstoßen, können zu falschen Ergebnissen führen.

Versiegelte Spalten

Versiegelte Spalten sind Spalten, die kryptografisch für die Verwendung in SELECT Anweisungen geschützt sind.

Versiegelte Spalten dürfen nur in den folgenden SQL-Klauseln und Funktionen verwendet werden:

- SELECT
- SELECT ... AS
- SELECT COUNT()

Note

SELECT COUNT(DISTINCT) wird nicht unterstützt.

Abfragen, die gegen diese Einschränkungen verstoßen, können zu falschen Ergebnissen führen.

Daten für eine sealed Spalte vor der Verschlüsselung auffüllen

Wenn Sie angeben, dass eine Spalte eine Spalte sein soll, fragt C3R Sie, welche Art von Polsterung Sie wählen sollen. sealed Das Auffüllen von Daten vor der Verschlüsselung ist optional. Ohne Auffüllung (ein Pad-Typ von `none`) gibt die Länge der verschlüsselten Daten die Größe des an. cleartext Unter bestimmten Umständen cleartext könnte die Größe von den Klartext offenlegen. Bei Padding (ein Pad-Typ von `fixed` oder `max`) werden alle Werte zunächst auf eine gemeinsame Größe aufgefüllt und dann verschlüsselt. Beim Padding gibt die Länge der verschlüsselten Daten keine Auskunft über die ursprüngliche cleartext Länge, es sei denn, es wird eine Obergrenze für die Größe angegeben.

Wenn Sie für eine Spalte eine Auffüllung wünschen und die maximale Bytelänge der Daten in dieser Spalte bekannt ist, verwenden Sie `fixed` Padding. Verwenden Sie einen `length` Wert, der mindestens so groß ist wie die Bytelänge des längsten Werts in dieser Spalte.

 Note

Wenn ein Wert länger als der angegebene Wert ist, tritt ein Fehler auf und die Verschlüsselung schlägt fehl. `length`

Wenn Sie für eine Spalte eine Auffüllung wünschen und die maximale Bytelänge der Daten in dieser Spalte nicht bekannt ist, verwenden Sie `max Padding`. In diesem Auffüllmodus werden alle Daten auf die Länge des längsten Werts zuzüglich zusätzlicher Byte aufgefüllt. `length`

 Note

Möglicherweise möchten Sie Daten stapelweise verschlüsseln oder Ihre Tabellen regelmäßig mit neuen Daten aktualisieren. Beachten Sie, dass beim `max` Auffüllen die Einträge auf die Länge (plus `length` Byte) des längsten Klartexteintrags in einem bestimmten Stapel aufgefüllt werden. Das bedeutet, dass die Länge des Chiffretextes von Stapel zu Stapel variieren kann. Wenn Sie also die maximale Bytelänge für eine Spalte kennen, sollten Sie stattdessen die Option verwenden. `fixed max`

CleartextSpalten

CleartextSpalten sind Spalten, die nicht kryptografisch für die Verwendung in JOIN Oder-Anweisungen geschützt sind. `SELECT`

CleartextSpalten können in jedem Teil der SQL-Abfrage verwendet werden.

Kryptografische Rechenparameter

[Kryptografische Rechenparameter sind für Kollaborationen verfügbar, bei denen Cryptographic Computing for Clean Rooms \(C3R\) beim Erstellen einer Kollaboration verwendet wird.](#) Sie können eine Kollaboration entweder mithilfe der AWS Clean Rooms Konsole oder mithilfe der API-Operation erstellen. `CreateCollaboration` In der Konsole können Sie Werte für die Parameter unter Verschlüsselungsparameter festlegen, nachdem Sie die Option Kryptografisches Rechnen Support aktiviert haben. Weitere Informationen finden Sie unter den folgenden Themen.

Themen

- [Parameter „Spalten zulassencleartext“](#)

- [Parameter „Duplikate zulassen“](#)
- [Parameter „Zulassen JOIN von Spalten mit unterschiedlichen Namen“](#)
- [Parameter „NULLWerte beibehalten“](#)

Parameter „Spalten zulassencleartext“

In der Konsole können Sie beim [Erstellen einer Kollaboration](#) den Parameter cleartextSpalten zulassen festlegen, um anzugeben, ob cleartext Daten in einer Tabelle mit verschlüsselten Daten zulässig sind.

In der folgenden Tabelle werden die Werte für den Parameter Allow cleartext columns beschrieben.

Parameterwert	Description
Nein	CleartextSpalten sind in der verschlüsselten Tabelle nicht zulässig. Alle Daten sind kryptografisch geschützt.
Ja	CleartextSpalten sind in der verschlüsselten Tabelle zulässig. CleartextSpalten sind nicht kryptografisch geschützt und werden als aufgenommen. cleartext Sie sollten sich notieren, was die Daten Ihrer Zeilen über die anderen cleartext Daten in der Tabelle aussagen könnten. Um SUM oder AVG für bestimmte Spalten auszuführen, müssen sich die Spalten in cleartext befinden.

Mithilfe der CreateCollaboration API-Operation können Sie für den dataEncryptionMetadata Parameter den Wert allowCleartext auf true oder festlegenfalse. Weitere Informationen zu API-Vorgängen finden Sie in der [AWS Clean Rooms API-Referenz](#).

CleartextSpalten entsprechen Spalten, die cleartext im tabellenspezifischen Schema als Spalten klassifiziert sind. Die Daten in diesen Spalten sind nicht verschlüsselt und können auf beliebige Weise verwendet werden. CleartextSpalten können nützlich sein, wenn die Daten nicht sensibel sind, and/or wenn mehr Flexibilität erforderlich ist, als es eine verschlüsselte sealed Spalte oder fingerprint Spalte zulässt.

Parameter „Duplikate zulassen“

In der Konsole können Sie beim [Erstellen einer Kollaboration](#) den Parameter Duplikate zulassen festlegen, um anzugeben, ob für JOIN Abfragen verschlüsselte Spalten doppelte Nichtwerte enthalten können. NULL

Important

Die Parameter „Duplikate [zulassenJOIN](#)“, „[Spalten mit unterschiedlichen Namen](#) zulassen“ und „[NULLWerte beibehalten](#)“ haben unterschiedliche, aber verwandte Auswirkungen.

In der folgenden Tabelle werden die Werte für den Parameter Duplikate zulassen beschrieben.

Parameterwert	Description
Nein	Wiederholte Werte sind in einer fingerprint Spalte nicht zulässig. Alle Werte in einer einzelnen fingerprint Spalte müssen eindeutig sein.
Ja	Wiederholte Werte sind in einer fingerprint Spalte zulässig. Wenn Sie Spalten mit wiederholten Werten verbinden müssen, setzen Sie diesen Wert auf Ja. Wenn diese Option auf Ja gesetzt ist, können Häufigkeitsmuster, die in den fingerprint Spalten der C3R-Tabelle oder der Ergebnisse erscheinen, zusätzliche Informationen über die Struktur der cleartext Daten enthalten.

Mithilfe der CreateCollaboration API-Operation können Sie für den dataEncryptionMetadata Parameter den Wert allowDuplicates auf true oder false festlegen. Weitere Informationen zu API-Vorgängen finden Sie in der [AWS Clean Rooms API-Referenz](#).

Wenn verschlüsselte Daten in JOIN Abfragen verwendet werden müssen, verlangt der C3R-Verschlüsselungscient standardmäßig, dass diese Spalten keine doppelten Werte enthalten. Diese Anforderung ist ein Versuch, den Datenschutz zu verbessern. Dieses Verhalten kann dazu beitragen,

dass wiederholte Muster in den Daten nicht beobachtbar sind. Wenn Sie jedoch mit verschlüsselten Daten in JOIN Abfragen arbeiten möchten und sich keine Gedanken über doppelte Werte machen, kann der Parameter Duplikate zulassen diese konservative Prüfung deaktivieren.

Parameter „Zulassen JOIN von Spalten mit unterschiedlichen Namen“

In der Konsole können Sie beim [Erstellen einer Kollaboration](#) den Parameter Spalten mit unterschiedlichen Namen zulassen JOIN festlegen, um anzugeben, ob JOIN Anweisungen zwischen Spalten mit unterschiedlichen Namen unterstützt werden.

Weitere Informationen finden Sie unter [Normalisierung der Namen der Spaltenüberschriften](#).

In der folgenden Tabelle werden die Werte für den Parameter Zulassen JOIN von Spalten mit unterschiedlichen Namen beschrieben.

Parameterwert	Description
Nein	Verknüpfungen von fingerprint Spalten mit unterschiedlichen Namen werden nicht unterstützt. JOINAnweisungen liefern nur genaue Ergebnisse für Spalten mit demselben Namen.  Important Der Wert „Nein“ erhöht die Informationssicherheit, erfordert jedoch, dass sich die Kollaborationsteilnehmer zuvor über die Spaltennamen einigen. Wenn zwei Spalten unterschiedliche Namen haben, wenn sie als fingerprint Spalten verschlüsselt sind und Spalten mit unterschiedlichen Namen zulassen JOIN auf Nein gesetzt ist, führen JOIN Anweisungen zu diesen Spalten zu keinen Ergebnissen. Das liegt daran, dass sie nach der Verschlüsselung keine Werte gemeinsam nutzen.
Ja	Verknüpfungen von fingerprint Spalten mit unterschiedlichen Namen werden unterstützt. Für zusätzliche Flexibilität können Benutzer diesen Wert auf Ja setzen, sodass JOIN Aussagen zu Spalten unabhängig von deren Spaltennamen möglich sind.

Parameterwert	Description
	Wenn dieser Wert auf Ja gesetzt ist, berücksichtigt der C3R-Verschlüsselungsclient den Spaltennamen beim Schutz von fingerprint Spalten nicht. Daher sind gemeinsame Werte in verschiedenen fingerprint Spalten in der C3R-Tabelle beobachtbar. Wenn eine Zeile beispielsweise denselben verschlüsselten JOIN Wert sowohl in einer Spalte als auch in einer City Spalte hat, kann es sinnvoll sein, daraus zu schließen, dass dieser Wert ist. State New York

Mithilfe der CreateCollaboration API-Operation können Sie für den dataEncryptionMetadata Parameter den Wert allowJoinsOnColumnsWithDifferentNames auf true oder false festlegen. Weitere Informationen zu API-Vorgängen finden Sie in der [AWS Clean Rooms API-Referenz](#).

Standardmäßig wird die fingerprint Spaltenverschlüsselung durch die targetHeader für diese Spalte eingestellte Einstellung beeinflusst [Schritt 4: Generieren Sie ein Verschlüsselungsschema für eine tabellarische Datei](#). Daher hat derselbe cleartext Wert in jeder fingerprint Spalte, für die er verschlüsselt ist, unterschiedliche verschlüsselte Darstellungen.

Dieser Parameter kann in einigen Fällen nützlich sein, um die Inferenz von cleartext Werten zu verhindern. StateEs kann beispielsweise verwendet werden, wenn derselbe verschlüsselte Wert in fingerprint Spalten angezeigt City wird, um vernünftigerweise auf den Wert schließen zu können. New York Die Verwendung dieses Parameters erfordert jedoch eine zusätzliche Abstimmung im Voraus, sodass alle Spalten, die in Abfragen verknüpft werden sollen, gemeinsame Namen haben.

Sie können den Parameter Spalten mit unterschiedlichen Namen zulassen JOIN verwenden, um diese Einschränkung zu lockern. Wenn der Parameterwert auf gesetzt istYes, können alle Spalten, für die verschlüsselt wurdeJOIN, unabhängig vom Namen zusammen verwendet werden.

Parameter „NULLWerte beibehalten“

In der Konsole können Sie beim [Erstellen einer Kollaboration](#) den Parameter NULLWerte beibehalten so einstellen, dass für diese Spalte kein Wert vorhanden ist.

In der folgenden Tabelle werden die Werte für den Parameter NULLWerte beibehalten beschrieben.

Parameterwert	Description
Nein	NULLWerte werden nicht beibehalten. NULLWerte werden nicht wie NULL in einer verschlüsselten Tabelle angezeigt . NULLWerte werden in einer C3R-Tabelle als eindeutige Zufallswerte angezeigt.
Ja	NULLWerte werden beibehalten. NULLWerte werden wie NULL in einer verschlüsselten Tabelle angezeigt. Wenn Sie die NULL SQL-Wertesemantik benötigen, können Sie diesen Wert auf Ja setzen. Daher werden NULL Einträge wie NULL in der C3R-Tabelle angezeigt, unabhängig davon, ob die Spalte verschlüsselt ist und unabhängig von der Parametereinstellung für Duplikate zulassen.

Mithilfe der `CreateCollaboration` API-Operation können Sie für den `dataEncryptionMetadata` Parameter den Wert auf `true` oder `false` festlegen. Weitere Informationen zu API-Vorgängen finden Sie in der [AWS Clean Rooms API-Referenz](#).

Wenn der Parameter `NULLWerte beibehalten` für die Zusammenarbeit auf `Nein` gesetzt ist:

1. NULLEinträge in `cleartext` Spalten sind unverändert.
2. NULLEinträge in verschlüsselten `fingerprint` Spalten werden als Zufallswerte verschlüsselt, um ihren Inhalt zu verbergen. Bei der Verknüpfung zu einer verschlüsselten Spalte mit NULL Einträgen in der `cleartext` Spalte ergeben sich keine Treffer für einen der NULL Einträge. Es werden keine Treffer erzielt, da sie jeweils ihren eigenen, eindeutigen zufälligen Inhalt erhalten.
3. NULLEinträge in verschlüsselten `sealed` Spalten sind verschlüsselt.

Wenn der Wert des Parameters `NULLWerte beibehalten` für die Kollaboration auf `Ja` gesetzt ist, bleiben die NULL Einträge aus allen Spalten unverändert, NULL unabhängig davon, ob die Spalte verschlüsselt ist.

Der Parameter `NULLWerte beibehalten` ist nützlich in Szenarien wie der Datenanreicherung, in denen Sie fehlende Informationen weitergeben möchten, ausgedrückt als NULL. Der Parameter `NULLWerte beibehalten` ist auch im `fingerprint` oder `HMAC`-Format nützlich, wenn Sie NULL Werte in der gewünschten Spalte haben oder. `JOIN GROUP BY`

Wenn der Wert der Parameter Duplikate zulassen und NULLWerte beibehalten auf Nein gesetzt ist, führt das Vorhandensein von mehr als einem NULL Eintrag in einer fingerprint Spalte zu einem Fehler und die Verschlüsselung wird gestoppt. Wenn der Wert eines der beiden Parameter auf Ja gesetzt ist, tritt kein solcher Fehler auf.

Optionale Flags in Cryptographic Computing für Clean Rooms

In den folgenden Abschnitten werden die optionalen Flags beschrieben, die Sie festlegen können, wenn Sie [Daten mit dem C3R-Verschlüsselungsclient verschlüsseln](#), um tabellarische Dateien anzupassen und zu testen.

Themen

- [--csvInputNULLValue-Flag](#)
- [--csvOutputNULLValue-Flag](#)
- [--enableStackTraces-Flag](#)
- [--dryRun-Flag](#)
- [--tempDir-Flag](#)

--csvInputNULLValue-Flag

Sie können das `--csvInputNULLValue` Flag verwenden, um benutzerdefinierte Kodierungen für NULL Einträge in den Eingabedaten anzugeben, wenn Sie Daten mit dem [C3R-Verschlüsselungsclient verschlüsseln](#).

In der folgenden Tabelle werden die Verwendung und die Parameter dieses Flags zusammengefasst.

Usage	Parameters
Optional. Benutzer können benutzerdefinierte Kodierungen für NULL Einträge in den Eingabedaten angeben.	Benutzerdefinierte Kodierung von NULL Werten in der CSV-Eingabedatei

Ein NULL Eintrag ist ein Eintrag, der als inhaltslos angesehen wird, insbesondere im Zusammenhang mit einem umfangreicheren Tabellenformat wie einer SQL-Tabelle. Obwohl .csv diese Charakterisierung aus historischen Gründen nicht ausdrücklich unterstützt, ist es üblich, einen

leeren Eintrag, der nur Leerraum enthält, als solche zu betrachten. NULL Daher ist dies das Standardverhalten des C3R-Verschlüsselungsclients und kann nach Bedarf angepasst werden.

--csvOutputNULLValue-Flag

Sie können das --csvOutputNULLValue Flag verwenden, um benutzerdefinierte Kodierungen für NULL Einträge in den Ausgabedaten anzugeben, wenn Sie [Daten mit dem C3R-Verschlüsselungsclient](#) verschlüsseln.

In der folgenden Tabelle werden die Verwendung und die Parameter dieses Flags zusammengefasst.

Usage	Parameters
Optional. Benutzer können in der generierten Ausgabedatei für NULL Einträge benutzerdefinierte Kodierungen angeben.	Benutzerdefinierte Kodierung von NULL Werten in der CSV-Ausgabedatei

Ein NULL Eintrag ist ein Eintrag, der als inhaltslos angesehen wird, insbesondere im Zusammenhang mit einem umfangreicheren Tabellenformat wie einer SQL-Tabelle. Obwohl .csv diese Charakterisierung aus historischen Gründen nicht ausdrücklich unterstützt, ist es üblich, einen leeren Eintrag, der nur Leerraum enthält, als solche zu betrachten. NULL Daher ist dies das Standardverhalten des C3R-Verschlüsselungsclients und kann nach Bedarf angepasst werden.

--enableStackTraces-Flag

Wenn Sie [Daten mit dem C3R-Verschlüsselungsclient verschlüsseln](#), verwenden Sie das --enableStackTraces Flag, um zusätzliche Kontextinformationen für die Fehlerberichterstattung bereitzustellen, wenn C3R auf einen Fehler stößt.

AWS sammelt keine Fehler. Wenn Sie auf einen Fehler stoßen, verwenden Sie den Stack-Trace, um den Fehler selbst zu beheben, oder senden Sie den Stack-Trace an, um Support Unterstützung zu erhalten.

In der folgenden Tabelle werden die Verwendung und die Parameter dieses Flags zusammengefasst.

Usage	Parameters
Optional. Wird verwendet, um zusätzliche Kontextinformationen für die Fehlerberichtersta	Keine

Usage	Parameters
ttung bereitzustellen, wenn der C3R-Versc hlüsselungsclient auf einen Fehler stößt.	

--dryRun-Flag

Die Befehle des [C3R-Verschlüsselungsclients verschlüsseln](#) und [entschlüsseln](#) enthalten ein optionales Flag. --dryRun Das Flag verwendet alle vom Benutzer angegebenen Argumente und überprüft sie auf Gültigkeit und Konsistenz.

Sie können das --dryRun Flag verwenden, um zu überprüfen, ob Ihre Schemadatei gültig ist und mit der entsprechenden Eingabedatei konsistent ist.

In der folgenden Tabelle werden die Verwendung und die Parameter dieses Flags zusammengefasst.

Usage	Parameters
Optional. Bewirkt, dass der C3R-Versc hlüsselungsclient Parameter analysiert und Dateien überprüft, aber keine Verschlüsselung oder Entschlüsselung durchführt.	Keine

--tempDir-Flag

Möglicherweise möchten Sie ein temporäres Verzeichnis verwenden, da verschlüsselte Dateien je nach ihren Einstellungen manchmal größer sein können als unverschlüsselte Dateien. Datensätze müssen außerdem pro Kollaboration verschlüsselt werden, damit sie korrekt funktionieren.

Wenn Sie [Daten mit C3R verschlüsseln](#), verwenden Sie das --tempDir Flag, um den Speicherort anzugeben, an dem temporäre Dateien während der Verarbeitung der Eingabe erstellt werden können.

In der folgenden Tabelle werden die Verwendung und die Parameter dieses Flags zusammengefasst.

Usage	Parameters
Benutzer können den Speicherort angeben, an dem temporäre Dateien während der Verarbeitung der Eingabe erstellt werden können.	Standardmäßig wird das temporäre Systemverzeichnis verwendet.

Abfragen mit Cryptographic Computing für Clean Rooms

Dieses Thema enthält Informationen zum Schreiben von Abfragen, die Datentabellen verwenden, die mithilfe von Cryptographic Computing verschlüsselt wurden. Clean Rooms

Themen

- [Abfragen, die sich wie folgt verzweigen NULL](#)
- [Zuordnen einer Quellspalte zu mehreren Zielspalten](#)
- [Verwenden Sie dieselben Daten für beide JOIN SELECT Abfragen](#)

Abfragen, die sich wie folgt verzweigen NULL

Eine Abfrageverzweigung für eine NULL Anweisung zu haben, bedeutet, eine Syntax wie zu verwenden `IF x IS NULL THEN 0 ELSE 1`.

Abfragen können sich immer auf NULL Anweisungen in cleartext Spalten verzweigen.

Abfragen können sich nur dann auf NULL Anweisungen in sealed Spalten und fingerprint Spalten stützen, wenn der Wert des Parameters `NULL-Werte beibehalten (preserveNulls)` auf `gesetzt ist true`.

Abfragen, die gegen diese Einschränkungen verstoßen, können zu falschen Ergebnissen führen.

Zuordnen einer Quellspalte zu mehreren Zielspalten

Eine Quellspalte kann mehreren Zielspalten zugeordnet werden. Beispielsweise möchten Sie vielleicht beides JOIN und SELECT eine Spalte gleichzeitig verwenden.

Weitere Informationen finden Sie unter [Verwenden Sie dieselben Daten für beide JOIN SELECT Abfragen](#).

Verwenden Sie dieselben Daten für beide JOIN SELECT Abfragen

Wenn die Daten in einer Spalte nicht vertraulich sind, können sie in einer cleartext Zielspalte erscheinen, sodass sie für jeden Zweck verwendet werden können.

Wenn Daten in einer Spalte vertraulich sind JOIN und sowohl für SELECT Abfragen als auch verwendet werden müssen, ordnen Sie diese Quellspalte zwei Zielspalten in der Ausgabedatei zu. Eine Spalte ist mit der type als fingerprint Spalte verschlüsselt, und eine Spalte ist mit der type als versiegelte Spalte verschlüsselt. Die interaktive Schemagenerierung des C3R-Verschlüsselungsclients schlägt Header-Suffixe von und vor. `_fingerprint` `_sealed` Diese Header-Suffixe können eine nützliche Konvention sein, um solche Spalten schnell zu unterscheiden.

Richtlinien für den C3R-Verschlüsselungsclient

Der C3R-Verschlüsselungsclient ist ein Tool, mit dem Unternehmen sensible Daten zusammenführen können, um aus Datenanalysen neue Erkenntnisse zu gewinnen. Das Tool schränkt kryptografisch ein, was von jeder Partei und AWS während des Prozesses gelernt werden kann. Dies ist zwar von entscheidender Bedeutung, aber der Prozess der kryptografischen Sicherung von Daten kann zu einem erheblichen Mehraufwand sowohl in Bezug auf Rechen- als auch Speicherressourcen führen. Daher ist es wichtig, die Kompromisse bei der Verwendung der einzelnen Einstellungen zu verstehen und zu verstehen, wie die Einstellungen optimiert und gleichzeitig die gewünschten kryptografischen Garantien beibehalten werden können. Dieses Thema konzentriert sich auf die Auswirkungen verschiedener Einstellungen im C3R-Verschlüsselungsclient und in den Schemas auf die Leistung.

Alle Verschlüsselungseinstellungen des C3R-Verschlüsselungsclients bieten unterschiedliche kryptografische Garantien. Die Einstellungen auf Kollaborationsebene sind standardmäßig am sichersten. Durch die Aktivierung zusätzlicher Funktionen bei gleichzeitiger Schaffung einer Zusammenarbeit werden die Datenschutzgarantien geschwächt, sodass Aktivitäten wie Frequenzanalysen anhand des Chiffretextes durchgeführt werden können. Weitere Informationen darüber, wie diese Einstellungen verwendet werden und welche Auswirkungen sie haben, finden Sie unter [the section called “Kryptografisches Rechnen”](#)

Themen

- [Auswirkungen auf die Leistung von Spaltentypen](#)
- [Behebung unerwarteter Zunahmen der Chiffretext-Größe](#)

Auswirkungen auf die Leistung von Spaltentypen

C3R verwendet drei Spaltentypen: cleartextfingerprint, undsealed. Jeder dieser Spaltentypen bietet unterschiedliche kryptografische Garantien und hat unterschiedliche Verwendungszwecke. In den folgenden Abschnitten werden die Auswirkungen des Spaltentyps auf die Leistung sowie die Auswirkungen der einzelnen Einstellungen auf die Leistung erörtert.

Themen

- [CleartextSpalten](#)
- [FingerprintSpalten](#)
- [SealedSpalten](#)

CleartextSpalten

CleartextSpalten werden gegenüber ihrem ursprünglichen Format nicht verändert und in keiner Weise kryptografisch verarbeitet. Dieser Spaltentyp kann nicht konfiguriert werden und beeinträchtigt weder die Speicher- noch die Rechenleistung.

FingerprintSpalten

FingerprintSpalten sollen verwendet werden, um Daten aus mehreren Tabellen zu verbinden. Zu diesem Zweck muss die resultierende Chiffretextgröße immer dieselbe sein. Diese Spalten werden jedoch von den Einstellungen auf Kollaborationsebene beeinflusst. FingerprintSpalten können sich unterschiedlich stark auf die Größe der Ausgabedatei auswirken, je nachdem, was in der Eingabe cleartext enthalten ist.

Themen

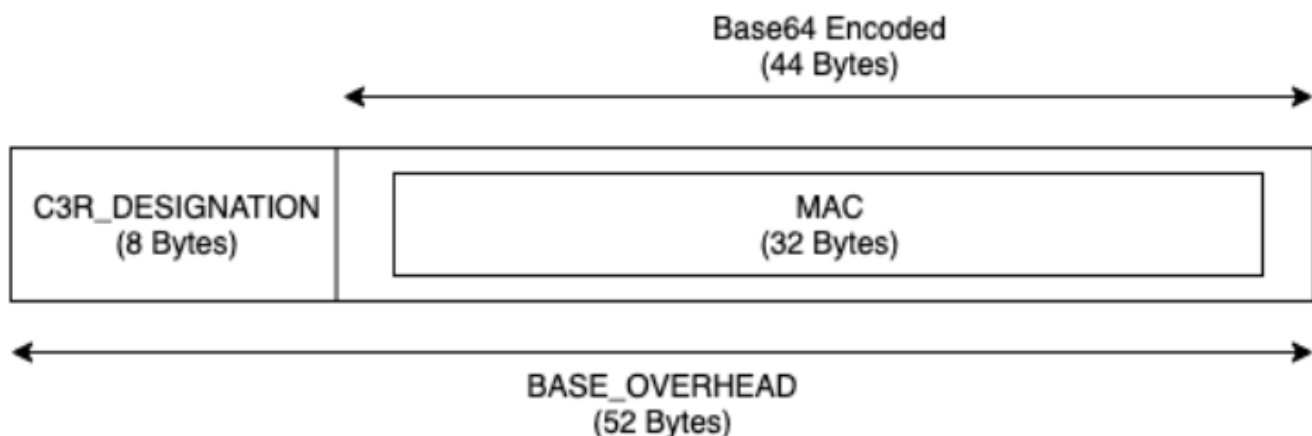
- [Basis-Overhead für fingerprint Spalten](#)
- [Einstellungen für die Zusammenarbeit für fingerprint Spalten](#)
- [Beispieldaten für eine fingerprint Spalte](#)
- [Problembehandlung bei fingerprint Spalten](#)

Basis-Overhead für fingerprint Spalten

Es gibt einen Basisgemeinkosten für fingerprint Spalten. Dieser Overhead ist konstant und ersetzt die Größe der cleartext Bytes.

Die Daten in den fingerprint Spalten werden mithilfe einer Hash-basierten HMAC-Funktion (Message Authentication Code) kryptografisch verarbeitet, die die Daten in einen 32-Byte-Nachrichtenauthentifizierungscode (MAC) umwandelt. Diese Daten werden dann über einen Base64-Encoder verarbeitet, wodurch die Bytegröße um etwa 33 Prozent erhöht wird. Ihm wird eine 8-Byte-C3R-Bezeichnung vorangestellt, um den Spaltentyp zu bezeichnen, zu dem die Daten gehören, und die Client-Version, die sie erzeugt hat. Das Endergebnis ist 52 Byte. Dieses Ergebnis wird dann mit der Zeilenanzahl multipliziert, um den gesamten Basis-Overhead zu erhalten (verwenden Sie die Anzahl der gesamten null Nichtwerte, wenn der Wert auf „true“ gesetzt `preserveNulls` ist).

Die folgende Abbildung zeigt, wie $BASE\_OVERHEAD = C3R\_DESIGNATION + (MAC * 1.33)$



Der ausgegebene Chiffretext in den fingerprint Spalten wird immer 52 Byte lang sein. Dies kann zu einer erheblichen Verringerung des Speicherplatzes führen, wenn die cleartext Eingabedaten im Durchschnitt mehr als 52 Byte umfassen (z. B. vollständige Straßenadressen). Dies kann eine erhebliche Speichererweiterung bedeuten, wenn die cleartext Eingabedaten im Durchschnitt weniger als 52 Byte enthalten (z. B. aufgrund des Alters des Kunden).

Einstellungen für die Zusammenarbeit für fingerprint Spalten

preserveNulls-Einstellung

Wenn die Einstellung `preserveNulls` auf Kollaborationsebene `false` (Standard) lautet, wird jeder `null` Wert durch eindeutige, zufällige 32 Byte ersetzt und so verarbeitet, als ob dies nicht der Fall wäre. Das Ergebnis ist, dass jeder `null` Wert jetzt 52 Byte groß ist. Dies kann zu erheblichen Speicheranforderungen für Tabellen führen, die nur sehr wenige Daten enthalten, verglichen mit der Einstellung, bei der `null` Werte übergeben werden. `true null`

Wenn Sie die Datenschutzgarantien dieser Einstellung nicht benötigen und es vorziehen, `null` Werte in Ihren Datensätzen beizubehalten, aktivieren Sie die `preserveNulls` Einstellung bei

der Erstellung der Kollaboration. Die `preserveNulls` Einstellung kann nach der Erstellung der Kollaboration nicht mehr geändert werden.

Beispieldaten für eine fingerprint Spalte

Im Folgenden finden Sie ein Beispiel für Eingabe- und Ausgabedaten für eine fingerprint Spalte mit Einstellungen zur Reproduktion. Andere Einstellungen auf Kollaborationsebene wirken sich wie `allowCleartext` und `allowDuplicates` nicht auf die Ergebnisse aus und können so eingestellt werden, als `true` `false` ob versucht wird, lokal zu reproduzieren.

Beispiel für ein geteiltes Geheimnis: `wJa1rXUtnFEMI/K7MDENG/bPxRfiCYEXAMPLEKEY`

Beispiel für eine Kollaborations-ID: `a1b2c3d4-5678-90ab-cdef-EXAMPLE11111`

`allowJoinsOnColumnsWithDifferentNames`: `True` Diese Einstellung hat keinen Einfluss auf die Leistungs- oder Speicheranforderungen. Diese Einstellung macht die Wahl des Spaltennamens jedoch irrelevant, wenn die Werte in den folgenden Tabellen wiedergegeben werden.

Beispiel 1

Input	<code>null</code>
<code>preserveNulls</code>	<code>TRUE</code>
Ausgabe	<code>null</code>
Deterministisch	<code>Yes</code>
Eingabe-Bytes	<code>0</code>
Ausgabe-Bytes	<code>0</code>

Beispiel 2

Input	<code>null</code>
<code>preserveNulls</code>	<code>FALSE</code>
Ausgabe	<code>01:hmac:31kFjthvV3IUu6mMvFc1a+XAHwgw/E1m0q4p3Yg25kk=</code>

Deterministisch	No
Eingabe-Bytes	0
Ausgabe-Bytes	52

Beispiel 3

Input	empty string
preserveNulls	-
Ausgabe	01:hmac:oKTgi3Gba+eUb3JteSz 2EMgXUkF1WgM77UP0Ydw5kPQ=
Deterministisch	Yes
Eingabe-Bytes	0
Ausgabe-Bytes	52

Beispiel 4

Input	abcdefghijklmnopqrstuvwxyz
preserveNulls	-
Ausgabe	01:hmac:kU/IqwG7FMmzzshr0B9 scomE0UJUEE7j9keTctp1Gww=
Deterministisch	Yes
Eingabe-Bytes	26
Ausgabe-Bytes	52

Beispiel 5

Input	abcdefghijklmnopqrstuvwxyzA BCDEFGHIJKLMNOPQRSTUVWXYZ01 23456789
preserveNulls	-
Ausgabe	01:hmac:ks3htnQbw2vdhCRFF6J NzW5LMndJaHG57uvE26mBtSs=
Deterministisch	Yes
Eingabe-Bytes	62
Ausgabe-Bytes	52

Problembehandlung bei fingerprint Spalten

Warum ist der Chiffretext in meinen fingerprint Spalten um ein Vielfaches größer als cleartext der Inhalt?

Der Chiffretext in einer fingerprint Spalte ist immer 52 Byte lang. Wenn Ihre Eingabedaten klein waren (z. B. das Alter der Kunden), wurde sie deutlich größer. Dies kann auch passieren, wenn die `preserveNulls` Einstellung auf `false` gesetzt ist.

Warum ist der Chiffretext in meinen fingerprint Spalten um ein Vielfaches kleiner als cleartext der Inhalt?

Der Chiffretext in einer fingerprint Spalte ist immer 52 Byte lang. Wenn Ihre Eingabedaten umfangreich sind (z. B. die vollständigen Straßenadressen von Kunden), ist die Größe deutlich geringer.

Woher weiß ich, ob ich die kryptografischen Garantien von benötige? **preserveNulls**

Leider lautet die Antwort, dass es darauf ankommt. Zumindest [the section called "Parameters"](#) sollte überprüft werden, wie die `preserveNulls` Einstellung Ihre Daten schützt. Wir empfehlen Ihnen jedoch, die Datenverarbeitungsanforderungen Ihres Unternehmens und alle Verträge, die für die jeweilige Zusammenarbeit gelten, zu beachten.

Warum muss ich den Overhead von Base64 auf mich nehmen?

Um die Kompatibilität mit tabellarischen Dateiformaten wie CSV zu gewährleisten, ist eine Base64-Kodierung erforderlich. Zwar unterstützen einige Dateiformate wie z. Parquet B. binäre Darstellungen von Daten, es ist jedoch wichtig, dass alle Teilnehmer einer Zusammenarbeit Daten auf die gleiche Weise darstellen, um korrekte Abfrageergebnisse zu gewährleisten.

SealedSpalten

SealedSpalten sollen für die Übertragung von Daten zwischen Mitgliedern einer Kollaboration verwendet werden. Der Geheimtext in diesen Spalten ist nicht deterministisch und hat je nach Konfiguration der Spalten erhebliche Auswirkungen sowohl auf die Leistung als auch auf den Speicherplatz. Diese Spalten können individuell konfiguriert werden und haben oft den größten Einfluss auf die Leistung des C3R-Verschlüsselungsclients und die daraus resultierende Größe der Ausgabedatei.

Themen

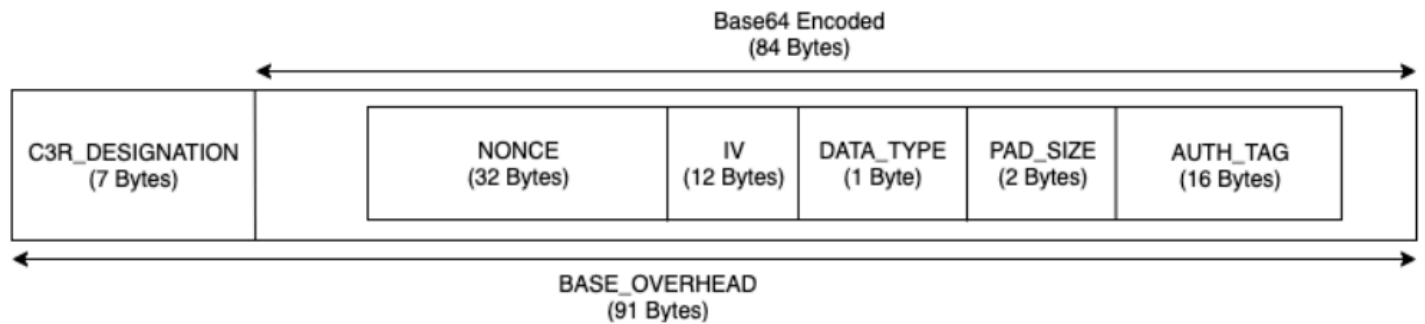
- [Basis-Overhead für Spalten sealed](#)
- [Einstellungen für die Zusammenarbeit für sealed Spalten](#)
- [sealedSpalten mit Schemaeinstellungen: Polstertypen](#)
- [Beispieldaten für eine Spalte sealed](#)
- [sealedSpalten zur Problembehandlung](#)

Basis-Overhead für Spalten sealed

Es gibt einen Basisgemeinkosten für sealed Spalten. Dieser Overhead ist konstant und kommt zu der Größe der Byte cleartext und der Füllmenge (falls vorhanden) hinzu.

Vor jeder Verschlüsselung wird den Daten in den sealed Spalten ein 1-Byte-Zeichen vorangestellt, das angibt, welcher Datentyp enthalten ist. Wenn Padding ausgewählt ist, werden die Daten aufgefüllt und mit 2 Byte angehängt, die die Pad-Größe angeben. Nachdem diese Byte hinzugefügt wurden, werden die Daten mithilfe von AES-GCM kryptografisch verarbeitet und mit IV (12 Byte), (32 Byte) und nonce (16 Byte) gespeichert. Auth Tag Diese Daten werden dann durch einen Base64-Encoder verarbeitet, wodurch die Bytegröße um etwa 33 Prozent erhöht wird. Den Daten wird eine 7-Byte-C3R-Bezeichnung vorangestellt, um anzugeben, zu welchem Spaltentyp die Daten gehören und welche Client-Version verwendet wurde, um sie zu erzeugen. Das Ergebnis ist ein endgültiger Basisaufwand von 91 Byte. Dieses Ergebnis kann dann mit der Zeilenanzahl multipliziert werden, um den gesamten Basis-Overhead zu erhalten (verwenden Sie die Anzahl der Gesamtwerte ungleich Null, wenn der Wert auf true gesetzt `preserveNulls` ist).

Die folgende Abbildung zeigt, wie $BASE\_OVERHEAD = C3R\_DESIGNATION + ((NONCE + IV + DATA\_TYPE + PAD\_SIZE + AUTH\_TAG) * 1.33)$



Einstellungen für die Zusammenarbeit für sealed Spalten

preserveNulls-Einstellung

Wenn die Einstellung auf Kollaborationsebene auf `false` (Standard) gesetzt `preserveNulls` ist, ist jeder `null` Wert einmalig, hat 32 Byte Zufallswerte und wird so verarbeitet, als ob dies nicht der Fall wäre. Das Ergebnis ist, dass jeder `null` Wert jetzt 91 Byte groß ist (mehr, wenn er aufgefüllt wird). Dies kann zu erheblichen Speicheranforderungen für Tabellen führen, die nur sehr wenige Daten enthalten, als wenn diese Einstellung aktiviert ist `true` und `null` Werte als `null` übergeben werden.

Wenn Sie die Datenschutzgarantien dieser Einstellung nicht benötigen und es vorziehen, `null` Werte in Ihren Datensätzen beizubehalten, aktivieren Sie die `preserveNulls` Einstellung bei der Erstellung der Kollaboration. Die `preserveNulls` Einstellung kann nach der Erstellung der Kollaboration nicht mehr geändert werden.

sealedSpalten mit Schemaeinstellungen: Polstertypen

Themen

- [Pad-Typ von none](#)
- [Pad-Typ von fixed](#)
- [Pad-Typ von max](#)

Pad-Typ von **none**

Durch die Auswahl des Padtyps von `none` wird dem zuvor beschriebenen Grund-Overhead keine Polsterung hinzugefügt cleartext und auch kein zusätzlicher Overhead hinzugefügt. Wenn keine Polsterung vorhanden ist, ergibt sich die platzsparendste Ausgabegröße. Es bietet jedoch nicht die

gleichen Datenschutzgarantien wie die Polstertypen **fixed** und **max**. Das liegt daran, dass die Größe des Basistextes anhand der Größe des Chiffretextes erkennbar **cleartext** ist.

Pad-Typ von **fixed**

Die Auswahl des Pad-Typs von **fixed** dient dem Schutz der Privatsphäre, um die Länge der in einer Spalte enthaltenen Daten zu verbergen. Dies wird erreicht, indem das gesamte Feld mit dem angegebenen Feld ausgefüllt wird **cleartext**, **pad_length** bevor es verschlüsselt wird. Alle Daten, die diese Größe überschreiten, führen dazu, dass der C3R-Verschlüsselungsclient fehlschlägt.

Da das Padding **cleartext** vor der Verschlüsselung hinzugefügt wird, hat AES-GCM eine 1-zu-1-Zuordnung von Chiffretext-Bytes. **cleartext** Durch die Base64-Kodierung werden 33 Prozent hinzukommen. Der zusätzliche Speicheraufwand des Padding kann berechnet werden, indem die durchschnittliche Länge von **cleartext** vom Wert von **pad_length** subtrahiert und mit 1,33 multipliziert wird. Das Ergebnis ist der durchschnittliche Mehraufwand für das Auffüllen pro Datensatz. Dieses Ergebnis kann dann mit der Anzahl der Zeilen multipliziert werden, um den gesamten Auffüllaufwand zu erhalten (verwenden Sie die Anzahl der gesamten **null** Nichtwerte, falls **preserveNulls** auf **true** gesetzt).

$$PADDING\_OVERHEAD = (PAD\_LENGTH - AVG\_CLEARTEXT\_LENGTH) * 1.33 * ROW\_COUNT$$

Es wird empfohlen, das Minimum auszuwählen **pad_length**, das den größten Wert in einer Spalte umfasst. Wenn der größte Wert beispielsweise 50 Byte beträgt, ist ein Wert **pad_length** von 50 ausreichend. Ein höherer Wert erhöht nur zusätzlichen Speicheraufwand.

Eine feste Polsterung erhöht keinen nennenswerten Rechenaufwand.

Pad-Typ von **max**

Die Auswahl des Pad-Typs von **max** dient dem Schutz der Privatsphäre, um die Länge der in einer Spalte enthaltenen Daten zu verbergen. Dies wird erreicht, indem der **cleartext** gesamte Wert mit dem größten Wert in der Spalte und dem zusätzlichen Wert aufgefüllt wird, **pad_length** bevor die Spalte verschlüsselt wird. Im Allgemeinen bietet **max** das Auffüllen die gleiche Sicherheit wie das Auffüllen **fixed** eines einzelnen Datensatzes, ermöglicht jedoch, dass der größte **cleartext** Wert in der Spalte nicht bekannt ist. Das Auffüllen **max** bietet jedoch möglicherweise nicht die gleichen Datenschutzgarantien wie **fixed** das Auffüllen bei Aktualisierungen, da der größte Wert in den einzelnen Datensätzen unterschiedlich sein kann.

Wir empfehlen, dass Sie bei der Verwendung von Padding einen zusätzlichen Wert **pad_length** von 0 wählen. **max** Bei dieser Länge werden alle Werte so aufgefüllt, dass sie dieselbe Größe wie der größte Wert in der Spalte haben. Ein höherer Wert erhöht nur zusätzlichen Speicheraufwand.

Wenn der größte cleartext Wert für eine bestimmte Spalte bekannt ist, empfehlen wir, stattdessen den fixed Pad-Typ zu verwenden. Die Verwendung von fixed Padding sorgt für Konsistenz zwischen aktualisierten Datensätzen. Die Verwendung max von Auffüllung führt dazu, dass jede Teilmenge der Daten mit dem größten Wert aufgefüllt wird, der in der Teilmenge enthalten war.

Beispieldaten für eine Spalte sealed

Im Folgenden finden Sie ein Beispiel für Eingabe- und Ausgabedaten für eine sealed Spalte mit Einstellungen zur Reproduktion. Andere Einstellungen auf Kollaborationsebene wie allowCleartextallowJoinsOnColumnsWithDifferentNames, und wirken sich allowDuplicates nicht auf die Ergebnisse aus und können so eingestellt werden, als true false ob versucht wird, sich lokal zu reproduzieren. Dies sind zwar die Grundeinstellungen für die Reproduktion, aber die sealed Spalte ist nicht deterministisch und die Werte ändern sich jedes Mal. Das Ziel besteht darin, die eingehenden Byte im Vergleich zu den ausgehenden Bytes anzuzeigen. Die pad_length Beispielwerte wurden bewusst ausgewählt. Sie zeigen, dass beim fixed Auffüllen die gleichen Werte wie max beim Auffüllen mit den empfohlenen pad_length Mindesteinstellungen erzielt werden oder wenn zusätzliche Polsterung gewünscht wird.

Beispiel für einen gemeinsamen geheimen Schlüssel: wJa1rXUtnFEMI/K7MDENG/
bPxRfiCYEXAMPLEKEY

Beispiel für eine Kollaborations-ID: a1b2c3d4-5678-90ab-cdef-EXAMPLE11111

Themen

- [Pad-Typ von none](#)
- [Pad-Typ von fixed \(Beispiel 1\)](#)
- [Pad-Typ von fixed \(Beispiel 2\)](#)
- [Pad-Typ von max \(Beispiel 1\)](#)
- [Pad-Typ von max \(Beispiel 2\)](#)

Pad-Typ von **none**

Beispiel 1

Input	null
preserveNulls	TRUE

Ausgabe	null
Deterministisch	Yes
Eingabe-Bytes	0
Ausgabe-Bytes	0

Beispiel 2

Input	null
preserveNulls	FALSE
Ausgabe	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc40TBqfRY Z98t5KU6aWfssGSPbNIJfG3iXmu 6cbCUrizuV
Deterministisch	No
Eingabe-Bytes	0
Ausgabe-Bytes	91

Beispiel 3

Input	empty string
preserveNulls	-
Ausgabe	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc40TBqfRY Z98t5KU6aWfstGSPeM6qR8DWC2P B2GM1X41YK
Deterministisch	No
Eingabe-Bytes	0

Ausgabe-Bytes	91

Beispiel 4

Input	abcdefghijklmnopqrstuvwxyz
preserveNulls	-
Ausgabe	01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc40TBqfRYZ98t5KU6aWfstEE1GKEPiRzyh0h7t60mWMLTWCv02ckr6pkx9sGL5VLDQeHzh6DmPpyWNuI=
Deterministisch	No
Eingabe-Bytes	26
Ausgabe-Bytes	127

Beispiel 5

Input	abcdefghijklmnopqrstuvwxyza BCDEFGHIJKLMNOPQRSTUVWXYZ01 23456789
preserveNulls	-
Ausgabe	01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc40TBqfRYZ98t5KU6aWfstEE1GKEPiRzyh0h7t60mWMLTWCv02ckr6plwtH/8tRFnn2rF91bcB9G4+n8GiRfJNmqdP4/Q0Q3cXb/pbvPcnnohrHIGSX54ua+1/JfcVjc=
Deterministisch	No

Eingabe-Bytes	62
Ausgabe-Bytes	175

Pad-Typ von **fixed** (Beispiel 1)

In diesem Beispiel `pad_length` ist es 62 und die größte Eingabe ist 62 Byte.

Beispiel 1

Input	<code>null</code>
<code>preserveNulls</code>	<code>TRUE</code>
Ausgabe	<code>null</code>
Deterministisch	<code>Yes</code>
Eingabe-Bytes	0
Ausgabe-Bytes	0

Beispiel 2

Input	<code>null</code>
<code>preserveNulls</code>	<code>FALSE</code>
Ausgabe	<code>01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc4OTBqfRYZ98t5KU6aWfssGSNWfMRp7nSb7SMX2s3JKLOhK1+7r75Tk+Mx9jy48Fcg1y0PvBqRSZ7oqy1V3UKfYTLEZb/hCz7oaIneVsrcoNpATs0GzbnLkor4L+/aSuA=</code>
Deterministisch	<code>No</code>
Eingabe-Bytes	0

Ausgabe-Bytes	175

Beispiel 3

Input	empty string
preserveNulls	-
Ausgabe	01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc40TBqfRYZ98t5KU6aWfstGSNWfMRp7nSb7SMX2s3JKL0hK1+7r75Tk+Mx9jy48Fcg1y0PvBqRSZ7oqy1V3UKfYTLEZb/hCz7oaIneVsrcolB53l07VZpA60wkuXu29CA=
Deterministisch	No
Eingabe-Bytes	0
Ausgabe-Bytes	175

Beispiel 4

Input	abcdefghijklmnopqrstuvwxyz
preserveNulls	-
Ausgabe	01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc40TBqfRYZ98t5KU6aWfstEE1GKEPiRzyh0h7t60mWMLTWCv02ckr6pkx9jy48Fcg1y0PvBqRSZ7oqy1V3UKfYTLEZb/hCz7oaIneVsrcutBAc0+Mb9tuU2KIIHH31AWg=
Deterministisch	No

Eingabe-Bytes	26
Ausgabe-Bytes	175

Beispiel 5

Input	abcdefghijklmnopqrstuvwxyzA BCDEFGHIJKLMNOPQRSTUVWXYZ01 23456789
preserveNulls	-
Ausgabe	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmlmDEyMzQ1Njc40TBqfRY Z98t5KU6aWfstEE1GKEPiRzyh0 h7t60mWMLTWcV02ckr6plwtH/8t RFnn2rF91bcB9G4+n8GiRfJNmqdP4/ QQQ3cXb/pbvPcnnohrHIGSX54ua+1/ JfcVjc=
Deterministisch	No
Eingabe-Bytes	62
Ausgabe-Bytes	175

Pad-Typ von **fixed** (Beispiel 2)

In diesem Beispiel `pad_length` ist es 162 und die größte Eingabe ist 62 Byte.

Beispiel 1

Input	null
preserveNulls	TRUE
Ausgabe	null
Deterministisch	Yes

Eingabe-Bytes	0
Ausgabe-Bytes	0

Beispiel 2

Input	null
preserveNulls	FALSE
Ausgabe	01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc40TBqfRYZ98t5KU6aWfssGSNWfMRp7nSb7SMX2s3JKLOhK1+7r75Tk+Mx9jy48Fcg1y0PvBqRSZ7oqy1V3UKfYTLEZb/hCz7oaIneVsrcnkB0xbLWD7zNdAqQGR0rXoSESdW0I0vpNoGcBfv4cJbG0A3h1DvtkSSVc2B8000GppzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn+8o4WtG/ClipNcjDXvXVtK4vfCohcCA6uwrmwv/xAySX+xcntotL703aBTBb
Deterministisch	No
Eingabe-Bytes	0
Ausgabe-Bytes	307

Beispiel 3

Input	empty string
preserveNulls	-
Ausgabe	01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc40TBqfRYZ98t5KU6aWfstGSNWfMRp7nSb7SMX2s3JKLOhK1+7r75Tk+Mx9jy48

	Fcg1y0PvBqRSZ7oqy1V3UKfYTLE Zb/hCz7oaIneVsrcnkB0xbLWD7z NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvtkSSVc2B8000Gp pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrmwv84lVaT9Yd+6oQx65/+gdVT
Deterministisch	No
Eingabe-Bytes	0
Ausgabe-Bytes	307

Beispiel 4

Input	abcdefghijklmnopqrstuvwxyz
preserveNulls	-
Ausgabe	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfstEE1GKEPiRzyh0 h7t60mWMLTWcV02ckr6pkx9jy48 Fcg1y0PvBqRSZ7oqy1V3UKfYTLE Zb/hCz7oaIneVsrcnkB0xbLWD7z NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvtkSSVc2B8000Gp pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrmwX5Hn1+Wyf06ks3QMaRDGSf
Deterministisch	No
Eingabe-Bytes	26
Ausgabe-Bytes	307

Beispiel 5

Input	abcdefghijklmnopqrstuvwxyzA BCDEFGHIJKLMNOPQRSTUVWXYZ01 23456789
preserveNulls	-
Ausgabe	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc40TBqfRY Z98t5KU6aWfstEE1GKEPiRzyh0 h7t60mWMLTWcV02ckr6plwtH/8t RFnn2rF91bcB9G4+n8GiRfJNmqd P4/Q0Q3cXb/pbvPcnkB0xbLWD7z NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvtkSSVc2B8000Gp pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrmwjkJXQZ0gPdeFX9Yr/8a1V5i
Deterministisch	No
Eingabe-Bytes	62
Ausgabe-Bytes	307

Pad-Typ von **max** (Beispiel 1)

In diesem Beispiel `pad_length` ist der Wert 0 und die größte Eingabe ist 62 Byte.

Beispiel 1

Input	null
preserveNulls	TRUE
Ausgabe	null
Deterministisch	Yes

Eingabe-Bytes	0
Ausgangs-Bytes	0

Beispiel 2

Input	null
preserveNulls	FALSE
Ausgabe	01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc40TBqfRYZ98t5KU6aWfssGSNWfMRp7nSb7SMX2s3JKL0hK1+7r75Tk+Mx9jy48Fcg1y0PvBqRSZ7oqy1V3UKfYTL EZb/hCz7oaIneVsrcoNpATs0GzbnLkor4L+/aSuA=
Deterministisch	No
Eingabe-Bytes	0
Ausgabe-Bytes	175

Beispiel 3

Input	empty string
preserveNulls	-
Ausgabe	01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc40TBqfRYZ98t5KU6aWfstGSNWfMRp7nSb7SMX2s3JKL0hK1+7r75Tk+Mx9jy48Fcg1y0PvBqRSZ7oqy1V3UKfYTL EZb/hCz7oaIneVsrcoLB53l07VZpA60wkuXu29CA=

Deterministisch	No
Eingabe-Bytes	0
Ausgabe-Bytes	175

Beispiel 4

Input	abcdefghijklmnopqrstuvwxyz
preserveNulls	-
Ausgabe	01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc40TBqfRYZ98t5KU6aWfstEE1GKEPiRzyh0h7t60mWMLTWCv02ckr6pkx9jy48Fcg1y0PvBqRSZ7oqy1V3UKfYTLEZb/hCz7oaIneVsrcutBAC0+Mb9tuU2KIIHH31AWg=
Deterministisch	No
Eingabe-Bytes	26
Ausgabe-Bytes	175

Beispiel 5

Input	abcdefghijklmnopqrstuvwxyzA BCDEFGHIJKLMNOPQRSTUVWXYZ01 23456789
preserveNulls	-
Ausgabe	01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc40TBqfRYZ98t5KU6aWfstEE1GKEPiRzyh0h7t60mWMLTWCv02ckr6plwtH/8t

	RFnn2rF91bcB9G4+n8GiRfJNmqdP4/ Q0Q3cXb/pbvPcnnohrHIGSX54ua+1/ JfcVjc=
Deterministisch	No
Eingabe-Bytes	62
Ausgabe-Bytes	175

Pad-Typ von **max** (Beispiel 2)

In diesem Beispiel `pad_length` ist es 100 und die größte Eingabe ist 62 Byte.

Beispiel 1

Input	null
<code>preserveNulls</code>	TRUE
Ausgabe	null
Deterministisch	Yes
Eingabe-Bytes	0
Ausgabe-Bytes	0

Beispiel 2

Input	null
<code>preserveNulls</code>	FALSE
Ausgabe	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmlmDEyMzQ1Njc40TBqfRY Z98t5KU6aWfssGSNWfMRp7nSb7S MX2s3JKL0hK1+7r75Tk+Mx9jy48 Fcg1y0PvBqRSZ7oqy1V3UKfYTLE Zb/hCz7oaIneVsrcnkB0xbLWD7z

	NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvtkSSVc2B8000Gp pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrmwv/xAySX+xcntotL703aBTBb
Deterministisch	No
Eingabe-Bytes	0
Ausgabe-Bytes	307

Beispiel 3

Input	empty string
preserveNulls	-
Ausgabe	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmn1MDEyMzQ1Njc40TBqfRY Z98t5KU6aWfstGSNWfMRp7nSb7S MX2s3JKL0hK1+7r75Tk+Mx9jy48 Fcgl1y0PvBqRSZ7oqy1V3UKfYTLE Zb/hCz7oaIneVsrcnkB0xbLWD7z NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvtkSSVc2B8000Gp pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrmwv841VaT9Yd+6oQx65/+gdVT
Deterministisch	No
Eingabe-Bytes	0
Ausgabe-Bytes	307

Beispiel 4

Input	abcdefghijklmnopqrstuvwxyz
preserveNulls	-
Ausgabe	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfstEE1GKEPiRzyh0 h7t60mWMLTWcV02ckr6pkx9jy48 Fcgl1y0PvBqRSZ7oqy1V3UKfYTLE Zb/hCz7oaIneVsircnkB0xbLWD7z NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvtkSSVc2B8000Gp pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrmtX5Hn1+Wyf06ks3QMaRDGSf
Deterministisch	No
Eingabe-Bytes	26
Ausgabe-Bytes	307

Beispiel 5

Input	abcdefghijklmnopqrstuvwxyzA BCDEFGHIJKLMNOPQRSTUVWXYZ01 23456789
preserveNulls	-
Ausgabe	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfstEE1GKEPiRzyh0 h7t60mWMLTWcV02ckr6plwtH/8t RFnn2rF91bcB9G4+n8GiRfJNmqd P4/Q0Q3cXb/pbvPcnkB0xbLWD7z


```
NdAqQGR0rXoSESdW0I0vpNoGcBf
v4cJbG0A3h1DvtkSSVc2B8000Gp
pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn
+8o4WtG/ClipNcjDXvXVtK4vfCohcCA6
uwrmwjkJXQZ0gPdeFX9Yr/8a1V5i
```

Deterministisch	No
Eingabe-Bytes	62
Ausgabe-Bytes	307

sealedSpalten zur Problembehandlung

Warum ist der Chiffretext in meinen sealed Spalten um ein Vielfaches größer als cleartext der Inhalt?

Das hängt von mehreren Faktoren ab. Zum einen ist der Chiffretext in einer Cleartext Spalte immer mindestens 91 Byte lang. Wenn Ihre Eingabedaten klein wären (z. B. das Alter der Kunden), würden sie deutlich an Größe zunehmen. Zweitens, wenn `preserveNulls` auf `false` eingestellt ist und Ihre Eingabedaten viele `null` Werte enthielten, wurde jeder dieser `null` Werte in 91 Byte Chiffretext umgewandelt. Und wenn Sie Padding verwenden, werden den cleartext Daten per Definition Byte hinzugefügt, bevor sie verschlüsselt werden.

Die meisten meiner Daten in einer sealed Spalte sind sehr klein, und ich muss Padding verwenden. Kann ich einfach die großen Werte entfernen und sie separat verarbeiten, um Platz zu sparen?

Wir empfehlen nicht, große Werte zu entfernen und separat zu verarbeiten. Dadurch ändern sich die Datenschutzgarantien, die der C3R-Verschlüsselungsclient bietet. Gehen Sie als Bedrohungsmodell davon aus, dass ein Beobachter beide verschlüsselten Datensätze sehen kann. Wenn der Beobachter feststellt, dass bei einer Teilmenge von Daten eine Spalte deutlich mehr oder weniger aufgefüllt ist als bei einer anderen Teilmenge, kann er Rückschlüsse auf die Größe der Daten in jeder Teilmenge ziehen. Nehmen wir beispielsweise an, dass eine `fullName` Spalte in einer Datei auf insgesamt 40 Byte aufgefüllt ist und in einer anderen Datei auf 800 Byte aufgefüllt wird. Ein Beobachter könnte davon ausgehen, dass ein Datensatz den längsten Namen der Welt (747 Byte) enthält.

Muss ich zusätzliche Polsterung bereitstellen, wenn ich den `max` Padding-Typ verwende?

Nein. Bei der Verwendung `max` von Innenabständen empfehlen wir `pad_length`, die auch als zusätzliche Polsterung bezeichnet wird, die über den größten Wert in der Spalte hinausgeht, auf 0 zu setzen.

Kann ich **`pad_length`** bei der Verwendung von **`fixed`** Padding einfach einen Wert vom Typ `L` wählen, damit ich mir keine Gedanken darüber machen muss, ob der größte Wert passt?

Ja, aber die große Länge des Pads ist ineffizient und beansprucht mehr Speicherplatz als nötig. Wir empfehlen Ihnen, zu überprüfen, wie groß der größte Wert ist, und den Wert `pad_length` auf diesen Wert einzustellen.

Woher weiß ich, ob ich die kryptografischen Garantien von benötige? **`preserveNulls`**

Leider lautet die Antwort, dass es darauf ankommt. Zumindest [Kryptografisches Rechnen für Clean Rooms](#) sollte überprüft werden, wie die `preserveNulls` Einstellung Ihre Daten schützt. Wir empfehlen Ihnen jedoch, die Datenverarbeitungsanforderungen Ihres Unternehmens und alle Verträge, die für die jeweilige Zusammenarbeit gelten, zu beachten.

Warum muss ich den Overhead von Base64 auf mich nehmen?

Um die Kompatibilität mit tabellarischen Dateiformaten wie CSV zu gewährleisten, ist eine Base64-Kodierung erforderlich. Obwohl einige Dateiformate, wie z. Parquet B., binäre Darstellungen von Daten unterstützen, ist es wichtig, dass alle Teilnehmer einer Zusammenarbeit Daten auf die gleiche Weise darstellen, um korrekte Abfrageergebnisse zu gewährleisten.

Behebung unerwarteter Zunahmen der Chiffretext-Größe

Nehmen wir an, Sie haben Ihre Daten verschlüsselt und die Größe der resultierenden Daten ist überraschend groß. Mithilfe der folgenden Schritte können Sie ermitteln, wo der Größenzuwachs stattgefunden hat und welche Maßnahmen Sie gegebenenfalls ergreifen können.

Identifizieren Sie, wo die Größenzunahme stattgefunden hat

Bevor Sie herausfinden können, warum Ihre verschlüsselten Daten deutlich größer sind als Ihre cleartext Daten, müssen Sie zunächst herausfinden, wo die Zunahme liegt. CleartextSpalten können bedenkenlos ignoriert werden, da sie unverändert sind. Sehen Sie sich die verbleibenden sealed Spalten fingerprint und die anderen an und wählen Sie eine aus, die aussagekräftig erscheint.

Identifizieren Sie den Grund für die Größenzunahme

Eine fingerprint Spalte oder eine sealed Spalte könnte zur Größenzunahme beitragen.

Themen

- [Kommt die Größenzunahme von einer fingerprint Spalte?](#)
- [Ist die Größenzunahme auf eine sealed Spalte zurückzuführen?](#)

Kommt die Größenzunahme von einer fingerprint Spalte?

Wenn es sich bei der Spalte, die am meisten zur Speichererweiterung beiträgt, um eine fingerprint Spalte handelt, liegt das wahrscheinlich daran, dass die cleartext Daten klein sind (z. B. das Alter des Kunden). Jeder resultierende fingerprint Chiffretext hat eine Länge von 52 Byte. Leider kann gegen dieses Problem nichts auf der Grundlage unternommen werden. column-by-column Weitere Informationen zu dieser Spalte, einschließlich der Auswirkungen auf die Speicheranforderungen, finden Sie unter. [Basis-Overhead für fingerprint Spalten](#)

Die andere mögliche Ursache für die Vergrößerung einer fingerprint Spalte ist die Einstellung für die Zusammenarbeit, `preserveNulls`. Wenn die Einstellung für die Zusammenarbeit deaktiviert `preserveNulls` ist (die Standardeinstellung), werden aus allen null Werten in fingerprint Spalten 52 Byte Chiffretext. In der aktuellen Zusammenarbeit kann dafür nichts unternommen werden. Die `preserveNulls` Einstellung wird bei der Erstellung einer Kollaboration festgelegt, und alle Mitarbeiter müssen dieselbe Einstellung verwenden, um korrekte Abfrageergebnisse sicherzustellen. Weitere Informationen zu dieser `preserveNulls` Einstellung und dazu, wie sich ihre Aktivierung auf die Datenschutzgarantien Ihrer Daten auswirkt, finden Sie unter. [the section called “Kryptografisches Rechnen”](#)

Ist die Größenzunahme auf eine sealed Spalte zurückzuführen?

Wenn es sich bei der Spalte, die am meisten zur Erhöhung des Speicherplatzes beiträgt, um eine sealed Spalte handelt, gibt es einige Details, die zur Vergrößerung beitragen könnten.

Wenn die cleartext Daten klein sind (z. B. das Alter des Kunden), ist jeder resultierende sealed Chiffretext mindestens 91 Byte lang. Leider kann nichts gegen dieses Problem unternommen werden. Weitere Informationen zu dieser Spalte, einschließlich der Auswirkungen auf die Speicheranforderungen, finden Sie unter. [Basis-Overhead für Spalten sealed](#)

Die zweite Hauptursache für die Erhöhung des Speicherplatzes in sealed Spalten ist die Polsterung. Beim Auffüllen werden zusätzliche Byte hinzugefügt, cleartext bevor es verschlüsselt wird, um die Größe einzelner Werte in einem Datensatz zu verbergen. Wir empfehlen Ihnen, das Padding auf den kleinstmöglichen Wert für Ihren Datensatz festzulegen. `pad_length` Für das fixed Padding muss mindestens so eingestellt werden, dass es den größtmöglichen Wert in der Spalte umfasst.

Eine höhere Einstellung als diese bietet keine zusätzlichen Datenschutzgarantien. Wenn Sie beispielsweise wissen, dass der größtmögliche Wert in einer Spalte 50 Byte sein kann, empfehlen wir, den Wert `pad_length` auf 50 Byte festzulegen. Wenn in der sealed Spalte jedoch max Padding verwendet wird, empfehlen wir, den Wert `pad_length` auf 0 Byte zu setzen. Dies liegt daran, dass max sich das Auffüllen auf das zusätzliche Auffüllen bezieht, das über den größten Wert in der Spalte hinausgeht.

Die letzte mögliche Ursache für die Vergrößerung einer sealed Spalte ist die Einstellung für die Zusammenarbeit, `preserveNulls`. Wenn die Einstellung für die Zusammenarbeit deaktiviert `preserveNulls` ist (die Standardeinstellung), werden aus allen `null` Werten in sealed Spalten 91 Byte Chiffretext. In der aktuellen Zusammenarbeit kann dafür nichts unternommen werden. Die `preserveNulls` Einstellung wird bei der Erstellung einer Kollaboration festgelegt, und alle Mitarbeiter müssen dieselbe Einstellung verwenden, um korrekte Abfrageergebnisse sicherzustellen. Weitere Informationen zu dieser Einstellung und zu den Auswirkungen ihrer Aktivierung auf die Datenschutzgarantien Ihrer Daten finden Sie unter [the section called “Kryptografisches Rechnen”](#)

Einrichten AWS Clean Rooms

In den folgenden Themen wird die Einrichtung erläutert AWS Clean Rooms.

Themen

- [Melden Sie sich an für AWS](#)
- [Richten Sie Servicerollen ein für AWS Clean Rooms](#)
- [Servicerollen für AWS Clean Rooms ML einrichten](#)

Melden Sie sich an für AWS

Bevor Sie eine oder eine andere nutzen AWS Clean Rooms können AWS-Service, müssen Sie sich AWS mit einem anmelden AWS-Konto.

Wenn Sie noch keine haben AWS-Konto, führen Sie die folgenden Schritte aus, um eine zu erstellen.

Um sich für eine anzumelden AWS-Konto

1. Öffnen Sie [https://portal.aws.amazon.com/billing/die Anmeldung](https://portal.aws.amazon.com/billing/die-Anmeldung).
2. Folgen Sie den Online-Anweisungen.

Während des Anmeldevorgangs erhalten Sie einen Telefonanruf mit einem Bestätigungscode, den Sie auf der Telefontastatur eingeben.

3. Wenn Sie sich für einen anmelden AWS-Konto, wird ein AWS-Konto Root-Benutzer erstellt. Der Root-Benutzer hat Zugriff auf alle AWS-Services und Ressourcen des Kontos. Als bewährte Sicherheitsmethode weisen Sie einem [Administratorbenutzer Administratorzugriff](#) zu und verwenden Sie nur den Root-Benutzer, um [Aufgaben auszuführen, die Root-Benutzerzugriff](#) erfordern.

Richten Sie Servicerollen ein für AWS Clean Rooms

In den folgenden Abschnitten werden die Rollen beschrieben, die zur Ausführung der einzelnen Aufgaben benötigt werden.

Themen

- [Erstellen Sie einen Administratorbenutzer](#)

- [Erstellen Sie eine IAM-Rolle für ein Kollaborationsmitglied](#)
- [Erstellen Sie eine Servicerolle zum Lesen von Daten aus Amazon S3](#)
- [Erstellen Sie eine Servicerolle zum Lesen von Daten aus Amazon Athena](#)
- [Erstellen Sie eine Servicerolle, um Daten aus Snowflake zu lesen](#)
- [Erstellen Sie eine Servicerolle, um Code aus einem S3-Bucket zu lesen \(PySpark Analysevorlagenrolle\)](#)
- [Erstellen Sie eine Servicerolle, um die Ergebnisse eines PySpark Jobs zu schreiben](#)
- [Erstellen Sie eine Servicerolle, um Ergebnisse zu erhalten](#)

Erstellen Sie einen Administratorbenutzer

Zur Verwendung AWS Clean Rooms müssen Sie einen Administratorbenutzer für sich selbst erstellen und den Administratorbenutzer einer Administratorgruppe hinzufügen.

Wählen Sie zum Erstellen eines Administratorbenutzers eine der folgenden Optionen aus.

Wählen Sie eine Möglichkeit zur Verwaltung Ihres Administrators aus.	Bis	Von	Sie können auch
Im IAM Identity Center (Empfohlen)	Verwendung von kurzfristigen Anmeldeinformationen für den Zugriff auf AWS. Dies steht im Einklang mit den bewährten Methoden für die Sicherheit. Weitere	Beachtung der Anweisungen unter Erste Schritte im AWS IAM Identity Center - Benutzerhandbuch.	Konfigurieren Sie den programmatischen Zugriff, indem Sie AWS CLI die Konfiguration für die Verwendung AWS IAM Identity Center im AWS Command Line Interface Benutzerhandbuch vornehmen.

Wählen Sie eine Möglichkeit zur Verwaltung Ihres Administrators aus.	Bis	Von	Sie können auch
	Informationen zu bewährten Methoden finden Sie unter Bewährte Methoden für die Sicherheit in IAM im IAM-Benutzerhandbuch.		
In IAM (Nicht empfohlen)	Verwendung von langfristigen Anmeldeinformationen für den Zugriff auf AWS.	Folgen Sie den Anleitungen unter IAM-Benutzer für den Notfallzugriff erstellen im IAM-Benutzerhandbuch.	Sie konfigurieren den programmgesteuerten Zugriff unter Verwendung der Informationen unter Verwalten der Zugriffsschlüssel für IAM-Benutzer im IAM-Benutzerhandbuch.

Erstellen Sie eine IAM-Rolle für ein Kollaborationsmitglied

Ein Mitglied ist ein AWS Kunde, der an einer Kollaboration teilnimmt.

Um eine IAM-Rolle für ein Kollaborationsmitglied zu erstellen

1. Folgen Sie dem Verfahren [Erstellen einer Rolle zum Delegieren von Berechtigungen an einen IAM-Benutzer](#) im AWS Identity and Access Management Benutzerhandbuch.
2. Wählen Sie für den Schritt Richtlinie erstellen im Richtlinien-Editor die Registerkarte JSON aus und fügen Sie dann je nach den Fähigkeiten, die dem Kollaborationsmitglied gewährt wurden, Richtlinien hinzu.

AWS Clean Rooms bietet die folgenden verwalteten Richtlinien auf der Grundlage gängiger Anwendungsfälle.

Wenn Sie ...	Dann benutze...
Sehen Sie sich die Ressourcen und Metadaten an	AWS verwaltete Richtlinie: <code>AWSCleanRoomsReadOnlyAccess</code>
Query	AWS verwaltete Richtlinie: <code>AWSCleanRoomsFullAccess</code>
Jobs abfragen und ausführen	AWS verwaltete Richtlinie: <code>AWSCleanRoomsFullAccess</code>
Ergebnisse abfragen und empfangen	AWS verwaltete Richtlinie: <code>AWSCleanRoomsFullAccess</code>
Ressourcen für die Zusammenarbeit verwalten, aber keine Abfragen durchführen	AWS verwaltete Richtlinie: <code>AWSCleanRoomsFullAccessNoQuerying</code>

Informationen zu den verschiedenen verwalteten Richtlinien, die von angeboten werden AWS Clean Rooms [AWS verwaltete Richtlinien für AWS Clean Rooms](#), finden Sie unter

Erstellen Sie eine Servicerolle zum Lesen von Daten aus Amazon S3

AWS Clean Rooms verwendet eine Servicerolle, um die Daten aus Amazon S3 zu lesen.

Es gibt zwei Möglichkeiten, diese Servicerolle zu erstellen.

- Wenn Sie über die erforderlichen IAM-Berechtigungen zum Erstellen einer Servicerolle verfügen, verwenden Sie die AWS Clean Rooms Konsole, um eine Servicerolle zu erstellen.
- Wenn Sie nicht über die `iam:AttachRolePolicy` erforderlichen Berechtigungen verfügen `iam:CreateRole` oder die IAM-Rollen manuell erstellen möchten, gehen Sie wie folgt vor:
`iam:CreatePolicy`
- Gehen Sie wie folgt vor, um eine Servicerolle mithilfe benutzerdefinierter Vertrauensrichtlinien zu erstellen.

- Bitten Sie Ihren Administrator, die Servicerolle mithilfe des folgenden Verfahrens zu erstellen.

Note

Sie oder Ihr IAM-Administrator sollten dieses Verfahren nur befolgen, wenn Sie nicht über die erforderlichen Berechtigungen verfügen, um mithilfe der AWS Clean Rooms Konsole eine Servicerolle zu erstellen.

So erstellen Sie eine Servicerolle zum Lesen von Daten aus Amazon S3 mithilfe benutzerdefinierter Vertrauensrichtlinien

1. Erstellen Sie eine Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien. Weitere Informationen finden Sie unter dem Verfahren [Erstellen einer Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien \(Konsole\)](#) im AWS Identity and Access Management Benutzerhandbuch.
2. Verwenden Sie die folgende benutzerdefinierte Vertrauensrichtlinie gemäß dem Verfahren [Erstellen einer Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien \(Konsole\)](#).

Note

Wenn Sie sicherstellen möchten, dass die Rolle nur im Rahmen einer bestimmten Kollaborationsmitgliedschaft verwendet wird, können Sie die Vertrauensrichtlinie weiter eingrenzen. Weitere Informationen finden Sie unter [Serviceübergreifende Confused-Deputy-Prävention](#).

JSON

```
{  
  
  "Version": "2012-10-17",  
  "Statement": [  
    {  
      "Sid": "RoleTrustPolicyForCleanRoomsService",  
      "Effect": "Allow",  
      "Principal": {  
        "Service": "cleanrooms.amazonaws.com"  
      },  
    },  
  ],  
}
```

```

        "Action": "sts:AssumeRole"
      }
    ]
  }

```

3. Verwenden Sie die folgende Berechtigungsrichtlinie gemäß dem Verfahren [Erstellen einer Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien \(Konsole\)](#).

Note

Die folgende Beispielrichtlinie unterstützt die Berechtigungen, die zum Lesen von AWS Glue Metadaten und den entsprechenden Amazon S3 S3-Daten erforderlich sind. Je nachdem, wie Sie Ihre Amazon S3 S3-Daten eingerichtet haben, müssen Sie diese Richtlinie jedoch möglicherweise ändern. Wenn Sie beispielsweise einen benutzerdefinierten KMS-Schlüssel für Ihre Amazon S3 S3-Daten eingerichtet haben, müssen Sie diese Richtlinie möglicherweise mit zusätzlichen AWS Key Management Service (AWS KMS) Berechtigungen ändern.

Ihre AWS Glue Ressourcen und die zugrunde liegenden Amazon S3 S3-Ressourcen müssen mit der AWS Clean Rooms Zusammenarbeit AWS-Region identisch sein.

JSON

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "NecessaryGluePermissions",
      "Effect": "Allow",
      "Action": [
        "glue:GetDatabase",
        "glue:GetDatabases",
        "glue:GetTable",
        "glue:GetTables",
        "glue:GetPartition",
        "glue:GetPartitions",
        "glue:BatchGetPartition"
      ],
      "Resource": [
        "arn:aws:glue:us-east-1:111122223333:database/databaseName",

```

```

        "arn:aws:glue:us-east-1:111122223333:table/databaseName/tableName",
        "arn:aws:glue:us-east-1:111122223333:catalog"
    ],
},
{
    "Effect": "Allow",
    "Action": [
        "glue:GetSchema",
        "glue:GetSchemaVersion"
    ],
    "Resource": [
        "*"
    ]
},
{
    "Sid": "NecessaryS3BucketPermissions",
    "Effect": "Allow",
    "Action": [
        "s3:GetBucketLocation",
        "s3:ListBucket"
    ],
    "Resource": [
        "arn:aws:s3:::bucket"
    ],
    "Condition": {
        "StringEquals": {
            "s3:ResourceAccount": [
                "444455556666"
            ]
        }
    }
},
{
    "Sid": "NecessaryS3ObjectPermissions",
    "Effect": "Allow",
    "Action": [
        "s3:GetObject"
    ],
    "Resource": [
        "arn:aws:s3:::bucket/prefix/*"
    ],
    "Condition": {
        "StringEquals": {

```

```

        "s3:ResourceAccount": [
            "444455556666"
        ]
    }
}

```

Note

Diese Richtlinie bezieht sich auf zwei verschiedene Arten AWS-Konto IDs , um eine AWS Clean Rooms Zusammenarbeit zu unterstützen, bei der die Metadaten des Datenkatalogs und der tatsächliche Datenspeicher von verschiedenen Parteien verwaltet werden:

- 111122223333 — Dies ist das Konto, dem die AWS Glue Datenkatalogressourcen (Datenbanken, Tabellen und Katalog) gehören. Die erste Anweisung gewährt Berechtigungen für den Zugriff auf Tabellenschemas, Partitionsinformationen und Metadaten aus dem Katalog dieses Kontos. AWS Glue
- 444455556666 — Dies ist das Konto, dem der Amazon S3 S3-Bucket gehört, der die eigentlichen Datendateien enthält. Die Amazon S3 S3-Berechtigungen (Anweisungen 3 und 4) sind aufgrund der `s3:ResourceAccount` Bedingung auf Buckets beschränkt, die diesem Konto gehören.

Diese Konfiguration unterstützt gängige Unternehmensdatenarchitekturen, bei denen ein Team den Datenkatalog und die Schemadefinitionen verwaltet, während ein anderes Team die zugrunde liegende Datenspeicherinfrastruktur besitzt. Diese `s3:ResourceAccount` Bedingung bietet eine zusätzliche Sicherheitsebene, indem sichergestellt wird, dass Amazon S3 S3-Operationen nur mit Buckets funktionieren, die dem angegebenen Konto gehören.

4. Ersetzen Sie jeden *placeholder* durch Ihre Informationen.
5. Folgen Sie weiterhin dem Verfahren zum [Erstellen einer Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien \(Konsole\)](#), um die Rolle zu erstellen.

Erstellen Sie eine Servicerolle zum Lesen von Daten aus Amazon Athena

AWS Clean Rooms verwendet eine Servicerolle, um die Daten von Amazon Athena zu lesen.

So erstellen Sie eine Servicerolle zum Lesen von Daten aus Athena mithilfe benutzerdefinierter Vertrauensrichtlinien

1. Erstellen Sie eine Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien. Weitere Informationen finden Sie unter dem Verfahren [Erstellen einer Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien \(Konsole\)](#) im AWS Identity and Access Management Benutzerhandbuch.
2. Verwenden Sie die folgende benutzerdefinierte Vertrauensrichtlinie gemäß dem Verfahren [Erstellen einer Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien \(Konsole\)](#).

Note

Wenn Sie sicherstellen möchten, dass die Rolle nur im Rahmen einer bestimmten Kollaborationsmitgliedschaft verwendet wird, können Sie die Vertrauensrichtlinie weiter eingrenzen. Weitere Informationen finden Sie unter [Serviceübergreifende Confused-Deputy-Prävention](#).

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "RoleTrustPolicyForCleanRoomsService",
      "Effect": "Allow",
      "Principal": {
        "Service": "cleanrooms.amazonaws.com"
      },
      "Action": "sts:AssumeRole"
    }
  ]
}
```

3. Verwenden Sie die folgende Berechtigungsrichtlinie gemäß dem Verfahren [Erstellen einer Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien \(Konsole\)](#).

 Note

Die folgende Beispielrichtlinie unterstützt die Berechtigungen, die zum Lesen von AWS Glue Metadaten und den entsprechenden Athena-Daten erforderlich sind. Je nachdem, wie Sie Ihre Amazon S3 S3-Daten eingerichtet haben, müssen Sie diese Richtlinie jedoch möglicherweise ändern. Wenn Sie beispielsweise bereits einen benutzerdefinierten KMS-Schlüssel für Ihre Amazon S3 S3-Daten eingerichtet haben, müssen Sie diese Richtlinie möglicherweise mit zusätzlichen AWS KMS Berechtigungen ändern.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "athena:GetWorkGroup",
        "athena:GetTableMetadata",
        "athena:GetDataCatalog",
        "athena:StartQueryExecution",
        "athena:GetQueryExecution",
        "athena:GetQueryResults"
      ],
      "Resource": [
        "arn:aws:athena:region:accountId:workgroup/workgroup",
        "arn:aws:athena:region:accountId:datacatalog/federatedCatalogName"
      ]
    },
    {
      "Effect": "Allow",
      "Action": [
        "glue:GetDatabase",
        "glue:GetTable",
        "glue:GetCatalog"
      ],
      "Resource": [
        "arn:aws:glue:region:accountId:catalog",

```

```

        "arn:aws:glue:region:accountId:catalog/federatedCatalogName",
        "arn:aws:glue:region:accountId:database/federatedCatalogName/databaseName",
        "arn:aws:glue:region:accountId:table/federatedCatalogName/databaseName/tableName"
    ],
    },
    {
        "Effect": "Allow",
        "Action": [
            "s3:GetObject",
            "s3:GetBucketLocation",
            "s3:AbortMultipartUpload",
            "s3:ListBucket",
            "s3:PutObject",
            "s3:ListMultipartUploadParts"
        ],
        "Resource": [
            "arn:aws:s3::athenaResultsBucket",
            "arn:aws:s3::athenaResultsBucket/*"
        ],
        "Condition": {
            "StringEquals": {
                "aws:ResourceAccount": "accountId"
            }
        }
    },
    {
        "Effect": "Allow",
        "Action": "lakeformation:GetDataAccess",
        "Resource": "*"
    }
]
}

```

4. Ersetzen Sie jeden *placeholder* durch Ihre Informationen.
5. Folgen Sie weiterhin dem Verfahren zum [Erstellen einer Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien \(Konsole\)](#), um die Rolle zu erstellen.

Lake Formation Formation-Berechtigungen einrichten

Wenn Sie Ressourcen abfragen, die mit Lake Formation Formation-Berechtigungen geschützt sind, muss die Servicerolle über die Zugriffsberechtigungen Select and Describe für die Datenbank table/view/catalog und Describe-Berechtigungen für die AWS Glue Datenbank verfügen.

Weitere Informationen finden Sie unter:

- [Verwenden Sie Athena, um Daten abzufragen, mit denen Sie AWS Lake Formation im Amazon Athena Athena-Benutzerhandbuch registriert](#) sind
- [Onboarding zu Lake Formation Formation-Berechtigungen](#) im AWS Lake Formation Developer Guide

Erstellen Sie eine Servicerolle, um Daten aus Snowflake zu lesen

AWS Clean Rooms verwendet eine Servicerolle, um Ihre Anmeldeinformationen abzurufen, damit Snowflake Ihre Daten aus dieser Quelle lesen kann.

Es gibt zwei Möglichkeiten, diese Servicerolle zu erstellen:

- Wenn Sie über die erforderlichen IAM-Berechtigungen zum Erstellen einer Servicerolle verfügen, verwenden Sie die AWS Clean Rooms Konsole, um eine Servicerolle zu erstellen.
- Wenn Sie nicht über die `iam:AttachRolePolicy` erforderlichen Berechtigungen verfügen `iam:CreateRole` oder die IAM-Rollen manuell erstellen möchten, gehen Sie wie folgt vor:
`iam:CreatePolicy`
 - Gehen Sie wie folgt vor, um eine Servicerolle mithilfe benutzerdefinierter Vertrauensrichtlinien zu erstellen.
 - Bitten Sie Ihren Administrator, die Servicerolle mithilfe des folgenden Verfahrens zu erstellen.

Note

Sie oder Ihr IAM-Administrator sollten dieses Verfahren nur befolgen, wenn Sie nicht über die erforderlichen Berechtigungen verfügen, um mithilfe der AWS Clean Rooms Konsole eine Servicerolle zu erstellen.

So erstellen Sie eine Servicerolle zum Lesen von Daten aus Snowflake mithilfe benutzerdefinierter Vertrauensrichtlinien

1. Erstellen Sie eine Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien. Weitere Informationen finden Sie unter dem Verfahren [Erstellen einer Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien \(Konsole\)](#) im AWS Identity and Access Management Benutzerhandbuch.
2. Verwenden Sie die folgende benutzerdefinierte Vertrauensrichtlinie gemäß dem Verfahren [Erstellen einer Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien \(Konsole\)](#).

 Note

Wenn Sie sicherstellen möchten, dass die Rolle nur im Rahmen einer bestimmten Kollaborationsmitgliedschaft verwendet wird, können Sie die Vertrauensrichtlinie weiter eingrenzen. Weitere Informationen finden Sie unter [Serviceübergreifende Confused-Deputy-Prävention](#).

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowIfSourceArnMatches",
      "Effect": "Allow",
      "Principal": {
        "Service": "cleanrooms.amazonaws.com"
      },
      "Action": "sts:AssumeRole",
      "Condition": {
        "ForAnyValue:ArnEquals": {
          "aws:SourceArn": [
            "arn:aws:cleanrooms:us-east-1:111122223333:membership/membershipId",
            "arn:aws:cleanrooms:us-east-1:444455556666:membership/queryRunnerMembershipId"
          ]
        }
      }
    }
  ]
}
```

```
]
}
```

Note

Diese Vertrauensrichtlinie bezieht sich auf zwei verschiedene Arten AWS-Konto IDs , um eine AWS Clean Rooms Zusammenarbeit zu unterstützen, bei der die Verantwortung für die Ausführung von Abfragen auf mehrere Parteien verteilt ist:

- 111122223333 — Dies ist das Konto, das eine Mitgliedschaft enthält, die an der Kollaboration teilnimmt. Diese Mitgliedschaft kann Datentabellen, Analyseregeln oder andere Kollaborationsressourcen besitzen, für die Rollenzugriff erforderlich ist.
- 444455556666 — Dies ist das Konto, das die Mitgliedschaft enthält, die für die Ausführung von Abfragen verantwortlich ist (der „Query Runner“). Diese Mitgliedschaft führt geschützte Abfragen aus und muss diese Rolle übernehmen, um auf die erforderlichen Rechen- und Datenressourcen zugreifen zu können.

Diese Konfiguration ermöglicht Szenarien, in denen eine Partei Daten oder Analysevorlagen bereitstellt, während eine andere Partei die eigentlichen Abfragen ausführt. Für beide Rollen sind unterschiedliche, aber sich ergänzende Berechtigungen für dieselbe Ausführungsrolle erforderlich. Diese `aws:SourceArn` Bedingung stellt sicher, dass nur AWS Clean Rooms Operationen, die aus diesen beiden spezifischen Mitgliedschaften stammen, diese Rolle übernehmen können, wodurch die Sicherheit gewährleistet und gleichzeitig die verteilte Auftragsausführung und das Ergebnismanagement unterstützt werden.

3. Verwenden Sie eine der folgenden Berechtigungsrichtlinien gemäß dem Verfahren [Erstellen einer Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien \(Konsole\)](#).

Berechtigungsrichtlinie für Geheimnisse, die mit einem kundeneigenen KMS-Schlüssel verschlüsselt wurden

JSON

```
{
  "Version": "2012-10-17",
```

```

    "Statement": [
      {
        "Action": "secretsmanager:GetSecretValue",
        "Resource": "arn:aws:secretsmanager:us-
east-1:111122223333:secret:secretIdentifier",
        "Effect": "Allow"
      },
      {
        "Sid": "AllowDecryptViaSecretsManagerForKey",
        "Action": "kms:Decrypt",
        "Resource": "arn:aws:kms:us-
east-1:444455556666:key/keyIdentifier",
        "Effect": "Allow",
        "Condition": {
          "StringEquals": {
            "kms:ViaService": "secretsmanager.us-
east-1.amazonaws.com",
            "kms:EncryptionContext:SecretARN":
"arn:aws:secretsmanager:us-east-1:111122223333:secret:secretIdentifier"
          }
        }
      }
    ]
  }
}

```

Note

Diese Richtlinie bezieht sich auf zwei unterschiedliche Verfahren AWS-Konto IDs , um ein kontenübergreifendes Verwaltungsszenario für geheime Daten zu unterstützen:

- 111122223333 — Dies ist das Konto, dem das Geheimnis gehört und in dem es gespeichert wird. Die erste Anweisung erteilt die Erlaubnis, den geheimen Wert von diesem Konto abzurufen.
- 444455556666 — Dies ist das Konto, dem der AWS KMS Schlüssel gehört, mit dem das Geheimnis verschlüsselt wurde. Die zweite Anweisung erteilt die Erlaubnis, das Geheimnis mithilfe des Schlüssels aus diesem Konto zu entschlüsseln. AWS KMS

Diese Konfiguration ist in Unternehmensumgebungen üblich, in denen:

- Geheimnisse werden zentral in einem Konto (Konto 1) verwaltet

- Verschlüsselungsschlüssel werden über ein separates Sicherheits- oder Shared-Services-Konto (Konto 2) verwaltet
- Die AWS KMS Schlüsselrichtlinie in Konto 2 muss es auch dem Dienst in Konto 1 ermöglichen, den Schlüssel für encryption/decryption Operationen zu verwenden

Die `kms:EncryptionContext:SecretARN` Bedingung stellt sicher, dass der AWS KMS Schlüssel nur zur Entschlüsselung dieses speziellen Geheimnisses verwendet werden kann, wodurch eine zusätzliche Sicherheitsebene für den kontoübergreifenden Zugriff geschaffen wird.

Berechtigungsrichtlinie für Geheimnisse, die mit einem verschlüsselt wurden Von AWS verwalteter Schlüssel

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": "secretsmanager:GetSecretValue",
      "Resource": "arn:aws:secretsmanager:us-  
east-1:111122223333:secret:secretIdentifier",
      "Effect": "Allow"
    }
  ]
}
```

4. Ersetzen Sie jeden *placeholder* durch Ihre Informationen.
5. Folgen Sie weiterhin dem Verfahren [Erstellen einer Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien \(Konsole\)](#), um die Rolle zu erstellen.

Erstellen Sie eine Servicerolle, um Code aus einem S3-Bucket zu lesen (PySpark Analysevorlagenrolle)

AWS Clean Rooms verwendet eine Servicerolle, um Code aus dem angegebenen S3-Bucket eines Kollaborationsmitglieds zu lesen, wenn eine PySpark Analysevorlage verwendet wird.

Um eine Servicerolle zum Lesen von Code aus einem S3-Bucket zu erstellen

1. Erstellen Sie eine Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien. Weitere Informationen finden Sie unter dem Verfahren [Erstellen einer Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien \(Konsole\)](#) im AWS Identity and Access Management Benutzerhandbuch.
2. Verwenden Sie die folgende benutzerdefinierte Vertrauensrichtlinie gemäß dem Verfahren [Erstellen einer Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien \(Konsole\)](#).

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "cleanrooms.amazonaws.com"
      },
      "Action": "sts:AssumeRole",
      "Condition": {
        "ForAnyValue:ArnEquals": {
          "aws:SourceArn": [
            "arn:aws:cleanrooms:us-east-1:111122223333:membership/jobRunnerMembershipId",
            "arn:aws:cleanrooms:us-east-1:444455556666:membership/analysisTemplateOwnerMembershipId"
          ]
        }
      }
    }
  ]
}
```

Note

Diese Vertrauensrichtlinie bezieht sich auf zwei verschiedene Arten AWS-Konto IDs , um ein Szenario der AWS Clean Rooms Zusammenarbeit mit mehreren Parteien zu unterstützen:

- 111122223333 — Dies ist das Konto, das die Mitgliedschaft enthält, die für die Ausführung von Abfragen verantwortlich ist (der „Job Runner“). Diese Mitgliedschaft führt die Analysejobs aus und muss diese Rolle übernehmen, um auf die erforderlichen Ressourcen zugreifen zu können.
- 444455556666 — Dies ist das Konto, dem die Analysevorlage und die zugehörige Mitgliedschaft gehören (der „Besitzer der Analysevorlage“). Diese Mitgliedschaft definiert, welche Abfragen ausgeführt werden können, und muss auch diese Rolle übernehmen, um die Analyse zu verwalten und auszuführen.

Diese Konfiguration ist typisch für AWS Clean Rooms Kollaborationen, bei denen mehrere Parteien an derselben Zusammenarbeit teilnehmen, wobei jede Partei ihre eigene AWS-Konto Mitgliedschaft hat. Sowohl der Abfrage-Executor als auch der Besitzer der Analysevorlage benötigen Zugriff auf gemeinsam genutzte Ressourcen. Die `aws:SourceArn` Bedingung stellt sicher, dass nur AWS Clean Rooms Operationen, die aus diesen beiden spezifischen Mitgliedschaften stammen, die Rolle übernehmen können, wodurch eine präzise Zugriffskontrolle für die Zusammenarbeit mit mehreren Parteien gewährleistet ist.

3. Verwenden Sie die folgende Berechtigungsrichtlinie gemäß dem Verfahren [Erstellen einer Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien \(Konsole\)](#).

Note

Die folgende Beispielrichtlinie unterstützt die Berechtigungen, die zum Lesen Ihres Codes aus Amazon S3 erforderlich sind. Je nachdem, wie Sie Ihre S3-Daten eingerichtet haben, müssen Sie diese Richtlinie jedoch möglicherweise ändern.

Ihre Amazon S3 S3-Ressourcen müssen sich in derselben Umgebung AWS-Region wie die AWS Clean Rooms Kollaboration befinden.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
```

```

        "Effect": "Allow",
        "Action": [
            "s3:GetObject",
            "s3:GetObjectVersion"
        ],
        "Resource": ["arn:aws:s3:::s3Path"],
        "Condition": {
            "StringEquals": {
                "s3:ResourceAccount": [
                    "s3BucketOwnerAccountId"
                ]
            }
        }
    }
}

```

4. Ersetzen Sie jede *placeholder* durch Ihre eigenen Informationen:

- *s3Path*— Der S3-Bucket-Speicherort Ihres Codes.
- *s3BucketOwnerAccountId*— Die AWS-Konto ID des S3-Bucket-Besitzers.
- *region* – Der Name des AWS-Region. Beispiel, **us-east-1**.
- *jobRunnerAccountId*— Die AWS-Konto ID des Mitglieds, das Abfragen und Jobs ausführen kann.
- *jobRunnerMembershipId*— Die Mitglieds-ID des Mitglieds, das Jobs abfragen und ausführen kann. Die Mitglieds-ID finden Sie auf der Registerkarte Details der Kollaboration. Dadurch AWS Clean Rooms wird sichergestellt, dass die Rolle nur übernommen wird, wenn dieses Mitglied die Analyse in dieser Kollaboration ausführt.
- *analysisTemplateAccountId*— Die AWS-Konto ID der Analysevorlage.
- *analysisTemplateOwnerMembershipId*— Die Mitglieds-ID des Mitglieds, dem die Analysevorlage gehört. Die Mitglieds-ID finden Sie auf der Registerkarte Details der Kollaboration.

5. Folgen Sie weiterhin dem Verfahren [Erstellen einer Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien \(Konsole\)](#), um die Rolle zu erstellen.

Erstellen Sie eine Servicerolle, um die Ergebnisse eines PySpark Jobs zu schreiben

AWS Clean Rooms verwendet eine Servicerolle, um die Ergebnisse eines PySpark Jobs in einen angegebenen S3-Bucket zu schreiben.

Um eine Servicerolle zu erstellen, um Ergebnisse eines PySpark Jobs zu schreiben

1. Erstellen Sie eine Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien. Weitere Informationen finden Sie unter dem Verfahren [Erstellen einer Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien \(Konsole\)](#) im AWS Identity and Access Management Benutzerhandbuch.
2. Verwenden Sie die folgende benutzerdefinierte Vertrauensrichtlinie gemäß dem Verfahren [Erstellen einer Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien \(Konsole\)](#).

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "cleanrooms.amazonaws.com"
      },
      "Action": "sts:AssumeRole",
      "Condition": {
        "ForAnyValue:ArnEquals": {
          "aws:SourceArn": [
            "arn:aws:cleanrooms:us-east-1:111122223333:membership/jobRunnerMembershipId",
            "arn:aws:cleanrooms:us-east-1:444455556666:membership/rrMembershipId"
          ]
        }
      }
    }
  ]
}
```


 Note

Diese Vertrauensrichtlinie bezieht sich auf zwei verschiedene AWS-Konto IDs , um eine AWS Clean Rooms Zusammenarbeit mit unterschiedlichen operativen Rollen zu unterstützen:

- 111122223333 — Dies ist das Konto, das die Mitgliedschaft enthält, die für die Ausführung von Analysejobs verantwortlich ist (der „Job Runner“). Diese Mitgliedschaft führt die Rechen-Workloads aus und muss diese Rolle übernehmen, um auf Verarbeitungsressourcen zugreifen zu können.
- 444455556666 — Dies ist das Konto, das die Mitgliedschaft mit den Zuständigkeiten des Ergebnisempfängers (RR) enthält. Diese Mitgliedschaft ist berechtigt, die Ergebnisse von Analysejobs zu empfangen und darauf zuzugreifen, und benötigt Rollenzugriff, um Ergebnisse an bestimmte Speicherorte zu schreiben.

Diese Konfiguration ermöglicht AWS Clean Rooms Szenarien, in denen eine Partei die rechnerische Analyse durchführt, während eine andere Partei die Ergebnisse erhält und verwaltet. Für beide Rollen sind unterschiedliche, aber sich ergänzende Berechtigungen für dieselbe Ausführungsrolle erforderlich. Diese `aws:SourceArn` Bedingung stellt sicher, dass nur AWS Clean Rooms Operationen, die aus diesen beiden spezifischen Mitgliedschaften stammen, diese Rolle übernehmen können, wodurch die Sicherheit gewährleistet und gleichzeitig die verteilte Auftragsausführung und das Ergebnismanagement unterstützt werden.

3. Verwenden Sie die folgende Berechtigungsrichtlinie gemäß dem Verfahren [Erstellen einer Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien \(Konsole\)](#).

 Note

Die folgende Beispielrichtlinie unterstützt die Berechtigungen, die zum Schreiben in Amazon S3 erforderlich sind. Je nachdem, wie Sie S3 eingerichtet haben, müssen Sie diese Richtlinie jedoch möglicherweise ändern.

Ihre Amazon S3 S3-Ressourcen müssen sich in derselben Umgebung AWS-Region wie die AWS Clean Rooms Kollaboration befinden.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:PutObject"
      ],
      "Resource": "arn:aws:s3:::bucket/optionalPrefix/*",
      "Condition": {
        "StringEquals": {
          "s3:ResourceAccount": [
            "s3BucketOwnerAccountId"
          ]
        }
      }
    },
    {
      "Effect": "Allow",
      "Action": [
        "s3:GetBucketLocation",
        "s3:ListBucket"
      ],
      "Resource": "arn:aws:s3:::bucket",
      "Condition": {
        "StringEquals": {
          "s3:ResourceAccount": [
            "s3BucketOwnerAccountId"
          ]
        }
      }
    }
  ]
}
```

4. Ersetzen Sie jede *placeholder* durch Ihre eigenen Informationen:

- *region* – Der Name des AWS-Region. Beispiel, **us-east-1**.
- *jobRunnerAccountId*— Die AWS-Konto ID, in der sich der S3-Bucket befindet.

- ***jobRunnerMembershipId***— Die Mitglieds-ID des Mitglieds, das Jobs abfragen und ausführen kann. Die Mitglieds-ID finden Sie auf der Registerkarte Details der Kollaboration. Dadurch AWS Clean Rooms wird sichergestellt, dass die Rolle nur übernommen wird, wenn dieses Mitglied die Analyse in dieser Kollaboration ausführt.
 - ***rrAccountId***— Die AWS-Konto ID, in der sich der S3-Bucket befindet.
 - ***rrMembershipId***— Die Mitglieds-ID des Mitglieds, das Ergebnisse erhalten kann. Die Mitglieds-ID finden Sie auf der Registerkarte Details der Kollaboration. Dadurch AWS Clean Rooms wird sichergestellt, dass die Rolle nur übernommen wird, wenn dieses Mitglied die Analyse in dieser Kollaboration ausführt.
 - ***bucket***— Der Name und der Speicherort des S3-Buckets.
 - ***optionalPrefix***— Ein optionales Präfix, wenn Sie Ihre Ergebnisse unter einem bestimmten S3-Präfix speichern möchten.
 - ***s3BucketOwnerAccountId***— Die AWS-Konto ID des S3-Bucket-Besitzers.
5. Folgen Sie weiterhin dem Verfahren [Erstellen einer Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien \(Konsole\)](#), um die Rolle zu erstellen.

Erstellen Sie eine Servicerolle, um Ergebnisse zu erhalten

Note

Wenn Sie das Mitglied sind, das nur Ergebnisse erhalten kann (in der Konsole ist Ihre Mitgliederfähigkeit nur Ergebnisse erhalten aktiviert), gehen Sie wie folgt vor.

Wenn Sie ein Mitglied sind, das Ergebnisse sowohl abfragen als auch empfangen kann (in der Konsole ist Ihre Fähigkeit als Mitglied sowohl Ergebnisse abfragen als auch Ergebnisse erhalten), können Sie dieses Verfahren überspringen.

Für Kollaborationsmitglieder, die nur Ergebnisse empfangen können, AWS Clean Rooms verwendet eine Servicerolle, um die Ergebnisse der abgefragten Daten in der Kollaboration in den angegebenen S3-Bucket zu schreiben.

Es gibt zwei Möglichkeiten, diese Servicerolle zu erstellen:

- Wenn Sie über die erforderlichen IAM-Berechtigungen zum Erstellen einer Servicerolle verfügen, verwenden Sie die AWS Clean Rooms Konsole, um eine Servicerolle zu erstellen.

- Wenn Sie nicht über die `iam:AttachRolePolicy` erforderlichen Berechtigungen verfügen `iam:CreateRole` oder die IAM-Rollen manuell erstellen möchten, gehen Sie wie folgt vor:
`iam:CreatePolicy`
- Gehen Sie wie folgt vor, um eine Servicerolle mithilfe benutzerdefinierter Vertrauensrichtlinien zu erstellen.
- Bitten Sie Ihren Administrator, die Servicerolle mithilfe des folgenden Verfahrens zu erstellen.

Note

Sie oder Ihr IAM-Administrator sollten dieses Verfahren nur befolgen, wenn Sie nicht über die erforderlichen Berechtigungen verfügen, um mithilfe der AWS Clean Rooms Konsole eine Servicerolle zu erstellen.

Um mithilfe benutzerdefinierter Vertrauensrichtlinien eine Servicerolle zu erstellen, um Ergebnisse zu erhalten

1. Erstellen Sie eine Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien. Weitere Informationen finden Sie unter dem Verfahren [Erstellen einer Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien \(Konsole\)](#) im AWS Identity and Access Management Benutzerhandbuch.
2. Verwenden Sie die folgende benutzerdefinierte Vertrauensrichtlinie gemäß dem Verfahren [Erstellen einer Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien \(Konsole\)](#).

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowIfExternalIdMatches",
      "Effect": "Allow",
      "Principal": {
        "Service": "cleanrooms.amazonaws.com"
      },
      "Action": "sts:AssumeRole",
      "Condition": {
        "ArnLike": {
```

```

        "sts:ExternalId":
        "arn:aws:*:region:*:dbuser:*/a1b2c3d4-5678-90ab-cdef-EXAMPLEaaaaa*"
    },
    {
        "Sid": "AllowIfSourceArnMatches",
        "Effect": "Allow",
        "Principal": {
            "Service": "cleanrooms.amazonaws.com"
        },
        "Action": "sts:AssumeRole",
        "Condition": {
            "ForAnyValue:ArnEquals": {
                "aws:SourceArn": [
                    "arn:aws:cleanrooms:us-east-1:555555555555:membership/a1b2c3d4-5678-90ab-cdef-EXAMPLEaaaaa"
                ]
            }
        }
    }
]
}

```

3. Verwenden Sie die folgende Berechtigungsrichtlinie gemäß dem Verfahren [Erstellen einer Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien \(Konsole\)](#).

Note

Die folgende Beispielrichtlinie unterstützt die Berechtigungen, die zum Lesen von AWS Glue Metadaten und den entsprechenden Amazon S3 S3-Daten erforderlich sind. Je nachdem, wie Sie Ihre S3-Daten eingerichtet haben, müssen Sie diese Richtlinie jedoch möglicherweise ändern.

Ihre AWS Glue Ressourcen und die zugrunde liegenden Amazon S3 S3-Ressourcen müssen mit der AWS Clean Rooms Zusammenarbeit AWS-Region identisch sein.

JSON

```
{
```

```

"Version": "2012-10-17",
"Statement": [
  {
    "Effect": "Allow",
    "Action": [
      "s3:GetBucketLocation",
      "s3:ListBucket"
    ],
    "Resource": [
      "arn:aws:s3:::bucket_name"
    ],
    "Condition": {
      "StringEquals": {
        "aws:ResourceAccount": "accountId"
      }
    }
  },
  {
    "Effect": "Allow",
    "Action": [
      "s3:PutObject"
    ],
    "Resource": [
      "arn:aws:s3:::bucket_name/optional_key_prefix/*"
    ],
    "Condition": {
      "StringEquals": {
        "aws:ResourceAccount": "accountId"
      }
    }
  }
]
}

```

4. Ersetzen Sie jede *placeholder* durch Ihre eigenen Informationen:

- *region* – Der Name des AWS-Region. Beispiel, **us-east-1**.
- *a1b2c3d4-5678-90ab-cdef-EXAMPLEaaaaa*— Die Mitglieds-ID des Mitglieds, das Abfragen durchführen kann. Die Mitglieds-ID finden Sie auf der Registerkarte Details der Kollaboration. Dadurch AWS Clean Rooms wird sichergestellt, dass die Rolle nur übernommen wird, wenn dieses Mitglied die Analyse in dieser Kollaboration ausführt.

- *arn:aws:cleanrooms:us-east-1:555555555555:membership/a1b2c3d4-5678-90ab-cdef-EXAMPLEaaaaa*— Der einzige Mitgliedschafts-ARN des Mitglieds, das Abfragen durchführen kann. Den Mitglieds-ARN finden Sie auf der Registerkarte Details der Kollaboration. Dadurch AWS Clean Rooms wird sichergestellt, dass die Rolle nur übernommen wird, wenn dieses Mitglied die Analyse in dieser Kollaboration ausführt.
- *bucket_name*— Der Amazon-Ressourcenname (ARN) des S3-Buckets. Den Amazon-Ressourcenamen (ARN) finden Sie auf der Registerkarte Eigenschaften des Buckets in Amazon S3.
- *accountId*— Die AWS-Konto ID, in der sich der S3-Bucket befindet.

bucket_name/optional_key_prefix— Der Amazon-Ressourcenname (ARN) des Ergebnisziels in Amazon S3. Den Amazon-Ressourcenamen (ARN) finden Sie auf der Registerkarte Eigenschaften des Buckets in Amazon S3.

5. Folgen Sie weiterhin dem Verfahren zum [Erstellen einer Rolle mithilfe benutzerdefinierter Vertrauensrichtlinien \(Konsole\)](#), um die Rolle zu erstellen.

Servicerollen für AWS Clean Rooms ML einrichten

Die Rollen, die für die Durchführung der Lookalike-Modellierung benötigt werden, unterscheiden sich von denen, die für die Verwendung eines benutzerdefinierten Modells erforderlich sind. In den folgenden Abschnitten werden die Rollen beschrieben, die zur Ausführung der einzelnen Aufgaben benötigt werden.

Themen

- [Richten Sie Servicerollen für die Lookalike-Modellierung ein](#)
- [Richten Sie Servicerollen für die benutzerdefinierte Modellierung ein](#)

Richten Sie Servicerollen für die Lookalike-Modellierung ein

Themen

- [Erstellen Sie eine Servicerolle zum Lesen von Trainingsdaten](#)
- [Erstellen Sie eine Servicerolle, um ein Lookalike-Segment zu schreiben](#)
- [Erstellen Sie eine Servicerolle zum Lesen von Startdaten](#)

Erstellen Sie eine Servicerolle zum Lesen von Trainingsdaten

AWS Clean Rooms verwendet eine Servicerolle, um Trainingsdaten zu lesen. Sie können diese Rolle mithilfe der Konsole erstellen, wenn Sie über die erforderlichen IAM-Berechtigungen verfügen. Wenn Sie keine `CreateRole` Berechtigungen haben, bitten Sie Ihren Administrator, die Servicerolle zu erstellen.

Um eine Servicerolle zum Trainieren eines Datensatzes zu erstellen

1. Melden Sie sich mit Ihrem Administratorkonto bei der IAM-Konsole (<https://console.aws.amazon.com/iam/>) an.
2. Wählen Sie unter Access management (Zugriffsverwaltung) Policies (Richtlinien) aus.
3. Wählen Sie Richtlinie erstellen aus.
4. Wählen Sie im Richtlinien-Editor die Registerkarte JSON aus und kopieren Sie dann die folgende Richtlinie und fügen Sie sie ein.

Note

Die folgende Beispielrichtlinie unterstützt die Berechtigungen, die zum Lesen von AWS Glue Metadaten und den entsprechenden Amazon S3 S3-Daten erforderlich sind. Je nachdem, wie Sie Ihre S3-Daten eingerichtet haben, müssen Sie diese Richtlinie jedoch möglicherweise ändern. Diese Richtlinie beinhaltet keinen KMS-Schlüssel zum Entschlüsseln von Daten.

Ihre AWS Glue Ressourcen und die zugrunde liegenden Amazon S3 S3-Ressourcen müssen mit der AWS Clean Rooms Zusammenarbeit AWS-Region identisch sein.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "glue:GetDatabase",
        "glue:GetDatabases",
        "glue:GetTable",
```



```

        "glue:GetTables",
        "glue:GetPartitions",
        "glue:GetPartition",
        "glue:BatchGetPartition",
        "glue:GetUserDefinedFunctions"
    ],
    "Resource": [
        "arn:aws:glue:us-east-1:111122223333:database/databases",
        "arn:aws:glue:us-east-1:111122223333:table/databases/tables",
        "arn:aws:glue:us-east-1:111122223333:catalog",
        "arn:aws:glue:us-east-1:111122223333:database/default"
    ]
},
{
    "Effect": "Allow",
    "Action": [
        "glue:CreateDatabase"
    ],
    "Resource": [
        "arn:aws:glue:us-east-1:111122223333:database/default"
    ]
},
{
    "Effect": "Allow",
    "Action": [
        "s3:ListBucket",
        "s3:GetBucketLocation"
    ],
    "Resource": [
        "arn:aws:s3::bucket"
    ],
    "Condition": {
        "StringEquals": {
            "s3:ResourceAccount": [
                "111122223333"
            ]
        }
    }
},
{
    "Effect": "Allow",
    "Action": [
        "s3:GetObject"
    ],

```

```

        "Resource": [
            "arn:aws:s3:::bucketFolders/*"
        ],
        "Condition": {
            "StringEquals": {
                "s3:ResourceAccount": [
                    "111122223333"
                ]
            }
        }
    }
}

```

Wenn Sie einen KMS-Schlüssel zum Entschlüsseln von Daten verwenden müssen, fügen Sie diese AWS KMS Anweisung zur vorherigen Vorlage hinzu:

```

{
    "Effect": "Allow",
    "Action": [
        "kms:Decrypt",
    ],
    "Resource": [
        "arn:aws:kms:region:accountId:key/keyId"
    ],
    "Condition": {
        "ArnLike": {
            "kms:EncryptionContext:aws:s3:arn":
                "arn:aws:s3:::bucketFolders*"
        }
    }
}

```

5. Ersetzen Sie jede *placeholder* durch Ihre eigenen Informationen:

- *region* – Der Name des AWS-Region. Beispiel, **us-east-1**.
- *accountId*— Die AWS-Konto ID, in der sich der S3-Bucket befindet.

- *database/databases*, *table/databases/tablescatalog*, und *database/default* — Der Speicherort der Trainingsdaten, auf die zugegriffen AWS Clean Rooms werden muss.
- *bucket* — Der Amazon-Ressourcenname (ARN) des S3-Buckets. Den Amazon-Ressourcenamen (ARN) finden Sie auf der Registerkarte Eigenschaften des Buckets in Amazon S3.
- *bucketFolders* — Der Name bestimmter Ordner im S3-Bucket, auf die zugegriffen AWS Clean Rooms werden muss.

6. Wählen Sie Weiter aus.
7. Geben Sie unter Überprüfen und erstellen einen Richtliniennamen und eine Beschreibung ein und überprüfen Sie die Zusammenfassung.
8. Wählen Sie Richtlinie erstellen aus.

Sie haben eine Richtlinie für erstellt AWS Clean Rooms.

9. Wählen Sie unter Access management (Zugriffsverwaltung) Roles (Rollen) aus.

Mit Rollen können Sie kurzfristige Anmeldeinformationen erstellen, was aus Sicherheitsgründen empfohlen wird. Sie können auch Benutzer auswählen, um langfristige Anmeldeinformationen zu erstellen.

10. Wählen Sie Rolle erstellen aus.
11. Wählen Sie im Assistenten zum Erstellen von Rollen unter Vertrauenswürdiger Entitätstyp die Option Benutzerdefinierte Vertrauensrichtlinie aus.
12. Kopieren Sie die folgende benutzerdefinierte Vertrauensrichtlinie und fügen Sie sie in den JSON-Editor ein.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowAssumeRole",
      "Effect": "Allow",
      "Principal": {
        "Service": "cleanrooms-ml.amazonaws.com"
      },
      "Action": "sts:AssumeRole",
      "Condition": {
```

```

        "StringEqualsIfExists": {
            "aws:SourceAccount": ["111122223333"]
        },
        "ArnLikeIfExists": {
            "aws:SourceArn": "arn:aws:cleanrooms-ml:us-
east-1:111122223333:training-dataset/*"
        }
    }
}
]
}

```

Das SourceAccount ist immer dein AWS-Konto. Der SourceArn kann auf einen bestimmten Trainingsdatensatz beschränkt werden, jedoch erst, nachdem dieser Datensatz erstellt wurde. Da Sie den ARN des Trainingsdatensatzes noch nicht kennen, wird hier der Platzhalter angegeben.

accountId ist die ID AWS-Konto, die die Trainingsdaten enthält.

13. Wählen Sie Weiter und geben Sie unter Berechtigungen hinzufügen den Namen der Richtlinie ein, die Sie gerade erstellt haben. (Möglicherweise müssen Sie die Seite neu laden.)
14. Aktivieren Sie das Kontrollkästchen neben dem Namen der Richtlinie, die Sie erstellt haben, und klicken Sie dann auf Weiter.
15. Geben Sie unter Name, review and create den Rollennamen und die Beschreibung ein.

Note

Der Rollename muss dem Muster in den passRole Berechtigungen entsprechen, die dem Mitglied erteilt wurden, das Ergebnisse abfragen und empfangen kann, und den Mitgliedsrollen.

- a. Überprüfen Sie die Option Vertrauenswürdige Entitäten auswählen und bearbeiten Sie sie gegebenenfalls.
- b. Überprüfen Sie die Berechtigungen unter Berechtigungen hinzufügen und bearbeiten Sie sie gegebenenfalls.
- c. Überprüfen Sie die Tags und fügen Sie bei Bedarf Stichwörter hinzu.

- d. Wählen Sie Rolle erstellen aus.

Sie haben die Servicerolle für erstellt AWS Clean Rooms.

Erstellen Sie eine Servicerolle, um ein Lookalike-Segment zu schreiben

AWS Clean Rooms verwendet eine Servicerolle, um Lookalike-Segmente in einen Bucket zu schreiben. Sie können diese Rolle mithilfe der Konsole erstellen, wenn Sie über die erforderlichen IAM-Berechtigungen verfügen. Wenn Sie keine `CreateRole` Berechtigungen haben, bitten Sie Ihren Administrator, die Servicerolle zu erstellen.

Um eine Servicerolle zu erstellen, um ein Lookalike-Segment zu schreiben

1. Melden Sie sich mit Ihrem Administratorkonto bei der IAM-Konsole (<https://console.aws.amazon.com/iam/>) an.
2. Wählen Sie unter Access management (Zugriffsverwaltung) Policies (Richtlinien) aus.
3. Wählen Sie Richtlinie erstellen aus.
4. Wählen Sie im Richtlinien-Editor die Registerkarte JSON aus und kopieren Sie dann die folgende Richtlinie und fügen Sie sie ein.

Note

Die folgende Beispielrichtlinie unterstützt die Berechtigungen, die zum Lesen von AWS Glue Metadaten und den entsprechenden Amazon S3 S3-Daten erforderlich sind. Je nachdem, wie Sie Ihre Amazon S3 S3-Daten eingerichtet haben, müssen Sie diese Richtlinie jedoch möglicherweise ändern. Diese Richtlinie beinhaltet keinen KMS-Schlüssel zum Entschlüsseln von Daten.

Ihre AWS Glue Ressourcen und die zugrunde liegenden Amazon S3 S3-Ressourcen müssen mit der AWS Clean Rooms Zusammenarbeit AWS-Region identisch sein.

JSON

```
{  
  "Version": "2012-10-17",  
  "Statement": [  
    {
```

```

        "Effect": "Allow",
        "Action": [
            "s3:ListBucket",
            "s3:GetBucketLocation"
        ],
        "Resource": [
            "arn:aws:s3:::buckets"
        ],
        "Condition": {
            "StringEquals": {
                "s3:ResourceAccount": [
                    "accountId"
                ]
            }
        }
    },
    {
        "Effect": "Allow",
        "Action": [
            "s3:PutObject"
        ],
        "Resource": [
            "arn:aws:s3:::bucketFolders/*"
        ],
        "Condition": {
            "StringEquals": {
                "s3:ResourceAccount": [
                    "accountId"
                ]
            }
        }
    }
}

```

Wenn Sie einen KMS-Schlüssel zum Verschlüsseln von Daten verwenden müssen, fügen Sie der Vorlage diese AWS KMS Anweisung hinzu:

```

{
    "Effect": "Allow",
    "Action": [
        "kms:Encrypt",

```

```

        "kms:GenerateDataKey*",
        "kms:ReEncrypt*",
    ],
    "Resource": [
        "arn:aws:kms:region:accountId:key/keyId"
    ],
    "Condition": {
        "ArnLike": {
            "kms:EncryptionContext:aws:s3:arn":
                "arn:aws:s3:::bucketFolders*"
        }
    }
}
]
}

```

5. Ersetzen Sie jede *placeholder* durch Ihre eigenen Informationen:

- *buckets*— Der Amazon-Ressourcenname (ARN) des S3-Buckets. Den Amazon-Ressourcenamen (ARN) finden Sie auf der Registerkarte Eigenschaften des Buckets in Amazon S3.
- *accountId*— Die AWS-Konto ID, in der sich der S3-Bucket befindet.
- *bucketFolders*— Der Name bestimmter Ordner im S3-Bucket, auf die zugegriffen AWS Clean Rooms werden muss.
- *region* – Der Name des AWS-Region. Beispiel, **us-east-1**.
- *keyId*— Der KMS-Schlüssel, der zur Verschlüsselung Ihrer Daten benötigt wird.

6. Wählen Sie Weiter aus.

7. Geben Sie unter Überprüfen und erstellen einen Richtliniennamen und eine Beschreibung ein und überprüfen Sie die Zusammenfassung.

8. Wählen Sie Richtlinie erstellen aus.

Sie haben eine Richtlinie für erstellt AWS Clean Rooms.

9. Wählen Sie unter Access management (Zugriffsverwaltung) Roles (Rollen) aus.

Mit Rollen können Sie kurzfristige Anmeldeinformationen erstellen, was aus Sicherheitsgründen empfohlen wird. Sie können auch Benutzer auswählen, um langfristige Anmeldeinformationen zu erstellen.

10. Wählen Sie Rolle erstellen aus.

11. Wählen Sie im Assistenten zum Erstellen von Rollen unter Vertrauenswürdiger Entitätstyp die Option Benutzerdefinierte Vertrauensrichtlinie aus.
12. Kopieren Sie die folgende benutzerdefinierte Vertrauensrichtlinie und fügen Sie sie in den JSON-Editor ein.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowAssumeRole",
      "Effect": "Allow",
      "Principal": {
        "Service": "cleanrooms-ml.amazonaws.com"
      },
      "Action": "sts:AssumeRole",
      "Condition": {
        "StringEqualsIfExists": {
          "aws:SourceAccount": ["111122223333"]
        },
        "ArnLikeIfExists": {
          "aws:SourceArn": "arn:aws:cleanrooms-ml:us-east-1:111122223333:configured-audience-model/*"
        }
      }
    }
  ]
}
```

Das SourceAccount ist immer dein AWS-Konto. Der SourceArn kann auf einen bestimmten Trainingsdatensatz beschränkt werden, jedoch erst, nachdem dieser Datensatz erstellt wurde. Da Sie den ARN des Trainingsdatensatzes noch nicht kennen, wird hier der Platzhalter angegeben.

13. Wählen Sie Weiter aus.
14. Aktivieren Sie das Kontrollkästchen neben dem Namen der Richtlinie, die Sie erstellt haben, und wählen Sie dann Weiter aus.

15. Geben Sie unter Name, review and create den Rollennamen und die Beschreibung ein.

 Note

Der Rollenname muss dem Muster in den `passRole` Berechtigungen entsprechen, die dem Mitglied erteilt wurden, das Ergebnisse abfragen und empfangen kann, und den Mitgliedsrollen.

- a. Überprüfen Sie die Option Vertrauenswürdige Entitäten auswählen und bearbeiten Sie sie gegebenenfalls.
- b. Überprüfen Sie die Berechtigungen unter Berechtigungen hinzufügen und bearbeiten Sie sie gegebenenfalls.
- c. Überprüfen Sie die Tags und fügen Sie bei Bedarf Stichwörter hinzu.
- d. Wählen Sie Rolle erstellen aus.

Sie haben die Servicerolle für erstellt AWS Clean Rooms.

Erstellen Sie eine Servicerolle zum Lesen von Startdaten

AWS Clean Rooms verwendet eine Servicerolle, um Seed-Daten zu lesen. Sie können diese Rolle mithilfe der Konsole erstellen, wenn Sie über die erforderlichen IAM-Berechtigungen verfügen. Wenn Sie keine `CreateRole` Berechtigungen haben, bitten Sie Ihren Administrator, die Servicerolle zu erstellen.

Um eine Servicerolle zum Lesen von Seed-Daten zu erstellen, die in einem S3-Bucket gespeichert sind.

1. Melden Sie sich mit Ihrem Administratorkonto bei der IAM-Konsole (<https://console.aws.amazon.com/iam/>) an.
2. Wählen Sie unter Access management (Zugriffsverwaltung) Policies (Richtlinien) aus.
3. Wählen Sie Richtlinie erstellen aus.
4. Wählen Sie im Richtlinien-Editor die Registerkarte JSON aus und kopieren Sie dann eine der folgenden Richtlinien und fügen Sie sie ein.

 Note

Die folgende Beispielrichtlinie unterstützt die Berechtigungen, die zum Lesen von AWS Glue Metadaten und den entsprechenden Amazon S3 S3-Daten erforderlich sind. Je nachdem, wie Sie Ihre Amazon S3 S3-Daten eingerichtet haben, müssen Sie diese Richtlinie jedoch möglicherweise ändern. Diese Richtlinie beinhaltet keinen KMS-Schlüssel zum Entschlüsseln von Daten.

Ihre AWS Glue Ressourcen und die zugrunde liegenden Amazon S3 S3-Ressourcen müssen mit der AWS Clean Rooms Zusammenarbeit AWS-Region identisch sein.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:ListBucket"
      ],
      "Resource": [
        "arn:aws:s3:::buckets"
      ],
      "Condition": {
        "StringEquals": {
          "s3:ResourceAccount": [
            "accountId"
          ]
        }
      }
    },
    {
      "Effect": "Allow",
      "Action": [
        "s3:GetObject"
      ],
      "Resource": [
        "arn:aws:s3:::bucketFolders/*"
      ],
    }
  ]
}
```

```

        "Condition": {
            "StringEquals": {
                "s3:ResourceAccount": [
                    "accountId"
                ]
            }
        }
    }
}

```

Note

Die folgende Beispielrichtlinie unterstützt die Berechtigungen, die erforderlich sind, um die Ergebnisse einer SQL-Abfrage zu lesen und diese als Eingabedaten zu verwenden. Je nachdem, wie Ihre Abfrage strukturiert ist, müssen Sie diese Richtlinie jedoch möglicherweise ändern. Diese Richtlinie beinhaltet keinen KMS-Schlüssel zum Entschlüsseln von Daten.

JSON

```

{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "AllowCleanRoomsStartQuery",
            "Effect": "Allow",
            "Action": [
                "cleanrooms:GetCollaborationAnalysisTemplate",
                "cleanrooms:GetSchema",
                "cleanrooms:StartProtectedQuery"
            ],
            "Resource": "*"
        },
        {
            "Sid": "AllowCleanRoomsGetAndUpdateQuery",
            "Effect": "Allow",
            "Action": [
                "cleanrooms:GetProtectedQuery",

```

```

        "cleanrooms:UpdateProtectedQuery"
      ],
      "Resource": [
        "arn:aws:cleanrooms:us-east-1:111122223333:membership/queryRunnerMembershipId"
      ]
    }
  ]
}

```

Wenn Sie einen KMS-Schlüssel zum Entschlüsseln von Daten verwenden müssen, fügen Sie der Vorlage diese AWS KMS Anweisung hinzu:

```

{
    "Effect": "Allow",
    "Action": [
        "kms:Decrypt",
        "kms:DescribeKey"
    ],
    "Resource": [
        "arn:aws:kms:region:accountId:key/keyId"
    ],
    "Condition": {
        "ArnLike": {
            "kms:EncryptionContext:aws:s3:arn":
"arn:aws:s3:::bucketFolders*"
        }
    }
}
]
}

```

5. Ersetzen Sie jede *placeholder* durch Ihre eigenen Informationen:

- *buckets*— Der Amazon-Ressourcenname (ARN) des S3-Buckets. Den Amazon-Ressourcenname (ARN) finden Sie auf der Registerkarte Eigenschaften des Buckets in Amazon S3.
- *accountId*— Die AWS-Konto ID, in der sich der S3-Bucket befindet.
- *bucketFolders*— Der Name bestimmter Ordner im S3-Bucket, auf die zugegriffen AWS Clean Rooms werden muss.

- **region** – Der Name des AWS-Region. Beispiel, **us-east-1**.
- **queryRunnerAccountId**— Die AWS-Konto ID des Kontos, das Abfragen ausführt.
- **queryRunnerMembershipId**— Die Mitglieds-ID des Mitglieds, das Abfragen durchführen kann. Die Mitglieds-ID finden Sie auf der Registerkarte Details der Kollaboration. Dadurch AWS Clean Rooms wird sichergestellt, dass die Rolle nur übernommen wird, wenn dieses Mitglied die Analyse in dieser Kollaboration ausführt.
- **keyId**— Der KMS-Schlüssel, der zur Verschlüsselung Ihrer Daten benötigt wird.

6. Wählen Sie Weiter aus.
7. Geben Sie unter Überprüfen und erstellen einen Richtliniennamen und eine Beschreibung ein und überprüfen Sie die Zusammenfassung.
8. Wählen Sie Richtlinie erstellen aus.

Sie haben eine Richtlinie für erstellt AWS Clean Rooms.

9. Wählen Sie unter Access management (Zugriffsverwaltung) Roles (Rollen) aus.

Mit Rollen können Sie kurzfristige Anmeldeinformationen erstellen, was aus Sicherheitsgründen empfohlen wird. Sie können auch Benutzer auswählen, um langfristige Anmeldeinformationen zu erstellen.

10. Wählen Sie Rolle erstellen aus.
11. Wählen Sie im Assistenten zum Erstellen von Rollen unter Vertrauenswürdiger Entitätstyp die Option Benutzerdefinierte Vertrauensrichtlinie aus.
12. Kopieren Sie die folgende benutzerdefinierte Vertrauensrichtlinie und fügen Sie sie in den JSON-Editor ein.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowAssumeRole",
      "Effect": "Allow",
      "Principal": {
        "Service": "cleanrooms-ml.amazonaws.com"
      },
      "Action": "sts:AssumeRole",
```

```
    "Condition": {
      "StringEqualsIfExists": {
        "aws:SourceAccount": ["111122223333"]
      },
      "ArnLikeIfExists": {
        "aws:SourceArn": "arn:aws:cleanrooms-m1:us-  
east-1:111122223333:audience-generation-job/*"
      }
    }
  }
}
```

Das SourceAccount ist immer dein AWS-Konto. Der SourceArn kann auf einen bestimmten Trainingsdatensatz beschränkt werden, jedoch erst, nachdem dieser Datensatz erstellt wurde. Da Sie den ARN des Trainingsdatensatzes noch nicht kennen, wird hier der Platzhalter angegeben.

13. Wählen Sie Weiter aus.
14. Aktivieren Sie das Kontrollkästchen neben dem Namen der Richtlinie, die Sie erstellt haben, und wählen Sie dann Weiter aus.
15. Geben Sie unter Name, review and create den Rollennamen und die Beschreibung ein.

 Note

Der Rollename muss dem Muster in den passRole Berechtigungen entsprechen, die dem Mitglied erteilt wurden, das Ergebnisse abfragen und empfangen kann, und den Mitgliedsrollen.

- a. Überprüfen Sie die Option Vertrauenswürdige Entitäten auswählen und bearbeiten Sie sie gegebenenfalls.
- b. Überprüfen Sie die Berechtigungen unter Berechtigungen hinzufügen und bearbeiten Sie sie gegebenenfalls.
- c. Überprüfen Sie die Tags und fügen Sie bei Bedarf Stichwörter hinzu.
- d. Wählen Sie Rolle erstellen aus.

Sie haben die Servicerolle für erstellt AWS Clean Rooms.

Richten Sie Servicerollen für die benutzerdefinierte Modellierung ein

Themen

- [Erstellen Sie eine Servicerolle für die benutzerdefinierte ML-Modellierung — ML-Konfiguration](#)
- [Erstellen Sie eine Servicerolle, um ein benutzerdefiniertes ML-Modell bereitzustellen](#)
- [Erstellen Sie eine Servicerolle, um einen Datensatz abzufragen](#)
- [Erstellen Sie eine Servicerolle, um eine konfigurierte Tabellenzuordnung zu erstellen](#)

Erstellen Sie eine Servicerolle für die benutzerdefinierte ML-Modellierung — ML-Konfiguration

AWS Clean Rooms verwendet eine Servicerolle, um zu steuern, wer eine benutzerdefinierte ML-Konfiguration erstellen kann. Sie können diese Rolle mithilfe der Konsole erstellen, wenn Sie über die erforderlichen IAM-Berechtigungen verfügen. Wenn Sie keine `CreateRole` Berechtigungen haben, bitten Sie Ihren Administrator, die Servicerolle zu erstellen.

Mit dieser Rolle können Sie die [MLConfigurationPut-Aktion](#) verwenden.

Um eine Servicerolle zu erstellen, um die Erstellung einer benutzerdefinierten ML-Konfiguration zu ermöglichen

1. Melden Sie sich mit Ihrem Administratorkonto bei der IAM-Konsole (<https://console.aws.amazon.com/iam/>) an.
2. Wählen Sie unter Access management (Zugriffsverwaltung) Policies (Richtlinien) aus.
3. Wählen Sie Richtlinie erstellen aus.
4. Wählen Sie im Richtlinien-Editor die Registerkarte JSON aus und kopieren Sie dann die folgende Richtlinie und fügen Sie sie ein.

Note

Die folgende Beispielrichtlinie unterstützt die Berechtigungen, die für den Zugriff auf und das Schreiben von Daten in einen S3-Bucket sowie für die Veröffentlichung von CloudWatch Metriken erforderlich sind. Je nachdem, wie Sie Ihre Amazon S3 S3-Daten eingerichtet haben, müssen Sie diese Richtlinie jedoch möglicherweise ändern. Diese Richtlinie beinhaltet keinen KMS-Schlüssel zum Entschlüsseln von Daten.

Ihre Amazon S3 S3-Ressourcen müssen sich in derselben Umgebung AWS-Region wie die AWS Clean Rooms Kollaboration befinden.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowS3ObjectWriteForExport",
      "Effect": "Allow",
      "Action": [
        "s3:PutObject"
      ],
      "Resource": [
        "arn:aws:s3:::bucket/*"
      ],
      "Condition": {
        "StringEquals": {
          "s3:ResourceAccount": [
            "111122223333"
          ]
        }
      }
    },
    {
      "Sid": "AllowS3KMSEncryptForExport",
      "Effect": "Allow",
      "Action": [
        "kms:Encrypt",
        "kms:GenerateDataKey*"
      ],
      "Resource": [
        "arn:aws:kms:us-east-1:111122223333:key/keyId"
      ],
      "Condition": {
        "StringLike": {
          "kms:EncryptionContext:aws:s3:arn":
            "arn:aws:s3:::bucket*"
        }
      }
    }
  ]
}
```



```

    },
    {
      "Sid": "AllowCloudWatchMetricsPublishingForTrainingJobs",
      "Action": "cloudwatch:PutMetricData",
      "Resource": "*",
      "Effect": "Allow",
      "Condition": {
        "StringLike": {
          "cloudwatch:namespace": "/aws/cleanroomsml/*"
        }
      }
    },
  ],
  {
    "Sid": "AllowCloudWatchLogsPublishingForTrainingOrInferenceJobs",
    "Effect": "Allow",
    "Action": [
      "logs:CreateLogGroup",
      "logs:CreateLogStream",
      "logs:DescribeLogStreams",
      "logs:PutLogEvents"
    ],
    "Resource": [
      "arn:aws:logs:us-east-1:111122223333:log-group:/aws/cleanroomsml/*"
    ]
  }
]
}

```

5. Ersetzen Sie jede *placeholder* durch Ihre eigenen Informationen:

- *bucket*— Der Amazon-Ressourcenname (ARN) des S3-Buckets. Den Amazon-Ressourcenamen (ARN) finden Sie auf der Registerkarte Eigenschaften des Buckets in Amazon S3.
- *region* – Der Name des AWS-Region. Beispiel, **us-east-1**.
- *accountId*— Die AWS-Konto ID, in der sich der S3-Bucket befindet.
- *keyId*— Der KMS-Schlüssel, der zur Verschlüsselung Ihrer Daten benötigt wird.

6. Wählen Sie Weiter aus.

7. Geben Sie unter Überprüfen und erstellen einen Richtliniennamen und eine Beschreibung ein und überprüfen Sie die Zusammenfassung.

8. Wählen Sie Richtlinie erstellen aus.

Sie haben eine Richtlinie für erstellt AWS Clean Rooms.

9. Wählen Sie unter Access management (Zugriffsverwaltung) Roles (Rollen) aus.

Mit Rollen können Sie kurzfristige Anmeldeinformationen erstellen, was aus Sicherheitsgründen empfohlen wird. Sie können auch Benutzer auswählen, um langfristige Anmeldeinformationen zu erstellen.

10. Wählen Sie Rolle erstellen aus.
11. Wählen Sie im Assistenten zum Erstellen von Rollen unter Vertrauenswürdiger Entitätstyp die Option Benutzerdefinierte Vertrauensrichtlinie aus.
12. Kopieren Sie die folgende benutzerdefinierte Vertrauensrichtlinie und fügen Sie sie in den JSON-Editor ein.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "cleanrooms-ml.amazonaws.com"
      },
      "Action": "sts:AssumeRole",
      "Condition": {
        "StringEquals": {
          "aws:SourceAccount": "111122223333"
        },
        "ArnLike": {
          "aws:SourceArn": "arn:aws:cleanrooms:us-east-1:111122223333:membership/membershipID"
        }
      }
    }
  ]
}
```

Das SourceAccount ist immer dein AWS-Konto. Der SourceArn kann auf einen bestimmten Trainingsdatensatz beschränkt werden, jedoch erst, nachdem dieser Datensatz erstellt wurde.

Da Sie den ARN des Trainingsdatensatzes noch nicht kennen, wird hier der Platzhalter angegeben.

13. Wählen Sie Weiter aus.
14. Aktivieren Sie das Kontrollkästchen neben dem Namen der Richtlinie, die Sie erstellt haben, und wählen Sie dann Weiter aus.
15. Geben Sie unter Name, review and create den Rollennamen und die Beschreibung ein.

 Note

Der Rollename muss dem Muster in den `passRole` Berechtigungen entsprechen, die dem Mitglied erteilt wurden, das Ergebnisse abfragen und empfangen kann, und den Mitgliedsrollen.

- a. Überprüfen Sie die Option Vertrauenswürdige Entitäten auswählen und bearbeiten Sie sie gegebenenfalls.
- b. Überprüfen Sie die Berechtigungen unter Berechtigungen hinzufügen und bearbeiten Sie sie gegebenenfalls.
- c. Überprüfen Sie die Tags und fügen Sie bei Bedarf Stichwörter hinzu.
- d. Wählen Sie Rolle erstellen aus.

Sie haben die Servicerolle für erstellt AWS Clean Rooms.

Erstellen Sie eine Servicerolle, um ein benutzerdefiniertes ML-Modell bereitzustellen

AWS Clean Rooms verwendet eine Servicerolle, um zu steuern, wer einen benutzerdefinierten ML-Modellalgorithmus erstellen kann. Sie können diese Rolle mithilfe der Konsole erstellen, wenn Sie über die erforderlichen IAM-Berechtigungen verfügen. Wenn Sie keine `CreateRole` Berechtigungen haben, bitten Sie Ihren Administrator, die Servicerolle zu erstellen.

Mit dieser Rolle können Sie die [CreateConfiguredModelAlgorithm](#)Aktion verwenden.

Um eine Servicerolle zu erstellen, die es einem Mitglied ermöglicht, ein benutzerdefiniertes ML-Modell bereitzustellen

1. Melden Sie sich mit Ihrem Administratorkonto bei der IAM-Konsole (<https://console.aws.amazon.com/iam/>) an.

2. Wählen Sie unter Access management (Zugriffsverwaltung) Policies (Richtlinien) aus.
3. Wählen Sie Richtlinie erstellen aus.
4. Wählen Sie im Richtlinien-Editor die Registerkarte JSON aus und kopieren Sie dann die folgende Richtlinie und fügen Sie sie ein.

Note

Die folgende Beispielrichtlinie unterstützt die Berechtigungen, die zum Abrufen des Docker-Images erforderlich sind, das den Modellalgorithmus enthält. Je nachdem, wie Sie Ihre Amazon S3 S3-Daten eingerichtet haben, müssen Sie diese Richtlinie jedoch möglicherweise ändern. Diese Richtlinie beinhaltet keinen KMS-Schlüssel zum Entschlüsseln von Daten.

Ihre Amazon S3 S3-Ressourcen müssen sich in derselben Umgebung AWS-Region wie die AWS Clean Rooms Kollaboration befinden.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowECRIImageDownloadForTrainingAndInferenceJobs",
      "Effect": "Allow",
      "Action": [
        "ecr:BatchGetImage",
        "ecr:BatchCheckLayerAvailability",
        "ecr:GetDownloadUrlForLayer"
      ],
      "Resource": "arn:aws:ecr:us-east-1:111122223333:repository/repoName"
    }
  ]
}
```

5. Ersetzen Sie jede *placeholder* durch Ihre eigenen Informationen:
- *region* – Der Name des AWS-Region. Beispiel, **us-east-1**.
 - *accountId*— Die AWS-Konto ID, in der sich der S3-Bucket befindet.

- *repoName*— Der Name des Repositorys, das Ihre Daten enthält.
6. Wählen Sie Weiter aus.
 7. Geben Sie unter Überprüfen und erstellen einen Richtliniennamen und eine Beschreibung ein, und überprüfen Sie die Zusammenfassung.
 8. Wählen Sie Richtlinie erstellen aus.

Sie haben eine Richtlinie für erstellt AWS Clean Rooms.

9. Wählen Sie unter Access management (Zugriffsverwaltung) Roles (Rollen) aus.

Mit Rollen können Sie kurzfristige Anmeldeinformationen erstellen, was aus Sicherheitsgründen empfohlen wird. Sie können auch Benutzer auswählen, um langfristige Anmeldeinformationen zu erstellen.

10. Wählen Sie Rolle erstellen aus.
11. Wählen Sie im Assistenten zum Erstellen von Rollen unter Vertrauenswürdiger Entitätstyp die Option Benutzerdefinierte Vertrauensrichtlinie aus.
12. Kopieren Sie die folgende benutzerdefinierte Vertrauensrichtlinie und fügen Sie sie in den JSON-Editor ein.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "cleanrooms-ml.amazonaws.com"
      },
      "Action": "sts:AssumeRole"
    }
  ]
}
```

Das SourceAccount ist immer dein AWS-Konto. Das SourceArn kann auf einen bestimmten Trainingsdatensatz beschränkt werden, aber erst, nachdem dieser Datensatz erstellt wurde. Da Sie den ARN des Trainingsdatensatzes noch nicht kennen, wird hier der Platzhalter angegeben.

13. Wählen Sie Weiter aus.
14. Aktivieren Sie das Kontrollkästchen neben dem Namen der Richtlinie, die Sie erstellt haben, und wählen Sie dann Weiter aus.
15. Geben Sie unter Name, review and create den Rollennamen und die Beschreibung ein.

 Note

Der Rollenname muss dem Muster in den `passRole` Berechtigungen entsprechen, die dem Mitglied erteilt wurden, das Ergebnisse abfragen und empfangen kann, und den Mitgliedsrollen.

- a. Überprüfen Sie die Option Vertrauenswürdige Entitäten auswählen und bearbeiten Sie sie gegebenenfalls.
- b. Überprüfen Sie die Berechtigungen unter Berechtigungen hinzufügen und bearbeiten Sie sie gegebenenfalls.
- c. Überprüfen Sie die Tags und fügen Sie bei Bedarf Stichwörter hinzu.
- d. Wählen Sie Rolle erstellen aus.

Sie haben die Servicerolle für erstellt AWS Clean Rooms.

Erstellen Sie eine Servicerolle, um einen Datensatz abzufragen

AWS Clean Rooms verwendet eine Servicerolle, um zu steuern, wer einen Datensatz abfragen kann, der für die benutzerdefinierte ML-Modellierung verwendet wird. Sie können diese Rolle mithilfe der Konsole erstellen, wenn Sie über die erforderlichen IAM-Berechtigungen verfügen. Wenn Sie keine `CreateRole` Berechtigungen haben, bitten Sie Ihren Administrator, die Servicerolle zu erstellen.

Mit dieser Rolle können Sie die Aktion „[MLInputKanal erstellen](#)“ verwenden.

Um eine Servicerolle zu erstellen, die es einem Mitglied ermöglicht, einen Datensatz abzufragen

1. Melden Sie sich mit Ihrem Administratorkonto bei der IAM-Konsole (<https://console.aws.amazon.com/iam/>) an.
2. Wählen Sie unter Access management (Zugriffsverwaltung) Policies (Richtlinien) aus.
3. Wählen Sie Richtlinie erstellen aus.

4. Wählen Sie im Richtlinien-Editor die Registerkarte JSON aus und kopieren Sie dann die folgende Richtlinie und fügen Sie sie ein.

 Note

Die folgende Beispielrichtlinie unterstützt die Berechtigungen, die für die Abfrage eines Datensatzes erforderlich sind, der für die benutzerdefinierte ML-Modellierung verwendet wird. Je nachdem, wie Sie Ihre Amazon S3 S3-Daten eingerichtet haben, müssen Sie diese Richtlinie jedoch möglicherweise ändern. Diese Richtlinie beinhaltet keinen KMS-Schlüssel zum Entschlüsseln von Daten.

Ihre Amazon S3 S3-Ressourcen müssen sich in derselben Umgebung AWS-Region wie die AWS Clean Rooms Kollaboration befinden.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowCleanRoomsStartQueryForMLInputChannel",
      "Effect": "Allow",
      "Action": "cleanrooms:StartProtectedQuery",
      "Resource": "*"
    },
    {
      "Sid": "AllowCleanroomsGetSchemaAndGetAnalysisTemplateForMLInputChannel",
      "Effect": "Allow",
      "Action": [
        "cleanrooms:GetSchema",
        "cleanrooms:GetCollaborationAnalysisTemplate"
      ],
      "Resource": "*"
    },
    {
      "Sid": "AllowCleanRoomsGetAndUpdateQueryForMLInputChannel",
      "Effect": "Allow",
      "Action": [
        "cleanrooms:GetProtectedQuery",
```

```

        "cleanrooms:UpdateProtectedQuery"
      ],
      "Resource": [
        "arn:aws:cleanrooms:us-east-1:111122223333:membership/queryRunnerMembershipId"
      ]
    }
  ]
}

```

5. Ersetzen Sie jede *placeholder* durch Ihre eigenen Informationen:
 - *region* – Der Name des AWS-Region. Beispiel, **us-east-1**.
 - *queryRunnerAccountId*— Die AWS-Konto ID des Kontos, das die Abfragen ausführt.
 - *queryRunnerMembershipId*— Die Mitglieds-ID des Mitglieds, das Abfragen durchführen kann. Die Mitglieds-ID finden Sie auf der Registerkarte Details der Kollaboration. Dadurch AWS Clean Rooms wird sichergestellt, dass die Rolle nur übernommen wird, wenn dieses Mitglied die Analyse in dieser Kollaboration ausführt.
6. Wählen Sie Weiter aus.
7. Geben Sie unter Überprüfen und erstellen einen Richtliniennamen und eine Beschreibung ein, und überprüfen Sie die Zusammenfassung.
8. Wählen Sie Richtlinie erstellen aus.

Sie haben eine Richtlinie für erstellt AWS Clean Rooms.
9. Wählen Sie unter Access management (Zugriffsverwaltung) Roles (Rollen) aus.

Mit Rollen können Sie kurzfristige Anmeldeinformationen erstellen, was aus Sicherheitsgründen empfohlen wird. Sie können auch Benutzer auswählen, um langfristige Anmeldeinformationen zu erstellen.
10. Wählen Sie Rolle erstellen aus.
11. Wählen Sie im Assistenten zum Erstellen von Rollen unter Vertrauenswürdiger Entitätstyp die Option Benutzerdefinierte Vertrauensrichtlinie aus.
12. Kopieren Sie die folgende benutzerdefinierte Vertrauensrichtlinie und fügen Sie sie in den JSON-Editor ein.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
        "Service": "cleanrooms-ml.amazonaws.com"
      },
      "Action": "sts:AssumeRole"
    }
  ]
}
```

Das SourceAccount ist immer dein AWS-Konto. Das SourceArn kann auf einen bestimmten Trainingsdatensatz beschränkt werden, aber erst, nachdem dieser Datensatz erstellt wurde. Da Sie den ARN des Trainingsdatensatzes noch nicht kennen, wird hier der Platzhalter angegeben.

13. Wählen Sie Weiter aus.
14. Aktivieren Sie das Kontrollkästchen neben dem Namen der Richtlinie, die Sie erstellt haben, und wählen Sie dann Weiter aus.
15. Geben Sie unter Name, review and create den Rollennamen und die Beschreibung ein.

 Note

Der Rollename muss dem Muster in den passRole Berechtigungen entsprechen, die dem Mitglied erteilt wurden, das Ergebnisse abfragen und empfangen kann, und den Mitgliedsrollen.

- a. Überprüfen Sie die Option Vertrauenswürdige Entitäten auswählen und bearbeiten Sie sie gegebenenfalls.
- b. Überprüfen Sie die Berechtigungen unter Berechtigungen hinzufügen und bearbeiten Sie sie gegebenenfalls.
- c. Überprüfen Sie die Tags und fügen Sie bei Bedarf Stichwörter hinzu.

- d. Wählen Sie Rolle erstellen aus.

Sie haben die Servicerolle für erstellt AWS Clean Rooms.

Erstellen Sie eine Servicerolle, um eine konfigurierte Tabellenzuordnung zu erstellen

AWS Clean Rooms verwendet eine Servicerolle, um zu steuern, wer eine konfigurierte Tabellenzuordnung erstellen kann. Sie können diese Rolle mithilfe der Konsole erstellen, wenn Sie über die erforderlichen IAM-Berechtigungen verfügen. Wenn Sie keine `CreateRole` Berechtigungen haben, bitten Sie Ihren Administrator, die Servicerolle zu erstellen.

Mit dieser Rolle können Sie die `CreateConfiguredTableAssociation` Aktion verwenden.

Um eine Servicerolle zu erstellen, um die Erstellung einer konfigurierten Tabellenzuordnung zu ermöglichen

1. Melden Sie sich mit Ihrem Administratorkonto bei der IAM-Konsole (<https://console.aws.amazon.com/iam/>) an.
2. Wählen Sie unter Access management (Zugriffsverwaltung) Policies (Richtlinien) aus.
3. Wählen Sie Richtlinie erstellen aus.
4. Wählen Sie im Richtlinien-Editor die Registerkarte JSON aus und kopieren Sie dann die folgende Richtlinie und fügen Sie sie ein.

Note

Die folgende Beispielrichtlinie unterstützt die Erstellung einer konfigurierten Tabellenzuordnung. Je nachdem, wie Sie Ihre Amazon S3 S3-Daten eingerichtet haben, müssen Sie diese Richtlinie jedoch möglicherweise ändern. Diese Richtlinie beinhaltet keinen KMS-Schlüssel zum Entschlüsseln von Daten. Ihre Amazon S3 S3-Ressourcen müssen sich in derselben Umgebung AWS-Region wie die AWS Clean Rooms Kollaboration befinden.

JSON

```
{  
  "Version": "2012-10-17",
```

```

"Statement": [
  {
    "Action": [
      "kms:Decrypt",
      "kms:DescribeKey"
    ],
    "Resource": "arn:aws:kms:us-east-1:111122223333:key/KMS-key-ID",
    "Effect": "Allow"
  },
  {
    "Action": [
      "s3:ListBucket",
      "s3:GetBucketLocation"
    ],
    "Resource": "arn:aws:s3::bucket-name",
    "Effect": "Allow"
  },
  {
    "Action": "s3:GetObject",
    "Resource": "arn:aws:s3::bucket-name/*",
    "Effect": "Allow"
  },
  {
    "Action": [
      "glue:GetDatabase",
      "glue:GetDatabases",
      "glue:GetTable",
      "glue:GetTables",
      "glue:GetPartitions",
      "glue:GetPartition",
      "glue:BatchGetPartition"
    ],
    "Resource": [
      "arn:aws:glue:us-east-1:111122223333:catalog",
      "arn:aws:glue:us-east-1:111122223333:database/Glue database name",
      "arn:aws:glue:us-east-1:111122223333:table/Glue database name/Glue table name"
    ],
    "Effect": "Allow"
  },
  {
    "Action": [
      "glue:GetSchema",

```

```

        "glue:GetSchemaVersion"
    ],
    "Resource": "*",
    "Effect": "Allow"
}
]
}

```

Platzhalter-Ressource ersetzen ARNs

Wenn Sie diese Richtlinie verwenden, müssen Sie die Platzhalter-Ressourcen-IDs durch die tatsächlichen ARNs Ressourcen ersetzen:

- AWS KMS Schlüsselressource: *KMS-key-ID* Ersetzen Sie es durch die tatsächliche AWS KMS Schlüssel-ID, die Ihre Amazon S3 S3-Daten verschlüsselt. Der Schlüssel muss sich in demselben Konto (111122223333) befinden, dem die Katalogressourcen gehören. AWS Glue
- Amazon S3 S3-Bucket-Ressourcen: *bucket-name* Ersetzen Sie es durch den tatsächlichen Namen des Amazon S3 S3-Buckets, der Ihre AWS Glue Tabellendaten enthält. Beachten Sie, dass der Amazon S3 S3-Bucket ARNs kein Konto enthält, IDs da Bucket-Namen weltweit einzigartig sind.
- AWS Glue Ressourcen: Ersetzen Sie die folgenden Platzhalter durch Ihre tatsächlichen Ressourcennamen:
 - *Glue database name*- Der Name Ihrer Datenbank AWS Glue
 - *Glue table name*- Der Name Ihrer AWS Glue Tabelle

Alle AWS Glue Ressourcen (Katalog, Datenbank und Tabelle) müssen sich in derselben AWS-Konto (111122223333) befinden, um konsistente Zugriffsberechtigungen zu gewährleisten. Bei diesem Konto sollte es sich um dasselbe Konto handeln, dem der AWS KMS Schlüssel gehört, der für die Datenverschlüsselung verwendet wird, wodurch eine einheitliche Sicherheitsgrenze für Ihre Datenressourcen geschaffen wird. AWS Clean Rooms

5. Ersetzen Sie jedes *placeholder* durch Ihre eigenen Informationen:

- *KMS key used to encrypt the Amazon S3 data*— Der KMS-Schlüssel, der zur Verschlüsselung der Amazon S3 S3-Daten verwendet wurde. Um die Daten zu entschlüsseln, müssen Sie denselben KMS-Schlüssel angeben, der zum Verschlüsseln der Daten verwendet wurde.
- *Amazon S3 bucket of AWS Glue table*— Der Name des Amazon S3 S3-Buckets, der die AWS Glue Tabelle enthält, die Ihre Daten enthält.
- *region* – Der Name des AWS-Region. Beispiel, **us-east-1**.
- *accountId*— Die AWS-Konto ID des Kontos, dem die Daten gehören.
- *AWS Glue database name*— Der Name der AWS Glue Datenbank, die Ihre Daten enthält.
- *AWS Glue table name*— Der Name der AWS Glue Tabelle, die Ihre Daten enthält.

6. Wählen Sie Weiter aus.
7. Geben Sie unter Überprüfen und erstellen einen Richtliniennamen und eine Beschreibung ein, und überprüfen Sie die Zusammenfassung.
8. Wählen Sie Richtlinie erstellen aus.

Sie haben eine Richtlinie für erstellt AWS Clean Rooms.

9. Wählen Sie unter Access management (Zugriffsverwaltung) Roles (Rollen) aus.

Mit Rollen können Sie kurzfristige Anmeldeinformationen erstellen, was aus Sicherheitsgründen empfohlen wird. Sie können auch Benutzer auswählen, um langfristige Anmeldeinformationen zu erstellen.

10. Wählen Sie Rolle erstellen aus.
11. Wählen Sie im Assistenten zum Erstellen von Rollen unter Vertrauenswürdiger Entitätstyp die Option Benutzerdefinierte Vertrauensrichtlinie aus.
12. Kopieren Sie die folgende benutzerdefinierte Vertrauensrichtlinie und fügen Sie sie in den JSON-Editor ein.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Principal": {
```

```
        "Service": "cleanrooms-m1.amazonaws.com"
      },
      "Action": "sts:AssumeRole"
    }
  ]
}
```

Das SourceAccount ist immer dein AWS-Konto. Das SourceArn kann auf einen bestimmten Trainingsdatensatz beschränkt werden, aber erst, nachdem dieser Datensatz erstellt wurde. Da Sie den ARN des Trainingsdatensatzes noch nicht kennen, wird hier der Platzhalter angegeben.

13. Wählen Sie Weiter aus.
14. Aktivieren Sie das Kontrollkästchen neben dem Namen der Richtlinie, die Sie erstellt haben, und wählen Sie dann Weiter aus.
15. Geben Sie unter Name, review and create den Rollennamen und die Beschreibung ein.

 Note

Der Rollename muss dem Muster in den passRole Berechtigungen entsprechen, die dem Mitglied erteilt wurden, das Ergebnisse abfragen und empfangen kann, und den Mitgliedsrollen.

- a. Überprüfen Sie die Option Vertrauenswürdige Entitäten auswählen und bearbeiten Sie sie gegebenenfalls.
- b. Überprüfen Sie die Berechtigungen unter Berechtigungen hinzufügen und bearbeiten Sie sie gegebenenfalls.
- c. Überprüfen Sie die Tags und fügen Sie bei Bedarf Stichwörter hinzu.
- d. Wählen Sie Rolle erstellen aus.

Sie haben die Servicerolle für erstellt AWS Clean Rooms.

Kooperationen und Mitgliedschaften in AWS Clean Rooms

Eine Kollaboration ist eine sichere logische Grenze, AWS Clean Rooms innerhalb derer Mitglieder Analysen an konfigurierten Tabellen durchführen können.

Jedes Mitglied AWS Clean Rooms kann eine Kollaboration erstellen.

Der Ersteller der Kollaboration kann ein einzelnes Mitglied bestimmen, das konfigurierte Tabellen analysiert und Ergebnisse erhält. Der Ersteller der Kollaboration möchte jedoch möglicherweise verhindern, dass das Mitglied, das die Analyse durchführen kann, Zugriff auf die Abfrageergebnisse hat. In diesem Fall kann der Ersteller der Kollaboration ein [Mitglied bestimmen, das Abfragen durchführen kann](#), oder [ein Mitglied, das Abfragen und Jobs ausführen kann, und ein anderes Mitglied, das Ergebnisse empfangen kann](#).

In den meisten Fällen [zahlt das Mitglied, das Abfragen durchführen kann, oder das Mitglied, das Jobs abfragen und ausführen kann, auch die Rechenkosten](#). Der Ersteller der Kollaboration kann jedoch ein anderes Mitglied so konfigurieren, dass es für die Bezahlung der Rechenkosten für Abfragen verantwortlich ist.

Informationen zum Erstellen einer Kollaboration mithilfe von finden Sie in der [AWS Clean Rooms API-Referenz](#). AWS SDKs

Topics

- [Eine Zusammenarbeit erstellen](#)
- [Eine Mitgliedschaft erstellen und einer Kollaboration beitreten](#)
- [Kollaborationen bearbeiten](#)
- [Anfragen ändern in AWS Clean Rooms](#)
- [Kollaborationen löschen](#)
- [Kollaborationen anzeigen](#)
- [Mitglieder zu einer Kollaboration einladen](#)
- [Mitglieder überwachen](#)
- [Mitglieder zu einer Kollaboration hinzufügen](#)
- [Mitglieder aus einer Kollaboration entfernen](#)
- [Verlassen einer Zusammenarbeit](#)

Eine Zusammenarbeit erstellen

Es gibt drei Möglichkeiten, eine Zusammenarbeit in zu erstellen AWS Clean Rooms.

Die einfachste Form ist die [Zusammenarbeit bei Abfragen](#). Diese Zusammenarbeit konzentriert sich auf die Analyse von SQL-Abfragen und verfolgt eine einfache Struktur mit zwei Hauptrollen: ein Mitglied, das Abfragen ausführen kann, und ein anderes, das Ergebnisse erhalten kann. Diese grundlegende Einrichtung der Zusammenarbeit eignet sich gut für einfache Datenanalyseaufgaben.

Die zweite Form, die [Zusammenarbeit für Abfragen und Jobs](#), erweitert die Funktionalität, indem sie sowohl SQL-Abfragen als auch PySpark Jobs einbezieht. Bei dieser Einrichtung der Zusammenarbeit wird dieselbe grundlegende Rollenstruktur beibehalten, die Berechtigungen werden jedoch um die Ausführung von Aufgaben erweitert. Eine wichtige Anforderung besteht darin, dass das Mitglied, das PySpark Analysevorlagen erstellt, auch die Ergebnisse erhält, sodass eine klare Rechenschaftspflicht im Analyseprozess gewährleistet ist.

Die dritte Form, die [Zusammenarbeit bei der ML-Modellierung, ist für](#) Arbeitsabläufe im Bereich maschinelles Lernen konzipiert. Diese Einrichtung für die Zusammenarbeit fügt zwei weitere Rollen hinzu: eine für Benutzer, die die Ergebnisse trainierter Modelle benötigen, und eine weitere für Benutzer, die die Ergebnisse der Verwendung dieser Modelle benötigen, um Vorhersagen zu treffen. Diese Einrichtung für die Zusammenarbeit hilft den Mitgliedern der Zusammenarbeit, gemeinsam an komplexen Datenprojekten zu arbeiten und gleichzeitig die Rollen und Berechtigungen aller Beteiligten klar zu definieren.

In den folgenden Themen wird erklärt, wie Sie Kollaborationen für Abfragen, Jobs und ML-Modellierung erstellen.

Themen

- [Eine Zusammenarbeit für Abfragen erstellen](#)
- [Eine Kollaboration für Anfragen und Jobs erstellen](#)
- [Eine Zusammenarbeit für ML-Modellierung erstellen](#)

Eine Zusammenarbeit für Abfragen erstellen

In diesem Verfahren führen Sie als [Ersteller der Kollaboration](#) die folgenden Aufgaben aus:

- [Erstellen Sie eine Kollaboration.](#)

- Laden Sie ein oder mehrere [Mitglieder](#) zur [Kollaboration](#) ein.
- Weisen Sie Mitgliedern Fähigkeiten zu, z. B. dem [Mitglied, das Abfragen durchführen kann](#), und dem [Mitglied, das Ergebnisse erhalten kann](#).

Wenn der Ersteller der Kollaboration auch das Mitglied ist, das Ergebnisse empfangen kann, gibt er das Ziel und das Format der Ergebnisse an. Sie bieten auch eine Servicerolle Amazon Resource Name (ARN), um die Ergebnisse in das Ergebnisziel zu schreiben.

- Konfigurieren Sie, [welches Mitglied für die Bezahlung der Rechenkosten im Rahmen der Zusammenarbeit verantwortlich ist](#).

Prüfen Sie bevor Sie beginnen ob die folgenden Voraussetzungen erfüllt sind:

- Sie haben den Namen und die AWS-Konto ID für jedes Mitglied, das Sie zur Kollaboration einladen möchten.
- Sie sind berechtigt, den Namen und die AWS-Konto ID jedes Mitglieds mit allen Mitgliedern der Kollaboration zu teilen.

 Note

Sie können keine weiteren Mitglieder hinzufügen, nachdem Sie die Kollaboration erstellt haben.

Informationen zum Erstellen einer Kollaboration mithilfe von finden Sie in der [AWS Clean Rooms API-Referenz](#). AWS SDKs

Um eine Kollaboration für Abfragen zu erstellen

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#) mit der AWS-Konto , die als Ersteller der Kollaboration fungiert.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie in der oberen rechten Ecke die Option Kollaboration erstellen aus.
4. Gehen Sie für Schritt 1: Zusammenarbeit definieren wie folgt vor:
 - a. Geben Sie für Details den Namen und die Beschreibung der Zusammenarbeit ein.

Diese Informationen sind für Mitglieder der Kollaboration sichtbar, die zur Teilnahme an der Kollaboration eingeladen wurden. Der Name und die Beschreibung helfen ihnen zu verstehen, worauf sich die Zusammenarbeit bezieht.

b. Für Mitglieder:

- i. Für Mitglied 1: Sie geben den Anzeigenamen Ihres Mitglieds so ein, wie er für die Kollaboration angezeigt werden soll.

 Note

Ihre AWS-Konto ID ist automatisch in der AWS-Konto Mitglieds-ID enthalten.

- ii. Geben Sie für Mitglied 2 den Anzeigenamen und die AWS-Konto Mitglieds-ID des Mitglieds ein, das Sie zur Kollaboration einladen möchten.

Der Anzeigename und die AWS-Konto Mitglieds-ID des Mitglieds sind für alle zu der Kollaboration eingeladenen Personen sichtbar. Nachdem Sie die Werte für diese Felder eingegeben und gespeichert haben, können Sie sie nicht mehr bearbeiten.

 Note

Sie müssen das Mitglied der Kollaboration darüber informieren, dass seine AWS-Konto Mitglieds-ID und sein Anzeigename für alle eingeladenen und aktiven Mitarbeiter in der Kollaboration sichtbar sind.

- iii. Wenn Sie ein weiteres Mitglied hinzufügen möchten, wählen Sie Weiteres Mitglied hinzufügen. Geben Sie dann den Anzeigenamen und die AWS-Konto Mitglieds-ID für jedes Mitglied ein, das Daten beitragen kann, die Sie zur Kollaboration einladen möchten.

- c. Wenn Sie die Analyseprotokollierung aktivieren möchten, aktivieren Sie das Kontrollkästchen Analyseprotokollierung aktivieren.

- Wählen Sie unter Unterstützte Protokolltypen das Kontrollkästchen Protokolle aus Abfragen aus.

Sie erhalten Protokolle, die aus SQL-Abfragen generiert wurden, in Ihrem Amazon CloudWatch Logs-Konto.

- d. Wenn Sie die detaillierte Überwachung aktivieren möchten, aktivieren Sie das Kontrollkästchen Detaillierte Überwachung aktivieren.

Der Analysis Runner und der konfigurierte Zahler können bei der Erstellung ihrer Mitgliedschaft festlegen, ob detaillierte Metriken aktiviert werden sollen. Wenn diese Option aktiviert ist, werden detaillierte Monitoring-Metriken veröffentlicht, CloudWatch um die operative Überwachung der Zusammenarbeit zu ermöglichen, einschließlich der Abfrageleistung und der Ressourcennutzung. Diese Metriken stehen dem Analysis Runner und dem konfigurierten Payor in ihren jeweiligen AWS-Konto s zur Verfügung.

Weitere Informationen zur Preisgestaltung finden Sie unter CloudWatch [CloudWatch Preisgestaltung](#).

- e. Wählen Sie unter Zulässige Bereiche für Abfrageergebnisse einen oder mehrere Bereiche aus, AWS-Regionen an die Sie Abfrageergebnisse senden möchten.

Standardmäßig ist nur die aktuelle Region (z. B. Nord-Virginia us-east-1) ausgewählt.

 Important

Wenn Sie die regionsübergreifende Bereitstellung von Abfrageergebnissen aktivieren, können Ihre Ergebnisse außerhalb der Quellregion verarbeitet und gespeichert werden.

Weitere Informationen zu Regionen finden Sie unter [Regionen und Endpunkte](#) in der. Allgemeine AWS-Referenz

- f. (Optional) Verwalten Sie den Zugriff auf Ihre Daten mit der Funktion Automatische Genehmigung von Änderungsanträgen, indem Sie konfigurieren, welche Einstellungen automatisch geändert werden können, ohne dass eine Änderungsanforderung manuell genehmigt werden muss. Standardmäßig können einige Einstellungen nur geändert werden, indem eine Änderungsanforderung eingereicht wird, die von allen Mitgliedern genehmigt werden muss, bevor sie wirksam wird.
 - Mitgliedern Fähigkeiten gewähren — Wählen Sie die Fähigkeiten aus, die Mitgliedern der Kollaboration ohne manuelle Genehmigung gewährt werden können. Mitglieder können jederzeit Daten beitragen.
 - Wähle Fähigkeiten:

- Daten beitragen (immer aktiviert)
- Ergebnisse erhalten
- Neue Mitglieder mit diesen Fähigkeiten automatisch genehmigen — Falls zulässig, werden alle Mitglieder, die mit den oben ausgewählten Fähigkeiten hinzugefügt wurden, sofort der Kollaboration beitreten. Mitglieder, die mit anderen Fähigkeiten hinzugefügt wurden, benötigen weiterhin eine manuelle Genehmigung, um beitreten zu können.
- Fähigkeiten, die automatisch entzogen werden können — Wähle die Fähigkeiten aus, die ohne manuelle Genehmigung entzogen werden können. Mitglieder können jederzeit Daten beisteuern.
- Wähle Fähigkeiten:
 - Daten beitragen (immer aktiviert)
 - Ergebnisse erhalten

Wenn Sie diese Option wählen, können Sie alle Änderungen an der Konfiguration der Zusammenarbeit über den Verlauf der Änderungsanträge verfolgen, der sich auf der Registerkarte Details der Kollaborationsdetails befindet.

g. (Optional) Wenn Sie die kryptografische Rechenfunktion aktivieren möchten, aktivieren Sie das Kontrollkästchen Kryptografisches Rechnen aktivieren.

i. Wählen Sie die folgenden Parameter für die kryptografische Abdeckung aus:

- Spalten zulassen plaintext

Wählen Sie Nein, wenn Sie vollständig verschlüsselte Tabellen benötigen.

Wählen Sie Ja, wenn cleartext Spalten in der verschlüsselten Tabelle zulässig sein sollen.

Damit SUM oder AVG für bestimmte Spalten ausgeführt werden kann, müssen sich die Spalten in befindencleartext.

- NULLWerte beibehalten

Wählen Sie Nein, wenn Sie NULL Werte nicht beibehalten möchten. NULLWerte werden nicht wie NULL in einer verschlüsselten Tabelle angezeigt.

Wählen Sie Ja, wenn Sie NULL Werte beibehalten möchten. NULLWerte werden wie NULL in einer verschlüsselten Tabelle angezeigt.

ii. Wählen Sie die folgenden Fingerprinting-Parameter:

- Duplikate zulassen

Wählen Sie Nein, wenn Sie nicht möchten, dass doppelte Einträge in einer fingerprint Spalte zulässig sind.

Wählen Sie Ja, wenn doppelte Einträge in einer fingerprint Spalte zulässig sein sollen.

- Zulassen JOIN von Spalten mit unterschiedlichen Namen

Wählen Sie Nein, wenn Sie fingerprint Spalten mit unterschiedlichen Namen nicht verbinden möchten.

Wählen Sie Ja, wenn Sie fingerprint Spalten mit unterschiedlichen Namen verbinden möchten.

Weitere Hinweise zu kryptografischen Berechnungsparametern finden Sie unter [Kryptografische Rechenparameter](#).

Weitere Hinweise zur Verschlüsselung Ihrer Daten für die Verwendung in finden Sie AWS Clean Rooms unter. [Vorbereiten verschlüsselter Datentabellen mit Cryptographic Computing für Clean Rooms](#)

 Note

Überprüfen Sie diese Konfigurationen sorgfältig, bevor Sie den nächsten Schritt ausführen. Nachdem Sie die Kollaboration erstellt haben, können Sie nur den Namen und die Beschreibung der Kollaboration bearbeiten und angeben, ob die Protokolle in Amazon CloudWatch Logs gespeichert sind.

- h. Wenn Sie Tags für die Kollaborationsressource aktivieren möchten, wählen Sie Neues Tag hinzufügen und geben Sie dann das Schlüssel - und Wertepaar ein.
- i. Wählen Sie Weiter aus.
5. Lassen Sie für Schritt 2: Spezifizieren Sie die Fähigkeiten der Mitglieder für Analysen mithilfe von Abfragen und Jobs unter Unterstützte Analysetypen das Kontrollkästchen Abfragen aktiviert und ergreifen Sie je nach Ziel die empfohlene Maßnahme.

Ihr Ziel	Empfohlene Aktion
Fragen Sie die Daten in der Zusammenarbeit ab und erhalten Sie die Ergebnisse	1. Wählen Sie sich selbst als das Mitglied aus, das Abfragen ausführen kann. 2. Wählen Sie sich selbst als Mitglied aus, das Analyseergebnisse aus der Drop-down-Liste erhalten kann.
Fragen Sie die Daten in der Kollaboration ab und weisen Sie einem anderen Mitglied die Ergebnisse zu	1. Wählen Sie sich selbst als das Mitglied aus, das Abfragen ausführen kann. 2. Wählen Sie aus der Drop-down-Liste das Mitglied aus, das Ergebnisse aus Analysen erhalten kann.
Empfangen Sie die Ergebnisse der Abfrage in der Kollaboration und weisen Sie ein anderes Mitglied mit der Abfrage der Daten zu	1. Wählen Sie aus der Dropdownliste das Mitglied aus, das Abfragen ausführen kann. 2. Wählen Sie sich selbst als Mitglied aus, das Ergebnisse aus Analysen aus der Drop-down-Liste erhalten kann.
Erstellen und verwalten Sie die Kollaboration, weisen Sie einem anderen Mitglied die Abfrage der Daten zu und weisen Sie ein anderes Mitglied dem Empfang der Ergebnisse zu	1. Wählen Sie aus der Dropdownliste das Mitglied aus, das Abfragen ausführen kann. 2. Wählen Sie aus der Drop-down-Liste das Mitglied aus, das Ergebnisse aus Analysen erhalten kann.

- a. Wenn Sie Clean Rooms ML für die ML-Modellierung mit speziell entwickelten Workflows verwenden,
 - i. (Optional) Wählen Sie aus der Dropdownliste das Mitglied aus, das Ergebnisse von trainierten Modellen erhalten kann.
 - ii. (Optional) Wählen Sie aus der Dropdownliste das Mitglied aus, das Ergebnisse von Modellinferenzen empfangen kann.

- b. Sehen Sie sich die Fähigkeiten der Mitglieder unter ID-Auflösung mit an. AWS Entity Resolution
 - c. Wählen Sie Weiter aus.
6. Führen Sie für Schritt 3: Zahlung konfigurieren für Analysen mithilfe von Abfragen je nach Ziel eine der folgenden Aktionen aus.

Ihr Ziel	Empfohlene Aktion
Weisen Sie dem Mitglied, das Abfragen ausführen kann, das Mitglied zu, das die Rechenkosten für Abfragen bezahlt	1. Wählen Sie für Analysen mithilfe von Abfragen das Mitglied, das für Abfragen bezahlt, genauso aus wie das Mitglied, das Abfragen ausführen kann. 2. Wählen Sie Weiter aus.
Weisen Sie ein anderes Mitglied zu, die Kosten für die Berechnung der Abfrage zu tragen	1. Für Analysen mithilfe von Abfragen wählen Sie sich selbst als das Mitglied, das für Abfragen bezahlt. 2. Wählen Sie Weiter aus.

Bei ML-Modellierung mit speziell entwickelten Workflows ist der Ersteller des konfigurierten Lookalike-Modells das Mitglied, das für die Lookalike-Modellierung bezahlt.

Bei der ID-Auflösung mit AWS Entity Resolution ist der Ersteller der ID-Zuordnungstabelle das Mitglied, das für die ID-Zuordnungstabelle bezahlt.

7. Wählen Sie für Schritt 4: Mitgliedschaft konfigurieren eine der folgenden Optionen:

Yes, join by creating membership now

1. Für Standardeinstellungen für Ergebniseinstellungen und für Einstellungen für Abfrageergebnisse, wenn Sie das Mitglied sind, das Ergebnisse empfangen kann,
 - a. Geben Sie für das Ergebnisziel in Amazon S3 das Amazon S3 S3-Ziel ein oder wählen Sie Browse S3, um einen S3-Bucket auszuwählen.
 - b. Wählen Sie für das Abfrageergebnisformat entweder CSV oder PARQUET.
 - c. (Nur Spark) Wählen Sie für die Ergebnisdateien entweder „Mehrfach“ oder „Einfach“.

- d. (Optional) Wenn Sie Anfragen, die bis zu 24 Stunden dauern, an Ihr S3-Ziel weiterleiten möchten, aktivieren Sie für den Zugriff auf Dienste das Kontrollkästchen Servicerolle hinzufügen, um Anfragen zu unterstützen, deren Bearbeitung bis zu 24 Stunden in Anspruch nimmt.

Umfangreiche Anfragen, deren Bearbeitung bis zu 24 Stunden in Anspruch nimmt, werden an Ihr S3-Ziel zugestellt.

Wenn Sie das Kontrollkästchen nicht aktivieren, werden nur Anfragen, die innerhalb von 12 Stunden abgeschlossen wurden, an Ihren S3-Standort zugestellt.

- e. Geben Sie die Dienstzugriffsberechtigungen an, indem Sie entweder Neue Servicerolle erstellen und verwenden oder Bestehende Servicerolle verwenden auswählen.

Wenn Sie sich dafür entscheiden...	Dann...
Erstellen und verwenden Sie eine neue Servicerolle	• AWS Clean Rooms erstellt eine Servicerolle mit der erforderlichen Richtlinie für diese Tabelle. • Der Standardname der Servicerolle lautet <code>cleanrooms-result-receiver-<timestamp></code> • Sie benötigen die erforderlichen Berechtigungen, um Rollen zu erstellen und Richtlinien anzuhängen.

Wenn Sie sich dafür entscheiden...	Dann...
Verwenden Sie eine vorhandene Servicerolle	i. Wählen Sie einen vorhandenen Servicerollennamen aus der Dropdownliste aus. Die Liste der Rollen wird angezeigt , wenn Sie berechtigt sind, Rollen aufzulisten. Wenn Sie nicht berechtigt sind, Rollen aufzulisten, können Sie den Amazon-Ressourcennamen (ARN) der Rolle eingeben, die Sie verwenden möchten. ii. Rufen Sie die Servicerolle auf, indem Sie auf den externen Link In IAM anzeigen klicken. Wenn keine vorhandenen Servicerollen vorhanden sind, ist die Option „Eine bestehende Servicerolle verwenden“ nicht verfügbar. Versucht standardmäßig AWS Clean Rooms nicht, die bestehende Rollenrichtlinie zu aktualisieren, um die erforderlichen Berechtigungen hinzuzufügen.

 Note

- AWS Clean Rooms erfordert Berechtigungen für Abfragen gemäß den Analyseregeln. Weitere Informationen zu Berechtigungen für AWS Clean Rooms finden Sie unter [AWS verwaltete Richtlinien für AWS Clean Rooms](#).
- Wenn die Rolle nicht über ausreichende Berechtigungen für verfügt AWS Clean Rooms, erhalten Sie eine Fehlermeldung, dass die Rolle nicht

über ausreichende Berechtigungen für verfügt AWS Clean Rooms. Die Rollenrichtlinie muss hinzugefügt werden, bevor Sie fortfahren können.

- Wenn Sie die Rollenrichtlinie nicht ändern können, erhalten Sie eine Fehlermeldung, dass die Richtlinie für die Servicerolle nicht gefunden werden AWS Clean Rooms konnte.

2. Wählen Sie für die Logs-Einstellungen eine der folgenden Optionen für die Protokollspeicherung in Amazon CloudWatch Logs:

 Note

Der Abschnitt Log-Einstellungen wird angezeigt, wenn Sie die Abfrageprotokollierung aktiviert haben.

- a. Wählen Sie Einschalten und die für Sie relevanten Abfrageprotokolle werden in Ihrem Amazon CloudWatch Logs-Konto gespeichert.

Jedes Mitglied kann nur Protokolle für Anfragen erhalten, die es initiiert hat oder die seine Daten enthalten.

Das Mitglied, das Ergebnisse erhalten kann, erhält auch Protokolle für alle Abfragen, die in einer Kollaboration ausgeführt werden, auch wenn in einer Abfrage nicht auf seine Daten zugegriffen wird.

Unter Unterstützte Protokolltypen ist das Kontrollkästchen Protokolle abfragen standardmäßig aktiviert.

 Note

Nachdem Sie die Abfrageprotokollierung aktiviert haben, kann es einige Minuten dauern, bis der Protokollspeicher eingerichtet ist und Protokolle in Amazon CloudWatch Logs empfangen werden. Während dieser kurzen Zeit führt das Mitglied, das Abfragen durchführen kann, möglicherweise Abfragen durch, ohne dass tatsächlich Protokolle gesendet werden.

- b. Wählen Sie Ausschalten und die für Sie relevanten Abfrageprotokolle werden nicht in Ihrem Amazon CloudWatch Logs-Konto gespeichert.

3. Wenn Sie Tags für die Mitgliedschaftsressource aktivieren möchten, wählen Sie Neues Tag hinzufügen und geben Sie dann das Schlüssel - und Wertepaar ein.
4. Wenn Sie das Mitglied sind, das für Query Compute bezahlt, geben Sie Ihre Zustimmung an, indem Sie das Kontrollkästchen Ich stimme zu, für die Rechenkosten in dieser Zusammenarbeit zu zahlen, aktivieren.

 Note

Sie müssen dieses Kontrollkästchen aktivieren, um fortzufahren.

Weitere Informationen zur Preisberechnung finden Sie unter [Preisgestaltung für AWS Clean Rooms](#).

Wenn Sie das [Mitglied sind, das die Kosten für die Berechnung von Abfragen bezahlt](#), aber nicht das [Mitglied, das Abfragen durchführen kann](#), wird empfohlen, ein Budget für die Konfiguration eines Budgets AWS Budgets zu verwenden AWS Clean Rooms und Benachrichtigungen zu erhalten, sobald das maximale Budget erreicht ist. Weitere Informationen zur Einrichtung eines Budgets finden Sie AWS Budgets im AWS Cost Management Benutzerhandbuch unter [Ihre Kosten verwalten mit](#). Weitere Informationen zum Einrichten von Benachrichtigungen finden Sie unter [Erstellen eines Amazon SNS SNS-Themas für Budgetbenachrichtigungen](#) im AWS Cost Management Benutzerhandbuch. Wenn das maximale Budget erreicht wurde, können Sie sich an das Mitglied wenden, das Anfragen stellen oder [die Kollaboration verlassen](#) kann. Wenn Sie die Kollaboration verlassen, dürfen keine Abfragen mehr ausgeführt werden, sodass Ihnen auch keine Kosten für die Berechnung von Abfragen mehr in Rechnung gestellt werden.

5. Wählen Sie Weiter aus.

Sowohl die Kollaboration als auch Ihre Mitgliedschaft werden erstellt.

Ihr Status in der Kollaboration ist aktiv.

No, I will create a membership later

1. Wählen Sie Weiter aus.

Nur die Kollaboration wird erstellt.

Ihr Status in der Kollaboration ist inaktiv.

8. Gehen Sie für Schritt 5: Überprüfen und erstellen wie folgt vor:

- a. Überprüfen Sie die Auswahlen, die Sie für die vorherigen Schritte getroffen haben, und bearbeiten Sie sie gegebenenfalls.
- b. Wählen Sie eine der Optionen aus.

Wenn Sie sich dafür entschieden haben...	Dann wähle...
Erstellen Sie eine Mitgliedschaft mit der Kollaboration (Ja, treten Sie bei, indem Sie jetzt eine Mitgliedschaft erstellen)	Erstellen Sie eine Zusammenarbeit und Mitgliedschaft
Erstellen Sie die Kollaboration und legen Sie zu diesem Zeitpunkt noch keine Mitgliedschaft an (Nein, ich werde später eine Mitgliedschaft erstellen)	Zusammenarbeit erstellen

Nachdem Ihre Kollaboration erfolgreich erstellt wurde, können Sie die Seite mit den Kollaborationsdetails unter Kollaborationen einsehen.

Sie sind jetzt bereit für:

- [Bereiten Sie Ihre Datentabelle für die Analyse vor AWS Clean Rooms](#). (Optional, wenn Sie Ihre eigenen Ereignisdaten analysieren oder Identitätsdaten abfragen möchten.)
- [Ordnen Sie die konfigurierte Tabelle Ihrer Kollaboration zu](#). (Optional, wenn Sie Ihre eigenen Ereignisdaten analysieren möchten.)
- [Fügen Sie eine Analyseregeln für die konfigurierte Tabelle](#) hinzu. (Optional, wenn Sie Ihre eigenen Ereignisdaten analysieren möchten.)
- [Erstellen Sie eine Mitgliedschaft und treten Sie einer Kollaboration](#) bei. (Optional, wenn Sie bereits eine Mitgliedschaft erstellt haben.)
- [Laden Sie Mitglieder ein, der Kollaboration beizutreten](#).

Eine Kollaboration für Anfragen und Jobs erstellen

In diesem Verfahren führen Sie als [Ersteller der Kollaboration](#) die folgenden Aufgaben aus:

- [Erstellen Sie eine Kollaboration](#).
- Laden Sie ein oder mehrere [Mitglieder](#) zur [Kollaboration](#) ein.
- Weisen Sie Mitgliedern Fähigkeiten zu, z. B. dem [Mitglied, das Abfragen und Jobs ausführen kann](#), und dem [Mitglied, das Ergebnisse erhalten kann](#).

Wenn der Ersteller der Kollaboration auch das Mitglied ist, das Ergebnisse empfangen kann, gibt er das Ziel und das Format der Ergebnisse an. Sie bieten auch eine Servicrolle Amazon Resource Name (ARN), um die Ergebnisse in das Ergebnisziel zu schreiben.

- Legen Sie fest, [welches Mitglied für die Bezahlung der Kosten für Abfragen und Datenverarbeitung im Rahmen der Zusammenarbeit verantwortlich ist](#).

Prüfen Sie bevor Sie beginnen ob die folgenden Voraussetzungen erfüllt sind:

- Sie haben den Namen und die AWS-Konto ID für jedes Mitglied, das Sie zur Kollaboration einladen möchten.
- Sie sind berechtigt, den Namen und die AWS-Konto ID jedes Mitglieds mit allen Mitgliedern der Kollaboration zu teilen.

 Note

Sie können keine weiteren Mitglieder hinzufügen, nachdem Sie die Kollaboration erstellt haben.

Informationen zum Erstellen einer Kollaboration mithilfe von finden Sie in der [AWS Clean Rooms API-Referenz](#). AWS SDKs

Um eine Kollaboration für Abfragen und Jobs zu erstellen

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#) mit der AWS-Konto , die als Ersteller der Kollaboration fungiert.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie in der oberen rechten Ecke die Option Kollaboration erstellen aus.
4. Gehen Sie für Schritt 1: Zusammenarbeit definieren wie folgt vor:
 - a. Geben Sie für Details den Namen und die Beschreibung der Zusammenarbeit ein.

Diese Informationen sind für Mitglieder der Kollaboration sichtbar, die zur Teilnahme an der Kollaboration eingeladen wurden. Der Name und die Beschreibung helfen ihnen zu verstehen, worauf sich die Zusammenarbeit bezieht.

b. Für Mitglieder:

- i. Für Mitglied 1: Sie geben den Anzeigenamen Ihres Mitglieds so ein, wie er für die Kollaboration angezeigt werden soll.

 Note

Ihre AWS-Konto ID ist automatisch in der AWS-Konto Mitglieds-ID enthalten.

- ii. Geben Sie für Mitglied 2 den Anzeigenamen und die AWS-Konto Mitglieds-ID des Mitglieds ein, das Sie zur Kollaboration einladen möchten.

Der Anzeigename und die AWS-Konto Mitglieds-ID des Mitglieds sind für alle zu der Kollaboration eingeladenen Personen sichtbar. Nachdem Sie die Werte für diese Felder eingegeben und gespeichert haben, können Sie sie nicht mehr bearbeiten.

 Note

Sie müssen das Mitglied der Kollaboration darüber informieren, dass seine AWS-Konto Mitglieds-ID und sein Anzeigename für alle eingeladenen und aktiven Mitarbeiter in der Kollaboration sichtbar sind.

- iii. Wenn Sie ein weiteres Mitglied hinzufügen möchten, wählen Sie Weiteres Mitglied hinzufügen. Geben Sie dann den Anzeigenamen und die AWS-Konto Mitglieds-ID für jedes Mitglied ein, das Daten beitragen kann, die Sie zur Kollaboration einladen möchten.

- c. Wenn Sie die Analyseprotokollierung aktivieren möchten, aktivieren Sie das Kontrollkästchen Analyseprotokollierung aktivieren und wählen Sie dann Unterstützte Protokolltypen aus.

- Wenn Sie aus SQL-Abfragen generierte Protokolle empfangen möchten, aktivieren Sie das Kontrollkästchen Protokolle aus Abfragen.
- Wenn Sie Protokolle empfangen möchten, die aus Aufträgen mit generiert wurden PySpark, aktivieren Sie das Kontrollkästchen Protokolle von Aufträgen.

- d. Wenn Sie die detaillierte Überwachung aktivieren möchten, aktivieren Sie das Kontrollkästchen **Detaillierte Überwachung aktivieren**.

Der Analysis Runner und der konfigurierte Zahler können bei der Erstellung ihrer Mitgliedschaft festlegen, ob detaillierte Metriken aktiviert werden sollen. Wenn diese Option aktiviert ist, werden detaillierte Monitoring-Metriken veröffentlicht, CloudWatch um die operative Überwachung der Zusammenarbeit zu ermöglichen, einschließlich der Abfrageleistung und der Ressourcennutzung. Diese Metriken stehen dem Analysis Runner und dem konfigurierten Payor in ihren jeweiligen AWS-Konten zur Verfügung.

Weitere Informationen zur Preisgestaltung finden Sie unter CloudWatch [CloudWatch Preisgestaltung](#).

- e. Wählen Sie unter **Zulässige Bereiche für Abfrageergebnisse** einen oder mehrere Bereiche aus, AWS-Regionen an die Sie Abfrageergebnisse senden möchten.

Standardmäßig ist nur die aktuelle Region (z. B. Nord-Virginia us-east-1) ausgewählt.

 Important

Wenn Sie die regionsübergreifende Bereitstellung von Abfrageergebnissen aktivieren, können Ihre Ergebnisse außerhalb der Quellregion verarbeitet und gespeichert werden.

Weitere Informationen zu Regionen finden Sie unter [Regionen und Endpunkte](#) in der Allgemeinen AWS-Referenz.

- f. (Optional) Verwalten Sie den Zugriff auf Ihre Daten mit der Funktion **Automatische Genehmigung von Änderungsanträgen**, indem Sie konfigurieren, welche Einstellungen automatisch geändert werden können, ohne dass eine Änderungsanforderung manuell genehmigt werden muss. Standardmäßig können einige Einstellungen nur geändert werden, indem eine Änderungsanforderung eingereicht wird, die von allen Mitgliedern genehmigt werden muss, bevor sie wirksam wird.
- **Mitgliedern Fähigkeiten gewähren** — Wählen Sie die Fähigkeiten aus, die Mitgliedern der Kollaboration ohne manuelle Genehmigung gewährt werden können. Mitglieder können jederzeit Daten beitragen.
 - **Wähle Fähigkeiten:**

- Daten beitragen (immer aktiviert)
- Ergebnisse erhalten
- Neue Mitglieder mit diesen Fähigkeiten automatisch genehmigen — Falls zulässig, werden alle Mitglieder, die mit den oben ausgewählten Fähigkeiten hinzugefügt wurden, sofort der Kollaboration beitreten. Mitglieder, die mit anderen Fähigkeiten hinzugefügt wurden, benötigen weiterhin eine manuelle Genehmigung, um beitreten zu können.
- Fähigkeiten, die automatisch entzogen werden können — Wähle die Fähigkeiten aus, die ohne manuelle Genehmigung entzogen werden können. Mitglieder können jederzeit Daten beisteuern.
- Wähle Fähigkeiten:
 - Daten beitragen (immer aktiviert)
 - Ergebnisse erhalten

Wenn Sie diese Option wählen, können Sie alle Änderungen an der Konfiguration der Zusammenarbeit über den Verlauf der Änderungsanträge verfolgen, der sich auf der Registerkarte Details der Kollaborationsdetails befindet.

- g. (Optional) Wenn Sie die kryptografische Rechenfunktion aktivieren möchten, aktivieren Sie das Kontrollkästchen Kryptografisches Rechnen aktivieren.

- i. Wählen Sie die folgenden Parameter für die kryptografische Abdeckung aus:

- Spalten zulassen plaintext

Wählen Sie Nein, wenn Sie vollständig verschlüsselte Tabellen benötigen.

Wählen Sie Ja, wenn cleartext Spalten in der verschlüsselten Tabelle zulässig sein sollen.

Damit SUM oder AVG für bestimmte Spalten ausgeführt werden kann, müssen sich die Spalten in befindencleartext.

- NULLWerte beibehalten

Wählen Sie Nein, wenn Sie NULL Werte nicht beibehalten möchten. NULLWerte werden nicht wie NULL in einer verschlüsselten Tabelle angezeigt.

Wählen Sie Ja, wenn Sie NULL Werte beibehalten möchten. NULLWerte werden wie NULL in einer verschlüsselten Tabelle angezeigt.

ii. Wählen Sie die folgenden Fingerprinting-Parameter:

- Duplikate zulassen

Wählen Sie Nein, wenn Sie nicht möchten, dass doppelte Einträge in einer fingerprint Spalte zulässig sind.

Wählen Sie Ja, wenn doppelte Einträge in einer fingerprint Spalte zulässig sein sollen.

- Zulassen JOIN von Spalten mit unterschiedlichen Namen

Wählen Sie Nein, wenn Sie fingerprint Spalten mit unterschiedlichen Namen nicht verbinden möchten.

Wählen Sie Ja, wenn Sie fingerprint Spalten mit unterschiedlichen Namen verbinden möchten.

Weitere Hinweise zu kryptografischen Berechnungsparametern finden Sie unter [Kryptografische Rechenparameter](#).

Weitere Hinweise zur Verschlüsselung Ihrer Daten für die Verwendung in finden Sie AWS Clean Rooms unter. [Vorbereiten verschlüsselter Datentabellen mit Cryptographic Computing für Clean Rooms](#)

 Note

Überprüfen Sie diese Konfigurationen sorgfältig, bevor Sie den nächsten Schritt ausführen. Nachdem Sie die Kollaboration erstellt haben, können Sie nur den Namen und die Beschreibung der Kollaboration bearbeiten und angeben, ob die Protokolle in Amazon CloudWatch Logs gespeichert sind.

h. Wenn Sie Tags für die Kollaborationsressource aktivieren möchten, wählen Sie Neues Tag hinzufügen und geben Sie dann das Schlüssel - und Wertepaar ein.

i. Wählen Sie Weiter aus.

5. Gehen Sie für Schritt 2: Fähigkeiten der Mitglieder angeben wie folgt vor:

a. Aktivieren Sie für Analysen mithilfe von Abfragen und Jobs unter Unterstützte Analysetypen das Kontrollkästchen Jobs.

Das Kontrollkästchen Abfragen ist standardmäßig aktiviert.

- i. Wählen Sie aus der Dropdownliste das Mitglied aus, das Abfragen und Jobs ausführen kann.
- ii. Wählen Sie aus der Drop-down-Liste das Mitglied aus, das Ergebnisse aus Analysen erhalten kann.

 Note

Das Mitglied, das die PySpark Analysevorlage erstellt, muss auch das Mitglied sein, das Ergebnisse erhält.

- b. Wenn Sie Clean Rooms ML für die ML-Modellierung mit speziell entwickelten Workflows verwenden,
 - i. (Optional) Wählen Sie aus der Dropdownliste das Mitglied aus, das Ergebnisse von trainierten Modellen erhalten kann.
 - ii. (Optional) Wählen Sie aus der Dropdownliste das Mitglied aus, das Ergebnisse von Modellinferenzen empfangen kann.
 - c. Sehen Sie sich die Fähigkeiten der Mitglieder unter ID-Auflösung mit an. AWS Entity Resolution
 - d. Wählen Sie Weiter aus.
6. Für Schritt 3: Zahlung konfigurieren,
- a. Wählen Sie für Analysen mithilfe von Abfragen und Aufträgen das Mitglied aus, das für Abfragen und Jobs bezahlt.
- Sie können das Mitglied, das Abfragen und Jobs ausführen kann, das Mitglied sein, das die Berechnungskosten für die Abfragen und Jobs bezahlt.
- Sie können ein anderes Mitglied mit der Bezahlung der Berechnungskosten für Abfragen und Jobs beauftragen.
- b. Bei ML-Modellierung mit speziell entwickelten Workflows ist der Ersteller des konfigurierten Lookalike-Modells das Mitglied, das für die Lookalike-Modellierung bezahlt.
 - c. Bei der ID-Auflösung mit AWS Entity Resolution ist der Ersteller der ID-Zuordnungstabelle das Mitglied, das für die ID-Zuordnungstabelle bezahlt.

d. Wählen Sie Weiter aus.

7. Wählen Sie für Schritt 4: Mitgliedschaft konfigurieren eine der folgenden Optionen:

Yes, join by creating membership now

1. Für Standardeinstellungen für Ergebniseinstellungen und für Einstellungen für Abfrageergebnisse, wenn Sie das Mitglied sind, das Ergebnisse empfangen kann,
 - a. Aktivieren Sie das Kontrollkästchen Standardeinstellungen für Abfragen festlegen. Geben Sie für das Ergebnisziel in Amazon S3 das Amazon S3 S3-Ziel ein oder wählen Sie Browse S3, um einen S3-Bucket auszuwählen.
 - b. Wählen Sie für das Abfrageergebnisformat entweder CSV oder PARQUET.
 - c. (Nur Spark) Wählen Sie für die Ergebnisdateien entweder „Mehrfach“ oder „Einfach“.
 - d. (Optional) Wenn Sie Anfragen, die bis zu 24 Stunden dauern, an Ihr S3-Ziel weiterleiten möchten, aktivieren Sie für den Zugriff auf Dienste das Kontrollkästchen Servicerolle hinzufügen, um Anfragen zu unterstützen, deren Bearbeitung bis zu 24 Stunden in Anspruch nimmt.

Umfangreiche Anfragen, deren Bearbeitung bis zu 24 Stunden in Anspruch nimmt, werden an Ihr S3-Ziel zugestellt.

Wenn Sie das Kontrollkästchen nicht aktivieren, werden nur Anfragen, die innerhalb von 12 Stunden abgeschlossen wurden, an Ihren S3-Standort zugestellt.

- e. Geben Sie die Dienstzugriffsberechtigungen an, indem Sie entweder Neue Servicerolle erstellen und verwenden oder Bestehende Servicerolle verwenden auswählen.

Wenn Sie sich dafür entscheiden...	Dann...
Erstellen und verwenden Sie eine neue Servicerolle	• AWS Clean Rooms erstellt eine Servicerolle mit der erforderlichen Richtlinie für diese Tabelle. • Der Standardname der Servicerolle lautet <code>cleanrooms-result-receiver-<timestamp></code> • Sie benötigen die erforderlichen Berechtigungen, um Rollen zu erstellen und Richtlinien anzuhängen.

Wenn Sie sich dafür entscheiden...	Dann...
Verwenden Sie eine vorhandene Servicerolle	i. Wählen Sie einen vorhandenen Servicerollennamen aus der Dropdownliste aus. Die Liste der Rollen wird angezeigt , wenn Sie berechtigt sind, Rollen aufzulisten. Wenn Sie nicht berechtigt sind, Rollen aufzulisten, können Sie den Amazon-Ressourcennamen (ARN) der Rolle eingeben, die Sie verwenden möchten. ii. Rufen Sie die Servicerolle auf, indem Sie auf den externen Link In IAM anzeigen klicken. Wenn keine vorhandenen Servicerollen vorhanden sind, ist die Option „Eine bestehende Servicerolle verwenden“ nicht verfügbar. Versucht standardmäßig AWS Clean Rooms nicht, die bestehende Rollenrichtlinie zu aktualisieren, um die erforderlichen Berechtigungen hinzuzufügen.

 Note

- AWS Clean Rooms erfordert Berechtigungen für Abfragen gemäß den Analyseregeln. Weitere Informationen zu Berechtigungen für AWS Clean Rooms finden Sie unter [AWS verwaltete Richtlinien für AWS Clean Rooms](#).
- Wenn die Rolle nicht über ausreichende Berechtigungen für verfügt AWS Clean Rooms, erhalten Sie eine Fehlermeldung, dass die Rolle nicht

über ausreichende Berechtigungen für verfügt AWS Clean Rooms. Die Rollenrichtlinie muss hinzugefügt werden, bevor Sie fortfahren können.

- Wenn Sie die Rollenrichtlinie nicht ändern können, erhalten Sie eine Fehlermeldung, dass die Richtlinie für die Servicerolle nicht gefunden werden AWS Clean Rooms konnte.

2. Für Job-Ergebnisse

Example

Beispiel: **s3://bucket/prefix**

- a. Aktivieren Sie das Kontrollkästchen Standardeinstellungen für Jobs festlegen und geben Sie dann das Ergebnisziel in Amazon S3 an, indem Sie das S3-Ziel eingeben, oder wählen Sie S3 durchsuchen, um aus einer Liste verfügbarer S3-Buckets auszuwählen.
 - b. Geben Sie die Dienstzugriffsberechtigungen an, indem Sie einen vorhandenen Servicerollennamen aus der Drop-down-Liste auswählen.
3. Wählen Sie für die Logs-Einstellungen eine der folgenden Optionen für die Protokollspeicherung in Amazon CloudWatch Logs:

Note

Der Abschnitt Log-Einstellungen wird angezeigt, wenn Sie die Abfrageprotokollierung aktiviert haben.

- a. Wählen Sie Einschalten und die für Sie relevanten Abfrageprotokolle werden in Ihrem Amazon CloudWatch Logs-Konto gespeichert.

Jedes Mitglied kann nur Protokolle für Anfragen erhalten, die es initiiert hat oder die seine Daten enthalten.

Das Mitglied, das Ergebnisse erhalten kann, erhält auch Protokolle für alle Anfragen, die in einer Kollaboration ausgeführt werden, auch wenn in einer Abfrage nicht auf seine Daten zugegriffen wird.

Wählen Sie unter Unterstützte Protokolltypen einen der Protokolltypen aus, die der Kollaborationsersteller zur Unterstützung ausgewählt hat:

Unter Unterstützte Protokolltypen sind die Kontrollkästchen Abfrageprotokolle und Jobprotokolle standardmäßig aktiviert.

 Note

Nachdem Sie die Analysis-Protokollierung aktiviert haben, kann es einige Minuten dauern, bis der Protokollspeicher eingerichtet ist und Protokolle in Amazon CloudWatch Logs empfangen werden. Während dieser kurzen Zeit führt das Mitglied, das Abfragen durchführen kann, möglicherweise Abfragen durch, bei denen keine Protokolle gesendet werden.

- b. Wählen Sie Ausschalten und die für Sie relevanten Abfrageprotokolle werden nicht in Ihrem Amazon CloudWatch Logs-Konto gespeichert.
4. Wenn Sie Mitgliedschafts-Tags für die Mitgliedschaftsressource aktivieren möchten, wählen Sie Neues Tag hinzufügen und geben Sie dann das Schlüssel - und Wertepaar ein.
5. Wenn Sie das Mitglied sind, das für Query Compute oder Job Compute oder beides bezahlt, geben Sie Ihre Zustimmung an, indem Sie das Kontrollkästchen Ich stimme zu, für die Rechenkosten in dieser Zusammenarbeit zu zahlen, aktivieren.

 Note

Sie müssen dieses Kontrollkästchen aktivieren, um fortzufahren.

Weitere Informationen zur Preisberechnung finden Sie unter [Preisgestaltung für AWS Clean Rooms](#).

Wenn Sie das [Mitglied sind, das die Kosten für die Berechnung von Abfragen bezahlt, aber nicht das Mitglied, das Abfragen durchführen kann](#), empfiehlt es sich, ein Budget AWS Budgets zu konfigurieren AWS Clean Rooms und Benachrichtigungen zu erhalten, sobald das maximale Budget erreicht ist. Weitere Informationen zur Einrichtung eines Budgets finden Sie AWS Budgets im AWS Cost Management Benutzerhandbuch unter [Ihre Kosten verwalten mit](#). Weitere Informationen zum Einrichten von Benachrichtigungen finden Sie unter [Erstellen eines Amazon SNS SNS-Themas für Budgetbenachrichtigungen](#) im AWS Cost Management Benutzerhandbuch. Wenn das maximale Budget erreicht wurde, können Sie sich an das Mitglied wenden, das Anfragen stellen oder [die Kollaboration verlassen](#) kann. Wenn Sie die Kollaboration verlassen, dürfen keine Abfragen mehr

ausgeführt werden, sodass Ihnen auch keine Kosten für die Berechnung von Abfragen mehr in Rechnung gestellt werden.

6. Wählen Sie Weiter aus.

Sowohl die Kollaboration als auch Ihre Mitgliedschaft werden erstellt.

Ihr Status in der Kollaboration ist aktiv.

No, I will create a membership later

1. Wählen Sie Weiter aus.

Nur die Kollaboration wird erstellt.

Ihr Status in der Kollaboration ist inaktiv.

8. Gehen Sie für Schritt 5: Überprüfen und erstellen wie folgt vor:

- a. Überprüfen Sie die Auswahlen, die Sie für die vorherigen Schritte getroffen haben, und bearbeiten Sie sie gegebenenfalls.
- b. Wählen Sie eine der Optionen aus.

Wenn Sie sich dafür entschieden haben...	Dann wähle...
Erstellen Sie eine Mitgliedschaft mit der Kollaboration (Ja, treten Sie bei, indem Sie jetzt eine Mitgliedschaft erstellen)	Erstellen Sie eine Zusammenarbeit und Mitgliedschaft
Erstellen Sie die Kollaboration und legen Sie zu diesem Zeitpunkt noch keine Mitgliedschaft an (Nein, ich werde später eine Mitgliedschaft erstellen)	Zusammenarbeit erstellen

Nachdem Ihre Kollaboration erfolgreich erstellt wurde, können Sie die Seite mit den Kollaborationsdetails unter Kollaborationen einsehen.

Sie sind jetzt bereit für:

- [Bereiten Sie Ihre Datentabelle für die Analyse vor AWS Clean Rooms](#). (Optional, wenn Sie Ihre eigenen Ereignisdaten analysieren oder Identitätsdaten abfragen möchten.)
- [Ordnen Sie die konfigurierte Tabelle Ihrer Kollaboration zu](#). (Optional, wenn Sie Ihre eigenen Ereignisdaten analysieren möchten.)
- [Fügen Sie eine Analyseregeln für die konfigurierte Tabelle hinzu](#). (Optional, wenn Sie Ihre eigenen Ereignisdaten analysieren möchten.)
- [Erstellen Sie eine Mitgliedschaft und treten Sie einer Kollaboration bei](#). (Optional, wenn Sie bereits eine Mitgliedschaft erstellt haben.)
- [Laden Sie Mitglieder ein, der Kollaboration beizutreten](#).

Eine Zusammenarbeit für ML-Modellierung erstellen

In diesem Verfahren führen Sie als [Ersteller der Kollaboration](#) die folgenden Aufgaben aus:

- [Erstellen Sie eine Kollaboration](#).
- Laden Sie ein oder mehrere [Mitglieder](#) zur [Kollaboration](#) ein.
- Weisen Sie Mitgliedern Fähigkeiten zu, wie z. B.
 - [Mitglied, das Fragen stellen kann](#)
 - [Mitglied, das Ergebnisse erhalten kann](#)
 - Mitglied, das Ergebnisse von trainierten Modellen erhalten kann
 - Mitglied, das Ergebnisse aus Modellinferenz generieren kann

Wenn der Ersteller der Kollaboration auch das Mitglied ist, das Ergebnisse empfangen kann, gibt er das Ziel und das Format der Ergebnisse an. Sie bieten auch eine Serviceroles Amazon Resource Name (ARN), um die Ergebnisse in das Ergebnisziel zu schreiben.

- Konfigurieren Sie, [welches Mitglied für die Bezahlung der Rechenkosten, der Modellschulung und der Modellinferenz im Rahmen der Zusammenarbeit verantwortlich](#) ist.

Prüfen Sie bevor Sie beginnen ob die folgenden Voraussetzungen erfüllt sind:

- Sie haben den Namen und die AWS-Konto ID für jedes Mitglied, das Sie zur Kollaboration einladen möchten.
- Sie sind berechtigt, den Namen und die AWS-Konto ID jedes Mitglieds mit allen Mitgliedern der Kollaboration zu teilen.

 Note

Sie können keine weiteren Mitglieder hinzufügen, nachdem Sie die Kollaboration erstellt haben.

Informationen zum Erstellen einer Kollaboration mithilfe von finden Sie in der [AWS Clean Rooms API-Referenz](#). AWS SDKs

Um eine Kollaboration für die ML-Modellierung zu erstellen

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#) mit der AWS-Konto , die als Ersteller der Zusammenarbeit fungiert.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie in der oberen rechten Ecke die Option Kollaboration erstellen aus.
4. Gehen Sie für Schritt 1: Zusammenarbeit definieren wie folgt vor:
 - a. Geben Sie für Details den Namen und die Beschreibung der Zusammenarbeit ein.

Diese Informationen sind für Mitglieder der Kollaboration sichtbar, die zur Teilnahme an der Kollaboration eingeladen wurden. Der Name und die Beschreibung helfen ihnen zu verstehen, worauf sich die Zusammenarbeit bezieht.

- b. Für Mitglieder:
 - i. Für Mitglied 1: Sie geben den Anzeigenamen Ihres Mitglieds so ein, wie er für die Kollaboration angezeigt werden soll.

 Note

Ihre AWS-Konto ID ist automatisch in der AWS-Konto Mitglieds-ID enthalten.

- ii. Geben Sie für Mitglied 2 den Anzeigenamen und die AWS-Konto Mitglieds-ID des Mitglieds ein, das Sie zur Kollaboration einladen möchten.

Der Anzeigename und die AWS-Konto Mitglieds-ID des Mitglieds sind für alle zu der Kollaboration eingeladenen Personen sichtbar. Nachdem Sie die Werte für diese Felder eingegeben und gespeichert haben, können Sie sie nicht mehr bearbeiten.

 Note

Sie müssen das Mitglied der Kollaboration darüber informieren, dass seine AWS-Konto Mitglieds-ID und sein Anzeigename für alle eingeladenen und aktiven Mitarbeiter in der Kollaboration sichtbar sind.

- iii. Wenn Sie ein weiteres Mitglied hinzufügen möchten, wählen Sie Weiteres Mitglied hinzufügen. Geben Sie dann den Anzeigenamen und die AWS-Konto Mitglieds-ID für jedes Mitglied ein, das Daten beitragen kann, die Sie zur Kollaboration einladen möchten.
- c. Wenn Sie die Analyseprotokollierung aktivieren möchten, aktivieren Sie das Kontrollkästchen Analyseprotokollierung aktivieren und wählen Sie dann unter Unterstützte Protokolltypen die Option Protokolle aus Abfragen aus.
- d. Wenn Sie die detaillierte Überwachung aktivieren möchten, aktivieren Sie das Kontrollkästchen Detaillierte Überwachung aktivieren.

Der Analysis Runner und der konfigurierte Zahler können bei der Erstellung ihrer Mitgliedschaft festlegen, ob detaillierte Metriken aktiviert werden sollen. Wenn diese Option aktiviert ist, werden detaillierte Monitoring-Metriken veröffentlicht, CloudWatch um die operative Überwachung der Zusammenarbeit zu ermöglichen, einschließlich der Abfrageleistung und der Ressourcennutzung. Diese Metriken stehen dem Analysis Runner und dem konfigurierten Payor in ihren jeweiligen AWS-Konto s zur Verfügung.

Weitere Informationen zur Preisgestaltung finden Sie unter CloudWatch [CloudWatch Preisgestaltung](#).

- e. Wählen Sie unter Zulässige Bereiche für Abfrageergebnisse einen oder mehrere Bereiche aus, AWS-Regionen an die Sie Abfrageergebnisse senden möchten.

Standardmäßig ist nur die aktuelle Region (z. B. Nord-Virginia us-east-1) ausgewählt.

 Important

Wenn Sie die regionsübergreifende Bereitstellung von Abfrageergebnissen aktivieren, können Ihre Ergebnisse außerhalb der Quellregion verarbeitet und gespeichert werden.

Weitere Informationen zu Regionen finden Sie unter [Regionen und Endpunkte](#) in der. Allgemeine AWS-Referenz

- f. (Optional) Verwalten Sie den Zugriff auf Ihre Daten mit der Funktion Automatische Genehmigung von Änderungsanträgen, indem Sie konfigurieren, welche Einstellungen automatisch geändert werden können, ohne dass eine Änderungsanforderung manuell genehmigt werden muss. Standardmäßig können einige Einstellungen nur geändert werden, indem eine Änderungsanforderung eingereicht wird, die von allen Mitgliedern genehmigt werden muss, bevor sie wirksam wird.
- Mitgliedern Fähigkeiten gewähren — Wählen Sie die Fähigkeiten aus, die Mitgliedern der Kollaboration ohne manuelle Genehmigung gewährt werden können. Mitglieder können jederzeit Daten beitragen.
 - Wähle Fähigkeiten:
 - Daten beitragen (immer aktiviert)
 - Ergebnisse erhalten
 - Neue Mitglieder mit diesen Fähigkeiten automatisch genehmigen — Falls zulässig, werden alle Mitglieder, die mit den oben ausgewählten Fähigkeiten hinzugefügt wurden, sofort der Kollaboration beitreten. Mitglieder, die mit anderen Fähigkeiten hinzugefügt wurden, benötigen weiterhin eine manuelle Genehmigung, um beitreten zu können.
 - Fähigkeiten, die automatisch entzogen werden können — Wähle die Fähigkeiten aus, die ohne manuelle Genehmigung entzogen werden können. Mitglieder können jederzeit Daten beisteuern.
 - Wähle Fähigkeiten:
 - Daten beitragen (immer aktiviert)
 - Ergebnisse erhalten

Wenn Sie diese Option wählen, können Sie alle Änderungen an der Konfiguration der Zusammenarbeit über den Verlauf der Änderungsanträge verfolgen, der sich auf der Registerkarte Details der Kollaborationsdetails befindet.

- g. (Optional) Wenn Sie die kryptografische Rechenfunktion aktivieren möchten, aktivieren Sie das Kontrollkästchen Kryptografisches Rechnen aktivieren.
- i. Wählen Sie die folgenden Parameter für die kryptografische Abdeckung aus:

- Spalten zulassen plaintext

Wählen Sie Nein, wenn Sie vollständig verschlüsselte Tabellen benötigen.

Wählen Sie Ja, wenn cleartext Spalten in der verschlüsselten Tabelle zulässig sein sollen.

Damit SUM oder AVG für bestimmte Spalten ausgeführt werden kann, müssen sich die Spalten in befundencleartext.

- NULLWerte beibehalten

Wählen Sie Nein, wenn Sie NULL Werte nicht beibehalten möchten. NULLWerte werden nicht wie NULL in einer verschlüsselten Tabelle angezeigt.

Wählen Sie Ja, wenn Sie NULL Werte beibehalten möchten. NULLWerte werden wie NULL in einer verschlüsselten Tabelle angezeigt.

ii. Wählen Sie die folgenden Fingerprinting-Parameter:

- Duplikate zulassen

Wählen Sie Nein, wenn Sie nicht möchten, dass doppelte Einträge in einer fingerprint Spalte zulässig sind.

Wählen Sie Ja, wenn doppelte Einträge in einer fingerprint Spalte zulässig sein sollen.

- Zulassen JOIN von Spalten mit unterschiedlichen Namen

Wählen Sie Nein, wenn Sie fingerprint Spalten mit unterschiedlichen Namen nicht verbinden möchten.

Wählen Sie Ja, wenn Sie fingerprint Spalten mit unterschiedlichen Namen verbinden möchten.

Weitere Hinweise zu kryptografischen Berechnungsparametern finden Sie unter [Kryptografische Rechenparameter](#).

Weitere Hinweise zur Verschlüsselung Ihrer Daten für die Verwendung in finden Sie AWS Clean Rooms unter. [Vorbereiten verschlüsselter Datentabellen mit Cryptographic Computing für Clean Rooms](#)

 Note

Überprüfen Sie diese Konfigurationen sorgfältig, bevor Sie den nächsten Schritt ausführen. Nachdem Sie die Kollaboration erstellt haben, können Sie nur den Namen und die Beschreibung der Kollaboration bearbeiten und angeben, ob die Protokolle in Amazon CloudWatch Logs gespeichert sind.

- h. Wenn Sie Tags für die Kollaborationsressource aktivieren möchten, wählen Sie Neues Tag hinzufügen und geben Sie dann das Schlüssel - und Wertepaar ein.
 - i. Wählen Sie Weiter aus.
5. Für Schritt 2: Geben Sie die Fähigkeiten der Mitglieder an
 - a. Lassen Sie für Analysen mithilfe von Abfragen und Jobs unter Unterstützte Analysetypen das Kontrollkästchen Abfragen aktiviert.
 - b. Wählen Sie unter Abfragen ausführen das Mitglied aus, das das Modelltraining initiieren soll
 - c. Wählen Sie unter Ergebnisse aus Analysen erhalten ein oder mehrere Mitglieder aus, die die Abfrageergebnisse erhalten sollen.
 - d. Für ML-Modellierung mit speziell entwickelten Workflows
 - i. Wählen Sie unter Ausgabe aus trainierten Modellen erhalten das Mitglied aus, das die Ergebnisse des trainierten Modells, einschließlich Modellartefakten und Metriken, erhalten soll.
 - ii. Wählen Sie unter Ausgabe aus Modellinferenz empfangen das Mitglied aus, das die Ergebnisse der Modellinferenz erhalten soll.
 - e. Sehen Sie sich die Fähigkeiten der Mitglieder unter ID-Auflösung mit an. AWS Entity Resolution
6. Für Schritt 3: Zahlung konfigurieren,
 - a. Führen Sie unter Analyse mithilfe von Abfragen für Pay for Queries eine der folgenden Aktionen aus:
 - Wenn dasselbe Mitglied für Abfragen bezahlen und diese ausführen soll, wählen Sie dasselbe Mitglied aus, das Sie für Abfragen ausführen ausgewählt haben.
 - Wählen Sie Ihr Mitgliedskonto aus, damit ein anderes Mitglied die Kosten für Abfragen trägt.

- b. Für die ML-Modellierung mit speziell entwickelten Workflows
 - Wählen Sie das Mitglied aus, das für das Modelltraining bezahlt.
 - c. Wählen Sie das Mitglied aus, das für den Inferenzjob bezahlt.
 - d. Für Pay for Lookalike-Modeling sind keine Maßnahmen erforderlich. Der Ersteller des konfigurierten Lookalike-Modells ist das Mitglied, das für die Lookalike-Modellierung bezahlt.
 - e. (Optional) Wählen Sie das Mitglied aus, das für die Generierung synthetischer Daten bezahlt.
 - f. Für die ID-Auflösung mit AWS Entity Resolution ist keine Aktion erforderlich. Der Ersteller der ID-Zuordnungstabelle ist das Mitglied, das für die ID-Zuordnungstabelle bezahlt.
7. Wählen Sie Weiter aus.
8. Wählen Sie für Schritt 4: Mitgliedschaft konfigurieren unter Collaboration-Mitgliedschaft eine der folgenden Optionen aus:

Yes, join by creating membership now

1. Für Standardeinstellungen für Ergebniseinstellungen und für Einstellungen für Abfrageergebnisse, wenn Sie das Mitglied sind, das Ergebnisse empfangen kann,
 - a. Aktivieren Sie das Kontrollkästchen Standardeinstellungen für Abfragen festlegen.
 - b. Geben Sie für das Ergebnisziel in Amazon S3 das Amazon S3 S3-Ziel ein oder wählen Sie Browse S3, um einen S3-Bucket auszuwählen.
 - c. Wählen Sie für das Abfrageergebnisformat entweder CSV oder PARQUET.
 - d. (Nur Spark) Wählen Sie für die Ergebnisdateien entweder „Mehrfach“ oder „Einfach“.
 - e. (Optional) Wenn Sie Anfragen, die bis zu 24 Stunden dauern, an Ihr S3-Ziel senden möchten, aktivieren Sie das Kontrollkästchen Servicерolle hinzufügen, um Anfragen zu unterstützen, deren Bearbeitung bis zu 24 Stunden dauert.

Große Anfragen, deren Bearbeitung bis zu 24 Stunden in Anspruch nimmt, werden an Ihr S3-Ziel zugestellt.

Wenn Sie das Kontrollkästchen nicht aktivieren, werden nur Anfragen, die innerhalb von 12 Stunden abgeschlossen wurden, an Ihren S3-Standort zugestellt.

- f. Geben Sie die Dienstzugriffsberechtigungen an, indem Sie entweder Neue Servicерolle erstellen und verwenden oder Bestehende Servicерolle verwenden auswählen.

Wenn Sie sich dafür entscheiden...	Dann...
Erstellen und verwenden Sie eine neue Servicerolle	• AWS Clean Rooms erstellt eine Servicerolle mit der erforderlichen Richtlinie für diese Tabelle. • Der Standardname der Servicerolle lautet <code>cleanrooms-result-receiver-<timestamp></code> • Sie benötigen die erforderlichen Berechtigungen, um Rollen zu erstellen und Richtlinien anzuhängen.

Wenn Sie sich dafür entscheiden...	Dann...
Verwenden Sie eine vorhandene Servicerolle	i. Wählen Sie einen vorhandenen Servicerollennamen aus der Dropdownliste aus. Die Liste der Rollen wird angezeigt , wenn Sie berechtigt sind, Rollen aufzulisten. Wenn Sie nicht berechtigt sind, Rollen aufzulisten, können Sie den Amazon-Ressourcennamen (ARN) der Rolle eingeben, die Sie verwenden möchten. ii. Rufen Sie die Servicerolle auf, indem Sie auf den externen Link In IAM anzeigen klicken. Wenn keine vorhandenen Servicerollen vorhanden sind, ist die Option „Eine bestehende Servicerolle verwenden“ nicht verfügbar. Versucht standardmäßig AWS Clean Rooms nicht, die bestehende Rollenrichtlinie zu aktualisieren, um die erforderlichen Berechtigungen hinzuzufügen.

 Note

- AWS Clean Rooms erfordert Berechtigungen für Abfragen gemäß den Analyseregeln. Weitere Informationen zu Berechtigungen für AWS Clean Rooms finden Sie unter [AWS verwaltete Richtlinien für AWS Clean Rooms](#).
- Wenn die Rolle nicht über ausreichende Berechtigungen für verfügt AWS Clean Rooms, erhalten Sie eine Fehlermeldung, dass die Rolle nicht

über ausreichende Berechtigungen für verfügt AWS Clean Rooms. Die Rollenrichtlinie muss hinzugefügt werden, bevor Sie fortfahren können.

- Wenn Sie die Rollenrichtlinie nicht ändern können, erhalten Sie eine Fehlermeldung, dass die Richtlinie für die Servicerolle nicht gefunden werden AWS Clean Rooms konnte.

2. Für die ML-Konfiguration

- a. Aktivieren Sie das Kontrollkästchen ML-Konfiguration erstellen und geben Sie dann das Ziel für die Modellausgabe auf Amazon S3 an, indem Sie das S3-Ziel eingeben, oder wählen Sie Browse S3, um aus einer Liste verfügbarer S3-Buckets auszuwählen.
 - b. Geben Sie die Dienstzugriffsberechtigungen an, indem Sie entweder „Neue Servicerolle erstellen und verwenden“ oder „Vorhandene Servicerolle verwenden“ wählen.
 - c. Wenn der S3-Bucket verschlüsselt ist, aktivieren Sie das Kontrollkästchen Der Ziel-Bucket ist mit einem KMS-Schlüssel verschlüsselt und geben Sie dann den Wert ein, AWS KMS key oder wählen Sie Create an AWS KMS key, um einen neuen KMS-Schlüssel zu erstellen.
3. Wenn Sie Mitgliedschafts-Tags für die Mitgliedschaftsressource aktivieren möchten, wählen Sie Neues Tag hinzufügen und geben Sie dann das Schlüssel-Wert-Paar ein.
 4. Wenn Sie das Mitglied sind, das für Query Compute bezahlt, geben Sie Ihre Zustimmung an, indem Sie das Kontrollkästchen Ich stimme zu, für die Rechenkosten in dieser Zusammenarbeit zu zahlen, aktivieren.

Note

Sie müssen dieses Kontrollkästchen aktivieren, um fortzufahren.

Weitere Informationen zur Preisberechnung finden Sie unter [Preisgestaltung für AWS Clean Rooms](#).

Wenn Sie das [Mitglied sind, das die Kosten für die Berechnung von Abfragen bezahlt, aber nicht das Mitglied, das Abfragen durchführen kann](#), empfiehlt es sich, ein Budget AWS Budgets zu konfigurieren AWS Clean Rooms und Benachrichtigungen zu erhalten, sobald das maximale Budget erreicht ist. Weitere Informationen zur Einrichtung eines Budgets finden Sie AWS Budgets im AWS Cost Management Benutzerhandbuch unter [Ihre Kosten verwalten mit](#). Weitere Informationen zum Einrichten von Benachrichtigungen finden Sie unter [Erstellen eines Amazon SNS SNS-Themas für Budgetbenachrichtigungen](#) im

AWS Cost Management Benutzerhandbuch. Wenn das maximale Budget erreicht wurde, können Sie sich an das Mitglied wenden, das Anfragen stellen oder [die Kollaboration verlassen](#) kann. Wenn Sie die Kollaboration verlassen, dürfen keine Abfragen mehr ausgeführt werden, sodass Ihnen auch keine Kosten für die Berechnung von Abfragen mehr in Rechnung gestellt werden.

5. Wählen Sie Weiter aus.

Sowohl die Kollaboration als auch Ihre Mitgliedschaft werden erstellt.

Ihr Status in der Kollaboration ist aktiv.

No, I will create a membership later

1. Wählen Sie Weiter aus.

Nur die Kollaboration wird erstellt.

Ihr Status in der Kollaboration ist inaktiv.

9. Gehen Sie für Schritt 5: Überprüfen und erstellen wie folgt vor:

- a. Überprüfen Sie die Auswahlen, die Sie für die vorherigen Schritte getroffen haben, und bearbeiten Sie sie gegebenenfalls.
- b. Wählen Sie eine der Optionen aus.

Wenn Sie sich dafür entschieden haben...	Dann wähle...
Erstellen Sie eine Mitgliedschaft mit der Kollaboration (Ja, treten Sie bei, indem Sie jetzt eine Mitgliedschaft erstellen)	Erstellen Sie eine Zusammenarbeit und Mitgliedschaft
Erstellen Sie die Kollaboration und legen Sie zu diesem Zeitpunkt noch keine Mitgliedschaft an (Nein, ich werde später eine Mitgliedschaft erstellen)	Zusammenarbeit erstellen

Eine Mitgliedschaft erstellen und einer Kollaboration beitreten

Eine Mitgliedschaft ist eine Ressource, die erstellt wird, wenn ein Mitglied einer Kollaboration in beitrifft AWS Clean Rooms.

Sie können einer Kollaboration beitreten als

- [Mitglied, das Fragen stellen kann](#)
- [Mitglied, das Abfragen und Jobs ausführen kann](#)
- [Mitglied, das Ergebnisse einer Abfrage oder eines Jobs erhalten kann](#)
- [Mitglied, das die Kosten für die Berechnung von Abfragen bezahlt](#)
- [Mitglied zahlt für Anfragen und Jobs](#)

Alle Mitglieder können Daten beisteuern.

Informationen zum Erstellen einer Mitgliedschaft und zum Beitritt zu einer Kollaboration mithilfe von finden Sie in der [AWS Clean Rooms API-Referenz](#). AWS SDKs

Bei diesem Verfahren [tritt das eingeladene Mitglied der Kollaboration bei, indem es eine Mitgliedschaftsressource erstellt](#).

Wenn das eingeladene Mitglied das Mitglied ist, das Ergebnisse erhalten kann, gibt es das Ziel und das Format der Ergebnisse an. Sie bieten auch einen ARN für die Servicерolle, um in das Ergebnisziel zu schreiben.

Wenn es sich bei dem eingeladenen Mitglied um das Mitglied handelt, das für die Bezahlung der Rechenkosten verantwortlich ist, akzeptiert es seine Zahlungsverpflichtungen, bevor es der Kollaboration beitrifft.

Um eine Mitgliedschaft zu erstellen und einer Kollaboration beizutreten

1. Melden Sie sich bei Ihrem Mitglied an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#) AWS-Konto.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie auf der Registerkarte „Zum Beitritt verfügbar“ für Kollaborationen, die zum Beitritt verfügbar sind, den Namen der Kollaboration aus.

4. Sehen Sie sich auf der Seite mit den Kollaborationsdetails im Abschnitt Übersicht die Kollaborationsdetails an, einschließlich Ihrer Mitgliedsdetails und einer Liste der anderen Mitglieder.

Vergewissern Sie sich, dass AWS-Konto IDs für jedes Mitglied der Kollaboration diejenigen sind, mit denen Sie die Zusammenarbeit eingehen möchten.

5. Wählen Sie Mitgliedschaft erstellen aus.
6. Sehen Sie sich auf der Seite Mitgliedschaft erstellen in der Übersicht den Namen der Kollaboration, die Beschreibung der Kollaboration, die AWS-Konto ID des Erstellers der Kollaboration, Ihre Mitgliedsdetails und die AWS-Konto ID des Mitglieds an, das für Anfragen bezahlt.
7. Wenn der Ersteller der Kollaboration sich dafür entschieden hat, die Analysis-Protokollierung zu aktivieren, wählen Sie eine der folgenden Optionen für die Protokollspeicherung in Amazon CloudWatch Logs:

Wenn Sie folgendes auswählen ...	Dann...
Einschalten	Die für Sie relevanten Protokolle werden in Amazon CloudWatch Logs gespeichert. Jedes Mitglied kann nur Protokolle für Anfragen erhalten, die es selbst initiiert hat oder die seine Daten enthalten. Das Mitglied, das Ergebnisse erhalten kann, erhält auch Protokolle für alle Analysen, die in einer Kollaboration ausgeführt werden, auch wenn bei einer Analyse nicht auf seine Daten zugegriffen wird. Wählen Sie unter Unterstützte Protokolltypen einen der Protokolltypen aus, die der Kollaborationsersteller für die Unterstützung ausgewählt hat: 1. Wenn Sie anhand von SQL-Abfragen generierte Protokolle erhalten möchten,

Wenn Sie folgendes auswählen ...	Dann...
	aktivieren Sie das Kontrollkästchen Protokolle aus Abfragen. 2. Wenn Sie Protokolle empfangen möchten, die aus Aufträgen mit generiert wurden PySpark, aktivieren Sie das Kontrollkästchen Protokolle von Aufträgen.
Schalten Sie aus	Die für Sie relevanten Abfrageprotokolle werden nicht in Ihrem Amazon CloudWatch Logs-Konto gespeichert.

 Note

Nachdem Sie die Analysis-Protokollierung aktiviert haben, kann es einige Minuten dauern, bis der Protokollspeicher eingerichtet ist und Protokolle in Amazon CloudWatch Logs empfangen werden. Während dieser kurzen Zeit führt das Mitglied, das Abfragen durchführen kann, möglicherweise Abfragen durch, bei denen keine Protokolle gesendet werden.

8. Wenn der Ersteller der Kollaboration die detaillierte Überwachung für diese Kollaboration aktiviert hat, wählen Sie aus, ob Sie detaillierte Messdaten zur Beobachtbarkeit in Ihrem CloudWatch Konto erhalten möchten.

Für eine detaillierte Überwachung:

Detaillierte Überwachungsoptionen

Option	Description
Einschalten	AWS Clean Rooms veröffentlicht detaillierte Monitoring-Metriken CloudWatch für diese Zusammenarbeit in Ihrem Konto. Sie können diese Messwerte für die Betriebsüberwachung verwenden, einschließlich der Abfrageleistung und der Ressourcennutzung. Es CloudWatch fallen zusätzliche Gebühren an. Weitere Informationen finden Sie unter CloudWatch – Preise.

Option	Description
Ausschalten	Es werden keine detaillierten Messwerte in Ihr CloudWatch Konto exportiert. Andere Mitglieder können weiterhin detaillierte Monitoring-Metriken in ihren eigenen Konten einsehen, wenn sie diese Option aktiviert haben.

 Note

Detaillierte Monitoring-Metriken sind nur für Mitglieder verfügbar, die Abfragen ausführen können (Analysis Runner), und für Mitglieder, die als Zahler für die Zusammenarbeit konfiguriert sind.

9. Wenn zu Ihren Mitgliederfähigkeiten die Option Ergebnisse erhalten gehört, sind die Standardeinstellungen für Ergebnisse wie folgt voreingestellt:
 - a. Aktivieren Sie für Query-Ergebnisse das Kontrollkästchen Standardeinstellungen für Abfragen festlegen und geben Sie dann das Ergebnisziel in Amazon S3 an, indem Sie das S3-Ziel eingeben, oder wählen Sie Browse S3, um aus einer Liste verfügbarer S3-Buckets auszuwählen.

Example

Beispiel: **s3://bucket/prefix**

- i. Wählen Sie für das Ergebnisformat entweder CSV oder PARQUET.
- ii. (Nur Spark) Wählen Sie für die Ergebnisdateien entweder Mehrfach oder Einfach aus.
- iii. (Optional) Wenn Sie Anfragen, die bis zu 24 Stunden dauern, an Ihr S3-Ziel weiterleiten möchten, aktivieren Sie für den Servicezugriff das Kontrollkästchen Servicerolle hinzufügen, um Anfragen zu unterstützen, deren Bearbeitung bis zu 24 Stunden in Anspruch nimmt.

Umfangreiche Anfragen, deren Bearbeitung bis zu 24 Stunden in Anspruch nimmt, werden an Ihr S3-Ziel zugestellt.

Wenn Sie das Kontrollkästchen nicht aktivieren, werden nur Anfragen, die innerhalb von 12 Stunden abgeschlossen wurden, an Ihren S3-Standort zugestellt.

 Note

Sie müssen entweder eine bestehende Servicerolle auswählen oder über die erforderlichen Berechtigungen verfügen, um eine neue zu erstellen. Weitere Informationen finden Sie unter [Erstellen Sie eine Servicerolle, um Ergebnisse zu erhalten](#).

- iv. Geben Sie die Dienstzugriffsberechtigungen an, indem Sie entweder Neue Servicerolle erstellen und verwenden oder Bestehende Servicerolle verwenden auswählen.

Create and use a new service role

- AWS Clean Rooms erstellt eine Servicerolle mit der erforderlichen Richtlinie für diese Tabelle.
- Der Standardname der Servicerolle lautet `cleanrooms-result-receiver-
<timestamp>`
- Sie müssen über die erforderlichen Berechtigungen verfügen, um Rollen zu erstellen und Richtlinien anzuhängen.

Use an existing service role

1. Wählen Sie einen vorhandenen Servicerollennamen aus der Dropdownliste aus.

Die Liste der Rollen wird angezeigt, wenn Sie berechtigt sind, Rollen aufzulisten.

Wenn Sie nicht berechtigt sind, Rollen aufzulisten, können Sie den Amazon-Ressourcennamen (ARN) der Rolle eingeben, die Sie verwenden möchten.

2. Rufen Sie die Servicerolle auf, indem Sie auf den externen Link In IAM anzeigen klicken.

Wenn keine vorhandenen Servicerollen vorhanden sind, ist die Option „Eine bestehende Servicerolle verwenden“ nicht verfügbar.

Versucht standardmäßig AWS Clean Rooms nicht, die bestehende Rollenrichtlinie zu aktualisieren, um die erforderlichen Berechtigungen hinzuzufügen.

 Note

- AWS Clean Rooms erfordert Berechtigungen für Abfragen gemäß den Analyseregeln. Weitere Informationen zu Berechtigungen für AWS Clean Rooms finden Sie unter [AWS verwaltete Richtlinien für AWS Clean Rooms](#).
- Wenn die Rolle nicht über ausreichende Berechtigungen für verfügt AWS Clean Rooms, erhalten Sie eine Fehlermeldung, dass die Rolle nicht über ausreichende Berechtigungen für verfügt AWS Clean Rooms. Die Rollenrichtlinie muss hinzugefügt werden, bevor Sie fortfahren können.
- Wenn Sie die Rollenrichtlinie nicht ändern können, erhalten Sie eine Fehlermeldung, dass die Richtlinie für die Servicerolle nicht gefunden werden AWS Clean Rooms konnte.

- b. Wählen Sie für Job-Ergebnisse das Kontrollkästchen Standardeinstellungen für Jobs festlegen und geben Sie dann das Ergebnisziel in Amazon S3 an, indem Sie das S3-Ziel eingeben, oder wählen Sie Browse S3, um aus einer Liste verfügbarer S3-Buckets auszuwählen.

Example

Beispiel: **s3://bucket/prefix**

- Geben Sie die Dienstzugriffsberechtigungen an, indem Sie einen vorhandenen Servicerollennamen aus der Dropdownliste auswählen.
10. Wenn Sie Tags für die Mitgliedschaftsressource aktivieren möchten, wählen Sie Neues Tag hinzufügen und geben Sie dann das Schlüssel - und Wertepaar ein.
11. Wenn der Ersteller der Kollaboration Sie als Mitglied bestimmt hat, das für Anfragen bezahlt oder für Anfragen und Jobs bezahlt, geben Sie Ihre Zustimmung an, indem Sie das Kontrollkästchen Ich stimme zu, für die Rechenkosten in dieser Zusammenarbeit zu zahlen, aktivieren.

 Note

Sie müssen dieses Kontrollkästchen aktivieren, um fortzufahren.

Weitere Informationen zur Preisberechnung finden Sie unter [Preisgestaltung für AWS Clean Rooms](#).

Wenn Sie das [Mitglied sind, das die Kosten für die Query-Compute bezahlt, oder das Mitglied, das für Abfragen und die Job-Computing-Kosten bezahlt, aber nicht das Mitglied sind, das Abfragen durchführen kann](#), wird empfohlen, ein Budget AWS Budgets zu konfigurieren AWS Clean Rooms und Benachrichtigungen zu erhalten, sobald das maximale Budget erreicht ist. Weitere Informationen zur Einrichtung eines Budgets finden Sie unter [Verwaltung Ihrer Kosten mit AWS Budgets](#) im AWS Cost Management Benutzerhandbuch. Weitere Informationen zum Einrichten von Benachrichtigungen finden Sie unter [Erstellen eines Amazon SNS SNS-Themas für Budgetbenachrichtigungen](#) im AWS Cost Management Benutzerhandbuch. Wenn das maximale Budget erreicht ist, können Sie sich an das Mitglied wenden, das Anfragen und Jobs ausführen oder [die Kollaboration verlassen](#) kann. Wenn Sie die Kollaboration verlassen, dürfen keine Abfragen mehr ausgeführt werden, sodass Ihnen auch keine Kosten für die Berechnung von Abfragen mehr in Rechnung gestellt werden.

12. Wenn Sie sicher sind, dass Sie eine Mitgliedschaft erstellen und der Kollaboration beitreten möchten, wählen Sie Mitgliedschaft erstellen aus.

Sie haben Lesezugriff auf die Metadaten der Kollaboration. Dazu gehören Informationen wie der Anzeigenname und die Beschreibung der Kollaboration sowie alle Namen und Namen AWS-Konto IDs anderer Mitglieder.

Sie sind jetzt bereit für:

- [Bereiten Sie Ihre Datentabelle für die Abfrage vor. AWS Clean Rooms](#) (Optional, wenn Sie Ihre eigenen Ereignisdaten oder Identitätsdaten abfragen möchten.)
- [Ordnen Sie die konfigurierte Tabelle Ihrer Kollaboration](#) zu, wenn Sie Ereignisdaten abfragen möchten.
- [Fügen Sie eine Analyseregeln für die konfigurierte Tabelle](#) hinzu, wenn Sie Ereignisdaten abfragen möchten.
- [Erstellen Sie einen neuen ID-Namespace und ordnen](#) Sie ihn zu, wenn Sie eine ID-Zuordnungstabelle für die Abfrage von Identitätsdaten erstellen möchten.

Informationen darüber, wie Sie eine Kollaboration verlassen können, finden Sie unter [Verlassen einer Zusammenarbeit](#).

Kollaborationen bearbeiten

Als Ersteller einer Kollaboration können Sie die verschiedenen Teile einer Kollaboration bearbeiten.

Informationen zum Bearbeiten einer Kollaboration mithilfe von AWS SDKs finden Sie in der [AWS Clean Rooms API-Referenz](#).

Themen

- [Bearbeiten Sie den Namen und die Beschreibung der Zusammenarbeit](#)
- [Aktualisieren Sie die Analyse-Engine für die Zusammenarbeit](#)
- [Schalten Sie den Protokollspeicher aus](#)
- [Bearbeiten Sie die Einstellungen für Kollaborationsprotokolle](#)
- [Tags für die Zusammenarbeit bearbeiten](#)
- [Bearbeiten Sie Mitgliedskennungen](#)
- [Ein neues Mitglied hinzufügen](#)
- [Bearbeiten Sie die Fähigkeiten vorhandener Mitglieder](#)
- [Einstellungen für die automatische Genehmigung der Zusammenarbeit bearbeiten](#)
- [Bearbeiten Sie die zugehörigen Tabellen-Tags](#)
- [Bearbeiten Sie die Tags der Analysevorlage](#)
- [Bearbeiten Sie unterschiedliche Datenschutzrichtlinientags](#)

Bearbeiten Sie den Namen und die Beschreibung der Zusammenarbeit

Nachdem Sie die Kollaboration erstellt haben, können Sie nur den Namen und die Beschreibung der Kollaboration bearbeiten.

Um den Namen und die Beschreibung der Kollaboration zu bearbeiten

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#).
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus, die Sie erstellt haben.
4. Wählen Sie auf der Seite mit den Kollaborationsdetails Aktionen und dann Zusammenarbeit bearbeiten aus.

5. Bearbeiten Sie auf der Seite Kollaboration bearbeiten unter Details den Namen und die Beschreibung der Kollaboration.
6. Wählen Sie Änderungen speichern aus.

Aktualisieren Sie die Analyse-Engine für die Zusammenarbeit

Nachdem Sie die Kollaboration erstellt haben, können Sie die Analyse-Engine von AWS Clean Rooms SQL auf Spark umstellen.

Note

Wenn Sie die Analyse-Engine von AWS Clean Rooms SQL zu Spark ändern, könnten bestehende Workflows beeinträchtigt werden.

Um die Collaboration Analytics Engine zu aktualisieren

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#) mit Ihrem AWS-Konto (falls Sie dies noch nicht getan haben).
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus, die Sie erstellt haben.
4. Wählen Sie auf der Seite mit den Kollaborationsdetails Aktionen und dann Zusammenarbeit bearbeiten aus.
5. Wählen Sie Änderungen speichern aus.

Schalten Sie den Protokollspeicher aus

Wenn Sie die Analyseprotokollierung aktiviert haben, können Sie bearbeiten, ob die Analyseprotokolle in Ihrem Amazon CloudWatch Logs-Konto gespeichert werden.

Um die Protokollspeicherung zu deaktivieren

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#) mit Ihrem AWS-Konto (falls Sie dies noch nicht getan haben).
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus, für die die Analyseprotokollierung aktiviert ist.

4. Wählen Sie auf der Seite mit den Kollaborationsdetails die Option Aktionen und dann Protokollspeicher deaktivieren aus.

 Note

Es wird eine Warnung angezeigt, die auf Folgendes hinweist:

- Neue Anfragen werden nicht mehr in Ihrem CloudWatch Konto protokolliert.
- Bestehende Protokolle werden gemäß Ihren aktuellen Aufbewahrungseinstellungen aufbewahrt.
- Wenn Sie die Protokollierung in future wieder aktivieren, gilt sie nur für Anfragen, die nach der Reaktivierung gestellt wurden.
- Diese Änderung wirkt sich nur auf Ihre Logs aus — die Logging-Einstellungen anderer Teammitglieder bleiben unverändert.

5. Wählen Sie Deaktivieren aus.

Bearbeiten Sie die Einstellungen für Kollaborationsprotokolle

Wenn Sie die Abfrageprotokollierung aktiviert haben, können Sie bearbeiten, ob die Abfrageprotokolle in Ihrem Amazon CloudWatch Logs-Konto gespeichert werden.

Um die Einstellungen für Kollaborationsprotokolle zu bearbeiten

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#) mit Ihrem AWS-Konto (falls Sie dies noch nicht getan haben).
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus, die Sie erstellt haben.
4. Führen Sie auf der Seite mit den Kollaborationsdetails einen der folgenden Schritte aus:
 - Wählen Sie Aktionen und dann Protokolleinstellungen bearbeiten aus.
 - Wählen Sie auf der Registerkarte Protokolle die Option Protokolleinstellungen bearbeiten aus.
5. Gehen Sie im Modal Protokolleinstellungen bearbeiten für die Protokollspeicherung in Amazon CloudWatch Logs wie folgt vor:
 - Wenn Sie nicht möchten, dass für Sie relevante Protokolle in Ihrem Amazon CloudWatch Logs-Konto gespeichert werden, wählen Sie Ausschalten.

- Wenn Sie möchten, dass für Sie relevante Protokolle in Ihrem Amazon CloudWatch Logs-Konto gespeichert werden, wählen Sie Einschalten.

Sie können nur Protokolle für Abfragen erhalten, die Sie initiiert haben oder die Ihre Daten enthalten.

Das Mitglied, das Ergebnisse erhalten kann, erhält auch Protokolle für alle Abfragen, die in einer Kollaboration ausgeführt werden, auch wenn in einer Abfrage nicht auf seine Daten zugegriffen wird.

1. Wählen Sie unter Unterstützte Protokolltypen einen der Protokolltypen aus, die der Kollaborationsersteller für die Unterstützung ausgewählt hat:

- Wenn Sie anhand von SQL-Abfragen generierte Protokolle erhalten möchten, aktivieren Sie das Kontrollkästchen Protokolle aus Abfragen.
- Wenn Sie Protokolle empfangen möchten, die aus Aufträgen mit generiert wurden PySpark, aktivieren Sie das Kontrollkästchen Protokolle von Aufträgen.

6. Wählen Sie Änderungen speichern aus.

Note

Nachdem Sie die Protokollierung aktiviert haben, kann es einige Minuten dauern, bis der Protokollspeicher eingerichtet ist und Protokolle in Amazon CloudWatch Logs empfangen werden. Während dieser kurzen Zeit führt das Mitglied, das Abfragen durchführen kann, möglicherweise Abfragen durch, bei denen keine Protokolle gesendet werden.

Tags für die Zusammenarbeit bearbeiten

Als Ersteller einer Kollaboration können Sie, nachdem Sie eine Kollaboration erstellt haben, die Tags auf der Kollaborationsressource verwalten.

Um die Kollaborations-Tags zu bearbeiten

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#) mit Ihrem AWS-Konto (falls Sie dies noch nicht getan haben).
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus, die Sie erstellt haben.

4. Wählen Sie eine der folgenden Optionen:

Wenn Sie ...	Dann...
Der Ersteller der Kollaboration und ein Mitglied der Kollaboration	Wählen Sie die Registerkarte Details.
Der Ersteller der Kollaboration, aber kein Mitglied der Kollaboration	Scrollen Sie auf der Seite nach unten zum Abschnitt Tags.

5. Für Details zur Zusammenarbeit wählen Sie Tags verwalten aus.

6. Auf der Seite Tags verwalten haben Sie folgende Möglichkeiten:

- Klicken Sie zum Entfernen eines Tags auf Remove (Entfernen).
- Um einen Tag hinzuzufügen, wählen Sie Add new tag (Neuen Tag hinzufügen).
- Um Ihre Änderungen zu speichern, wählen Sie Änderungen speichern

Bearbeiten Sie Mitgliedskennungen

Als Ersteller einer Kollaboration können Sie, nachdem Sie eine Kollaboration erstellt haben, die Tags in der Mitgliedschaftsressource verwalten.

Um die Mitgliedschafts-Tags zu bearbeiten

1. Melde dich bei der an AWS-Managementkonsole und öffne die [AWS Clean Rooms Konsole](#) mit deinem AWS-Konto (falls du das noch nicht getan hast).
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus, die Sie erstellt haben.
4. Wählen Sie die Registerkarte Details.
5. Wählen Sie für Mitgliedschaftsdetails die Option Stichwörter verwalten aus.
6. Auf der Seite Mitgliedschafts-Tags verwalten können Sie Folgendes tun:
 - Klicken Sie zum Entfernen eines Tags auf Remove (Entfernen).
 - Um einen Tag hinzuzufügen, wählen Sie Add new tag (Neuen Tag hinzufügen).
 - Klicken Sie auf Save changes (Änderungen speichern), um die Änderungen zu speichern.

Ein neues Mitglied hinzufügen

Weitere Informationen finden Sie unter [Fügen Sie einer Kollaboration ein neues Mitglied hinzu](#).

Bearbeiten Sie die Fähigkeiten vorhandener Mitglieder

Weitere Informationen finden Sie unter [Aktualisiere die Fähigkeiten vorhandener Mitglieder](#).

Einstellungen für die automatische Genehmigung der Zusammenarbeit bearbeiten

Weitere Informationen finden Sie unter [Einstellungen für die automatische Genehmigung der Zusammenarbeit bearbeiten](#).

Bearbeiten Sie die zugehörigen Tabellen-Tags

Als Ersteller einer Kollaboration können Sie, nachdem Sie Tabellen mit einer Kollaboration verknüpft haben, die Tags in der zugehörigen Tabellenressource verwalten.

Um die zugehörigen Tabellen-Tags zu bearbeiten

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#) mit Ihrem AWS-Konto (falls Sie dies noch nicht getan haben).
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus, die Sie erstellt haben.
4. Wählen Sie die Registerkarte Tables (Tabellen).
5. Wählen Sie für Von Ihnen zugeordnete Tabellen eine Tabelle aus.
6. Wählen Sie auf der konfigurierten Tabellendetailseite für Tags die Option Tags verwalten aus.

Auf der Seite Tags verwalten haben Sie folgende Möglichkeiten:

- Klicken Sie zum Entfernen eines Tags auf Remove (Entfernen).
- Um einen Tag hinzuzufügen, wählen Sie Add new tag (Neuen Tag hinzufügen).
- Klicken Sie auf Save changes (Änderungen speichern), um die Änderungen zu speichern.

Bearbeiten Sie die Tags der Analysevorlage

Als Ersteller einer Kollaboration können Sie, nachdem Sie eine Kollaboration erstellt haben, die Tags in der Analysevorlagenressource verwalten.

Um die Mitgliedschafts-Tags zu bearbeiten

1. Melde dich bei der an AWS-Managementkonsole und öffne die [AWS Clean Rooms Konsole](#) mit deinem AWS-Konto (falls du das noch nicht getan hast).
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus, die Sie erstellt haben.
4. Wählen Sie die Registerkarte Templates (Vorlagen) aus.
5. Wählen Sie im Abschnitt Von Ihnen erstellte Analysevorlagen die Analysevorlage aus.
6. Scrollen Sie auf der Detailseite der Analysevorlagentabelle nach unten zum Abschnitt Tags.
7. Wählen Sie Tags verwalten aus.
8. Auf der Seite Tags verwalten haben Sie folgende Möglichkeiten:
 - Klicken Sie zum Entfernen eines Tags auf Remove (Entfernen).
 - Um einen Tag hinzuzufügen, wählen Sie Add new tag (Neuen Tag hinzufügen).
 - Klicken Sie auf Save changes (Änderungen speichern), um die Änderungen zu speichern.

Bearbeiten Sie unterschiedliche Datenschutzrichtlinientags

Als Ersteller einer Kollaboration können Sie, nachdem Sie eine Kollaboration erstellt haben, die Tags in der Analysevorlagenressource verwalten.

Um die Mitgliedschafts-Tags zu bearbeiten

1. Melde dich bei der an AWS-Managementkonsole und öffne die [AWS Clean Rooms Konsole](#) mit deinem AWS-Konto (falls du das noch nicht getan hast).
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus, die die differenzielle Datenschutzrichtlinie enthält, die Sie bearbeiten möchten.
4. Wählen Sie die Registerkarte Tables (Tabellen).
5. Wählen Sie auf der Registerkarte Tabellen die Option Tags verwalten aus.

6. Auf der Seite Tags verwalten haben Sie folgende Möglichkeiten:

- Klicken Sie zum Entfernen eines Tags auf Remove (Entfernen).
- Um einen Tag hinzuzufügen, wählen Sie Add new tag (Neuen Tag hinzufügen).
- Klicken Sie auf Save changes (Änderungen speichern), um die Änderungen zu speichern.

Anfragen ändern in AWS Clean Rooms

Änderungsanfragen ermöglichen es Ihnen, Änderungen an bestehenden Kollaborationseinstellungen zur Genehmigung durch andere Mitglieder der Kollaboration vorzuschlagen. Mit Änderungsanfragen können Sie eine Anfrage einreichen, um neue Mitglieder hinzuzufügen, die Fähigkeiten vorhandener Mitglieder zu aktualisieren und die Einstellungen für die automatische Genehmigung von Kollaborationen zu ändern. Alle Mitglieder der Kollaboration müssen Änderungsanträge genehmigen, damit die vorgeschlagenen Änderungen wirksam werden.

Änderungsanfragen sind kollaborationsspezifisch und können vom Ersteller der Kollaboration eingereicht werden.

Sie können eine Änderungsanfrage auf folgende Weise einreichen:

- Fügen Sie einer Kollaboration ein neues Mitglied hinzu
- Aktualisieren Sie die Fähigkeiten vorhandener Mitglieder
- Einstellungen für die automatische Genehmigung der Zusammenarbeit bearbeiten

Note

Sie müssen der Ersteller der Kollaboration sein, um eine Änderungsanfrage einreichen zu können.

Fügen Sie einer Kollaboration ein neues Mitglied hinzu

Um einer Kollaboration ein neues Mitglied hinzuzufügen, müssen Sie der Ersteller der Kollaboration sein. Das Hinzufügen eines neuen Mitglieds zu einer Kollaboration erfordert die manuelle Genehmigung durch bestehende Mitglieder der Kollaboration und es wird eine Änderungsanfrage eingereicht.

Du kannst einer Kollaboration neue Mitglieder hinzufügen, indem du die folgenden Schritte befolgst:

Um ein neues Mitglied zu einer Kollaboration hinzuzufügen

1. Melden Sie sich bei Ihrem Mitglied an AWS-Managementkonsole und öffnen Sie die [AWS Clean RoomsKonsole](#) AWS-Konto.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie Ihre Kollaboration aus, um zu Ihren Kollaborationsdetails zu gelangen.
4. Navigieren Sie auf der Kollaborationsseite zum Tab Mitglieder.
5. Wählen Sie in der Tabelle Mitglieder die Option Mitglieder bearbeiten aus.
6. Wählen Sie Weiteres Mitglied hinzufügen aus.
7. Geben Sie die neuen Mitgliedsinformationen ein:
 - Anzeigename des Mitglieds
 - AWS-KontoMitglieds-ID
 - Geben Sie an, ob das Mitglied Ergebnisse erhalten kann. Markieren Sie das Kästchen, um diesem Mitglied die Fähigkeit zu gewähren.
8. Wählen Sie Änderungen speichern aus.
9. Bestätigen Sie die Einreichung Ihres Änderungsantrags. Bestätigen Sie im Bestätigungsmodus die Änderungen und wählen Sie Änderungsantrag einreichen aus.

Note

Wenn in Ihrer Zusammenarbeit automatisch genehmigte Änderungstypen unterstützt werden, ist eine manuelle Genehmigung von Änderungsanträgen möglicherweise nicht erforderlich. Im Bereich „Übersicht“ der Kollaboration können Sie überprüfen, für welche Änderungstypen keine Änderungsanforderung erforderlich ist. Weitere Informationen finden Sie unter [the section called “Einstellungen für die automatische Genehmigung der Zusammenarbeit bearbeiten”](#).

Aktualisiere die Fähigkeiten vorhandener Mitglieder

Um die Fähigkeiten vorhandener Kollaborationsmitglieder zu aktualisieren, müssen Sie der Ersteller der Kollaboration sein. Für die Aktualisierung der Fähigkeiten vorhandener Mitglieder einer

Kollaboration ist eine manuelle Genehmigung durch die vorhandenen Mitglieder erforderlich, und es wird eine Änderungsanfrage eingereicht.

Die Fähigkeiten der Mitglieder, die aktualisiert werden können, sind:

- Kann Ergebnisse erhalten

Gehen Sie wie folgt vor, um die Mitgliederfähigkeiten vorhandener Kollaborationsmitglieder zu aktualisieren:

Um die Fähigkeiten vorhandener Mitglieder zu aktualisieren

1. Melden Sie sich bei Ihrem Mitglied an AWS-Managementkonsole und öffnen Sie die [AWS Clean RoomsKonsole](#) AWS-Konto.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie Ihre Kollaboration aus, um zu Ihren Kollaborationsdetails zu gelangen.
4. Navigieren Sie auf der Kollaborationsseite zum Tab Mitglieder.
5. Wählen Sie in der Tabelle Mitglieder die Option Mitglieder bearbeiten aus.
6. Geben Sie die Fähigkeit des Mitglieds an, das geändert werden kann.
7. Wählen Sie Änderungen speichern aus.
8. Bestätigen Sie die Einreichung Ihrer Änderungsanfrage. Bestätigen Sie im Bestätigungsmodus die Änderungen und wählen Sie Änderungsantrag einreichen aus.

 Note

Wenn in Ihrer Zusammenarbeit automatisch genehmigte Änderungstypen unterstützt werden, ist eine manuelle Genehmigung von Änderungsanträgen möglicherweise nicht erforderlich. Im Bereich „Übersicht“ der Kollaboration können Sie überprüfen, für welche Änderungstypen keine Änderungsanforderung erforderlich ist. Weitere Informationen finden Sie unter [the section called “Einstellungen für die automatische Genehmigung der Zusammenarbeit bearbeiten”](#).

Einstellungen für die automatische Genehmigung der Zusammenarbeit bearbeiten

Um die Kollaborationseinstellungen für automatische Genehmigungen zu bearbeiten, müssen Sie der Ersteller der Kollaboration sein und eine Änderungsanfrage zur Genehmigung durch andere Kollaborationsmitglieder einreichen.

Gehen Sie wie folgt vor, um die Einstellungen für die automatische Genehmigung für eine Kollaboration zu bearbeiten:

So bearbeiten Sie die Einstellungen für die automatische Genehmigung von Kollaborationen

1. Melden Sie sich bei Ihrem Mitglied AWS-Konto an AWS-Managementkonsole und öffnen Sie die [AWS Clean RoomsKonsole](#).
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie Ihre Kollaboration aus, um zu Ihren Kollaborationsdetails zu gelangen.
4. Wählen Sie auf der Kollaborationsseite die Schaltfläche Aktionen und dann auto Genehmigungen bearbeiten aus.
5. So gewähren Sie bestehenden Mitgliedern Funktionen, ohne eine Änderungsanfrage manuell zu genehmigen:

1. Navigiere zum Abschnitt „Fähigkeiten für Mitglieder gewähren“.
2. Geben Sie an, welche Fähigkeiten Mitgliedern automatisch gewährt werden können.

Note

Standardmäßig können alle Mitglieder der Kollaboration „Daten beisteuern“.

3. (Optional) Damit neue Mitglieder sofort einer Kollaboration mit den angegebenen Fähigkeiten beitreten können, ohne Änderungsanträge manuell genehmigen zu müssen, wählen Sie Neue Mitglieder mit diesen Fähigkeiten automatisch genehmigen.
6. Gehen Sie wie folgt vor, um das Entfernen von Fähigkeiten vorhandener Mitglieder ohne Änderungsanfrage zu ermöglichen:
 1. Navigiere zum Abschnitt Fähigkeiten, die automatisch entzogen werden können.
 2. Gib an, welche Fähigkeiten von Mitgliedern automatisch entfernt werden können.
7. Wählen Sie Änderungen speichern aus.

8. Bestätigen Sie die Einreichung Ihres Änderungsantrags. Bestätigen Sie im Bestätigungsmodus die Änderungen und wählen Sie Änderungsantrag einreichen aus.

Kollaborationen löschen

Als Ersteller einer Kollaboration können Sie eine von Ihnen erstellte Kollaboration löschen.

Note

Wenn Sie eine Kollaboration löschen, können Sie und alle Mitglieder keine Abfragen ausführen, keine Ergebnisse empfangen oder Daten beitragen. Jedes Mitglied der Kollaboration hat im Rahmen seiner Mitgliedschaft weiterhin Zugriff auf seine eigenen Daten.

Um eine Kollaboration zu löschen

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#).
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus, die Sie löschen möchten.
4. Wählen Sie unter Aktionen die Option Kollaboration löschen aus.
5. Bestätigen Sie den Löschvorgang und wählen Sie dann Löschen.

Kollaborationen anzeigen

Als Ersteller einer Kollaboration können Sie sich alle Kollaborationen ansehen, die Sie erstellt haben.

Um Kollaborationen anzusehen

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#).
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Sehen Sie sich auf der Seite Kollaborationen unter Zuletzt verwendet die letzten 5 verwendeten Kollaborationen an.
4. Sehen Sie sich auf der Registerkarte Mit aktiver Mitgliedschaft die Liste der Kollaborationen mit aktiver Mitgliedschaft an.

Sie können nach dem Namen, dem Erstellungsdatum der Mitgliedschaft und Ihren Mitgliedsdetails sortieren.

Sie können die Suchleiste verwenden, um nach einer Kollaboration zu suchen.

5. Sehen Sie sich auf der Registerkarte „Für den Beitritt verfügbar“ die Liste der Kollaborationen an, denen Sie beitreten können.
6. Sehen Sie sich auf der Registerkarte Nicht mehr verfügbar die Liste der gelöschten Kollaborationen und Mitgliedschaften für Kollaborationen an, die nicht mehr verfügbar sind (entfernte Mitgliedschaften).

Mitglieder zu einer Kollaboration einladen

Als Ersteller einer Kollaboration können Sie, nachdem Sie eine Kollaboration erstellt haben, einen Einladungslink an die auf der Registerkarte Mitglieder aufgeführten Mitglieder senden.

Um Mitglieder zu einer Kollaboration einzuladen

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#).
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus, die Sie erstellt haben.
4. Wählen Sie den Tab Mitglieder.
5. Wählen Sie in der Tabelle Mitglieder die Schaltfläche Einladungslink kopieren aus.

Der Einladungslink wird kopiert.

6. Fügen Sie den Einladungslink in die sichere Kommunikationsmethode Ihrer Wahl ein und senden Sie ihn an jedes Mitglied der Kollaboration.

Mitglieder überwachen

Als Ersteller einer Kollaboration können Sie, nachdem Sie eine Kollaboration erstellt haben, den Status aller Mitglieder auf der Registerkarte Mitglieder überwachen. Dies trägt dazu bei, eine angemessene Zugriffskontrolle und ein angemessenes Kostenmanagement in Ihrer gesamten Zusammenarbeit sicherzustellen.

Um den Status eines Mitglieds zu überwachen

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#).
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus, die Sie erstellt haben.
4. Wählen Sie den Tab Mitglieder.
5. Überprüfen Sie in der Tabelle Mitglieder den Status der einzelnen Mitglieder.
6. Prüfen Sie in der Tabelle mit den Fähigkeiten der Mitglieder, welche Mitglieder Abfragen durchführen, Ergebnisse abrufen, Daten beitragen und andere Aufgaben ausführen können.
7. Prüfen Sie in der Tabelle mit den Zahlungskonfigurationen, welche Mitglieder für Abfragen, ID-Zuordnungstabellen und ML-Modellierung zahlen.

Mitglieder zu einer Kollaboration hinzufügen

Voraussetzungen

- Und AWS-Konto mit Berechtigungen zur Verwaltung von Kollaborationen
- Die Anzahl AWS-Konto IDs der Mitglieder, die Sie hinzufügen möchten

Um ein Mitglied zu einer Kollaboration hinzuzufügen

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean RoomsKonsole](#).
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus, zu der Sie Mitglieder hinzufügen möchten.
4. Wählen Sie den Tab Mitglieder.
5. Wählen Sie Edit members (Mitglieder bearbeiten) aus.
6. Wählen Sie Weiteres Mitglied hinzufügen und geben Sie die folgenden Informationen ein:
 - Anzeigename des Mitglieds
 - AWS-KontoMitglieds-ID
 - Geben Sie an, ob sie Ergebnisse erhalten können

7. Wählen Sie Änderungen speichern aus.
8. Vergewissern Sie sich im Bestätigungsmodus, dass die Informationen korrekt sind, und wählen Sie Änderungsantrag einreichen aus.
 - Wenn für die Änderungsanfrage die Zustimmung anderer Mitglieder erforderlich ist, müssen alle vorhandenen Mitglieder die Änderungsanfrage genehmigen, bevor das neue Mitglied hinzugefügt wird. Weitere Informationen zu Änderungsanträgen finden Sie unter [the section called “Anfragen ändern”](#).
 - Wenn die Änderungstypen der automatischen Genehmigung für neue Mitglieder mit den angegebenen Fähigkeiten unterstützt werden, wird diese Änderung sofort wirksam. Weitere Informationen finden Sie unter [the section called “Einstellungen für die automatische Genehmigung der Zusammenarbeit bearbeiten”](#).
9. Vergewissern Sie sich auf der Seite mit den Kollaborationsdetails unter dem Tab Mitglieder, dass der Mitgliedsstatus der hinzugefügten Mitglieder als Eingeladen angezeigt wird.

Nach Abschluss dieser Schritte können die eingeladenen Mitglieder der Kollaboration beitreten. Weitere Informationen zum Beitritt zu einer Kollaboration finden Sie unter [Eine Mitgliedschaft erstellen und einer Kollaboration beitreten](#)

Mitglieder aus einer Kollaboration entfernen

Note

Bevor Sie beginnen, beachten Sie, dass das Entfernen eines Mitglieds:

- Entfernt alle zugehörigen Datensätze aus der Kollaboration
- Wenn das [Mitglied die Rechenkosten für Abfragen bezahlt](#), stoppt diese Aktion die gesamte Abfrageausführung in der Kollaboration.

Durch das Entfernen eines Mitglieds werden auch alle zugehörigen Datensätze aus der Kollaboration entfernt.

Voraussetzungen

- Sie müssen ein Ersteller einer Kollaboration sein

- Sie können Ihr eigenes Konto nicht entfernen

Um ein Mitglied aus einer Kollaboration zu entfernen

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#).
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus, die Sie ändern möchten.
4. Wählen Sie den Tab Mitglieder.
5. Wählen Sie das Optionsfeld neben dem Mitglied aus, das Sie entfernen möchten.
6. Wählen Sie Remove (Entfernen) aus.
7. Geben Sie im Bestätigungsdialogfeld einen Text ein, **confirm** um das Entfernen zu bestätigen.

 Note

Nachdem Sie ein Mitglied entfernt haben, werden auch alle mit seinem Konto verknüpften Datensätze aus der Kollaboration entfernt.

 Important

Wenn Sie das Mitglied entfernen, das die Rechenkosten für Abfragen bezahlt, können in der Kollaboration keine weiteren Abfragen ausgeführt werden, bis Sie ein neues zahlendes Mitglied benannt haben.

Verlassen einer Zusammenarbeit

Als Mitglied einer Kollaboration können Sie eine Kollaboration verlassen, indem Sie Ihre Mitgliedschaft löschen. Wenn Sie der Ersteller der Kollaboration sind, können Sie eine Kollaboration nur verlassen, indem Sie [die Kollaboration löschen](#).

 Note

Wenn Sie Ihre Mitgliedschaft löschen, verlassen Sie die Kollaboration und können ihr nicht wieder beitreten. Wenn Sie das [Mitglied sind, das die Kosten für die Datenverarbeitung](#)

[bei Abfragen bezahlt](#), und Sie Ihre Mitgliedschaft löschen, dürfen keine Abfragen mehr ausgeführt werden.

Um eine Kollaboration zu verlassen

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#).
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie unter Mit aktiver Mitgliedschaft die Kollaboration aus, bei der Sie Mitglied sind.
4. Wählen Sie Aktionen.
5. Wählen Sie Mitgliedschaft löschen.
6. Bestätigen Sie im Dialogfeld die Entscheidung, die Kollaboration zu verlassen, indem Sie etwas **confirm** in das Texteingabefeld eingeben, und wählen Sie dann Leeren und Mitgliedschaft löschen aus.

Auf der Konsole wird eine Meldung angezeigt, dass die Mitgliedschaft gelöscht wurde.

Für den Ersteller der Kollaboration wird der Mitgliedsstatus „Links“ angezeigt.

Bereiten Sie Datentabellen vor in AWS Clean Rooms

Note

Die Vorbereitung von Datentabellen kann vor oder nach dem Beitritt zu einer Kollaboration erfolgen. Nachdem eine Tabelle vorbereitet wurde, können Sie sie in mehreren Kollaborationen wiederverwenden, sofern Ihre Datenschutzerfordernisse für diese Tabelle dieselben sind.

Als Mitglied der Kollaboration müssen Sie Ihre Datentabellen vorbereiten, bevor sie AWS Clean Rooms von dem Mitglied der Kollaboration, das Abfragen durchführen kann, abgefragt werden können.

Bei den Datentabellen, die Sie für Abfragen verwenden, AWS Clean Rooms handelt es sich im Allgemeinen um dieselben Datentabellentypen, die Sie für andere Anwendungen verwenden. Beispielsweise werden dieselben Datasetypen mit Amazon Athena, Amazon EMR, Amazon Redshift Spectrum und Amazon Quick verwendet.

Sie können die Daten in ihrem ursprünglichen Format direkt aus einer der folgenden Datenquellen abfragen:

- Amazon Simple Storage Service (Amazon-S3)
- Amazon Athena
- Snowflake

AWS Clean Rooms greift zur Laufzeit der Abfrage auf den Datensatz zu und stellt so sicher, dass Mitglieder, die Abfragen können, immer auf die meisten up-to-date Daten zugreifen können. Alle Daten, die vorübergehend in eine AWS Clean Rooms Kollaboration eingelesen werden, werden nach Abschluss der Abfrage gelöscht. Die Abfrageergebnisse werden in Ihren Amazon S3 S3-Bucket geschrieben.

Wenn Ihr Anwendungsfall das Abfragen von Identitätsdaten beinhaltet, finden Sie weitere Informationen unter [AWS Entity Resolution in AWS Clean Rooms](#).

Topics

- [Datenformate für AWS Clean Rooms](#)

- [Apache Iceberg Tische in AWS Clean Rooms](#)
- [Vorbereiten von Datentabellen für Abfragen in AWS Clean Rooms](#)
- [Vorbereiten verschlüsselter Datentabellen mit Cryptographic Computing für Clean Rooms](#)
- [Datentabellen mit dem C3R-Verschlüsselungsclient entschlüsseln](#)

Datenformate für AWS Clean Rooms

Um Daten zu analysieren, müssen die Datensätze in einem Format vorliegen, das dies AWS Clean Rooms unterstützt.

Themen

- [Unterstützte Datenformate für Jobs PySpark](#)
- [Unterstützte Datenformate für SQL-Abfragen](#)
- [Unterstützte Datentypen](#)
- [Arten der Dateikomprimierung für AWS Clean Rooms](#)
- [Serverseitige Verschlüsselung für AWS Clean Rooms](#)

Unterstützte Datenformate für Jobs PySpark

AWS Clean Rooms unterstützt die folgenden strukturierten Formate für die Ausführung von PySpark Jobs.

- Parquet
- OpenCSV
- JSON

Unterstützte Datenformate für SQL-Abfragen

AWS Clean Rooms unterstützt die folgenden strukturierten Formate für die Ausführung von SQL-Abfragen.

- [Apache Iceberg-Tabellen](#)
- Parquet
- OpenCSV

- JSON

 Note

Ein timestamp Wert in einer Textdatei muss das folgende Format yyyy-MM-dd HH:mm:ss.SSSSSS haben. Zum Beispiel:2017-05-01 11:30:59.000000.

Wir empfehlen die Verwendung eines spaltenförmigen Speicherdateiformats wie Apache Parquet. Mit einem spaltenbasierten Speicherdateiformat können Sie Datenbewegungen minimieren, indem Sie nur die Spalten auswählen, die Sie benötigen. Für eine optimale Leistung sollten große Objekte in Objekte mit einer Größe von 100 MB bis 1 GB aufgeteilt werden.

Unterstützte Datentypen

AWS Clean Rooms unterstützt die folgenden Datentypen.

- ARRAY
- BIGINT
- BOOLEAN
- BYTE
- CHAR
- DATE
- DECIMAL
- FLOAT
- INTEGER
- INTERVAL
- LONG
- MAP
- REAL
- SHORT
- SMALLINT
- STRUCT
- TIME

- `TIMESTAMP_LTZ`
- `TIMESTAMP_NTZ`
- `TINYINT`
- `VARCHAR`

Weitere Informationen finden Sie unter [Datentypen](#) in der SQL-Referenz.AWS Clean Rooms

Arten der Dateikomprimierung für AWS Clean Rooms

Um Speicherplatz zu reduzieren, die Leistung zu verbessern und die Kosten zu minimieren, empfehlen wir dringend, Ihre Datensätze zu komprimieren.

AWS Clean Rooms erkennt Dateikomprimierungstypen anhand der Dateierweiterung und unterstützt die in der folgenden Tabelle aufgeführten Komprimierungstypen und -erweiterungen.

Komprimierungsalgorithmus	Dateierweiterung
GZIP	.gz
Bzip2	.bz2
Snappy	.snappy

Sie können die Komprimierung auf verschiedenen Ebenen anwenden. Zumeist komprimieren Sie eine ganze Datei oder einzelne Blöcke innerhalb einer Datei. Das Komprimieren von Spaltenformaten auf Dateiebene bringt keine Leistungsvorteile.

Serverseitige Verschlüsselung für AWS Clean Rooms

Note

Die serverseitige Verschlüsselung ersetzt nicht die kryptografische Datenverarbeitung in den Anwendungsfällen, in denen sie erforderlich ist.

AWS Clean Rooms entschlüsselt transparent Datensätze, die mit den folgenden Verschlüsselungsoptionen verschlüsselt wurden:

- SSE-S3 — Serverseitige Verschlüsselung mit einem AES-256-Verschlüsselungsschlüssel, der von Amazon S3 verwaltet wird
- SSE-KMS — Serverseitige Verschlüsselung mit Schlüsseln, die verwaltet werden von AWS Key Management Service

Um SSE-S3 verwenden zu können, muss die AWS Clean Rooms Servicerolle, die verwendet wird, um die konfigurierte Tabelle der Kollaboration zuzuordnen, über KMS-Decrypt-Berechtigungen verfügen. Um SSE-KMS verwenden zu können, muss die KMS-Schlüsselrichtlinie auch die Entschlüsselung der Servicerolle zulassen. AWS Clean Rooms

AWS Clean Rooms unterstützt keine clientseitige Amazon S3 S3-Verschlüsselung. Weitere Informationen zur serverseitigen Verschlüsselung finden Sie unter [Schutz von Daten mithilfe serverseitiger Verschlüsselung](#) im Amazon Simple Storage Service-Benutzerhandbuch.

Apache Iceberg Tische in AWS Clean Rooms

Apache Iceberg ist ein Open-Source-Tabellenformat für Data Lakes. AWS Clean Rooms kann die in gespeicherten Statistiken verwenden Apache Iceberg Metadaten zur Optimierung von Abfrageplänen und zur Reduzierung der Anzahl von Dateiscans bei der Verarbeitung von Abfragen im Reinraum. Weitere Informationen finden Sie in der [Apache Iceberg-Dokumentation](#).

Beachten Sie bei der Verwendung AWS Clean Rooms mit Iceberg-Tabellen Folgendes:

- Apache Iceberg-Tabellen für S3 — Apache Iceberg Tabellen müssen in der [Implementierung des AWS Glue Data Catalog Open-Source-Glue-Katalogs](#) definiert werden.
- Apache Iceberg-Tabellen für Athena — [Weitere Informationen finden Sie unter -iceberg.html https://docs.aws.amazon.com/athena/latest/ug/querying](#)
- Apache Iceberg-Tabellen für Snowflake — [Weitere Informationen finden Sie unter user-guide/tables-iceberg https://docs.snowflake.com/en/](#)
- Parquet-Dateiformat — unterstützt AWS Clean Rooms nur Iceberg-Tabellen im Parquet-Datendateiformat.
- GZIP- und Snappy-Komprimierung — AWS Clean Rooms unterstützt Parquet mit GZIP und Snappy Komprimierung.
- Iceberg-Versionen — AWS Clean Rooms unterstützt das Ausführen von Abfragen für Iceberg-Tabellen der Versionen 1 und 2.

- Partitionen — Sie müssen keine Partitionen manuell hinzufügen Apache Iceberg Tabellen in AWS Glue. AWS Clean Rooms erkennt neue Partitionen in Apache Iceberg Tabellen automatisch und es ist kein manueller Vorgang erforderlich, um Partitionen in der Tabellendefinition zu aktualisieren. Iceberg-Partitionen erscheinen als reguläre Spalten im AWS Clean Rooms Tabellenschema und nicht separat als Partitionsschlüssel im konfigurierten Tabellenschema.

- Einschränkungen

- Nur neue Iceberg-Tabellen

Apache Iceberg Tabellen wurden konvertiert von Apache Parquet Tabellen werden nicht unterstützt.

- Zeitreiseabfragen

AWS Clean Rooms unterstützt keine Zeitreiseabfragen mit Apache Iceberg Tabellen.

- Athena-Motorversion 2

Iceberg Tabellen, die mit der Athena-Engine Version 2 erstellt wurden, werden nicht unterstützt.

- Dateiformate

Avro und ORC-Dateiformate (Optimized Row Columnar) werden nicht unterstützt.

- Komprimierung

Zstandard (Zstd) -Komprimierung für Parquet wird nicht unterstützt.

Unterstützte Datentypen für Iceberg-Tabellen

AWS Clean Rooms kann abfragen Iceberg Tabellen, die die folgenden Datentypen enthalten:

- BOOLEAN
- DATE
- DECIMAL
- DOUBLE
- FLOAT
- INT
- LIST
- LONG

- MAP
- STRING
- STRUCT
- TIMESTAMP WITHOUT TIME ZONE

Weitere Informationen zu Iceberg-Datentypen finden Sie unter [Schemata für Iceberg](#) in der Apache-Iceberg-Dokumentation.

Vorbereiten von Datentabellen für Abfragen in AWS Clean Rooms

Wenn Ihr Anwendungsfall nicht erfordert, dass Sie Ihre eigenen Daten mitbringen, können Sie dieses Verfahren überspringen.

Wenn Ihr Anwendungsfall das Abfragen von Identitätsdaten beinhaltet, finden Sie weitere Informationen unter [AWS Entity Resolution in AWS Clean Rooms](#).

Weitere Informationen zu den Datenformaten, die Sie verwenden können, finden Sie unter [Datenformate für AWS Clean Rooms](#).

Themen

- [Vorbereiten von Datentabellen in Amazon S3](#)
- [Vorbereiten von Datentabellen in Amazon Athena](#)
- [Datentabellen in Snowflake vorbereiten](#)

Vorbereiten von Datentabellen in Amazon S3

Sie können Datentabellen analysieren, die in Amazon S3 katalogisiert AWS Glue und gespeichert wurden. Wenn Ihre Datentabellen bereits katalogisiert sind AWS Glue, fahren Sie mit fort. [Erstellen einer konfigurierten Tabelle in AWS Clean Rooms](#)

Important

Wenn Sie Ihre Daten in Amazon S3 für die Verwendung mit vorbereiten AWS Clean Rooms, müssen Sie sicherstellen, dass Ihr AWS Glue Tabellenspeicherort genau dem Verzeichnispfad entspricht, in dem Ihre Datendateien gespeichert sind.

Beispiel: Wenn Ihre Daten in `s3://mybucket/folder/subfolder/data.parquet`` gespeichert sind, muss Ihr AWS Glue Tabellenspeicherort auf `zeigens3://mybucket/`

folder/subfolder/. Wenn Sie den Speicherort der Tabelle auf ein übergeordnetes Verzeichnis (s3://mybucket/folder/`) setzen, erscheint die Tabelle bei der Abfrage leer.

Die Vorbereitung Ihrer Datentabellen in Amazon S3 umfasst die folgenden Schritte:

Themen

- [Schritt 1: Erfüllen der Voraussetzungen](#)
- [Schritt 2: \(Optional\) Bereiten Sie Ihre Daten für die kryptografische Datenverarbeitung vor](#)
- [Schritt 3: Laden Sie Ihre Datentabelle auf Amazon S3 hoch](#)
- [Schritt 4: Erstellen Sie eine AWS Glue Tabelle](#)
- [Schritt 5: Nächste Schritte](#)

Schritt 1: Erfüllen der Voraussetzungen

Um Ihre Datentabellen für die Verwendung mit vorzubereiten AWS Clean Rooms, müssen Sie die folgenden Voraussetzungen erfüllen:

- Ihre Datentabellen werden in einem der [unterstützten Datenformate für](#) gespeichert AWS Clean Rooms.
- Ihre Datentabellen sind katalogisiert AWS Glue und verwenden die [unterstützten Datentypen für AWS Clean Rooms](#).
- Alle Ihre Datentabellen werden in Amazon Simple Storage Service (Amazon S3) in derselben Datei gespeichert, AWS-Region in der die Zusammenarbeit erstellt wurde.
- Sie AWS Glue Data Catalog müssen sich in derselben Region wie die Kollaboration befinden.
- Das AWS Glue Data Catalog ist dasselbe AWS-Konto wie die Mitgliedschaft.
- Der Amazon S3 S3-Bucket ist nicht bei registriert AWS Lake Formation.

Schritt 2: (Optional) Bereiten Sie Ihre Daten für die kryptografische Datenverarbeitung vor

(Optional) Wenn Sie kryptografische Berechnungen verwenden und Ihre Datentabelle vertrauliche Informationen enthält, die Sie verschlüsseln möchten, müssen Sie die Datentabelle mit dem C3R-Verschlüsselungssclient verschlüsseln.

Gehen Sie wie unter beschrieben vor, um Ihre Daten für die kryptografische Datenverarbeitung vorzubereiten. [Vorbereiten verschlüsselter Datentabellen mit Cryptographic Computing für Clean Rooms](#)

Schritt 3: Laden Sie Ihre Datentabelle auf Amazon S3 hoch

Note

Wenn Sie beabsichtigen, verschlüsselte Datentabellen in der Zusammenarbeit zu verwenden, müssen Sie zuerst die Daten für kryptografische Berechnungen verschlüsseln, bevor Sie Ihre Datentabelle auf Amazon S3 hochladen. Weitere Informationen finden Sie unter [Vorbereiten verschlüsselter Datentabellen mit Cryptographic Computing für Clean Rooms](#).

So laden Sie Ihre Datentabelle auf Amazon S3 hoch

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die Amazon S3 S3-Konsole unter <https://console.aws.amazon.com/s3/>.
2. Wählen Sie Buckets und dann einen Bucket aus, in dem Sie Ihre Datentabelle speichern möchten.
3. Wählen Sie Hochladen und folgen Sie dann den Anweisungen.
4. Wählen Sie die Registerkarte Objekte, um das Präfix anzuzeigen, in dem Ihre Daten gespeichert sind. Notieren Sie sich den Namen des Ordners.

Sie können den Ordner auswählen, um die Daten anzuzeigen.

Schritt 4: Erstellen Sie eine AWS Glue Tabelle

Wenn Sie bereits über eine AWS Glue Datentabelle verfügen, können Sie diesen Schritt überspringen.

In diesem Schritt richten Sie einen Crawler ein AWS Glue , der alle Dateien in Ihrem S3-Bucket crawlt und eine AWS Glue Tabelle erstellt. Weitere Informationen finden Sie [im AWS Glue Benutzerhandbuch unter Definieren von Crawlern](#).AWS Glue

Weitere Informationen zu unterstützten AWS Glue Data Catalog Datentypen finden Sie unter [Unterstützte Datentypen](#).

 Note

AWS Clean Rooms unterstützt derzeit keine S3-Buckets, bei AWS Lake Formation denen registriert ist.

Das folgende Verfahren beschreibt, wie Sie eine AWS Glue Tabelle erstellen. Wenn Sie ein AWS Glue Data Catalog verschlüsseltes Objekt mit einem Schlüssel AWS Key Management Service (AWS KMS) verwenden möchten, müssen Sie die KMS-Schlüsselberechtigungsrichtlinie so konfigurieren, dass der Zugriff auf diese verschlüsselte Tabelle möglich ist. Weitere Informationen finden Sie unter [Verschlüsselung in AWS Glue einrichten](#) im AWS Glue Entwicklerhandbuch.

Um eine AWS Glue Tabelle zu erstellen

1. Folgen Sie dem Verfahren [Arbeiten mit Crawlern auf der AWS Glue Konsole](#) im AWS Glue Benutzerhandbuch.
2. Notieren Sie sich den AWS Glue Datenbanknamen und den AWS Glue Tabellennamen.

Schritt 5: Nächste Schritte

Nachdem Sie Ihre Datentabellen in Amazon S3 vorbereitet haben, können Sie:

- [Erstellen Sie eine konfigurierte Tabelle](#)
- [Erstellen Sie ein ML-Modell](#)

Die Tabellen können abgefragt werden nach:

- Der Ersteller der Kollaboration hat eine Kollaboration in eingerichtet. AWS Clean Rooms Weitere Informationen finden Sie unter [Eine Zusammenarbeit erstellen](#).
- Der Ersteller der Kollaboration hat die Kollaborations-ID an Sie als Teilnehmer der Kollaboration gesendet.

Vorbereiten von Datentabellen in Amazon Athena

Sie können Datentabellen abfragen, die als AWS Glue Data Catalog (GDC) Views in Amazon Athena erstellt wurden.

Eine GDC-Ansicht ist eine virtuelle Tabelle, die aus einer oder mehreren zugrunde liegenden Tabellen erstellt wurde. AWS Glue Es muss mit Athena SQL im `AwsGlueCatalog` Athena-Katalog erstellt werden.

Die Vorbereitung Ihrer Datentabellen in Amazon Athena umfasst die folgenden Schritte:

Themen

- [Schritt 1: Erfüllen der Voraussetzungen](#)
- [Schritt 2: \(Optional\) Bereiten Sie Ihre Daten für die kryptografische Datenverarbeitung vor](#)
- [Schritt 3: Die nächsten Schritte](#)

Schritt 1: Erfüllen der Voraussetzungen

Um Ihre Datentabellen für die Verwendung mit vorzubereiten AWS Clean Rooms, müssen Sie die folgenden Voraussetzungen erfüllen:

- Ihre Datentabellen werden in einem der [unterstützten Datenformate für](#) gespeichert AWS Clean Rooms.
- Ihre Datentabellen verwenden die [unterstützten Datentypen für AWS Clean Rooms](#).
- Sie haben mit Athena SQL im Athena-Katalog eine GDC-Ansicht für Ihre AWS Glue Tabelle erstellt. `AwsDataCatalog`

Die Ansicht wird angezeigt in:

- Die Athena-Konsole (unter dem `AwsDataCatalog`) als Ansicht: <https://console.aws.amazon.com/athena/>
- Die AWS Glue Konsole als AWS Glue Tabelle: <https://console.aws.amazon.com/glue/>

Weitere Informationen finden Sie unter [Verwenden von Datenkatalogansichten in Athena](#) im Amazon Athena Athena-Benutzerhandbuch.

Note

Sie benötigen die entsprechenden Berechtigungen, um Ansichten in Athena und AWS Glue zu erstellen. Stellen Sie außerdem sicher, dass Sie Zugriff auf die zugrunde liegenden Tabellen haben, auf die in Ihrer View-Definition verwiesen wird.

AWS Clean Rooms unterstützt nur den AWS Glue Katalogtyp für Athena, keine Lambda- oder Hive-Katalogtypen.

- Ihre Datentabellen oder GDC-Ansichten sind katalogisiert und dort registriert. AWS Glue AWS Lake Formation
- Sie haben in Amazon S3 einen separaten Ausgabe-Bucket erstellt, um die Athena-Ergebnisse zu erhalten.
- Sie haben eine Servicerolle eingerichtet, um die Daten von Amazon Athena zu lesen. Weitere Informationen finden Sie unter [Erstellen Sie eine Servicerolle zum Lesen von Daten aus Amazon Athena](#).
- Die Servicerolle verfügt über die Zugriffsberechtigungen Lake Formation Select und Describe für die GDC-Ansicht oder -Tabelle.

Schritt 2: (Optional) Bereiten Sie Ihre Daten für die kryptografische Datenverarbeitung vor

(Optional) Wenn Sie kryptografische Berechnungen verwenden und Ihre Datentabelle vertrauliche Informationen enthält, die Sie verschlüsseln möchten, müssen Sie die Datentabelle mit dem C3R-Verschlüsselungssclient verschlüsseln.

Gehen Sie wie unter beschrieben vor, um Ihre Daten für die kryptografische Datenverarbeitung vorzubereiten. [Vorbereiten verschlüsselter Datentabellen mit Cryptographic Computing für Clean Rooms](#)

Schritt 3: Die nächsten Schritte

Nachdem Sie Ihre Datentabellen in Amazon Athena vorbereitet haben, können Sie:

- [Erstellen Sie eine konfigurierte Tabelle](#)
- [Erstellen Sie ein ML-Modell](#)

Die Tabellen können abgefragt werden nach:

- Der Ersteller der Kollaboration hat eine Kollaboration in eingerichtet. AWS Clean Rooms Weitere Informationen finden Sie unter [Eine Zusammenarbeit erstellen](#).
- Der Ersteller der Kollaboration hat die Kollaborations-ID an Sie als Teilnehmer der Kollaboration gesendet.

Datentabellen in Snowflake vorbereiten

Sie können Datentabellen abfragen, die im Snowflake Data Warehouse gespeichert wurden.

Die Vorbereitung Ihrer Datentabellen in Snowflake umfasst die folgenden Schritte:

Themen

- [Schritt 1: Erfüllen der Voraussetzungen](#)
- [Schritt 2: \(Optional\) Bereiten Sie Ihre Daten für kryptografische Berechnungen vor](#)
- [Schritt 3: Erstellen Sie ein Geheimnis AWS Secrets Manager](#)
- [Schritt 4: Nächste Schritte](#)

Schritt 1: Erfüllen der Voraussetzungen

Um Ihre Datentabellen für die Verwendung mit vorzubereiten AWS Clean Rooms, müssen Sie die folgenden Voraussetzungen erfüllen:

- Ihnen wurden die AWS-Konto entsprechenden Berechtigungen zum Lesen Ihrer Datentabellen erteilt. Weitere Informationen finden Sie unter [Erstellen Sie eine Servicerolle, um Daten aus Snowflake zu lesen](#).
- Ihre Datentabellen werden in einem der [unterstützten Datenformate für](#) gespeichert AWS Clean Rooms.
- Ihre Datentabellen verwenden die [unterstützten Datentypen für AWS Clean Rooms](#).
- Ihre Datentabelle wird in einem Snowflake-Warehouse gespeichert. Weitere Informationen finden Sie in der [Snowflake-Dokumentation](#).
- Sie haben einen neuen Snowflake-Benutzer mit Leseberechtigungen für die Snowflake-Tabelle eingerichtet, die Sie Ihrer Kollaboration zuordnen werden.

Schritt 2: (Optional) Bereiten Sie Ihre Daten für kryptografische Berechnungen vor

(Optional) Wenn Sie kryptografische Berechnungen verwenden und Ihre Datentabelle vertrauliche Informationen enthält, die Sie verschlüsseln möchten, müssen Sie die Datentabelle mit dem C3R-Verschlüsselungssclient verschlüsseln.

Gehen Sie wie unter beschrieben vor, um Ihre Daten für die kryptografische Datenverarbeitung vorzubereiten. [Vorbereiten verschlüsselter Datentabellen mit Cryptographic Computing für Clean Rooms](#)

Schritt 3: Erstellen Sie ein Geheimnis AWS Secrets Manager

Um von aus eine Verbindung zu Snowflake herzustellen AWS Clean Rooms, müssen Sie Ihre Snowflake-Anmeldeinformationen in einem AWS Secrets Manager Secret erstellen und speichern und dieses Secret dann einer Snowflake-Tabelle unter zuordnen. AWS Clean Rooms

Note

Wir empfehlen Ihnen, einen neuen Benutzer zu erstellen, der ausschließlich für bestimmt ist. AWS Clean Rooms Dieser Benutzer sollte nur eine Rolle mit Leseberechtigungen für die Daten haben, auf die Sie zugreifen AWS Clean Rooms möchten.

Um ein AWS Secrets Manager Geheimnis zu erstellen

1. Generieren Sie in Snowflake einen Benutzer `snowflakeUser` und richten Sie die Schlüsselpaar-Authentifizierung ein.

Note

Im November 2025 wird Snowflake dazu übergehen, nur die Schlüsselpaar-Authentifizierung zu unterstützen. Diese Änderung wird sich auf die aktuelle AWS Clean Rooms Integration mit Snowflake auswirken, die die Authentifizierung mit Benutzername und Passwort verwendet. Nach diesem Datum ist für Snowflake-Verbindungen eine Schlüsselpaar-Authentifizierung mit einem AWS Clean Rooms privaten Snowflake-PEM-Schlüssel (Privacy Enhanced Mail) erforderlich.

2. Ermitteln Sie, mit welchem Snowflake-Warehouse dieser Benutzer interagieren wird, `snowflakeWarehouse` Stellen Sie es entweder `snowflakeUser` in Snowflake als `DEFAULT_WAREHOUSE` für ein oder merken Sie es sich für den nächsten Schritt.
3. Erstellen Sie in [AWS Secrets Manager](#) ein Secret mit Ihren Snowflake-Anmeldeinformationen. Um ein Geheimnis in Secrets Manager zu erstellen, folgen Sie dem Tutorial, das im AWS Secrets Manager Benutzerhandbuch unter [Create an AWS Secrets Manager Secret](#) verfügbar ist. Nachdem Sie das Geheimnis erstellt haben, behalten Sie den geheimen Namen `secretName` für den nächsten Schritt bei.
 - Wenn Sie Schlüssel/Wert-Paare auswählen, erstellen Sie ein Paar für `snowflakeUser` mit dem Schlüssel. `sfUser`

- Erstellen Sie bei der Auswahl von Schlüssel/Wert-Paaren ein Paar für Ihren privaten Snowflake-PEM-Schlüssel mit dem Schlüssel. `pem_private_key`
- Erstellen Sie bei der Auswahl von Schlüssel/Wert-Paaren ein Paar für mit dem Schlüssel. `snowflakeWarehouse sfWarehouse`

Dies ist nicht erforderlich, wenn in Snowflake ein Standard festgelegt ist.

- Erstellen Sie bei der Auswahl von Schlüssel/Wert-Paaren ein Paar für `snowflakeRole` mit dem Schlüssel. `sfRole`

Schritt 4: Nächste Schritte

Nachdem Sie Ihre Datentabellen in Snowflake vorbereitet haben, können Sie:

- [Erstellen Sie eine konfigurierte Tabelle](#)
- [Erstellen Sie ein ML-Modell](#)

Die Tabellen können abgefragt werden nach:

- Der Ersteller der Kollaboration hat eine Kollaboration in eingerichtet. AWS Clean Rooms Weitere Informationen finden Sie unter [Eine Zusammenarbeit erstellen](#).
- Der Ersteller der Kollaboration hat die Kollaborations-ID an Sie als Teilnehmer der Kollaboration gesendet.

Vorbereiten verschlüsselter Datentabellen mit Cryptographic Computing für Clean Rooms

Cryptographic Computing for Clean Rooms (C3R) ist eine Fähigkeit in. AWS Clean Rooms Sie können C3R verwenden, um kryptografisch einzuschränken, was von jeder Partei und in einer Zusammenarbeit gelernt werden kann. AWS AWS Clean Rooms

Sie können die Datentabelle mit dem C3R-Verschlüsselungscient, einem clientseitigen Verschlüsselungstool, verschlüsseln, bevor Sie die Datentabelle in Ihre Datenquelle hochladen: Amazon Simple Storage Service (Amazon S3), Amazon Athena oder Snowflake.

Weitere Informationen finden Sie unter [Kryptografisches Rechnen für Clean Rooms](#).

Die Vorbereitung verschlüsselter Datentabellen mit C3R umfasst die folgenden Schritte:

Schritte

- [Schritt 1: Erfüllen der Voraussetzungen](#)
- [Schritt 2: Laden Sie den C3R-Verschlüsselungssclient herunter](#)
- [Schritt 3: \(Optional\) Verfügbare Befehle im C3R-Verschlüsselungssclient anzeigen](#)
- [Schritt 4: Generieren Sie ein Verschlüsselungsschema für eine tabellarische Datei](#)
- [Schritt 5: Erstellen Sie einen gemeinsamen geheimen Schlüssel](#)
- [Schritt 6: Speichern Sie den gemeinsamen geheimen Schlüssel in einer Umgebungsvariablen](#)
- [Schritt 7: Daten verschlüsseln](#)
- [Schritt 8: Überprüfen Sie die Datenverschlüsselung](#)
- [\(Optional\) Erstellen Sie ein Schema \(fortgeschrittene Benutzer\)](#)

Schritt 1: Erfüllen der Voraussetzungen

Um Ihre Datentabellen für die Verwendung mit C3R vorzubereiten, müssen Sie die folgenden Voraussetzungen erfüllen:

- Sie können auf das Clean Rooms Repository Cryptographic Computing for zugreifen unter: GitHub
<https://github.com/aws/c3r>
- Sie haben AWS Anmeldeinformationen für die Verwendung des C3R-Verschlüsselungssclients eingerichtet. Diese Anmeldeinformationen werden vom C3R-Verschlüsselungssclient für schreibgeschützte API-Aufrufe zum Abrufen von Metadaten für die Zusammenarbeit AWS Clean Rooms verwendet. Weitere Informationen finden Sie unter [Konfiguration von AWS CLI im AWS Command Line Interface](#) Benutzerhandbuch für Version 2.
- Sie haben Java Runtime Environment (JRE) 11 oder höher auf Ihrem Computer installiert.
 - [Die empfohlene Version Java Runtime Environment, Amazon Corretto 11 oder höher, kann von <https://aws.amazon.com/corretto> heruntergeladen werden.](#)
 - Java Development Kit (JDK) enthält eine entsprechende JRE Version derselben Version. Die zusätzlichen Funktionen von JDK werden jedoch nicht benötigt, um den Verschlüsselungssclient Cryptographic Computing for Clean Rooms (C3R) auszuführen.
- Ihre tabellarischen Datendateien (.csv) oder Dateien (. Parquet parquet) werden lokal gespeichert.
- Sie oder ein anderes Mitglied der Kollaboration haben die Möglichkeit, einen gemeinsamen geheimen Schlüssel zu erstellen. Weitere Informationen finden Sie unter [Schritt 5: Erstellen Sie einen gemeinsamen geheimen Schlüssel](#).

- Der Ersteller der Kollaboration hat eine Kollaboration erstellt, in der AWS Clean Rooms Cryptographic Computing für die Zusammenarbeit aktiviert ist. Weitere Informationen finden Sie unter [Eine Zusammenarbeit erstellen](#).
- Der Kollaborationsersteller hat die Kollaborations-ID an Sie als Teilnehmer der Kollaboration gesendet. Der Amazon Resource Name (ARN) der Kollaboration ist in der gesendeten Einladung enthalten, die die Kollaborations-ID enthält.

Schritt 2: Laden Sie den C3R-Verschlüsselungsclient herunter

Um den C3R-Verschlüsselungsclient herunterzuladen von GitHub

1. [Gehen Sie zum Clean RoomsAWSGitHub Repository Cryptographic Computing for: c3r https://github.com/aws/](https://github.com/aws/CleanRoomsAWSGitHubRepositoryCryptographicComputingfor:c3r)
2. Wählen Sie die Dateien aus und laden Sie sie herunter.

Der Quellcode, die Lizenzen und das zugehörige Material können geklont oder als heruntergeladen werden. zipDatei von der Landingpage des GitHub Repositorys. (Siehe die Code-Schaltfläche oben rechts in der Inhaltsliste des Repositorys).

Der zuletzt signierte C3R-Verschlüsselungsclient Java Executable File (d. h. die Anwendung mit der Befehlszeilenschnittstelle) befindet sich auf der Releases-Seite des GitHub Repositorys.

Das C3R-Verschlüsselungsclientpaket für Apache Spark (`c3r-cli-spark`) ist eine Version der `c3r-cli`, die als Job an einen laufenden Apache Spark-Server gesendet werden muss. Weitere Informationen finden Sie unter [C3R auf Apache Spark ausführen](#).

Schritt 3: (Optional) Verfügbare Befehle im C3R-Verschlüsselungsclient anzeigen

Gehen Sie wie folgt vor, um sich mit den verfügbaren Befehlen im C3R-Verschlüsselungsclient vertraut zu machen.

Um alle verfügbaren Befehle im C3R-Verschlüsselungsclient anzuzeigen

1. Navigieren Sie über eine Befehlszeilenschnittstelle (CLI) zu dem Ordner, der die heruntergeladene `c3r-cli.jar` Datei enthält.
2. Führen Sie den folgenden Befehl aus: `java -jar c3r-cli.jar`

3. Sehen Sie sich die Liste der verfügbaren Befehle und Optionen an.

Schritt 4: Generieren Sie ein Verschlüsselungsschema für eine tabellarische Datei

Um Daten zu verschlüsseln, ist ein Verschlüsselungsschema erforderlich, das beschreibt, wie die Daten verwendet werden. In diesem Abschnitt wird beschrieben, wie der C3R-Verschlüsselungsclient bei der Generierung eines Verschlüsselungsschemas für eine CSV-Datei mit einer Kopfzeile oder einer Parquet Datei hilft.

Sie müssen dies nur einmal pro Datei tun. Sobald das Schema existiert, kann es wiederverwendet werden, um dieselbe Datei (oder eine beliebige Datei mit identischen Spaltennamen) zu verschlüsseln. Wenn sich die Spaltennamen oder das gewünschte Verschlüsselungsschema ändern, müssen Sie die Schemadatei aktualisieren. Weitere Informationen finden Sie unter [\(Optional\) Erstellen Sie ein Schema \(fortgeschrittene Benutzer\)](#).

Important

Es ist von größter Bedeutung, dass alle beteiligten Parteien denselben gemeinsamen geheimen Schlüssel verwenden. Die beteiligten Parteien sollten auch die Spaltennamen so koordinieren, dass sie übereinstimmen, ob sie bei Abfragen JOIN bearbeitet oder auf andere Weise auf Gleichheit verglichen werden. Andernfalls könnten die SQL-Abfragen zu unerwarteten oder falschen Ergebnissen führen. Dies ist jedoch nicht erforderlich, wenn der Kollaborationsersteller die `allowJoinsOnColumnsWithDifferentNames` Verschlüsselungseinstellung bei der Erstellung der Kollaboration aktiviert hat. Weitere Informationen zu verschlüsselungsrelevanten Einstellungen finden Sie unter [Kryptografische Rechenparameter](#)

Wenn der C3R-Verschlüsselungsclient im Schemamodus ausgeführt wird, durchsucht er die Eingabedatei spaltenweise und fragt Sie, ob und wie diese Spalte behandelt werden soll. Wenn die Datei viele Spalten enthält, die für die verschlüsselte Ausgabe nicht benötigt werden, kann die interaktive Schemagenerierung mühsam werden, da Sie jede unerwünschte Spalte überspringen müssen. Um dies zu vermeiden, könnten Sie manuell ein Schema schreiben oder eine vereinfachte Version der Eingabedatei erstellen, die nur die gewünschten Spalten enthält. Dann könnte der interaktive Schema-Generator für diese reduzierte Datei ausgeführt werden. Der

C3R-Verschlüsselungsclient gibt Informationen über die Schemadatei aus und fragt Sie, wie die Quellspalten (wenn überhaupt) in der Zielausgabe enthalten oder verschlüsselt werden sollen.

Für jede Quellspalte in der Eingabedatei werden Sie aufgefordert, Folgendes einzugeben:

1. Wie viele Zielspalten sollen generiert werden
2. Wie soll jede Zielspalte verschlüsselt werden (wenn überhaupt)
3. Der Name jeder Zielspalte
4. Wie Daten vor der Verschlüsselung aufgefüllt werden sollen, wenn die Spalte als sealed Spalte verschlüsselt wird

 Note

Wenn Sie Daten für eine Spalte verschlüsseln, die als Spalte verschlüsselt wurde, müssen Sie festlegen, welche Daten aufgefüllt werden müssen. sealed Der C3R-Verschlüsselungsclient schlägt bei der Schemagenerierung ein Standard-Padding vor, bei dem alle Einträge in einer Spalte auf dieselbe Länge aufgefüllt werden. Beachten Sie bei der Bestimmung der Länge für `fixed`, dass das Auffüllen in Byte und nicht in Bits erfolgt.

Im Folgenden finden Sie eine Entscheidungstabelle für die Erstellung des Schemas.

Schema-Entscheidungstabelle

Entscheidung	Anzahl der Zielspalten aus der Quellspalte <' name-of-column '>?	Zielspaltenentyp: [c]cleartext, [f] fingerprint oder [s]? sealed	Headername der Zielspalte <default 'name-of-column'>	Fügen Sie der <suffix>Kopfzeile ein Suffix hinzu, um anzugeben, wie sie verschlüsselt wurde, [y] ja oder [n] nein <default 'yes'>	<' name-of-column _sealed'> Polstertyp: [n] eins, [f] fest oder [m] max <default 'max'>
Lassen Sie die Spalte unverschlüsselt.	1	c	Nicht zutreffend	Nicht zutreffend	Nicht zutreffend
Verschlüsseln Sie die Spalte als fingerprint Spalte.	1	f	Wählen Sie Standard oder geben Sie einen neuen Header-Namen ein.	Geben Sie ein, um Standard (_fingerprint) zu wählen, oder geben Sie die Eingabetaste ein.	Nicht zutreffend
Verschlüsseln Sie die Spalte als sealed Spalte.	1	s	Wählen Sie Standard oder geben Sie einen neuen Header-Namen ein.	Geben Sie ein, um Standard (_sealed) zu wählen, oder geben Sie die Eingabetaste ein.	Wählen Sie den Polstertyp. Weitere Informationen finden Sie unter

Entscheidung	Anzahl der Zielspalten aus der Quellspalte <' name-of-column '>?	Zielspaltenentyp: [c]cleartext, [f] fingerprint oder [s]? sealed	Headername der Zielspalte <default 'name-of-column'>	Fügen Sie der <suffix>Kopfzeile ein Suffix hinzu, um anzugeben, wie sie verschlüsselt wurde, [y] ja oder [n] nein <default 'yes'>	<' name-of-column _sealed'> Polstertyp: [n] eins, [f] fest oder [m] max <default 'max'>
					(Optional) Erstellen Sie ein Schema (fortgeschrittene Benutzer).
Verschlüsseln Sie die Spalte sowohl als auch fingerprint. sealed	2	Geben Sie die erste Zielspalte ein: f. Geben Sie die zweite Zielspalte ein: s.	Wählen Sie die Zielüberschriften für jede Zielspalte aus.	Geben Sie ein, um Standard zu wählen, oder geben Sie ein n.	Wählen Sie den Fülltyp (nur für sealed Spalten). Weitere Informationen finden Sie unter (Optional) Erstellen Sie ein Schema (fortgeschrittene Benutzer).

Im Folgenden finden Sie zwei Beispiele für die Erstellung von Verschlüsselungsschemas. Der genaue Inhalt Ihrer Interaktion hängt von der Eingabedatei und den Antworten ab, die Sie geben.

Beispiele

- [Beispiel: Generieren Sie ein Verschlüsselungsschema für eine fingerprint Spalte und eine cleartext Spalte](#)
- [Beispiel: Generieren Sie ein Verschlüsselungsschema mit sealed Spaltenfingerprint, und cleartext](#)

Beispiel: Generieren Sie ein Verschlüsselungsschema für eine fingerprint Spalte und eine cleartext Spalte

In diesem Beispiel gibt `ads.csv` es für nur zwei Spalten: `username` und `ad_variant`. Für diese Spalten wollen wir Folgendes:

- Damit die `username` Spalte als `fingerprint` Spalte verschlüsselt wird
- Damit die `ad_variant` Spalte eine `cleartext` Spalte ist

Um ein Verschlüsselungsschema für eine `fingerprint` Spalte und eine `cleartext` Spalte zu generieren

1. (Optional) Um sicherzustellen, dass die `c3r-cli.jar` Datei und die zu verschlüsselnde Datei vorhanden sind:
 - a. Navigieren Sie zum gewünschten Verzeichnis und führen Sie es aus `ls` (falls Sie ein Mac oder Unix/verwenden Linux) oder `dir` wenn Sie Windows verwenden.
 - b. Sehen Sie sich die Liste der tabellarischen Datendateien an (z. B. CSV) und wählen Sie eine zu verschlüsselnde Datei aus.

In diesem Beispiel `ads.csv` ist das die Datei, die wir verschlüsseln möchten.

2. Führen Sie in der CLI den folgenden Befehl aus, um interaktiv ein Schema zu erstellen.

```
java -jar c3r-cli.jar schema ads.csv --interactive --output=ads.json
```

Note

- Sie können ausführen `java --jar PATH/T0/c3r-cli.jar`. Oder, wenn Sie `PATH/T0/c3r-cli.jar` zu Ihrer `CLASSPATH`-Umgebungsvariablen etwas hinzugefügt haben, können Sie auch den Klassennamen ausführen.

Der C3R-Verschlüsselungsclient sucht im CLASSPATH danach (z. B.). `java com.amazon.psion.cli.Main`

- Das `--interactive` Flag wählt den interaktiven Modus für die Entwicklung des Schemas aus. Dadurch wird der Benutzer durch einen Assistenten zum Erstellen des Schemas geführt. Benutzer mit fortgeschrittenen Kenntnissen können ihr eigenes Schema-JSON erstellen, ohne den Assistenten zu verwenden. Weitere Informationen finden Sie unter [\(Optional\) Erstellen Sie ein Schema \(fortgeschrittene Benutzer\)](#).
- Das `--output` Flag legt einen Ausgabenamen fest. Wenn Sie das `--output` Flag nicht angeben, versucht der C3R-Verschlüsselungsclient, einen Standardausgabennamen zu wählen (z. B. `<input>.out.csv` oder für das Schema `<input>.json`).

3. Geben Sie für Number of target columns from source column 'username'? ein **1** und drücken Sie dann die Eingabetaste.
4. Geben Sie für Target column type: [c]leartext, [f]ingerprint, or [s]ealed? ein **f** und drücken Sie dann die Eingabetaste.
5. Drücken Sie für Target column headername <default 'username'> die Eingabetaste.

Der Standardname 'username' wird verwendet.

6. Geben Sie für Add suffix '_fingerprint' to header to indicate how it was encrypted, [y]es or [n]o <default 'yes'> ein **y** und drücken Sie dann die Eingabetaste.

Note

Der interaktive Modus schlägt Suffixe vor, die zu den verschlüsselten Spaltenüberschriften hinzugefügt werden sollen (`_fingerprint` für fingerprint Spalten und `_sealed` für sealed Spalten). Die Suffixe können hilfreich sein, wenn Sie Aufgaben wie das Hochladen von Daten in Kollaborationen oder das Erstellen von Kollaborationen ausführen. AWS-Services AWS Clean Rooms Anhand dieser Suffixe kann angegeben werden, was mit den verschlüsselten Daten in den einzelnen Spalten geschehen kann. Zum Beispiel funktionieren Dinge nicht, wenn Sie eine Spalte als sealed Spalte (`_sealed`) verschlüsseln und versuchen, dies zu JOIN tun, oder wenn Sie es umgekehrt versuchen.

7. Geben Sie für Number of target columns from source column 'ad_variant'? die Eingabe ein **1** und drücken Sie dann die Eingabetaste.
8. Geben Sie für Target column type: [c]leartext, [f]ingerprint, or [s]ealed? ein **c** und drücken Sie dann die Eingabetaste.
9. Drücken Sie für Target column headername <default 'username'> die Eingabetaste.

Der Standardname 'ad_variant' wird verwendet.

Das Schema wird in eine neue Datei mit dem Namen geschriebenads.json.

 Note

Sie können das Schema anzeigen, indem Sie es in einem beliebigen Texteditor öffnen, z. B. Notepad on Windows oder TextEdit on macOS.

10. Sie sind jetzt bereit, [Daten zu verschlüsseln](#).

Beispiel: Generieren Sie ein Verschlüsselungsschema mit sealed Spaltenfingerprint, und cleartext

In diesem Beispiel gibt sales.csv es für drei Spalten: usernamepurchased, undproduct. Für diese Spalten wollen wir Folgendes:

- Damit die product Spalte eine sealed Spalte ist
- Damit die username Spalte als fingerprint Spalte verschlüsselt wird
- Damit die purchased Spalte eine cleartext Spalte ist

Um ein Verschlüsselungsschema mit sealedfingerprint, und cleartext Spalten zu generieren

1. (Optional) Um sicherzustellen, dass die c3r-cli.jar Datei und die zu verschlüsselnde Datei vorhanden sind:
 - a. Navigieren Sie zum gewünschten Verzeichnis und führen Sie es aus ls (falls Sie ein Mac oder Unix/verwenden Linux) oder dir wenn Sie Windows verwenden.
 - b. Sehen Sie sich die Liste der tabellarischen Datendateien (.csv) an und wählen Sie eine zu verschlüsselnde Datei aus.

In diesem Beispiel `sales.csv` ist das die Datei, die wir verschlüsseln möchten.

2. Führen Sie in der CLI den folgenden Befehl aus, um interaktiv ein Schema zu erstellen.

```
java -jar c3r-cli.jar schema sales.csv --interactive --  
output=sales.json
```

 Note

- Das `--interactive` Flag wählt den interaktiven Modus für die Entwicklung des Schemas aus. Dadurch wird der Benutzer durch einen geführten Arbeitsablauf zur Erstellung des Schemas geführt.
- Wenn Sie ein erfahrener Benutzer sind, können Sie Ihr eigenes JSON-Schema erstellen, ohne den geführten Workflow zu verwenden. Weitere Informationen finden Sie unter [\(Optional\) Erstellen Sie ein Schema \(fortgeschrittene Benutzer\)](#).
- Informationen zu CSV-Dateien ohne Spaltenüberschriften finden Sie im `--noHeaders` Flag für den Schemabefehl, der in der CLI verfügbar ist.
- Das `--output` Flag legt einen Ausgabenamen fest. Wenn Sie das `--output` Flag nicht angeben, versucht der C3R-Verschlüsselungssclient, einen Standardausgabenenamen zu wählen (z. B. `<input>.out` oder für das Schema `<input>.json`).

3. Geben Sie für `Number of target columns from source column 'username'?` ein **1** und drücken Sie dann die Eingabetaste.
4. Geben Sie für `Target column type: [c]leartext, [f]ingerprint, or [s]ealed?` ein **f** und drücken Sie dann die Eingabetaste.
5. Drücken Sie für `Target column headername <default 'username'>` die Eingabetaste.

Der Standardname 'username' wird verwendet.

6. Geben Sie für `Add suffix '_fingerprint' to header to indicate how it was encrypted, [y]es or [n]o <default 'yes'>` ein **y** und drücken Sie dann die Eingabetaste.
7. Geben Sie für `Number of target columns from source column 'purchased'?` ein **1** und drücken Sie dann die Eingabetaste.
8. Geben Sie für `Target column type: [c]leartext, [f]ingerprint, or [s]ealed?` ein **c** und drücken Sie dann die Eingabetaste.

- Drücken Sie für Target column headername <default 'purchased'> die Eingabetaste.

Der Standardname 'purchased' wird verwendet.

- Geben Sie für Number of target columns from source column 'product'? ein **1** und drücken Sie dann die Eingabetaste.
- Geben Sie für Target column type: [c]leartext, [f]ingerprint, or [s]ealed? ein **s** und drücken Sie dann die Eingabetaste.
- Drücken Sie für Target column headername <default 'product'> die Eingabetaste.

Der Standardname 'product' wird verwendet.

- Drücken Sie für die Eingabetaste **'product_sealed' padding type: [n]one, [f]ixed, or [m]ax <default 'max'>**, um die Standardeinstellung auszuwählen.
- Byte-length beyond max length to pad cleartext to in 'product_sealed' <default '0'>? Drücken Sie die Eingabetaste, um die Standardeinstellung auszuwählen.

Das Schema wird in eine neue Datei namens geschriebensales.json.

- Sie sind jetzt bereit, [Daten zu verschlüsseln](#).

Schritt 5: Erstellen Sie einen gemeinsamen geheimen Schlüssel

Um die Datentabellen zu verschlüsseln, müssen sich die Teilnehmer der Zusammenarbeit auf einen gemeinsamen geheimen Schlüssel einigen und diesen auf sichere Weise gemeinsam nutzen.

Der gemeinsame geheime Schlüssel muss mindestens 256 Bit (32 Byte) lang sein. Sie können einen größeren Schlüssel angeben, der Ihnen jedoch keine zusätzliche Sicherheit bietet.

Important

Denken Sie daran, dass der Schlüssel und die Kollaborations-ID, die für die Verschlüsselung und Entschlüsselung verwendet werden, für alle Kollaborationsteilnehmer identisch sein müssen.

Die folgenden Abschnitte enthalten Beispiele für Konsolenbefehle zum Generieren eines gemeinsamen geheimen Schlüssels, der secret.key im aktuellen Arbeitsverzeichnis des jeweiligen Terminals gespeichert wird.

Themen

- [Beispiel: Schlüsselgenerierung mit OpenSSL](#)
- [Beispiel: Schlüsselgenerierung bei der Windows Verwendung PowerShell](#)

Beispiel: Schlüsselgenerierung mit OpenSSL

Führen Sie für eine allgemeine Kryptografiebibliothek den folgenden Befehl aus, um einen gemeinsamen geheimen Schlüssel zu erstellen.

```
openssl rand 32 > secret.key
```

Wenn Sie sie verwenden Windows und noch nicht OpenSSL installiert haben, können Sie Schlüssel anhand des Beispiels generieren, das unter [Beispiel: Schlüsselgenerierung bei Windows Verwendung PowerShell](#) beschrieben ist.

Beispiel: Schlüsselgenerierung bei der Windows Verwendung PowerShell

Führen Sie für eine Terminalanwendung PowerShell, die auf verfügbar ist Windows, den folgenden Befehl aus, um einen gemeinsamen geheimen Schlüssel zu erstellen.

```
$bs = New-Object Byte[](32);  
[Security.Cryptography.RandomNumberGenerator]::Create().GetBytes($bs); Set-  
Content 'secret.key' -Encoding Byte -Value $bs
```

Schritt 6: Speichern Sie den gemeinsamen geheimen Schlüssel in einer Umgebungsvariablen

Eine Umgebungsvariable ist eine bequeme und erweiterbare Möglichkeit für Benutzer, einen geheimen Schlüssel aus verschiedenen Schlüsselspeichern bereitzustellen AWS Secrets Manager und ihn an den C3R-Verschlüsselungsclient weiterzuleiten.

Der C3R-Verschlüsselungsclient kann Schlüssel verwenden, die in gespeichert sind, AWS-Services wenn Sie die verwenden AWS CLI , um diese Schlüssel in der entsprechenden Umgebungsvariablen zu speichern. Der C3R-Verschlüsselungsclient kann beispielsweise einen Schlüssel von verwenden. AWS Secrets Manager Weitere Informationen finden Sie unter [Geheimnisse erstellen und verwalten mit AWS Secrets Manager](#) im AWS Secrets Manager Benutzerhandbuch.

 Note

Bevor Sie jedoch ein AWS-Service solches als AWS Secrets Manager Aufbewahrung Ihrer C3R-Schlüssel verwenden, stellen Sie sicher, dass Ihr Anwendungsfall dies zulässt. In bestimmten Anwendungsfällen kann es erforderlich sein, dass Ihnen der Schlüssel vorenthalten wird. AWS Dadurch soll sichergestellt werden, dass die verschlüsselten Daten und der Schlüssel niemals von derselben dritten Partei verwaltet werden.

Die einzigen Voraussetzungen für einen gemeinsamen geheimen Schlüssel sind, dass der gemeinsame geheime Schlüssel base64 -kodierte und in der Umgebungsvariablen gespeichert ist. C3R_SHARED_SECRET

In den folgenden Abschnitten werden die Konsolenbefehle zum Konvertieren einer `secret.key` Datei in eine Umgebungsvariable base64 und zum Speichern dieser Datei als Umgebungsvariable beschrieben. Die `secret.key` Datei könnte aus jedem der unter aufgeführten Befehle generiert worden sein [Schritt 5: Erstellen Sie einen gemeinsamen geheimen Schlüssel](#) und ist nur eine Beispielquelle.

Speichern Sie den Schlüssel in einer Umgebungsvariablen, wenn Windows Sie PowerShell

Führen Sie den folgenden Befehl ausPowerShell, um bei der Windows Verwendung in die Umgebungsvariable zu konvertieren base64 und sie festzulegen.

```
$Bytes=[IO.File]::ReadAllBytes((Get-Location).ToString()+'\secret.key');  
$env:C3R_SHARED_SECRET=[Convert]::ToBase64String($Bytes)
```

Speichern Sie den Schlüssel in einer Umgebungsvariablen auf Linux oder macOS

Führen Sie den folgenden Befehl ausmacOS, um in Linux oder zu konvertieren base64 und die Umgebungsvariable auf oder zu setzen.

```
export C3R_SHARED_SECRET="$(cat secret.key | base64)"
```

Schritt 7: Daten verschlüsseln

Um diesen Schritt auszuführen, müssen Sie die AWS Clean Rooms Kollaborations-ID und den gemeinsamen geheimen Schlüssel erwerben. Weitere Informationen finden Sie unter [Voraussetzungen](#).

Im folgenden Beispiel führen wir die Verschlüsselung unter `ads.csv` Verwendung des von uns erstellten Schemas namens `ausads.json`.

So verschlüsseln Sie Daten

1. Speichern Sie den gemeinsamen geheimen Schlüssel für die Zusammenarbeit in [Schritt 6: Speichern Sie den gemeinsamen geheimen Schlüssel in einer Umgebungsvariablen](#).
2. Geben Sie in der Befehlszeile den folgenden Befehl ein.

```
java -jar c3r-cli.jar encrypt <name of input .csv file> --schema=<name of schema .json file> --id=<collaboration id> --output=<name of output.csv file> <optional flags>
```

3. Geben Sie für *<name of input .csv file>* den Namen der CSV-Eingabedatei ein.
4. Geben Sie für `schema=` den Namen der JSON-Verschlüsselungsschemadatei ein.
5. Geben Sie für `id=` die Kollaborations-ID ein.
6. Geben Sie für `output=` den Namen der Ausgabedatei ein (z. B. `ads-output.csv`).
7. Fügen Sie alle Befehlszeilen-Flags ein, die unter [Kryptografische Rechenparameter](#) und beschrieben sind [Optionale Flags in Cryptographic Computing für Clean Rooms](#).
8. Führen Sie den Befehl aus.

Im Beispiel für `ads.csv` führen wir den folgenden Befehl aus.

```
java -jar c3r-cli.jar encrypt ads.csv --schema=ads.json --id=123e4567-e89b-42d3-a456-556642440000 --output=ads-output.csv
```

Im Beispiel für `sales.csv` führen wir den folgenden Befehl aus.

```
java -jar c3r-cli.jar encrypt sales.csv --schema=sales.json --id=123e4567-e89b-42d3-a456-556642440000
```

 Note

In diesem Beispiel geben wir keinen Namen für die Ausgabedatei (`--output=sales-output.csv`) an. Infolgedessen `name-of-file.out.csv` wurde der Standardname der Ausgabedatei generiert.

Sie sind jetzt bereit, die verschlüsselten Daten zu überprüfen.

Schritt 8: Überprüfen Sie die Datenverschlüsselung

Um zu überprüfen, ob die Daten verschlüsselt wurden

1. Zeigen Sie die verschlüsselte Datendatei an (z. B. `sales-output.csv`).
2. Überprüfen Sie die folgenden Spalten:
 - a. Spalte 1 — Verschlüsselt (z. B. `username_fingerprint`).

Für die fingerprint Spalten (HMAC) gibt es nach dem Versions- und Typpräfix (z. B. `01:hmac:`) 44 Zeichen mit Base64-codierten Daten.

- b. Spalte 2 — Nicht verschlüsselt (zum Beispiel). `purchased`
- c. Spalte 3 — Verschlüsselt (zum Beispiel `product_sealed`).

Bei verschlüsselten (SELECT) Spalten ist die Länge der Spalte cleartext plus jegliches Auffüllen nach dem Versions- und Typpräfix (z. B. `01:enc:`) direkt proportional zur Länge der cleartext verschlüsselten Spalte. Das heißt, die Länge entspricht der Größe der Eingabe plus etwa 33 Prozent Mehraufwand aufgrund der Kodierung.

Sie sind jetzt bereit für:

1. [Laden Sie die verschlüsselten Daten auf S3 hoch](#).
2. [Erstellen Sie eine AWS Glue Tabelle](#).
3. [Erstellen Sie eine konfigurierte Tabelle in AWS Clean Rooms](#).

Der C3R-Verschlüsselungsclient erstellt temporäre Dateien, die keine unverschlüsselten Daten enthalten (es sei denn, diese Daten würden auch in der endgültigen Ausgabe unverschlüsselt sein). Einige verschlüsselte Werte werden jedoch möglicherweise nicht richtig aufgefüllt.

Fingerabdruckspalten können doppelte Werte enthalten, auch wenn die Einstellung für die Zusammenarbeit `allowRepeatedFingerprintValue` aktiviert ist `false`. Dieses Problem tritt auf, weil die temporäre Datei geschrieben wurde, bevor die richtigen Fülllängen und Eigenschaften zum Entfernen von Duplikaten überprüft wurden.

Wenn der C3R-Verschlüsselungsclient ausfällt oder während der Verschlüsselung unterbrochen wird, stoppt er möglicherweise, nachdem die temporäre Datei geschrieben wurde, aber bevor diese Eigenschaften überprüft und die temporären Dateien gelöscht wurden. Daher befinden sich diese temporären Dateien möglicherweise immer noch auf der Festplatte. Wenn dies der Fall ist, schützt der Inhalt dieser Dateien die Klartextdaten nicht auf demselben Niveau wie die Ausgabe. Insbesondere könnten diese temporären Dateien Klartextdaten für statistische Analysen preisgeben, die sich nicht auf die endgültige Ausgabe auswirken würden. Der Benutzer sollte diese Dateien (insbesondere eine SQLite Datenbank) löschen, um zu verhindern, dass diese Dateien in unbefugte Hände gelangen.

(Optional) Erstellen Sie ein Schema (fortgeschrittene Benutzer)

Das manuelle Erstellen eines Schemas ist für fortgeschrittene Benutzer vorgesehen.

Im Folgenden finden Sie eine Beschreibung des JSON-Schemadateiformats für Eingabedateien mit oder ohne Spaltenüberschriften. Fortgeschrittene Benutzer können das Schema bei Bedarf direkt schreiben oder ändern.

Note

Der C3R-Verschlüsselungsclient kann Sie bei der Erstellung eines Schemas entweder durch den unter beschriebenen interaktiven Prozess [Beispiel: Generieren Sie ein Verschlüsselungsschema mit sealed Spaltenfingerprint, und cleartext](#) oder durch die Erstellung einer Stub-Vorlage unterstützen.

Schemas für zugeordnete und positionierte Tabellen

Im folgenden Abschnitt werden zwei Arten von Tabellenschemas beschrieben:

- **Zugeordnetes Tabellenschema** — Dieses Schema wird für die Verschlüsselung von CSV-Dateien mit einer Kopfzeile und Dateien verwendet. Apache Parquet
- **Positionstabellenschema** — Dieses Schema wird zum Verschlüsseln von CSV-Dateien ohne Kopfzeile verwendet.

Der C3R-Verschlüsselungsclient kann eine tabellarische Datei für eine Zusammenarbeit verschlüsseln. Dazu muss er über eine entsprechende Schemadatei verfügen, die angibt, wie die verschlüsselte Ausgabe aus der Eingabe abgeleitet werden soll.

Der C3R-Verschlüsselungsclient kann helfen, ein Schema für eine INPUT Datei zu generieren, indem er den Befehl `C3R-Verschlüsselungsclient Schema` in der Befehlszeile ausführt. Ein Beispiel für einen Befehl ist `java -jar c3r-cli.jar schema --interactive INPUT`

Das Schema spezifiziert die folgenden Informationen:

1. Welche Quellspalten werden anhand ihrer Header-Namen (zugeordnete Schemas) oder ihrer Position (Positionsschemas) welchen transformierten Spalten in der Ausgabedatei zugeordnet
2. Welche Zielspalten sollen erhalten bleiben cleartext
3. Welche Zielspalten sollen für SELECT Abfragen verschlüsselt werden
4. Welche Zielspalten sollen für JOIN Abfragen verschlüsselt werden

Diese Informationen sind in einer tabellenspezifischen JSON-Schemadatei kodiert, die aus einem einzigen Objekt besteht, dessen `headerRow` Feld ein boolescher Wert ist. Der Wert muss `true` für Parquet Dateien und CSV-Dateien mit einer Kopfzeile gelten, andernfalls `false`

Zugeordnetes Tabellenschema

Das zugeordnete Schema hat die folgende Form.

```
{
  "headerRow": true,
  "columns": [
    {
      "sourceHeader": STRING,
      "targetHeader": STRING,
      "type": TYPE,
      "pad": PAD
    },
    ...
  ]
}
```

Falls `headerRow` `true` ist, ist das nächste Feld im Objekt `columns`, das eine Reihe von Spaltenschemas enthält, die Quellkopfzeilen Zielüberschriften zuordnen (d. h. JSON-Objekte, die beschreiben, was die Ausgabespalten enthalten sollen).

- **sourceHeader**— Der STRING Header-Name der Quellspalte, aus der die Daten abgeleitet wurden.

 Note

Dieselbe Quellspalte kann für mehrere Zielspalten verwendet werden.
Eine Spalte aus der Eingabedatei, die nicht **sourceHeader** irgendwo im Schema aufgeführt ist, erscheint nicht in der Ausgabedatei.

- **targetHeader**— Der STRING Header-Name der entsprechenden Spalte in der Ausgabedatei.

 Note

Dieses Feld ist für zugeordnete Schemas optional. Wenn dieses Feld weggelassen wird, **sourceHeader** wird das für den Header-Namen in der Ausgabe wiederverwendet.
Entweder **_fingerprint** oder **_sealed** wird angehängt, wenn es sich bei der Ausgabespalte um eine fingerprint Spalte bzw. sealed Spalte handelt.

- **type**— Die TYPE der Zielspalte in der Ausgabedatei. Das heißt, eine von **cleartextsealed**, oder **fingerprint** hängt davon ab, wie die Spalte in der Kollaboration verwendet wird.
- **pad**— Ein Feld eines Spaltenschemaobjekts, das nur vorhanden ist, wenn es vorhanden TYPE **istsealed**. Sein entsprechender Wert von PAD ist ein Objekt, das beschreibt, wie die Daten aufgefüllt werden sollen, bevor sie verschlüsselt werden.

```
{
  "type": PAD_TYPE,
  "length": INT
}
```

Sie geben die Auffüllung vor der Verschlüsselung an **type** und **length** werden wie folgt verwendet:

- **PAD_TYPE**as **none** — Auf die Daten der Spalte wird kein Auffüllen angewendet und das **length** Feld ist nicht zutreffend (d. h. es wird weggelassen).
- **PAD_TYPE**as **fixed** — Die Daten der Spalte werden auf die angegebene Anzahl **length** von Byte aufgefüllt.
- **PAD_TYPE**as **max** — Die Daten der Spalte werden auf die Größe der Bytelänge des längsten Werts zuzüglich weiterer **length** Byte aufgefüllt.

Im Folgenden finden Sie ein Beispiel für ein zugeordnetes Schema mit einer Spalte jedes Typs.

```
{
  "headerRow": true,
  "columns": [
    {
      "sourceHeader": "FullName",
      "targetHeader": "name",
      "type": "cleartext"
    },
    {
      "sourceHeader": "City",
      "targetHeader": "city_sealed",
      "type": "sealed",
      "pad": {
        "type": "max",
        "length": 16
      }
    },
    {
      "sourceHeader": "PhoneNumber",
      "targetHeader": "phone_number_fingerprint",
      "type": "fingerprint"
    },
    {
      "sourceHeader": "PhoneNumber",
      "targetHeader": "phone_number_sealed",
      "type": "sealed",
      "pad": {
        "type": "fixed",
        "length": 20
      }
    }
  ]
}
```

Als komplexeres Beispiel finden Sie im Folgenden eine CSV-Beispieldatei mit Headern.

```
FirstName,LastName,Address,City,State,PhoneNumber,Title,Level,Notes
Jorge,Souza,12345 Mills Rd,Anytown,SC,703-555-1234,CEO,10,
Paulo,Santos,0 Street,Anytown,MD,404-555-111,CIO,9,This is a really long note that
could really be a paragraph
Mateo,Jackson,1 Two St,Anytown,NY,304-555-1324,C00,9,""
```

```
Terry,Whitlock4 N St,Anytown,VA,407-555-8888,EA,7,Secret notes
Diego,Ramirez,9 Hollows Rd,Anytown,VA,407-555-1222,SDE I,4,null
John,Doe,8 Hollows Rd,Anytown,VA,407-555-4321,SDE I,4,Jane's younger brother
Jane,Doe,8 Hollows Rd,Anytown,VA,407-555-4322,SDE II,5,John's older sister
```

Im folgenden Beispiel für ein zugeordnetes Schema sind die Spalten `FirstName` und `LastName` Spalten. `cleartext` Die State Spalte ist als fingerprint Spalte und als sealed Spalte mit einer Auffüllung von verschlüsselt. `none` Die übrigen Spalten werden weggelassen.

```
{
  "headerRow": true,
  "columns": [
    {
      "sourceHeader": "FirstName",
      "targetHeader": "GivenName",
      "type": "cleartext"
    },
    {
      "sourceHeader": "LastName",
      "targetHeader": "Surname",
      "type": "cleartext"
    },
    {
      "sourceHeader": "State",
      "targetHeader": "State_Join",
      "type": "fingerprint"
    },
    {
      "sourceHeader": "State",
      "targetHeader": "State",
      "type": "sealed",
      "pad": {
        "type": "none"
      }
    }
  ]
}
```

Im Folgenden finden Sie die CSV-Datei, die sich aus dem zugewiesenen Schema ergibt.

```
givenname,surname,state_fingerprint,state
```

```
John,Doe,01:hmac:UK8s8Cn/WR2J0/To2dTxD73aDEe2ZUXeSHy3Tv
+1Mk=,01:enc:FQ3n3Ahv9BQQNWQGcugeHzHYzEZE1vapHa2Uu4SRgSAtZ3q0bjPA4TcsHt
+B0kMKBcnHWI13BeGG/SBqmj7vKpI=
Paulo,Santos,01:hmac:CHF4eIrtTNgAooU9v4h9Qjc
+txBnMidQTjdjWuaDTTA=,01:enc:KZ5n5GtaXACco65AXk48BQ02durDNR2ULc4YxmMC8NaZZKKJiksU1IwFadAvV4iBQ1
Mateo,Jackson,01:hmac:iIRnjfNBzryusIJ1w35lgNzeY1RQ1bSfq6PDHW8Xrbk=,01:enc:mLKpS5HI0SgphdEsrzhd
eN9nB02gAbIygt40Fn4La1Yn9Xyj/XUWXlmn8zFe2T4kyDTD8kG0vpQEUGxAUFk=
Diego,Ramirez,01:hmac:UK8s8Cn/WR2J0/To2dTxD73aDEe2ZUXeSHy3Tv+1Mk=,01:enc:rmZhT98Zm
+IIGw1UTjMIJP4IrW/AA1tBLMXcHvnYfRgmWP623VFQ6aUnhsb2MDqEw4G5Uwg5rKKZepUxx5uKbfbk=
Jorge,Souza,01:hmac:3BxJdXiFFyZ8HBbYNqqEhBVqhN0d7s2ZiKUe7QiTy08=,01:enc:vVaqWC1VRbhvkf8gnuR7q0z
Terry,Whitlock01:hmac:UK8s8Cn/WR2J0/To2dTxD73aDEe2ZUXeSHy3Tv+1Mk=,01:enc:3c9VEwb0D0/
xbQjdGuccLvI7oZTBdPU+SyrJIyr2kudfAxbuMQ2uRdU/q7rbgyJjxZS8M2U35ILJf/1DgTyg7cM=
Jane,Doe,01:hmac:UK8s8Cn/WR2J0/To2dTxD73aDEe2ZUXeSHy3Tv+1Mk=,01:enc:9RWv46YLveykeNZ/
G0Nd1YFg+AVd0nu05hHyAYTQkPLHnyX+0/jbzD/g9ZT8GCgVE9aB5bV4ooJIXHGBVMXcjrQ=
```

Positionstabellenschema

Das Positionsschema hat die folgende Form.

```
{
  "headerRow": false,
  "columns": [
    [
      {
        "targetHeader": STRING,
        "type": TYPE,
        "pad": PAD
      },
      {
        "targetHeader": STRING,
        "type": TYPE,
        "pad": PAD
      }
    ],
    [],
    ...
  ]
}
```

Falls `headerRow` `false` ist, ist das nächste Feld im Objekt `columns`, das eine Reihe von Einträgen enthält. Jeder Eintrag ist selbst ein Array von null oder mehr positionellen Spaltenschemas (kein `sourceHeader` Feld). Dabei handelt es sich um JSON-Objekte, die beschreiben, was die Ausgabe enthalten soll.

- **sourceHeader**— Der STRING Header-Name der Quellspalte, aus der die Daten abgeleitet werden.

 Note

Dieses Feld muss in Positionsschemas weggelassen werden. In Positionsschemas wird die Quellspalte aus dem entsprechenden Index der Spalte in der Schemadatei abgeleitet.

- **targetHeader**— Der STRING Header-Name der entsprechenden Spalte in der Ausgabedatei.

 Note

Dieses Feld ist für Positionsschemas erforderlich.

- **type**— Die TYPE der Zielspalte in der Ausgabedatei. Das heißt, eine von `cleartextsealed`, oder `fingerprint` hängt davon ab, wie die Spalte in der Kollaboration verwendet wird.
- **pad**— Ein Feld eines Spaltenschemaobjekts, das nur vorhanden ist, wenn es vorhanden TYPE `istsealed`. Sein entsprechender Wert von PAD ist ein Objekt, das beschreibt, wie die Daten aufgefüllt werden sollen, bevor sie verschlüsselt werden.

```
{
  "type": PAD_TYPE,
  "length": INT
}
```

Sie geben die Auffüllung vor der Verschlüsselung an `type` und `length` werden wie folgt verwendet:

- `PAD_TYPE`as `none` — Auf die Daten der Spalte wird kein Auffüllen angewendet und das `length` Feld ist nicht zutreffend (d. h. es wird weggelassen).
- `PAD_TYPE`as `fixed` — Die Daten der Spalte werden auf die angegebene Anzahl `length` von Byte aufgefüllt.
- `PAD_TYPE`as `max` — Die Daten der Spalte werden auf die Größe der Bytelänge des längsten Werts zuzüglich weiterer `length` Byte aufgefüllt.

 Note

`fixedist` nützlich, wenn Sie im Voraus eine Obergrenze für die Bytegröße der Spaltendaten kennen. Ein Fehler wird ausgelöst, wenn Daten in dieser Spalte länger als angegeben sind `length`.

`maxist` praktisch, wenn die genaue Größe der Eingabedaten unbekannt ist, da es unabhängig von der Größe der Daten funktioniert. `max` erfordert jedoch zusätzliche Verarbeitungszeit, da die Daten zweimal verschlüsselt werden. `max` verschlüsselt die Daten einmal, wenn sie in die temporäre Datei eingelesen werden, und einmal, nachdem der längste Dateneintrag in der Spalte bekannt ist.

Außerdem wird die Länge des längsten Werts zwischen Aufrufen des Clients nicht gespeichert. Wenn Sie planen, Ihre Daten stapelweise oder regelmäßig neue Daten zu verschlüsseln, beachten Sie, dass die daraus resultierenden Chiffretext-Längen je nach Batch variieren können.

Im Folgenden finden Sie ein Beispiel für ein Positionsschema.

```
{
  "headerRow": false,
  "columns": [
    [
      {
        "targetHeader": "name",
        "type": "cleartext"
      }
    ],
    [
      {
        "targetHeader": "city_sealed",
        "type": "sealed",
        "pad": {
          "type": "max",
          "length": 16
        }
      }
    ],
    [
      {
        "targetHeader": "phone_number_fingerprint",
```



```

    "type": "fingerprint"
  },
  {
    "targetHeader": "phone_number_sealed",
    "type": "sealed",
    "pad": {
      "type": "fixed",
      "length": 20
    }
  }
]
}

```

Im Folgenden finden Sie ein Beispiel für eine CSV-Beispieldatei, falls sie nicht die erste Zeile mit den Überschriften hatte.

```

Jorge,Souza,12345 Mills Rd,Anytown,SC, 703 -555 -1234,CEO, 10,
Paulo,Santos, 0 Street,Anytown,MD, 404-555-111,CIO, 9,This is a really long note that
could really be a paragraph
Mateo,Jackson, 1 Two St,Anytown,NY, 304-555-1324,C00, 9, ""
Terry,Whitlock, 4 N St,Anytown,VA, 407-555-8888,EA, 7,Secret notes
Diego,Ramirez, 9 Hollows Rd,Anytown,VA, 407-555-1222,SDE I, 4,null
John,Doe, 8 Hollows Rd,Anytown,VA, 407-555-4321,SDE I, 4,Jane's younger brother
Jane,Doe, 8 Hollows Rd,Anytown,VA, 407-555-4322,SDE II, 5,John's older sister

```

Das Positionsschema hat die folgende Form.

```

{
  "headerRow": false,
  "columns": [
    [
      {
        "targetHeader": "GivenName",
        "type": "cleartext"
      }
    ],
    [
      {
        "targetHeader": "Surname",
        "type": "cleartext"
      }
    ]
  ],

```

```
[
  [],
  [
    {
      "targetHeader": "State_Join",
      "type": "fingerprint"
    },
    {
      "targetHeader": "State",
      "type": "sealed",
      "pad": {
        "type": "none"
      }
    }
  ],
  [],
  [],
  [],
  []
]
```

Das vorherige Schema erzeugt die folgende Ausgabedatei mit einer Kopfzeile, die die angegebenen Ziel-Header enthält.

```
givenname,surname,state_fingerprint,state
Mateo,Jackson,01:hmac:iIRnjfNBzryusIJ1w35lgNzeY1RQ1bSfq6PDHW8Xrbk=,01:enc:ENS6QD3cMV19vQEGfe9MM
Q8m/Y5SA89dJwKpT5rGPP8e36h6klwDoslpFzGvU0=
Jorge,Souza,01:hmac:3BxJdXiFFyZ8HBbYNqqEhBVqhN0d7s2ZiKUe7QiTy08=,01:enc:LKo0zirq2+
+XEIIIMNRjAsGmdyWUDwYaum0B+IFP+rUf1BNeZDJjtFe1Z+zbZfXQWwJy52Rt7HqvAb2WIK1oMmk=
Paulo,Santos,01:hmac:CHF4eIrtTNgAooU9v4h9Qjc
+txBnMidQTjdjWuaDTTA=,01:enc:MyQKyWxJ9kvK1xDQQtX1UNwv3F+yrBRr0xrUY/1BGg5Kfg0n9pK+MZ7g
+ZNqZEPcPz4lht1u0t/wbTaqz0CLXFQ=
Jane,Doe,01:hmac:UK8s8Cn/WR2J0/To2dTxWD73aDEe2ZUXeSHy3Tv+1Mk=,01:enc:Pd8sbITBfb0/
ttUB4svVsgoYkDfnDvgkvxzeCi0Yxq54rLSwccy1o3/B50C3cpkkn56dovCwzgmPNwrmCmYtb4=
Terry,Whitlock01:hmac:UK8s8Cn/WR2J0/To2dTxWD73aDEe2ZUXeSHy3Tv
+1Mk=,01:enc:Qmtzu3B3GAXKh2KkRYTiEAaMopYedsSdF2e/
ADUiBQ9kv2CxCkPzWyYTD3ztmKPMka19dHre5VhUHNp030+j1AQ8=
Diego,Ramirez,01:hmac:UK8s8Cn/WR2J0/To2dTxWD73aDEe2ZUXeSHy3Tv+1Mk=,01:enc:ysdg
+GHKdeZrS/geBIoo0EPLHG68MsWpx1dh3xjb+fG5rmFmqUcJLNuuYBHhHA1xchM2WVeV1fmHkBX3mvZNvkc=
John,Doe,01:hmac:UK8s8Cn/WR2J0/To2dTxWD73aDEe2ZUXeSHy3Tv+1Mk=,01:enc:9uX0wZu07kAPAx
+Hf6uvQownkWqFSKtWS7gQIJSe5aXFquKWCK6yZN0X5Ea2N3bn03Uj1kh0agDwoiP9FRZGJA4=
```

Datentabellen mit dem C3R-Verschlüsselungsclient entschlüsseln

Folgen Sie diesem Verfahren für Kollaborationen, bei denen Cryptographic Computing für Clean Rooms und den C3R-Verschlüsselungsclient zum Verschlüsseln von Datentabellen. Verwenden Sie dieses Verfahren, nachdem Sie [Daten in der Zusammenarbeit abgefragt](#) haben.

Der gemeinsame geheime Schlüssel und die Kollaborations-ID sind für dieses Verfahren erforderlich.

Das Mitglied, das Ergebnisse erhalten kann, entschlüsselt die Daten mit demselben gemeinsamen geheimen Schlüssel und derselben Kollaborations-ID, die zur Verschlüsselung der Daten für die Kollaboration verwendet wurden.

Note

AWS Clean Rooms Kollaborationen schränken bereits ein, wer Abfrageergebnisse ausführen und anzeigen kann. Um die Entschlüsselung durchzuführen, benötigt jeder, der Zugriff auf diese Ergebnisse hat, denselben gemeinsamen geheimen Schlüssel und dieselbe Kollaborations-ID, mit der die Daten verschlüsselt wurden.

Um eine verschlüsselte Datentabelle zu entschlüsseln

1. (Optional) [Sehen Sie sich die verfügbaren Befehle im C3R-Verschlüsselungsclient](#) an.
 2. (Optional) Navigieren Sie zum gewünschten Verzeichnis und führen Sie `ls` (macOS) oder `dir` (Windows).
- Vergewissern Sie sich, dass `c3r-cli.jar` Datei und Datendatei mit verschlüsselten Abfrageergebnissen befinden sich im gewünschten Verzeichnis.

Note

Wenn Abfrageergebnisse von der AWS Clean Rooms Konsolenoberfläche heruntergeladen werden, befinden sie sich wahrscheinlich im Download-Ordner Ihres Benutzerkontos. (Zum Beispiel der Ordner Downloads in Ihrem Benutzerverzeichnis auf Windows and macOS.) Wir empfehlen, dass Sie die Datei mit den Abfrageergebnissen in denselben Ordner verschieben wie `c3r-cli.jar`.

3. Speichern Sie den gemeinsamen geheimen Schlüssel in der C3R_SHARED_SECRET Umgebungsvariablen. Weitere Informationen finden Sie unter [Schritt 6: Speichern Sie den gemeinsamen geheimen Schlüssel in einer Umgebungsvariablen](#).

4. Führen Sie von AWS Command Line Interface (AWS CLI) aus den folgenden Befehl aus.

```
java -jar c3r-cli.jar decrypt <name of input .csv file> --id=<collaboration id> --  
output=<output file name>
```

5. Ersetzen Sie jeden *user input placeholder* durch Ihre eigenen Informationen:
 - a. Geben Sie für id= die Kollaborations-ID ein.
 - b. Geben Sie für output= den Namen der Ausgabedatei ein (z. B. results-decrypt.csv).

Wenn Sie keinen Ausgabenamen angeben, wird im Terminal ein Standardname angezeigt.

- c. Zeigen Sie die entschlüsselten Daten in der angegebenen Ausgabedatei mit Ihrer bevorzugten CSV-Datei an oder Parquet Anwendung anzeigen (z. B. Microsoft Excel, ein Texteditor oder eine andere Anwendung).

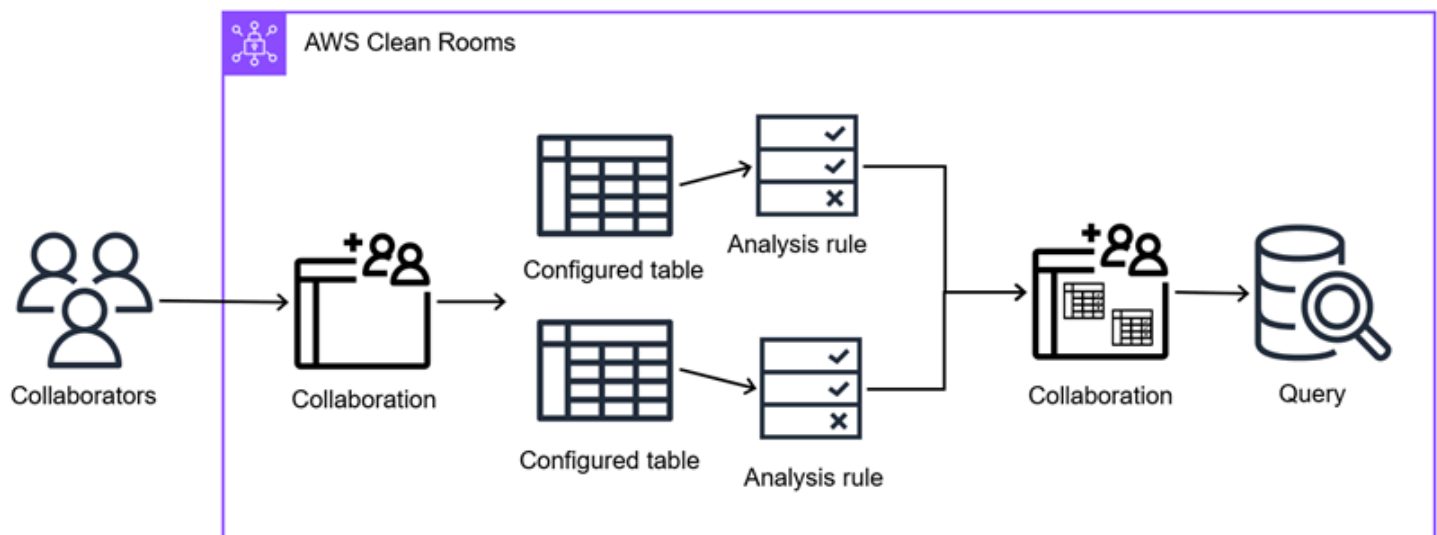
Konfigurierte Tabellen in AWS Clean Rooms

Eine konfigurierte Tabelle ist ein Verweis auf eine vorhandene Tabelle in einer Datenquelle. Sie enthält eine Analyseregeln, die festlegt, wie die Daten abgefragt werden können, AWS Clean Rooms und kann ein Datenzugriffsbudget zur Steuerung der Tabellennutzung beinhalten. Konfigurierte Tabellen können einer oder mehreren Kollaborationen zugeordnet werden.

Mit AWS Clean Rooms können Sie Aggregationsanalysen für Ereignisdaten durchführen, z. B. die Anzahl der Käufe im Vergleich zur Anzahl der Käufe. Sie können auch Listenanalysen für Ereignisdaten durchführen, z. B. die Anreicherung sich überschneidender Kundendaten von Segmentdaten bis hin zu CRM-Daten. Sie können auch benutzerdefinierte Abfragen durchführen und unterschiedliche Datenschutzeinstellungen für Veranstaltungsdaten wie Zuschauerdaten und Segmentattribute festlegen. Für jeden dieser Analysetypen können Sie Budgets für den Datenzugriff festlegen, um zu überwachen und zu kontrollieren, auf welche Menge Ihrer Daten über Abfragen zugegriffen wird.

Zunächst erstellen Sie eine Kollaboration in AWS Clean Rooms und fügen die hinzu, die AWS-Konten Sie einladen möchten, oder treten einer Kollaboration bei, zu der Sie eingeladen wurden, indem Sie eine Mitgliedschaft erstellen. Als Nächstes erstellen Sie und das andere Mitglied der Kollaboration konfigurierte Tabellen. Sie fügen beide den konfigurierten Tabellen (Aggregation, Liste oder benutzerdefiniert) eine Analyseregeln hinzu und legen optional Budgets für den Datenzugriff fest. Anschließend ordnen Sie die konfigurierten Tabellen der Kollaboration zu. Schließlich führt das Mitglied, das Abfragen durchführen kann, eine Abfrage für die beiden Datentabellen aus, wodurch das Datenzugriffsbudget bei der Ausführung der Abfragen aufgebraucht wird.

Das folgende Diagramm fasst zusammen, wie Sie mit Ereignisdaten in AWS Clean Rooms arbeiten.



Topics

- [Erstellen einer konfigurierten Tabelle in AWS Clean Rooms](#)
- [Hinzufügen einer Analyseregeln zu einer konfigurierten Tabelle](#)
- [Eine konfigurierte Tabelle einer Kooperation zuordnen](#)
- [Ein Datenzugriffsbudget konfigurieren](#)
- [Eine Regel für die Kollaborationsanalyse zu einer konfigurierten Tabelle hinzufügen](#)
- [Konfiguration der differenzierten Datenschutzrichtlinie \(optional\)](#)
- [Tabellen und Analyseregeln anzeigen](#)
- [Eine konfigurierte Tabelle bearbeiten](#)
- [Bearbeiten konfigurierter Tabellen-Tags](#)
- [Bearbeitung der konfigurierten Tabellenanalyseregeln](#)
- [Löschen der konfigurierten Tabellenanalyseregeln](#)
- [In der konfigurierten Tabelle sind keine Spalten zulässig](#)
- [Bearbeiten konfigurierter Tabellenzuordnungen](#)
- [Aufheben der Zuordnung konfigurierter Tabellen](#)

Erstellen einer konfigurierten Tabelle in AWS Clean Rooms

Eine konfigurierte Tabelle ist ein Verweis auf eine vorhandene Tabelle in einer Datenquelle. Sie enthält eine Analyseregeln, die festlegt, wie die Daten abgefragt werden können. AWS Clean Rooms Konfigurierte Tabellen können einer oder mehreren Kollaborationen zugeordnet werden.

Informationen zum Erstellen einer konfigurierten Tabelle mithilfe von finden Sie in der [AWS Clean Rooms API-Referenz](#). [AWS SDKs](#)

Themen

- [Eine konfigurierte Tabelle erstellen — Amazon S3 S3-Datenquelle](#)
- [Eine konfigurierte Tabelle erstellen — Amazon Athena Athena-Datenquelle](#)
- [Eine konfigurierte Tabelle erstellen — Snowflake-Datenquelle](#)

Eine konfigurierte Tabelle erstellen — Amazon S3 S3-Datenquelle

In diesem Verfahren führt das [Mitglied](#) die folgenden Aufgaben aus:

- Konfiguriert eine vorhandene AWS Glue Tabelle zur Verwendung in AWS Clean Rooms. (Dieser Schritt kann vor oder nach dem Beitritt zu einer Kollaboration durchgeführt werden, es sei denn, Sie verwenden Cryptographic Computing für Clean Rooms.)

Note

AWS Clean Rooms unterstützt AWS Glue Tabellen. Weitere Hinweise zur Eingabe Ihrer Daten finden Sie unter [Schritt 3: Laden Sie Ihre Datentabelle auf Amazon S3 hoch](#). AWS Glue

- Benennt die [konfigurierte Tabelle](#) und wählt aus, welche Spalten in der Kollaboration verwendet werden sollen.

Das folgende Verfahren setzt voraus, dass:

- Das Kollaborationsmitglied hat [seine Datentabellen bereits auf Amazon S3 hochgeladen](#) und [eine AWS Glue Tabelle erstellt](#).

Note

Das Ergebnisziel in Amazon S3 darf sich nicht im selben S3-Bucket wie jede Datenquelle befinden.

- (Optional) Nur für [verschlüsselte](#) Datentabellen hat das Kollaborationsmitglied bereits [verschlüsselte Datentabellen mit dem C3R-Verschlüsselungsclient erstellt](#).

Sie können die von bereitgestellte Statistikgenerierung verwenden, AWS Glue um Statistiken auf Spaltenebene für Tabellen zu berechnen. AWS Glue Data Catalog Nach der AWS Glue Generierung von Statistiken für Tabellen im Datenkatalog verwendet Amazon Redshift Spectrum diese Statistiken automatisch, um den Abfrageplan zu optimieren. Weitere Informationen zur Berechnung von Statistiken auf Spaltenebene mithilfe von AWS Glue Daten finden Sie unter [Optimieren der Abfrageleistung mithilfe von Spaltenstatistiken](#) im AWS Glue Benutzerhandbuch. Weitere Informationen AWS Glue dazu finden Sie im [AWS Glue Developer Guide](#).

So erstellen Sie eine konfigurierte Tabelle — Amazon S3 S3-Datenquelle

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Tables (Tabellen) aus.
3. Wählen Sie in der oberen rechten Ecke die Option Neue Tabelle konfigurieren aus.
4. Wählen Sie als Datenquelle unter AWS Datenquellen Amazon S3 aus.
5. In der Amazon S3 S3-Tabelle:
 - a. Wählen Sie die Region aus, in der die S3-Tabelle gehostet wird.

Standardmäßig ist die aktuelle Region (z. B. Nord-Virginia us-east-1) ausgewählt.

 Warning

Wenn sich Ihre Amazon S3 S3-Datenquelle in einer anderen Region als Ihrem Verarbeitungsstandort befindet, kann die Datenverarbeitung vorübergehend außerhalb der Quellregion erfolgen. Bevor Sie fortfahren, stellen Sie sicher, dass die regionsübergreifende Übertragung von Daten Ihren Anforderungen an die Datenhoheit, den Richtlinien zur Einhaltung gesetzlicher Vorschriften und den Datenverwaltungsstandards entspricht.

Weitere Informationen zu Regionen finden Sie unter [Regionen und Endpunkte](#) in der Allgemeine AWS-Referenz

- b. Wählen Sie die Datenbank aus der Dropdownliste aus.
- c. Wählen Sie die Tabelle, die Sie konfigurieren möchten, aus der Dropdownliste aus.

 Note

Um zu überprüfen, ob es sich um die richtige Tabelle handelt, führen Sie einen der folgenden Schritte aus:

- Wählen Sie Anzeigen in AWS Glue.
- Aktivieren Sie „Schema anzeigen von“ AWS Glue, um das Schema anzuzeigen.

 Important

Bei AWS Glue Tabellen, in denen die Daten im CSV-Format vorliegen, müssen die Spaltennamen und die Reihenfolge im Glue-Schema genau mit den CSV-Daten übereinstimmen. Wenn sie nicht aufeinander abgestimmt sind, wird die Liste der zulässigen Spalten für die konfigurierte Tabelle möglicherweise nicht richtig durchgesetzt.

6. Für Spalten und Analysemethoden, die in Kollaborationen zulässig sind,
 - a. Für welche Spalten möchten Sie in Kollaborationen zulassen?
 - Wählen Sie Alle Spalten aus, damit alle Spalten in der Kollaboration abgefragt werden können.
 - Wählen Sie Benutzerdefinierte Liste aus, damit eine oder mehrere Spalten aus der Dropdownliste Zulässige Spalten angeben in der Kollaboration abgefragt werden können.
 - b. Für Zulässige Analysemethoden
 - i. Wählen Sie Direkte Abfrage, damit SQL-Abfragen direkt in dieser Tabelle ausgeführt werden können
 - ii. Wählen Sie Direkter Job, damit PySpark Jobs direkt in dieser Tabelle ausgeführt werden können.

Example Beispiel

Wenn Sie beispielsweise Mitgliedern der Kollaboration ermöglichen möchten, sowohl direkte SQL-Abfragen als auch PySpark Jobs für alle Spalten auszuführen, wählen Sie Alle Spalten, Direkte Abfrage und Direkter Job aus.

7. Einzelheiten zur konfigurierten Tabelle finden Sie unter

- a. Geben Sie einen Namen für die konfigurierte Tabelle ein.

Sie können den Standardnamen verwenden oder diese Tabelle umbenennen.

- b. Geben Sie eine Beschreibung der Tabelle ein.

Die Beschreibung hilft dabei, zwischen anderen konfigurierten Tabellen mit ähnlichen Namen zu unterscheiden.

8. Wenn Sie Tags für die konfigurierte Tabellenressource aktivieren möchten, wählen Sie Neues Tag hinzufügen und geben Sie dann das Schlüssel - und Wertepaar ein.

9. Wählen Sie Neue Tabelle konfigurieren.

Nachdem Sie eine konfigurierte Tabelle erstellt haben, können Sie:

- [Fügen Sie der konfigurierten Tabelle eine Analyseregeln hinzu](#)
- [Ordnen Sie die konfigurierte Tabelle einer Kollaboration zu](#)

Eine konfigurierte Tabelle erstellen — Amazon Athena Athena-Datenquelle

Mit der Amazon Athena Athena-Datenquellenoption können Sie Daten abfragen, die in Amazon S3 gespeichert, im AWS Glue Datenkatalog oder in Verbundkatalogen katalogisiert sind und deren Zugriff über gesteuert wird. AWS Lake Formation Sowohl Tabellen als auch Ansichten werden unterstützt. AWS Glue Data Catalog Ressourcenlinks von Lake Formation können verwendet werden, um Tabellen und Ansichten für AWS-Konten das AWS Clean Rooms Mitgliedskonto gemeinsam AWS-Regionen zu nutzen, das sie zu einer AWS Clean Rooms Kollaboration verbindet.

 Note

Nur Amazon S3-basierte Datensätze können über die Athena-Datenquellenintegration abgefragt werden.

In diesem Verfahren führt das Mitglied die folgenden Aufgaben aus:

- Konfiguriert eine vorhandene Tabelle oder Ansicht in der AWS Glue Data Catalog zur Verwendung von AWS Clean Rooms
- Benennt die konfigurierte Tabelle und wählt aus, welche Spalten in der Kollaboration verwendet werden sollen.

Das folgende Verfahren setzt voraus, dass:

- Das Kollaborationsmitglied hat die AWS Glue Data Catalog Datenbank und die Tabelle oder die GDC-Ansicht bereits erstellt.

Um eine konfigurierte Tabelle zu erstellen — Athena-Datenquelle

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Tables (Tabellen) aus.
3. Wählen Sie in der oberen rechten Ecke die Option Neue Tabelle konfigurieren aus.
4. Wählen Sie als Datenquelle unter AWS Datenquellen Amazon Athena aus.
5. Unter der Amazon Athena Athena-Tabelle:
 - a. Wählen Sie die Region aus, in der die Amazon Athena Athena-Tabelle gehostet wird.

Standardmäßig ist die aktuelle Region (z. B. Nord-Virginia us-east-1) ausgewählt.

 Warning

Wenn sich Ihre Amazon Athena Athena-Datenquelle in einer anderen Region als Ihrem Verarbeitungsstandort befindet, kann die Datenverarbeitung vorübergehend außerhalb der Quellregion erfolgen. Bevor Sie fortfahren, stellen Sie sicher, dass die regionsübergreifende Übertragung von Daten Ihren Anforderungen an die

Datenhoheit, den Richtlinien zur Einhaltung gesetzlicher Vorschriften und den Datenverwaltungsstandards entspricht.

Weitere Informationen zu Regionen finden Sie unter [Regionen und Endpunkte](#) in der. Allgemeine AWS-Referenz

- b. Wählen Sie den Katalog aus der Dropdownliste aus.

Standardmäßig ist AWS Glue Datenkatalog ausgewählt.

- AWS Glue Datenkatalog — Der Standardkatalog für Tabellen in AWS Glue.
- Verbundkatalog — Verfügbar, wenn Sie Catalog Federation so konfiguriert haben, dass eine Verbindung zu AWS Glue Remote-REST-Katalogen von Apache Iceberg hergestellt wird. Weitere Informationen finden Sie unter [Catalog Federation im Developer Guide](#).AWS Lake Formation

- c. Wählen Sie die Datenbank aus der Dropdownliste aus.
- d. Wählen Sie die Tabelle, die Sie konfigurieren möchten, aus der Dropdownliste aus.

Note

Um zu überprüfen, ob es sich um die richtige Tabelle handelt, führen Sie einen der folgenden Schritte aus:

- Wählen Sie „Anzeigen in“ AWS Glue oder „Anzeigen in“ AWS Lake Formation(abhängig von Ihrem Katalogtyp).
- Aktivieren Sie „Schema anzeigen von“ AWS Glue, um das Schema anzuzeigen.

6. Für Amazon Athena Athena-Konfigurationen

- a. Wählen Sie eine Arbeitsgruppe aus der Drop-down-Liste aus.
- b. Wählen Sie für den S3-Ausgabespeicherort eine empfohlene Aktion aus, die auf einem der folgenden Szenarien basiert.

Szenario	Empfohlene Aktion
Ihre Arbeitsgruppe hat keinen Standardausgabespeicherort.	Geben Sie den S3-Ausgabespeicherort ein oder wählen Sie Browse S3.
Ihre Arbeitsgruppe erzwingt Ihren Standardausgabeort.	Der S3-Ausgabespeicherort wird automatisch ausgewählt und kann nicht geändert werden.
Ihre Arbeitsgruppe erzwingt Ihren Standardausgabeort nicht.	Geben Sie den S3-Ausgabespeicherort ein oder wählen Sie „S3 durchsuchen“.

7. Wählen Sie für Spalten, die in Kollaborationen zulässig sind, eine Option aus, die auf Ihrem Ziel basiert.

Dein Ziel	Empfohlene Option
Erlaube die Verwendung aller Spalten in AWS Clean Rooms (vorbehaltlich der Analyseregeln)	Alle Spalten
Lassen Sie eine oder mehrere Spalten aus der Dropdownliste Zulässige Spalten angeben zu	Benutzerdefinierte Liste

8. Einzelheiten zur konfigurierten Tabelle finden Sie

- a. Geben Sie einen Namen für die konfigurierte Tabelle ein.

Sie können den Standardnamen verwenden oder diese Tabelle umbenennen.

- b. Geben Sie eine Beschreibung der Tabelle ein.

Die Beschreibung hilft dabei, zwischen anderen konfigurierten Tabellen mit ähnlichen Namen zu unterscheiden.

- c. Wenn Sie Tags für die konfigurierte Tabellenressource aktivieren möchten, wählen Sie Neues Tag hinzufügen und geben Sie dann das Schlüssel - und Wertepaar ein.

9. Wählen Sie Neue Tabelle konfigurieren.

Nachdem Sie eine konfigurierte Tabelle erstellt haben, können Sie:

- [Fügen Sie der konfigurierten Tabelle eine Analyseregeln hinzu](#)
- [Ordnen Sie die konfigurierte Tabelle einer Kollaboration zu](#)

Eine konfigurierte Tabelle erstellen — Snowflake-Datenquelle

In diesem Verfahren führt das [Mitglied](#) die folgenden Aufgaben aus:

- Konfiguriert eine vorhandene Snowflake-Tabelle zur Verwendung in. AWS Clean Rooms(Dieser Schritt kann vor oder nach dem Beitritt zu einer Kollaboration durchgeführt werden, es sei denn, Sie verwenden Cryptographic Computing für.) Clean Rooms
- Benennt die [konfigurierte Tabelle](#) und wählt aus, welche Spalten in der Kollaboration verwendet werden sollen.

Das folgende Verfahren setzt voraus, dass:

- Das Mitglied der Kollaboration hat seine Datentabellen bereits auf Snowflake hochgeladen.
- (Optional) Nur für [verschlüsselte](#) Datentabellen hat das Kollaborationsmitglied bereits [verschlüsselte Datentabellen mit dem C3R-Verschlüsselungsclient erstellt](#).

Um eine konfigurierte Tabelle zu erstellen — Snowflake-Datenquelle

1. [Melden Sie sich bei /cleanrooms an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole. https://console.aws.amazon.com](#)
2. Wählen Sie im linken Navigationsbereich Tables (Tabellen) aus.
3. Wählen Sie in der oberen rechten Ecke die Option Neue Tabelle konfigurieren aus.
4. Wählen Sie für Datenquelle unter Clouds und Datenquellen von Drittanbietern die Option Snowflake aus.
5. Geben Sie die Snowflake-Anmeldeinformationen mithilfe eines vorhandenen geheimen ARN an oder speichern Sie ein neues Geheimnis für diese Tabelle.

Use existing secret ARN

1. Wenn Sie einen geheimen ARN haben, geben Sie ihn in das Feld Geheimer ARN ein.

Sie können Ihren geheimen ARN nachschlagen, indem Sie Gehe zu wählen AWS Secrets Manager.

2. Wenn Sie über ein vorhandenes Geheimnis aus einer anderen Tabelle verfügen, wählen Sie Geheime ARN aus vorhandener Tabelle importieren.



Note

Der geheime ARN kann kontoübergreifend sein.

Store a new secret for this table

1. Geben Sie die folgenden Snowflake-Anmeldeinformationen ein:
 - Snowflake-Benutzername
 - Snowflake-Lagerhaus
 - Die Rolle der Schneeflocke
 - Privater Schlüssel für Snowflake Privacy Enhanced Mail (PEM)
2. Führen Sie zur Verschlüsselung einen der folgenden Schritte aus:
 - Um den Von AWS verwalteter Schlüssel (Standard) zu verwenden, lassen Sie das Kontrollkästchen Verschlüsselungseinstellungen anpassen deaktiviert.
 - Um eine benutzerdefinierte zu verwenden AWS KMS key:
 - Aktivieren Sie das Kontrollkästchen Verschlüsselungseinstellungen anpassen.
 - Geben Sie für KMS-Schlüssel den Schlüssel-ARN ein oder wählen Sie einen aus der Liste aus.
3. Geben Sie einen geheimen Namen ein, damit Sie Ihre Anmeldeinformationen später leichter finden können.
6. Geben Sie für Snowflake-Tabellen- und Schemadetails die Details manuell ein oder importieren Sie sie automatisch.

Enter the details manually

1. Geben Sie die Snowflake-Konto-ID ein.

Weitere Informationen finden Sie unter [Kontokennungen](#) in der Snowflake-Dokumentation.

Ihre Konto-ID muss das für Snowflake-Treiber verwendete Format haben. Sie müssen den Punkt (.) durch einen Bindestrich (-) ersetzen, sodass der Bezeichner als formatiert wird.

<orgname>-<account_name>

2. Geben Sie die Snowflake-Datenbank ein.

Weitere Informationen finden Sie unter [Snowflake-Datenbank in der Snowflake-Dokumentation](#).

3. Geben Sie den Namen des Snowflake-Schemas ein.
4. Geben Sie den Namen der Snowflake-Tabelle ein.

Weitere Informationen finden Sie unter [Grundlegendes zu Snowflake-Tabellenstrukturen in der Snowflake-Dokumentation](#).

5. Geben Sie für das Schema den Spaltennamen ein und wählen Sie den Datentyp aus der Dropdownliste aus.
6. Wählen Sie Spalte hinzufügen, um weitere Spalten hinzuzufügen.
 - Wenn Sie einen Objektdatentyp wählen, geben Sie das Objektschema an.

Example Beispiel für ein Objektschema

```
name STRING,  
location OBJECT(  
    x INT,  
    y INT,  
    metadata OBJECT(uuid STRING)  
),  
history ARRAY(TEXT)
```

- Wenn Sie einen Array-Datentyp wählen, geben Sie das Array-Schema an.

Example Beispiel für ein Array-Schema

```
OBJECT(x INT, y INT)
```

- Wenn Sie einen Map-Datentyp wählen, geben Sie das Map-Schema an.

Example Beispiel für ein Kartenschema

```
STRING, OBJECT(x INT, y INT)
```


Automatically import the details

1. Exportieren Sie Ihre COLUMNS-Ansicht aus Snowflake als CSV-Datei.

Weitere Informationen zur Snowflake-COLUMNS-Ansicht finden Sie unter [COLUMNS-Ansicht in der Snowflake-Dokumentation](#).

2. Wählen Sie Aus Datei importieren, um die CSV-Datei zu importieren, und geben Sie weitere Informationen an.

Der Datenbankname, der Schemaname, der Tabellename, die Spaltennamen und die Datentypen werden automatisch importiert.

- Wenn Sie einen Objektdatentyp wählen, geben Sie das Objektschema an.
- Wenn Sie einen Array-Datentyp wählen, geben Sie das Array-Schema an.
- Wenn Sie einen Map-Datentyp wählen, geben Sie das Map-Schema an.

3. Geben Sie die Snowflake-Konto-ID ein.

Weitere Informationen finden Sie unter [Kontokennungen](#) in der Snowflake-Dokumentation.

Note

Nur katalogisierte S3-Tabellen AWS Glue können verwendet werden, um das Tabellenschema automatisch abzurufen.

7. Wählen Sie für Spalten, die in Kollaborationen zulässig sind, eine Option aus, die Ihrem Ziel entspricht.

Dein Ziel	Empfohlene Option
Erlaube die Verwendung aller Spalten in AWS Clean Rooms (vorbehaltlich der Analyseregeln)	Alle Spalten
Lassen Sie eine oder mehrere Spalten aus der Dropdownliste Zulässige Spalten angeben zu	Benutzerdefinierte Liste

8. Einzelheiten zur konfigurierten Tabelle finden Sie

- a. Geben Sie einen Namen für die konfigurierte Tabelle ein.

Sie können den Standardnamen verwenden oder diese Tabelle umbenennen.

- b. Geben Sie eine Beschreibung der Tabelle ein.

Die Beschreibung hilft dabei, zwischen anderen konfigurierten Tabellen mit ähnlichen Namen zu unterscheiden.

- c. Wenn Sie Tags für die konfigurierte Tabellenressource aktivieren möchten, wählen Sie Neues Tag hinzufügen und geben Sie dann das Schlüssel - und Wertepaar ein.

9. Wählen Sie Neue Tabelle konfigurieren.

Nachdem Sie eine konfigurierte Tabelle erstellt haben, können Sie:

- [Fügen Sie der konfigurierten Tabelle eine Analyseregeln hinzu](#)
- [Ordnen Sie die konfigurierte Tabelle einer Kollaboration zu](#)

Hinzufügen einer Analyseregeln zu einer konfigurierten Tabelle

In den folgenden Abschnitten wird beschrieben, wie Sie Ihrer konfigurierten Tabelle eine Analyseregeln hinzufügen. Durch die Definition der Analyseregeln können Sie das Mitglied, das Abfragen durchführen kann, autorisieren, Abfragen auszuführen, die einer bestimmten Analyseregeln entsprechen, die von unterstützt wird. AWS Clean Rooms

AWS Clean Rooms unterstützt die folgenden Arten von Analyseregeln:

- [Regel für die Aggregationsanalyse](#)
- [Regel für die Listenanalyse](#)
- [Benutzerdefinierte Analyseregeln in AWS Clean Rooms](#)

Pro konfiguierter Tabelle kann es nur eine Analyseregeln geben. Sie können die Analyseregeln jederzeit konfigurieren, bevor Sie Ihre konfigurierten Tabellen der Kollaboration zuordnen.

 Important

Wenn Sie Cryptographic Computing für verwenden Clean Rooms und in der Kollaboration verschlüsselte Datentabellen verwenden, sollte die Analyseregeln, die Sie der verschlüsselten konfigurierten Tabelle hinzufügen, mit der Art der Verschlüsselung der Daten übereinstimmen. Wenn Sie beispielsweise die Daten für SELECT (Aggregationsanalyseregeln) verschlüsselt haben, sollten Sie die Analyseregeln für JOIN (Listenanalyseregeln) nicht hinzufügen.

Topics

- [Hinzufügen einer Aggregationsanalyseregeln zu einer Tabelle \(geführter Ablauf\)](#)
- [Hinzufügen einer Listenanalyseregeln zu einer Tabelle \(geführter Ablauf\)](#)
- [Hinzufügen einer benutzerdefinierten Analyseregeln zu einer Tabelle \(geführter Ablauf\)](#)
- [Analyseregeln zu einer Tabelle hinzufügen \(JSON-Editor\)](#)
- [Nächste Schritte](#)

Hinzufügen einer Aggregationsanalyseregeln zu einer Tabelle (geführter Ablauf)

Die Aggregationsanalyseregeln ermöglicht Abfragen, bei denen Statistiken aggregiert werden, ohne Informationen auf Zeilenebene mithilfe von COUNTSUM, und AVG Funktionen entlang optionaler Dimensionen preiszugeben.

In diesem Verfahren wird beschrieben, wie Sie Ihrer konfigurierten Tabelle mithilfe der Option Guided Flow in der Konsole eine Aggregationsanalyseregeln hinzufügen. AWS Clean Rooms

 Note

Konfigurierte Tabellen, die Nicht-S3-Datenquellen verwenden, unterstützen nur [benutzerdefinierte Analyseregeln](#).

Um die Aggregationsanalyseregel zu einer Tabelle hinzuzufügen (geführter Ablauf)

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
 2. Wählen Sie im linken Navigationsbereich Tables (Tabellen) aus.
 3. Wählen Sie die konfigurierte Tabelle aus.
 4. Wählen Sie auf der Detailseite der konfigurierten Tabelle die Option Analyseregel konfigurieren aus.
 5. Wählen Sie unter Schritt 1: Analyseregeltyp auswählen und unter Analyseregeltyp die Option Aggregation aus.
 6. Wählen Sie unter Erstellungsmethode die Option Geführter Ablauf und dann Weiter aus.
 7. Gehen Sie unter Schritt 2: Abfragesteuerelemente angeben für Aggregatfunktionen wie folgt vor:
 - a. Wählen Sie eine Aggregatfunktion aus der Dropdownliste aus:
 - ZÄHLEN
 - ZÄHLE UNTERSCHIEDLICH
 - SUM
 - DIFFERENZIERTE SUMME
 - AVG
 - b. Wählen Sie in der Dropdownliste Spalten aus, welche Spalten in der Aggregatfunktion verwendet werden können.
 - c. (Optional) Wählen Sie Weitere Funktion hinzufügen, um eine weitere Aggregatfunktion hinzuzufügen und dieser Funktion eine oder mehrere Spalten zuzuordnen.
-  **Note**

Es ist mindestens eine Aggregatfunktion erforderlich.
- d. (Optional) Wählen Sie Entfernen, um eine Aggregatfunktion zu entfernen.
8. Für Join-Steuerelemente
 - a. Wählen Sie eine Option für die automatische Abfrage der Tabelle zulassen aus:

Wenn Sie folgendes auswählen ...	Dann...
Nein, es können nur Überschneidungen abgefragt werden	Die Tabelle kann nur abgefragt werden, wenn sie mit einer Tabelle verknüpft ist, die dem Mitglied gehört, das Abfragen ausführen kann.
Ja	Die Tabelle kann eigenständig oder wenn sie mit anderen Tabellen verknüpft ist, abgefragt werden.

- b. Wählen Sie unter Verbindungsspalten angeben die Spalten aus, deren Verwendung in der INNER JOIN Anweisung zulässig sein soll.

Dies ist optional, wenn Sie im vorherigen Schritt Ja ausgewählt haben.

- c. Wählen Sie unter Zulässige Operatoren für den Abgleich angeben aus, welche Operatoren gegebenenfalls für den Abgleich mehrerer Join-Spalten verwendet werden können. Wenn Sie zwei oder mehr JOIN Spalten auswählen, ist einer dieser Operatoren erforderlich.

Wenn Sie folgendes auswählen ...	Dann...
UND	Sie können AND in das INNER JOIN Spiel Bedingungen einbeziehen, um eine Spalte mit einer anderen Spalte zwischen Tabellen zu verbinden.
ODER	Sie können OR in die INNER JOIN Abgleichsbedingungen aufnehmen, um mehrere Spaltenübereinstimmungen zwischen Tabellen zu kombinieren. Dieser logische Operator ist nützlich, um eine höhere Trefferquote zu erzielen.

9. (Optional) Wählen Sie für Dimensionssteuerelemente in der Dropdownliste Dimensionsspalten angeben aus, welche Spalten Sie in der SELECT-Anweisung sowie in den ORDER BY Teilen WHERE GROUPBY, und der Abfrage verwenden lassen möchten.

Note

Aggregatfunktion oder Join-Spalten können nicht als Dimensionsspalten verwendet werden.

10. Wählen Sie für Skalarfunktionen eine Option für Welche Skalarfunktionen möchten Sie zulassen?

Wenn Sie folgendes auswählen ...	Dann...
Alle werden derzeit unterstützt von AWS Clean Rooms	Sie erlauben alle Skalarfunktionen, die derzeit von AWS Clean Rooms unterstützt werden. Sie können Liste anzeigen wählen, um die gesamte Liste der in unterstützten Skalarfunktionen anzuzeigen. AWS Clean Rooms
Eine benutzerdefinierte Liste	Sie können anpassen, welche Skalarfunktionen zulässig sind. Wählen Sie eine oder mehrere Optionen aus der Dropdownliste Zulässige Skalarfunktionen angeben aus.
Keine	Sie möchten keine Skalarfunktionen zulassen.

Weitere Informationen finden Sie unter [Skalarfunktionen](#).

11. Wählen Sie Weiter aus.
12. Gehen Sie unter Schritt 3: Steuerelemente für Abfrageergebnisse angeben für Aggregationseinschränkungen wie folgt vor:
- Wählen Sie die Dropdownliste für jeden Spaltennamen aus.

- b. Wählen Sie die Dropdownliste für jede Mindestanzahl unterschiedlicher Werte aus, die erfüllt sein müssen, damit jede Ausgabezeile zurückgegeben wird, nachdem die COUNT DISTINCT Funktion darauf angewendet wurde.
 - c. Wählen Sie Einschränkung hinzufügen, um weitere Aggregationseinschränkungen hinzuzufügen.
 - d. (Optional) Wählen Sie Entfernen, um eine Aggregationsbeschränkung zu entfernen.
13. Wählen Sie für Zusätzliche Analysen, die auf die Ausgabe angewendet werden, eine Option aus, die Ihrem Ziel entspricht.

Ihr Ziel	Empfohlene Option
Erlaube nur direkte Abfragen für diese Tabelle. Verweigern Sie die Ausführung zusätzlicher Analysen anhand von Abfrageergebnissen. Die Tabelle kann nur für direkte Abfragen verwendet werden.	Nicht erlaubt
Direkte Abfragen und zusätzliche Analysen in dieser Tabelle sind zulässig, aber nicht erforderlich.	Erlaubt
Erfordert, dass die Tabelle nur in direkten Abfragen verwendet werden kann, die mit einer der erforderlichen zusätzlichen Analysen verarbeitet werden. Direkte Abfragen für diese Tabelle müssen weiter verarbeitet werden, bevor sie zurückgegeben werden können.	Erforderlich

14. Wählen Sie Weiter aus.
15. Überprüfen Sie unter Schritt 4: Überprüfen und konfigurieren die Auswahlen, die Sie für die vorherigen Schritte getroffen haben, bearbeiten Sie sie gegebenenfalls und wählen Sie dann Analyseregeln konfigurieren aus.

Es wird eine Bestätigungsmeldung angezeigt, dass Sie erfolgreich eine Aggregationsanalyseregeln für die Tabelle konfiguriert haben.

Hinzufügen einer Listenanalyiseregeln zu einer Tabelle (geführter Ablauf)

Die Listenanalyiseregeln ermöglicht Abfragen, die Listen auf Zeilenebene ausgeben, in denen die Überschneidung zwischen der zugehörigen Tabelle und einer Tabelle des Mitglieds, das Abfragen durchführen kann, dargestellt wird.

Dieses Verfahren beschreibt den Vorgang des Hinzufügens der Listenanalyiseregeln zu Ihrer konfigurierten Tabelle mithilfe der Option Geführter Ablauf in der AWS Clean Rooms Konsole.

Note

Konfigurierte Tabellen, die Nicht-S3-Datenquellen verwenden, unterstützen nur [benutzerdefinierte Analyiseregeln](#).

Um einer Tabelle eine Regel für die Listenanalyse hinzuzufügen (geführter Ablauf)

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Tables (Tabellen) aus.
3. Wählen Sie die konfigurierte Tabelle aus.
4. Wählen Sie auf der Detailseite der konfigurierten Tabelle die Option Analyiseregeln konfigurieren aus.
5. Wählen Sie unter Schritt 1: Analyiseregeltyp auswählen und unter Analyiseregeltyp die Option Liste aus.
6. Wählen Sie unter Erstellungsmethode die Option Geführter Ablauf und dann Weiter aus.
7. Gehen Sie unter Schritt 2: Abfragesteuerelemente angeben für Join-Steuerelemente wie folgt vor:
 - a. Wählen Sie unter Join-Spalten angeben die Spalten aus, deren Verwendung in der INNER JOIN Anweisung zulässig sein soll.
 - b. Wählen Sie unter Zulässige Operatoren für den Abgleich angeben aus, welche Operatoren gegebenenfalls für den Abgleich mehrerer Join-Spalten verwendet werden können. Wenn Sie zwei oder mehr JOIN Spalten auswählen, ist einer dieser Operatoren erforderlich.

Wenn Sie folgendes auswählen ...	Dann...
UND	Sie können AND in das INNER JOIN Spiel Bedingungen einbeziehen, um eine Spalte mit einer anderen Spalte zwischen Tabellen zu verbinden.
ODER	Sie können OR in die INNER JOIN Abgleichsbedingungen aufnehmen, um mehrere Spaltenübereinstimmungen zwischen Tabellen zu kombinieren. Dieser logische Operator ist nützlich, um eine höhere Trefferquote zu erzielen.

8. (Optional) Wählen Sie für Listensteuerelemente in der Dropdownliste „Listenspalten angeben“ aus, welche Spalten in der Abfrageausgabe (d. h. in der SELECT Anweisung) oder zum Filtern von Ergebnissen (d. h. in der WHERE Anweisung) verwendet werden sollen.
9. Wählen Sie Weiter aus.
10. Wählen Sie unter Schritt 3: Steuerelemente für Abfrageergebnisse angeben für Zusätzliche Analysen, die auf die Ausgabe angewendet werden, eine Option aus, die Ihrem Ziel entspricht.

Ihr Ziel	Empfohlene Option
Erlaube nur direkte Abfragen für diese Tabelle. Verweigern Sie die Ausführung zusätzlicher Analysen anhand von Abfrageergebnissen. Die Tabelle kann nur für direkte Abfragen verwendet werden.	Nicht erlaubt
Direkte Abfragen und zusätzliche Analysen in dieser Tabelle sind zulässig, aber nicht erforderlich.	Erlaubt
Erfordert, dass die Tabelle nur in direkten Abfragen verwendet werden kann, die mit einer der erforderlichen zusätzlichen	Erforderlich

Ihr Ziel	Empfohlene Option
Analysen verarbeitet werden. Direkte Abfragen für diese Tabelle müssen weiter verarbeitet werden, bevor sie zurückgegeben werden können.	

- Überprüfen Sie unter Schritt 4: Überprüfen und konfigurieren die Auswahlen, die Sie für die vorherigen Schritte getroffen haben, bearbeiten Sie sie gegebenenfalls und wählen Sie dann Analyseregeln konfigurieren aus.

Es wird eine Bestätigungsmeldung angezeigt, dass Sie eine Listenanalyseregeln für die Tabelle erfolgreich konfiguriert haben.

Hinzufügen einer benutzerdefinierten Analyseregeln zu einer Tabelle (geführter Ablauf)

Die benutzerdefinierte Analyseregeln ermöglicht benutzerdefinierte SQL-Abfragen oder PySpark Jobs in einer konfigurierten Tabelle. Die benutzerdefinierte Analyseregeln ist erforderlich, wenn Sie Folgendes verwenden:

- [Analysevorlagen](#), die einen bestimmten Satz vorab genehmigter SQL-Abfragen oder PySpark - Jobs oder eine bestimmte Gruppe von Konten ermöglichen, die Abfragen bereitstellen können, die Ihre Daten verwenden.
- [AWS Clean Rooms Differenzierter Datenschutz](#) zum Schutz vor Versuchen zur Benutzeridentifikation.
- Nicht-S3-Datenquellen wie Amazon Athena oder Snowflake.

Dieses Verfahren beschreibt den Vorgang des Hinzufügens der benutzerdefinierten Analyseregeln zu Ihrer konfigurierten Tabelle mithilfe der Option Guided Flow in der Konsole. AWS Clean Rooms

So fügen Sie einer Tabelle eine benutzerdefinierte Analyseregeln hinzu (geführter Ablauf)

- Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
- Wählen Sie im linken Navigationsbereich Tables (Tabellen) aus.
- Wählen Sie die konfigurierte Tabelle aus.

4. Wählen Sie auf der Detailseite der konfigurierten Tabelle die Option Analyseregel konfigurieren aus.
5. Wählen Sie unter Schritt 1: Analyseregeltyp auswählen und unter Analyseregeltyp die Option Benutzerdefiniert aus.
6. Wählen Sie unter Erstellungsmethode die Option Geführter Ablauf und dann Weiter aus.
7. Wählen Sie unter Schritt 2: Analysekontrollen angeben für Direkte Analysekontrollen eine Option aus, die Ihrem Ziel entspricht.

Ihr Ziel	Empfohlene Aktion
Überprüfen Sie jede neue Analyse, bevor sie auf dieser konfigurierten Tabelle ausgeführt werden darf	1. Wählen Sie unter Analysevorlagen, die ausgeführt werden dürfen, die Option Analysevorlage hinzufügen aus. 2. Wählen Sie die entsprechende Vorlage für Zusammenarbeit und Analyse aus den Drop-down-Listen aus. 3. Wählen Sie Weiter aus.
Erlauben Sie bestimmten Mitarbeitern, Analysen eines ausgewählten Typs ohne Überprüfung anhand dieser Tabelle durchzuführen	1. Unter Analysetyp a. Wählen Sie Beliebige Abfrage aus, um jede Abfrage zuzulassen, die von AWS-Konto Ihnen angegeben wurde. b. Wählen Sie Beliebige Abfrage, um jeden Job zuzulassen, der von AWS-Konto Ihnen angegeben wurde. 2. Wählen Sie unter AWS-Konten Erlaubt, jede Analyse zu erstellen die Option Hinzufügen aus AWS-Konto. 3. Geben Sie eine ID ein AWS-Konto oder wählen Sie eine AWS-Konto ID aus der Drop-down-Liste aus. 4. (Optional) Wählen Sie Weitere hinzufügen AWS-Konto, um eine weitere AWS-Konto hinzuzufügen.

Ihr Ziel	Empfohlene Aktion
	5. Wählen Sie Weiter aus.

8. Unter Schritt 3: Geben Sie die Kontrollen für Analyseergebnisse an,
- Beachten Sie bei Kontrollen für Auftragsergebnisse, dass keine zusätzlichen Ergebniskontrollen unterstützt werden.
 - Wählen Sie unter Steuerelemente für Abfrageergebnisse für Spalten, die in der Ausgabe nicht zulässig sind, je nach Ziel die Spalten aus, die in der Abfrageausgabe zulässig sein sollen.

Ihr Ziel	Empfohlene Aktion
Lassen Sie zu, dass alle Spalten in Abfrageausgaben zurückgegeben werden	1. Wählen Sie Keine 2. Fahren Sie mit Zusätzliche Analysen fort, die auf die Ausgabe angewendet wurden.
Nicht zulassen, dass bestimmte Spalten in Abfrageausgaben zurückgegeben werden	1. Wählen Sie Benutzerdefinierte Liste 2. Wählen Sie unter Unzulässige Spalten angeben die Spalten aus, die Sie aus den Abfrageausgaben entfernen möchten.

- Wählen Sie für Zusätzliche Analysen, die auf die Ausgabe angewendet werden, je nach Ihrem Ziel aus, ob zusätzliche Analysen auf die Abfrageausgabe angewendet werden können.

Ihr Ziel	Empfohlene Option
• Erlaube nur direkte Abfragen für diese Tabelle. • Verweigern Sie die Ausführung zusätzlicher Analysen anhand von Abfrageergebnissen.	Nicht erlaubt

Ihr Ziel	Empfohlene Option
Die Tabelle kann nur für direkte Abfragen verwendet werden.	
Direkte Abfragen und zusätzliche Analysen in dieser Tabelle sind zulässig, aber nicht erforderlich.	Sobald Sie die Details auf dieser Seite überprüft haben, klicken Sie auf
- Erfordern Sie, dass die Tabelle nur in direkten Abfragen verwendet werden kann, die mit einer der erforderlichen zusätzlichen Analysen verarbeitet werden. - Direkte Abfragen für diese Tabelle müssen weiter verarbeitet werden, bevor sie zurückgegeben werden können.	Erforderlich

d. Wählen Sie Weiter aus.

9. (Optional) Stellen Sie unter Schritt 4: Differentiellen Datenschutz einrichten fest, ob Sie den differenziellen Datenschutz ein- oder ausschalten möchten.

Differential Privacy ist eine mathematisch erprobte Technik, mit der Sie Ihre Daten vor Reidentifikationsangriffen schützen können.

 Note

AWS Clean Rooms Differential Privacy ist nur für Kollaborationen verfügbar, bei denen die Daten in Amazon S3 gespeichert sind.

Wählen Sie für Differential Privacy je nach Ihrem Ziel aus, ob Sie den differenziellen Datenschutz ein- oder ausschalten möchten.

Ihr Ziel	Empfohlene Aktion
Sie benötigen keinen Schutz vor Versuchen, sich erneut zu identifizieren	- Wählen Sie Deaktivieren aus. - Wählen Sie Weiter aus.

Ihr Ziel	Empfohlene Aktion
Ihre Tabelle enthält keine Daten auf Benutzerebene	
- Sie benötigen Schutz vor Versuchen, sich erneut zu identifizieren - Ihre Tabelle enthält Daten auf Benutzerebene	- Wählen Sie Turn on (Einschalten) aus. - Wählen Sie die Spalte Benutzer-ID aus, die die eindeutige Kennung Ihrer Benutzer enthält, z. B. die <code>user_id</code> Spalte, deren Privatsphäre Sie schützen möchten. Um den differenziellen Datenschutz für zwei oder mehr Tabellen in einer Kollaboration zu aktivieren, müssen Sie in beiden Analyseregeln dieselbe Spalte wie die Benutzer-ID-Spalte konfigurieren, um eine konsistente Definition von Benutzern in allen Tabellen zu gewährleisten. Im Falle einer Fehlkonfiguration erhält das Mitglied, das Abfragen durchführen kann, eine Fehlermeldung, dass zwei Spalten zur Auswahl stehen, um die Anzahl der Benutzerbeiträge (z. B. die Anzahl der von einem Nutzer getätigten Anzeigenimpressionen) während der Ausführung der Abfrage zu berechnen. - Wählen Sie Weiter aus.

10. Überprüfen Sie unter Schritt 5: Überprüfen und konfigurieren die Auswahlen, die Sie für die vorherigen Schritte getroffen haben, bearbeiten Sie sie gegebenenfalls und wählen Sie dann Analyseregeln konfigurieren aus.

Es wird eine Bestätigungsmeldung angezeigt, dass Sie erfolgreich eine benutzerdefinierte Analyseregeln für die Tabelle konfiguriert haben.

Analyseregel zu einer Tabelle hinzufügen (JSON-Editor)

Das folgende Verfahren zeigt, wie Sie mithilfe der JSON-Editor-Option in der AWS Clean Rooms Konsole eine Analyseregel zu einer Tabelle hinzufügen.

Note

Konfigurierte Tabellen, die Nicht-S3-Datenquellen verwenden, unterstützen nur [benutzerdefinierte Analyseregeln](#).

Um einer Tabelle eine Aggregations-, Liste- oder benutzerdefinierte Analyseregel hinzuzufügen (JSON-Editor)

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Tables (Tabellen) aus.
3. Wählen Sie die konfigurierte Tabelle aus.
4. Wählen Sie auf der Detailseite der konfigurierten Tabelle die Option Analyseregel konfigurieren aus.
5. Wählen Sie unter Schritt 1: Analyseregeltyp auswählen unter Analyseregeltyp entweder die Option Aggregation, Liste oder Benutzerdefiniert aus.
6. Wählen Sie unter Erstellungsmethode die Option JSON-Editor und dann Weiter aus.
7. Unter Schritt 2: Steuerelemente angeben können Sie wählen, ob Sie eine Abfragestruktur (Vorlage einfügen) oder eine Datei einfügen möchten (Aus Datei importieren).

Wenn Sie folgendes auswählen ...	Dann...
Vorlage einfügen	1. Geben Sie die Parameter für die ausgewählte Analyseregel in der Definition der Analyseregeln an. 2. Sie können Strg + Leertaste drücken, um die automatische Vervollständigung zu aktivieren.

Wenn Sie folgendes auswählen ...	Dann...
	Weitere Informationen zu den Regelparametern für die Aggregationsanalyse finden Sie unter. Regel für die Aggregationsanalyse — Steuerelemente abfragen Weitere Informationen zu Regelparametern für Listenanalysen finden Sie unter Regel für die Listenanalyse — Steuerelemente abfragen.
Aus einer Datei importieren	1. Wählen Sie Ihre JSON-Datei von Ihrem lokalen Laufwerk aus. 2. Klicken Sie auf Open. In der Analyseregeldefinition wird die Analyseregel aus der hochgeladenen Datei angezeigt.

8. Wählen Sie Weiter aus.
9. Überprüfen Sie unter Schritt 3: Überprüfen und konfigurieren die Auswahlen, die Sie für die vorherigen Schritte getroffen haben, bearbeiten Sie sie gegebenenfalls und wählen Sie dann Analyseregel konfigurieren aus.

Sie erhalten eine Bestätigungsnachricht, dass Sie eine Analyseregel für die Tabelle erfolgreich konfiguriert haben.

Nächste Schritte

Nachdem Sie eine Analyseregel für Ihre konfigurierte Tabelle konfiguriert haben, können Sie:

- [Ordnen Sie eine konfigurierte Tabelle einer Kollaboration zu](#)
- [Fragen Sie die Datentabellen](#) ab (als Mitglied, das Abfragen durchführen kann)

Eine konfigurierte Tabelle einer Kooperation zuordnen

Nachdem Sie eine konfigurierte Tabelle erstellt und ihr eine Analyseregeln hinzugefügt haben, können Sie sie einer Kollaboration zuordnen und AWS Clean Rooms eine Servicerolle für den Zugriff auf Ihre AWS Glue Tabellen zuweisen.

Note

Diese Servicerolle hat Berechtigungen für die Tabellen. Die Servicerolle kann nur übernommen werden AWS Clean Rooms , wenn zulässige Abfragen im Namen des Mitglieds ausgeführt werden, das Abfragen durchführen kann. Keine Kollaborationsmitglieder (außer dem Datenbesitzer) haben Zugriff auf die zugrunde liegenden Tabellen in der Kollaboration. Der Datenbesitzer kann den differenziellen Datenschutz aktivieren, um seine Tabellen für Abfragen durch andere Mitglieder verfügbar zu machen.

Budget für den Datenzugriff

Wenn Sie eine konfigurierte Tabelle zuordnen, können Sie ein Datenzugriffsbudget anwenden. Ein Datenzugriffsbudget steuert, wie oft eine Tabelle für Abfragen, Jobs und ML-Eingabekanäle in einer Kollaboration verwendet werden kann. Diese Budgets helfen Unternehmen dabei, die Ressourcennutzung zu verwalten und die Kosten zu kontrollieren, indem sie die Tabellennutzung einschränken.

Jedes Mal, wenn eine Tabelle in einer Abfrage, einem Job oder einem ML-Eingabekanal verwendet wird, wird das Budget für diese Tabelle um eins reduziert. Wenn das Budget Null erreicht, kann die Tabelle nicht in SQL-Abfragen, Pyspark-Jobs oder als Teil von aus der Tabelle abgeleiteten ML-Eingabekanälen verwendet werden.

Sie können ein Budget pro Zeitraum einrichten, das regelmäßig aktualisiert wird, ein Lebenszeitbudget für die Gesamtnutzung oder beides. Standardmäßig ist die Tabellennutzung unbegrenzt.

- Budget pro Zeitraum — Eine erneuerbare Zuweisung, die begrenzt, wie oft diese Tabelle innerhalb eines bestimmten Zeitraums verwendet werden kann. Sie können den Zeitraum auf täglich, wöchentlich oder monatlich festlegen. Dieses Budget kann so eingestellt werden, dass es täglich, wöchentlich oder monatlich automatisch aktualisiert wird.

- **Lebenszeitbudget** — Eine laufende Zuweisung, die begrenzt, wie oft diese Tabelle insgesamt verwendet werden kann.

Ordnen Sie eine konfigurierte Tabelle zu

In den folgenden Themen wird beschrieben, wie Sie mithilfe der AWS Clean Rooms Konsole eine konfigurierte Tabelle zuordnen und ein Datenzugriffsbudget auf eine Zusammenarbeit anwenden.

Informationen dazu, wie Sie Ihre konfigurierten Tabellen mit der Kollaboration verknüpfen AWS SDKs, finden Sie in der [AWS Clean Rooms API-Referenz](#).

Schritt 1: Erfüllen der Voraussetzungen

Um eine konfigurierte Tabelle zuzuordnen, müssen Sie die folgenden Voraussetzungen erfüllen:

- Eine AWS Glue Tabelle, die auf einen Amazon S3 S3-Ordnerspeicherort verweist (keine einzelne Datei)
- Für verschlüsselte AWS Glue Tabellen:
 - Eine Servicerolle mit Berechtigungen zur Verwendung von AWS KMS Schlüsseln zum Entschlüsseln von Tabellen AWS Glue
 - Für AWS KMS-verschlüsselte Amazon S3 S3-Datensätze: Die Servicerolle muss auch über Berechtigungen verfügen, um den AWS KMS Schlüssel zum Entschlüsseln von Amazon S3 S3-Daten zu verwenden

Informationen zur Konfiguration der Verschlüsselung finden Sie unter [Verschlüsselung einrichten](#) [AWS Glue im Entwicklerhandbuch](#).AWS Glue

So verifizieren Sie den Speicherort Ihrer AWS Glue Tabelle:

1. Öffnen Sie die AWS Glue Konsole unter <https://console.aws.amazon.com/glue/>
2. Sehen Sie sich Ihre Tabellendetails an und vergewissern Sie sich, dass der Speicherort auf einen S3-Ordner verweist

Schritt 2: Ordnen Sie eine konfigurierte Tabelle zu

Um eine konfigurierte Tabelle zuzuordnen

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie die Methode, um Ihre Tabelle zu verknüpfen:
 - a. Auf der Detailseite der konfigurierten Tabelle:
 - i. Wählen Sie im linken Navigationsbereich Tables (Tabellen) aus.
 - ii. Wählen Sie die konfigurierte Tabelle aus.
 - iii. Wählen Sie auf der Detailseite der konfigurierten Tabelle die Option Mit Kollaboration verknüpfen aus.
 - iv. Wählen Sie im Dialogfeld Tabelle mit Kollaboration verknüpfen die Option Kollaboration aus der Dropdownliste aus.
 - b. Auf der Detailseite der Zusammenarbeit:
 - i. Wählen Sie im linken Navigationsbereich Collaborations aus.
 - ii. Wählen Sie die Zusammenarbeit aus.
 - iii. Wählen Sie auf der Registerkarte Tabellen die Option Tabelle zuordnen aus.
3. Führen Sie auf der Seite Tabelle zuordnen einen der folgenden Schritte aus:
 - Wählen Sie eine vorhandene konfigurierte Tabelle — Wählen Sie aus der Dropdownliste den Namen der konfigurierten Tabelle aus, die Sie der Kollaboration zuordnen möchten.
 - Neue Tabelle konfigurieren — Wählen Sie Neue Tabelle konfigurieren und folgen Sie den Anweisungen auf der Seite Neue Tabelle konfigurieren.
 - Schema und Analyseregeln für die konfigurierte Tabelle anzeigen — Aktivieren Sie die Option Schema- und Analyseregeln anzeigen.
4. Einzelheiten zur Tabellenverknüpfung finden Sie unter
 - a. Geben Sie einen Namen für die zugehörige Tabelle ein.

Sie können den Standardnamen verwenden oder diese Tabelle umbenennen.
 - b. (Optional) Geben Sie eine Beschreibung der Tabelle ein.

Die Beschreibung hilft beim Schreiben von Abfragen.

5. Geben Sie die Dienstzugriffsberechtigungen an, indem Sie entweder Neue Servicerolle erstellen und verwenden oder Bestehende Servicerolle verwenden auswählen.

 Note

Wenn Sie eine konfigurierte Tabelle verknüpfen, die von Amazon Athena unterstützt wird, wählen Sie einen vorhandenen Servicerollennamen aus der Dropdownliste aus. Stellen Sie sicher, dass die Servicerolle über IAM- und, falls erforderlich, Lake Formation Formation-Berechtigungen für den Datensatz verfügt.

Wenn Sie folgendes auswählen ...	Dann...
Erstellen und verwenden Sie eine neue Servicerolle	• AWS Clean Rooms erstellt eine Servicerolle mit der erforderlichen Richtlinie für diese Tabelle. • Der Standardname der Servicerolle lautet <code>cleanrooms-<timestamp></code> • Sie benötigen die erforderlichen Berechtigungen, um Rollen zu erstellen und Richtlinien anzuhängen. • Wenn Ihre Eingabedaten verschlüsselt sind, können Sie Diese Daten sind mit einem KMS-Schlüssel verschlüsselt auswählen und dann einen eingeben AWS KMS key , der zur Entschlüsselung Ihrer Dateneingabe verwendet wird.

Wenn Sie folgendes auswählen ...	Dann...
Verwenden Sie eine vorhandene Servicerolle	1. Wählen Sie einen vorhandenen Servicero llenamen aus der Dropdownliste aus. Die Liste der Rollen wird angezeigt, wenn Sie berechtigt sind, Rollen aufzulisten. Wenn Sie nicht berechtigt sind, Rollen aufzulisten, können Sie den Amazon- Ressourcennamen (ARN) der Rolle eingeben, die Sie verwenden möchten. 2. Rufen Sie die Servicerolle auf, indem Sie auf den externen Link In IAM anzeigen klicken. Wenn keine vorhandenen Servicero llen vorhanden sind, ist die Option „Eine bestehende Servicerolle verwenden“ nicht verfügbar. Versucht standardmäßig AWS Clean Rooms nicht, die bestehende Rollenric htlinie zu aktualisieren, um die erforderl ichen Berechtigungen hinzuzufügen. 3. (Optional) Aktivieren Sie das Kontrollk ästchen Eine vorkonfigurierte Richtlinie mit den erforderlichen Berechtigungen zu dieser Rolle hinzufügen, um der Rolle die erforderlichen Berechtigungen hinzuzufü gen. Sie müssen über Berechtigungen verfügen, um Rollen zu ändern und Richtlinien zu erstellen.

 Note

- AWS Clean Rooms erfordert Berechtigungen für Abfragen gemäß den Analyseregeln. Weitere Informationen zu Berechtigungen für AWS Clean Rooms finden Sie unter [AWS verwaltete Richtlinien für AWS Clean Rooms](#).
- Wenn die Rolle nicht über ausreichende Berechtigungen für verfügt AWS Clean Rooms, erhalten Sie eine Fehlermeldung, dass die Rolle nicht über ausreichende Berechtigungen für verfügt AWS Clean Rooms. Die Rollenrichtlinie muss hinzugefügt werden, bevor Sie fortfahren können.
- Wenn Sie die Rollenrichtlinie nicht ändern können, erhalten Sie eine Fehlermeldung, dass die Richtlinie für die Servicerolle nicht gefunden werden AWS Clean Rooms konnte.

6. Wenn Sie konfigurierte Tabellenzuordnungs-Tags für die konfigurierte Tabellenzuordnungsressource aktivieren möchten, wählen Sie Neues Tag hinzufügen und geben Sie dann das Schlüssel-Wert-Paar ein.
7. Wählen Sie Weiter aus.
8. Wählen Sie auf der Seite „Regel für die Kollaborationsanalyse konfigurieren“ eine der folgenden Optionen aus:
 - Ja, jetzt eine Regel für die Kollaborationsanalyse erstellen — Ordnet Ihre Tabelle dieser Kollaboration zu und erstellt eine Regel für die Kollaborationsanalyse
 - Nein, ich werde später eine Regel für die Kollaborationsanalyse erstellen — Ordnet Ihre Tabelle nur dieser Kollaboration zu. Sie können später eine Regel für die Kollaborationsanalyse erstellen.
9. Wenn Sie Ja, erstellen Sie jetzt eine Regel für die Kollaborationsanalyse auswählen, wählen Sie in der Dropdownliste unter Ergebnislieferung die Option Mitglieder, die Ergebnisse für die Abfrageausgabe erhalten dürfen.
10. Wählen Sie Weiter aus.
11. Wählen Sie auf der Seite Datenzugriffsbudget hinzufügen unter Konfiguration des Datenzugriffsbudgets eine der folgenden Optionen aus:

- Ja, fügen Sie jetzt ein Datenzugriffsbudget hinzu — Ordnet Ihre Tabelle dieser Kollaboration zu und fügt ein Datenzugriffsbudget hinzu. Sie können entweder ein Periodenbudget, ein Lebenszeitbudget oder beides auswählen.
- Nein, ich werde später ein Budget für den Datenzugriff hinzufügen — Ordnet Ihre Tabelle nur dieser Kollaboration zu. Sie können später ein Budget für den Datenzugriff hinzufügen.

Wenn Sie Nein, ich füge später ein Datenzugriffsbudget hinzu auswählen, fahren Sie mit Schritt 15 fort.

12. Wenn Sie Ja, jetzt ein Datenzugriffsbudget hinzufügen wählen, wählen Sie eine der folgenden Budgetkonfigurationen aus:

Nur Budget pro Periode	Nur lebenslanges Budget	Budgets sowohl pro Zeitraum als auch auf Lebenszeit
1. Lassen Sie die Option Budget pro Zeitraum hinzufügen ausgewählt. 2. Geben Sie einen Budgetbetrag pro Periode zwischen 1 und 1.000.000 ein. 3. Wählen Sie für Zeitraum die Option Täglich, Wöchentlich oder Monatlich aus. 4. (Optional) Lassen Sie die Option Budget wöchentlich automatisch aktualisieren aktiviert, um die Zuteilung zu erneuern. 5. Deaktivieren Sie die Option Lebenszeitbudget hinzufügen.	1. Deaktivieren Sie die Option Budget pro Zeitraum hinzufügen. 2. Wählen Sie Lebenszeitbudget hinzufügen aus. 3. Geben Sie einen Betrag für das Lifetime-Budget zwischen 1 und 1.000.000 ein.	1. Lassen Sie die Option Budget pro Zeitraum hinzufügen ausgewählt. 2. Geben Sie einen Budgetbetrag pro Periode zwischen 1 und 1.000.000 ein. 3. Wählen Sie für Zeitraum die Option Täglich, Wöchentlich oder Monatlich aus. 4. Lassen Sie die Option Budget wöchentlich automatisch aktualisieren aktiviert. 5. Wählen Sie Lebenszeitbudget hinzufügen aus. 6. Geben Sie einen Betrag für das Lifetime-Budget

Nur Budget pro Periode	Nur lebenslanges Budget	Budgets sowohl pro Zeitraum als auch auf Lebenszeit
		zwischen 1 und 1.000.000 ein.

13. Überprüfen Sie Ihre Auswahl unter Übersicht über das Datenzugriffsbudget.

Example Beispiel

Wenn Sie beispielsweise einen Budgetbetrag pro Zeitraum von 1.000 ausgewählt haben, den Zeitraum auf Wöchentlich setzen, das Kontrollkästchen Budget automatisch wöchentlich aktualisieren aktiviert lassen und das Lebenszeitbudget auf 1.000.000 festgelegt haben, wird in der Access-Budgetübersicht die folgende Meldung angezeigt: Jede Woche kann diese Tabelle bis zu 1.000 Mal für die Ausführung von Abfragen oder Aufträgen verwendet werden. Dieses Budget ist so eingestellt, dass es jeden Sonntag um 00:00 Uhr UTC automatisch aktualisiert wird. Es wird so lange aktualisiert, bis diese Tabelle ihr Lebenszeitbudget von 1.000.000 Benutzern erreicht hat.

14. (Optional) Wenn Sie Budget-Tags für den Datenzugriff für die Access-Budgetressource aktivieren möchten, wählen Sie Neues Tag hinzufügen aus und geben Sie ein Schlüssel- und Wertepaar ein.
15. Wählen Sie Weiter aus.
16. Überprüfen Sie die Informationen auf der Seite Überprüfen und erstellen.
 - a. Wenn Sie Abschnitte bearbeiten müssen, wählen Sie Bearbeiten.
 - b. Bearbeiten Sie Ihre Konfigurationen und wählen Sie dann Weiter.
17. Wählen Sie Tabelle zuordnen aus.

Schritt 3: Die nächsten Schritte

Nachdem Sie Ihre konfigurierte Datentabelle der Kollaboration zugeordnet haben, können Sie:

- [Fügen Sie der konfigurierten Tabelle eine Regel für die Kollaborationsanalyse](#) hinzu
- [Bearbeiten Sie die Kollaboration](#), wenn Sie der Kollaborationsersteller sind
- [Fragen Sie die Datentabellen](#) ab (als Mitglied, das Abfragen durchführen kann)

Ein Datenzugriffsbudget konfigurieren

Ein Mitarbeiter kann ein Datenzugriffsbudget anzeigen, hinzufügen, bearbeiten und löschen, um festzulegen, wie oft eine Tabelle in einem Workflow verwendet werden kann. Verwenden Sie diese Budgets, um Daten und Kosten zu verwalten.

Jedes Mal, wenn eine Tabelle abgefragt oder ein Job unter Verwendung eines von einer Tabelle abgeleiteten ML-Eingabekanals ausgeführt wird, wird das Budget für diese Tabelle um eins reduziert. Wenn das Budget Null erreicht, kann die Tabelle nicht abgefragt werden, und ML-Jobs können nicht mit aus der Tabelle abgeleiteten ML-Eingabekanälen ausgeführt werden.

Sie können ein Budget pro Zeitraum einrichten, das regelmäßig aktualisiert wird, ein Lebenszeitbudget für die Gesamtnutzung oder beides. Standardmäßig ist die Tabellennutzung unbegrenzt.

- **Budget pro Zeitraum** — Eine erneuerbare Zuweisung, die begrenzt, wie oft diese Tabelle innerhalb eines bestimmten Zeitraums verwendet werden kann. Sie können den Zeitraum auf täglich, wöchentlich oder monatlich festlegen. Dieses Budget kann so eingestellt werden, dass es täglich, wöchentlich oder monatlich automatisch aktualisiert wird.
- **Lebenszeitbudget** — Eine laufende Zuweisung, die begrenzt, wie oft diese Tabelle insgesamt verwendet werden kann.

Topics

- [Ein Budget für den Datenzugriff anzeigen](#)
- [Hinzufügen eines Datenzugriffsbudgets zu einer vorhandenen verknüpften Tabelle](#)
- [Ein Datenzugriffsbudget bearbeiten](#)
- [Löschen eines Datenzugriffsbudgets](#)

Ein Budget für den Datenzugriff anzeigen

Sie können ein Datenzugriffsbudget auf der Registerkarte „Tabellen“ oder auf der Seite mit den Tabellendetails anzeigen.

Um ein Datenzugriffsbudget anzuzeigen

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.

2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Zusammenarbeit aus.
4. Wählen Sie die Registerkarte Tables (Tabellen).
5. Führen Sie eine der folgenden Aktionen aus:
 - Wählen Sie in der Spalte Verbleibendes Budget für den Datenzugriff das Budget aus, um die Details anzuzeigen.
 - Wählen Sie eine Tabelle aus und scrollen Sie auf der Seite mit den Tabellendetails nach unten, um den Abschnitt mit den Budgetdetails für den Datenzugriff anzuzeigen.

Hinzufügen eines Datenzugriffsbudgets zu einer vorhandenen verknüpften Tabelle

Als Mitglied einer Kollaboration können Sie einer vorhandenen verknüpften Tabelle ein Datenzugriffsbudget hinzufügen.

Um einer vorhandenen verknüpften Tabelle ein Datenzugriffsbudget hinzuzufügen

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Zusammenarbeit aus.
4. Wählen Sie die Registerkarte Tables (Tabellen).
5. Wählen Sie das Optionsfeld neben der Tabelle aus, zu der Sie ein Datenzugriffsbudget hinzufügen möchten.
6. Wählen Sie in der Dropdownliste Aktionen unter Datenzugriffsbudget die Option Hinzufügen aus (falls noch kein Budget vorhanden ist).
7. Wählen Sie eine der folgenden Budgetkonfigurationen aus:

Nur Budget pro Periode	Nur lebenslanges Budget	Budgets sowohl pro Zeitraum als auch auf Lebenszeit
1. Lassen Sie die Option Budget pro Zeitraum hinzufügen ausgewählt.	1. Deaktivieren Sie die Option Budget pro Zeitraum hinzufügen.	1. Lassen Sie die Option Budget pro Zeitraum hinzufügen ausgewählt.

Nur Budget pro Periode	Nur lebenslanges Budget	Budgets sowohl pro Zeitraum als auch auf Lebenszeit
2. Geben Sie einen Budgetbetrag pro Periode zwischen 1 und 1.000.000 ein. 3. Wählen Sie für Zeitraum die Option Täglich, Wöchentlich oder Monatlich aus. 4. (Optional) Lassen Sie die Option Budget wöchentlich automatisch aktualisieren aktiviert, um die Zuteilung zu erneuern. 5. Deaktivieren Sie die Option Lebenszeitbudget hinzufügen.	2. Wählen Sie Lebenszeitbudget hinzufügen aus. 3. Geben Sie einen Betrag für das Lifetime-Budget zwischen 1 und 1.000.000 ein.	2. Geben Sie einen Budgetbetrag pro Periode zwischen 1 und 1.000.000 ein. 3. Wählen Sie für Zeitraum die Option Täglich, Wöchentlich oder Monatlich aus. 4. Lassen Sie die Option Budget wöchentlich automatisch aktualisieren aktiviert. 5. Wählen Sie Lebenszeitbudget hinzufügen aus. 6. Geben Sie einen Betrag für das Lifetime-Budget zwischen 1 und 1.000.000 ein.

8. Überprüfen Sie Ihre Auswahl unter Übersicht über das Datenzugriffsbudget.

9. Example Beispiel

Wenn Sie beispielsweise einen Budgetbetrag pro Zeitraum von 1.000 ausgewählt haben, den Zeitraum auf Wöchentlich setzen, das Kontrollkästchen Budget automatisch wöchentlich aktualisieren aktiviert lassen und das Lebenszeitbudget auf 1.000.000 festgelegt haben, wird in der Access-Budgetübersicht die folgende Meldung angezeigt: Jede Woche kann diese Tabelle bis zu 1.000 Mal für die Ausführung von Abfragen oder Aufträgen verwendet werden. Dieses Budget ist so eingestellt, dass es jeden Sonntag um 00:00 Uhr UTC automatisch aktualisiert wird. Es wird so lange aktualisiert, bis diese Tabelle ihr Lebenszeitbudget von 1.000.000 Benutzern erreicht hat.

10. (Optional) Wenn Sie Budget-Tags für den Datenzugriff für die Access-Budgetressource aktivieren möchten, wählen Sie Neues Tag hinzufügen aus und geben Sie ein Schlüssel- und Wertepaar ein.

11. Wählen Sie Datenzugriffsbudget hinzufügen aus.

Ein Datenzugriffsbudget bearbeiten

Als Collaboration-Mitglied können Sie das Datenzugriffsbudget bearbeiten. Wenn Sie ein Datenzugriffsbudget bearbeiten, wird der aktuelle Budgetsaldo zurückgesetzt.

Sie können ein Datenzugriffsbudget auf der Registerkarte Tabellen oder auf der Seite mit den Tabellendetails bearbeiten.

Tables tab

Um ein Datenzugriffsbudget auf der Registerkarte „Tabellen“ zu bearbeiten

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Zusammenarbeit aus.
4. Wählen Sie die Registerkarte Tables (Tabellen).
5. Wählen Sie das Optionsfeld neben der Tabelle aus, die Sie bearbeiten möchten.
6. Wählen Sie in der Dropdownliste Aktionen unter Datenzugriffsbudget die Option Bearbeiten aus.
7. Aktualisieren Sie auf der Seite Datenzugriffsbudget bearbeiten die Budgetinformationen pro Zeitraum oder Lebenszeitbudget.
8. Sehen Sie sich die Übersicht über das Datenzugriffsbudget an, um zu überprüfen, ob die von Ihnen vorgenommenen Änderungen korrekt sind.
9. Wählen Sie Änderungen speichern aus.

Table details page

Um ein Datenzugriffsbudget auf der Seite mit den Tabellendetails zu bearbeiten

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Zusammenarbeit aus.

4. Wählen Sie die Registerkarte Tables (Tabellen).
5. Wählen Sie eine -Tabelle aus.
6. Scrollen Sie auf der Seite mit den Tabellendetails nach unten zum Abschnitt Details zum Datenzugriffsbudget.
7. Wählen Sie in der Dropdown-Liste Aktionen die Option Bearbeiten aus.
8. Aktualisieren Sie auf der Seite Datenzugriffsbudget bearbeiten die Budgetinformationen pro Zeitraum oder Lebenszeitbudget.
9. Wählen Sie Änderungen speichern aus.

Löschen eines Datenzugriffsbudgets

Sie können ein Datenzugriffsbudget auf der Registerkarte Tabellen oder auf der Seite mit den Tabellendetails löschen.

Tables tab

Um ein Datenzugriffsbudget auf der Registerkarte „Tabellen“ zu löschen

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Zusammenarbeit aus.
4. Wählen Sie die Registerkarte Tables (Tabellen).
5. Wählen Sie das Optionsfeld neben der Tabelle aus, die Sie löschen möchten.
6. Wählen Sie in der Dropdownliste Aktionen unter Datenzugriffsbudget die Option Löschen aus.

Important

Sie können diese Aktion nicht rückgängig machen und Ihr Datenzugriffsbudget wird auf unbegrenzt zurückgesetzt.

7. Wenn Sie sicher sind, dass Sie das Datenzugriffsbudget löschen möchten, wählen Sie Löschen.

Table details page

Um ein Datenzugriffsbudget von der Tabellendetailseite zu löschen

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Zusammenarbeit aus.
4. Wählen Sie die Registerkarte Tables (Tabellen).
5. Wählen Sie eine -Tabelle aus.
6. Scrollen Sie auf der Seite mit den Tabellendetails nach unten zum Abschnitt Details zum Datenzugriffsbudget.
7. Wählen Sie in der Dropdown-Liste Aktionen die Option Löschen.



Important

Sie können diese Aktion nicht rückgängig machen und Ihr Datenzugriffsbudget wird auf unbegrenzt zurückgesetzt.

8. Wenn Sie sicher sind, dass Sie das Datenzugriffsbudget löschen möchten, wählen Sie Löschen.

Eine Regel für die Kollaborationsanalyse zu einer konfigurierten Tabelle hinzufügen

Mit der Regel für die Kollaborationsanalyse können Sie Kontrollen angeben, die für diese Zusammenarbeit spezifisch sind. Diese Steuerelemente bestimmen zusammen mit der konfigurierten Tabellenanalyseregel, wie diese Tabelle innerhalb dieser Zusammenarbeit analysiert werden kann.

Sie fügen einer konfigurierten Tabelle eine Regel für die Kollaborationsanalyse hinzu, nachdem Sie [eine konfigurierte Tabelle erstellt](#), [eine Analyseregel hinzugefügt](#) und [sie einer Kollaboration zugeordnet](#) haben. Sie müssen eine Regel für die Kollaborationsanalyse hinzufügen, wenn die Tabelle so konfiguriert ist, dass sie direkte Analysen unterstützt oder zusätzliche Analysen ermöglicht.

- **Direkte Analyse** — Die Tabelle kann in Abfragen verwendet werden, mit denen sie direkt analysiert wird. Zum Beispiel in einer Abfrage, die eine aggregierte Messanalyse oder eine Liste von Identifikatoren zur Aktivierung ausgibt.
- **Zusätzliche Analyse** — Die Tabelle kann neben Abfragen, mit denen sie direkt analysiert wird, auch als Eingabe für zusätzliche Analysen verwendet werden. Die Tabelle kann beispielsweise in einer Abfrage verwendet werden, die als Ausgangswert für ein Lookalike-ML-Modell dient, oder als ML-Eingabekanal für ein benutzerdefiniertes ML-Modell.

Um die Regel für die Kollaborationsanalyse zu einer Tabelle hinzuzufügen

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#) mit Ihrem AWS-Konto (falls Sie dies noch nicht getan haben).
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus.
4. Sehen Sie sich auf der Registerkarte Tabellen unter Von Ihnen zugeordnete Tabellen die konfigurierte Tabelle an, die Sie der Kollaboration zugeordnet haben.
 - Wenn der Status „Direkte Analyse“ oder „Zusätzliche Analyse“ den Status „Bereit“ hat, kann die Tabelle abgefragt werden.
 - Wenn der Status Direkte Analyse oder Zusätzliche Analyse den Status Nicht bereit hat, wählen Sie den Status aus, und klicken Sie dann im Dialogfeld auf Konfigurieren.
5. Erweitern Sie auf der Seite Regel für die Kollaborationsanalyse konfigurieren die Option Regel für die Analyse konfigurierter Tabellen anzeigen, um die Details anzuzeigen.
6. Wählen Sie unter Zulässige zusätzliche Analysen die Option aus, die Ihrem Ziel entspricht.

Ihr Ziel	Empfohlene Option
Lassen Sie alle zusätzlichen Analysen auf dem Tisch zu.	Alle
Erlauben Sie nur zusätzliche Analysen in der Tabelle, die von einem bestimmten Mitglied erstellt wurden.	Beliebige von bestimmten Mitgliedern
Erlaube nur spezifische Analysen in der Tabelle.	Benutzerdefinierte Liste

7. Geben Sie für die Bereitstellung von Ergebnissen in der Dropdownliste an, wer Ergebnisse von Mitgliedern erhalten kann, denen der Empfang von Ergebnissen für die Abfrageausgabe gestattet ist.
8. Wählen Sie Analyseregeln konfigurieren aus.

Konfiguration der differenzierten Datenschutzrichtlinie (optional)

Note

AWS Clean Rooms Differential Privacy ist nur für Kollaborationen verfügbar, bei denen die Daten in Amazon S3 gespeichert sind.

Dieses Verfahren beschreibt den Prozess der Konfiguration der differenziellen Datenschutzrichtlinie in einer Kollaboration mithilfe der Option Guided Flow in der AWS Clean Rooms Konsole. Dies ist ein einmaliger Schritt für alle Tabellen mit differenziellem Datenschutz.

So konfigurieren Sie differenzielle Datenschutzeinstellungen (geführter Ablauf)

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus.
4. Wählen Sie auf der Kollaborationsseite auf der Registerkarte Tabellen die Option Differenzielle Datenschutzrichtlinie konfigurieren aus.
5. Wählen Sie auf der Seite Differentielle Datenschutzrichtlinie konfigurieren Werte für die folgenden Eigenschaften aus:
 - Budget für den Datenschutz
 - Aktualisieren Sie das Datenschutzbudget monatlich
 - Pro Abfrage wurde ein Rauschen hinzugefügt

Sie können die Standardwerte verwenden oder benutzerdefinierte Werte eingeben, die Ihren speziellen Anwendungsfall unterstützen. Nachdem Sie die Werte für das Datenschutzbudget und die pro Abfrage hinzugefügten Störungen ausgewählt haben, können Sie eine Vorschau des

resultierenden Dienstprogramms im Hinblick auf die Anzahl der Aggregationen anzeigen, die für alle Abfragen Ihrer Daten möglich sind.

6. Wählen Sie Konfigurieren aus.

Es wird eine Bestätigungsnachricht angezeigt, dass Sie die differenzielle Datenschutzrichtlinie für die Zusammenarbeit erfolgreich konfiguriert haben.

Nachdem Sie den differenziellen Datenschutz konfiguriert haben, können Sie:

- [Fragen Sie die Datentabellen](#) ab (als Mitglied, das Abfragen durchführen kann)
- [Kollaborationen](#) (wenn Sie die Kollaboration erstellt haben)

Differenzielle Nutzungsprotokolle zum Datenschutz anzeigen

Als Collaboration-Mitglied, das Daten mit Differential Privacy schützt, können Sie, nachdem Sie eine Collaboration mit Differential Privacy erstellt haben, die Nutzung des Datenschutzbudgets überwachen.

Um zu sehen, wie viele Aggregationen ausgeführt wurden und wie viel des Datenschutzbudgets aufgebraucht wurde

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus.
4. Wählen Sie die Registerkarte Tables (Tabellen).
5. Wählen Sie Nutzungsprotokolle anzeigen (blauer Text).
6. Sehen Sie sich die Nutzungsdetails an, einschließlich des Datenschutzbudgets und der Anzahl der bereitgestellten Funktionen.

Eine differenzierte Datenschutzrichtlinie bearbeiten

Nachdem Sie die differenzielle Datenschutzrichtlinie konfiguriert haben, können Sie sie jederzeit aktualisieren, um Ihren Datenschutzerfordernissen besser gerecht zu werden.

Um die differenzielle Datenschutzrichtlinie zu bearbeiten

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus.
4. Wählen Sie auf der Registerkarte Tabellen der Kollaborationsseite unter Von Ihnen zugeordnete Tabellen die Option Bearbeiten aus.
5. Wählen Sie auf der Seite Differentiellen Datenschutz bearbeiten neue Werte für die folgenden Eigenschaften aus:
 - **Datenschutzbudget** — Bewegen Sie den Schieberegler, um das Budget zu einem beliebigen Zeitpunkt während einer Zusammenarbeit entweder zu erhöhen oder zu verringern. Sie können das Budget nicht verringern, nachdem das Mitglied, das Abfragen durchführen kann, mit der Abfrage Ihrer Daten begonnen hat. Wenn das Datenschutzbudget erhöht wird, AWS Clean Rooms wird das vorhandene Budget weiter verwendet, bis es vollständig aufgebraucht ist, bevor das neu hinzugefügte Datenschutzbudget verwendet wird.
 - **Pro Abfrage hinzugefügtes Rauschen** — Bewegen Sie den Schieberegler, um das pro Abfrage hinzugefügte Rauschen zu einem beliebigen Zeitpunkt während einer Zusammenarbeit entweder zu erhöhen oder zu verringern.

Note

Mithilfe interaktiver Beispiele können Sie untersuchen, wie sich unterschiedliche Werte für Datenschutzbudget und hinzugefügtes Rauschen pro Abfrage auf die Anzahl der Aggregatfunktionen auswirken, die Sie ausführen können.

Sie können den Wert der Aktualisierung des Datenschutzbudgets nicht ändern. Um Ihre Auswahl zu ändern, müssen Sie die differenzielle Datenschutzrichtlinie löschen und eine neue erstellen.

6. Wählen Sie **Änderungen speichern** aus.

Sie erhalten eine Bestätigungsnachricht, dass Sie die differenzielle Datenschutzrichtlinie erfolgreich bearbeitet haben.

Löschen einer differenzierten Datenschutzrichtlinie

Sie können die differenzielle Datenschutzrichtlinie auf der Registerkarte Tabellen einer Kollaboration löschen.

Um die differenzielle Datenschutzrichtlinie zu löschen

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus.
4. Wählen Sie auf der Kollaborationsseite auf der Registerkarte Tabellen neben Differenzielle Datenschutzrichtlinie die Option Löschen aus.
5. Wenn Sie sicher sind, dass Sie die differenzielle Datenschutzrichtlinie löschen möchten, wählen Sie Löschen aus.

Nach dem Löschen einer differenzierten Datenschutzrichtlinie können Sie in dieser Richtlinie nicht mehr auf die Nutzungsprotokolle des Datenschutzbudgets zugreifen. Tabellen mit aktiviertem differenziellen Datenschutz können nicht abgefragt werden, wenn die differenzielle Datenschutzrichtlinie gelöscht wird.

Anzeige der berechneten differenziellen Datenschutzparameter

Benutzer mit Erfahrung im Bereich Differential Privacy können die berechneten differenziellen Datenschutzparameter auf der Registerkarte Analyse einer Kollaboration einsehen.

Um die berechneten differenziellen Datenschutzparameter einzusehen

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus.
4. Wählen Sie auf der Registerkarte Analyse im Abschnitt Ergebnisse die Option Berechnete differenzielle Datenschutzparameter anzeigen aus.

In der Tabelle Berechnete differenzielle Datenschutzparameter können Sie die Sensitivitätswerte von Aggregatfunktionen sehen. Diese Werte sind als der Höchstwert definiert, um den sich das Ergebnis einer Funktion ändern kann, wenn die Datensätze eines einzelnen Benutzers hinzugefügt, entfernt oder geändert werden. Die Liste enthält die folgenden unterschiedlichen Datenschutzparameter:

- Das Benutzerbeitragslimit (User Contribution Limit, UCL) ist die maximale Anzahl von Zeilen, die ein Benutzer zu einer SQL-Abfrage beigetragen hat. Wenn Sie beispielsweise die Gesamtzahl der übereinstimmenden Impressionen in einer bestimmten Kampagne zählen möchten, bei der jeder Nutzer mehrere Impressionen haben kann, muss AWS Clean Rooms Differential Privacy die Anzahl der Impressionen eines einzelnen Benutzers begrenzen, um sicherzustellen, dass die Berechnung des unterschiedlichen Datenschutzes korrekt ist. Mit anderen Worten, wenn ein Benutzer mehr Impressionen als die Grenze hat, nimmt er AWS Clean Rooms automatisch eine einheitliche Zufallsstichprobe der Impressionen dieses Benutzers gemäß dem berechneten UCL-Wert und schließt die verbleibenden Impressionen dieses Benutzers bei der Ausführung der Abfrage aus. Der UCL-Wert entspricht 1, wenn Sie die Anzahl der eindeutigen Benutzer zählen. Das liegt daran, dass durch das Hinzufügen, Entfernen oder Ändern eines einzelnen Benutzers die Anzahl der einzelnen Benutzer um höchstens 1 geändert werden kann.
- Der Mindestwert ist die Untergrenze eines Ausdrucks, der in einer Aggregatfunktion wie verwendet wird `sum()`. Wenn es sich bei dem Ausdruck beispielsweise um eine Spalte handelt, die als bekannt ist `purchase_value`, ist der Mindestwert die Untergrenze der Spalte.
- Der Höchstwert ist die Obergrenze eines Ausdrucks, der in einer Aggregatfunktion wie verwendet wird `sum()`. Wenn es sich bei dem Ausdruck beispielsweise um eine Spalte handelt, die als bezeichnet wird `purchase_value`, ist der Höchstwert die Obergrenze der Spalte.

In der Tabelle Berechnete differenzielle Datenschutzparameter können Sie diese Parameter verwenden, um das Gesamtvolumen des Rauschens in den Abfrageergebnissen besser zu verstehen. Wenn die konfigurierte Anzahl der pro Abfrage hinzugefügten Störungen beispielsweise 30 Benutzer umfasst und eine `COUNT DISTINCT (user_id)` Abfrage ausgeführt wird, fügt AWS Clean Rooms Differential Privacy zufälliges Rauschen hinzu, das mit hoher Wahrscheinlichkeit zwischen -30 und 30 liegt, da die Sensitivität von 1 `COUNT DISTINCT` ist. Bei einer `COUNT` Abfrage mit derselben Konfiguration fügt AWS Clean Rooms Differential Privacy statistisches Rauschen hinzu, das nach dem Benutzerbeitragslimit skaliert wird, da ein einzelner Benutzer mehrere Zeilen zum Abfrageergebnis beitragen könnte. Bei einer `SUM` Abfrage wie `SUM (purchase_value)` bei der alle Spaltenwerte positiv sind, wird das Gesamtrauschen durch das Benutzerbeitragslimit multipliziert mit dem Höchstwert skaliert. AWS Clean Rooms Differential Privacy berechnet automatisch die Sensitivitätsparameter, um das Rauschen während der Laufzeit der Abfrage zu erhöhen, wodurch

das Datenschutzbudget aufgebraucht wird. Das Budget für den Datenschutz muss aufgebraucht werden, da die Sensitivitätsparameter datenabhängig sind.

Tabellen und Analyseregeln anzeigen

Um Tabellen anzuzeigen, die der Kollaboration und den Analyseregeln zugeordnet sind

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus.
4. Wählen Sie die Registerkarte Tables (Tabellen).
5. Wählen Sie eine der folgenden Optionen:
 - a. Um Ihre der Kollaboration zugehörigen Tabellen anzuzeigen, wählen Sie unter Von Ihnen zugeordnete Tabellen eine Tabelle aus (blauer Text).
 - b. Um andere der Kollaboration zugeordnete Tabellen anzuzeigen, wählen Sie unter Von Mitarbeitern zugeordnete Tabellen eine Tabelle aus (blauer Text).
6. Sehen Sie sich die Tabellendetails und Analyseregeln auf der Seite mit den Tabellendetails an.

Eine konfigurierte Tabelle bearbeiten

Voraussetzungen:

- Und AWS-Konto mit Zugriff auf AWS Clean Rooms

In den folgenden Abschnitten wird erklärt, wie Sie den Namen, die Beschreibung und die Konfigurationsdetails von Tabellen für Amazon S3-, Amazon Athena- und Snowflake-Datenquellen bearbeiten.

[Informationen zum Bearbeiten einer konfigurierten Tabelle mithilfe von finden Sie in der AWS SDKs API-Referenz.AWS Clean Rooms](#)

Um eine konfigurierte Tabelle zu bearbeiten

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.

2. Wählen Sie im linken Navigationsbereich Tables (Tabellen) aus.
3. Wählen Sie die konfigurierte Tabelle aus, die Sie erstellt haben.
4. Wählen Sie auf der Detailseite der konfigurierten Tabelle die Option Bearbeiten aus.
5. Bearbeiten Sie Ihre Konfiguration.
6. Wählen Sie Änderungen speichern aus.

Bearbeiten konfigurierter Tabellen-Tags

Als Mitglied der Kollaboration können Sie, nachdem Sie eine konfigurierte Tabelle erstellt haben, die Tags in der konfigurierten Tabellenressource auf der Registerkarte Konfigurierte Tabellen verwalten.

Um die konfigurierten Tabellen-Tags zu bearbeiten

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Tables (Tabellen) aus.
3. Wählen Sie die konfigurierte Tabelle aus, die Sie erstellt haben.
4. Scrollen Sie auf der Detailseite der konfigurierten Tabelle nach unten zum Abschnitt Tags.
5. Wählen Sie Tags verwalten aus.
6. Auf der Seite Tags verwalten haben Sie folgende Möglichkeiten:
 - Klicken Sie zum Entfernen eines Tags auf Remove (Entfernen).
 - Um einen Tag hinzuzufügen, wählen Sie Add new tag (Neuen Tag hinzufügen).
 - Klicken Sie auf Save changes (Änderungen speichern), um die Änderungen zu speichern.

Bearbeitung der konfigurierten Tabellenanalyseregeln

Um die konfigurierte Tabellenanalyseregeln zu bearbeiten

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Tables (Tabellen) aus.
3. Wählen Sie die konfigurierte Tabelle aus, die Sie erstellt haben.

4. Scrollen Sie auf der Detailseite der konfigurierten Tabelle entweder nach unten zum Abschnitt Aggregationsanalyiseregeln, Listenanalyiseregeln oder Benutzerdefinierte Analyiseregeln. (Ihre Auswahl hängt davon ab, welche Art von Analyiseregeln Sie für die konfigurierte Tabelle ausgewählt haben.)
5. Wählen Sie Bearbeiten aus.
6. Auf der Seite Analyiseregeln bearbeiten können Sie:
 - Ändern Sie die Definition der Analyiseregeln wie folgt:
 - Ändern des JSON-Editors.
 - Wählen Sie Aus Datei importieren, um eine neue Analyiseregeldefinition hochzuladen.
 - Wählen Sie aus den folgenden Optionen eine Vorschau dessen, was Mitglieder in einer Kollaboration sehen werden:
 - Tabellen-Ansicht
 - JSON
 - Beispiel für eine Abfrage
7. Wählen Sie Änderungen speichern aus, um Ihre Änderungen zu speichern.

Löschen der konfigurierten Tabellenanalyiseregeln

Warning

Diese Aktion kann nicht rückgängig gemacht werden und wirkt sich auf alle zugehörigen Ressourcen aus.

Um die konfigurierte Tabellenanalyiseregeln zu löschen

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Tables (Tabellen) aus.
3. Wählen Sie die konfigurierte Tabelle aus, die Sie erstellt haben.
4. Scrollen Sie auf der Detailseite der konfigurierten Tabelle entweder nach unten zum Abschnitt Aggregationsanalyiseregeln, Listenanalyiseregeln oder Benutzerdefinierte Analyiseregeln. (Ihre Auswahl hängt davon ab, welche Art von Analyiseregeln Sie für die konfigurierte Tabelle ausgewählt haben.)

5. Wählen Sie Löschen aus.
6. Wenn Sie sicher sind, dass Sie die Analyseregeln löschen möchten, wählen Sie Löschen.

In der konfigurierten Tabelle sind keine Spalten zulässig

Die Konfiguration unzulässiger Ausgabespalten ist ein Steuerelement in der Regel für AWS Clean Rooms benutzerdefinierte Analysen, mit dem Sie die Liste der Spalten (falls vorhanden) definieren können, die nicht in das Abfrageergebnis projiziert werden dürfen. Die Spalten, auf die in dieser Liste verwiesen wird, gelten als „unzulässige Ausgabespalten“. Das bedeutet, dass jeder Verweis auf eine solche Spalte durch Transformation, Aliasing oder auf andere Weise in der endgültigen SELECT-Projektion (Projektion) der Abfrage möglicherweise nicht vorhanden ist.

Die Funktion verhindert zwar, dass Spalten direkt in die Ausgabe projiziert werden, sie verhindert jedoch nicht vollständig, dass zugrunde liegende Werte indirekt durch andere Mechanismen abgeleitet werden. Diese Spalten können weiterhin in einer Projektionsklausel (z. B. in einer Unterabfrage oder einem Common Table Expression (CTE)) verwendet werden, solange in der allerletzten Projektion nicht auf sie verwiesen wird.

Die Konfiguration der unzulässigen Ausgabespalten bietet Ihnen die Flexibilität, die Kontrolle über Ihre Tabelle anzuwenden und zu kodifizieren, und zwar in Kombination mit Überprüfungen auf Analysevorlagenebene, die auf Anwendungsfällen und entsprechenden Datenschutzerfordernissen basieren.

Weitere Informationen zum Einstellen dieser Konfiguration finden Sie unter [Hinzufügen einer benutzerdefinierten Analyseregeln zu einer Tabelle \(geführter Ablauf\)](#)

Beispiele

Die folgenden Beispiele zeigen, wie das Steuerelement „Unzulässige Ausgabespalten“ angewendet wird.

- Mitglied A arbeitet mit Mitglied B zusammen.
- Mitglied B ist Mitglied, das Abfragen ausführen kann.
- Mitglied A definiert eine Tabelle Benutzer mit den Spalten Alter, Geschlecht, E-Mail und Name. Die Spalten Alter und Name sind unzulässige Ausgabespalten.
- Mitglied B definiert eine Tabelle „Haustiere“ mit einem ähnlichen Satz von Spalten „Alter“, „Geschlecht“ und „Eigentümername“. Sie legen jedoch keine Einschränkungen für die

Ausgabespalten fest, was bedeutet, dass alle Spalten in der Tabelle in der Abfrage frei projiziert werden können.

Wenn Mitglied B die folgende Abfrage ausführt, wird sie blockiert, da unzulässige Ausgabespalten nicht direkt projiziert werden können:

```
SELECT
  age
FROM
  users
```

Wenn Mitglied B die folgende Abfrage ausführt, wird sie blockiert, da unzulässige Ausgabespalten nicht implizit über Project Star projiziert werden können:

```
SELECT
  *
FROM
  users
```

Wenn Mitglied B die folgende Abfrage ausführt, wird sie blockiert, da Transformationen unzulässiger Ausgabespalten nicht projiziert werden können:

```
SELECT
  COUNT(age)
FROM
  users
```

Wenn Mitglied B die folgende Abfrage ausführt, wird sie blockiert, da auf unzulässige Ausgabespalten in der endgültigen Projektion nicht mit einem Alias verwiesen werden kann:

```
SELECT
  count_age
FROM
  (SELECT COUNT(age) AS count_age FROM users)
```

Wenn Mitglied B die folgende Abfrage ausführt, wird sie blockiert, da transformierte eingeschränkte Spalten in der Ausgabe projiziert werden:

```
SELECT
  CONCAT(name, email)
FROM
  users
```

Wenn Mitglied B die folgende Abfrage ausführt, wird sie blockiert, da in CTE definierte unzulässige Ausgabespalten in der endgültigen Projektion nicht referenziert werden können:

```
WITH cte AS (
  SELECT
    age AS age_alias
  FROM
    users
)
SELECT age_alias FROM cte
```

Wenn Mitglied B die folgende Abfrage ausführt, wird sie blockiert, da unzulässige Ausgabespalten in der endgültigen Projektion nicht als Sortier- oder Partitionsschlüssel verwendet werden können:

```
SELECT
  LISTAGG(gender) WITHIN GROUP (ORDER BY age) OVER (PARTITION BY age)
FROM
  users
```

Wenn Mitglied B die folgende Abfrage ausführt, ist sie erfolgreich, da Spalten, die Teil der unzulässigen Ausgabespalten sind, weiterhin für andere Konstrukte in der Abfrage verwendet werden können, z. B. in Join- oder Filterklauseln.

```
SELECT
  u.name,
  p.gender,
  p.age
FROM
  users AS u
JOIN
```

```
pets AS p
ON
  u.name = p.owner_name
```

Im gleichen Szenario kann Mitglied B auch die Namensspalte in Benutzern als Filter- oder Sortierschlüssel verwenden:

```
SELECT
  u.email,
  u.gender
FROM
  users AS u
WHERE
  u.name = 'Mike'
ORDER BY
  u.name
```

Darüber hinaus können die unzulässigen Ausgabespalten von Benutzern in Zwischenprojektionen verwendet werden, z. B. in Unterabfragen und in folgenden CTEs Fällen:

```
WITH cte AS (
  SELECT
    u.gender,
    u.id,
    u.first_name
  FROM
    users AS u
)
SELECT
  first_name
FROM
  (SELECT cte.gender, cte.id, cte.first_name FROM cte)
```

Bearbeiten konfigurierter Tabellenzuordnungen

Als Mitglied einer Kollaboration können Sie die konfigurierten Tabellenzuordnungen bearbeiten, die Sie erstellt haben.

Um konfigurierte Tabellenzuordnungen zu bearbeiten

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus.
4. Wählen Sie den Tab Tabellen.
5. Wählen Sie für von Ihnen zugeordnete Tabellen eine Tabelle aus.
6. Scrollen Sie auf der Seite mit den Tabellendetails nach unten, um die Details zur Tabellenverknüpfung anzuzeigen.
7. Wählen Sie Bearbeiten aus.
8. Aktualisieren Sie auf der Seite „Konfigurierte Tabellenzuordnungen bearbeiten“ die Beschreibung oder die Informationen zum Zugriff auf den Dienst.
9. Wählen Sie Änderungen speichern aus.

Aufheben der Zuordnung konfigurierter Tabellen

Als Mitglied einer Kollaboration können Sie die Zuordnung einer konfigurierten Tabelle zur Kollaboration aufheben. Diese Aktion verhindert, dass das Mitglied, das Abfragen durchführen kann, die Tabelle abfragt.

Um die Zuordnung einer konfigurierten Tabelle aufzuheben

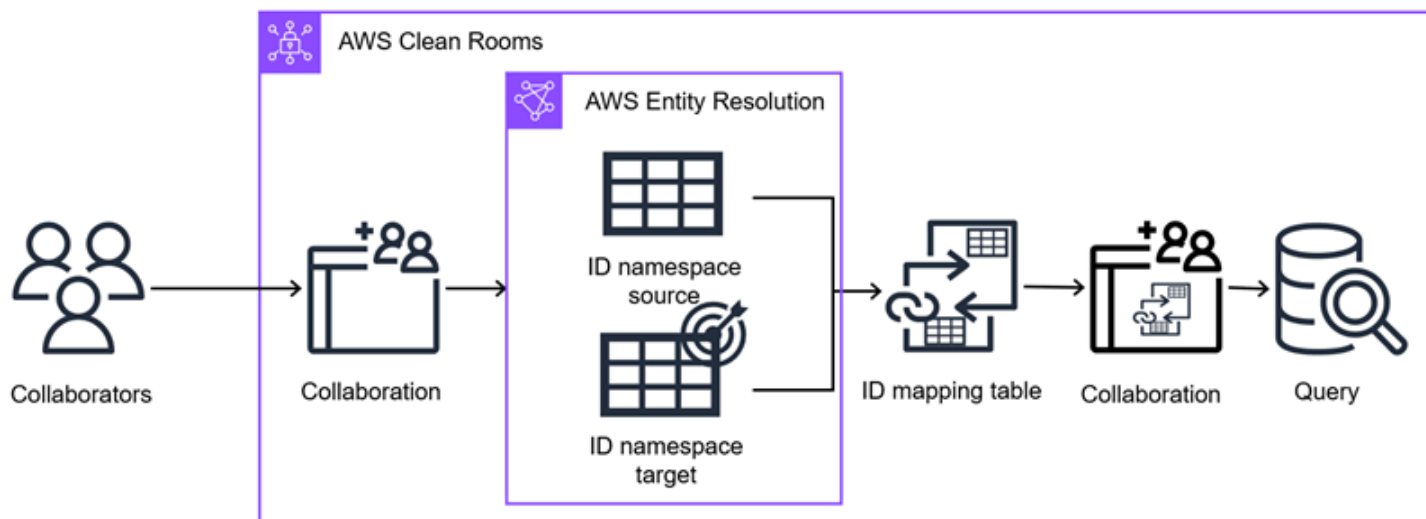
1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus.
4. Wählen Sie den Tab Tabellen.
5. Wählen Sie für von Ihnen zugeordnete Tabellen das Optionsfeld neben der Tabelle aus, deren Zuordnung Sie aufheben möchten.
6. Wählen Sie Disassociate (Zuordnung aufheben) aus.
7. Bestätigen Sie im Dialogfeld die Entscheidung, die Zuordnung der konfigurierten Tabelle aufzuheben, und verhindern Sie, dass das Mitglied, das Abfragen durchführen kann, die Tabelle abfragt, indem Sie die Option Zuordnung aufheben wählen.

AWS Entity Resolution in AWS Clean Rooms

Mit AWS Entity Resolution in AWS Clean Rooms können Sie Daten von einer Quelle in ein Ziel übersetzen, eine ID-Zuordnungstabelle mit den übersetzten Daten füllen und die Daten abfragen.

Zunächst erstellen Sie eine Kollaboration in AWS Clean Rooms und fügen die Kollaboration hinzu, die AWS-Konten Sie einladen möchten, oder treten einer Kollaboration bei, zu der Sie eingeladen wurden, indem Sie eine Mitgliedschaft erstellen. Als Nächstes führen Sie die ID-Zuordnung für zwei Datentabellen durch. Dazu verknüpfen Sie entweder eine vorhandene ID-Namespace-Quelle oder erstellen eine neue in. AWS Entity Resolution Das andere Mitglied der Kollaboration ordnet ein vorhandenes ID-Namespace-Ziel zu oder erstellt ein neues ID-Namespace-Ziel. Anschließend erstellen Sie eine ID-Zuordnungstabelle aus den beiden zugehörigen ID-Namespace-Quellen und füllen sie auf. Schließlich führt das Mitglied, das Abfragen durchführen kann, eine Abfrage für die beiden Datentabellen durch, indem es die ID-Zuordnungstabelle verknüpft.

Das folgende Diagramm fasst zusammen, wie Sie mit AWS Entity Resolution in AWS Clean Rooms arbeiten.



Note

Der derzeit unterstützte Transcodierungsdienstanbieter ist LiveRamp, der in den folgenden Ländern verfügbar ist AWS-Regionen: USA Ost (Nord-Virginia), USA Ost (Ohio) und USA West (Oregon).

- [ID-Namespaces in AWS Clean Rooms](#)
- [ID-Zuordnungstabellen in AWS Clean Rooms](#)

ID-Namespaces in AWS Clean Rooms

Ein ID-Namespace ist ein Wrapper, der Ihre Identitätstabelle umschließt und es Ihnen ermöglicht, Metadaten bereitzustellen, die Ihren Datensatz erläutern und erläutern, wie Sie ihn in einem ID-Mapping-Workflow verwenden können. Ein ID-Mapping-Workflow ist ein Datenverarbeitungsjob, der Daten aus einer Eingabedatenquelle einem Eingabedatenziel auf der Grundlage der angegebenen ID-Zuordnungsmethode zuordnet. Er erzeugt eine ID-Zuordnungstabelle.

Es gibt zwei Arten von ID-Namespaces: Quelle und Ziel. Die Quelle enthält Konfigurationen für die Quelldaten, die in einem ID-Mapping-Workflow verarbeitet werden. Das Ziel enthält eine Konfiguration der Zieldaten, in die alle Quellen aufgelöst werden. Um die Eingabedaten zu definieren, die Sie über zwei auflösen möchten AWS-Konten, erstellen Sie eine ID-Namespace-Quelle und ein ID-Namespace-Ziel, um Ihre Daten von einem Satz (Quelle) in einen anderen (Ziel) zu übersetzen.

Sie können entweder einen neuen ID-Namespace erstellen oder einen vorhandenen zuordnen. Weitere Informationen zum Erstellen eines ID-Namespace in AWS Entity Resolution finden Sie unter [Definieren von Eingabedaten mithilfe eines ID-Namespaces](#) im AWS Entity Resolution Benutzerhandbuch.

Topics

- [Einen neuen ID-Namespace erstellen und zuordnen](#)
- [Zuordnen eines vorhandenen ID-Namespaces](#)
- [ID-Namespace-Zuordnungen bearbeiten](#)
- [Zuordnung von ID-Namespace-Zuordnungen aufheben](#)

Einen neuen ID-Namespace erstellen und zuordnen

Jedes Mitglied der Kollaboration muss entweder einen ID-Namespace Source oder einen ID-Namespace Target erstellen und zuordnen, bevor eine ID-Zuordnungstabelle zur Abfrage von Identitätsdaten erstellt wird.

Wenn Sie bereits einen ID-Namespace in erstellt haben AWS Entity Resolution, fahren Sie mit fort. [Zuordnen eines vorhandenen ID-Namespaces](#)

Um einen neuen ID-Namespace zu erstellen und zuzuordnen

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole unter <https://console.aws.amazon.com/cleanrooms/>.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus.
4. Wählen Sie auf der Registerkarte Entitätsauflösung die Option Associate ID Namespace aus.
5. Wählen Sie auf der Seite „ID-Namespace zuordnen“ für Daten zur Entitätsauflösung die Option ID-Namespace erstellen aus.

Die AWS Entity Resolution Konsole wird auf einer neuen Registerkarte angezeigt.

6. Folgen Sie den Anweisungen auf der Seite „ID-Namespace erstellen“ in der AWS Entity Resolution Konsole.
 - a. Geben Sie für Details den ID-Namespace-Namen und die Beschreibung ein und wählen Sie den ID-Namespace-Typ (entweder Quelle oder Ziel) aus.
 - b. Wählen Sie für die ID-Namespace-Methode entweder die regelbasierte Methode für den regelbasierten Abgleich oder die Provider-Services für die Transcodierung durch Dritte.
 - c. Geben Sie je nach der ausgewählten ID-Namespace-Methode den Dateneingabetyp an.
 - d. Wählen Sie „ID-Namespace erstellen“.
7. Gehen Sie zurück zur AWS Clean Rooms Konsole.
8. Wählen Sie auf der Seite „ID-Namespace zuordnen“ für Daten zur Entitätsauflösung die Quelle oder das Ziel des AWS Entity Resolution ID-Namespacees, das Sie der Kollaboration zuordnen möchten, aus der Dropdownliste aus.
9. Gehen Sie wie folgt vor, um Informationen zur Zuordnung zu erhalten.
 - a. Geben Sie einen Namen für den zugehörigen ID-Namespace ein.

Sie können den Standardnamen verwenden oder diesen ID-Namespace umbenennen.

- b. (Optional) Geben Sie eine Beschreibung des ID-Namespacees ein.

Die Beschreibung hilft beim Schreiben von Abfragen.

10. Geben Sie die AWS Clean Rooms Zugriffsberechtigungen an, indem Sie eine Option auswählen und dann die empfohlene Maßnahme ergreifen.

Option	Empfohlene Aktion
AWS Clean Rooms Erlaubt das Hinzufügen und Verwalten von Berechtigungsrichtlinien	AWS Clean Rooms erstellt eine Servicerolle mit der für diese Zuordnung erforderlichen Richtlinie.
Manuelles Hinzufügen und Verwalten von Berechtigungen	Führen Sie eine der folgenden Aktionen aus: • Überprüfen Sie die Ressourcenrichtlinie und fügen Sie der Richtlinie die erforderlichen Berechtigungen hinzu. • Verwenden Sie eine bestehende Richtlinie, indem Sie Richtlinienerklärung hinzufügen wählen. Sie müssen über die erforderlichen Berechtigungen verfügen, um Rollen zu ändern und Richtlinien zu erstellen.  Note Wenn Sie die Rollenrichtlinie nicht ändern können, erhalten Sie eine Fehlermeldung, dass die Richtlinie für die Servicerolle nicht AWS Clean Rooms gefunden werden konnte.

11. (Optional) Ändern Sie bei Konfigurationen mit erweiterten ID-Zuordnungstabellen den Standardschutz für die Spalte, die aus dem ID-Namespace stammt.

Die ID-Zuordnungstabelle ist standardmäßig so konfiguriert, dass sowohl für die Spalte als auch für die `sourceID` Spalte nur ein `INNER JOIN` Zugriff zulässig ist. `targetID` Sie können diese Konfiguration so ändern, dass die Spalte, die aus diesem ID-Namespace stammt (entweder `sourceID` oder `targetID`), an beliebiger Stelle in der Abfrage zulässig ist.

Ihr Ziel	Empfohlene Option
Kategorisieren Sie die Spalte als „Join-Spalte“ und lassen Sie sie nur in einer INNER JOIN Klausel zu	Ja
Kategorisieren Sie die Spalte als „Dimensionsspalte“ und lassen Sie sie an einer beliebigen Stelle in der Abfrage zu, einschließlich einer JOIN Klausel WHERE und der GROUP BY Anweisungen der Abfrage. SELECT	Nein, an beliebiger Stelle in der Abfrage zulassen

- (Optional) Wenn Sie Tags für die ID-Namespace-Ressource aktivieren möchten, wählen Sie Neues Tag hinzufügen aus und geben Sie dann das Schlüssel - und Wertepaar ein.
- Wählen Sie Associate aus.
- Sehen Sie sich auf der Registerkarte Entitätsauflösung in der Tabelle Zugeordnete ID-Namespace den zugehörigen ID-Namespace an und überprüfen Sie, ob der ID-Namespace-Typ korrekt ist (Quelle oder Ziel).

Nachdem alle Mitglieder der Kollaboration ihre ID-Namespace zugeordnet haben, können Sie [eine ID-Zuordnungstabelle erstellen](#) und die Daten abfragen.

Zuordnen eines vorhandenen ID-Namespace

In diesem Verfahren ordnet jedes Mitglied der Kollaboration entweder seine bestehende ID-Namespace-Quelle oder sein ID-Namespace-Ziel zu.

Um einen vorhandenen ID-Namespace zuzuordnen

- Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole unter <https://console.aws.amazon.com/cleanrooms/>.
- Wählen Sie im linken Navigationsbereich Collaborations aus.
- Wählen Sie die Kollaboration aus.
- Wählen Sie auf der Registerkarte Entitätsauflösung die Option Associate ID Namespace aus.

5. Wählen Sie auf der Seite „ID-Namespace zuordnen“ für Daten zur Entitätsauflösung die Quelle oder das Ziel des AWS Entity Resolution ID-Namespacees, das Sie der Kollaboration zuordnen möchten, aus der Dropdownliste aus.
6. Gehen Sie wie folgt vor, um Informationen zur Zuordnung zu erhalten.

- a. Geben Sie einen Namen für den zugehörigen ID-Namespace ein.

Sie können den Standardnamen verwenden oder diesen ID-Namespace umbenennen.

- b. (Optional) Geben Sie eine Beschreibung des ID-Namespacees ein.

Die Beschreibung hilft beim Schreiben von Abfragen.

7. Geben Sie die AWS Clean Rooms Zugriffsberechtigungen an, indem Sie eine Option auswählen und dann die empfohlene Maßnahme ergreifen.

Option	Empfohlene Aktion
AWS Clean Rooms Erlaubt das Hinzufügen und Verwalten von Berechtigungsrichtlinien	AWS Clean Rooms erstellt eine Servicerolle mit der für diese Zuordnung erforderlichen Richtlinie.
Manuelles Hinzufügen und Verwalten von Berechtigungen	Führen Sie eine der folgenden Aktionen aus: - Überprüfen Sie die Ressourcenrichtlinie und fügen Sie der Richtlinie die erforderlichen Berechtigungen hinzu. - Verwenden Sie eine bestehende Richtlinie, indem Sie Richtlinienerklärung hinzufügen wählen. Sie müssen über die erforderlichen Berechtigungen verfügen, um Rollen zu ändern und Richtlinien zu erstellen.  Note Wenn Sie die Rollenrichtlinie nicht ändern können, erhalten Sie eine

Option	Empfohlene Aktion
	Fehlermeldung, dass die Richtlinie für die Servicerolle nicht AWS Clean Rooms gefunden werden konnte.

8. (Optional) Ändern Sie bei Konfigurationen mit erweiterten ID-Zuordnungstabellen den Standardschutz für die Spalte, die aus dem ID-Namespace stammt.

Die ID-Zuordnungstabelle ist standardmäßig so konfiguriert, dass sowohl für die Spalte als auch für die `sourceID` Spalte nur ein `INNER JOIN` Zugriff zulässig ist. `targetID` Sie können diese Konfiguration so ändern, dass die Spalte, die aus diesem ID-Namespace stammt (entweder `sourceID` oder `targetID`), an beliebiger Stelle in der Abfrage zulässig ist.

Ihr Ziel	Empfohlene Option
Kategorisieren Sie die Spalte als „Join-Spalte“ und lassen Sie sie nur in einer <code>INNER JOIN</code> Klausel zu.	Ja
Kategorisieren Sie die Spalte als „Dimensionspalte“ und lassen Sie sie an einer beliebigen Stelle in der Abfrage zu, einschließlich einer <code>JOIN</code> Klausel, <code>SELECT</code> , <code>WHERE</code> , und <code>GROUP BY</code> Anweisungen der Abfrage.	Nein, lassen Sie eine beliebige Stelle in der Abfrage zu

9. (Optional) Wenn Sie Tags für die ID-Namespace-Ressource aktivieren möchten, wählen Sie Neues Tag hinzufügen aus und geben Sie dann das Schlüssel - und Wertepaar ein.
10. Wählen Sie Associate aus.
11. Sehen Sie sich auf der Registerkarte Entitätsauflösung in der Tabelle Zugeordnete ID-Namespace den zugehörigen ID-Namespace an und überprüfen Sie, ob der ID-Namespace-Typ korrekt ist (Quelle oder Ziel).

Nachdem alle Mitglieder der Kollaboration ihre ID-Namespace zugeordnet haben, können Sie [eine ID-Zuordnungstabelle erstellen](#) und die Daten abfragen.

ID-Namespace-Zuordnungen bearbeiten

Als Mitglied einer Kollaboration können Sie die von Ihnen erstellten ID-Namespace-Zuordnungen bearbeiten.

Um eine ID-Namespace-Zuordnung zu bearbeiten

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole unter <https://console.aws.amazon.com/cleanrooms/>.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus.
4. Wählen Sie die Registerkarte Entitätsauflösung.
5. Wählen Sie für Zugeordnete ID-Namespace einen ID-Namespace aus.
6. Scrollen Sie auf der Seite mit den ID-Namespace-Details nach unten, um die Details zur ID-Namespace-Zuordnung anzuzeigen.
7. Wählen Sie Bearbeiten aus.
8. Bearbeiten Sie auf der Seite ID-Namespace-Zuordnungen bearbeiten eine der folgenden Optionen:
 - a. Für Zuordnungsdetails aktualisieren Sie den Namen oder die Beschreibung.
 - b. (Optional) Ändern Sie für erweiterte Konfigurationen von ID-Zuordnungstabellen den Standardschutz für die Spalte, die aus dem ID-Namespace stammt.

Die ID-Zuordnungstabelle ist standardmäßig so konfiguriert, dass sowohl für die Spalte als auch für die `sourceID` Spalte nur ein `INNER JOIN` Zugriff zulässig ist. `targetID` Sie können diese Konfiguration so ändern, dass die Spalte, die aus diesem ID-Namespace stammt (entweder `sourceID` oder `targetID`), an beliebiger Stelle in der Abfrage zulässig ist.

Ihr Ziel	Empfohlene Option
Kategorisieren Sie die Spalte als „Join-Spalte“ und lassen Sie sie nur in einer <code>INNER JOIN</code> Klausel zu	Ja

Ihr Ziel	Empfohlene Option
Kategorisieren Sie die Spalte als „Dimensionsspalte“ und lassen Sie sie an einer beliebigen Stelle in der Abfrage zu, einschließlich einer JOIN Klausel WHERE und der GROUP BY Anweisungen der Abfrage. SELECT	Nein, an beliebiger Stelle in der Abfrage zulassen

9. Wählen Sie **Änderungen speichern** aus.

Zuordnung von ID-Namespace-Zuordnungen aufheben

Als Mitglied einer Kollaboration können Sie die Zuordnung eines ID-Namespace zur Kollaboration aufheben. Diese Aktion verhindert, dass das Mitglied, das Abfragen durchführen kann, die Tabelle abfragt.

Warning

Beim Trennen der Zuordnung einer ID-Namespace-Zuordnung zu einer Kollaboration werden alle Daten aus abgeleiteten ID-Zuordnungstabellen gelöscht, sodass sie nicht mehr abfragbar sind.

Wenn Ihre ID-Namespace-Zuordnung beispielsweise als QUELLE in drei verschiedenen ID-Zuordnungstabellen verwendet wurde, werden alle Daten aus diesen ID-Zuordnungstabellen gelöscht, wenn Sie die Zuordnung Ihrer ID-Namespace-Zuordnung aufheben.

Um die Zuordnung einer ID-Namespace-Zuordnung aufzuheben

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole unter. <https://console.aws.amazon.com/cleanrooms/>
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus.
4. Wählen Sie die Registerkarte Entitätsauflösung.
5. Wählen Sie unter Zugeordnete ID-Namespace das Optionsfeld neben dem ID-Namespace aus, dessen Zuordnung Sie aufheben möchten.
6. Wählen Sie Disassociate (Zuordnung aufheben) aus.

7. Bestätigen Sie im Dialogfeld Ihre Entscheidung, die Verbindung zum ID-Namespace zu trennen, indem Sie „Zuordnung trennen“ wählen. Diese Aktion verhindert, dass jedes Mitglied, das Abfragen durchführen kann, auf die ID-Zuordnungstabelle zugreifen kann.

Wenn ein Mitglied der Kollaboration einen der ID-Namespace entfernt, können Sie die ID-Zuordnungstabelle nicht erneut auffüllen, wenn die Quelle die Kollaboration verlassen hat.

Auch wenn die ID-Zuordnungstabelle zuvor gefüllt wurde, bedeutet das Aufheben der Zuordnung des ID-Namespace, dass Sie für diese Tabelle keine Abfragen mehr ausführen können.

ID-Zuordnungstabellen in AWS Clean Rooms

Eine ID-Zuordnungstabelle ist eine Ressource AWS Clean Rooms , die die Identitätszuweisung mehrerer Parteien in einer Zusammenarbeit ermöglicht.

Bevor Sie eine ID-Zuordnungstabelle erstellen, müssen Sie zunächst sowohl Quell- als auch Zieldaten als ID-Namespace konfiguriert haben.

Nachdem Sie eine ID-Zuordnungstabelle erstellt haben, verwenden Sie einen ID-Mapping-Workflow, um den Quell-ID-Namespace in den Ziel-ID-Namespace zu übersetzen. Sie können dazu entweder eine regelbasierte Methode oder eine Transcodierungsmethode für Providerdienste verwenden.

Ein ID-Mapping-Workflow ist ein Datenverarbeitungsjob, der Daten aus einer Eingabedatenquelle einem Eingabedatenziel auf der Grundlage der angegebenen ID-Zuordnungs-Workflow-Methode zuordnet. Dieser Workflow füllt eine ID-Zuordnungstabelle auf.

Note

ID-Zuordnungstabellen können nur aus Datensätzen erstellt werden, die in Amazon S3 gespeichert und in Tabellen gecrawlt wurden. AWS Glue

Es gibt zwei Workflow-Methoden für die ID-Zuordnung: die regelbasierte ID-Zuordnung oder die ID-Zuordnung von Providerdiensten:

- Regelbasierte ID-Zuordnung – Sie benutzen Abgleichsregeln, um unternehmenseigene Daten von einer Quelle zu einem Ziel zu übersetzen.
- ID-Zuordnung für Providerdienste — Sie verwenden den LiveRamp Provider-Service, um Daten von Drittanbietern von einer Quelle in ein Ziel zu übersetzen.

 Note

Der derzeit unterstützte Transcodierungsdienstleister ist LiveRamp. Jedes Mitglied der Kollaboration, das ein Abonnement bei LiveRamp Through hat, AWS Data Exchange kann die ID-Zuordnungstabelle erstellen. Wenn Sie bereits ein Abonnement für LiveRamp, aber nicht über, haben, wenden Sie sich an AWS Data Exchange, LiveRamp um ein privates Angebot zu erhalten. Weitere Informationen finden [Sie unter Abonnieren eines Anbieterdienstes AWS Data Exchange](#) im AWS Entity Resolution Benutzerhandbuch.

Themen

- [Eine neue ID-Zuordnungstabelle erstellen und auffüllen](#)
- [Auffüllen einer vorhandenen ID-Zuordnungstabelle](#)
- [Bearbeiten einer ID-Zuordnungstabelle](#)
- [Löschen einer ID-Zuordnungstabelle](#)

Eine neue ID-Zuordnungstabelle erstellen und auffüllen

Voraussetzungen

Bevor Sie eine ID-Zuordnungstabelle erstellen, stellen Sie sicher, dass Sie über Folgendes verfügen:

- Quelle und Ziel eines zugehörigen ID-Namespace
- Die ID-Namespace, die entweder für die regelbasierte ID-Zuordnung oder die ID-Zuordnung für Provider-Dienste konfiguriert sind

Sie können zwei Arten von ID-Zuordnungstabellen erstellen:

- Regelbasiert — Verwendet Abgleichsregeln, um Daten von Erstanbietern zu übersetzen
- Provider-Dienste — Wird zur Übersetzung von Ramp verwendet LiveRamp IDs

Nachdem Sie eine ID-Zuordnungstabelle erstellt haben, können Sie sie entweder sofort auffüllen, indem Sie einen ID-Zuordnungs-Workflow ausführen, oder Sie können warten, bis die Tabelle später gefüllt wird.

Nachdem die ID-Zuordnungstabelle erfolgreich gefüllt wurde, können Sie eine Join-Abfrage mit mehreren Tabellen für die ID-Zuordnungstabelle ausführen, um die Daten `sourceId` mit den zu verknüpfen `targetId` und die Daten zu analysieren.

Themen

- [Erstellen Sie eine ID-Zuordnungstabelle \(regelbasiert\)](#)
- [Erstellen Sie eine ID-Zuordnungstabelle \(Provider-Services\)](#)

Erstellen Sie eine ID-Zuordnungstabelle (regelbasiert)

In diesem Thema wird der Prozess der Erstellung einer ID-Zuordnungstabelle beschrieben, die Abgleichsregeln verwendet, um Erstanbieterdaten von einer Quelle in ein Ziel zu übersetzen.

Wenn Sie eine regelbasierte ID-Zuordnungstabelle erstellen, können Sie festlegen, dass nur neue, aktualisierte oder gelöschte Datensätze im Workflow verarbeitet werden, indem Sie die inkrementelle Verarbeitung aktivieren.

Um eine neue ID-Zuordnungstabelle mit der regelbasierten Methode zu erstellen und aufzufüllen

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole unter. <https://console.aws.amazon.com/cleanrooms/>
2. Wählen Sie im Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration und dann die Registerkarte Entitätsauflösung aus.
4. Wählen Sie „ID-Zuordnungstabelle erstellen“.
5. Führen Sie unter Einstellungen für die ID-Zuordnung einen der folgenden Schritte aus:
 - Um einen neuen Workflow zu erstellen, lassen Sie die Option Neuen ID-Zuordnungs-Workflow erstellen ausgewählt.
 - Um einen vorhandenen Workflow zu verwenden, deaktivieren Sie das Kontrollkästchen und wählen Sie einen Workflow aus der Dropdownliste aus. Fahren Sie dann mit Schritt 9 fort.
6. Zeigen Sie unter Identitätsdaten die Quelle und das Ziel an oder konfigurieren Sie sie.
 - Für ein einzelnes ID-Namespace-Paar: Überprüfen Sie die vorgewählte Quelle und das Ziel.
 - Für mehrere ID-Namespace: Wählen Sie Quelle und Ziel aus den Drop-down-Listen aus.
7. (Optional) Wählen Sie Inkrementelle Verarbeitung aktivieren aus, um nur neue, aktualisierte oder gelöschte Datensätze im Workflow zu verarbeiten.

AWS Entity Resolution verarbeitet nur neue, aktualisierte oder gelöschte Datensätze im Quell- oder Ziel-ID-Namespace, anstatt die gesamte ID-Zuordnungstabelle neu zu erstellen.

Wenn Sie diese Option nicht ausgewählt lassen, AWS Entity Resolution wird der standardmäßige ID-Mapping-Workflow für die ID-Zuordnungstabelle für die Batchverarbeitung ausgeführt.

8. Konfigurieren Sie unter Regelparameter Folgendes:

- Regelkontrollen — Wählen Sie aus, ob das Ziel oder die Quelle passende Regeln bereitstellt.

Sie können die Regeln anzeigen, indem Sie die Option Regeln anzeigen aktivieren.

Regelsteuerungen müssen zwischen dem Quell- und dem Ziel-ID-Namespace kompatibel sein, um in einem ID-Mapping-Workflow verwendet werden zu können. Wenn beispielsweise ein Quell-ID-Namespace die Regeln auf das Ziel beschränkt, der Ziel-ID-Namespace die Regeln jedoch auf die Quelle beschränkt, führt dies zu einem Fehler.

- Der Vergleichstyp wird automatisch auf Mehrere Eingabefelder gesetzt.

Dies liegt daran, dass beide Teilnehmer diese Option zuvor ausgewählt hatten.

- Abgleich von Datensätzen
 - Eine Quelle zu einem Ziel — Speichert einen passenden Datensatz pro Ziel
 - Viele Quellen für ein Ziel — Speichert alle passenden Datensätze pro Ziel

 Note

Die für die Quell- und Ziel-ID-Namespace angegebenen Einschränkungen müssen kompatibel sein.

9. Für Details zur ID-Zuordnung konfigurieren Sie Folgendes:

- a. Geben Sie einen Namen für die ID-Zuordnungstabelle ein, oder behalten Sie den Standardnamen bei.
- b. (Optional) Geben Sie eine Beschreibung der ID-Zuordnungstabelle ein.

Die Beschreibung hilft beim Schreiben von Abfragen.

10. Wählen Sie für AWS Clean Rooms den Zugriff eine Option aus:

- Hinzufügen und Verwalten von Berechtigungsrichtlinien zulassen AWS Clean Rooms — Erstellt automatisch eine Servicerolle.
- Manuelles Hinzufügen und Verwalten von Berechtigungen — Überprüfen und ändern Sie entweder die Ressourcenrichtlinie oder wählen Sie Richtlinienerklärung hinzufügen.

 Note

Wenn Sie die Rollenrichtlinie nicht ändern können, erhalten Sie eine Fehlermeldung, dass die Richtlinie für die Servicerolle nicht gefunden werden AWS Clean Rooms konnte.

11. Wählen Sie für AWS Entity Resolution den Zugriff eine Option aus:

Dieser Abschnitt ist nur sichtbar, wenn Sie eine neue ID-Zuordnungstabelle erstellen.

- Erstellen und verwenden Sie eine neue Servicerolle
 - Der Standardname der Servicerolle lautet `entityresolution-id-mapping-workflow-<timestamp>`
 - (Optional) Wählen Sie für verschlüsselte Daten die Option Diese Daten werden mit einem KMS-Schlüssel verschlüsselt aus und geben Sie den AWS KMS Schlüssel ein.
- Verwenden Sie eine vorhandene Servicerolle
 - Wählen Sie einen vorhandenen Servicerollennamen aus der Dropdownliste aus oder geben Sie einen Rollen-ARN ein.

Die Liste der Rollen wird angezeigt, wenn Sie berechtigt sind, Rollen aufzulisten.

Rufen Sie die Servicerolle auf, indem Sie auf den externen Link In IAM anzeigen klicken.

Wenn keine vorhandenen Servicerollen vorhanden sind, ist die Option „Eine bestehende Servicerolle verwenden“ nicht verfügbar.

Versucht standardmäßig AWS Clean Rooms nicht, die bestehende Rollenrichtlinie zu aktualisieren, um die erforderlichen Berechtigungen hinzuzufügen.

- (Optional) Wählen Sie Eine vorkonfigurierte Richtlinie mit den erforderlichen Berechtigungen zu dieser Rolle hinzufügen aus, um der Rolle die erforderlichen Berechtigungen zuzuweisen.

Sie benötigen Berechtigungen, um Rollen zu ändern und Richtlinien zu erstellen.

12. (Optional) Konfigurieren Sie unter Zusätzliche Einstellungen:

a. Einstellungen für die ID-Zuordnungstabelle

- Um die benutzerdefinierte Verschlüsselung zu aktivieren, wählen Sie Verschlüsselungseinstellungen anpassen und geben Sie einen AWS KMS Schlüssel ein.

Note

Dieser KMS-Schlüssel muss die erforderlichen Berechtigungen für die Verwendung innerhalb AWS Entity Resolution einer KMS-Schlüsselrichtlinie gewähren. `cleanrooms.amazonaws.com` Weitere Informationen zu den erforderlichen Berechtigungen für die Arbeit mit Verschlüsselungen mit einem ID-Mapping-Workflow finden [Sie unter Erstellen einer Workflow-Jobrolle für AWS Entity Resolution](#) im AWS Entity Resolution Benutzerhandbuch.

- Um Tags hinzuzufügen, wählen Sie Neues Tag hinzufügen und geben Sie Schlüssel-Wert-Paare ein

b. Workflow-Einstellungen für die ID-Zuordnung (nur für neue Workflows):

- Um unterschiedliche Namen zu verwenden, deaktivieren Sie die Option Gleiche Namen und Beschreibung für die ID-Zuordnungstabelle beibehalten und geben Sie neue Werte ein.
- Um Tags hinzuzufügen, wählen Sie Neues Tag hinzufügen und geben Sie Schlüssel-Wert-Paare ein

13. Wählen Sie eine der folgenden Optionen:

- ID-Zuordnungstabelle erstellen — Erstellt eine leere Tabelle, die Sie später auffüllen können ([Auffüllen einer vorhandenen ID-Zuordnungstabelle](#))
- ID-Zuordnungstabelle erstellen und auffüllen — Erstellt die Tabelle und füllt sie sofort auf (kann mehrere Stunden dauern)

Der Workflow-Prozess für die ID-Zuordnung beginnt. Während dieses Vorgangs wird die ID-Zuordnungstabelle mit übersetzten Daten gefüllt IDs. Die Bearbeitung des Workflows zur ID-Zuordnung kann einige Stunden dauern.

Nachdem die ID-Zuordnungstabelle erfolgreich gefüllt wurde, können Sie [die ID-Zuordnungstabelle abfragen](#), um sie `sourceId` mit den Daten zu verknüpfen `targetId` und zu analysieren.

Erstellen Sie eine ID-Zuordnungstabelle (Provider-Services)

In diesem Thema wird der Vorgang zum Erstellen einer ID-Zuordnungstabelle beschrieben, die einen Anbieterdienst (LiveRamp) verwendet. Die LiveRamp Providerdienste übersetzen einen IDs Quell-Ramp-Satz in einen anderen, wobei entweder ein verwalteter oder ein abgeleiteter IDs Ramp-Satz verwendet wird.

Um eine neue ID-Zuordnungstabelle mit der Provider Services-Methode zu erstellen

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole unter <https://console.aws.amazon.com/cleanrooms/>.
2. Wählen Sie im Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration und dann die Registerkarte Entitätsauflösung aus.
4. Wählen Sie „ID-Zuordnungstabelle erstellen“.
5. Führen Sie unter Einstellungen für die ID-Zuordnung einen der folgenden Schritte aus:
 - Um einen neuen Workflow zu erstellen, lassen Sie die Option Neuen ID-Zuordnungs-Workflow erstellen ausgewählt.
 - Um einen vorhandenen Workflow zu verwenden, deaktivieren Sie das Kontrollkästchen und wählen Sie einen Workflow aus der Dropdownliste aus. Fahren Sie dann mit Schritt 9 fort.
6. Zeigen Sie unter Identitätsdaten die Quelle und das Ziel an oder konfigurieren Sie sie.
 - Für ein einzelnes ID-Namespace-Paar: Überprüfen Sie die vorgewählte Quelle und das Ziel.
 - Für mehrere ID-Namespace: Wählen Sie Quelle und Ziel aus den Drop-down-Listen aus.
7. Stellen Sie unter Methode sicher, dass die ausgewählte Workflow-Methode für die ID-Zuordnung transkodiert ist. LiveRamp
8. Führen LiveRamp Sie für Konfigurationen einen der folgenden Schritte aus:
 - Geben Sie den LiveRamp ID-Manager-ARN und den LiveRamp Secret Manager-ARN ein.
 - Wählen Sie Aus vorhandenem Workflow importieren aus.
9. Für Details zur ID-Zuordnung konfigurieren Sie Folgendes:

- a. Geben Sie einen Namen für die ID-Zuordnungstabelle ein, oder behalten Sie den Standardnamen bei.
- b. (Optional) Geben Sie eine Beschreibung der ID-Zuordnungstabelle ein.

Die Beschreibung hilft beim Schreiben von Abfragen.

10. Wählen Sie für AWS Clean Rooms den Zugriff eine Option aus:

- Hinzufügen und Verwalten von Berechtigungsrichtlinien zulassen AWS Clean Rooms — Erstellt automatisch eine Servicerolle.
- Manuelles Hinzufügen und Verwalten von Berechtigungen — Überprüfen und ändern Sie entweder die Ressourcenrichtlinie oder wählen Sie Richtlinienerklärung hinzufügen.

 Note

Wenn Sie die Rollenrichtlinie nicht ändern können, erhalten Sie eine Fehlermeldung, dass die Richtlinie für die Servicerolle nicht gefunden werden AWS Clean Rooms konnte.

11. Wählen Sie für AWS Entity Resolution den Zugriff eine Option aus:

Dieser Abschnitt ist nur sichtbar, wenn Sie eine neue ID-Zuordnungstabelle erstellen.

- Erstellen und verwenden Sie eine neue Servicerolle
 - Der Standardname der Servicerolle lautet `entityresolution-id-mapping-workflow-<timestamp>`
 - (Optional) Wählen Sie für verschlüsselte Daten die Option Diese Daten werden mit einem KMS-Schlüssel verschlüsselt aus und geben Sie den AWS KMS Schlüssel ein.
- Verwenden Sie eine vorhandene Servicerolle
 - Wählen Sie einen vorhandenen Servicerollennamen aus der Dropdownliste aus oder geben Sie einen Rollen-ARN ein.

Die Liste der Rollen wird angezeigt, wenn Sie berechtigt sind, Rollen aufzulisten.

Rufen Sie die Servicerolle auf, indem Sie auf den externen Link In IAM anzeigen klicken.

Wenn keine vorhandenen Servicerollen vorhanden sind, ist die Option „Eine bestehende Servicerolle verwenden“ nicht verfügbar.

Versucht standardmäßig AWS Clean Rooms nicht, die bestehende Rollenrichtlinie zu aktualisieren, um die erforderlichen Berechtigungen hinzuzufügen.

- (Optional) Wählen Sie Eine vorkonfigurierte Richtlinie mit den erforderlichen Berechtigungen zu dieser Rolle hinzufügen aus, um der Rolle die erforderlichen Berechtigungen zuzuweisen.

Sie benötigen Berechtigungen, um Rollen zu ändern und Richtlinien zu erstellen.

12. (Optional) Konfigurieren Sie unter Zusätzliche Einstellungen:

a. Einstellungen für die ID-Zuordnungstabelle

- Um die benutzerdefinierte Verschlüsselung zu aktivieren, wählen Sie Verschlüsselungseinstellungen anpassen und geben Sie einen AWS KMS Schlüssel ein.

Note

Dieser KMS-Schlüssel muss die erforderlichen Berechtigungen für die Verwendung innerhalb AWS Entity Resolution einer KMS-Schlüsselrichtlinie gewähren. `cleanrooms.amazonaws.com` Weitere Informationen zu den erforderlichen Berechtigungen für die Arbeit mit Verschlüsselungen mit einem ID-Mapping-Workflow finden [Sie unter Erstellen einer Workflow-Jobrolle für AWS Entity Resolution](#) im AWS Entity Resolution Benutzerhandbuch.

- Um Tags hinzuzufügen, wählen Sie Neues Tag hinzufügen und geben Sie Schlüssel-Wert-Paare ein.

b. Workflow-Einstellungen für die ID-Zuordnung (nur für neue Workflows):

- Um unterschiedliche Namen zu verwenden, deaktivieren Sie die Option Gleiche Namen und Beschreibung für die ID-Zuordnungstabelle beibehalten und geben Sie neue Werte ein.
- Um Tags hinzuzufügen, wählen Sie Neues Tag hinzufügen und geben Sie Schlüssel-Wert-Paare ein.

13. Wählen Sie eine der folgenden Optionen:

- ID-Zuordnungstabelle erstellen — Erstellt eine leere Tabelle, die Sie später auffüllen können ([Auffüllen einer vorhandenen ID-Zuordnungstabelle](#))
- ID-Zuordnungstabelle erstellen und auffüllen — Erstellt die Tabelle und füllt sie sofort auf (kann mehrere Stunden dauern)

Der Workflow-Prozess für die ID-Zuordnung beginnt. Während dieses Vorgangs wird die ID-Zuordnungstabelle mit übersetzten Daten gefüllt. Die Bearbeitung des Workflows zur ID-Zuordnung kann einige Stunden dauern.

Nachdem die ID-Zuordnungstabelle erfolgreich gefüllt wurde, können Sie [die ID-Zuordnungstabelle abfragen](#), um sie `sourceId` mit den Daten zu verknüpfen `targetId` und zu analysieren.

Auffüllen einer vorhandenen ID-Zuordnungstabelle

Verwenden Sie diesen Workflow, wenn einem ID-Namespace neue Daten hinzugefügt werden. Wenn Sie sich bei der [Erstellung der ID-Zuordnungstabelle dafür entschieden haben, die](#) inkrementelle Verarbeitung zu aktivieren, können Sie nur neue, aktualisierte oder gelöschte Datensätze im Quell- oder Ziel-ID-Namespace verarbeiten, anstatt die gesamte ID-Zuordnungstabelle neu zu erstellen.

Um eine bestehende ID-Zuordnungstabelle aufzufüllen

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole unter <https://console.aws.amazon.com/cleanrooms/>.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Zusammenarbeit aus.
4. Gehen Sie zur Registerkarte Entitätsauflösung.
5. Wählen Sie im Abschnitt ID-Zuordnungstabellen eine ID-Zuordnungstabelle aus.
6. Wenn Sie die inkrementelle Verarbeitung für diese ID-Zuordnungstabelle nicht aktiviert haben, wählen Sie eine ID-Zuordnungstabelle und dann Auffüllen aus.
7. Wenn Sie die inkrementelle Verarbeitung für diese ID-Zuordnungstabelle aktiviert haben, wählen Sie Auffüllen mit und wählen Sie dann eine der folgenden Optionen:

- Inkrementelle Verarbeitung — Verarbeitet nur neue, aktualisierte oder gelöschte Datensätze im Quell- oder Ziel-ID-Namespace.

Empfohlen für häufige Updates, tägliche Datenläufe oder Datensynchronisierung in Echtzeit.

- Stapelverarbeitung — Verarbeitet die gesamte ID-Zuordnungstabelle.

Empfohlen für die Ersteinrichtung, regelmäßige vollständige Aktualisierungen oder wenn signifikante Änderungen sowohl im Quell- als auch im Ziel-ID-Namespace auftreten.

- Verarbeitung nur löschen — Verarbeitet nur gelöschte Datensätze aus dem Quell-ID-Namespace und aktualisiert den Ziel-ID-Namespace entsprechend.

Wird für die schnelle Synchronisation von Löschungen empfohlen.

8. Der Workflow-Prozess für die ID-Zuordnung beginnt.

Während dieses Vorgangs wird die ID-Zuordnungstabelle mit transcodierten Daten IDs gefüllt. Die Bearbeitung des Workflows zur ID-Zuordnung kann einige Stunden dauern.

Nachdem die ID-Zuordnungstabelle erfolgreich gefüllt wurde, können Sie [die ID-Zuordnungstabelle abfragen](#), um sie `sourceId` mit der zu verknüpfentargetId.

Bearbeiten einer ID-Zuordnungstabelle

Als Mitglied einer Kollaboration können Sie die von Ihnen erstellte ID-Zuordnungstabelle bearbeiten.

Um eine ID-Zuordnungstabelle zu bearbeiten

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#) mit Ihrem AWS-Konto (falls Sie dies noch nicht getan haben).
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Zusammenarbeit aus.
4. Wählen Sie die Registerkarte Entitätsauflösung.
5. Wählen Sie für ID-Zuordnungstabellen eine Tabelle aus.
6. Scrollen Sie auf der Detailseite der ID-Zuordnungstabelle nach unten, um die Details der ID-Zuordnungstabelle anzuzeigen.
7. Wählen Sie Bearbeiten aus.
8. Aktualisieren Sie auf der Seite ID-Zuordnungstabelle bearbeiten die Beschreibung oder die Informationen zum Dienstzugriff.
9. Wählen Sie Änderungen speichern aus.

Löschen einer ID-Zuordnungstabelle

Als Mitglied einer Kollaboration können Sie eine von Ihnen erstellte ID-Zuordnungstabelle löschen. Diese Aktion verhindert, dass das Mitglied, das Abfragen durchführen kann, die Tabelle abfragt.

 Warning

Durch das Löschen einer Zuordnungstabelle werden alle aufgefüllten Daten dauerhaft entfernt.

Um eine ID-Zuordnungstabelle zu löschen

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#) mit Ihrem AWS-Konto (falls Sie dies noch nicht getan haben).
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Zusammenarbeit aus.
4. Wählen Sie die Registerkarte Entitätsauflösung.
5. Wählen Sie für ID-Zuordnungstabellen eine Tabelle aus.
6. Scrollen Sie auf der Detailseite der ID-Zuordnungstabelle nach unten, um die ID-Zuordnungstabellen anzuzeigen.
7. Wählen Sie eine ID-Zuordnungstabelle aus und klicken Sie dann auf Löschen.
8. Wenn Sie sicher sind, dass Sie die ID-Zuordnungstabelle löschen möchten, wählen Sie Löschen.

Analysevorlagen in AWS Clean Rooms

Analysevorlagen funktionieren mit dem [Benutzerdefinierte Analyseregeln in AWS Clean Rooms](#). Mit einer Analysevorlage können Sie Parameter definieren, die Ihnen helfen, dieselbe Abfrage wiederzuverwenden. AWS Clean Rooms unterstützt eine Teilmenge der Parametrisierung mit Literalwerten.

Analysevorlagen sind kollaborationsspezifisch. Für jede Kollaboration können Mitglieder nur die Abfragen in dieser Kollaboration sehen. Wenn Sie beabsichtigen, Differential Privacy in einer Kollaboration zu verwenden, sollten Sie sicherstellen, dass Ihre Analysevorlagen mit der [allgemeinen Abfragestruktur](#) von AWS Clean Rooms Differential Privacy kompatibel sind.

Sie können eine Analysevorlage auf zwei Arten erstellen: mit SQL-Code oder mit Python-Code für Spark.

Topics

- [SQL-Analysevorlagen](#)
- [PySpark Analysevorlagen](#)
- [PySpark Analysevorlagen zur Fehlerbehebung](#)

SQL-Analysevorlagen

Mit SQL-Analysevorlagen können Sie Daten aus verschiedenen Datensätzen innerhalb einer Zusammenarbeit abfragen und analysieren. Sie können diese Vorlagen verwenden, um verschiedene Arten von Analysen durchzuführen, z. B. um Zielgruppenüberschneidungen zu identifizieren und aggregierte Metriken zu berechnen.

Mit SQL-Analysevorlagen können Sie:

- Standard-SQL-Abfragen schreiben
- Fügen Sie Parameter hinzu, um Ihre Abfragen dynamisch zu gestalten
- Steuern Sie den Zugriff auf bestimmte Spalten und Tabellen
- Legen Sie Aggregationsanforderungen für sensible Daten fest
- Definieren Sie Eingabedaten für die Generierung von synthetischen Datensätzen mit verbessertem Datenschutz für benutzerdefinierte Modelle für maschinelles Lernen (ML)

Topics

- [Erstellen einer SQL-Analysevorlage](#)
- [Überprüfen einer SQL-Analysevorlage](#)

Erstellen einer SQL-Analysevorlage

Voraussetzungen

Bevor Sie eine SQL-Analysevorlage erstellen, müssen Sie über Folgendes verfügen:

- Eine aktive AWS Clean Rooms Zusammenarbeit
- Zugriff auf mindestens eine konfigurierte Tabelle in der Kollaboration

Hinweise zur Konfiguration von Tabellen in AWS Clean Rooms finden Sie unter [Erstellen einer konfigurierten Tabelle in AWS Clean Rooms](#).

- Berechtigungen zum Erstellen von Analysevorlagen
- Grundkenntnisse der SQL-Abfragesyntax

Das folgende Verfahren beschreibt den Prozess der Erstellung einer SQL-Analysevorlage mithilfe der [AWS Clean Rooms Konsole](#).

Informationen zum Erstellen einer SQL-Analysevorlage mithilfe von finden Sie in der [AWS Clean Rooms API-Referenz](#). AWS SDKs

Um eine SQL-Analysevorlage zu erstellen

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#) mit der AWS-Konto , die als Ersteller der Kollaboration fungiert.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus.
4. Gehen Sie auf der Registerkarte Vorlagen zum Abschnitt Von Ihnen erstellte Analysevorlagen.
5. Wählen Sie Analysevorlage erstellen aus.
6. Auf der Seite Analysevorlage erstellen für Details
 - a. Geben Sie einen Namen für die Analysevorlage ein.
 - b. (Optional) Geben Sie eine Beschreibung ein.

- c. Lassen Sie für Format die Option SQL ausgewählt.
7. Sehen Sie sich unter Tabellen die konfigurierten Tabellen an, die der Kollaboration zugeordnet sind.
8. Zur Definition
 - a. Geben Sie die Definition für die Analysevorlage ein.
 - b. Wählen Sie Import aus, um eine Definition zu importieren.
 - c. (Optional) Geben Sie einen Parameter im SQL-Editor an, indem Sie vor dem Parameternamen einen Doppelpunkt (:) eingeben.

Beispiel:

```
WHERE table1.date + :date_period > table1.date
```

9. Wenn Sie zuvor Parameter hinzugefügt haben, wählen Sie unter Parameter — optional für jeden Parameternamen den Typ und den Standardwert (optional) aus.
10. Wenn Sie synthetische Daten für das Modelltraining generieren möchten, aktivieren Sie für synthetische Daten das Kontrollkästchen Synthetische Analysevorlagenausgabe erforderlich.

Weitere Informationen finden Sie unter Generierung [synthetischer Datensätze mit verbessertem Datenschutz](#).

- a. Wählen Sie für die Spaltenklassifizierung eine Spalte aus der Dropdownliste aus. Es sind mindestens fünf Spalten erforderlich.
 - i. Wählen Sie eine Klassifizierung aus der Drop-down-Liste aus. Dadurch wird der Datentyp für jede Spalte identifiziert.

Zu den Klassifizierungstypen gehören:

- Numerisch — Kontinuierliche numerische Werte wie Messungen oder Zählungen
 - Kategorisch — Diskrete Werte oder Kategorien wie Beschriftungen oder Typen
- ii. Um eine Spalte zu entfernen, wählen Sie Entfernen aus.
 - iii. Um eine weitere Spalte hinzuzufügen, wählen Sie Weitere Spalte hinzufügen aus. Wählen Sie die Spalte und die Klassifizierung aus den Drop-down-Listen aus.
 - iv. Wählen Sie für Prädiktiver Wert eine Spalte aus der Dropdownliste aus. Dies ist die Spalte, die das benutzerdefinierte Modell für die Vorhersage verwendet, nachdem es mit dem synthetischen Datensatz trainiert wurde.

- b. In den erweiterten Einstellungen können Sie die Datenschutzzstufe und den Datenschuttschwellenwert festlegen. Passen Sie die Einstellungen an Ihre Bedürfnisse an.
- Geben Sie unter Datenschutzzstufe einen Epsilon-Wert ein, um zu bestimmen, wie viel Rauschen das synthetische Modell zum Schutz der Privatsphäre in Ihrem generierten Datensatz hinzufügt. Der Wert muss zwischen 0,0001 und 10 liegen.
 - Niedrigere Werte erhöhen das Rauschen und sorgen für einen besseren Datenschutz, verringern aber möglicherweise den Nutzen eines nachgeschalteten benutzerdefinierten Modells, das mit diesen Daten trainiert wurde.
 - Höhere Werte sorgen für weniger Rauschen und bieten mehr Korrektheit, beeinträchtigen aber möglicherweise den Datenschutz.

Geben Sie als Datenschuttschwellenwert die höchste zulässige Wahrscheinlichkeit ein, mit der Mitglieder des ursprünglichen Datensatzes durch einen Angriff auf die Mitgliedschaft identifiziert werden könnten. Der Wert muss zwischen 50,0 und 100 liegen.

- Werte von 50% deuten darauf hin, dass ein Angriff auf die Zugehörigkeit Mitglieder nicht besser von Nichtmitgliedern unterscheiden kann als durch eine zufällige Vermutung.
- Wenn Sie keine Datenschutzbegrenzung wünschen, geben Sie 100% ein.

Der optimale Wert hängt von Ihrem speziellen Anwendungsfall und Ihren Datenschutzanforderungen ab. Wenn der Datenschuttschwellenwert überschritten wird, schlägt die Erstellung des ML-Eingangskanals fehl und Sie können den synthetischen Datensatz nicht zum Trainieren eines Modells verwenden.

 Warning

Die Generierung synthetischer Daten schützt davor, auf individuelle Merkmale zu schließen, unabhängig davon, ob bestimmte Personen im ursprünglichen Datensatz vorhanden sind oder Lernattribute dieser Personen vorhanden sind. Es verhindert jedoch nicht, dass wörtliche Werte aus dem ursprünglichen Datensatz, einschließlich persönlich identifizierbarer Informationen (PII), im synthetischen Datensatz erscheinen.

Wir empfehlen, Werte im Eingabedatensatz zu vermeiden, die nur einer betroffenen Person zugeordnet sind, da diese eine betroffene Person neu identifizieren können. Wenn beispielsweise nur ein Benutzer in einer Postleitzahl wohnt, würde das Vorhandensein dieser Postleitzahl im synthetischen Datensatz bestätigen, dass sich der Benutzer im ursprünglichen Datensatz befand. Techniken wie das Kürzen von Werten mit hoher Genauigkeit oder das Ersetzen ungewöhnlicher Kataloge durch andere können verwendet werden, um dieses Risiko zu minimieren. Diese Transformationen können Teil der Abfrage sein, mit der der ML-Eingabekanal erstellt wurde.

11. Wenn Sie Tags für die Ressource aktivieren möchten, wählen Sie Neues Tag hinzufügen und geben Sie dann das Schlüssel - und Wertepaar ein.
12. Wählen Sie Erstellen aus.
13. Sie sind jetzt bereit, Ihr Kollaborationsmitglied darüber zu informieren, dass es [eine Analysevorlage überprüfen](#) kann. (Optional, wenn Sie Ihre eigenen Daten abfragen möchten.)

Überprüfen einer SQL-Analysevorlage

Nachdem ein Mitglied der Kollaboration eine SQLanalysis Vorlage erstellt hat, können Sie sie überprüfen und genehmigen. Nachdem die Analysevorlage genehmigt wurde, kann sie in einer Abfrage in verwendet werden AWS Clean Rooms.

Note

Wenn Sie Ihren Analysecode in eine Kollaboration integrieren, sollten Sie Folgendes beachten:

- AWS Clean Rooms validiert oder garantiert nicht das Verhalten des Analysecodes.
 - Wenn Sie ein bestimmtes Verhalten sicherstellen müssen, überprüfen Sie den Code Ihres Kooperationspartners direkt oder arbeiten Sie mit einem vertrauenswürdigen externen Prüfer zusammen, um ihn zu überprüfen.
- Im gemeinsamen Sicherheitsmodell:
 - Sie (der Kunde) sind für die Sicherheit des Codes verantwortlich, der in der Umgebung ausgeführt wird.
 - AWS Clean Rooms ist für die Sicherheit der Umgebung verantwortlich und stellt sicher, dass
 - nur der genehmigte Code läuft

- nur spezifizierte konfigurierte Tabellen sind zugänglich
- Das einzige Ausgabeziel ist der S3-Bucket des Ergebnisempfängers.

Um eine SQL-Analysevorlage mit der AWS Clean Rooms Konsole zu überprüfen

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#) mit der AWS-Konto , die als Ersteller der Kollaboration fungiert.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus.
4. Gehen Sie auf der Registerkarte Vorlagen zum Abschnitt Analysevorlagen, die von anderen Mitgliedern erstellt wurden.
5. Wählen Sie die Analysevorlage mit dem Status Kann ausgeführt werden auf Nein, Ihre Überprüfung ist erforderlich.
6. Wählen Sie Überprüfen aus.
7. Überprüfen Sie die Übersicht, die Definition und die Parameter der Analyseregeln (falls vorhanden).
8. Überprüfen Sie die konfigurierten Tabellen, die unter In der Definition referenzierte Tabellen aufgeführt sind.

Der Status neben jeder Tabelle lautet Vorlage nicht zulässig.

9. Wählen Sie eine -Tabelle aus.

Wenn Sie	Wählen Sie dann
Genehmigen Sie die Analysevorlage	Vorlage auf der Tabelle zulassen. Bestätigen Sie Ihre Zustimmung, indem Sie Zulassen wählen.
Genehmigen Sie die Analysevorlage nicht	Nicht zulassen

Sie sind jetzt bereit, die konfigurierte Tabelle mithilfe einer SQL-Analysevorlage abzufragen. Weitere Informationen finden Sie unter [Ausführen von SQL-Abfragen](#).

PySpark Analysevorlagen

PySpark Analysevorlagen erfordern ein Python-Benutzerskript und eine optionale virtuelle Umgebung, um benutzerdefinierte Bibliotheken und Open-Source-Bibliotheken verwenden zu können. Diese Dateien werden Artefakte genannt.

Bevor Sie eine Analysevorlage erstellen, erstellen Sie zuerst die Artefakte und speichern sie dann in einem Amazon S3 S3-Bucket. AWS Clean Rooms verwendet diese Artefakte bei der Ausführung von Analyseaufträgen. AWS Clean Rooms greift nur bei der Ausführung eines Jobs auf die Artefakte zu.

Bevor Code auf einer PySpark Analysevorlage ausgeführt wird, werden Artefakte wie AWS Clean Rooms folgt validiert:

- Überprüfung der spezifischen S3-Objektversion, die bei der Erstellung der Vorlage verwendet wurde
- Überprüfung des SHA-256-Hashs des Artefakts
- Fehler bei einem Job, bei dem Artefakte geändert oder entfernt wurden

Note

Die maximale Größe aller kombinierten Artefakte für eine bestimmte PySpark Analysevorlage AWS Clean Rooms beträgt 1 GB.

Sicherheit für PySpark Analysevorlagen

Um eine sichere Datenverarbeitungsumgebung zu gewährleisten, AWS Clean Rooms verwendet es eine zweistufige Rechenarchitektur, um Benutzercode vom Systembetrieb zu isolieren. Diese Architektur basiert auf der Amazon EMR Serverless Fine Grained Access Control-Technologie, auch bekannt als Membrane. Weitere Informationen finden Sie unter [Membrane — Sichere und leistungsstarke Datenzugriffskontrollen in Apache Spark bei Vorhandensein von imperativem Code](#).

Die Komponenten der Rechenumgebung sind in einen separaten Benutzerbereich und einen Systembereich unterteilt. Der Benutzerbereich führt den PySpark Code in der PySpark Analysevorlage aus. AWS Clean Rooms verwendet den Systemspeicher, um die Ausführung des Jobs zu ermöglichen, einschließlich der Verwendung von Servicerollen, die von Kunden bereitgestellt werden, um Daten zur Ausführung des Jobs zu lesen und die Spalte Allowlist zu implementieren. Aufgrund dieser Architektur wird der PySpark Code eines Kunden, der sich auf den Systemspeicher

auswirkt und der eine geringe Anzahl von Spark-SQL und enthalten könnte PySpark DataFrames APIs, blockiert.

PySpark Einschränkungen in AWS Clean Rooms

Wenn Kunden eine genehmigte PySpark Analysevorlage einreichen, AWS Clean Rooms wird diese in einer eigenen sicheren Computerumgebung ausgeführt, auf die kein Kunde zugreifen kann. Die Rechenumgebung implementiert eine Rechenarchitektur mit einem Benutzerbereich und einem Systembereich, um eine sichere Computerumgebung zu gewährleisten. Weitere Informationen finden Sie unter [Sicherheit für PySpark Analysevorlagen](#).

Beachten Sie die folgenden Einschränkungen, bevor Sie PySpark in verwenden AWS Clean Rooms.

Einschränkungen

- Es werden nur DataFrame Ausgaben unterstützt
- Eine einzige Spark-Sitzung pro Jobausführung

Nicht unterstützte Funktionen

- Datenverwaltung
 - Iceberg-Tabellenformate
 - LakeFormation verwaltete Tabellen
 - Resilient Distributed Datasets (RDD)
 - Spark-Streaming
 - Zugriffskontrolle für verschachtelte Spalten
- Benutzerdefinierte Funktionen und Erweiterungen
 - Benutzerdefinierte Tabellenfunktionen () UDTFs
 - Bienenstock UDFs
 - Benutzerdefinierte Klassen in benutzerdefinierten Funktionen
 - Benutzerdefinierte Datenquellen
 - Zusätzliche JAR-Dateien für:
 - Spark-Erweiterungen
 - Konnektoren
 - Metastore-Konfigurationen

- Überwachung und Analyse
 - Spark-Protokollierung
 - Spark-Benutzeroberfläche
 - ANALYZE TABLE-Befehle

Important

Diese Einschränkungen wurden eingeführt, um die Sicherheitsisolierung zwischen Benutzer- und Systembereichen aufrechtzuerhalten.

Alle Einschränkungen gelten unabhängig von der Konfiguration der Zusammenarbeit.

Zukünftige Updates bieten möglicherweise Unterstützung für zusätzliche Funktionen, die auf Sicherheitsbewertungen basieren.

Best Practices

Wir empfehlen die folgenden bewährten Methoden bei der Erstellung von PySpark Analysevorlagen.

- Denken Sie bei der Gestaltung Ihrer Analysevorlagen [PySpark Einschränkungen in AWS Clean Rooms](#) daran.
- Testen Sie Ihren Code zunächst in einer Entwicklungsumgebung.
- Verwenden Sie ausschließlich unterstützte DataFrame Operationen.
- Planen Sie Ihre Ausgabestruktur so, dass sie mit DataFrame Einschränkungen funktioniert.

Wir empfehlen die folgenden bewährten Methoden für die Verwaltung von Artefakten

- Bewahren Sie alle Artefakte der PySpark Analysevorlage in einem speziellen S3-Bucket oder -Präfix auf.
- Verwenden Sie klare Versionsnamen für verschiedene Artefaktversionen.
- Erstellen Sie neue Analysevorlagen, wenn Artefaktaktualisierungen erforderlich sind.
- Führen Sie eine Bestandsaufnahme darüber, welche Vorlagen welche Artefaktversionen verwenden.

Weitere Informationen zum Schreiben von Spark-Code finden Sie im Folgenden:

- [Apache Spark-Beispiele](#)
- [Schreiben Sie eine Spark-Anwendung](#) im Amazon EMR Release Guide
- [Tutorial: Schreiben eines Skripts AWS Glue für Spark](#) im AWS Glue Benutzerhandbuch

In den folgenden Themen wird erklärt, wie Sie Python-Benutzerskripte und -Bibliotheken erstellen, bevor Sie die Analysevorlage erstellen und überprüfen.

Themen

- [Ein Benutzerskript erstellen](#)
- [Arbeiten mit Parametern in PySpark Analysevorlagen](#)
- [Erstellen einer virtuellen Umgebung \(optional\)](#)
- [Speichern eines Benutzerskripts und einer virtuellen Umgebung in S3](#)
- [Eine PySpark Analysevorlage erstellen](#)
- [Überprüfen einer PySpark Analysevorlage](#)

Ein Benutzerskript erstellen

Das Benutzerskript muss eine Einstiegsfunktion (mit anderen Worten, einen Handler) enthalten. Sie können Ihre Benutzerskriptdatei mit einem beliebigen gültigen Python-Dateinamen benennen.

Das folgende Verfahren beschreibt, wie Sie ein Benutzerskript erstellen, um die Kernfunktionen Ihrer PySpark Analyse zu definieren.

Voraussetzungen

- PySpark 1.0 (entspricht Python 3.11 und Spark 3.5.3)
- Datensätze in Amazon S3 können in der von Ihnen definierten Spark-Sitzung nur als konfigurierte Tabellenzuordnungen gelesen werden.
- Ihr Code kann Amazon S3 nicht direkt aufrufen und AWS Glue
- Ihr Code kann keine Netzwerkanrufe tätigen

Um ein Benutzerskript zu erstellen

1. Öffnen Sie einen Texteditor oder eine integrierte Entwicklungsumgebung (IDE) Ihrer Wahl.

Sie können jeden Texteditor oder jede IDE (wie Visual Studio Code oder Notepad++) verwenden PyCharm, die Python-Dateien unterstützt.

2. Erstellen Sie eine neue Python-Datei mit einem Namen Ihrer Wahl (z. B. `my_analysis.py`).
3. Definieren Sie eine Einstiegsfunktion, die einen Kontextobjektparameter akzeptiert.

```
def entrypoint(context)
```

Der `context` Objektparameter ist ein Wörterbuch, das Zugriff auf wichtige Spark-Komponenten, referenzierte Tabellen und Analyseparameter bietet. Sie enthält Folgendes:

Zugriff auf Spark-Sitzungen über `context['sparkSession']`

Referenzierte Tabellen über `context['referencedTables']`

Analyseparameter über `context['analysisParameters']` (falls Parameter in der Vorlage definiert sind)

4. Definieren Sie die Ergebnisse der Entrypoint-Funktion:

```
return results
```

Das `results` muss ein Objekt, das ein Ergebniswörterbuch mit Dateinamen enthält, an eine Ausgabe zurückgeben. `DataFrame`

Note

AWS Clean Rooms schreibt die `DataFrame` Objekte automatisch in den S3-Bucket des Ergebnisempfängers.

5. Sie sind jetzt bereit für:
 - a. Speichern Sie dieses Benutzerskript in S3. Weitere Informationen finden Sie unter [Speichern eines Benutzerskripts und einer virtuellen Umgebung in S3](#).
 - b. Erstellen Sie die optionale virtuelle Umgebung, um alle zusätzlichen Bibliotheken zu unterstützen, die für Ihr Benutzerskript erforderlich sind. Weitere Informationen finden Sie unter [Erstellen einer virtuellen Umgebung \(optional\)](#).

Example Beispiel 1

<caption>The following example demonstrates a generic user script for a PySpark analysis template.</caption>

```
# File name: my_analysis.py

def entrypoint(context):
    try:
        # Access Spark session
        spark = context['sparkSession']

        # Access input tables
        input_table1 = context['referencedTables']['table1_name']
        input_table2 = context['referencedTables']['table2_name']

        # Example data processing operations
        output_df1 = input_table1.select("column1", "column2")
        output_df2 = input_table2.join(input_table1, "join_key")
        output_df3 = input_table1.groupBy("category").count()

        # Return results - each key creates a separate output folder
        return {
            "results": {
                "output1": output_df1,          # Creates output1/ folder
                "output2": output_df2,          # Creates output2/ folder
                "analysis_summary": output_df3 # Creates analysis_summary/ folder
            }
        }

    except Exception as e:
        print(f"Error in main function: {str(e)}")
        raise e
```

Die Ordnerstruktur dieses Beispiels sieht wie folgt aus:

```
analysis_results/
#
### output1/ # Basic selected columns
# ### part-000000.parquet
# ### _SUCCESS
#
### output2/ # Joined data
```

```
# ### part-00000.parquet
# ### _SUCCESS
#
### analysis_summary/ # Aggregated results
### part-00000.parquet
### _SUCCESS
```

Example Beispiel 2

<caption>The following example demonstrates a more complex user script for a PySpark analysis template.</caption>

```
def entrypoint(context):
    try:
        # Get DataFrames from context
        emp_df = context['referencedTables']['employees']
        dept_df = context['referencedTables']['departments']

        # Apply Transformations
        emp_dept_df = emp_df.join(
            dept_df,
            on="dept_id",
            how="left"
        ).select(
            "emp_id",
            "name",
            "salary",
            "dept_name"
        )

        # Return Dataframes
        return {
            "results": {
                "outputTable": emp_dept_df
            }
        }

    except Exception as e:
        print(f"Error in entrypoint function: {str(e)}")
        raise e
```

Arbeiten mit Parametern in PySpark Analysevorlagen

Parameter erhöhen die Flexibilität Ihrer PySpark Analysevorlagen, da bei der Einreichung des Jobs unterschiedliche Werte angegeben werden können. Auf Parameter kann über das Kontextobjekt zugegriffen werden, das an Ihre Einstiegsfunktion übergeben wurde.

Note

Parameter sind vom Benutzer bereitgestellte Zeichenketten, die beliebigen Inhalt enthalten können.

- Überprüfen Sie den Code, um sicherzustellen, dass die Parameter sicher behandelt werden, um unerwartetes Verhalten in Ihrer Analyse zu verhindern.
- Entwerfen Sie die Parameterbehandlung so, dass sie sicher funktioniert, unabhängig davon, welche Parameterwerte zum Zeitpunkt der Einreichung angegeben werden.

Zugreifen auf Parameter

Parameter sind im `context['analysisParameters']` Wörterbuch verfügbar. Alle Parameterwerte sind Zeichenketten.

Example Sicherer Zugriff auf Parameter

```
def entrypoint(context):  
    # Access parameters from context  
    parameters = context['analysisParameters']  
    threshold = parameters['threshold']  
    table_name = parameters['table_name']  
  
    # Continue with analysis using parameters  
    spark = context['sparkSession']  
    input_df = context['referencedTables'][table_name]  
  
    # Convert threshold value  
    threshold_val = int(threshold)  
  
    # Use parameter in DataFrame operation  
    filtered_df = input_df.filter(input_df.amount > threshold_val)  
  
    return {
```

```
"results": {  
    "output": filtered_df  
}
```

Bewährte Methoden für die Parametersicherheit

Warning

Parameter sind vom Benutzer bereitgestellte Zeichenketten, die beliebigen Inhalt enthalten können. Sie müssen mit Parametern sicher umgehen, um Sicherheitslücken in Ihrem Analysecode zu vermeiden.

Unsichere Muster bei der Parameterbehandlung sollten Sie Folgendes vermeiden:

- Parameter als Code ausführen — Verwenden Sie niemals `eval()` oder `exec()` Parameterwerte

```
# UNSAFE - Don't do this  
eval(parameters['expression']) # Can execute arbitrary code
```

- Interpolation von SQL-Zeichenketten — Verketteten Sie niemals Parameter direkt zu SQL-Zeichenfolgen

```
# UNSAFE - Don't do this  
sql = f"SELECT * FROM table WHERE column = '{parameters['value']}'" # SQL injection risk
```

- Unsichere Dateipfadoperationen — Verwenden Sie Parameter niemals direkt in Dateisystemoperationen ohne Überprüfung

```
# UNSAFE - Don't do this  
file_path = f"/data/{parameters['filename']}" # Path traversal risk
```

Muster für die sichere Handhabung von Parametern:

- Verwenden Sie Parameter in DataFrame Operationen — Spark DataFrames geht sicher mit Parameterwerten um


```
# SAFE - Use parameters in DataFrame operations
threshold = int(parameters['threshold'])
filtered_df = input_df.filter(input_df.value > threshold)
```

- Parameterwerte validieren — Stellen Sie vor der Verwendung sicher, dass die Parameter den erwarteten Formaten entsprechen

```
# SAFE - Validate parameters before use
def validate_date(date_str):
    try:
        from datetime import datetime
        datetime.strptime(date_str, '%Y-%m-%d')
        return True
    except ValueError:
        return False

date_param = parameters['date_filter'] or '2024-01-01'
if not validate_date(date_param):
    raise ValueError(f"Invalid date format: {date_param}")
```

- Zulässige Listen für Parameterwerte verwenden — Überprüfen Sie die Parameter nach Möglichkeit anhand von zweifelsfrei funktionierenden Werten

```
# SAFE - Use allowlists
allowed_columns = ['column1', 'column2', 'column3']
column_param = parameters['column_name']
if column_param not in allowed_columns:
    raise ValueError(f"Invalid column: {column_param}")
```

- Typkonvertierung mit Fehlerbehandlung — Konvertiere Zeichenkettenparameter sicher in erwartete Typen

```
# SAFE - Convert with error handling
try:
    batch_size = int(parameters['batch_size'] or '1000')
    if batch_size <= 0 or batch_size > 10000:
        raise ValueError(f"Batch size must be between 1 and 10000")
except ValueError as e:
    print(f"Invalid parameter: {e}")
    raise
```

⚠ Important

Denken Sie daran, dass Parameter die Codeüberprüfung umgehen, wenn Job Runner unterschiedliche Werte angeben. Gestalten Sie Ihre Parameterbehandlung so, dass sie unabhängig davon, welche Parameterwerte bereitgestellt werden, sicher funktioniert.

Vollständiges Beispiel für Parameter

Example Sicheres Verwenden von Parametern in einem PySpark Skript

```
def entrypoint(context):
    try:
        # Access Spark session and tables
        spark = context['sparkSession']
        input_table = context['referencedTables']['sales_data']

        # Access parameters - fail fast if analysisParameters missing
        parameters = context['analysisParameters']

        # Validate and convert numeric parameter (handles empty strings with default)
        try:
            threshold = int(parameters['threshold'] or '100')
            if threshold <= 0:
                raise ValueError("Threshold must be positive")
        except (ValueError, TypeError) as e:
            print(f"Invalid threshold parameter: {e}")
            raise

        # Validate date parameter (handles empty strings with default)
        date_filter = parameters['start_date'] or '2024-01-01'
        from datetime import datetime
        try:
            datetime.strptime(date_filter, '%Y-%m-%d')
        except ValueError:
            raise ValueError(f"Invalid date format: {date_filter}")

        # Use parameters safely in DataFrame operations
        filtered_df = input_table.filter(
            (input_table.amount > threshold) &
            (input_table.date >= date_filter)
        )
```

```
result_df = filtered_df.groupBy("category").agg(
    {"amount": "sum"}
)

return {
    "results": {
        "filtered_results": result_df
    }
}

except Exception as e:
    print(f"Error in analysis: {str(e)}")
    raise
```

Erstellen einer virtuellen Umgebung (optional)

Wenn Ihr Benutzerskript zusätzliche Bibliotheken benötigt, haben Sie die Möglichkeit, eine virtuelle Umgebung zum Speichern dieser Bibliotheken zu erstellen. Wenn Sie keine zusätzlichen Bibliotheken benötigen, können Sie diesen Schritt überspringen.

Wenn Sie mit Bibliotheken arbeiten, die native Erweiterungen haben, denken Sie PySpark daran, dass sie unter Linux mit ARM64 Architektur AWS Clean Rooms funktioniert.

Das folgende Verfahren zeigt, wie Sie mit einem einfachen CLI-Befehl eine virtuelle Umgebung erstellen.

Um eine virtuelle Umgebung zu erstellen

1. Öffnen Sie ein Terminal oder eine Befehlszeile.
2. Fügen Sie den folgenden Inhalt hinzu:

```
# create and activate a python virtual environment
python3 -m venv pyspark_venvsource
source pyspark_venvsource/bin/activate

# install the python packages
pip3 install pandas # add packages here

# package the virtual environment into an archive
pip3 install venv-pack
venv-pack -f -o pyspark_venv.tar.gz
```

```
# optionally, remove the virtual environment directory
deactivate
rm -fr pyspark_venvsource
```

3. Sie sind jetzt bereit, diese virtuelle Umgebung in S3 zu speichern. Weitere Informationen finden Sie unter [Speichern eines Benutzerskripts und einer virtuellen Umgebung in S3](#).

Weitere Informationen zur Arbeit mit Docker und Amazon ECR finden Sie im [ECRUser Amazon-Leitfaden](#).

Speichern eines Benutzerskripts und einer virtuellen Umgebung in S3

Das folgende Verfahren erklärt, wie ein Benutzerskript und eine optionale virtuelle Umgebung in Amazon S3 gespeichert werden. Schließen Sie diesen Schritt ab, bevor Sie eine PySpark Analysevorlage erstellen.

Important

Ändern oder entfernen Sie keine Artefakte (Benutzerskripte oder virtuelle Umgebungen), nachdem Sie eine Analysevorlage erstellt haben.

Dadurch wird:

- Verursacht, dass alle future Analysejobs, die diese Vorlage verwenden, fehlschlagen.
- Erfordert die Erstellung einer neuen Analysevorlage mit neuen Artefakten.
- Wirkt sich nicht auf zuvor abgeschlossene Analyseaufträge aus

Voraussetzungen

- Und AWS-Konto mit den entsprechenden Berechtigungen
- Eine Benutzerskriptdatei (`wiemy_analysis.py`)
- (Optional, falls vorhanden) Ein virtuelles Umgebungspaket (`.tar.gz` Datei)
- Zugriff zum Erstellen oder Ändern von IAM-Rollen

Console

Um ein Benutzerskript und eine virtuelle Umgebung in S3 mithilfe der Konsole zu speichern:

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die Amazon S3 S3-Konsole unter <https://console.aws.amazon.com/s3/>.
2. Erstellen Sie einen neuen S3-Bucket oder verwenden Sie einen vorhandenen.
3. Aktivieren Sie die Versionierung für den Bucket.
 - a. Wählen Sie Ihren Bucket aus.
 - b. Wählen Sie Properties (Eigenschaften).
 - c. Wählen Sie im Abschnitt Bucket-Versionierung die Option Bearbeiten aus.
 - d. Wählen Sie Aktivieren und speichern Sie die Änderungen.
4. Laden Sie Ihre Artefakte hoch und aktivieren Sie den SHA-256-Hash.
 - a. Navigiere zu deinem Bucket.
 - b. Klicken Sie auf Upload.
 - c. Wählen Sie Dateien hinzufügen und fügen Sie Ihre Benutzerskriptdatei hinzu.
 - d. (Optional, falls vorhanden) Fügen Sie Ihre .tar.gz-Datei hinzu.
 - e. Erweitern Sie Eigenschaften.
 - f. Wählen Sie unter Prüfsummen für Prüfsummenfunktion die Option aus. SHA256
 - g. Klicken Sie auf Upload.
5. Sie sind jetzt bereit, eine PySpark Analysevorlage zu erstellen.

CLI

Um das Benutzerskript und die virtuelle Umgebung in S3 zu speichern, verwenden Sie AWS CLI:

1. Führen Sie den folgenden Befehl aus:

```
aws s3 cp --checksum-algorithm sha256 pyspark_venv.tar.gz s3://ARTIFACT-BUCKET/  
EXAMPLE-PREFIX/
```

2. Sie sind jetzt bereit, eine PySpark Analysevorlage zu erstellen.

 Note

Wenn Sie das Skript oder die virtuelle Umgebung aktualisieren müssen:

1. Laden Sie die neue Version als separates Objekt hoch.
2. Erstellen Sie eine neue Analysevorlage mit den neuen Artefakten.
3. Verwerfen Sie die alte Vorlage.
4. Behalten Sie die ursprünglichen Artefakte in S3 bei, falls die alte Vorlage möglicherweise noch benötigt wird.

Eine PySpark Analysevorlage erstellen

 Note

Parameter sind vom Benutzer bereitgestellte Zeichenfolgen, die beliebigen Inhalt enthalten können.

- Überprüfen Sie den Code, um sicherzustellen, dass die Parameter sicher behandelt werden, um unerwartetes Verhalten in Ihrer Analyse zu verhindern.
- Entwerfen Sie die Parameterbehandlung so, dass sie sicher funktioniert, unabhängig davon, welche Parameterwerte zum Zeitpunkt der Einreichung angegeben werden.

Voraussetzungen

Bevor Sie eine PySpark Analysevorlage erstellen, müssen Sie über Folgendes verfügen:

- Eine Mitgliedschaft in einer aktiven AWS Clean Rooms Kollaboration
- Zugriff auf mindestens eine konfigurierte Tabelle in der aktiven Kollaboration
- Berechtigungen zum Erstellen von Analysevorlagen
- Ein Python-Benutzerskript und eine virtuelle Umgebung, die in S3 erstellt und gespeichert wurden
 - Für den S3-Bucket ist die Versionierung aktiviert. Weitere Informationen finden Sie unter [Verwenden der Versionierung in S3-Buckets](#)
 - Der S3-Bucket kann SHA-256-Prüfsummen für hochgeladene Artefakte berechnen. [Weitere Informationen finden Sie unter Prüfsummen verwenden](#)

- Berechtigungen zum Lesen von Code aus einem S3-Bucket

Hinweise zum Erstellen der erforderlichen Servicerolle finden Sie unter [Erstellen Sie eine Servicerolle, um Code aus einem S3-Bucket zu lesen \(PySpark Analysevorlagenrolle\)](#).

Das folgende Verfahren beschreibt den Prozess der Erstellung einer PySpark Analysevorlage mithilfe der [AWS Clean Rooms Konsole](#). Es wird davon ausgegangen, dass Sie bereits ein Benutzerskript und virtuelle Umgebungsdateien erstellt und Ihr Benutzerskript und Ihre virtuellen Umgebungsdateien in einem Amazon S3 S3-Bucket gespeichert haben.

 Note

Das Mitglied, das die PySpark Analysevorlage erstellt, muss auch das Mitglied sein, das die Ergebnisse erhält.

Informationen zum Erstellen einer PySpark Analysevorlage mithilfe von finden Sie in der [AWS Clean Rooms API-Referenz](#). AWS SDKs

Um eine PySpark Analysevorlage zu erstellen

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#) mit der AWS-Konto , die als Ersteller der Kollaboration fungiert.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus.
4. Gehen Sie auf der Registerkarte Vorlagen zum Abschnitt Von Ihnen erstellte Analysevorlagen.
5. Wählen Sie Analysevorlage erstellen aus.
6. Auf der Seite Analysevorlage erstellen für Details
 - a. Geben Sie einen Namen für die Analysevorlage ein.
 - b. (Optional) Geben Sie eine Beschreibung ein.
 - c. Wählen Sie für Format die PySparkOption aus.
7. Für Definition
 - a. Überprüfen Sie die Voraussetzungen und stellen Sie sicher, dass alle Voraussetzungen erfüllt sind, bevor Sie fortfahren.
 - b. Geben Sie als Einstiegspunktdatei den S3-Bucket ein oder wählen Sie Browse S3.

- c. (Optional) Geben Sie für die Bibliotheksdatei den S3-Bucket ein oder wählen Sie Browse S3 aus.
8. Für Parameter — optional, wenn Sie Parameter hinzufügen möchten, um Ihre Analysevorlage wiederverwendbar zu machen:
 - a. Wählen Sie Parameter hinzufügen aus.
 - b. Geben Sie einen Parameternamen ein.

Parameternamen müssen mit einem Buchstaben oder Unterstrich beginnen, gefolgt von alphanumerischen Zeichen oder Unterstrichen.

- c. Für Type wird STRING automatisch als einziger unterstützter Typ für PySpark Analysevorlagen ausgewählt.
 - d. (Optional) Geben Sie einen Standardwert für den Parameter ein.

Wenn Sie einen Standardwert angeben, können Job-Runner diesen Wert verwenden, wenn sie Jobs ausführen, ohne explizit einen Parameterwert anzugeben.

- e. Um weitere Parameter hinzuzufügen, wählen Sie Anderen Parameter hinzufügen und wiederholen Sie die vorherigen Schritte.

 Note

Sie können bis zu 50 Parameter pro PySpark Analysevorlage definieren. Jeder Parameterwert kann bis zu 1.000 Zeichen lang sein.

9. Für Tabellen, auf die in der Definition verwiesen wird,
 - Wenn alle in der Definition referenzierten Tabellen der Kollaboration zugeordnet wurden:
 - Lassen Sie das Kontrollkästchen Alle in der Definition referenzierten Tabellen wurden der Kollaboration zugeordnet aktiviert.
 - Wählen Sie unter Mit der Kollaboration verknüpfte Tabellen alle verknüpften Tabellen aus, auf die in der Definition verwiesen wird.
 - Wenn nicht alle Tabellen, auf die in der Definition verwiesen wird, der Kollaboration zugeordnet wurden:
 - Deaktivieren Sie das Kontrollkästchen Alle in der Definition referenzierten Tabellen wurden der Kollaboration zugeordnet.

- Wählen Sie unter Der Kollaboration zugeordnete Tabellen alle verknüpften Tabellen aus, auf die in der Definition verwiesen wird.
- Geben Sie unter Tabellen, die später verknüpft werden, einen Tabellennamen ein.
- Wählen Sie Andere Tabelle auflisten aus, um eine weitere Tabelle aufzulisten.

10. Wählen Sie für die Konfiguration von Fehlermeldungen eine der folgenden Optionen:

- Grundlegende Fehlermeldungen — gibt grundlegende Fehlermeldungen zurück, ohne die zugrunde liegenden Daten offenzulegen. Empfohlen für Produktionsworkloads.
- Detaillierte Fehlermeldungen — gibt detaillierte Fehlermeldungen für eine schnellere Problembehebung zurück. In Entwicklungs- und Testumgebungen empfohlen. Kann sensible Daten, einschließlich persönlich identifizierbarer Informationen (PII), offenlegen.

 Note

Wenn Sie detaillierte Fehlermeldungen verwenden, müssen alle Mitglieder des Datenanbieters diese Einstellung für die Vorlage genehmigen.

11. Geben Sie die Dienstzugriffsberechtigungen an, indem Sie einen vorhandenen Servicerollennamen aus der Dropdownliste auswählen.

1. Die Liste der Rollen wird angezeigt, wenn Sie berechtigt sind, Rollen aufzulisten.

Wenn Sie nicht berechtigt sind, Rollen aufzulisten, können Sie den Amazon-Ressourcennamen (ARN) der Rolle eingeben, die Sie verwenden möchten.

2. Rufen Sie die Servicerolle auf, indem Sie auf den externen Link In IAM anzeigen klicken.

Wenn keine vorhandenen Servicerollen vorhanden sind, ist die Option „Eine bestehende Servicerolle verwenden“ nicht verfügbar.

Versucht standardmäßig AWS Clean Rooms nicht, die bestehende Rollenrichtlinie zu aktualisieren, um die erforderlichen Berechtigungen hinzuzufügen.

 Note

- AWS Clean Rooms erfordert Berechtigungen für Abfragen gemäß den Analyseregeln. Weitere Informationen zu Berechtigungen für AWS Clean Rooms finden Sie unter [AWS verwaltete Richtlinien für AWS Clean Rooms](#).
- Wenn die Rolle nicht über ausreichende Berechtigungen für verfügt AWS Clean Rooms, erhalten Sie eine Fehlermeldung, dass die Rolle nicht über ausreichende Berechtigungen für verfügt AWS Clean Rooms. Die Rollenrichtlinie muss hinzugefügt werden, bevor Sie fortfahren können.
- Wenn Sie die Rollenrichtlinie nicht ändern können, erhalten Sie eine Fehlermeldung, dass die Richtlinie für die Servicerolle nicht gefunden werden AWS Clean Rooms konnte.

12. Wenn Sie Tags für die konfigurierte Tabellenressource aktivieren möchten, wählen Sie Neues Tag hinzufügen aus und geben Sie dann das Schlüssel - und Wertepaar ein.
13. Wählen Sie Erstellen aus.
14. Sie sind jetzt bereit, Ihr Kollaborationsmitglied darüber zu informieren, dass es [eine Analysevorlage überprüfen](#) kann. (Optional, wenn Sie Ihre eigenen Daten abfragen möchten.)

 Important

Ändern oder entfernen Sie keine Artefakte (Benutzerskripte oder virtuelle Umgebungen), nachdem Sie eine Analysevorlage erstellt haben.

Dadurch wird:

- Verursacht, dass alle future Analysejobs, die diese Vorlage verwenden, fehlschlagen.
- Erfordert die Erstellung einer neuen Analysevorlage mit neuen Artefakten.
- Wirkt sich nicht auf zuvor abgeschlossene Analyseaufträge aus.

Überprüfen einer PySpark Analysevorlage

Wenn ein anderes Mitglied in Ihrer Kollaboration eine Analysevorlage erstellt, müssen Sie diese überprüfen und genehmigen, bevor sie verwendet werden kann.

Das folgende Verfahren zeigt Ihnen, wie Sie eine PySpark Analysevorlage einschließlich ihrer Regeln, Parameter und referenzierten Tabellen überprüfen. Als Mitglied der Kollaboration werden Sie beurteilen, ob die Vorlage Ihren Vereinbarungen zur gemeinsamen Nutzung von Daten und Ihren Sicherheitsanforderungen entspricht.

Nachdem die Analysevorlage genehmigt wurde, kann sie in einem Job in AWS Clean Rooms verwendet werden.

Note

Wenn Sie Ihren Analysecode in eine Kollaboration einbringen, sollten Sie Folgendes beachten:

- AWS Clean Rooms validiert oder garantiert nicht das Verhalten des Analysecodes.
 - Wenn Sie ein bestimmtes Verhalten sicherstellen müssen, überprüfen Sie den Code Ihres Kooperationspartners direkt oder arbeiten Sie mit einem vertrauenswürdigen externen Prüfer zusammen, um ihn zu überprüfen.
- AWS Clean Rooms garantiert, dass die SHA-256-Hashes des in der PySpark Analysevorlage aufgeführten Codes mit dem Code übereinstimmen, der in der Analyseumgebung ausgeführt wird. PySpark
- AWS Clean Rooms führt keine Prüfungen oder Sicherheitsanalysen zusätzlicher Bibliotheken durch, die Sie in die Umgebung einbringen.
- Im gemeinsamen Sicherheitsmodell:
 - Sie (der Kunde) sind für die Sicherheit des Codes verantwortlich, der in der Umgebung ausgeführt wird.
 - Sie (der Kunde) sind dafür verantwortlich, die entsprechende Konfiguration für Fehlermeldungen für die Umgebung festzulegen.
 - AWS Clean Rooms ist für die Sicherheit der Umgebung verantwortlich und stellt sicher, dass
 - nur der genehmigte Code läuft
 - nur spezifizierte konfigurierte Tabellen sind zugänglich
 - Das einzige Ausgabeziel ist der S3-Bucket des Ergebnisempfängers.

AWS Clean Rooms generiert SHA-256-Hashes des Benutzerskripts und der virtuellen Umgebung zur Überprüfung. Auf das eigentliche Benutzerskript und die Bibliotheken kann jedoch nicht direkt von dort aus zugegriffen werden. AWS Clean Rooms

Um zu überprüfen, ob das Benutzerskript und die gemeinsam genutzten Bibliotheken mit denen identisch sind, auf die in der Analysevorlage verwiesen wird, können Sie einen SHA-256-Hash der gemeinsam genutzten Dateien erstellen und ihn mit dem Hash der Analysevorlage vergleichen, der von erstellt wurde. AWS Clean Rooms Die Hashes der Codeausführung werden auch in den Job-Logs aufgeführt.

Voraussetzungen

- Linux/Unix-Betriebssystem oder Windows-Subsystem für Linux (WSL)
- Benutzerskriptdatei, die Sie hashen möchten
 - Bitten Sie den Ersteller der Analysevorlage, die Datei über einen sicheren Kanal weiterzugeben.
- Der Hash der Analysevorlage, erstellt von AWS Clean Rooms

Um eine PySpark Analysevorlage mit der AWS Clean Rooms Konsole zu überprüfen

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#) mit der AWS-Konto , die als Ersteller der Kollaboration fungiert.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus.
4. Gehen Sie auf der Registerkarte Vorlagen zum Abschnitt Analysevorlagen, die von anderen Mitgliedern erstellt wurden.
5. Wählen Sie die Analysevorlage mit dem Status Kann ausgeführt werden auf Nein, Ihre Überprüfung ist erforderlich.
6. Wählen Sie Überprüfen aus.
7. Überprüfen Sie die Übersicht, die Definition und die Parameter der Analyseregeln (falls vorhanden).

Note

Mithilfe von Parametern können Analyseläufer zum Zeitpunkt der Übermittlung unterschiedliche Werte einreichen. Wenn eine Analysevorlage Parameter unterstützt,

überprüfen Sie, wie die Parameterwerte im Code Ihres Kooperationspartners verwendet werden, um sicherzustellen, dass er Ihren Anforderungen entspricht.

8. Stellen Sie sicher, dass das gemeinsam genutzte Benutzerskript und die Bibliotheken mit denen identisch sind, auf die in der Analysevorlage verwiesen wird.

- a. Erstellen Sie einen SHA-256-Hash der gemeinsam genutzten Dateien und vergleichen Sie ihn mit dem Hash der Analysevorlage, der von erstellt wurde. AWS Clean Rooms

Sie können einen Hash generieren, indem Sie zu dem Verzeichnis navigieren, das Ihre Benutzerskriptdatei enthält, und dann den folgenden Befehl ausführen:

```
sha256sum your_script_filename.py
```

Beispielausgabe:

```
e3b0c44298fc1c149afb4c8996fb92427ae41e4649b934ca495991b7852b855 my_analysis.py
```

- b. Alternativ können Sie die Amazon S3-Prüfsummenfunktionen verwenden. Weitere Informationen finden Sie unter [Überprüfen der Objektintegrität in Amazon S3](#) im Amazon S3 S3-Benutzerhandbuch.
 - c. Eine weitere Alternative besteht darin, die Hashes des ausgeführten Codes in den Jobprotokollen einzusehen.
9. Sehen Sie sich die konfigurierten Tabellen an, die unter In der Definition referenzierte Tabellen aufgeführt sind.

Der Status neben jeder Tabelle lautet Vorlage nicht zulässig.

10. Wählen Sie eine -Tabelle aus.

- a. Um die Analysevorlage zu genehmigen, wählen Sie Vorlage in Tabelle zulassen aus. Bestätigen Sie Ihre Genehmigung, indem Sie Zulassen wählen.
 - b. Um die Genehmigung abzulehnen, wählen Sie Ablehnen.

Wenn Sie sich dafür entschieden haben, die Analysevorlage zu genehmigen, kann das Mitglied, das Jobs ausführen kann, nun mithilfe einer PySpark Analysevorlage einen PySpark Job in einer konfigurierten Tabelle ausführen. Weitere Informationen finden Sie unter [PySpark Jobs werden ausgeführt](#).

PySpark Analysevorlagen zur Fehlerbehebung

Wenn Sie Jobs mithilfe von PySpark Analysevorlagen ausführen, kann es bei der Initialisierung oder Ausführung von Jobs zu Fehlern kommen. Diese Fehler beziehen sich in der Regel auf die Skriptkonfiguration, die Datenzugriffsberechtigungen oder die Einrichtung der Umgebung.

Weitere Informationen zu PySpark Einschränkungen finden Sie unter [PySpark Einschränkungen in AWS Clean Rooms](#).

Themen

- [Problembehandlung bei Ihrem Code](#)
- [Der Job mit der Analysevorlage wird nicht gestartet](#)
- [Der Job für die Analysevorlage wird gestartet, schlägt aber bei der Verarbeitung fehl](#)
- [Die Einrichtung der virtuellen Umgebung schlägt fehl](#)

Problembehandlung bei Ihrem Code

Um Ihnen bei der Entwicklung und Problembehebung Ihres Codes zu helfen, empfehlen wir Ihnen, selbst zu simulieren AWS Clean Rooms , AWS-Konto indem Sie detaillierte Fehlermeldungen aktivieren und Jobs mit Ihren eigenen Testdaten ausführen.

Sie können es auch mit den folgenden Schritten PySpark AWS Clean Rooms in Amazon EMR Serverless simulieren. Es wird kleine Unterschiede zu PySpark AWS Clean Rooms haben, behandelt aber hauptsächlich, wie Ihr Code ausgeführt werden kann.

Um PySpark AWS Clean Rooms in EMR Serverless zu simulieren

1. Erstellen Sie einen Datensatz in Amazon S3, katalogisieren Sie ihn in der AWS Glue Data Catalog und richten Sie Lake Formation Formation-Berechtigungen ein.
2. Registrieren Sie den S3-Standort mit einer benutzerdefinierten Rolle bei Lake Formation.
3. Erstellen Sie eine Amazon EMR Studio-Instance, falls Sie noch keine haben (Amazon EMR Studio ist erforderlich, um Amazon EMR Serverless zu verwenden).
4. Eine serverlose EMR-App erstellen
 - Wählen Sie die Release-Version emr-7.7.0 aus.

- Wählen Sie Architektur aus. ARM64
 - Entscheiden Sie sich für Benutzerdefinierte Einstellungen verwenden.
 - Deaktivieren Sie die vorinitialisierte Kapazität.
 - Wenn Sie interaktiv arbeiten möchten, wählen Sie Interaktiver Endpunkt > Endpunkt für EMR Studio aktivieren.
 - Wählen Sie Zusätzliche Konfigurationen > Lake Formation für eine detaillierte Zugriffskontrolle verwenden.
 - Erstellen Sie die Anwendung.
5. Verwenden Sie EMR-S entweder über EMR-Studio-Notebooks oder die API. `StartJobRun`

Der Job mit der Analysevorlage wird nicht gestartet

Häufige Ursachen

Jobs mit Analysevorlagen können aufgrund von drei Hauptkonfigurationsproblemen sofort beim Start fehlschlagen:

- Falsche Skriptbenennung, die nicht dem erforderlichen Format entspricht
- Fehlende oder falsch formatierte Einstiegspunktfunktion im Benutzerskript
- Inkompatible Python-Version in der virtuellen Umgebung

Auflösung

Um dies zu lösen:

1. Überprüfen Sie Ihr Benutzerskript:
 - Überprüfen Sie, ob Ihr Benutzerskript einen gültigen Python-Dateinamen hat.

Gültige Python-Dateinamen verwenden Kleinbuchstaben, Unterstriche zur Trennung von Wörtern und die Erweiterung.py.
2. Überprüfen Sie die Einstiegspunktfunktion. Wenn Ihr Benutzerskript keine Einstiegsfunktion hat, fügen Sie eine hinzu.
 - a. Öffnen Sie Ihr Benutzerskript.
 - b. Fügen Sie diese Einstiegsfunktion hinzu:

```
def entrypoint(context):  
    # Your analysis code here
```

- c. Stellen Sie sicher, dass der Funktionsname genau wie geschrieben ist. `entrypoint`
 - d. Stellen Sie sicher, dass die Funktion den `context` Parameter akzeptiert.
3. Überprüfen Sie die Kompatibilität der Python-Version:
- a. Stellen Sie sicher, dass Ihre virtuelle Umgebung Python 3.9 oder 3.11 verwendet.
 - b. Führen Sie Folgendes aus, um Ihre Version zu überprüfen: `python --version`
 - c. Aktualisieren Sie bei Bedarf Ihre virtuelle Umgebung:

```
conda create -n analysis-env python=3.9  
conda activate analysis-env
```

Prävention

- Verwenden Sie den bereitgestellten Startercode für die Analysevorlage, der die richtige Dateistruktur enthält.
- Richten Sie eine dedizierte virtuelle Umgebung mit Python 3.9 oder 3.11 für alle Analysevorlagen ein.
- Testen Sie Ihre Analysevorlage lokal mit dem Tool zur Vorlagenvalidierung, bevor Sie Jobs einreichen.
- Implementieren Sie CI/CD Prüfungen, um die Anforderungen an die Skriptbenennung und die Einstiegsfunktionen zu überprüfen.

Der Job für die Analysevorlage wird gestartet, schlägt aber bei der Verarbeitung fehl

Häufige Ursachen

Analyseaufträge können aus folgenden Sicherheits- und Formatierungsgründen während der Ausführung fehlschlagen:

- Unbefugte Direktzugriffsversuche auf AWS Dienste wie Amazon S3 oder AWS Glue

- Rückgabe der Ausgabe in falschen Formaten, die nicht den erforderlichen DataFrame Spezifikationen entsprechen
- Blockierte Netzwerkanrufe aufgrund von Sicherheitseinschränkungen in der Ausführungsumgebung

Auflösung

Um das Problem zu lösen

1. Direkten AWS Servicezugriff entfernen:
 - a. Suchen Sie in Ihrem Code nach direkten AWS Serviceimporten und Aufrufen.
 - b. Ersetzen Sie den direkten S3-Zugriff durch die bereitgestellten Spark-Sitzungsmethoden.
 - c. Verwenden Sie nur vorkonfigurierte Tabellen über die Kollaborationsoberfläche.
2. Formatieren Sie die Ausgaben korrekt:
 - a. Stellen Sie sicher, dass alle Ausgaben von Spark stammen DataFrames.
 - b. Aktualisieren Sie Ihre Rücksendeerklärung so, dass sie diesem Format entspricht:

```
return {  
    "results": {  
        "output1": dataframe1  
    }  
}
```

- c. Entfernen Sie alle Objekte, die nicht DataFrame zurückgegeben wurden.
3. Netzwerkanrufe entfernen:
 - a. Identifizieren und entfernen Sie alle externen API-Aufrufe.
 - b. Entfernen Sie alle URLLIB, Anfragen oder ähnliche Netzbibliothek.
 - c. Entfernen Sie alle Socket-Verbindungen oder den HTTP-Client-Code.

Prävention

- Verwenden Sie den mitgelieferten Code-Linter, um nach nicht autorisierten AWS Importen und Netzwerkaufrufen zu suchen.

- Testen Sie Jobs in der Entwicklungsumgebung, in der Sicherheitseinschränkungen der Produktion entsprechen.
- Folgen Sie dem Validierungsprozess für das Ausgabeschema, bevor Sie Jobs bereitstellen.
- Lesen Sie die Sicherheitsrichtlinien für genehmigte Dienstzugriffsmuster.

Die Einrichtung der virtuellen Umgebung schlägt fehl

Häufige Ursachen

Konfigurationsfehler in der virtuellen Umgebung treten häufig aufgrund folgender Ursachen auf:

- Nicht übereinstimmende CPU-Architektur zwischen Entwicklungs- und Ausführungsumgebungen
- Probleme mit der Formatierung von Python-Code, die eine korrekte Initialisierung der Umgebung verhindern
- Falsche Konfiguration des Basis-Images in den Container-Einstellungen

Auflösung

Um das Problem zu lösen

1. Konfigurieren Sie die richtige Architektur:

- a. Überprüfen Sie Ihre aktuelle Architektur mit `uname -m`.
- b. Aktualisieren Sie Ihr Dockerfile, um Folgendes anzugeben: ARM64

```
FROM --platform=linux/arm64 public.ecr.aws/amazonlinux/amazonlinux:2023-minimal
```

- c. Erstellen Sie Ihren Container neu mit `docker build --platform=linux/arm64`.

2. Python-Einrückung korrigieren:

- a. Führen Sie einen Python-Codeformatierer wie `black` in Ihren Codedateien aus.
- b. Stellen Sie sicher, dass Leerzeichen oder Tabulatoren (nicht beide) einheitlich verwendet werden.
- c. Überprüfen Sie die Einrückung aller Codeblöcke:

```
def my_function():  
    if condition:
```

```
do_something()  
return result
```

- d. Verwenden Sie eine IDE mit Python-Einrückungshervorhebung.
3. Überprüfen Sie die Umgebungskonfiguration:
 - a. Ausführen `python -m py_compile your_script.py`, um nach Syntaxfehlern zu suchen.
 - b. Testen Sie die Umgebung vor der Bereitstellung lokal.
 - c. Stellen Sie sicher, dass alle Abhängigkeiten unter `requirements.txt` aufgeführt sind.

Prävention

- Verwenden Sie Visual Studio Code oder PyCharm mit Python-Formatierungs-Plugins
- Konfigurieren Sie Pre-Commit-Hooks, um Codeformatierer automatisch auszuführen
- Erstellen und testen Sie Umgebungen lokal mit dem bereitgestellten Basis-Image ARM64
- Implementieren Sie die automatische Überprüfung des Codestils in Ihrer CI/CD Pipeline

Analysieren Sie Daten in einer Zusammenarbeit

In können Sie Daten analysieren AWS Clean Rooms, indem Sie Abfragen oder Jobs ausführen.

Eine Abfrage ist eine Methode, um auf konfigurierte Tabellen in einer Kollaboration zuzugreifen und diese zu analysieren, wobei ein unterstützter Satz von Funktionen, Klassen und Variablen verwendet wird. Die derzeit unterstützte Abfragesprache AWS Clean Rooms ist SQL. Es gibt drei Möglichkeiten, eine Abfrage auszuführen AWS Clean Rooms: Schreiben Sie SQL-Code, verwenden Sie eine zugelassene SQL-Analysevorlage oder verwenden Sie die Analysis Builder-Benutzeroberfläche.

Ein Job ist eine Methode zum Zugreifen auf konfigurierte Tabellen und zum Analysieren von konfigurierten Tabellen in einer Kollaboration mithilfe eines unterstützten Satzes von Funktionen, Klassen und Variablen. Der derzeit unterstützte Jobtyp in AWS Clean Rooms ist PySpark. Es gibt eine Möglichkeit, einen Job auszuführen AWS Clean Rooms: mithilfe einer genehmigten PySpark Analysevorlage.

In den folgenden Themen wird beschrieben, wie Sie Daten analysieren, AWS Clean Rooms indem Sie SQL-Abfragen oder PySpark -Jobs ausführen.

Topics

- [Ausführen von SQL-Abfragen](#)
- [PySpark Jobs werden ausgeführt](#)

Ausführen von SQL-Abfragen

Note

Sie können Abfragen nur ausführen, wenn das Mitglied, das für die Berechnung von Abfragen verantwortlich ist, der Kollaboration als aktives Mitglied beigetreten ist.

Als [Mitglied, das Abfragen durchführen kann](#), können Sie eine SQL-Abfrage wie folgt ausführen:

- Manuelles Erstellen einer SQL-Abfrage mithilfe des SQL-Code-Editors.
- Verwendung einer genehmigten [SQL-Analysevorlage](#).
- Verwenden der Analysis Builder-Benutzeroberfläche, um eine Abfrage zu erstellen, ohne SQL-Code schreiben zu müssen.

Wenn das Mitglied, das Abfragen durchführen kann, eine SQL-Abfrage für die Tabellen in der Kollaboration ausführt, AWS Clean Rooms übernimmt es die entsprechenden Rollen, um in seinem Namen auf die Tabellen zuzugreifen. AWS Clean Rooms wendet die Analyseregeln nach Bedarf auf die Eingabeabfrage und ihre Ausgabe an.

Die Analyseregeln und Ausgabebeschränkungen werden automatisch durchgesetzt. AWS Clean Rooms gibt nur die Ergebnisse zurück, die den definierten Analyseregeln entsprechen.

AWS Clean Rooms unterstützt SQL-Abfragen, die sich von anderen Abfrage-Engines unterscheiden können. Spezifikationen finden Sie in der [AWS Clean Rooms SQL-Referenz](#). Wenn Sie Abfragen für Datentabellen ausführen möchten, die mit Differential Privacy geschützt sind, sollten Sie sicherstellen, dass Ihre Abfragen mit der [allgemeinen Abfragestruktur](#) von AWS Clean Rooms Differential Privacy kompatibel sind.

Note

Wenn Sie [Cryptographic Computing für](#) verwenden Clean Rooms, generieren nicht alle SQL-Operationen gültige Ergebnisse. Sie können beispielsweise COUNT auf eine verschlüsselte Spalte zugreifen, aber die Ausführung SUM auf verschlüsselte Zahlen führt zu Fehlern. Darüber hinaus können Abfragen auch zu falschen Ergebnissen führen. Beispielsweise führen Abfragen, bei denen Spalten SUM versiegelt wurden, zu Fehlern. Eine GROUP BY Abfrage über versiegelte Spalten scheint jedoch erfolgreich zu sein, erzeugt aber andere Gruppen als die, die bei einer GROUP BY Abfrage über den Klartext erzeugt werden.

Dem [Mitglied, das die Rechenkosten für Abfragen bezahlt](#), werden die im Rahmen der Kollaboration ausgeführten Abfragen in Rechnung gestellt.

Das Mitglied, das Abfragen durchführen kann, kann mehrere [Mitglieder auswählen, die Ergebnisse erhalten können](#), um die Ergebnisse einer einzelnen Abfrage zu erhalten. Weitere Informationen finden Sie unter [Abfragen konfigurierter Tabellen mit dem SQL-Code-Editor](#). Allgemeine Informationen zum Empfangen von Abfrageergebnissen finden Sie unter [Empfangen und Verwenden von Analyseergebnissen](#).

Voraussetzungen

Bevor Sie eine SQL-Abfrage ausführen, stellen Sie sicher, dass Sie über Folgendes verfügen:

- Eine aktive Mitgliedschaft bei AWS Clean Rooms Collaboration

- Zugriff auf mindestens eine konfigurierte Tabelle in der Kollaboration
- Bestätigung, dass das Mitglied, das für die Berechnung der Abfragekosten verantwortlich ist, ein aktives Mitglied der Kollaboration ist

Informationen dazu, wie Sie Daten abfragen oder Abfragen anzeigen können, indem Sie den AWS Clean Rooms [StartProtectedQuery API-Vorgang](#) direkt aufrufen oder den verwenden AWS SDKs, finden Sie in der [AWS Clean Rooms API-Referenz](#).

Hinweise zur Abfrageprotokollierung finden Sie unter [Analyse Einloggen AWS Clean Rooms](#).

 Note

Wenn Sie eine Abfrage für [verschlüsselte](#) Datentabellen ausführen, werden die Ergebnisse der verschlüsselten Spalten verschlüsselt.

Konfiguration der Spark-Eigenschaften für SQL-Abfragen

AWS Clean Rooms ermöglicht es Ihnen, das Spark-Laufzeitverhalten optional anzupassen, indem Sie die unterstützten Spark-Eigenschaften für SQL-Abfragen konfigurieren. Mit diesen Eigenschaften können Sie die Leistung, die Speichernutzung und die Parameter für die Abfrageausführung fein abstimmen. Mit dieser Funktion haben Sie mehr Kontrolle darüber, wie Ihre SPARK-basierten Abfragen verarbeitet werden, und können so eine Optimierung auf der Grundlage Ihrer spezifischen Workload-Anforderungen vornehmen.

Sie können jetzt Einstellungen wie Shuffle-Partitionen, Schwellenwerte für Broadcast-Joins und Parameter für die Ausführung adaptiver Abfragen direkt von der Konsole aus anpassen. AWS Clean Rooms Diese Funktion ist besonders nützlich für komplexe Abfragen oder große Datenmengen, bei denen Standardkonfigurationen möglicherweise nicht optimal sind. Durch die Feinabstimmung dieser Spark-Eigenschaften können Sie potenziell die Abfrageleistung verbessern, den Ressourcenverbrauch reduzieren und die Speichernutzung für Ihre Spark-basierten Kollaborationsanalysen besser verwalten.

Um diese Funktion nutzen zu können, finden Sie in der Abfrageschnittstelle einen neuen Abschnitt mit Spark-Eigenschaften. Sie können aus einer Liste unterstützter Eigenschaften auswählen und benutzerdefinierte Werte angeben. [Sie können Spark-Eigenschaften auch programmgesteuert mithilfe der StartProtectedQuery API konfigurieren](#). Diese erweiterte Konfigurationsoption ermöglicht

es Datenanalysten und Ingenieuren, ihre Abfragen zu optimieren, um die Effizienz und Skalierbarkeit zu verbessern.

Weitere Informationen zu Spark-Eigenschaften, einschließlich Standardwerten, finden Sie unter [Spark-Eigenschaften](#) in der Apache Spark-Dokumentation.

In den folgenden Themen wird erklärt, wie Daten in einer Kollaboration mithilfe der AWS Clean Rooms Konsole abgefragt werden.

Topics

- [Abfragen konfigurierter Tabellen mit dem SQL-Code-Editor](#)
- [Abfragen von ID-Zuordnungstabellen mit dem SQL-Code-Editor](#)
- [Abfragen konfigurierter Tabellen mithilfe einer SQL-Analysevorlage](#)
- [Abfragen mit dem Analysis Builder](#)
- [Überblick über die Auswirkungen des unterschiedlichen Datenschutzes](#)
- [Anzeige kürzlicher Abfragen](#)
- [Anzeigen von Abfragedetails](#)

Abfragen konfigurierter Tabellen mit dem SQL-Code-Editor

Als Mitglied, das Abfragen durchführen kann, können Sie eine Abfrage manuell erstellen, indem Sie SQL-Code in den SQL-Code-Editor schreiben. Der SQL-Code-Editor befindet sich in der AWS Clean Rooms Konsole auf der Registerkarte Analyse im Abschnitt Analyse.

Der SQL-Code-Editor wird standardmäßig angezeigt. Wenn Sie den Analysis Builder zum Erstellen von Abfragen verwenden möchten, finden Sie weitere Informationen unter [Abfragen mit dem Analysis Builder](#).

Important

Wenn Sie mit dem Schreiben einer SQL-Abfrage im Code-Editor beginnen und dann die Analysis Builder-Benutzeroberfläche einschalten, wird Ihre Abfrage nicht gespeichert.

AWS Clean Rooms unterstützt viele SQL-Befehle, Funktionen und Bedingungen. Weitere Informationen finden Sie in der [AWS Clean Rooms SQL-Referenz](#).

 Tip

Wenn während der Ausführung einer Abfrage eine geplante Wartung stattfindet, wird die Abfrage beendet und ein Rollback durchgeführt. Sie müssen die Abfrage neu starten.

Um konfigurierte Tabellen mit dem SQL-Code-Editor abzufragen

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus, für die der Status Ihrer Mitgliederfähigkeiten den Status Abfragen ausführen lautet.
4. Sehen Sie sich auf der Registerkarte Analyse unter Tabellen die Liste der Tabellen und den zugehörigen Analyseregeltyp an (Aggregationsanalyseregeln, Listenanalyseregeln oder Benutzerdefinierte Analyseregeln).

 Note

Wenn Sie die erwarteten Tabellen in der Liste nicht sehen, kann das folgende Gründe haben:

- Die Tabellen wurden nicht [verknüpft](#).
- Für die Tabellen ist keine [Analyseregel konfiguriert](#).

5. (Optional) Um das Schema und die Analyseregelnsteuerelemente der Tabelle anzuzeigen, erweitern Sie die Tabelle, indem Sie das Pluszeichen (+) auswählen.
6. Wählen Sie im Abschnitt Analyse für den Analysemodus die Option SQL-Code schreiben aus.

 Note

Im Abschnitt Analyse wird nur angezeigt, ob das Mitglied, das Ergebnisse erhalten kann, und das Mitglied, das für die Bezahlung der Abfrage-Rechenkosten verantwortlich ist, der Kollaboration als aktives Mitglied beigetreten sind.

7. Erstellen Sie die Abfrage, indem Sie die Abfrage in den SQL-Code-Editor eingeben.

Weitere Informationen zu unterstützten SQL-Befehlen und -Funktionen finden Sie in der [AWS Clean Rooms SQL-Referenz](#).

Sie können auch die folgenden Optionen verwenden, um Ihre Abfrage zu erstellen.

Use an example query

Um eine Beispielabfrage zu verwenden

1. Wählen Sie die drei vertikalen Punkte neben der Tabelle aus.
2. Wählen Sie unter In Editor einfügen die Option Beispielabfrage aus.

 Note

Wenn Sie eine Beispielabfrage einfügen, wird sie an die Abfrage angehängt, die sich bereits im Editor befindet.

Das Abfragebeispiel wird angezeigt. Alle unter Tabellen aufgeführten Tabellen sind in der Abfrage enthalten.

3. Bearbeiten Sie die Platzhalterwerte in der Abfrage.

Insert column names or functions

Um einen Spaltennamen oder eine Funktion einzufügen

1. Wählen Sie die drei vertikalen Punkte neben einer Spalte aus.
2. Wählen Sie unter In Editor einfügen die Option Spaltenname aus.
3. Um eine Funktion, die für eine Spalte zulässig ist, manuell einzufügen,
 - a. Wählen Sie die drei vertikalen Punkte neben einer Spalte aus.
 - b. Wählen Sie In Editor einfügen aus.
 - c. Wählen Sie den Namen der zulässigen Funktion aus (z. B. INNER JOIN SUM, SUM DISTINCT, oder COUNT).
4. Drücken Sie Strg + Leertaste, um die Tabellenschemas im Code-Editor anzuzeigen.

 Note

Mitglieder, die Abfragen durchführen können, können die Partitionsspalten in jeder konfigurierten Tabellenverknüpfung anzeigen und verwenden. Stellen Sie sicher, dass die Partitionsspalte in der AWS Glue Tabelle, die der konfigurierten Tabelle zugrunde liegt, als Partitionsspalte gekennzeichnet ist.

5. Bearbeiten Sie die Platzhalterwerte in der Abfrage.

8. Geben Sie den unterstützten Arbeitertyp und die Anzahl der Mitarbeiter an.

Sie können den Instanztyp und die Anzahl der Instanzen (Worker) wählen, um Ihre SQL-Abfragen auszuführen.

Für CR.1X können Sie bis zu 128 Worker oder mindestens 4 Worker auswählen.

Für CR.4X können Sie bis zu 32 Arbeiter oder mindestens 4 Arbeiter auswählen.

Ermitteln Sie anhand der folgenden Tabelle den Typ und die Anzahl der Mitarbeiter, die Sie für Ihren Anwendungsfall benötigen.

Worker type (Worker-Typ)	vCPU	Speicher (GB)	Speicher (GB)	Number of workers (Anzahl der Worker)	Gesamtzahl der Verarbeitungseinheiten für Reinräume (CRPU)
CR.1X (Standard)	4	30	100	4	8
				128	256
CR.4X	16	120	400	4	32
				32	256

 Note

Verschiedene Arten von Arbeitnehmern und die Anzahl der Arbeitnehmer sind mit Kosten verbunden. Weitere Informationen zu den Preisen finden Sie unter [AWS Clean Rooms Preisgestaltung](#).

9. Geben Sie unter Ergebnisse senden an an an, wer Ergebnisse erhalten kann.

 Note

Um Ergebnisse zu erhalten, muss das Kollaborationsmitglied als Empfänger von Ergebnissen konfiguriert sein und ein aktiver Teilnehmer an der Kollaboration sein (Status: Aktiv)

10. (Mitglied, das nur Abfragen durchführen kann) Das Kontrollkästchen Verwenden Sie Ihre Standardergbniseinstellungen ist standardmäßig aktiviert. Lassen Sie dieses Kontrollkästchen aktiviert, wenn Sie Ihre Standardergbniseinstellungen beibehalten möchten.

Wenn Sie andere Ergebniseinstellungen für diese Abfrage angeben möchten, deaktivieren Sie das Kontrollkästchen Verwenden Sie Ihre Standardergbniseinstellungen, und wählen Sie dann Folgendes aus.

- a. Ergebnisformat (CSV oder PARQUET)
- b. Ergebnisdateien (einzeln oder mehrfach)
- c. Ziel der Ergebnisse in Amazon S3

Jedes Mitglied, das Ergebnisse erhalten kann, kann in Amazon S3 ein anderes Ergebnisformat, andere Ergebnisdateien und ein anderes Ergebnisziel angeben.

11. So geben Sie Spark-Eigenschaften an:

- a. Erweitern Sie Spark-Eigenschaften.
- b. Wählen Sie Spark-Eigenschaften hinzufügen.
- c. Wählen Sie im Dialogfeld „Spark-Eigenschaften“ einen Eigenschaftsnamen aus der Dropdownliste aus und geben Sie einen Wert ein.

Die folgende Tabelle enthält eine Definition für jede Eigenschaft.

Weitere Informationen zu Spark-Eigenschaften finden Sie unter [Spark-Eigenschaften](#) in der Apache Spark-Dokumentation.

Eigenschaftensname	Description	Standardwert
spark.task.maxFailures	Steuert, wie oft eine Aufgabe hintereinander fehlschlagen kann, bevor der Job fehlschlägt. Erfordert einen Wert größer oder gleich 1. Die Anzahl der zulässigen Wiederholungen entspricht diesem Wert minus 1. Die Anzahl der Fehlschläge wird zurückgesetzt, wenn ein Versuch erfolgreich ist. Fehler bei verschiedenen Aufgaben summieren sich nicht bis zu diesem Limit.	4
spark.sql.files.maxPartitionBytes	Legt die maximale Anzahl von Byte fest, die beim Lesen aus dateibasierten Quellen wie Parquet, JSON und ORC in eine einzelne Partition gepackt werden sollen.	128 MB
spark.hadoop.fs.s3.maxRetries	Legt die maximale Anzahl von Wiederholungsversuchen für Amazon S3 S3-Dateioperationen fest.	
spark.network.timeout	Legt das Standard-Timeout für alle Netzwerkinteraktionen fest. Setzt die folgenden Timeout-Einstellungen außer Kraft, wenn sie nicht konfiguriert sind: - spark.storage.blockManagerHeartbeatTimeoutMs - spark.shuffle.io.ConnectionTimeout - spark.rpc.AskTimeout - spark.rpc.LookupTimeout	120 s

Eigenschaftename	Description	Standardwert
<code>spark.rdd.com/pressure</code>	Gibt an, ob serialisierte RDD-Partitionen mit <code>spark.io.compression.codec</code> komprimiert werden sollen. Gilt für <code>StorageLevel .MEMORY_ONLY_SER</code> in Java und Scala oder <code>.MEMORY_ONLY</code> in Python. <code>StorageLevel</code> Reduziert den Speicherplatz, erfordert jedoch zusätzliche CPU-Verarbeitungszeit.	FALSE
<code>spark.shuffle.spill.com/press</code>	Gibt an, ob Shuffle-Spill-Daten mit <code>spark.io.compression.codec</code> komprimiert werden sollen.	TRUE
<code>spark.sql.adaptive.advisoryPartitionSizeInBytes</code>	Legt die Zielgröße in Byte für Shuffle-Partitionen während der adaptiven Optimierung fest, wenn <code>spark.sql.adaptive.enabled</code> den Wert <code>true</code> hat. Steuert die Partitionsgröße beim Zusammenführen kleiner Partitionen oder beim Teilen schiefer Partitionen.	(Wert von <code>spark.sql.adaptive.shuffle.targetPostShuffleInputSize</code>)
<code>spark.sql.adaptive.autoBroadcastJoinSchwellenwert</code>	Legt die maximale Tabellengröße in Byte für die Übertragung an Worker-Knoten bei Joins fest. Gilt nur im adaptiven Framework. Verwendet denselben Standardwert wie <code>spark.sql.autoBroadcastJoinSchwellenwert</code> . Auf -1 setzen, um die Übertragung zu deaktivieren.	(Keine)
<code>spark.sql.adaptive.CoalescePartitions.Enabled</code>	Gibt an, ob zusammenhängende Shuffle-Partitionen, die auf <code>spark.sql.adaptive</code> basieren, zusammengeführt werden sollen. <code>advisoryPartitionSizeInBytes</code> um die Aufgabengröße zu optimieren. Erfordert, dass <code>spark.sql.adaptive.enabled</code> den Wert <code>true</code> hat.	TRUE

Eigenschaftensname	Description	Standardwert
spark.sql.adaptive. .CoalescePartitions. initialPartitionNum	Definiert die anfängliche Anzahl von Shuffle-Partitionen vor dem Zusammenführen. Erfordert, dass sowohl spark.sql.adaptive.enabled als auch spark.sql.adaptive.CoalescePartitions.Enabled den Wert true haben. Der Standardwert ist der Wert von spark.sql.shuffle.partitions.	(Keine)
spark.sql.adaptive. .CoalescePartitions. minPartitionSize	Legt die Mindestgröße für zusammenggeführte Shuffle-Partitionen fest, um zu verhindern, dass Partitionen während der adaptiven Optimierung zu klein werden.	1 MB
spark.sql.adaptive. .CoalescePartitions. ParallelismFirst	Gibt an, ob Partitionsgrößen auf der Grundlage der Cluster-Parallelität und nicht auf der Grundlage von spark.sql.adaptive.advisoryPartitionSizeInBytes während der Zusammenführung von Partitionen. Generiert kleinere Partitionsgrößen als die konfigurierte Zielgröße, um die Parallelität zu maximieren. Wir empfehlen, diesen Wert bei stark frequentierten Clustern auf „False“ zu setzen, um die Ressourcennutzung zu verbessern und übermäßig kleine Aufgaben zu vermeiden.	TRUE
spark.sql.adaptive. .enabled	Gibt an, ob die adaptive Abfrageausführung aktiviert werden soll, um Abfragepläne während der Abfrageausführung auf der Grundlage genauer Laufzeitstatistiken erneut zu optimieren.	TRUE
spark.sql.adaptive. forceOptimizeSkewed Join	Gibt an, ob die Aktivierung erzwungen werden soll, OptimizeSkewedJoin auch wenn dadurch zusätzlicher Shuffle eingeführt wird.	FALSE

Eigenschaftename	Description	Standardwert
spark.sql.adaptive.localShuffleReader.aktiviert	Gibt an, ob lokale Shuffle-Reader verwendet werden sollen, wenn eine Shuffle-Partitionierung nicht erforderlich ist, z. B. nach der Konvertierung von Sort-Merge-Joins in Broadcast-Hash-Joins. Erfordert, dass spark.sql.adaptive.enabled den Wert true hat.	TRUE
spark.sql.adaptive.maxShuffledHashJoinLocalMapThreshold	Legt die maximale Partitionsgröße in Byte für die Erstellung lokaler Hash-Maps fest. Priorisiert gemischte Hash-Joins gegenüber Sort-Merge-Joins, wenn: - Dieser Wert entspricht oder übersteigt spark.sql.adaptive.advisoryPartitionSizeInBytes - Alle Partitionsgrößen liegen innerhalb dieser Grenze Überschreibt spark.sql.join.preferSortMergeJoin-Einstellung.	0 Byte
spark.sql.adaptive.optimizeSkewsInRebalancePartitions.aktiviert	Gibt an, ob schiefe Shuffle-Partitionen optimiert werden sollen, indem sie auf der Grundlage von spark.sql.adaptive in kleinere Partitionen aufgeteilt werden. advisoryPartitionSizeInBytes. Erfordert, dass spark.sql.adaptive.enabled wahr ist.	TRUE
spark.sql.adaptive.rebalancePartitionsSmallPartitionFactor	Definiert den Größenschwellenwert für das Zusammenführen von Partitionen beim Teilen. Partitionen, die kleiner sind als dieser Faktor multipliziert mit spark.sql.adaptive.advisoryPartitionSizeInBytes werden zusammengeführt.	0.2

Eigenschaftename	Description	Standardwert
spark.sql.Adaptive .SkewJoin.Enabled	Gibt an, ob Datenverzerrungen in gemischten Verknüpfungen behandelt werden sollen, indem schiefe Partitionen aufgeteilt und optional repliziert werden. Gilt für Sort-Merge- und Shuffled-Hash-Joins. Erfordert, dass spark.sql.adaptive.enabled wahr ist.	TRUE
spark.sql.adaptive .SkewJoin. skewedPartitionFactor	Bestimmt den Größenfaktor, der die Partition schiefe bestimmt. Eine Partition ist schief, wenn ihre Größe beide Werte überschreitet: - Dieser Faktor wird mit der mittleren Partitionsgröße multipliziert - Der Wert von spark.sql.adaptive.SkewJoin.skewedPartitionThresholdInBytes	5
spark.sql.Adaptive .SkewJoin. skewedPartitionThresholdInBytes	Legt den Größenschwellenwert in Byte zur Identifizierung schiefer Partitionen fest. Eine Partition ist schief, wenn ihre Größe beide Werte überschreitet: - Dieser Schwellenwert - Die mittlere Partitionsgröße multipliziert mit spark.sql.adaptive.SkewJoin.skewedPartitionFactor Wir empfehlen, diesen Wert größer als spark.sql.adaptive festzulegen. advisoryPartitionSizeInBytes.	256 MB
spark.sql. autoBroadcastJoinSchwellenwert	Legt die maximale Tabellengröße in Byte für die Übertragung an Worker-Knoten bei Joins fest. Auf -1 setzen, um die Übertragung zu deaktivieren.	10 MB

Eigenschaftename	Description	Standardwert
spark.sql.BroadcastTimeout	Steuert den Timeout-Zeitraum in Sekunden für die Broadcast-Operationen bei Broadcast-Joins.	300 Sekunden
spark.sql.cbo.enabled	Gibt an, ob die kostenbasierte Optimierung (CBO) für die Schätzung von Planstatistiken aktiviert werden soll.	FALSE
spark.sql.CBO.JoinReorder.DP.Star.Filter	Gibt an, ob bei der kostenbasierten Join-Aufzählung Heuristiken des Star-Join-Filters angewendet werden sollen.	FALSE
spark.sql.CBO.JoinReorder.DP.Threshold	Legt die maximale Anzahl verbundener Knoten fest, die im dynamischen Programmieralgorithmus zulässig sind.	12
spark.sql.CBO.JoinReorder.Enabled	Gibt an, ob die Neuordnung von Verknüpfungen bei der kostenbasierten Optimierung (CBO) aktiviert werden soll.	FALSE
spark.sql.CBO.PlanStats.Enabled	Gibt an, ob bei der Generierung logischer Pläne Zeilenanzahl und Spaltenstatistiken aus dem Katalog abgerufen werden sollen.	FALSE
spark.sql.cbo.starSchemaDetection	Gibt an, ob die Neuordnung von Verknüpfungen auf der Grundlage der Star-Schemaerkennung aktiviert werden soll.	FALSE
spark.sql.files.maxPartitionNum	Legt die maximale Zielanzahl von geteilten Dateipartitionen für dateibasierte Quellen (Parquet, JSON und ORC) fest. Skaliert Partitionen neu, wenn die anfängliche Anzahl diesen Wert überschreitet. Dies ist ein empfohlenes Ziel, kein garantiertes Limit.	(Keine)

Eigenschaftename	Description	Standardwert
spark.sql.files. maxRecord sPerDatei	Legt die maximale Anzahl von Datensätzen fest, die in eine einzelne Datei geschrieben werden sollen. Wenn der Wert auf Null oder einen negativen Wert gesetzt ist, gilt kein Limit.	0
spark.sql.files. minPartitionNum	Legt die Mindestanzahl von geteilten Dateipartitionen für dateibasierte Quellen (Parquet, JSON und ORC) fest. Die Standardeinstellung ist spark.sql. leafNodeDefaultParallelität. Dies ist ein empfohlenes Ziel, kein garantiertes Limit.	(Keine)
spark.sql. inMemoryC olumnarSpeicher. Batchgröße	Steuert die Batchgröße für das spaltenförmige Caching. Eine Erhöhung der Größe verbessert die Speichernutzung und Komprimierung, erhöht jedoch das Fehlerrisiko out-of-memory.	10000
spark.sql. inMemoryC olumnarSpeicher. Komprimiert	Gibt an, ob auf der Grundlage von Datenstatistiken automatisch Komprimierungscodecs für Spalten ausgewählt werden sollen.	TRUE
spark.sql. inMemoryC olumnarAu fbewahrung. enableVectorizedRe ader	Gibt an, ob vektorisiertes Lesen für das spaltenförmige Caching aktiviert werden soll.	TRUE
spark.sql.legacy. allowHash OnMapType	Gibt an, ob Hash-Operationen für Map-Datenstrukturen zulässig sind. Diese Legacy-Einstellung gewährleistet die Kompatibilität mit der Map-Typ-Behandlung älterer Spark-Versionen.	

Eigenschaftensname	Description	Standardwert
spark.sql.legacy.allowNegativeScaleOfDecimal	Gibt an, ob negative Skalenwerte in Dezimaltypdefinitionen zulässig sind. Diese ältere Einstellung gewährleistet die Kompatibilität mit älteren Spark-Versionen, die negative Dezimalskalen unterstützten.	
spark.sql.legacy.castComplexTypesToString.aktiviert	Gibt an, ob veraltetes Verhalten für die Umwandlung komplexer Typen in Zeichenketten aktiviert werden soll. Behält die Kompatibilität mit den Typkonvertierungsregeln älterer Spark-Versionen bei.	
spark.sql.legacy.charVarcharAsString.enfolge	Gibt an, ob die Typen CHAR und VARCHAR als STRING-Typen behandelt werden sollen. Diese Legacy-Einstellung bietet Kompatibilität mit der Verarbeitung von String-Typen in älteren Spark-Versionen.	
spark.sql.legacy.createEmptyCollectionUsingStringType	Gibt an, ob leere Sammlungen mithilfe von Zeichenkettenelementen erstellt werden sollen. Diese Legacy-Einstellung gewährleistet die Kompatibilität mit dem Verhalten älterer Spark-Versionen bei der Initialisierung von Sammlungen.	
spark.sql.legacy.exponentLiteralAsDecimal. Aktiviert	Gibt an, ob exponentielle Literale als Dezimaltypen interpretiert werden sollen. Diese Legacy-Einstellung gewährleistet die Kompatibilität mit der numerischen Literalverarbeitung älterer Spark-Versionen.	
spark.sql.legacy.json.allowEmptyString. aktiviert	Gibt an, ob leere Zeichenfolgen bei der JSON-Verarbeitung zulässig sind. Diese Legacy-Einstellung gewährleistet die Kompatibilität mit dem JSON-Parsing-Verhalten älterer Spark-Versionen.	

Eigenschaftensname	Description	Standardwert
spark.sql.legacy.parquet.int96RebaseModelRead	Gibt an, ob beim Lesen von Parquet-Dateien der alte Timestamp-Rebase-Modus verwendet werden soll. INT96 Diese Legacy-Einstellung gewährleistet die Kompatibilität mit der Zeitstempelverarbeitung älterer Spark-Versionen.	
spark.sql.legacy.timeParserPolicy	Steuert das Zeitanalyseverhalten aus Gründen der Abwärtskompatibilität. Diese ältere Einstellung bestimmt, wie Zeitstempel und Datumsangaben anhand von Zeichenketten analysiert werden.	
spark.sql.Legacy.TypeCoercion.datetimeToString.aktiviert	Gibt an, ob bei der Konvertierung von Datetime-Werten in Zeichenfolgen das Zwangsverhalten älterer Typen aktiviert werden soll. Behält die Kompatibilität mit den Datetime-Konvertierungsregeln älterer Spark-Versionen bei.	
spark.sql. maxSinglePartitionByte	Legt die maximale Partitionsgröße in Byte fest. Der Planer führt Shuffle-Operationen für größere Partitionen ein, um die Parallelität zu verbessern.	128 m

Eigenschaftensname	Description	Standardwert
spark.sql.Metadatancache TTLSeconds	Steuert die (TTL) für time-to-live Metadaten -Caches. Gilt für Partitionsdatei-Metadaten und Sitzungskatalog-Caches. Erfordert: • Ein positiver Wert größer als Null • spark.sql.CatalogImplementation ist auf Hive gesetzt • spark.sql.hive. filesourcePartitionFileCacheSize größer als Null • spark.sql.hive. manageFilesourcePartitions auf true gesetzt	-1000 ms
spark.sql.optimizer.collapseProjectAlwaysInline	Gibt an, ob benachbarte Projektionen und Inline-Ausdrücke ausgeblendet werden sollen, auch wenn dies zu Duplikaten führt.	FALSE
spark.sql.optimizer.dynamicPartitionPruning. aktiviert	Gibt an, ob Prädikate für Partitionsspalten generiert werden sollen, die als Join-Schlüssel verwendet werden.	TRUE
spark.sql.optimizer.enableCsvExpressionOptimierung	Gibt an, ob CSV-Ausdrücke im SQL-Optimizer optimiert werden sollen, indem unnötige Spalten aus from_csv-Vorgängen entfernt werden.	TRUE
spark.sql.optimizer.enableJsonExpressionOptimierung	Gibt an, ob JSON-Ausdrücke im SQL-Optimizer wie folgt optimiert werden sollen: • Löschen unnötiger Spalten aus from_json-Vorgängen • Vereinfachung der Kombinationen from_json und to_json • Optimierung von named_struct-Operationen	TRUE

Eigenschaftename	Description	Standardwert
spark.sql.Optimizer.ExcludedRules	Definiert zu deaktivierende Optimizer-Regeln, identifiziert durch kommasetrennte Regelnamen. Einige Regeln können nicht deaktiviert werden, da sie aus Gründen der Richtigkeit erforderlich sind. Der Optimizer protokolliert, welche Regeln erfolgreich deaktiviert wurden.	(Keine)
spark.sql.optimizer.runtime.BloomFilter.applicationSideScanSizeThreshold	Legt die minimale aggregierte Scangröße in Byte fest, die erforderlich ist, um einen Bloom-Filter auf der Anwendungsseite einzufügen.	10 GB
spark.sql.Optimizer.Runtime.BloomFilter.creationSideThreshold	Definiert den maximalen Größenschwellenwert für die Injektion eines Bloom-Filters auf der Erstellungsseite.	10 MB
spark.sql.Optimizer.Runtime.BloomFilter.Enabled	Gibt an, ob ein Bloom-Filter eingefügt werden soll, um Shuffle-Daten zu reduzieren, wenn eine Seite einer Shuffle-Verknüpfung über ein selektives Prädikat verfügt.	TRUE
spark.sql.optimizer.runtime.BloomFilter.expectedNumItems	Definiert die Standardanzahl erwarteter Elemente im Runtime-Bloom-Filter.	1000000
spark.sql.optimizer.runtime.BloomFilter.maxNumBits	Legt die maximale Anzahl von Bits fest, die im Runtime-Bloom-Filter zulässig sind.	67108864
spark.sql.optimizer.runtime.BloomFilter.maxNumItems	Legt die maximale Anzahl erwarteter Elemente fest, die im Runtime-Bloom-Filter zulässig sind.	4000000

Eigenschaftensname	Description	Standardwert
spark.sql.optimize r.runtime.BloomFilter.Number.Threshold	Schränkt die maximale Anzahl von Nicht-DPP-Laufzeitfiltern pro Abfrage ein, um Fehler im Treiber zu verhindern. out-of-memory	10
spark.sql.optimize r.runtime.BloomFilter.NumBits	Definiert die Standardanzahl von Bits, die im Runtime-Bloom-Filter verwendet werden.	8388608
spark.sql.optimize r.runtime.rowlevelOperationGroupFilter. Aktiviert	Gibt an, ob die Laufzeitgruppenfilterung für Operationen auf Zeilenebene aktiviert werden soll. Ermöglicht Datenquellen: • Löschen ganzer Datengruppen (wie Dateien oder Partitionen) mithilfe von Datenquellenfiltern • Führen Sie Laufzeitabfragen aus, um übereinstimmende Datensätze zu identifizieren • Verwerfen Sie unnötige Gruppen, um teure Neuschreibungen zu vermeiden Einschränkungen: • Nicht alle Ausdrücke können in Datenquellenfilter konvertiert werden • Einige Ausdrücke erfordern eine Spark-Auswertung (z. B. Unterabfragen)	TRUE
spark.sql.Optimize r.RuntimeFilter. semiJoinReduction. aktiviert	Gibt an, ob ein Semi-Join eingefügt werden soll, um die Anzahl der Shuffle-Daten zu reduzieren, wenn eine Seite einer Shuffle-Verknüpfung über ein selektives Prädikat verfügt.	FALSE

Eigenschaftensname	Description	Standardwert
spark.sql.parquet. AggregatePushDown	Gibt an, ob Aggregate zur Optimierung nach Parquet übertragen werden sollen. Unterstützt: • MIN und MAX für die Typen Boolean, Integer, Float und Date • COUNT für alle Datentypen Löst eine Ausnahme aus, wenn Statistiken in einer Fußzeile einer Parquet-Datei fehlen.	FALSE
spark.sql.parquet. columnarReaderBatchGröße	Steuert die Anzahl der Zeilen in jedem Stapel von vektorisierten Parquet-Readern. Wählen Sie einen Wert, der Leistungsaufwand und Speichernutzung ausbalanciert, um Fehler zu vermeiden out-of-memory.	4096
spark.sql.session. TimeZone	Definiert die Sitzungszeitzone für die Verarbeitung von Zeitstempeln in Zeichenfolgenliteralen und die Konvertierung von Java-Objekten. Akzeptiert: • Regionsbasiertes IDs area/city Format (z. B. America/Los_Angeles) • Zonenversätze im Format (+/-) HH, (+/-) HH:mm oder (+/-) HH:mm:ss (z. B. -08 oder + 01:00) • UTC oder Z als Aliase für + 00:00	(Wert der lokalen Zeitzone)
spark.sql.shuffle. partitions	Legt die Standardanzahl von Partitionen für das Mischen von Daten bei Verknüpfungen oder Aggregationen fest. Kann zwischen Neustarts strukturierter Streaming-Abfragen von derselben Checkpoint-Position aus nicht geändert werden.	200

Eigenschaftename	Description	Standardwert
spark.sql.shuffledHashJoinFaktor	Definiert den Multiplikationsfaktor, der verwendet wird, um die Eignung für einen Shuffle-Hash-Join zu bestimmen. Ein Shuffle-Hash-Join wird ausgewählt, wenn die Datengröße der kleinen Seite multipliziert mit diesem Faktor kleiner als die Datengröße der großen Seite ist.	3
spark.sql.sources.parallelPartitionDiscovery.Schwellenwert	Legt die maximale Anzahl von Pfaden für die treiberseitige Dateiauflistung mit dateibasierten Quellen (Parquet, JSON und ORC) fest. Wenn sie bei der Partitionserkennung überschritten werden, werden Dateien mithilfe eines separaten verteilten Spark-Jobs aufgelistet.	32
spark.sql.statistics.histogram.enabled	Gibt an, ob bei der Berechnung der Spaltenstatistiken Histogramme mit gleicher Höhe generiert werden sollen, um die Schätzgenauigkeit zu verbessern. Erfordert einen zusätzlichen Tabellenscan, der über den für einfache Spaltenstatistiken erforderlichen hinausgeht.	FALSE

12. Klicken Sie auf Ausführen.

 Note

Sie können die Abfrage nicht ausführen, wenn das Mitglied, das Ergebnisse empfangen kann, die Einstellungen für die Abfrageergebnisse nicht konfiguriert hat.

13. Die Ergebnisse anzeigen.

Weitere Informationen finden Sie unter [Empfangen und Verwenden von Analyseergebnissen](#).

14. Passen Sie die Parameter weiter an und führen Sie Ihre Abfrage erneut aus, oder klicken Sie auf die Schaltfläche +, um eine neue Abfrage auf einer neuen Registerkarte zu starten.

 Note

AWS Clean Rooms zielt darauf ab, klare Fehlermeldungen bereitzustellen. Wenn eine Fehlermeldung nicht genügend Details enthält, um Ihnen bei der Fehlerbehebung zu helfen, wenden Sie sich an das Account-Team. Geben Sie ihnen eine Beschreibung, wie der Fehler aufgetreten ist, und geben Sie ihnen die Fehlermeldung (einschließlich aller Identifikatoren). Weitere Informationen finden Sie unter [Problembehebung AWS Clean Rooms](#).

Abfragen von ID-Zuordnungstabellen mit dem SQL-Code-Editor

Das folgende Verfahren beschreibt, wie Sie eine Join-Abfrage mit mehreren Tabellen für die ID-Zuordnungstabelle ausführen, um die `sourceId` mit der `targetId` zu verknüpfen.

Bevor Sie die ID-Zuordnungstabelle abfragen, muss die ID-Zuordnungstabelle erfolgreich gefüllt werden.

Um ID-Zuordnungstabellen mit dem SQL-Code-Editor abzufragen

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus, für die der Status Ihrer Mitgliederfähigkeiten den Status Abfragen ausführen lautet.
4. Gehen Sie auf der Registerkarte Analyse zum Abschnitt Analyse.

 Note

Im Abschnitt Analyse wird nur angezeigt, ob das Mitglied, das Ergebnisse erhalten kann, und das Mitglied, das für die Bezahlung der Abfrage-Rechenkosten verantwortlich ist, der Kollaboration als aktives Mitglied beigetreten sind.

5. Sehen Sie sich auf der Registerkarte Analyse unter Tabellen die Liste der ID-Zuordnungstabellen (unter Verwaltet von AWS Clean Rooms) und den zugehörigen Analyseregeltyp (Analyseregel für ID-Zuordnungstabellen) an.

Note

Wenn Sie die erwarteten ID-Zuordnungstabellen nicht in der Liste sehen, liegt das möglicherweise daran, dass die ID-Zuordnungstabellen nicht erfolgreich gefüllt wurden. Weitere Informationen finden Sie unter [Auffüllen einer vorhandenen ID-Zuordnungstabelle](#).

- Erstellen Sie die Abfrage, indem Sie die Abfrage in den SQL-Code-Editor eingeben.

(Optional) Wenn Sie eine Beispielabfrage verwenden möchten

- Wählen Sie die drei vertikalen Punkte neben der Tabelle aus.
- Wählen Sie unter In Editor einfügen die Option Beispiel für eine JOIN-Anweisung aus.

Note

Wenn Sie eine JOIN-Beispielanweisung einfügen, wird die Abfrage angehängt, die sich bereits im Editor befindet.

Das Beispiel für eine JOIN-Anweisung wird angezeigt.

- Bearbeiten Sie die Platzhalterwerte in der Abfrage.

(Optional) Wenn Sie einen Tabellennamen einfügen möchten

- Wählen Sie die drei vertikalen Punkte neben einer Spalte aus.
- Wählen Sie unter In Editor einfügen die Option Tabellennamen aus.
- Bearbeiten Sie die Platzhalterwerte in der Abfrage.

- Geben Sie den unterstützten Arbeitertyp und die Anzahl der Mitarbeiter an.

Ermitteln Sie anhand der folgenden Tabelle den Typ und die Anzahl der Mitarbeiter, die Sie für Ihren Anwendungsfall benötigen.

Worker type (Worker-Typ)	vCPU	Speicher (GB)	Speicher (GB)	Number of workers (Anzahl der Worker)	Gesamtzahl der Verarbeitungseinheiten für Reinräume (CRPU)
CR.1X (Standard)	4	30	100	4	8
				128	256
CR.4X	16	120	400	4	32
				32	256

 Note

Verschiedene Arten von Arbeitnehmern und die Anzahl der Arbeitnehmer sind mit Kosten verbunden. Weitere Informationen zu den Preisen finden Sie unter [AWS Clean Rooms Preisgestaltung](#).

8. Klicken Sie auf Ausführen.

 Note

Sie können die Abfrage nicht ausführen, wenn das Mitglied, das Ergebnisse erhalten kann, die Einstellungen für die Abfrageergebnisse nicht konfiguriert hat.

9. Die Ergebnisse anzeigen.

Weitere Informationen finden Sie unter [Empfangen und Verwenden von Analyseergebnissen](#).

10. Passen Sie die Parameter weiter an und führen Sie Ihre Abfrage erneut aus, oder klicken Sie auf die Schaltfläche +, um eine neue Abfrage auf einer neuen Registerkarte zu starten.

 Note

AWS Clean Rooms zielt darauf ab, klare Fehlermeldungen bereitzustellen. Wenn eine Fehlermeldung nicht genügend Details enthält, um Ihnen bei der Fehlerbehebung zu helfen, wenden Sie sich an das Account-Team. Geben Sie ihnen eine Beschreibung, wie der Fehler aufgetreten ist, und geben Sie ihnen die Fehlermeldung (einschließlich aller Identifikatoren). Weitere Informationen finden Sie unter [Problembehebung AWS Clean Rooms](#).

Abfragen konfigurierter Tabellen mithilfe einer SQL-Analysevorlage

Dieses Verfahren zeigt, wie Sie eine Analysevorlage in der AWS Clean Rooms Konsole verwenden, um konfigurierte Tabellen mit der benutzerdefinierten Analyseregeln abzufragen.

So verwenden Sie eine SQL-Analysevorlage, um konfigurierte Tabellen mit der benutzerdefinierten Analyseregeln abzufragen

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus, für die der Status Ihrer Mitgliederfähigkeiten den Status Abfragen ausführen lautet.
4. Sehen Sie sich auf der Registerkarte Analyse im Abschnitt Tabellen die Tabellen und den zugehörigen Analyseregeltyp an (benutzerdefinierte Analyseregeln).

 Note

Wenn Sie die erwarteten Tabellen nicht in der Liste sehen, kann das folgende Gründe haben:

- Die Tabellen wurden nicht [verknüpft](#).
- Für die Tabellen ist keine [Analyseregel konfiguriert](#).

5. Wählen Sie im Abschnitt Analyse für den Analysemodus die Option Analysevorlagen ausführen aus und wählen Sie dann die Analysevorlage aus der Dropdownliste aus.
6. Die Parameter aus der SQL-Analysevorlage werden automatisch in die Definition übernommen.
7. Geben Sie den unterstützten Worker-Typ und die Anzahl der Worker an.

Ermitteln Sie anhand der folgenden Tabelle den Typ und die Anzahl der Mitarbeiter, die Sie für Ihren Anwendungsfall benötigen.

Worker type (Worker-Typ)	vCPU	Speicher (GB)	Speicher (GB)	Number of workers (Anzahl der Worker)	Gesamtzahl der Verarbeitungseinheiten für Reinräume (CRPU)
CR.1X (Standard)	4	30	100	4	8
				128	256
CR.4X	16	120	400	4	32
				32	256

 Note

Verschiedene Arten von Arbeitnehmern und die Anzahl der Arbeitnehmer sind mit Kosten verbunden. Weitere Informationen zu den Preisen finden Sie unter [AWS Clean Rooms Preisgestaltung](#).

8. Geben Sie die unterstützten Spark-Eigenschaften an.
 - a. Wählen Sie Spark-Eigenschaften hinzufügen aus.
 - b. Wählen Sie im Dialogfeld „Spark-Eigenschaften“ einen Eigenschaftsnamen aus der Dropdownliste aus und geben Sie einen Wert ein.

Die folgende Tabelle enthält eine Definition für jede Eigenschaft.

Weitere Informationen zu Spark-Eigenschaften finden Sie unter [Spark-Eigenschaften](#) in der Apache Spark-Dokumentation.

Eigenschaftename	Description	Standardwert
spark.task.maxFailures	Steuert, wie oft eine Aufgabe hintereinander fehlschlagen kann, bevor der Job fehlschlägt. Erfordert einen Wert größer oder gleich 1. Die Anzahl der zulässigen Wiederholungen entspricht diesem Wert minus 1. Die Anzahl der Fehlschläge wird zurückgesetzt, wenn ein Versuch erfolgreich ist. Fehler bei verschiedenen Aufgaben summieren sich nicht bis zu diesem Limit.	4
spark.sql.files.maxPartitionBytes	Legt die maximale Anzahl von Byte fest, die beim Lesen aus dateibasierten Quellen wie Parquet, JSON und ORC in eine einzelne Partition gepackt werden sollen.	128 MB
spark.hadoop.fs.s3.maxVersuchtEsErneut	Legt die maximale Anzahl von Wiederholungsversuchen für Amazon S3 S3-Dateioperationen fest.	
spark.network.timeout	Legt das Standard-Timeout für alle Netzwerkinteraktionen fest. Setzt die folgenden Timeout-Einstellungen außer Kraft, wenn sie nicht konfiguriert sind: - spark.storage.blockManagerHeartbeatTimeoutMs - spark.shuffle.io.ConnectionTimeout - spark.rpc.AskTimeout - spark.rpc.LookupTimeout	120 s

Eigenschaftename	Description	Standardwert
<code>spark.rdd.compression.enabled</code>	Gibt an, ob serialisierte RDD-Partitionen mit <code>spark.io.compression.codec</code> komprimiert werden sollen. Gilt für <code>StorageLevel .MEMORY_ONLY_SER</code> in Java und Scala oder <code>.MEMORY_ONLY</code> in Python. <code>StorageLevel</code> Reduziert den Speicherplatz, erfordert jedoch zusätzliche CPU-Verarbeitungszeit.	FALSE
<code>spark.shuffle.spill.compression.enabled</code>	Gibt an, ob Shuffle-Spill-Daten mit <code>spark.io.compression.codec</code> komprimiert werden sollen.	TRUE
<code>spark.sql.adaptive.advisoryPartitionSizeInBytes</code>	Legt die Zielgröße in Byte für Shuffle-Partitionen während der adaptiven Optimierung fest, wenn <code>spark.sql.adaptive.enabled</code> den Wert <code>true</code> hat. Steuert die Partitionsgröße beim Zusammenführen kleiner Partitionen oder beim Teilen schiefer Partitionen.	(Wert von <code>spark.sql.adaptive.shuffle.targetPostShuffleInputSize</code>)
<code>spark.sql.adaptive.autoBroadcastJoinSwitch</code>	Legt die maximale Tabellengröße in Byte für die Übertragung an Worker-Knoten bei Joins fest. Gilt nur im adaptiven Framework. Verwendet denselben Standardwert wie <code>spark.sql.adaptive.autoBroadcastJoinSwitch</code> . Auf -1 setzen, um die Übertragung zu deaktivieren.	(Keine)
<code>spark.sql.adaptive.coalescePartitions.enabled</code>	Gibt an, ob zusammenhängende Shuffle-Partitionen, die auf <code>spark.sql.adaptive.advisoryPartitionSizeInBytes</code> basieren, zusammengeführt werden sollen. <code>spark.sql.adaptive.advisoryPartitionSizeInBytes</code> um die Aufgabengröße zu optimieren. Erfordert, dass <code>spark.sql.adaptive.enabled</code> den Wert <code>true</code> hat.	TRUE

Eigenschaftename	Description	Standardwert
spark.sql.adaptive. .CoalescePartitions. initialPartitionNum	Definiert die anfängliche Anzahl von Shuffle-Partitionen vor dem Zusammenführen. Erfordert, dass sowohl spark.sql.adaptive.enabled als auch spark.sql.adaptive.CoalescePartitions.Enabled den Wert true haben. Der Standardwert ist der Wert von spark.sql.shuffle.partitions.	(Keine)
spark.sql.adaptive. .CoalescePartitions. minPartitionSize	Legt die Mindestgröße für zusammenggeführte Shuffle-Partitionen fest, um zu verhindern, dass Partitionen während der adaptiven Optimierung zu klein werden.	1 MB
spark.sql.adaptive. .CoalescePartitions. ParallelismFirst	Gibt an, ob Partitionsgrößen auf der Grundlage der Cluster-Parallelität und nicht auf der Grundlage von spark.sql.adaptive.advisoryPartitionSizeInBytes während der Zusammenführung von Partitionen. Generiert kleinere Partitionsgrößen als die konfigurierte Zielgröße, um die Parallelität zu maximieren. Wir empfehlen, diesen Wert bei stark frequentierten Clustern auf „False“ zu setzen, um die Ressourcennutzung zu verbessern und übermäßig kleine Aufgaben zu vermeiden.	TRUE
spark.sql.adaptive. .enabled	Gibt an, ob die adaptive Abfrageausführung aktiviert werden soll, um Abfragepläne während der Abfrageausführung auf der Grundlage genauer Laufzeitstatistiken erneut zu optimieren.	TRUE
spark.sql.adaptive. forceOptimizeSkewed Join	Gibt an, ob die Aktivierung erzwungen werden soll, OptimizeSkewedJoin auch wenn dadurch zusätzlicher Shuffle eingeführt wird.	FALSE

Eigenschaftename	Description	Standardwert
spark.sql.adaptive.localShuffleReader.aktiviert	Gibt an, ob lokale Shuffle-Reader verwendet werden sollen, wenn eine Shuffle-Partitionierung nicht erforderlich ist, z. B. nach der Konvertierung von Sort-Merge-Joins in Broadcast-Hash-Joins. Erfordert, dass spark.sql.adaptive.enabled den Wert true hat.	TRUE
spark.sql.adaptive.maxShuffledHashJoinLocalMapThreshold	Legt die maximale Partitionsgröße in Byte für die Erstellung lokaler Hash-Maps fest. Priorisiert gemischte Hash-Joins gegenüber Sort-Merge-Joins, wenn: - Dieser Wert entspricht oder übersteigt spark.sql.adaptive.advisoryPartitionSizeInBytes - Alle Partitionsgrößen liegen innerhalb dieser Grenze Überschreibt spark.sql.join.preferSortMergeJoin-Einstellung.	0 Byte
spark.sql.adaptive.optimizeSkewsInRebalancePartitions.aktiviert	Gibt an, ob schiefe Shuffle-Partitionen optimiert werden sollen, indem sie auf der Grundlage von spark.sql.adaptive in kleinere Partitionen aufgeteilt werden. advisoryPartitionSizeInBytes. Erfordert, dass spark.sql.adaptive.enabled wahr ist.	TRUE
spark.sql.adaptive.rebalancePartitionsSmallPartitionFactor	Definiert den Größenschwellenwert für das Zusammenführen von Partitionen beim Teilen. Partitionen, die kleiner sind als dieser Faktor multipliziert mit spark.sql.adaptive.advisoryPartitionSizeInBytes werden zusammengeführt.	0.2

Eigenschaftename	Description	Standardwert
spark.sql.Adaptive .SkewJoin.Enabled	Gibt an, ob Datenverzerrungen in gemischten Verknüpfungen behandelt werden sollen, indem schiefe Partitionen aufgeteilt und optional repliziert werden. Gilt für Sort-Merge- und Shuffled-Hash-Joins. Erfordert, dass spark.sql.adaptive.enabled wahr ist.	TRUE
spark.sql.adaptive .SkewJoin. skewedPartitionFactor	Bestimmt den Größenfaktor, der die Partition schiefe bestimmt. Eine Partition ist schief, wenn ihre Größe beide Werte überschreitet: - Dieser Faktor wird mit der mittleren Partitionsgröße multipliziert - Der Wert von spark.sql.adaptive.SkewJoin.skewedPartitionThresholdInBytes	5
spark.sql.Adaptive .SkewJoin. skewedPartitionThresholdInBytes	Legt den Größenschwellenwert in Byte zur Identifizierung schiefer Partitionen fest. Eine Partition ist schief, wenn ihre Größe beide Werte überschreitet: - Dieser Schwellenwert - Die mittlere Partitionsgröße multipliziert mit spark.sql.adaptive.SkewJoin.skewedPartitionFactor Wir empfehlen, diesen Wert größer als spark.sql.adaptive festzulegen. advisoryPartitionSizeInBytes.	256 MB
spark.sql. autoBroadcastJoinSchwellenwert	Legt die maximale Tabellengröße in Byte für die Übertragung an Worker-Knoten bei Joins fest. Auf -1 setzen, um die Übertragung zu deaktivieren.	10 MB

Eigenschaftename	Description	Standardwert
spark.sql.BroadcastTimeout	Steuert den Timeout-Zeitraum in Sekunden für die Broadcast-Operationen bei Broadcast-Joins.	300 Sekunden
spark.sql.cbo.enabled	Gibt an, ob die kostenbasierte Optimierung (CBO) für die Schätzung von Planstatistiken aktiviert werden soll.	FALSE
spark.sql.CBO.JoinReorder.DP.Star.Filter	Gibt an, ob bei der kostenbasierten Join-Aufzählung Heuristiken des Star-Join-Filters angewendet werden sollen.	FALSE
spark.sql.CBO.JoinReorder.DP.Threshold	Legt die maximale Anzahl verbundener Knoten fest, die im dynamischen Programmieralgorithmus zulässig sind.	12
spark.sql.CBO.JoinReorder.Enabled	Gibt an, ob die Neuordnung von Verknüpfungen bei der kostenbasierten Optimierung (CBO) aktiviert werden soll.	FALSE
spark.sql.CBO.PlanStats.Enabled	Gibt an, ob bei der Generierung logischer Pläne Zeilenanzahl und Spaltenstatistiken aus dem Katalog abgerufen werden sollen.	FALSE
spark.sql.cbo.starSchemaDetection	Gibt an, ob die Neuordnung von Verknüpfungen auf der Grundlage der Star-Schemaerkennung aktiviert werden soll.	FALSE
spark.sql.files.maxPartitionNum	Legt die maximale Zielanzahl von geteilten Dateipartitionen für dateibasierte Quellen (Parquet, JSON und ORC) fest. Skaliert Partitionen neu, wenn die anfängliche Anzahl diesen Wert überschreitet. Dies ist ein empfohlenes Ziel, kein garantiertes Limit.	(Keine)

Eigenschaftensname	Description	Standardwert
spark.sql.files. maxRecord sPerDatei	Legt die maximale Anzahl von Datensätzen fest, die in eine einzelne Datei geschrieben werden sollen. Wenn der Wert auf Null oder einen negativen Wert gesetzt ist, gilt kein Limit.	0
spark.sql.files. minPartitionNum	Legt die Mindestanzahl von geteilten Dateipartitionen für dateibasierte Quellen (Parquet, JSON und ORC) fest. Die Standardeinstellung ist spark.sql. leafNodeDefaultParallelität. Dies ist ein empfohlenes Ziel, kein garantiertes Limit.	(Keine)
spark.sql. inMemoryC olumnarSpeicher. Batchgröße	Steuert die Batchgröße für das spaltenförmige Caching. Eine Erhöhung der Größe verbessert die Speichernutzung und Komprimierung, erhöht jedoch das Fehlerrisiko out-of-memory.	10000
spark.sql. inMemoryC olumnarSpeicher. Komprimiert	Gibt an, ob auf der Grundlage von Datenstatistiken automatisch Komprimierungscodecs für Spalten ausgewählt werden sollen.	TRUE
spark.sql. inMemoryC olumnarAu fbewahrung. enableVectorizedReader	Gibt an, ob vektorisiertes Lesen für das spaltenförmige Caching aktiviert werden soll.	TRUE
spark.sql.legacy. allowHash OnMapType	Gibt an, ob Hash-Operationen für Map-Datenstrukturen zulässig sind. Diese Legacy-Einstellung gewährleistet die Kompatibilität mit der Map-Typ-Behandlung älterer Spark-Versionen.	

Eigenschaftename	Description	Standardwert
spark.sql.legacy.allowNegativeScaleOfDecimal	Gibt an, ob negative Skalenwerte in Dezimaltypdefinitionen zulässig sind. Diese ältere Einstellung gewährleistet die Kompatibilität mit älteren Spark-Versionen, die negative Dezimalskalen unterstützten.	
spark.sql.legacy.castComplexTypesToString.aktiviert	Gibt an, ob veraltetes Verhalten für die Umwandlung komplexer Typen in Zeichenketten aktiviert werden soll. Behält die Kompatibilität mit den Typkonvertierungsregeln älterer Spark-Versionen bei.	
spark.sql.legacy.charVarcharAsString.enfolge	Gibt an, ob die Typen CHAR und VARCHAR als STRING-Typen behandelt werden sollen. Diese Legacy-Einstellung bietet Kompatibilität mit der Verarbeitung von String-Typen in älteren Spark-Versionen.	
spark.sql.legacy.createEmptyCollectionUsingStringType	Gibt an, ob leere Sammlungen mithilfe von Zeichenkettenelementen erstellt werden sollen. Diese Legacy-Einstellung gewährleistet die Kompatibilität mit dem Verhalten älterer Spark-Versionen bei der Initialisierung von Sammlungen.	
spark.sql.legacy.exponentLiteralAsDecimal. Aktiviert	Gibt an, ob exponentielle Literale als Dezimaltypen interpretiert werden sollen. Diese Legacy-Einstellung gewährleistet die Kompatibilität mit der numerischen Literalverarbeitung älterer Spark-Versionen.	
spark.sql.legacy.json.allowEmptyString. aktiviert	Gibt an, ob leere Zeichenfolgen bei der JSON-Verarbeitung zulässig sind. Diese Legacy-Einstellung gewährleistet die Kompatibilität mit dem JSON-Parsing-Verhalten älterer Spark-Versionen.	

Eigenschaftename	Description	Standardwert
spark.sql.legacy.parquet.int96RebaseModelRead	Gibt an, ob beim Lesen von Parquet-Dateien der alte Timestamp-Rebase-Modus verwendet werden soll. INT96 Diese Legacy-Einstellung gewährleistet die Kompatibilität mit der Zeitstempelverarbeitung älterer Spark-Versionen.	
spark.sql.legacy.timeParserPolicy	Steuert das Zeitanalyseverhalten aus Gründen der Abwärtskompatibilität. Diese ältere Einstellung bestimmt, wie Zeitstempel und Datumsangaben anhand von Zeichenketten analysiert werden.	
spark.sql.Legacy.TypeCoercion.datetimeToString.aktiviert	Gibt an, ob bei der Konvertierung von Datetime-Werten in Zeichenfolgen das Zwangsverhalten älterer Typen aktiviert werden soll. Behält die Kompatibilität mit den Datetime-Konvertierungsregeln älterer Spark-Versionen bei.	
spark.sql. maxSinglePartitionByte	Legt die maximale Partitionsgröße in Byte fest. Der Planer führt Shuffle-Operationen für größere Partitionen ein, um die Parallelität zu verbessern.	128 m

Eigenschaftename	Description	Standardwert
spark.sql.Metadatancache TTLSeconds	Steuert die (TTL) für time-to-live Metadaten -Caches. Gilt für Partitionsdatei-Metadaten und Sitzungskatalog-Caches. Erfordert: • Ein positiver Wert größer als Null • spark.sql.CatalogImplementation ist auf Hive gesetzt • spark.sql.hive. filesourcePartitionFileCacheSize größer als Null • spark.sql.hive. manageFilesourcePartitions auf true gesetzt	-1000 ms
spark.sql.optimizer.collapseProjectAlwaysInline	Gibt an, ob benachbarte Projektionen und Inline-Ausdrücke ausgeblendet werden sollen, auch wenn dies zu Duplikaten führt.	FALSE
spark.sql.optimizer.dynamicPartitionPruning. aktiviert	Gibt an, ob Prädikate für Partitionsspalten generiert werden sollen, die als Join-Schlüssel verwendet werden.	TRUE
spark.sql.optimizer.enableCsvExpressionOptimierung	Gibt an, ob CSV-Ausdrücke im SQL-Optimizer optimiert werden sollen, indem unnötige Spalten aus from_csv-Vorgängen entfernt werden.	TRUE
spark.sql.optimizer.enableJsonExpressionOptimierung	Gibt an, ob JSON-Ausdrücke im SQL-Optimizer wie folgt optimiert werden sollen: • Löschen unnötiger Spalten aus from_json-Vorgängen • Vereinfachung der Kombinationen from_json und to_json • Optimierung von named_struct-Operationen	TRUE

Eigenschaftename	Description	Standardwert
spark.sql.Optimizer.ExcludedRules	Definiert zu deaktivierende Optimizer-Regeln, identifiziert durch kommasetrennte Regelnamen. Einige Regeln können nicht deaktiviert werden, da sie aus Gründen der Richtigkeit erforderlich sind. Der Optimizer protokolliert, welche Regeln erfolgreich deaktiviert wurden.	(Keine)
spark.sql.optimizer.runtime.BloomFilter.applicationSideScanSizeThreshold	Legt die minimale aggregierte Scangröße in Byte fest, die erforderlich ist, um einen Bloom-Filter auf der Anwendungsseite einzufügen.	10 GB
spark.sql.Optimizer.Runtime.BloomFilter.creationSideThreshold	Definiert den maximalen Größenschwellenwert für die Injektion eines Bloom-Filters auf der Erstellungsseite.	10 MB
spark.sql.Optimizer.Runtime.BloomFilter.Enabled	Gibt an, ob ein Bloom-Filter eingefügt werden soll, um Shuffle-Daten zu reduzieren, wenn eine Seite einer Shuffle-Verknüpfung über ein selektives Prädikat verfügt.	TRUE
spark.sql.optimizer.runtime.BloomFilter.expectedNumItems	Definiert die Standardanzahl erwarteter Elemente im Runtime-Bloom-Filter.	1000000
spark.sql.optimizer.runtime.BloomFilter.maxNumBits	Legt die maximale Anzahl von Bits fest, die im Runtime-Bloom-Filter zulässig sind.	67108864
spark.sql.optimizer.runtime.BloomFilter.maxNumItems	Legt die maximale Anzahl erwarteter Elemente fest, die im Runtime-Bloom-Filter zulässig sind.	4000000

Eigenschaftename	Description	Standardwert
spark.sql.optimize r.runtime.BloomFilter.Number.Threshold	Schränkt die maximale Anzahl von Nicht-DPP-Laufzeitfiltern pro Abfrage ein, um Fehler im Treiber zu verhindern. out-of-memory	10
spark.sql.optimize r.runtime.BloomFilter.NumBits	Definiert die Standardanzahl von Bits, die im Runtime-Bloom-Filter verwendet werden.	8388608
spark.sql.optimize r.runtime.rowlevelOperationGroupFilter. Aktiviert	Gibt an, ob die Laufzeitgruppenfilterung für Operationen auf Zeilenebene aktiviert werden soll. Ermöglicht Datenquellen: • Löschen ganzer Datengruppen (wie Dateien oder Partitionen) mithilfe von Datenquellenfiltern • Führen Sie Laufzeitabfragen aus, um übereinstimmende Datensätze zu identifizieren • Verwerfen Sie unnötige Gruppen, um teure Neuschreibungen zu vermeiden Einschränkungen: • Nicht alle Ausdrücke können in Datenquellenfilter konvertiert werden • Einige Ausdrücke erfordern eine Spark-Auswertung (z. B. Unterabfragen)	TRUE
spark.sql.Optimize r.RuntimeFilter. semiJoinReduction. aktiviert	Gibt an, ob ein Semi-Join eingefügt werden soll, um die Anzahl der Shuffle-Daten zu reduzieren, wenn eine Seite einer Shuffle-Verknüpfung über ein selektives Prädikat verfügt.	FALSE

Eigenschaftensname	Description	Standardwert
spark.sql.parquet. AggregatePushDown	Gibt an, ob Aggregate zur Optimierung nach Parquet übertragen werden sollen. Unterstützt: • MIN und MAX für die Typen Boolean, Integer, Float und Date • COUNT für alle Datentypen Löst eine Ausnahme aus, wenn Statistiken in einer Fußzeile einer Parquet-Datei fehlen.	FALSE
spark.sql.parquet. columnarReaderBatchGröße	Steuert die Anzahl der Zeilen in jedem Stapel von vektorisierten Parquet-Readern. Wählen Sie einen Wert, der Leistungsaufwand und Speichernutzung ausbalanciert, um Fehler zu vermeiden out-of-memory.	4096
spark.sql.session. TimeZone	Definiert die Sitzungszeitzone für die Verarbeitung von Zeitstempeln in Zeichenfolgenliteralen und die Konvertierung von Java-Objekten. Akzeptiert: • Regionsbasiertes IDs area/city Format (z. B. America/Los_Angeles) • Zonenversätze im Format (+/-) HH, (+/-) HH:mm oder (+/-) HH:mm:SS (z. B. -08 oder + 01:00) • UTC oder Z als Aliase für + 00:00	(Wert der lokalen Zeitzone)
spark.sql.shuffle. partitions	Legt die Standardanzahl von Partitionen für das Mischen von Daten bei Verknüpfungen oder Aggregationen fest. Kann zwischen Neustarts strukturierter Streaming-Abfragen von derselben Checkpoint-Position aus nicht geändert werden.	200

Eigenschaftensname	Description	Standardwert
spark.sql.shuffledHashJoinFaktor	Definiert den Multiplikationsfaktor, der verwendet wird, um die Eignung für einen Shuffle-Hash-Join zu bestimmen. Ein Shuffle-Hash-Join wird ausgewählt, wenn die Datengröße der kleinen Seite multipliziert mit diesem Faktor kleiner als die Datengröße der großen Seite ist.	3
spark.sql.sources.parallelPartitionDiscovery.Schwellenwert	Legt die maximale Anzahl von Pfaden für die treiberseitige Dateiauflistung mit dateibasierten Quellen (Parquet, JSON und ORC) fest. Wenn sie bei der Partitionserkennung überschritten werden, werden Dateien mithilfe eines separaten verteilten Spark-Jobs aufgelistet.	32
spark.sql.statistics.histogram.enabled	Gibt an, ob bei der Berechnung der Spaltenstatistiken Histogramme mit gleicher Höhe generiert werden sollen, um die Schätzgenauigkeit zu verbessern. Erfordert einen zusätzlichen Tabellenscan, der über den für einfache Spaltenstatistiken erforderlichen hinausgeht.	FALSE

9. Klicken Sie auf Ausführen.

 Note

Sie können die Abfrage nicht ausführen, wenn das Mitglied, das Ergebnisse empfangen kann, die Einstellungen für die Abfrageergebnisse nicht konfiguriert hat.

10. Passen Sie die Parameter weiter an und führen Sie Ihre Abfrage erneut aus, oder klicken Sie auf die Schaltfläche +, um eine neue Abfrage auf einer neuen Registerkarte zu starten.

Abfragen mit dem Analysis Builder

Sie können den Analysis Builder verwenden, um Abfragen zu erstellen, ohne SQL-Code schreiben zu müssen. Mit dem Analysis Builder können Sie eine Abfrage für eine Kollaboration erstellen, die Folgendes bietet:

- Eine einzelne Tabelle, die die [Aggregationsanalyiseregel](#) verwendet, ohne dass JOIN erforderlich ist
- Zwei Tabellen (eine von jedem Mitglied), die beide die [Aggregationsanalyiseregel](#) verwenden
- Zwei Tabellen (eine für jedes Mitglied), die beide die [Listenanalyiseregel](#) verwenden
- Zwei Tabellen (eine für jedes Mitglied), die beide die Aggregationsanalyiseregel verwenden, und zwei Tabellen (eine für jedes Mitglied), die beide die Listenanalyiseregel verwenden

Wenn Sie SQL-Abfragen manuell schreiben möchten, finden Sie weitere Informationen unter [Abfragen konfigurierter Tabellen mit dem SQL-Code-Editor](#).

Der Analysis Builder wird in der AWS Clean Rooms Konsole auf der Registerkarte Analyse im Abschnitt Analyse als Benutzeroberflächenoption für Analysis Builder angezeigt.

Important

Wenn Sie die Analysis Builder-Benutzeroberfläche aktivieren, mit der Erstellung einer Abfrage im Analysis Builder beginnen und dann die Analysis Builder-Benutzeroberfläche ausschalten, wird Ihre Abfrage nicht gespeichert.

Tip

Wenn während der Ausführung einer Abfrage eine geplante Wartung stattfindet, wird die Abfrage beendet und ein Rollback durchgeführt. Sie müssen die Abfrage neu starten.

In den folgenden Themen wird die Verwendung des Analysis Builder erläutert.

Topics

- [Verwenden Sie den Analysis Builder, um eine einzelne Tabelle abzufragen \(Aggregation\)](#)

- [Verwenden Sie den Analysis Builder, um zwei Tabellen \(Aggregation oder Liste\) abzufragen](#)

Verwenden Sie den Analysis Builder, um eine einzelne Tabelle abzufragen (Aggregation)

Dieses Verfahren zeigt, wie Sie die Analysis Builder-Benutzeroberfläche in der AWS Clean Rooms Konsole verwenden, um eine Abfrage zu erstellen. Die Abfrage ist für eine Kollaboration vorgesehen, die über eine einzelne Tabelle verfügt, die die [Aggregationsanalyseregeln](#) verwendet, ohne dass JOIN dies erforderlich ist.

Um den Analysis Builder für die Abfrage einer einzelnen Tabelle zu verwenden

Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.

1. Wählen Sie im linken Navigationsbereich Collaborations aus.
2. Wählen Sie die Kollaboration aus, für die der Status Ihrer Mitgliederfähigkeiten den Status Abfragen ausführen lautet.
3. Sehen Sie sich auf der Registerkarte Analyse unter Tabellen die Tabelle und den zugehörigen Analyseregeltyp an. (Der Analyseregeltyp sollte die Aggregationsanalyseregeln sein.)

Note

Wenn Sie die erwartete Tabelle nicht sehen, kann das folgende Gründe haben:

- Die Tabelle wurde nicht [verknüpft](#).
- Für die Tabelle ist keine [Analyseregel konfiguriert](#).

4. Aktivieren Sie im Abschnitt Analyse die Benutzeroberfläche von Analysis Builder.
5. Erstellen Sie eine Abfrage.

Wenn Sie alle Aggregationsmetriken sehen möchten, fahren Sie mit Schritt 9 fort.

- a. Überprüfen Sie unter Metriken auswählen die aggregierten Metriken, die standardmäßig vorausgewählt wurden, und entfernen Sie bei Bedarf alle Metriken.
- b. (Optional) Wählen Sie unter Segmente hinzufügen — optional einen oder mehrere Parameter aus.

 Note

Segmente hinzufügen — optional wird nur angezeigt, wenn Dimensionen für die Tabelle angegeben sind.

- c. (Optional) Wählen Sie unter Filter hinzufügen — optional die Option Filter hinzufügen und wählen Sie dann einen Parameter, einen Operator und einen Wert aus.

Um weitere Filter hinzuzufügen, wählen Sie „Weiteren Filter hinzufügen“.

Um einen Filter zu entfernen, wählen Sie Entfernen.

 Note

ORDER BY wird für Aggregationsabfragen nicht unterstützt.
Nur der AND Operator wird in Filtern unterstützt.

- d. (Optional) Geben Sie unter Beschreibung hinzufügen — optional eine Beschreibung ein, um die Abfrage in der Abfrageliste leichter identifizieren zu können.

6. Erweitern Sie Vorschau-SQL-Code.

- a. Zeigen Sie den SQL-Code an, der mit dem Analysis Builder generiert wurde.
- b. Um den SQL-Code zu kopieren, wählen Sie Kopieren.
- c. Um den SQL-Code zu bearbeiten, wählen Sie Im SQL-Code-Editor bearbeiten.

7. Klicken Sie auf Ausführen.

 Note

Sie können die Abfrage nicht ausführen, wenn das Mitglied, das Ergebnisse empfangen kann, die Einstellungen für die Abfrageergebnisse nicht konfiguriert hat.

8. Passen Sie die Parameter weiter an und führen Sie Ihre Abfrage erneut aus, oder klicken Sie auf die Schaltfläche +, um eine neue Abfrage auf einer neuen Registerkarte zu starten.

 Note

AWS Clean Rooms zielt darauf ab, klare Fehlermeldungen bereitzustellen. Wenn eine Fehlermeldung nicht genügend Details enthält, um Ihnen bei der Fehlerbehebung zu helfen, wenden Sie sich an das Account-Team. Geben Sie ihnen eine Beschreibung, wie der Fehler aufgetreten ist, und geben Sie ihnen die Fehlermeldung (einschließlich aller Identifikatoren). Weitere Informationen finden Sie unter [Problembehebung AWS Clean Rooms](#).

Verwenden Sie den Analysis Builder, um zwei Tabellen (Aggregation oder Liste) abzufragen

In diesem Verfahren wird beschrieben, wie Sie den Analysis Builder in der AWS Clean Rooms Konsole verwenden, um eine Abfrage für eine Kollaboration zu erstellen, die:

- Zwei Tabellen (eine für jedes Mitglied), die beide die [Aggregationsanalyseregeln](#) verwenden
- Zwei Tabellen (eine für jedes Mitglied), die beide die [Listenanalyseregeln](#) verwenden
- Zwei Tabellen (eine für jedes Mitglied), die beide die Aggregationsanalyseregeln verwenden, und zwei Tabellen (eine für jedes Mitglied), die beide die Listenanalyseregeln verwenden

So verwenden Sie den Analysis Builder, um zwei Tabellen abzufragen

Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.

1. Wählen Sie im linken Navigationsbereich Collaborations aus.
2. Wählen Sie die Kollaboration aus, deren Status Ihre Mitgliederfähigkeiten den Status Abfragen ausführen lautet.
3. Sehen Sie sich auf der Registerkarte Analyse unter Tabellen die beiden Tabellen und den zugehörigen Analyseregeltyp an (Aggregationsanalyseregeln oder Listenanalyseregeln).

 Note

Wenn Sie die erwarteten Tabellen nicht in der Liste sehen, kann das folgende Gründe haben:

- Die Tabellen wurden nicht [verknüpft](#).

- Für die Tabellen ist keine [Analyseregel konfiguriert](#).

4. Aktivieren Sie im Abschnitt Analyse die Benutzeroberfläche von Analysis Builder.
5. Erstellen Sie eine Abfrage.

Wenn die Kollaboration zwei Tabellen enthält, die die Aggregationsanalyseregel verwenden, und zwei Tabellen, die die Analyseregel „Liste“ verwenden, wählen Sie zuerst Aggregation oder Liste aus, und folgen Sie dann den Anweisungen, die auf der ausgewählten Analyseregel basieren.

Wenn die beiden Tabellen die Aggregationsanalyseregel verwenden

1. Überprüfen Sie unter Metriken auswählen die aggregierten Metriken, die standardmäßig vorausgewählt wurden, und entfernen Sie bei Bedarf alle Metriken.
2. Wählen Sie für Datensätze zuordnen einen oder mehrere Datensätze aus.

Note

Wenn Sie den Analysegenerator verwenden, können Sie nur für ein einzelnes Spaltenpaar einen Abgleich durchführen.

3. (Optional) Wählen Sie unter Segmente hinzufügen — optional einen oder mehrere Parameter aus.

Note

Segmente hinzufügen — optional wird nur angezeigt, wenn Dimensionen für die Tabelle angegeben sind.

Wenn die beiden Tabellen die Listenanalyseregel verwenden

1. Überprüfen Sie unter Attribute auswählen die Listenattribute, die standardmäßig vorausgewählt wurden, und entfernen Sie bei Bedarf alle Metriken.
2. Wählen Sie für Datensätze zuordnen einen oder mehrere Datensätze aus.

Note

Wenn Sie den Analysegenerator verwenden, können Sie nur für ein einzelnes Spaltenpaar einen Abgleich durchführen.

3. (Optional) Wählen Sie unter Filter hinzufügen — optional die Option Filter hinzufügen und wählen Sie dann einen Parameter, einen Operator und einen Wert aus.

Um weitere Filter hinzuzufügen, wählen Sie „Weiteren Filter hinzufügen“.

Um einen Filter zu entfernen, wählen Sie Entfernen.

Wenn die beiden Tabellen die Aggregationsanalyseregel verwenden

4. (Optional) Wählen Sie unter Filter hinzufügen — optional die Option Filter hinzufügen und wählen Sie dann einen Parameter, einen Operator und einen Wert aus.

Um weitere Filter hinzuzufügen, wählen Sie „Weiteren Filter hinzufügen“.

Um einen Filter zu entfernen, wählen Sie Entfernen.

Note

ORDER BY wird für Aggregationsabfragen nicht unterstützt.
In Filtern wird nur der AND Operator unterstützt.

5. (Optional) Geben Sie unter Beschreibung hinzufügen — optional eine Beschreibung ein, um die Abfrage in der Liste der letzten Abfragen leichter identifizieren zu können.

Wenn die beiden Tabellen die Listenanalyseregel verwenden

Note

LIMIT wird für Listenabfragen nicht unterstützt.
In Filtern wird nur der AND Operator unterstützt.

4. (Optional) Geben Sie unter Beschreibung hinzufügen — optional eine Beschreibung ein, um die Abfrage in der Liste der letzten Abfragen leichter identifizieren zu können.

6. Erweitern Sie Vorschau-SQL-Code.

- Zeigen Sie den SQL-Code an, der mit dem Analysis Builder generiert wurde.
- Um den SQL-Code zu kopieren, wählen Sie Kopieren.
- Um den SQL-Code zu bearbeiten, wählen Sie Im SQL-Code-Editor bearbeiten.

7. Klicken Sie auf Ausführen.

Note

Sie können die Abfrage nicht ausführen, wenn das Mitglied, das Ergebnisse empfangen kann, die Einstellungen für die Abfrageergebnisse nicht konfiguriert hat

8. Passen Sie die Parameter weiter an und führen Sie Ihre Abfrage erneut aus, oder klicken Sie auf die Schaltfläche +, um eine neue Abfrage auf einer neuen Registerkarte zu starten.

 Note

AWS Clean Rooms zielt darauf ab, klare Fehlermeldungen bereitzustellen. Wenn eine Fehlermeldung nicht genügend Details enthält, um Ihnen bei der Fehlerbehebung zu helfen, wenden Sie sich an das Account-Team. Geben Sie ihnen eine Beschreibung, wie der Fehler aufgetreten ist, und geben Sie ihnen die Fehlermeldung (einschließlich aller Identifikatoren). Weitere Informationen finden Sie unter [Problembehebung AWS Clean Rooms](#).

Überblick über die Auswirkungen des unterschiedlichen Datenschutzes

Im Allgemeinen ändert sich das Schreiben und Ausführen von Abfragen nicht, wenn Differential Privacy aktiviert ist. Sie können jedoch keine Abfrage ausführen, wenn nicht genügend Datenschutzbudget übrig ist. Wenn Sie Abfragen ausführen und das Datenschutzbudget verbrauchen, können Sie ungefähr sehen, wie viele Aggregationen Sie ausführen können und wie sich dies auf future Abfragen auswirken könnte.

Um die Auswirkungen des unterschiedlichen Datenschutzes in einer Zusammenarbeit zu untersuchen

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus, deren Status Ihre Mitgliedsdetails auf Abfragen ausführen lautet.
4. Sehen Sie sich auf der Registerkarte Analyse unter Tabellen das verbleibende Datenschutzbudget an. Dies wird als geschätzte Anzahl der verbleibenden Aggregationsfunktionen und des verwendeten Dienstprogramms (in Prozent) angezeigt.

 Note

Die geschätzte Anzahl der verbleibenden Aggregatfunktionen und der Prozentsatz des verwendeten Dienstprogramms werden nur für das Mitglied angezeigt, das Abfragen durchführen kann.

5. Wählen Sie „Auswirkung anzeigen“, um zu sehen, wie viel Rauschen in die Ergebnisse eingedrungen ist und wie viele Aggregationsfunktionen Sie ungefähr ausführen können.

Anzeige kürzlicher Abfragen

Auf der Registerkarte Analyse können Sie sich die Abfragen ansehen, die in den letzten 90 Tagen ausgeführt wurden.

 Note

Wenn Sie als Mitglied nur Contribute-Daten haben und Sie nicht das [Mitglied sind, das für die Berechnung von Abfragen bezahlt](#), wird die Registerkarte Analyse nicht in der Konsole angezeigt.

Um aktuelle Abfragen zu sehen

Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.

1. Wählen Sie im linken Navigationsbereich Collaborations aus.
2. Wählen Sie eine Zusammenarbeit aus.
3. Wählen Sie auf der Registerkarte Analyse unter Analysen die Option Alle Abfragen aus der Dropdownliste aus und sehen Sie sich die Abfragen an, die in den letzten 90 Tagen ausgeführt wurden.
4. Um die letzten Abfragen nach Status zu sortieren, wählen Sie einen Status aus der Dropdownliste Alle Status aus.

Die Status lauten: Eingereicht, Gestartet, Storniert, Erfolgreich, Fehlgeschlagen und Zeitlimit überschritten.

Anzeigen von Abfragedetails

Sie können die Abfragedetails als Mitglied anzeigen, das Abfragen ausführen kann, oder als Mitglied, das Ergebnisse erhalten kann.

Um die Details der Abfrage anzuzeigen

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie eine Zusammenarbeit aus.
4. Führen Sie auf der Registerkarte Analyse einen der folgenden Schritte aus:
 - Wählen Sie das Optionsfeld für die spezifische Abfrage, die Sie anzeigen möchten, und klicken Sie dann auf Details anzeigen.
 - Wählen Sie die ID der geschützten Abfrage aus.
5. Auf der Seite mit den Abfragedetails
 - Wenn Sie das Mitglied sind, das Abfragen ausführen kann, sehen Sie sich die Abfragedetails, den SQL-Text und die Ergebnisse an.

Es wird eine Meldung angezeigt, die bestätigt, dass die Abfrageergebnisse an das Mitglied übermittelt wurden, das Ergebnisse empfangen kann.
 - Wenn Sie das Mitglied sind, das Ergebnisse erhalten kann, sehen Sie sich die Abfragedetails und Ergebnisse an.

PySpark Jobs werden ausgeführt

Als [Mitglied, das Abfragen durchführen kann](#), können Sie mithilfe einer genehmigten PySpark [Analysevorlage](#) einen PySpark Job in einer konfigurierten Tabelle ausführen.

Voraussetzungen

Bevor Sie einen PySpark Job ausführen können, müssen Sie über Folgendes verfügen:

- Eine aktive Mitgliedschaft bei AWS Clean Rooms Collaboration
- Zugriff auf mindestens eine Analysevorlage in der Kollaboration
- Zugriff auf mindestens eine konfigurierte Tabelle in der Kollaboration

- Berechtigungen zum Schreiben der Ergebnisse eines PySpark Jobs in einen angegebenen S3-Bucket

Hinweise zum Erstellen der erforderlichen Servicerolle finden Sie unter [Erstellen Sie eine Servicerolle, um die Ergebnisse eines PySpark Jobs zu schreiben](#).

- Das Mitglied, das für die Bezahlung der Rechenkosten verantwortlich ist, ist der Kollaboration als aktives Mitglied beigetreten

Informationen dazu, wie Sie Daten abfragen oder Abfragen anzeigen können, indem Sie den AWS Clean Rooms StartProtectedJob API-Vorgang direkt aufrufen oder den verwenden AWS SDKs, finden Sie in der [AWS Clean Rooms API-Referenz](#).

Hinweise zur Auftragsprotokollierung finden Sie unter [Analyse Einloggen AWS Clean Rooms](#).

Hinweise zum Empfangen von Auftragsergebnissen finden Sie unter [Empfangen und Verwenden von Analyseergebnissen](#).

In den folgenden Themen wird erklärt, wie Sie in einer Kollaboration mithilfe der AWS Clean Rooms Konsole einen PySpark Job für eine konfigurierte Tabelle ausführen.

Topics

- [Ausführen eines PySpark Jobs in einer konfigurierten Tabelle mithilfe einer PySpark Analysevorlage](#)
- [Aktuelle Jobs anzeigen](#)
- [Anzeigen von Auftragsdetails](#)

Ausführen eines PySpark Jobs in einer konfigurierten Tabelle mithilfe einer PySpark Analysevorlage

Dieses Verfahren zeigt, wie Sie eine PySpark Analysevorlage in der AWS Clean Rooms Konsole verwenden, um konfigurierte Tabellen mit der benutzerdefinierten Analyseregeln zu analysieren.

So führen Sie mithilfe einer PySpark Analysevorlage einen PySpark Job für eine konfigurierte Tabelle aus

Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.

1. Wählen Sie im linken Navigationsbereich Collaborations aus.
2. Wählen Sie die Kollaboration aus, für die der Status Ihrer Mitgliederfähigkeiten den Status Jobs ausführen lautet.
3. Sehen Sie sich auf der Registerkarte Analyse im Abschnitt Tabellen die Tabellen und den zugehörigen Analyseregeltyp an (benutzerdefinierte Analyseregeln).

 Note

Wenn Sie die erwarteten Tabellen nicht in der Liste sehen, kann das folgende Gründe haben:

- Die Tabellen wurden nicht [verknüpft](#).
- Für die Tabellen ist keine [Analyseregel konfiguriert](#).

4. Wählen Sie im Abschnitt Analyse für den Analysemodus die Option Analysevorlagen ausführen aus.
5. Wählen Sie die PySpark Analysevorlage aus der Dropdownliste Analysevorlage aus.

Die Parameter aus der PySpark Analysevorlage werden automatisch in die Definition übernommen.

6. Wenn für die Analysevorlage Parameter definiert sind, geben Sie unter Parameter Werte für die Parameter ein:
 - a. Sehen Sie sich für jeden Parameter den Parameternamen und den Standardwert an (falls konfiguriert).
 - b. Geben Sie für jeden Parameter, den Sie überschreiben möchten, einen Wert ein.

 Note

Wenn Sie keinen Wert angeben, aber ein Standardwert vorhanden ist, wird der Standardwert verwendet.

⚠ Important

Parameterwerte können bis zu 1.000 Zeichen lang sein und unterstützen die UTF-8-Kodierung. Alle Parameterwerte werden als Zeichenketten behandelt und über das Kontextobjekt an Ihr Benutzerskript übergeben.

Stellen Sie sicher, dass Ihr Benutzerskript Parameterwerte validiert und sicher verarbeitet. Weitere Hinweise zur sicheren Handhabung von Parametern finden Sie unter [Arbeiten mit Parametern in PySpark Analysevorlagen](#).

7. Geben Sie den unterstützten Worker-Typ und die Anzahl der Worker an.

Ermitteln Sie anhand der folgenden Tabelle den Typ und die Anzahl der Mitarbeiter, die Sie für Ihren Anwendungsfall benötigen.

Worker type (Worker-Typ)	vCPU	Speicher (GB)	Speicher (GB)	Number of workers (Anzahl der Worker)	Gesamtzahl der Verarbeitungseinheiten für Reinräume (CRPU)
CR.1X (Standard)	4	30	100	4	8
				128	256
CR.4X	16	120	400	4	32
				32	256

ℹ Note

Verschiedene Arten von Arbeitnehmern und die Anzahl der Arbeitnehmer sind mit Kosten verbunden. Weitere Informationen zu den Preisen finden Sie unter [AWS Clean Rooms Preisgestaltung](#).

8. Klicken Sie auf Ausführen.

 Note

Sie können den Job nicht ausführen, wenn das Mitglied, das Ergebnisse erhalten kann, die Einstellungen für die Job-Ergebnisse nicht konfiguriert hat.

9. Passen Sie die Parameter weiter an und führen Sie Ihren Job erneut aus, oder klicken Sie auf die Schaltfläche +, um einen neuen Job auf einer neuen Registerkarte zu starten.

Aktuelle Jobs anzeigen

Auf der Registerkarte Analyse können Sie sich die Jobs ansehen, die in den letzten 90 Tagen ausgeführt wurden.

 Note

Wenn Sie als Mitglied nur Contribute-Daten haben und Sie nicht das [Mitglied sind, das die Kosten für die Auftragsverarbeitung bezahlt](#), wird die Registerkarte Analyse nicht in der Konsole angezeigt.

Um aktuelle Jobs zu sehen

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie eine Zusammenarbeit aus.
4. Wählen Sie auf der Registerkarte Analyse unter Analysen die Option Alle Jobs aus der Dropdownliste aus und sehen Sie sich die Jobs an, die in den letzten 90 Tagen ausgeführt wurden.
5. Um die letzten Jobs nach Status zu sortieren, wählen Sie einen Status aus der Dropdownliste Alle Status aus.

Die Status lauten: Eingereicht, Gestartet, Storniert, Erfolgreich, Fehlgeschlagen und Zeitlimit überschritten.

Anzeigen von Auftragsdetails

Sie können sich die Jobdetails als Mitglied ansehen, das Jobs ausführen kann, oder als Mitglied, das Ergebnisse erhalten kann.

Um die Details des Jobs einzusehen

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie eine Zusammenarbeit aus.
4. Wählen Sie auf der Registerkarte Analyse unter Analysen die Option Alle Jobs aus der Dropdownliste aus, und führen Sie dann einen der folgenden Schritte aus:
 - Wählen Sie das Optionsfeld für den spezifischen Job, den Sie anzeigen möchten, und wählen Sie dann Details anzeigen aus.
 - Wählen Sie die geschützte Job-ID aus.
5. Auf der Seite mit den Jobdetails
 - Wenn Sie das Mitglied sind, das Jobs ausführen kann, sehen Sie sich die Jobdetails, den Job und die Ergebnisse an.

Es wird eine Meldung angezeigt, in der bestätigt wird, dass die Auftragsergebnisse an das Mitglied übermittelt wurden, das Ergebnisse erhalten kann.
 - Wenn Sie das Mitglied sind, das Ergebnisse erhalten kann, sehen Sie sich die Jobdetails und Ergebnisse an.

Empfangen und Verwenden von Analyseergebnissen

[Mitglieder, die Ergebnisse erhalten können](#), können die Abfrageergebnisse entweder in der AWS Clean Rooms Konsole oder in den Amazon S3 S3-Buckets überprüfen, die sie bei ihrem Beitritt zur Kollaboration angegeben haben.

Note

Nur bei verschlüsselten Datentabellen entschlüsselt das Mitglied, das Ergebnisse empfangen kann, die Abfrageergebnisse, indem es den C3R-Verschlüsselungsclient im Entschlüsselungsmodus ausführt.

Das Ergebnisziel in Amazon S3 darf sich nicht im selben S3-Bucket wie jede Datenquelle befinden.

In den folgenden Themen wird erklärt, wie Sie Analyseergebnisse mithilfe der AWS Clean Rooms Konsole abrufen können.

Topics

- [Abfrageergebnisse werden empfangen](#)
- [Empfangen von Auftragsergebnissen](#)
- [Standardwerte für Einstellungen für Abfrageergebnisse bearbeiten](#)
- [Bearbeiten der Standardwerte für die Einstellungen für die Auftragsergebnisse](#)
- [Verwenden der Abfrageausgabe in anderen AWS-Services](#)

Informationen dazu, wie Sie Daten abfragen oder Abfragen anzeigen können, indem Sie die AWS Clean Rooms API direkt aufrufen oder die verwenden AWS SDKs, finden Sie in der [AWS Clean Rooms API-Referenz](#).

Hinweise zur Abfrageprotokollierung finden Sie unter [Analyse Einloggen AWS Clean Rooms](#).

Note

Wenn Sie eine Abfrage für verschlüsselte Datentabellen ausführen, werden die Ergebnisse der verschlüsselten Spalten verschlüsselt.

Abfrageergebnisse werden empfangen

Note

Das Ergebnisziel in Amazon S3 darf sich nicht im selben S3-Bucket wie jede Datenquelle befinden.

Die Ergebnisse der Abfrage befinden sich im Abschnitt Standardwerte für Ergebniseinstellungen auf der Registerkarte Analyse in der AWS Clean Rooms Konsole.

Um Abfrageergebnisse zu erhalten

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#) mit Ihrem AWS-Konto (falls Sie dies noch nicht getan haben).
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus, für die der Status Ihrer Mitgliederfähigkeiten den Status Ergebnisse erhalten lautet.
4. Um die Abfrageergebnisse direkt von zu erhalten AWS Clean Rooms, wählen Sie auf der Registerkarte Analyse unter Analysen die Option Alle Abfragen aus der Dropdownliste aus und wählen Sie dann in der Spalte Geschützte Abfrage-ID die Abfrage aus.
5. Führen Sie auf der Seite mit den Abfragedetails unter Ergebnisse einen der folgenden Schritte aus:

Vorhaben	Wählen Sie dann...
Kopieren Sie die Ergebnisse.	Copy
Laden Sie die Ergebnisse herunter.	Download

Note

Standardmäßig entspricht der Name der heruntergeladenen Datei dem NamenQuery id, der beim

Vorhaben	Wählen Sie dann...
Sehen Sie sich die Ergebnisse in Amazon S3 an.	Ausführen der Abfrage angezeigt wurde. AWS Clean Rooms In Amazon S3 anzeigen Die Amazon S3 S3-Konsole wird auf einer separaten Registerkarte geöffnet.

6. Wenn Sie verschlüsselte Daten verwenden, können Sie die Datentabellen jetzt [entschlüsseln](#).

Weitere Informationen finden Sie unter [Datentabellen mit dem C3R-Verschlüsselungsclient entschlüsseln](#).

Empfangen von Auftragsergebnissen

Note

Das Ergebnisziel in Amazon S3 darf sich nicht im selben S3-Bucket wie jede Datenquelle befinden.

Die Ergebnisse des Jobs befinden sich in der AWS Clean Rooms Konsole im Abschnitt Standardeinstellungen für Ergebniseinstellungen auf der Registerkarte Analyse.

Um Job-Ergebnisse zu erhalten

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#) mit Ihrem AWS-Konto (falls Sie dies noch nicht getan haben).
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus, für die der Status Ihrer Mitgliederfähigkeiten den Status Ergebnisse erhalten lautet.
4. Um die Job-Ergebnisse direkt von zu erhalten AWS Clean Rooms, wählen Sie auf der Registerkarte Analyse unter Analysen die Option Alle Jobs aus der Dropdownliste aus und wählen Sie dann in der Spalte Geschützte Job-ID den Job aus.
5. Kopieren Sie auf der Seite mit den Jobdetails unter Ergebnisse die Job-ID.

Kehren Sie zur Registerkarte Analyse zurück und erweitern Sie die Standardeinstellungen für die Ergebniseinstellungen.

Wählen Sie unter Ergebnisziel den Link aus, um die Ergebnisse in Amazon S3 anzuzeigen.

Die Amazon S3 S3-Konsole wird auf einer separaten Registerkarte geöffnet.

Fügen Sie in Amazon S3 die Job-ID in die Suchleiste ein und drücken Sie die Eingabetaste.

Der Ordner mit den Ergebnissen wird angezeigt. Wählen Sie den Ordner aus, um die Job-Ergebnisse anzuzeigen.

Standardwerte für Einstellungen für Abfrageergebnisse bearbeiten

Note

Das Ergebnisziel in Amazon S3 darf sich nicht im selben S3-Bucket wie jede Datenquelle befinden.

Als Mitglied, das Ergebnisse empfangen kann, können Sie die Standardwerte für die Einstellungen für Abfrageergebnisse in der AWS Clean Rooms Konsole bearbeiten.

Um die Standardwerte für die Einstellungen für Abfrageergebnisse zu bearbeiten

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#) mit Ihrem AWS-Konto (falls Sie dies noch nicht getan haben).
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie die Kollaboration aus, für die der Status Ihrer Mitgliederfähigkeiten den Status Ergebnisse erhalten lautet.
4. Wählen Sie auf der Registerkarte Analyse unter Standardeinstellungen für Ergebniseinstellungen die Option Bearbeiten aus.
5. Ändern Sie auf der Seite Standardwerte für Ergebniseinstellungen bearbeiten die folgenden Optionen nach Bedarf:
 - a. Ändern Sie unter Ergebnisse abfragen das Ergebnisziel in Amazon S3, das Ergebnisformat oder die Ergebnisdateien.

- b. (Optional) Wenn Sie Anfragen, die bis zu 24 Stunden dauern, an Ihr S3-Ziel senden möchten, aktivieren Sie für den Servicezugriff das Kontrollkästchen Servicerolle hinzufügen, um Anfragen zu unterstützen, deren Bearbeitung bis zu 24 Stunden in Anspruch nimmt.

Umfangreiche Anfragen, deren Bearbeitung bis zu 24 Stunden in Anspruch nimmt, werden an Ihr S3-Ziel zugestellt.

Wenn Sie das Kontrollkästchen nicht aktivieren, werden nur Anfragen, die innerhalb von 12 Stunden abgeschlossen wurden, an Ihren S3-Standort zugestellt.

- Geben Sie die Dienstzugriffsberechtigungen an, indem Sie entweder Neue Servicerolle erstellen und verwenden oder Bestehende Servicerolle verwenden auswählen.

Create and use a new service role

- AWS Clean Rooms erstellt eine Servicerolle mit der erforderlichen Richtlinie für diese Tabelle.
- Der Standardname der Servicerolle lautet `cleanrooms-query-receiver-
<timestamp>`
- Sie benötigen die erforderlichen Berechtigungen, um Rollen zu erstellen und Richtlinien anzuhängen.

Use an existing service role

1. Wählen Sie einen vorhandenen Servicerollennamen aus der Dropdownliste aus.

Die Liste der Rollen wird angezeigt, wenn Sie berechtigt sind, Rollen aufzulisten.

Wenn Sie nicht berechtigt sind, Rollen aufzulisten, können Sie den Amazon-Ressourcennamen (ARN) der Rolle eingeben, die Sie verwenden möchten.

2. Rufen Sie die Servicerolle auf, indem Sie auf den externen Link In IAM anzeigen klicken.

Wenn keine vorhandenen Servicerollen vorhanden sind, ist die Option „Eine bestehende Servicerolle verwenden“ nicht verfügbar.

Versucht standardmäßig AWS Clean Rooms nicht, die bestehende Rollenrichtlinie zu aktualisieren, um die erforderlichen Berechtigungen hinzuzufügen.

 Note

- AWS Clean Rooms erfordert Berechtigungen für Abfragen gemäß den Analyseregeln. Weitere Informationen zu Berechtigungen für AWS Clean Rooms finden Sie unter [AWS verwaltete Richtlinien für AWS Clean Rooms](#).
- Wenn die Rolle nicht über ausreichende Berechtigungen für verfügt AWS Clean Rooms, erhalten Sie eine Fehlermeldung, dass die Rolle nicht über ausreichende Berechtigungen für verfügt AWS Clean Rooms. Die Rollenrichtlinie muss hinzugefügt werden, bevor Sie fortfahren können.
- Wenn Sie die Rollenrichtlinie nicht ändern können, erhalten Sie eine Fehlermeldung, dass die Richtlinie für die Servicerolle nicht gefunden werden AWS Clean Rooms konnte.

6. Wählen Sie **Änderungen speichern** aus.
7. Die aktualisierten Einstellungen für die Abfrageergebnisse werden auf der Detailseite der Zusammenarbeit angezeigt.

Bearbeiten der Standardwerte für die Einstellungen für die Auftragsergebnisse

 Note

Das Ergebnisziel in Amazon S3 darf sich nicht im selben S3-Bucket wie jede Datenquelle befinden.

Als Mitglied, das Ergebnisse empfangen kann, können Sie die Standardwerte für die Einstellungen für die Auftragsergebnisse in der AWS Clean Rooms Konsole bearbeiten.

Um die Standardwerte für die Einstellungen für die Arbeitsergebnisse zu bearbeiten

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#) mit Ihrem AWS-Konto (falls Sie dies noch nicht getan haben).
2. Wählen Sie im linken Navigationsbereich **Collaborations** aus.

3. Wählen Sie die Kollaboration aus, für die der Status Ihrer Mitgliederfähigkeiten den Status Ergebnisse erhalten lautet.
4. Wählen Sie auf der Registerkarte Analyse unter Standardeinstellungen für Ergebniseinstellungen die Option Bearbeiten aus.
5. Ändern Sie auf der Seite Standardwerte für Ergebniseinstellungen bearbeiten die folgenden Optionen nach Bedarf:
 - a. Ändern Sie unter Auftragsergebnisse das Ergebnisziel in Amazon S3.
 - b. Ändern Sie unter Servicezugriff den Namen der vorhandenen Servicerolle.
6. Wählen Sie Änderungen speichern aus.
7. Die aktualisierten Einstellungen für die Auftragsergebnisse werden auf der Detailseite der Zusammenarbeit angezeigt.

Verwenden der Abfrageausgabe in anderen AWS-Services

Die SQL-Abfrageausgabe kann für die Ausgangsdaten für ein Clean Rooms-ML-Modell verwendet werden. Weitere Informationen finden Sie unter [AWS Clean Rooms ML](#).

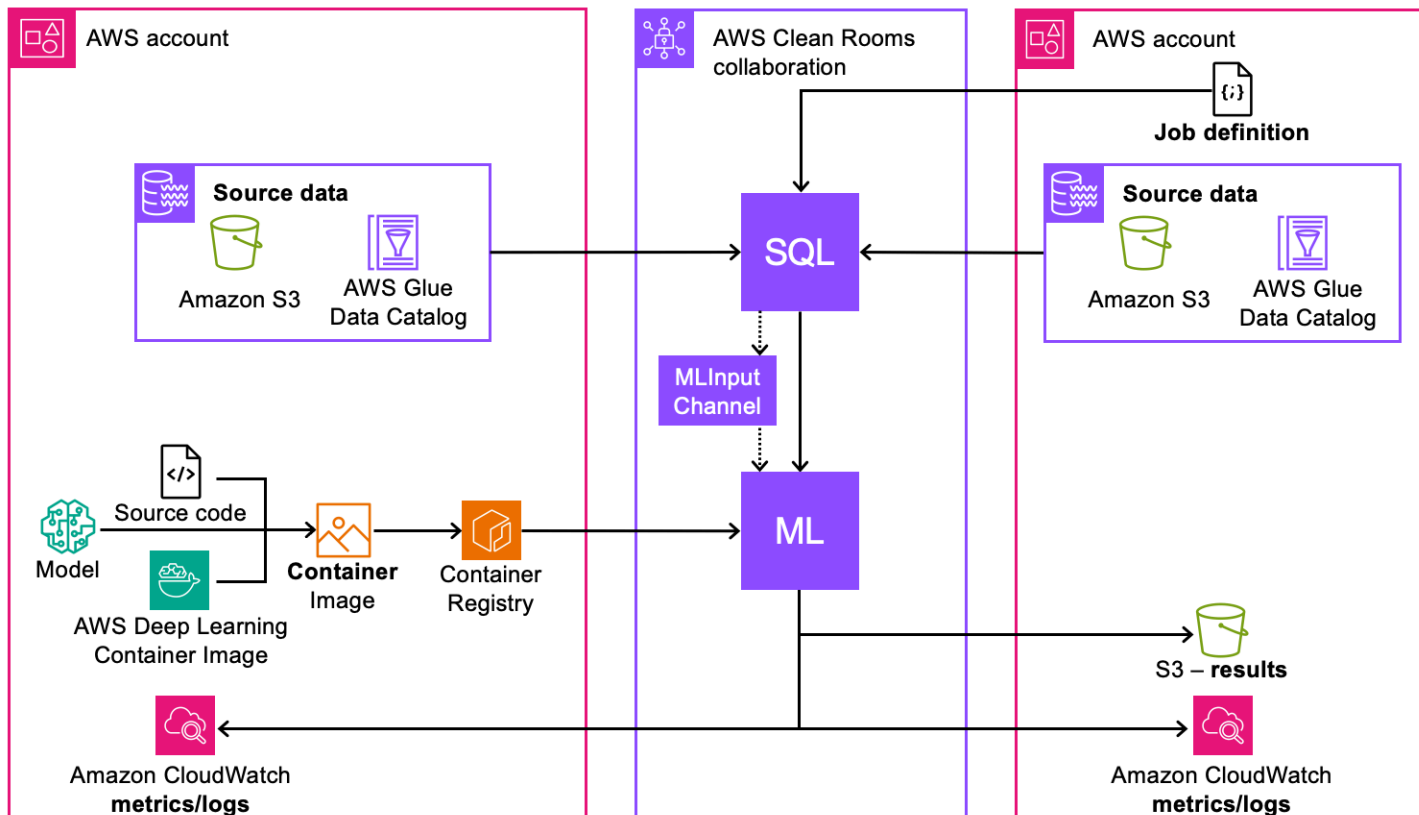
Die Abfrageausgabe von AWS Clean Rooms ist auf der Konsole verfügbar (wenn die Konsole zum Ausführen von Abfragen verwendet wird) und in einem angegebenen Amazon S3 S3-Bucket heruntergeladen. Von dort aus können Sie die Abfrageausgabe in anderen Programmen AWS-Services wie Amazon Quick und Amazon SageMaker AI verwenden, je nachdem, wie diese Dienste Daten aus Amazon S3 verwenden.

Weitere Informationen zu Amazon Quick finden Sie in der [Amazon Quick-Dokumentation](#).

Weitere Informationen zu Amazon SageMaker AI finden Sie in der [Amazon SageMaker AI-Dokumentation](#).

Benutzerdefinierte Modellierung in AWS Clean Rooms ML

Aus technischer Sicht beschreibt das folgende Diagramm, wie die benutzerdefinierte ML-Modellierung in AWS Clean Rooms ML funktioniert.



So funktioniert die benutzerdefinierte ML-Modellierung in Clean Rooms ML:

1. Konfiguration der Datenquelle

- Quelldaten können im Amazon S3 S3-Katalog AWS Glue Data Catalog, im oder Snowflake gespeichert werden
- AWS Glue Data Catalog wird zum Organisieren und Katalogisieren verwendet
- Daten von mehreren AWS-Konten können innerhalb derselben Zusammenarbeit verwendet werden

2. SQL-Abfrage und Datenverarbeitung

- SQL-Abfragen werden verwendet, um auf die Quelldaten zuzugreifen und sie zu verarbeiten
- Die Abfragen laufen innerhalb der Grenzen der AWS Clean Rooms Zusammenarbeit
- Verarbeitete Daten werden für das Modelltraining in ML-Eingangskanäle eingespeist

3. Entwicklung von ML-Modellen

- Der Quellcode für das Modell kann mithilfe von AWS Deep Learning-Container-Images entwickelt werden
- Benutzerdefinierte Container-Images müssen erstellt und in der Amazon Elastic Container Registry gespeichert werden

4. Komponenten der Infrastruktur

- Amazon Elastic Container Registry speichert und verwaltet die ML-Modellcontainer
- Die ML-Verarbeitung erfolgt innerhalb der sicheren AWS Clean Rooms Kollaborationsumgebung

5. Überwachung und Protokollierung

- Amazon CloudWatch stellt Metriken und Protokolle für beide Kooperationspartner bereit.
- Die Überwachung ist für alle an der Zusammenarbeit AWS-Konten Beteiligten verfügbar
- Leistungskennzahlen und Betriebsprotokolle sind für relevante Parteien zugänglich

6. Verwaltung der Ergebnisse

- Der Zugriff auf Ergebnisse wird entsprechend den Kollaborationsberechtigungen gesteuert

Bevor Sie beginnen, finden Sie weitere [Richtlinien für die Modellerstellung für den Trainingscontainer](#) Informationen unter [Voraussetzungen für die benutzerdefinierte ML-Modellierung](#) und.

Generierung synthetischer Datensätze mit verbessertem Datenschutz

Ein synthetischer Datensatz hat ähnliche statistische Eigenschaften wie der ursprüngliche Datensatz, auf dem er basiert, enthält jedoch nicht die realen Beobachtungen, die im ursprünglichen Datensatz enthalten sind. Durch die Verwendung von synthetischen Datensätzen mit verbessertem Datenschutz können Sie neue Anwendungsfälle für das Modelltraining mit maschinellem Lernen (ML) erschließen, die bisher aufgrund von Datenschutzproblemen verhindert wurden. Wenn Sie einen ML-Eingabekanal erstellen, können Sie synthetische Daten generieren, um vertrauliche Informationen beim Training von ML-Modellen zu schützen.

Wenn Sie eine Vorlage mit synthetischen Daten erstellen, müssen Sie:

- Erfordert, dass die Vorlagenausgabe synthetisch ist
- Klassifizieren Sie die Spalten des Ausgabeschemas als numerisch oder kategorisch
- Passen Sie synthetische Daten an die Bedürfnisse Ihres Unternehmens an

- Passen Sie die Datenschutzeinstellungen an:
 - Datenschutzstufe festlegen (Epsilon)
 - Datenschutzwert konfigurieren

Warning

Die Generierung synthetischer Daten schützt davor, auf individuelle Merkmale zu schließen, unabhängig davon, ob bestimmte Personen im ursprünglichen Datensatz vorhanden sind oder Lernattribute dieser Personen vorhanden sind. Es verhindert jedoch nicht, dass wörtliche Werte aus dem ursprünglichen Datensatz, einschließlich persönlich identifizierbarer Informationen (PII), im synthetischen Datensatz erscheinen.

Wir empfehlen, Werte im Eingabedatensatz zu vermeiden, die nur einer betroffenen Person zugeordnet sind, da diese eine betroffene Person neu identifizieren können. Wenn beispielsweise nur ein Benutzer in einer Postleitzahl wohnt, würde das Vorhandensein dieser Postleitzahl im synthetischen Datensatz bestätigen, dass sich der Benutzer im ursprünglichen Datensatz befand. Techniken wie das Kürzen hochgenauer Werte oder das Ersetzen ungewöhnlicher Kataloge durch andere können verwendet werden, um dieses Risiko zu minimieren. Diese Transformationen können Teil der Abfrage sein, die zur Erstellung des ML-Eingabekanals verwendet wird.

Weitere Informationen zum Generieren synthetischer Daten für das Training benutzerdefinierter Modelle finden Sie unter [Erstellen einer SQL-Analysevorlage](#).

Analysevorlagen mit synthetischen Ausgaben können nur zur Erstellung von ML-Eingangskanälen verwendet werden. Weitere Informationen finden Sie unter [Erstellen eines ML-Eingabekanals in AWS Clean Rooms ML](#).

Überlegungen zur Generierung synthetischer Daten

Mit AWS Clean Rooms ML können Mitglieder der Kollaboration einen synthetischen Datensatz erstellen, der die Subjekte des ursprünglichen Datensatzes unwiderruflich aus ihren kollektiven Datensätzen anonymisiert, um ein benutzerdefiniertes Modell für maschinelles Lernen zu trainieren. Bei der Erstellung der Zusammenarbeit müssen Sie die Zahlungsinformationen so konfigurieren, dass angegeben wird, wer für die Generierung synthetischer Daten bezahlt. Im Folgenden finden Sie die wichtigsten Schritte zum Generieren eines synthetischen Datensatzes und zum Trainieren eines benutzerdefinierten Modells für maschinelles Lernen:

1. Ein Mitglied der Kollaboration erstellt eine Analysevorlage, die Folgendes beinhaltet:
 - Das SQL, das zur Definition des zu synthetisierenden Datensatzes benötigt wird.
 - Datenschutzbezogene Konfigurationen, die verwendet werden, um sicherzustellen, dass die synthetischen Daten den Compliance-Anforderungen der Datenanbieter entsprechen.
2. Sobald alle Datenanbieter die Analysevorlage genehmigt haben, erstellt der Collaboration Query Runner anhand der Vorlage einen Eingabekanal für maschinelles Lernen (ML).
3. Clean Rooms ML generiert den synthetischen Datensatz und überprüft, ob er die in der Analysevorlage angegebenen Datenschutzwerte erfüllt.
4. Wenn alle Schwellenwerte erfüllt sind, wird der ML-Eingabekanal mit dem synthetischen Datensatz gefüllt.
5. Kunden können diesen ML-Eingangskanal dann verwenden, um das mit der Zusammenarbeit verknüpfte benutzerdefinierte ML-Modell zu trainieren.

Wichtige Überlegungen:

- Bei synthetischen Daten, die in Clean Rooms ML generiert wurden, werden keine individuellen Werte entfernt, geschwärzt, verschleiert oder bereinigt, einschließlich personenbezogener Daten (PII), die im ursprünglichen Datensatz gefunden wurden. Der synthetische Datensatz wird generiert, indem Werte, aber nicht ganze Datensätze, aus dem ursprünglichen Datensatz entnommen werden.
- Wenn der ursprüngliche Datensatz ähnliche Zeilen enthält, ist es möglich, dass die synthetischen Daten Zeilen enthalten, die identisch mit den Zeilen im ursprünglichen Datensatz aussehen.

Vorbereitung des Datensatzes:

- Vermeiden Sie Spalten mit einer deutlich unausgewogenen Klassenverteilung. Dies ist besonders wichtig für den prognostizierten Wert oder die Spalte „Y“. Extreme Ungleichgewichte beeinträchtigen die allgemeine Privatsphäre des synthetischen Datensatzes.
- Clean Rooms ML unterstützt nicht die Generierung synthetischer Daten aus Zeitreihendaten, bei denen es wichtig ist, die Korrelationen zwischen sequentiellen Datensätzen aufrechtzuerhalten.
- Clean Rooms ML unterstützt nicht die Generierung synthetischer Daten aus Text oder unstrukturierten Daten.
- Die folgenden Datentypen werden unterstützt:

Name
des
Datentyps

BIGINT

BOOLEAN

CHAR

DATE

DECIMAL

FLOAT

INTEGER

LONG

REAL

SHORT

SMALLINT

TIME

TIMESTAMP
_LTZ

TIMESTAMP
_NTZ

TINYINT

VARCHAR

Einschränkungen:

- Bei der Generierung synthetischer Daten ist die maximale Anzahl prädiktiver Spalten eins.
- Wenn die Zielspalte kategorisch ist, beträgt die maximale Anzahl von Kategorien im Originaldatensatz 100.
- Im ursprünglichen Datensatz muss die Anzahl der Zeilen zwischen 1.500 und 2,5 Millionen liegen und die maximale Anzahl von Spalten beträgt 1.000. Für Werte ungleich Null in der Zielspalte beträgt die Mindestanzahl von Zeilen 1.000.

Kennzahlen zum Datenschutz:

- Clean Rooms ML bietet einen Datenschutz-Score, der misst, wie geschützt die generierten synthetischen Daten vor Angriffen auf Mitgliedschaftsinformationen sind (MIAs). Der Dienst verwendet 5% der Originaldaten aus dem Syntheseprozess, um diesen Wert zu berechnen.
- Werte von fast 50% werden als gut angesehen. Höhere Werte bedeuten weniger Schutz vor MIAs. Werte, die deutlich unter 50% liegen, sind selten und können darauf zurückzuführen sein, dass Muster aus den Originaldaten in den synthetisierten Daten nicht abgebildet sind.

Benutzerdefiniertes Downstream-Modell:

- In Clean Rooms ML generierte synthetische Daten eignen sich am besten für das Training von binären Klassifikationsmodellen und Klassifikationsmodellen mit mehreren Klassen mit bis zu fünf Klassen.
- Das Training von Regressionsmodellen mit synthetischen Daten, die in Clean Rooms ML generiert wurden, kann zu einer geringen Modellgenauigkeit führen, die anhand des Root Mean Square Error (RMSE) gemessen wird.

Aufbau und Beitritt zur Zusammenarbeit in AWS Clean Rooms ML

Der Ersteller der Kollaboration ist dafür verantwortlich, die Kollaboration zu erstellen, Mitglieder einzuladen und ihnen Rollen zuzuweisen. Die eingeladenen Mitglieder treten der Kollaboration bei und legen die Ergebniseinstellungen und Zieleinstellungen für die Artefakte des trainierten Modells fest und übernehmen die Zahlungsverpflichtungen, je nachdem, wie die Kollaboration eingerichtet ist.

Aufbau einer Zusammenarbeit für maschinelles Lernen

Das folgende Verfahren zeigt, wie Sie eine Kollaboration für maschinelles Lernen einrichten, ein oder mehrere Mitglieder einladen und Mitglieder zuweisen, die mit dem Modelltraining beginnen,

Ergebnisse erhalten, trainierte Modellergebnisse, einschließlich Modellartefakte und Metriken, erhalten und Modellinferenzergebnisse erhalten können. Der Ersteller der Kollaboration weist außerdem ein Mitglied zu, das die Kosten für die Berechnung von Abfragen, das Modelltraining und die Modellinferenz trägt.

Console

Um eine Kollaboration für maschinelles Lernen zu erstellen (Konsole)

1. [Erstellen Sie eine Kollaboration und laden Sie ein oder mehrere Mitglieder ein, der Kollaboration beizutreten](#)
2. Weisen Sie Mitgliedern mithilfe von Abfragen und Jobs die folgenden Analysefähigkeiten zu:
 - Weisen Sie dem Mitglied, das mit dem Modelltraining beginnen wird, Abfragen ausführen zu.
 - Weisen Sie den Mitgliedern, die Abfrageergebnisse erhalten sollen, die Option Ergebnisse aus der Analyse empfangen zu.
3. Weisen Sie Mitgliedern mithilfe speziell entwickelter Workflows die folgenden Fähigkeiten für die ML-Modellierung zu:
 - Weisen Sie dem Mitglied, das die Ergebnisse des trainierten Modells, einschließlich Modellartefakten und Metriken, erhalten soll, die Empfangsausgabe von trainierten Modellen zu.
 - Weisen Sie dem Mitglied, das die Ergebnisse der Modellinferenz erhalten soll, die Ergebnisse der Modellinferenz empfangen zu.
4. Geben Sie unter Zahlung konfigurieren die Mitglieder an, die für die Kosten für die Abfrageberechnung, das Modelltraining, die Modellinferenz und die Generierung synthetischer Datensätze zahlen sollen.

Ein synthetischer Datensatz hat ähnliche statistische Eigenschaften wie der ursprüngliche Datensatz, auf dem er basiert, enthält jedoch nicht die realen Beobachtungen, die im ursprünglichen Datensatz vorhanden sind. Durch die Verwendung von synthetischen Datensätzen mit verbessertem Datenschutz können Sie neue Anwendungsfälle für das Training mit ML-Modellen erschließen, die Datenschutzprobleme bisher verhindert haben. Weitere Informationen finden Sie unter Generierung synthetischer Daten mit verbessertem [Datenschutz](#).

Jede dieser Kosten kann demselben oder unterschiedlichen Mitgliedern zugewiesen werden. Wenn es sich bei einem eingeladenen Mitglied um das Mitglied handelt, das für die Zahlung der Zahlungskosten verantwortlich ist, muss es seine Zahlungsverpflichtungen akzeptieren, bevor es der Zusammenarbeit beiträgt.

5. Bei der Option Mitgliedschaft konfigurieren kann der Ersteller der Kollaboration entscheiden, ob er der Mitgliedschaft jetzt beitreten oder später eine Mitgliedschaft erstellen möchte. Der Ersteller der Kollaboration muss dann die ML-Konfiguration einrichten.
 - a. Wenn der Ersteller der Kollaboration auch der Empfänger der Ergebnisse ist, muss er auch das Ziel und das Format der Abfrageergebnisse in den Standardeinstellungen für die Ergebniseinstellungen angeben.
 - b. Die ML-Konfiguration bietet Clean Rooms ML die Rolle, Metriken auf einem AWS-Konto zu veröffentlichen. Wenn der Kollaborationsersteller auch trainierte Modellartefakte erhält, kann er den Amazon S3 S3-Bucket angeben, der für den Empfang von Ergebnissen verwendet wird.
 - c. Wählen Sie im Abschnitt ML-Konfigurationen die Option ML-Konfiguration erstellen aus und geben Sie dann das Model-Ausgabeziel auf Amazon S3 und die Service-Zugriffsrolle an, die für den Zugriff auf diesen Speicherort erforderlich ist.
 - d. Wenn der Gründer der Kollaboration das Mitglied ist, das für die Zahlung der Zahlungskosten verantwortlich ist, muss er vor der Erstellung der Kollaboration seine Zahlungsverpflichtungen akzeptieren.

API

Um eine Kollaboration für maschinelles Lernen (API) zu erstellen

1. [Erstellen Sie eine Kollaboration und laden Sie ein oder mehrere Mitglieder ein, der Kollaboration beizutreten](#)
2. Weisen Sie Mitgliedern der Kollaboration die folgenden Rollen zu:
 - CAN_QUERY- wird dem Mitglied zugewiesen, das mit dem Modelltraining und der Inferenz beginnen wird.
 - CAN_RECEIVE_MODEL_OUTPUT- wird den Mitgliedern zugewiesen, die die Ergebnisse des trainierten Modells erhalten.
 - CAN_RECEIVE_INFERENCE_OUTPUT- wird den Mitgliedern zugewiesen, die die Ergebnisse der Modellinferenz erhalten.

Wenn der Kollaborationsersteller auch der Empfänger der Ergebnisse ist, muss er bei der Erstellung der Kollaboration auch das Ziel und das Format der Abfrageergebnisse angeben. Sie weisen auch eine Servicerolle Amazon Resource Name (ARN) zu, um die Ergebnisse in das Ziel der Abfrageergebnisse zu schreiben.

3. Geben Sie die Mitglieder an, die für die Kosten für die Berechnung von Abfragen, das Modelltraining und die Modellinferenz aufkommen sollen. Jede dieser Kosten kann denselben oder unterschiedlichen Mitgliedern zugewiesen werden. Wenn es sich bei einem eingeladenen Mitglied um das Mitglied handelt, das für die Zahlung der Zahlungskosten verantwortlich ist, muss es seine Zahlungsverpflichtungen akzeptieren, bevor es der Zusammenarbeit beitrifft.
4. Der folgende Code erstellt eine Kollaboration, lädt ein Mitglied ein, das Abfragen ausführen und Ergebnisse empfangen kann, und gibt den Kollaborationsersteller als Empfänger der Modellartefakte an.

```
import boto3
acr_client= boto3.client('cleanrooms')

collaboration = a_acr_client.create_collaboration(
    members=[
        {
            'accountId': 'invited_member_accountId',
            'memberAbilities':["CAN_QUERY","CAN_RECEIVE_RESULTS"],
            'displayName': 'member_display_name'
        }
    ],
    name='collaboration_name',
    description=collaboration_description,
    creatorMLMemberAbilities= {
        'customMLMemberAbilities':["CAN_RECEIVE_MODEL_OUTPUT",
"CAN_RECEIVE_INFERENCE_OUTPUT"],
    },
    creatorDisplayName='creator_display_name',
    queryLogStatus="ENABLED",
    analyticsEngine="SPARK",
    creatorPaymentConfiguration={
        "queryCompute": {
            "isResponsible": True
        },
        "machineLearning": {
            "modelTraining": {
                "isResponsible": True
            }
        }
    }
)
```

```

        },
        "modelInference": {
            "isResponsible": True
        }
    }
}
)

collaboration_id = collaboration['collaboration']['id']
print("collaborationId: {collaboration_id}")

member_membership = a_acr_client.create_membership(
    collaborationIdentifier = collaboration_id,
    queryLogStatus = 'ENABLED',
    paymentConfiguration={
        "queryCompute": {
            "isResponsible": True
        },
        "machineLearning": {
            "modelTraining": {
                "isResponsible": True
            },
            "modelInference": {
                "isResponsible": True
            }
        }
    }
)

```

5. Der Ersteller der Kollaboration muss dann die ML-Konfiguration einrichten. Die ML-Konfiguration bietet Clean Rooms ML die Rolle, Metriken und Protokolle auf einem zu veröffentlichen AWS-Konto. Wenn der Ersteller der Kollaboration auch Ergebnisse erhält (Modellartefakte oder Inferenzergebnisse), kann er den Amazon S3 S3-Bucket angeben, der für den Empfang von Ergebnissen verwendet wird.

```

import boto3
acr_ml_client= boto3.client('cleanroomsmachinelearning')

acr_ml_client.put_ml_configuration(
    membershipIdentifier=membership_id,
    defaultOutputLocation={
        'roleArn': 'arn:aws:iam::account:role/roleName',
        'destination':{

```

```
        's3Destination':{
            's3Uri':"s3://bucketName/prefix"
        }
    }
}
```

Einer Kollaboration beitreten

Nachdem der Ersteller der Kollaboration seine Aufgaben erledigt hat, müssen die eingeladenen Mitglieder ihre Aufgaben erledigen.

Console

Um eine Mitgliedschaft zu erstellen und einer Kollaboration beizutreten (Konsole)

1. Das eingeladene Mitglied [erstellt eine Mitgliedschaft und tritt der Kollaboration](#) bei.
2. Wenn es sich bei dem eingeladenen Mitglied um das Mitglied handelt, das für die Zahlung der Kosten, einschließlich der Kosten für die Berechnung von Abfragen, das Modelltraining und die Modellinferenz verantwortlich ist, muss es seine Zahlungsverpflichtungen akzeptieren, bevor es der Kollaboration beitreten kann.
3. Das eingeladene Mitglied richtet die ML-Konfiguration ein, die Clean Rooms ML die Rolle einräumt, Modellmetriken auf einem AWS-Konto zu veröffentlichen. Wenn sie auch das Mitglied sind, das trainierte Modellartefakte erhält, muss sie einen Amazon S3 S3-Bucket angeben, in dem trainierte Modellartefakte gespeichert werden.

API

Um eine Mitgliedschaft zu erstellen und einer Kollaboration beizutreten (API)

1. Wenn das eingeladene Mitglied das Mitglied ist, das Ergebnisse erhalten kann, gibt es das Ziel und das Format der Abfrageergebnisse an. Sie bieten auch eine Dienstrolle (ARN), die es dem Dienst ermöglicht, in das Ziel der Abfrageergebnisse zu schreiben.

Wenn es sich bei dem eingeladenen Mitglied um das Mitglied handelt, das für die Bezahlung der Kosten, einschließlich der Kosten für die Abfrageberechnung, das Modelltraining und die Modellinferenz verantwortlich ist, muss es seine Zahlungsverpflichtungen akzeptieren, bevor es der Kollaboration beitreten kann.

Handelt es sich bei dem eingeladenen Mitglied um das Mitglied, das für die Bezahlung von Modellschulungen und Modellinferenzen für benutzerdefinierte Modelle verantwortlich ist, muss es seinen Zahlungsverpflichtungen nachkommen, bevor es der Zusammenarbeit beitrifft.

Der folgende Code erstellt eine Mitgliedschaft mit aktivierter Abfrageprotokollierung.

```
import boto3
acr_client= boto3.client('cleanrooms')

acr_client.create_membership(
    membershipIdentifier='membership_id',
    queryLogStatus='ENABLED'
)
```

2. Das eingeladene Mitglied richtet die ML-Konfiguration ein, die Clean Rooms ML die Rolle einräumt, Modellmetriken auf einem zu veröffentlichen AWS-Konto. Wenn sie auch das Mitglied sind, das trainierte Modellartefakte erhält, muss sie einen Amazon S3 S3-Bucket bereitstellen, in dem trainierte Modellartefakte gespeichert werden.

```
import boto3
acr_ml_client= boto3.client('cleanroomsmml')

acr_ml_client.put_ml_configuration(
    membershipIdentifier='membership_id',
    defaultOutputLocation={
        'roleArn':"arn:aws:iam::account:role/role_name",
        'destination':{
            's3Destination':{
                's3Uri':"s3://bucket_name/prefix"
            }
        }
    }
)
```

Bereitstellung von Trainingsdaten in AWS Clean Rooms ML

Nachdem der Ersteller der Kollaboration die Kollaboration erstellt hat und eingeladene Mitglieder ihr beigetreten sind, sind Sie bereit, Trainingsdaten zur Kollaboration beizutragen. Jedes Mitglied kann Trainingsdaten beisteuern.

Console

Um Trainingsdaten beizutragen (Konsole)

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Tables (Tabellen) aus.
3. Wählen Sie auf der Seite Tabellen die Option Neue Tabelle konfigurieren aus.
4. Wählen Sie unter Neue Tabelle konfigurieren für Datenquelle Amazon S3, Amazon Athena oder Snowflake aus und führen Sie je nach Datenquelle die folgenden Schritte aus:

Wenn Sie verwenden	Empfehlung
Amazon S3	1. Wählen Sie eine Datenbank aus der Dropdownliste und wählen Sie dann die Tabelle aus der Datenbank aus. 2. Wählen Sie für Spalten, die in Kollaborationen zulässig sind, entweder Alle Spalten oder Benutzerdefinierte Liste aus. 3. Geben Sie für Details zur konfigurierten Tabelle den Namen und optional eine Beschreibung für diese Tabelle an. 4. Wenn Sie Modellmetriken melden möchten, geben Sie den Namen der Metriken und die Regex-Anweisung ein, mit der die Ausgabeprotokolle nach der Metrik durchsucht werden. 5. Wenn Sie Tags für die konfigurierte Tabellenressource aktivieren möchten, wählen Sie Neues Tag hinzufügen aus und geben Sie dann das Schlüssel - und Wertepaar ein.

Wenn Sie verwenden	Empfehlung
Amazon Athena	1. Wählen Sie eine Datenbank aus der Dropdownliste und wählen Sie dann die Tabelle aus der Datenbank aus. 2. Wählen Sie für Spalten, die in Kollaborationen zulässig sind, entweder Alle Spalten oder Benutzerdefinierte Liste aus. 3. Geben Sie für Details zur konfigurierten Tabelle den Namen und optional eine Beschreibung für diese Tabelle an. 4. Wenn Sie Modellmetriken melden möchten, geben Sie den Namen der Metriken und die Regex-Anweisung ein, mit der die Ausgabeprotokolle nach der Metrik durchsucht werden. 5. Wenn Sie Tags für die konfigurierte Tabellenressource aktivieren möchten, wählen Sie Neues Tag hinzufügen aus und geben Sie dann das Schlüssel - und Wertepaar ein.

Wenn Sie verwenden	Empfehlung
Snowflake	1. Geben Sie die Snowflake-Anmeldeinformationen mithilfe eines vorhandenen geheimen ARN an oder speichern Sie ein neues Geheimnis für diese Tabelle. 2. Geben Sie für Snowflake-Tabellen- und Schemadetails die Details manuell ein oder importieren Sie die Details automatisch. 3. Geben Sie für das Schema den Spaltennamen ein und wählen Sie den Datentyp aus der Dropdownliste aus. 4. Wenn Sie Tags für die konfigurierte Tabellenressource aktivieren möchten, wählen Sie Neues Tag hinzufügen und geben Sie dann das Schlüssel - und Wertepaar ein.

5. Wählen Sie Neue Tabelle konfigurieren.
6. Wählen Sie auf der Seite mit den Tabellendetails die Option Analyseregeln konfigurieren aus, um eine benutzerdefinierte Analyseregeln für diese Tabelle zu konfigurieren. Eine benutzerdefinierte Analyseregeln schränkt den Zugriff auf Ihre Daten ein. Sie können entweder eine bestimmte Gruppe von vorab autorisierten Abfragen Ihrer Daten zulassen oder einer bestimmten Gruppe von Konten erlauben, Ihre Daten abzufragen.
 - a. Wählen Sie als Regeltyp für die Analyse die Option Benutzerdefiniert und für Erstellungsmethode die Option Geführter Ablauf aus.
 - b. Wählen Sie Weiter aus.
 - c. Wählen Sie unter „Analysekontrollen angeben“ zwischen „Jede neue Analyse überprüfen“ und „Alle Analysen durch bestimmte Mitarbeiter zulassen“.
 - d. Wählen Sie Weiter aus.
 - e. (Optional) Geben Sie unter „Steuerelemente für Analyseergebnisse angeben“ für Spalten, die in der Ausgabe nicht zulässig sind, an, ob Sie Spalten von der Ausgabe ausschließen möchten. Wenn Sie „Keine“ wählen, werden keine Spalten von der

Ausgabe ausgeschlossen. Wenn Sie Benutzerdefinierte Liste wählen, können Sie bestimmte Spalten angeben, die aus der Ausgabe entfernt werden.

- f. Geben Sie für zusätzliche Analysen, die auf die Ausgabe angewendet werden, an, ob Sie eine zusätzliche Analyse zulassen, verweigern oder erfordern möchten, bevor die Ergebnisse generiert werden.
 - g. Wählen Sie Weiter aus.
 - h. (Optional) Wählen Sie unter Differentiellen Datenschutz festlegen die Option Ausschalten aus.
 - i. Wählen Sie Weiter aus.
 - j. Überprüfen Sie die Informationen auf der Seite Überprüfen und konfigurieren und wählen Sie dann Analyseregeln konfigurieren aus.
7. Wählen Sie auf der Seite mit den Tabellendetails die Option Mit Kollaboration verknüpfen aus.
 8. Wählen Sie im Dialogfeld Tabelle zuordnen die Kollaboration aus, der Sie diese Tabelle zuordnen möchten, und wählen Sie Kollaboration auswählen aus.
 9. Überprüfen und verifizieren Sie auf der Seite Tabelle zuordnen die Informationen in den Bereichen Tabellenzuordnungsdetails, Servicezugriff und Tags. Wählen Sie Tabelle zuordnen aus.
 10. Wählen Sie in der Tabelle „Mit Ihrer Tabelle verknüpfte Tabellen“ das Optionsfeld neben der Tabelle aus, die Sie gerade verknüpft haben. Wählen Sie im Menü Aktionen in der Regelgruppe Kollaborationsanalyse die Option Konfigurieren aus.
 11. Wählen Sie auf der Seite Regel für Kollaborationsanalyse konfigurieren unter Zulässige zusätzliche Analysen aus, ob Mitglieder der Kollaboration oder bestimmte Mitglieder der Kollaboration zusätzliche Analysen durchführen können.

Wählen Sie für die Bereitstellung von Ergebnissen aus, welche Mitglieder Ergebnisse aus Abfrageausgaben erhalten dürfen.

12. Wählen Sie Analyseregeln konfigurieren aus.

API

Um Trainingsdaten beizutragen (API)

1. Konfigurieren Sie eine vorhandene AWS Glue Tabelle für die Verwendung in, AWS Clean Rooms indem Sie die Tabelle und die Spalten angeben, die verwendet werden können.

Führen Sie den folgenden Code mit Ihren spezifischen Parametern aus.

```
import boto3
acr_client= boto3.client('cleanrooms')

acr_client.create_configured_table(
    name='configured_table_name',
    tableReference= {
        'glue': {
            'tableName': 'glue_table_name',
            'databaseName': 'glue_database_name'
        }
    },
    analysisMethod="DIRECT_QUERY",
    allowedColumns=["column1", "column2", "column3",...]
)
```

2. Konfigurieren Sie eine benutzerdefinierte Analyseregeln, die den Zugriff auf Ihre Daten einschränkt. Sie können entweder eine bestimmte Gruppe von vorab autorisierten Abfragen Ihrer Daten zulassen oder einer bestimmten Gruppe von Konten erlauben, Ihre Daten abzufragen.

Führen Sie den folgenden Code mit Ihren spezifischen Parametern aus.

```
import boto3
acr_client= boto3.client('cleanrooms')

acr_client.create_configured_table_analysis_rule(
    configuredTableIdentifier='configured_table_id',
    analysisRuleType='CUSTOM',
    analysisRulePolicy= {
        'v1': {
            'custom': {
                'allowedAnalyses': ['ANY_QUERY'],
                'allowedAnalysisProviders': ['query_runner_account'],
                'additionalAnalyses': "REQUIRED"
            }
        }
    }
)
```

In diesem Beispiel darf ein bestimmtes Konto jede Abfrage der Daten ausführen, und eine zusätzliche Analyse ist erforderlich.

3. Ordnen Sie der Kollaboration eine konfigurierte Tabelle zu und weisen Sie den AWS Glue Tabellen eine Dienstzugriffsrolle zu.

Führen Sie den folgenden Code mit Ihren spezifischen Parametern aus.

```
import boto3
acr_client= boto3.client('cleanrooms')

acr_client.create_configured_table_association(
    name='configured_table_association_name',
    membershipIdentifier='membership_id',
    configuredTableIdentifier='configured_table_id',
    roleArn='arn:aws:iam::account:role/role_name'
)
```

Note

Diese Servicerolle hat Berechtigungen für die Tabellen. Die Servicerolle kann nur übernommen werden AWS Clean Rooms , wenn zulässige Abfragen im Namen des Mitglieds ausgeführt werden, das Abfragen durchführen kann. Keine Kollaborationsmitglieder (außer dem Datenbesitzer) haben Zugriff auf die zugrunde liegenden Tabellen in der Kollaboration. Der Datenbesitzer kann den differenziellen Datenschutz deaktivieren, um seine Tabellen für Abfragen durch andere Mitglieder verfügbar zu machen.

4. Fügen Sie abschließend der konfigurierten Tabellenzuordnung eine Analyseregeln hinzu.

Führen Sie den folgenden Code mit Ihren spezifischen Parametern aus.

```
import boto3
acr_client= boto3.client('cleanrooms')

acr_client.create_configured_table_association_analysis_rule(

    configuredTableAssociationIdentifier='configured_table_association_identifier',
    membershipIdentifier='membership_id',
    configuredTableIdentifier='configured_table_id',
```

```

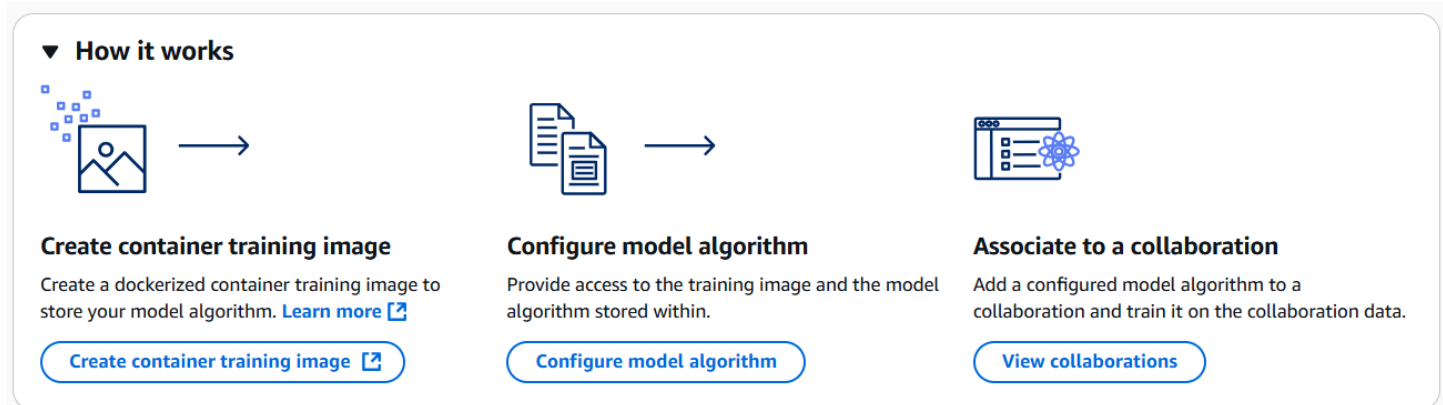
analysisRuleType = 'CUSTOM',
analysisRulePolicy= {
  'v1': {
    'custom': {
      'allowedAdditionalAnalyses':
['configured_model_algorithm_association_arns'],
      'allowedResultReceivers': ['query_runner_account']
    }
  }
}
)

```

Konfiguration eines Modellalgorithmus in AWS Clean Rooms ML

Nachdem Sie [ein Container-Training-Image erstellt](#) haben, müssen Sie Ihren Modellalgorithmus konfigurieren. Durch die Konfiguration eines Modellalgorithmus kann er einer Kollaboration zugeordnet werden.

Die folgende Abbildung zeigt die Konfiguration eines Modellalgorithmus als einen Schritt, der ausgeführt wird, nachdem Sie das Container-Trainingsimage erstellt haben und bevor Sie es der Kollaboration zuordnen.



Console

Um einen benutzerdefinierten ML-Modellalgorithmus zu konfigurieren (Konsole)

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich die Option Benutzerdefinierte ML-Modelle aus.

3. Wählen Sie auf der Seite Benutzerdefinierte ML-Modelle die Option Modellalgorithmus konfigurieren aus.
4. Geben Sie auf der Seite Modellalgorithmus konfigurieren unter Details zum Modellalgorithmus einen Namen und optional eine Beschreibung ein.
5. Wenn Sie ein Modelltraining durchführen möchten, gehen Sie für Informationen zum ECR-Container des Trainingsbilds wie folgt vor:
 - a. Aktivieren Sie das Kontrollkästchen Trainingsbild-URI angeben.
 - b. Wählen Sie in der Dropdownliste das Repository aus, das das Trainingsmodell, den Inferenzcontainer oder beides enthält.
 - c. Wählen Sie das Bild aus.
 - d. (Optional) Geben Sie den Wert für die Einstiegspunkte ein, um auf das Trainingsbild zuzugreifen.
 - e. (Optional) Geben Sie den Wert für die Argumente ein.
6. (Optional) Wenn Sie Modellmetriken melden möchten, geben Sie für Trainingsmetriken den Namen der Metriken und die Regex-Anweisung ein, mit der die Ausgabeprotokolle nach der Metrik durchsucht werden.
7. Wenn Sie eine Modellinferenz durchführen möchten, geben Sie für Inference Image ECR Container-Details Folgendes ein:
 - a. Aktivieren Sie das Kontrollkästchen „URI für Inferenzbild angeben“.
 - b. Wählen Sie das Repository aus der Drop-down-Liste aus.
 - c. Wählen Sie das Bild aus.
8. Wählen Sie für den Dienstzugriff den Namen der vorhandenen Servicerolle aus, der für den Zugriff auf diese Tabelle verwendet werden soll.
9. Wählen Sie für Verschlüsselung die Option Verschlüsselungseinstellungen anpassen aus, um Ihren eigenen KMS-Schlüssel und zugehörige Informationen anzugeben. Andernfalls verwaltet Clean Rooms ML die Verschlüsselung.
10. Wenn Sie Tags aktivieren möchten, wählen Sie Neues Tag hinzufügen und geben Sie dann das Schlüssel - und Wertepaar ein.
11. Wählen Sie Modellalgorithmus konfigurieren aus.

API

Um einen benutzerdefinierten ML-Modellalgorithmus (API) zu konfigurieren

1. Erstellen Sie ein SageMaker KI-kompatibles Docker-Image. Clean Rooms ML unterstützt nur SageMaker KI-kompatible Docker-Images.
2. Nachdem Sie ein SageMaker KI-kompatibles Docker-Image erstellt haben, verwenden Sie Amazon ECR, um ein Trainings-Image zu erstellen. Folgen Sie den Anweisungen im [Amazon Elastic Container Registry User Guide](#), um ein Container-Training-Image zu erstellen.
3. Konfigurieren Sie den Modellalgorithmus für die Verwendung in Clean Rooms ML. Sie müssen die folgenden Informationen angeben:
 - Der Amazon ECR-Repository-Link und zusätzliche Argumente zum Trainieren des Modells und zum Ausführen von Inferenzen. Clean Rooms ML unterstützt die Ausführung von Batch-Transformationsjobs in einem Inferenzcontainer.
 - Eine Dienstzugriffsrolle, die Clean Rooms ML den Zugriff auf das Repository ermöglicht.
 - (Optional) Ein Inferenzcontainer. Sie können dies zwar in einem separaten konfigurierten Modellalgorithmus bereitstellen, wir empfehlen jedoch, ihn in diesem Schritt bereitzustellen, sodass sowohl der Trainings- als auch der Inferenzcontainer als Teil derselben Ressource verwaltet werden.

Führen Sie den folgenden Code mit Ihren spezifischen Parametern aus.

```
import boto3
acr_ml_client= boto3.client('cleanroomsm1')

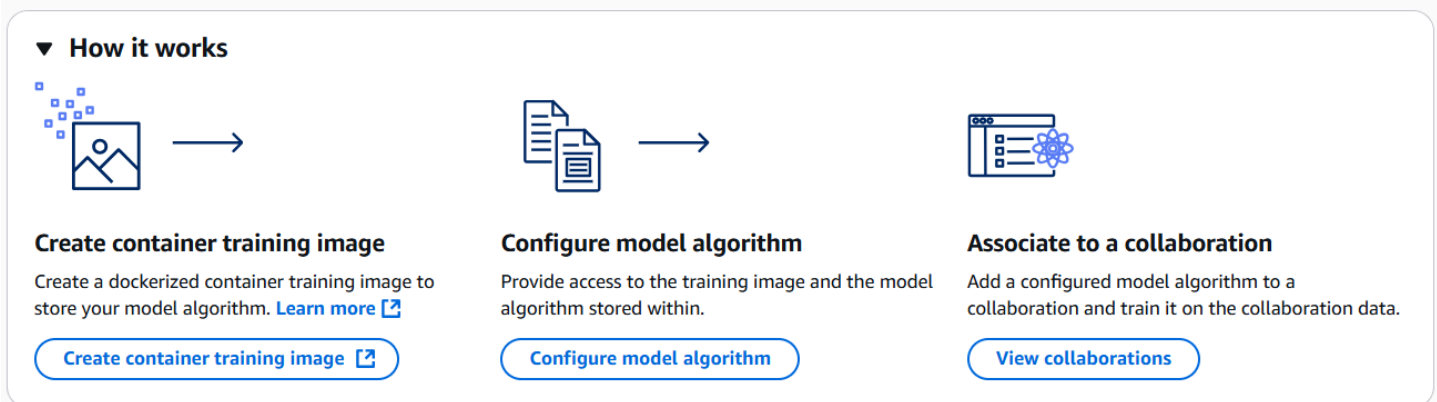
acr_ml_client.create_configured_model_algorithm(
    name='configured_model_algorithm_name',
    trainingContainerConfig={
        'imageUri': 'account.dkr.ecr.region.amazonaws.com/image_name:tag',
        'metricDefinitions': [
            {
                'name': 'custom_metric_name_1',
                'regex': 'custom_metric_regex_1'
            }
        ]
    },
    inferenceContainerConfig={
        'imageUri': 'account.dkr.ecr.region.amazonaws.com/image_name:tag',
    },
    roleArn='arn:aws:iam::account:role/role_name'
```

)

Zuordnen des konfigurierten Modellalgorithmus in AWS Clean Rooms ML

Nachdem Sie den Modellalgorithmus konfiguriert haben, können Sie den Modellalgorithmus einer Kollaboration zuordnen. Durch die Zuordnung eines Modellalgorithmus steht der Modellalgorithmus allen Mitgliedern der Kollaboration zur Verfügung.

Die folgende Abbildung zeigt die Zuordnung des konfigurierten Modellalgorithmus als letzten Schritt nach der Erstellung des Container-Training-Images und der Konfiguration eines Modellalgorithmus.



Console

Note

Nachdem der Modellalgorithmus verknüpft wurde, kann er nicht mehr bearbeitet werden. Um Änderungen vorzunehmen, können Sie den zugehörigen Modellalgorithmus löschen und einen neuen zuordnen.

Um einen benutzerdefinierten ML-Modellalgorithmus zuzuordnen (Konsole)

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich die Option Benutzerdefinierte ML-Modelle aus.

3. Wählen Sie auf der Seite Benutzerdefinierte ML-Modelle den konfigurierten Modellalgorithmus aus, den Sie einer Kollaboration zuordnen möchten, und wählen Sie dann Mit Kollaboration verknüpfen aus.
4. Wählen Sie im Fenster Konfigurierten Modellalgorithmus zuordnen die Kollaboration aus, der Sie eine Verbindung herstellen möchten.
5. Wählen Sie Kollaboration auswählen aus.
6. Geben Sie auf der Seite Modellalgorithmus zuordnen unter Details zur Zuordnung des Modellalgorithmus einen Namen und optional eine Beschreibung ein.
7. Wählen Sie für Modellalgorithmus einen konfigurierten Modellalgorithmus aus.
8. Exportieren Sie Datenschutzkonfigurationen für trainierte Modelle
 - a. Um Modelldateien zu exportieren, aktivieren Sie das Kontrollkästchen Modelldateien.
 - b. Um Ausgabedateien zu exportieren, aktivieren Sie das Kontrollkästchen Ausgabedateien.
 - c. Geben Sie einen Wert für die maximale Größe für die exportierten Daten ein. Der Wert muss zwischen 0,01 und 10 liegen.
9. (Optional) Wenn Sie entweder vollständige Fehlerprotokolle oder kürzere Fehlerzusammenfassungen an die Mitglieder senden möchten, finden Sie unter Konfiguration für den Datenschutz unter Trained Model Inference Job Privacy
 - a. Wählen Sie unter Vollständige Protokolle ein oder mehrere Konten IDs aus der Dropdownliste aus.
 - b. (Optional) Wenn Sie Protokolle senden möchten, die einem Filtermuster entsprechen, geben Sie ein Filtermuster ein.
 - c. (Optional) Wenn Sie ein weiteres Konto und ein optionales Filtermuster hinzufügen möchten, wählen Sie Protokollrichtlinie hinzufügen.
 - d. Wählen Sie unter Fehlerzusammenfassungen ein oder mehrere Konten IDs aus der Dropdownliste aus.
 - e. (Optional) Wählen Sie eine oder mehrere Entitäten aus, die geschwärzt werden sollen, um anzugeben, welche Entitäten aus dem Fehlerprotokoll oder den Fehlerzusammenfassungen gelöscht werden sollen.
 - PII — Personenbezogene Daten unkenntlich machen
 - Zahlen — Zahlen redigieren

- Benutzerdefiniert — auf der Grundlage des benutzerdefinierten Schwärzungsmusters schwärzen
 - i. Wenn Sie im vorherigen Schritt Benutzerdefiniert ausgewählt haben, geben Sie ein Benutzerdefiniertes Schwärzungsmuster ein. Dadurch werden Informationen protokolliert, die diesem Muster entsprechen.
 - ii. (Optional) Wenn Sie ein weiteres benutzerdefiniertes Schwärzungsmuster hinzufügen möchten, wählen Sie Weiteres benutzerdefiniertes Muster hinzufügen.
- 10. (Optional) Wenn Sie Metriken für trainierte Modelle konfigurieren möchten, wählen Sie unter Konfiguration der trainierten Modellmetriken einen Geräuschpegel aus der Dropdownliste aus.

Sie können „Keine“, „Niedrig“, „Mittel“ und „Hoch“ auswählen.
- 11. (Optional) Wenn Sie die maximale Größe der Artefakte festlegen möchten, geben Sie unter Konfiguration der Artefakte den Wert Max. Artefaktgröße ein. Der Wert muss zwischen 0,01 und 10 liegen.
- 12. (Optional) Wenn Sie Tags aktivieren möchten, wählen Sie Neues Tag hinzufügen aus und geben Sie dann das Schlüssel - und Wertepaar ein.
- 13. Wählen Sie Associate aus.

API

Um einen benutzerdefinierten ML-Modellalgorithmus (API) zuzuordnen

Führen Sie den folgenden Code mit Ihren spezifischen Parametern aus.

Sie stellen auch eine Datenschutzrichtlinie bereit, die festlegt, wer Zugriff auf die verschiedenen Protokolle hat, die es Kunden ermöglicht, Regex zu definieren und wie viele Daten aus den Trainingsmodellausgaben oder Inferenzergebnissen exportiert werden können.

Note

Konfigurierte Modellalgorithmuszuordnungen sind unveränderlich.

```
import boto3
acr_ml_client= boto3.client('cleanroomsm1')
```

```

acr_ml_client.create_configured_model_algorithm_association(
    name='configured_model_algorithm_association_name',
    description='purpose of the association',
    configuredModelAlgorithmArn='arn:aws:cleanrooms-ml:region:account:membership/
membershipIdentifier/configured-model-algorithm/identifier',
    privacyConfiguration={
        "policies": {
            "trainedModelExports": {
                "filesToExport": ['files to export'],
                "containerLogs": [
                    {
                        "allowedAccountIds": ['member_account_id'],
                        "filterPattern": ['filter pattern'],
                        "logRedactionConfiguration": {
                            "entitiesToRedact": [
                                'ALL_PERSONALLY_IDENTIFIABLE_INFORMATION',
                                'NUMBERS',
                                'CUSTOM'
                            ],
                            "customEntityConfig": {
                                "customDataIdentifiers": [
                                    'custom_regex_1',
                                    'custom_regex_2'
                                ]
                            }
                        }
                    }
                ]
            },
            "containerMetrics": {
                "noiseLevel": 'noise value'
            },
            "maxArtifactSize": {
                "unit": 'unit',
                "value": 'number'
            }
        },
        "trainedModelInferenceJobs": {
            "containerLogs": [
                {
                    "allowedAccountIds": ['member_account_id'],
                    "filterPattern": ['filter pattern'],
                    "logRedactionConfiguration": {
                        "entitiesToRedact": [

```

```

        'ALL_PERSONALLY_IDENTIFIABLE_INFORMATION',
        'NUMBERS',
        'CUSTOM'
    ],
    "customEntityConfig": {
        "customDataIdentifiers": [
            'custom_regex_1',
            'custom_regex_2'
        ]
    }
}
},
"maxOutputSize": {
    "unit": 'unit',
    "value": 'number'
}
}
},
tags={
    'tag': 'tag'
}
)

```

Nachdem der konfigurierte Modellalgorithmus der Kollaboration zugeordnet wurde, müssen Anbieter von Trainingsdaten ihrer Tabelle eine Regel für die Kollaborationsanalyse hinzufügen. Diese Regel ermöglicht es der konfigurierten Modellalgorithmus-Assoziation, auf ihre konfigurierte Tabelle zuzugreifen. Alle Anbieter von Trainingsdaten, die Beiträge leisten, müssen den folgenden Code ausführen:

```

import boto3
acr_client= boto3.client('cleanrooms')

acr_client.create_configured_table_association_analysis_rule(
    membershipIdentifier= 'membership_id',
    configuredTableAssociationIdentifier= 'configured_table_association_id',
    analysisRuleType= 'CUSTOM',
    analysisRulePolicy = {
        'v1': {
            'custom': {
                'allowedAdditionalAnalyses': ['arn:aws:cleanrooms-
ml:region:*:membership/*/configured-model-algorithm-association/*'],

```

```
        'allowedResultReceivers': []  
    }  
}  
}  
)
```

Note

Da konfigurierte Modellalgorithmuszuordnungen unveränderlich sind, empfehlen wir Anbietern von Trainingsdaten, die Modelle zur Verwendung zulassen möchten, in den `allowedAdditionalAnalyses` ersten Iterationen der benutzerdefinierten Modellkonfiguration Platzhalter zu verwenden. Auf diese Weise können Modellanbieter ihren Code iterieren, ohne dass andere Schulungsanbieter die Zuordnung erneut vornehmen müssen, bevor sie ihren aktualisierten Modellcode mit Daten trainieren.

Erstellen eines ML-Eingabekanals in AWS Clean Rooms ML

Voraussetzungen:

- Und AWS-Konto mit Zugriff auf AWS Clean Rooms
- Eine Zusammenarbeit, AWS Clean Rooms in der Sie den ML-Eingangskanal erstellen möchten
- Berechtigungen zum Abfragen von Daten und zum Erstellen von ML-Eingabekanälen in der Kollaboration.
- (Optional) Ein vorhandener Modellalgorithmus, der dem ML-Eingabekanal zugeordnet werden soll, oder Berechtigungen zum Erstellen eines neuen
- (Optional) Tabellen mit Analyseregeln, die für das angegebene Modell ausgeführt werden können.
- (Optional) Eine vorhandene SQL-Abfrage- oder Analysevorlage, die für die Generierung des Datensatzes verwendet werden soll
- (Optional) Eine vorhandene Servicerolle mit entsprechenden Berechtigungen oder Berechtigungen zum Erstellen einer neuen Servicerolle
- (Optional) Ein benutzerdefinierter AWS KMS Schlüssel, wenn Sie Ihren eigenen Verschlüsselungsschlüssel verwenden möchten
- Entsprechende Berechtigungen zum Erstellen und Verwalten von ML-Modellen in der Zusammenarbeit

Ein ML-Eingabekanal ist ein Datensatz, der anhand einer bestimmten Datenabfrage erstellt wird. Mitglieder mit der Fähigkeit, Daten abzufragen, können ihre Daten für Training und Inferenz vorbereiten, indem sie einen ML-Eingabekanal erstellen. Durch die Erstellung eines ML-Eingangskanals können diese Daten in verschiedenen Trainingsmodellen innerhalb derselben Zusammenarbeit verwendet werden. Sie sollten separate ML-Eingangskanäle für Training und Inferenz erstellen.

Um einen ML-Eingabekanal zu erstellen, müssen Sie die SQL-Abfrage angeben, die zur Abfrage der Eingabedaten verwendet wird, und den ML-Eingabekanal erstellen. Die Ergebnisse dieser Abfrage werden niemals an ein Mitglied weitergegeben und bleiben innerhalb der Grenzen von Clean Rooms ML. Die Referenz Amazon Resource Name (ARN) wird in den nächsten Schritten verwendet, um ein Modell zu trainieren oder Inferenz auszuführen.

Console

Um einen ML-Eingangskanal (Konsole) zu erstellen

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie auf der Seite Collaborations die Kollaboration aus, für die Sie einen ML-Eingabekanal erstellen möchten.
4. Wählen Sie nach dem Öffnen der Kollaboration die Registerkarte ML-Modelle aus.
5. Wählen Sie unter Benutzerdefinierte ML-Modelle im Abschnitt ML-Eingangskanäle die Option ML-Eingangskanal erstellen aus.
6. Gehen Sie auf der Seite „ML-Eingangskanal erstellen“ für Details zum ML-Eingangskanal wie folgt vor:
 - a. Geben Sie unter Name einen eindeutigen Namen für Ihren Kanal ein.
 - b. (Optional) Geben Sie unter Beschreibung eine Beschreibung Ihres Kanals ein.
 - c. Wählen Sie unter Assoziierter Modellalgorithmus den zu verwendenden Algorithmus aus.

Wählen Sie Modellalgorithmus zuordnen aus, um einen neuen Modellalgorithmus hinzuzufügen.

7. Wählen Sie für Datensatz eine Methode zum Generieren des Trainingsdatensatzes aus:

- Wählen Sie SQL-Abfrage, um die Ergebnisse einer SQL-Abfrage als Trainingsdatensatz zu verwenden.

Wenn Sie SQL-Abfrage gewählt haben, geben Sie Ihre Abfrage in das SQL-Abfragefeld ein.

(Optional) Um eine Abfrage zu importieren, die Sie kürzlich verwendet haben, wählen Sie Aus letzten Abfragen importieren aus.

- Wählen Sie Analysevorlage aus, um die Ergebnisse einer Analysevorlage als Trainingsdatensatz zu verwenden.

 Warning

Die Generierung synthetischer Daten schützt vor dem Ableiten einzelner Attribute, unabhängig davon, ob bestimmte Personen im ursprünglichen Datensatz vorhanden sind oder Lernattribute dieser Personen vorhanden sind. Es verhindert jedoch nicht, dass wörtliche Werte aus dem ursprünglichen Datensatz, einschließlich persönlich identifizierbarer Informationen (PII), im synthetischen Datensatz erscheinen.

Wir empfehlen, Werte im Eingabedatensatz zu vermeiden, die nur einer betroffenen Person zugeordnet sind, da diese eine betroffene Person neu identifizieren können. Wenn beispielsweise nur ein Benutzer in einer Postleitzahl wohnt, würde das Vorhandensein dieser Postleitzahl im synthetischen Datensatz bestätigen, dass sich der Benutzer im ursprünglichen Datensatz befand. Techniken wie das Kürzen hochgenauer Werte oder das Ersetzen ungewöhnlicher Kataloge durch andere können verwendet werden, um dieses Risiko zu minimieren. Diese Transformationen können Teil der Abfrage sein, die zur Erstellung des ML-Eingabekanals verwendet wird.

- a. Wenn keine Tabellen verknüpft sind, wählen Sie Tabelle zuordnen aus, um Tabellen mit einer Analyseregeln hinzuzufügen, die für das angegebene Modell ausgeführt werden kann.
- b. Wählen Sie den Worker-Typ aus, der bei der Erstellung dieses Datenkanals verwendet werden soll. Der Standard-Worker-Typ ist CR.1X. Geben Sie die Anzahl der zu

verwendenden Worker an. Die Standardnummer der Arbeiter ist 16. So geben Sie Spark-Eigenschaften an:

- i. Erweitern Sie Spark-Eigenschaften.
- ii. Wählen Sie Spark-Eigenschaften hinzufügen.
- iii. Wählen Sie im Dialogfeld „Spark-Eigenschaften“ einen Eigenschaftsnamen aus der Dropdownliste aus und geben Sie einen Wert ein.

Die folgende Tabelle enthält eine Definition für jede Eigenschaft.

Weitere Informationen zu Spark-Eigenschaften finden Sie unter [Spark-Eigenschaften](#) in der Apache Spark-Dokumentation.

Eigenschaftename	Description	Standardwert
spark.task.maxFailures	Steuert, wie oft eine Aufgabe hintereinander fehlschlagen kann, bevor der Job fehlschlägt. Erfordert einen Wert größer oder gleich 1. Die Anzahl der zulässigen Wiederholungen entspricht diesem Wert minus 1. Die Anzahl der Fehlschläge wird zurückgesetzt, wenn ein Versuch erfolgreich ist. Fehler bei verschiedenen Aufgaben summieren sich nicht bis zu diesem Limit.	4
spark.sql.files.maxPartitionBytes	Legt die maximale Anzahl von Bytes fest, die beim Lesen aus dateibasierten Quellen wie Parquet, JSON und ORC in eine einzelne Partition gepackt werden sollen.	128 MB
spark.hadoop.fs.s3.maxVersuch es erneut	Legt die maximale Anzahl von Wiederholungsversuchen für Amazon S3 S3-Dateioperationen fest.	

Eigenschaftename	Description	Standardwert
spark.network.timeout	Legt das Standard-Timeout für alle Netzwerkinteraktionen fest. Setzt die folgenden Timeout-Einstellungen außer Kraft, wenn sie nicht konfiguriert sind: • spark.storage.blockManagerHeartbeatTimeoutMs • spark.shuffle.io.ConnectionTimeout • spark.rpc.AskTimeout • spark.rpc.LookupTimeout	120 s
spark.rdd.compress	Gibt an, ob serialisierte RDD-Partitionen mit spark.io.compression.codec komprimiert werden sollen. Gilt für StorageLevel .MEMORY_ONLY_SER in Java und Scala oder .MEMORY_ONLY in Python. StorageLevel Reduziert den Speicherplatz, erfordert jedoch zusätzliche CPU-Verarbeitungszeit.	FALSE
spark.shuffle.spill.compress	Gibt an, ob Shuffle-Spill-Daten mit spark.io.compression.codec komprimiert werden sollen.	TRUE
spark.sql.adaptive.advisoryPartitionSizeInBytes	Legt die Zielgröße in Byte für Shuffle-Partitionen während der adaptiven Optimierung fest, wenn spark.sql.adaptive.enabled den Wert true hat. Steuert die Partitionsgröße beim Zusammenführen kleiner Partitionen oder beim Teilen schiefer Partitionen.	(Wert von spark.sql.adaptive.shuffle.targetPostShuffleInputSize)

Eigenschaftename	Description	Standardwert
spark.sql.adaptive.autoBroadcastJoinSchwellenwert	Legt die maximale Tabellengröße in Byte für die Übertragung an Worker-Knoten bei Joins fest. Gilt nur im adaptiven Framework. Verwendet denselben Standardwert wie spark.sql.adaptive.autoBroadcastJoinSchwellenwert. Auf -1 setzen, um die Übertragung zu deaktivieren.	(Keine)
spark.sql.adaptive.CoalescePartitions.Enabled	Gibt an, ob zusammenhängende Shuffle-Partitionen, die auf spark.sql.adaptive basieren, zusammengeführt werden sollen. advisoryPartitionSizeInBytes um die Aufgabengröße zu optimieren. Erfordert, dass spark.sql.adaptive.enabled den Wert true hat.	TRUE
spark.sql.adaptive.CoalescePartitions.initialPartitionNum	Definiert die anfängliche Anzahl von Shuffle-Partitionen vor dem Zusammenführen. Erfordert, dass sowohl spark.sql.adaptive.enabled als auch spark.sql.adaptive.CoalescePartitions.Enabled den Wert true haben. Der Standardwert ist der Wert von spark.sql.shuffle.partitions.	(Keine)
spark.sql.adaptive.CoalescePartitions.minPartitionSize	Legt die Mindestgröße für zusammengeführte Shuffle-Partitionen fest, um zu verhindern, dass Partitionen während der adaptiven Optimierung zu klein werden.	1 MB

Eigenschaftename	Description	Standardwert
spark.sql.adaptive. CoalescePartitions.ParallelismFirst	Gibt an, ob Partitionsgrößen auf der Grundlage der Cluster-Parallelität und nicht auf der Grundlage von spark.sql.adaptive berechnet werden sollen. advisoryPartitionSizeInBytes während der Zusammenführung von Partitionen. Generiert kleinere Partitionsgrößen als die konfigurierte Zielgröße, um die Parallelität zu maximieren. Wir empfehlen, diesen Wert bei stark frequentierten Clustern auf „False“ zu setzen, um die Ressourcennutzung zu verbessern und übermäßig kleine Aufgaben zu vermeiden.	TRUE
spark.sql.adaptive. .enabled	Gibt an, ob die adaptive Abfrageausführung aktiviert werden soll, um Abfragepläne während der Abfrageausführung auf der Grundlage genauer Laufzeitstatistiken erneut zu optimieren.	TRUE
spark.sql.adaptive. .forceOptimizeSkewedBeitreten	Gibt an, ob die Aktivierung erzwungen werden soll, OptimizeSkewedJoin auch wenn dadurch zusätzlicher Shuffle eingeführt wird.	FALSE
spark.sql.adaptive. localShuffleReader. . aktiviert	Gibt an, ob lokale Shuffle-Reader verwendet werden sollen, wenn eine Shuffle-Partitionierung nicht erforderlich ist, z. B. nach der Konvertierung von Sort-Merge-Joins in Broadcast-Hash-Joins. Erfordert, dass spark.sql.adaptive.enabled den Wert true hat.	TRUE

Eigenschaftename	Description	Standardwert
<code>spark.sql.adaptive.maxShuffledHashJoinLocalMapThreshold</code>	Legt die maximale Partitionsgröße in Byte für die Erstellung lokaler Hash-Maps fest. Priorisiert gemischte Hash-Joins gegenüber Sort-Merge-Joins, wenn: - Dieser Wert entspricht oder übersteigt <code>spark.sql.adaptive.advisoryPartitionSizeInBytes</code> - Alle Partitionsgrößen liegen innerhalb dieser Grenze Überschreibt <code>spark.sql.join.preferSortMergeJoin</code>-Einstellung.	0 Byte
<code>spark.sql.adaptive.optimizeShufflePartitionsInRebalancePartitions</code> aktiviert	Gibt an, ob schiefe Shuffle-Partitionen optimiert werden sollen, indem sie auf der Grundlage von <code>spark.sql.adaptive.advisoryPartitionSizeInBytes</code> in kleinere Partitionen aufgeteilt werden. Erfordert, dass <code>spark.sql.adaptive.enabled</code> wahr ist.	TRUE
<code>spark.sql.adaptive.rebalancePartitionsSmallPartitionFactor</code>	Definiert den Größenschwellenwert für das Zusammenführen von Partitionen beim Teilen. Partitionen, die kleiner sind als dieser Faktor multipliziert mit <code>spark.sql.adaptive.advisoryPartitionSizeInBytes</code> werden zusammengeführt.	0.2

Eigenschaftename	Description	Standardwert
spark.sql.Adaptive.SkewJoin.Enabled	Gibt an, ob Datenverzerrungen in gemischten Verknüpfungen behandelt werden sollen, indem schiefe Partitionen aufgeteilt und optional repliziert werden. Gilt für Sort-Merge- und Shuffled-Hash-Joins. Erfordert, dass spark.sql.adaptive.enabled wahr ist.	TRUE
spark.sql.adaptive.SkewJoin.skewedPartitionFactor	Bestimmt den Größenfaktor, der die Partitionsneigung bestimmt. Eine Partition ist schief, wenn ihre Größe beide Werte überschreitet: - Dieser Faktor wird mit der mittleren Partitionsgröße multipliziert - Der Wert von spark.sql.adaptive.SkewJoin.skewedPartitionThresholdInBytes	5
spark.sql.Adaptive.SkewJoin.skewedPartitionThresholdInBytes	Legt den Größenschwellenwert in Byte zur Identifizierung schiefer Partitionen fest. Eine Partition ist schief, wenn ihre Größe beide Werte überschreitet: - Dieser Schwellenwert - Die mittlere Partitionsgröße multipliziert mit spark.sql.adaptive.SkewJoin.skewedPartitionFactor Wir empfehlen, diesen Wert größer als spark.sql.adaptive festzulegen. advisoryPartitionSizeInBytes.	256 MB

Eigenschaftename	Description	Standardwert
spark.sql.autoBroadcastJoinSwitch	Legt die maximale Tabellengröße in Byte für die Übertragung an Worker-Knoten bei Joins fest. Auf -1 setzen, um die Übertragung zu deaktivieren.	10 MB
spark.sql.BroadcastTimeout	Steuert den Timeout-Zeitraum in Sekunden für die Broadcast-Operationen bei Broadcast-Joins.	300 Sekunden
spark.sql.cbo.enabled	Gibt an, ob die kostenbasierte Optimierung (CBO) für die Schätzung von Planstatistiken aktiviert werden soll.	FALSE
spark.sql.CBO.JoinReorder.DP.Star.Filter	Gibt an, ob bei der kostenbasierten Join-Aufzählung Heuristiken des Star-Join-Filters angewendet werden sollen.	FALSE
spark.sql.CBO.JoinReorder.DP.Threshold	Legt die maximale Anzahl verbundener Knoten fest, die im dynamischen Programmieralgorithmus zulässig sind.	12
spark.sql.CBO.JoinReorder.Enabled	Gibt an, ob die Neuordnung von Verknüpfungen bei der kostenbasierten Optimierung (CBO) aktiviert werden soll.	FALSE
spark.sql.CBO.PlanStats.Enabled	Gibt an, ob bei der Generierung logischer Pläne Zeilenanzahl und Spaltenstatistiken aus dem Katalog abgerufen werden sollen.	FALSE
spark.sql.cbo.starSchemaDetection	Gibt an, ob die Neuordnung von Verknüpfungen auf der Grundlage der Star-Schemaerkennung aktiviert werden soll.	FALSE

Eigenschaftename	Description	Standardwert
spark.sql.files.maxPartitionNum	Legt die maximale Zielanzahl von geteilten Dateipartitionen für dateibasierte Quellen (Parquet, JSON und ORC) fest. Skaliert Partitionen neu, wenn die anfängliche Anzahl diesen Wert überschreitet. Dies ist ein empfohlenes Ziel, kein garantiertes Limit.	(Keine)
spark.sql.files.maxRecordsPerDatei	Legt die maximale Anzahl von Datensätzen fest, die in eine einzelne Datei geschrieben werden sollen. Wenn der Wert auf Null oder einen negativen Wert gesetzt ist, gilt kein Limit.	0
spark.sql.files.minPartitionNum	Legt die Mindestanzahl von geteilten Dateipartitionen für dateibasierte Quellen (Parquet, JSON und ORC) fest. Die Standardeinstellung ist spark.sql.leafNodeDefaultParallelität. Dies ist ein empfohlenes Ziel, kein garantiertes Limit.	(Keine)
spark.sql.inMemoryColumnarSpeicher.Batchgröße	Steuert die Batchgröße für das spaltenförmige Caching. Eine Erhöhung der Größe verbessert die Speichernutzung und Komprimierung, erhöht jedoch das Fehlerrisiko out-of-memory.	10000
spark.sql.inMemoryColumnarSpeicher.Komprimiert	Gibt an, ob auf der Grundlage von Datenstatistiken automatisch Komprimierungscodecs für Spalten ausgewählt werden sollen.	TRUE

Eigenschaftennamenname	Description	Standardwert
spark.sql.inMemoryColumnarStorage.enableVectorizedReader	Gibt an, ob vektorisiertes Lesen für das spaltenförmige Caching aktiviert werden soll.	TRUE
spark.sql.legacy.allowHashOnMapType	Gibt an, ob Hash-Operationen für Map-Datenstrukturen zulässig sind. Diese Legacy-Einstellung gewährleistet die Kompatibilität mit der Map-Typ-Behandlung älterer Spark-Versionen.	
spark.sql.legacy.allowNegativeScaleOfDecimal	Gibt an, ob negative Skalenwerte in Dezimaltypdefinitionen zulässig sind. Diese Legacy-Einstellung gewährleistet die Kompatibilität mit älteren Spark-Versionen, die negative Dezimalskalen unterstützten.	
spark.sql.legacy.castComplexTypesToString.aktiviert	Gibt an, ob veraltetes Verhalten für die Umwandlung komplexer Typen in Zeichenketten aktiviert werden soll. Behält die Kompatibilität mit den Typkonvertierungsregeln älterer Spark-Versionen bei.	
spark.sql.legacy.charVarcharAsString	Gibt an, ob die Typen CHAR und VARCHAR als STRING-Typen behandelt werden sollen. Diese Legacy-Einstellung bietet Kompatibilität mit der Verarbeitung von String-Typen in älteren Spark-Versionen.	

Eigenschaftename	Description	Standardwert
<code>spark.sql.legacy.createEmptyCollectionUsingStringType</code>	Gibt an, ob leere Sammlungen mithilfe von Zeichenkettenelementen erstellt werden sollen. Diese Legacy-Einstellung gewährleistet die Kompatibilität mit dem Verhalten älterer Spark-Versionen bei der Initialisierung von Sammlungen.	
<code>spark.sql.legacy.exponentLiteralAsDezimal</code> . Aktiviert	Gibt an, ob exponentielle Literale als Dezimaltypen interpretiert werden sollen. Diese Legacy-Einstellung gewährleistet die Kompatibilität mit der numerischen Literalverarbeitung älterer Spark-Versionen.	
<code>spark.sql.legacy.json.allowEmptyString</code> . aktiviert	Gibt an, ob leere Zeichenfolgen bei der JSON-Verarbeitung zulässig sind. Diese Legacy-Einstellung gewährleistet die Kompatibilität mit dem JSON-Parsing-Verhalten älterer Spark-Versionen.	
<code>spark.sql.legacy.parquet.int96RebaseModelRead</code>	Gibt an, ob beim Lesen von Parquet-Dateien der alte Timestamp-Rebase-Modus verwendet werden soll. INT96 Diese Legacy-Einstellung gewährleistet die Kompatibilität mit der Zeitstempelverarbeitung älterer Spark-Versionen.	
<code>spark.sql.legacy.timeParserPolicy</code>	Steuert das Zeitanalyseverhalten aus Gründen der Abwärtskompatibilität. Diese ältere Einstellung bestimmt, wie Zeitstempel und Datumsangaben anhand von Zeichenketten analysiert werden.	

Eigenschaftename	Description	Standardwert
spark.sql.Legacy.TypeCoercion.datetimeToString.aktiviert	Gibt an, ob bei der Konvertierung von Datetime-Werten in Zeichenfolgen das Zwangsverhalten älterer Typen aktiviert werden soll. Behält die Kompatibilität mit den Datetime-Konvertierungsregeln älterer Spark-Versionen bei.	
spark.sql.maxSinglePartitionByte	Legt die maximale Partitionsgröße in Byte fest. Der Planer führt Shuffle-Operationen für größere Partitionen ein, um die Parallelität zu verbessern.	128 m
spark.sql.MetadatencacheTTLSeconds	Steuert die (TTL) für time-to-live Metadaten-Caches. Gilt für Partition, sdatei-Metadaten und Sitzungskatalog-Caches. Erfordert: • Ein positiver Wert größer als Null • spark.sql.CatalogImplementation ist auf Hive gesetzt • spark.sql.hive.filesourcePartitionFileCacheSize größer als Null • spark.sql.hive.manageFilesourcePartitions auf true gesetzt	-1000 ms
spark.sql.optimizer.collapseProjectAlwaysInline	Gibt an, ob benachbarte Projektionen und Inline-Ausdrücke ausgeblendet werden sollen, auch wenn dies zu Duplikaten führt.	FALSE
spark.sql.optimizer.dynamicPartitionPruning.aktiviert	Gibt an, ob Prädikate für Partitionsspalten generiert werden sollen, die als Join-Schlüssel verwendet werden.	TRUE

Eigenschaftename	Description	Standardwert
spark.sql.optimizer.enableCsvExpressionOptimierung	Gibt an, ob CSV-Ausdrücke im SQL-Optimizer optimiert werden sollen, indem unnötige Spalten aus from_csv-Vorgängen entfernt werden.	TRUE
spark.sql.optimizer.enableJsonExpressionOptimierung	Gibt an, ob JSON-Ausdrücke im SQL-Optimizer wie folgt optimiert werden sollen: • Löschen unnötiger Spalten aus from_json-Vorgängen • Vereinfachung der Kombinationen from_json und to_json • Optimierung von named_struct-Operationen	TRUE
spark.sql.Optimizer.ExcludedRules	Definiert zu deaktivierende Optimizer-Regeln, identifiziert durch kommagetrennte Regelnamen. Einige Regeln können nicht deaktiviert werden, da sie aus Gründen der Richtigkeit erforderlich sind. Der Optimizer protokolliert, welche Regeln erfolgreich deaktiviert wurden.	(Keine)
spark.sql.optimizer.runtime.BloomFilter.applicationSideScanSizeThreshold	Legt die minimale aggregierte Scangröße in Byte fest, die erforderlich ist, um einen Bloom-Filter auf der Anwendungsseite einzufügen.	10 GB
spark.sql.Optimizer.Runtime.BloomFilter.creationSizeThreshold	Definiert den maximalen Größenschwellenwert für die Injektion eines Bloom-Filters auf der Erstellungsseite.	10 MB

Eigenschaftename	Description	Standardwert
<code>spark.sql.Optimizer.Runtime.BloomFilter.Enabled</code>	Gibt an, ob ein Bloom-Filter eingefügt werden soll, um Shuffle-Daten zu reduzieren, wenn eine Seite einer Shuffle-Verknüpfung über ein selektives Prädikat verfügt.	TRUE
<code>spark.sql.optimize r.runtime.BloomFilter. expectedNumItems</code>	Definiert die Standardanzahl erwarteter Elemente im Runtime-Bloom-Filter.	1000000
<code>spark.sql.optimize r.runtime.BloomFilter. maxNumBits</code>	Legt die maximale Anzahl von Bits fest, die im Runtime-Bloom-Filter zulässig sind.	67108864
<code>spark.sql.optimize r.runtime.BloomFilter. maxNumItems</code>	Legt die maximale Anzahl erwarteter Elemente fest, die im Runtime-Bloom-Filter zulässig sind.	4000000
<code>spark.sql.optimize r.runtime.BloomFilter. NumberThreshold</code>	Schränkt die maximale Anzahl von Nicht-DPP-Laufzeitfiltern pro Abfrage ein, um Fehler im Treiber zu verhindern. out-of-memory	10
<code>spark.sql.optimize r.runtime.BloomFilter. NumBits</code>	Definiert die Standardanzahl von Bits, die im Runtime-Bloom-Filter verwendet werden.	8388608

Eigenschaftename	Description	Standardwert
spark.sql.optimizer.runtime.rowlevelOperationGroupFilter. Aktiviert	Gibt an, ob die Laufzeitgruppenfilterung für Operationen auf Zeilenebene aktiviert werden soll. Ermöglicht Datenquellen: • Löschen ganzer Datengruppen (wie Dateien oder Partitionen) mithilfe von Datenquellenfiltern • Führen Sie Laufzeitabfragen aus, um übereinstimmende Datensätze zu identifizieren • Verwerfen Sie unnötige Gruppen, um teure Neuschreibungen zu vermeiden Einschränkungen: • Nicht alle Ausdrücke können in Datenquellenfilter konvertiert werden • Einige Ausdrücke erfordern eine Spark-Auswertung (z. B. Unterabfragen)	TRUE
spark.sql.Optimizer.RuntimeFilter.semiJoinReduction. aktiviert	Gibt an, ob ein Semi-Join eingefügt werden soll, um die Anzahl der Shuffle-Daten zu reduzieren, wenn eine Seite einer Shuffle-Verknüpfung über ein selektives Prädikat verfügt.	FALSE

Eigenschaftename	Description	Standardwert
spark.sql.parquet. Aggregate PushDown	Gibt an, ob Aggregate zur Optimierung nach Parquet übertragen werden sollen. Unterstützt: • MIN und MAX für die Typen Boolean, Integer, Float und Date • COUNT für alle Datentypen Löst eine Ausnahme aus, wenn Statistiken in einer Fußzeile einer Parquet-Datei fehlen.	FALSE
spark.sql.parquet. columnarReaderBatchGröße	Steuert die Anzahl der Zeilen in jedem Stapel von vektorisierten Parquet-Readern. Wählen Sie einen Wert, der Leistungsaufwand und Speichernutzung ausbalanciert, um Fehler zu vermeiden out-of-memory.	4096
spark.sql.session. TimeZone	Definiert die Sitzungszeitzone für die Verarbeitung von Zeitstempeln in Zeichenfolgenliteralen und die Konvertierung von Java-Objekten. Akzeptiert: • Regionsbasiertes IDs area/city Format (z. B. America/Los_Angeles) • Zonenversätze im Format (+/-) HH, (+/-) HH:mm oder (+/-) HH:mm:ss (z. B. -08 oder + 01:00) • UTC oder Z als Aliase für + 00:00	(Wert der lokalen Zeitzone)

Eigenschaftename	Description	Standardwert
spark.sql.shuffle.partitions	Legt die Standardanzahl von Partitionen für das Mischen von Daten bei Verknüpfungen oder Aggregationen fest. Kann zwischen Neustarts strukturierter Streaming-Abfragen von derselben Checkpoint-Position aus nicht geändert werden.	200
spark.sql.shuffleJoinFactor	Definiert den Multiplikationsfaktor, der verwendet wird, um die Eignung für einen Shuffle-Hash-Join zu bestimmen. Ein Shuffle-Hash-Join wird ausgewählt, wenn die Datengröße der kleinen Seite multipliziert mit diesem Faktor kleiner als die Datengröße der großen Seite ist.	3
spark.sql.sources.parallelPartitionDiscovery.Schwellenwert	Legt die maximale Anzahl von Pfaden für die treiberseitige Dateiauflistung mit dateibasierten Quellen (Parquet, JSON und ORC) fest. Wenn sie bei der Partitionserkennung überschritten werden, werden Dateien mithilfe eines separaten verteilten Spark-Jobs aufgelistet.	32
spark.sql.statistics.histogram.enabled	Gibt an, ob bei der Berechnung der Spaltenstatistiken Histogramme mit gleicher Höhe generiert werden sollen, um die Schätzgenauigkeit zu verbessern. Erfordert einen zusätzlichen Tabellenscan, der über den für einfache Spaltenstatistiken erforderlichen hinausgeht.	FALSE

- c. Geben Sie für Datenspeicherung in Tagen die Anzahl der Tage ein, für die die Daten aufbewahrt werden sollen.

- d. Wählen Sie als Ergebnisformat entweder CSV oder Parquet als Datenformat, das der ML-Eingangskanal verwenden soll.
8. Wählen Sie für Dienstzugriff den Namen der vorhandenen Servicerolle aus, der für den Zugriff auf diese Tabelle verwendet werden soll, oder wählen Sie Neue Servicerolle erstellen und verwenden aus.
9. Wählen Sie für Verschlüsselung die Option Geheimnis mit einem benutzerdefinierten KMS-Schlüssel verschlüsseln aus, um Ihren eigenen KMS-Schlüssel und zugehörige Informationen anzugeben. Andernfalls verwaltet Clean Rooms ML die Verschlüsselung.
10. Wählen Sie „ML-Eingangskanal erstellen“.

Die Erstellung des ML-Eingangskanals dauert einige Minuten. Auf der Registerkarte ML-Modelle finden Sie eine Liste der ML-Eingangskanäle.

Note

Nachdem der ML-Eingangskanal erstellt wurde, können Sie ihn nicht mehr bearbeiten.

API

Um einen ML-Eingangskanal (API) zu erstellen

Führen Sie den folgenden Code mit Ihren spezifischen Parametern aus:

```
import boto3
acr_client = boto3.client('cleanroomsmml')

acr_client.create_ml_input_channel(
    name="ml_input_channel_name",
    membershipIdentifier='membership_id',

    configuredModelAlgorithmAssociations=[configured_model_algorithm_association_arn],
    retentionInDays=1,
    inputChannel={
        "dataSource": {
            "protectedQueryInputParameters": {
                "sqlParameters": {
                    "queryString": "select * from table",
                    "computeConfiguration": {
```

```

        "worker": {
            "type": "CR.1X",
            "number": 16,
            "properties": {
                "spark": {
                    "spark configuration key": "spark configuration
value",
                }
            }
        },
        "resultFormat": "PARQUET"
    },
    "roleArn": "arn:aws:iam::111122223333:role/role_name"
}
)
channel_arn = resp['ML Input Channel ARN']

```

Erstellen eines trainierten Modells in AWS Clean Rooms ML

Voraussetzungen:

- Und AWS-Konto mit Zugriff auf AWS Clean Rooms
- Eine Zusammenarbeit wurde eingerichtet in AWS Clean Rooms
- Ein konfigurierter Modellalgorithmus, der mit der Zusammenarbeit verknüpft ist
- Mindestens ein konfigurierter ML-Eingangskanal
- Entsprechende Berechtigungen zum Erstellen und Verwalten von ML-Modellen in der Zusammenarbeit

Nachdem Sie den konfigurierten Modellalgorithmus einer Kollaboration zugeordnet und anschließend einen ML-Eingangskanal erstellt und konfiguriert haben, können Sie ein trainiertes Modell erstellen. Ein trainiertes Modell wird von Mitgliedern einer Kollaboration verwendet, um ihre Daten gemeinsam zu analysieren.

Mit dem folgenden Verfahren können Sie ein trainiertes Modell erstellen.

Alternativ können Sie inkrementelles Training verwenden, um ein vorhandenes Modell mit neuen Daten zu verbessern, oder verteiltes Training, um Modelle über mehrere Recheninstanzen hinweg zu trainieren.

Themen

- [Nutzung von inkrementellen Schulungen in AWS Clean Rooms ML](#)
- [Nutzung verteilter Schulungen in AWS Clean Rooms ML](#)

Console

Um ein trainiertes Modell zu erstellen (Konsole)

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie auf der Seite Kollaborationen die Kollaboration aus, für die Sie ein trainiertes Modell erstellen möchten.
4. Wählen Sie nach dem Öffnen der Kollaboration die Registerkarte ML-Modelle aus.
5. Wählen Sie unter Benutzerdefinierte ML-Modelle im Abschnitt Trainierte Modelle die Option Trainiertes Modell erstellen aus.
6. Geben Sie auf der Seite Trainiertes Modell erstellen für Algorithmus für Zugeordnetes Modell den Algorithmus an.
7. Geben Sie für Details zum trainierten Modell Folgendes ein:
 - a. Geben Sie unter Name einen eindeutigen Namen für das Modell in der Kollaboration ein.
 - b. (Optional) Geben Sie unter Beschreibung eine Beschreibung des trainierten Modells ein.
 - c. Wählen Sie für den Eingabemodus Trainingsdaten eine der folgenden Optionen:
 - Wählen Sie Datei aus, wenn Sie über einen kleineren Datensatz verfügen, der auf das ML-Speichervolumen passt, und Sie für Ihr Trainingsskript den herkömmlichen Dateisystemzugriff bevorzugen.
 - Wählen Sie Pipe für große Datensätze, um Daten direkt aus S3 zu streamen, sodass Sie nicht alles auf die Festplatte herunterladen müssen, was die Trainingsgeschwindigkeit verbessern und den Speicherbedarf reduzieren kann.

- Wählen Sie aus, FastFileob Sie die Vorteile des Streamings von S3 mit dem Dateisystemzugriff kombinieren möchten, insbesondere für sequentiell gelesene Daten oder wenn Sie mit weniger Dateien arbeiten müssen, um die Startzeiten zu verkürzen.

8. Gehen Sie wie folgt vor, um Einzelheiten zum ML-Eingangskanal zu erhalten:

- a. Geben Sie für den ML-Eingangskanal den ML-Eingangskanal an, der Daten für den Modellalgorithmus bereitstellt.

Um einen weiteren Kanal hinzuzufügen, wählen Sie „Weiteren ML-Eingangskanal hinzufügen“. Sie können bis zu 19 zusätzliche ML-Eingangskanäle hinzufügen.

- b. Geben Sie als Kanalname den Namen des ML-Eingangskanals ein.
- c. Wählen Sie für den Amazon S3 S3-Datenverteilungstyp eine der folgenden Optionen:
 - Wählen Sie Vollständig repliziert aus, um jeder Trainingsinstanz eine vollständige Kopie Ihres Datensatzes zu geben. Dies funktioniert am besten, wenn Ihr Datensatz klein genug ist, um in den Arbeitsspeicher zu passen, oder wenn jede Instanz Zugriff auf alle Daten benötigt.
 - Wählen Sie Sharded by S3 Key aus, um Ihren Datensatz anhand von S3-Schlüsseln auf Trainingsinstanzen aufzuteilen. Jede Instanz empfängt etwa $1/n$ der gesamten S3-Objekte, wobei 'n' für die Anzahl der Instanzen steht. Dies funktioniert am besten für große Datensätze, die Sie parallel verarbeiten möchten.

 Note

Berücksichtigen Sie bei der Auswahl eines Verteilungstyps Ihre Datensatzgröße und Ihre Trainingsanforderungen. Vollständig repliziert bietet vollständigen Datenzugriff, benötigt jedoch mehr Speicherplatz, während Sharded by S3 Key die verteilte Verarbeitung großer Datensätze ermöglicht.

9. Wählen Sie für Maximale Trainingsdauer die maximale Dauer aus, für die Sie Ihr Modell trainieren möchten.
10. Geben Sie für Hyperparameter alle algorithmusspezifischen Parameter und ihre beabsichtigten Werte an. Hyperparameter sind spezifisch für das trainierte Modell und werden zur Feinabstimmung des Modelltrainings verwendet.
11. Geben Sie für Umgebungsvariablen alle algorithmusspezifischen Variablen und ihre beabsichtigten Werte an. Umgebungsvariablen werden im Docker-Container festgelegt.

12. Um eine benutzerdefinierte Verschlüsselung zu verwenden AWS KMS key, aktivieren Sie das Kontrollkästchen Geheimnis mit einem benutzerdefinierten KMS-Schlüssel verschlüsseln.
13. Geben Sie für die EC2-Ressourcenkonfiguration Informationen zu den Rechenressourcen an, die für das Modelltraining verwendet werden.
 - a. Wählen Sie unter Instanztyp den Instanztyp aus, den Sie ausführen möchten.
 - b. Geben Sie unter Anzahl der Instanzen die Anzahl der Instanzen ein.
 - c. Geben Sie für Volumengröße in GB die Größe des ML-Speichervolumens ein.
14. Wählen Sie Trainiertes Modell erstellen aus.

API

Um ein trainiertes Modell (API) zu erstellen

Das Mitglied, das ein Modell trainieren kann, beginnt mit dem Training, indem es den ML-Eingangskanal und den Modellalgorithmus auswählt.

Führen Sie den folgenden Code mit Ihren spezifischen Parametern aus:

```
import boto3
acr_ml_client= boto3.client('cleanroomsm1')

acr_ml_client.create_trained_model(
    membershipIdentifier= 'membership_id',
    configuredModelAlgorithmAssociationArn = 'arn:aws:cleanrooms-
ml:region:account:membership/membershipIdentifier/configured-model-algorithm-
association/identifier',
    name='trained_model_name',
    resourceConfig={
        'instanceType': "ml.m5.xlarge",
        'volumeSizeInGB': 1
    },
    dataChannels=[
        {
            "mlInputChannelArn": channel_arn_1,
            "channelName": "channel_name"
        },
        {
            "mlInputChannelArn": channel_arn_2,
            "channelName": "channel_name"
        }
    ]
)
```

```
}  
]  
)
```

Note

Nachdem das trainierte Modell erstellt wurde, können Sie es nicht mehr bearbeiten. Um Änderungen vorzunehmen, löschen Sie das trainierte Modell und erstellen Sie ein neues.

Nutzung von inkrementellen Schulungen in AWS Clean Rooms ML

Voraussetzungen:

- Und AWS-Konto mit Zugriff auf AWS Clean Rooms
- Ein vorhandenes trainiertes Modell in einer Zusammenarbeit
- Ein neuer oder aktualisierter Datensatz für inkrementelles Training
- Entsprechende Berechtigungen zum Erstellen und Verwalten von ML-Modellen in der Zusammenarbeit
- Vertrautheit mit den Hyperparametern und der Konfiguration des vorhandenen Modells

Beim inkrementellen Training können Sie die Artefakte eines vorhandenen Modells und eines aktualisierten Datensatzes verwenden, um ein neues Modell zu trainieren. Durch das inkrementelle Training sparen Sie sowohl Zeit als auch Ressourcen.

Mit dem inkrementellen Training haben Sie folgende Möglichkeiten:

- Trainieren Sie ein neues Modell mithilfe eines erweiterten Datensatzes, dem ein Muster zugrunde liegt, das im vorherigen Training nicht berücksichtigt wurde.
- Trainieren Sie mehrere Varianten eines Modells, entweder mit unterschiedlichen Hyperparametern oder mit unterschiedlichen Datensätzen.

Console

Um einen inkrementellen Trainingsjob auszuführen (Konsole)

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie auf der Seite Kollaborationen die Kollaboration aus, in der die Modellartefakte vorhanden sind, die Sie für das inkrementelle Training verwenden möchten.
4. Wählen Sie nach dem Öffnen der Kollaboration die Registerkarte ML-Modelle aus.
5. Wählen Sie unter Benutzerdefinierte ML-Modelle im Abschnitt Trainierte Modelle das Optionsfeld neben dem trainierten Modell aus, das Sie inkrementell trainieren möchten.
6. Auf der Übersichtsseite unter Versionen
 - a. Wählen Sie das Optionsfeld neben dem trainierten Modell aus, das Sie schrittweise trainieren möchten.
 - b. Wählen Sie „Aus Version trainieren“.
7. Wählen Sie auf der Seite Trainiertes Modell aus Version erstellen für Trainierte Modellversion die Version aus.

Die Basismodellversion wird automatisch ausgewählt. Sie können diese Version ändern, falls andere Versionen existieren.

8. Geben Sie für Details zum trainierten Modell Folgendes ein:
 - a. Geben Sie unter Name einen eindeutigen Namen für das Modell in der Kollaboration ein.
 - b. (Optional) Geben Sie unter Beschreibung eine Beschreibung des trainierten Modells ein.
 - c. Wählen Sie für den Eingabemodus Trainingsdaten eine der folgenden Optionen:
 - Wählen Sie Datei aus, wenn Sie über einen kleineren Datensatz verfügen, der auf das ML-Speichervolumen passt, und Sie für Ihr Trainingsskript den herkömmlichen Dateisystemzugriff bevorzugen.
 - Wählen Sie Pipe für große Datensätze, um Daten direkt aus S3 zu streamen, sodass Sie nicht alles auf die Festplatte herunterladen müssen, was die Trainingsgeschwindigkeit verbessern und den Speicherbedarf reduzieren kann.

- Wählen Sie aus, FastFileob Sie die Vorteile des Streamings von S3 mit dem Dateisystemzugriff kombinieren möchten, insbesondere für sequentiell gelesene Daten oder wenn Sie mit weniger Dateien arbeiten müssen, um die Startzeiten zu verkürzen.
- d. Geben Sie unter Name des inkrementellen Trainingskanals einen Namen für den inkrementellen Trainingskanal ein

 Note

Wenn Sie den Namen des inkrementellen Trainingskanals ohne Versions-ID angeben, verwendet das System das Basismodell für das inkrementelle Training.

9. Gehen Sie wie folgt vor, um Details zum ML-Eingangskanal zu erhalten:

- a. Geben Sie für den ML-Eingangskanal den ML-Eingangskanal an, der Daten für den Modellalgorithmus bereitstellt.

Um einen weiteren Kanal hinzuzufügen, wählen Sie „Weiteren ML-Eingangskanal hinzufügen“. Sie können bis zu 19 zusätzliche ML-Eingangskanäle hinzufügen.

- b. Geben Sie als Kanalname den Namen des ML-Eingangskanals ein.
- c. Wählen Sie für den Amazon S3 S3-Datenverteilungstyp eine der folgenden Optionen:
- Wählen Sie Vollständig repliziert aus, um jeder Trainingsinstanz eine vollständige Kopie Ihres Datensatzes zu geben. Dies funktioniert am besten, wenn Ihr Datensatz klein genug ist, um in den Arbeitsspeicher zu passen, oder wenn jede Instanz Zugriff auf alle Daten benötigt.
 - Wählen Sie Sharded by S3 Key aus, um Ihren Datensatz anhand von S3-Schlüsseln auf Trainingsinstanzen aufzuteilen. Jede Instanz empfängt etwa $1/n$ der gesamten S3-Objekte, wobei 'n' für die Anzahl der Instanzen steht. Dies funktioniert am besten für große Datensätze, die Sie parallel verarbeiten möchten.

 Note

Berücksichtigen Sie bei der Auswahl eines Verteilungstyps Ihre Datensatzgröße und Ihre Trainingsanforderungen. Vollständig repliziert bietet vollständigen Datenzugriff, benötigt jedoch mehr Speicherplatz, während Sharded by S3 Key die verteilte Verarbeitung großer Datensätze ermöglicht.

10. Wählen Sie für Maximale Trainingsdauer die maximale Dauer aus, für die Sie Ihr Modell trainieren möchten.
11. Geben Sie für Hyperparameter alle algorithmusspezifischen Parameter und ihre beabsichtigten Werte an. Hyperparameter sind spezifisch für das trainierte Modell und werden zur Feinabstimmung des Modelltrainings verwendet.
12. Geben Sie für Umgebungsvariablen alle algorithmusspezifischen Variablen und ihre beabsichtigten Werte an. Umgebungsvariablen werden im Docker-Container festgelegt.
13. Um eine benutzerdefinierte Verschlüsselung zu verwenden AWS KMS key, aktivieren Sie das Kontrollkästchen Geheimnis mit einem benutzerdefinierten KMS-Schlüssel verschlüsseln.
14. Geben Sie für die EC2-Ressourcenkonfiguration Informationen zu den Rechenressourcen an, die für das Modelltraining verwendet werden.
 - a. Wählen Sie unter Instanztyp den Instanztyp aus, den Sie ausführen möchten.
 - b. Geben Sie unter Anzahl der Instanzen die Anzahl der Instanzen ein.
 - c. Geben Sie für Volumengröße in GB die Größe des ML-Speichervolumens ein.
15. Wählen Sie „Trainiertes Modell aus Version erstellen“.

API

Um einen inkrementellen Trainingsjob (API) auszuführen

Führen Sie den folgenden Code mit Ihren spezifischen Parametern aus:

```
import boto3
acr_ml_client= boto3.client('cleanroomsm1')

acr_ml_client.create_trained_model(
    membershipIdentifier= 'membership_id',
    configuredModelAlgorithmAssociationArn = 'arn:aws:cleanrooms-
ml:region:account:membership/membershipIdentifier/configured-model-algorithm-
association/identifier',
    name='trained_model_name',
    resourceConfig={
        'instanceType': 'ml.m5.xlarge',
        'volumeSizeInGB': 1
    },
    incrementalTrainingDataChannels=[
        {
```

```
        'trainedModelArn': trained_model_arn,  
        'channelName': 'channel_name'  
    },  
    ]  
    dataChannels=[  
        {  
            'mlInputChannelArn': channel_arn_1,  
            'channelName': 'channel_name'  
        },  
        {  
            'mlInputChannelArn': channel_arn_2,  
            'channelName': 'channel_name'  
        }  
    ]  
    )
```

 Note

Limit: Insgesamt maximal 20 Kanäle (einschließlich dataChannels sowohl als auch incrementalTrainingDataChannels).

 Note

Nachdem das trainierte Modell erstellt wurde, können Sie es nicht mehr bearbeiten. Um Änderungen vorzunehmen, löschen Sie das trainierte Modell und erstellen Sie ein neues.

Nutzung verteilter Schulungen in AWS Clean Rooms ML

Voraussetzungen:

- Und AWS-Konto mit Zugriff auf AWS Clean Rooms
- Eine Zusammenarbeit wurde eingerichtet in AWS Clean Rooms
- Ein konfigurierter Modellalgorithmus, der verteiltes Training unterstützt
- Ein großer Datensatz, der für die verteilte Verarbeitung geeignet ist
- Entsprechende Berechtigungen zum Erstellen und Verwalten von ML-Modellen in der Zusammenarbeit

- Ausreichendes Amazon EC2 EC2-Kontingent zum Ausführen mehrerer Instances für verteilte Schulungen

Verteiltes Training nutzt die Leistung vieler parallel arbeitender Rechenknoten, um große Datenmengen zu verarbeiten und Modellparameter effizient zu aktualisieren.

Weitere Informationen zu verteilten Schulungen finden Sie unter [Konzepte für verteilte Schulungen](#) im Amazon SageMaker AI Developer Guide.

Console

So führen Sie einen verteilten Trainingsjob aus (Konsole)

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie auf der Seite Kollaborationen die Kollaboration aus, für die Sie ein trainiertes Modell erstellen möchten.
4. Wählen Sie nach dem Öffnen der Kollaboration die Registerkarte ML-Modelle aus.
5. Wählen Sie unter Benutzerdefinierte ML-Modelle im Abschnitt Trainierte Modelle die Option Trainiertes Modell erstellen aus.
6. Geben Sie auf der Seite Trainiertes Modell erstellen für Algorithmus für Zugeordnetes Modell den Algorithmus an
7. Geben Sie für Details zum trainierten Modell Folgendes ein:
 - a. Geben Sie unter Name einen eindeutigen Namen für das Modell in der Kollaboration ein.
 - b. (Optional) Geben Sie unter Beschreibung eine Beschreibung des trainierten Modells ein.
 - c. Wählen Sie für den Eingabemodus Trainingsdaten eine der folgenden Optionen:
 - Wählen Sie Datei aus, wenn Sie über einen kleineren Datensatz verfügen, der auf das ML-Speichervolumen passt, und Sie für Ihr Trainingsskript den herkömmlichen Dateisystemzugriff bevorzugen.
 - Wählen Sie Pipe für große Datensätze, um Daten direkt aus S3 zu streamen, sodass Sie nicht alles auf die Festplatte herunterladen müssen, was die Trainingsgeschwindigkeit verbessern und den Speicherbedarf reduzieren kann.

- Wählen Sie aus, FastFileob Sie die Vorteile des Streamings von S3 mit dem Dateisystemzugriff kombinieren möchten, insbesondere für sequentiell gelesene Daten oder wenn Sie mit weniger Dateien arbeiten müssen, um die Startzeiten zu verkürzen.
8. Gehen Sie wie folgt vor, um Einzelheiten zum ML-Eingangskanal zu erhalten:

- a. Geben Sie für den ML-Eingangskanal den ML-Eingangskanal an, der Daten für den Modellalgorithmus bereitstellt.

Um einen weiteren Kanal hinzuzufügen, wählen Sie „Weiteren ML-Eingangskanal hinzufügen“. Sie können bis zu 19 zusätzliche ML-Eingangskanäle hinzufügen.

- b. Geben Sie als Kanalname den Namen des ML-Eingangskanals ein.
- c. Wählen Sie für den Amazon S3 S3-Datenverteilungstyp eine der folgenden Optionen:
- Wählen Sie Vollständig repliziert aus, um jeder Trainingsinstanz eine vollständige Kopie Ihres Datensatzes zu geben. Dies funktioniert am besten, wenn Ihr Datensatz klein genug ist, um in den Arbeitsspeicher zu passen, oder wenn jede Instanz Zugriff auf alle Daten benötigt.
 - Wählen Sie Sharded by S3 Key aus, um Ihren Datensatz anhand von S3-Schlüsseln auf Trainingsinstanzen aufzuteilen. Jede Instanz empfängt etwa $1/n$ der gesamten S3-Objekte, wobei 'n' für die Anzahl der Instanzen steht. Dies funktioniert am besten für große Datensätze, die Sie parallel verarbeiten möchten.

 Note

Berücksichtigen Sie bei der Auswahl eines Verteilungstyps Ihre Datensatzgröße und Ihre Trainingsanforderungen. Vollständig repliziert bietet vollständigen Datenzugriff, benötigt jedoch mehr Speicherplatz, während Sharded by S3 Key die verteilte Verarbeitung großer Datensätze ermöglicht.

9. Wählen Sie für Maximale Trainingsdauer die maximale Dauer aus, für die Sie Ihr Modell trainieren möchten.
10. Geben Sie für Hyperparameter alle algorithmusspezifischen Parameter und ihre beabsichtigten Werte an. Hyperparameter sind spezifisch für das trainierte Modell und werden zur Feinabstimmung des Modelltrainings verwendet.
11. Geben Sie für Umgebungsvariablen alle algorithmusspezifischen Variablen und ihre beabsichtigten Werte an. Umgebungsvariablen werden im Docker-Container festgelegt.

12. Um eine benutzerdefinierte Verschlüsselung zu verwenden AWS KMS key, aktivieren Sie das Kontrollkästchen Geheimnis mit einem benutzerdefinierten KMS-Schlüssel verschlüsseln.
13. Geben Sie für die EC2-Ressourcenkonfiguration Informationen zu den Rechenressourcen an, die für das Modelltraining verwendet werden.

- a. Wählen Sie unter Instanztyp den Instanztyp aus, den Sie ausführen möchten.

Folgende Instanztypen für verteiltes Training werden unterstützt:

- ml.m5.4xlarge
- ml.m5.12xlarge
- ml.m5.2xlarge
- ml.g5.12xlarge
- ml.g5.24xlarge

- b. Geben Sie im Feld Anzahl der Instanzen die Anzahl der Instanzen ein.
- c. Geben Sie für Volumengröße in GB die Größe des ML-Speichervolumens ein.

14. Wählen Sie Trainiertes Modell erstellen aus.

API

Um einen verteilten Trainingsjob (API) auszuführen

Führen Sie den folgenden Code mit Ihren spezifischen Parametern aus:

```
import boto3
acr_ml_client= boto3.client('cleanroomsm1')

acr_ml_client.create_trained_model(
    membershipIdentifier= 'membership_id',
    configuredModelAlgorithmAssociationArn = 'arn:aws:cleanrooms-
ml:region:account:membership/membershipIdentifier/configured-model-algorithm-
association/identifier',
    name='trained_model_name',
    trainingInputMode: "File",
    resourceConfig={
        'instanceCount': "3"
        'instanceType': "ml.m5.xlarge",
        'volumeSizeInGB': 3
    },
    ),
```

```
dataChannels=[
  {
    "mlInputChannelArn": channel_arn_1,
    "channelName": "channel_name",
    "S3DataDistributionType": "FullyReplicated"
  }
]
```

Note

Nachdem das trainierte Modell erstellt wurde, können Sie es nicht mehr bearbeiten. Um Änderungen vorzunehmen, löschen Sie das trainierte Modell und erstellen Sie ein neues.

Exportieren von Modellartefakten aus AWS Clean Rooms ML

Diese Aufgabe ist optional und sollte abgeschlossen sein, wenn Sie einem CAN_RECEIVE_MODEL_OUTPUT Mitglied der Kollaboration die Fähigkeit „Mitglied“ zugewiesen haben.

Nach Abschluss des Modelltrainings kann das Mitglied, das das Modell trainiert hat, den Export von Modellartefakten initiieren. Das Mitglied, das das Modell trainiert hat, entscheidet, wer Modellartefakte erhält, vorausgesetzt, dieses Mitglied kann Ergebnisse und eine gültige ML-Konfiguration erhalten.

Console

Um einen benutzerdefinierten ML-Modellalgorithmus zu konfigurieren (Konsole)

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie auf der Seite Kollaborationen die Kollaboration aus, die das benutzerdefinierte Modell enthält, das Sie exportieren möchten.
4. Wählen Sie nach dem Öffnen der Kollaboration die Registerkarte ML-Modelle und anschließend Ihr Modell aus der Tabelle Benutzerdefiniertes trainiertes Modell aus

5. Klicken Sie auf der Detailseite für das benutzerdefinierte trainierte Modell auf Modellausgabe exportieren.
6. Geben Sie für Modellausgabe exportieren für Modellausgabedetails exportieren einen Namen und optional eine Beschreibung ein.

Wählen Sie in der Dropdownliste Modellausgabe, die an Mitglieder der Kollaboration exportiert wurde, welches Mitglied die Modellartefakte erhalten soll.

7. Wählen Sie Export aus.

Die Ergebnisse werden in den folgenden Pfad am Amazon S3 S3-Speicherort exportiert, der in der ML-Konfiguration angegeben wurde: `yourSpecifiedS3Path/collaborationIdentifier/trainedModelName/callerAccountId/jobName`. Es werden nur die zu exportierenden Dateien bis zur angegebenen maximalen Dateigröße exportiert, die Sie bei der Zuordnung des konfigurierten Modellalgorithmus ausgewählt haben.

API

Um einen benutzerdefinierten ML-Modellalgorithmus (API) zu konfigurieren

Initiieren Sie den Modellexport, indem Sie den folgenden Code ausführen:

```
import boto3
acr_ml_client= boto3.client('cleanroomsm1')

acr_ml_client.start_trained_model_export_job(
    membershipIdentifier='membership_id',
    trainedModelArn='arn:aws:cleanrooms-ml:region:account:membership/
membershipIdentifier/trained-model/identifier',
    outputConfiguration={
        'member': {
            'accountId': 'model_output_receiver_account'
        }
    },
    name='export_job_name'
)
```

Die Ergebnisse werden in den folgenden Pfad am Amazon S3 S3-Speicherort exportiert, der in der ML-Konfiguration angegeben wurde: `yourSpecifiedS3Path/collaborationIdentifier/trainedModelName/callerAccountId/jobName`. Nur

die `filesToExport` bis zu den `maxSize` angegebenen Daten, die Sie bei der Zuordnung des konfigurierten Modellalgorithmus ausgewählt haben, werden exportiert.

Ausführen von Inferenzen auf einem trainierten Modell in AWS Clean Rooms ML

Mitglieder, die Abfragen ausführen können, können auch Inferenzaufgaben starten, sobald die Trainingsaufgabe abgeschlossen ist. Sie wählen den Inferenzdatensatz aus, für den sie die Inferenz ausführen möchten, und verweisen auf die trainierten Modellausgaben, mit denen sie den Inferenzcontainer ausführen möchten.

Dem Mitglied, das die Ergebnisse der Inferenz erhalten soll, muss die Fähigkeit „Mitglied“ gewährt werden. `CAN_RECEIVE_INFERENCE_OUTPUT`

Console

Um einen Model-Inferenz-Job zu erstellen (Konsole)

1. Melden Sie sich bei <https://console.aws.amazon.com/cleanrooms> an AWS-Managementkonsole und öffnen Sie die AWS Clean Rooms Konsole.
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie auf der Seite Kollaborationen die Kollaboration aus, die das benutzerdefinierte Modell enthält, für das Sie einen Inferenzjob erstellen möchten.
4. Wählen Sie nach dem Öffnen der Kollaboration die Registerkarte ML-Modelle und dann Ihr Modell aus der Tabelle Benutzerdefiniertes trainiertes Modell aus.
5. Klicken Sie auf der Detailseite für das benutzerdefinierte trainierte Modell auf Inferenzjob starten.
6. Geben Sie für Inferenzjob starten für Inferenzjobdetails einen Namen und optional eine Beschreibung ein.

Geben Sie die folgenden Informationen ein:

- Zugeordneter Modellalgorithmus — Der zugehörige Modellalgorithmus, der während des Inferenzjobs verwendet wird.
- Details zum ML-Eingangskanal — Der ML-Eingangskanal, der die Daten für diesen Inferenzjob bereitstellt.

- Transformationsressourcen — Die Recheninstanz, die zur Ausführung der Transformationsfunktion des Inferenzjobs verwendet wird.
- Ausgabekonfiguration — Wer erhält die Ausgabe des Inferenzjobs und den MIME-Typ der Ausgabe.
- Verschlüsselung — Wählen Sie die Verschlüsselungseinstellungen anpassen, um Ihren eigenen KMS-Schlüssel und zugehörige Informationen anzugeben. Andernfalls verwaltet Clean Rooms ML die Verschlüsselung.
- Auftragsdetails transformieren — Die maximale Nutzlast des Inferenzjobs in MB.
- Umgebungsvariablen — Alle Umgebungsvariablen, die für den Zugriff auf das Container-Image des Inferenzjobs erforderlich sind.

7. Wählen Sie Inferenzjob starten aus.

Die Ergebnisse werden in den folgenden Pfad am Amazon S3 S3-Speicherort exportiert, der in der ML-Konfiguration angegeben wurde: `yourSpecifiedS3Path/collaborationIdentifier/trainedModelName/callerAccountId/jobName`.

API

Um einen Model-Inferenz-Job (API) zu erstellen

Initiieren Sie den Inferenzjob, indem Sie den folgenden Code ausführen:

```
import boto3
acr_ml_client= boto3.client('cleanroomsm1')

acr_ml_client.start_trained_model_inference_job(
    name="inference_job",
    membershipIdentifier='membership_id',
    trainedModelArn='arn:aws:cleanrooms-ml:region:account:membership/
membershipIdentifier/trained-model/identifier',

    dataSource={
        "mlInputChannelArn": 'channel_arn_3'
    },
    resourceConfig={'instanceType': 'ml.m5.xlarge'},
    outputConfiguration={
        'accept': 'text/csv',
        'members': [
            {
```

```
        "accountId": 'member_account_id'  
      }  
    ]  
  }  
)
```

Die Ergebnisse werden in den folgenden Pfad am Amazon S3 S3-Speicherort exportiert, der in der ML-Konfiguration angegeben wurde: `yourSpecifiedS3Path/collaborationIdentifier/trainedModelName/callerAccountId/jobName`.

Erstellen Sie AWS Clean Rooms-ML-Modelle als Anbieter von Trainingsdaten

Ein Lookalike-Modell ist ein Modell der Daten eines Trainingsdatenanbieters, das es einem Anbieter von Ausgangsdaten ermöglicht, ein ähnliches Segment der Daten eines Trainingsdatenanbieters zu erstellen, das seinen Ausgangsdaten am ähnlichsten ist. Um ein Lookalike-Modell zu erstellen, das in einer Zusammenarbeit verwendet werden kann, müssen Sie Ihre Trainingsdaten importieren, ein Lookalike-Modell erstellen, dieses Lookalike-Modell konfigurieren und es dann einer Kollaboration zuordnen.

Um mit Lookalike-Modellen zu arbeiten, müssen zwei Parteien, ein Anbieter von Trainingsdaten und ein Anbieter von Ausgangsdaten, nacheinander zusammenarbeiten, um ihre Daten in eine Zusammenarbeit AWS Clean Rooms einzubringen. Dies ist der Workflow, den der Trainingsdatenanbieter zuerst abschließen muss:

1. Die Daten des Trainingsdatenanbieters müssen in einer AWS Glue Datenkatalogtabelle mit Interaktionen zwischen Benutzern und Elementen gespeichert werden. Die Trainingsdaten müssen mindestens eine Benutzer-ID-Spalte, eine Interaktions-ID-Spalte und eine Zeitstempelspalte enthalten.
2. Der Trainingsdatenanbieter registriert die Trainingsdaten bei AWS Clean Rooms.
3. Der Trainingsdatenanbieter erstellt ein Lookalike-Modell, das mit mehreren Startdatenanbietern gemeinsam genutzt werden kann. Das Lookalike-Modell ist ein tiefes neuronales Netzwerk, dessen Training bis zu 24 Stunden dauern kann. Es wird nicht automatisch neu trainiert und wir empfehlen, dass Sie das Modell wöchentlich neu trainieren.
4. Der Anbieter von Trainingsdaten konfiguriert das Lookalike-Modell, einschließlich der Frage, ob Relevanzkennzahlen und der Amazon S3 S3-Speicherort der Ausgabesegmente geteilt werden sollen. Der Anbieter von Trainingsdaten kann mehrere konfigurierte Lookalike-Modelle aus einem einzigen Lookalike-Modell erstellen.
5. Der Anbieter von Trainingsdaten ordnet das konfigurierte Zielgruppenmodell einer Zusammenarbeit zu, die mit einem Startdatenanbieter geteilt wird.

Nachdem der Trainingsdatenanbieter das ML-Modell erstellt hat, [kann der Seed-Datenanbieter das Lookalike-Segment erstellen und exportieren](#).

Topics

- [Trainingsdaten importieren](#)
- [Ein Lookalike-Modell erstellen](#)
- [Konfiguration eines Lookalike-Modells](#)
- [Zuordnen eines konfigurierten Lookalike-Modells](#)
- [Aktualisierung eines konfigurierten Lookalike-Modells](#)

Trainingsdaten importieren

Note

Sie können nur einen Trainingsdatensatz zur Verwendung in einem Clean Rooms ML-Lookalike-Modell bereitstellen, dessen Daten in Amazon S3 gespeichert sind. Sie können jedoch die Ausgangsdaten für ein Lookalike-Modell mithilfe von SQL bereitstellen, das auf Daten läuft, die in einer beliebigen unterstützten Datenquelle gespeichert sind.

Bevor Sie ein Lookalike-Modell erstellen, müssen Sie die AWS Glue Tabelle angeben, die die Trainingsdaten enthält. Clean Rooms ML speichert keine Kopie dieser Daten, sondern lediglich Metadaten, die den Zugriff auf die Daten ermöglichen.

Um Trainingsdaten zu importieren AWS Clean Rooms

1. Melde dich bei der an AWS-Managementkonsole und öffne die [AWS Clean Rooms Konsole](#) mit deinem AWS-Konto (falls du das noch nicht getan hast).
2. Wählen Sie im linken Navigationsbereich AWS ML-Modelle aus.
3. Wählen Sie auf der Registerkarte Trainingsdatensätze die Option Trainingsdatensatz erstellen aus.
4. Geben Sie auf der Seite Trainingsdatensatz erstellen für Details zum Trainingsdatensatz einen Namen und optional eine Beschreibung ein.
5. Wählen Sie die Trainingsdatenquelle aus, indem Sie die Datenbank und die Tabelle, die Sie konfigurieren möchten, aus den Dropdownlisten auswählen.

Note

Um zu überprüfen, ob es sich um die richtige Tabelle handelt, führen Sie einen der folgenden Schritte aus:

- Wählen Sie Anzeigen in AWS Glue.
- Aktivieren Sie „Schema anzeigen“, um das Schema anzuzeigen.

6. Wählen Sie für Trainingsdetails die Spalten Benutzer-ID, Artikel-ID und Timestamp aus den Drop-down-Listen aus. Die Trainingsdaten müssen diese drei Spalten enthalten. Sie können auch alle anderen Spalten auswählen, die Sie in die Trainingsdaten aufnehmen möchten.

Die Daten in der Timestamp-Spalte müssen im Format Unix-Epochezeit in Sekunden vorliegen.

7. (Optional) Wenn Sie weitere Spalten trainieren möchten, wählen Sie den Spaltennamen und den Typ aus den Dropdownlisten aus.
8. Unter Dienstzugriff müssen Sie eine Servicerolle angeben, die auf Ihre Daten zugreifen kann, und einen KMS-Schlüssel angeben, falls Ihre Daten verschlüsselt sind. Wählen Sie Neue Servicerolle erstellen und verwenden aus. Clean Rooms ML erstellt dann automatisch eine Servicerolle und fügt die erforderlichen Berechtigungsrichtlinien hinzu. Wählen Sie Bestehende Servicerolle verwenden und geben Sie diese in das Feld Servicerollenname ein, wenn Sie über eine bestimmte Servicerolle verfügen, die Sie verwenden möchten.

Wenn Ihre Daten verschlüsselt sind, geben Sie Ihren KMS-Schlüssel in das AWS KMS keyFeld ein oder klicken Sie auf Erstellen, AWS KMS key um einen neuen KMS-Schlüssel zu generieren.

9. Wenn Sie Tags für den Trainingsdatensatz aktivieren möchten, wählen Sie Neues Tag hinzufügen und geben Sie dann das Schlüssel-Wert-Paar ein.
10. Wählen Sie Trainingsdatensatz erstellen aus.

Die entsprechende API-Aktion finden Sie unter [CreateTrainingDataset](#).

Ein Lookalike-Modell erstellen

Nachdem Sie einen Trainingsdatensatz erstellt haben, sind Sie bereit, ein Lookalike-Modell zu erstellen. Sie können viele Lookalike-Modelle aus einem einzigen Trainingsdatensatz erstellen.

Sie müssen eine Standarddatenbank in Ihrer Rolle erstellen AWS Glue Data Catalog oder die `glue:createDatabase` Berechtigung in der angegebenen Rolle angeben.

So erstellen Sie ein Lookalike-Modell in AWS Clean Rooms

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#) mit Ihrem AWS-Konto (falls Sie dies noch nicht getan haben).

2. Wählen Sie im linken Navigationsbereich AWS ML-Modelle aus.
3. Wählen Sie auf der Registerkarte Lookalike-Modelle die Option Lookalike-Modell erstellen aus.
4. Geben Sie auf der Seite Lookalike-Modell erstellen für Details zum Lookalike-Modell einen Namen und optional eine Beschreibung ein.
 - a. Wählen Sie den Trainingsdatensatz, den Sie modellieren möchten, aus der Dropdownliste aus.

 Note

Um zu überprüfen, ob es sich um den richtigen Trainingsdatensatz handelt, aktivieren Sie die Option Details zum Trainingsdatensatz anzeigen, um die Details anzuzeigen.

Um einen neuen Trainingsdatensatz zu erstellen, wählen Sie Trainingsdatensatz erstellen.

- b. (Optional) Rufen Sie ein Trainingsfenster auf.
5. Wenn Sie benutzerdefinierte Verschlüsselungseinstellungen für das Lookalike-Modell aktivieren möchten, wählen Sie Verschlüsselungseinstellungen anpassen und geben Sie dann den KMS-Schlüssel ein.
6. Wenn Sie Tags für das Lookalike-Modell aktivieren möchten, wählen Sie Neues Tag hinzufügen und geben Sie dann das Schlüssel-Wert-Paar ein.
7. Wählen Sie Lookalike-Modell erstellen aus.

 Note

Das Modelltraining kann mehrere Stunden bis 2 Tage dauern.

Die entsprechende API-Aktion finden Sie unter [CreateAudienceModel](#).

Konfiguration eines Lookalike-Modells

Nachdem Sie ein Lookalike-Modell erstellt haben, können Sie es für die Verwendung in einer Kollaboration konfigurieren. Sie können mehrere konfigurierte Lookalike-Modelle aus einem einzigen Lookalike-Modell erstellen.

So konfigurieren Sie ein Lookalike-Modell in AWS Clean Rooms

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#) mit Ihrem AWS-Konto (falls Sie dies noch nicht getan haben).
2. Wählen Sie im linken Navigationsbereich AWS ML-Modelle aus.
3. Wählen Sie auf der Registerkarte Konfigurierte Lookalike-Modelle die Option Lookalike-Modell konfigurieren aus.
4. Geben Sie auf der Seite Lookalike-Modell konfigurieren für Details zum konfigurierten Lookalike-Modell einen Namen und optional eine Beschreibung ein.
 - a. Wählen Sie das Lookalike-Modell, das Sie konfigurieren möchten, aus der Dropdownliste aus.

Note

Um zu überprüfen, ob es sich um das richtige Lookalike-Modell handelt, aktivieren Sie die Option Details des Lookalike-Modells anzeigen, um die Details anzuzeigen. Um ein neues Lookalike-Modell zu erstellen, wählen Sie „Lookalike-Modell erstellen“.

- b. Wählen Sie die gewünschte Mindestgröße für den passenden Samen aus. Dies ist die Mindestanzahl von Benutzern in den Daten des Seed-Datenanbieters, die sich mit den Benutzern in den Trainingsdaten überschneiden. Dieser Wert muss größer als 0 sein.
5. Damit Metriken mit anderen Mitgliedern geteilt werden können, wählen Sie aus, ob der Seed-Datenanbieter in Ihrer Zusammenarbeit Modellmetriken, einschließlich Relevanzbewertungen, erhalten soll.
 6. Geben Sie für Zielort des Lookalike-Segments den Amazon S3 S3-Bucket ein, in den das Lookalike-Segment exportiert wird. Dieser Bucket muss sich in derselben Region wie Ihre anderen Ressourcen befinden.
 7. Wählen Sie für Dienstzugriff den Namen der vorhandenen Servicerolle aus, der für den Zugriff auf diese Tabelle verwendet werden soll.
 8. Geben Sie für die erweiterte Konfiguration der Partitionsgröße den Typ Zielgruppengröße entweder als Absolute Zahl oder als Prozentsatz an.
 9. Wenn Sie Tags für die konfigurierte Tabellenressource aktivieren möchten, wählen Sie Neues Tag hinzufügen und geben Sie dann das Schlüssel - und Wertepaar ein.
 10. Wählen Sie Lookalike-Modell konfigurieren aus.

Die entsprechende API-Aktion finden Sie unter [CreateConfiguredAudienceModel](#).

Zuordnen eines konfigurierten Lookalike-Modells

Nachdem Sie ein Lookalike-Modell konfiguriert haben, können Sie es einer Kollaboration zuordnen.

Um ein konfiguriertes Lookalike-Modell zuzuordnen AWS Clean Rooms

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#) mit Ihrem AWS-Konto (falls Sie dies noch nicht getan haben).
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie auf der Registerkarte Mit aktiver Mitgliedschaft eine Kollaboration aus.
4. Wählen Sie auf der Registerkarte ML-Modelle unter eady-to-useR-Lookalike-Modelle die Option Lookalike-Modell zuordnen aus.
5. Gehen Sie auf der Seite Konfiguriertes Lookalike-Modell zuordnen für Details zur Zuordnung konfigurierter Lookalike-Modelle wie folgt vor:
 - a. Geben Sie einen Namen für das zugehörige konfigurierte Zielgruppenmodell ein.
 - b. Geben Sie eine Beschreibung der Tabelle ein.

Die Beschreibung hilft dabei, zwischen anderen zugehörigen konfigurierten Zielgruppenmodellen mit ähnlichen Namen zu unterscheiden.

6. Wählen Sie für Konfiguriertes Lookalike-Modell ein konfiguriertes Lookalike-Modell aus der Drop-down-Liste aus.
7. Wählen Sie Associate aus.

Die entsprechende API-Aktion finden Sie unter [CreateConfiguredAudienceModelAssociation](#)

Aktualisierung eines konfigurierten Lookalike-Modells

Nachdem Sie ein konfiguriertes Lookalike-Modell zugeordnet haben, können Sie es aktualisieren, um Informationen wie den Namen, die zu teilenden Metriken oder den Amazon S3 S3-Ausgabeort zu ändern.

So aktualisieren Sie ein zugeordnetes konfiguriertes Lookalike-Modell in AWS Clean Rooms

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#) mit Ihrem AWS-Konto (falls Sie dies noch nicht getan haben).
2. Wählen Sie im linken Navigationsbereich AWS ML-Modelle aus.
3. Wählen Sie auf der Registerkarte Konfigurierte Lookalike-Modelle unter Ready-to-use Lookalike-Modelle ein konfiguriertes Lookalike-Modell aus und wählen Sie Bearbeiten aus.
4. Gehen Sie auf der Seite Bearbeiten für Details zur Zuordnung konfigurierter Lookalike-Modelle wie folgt vor:
 - a. Aktualisieren Sie den Namen und optional die Beschreibung.
 - b. Wählen Sie das Lookalike-Modell, das Sie konfigurieren möchten, aus der Dropdownliste aus.
 - c. Wählen Sie die gewünschte Mindestgröße für die passende Samengröße aus. Dies ist die Mindestanzahl von Benutzern in den Daten des Seed-Datenanbieters, die sich mit den Benutzern in den Trainingsdaten überschneiden. Dieser Wert muss größer als 0 sein.
5. Damit Metriken mit anderen Mitgliedern geteilt werden können, wählen Sie aus, ob der Seed-Datenanbieter in Ihrer Zusammenarbeit Modellmetriken, einschließlich Relevanzbewertungen, erhalten soll.
6. Geben Sie für Zielort des Lookalike-Segments den Amazon S3 S3-Bucket ein, in den das Lookalike-Segment exportiert wird. Dieser Bucket muss sich in derselben Region wie Ihre anderen Ressourcen befinden.
7. Wählen Sie für Dienstzugriff den Namen der vorhandenen Servicerolle aus, der für den Zugriff auf diese Tabelle verwendet werden soll.
8. Wählen Sie unter Erweiterte Konfiguration der Partitionsgröße aus, wie Sie die Zielgruppen-Bin-Größen konfigurieren möchten.
9. Wählen Sie Änderungen speichern aus.

Die entsprechende API-Aktion finden Sie unter [UpdateConfiguredAudienceModel](#).

Erstellung von AWS Clean Rooms-ML-Modellen als Seed-Datenanbieter

Nachdem der Trainingsdatenanbieter das ML-Modell erstellt hat, kann der Seed-Datenanbieter das Lookalike-Segment erstellen und exportieren. Das Lookalike-Segment ist eine Teilmenge der Trainingsdaten, die den Ausgangsdaten am ähnlichsten ist.

Dies ist der Workflow, den der Seed-Datenanbieter abschließen muss:

1. Die Daten des Seed-Datenanbieters können in einem Amazon S3 S3-Bucket gespeichert werden oder aus den Ergebnissen einer Abfrage stammen.
2. Der Seed-Datenanbieter eröffnet die Zusammenarbeit, die er mit dem Trainingsdatenanbieter teilt.
3. Der Seed-Datenanbieter erstellt auf der Registerkarte Clean Rooms ML der Kollaborationsseite ein ähnliches Segment.
4. Der Seed-Datenanbieter kann die Relevanzkennzahlen auswerten, sofern sie geteilt wurden, und das Lookalike-Segment zur externen Verwendung exportieren. AWS Clean Rooms

Topics

- [Ein Lookalike-Segment erstellen](#)
- [Exportieren eines Lookalike-Segments](#)

Ein Lookalike-Segment erstellen

Note

Sie können nur einen Trainingsdatensatz zur Verwendung in einem Clean Rooms ML-Lookalike-Modell bereitstellen, dessen Daten in Amazon S3 gespeichert sind. Sie können jedoch die Ausgangsdaten für ein Lookalike-Modell mithilfe von SQL bereitstellen, das auf Daten läuft, die in einer beliebigen unterstützten Datenquelle gespeichert sind.

Ein Lookalike-Segment ist eine Teilmenge der Trainingsdaten, die den Ausgangsdaten am ähnlichsten ist.

Um ein Lookalike-Segment zu erstellen AWS Clean Rooms

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#) mit Ihrem AWS-Konto (falls Sie dies noch nicht getan haben).
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie auf der Registerkarte Mit aktiver Mitgliedschaft eine Kollaboration aus.
4. Wählen Sie auf der Registerkarte ML-Modelle die Option Lookalike-Segment erstellen aus.
5. Wählen Sie auf der Seite Lookalike-Segment erstellen unter Zugeordnetes konfiguriertes Lookalike-Modell das zugehörige konfigurierte Lookalike-Modell aus, das für dieses Lookalike-Segment verwendet werden soll.
6. Geben Sie für Details zum Lookalike-Segment einen Namen und optional eine Beschreibung ein.
7. Wählen Sie für Seed-Profile Ihre Seed-Methode aus, indem Sie eine Option auswählen und dann die empfohlene Maßnahme ergreifen.

Option	Empfohlene Aktion
Amazon S3 S3-Pfad	1. Wählen Sie einen Amazon S3 S3-Standort aus. 2. (Optional) Wählen Sie „Ausgangsprofile in die Ausgabe einbeziehen“.
SQL-Abfrage	Schreiben Sie eine SQL-Abfrage und verwenden Sie ihre Ergebnisse als Ausgangsdaten.
Analysevorlage	Wählen Sie eine Analysevorlage aus der Drop-down-Liste und verwenden Sie die Ergebnisse, die mit einer Analysevorlage erstellt wurden.

8. Wählen Sie den Worker-Typ aus, der bei der Erstellung dieser Datenquelle verwendet werden soll. Der Standard-Worker-Typ ist CR.1X. Geben Sie die Anzahl der zu verwendenden Worker an. Die Standardeinstellung ist Arbeiter Nummer 16. So geben Sie Spark-Eigenschaften an:
 - a. Erweitern Sie Spark-Eigenschaften.

- b. Wählen Sie Spark-Eigenschaften hinzufügen.
- c. Wählen Sie im Dialogfeld „Spark-Eigenschaften“ einen Eigenschaftsnamen aus der Dropdownliste aus und geben Sie einen Wert ein.

Die folgende Tabelle enthält eine Definition für jede Eigenschaft.

Weitere Informationen zu Spark-Eigenschaften finden Sie unter [Spark-Eigenschaften](#) in der Apache Spark-Dokumentation.

Eigenschaftensname	Description	Standardwert
spark.task.maxFailures	Steuert, wie oft eine Aufgabe hintereinander fehlschlagen kann, bevor der Job fehlschlägt. Erfordert einen Wert größer oder gleich 1. Die Anzahl der zulässigen Wiederholungen entspricht diesem Wert minus 1. Die Anzahl der Fehlschläge wird zurückgesetzt, wenn ein Versuch erfolgreich ist. Fehler bei verschiedenen Aufgaben summieren sich nicht bis zu diesem Limit.	4
spark.sql.files.maxPartitionBytes	Legt die maximale Anzahl von Byte fest, die beim Lesen aus dateibasierten Quellen wie Parquet, JSON und ORC in eine einzelne Partition gepackt werden sollen.	128 MB
spark.hadoop.fs.s3.maxRetries	Legt die maximale Anzahl von Wiederholungsversuchen für Amazon S3 S3-Dateioperationen fest.	

Eigenschaftename	Description	Standardwert
spark.network.timeout	Legt das Standard-Timeout für alle Netzwerkinteraktionen fest. Setzt die folgenden Timeout-Einstellungen außer Kraft, wenn sie nicht konfiguriert sind: • spark.storage.blockManagerHeartbeatTimeoutMs • spark.shuffle.io.ConnectionTimeout • spark.rpc.AskTimeout • spark.rpc.LookupTimeout	120 s
spark.rdd.compress	Gibt an, ob serialisierte RDD-Partitionen mit spark.io.compression.codec komprimiert werden sollen. Gilt für StorageLevel .MEMORY_ONLY_SER in Java und Scala oder .MEMORY_ONLY in Python. StorageLevel Reduziert den Speicherplatz, erfordert jedoch zusätzliche CPU-Verarbeitungszeit.	FALSE
spark.shuffle.spill.compress	Gibt an, ob Shuffle-Spill-Daten mit spark.io.compression.codec komprimiert werden sollen.	TRUE
spark.sql.adaptive.advisoryPartitionSizeInBytes	Legt die Zielgröße in Byte für Shuffle-Partitionen während der adaptiven Optimierung fest, wenn spark.sql.adaptive.enabled den Wert true hat. Steuert die Partitionsgröße beim Zusammenführen kleiner Partitionen oder beim Teilen schiefer Partitionen.	(Wert von spark.sql.adaptive.shuffle.targetPostShuffleInputSize)

Eigenschaftensname	Description	Standardwert
spark.sql.adaptive.autoBroadcastJoinSchwellenwert	Legt die maximale Tabellengröße in Byte für die Übertragung an Worker-Knoten bei Joins fest. Gilt nur im adaptiven Framework . Verwendet denselben Standardwert wie spark.sql. autoBroadcastJoinSchwellenwert. Auf -1 setzen, um die Übertragung zu deaktivieren.	(Keine)
spark.sql.adaptive.CoalescePartitions.Enabled	Gibt an, ob zusammenhängende Shuffle-Partitionen, die auf spark.sql.adaptive basieren, zusammengeführt werden sollen. advisoryPartitionSizeInBytes um die Aufgabengröße zu optimieren. Erfordert, dass spark.sql.adaptive.enabled den Wert true hat.	TRUE
spark.sql.adaptive.CoalescePartitions.initialPartitionNum	Definiert die anfängliche Anzahl von Shuffle-Partitionen vor dem Zusammenführen. Erfordert, dass sowohl spark.sql.adaptive.enabled als auch spark.sql.adaptive.CoalescePartitions.Enabled den Wert true haben. Der Standardwert ist der Wert von spark.sql.shuffle.partitions.	(Keine)
spark.sql.adaptive.CoalescePartitions.minPartitionSize	Legt die Mindestgröße für zusammengeführte Shuffle-Partitionen fest, um zu verhindern, dass Partitionen während der adaptiven Optimierung zu klein werden.	1 MB

Eigenschaftensname	Description	Standardwert
spark.sql.adaptive. CoalescePartitions.ParallelismFirst	Gibt an, ob Partitionsgrößen auf der Grundlage der Cluster-Parallelität und nicht auf der Grundlage von spark.sql.adaptive.advisoryPartitionSizeInBytes während der Zusammenführung von Partitionen. Generiert kleinere Partitionsgrößen als die konfigurierte Zielgröße, um die Parallelität zu maximieren. Wir empfehlen, diesen Wert bei stark frequentierten Clustern auf „False“ zu setzen, um die Ressourcennutzung zu verbessern und übermäßig kleine Aufgaben zu vermeiden.	TRUE
spark.sql.adaptive. .enabled	Gibt an, ob die adaptive Abfrageausführung aktiviert werden soll, um Abfragepläne während der Abfrageausführung auf der Grundlage genauer Laufzeitstatistiken erneut zu optimieren.	TRUE
spark.sql.adaptive. forceOptimizeSkewedBeitreten	Gibt an, ob die Aktivierung erzwungen werden soll, OptimizeSkewedJoin auch wenn dadurch zusätzlicher Shuffle eingeführt wird.	FALSE
spark.sql.adaptive. localShuffleReader. aktiviert	Gibt an, ob lokale Shuffle-Reader verwendet werden sollen, wenn eine Shuffle-Partitionierung nicht erforderlich ist, z. B. nach der Konvertierung von Sort-Merge-Joins in Broadcast-Hash-Joins. Erfordert, dass spark.sql.adaptive.enabled den Wert true hat.	TRUE

Eigenschaftename	Description	Standardwert
spark.sql.adaptive.maxShuffledHashJoinLocalMapThreshold	Legt die maximale Partitionsgröße in Byte für die Erstellung lokaler Hash-Maps fest. Priorisiert gemischte Hash-Joins gegenüber Sort-Merge-Joins, wenn: - Dieser Wert entspricht oder übersteigt spark.sql.adaptive.advisoryPartitionSizeInBytes - Alle Partitionsgrößen liegen innerhalb dieser Grenze Überschreibt spark.sql.join.preferSortMergeJoin-Einstellung.	0 Byte
spark.sql.adaptive.optimizeSkewsInRebalancePartitions.aktiviert	Gibt an, ob schiefe Shuffle-Partitionen optimiert werden sollen, indem sie auf der Grundlage von spark.sql.adaptive in kleinere Partitionen aufgeteilt werden. advisoryPartitionSizeInBytes. Erfordert, dass spark.sql.adaptive.enabled wahr ist.	TRUE
spark.sql.adaptive.rebalancePartitionsSmallPartitionFactor	Definiert den Größenschwellenwert für das Zusammenführen von Partitionen beim Teilen. Partitionen, die kleiner sind als dieser Faktor multipliziert mit spark.sql.adaptive.advisoryPartitionSizeInBytes werden zusammengeführt.	0.2
spark.sql.Adaptive.SkewJoin.Enabled	Gibt an, ob Datenverzerrungen in gemischten Verknüpfungen behandelt werden sollen, indem schiefe Partitionen aufgeteilt und optional repliziert werden. Gilt für Sort-Merge- und Shuffled-Hash-Joins. Erfordert, dass spark.sql.adaptive.enabled wahr ist.	TRUE

Eigenschaftename	Description	Standardwert
spark.sql.adaptive.SkewJoin.skewedPartitionFactor	Bestimmt den Größenfaktor, der die Partition sneigung bestimmt. Eine Partition ist schief, wenn ihre Größe beide Werte überschreitet: - Dieser Faktor wird mit der mittleren Partitionsgröße multipliziert - Der Wert von spark.sql.adaptive.SkewJoin.skewedPartitionThresholdInBytes	5
spark.sql.Adaptive.SkewJoin.skewedPartitionThresholdInBytes	Legt den Größenschwellenwert in Byte zur Identifizierung schiefer Partitionen fest. Eine Partition ist schief, wenn ihre Größe beide Werte überschreitet: - Dieser Schwellenwert - Die mittlere Partitionsgröße multipliziert mit spark.sql.adaptive.SkewJoin.skewedPartitionFactor Wir empfehlen, diesen Wert größer als spark.sql.adaptive festzulegen. advisoryPartitionSizeInBytes.	256 MB
spark.sql. autoBroadcastJoinSchwellenwert	Legt die maximale Tabellengröße in Byte für die Übertragung an Worker-Knoten bei Joins fest. Auf -1 setzen, um die Übertragung zu deaktivieren.	10 MB
spark.sql.BroadcastTimeout	Steuert den Timeout-Zeitraum in Sekunden für die Broadcast-Operationen bei Broadcast-Joins.	300 Sekunden
spark.sql.cbo.enabled	Gibt an, ob die kostenbasierte Optimierung (CBO) für die Schätzung von Planstatistiken aktiviert werden soll.	FALSE

Eigenschaftename	Description	Standardwert
spark.sql.CBO.JoinReorder.DP.Star.Filter	Gibt an, ob bei der kostenbasierten Join-Aufzählung Heuristiken des Star-Join-Filters angewendet werden sollen.	FALSE
spark.sql.CBO.JoinReorder.DP.Threshold	Legt die maximale Anzahl verbundener Knoten fest, die im dynamischen Programmieralgorithmus zulässig sind.	12
spark.sql.CBO.JoinReorder.Enabled	Gibt an, ob die Neuordnung von Verknüpfungen bei der kostenbasierten Optimierung (CBO) aktiviert werden soll.	FALSE
spark.sql.CBO.PlanStats.Enabled	Gibt an, ob bei der Generierung logischer Pläne Zeilenanzahlen und Spaltenstatistiken aus dem Katalog abgerufen werden sollen.	FALSE
spark.sql.cbo.starSchemaDetection	Gibt an, ob die Neuordnung von Verknüpfungen auf der Grundlage der Star-Schemamarkierung aktiviert werden soll.	FALSE
spark.sql.files.maxPartitionNum	Legt die maximale Zielanzahl von geteilten Dateipartitionen für dateibasierte Quellen (Parquet, JSON und ORC) fest. Skaliert Partitionen neu, wenn die anfängliche Anzahl diesen Wert überschreitet. Dies ist ein empfohlenes Ziel, kein garantiertes Limit.	(Keine)
spark.sql.files.maxRecordSizePerDatei	Legt die maximale Anzahl von Datensätzen fest, die in eine einzelne Datei geschrieben werden sollen. Wenn der Wert auf Null oder einen negativen Wert gesetzt ist, gilt kein Limit.	0

Eigenschaftensname	Description	Standardwert
spark.sql.files.minPartitionNum	Legt die Mindestanzahl von geteilten Dateipartitionen für dateibasierte Quellen (Parquet, JSON und ORC) fest. Die Standardeinstellung ist spark.sql. leafNodeDefaultParallelität. Dies ist ein empfohlenes Ziel, kein garantiertes Limit.	(Keine)
spark.sql.inMemoryColumnarSpeicher.Batchgröße	Steuert die Batchgröße für das spaltenförmige Caching. Eine Erhöhung der Größe verbessert die Speichernutzung und Komprimierung, erhöht jedoch das Fehlerrisiko out-of-memory.	10000
spark.sql.inMemoryColumnarSpeicher.Komprimiert	Gibt an, ob auf der Grundlage von Datenstatistiken automatisch Komprimierungscodecs für Spalten ausgewählt werden sollen.	TRUE
spark.sql.inMemoryColumnarAufbewahrung.enableVectorizedReader	Gibt an, ob vektorisiertes Lesen für das spaltenförmige Caching aktiviert werden soll.	TRUE
spark.sql.legacy.allowHashOnMapType	Gibt an, ob Hash-Operationen für Map-Datenstrukturen zulässig sind. Diese Legacy-Einstellung gewährleistet die Kompatibilität mit der Map-Typ-Behandlung älterer Spark-Versionen.	
spark.sql.legacy.allowNegativeScaleOfDecimal	Gibt an, ob negative Skalenwerte in Dezimaltypdefinitionen zulässig sind. Diese ältere Einstellung gewährleistet die Kompatibilität mit älteren Spark-Versionen, die negative Dezimalskalen unterstützten.	

Eigenschaftensname	Description	Standardwert
spark.sql.legacy.castComplexTypesToString.aktiviert	Gibt an, ob veraltetes Verhalten für die Umwandlung komplexer Typen in Zeichenketten aktiviert werden soll. Behält die Kompatibilität mit den Typkonvertierungsregeln älterer Spark-Versionen bei.	
spark.sql.legacy.charVarcharAsString.enfolge	Gibt an, ob die Typen CHAR und VARCHAR als STRING-Typen behandelt werden sollen. Diese Legacy-Einstellung bietet Kompatibilität mit der Verarbeitung von String-Typen in älteren Spark-Versionen.	
spark.sql.legacy.createEmptyCollectionUsingStringType	Gibt an, ob leere Sammlungen mithilfe von Zeichenkettenelementen erstellt werden sollen. Diese Legacy-Einstellung gewährleistet die Kompatibilität mit dem Verhalten älterer Spark-Versionen bei der Initialisierung von Sammlungen.	
spark.sql.legacy.exponentLiteralAsDecimal. Aktiviert	Gibt an, ob exponentielle Literale als Dezimaltypen interpretiert werden sollen. Diese Legacy-Einstellung gewährleistet die Kompatibilität mit der numerischen Literalverarbeitung älterer Spark-Versionen.	
spark.sql.legacy.json.allowEmptyString. aktiviert	Gibt an, ob leere Zeichenfolgen bei der JSON-Verarbeitung zulässig sind. Diese Legacy-Einstellung gewährleistet die Kompatibilität mit dem JSON-Parsing-Verhalten älterer Spark-Versionen.	

Eigenschaftename	Description	Standardwert
spark.sql.legacy.parquet.int96RebaseModelRead	Gibt an, ob beim Lesen von Parquet-Dateien der alte Timestamp-Rebase-Modus verwendet werden soll. INT96 Diese Legacy-Einstellung gewährleistet die Kompatibilität mit der Zeitstempelverarbeitung älterer Spark-Versionen.	
spark.sql.legacy.timeParserPolicy	Steuert das Zeitanalyseverhalten aus Gründen der Abwärtskompatibilität. Diese ältere Einstellung bestimmt, wie Zeitstempel und Datumsangaben anhand von Zeichenketten analysiert werden.	
spark.sql.Legacy.TypeCoercion.datetimeToString.aktiviert	Gibt an, ob bei der Konvertierung von Datetime-Werten in Zeichenfolgen das Zwangsverhalten älterer Typen aktiviert werden soll. Behält die Kompatibilität mit den Datetime-Konvertierungsregeln älterer Spark-Versionen bei.	
spark.sql. maxSinglePartitionByte	Legt die maximale Partitionsgröße in Byte fest. Der Planer führt Shuffle-Operationen für größere Partitionen ein, um die Parallelität zu verbessern.	128 m

Eigenschaftensname	Description	Standardwert
spark.sql.MetadataCache TTLSeconds	Steuert die (TTL) für time-to-live Metadaten-Caches. Gilt für Partitionsdatei-Metadaten und Sitzungskatalog-Caches. Erfordert: • Ein positiver Wert größer als Null • spark.sql.CatalogImplementation ist auf Hive gesetzt • spark.sql.hive.filesourcePartitionFileCacheSize größer als Null • spark.sql.hive.manageFilesourcePartitions auf true gesetzt	-1000 ms
spark.sql.optimizer.collapseProjectAlwaysInline	Gibt an, ob benachbarte Projektionen und Inline-Ausdrücke ausgeblendet werden sollen, auch wenn dies zu Duplikaten führt.	FALSE
spark.sql.optimizer.dynamicPartitionPruning. aktiviert	Gibt an, ob Prädikate für Partitionsspalten generiert werden sollen, die als Join-Schlüssel verwendet werden.	TRUE
spark.sql.optimizer.enableCsvExpressionOptimierung	Gibt an, ob CSV-Ausdrücke im SQL-Optimizer optimiert werden sollen, indem unnötige Spalten aus from_csv-Vorgängen entfernt werden.	TRUE
spark.sql.optimizer.enableJsonExpressionOptimierung	Gibt an, ob JSON-Ausdrücke im SQL-Optimizer wie folgt optimiert werden sollen: • Löschen unnötiger Spalten aus from_json-Vorgängen • Vereinfachung der Kombinationen from_json und to_json • Optimierung von named_struct-Operationen	TRUE

Eigenschaftename	Description	Standardwert
spark.sql.Optimizer.ExcludedRules	Definiert zu deaktivierende Optimizer-Regeln, identifiziert durch kommasetrennte Regelnamen. Einige Regeln können nicht deaktiviert werden, da sie aus Gründen der Richtigkeit erforderlich sind. Der Optimizer protokolliert, welche Regeln erfolgreich deaktiviert wurden.	(Keine)
spark.sql.optimizer.runtime.BloomFilter.applicationSideScanSizeThreshold	Legt die minimale aggregierte Scangröße in Byte fest, die erforderlich ist, um einen Bloom-Filter auf der Anwendungsseite einzufügen.	10 GB
spark.sql.Optimizer.Runtime.BloomFilter.creationSideThreshold	Definiert den maximalen Größenschwellenwert für die Injektion eines Bloom-Filters auf der Erstellungsseite.	10 MB
spark.sql.Optimizer.Runtime.BloomFilter.Enabled	Gibt an, ob ein Bloom-Filter eingefügt werden soll, um Shuffle-Daten zu reduzieren, wenn eine Seite einer Shuffle-Verknüpfung über ein selektives Prädikat verfügt.	TRUE
spark.sql.optimizer.runtime.BloomFilter.expectedNumItems	Definiert die Standardanzahl erwarteter Elemente im Runtime-Bloom-Filter.	1000000
spark.sql.optimizer.runtime.BloomFilter.maxNumBits	Legt die maximale Anzahl von Bits fest, die im Runtime-Bloom-Filter zulässig sind.	67108864
spark.sql.optimizer.runtime.BloomFilter.maxNumItems	Legt die maximale Anzahl erwarteter Elemente fest, die im Runtime-Bloom-Filter zulässig sind.	4000000

Eigenschaftename	Description	Standardwert
spark.sql.optimize r.runtime.BloomFilter.Number.Threshold	Schränkt die maximale Anzahl von Nicht-DPP-Laufzeitfiltern pro Abfrage ein, um Fehler im Treiber zu verhindern. out-of-memory	10
spark.sql.optimize r.runtime.BloomFilter.NumBits	Definiert die Standardanzahl von Bits, die im Runtime-Bloom-Filter verwendet werden.	8388608
spark.sql.optimize r.runtime.rowlevelOperationGroupFilter. Aktiviert	Gibt an, ob die Laufzeitgruppenfilterung für Operationen auf Zeilenebene aktiviert werden soll. Ermöglicht Datenquellen: • Löschen ganzer Datengruppen (wie Dateien oder Partitionen) mithilfe von Datenquellenfiltern • Führen Sie Laufzeitabfragen aus, um übereinstimmende Datensätze zu identifizieren • Verwerfen Sie unnötige Gruppen, um teure Neuschreibungen zu vermeiden Einschränkungen: • Nicht alle Ausdrücke können in Datenquellenfilter konvertiert werden • Einige Ausdrücke erfordern eine Spark-Auswertung (z. B. Unterabfragen)	TRUE
spark.sql.Optimize r.RuntimeFilter. semiJoinReduction. aktiviert	Gibt an, ob ein Semi-Join eingefügt werden soll, um die Anzahl der Shuffle-Daten zu reduzieren, wenn eine Seite einer Shuffle-Verknüpfung über ein selektives Prädikat verfügt.	FALSE

Eigenschaftensname	Description	Standardwert
spark.sql.parquet. AggregatePushDown	Gibt an, ob Aggregate zur Optimierung nach Parquet übertragen werden sollen. Unterstützt: • MIN und MAX für die Typen Boolean, Integer, Float und Date • COUNT für alle Datentypen Löst eine Ausnahme aus, wenn Statistiken in einer Fußzeile einer Parquet-Datei fehlen.	FALSE
spark.sql.parquet. columnarReaderBatchGröße	Steuert die Anzahl der Zeilen in jedem Stapel von vektorisierten Parquet-Readern. Wählen Sie einen Wert, der Leistungsaufwand und Speichernutzung ausbalanciert, um Fehler zu vermeiden out-of-memory.	4096
spark.sql.session. TimeZone	Definiert die Sitzungszeitzone für die Verarbeitung von Zeitstempeln in Zeichenfolgenliteralen und die Konvertierung von Java-Objekten. Akzeptiert: • Regionsbasiertes IDs area/city Format (z. B. America/Los_Angeles) • Zonenversätze im Format (+/-) HH, (+/-) HH:mm oder (+/-) HH:mm:SS (z. B. -08 oder + 01:00) • UTC oder Z als Aliase für + 00:00	(Wert der lokalen Zeitzone)
spark.sql.shuffle. partitions	Legt die Standardanzahl von Partitionen für das Mischen von Daten bei Verknüpfungen oder Aggregationen fest. Kann zwischen Neustarts strukturierter Streaming-Abfragen von derselben Checkpoint-Position aus nicht geändert werden.	200

Eigenschaftename	Description	Standardwert
spark.sql.shuffledHashJoinFaktor	Definiert den Multiplikationsfaktor, der verwendet wird, um die Eignung für einen Shuffle-Hash-Join zu bestimmen. Ein Shuffle-Hash-Join wird ausgewählt, wenn die kleine Datengröße multipliziert mit diesem Faktor kleiner ist als die große Datengröße.	3
spark.sql.sources.parallelPartitionDiscovery.Schwellenwert	Legt die maximale Anzahl von Pfaden für die treiberseitige Dateiauflistung mit dateibasierten Quellen (Parquet, JSON und ORC) fest. Wenn sie bei der Partitionserkennung überschritten werden, werden Dateien mithilfe eines separaten verteilten Spark-Jobs aufgelistet.	32
spark.sql.statistics.histogram.enabled	Gibt an, ob bei der Berechnung der Spaltenstatistiken Histogramme mit gleicher Höhe generiert werden sollen, um die Schätzgenauigkeit zu verbessern. Erfordert einen zusätzlichen Tabellenscan, der über den für einfache Spaltenstatistiken erforderlichen hinausgeht.	FALSE

9. Wählen Sie für Dienstzugriff den Namen der vorhandenen Servicerolle aus, der für den Zugriff auf diese Tabelle verwendet werden soll.
10. Wenn Sie Tags für den Trainingsdatensatz aktivieren möchten, wählen Sie Neues Tag hinzufügen und geben Sie dann das Schlüssel - und Wertepaar ein.
11. Wählen Sie Lookalike-Segment erstellen aus.

Die entsprechende API-Aktion finden Sie unter [StartAudienceGenerationJob](#).

Exportieren eines Lookalike-Segments

Nachdem Sie ein Lookalike-Segment erstellt haben, können Sie diese Daten in einen Amazon S3 S3-Bucket exportieren.

Um ein Lookalike-Segment zu exportieren in AWS Clean Rooms

1. Melden Sie sich bei der an AWS-Managementkonsole und öffnen Sie die [AWS Clean Rooms Konsole](#) mit Ihrem AWS-Konto (falls Sie dies noch nicht getan haben).
2. Wählen Sie im linken Navigationsbereich Collaborations aus.
3. Wählen Sie auf der Registerkarte Mit aktiver Mitgliedschaft eine Kollaboration aus.
4. Wählen Sie auf der Registerkarte ML-Modelle ein Lookalike-Segment aus und klicken Sie auf Exportieren.
5. Geben Sie für Lookalike-Modell exportieren unter Details des Lookalike-Modells exportieren einen Namen und optional eine Beschreibung ein.
6. Wählen Sie unter Segmentgröße die gewünschte Größe für das exportierte Segment aus.
7. Wählen Sie Export aus.

Die entsprechende API-Aktion finden Sie unter [StartAudienceExportJob](#).

Problembehebung AWS Clean Rooms

In diesem Abschnitt werden einige häufig auftretende Probleme beschrieben, die bei der Verwendung auftreten können, AWS Clean Rooms und deren Behebung.

Problembereiche

- [Auf eine oder mehrere Tabellen, auf die in der Abfrage verwiesen wird, kann über die zugehörige Dienstrolle nicht zugegriffen werden. Der table/role Besitzer muss der Servicerolle Zugriff auf die Tabelle gewähren.](#)
- [Einer der zugrunde liegenden Datensätze hat ein nicht unterstütztes Dateiformat.](#)
- [Die Abfrageergebnisse entsprechen nicht den Erwartungen, wenn Sie Cryptographic Computing for Clean Rooms verwenden.](#)
- [AWS Clean Rooms Spark SQL: Fehlende Partitionsdaten](#)

Auf eine oder mehrere Tabellen, auf die in der Abfrage verwiesen wird, kann über die zugehörige Dienstrolle nicht zugegriffen werden. Der table/role Besitzer muss der Servicerolle Zugriff auf die Tabelle gewähren.

- Stellen Sie sicher, dass die Berechtigungen für die Servicerolle wie erforderlich eingerichtet sind. Weitere Informationen finden Sie unter [Einrichten AWS Clean Rooms](#).

Einer der zugrunde liegenden Datensätze hat ein nicht unterstütztes Dateiformat.

- Stellen Sie sicher, dass Ihr Datensatz in einem der unterstützten Dateiformate vorliegt:
 - Parquet
 - RCFile
 - TextFile
 - SequenceFile
 - RegexSerde

- OpenCSV
- AVRO
- JSON

Weitere Informationen finden Sie unter [Datenformate für AWS Clean Rooms](#).

Die Abfrageergebnisse entsprechen nicht den Erwartungen, wenn Sie Cryptographic Computing for Clean Rooms verwenden.

Wenn Sie Cryptographic Computing for Clean Rooms (C3R) verwenden, stellen Sie sicher, dass Ihre Abfrage verschlüsselte Spalten korrekt verwendet:

- Die sealed Spalten werden nur in SELECT Klauseln verwendet.
- Die fingerprint Spalten werden nur in JOIN Klauseln (und GROUP BY Klauseln unter bestimmten Bedingungen) verwendet.
- Dass Sie nur JOINing fingerprint Spalten mit demselben Namen sind, wenn die Einstellungen für die Zusammenarbeit dies erfordern.

Weitere Informationen erhalten Sie unter [the section called “Kryptografisches Rechnen”](#) und [the section called “Spaltentypen”](#).

AWS Clean Rooms Spark SQL: Fehlende Partitionsdaten

Alle Partitionen in der AWS Glue Data Catalog müssen auch Daten in S3 enthalten. Die Engine verwendet die Spark-Einstellung `spark.sql.files.ignoreMissingFiles=False`

Weitere Informationen finden Sie unter <https://spark.apache.org/docs/latest/sql-data-sources-generic-options.html#ignore-missing-files>.

Wenn Sie auf diesen Fehler stoßen, erhalten Sie die folgende Fehlermeldung: "Missing partition data: One of the configured tables is partitioned and one or more of the partitions does not have data".

Vergleichen Sie Ihre in Amazon S3 vorhandenen Daten mit den in der AWS Glue Data Catalog Tabelle aufgeführten Partitionen. Löschen Sie Partitionen, die keine entsprechenden Daten in S3 enthalten.

Sicherheit in AWS Clean Rooms

Cloud-Sicherheit AWS hat höchste Priorität. Als AWS Kunde profitieren Sie von einer Rechenzentrums- und Netzwerkarchitektur, die darauf ausgelegt sind, die Anforderungen der sicherheitssensibelsten Unternehmen zu erfüllen.

Sicherheit ist eine gemeinsame Verantwortung von Ihnen AWS und Ihnen. Das [Modell der übergreifenden Verantwortlichkeit](#) beschreibt dies als Sicherheit der Cloud und Sicherheit in der Cloud:

- Sicherheit der Cloud — AWS ist verantwortlich für den Schutz der Infrastruktur, die AWS Dienste in der AWS Cloud ausführt. AWS bietet Ihnen auch Dienste, die Sie sicher nutzen können. Externe Prüfer testen und verifizieren regelmäßig die Wirksamkeit unserer Sicherheitsmaßnahmen im Rahmen der [AWS](#). Weitere Informationen zu den Compliance-Programmen, für die gelten AWS Clean Rooms, finden Sie unter [AWS-Services in Umfang nach Compliance-Programm](#).
- Sicherheit in der Cloud — Ihre Verantwortung richtet sich nach dem AWS Service, den Sie nutzen. Sie sind auch für andere Faktoren verantwortlich, etwa für die Vertraulichkeit Ihrer Daten, für die Anforderungen Ihres Unternehmens und für die geltenden Gesetze und Vorschriften.

Diese Dokumentation hilft Ihnen zu verstehen, wie Sie das Modell der gemeinsamen Verantwortung bei der Nutzung anwenden können AWS Clean Rooms. Es zeigt Ihnen, wie Sie die Konfiguration vornehmen AWS Clean Rooms, um Ihre Sicherheits- und Compliance-Ziele zu erreichen. Sie erfahren auch, wie Sie andere AWS Dienste nutzen können, die Sie bei der Überwachung und Sicherung Ihrer AWS Clean Rooms Ressourcen unterstützen.

Inhalt

- [Datenschutz in AWS Clean Rooms](#)
- [Verwenden von serviceverknüpften Rollen für AWS Clean Rooms](#)
- [Aufbewahrung von Daten in AWS Clean Rooms](#)
- [Bewährte Methoden für die Zusammenarbeit bei Daten in AWS Clean Rooms](#)
- [Identity and Access Management für AWS Clean Rooms](#)
- [Konformitätsprüfung für AWS Clean Rooms](#)
- [Resilienz in AWS Clean Rooms](#)
- [Sicherheit der Infrastruktur in AWS Clean Rooms](#)

- [Zugriff AWS Clean Rooms oder AWS Clean Rooms ML über einen Schnittstellen-Endpunkt \(\)AWS PrivateLink](#)

Datenschutz in AWS Clean Rooms

Das [Modell der AWS gemeinsamen Verantwortung](#) und geteilter Verantwortung gilt für den Datenschutz in AWS Clean Rooms. Wie in diesem Modell beschrieben, AWS ist verantwortlich für den Schutz der globalen Infrastruktur, auf der alle Systeme laufen AWS Cloud. Sie sind dafür verantwortlich, die Kontrolle über Ihre in dieser Infrastruktur gehosteten Inhalte zu behalten. Sie sind auch für die Sicherheitskonfiguration und die Verwaltungsaufgaben für die von Ihnen verwendeten AWS-Services verantwortlich. Weitere Informationen zum Datenschutz finden Sie unter [Häufig gestellte Fragen zum Datenschutz](#). Informationen zum Datenschutz in Europa finden Sie im Blog-Beitrag [AWS -Modell der geteilten Verantwortung und in der DSGVO](#) im AWS -Sicherheitsblog.

Aus Datenschutzgründen empfehlen wir, dass Sie AWS-Konto Anmeldeinformationen schützen und einzelne Benutzer mit AWS IAM Identity Center oder AWS Identity and Access Management (IAM) einrichten. So erhält jeder Benutzer nur die Berechtigungen, die zum Durchführen seiner Aufgaben erforderlich sind. Außerdem empfehlen wir, die Daten mit folgenden Methoden schützen:

- Verwenden Sie für jedes Konto die Multi-Faktor-Authentifizierung (MFA).
- Wird verwendet SSL/TLS , um mit AWS Ressourcen zu kommunizieren. Wir benötigen TLS 1.2 und empfehlen TLS 1.3.
- Richten Sie die API und die Protokollierung von Benutzeraktivitäten mit ein AWS CloudTrail. Informationen zur Verwendung von CloudTrail Pfaden zur Erfassung von AWS Aktivitäten finden Sie unter [Arbeiten mit CloudTrail Pfaden](#) im AWS CloudTrail Benutzerhandbuch.
- Verwenden Sie AWS Verschlüsselungslösungen zusammen mit allen darin enthaltenen Standardsicherheitskontrollen AWS-Services.
- Verwenden Sie erweiterte verwaltete Sicherheitsservices wie Amazon Macie, die dabei helfen, in Amazon S3 gespeicherte persönliche Daten zu erkennen und zu schützen.
- Wenn Sie für den Zugriff AWS über eine Befehlszeilenschnittstelle oder eine API FIPS 140-3-validierte kryptografische Module benötigen, verwenden Sie einen FIPS-Endpunkt. Weitere Informationen über verfügbare FIPS-Endpunkte finden Sie unter [Federal Information Processing Standard \(FIPS\) 140-3](#).

Wir empfehlen dringend, in Freitextfeldern, z. B. im Feld Name, keine vertraulichen oder sensiblen Informationen wie die E-Mail-Adressen Ihrer Kunden einzugeben. Dies gilt auch, wenn Sie mit

der Konsole, der AWS Clean Rooms API oder auf andere AWS-Services Weise arbeiten oder diese verwenden. AWS CLI AWS SDKs Alle Daten, die Sie in Tags oder Freitextfelder eingeben, die für Namen verwendet werden, können für Abrechnungs- oder Diagnoseprotokolle verwendet werden. Wenn Sie eine URL für einen externen Server bereitstellen, empfehlen wir dringend, keine Anmeldeinformationen zur Validierung Ihrer Anforderung an den betreffenden Server in die URL einzuschließen.

Verschlüsselung im Ruhezustand

AWS Clean Rooms verschlüsselt immer alle Dienstmetadaten im Ruhezustand, ohne dass eine zusätzliche Konfiguration erforderlich ist. Diese Verschlüsselung erfolgt automatisch, wenn Sie sie verwenden AWS Clean Rooms.

Clean Rooms ML verschlüsselt alle im Service gespeicherten Daten im Ruhezustand mit AWS KMS. Wenn Sie Ihren eigenen KMS-Schlüssel angeben, werden die Inhalte Ihrer Lookalike-Modelle und Jobs zur Generierung von Lookalike-Segmenten im Ruhezustand mit Ihrem KMS-Schlüssel verschlüsselt.

Wenn Sie AWS Clean Rooms benutzerdefinierte ML-Modelle verwenden, verschlüsselt der Dienst alle Daten, die im Ruhezustand gespeichert sind, mit. AWS KMS AWS Clean Rooms unterstützt die Verwendung symmetrischer, vom Kunden verwalteter Schlüssel, die Sie erstellen, besitzen und verwalten, um Daten im Ruhezustand zu verschlüsseln. Wenn vom Kunden verwaltete Schlüssel nicht angegeben AWS-eigene Schlüssel sind, werden diese standardmäßig verwendet.

AWS Clean Rooms verwendet Zuschüsse und wichtige Richtlinien für den Zugriff auf vom Kunden verwaltete Schlüssel. Sie können den Zugriff auf die Genehmigung jederzeit widerrufen oder den Zugriff des Services auf den vom Kunden verwalteten Schlüssel entfernen. Wenn Sie dies tun, können Sie auf AWS Clean Rooms keine der mit dem vom Kunden verwalteten Schlüssel verschlüsselten Daten zugreifen, was sich auf Vorgänge auswirkt, die von diesen Daten abhängig sind. Wenn Sie beispielsweise versuchen, ein trainiertes Modell aus einem verschlüsselten ML-Eingabekanal zu erstellen, AWS Clean Rooms auf den kein Zugriff möglich ist, würde der Vorgang einen `ValidationException` Fehler zurückgeben.

Note

Sie können die Verschlüsselungsoptionen in Amazon S3 verwenden, um Ihre Daten im Ruhezustand zu schützen.

Weitere Informationen finden Sie unter [Angabe der Amazon S3 S3-Verschlüsselung](#) im Amazon S3 S3-Benutzerhandbuch.

Wenn Sie darin eine ID-Zuordnungstabelle verwenden AWS Clean Rooms, verschlüsselt der Service alle Daten, die im Ruhezustand gespeichert sind, mit AWS KMS. Wenn Sie Ihren eigenen KMS-Schlüssel angeben, wird der Inhalt Ihrer ID-Zuordnungstabelle im Ruhezustand mit Ihrem KMS-Schlüssel über AWS Entity Resolution verschlüsselt. Weitere Informationen zu den erforderlichen Berechtigungen für die Arbeit mit Verschlüsselungen mit einem ID-Mapping-Workflow finden Sie unter [Erstellen einer Workflow-Jobrolle für AWS Entity Resolution](#) im AWS Entity Resolution Benutzerhandbuch.

Verschlüsselung während der Übertragung

AWS Clean Rooms verwendet Transport Layer Security (TLS) für die Verschlüsselung bei der Übertragung. Die Kommunikation mit AWS Clean Rooms erfolgt immer über HTTPS, sodass Ihre Daten bei der Übertragung immer verschlüsselt werden, unabhängig davon, ob sie in Amazon S3, Amazon Athena oder Snowflake gespeichert sind. Dies schließt alle Daten ein, die bei der Verwendung von Clean Rooms ML übertragen werden.

Verschlüsselung der zugrunde liegenden Daten

Weitere Hinweise zur Verschlüsselung der zugrunde liegenden Daten finden Sie unter [Kryptografisches Rechnen für Clean Rooms](#)

Schlüsselrichtlinie

Schlüsselrichtlinien steuern den Zugriff auf den vom Kunden verwalteten Schlüssel. Jeder vom Kunden verwaltete Schlüssel muss über genau eine Schlüsselrichtlinie verfügen, die aussagt, wer den Schlüssel wie verwenden kann. Wenn Sie Ihren kundenseitig verwalteten Schlüssel erstellen, können Sie eine Schlüsselrichtlinie angeben. Weitere Informationen finden Sie im AWS Key Management Service Entwicklerhandbuch unter Verwaltung des Zugriffs auf vom Kunden verwaltete Schlüssel.

Um Ihren vom Kunden verwalteten Schlüssel mit Ihren AWS Clean Rooms benutzerdefinierten ML-Modellen zu verwenden, müssen die folgenden API-Operationen in der Schlüsselrichtlinie zulässig sein:

- `kms:DescribeKey`— Stellt dem Kunden verwaltete Schlüsseldetails zur Verfügung, damit AWS Clean Rooms der Schlüssel validiert werden kann.
- `kms:Decrypt`— Ermöglicht den Zugriff auf die verschlüsselten Daten AWS Clean Rooms , um sie zu entschlüsseln und sie für verwandte Aufgaben zu verwenden.
- `kms:CreateGrant`- Clean Rooms ML verschlüsselt in Amazon ECR gespeicherte Schulungs- und Inferenzbilder, indem Zuschüsse für Amazon ECR erstellt werden. Weitere Informationen finden Sie unter [Verschlüsselung im Ruhezustand in Amazon ECR](#). Clean Rooms ML verwendet Amazon SageMaker AI auch zur Ausführung von Trainings- und Inferenzjobs und erstellt Zuschüsse für SageMaker KI, um die an die Instances angehängten Amazon EBS-Volumes sowie die Ausgabedaten in Amazon S3 zu verschlüsseln. Weitere Informationen finden Sie unter [Schützen von Daten im Ruhezustand mithilfe von Verschlüsselung in Amazon SageMaker AI](#).
- `kms:GenerateDataKey`- Clean Rooms ML verschlüsselt Daten im Ruhezustand, die in Amazon S3 gespeichert sind, mithilfe serverseitiger Verschlüsselung mit AWS KMS keys Weitere Informationen finden Sie unter [Serverseitige Verschlüsselung mit AWS KMS keys \(SSE-KMS\) in Amazon S3 verwenden](#).

Im Folgenden finden Sie Beispiele für Richtlinienerklärungen, die Sie AWS Clean Rooms für die folgenden Ressourcen hinzufügen können:

ML-Eingangskanal mit synthetischen Daten

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "Allow access to principals authorized to use AWS Clean Rooms ML",
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::666666666666:role/ExampleRole"
      },
      "Action": [
        "kms:GenerateDataKey",
        "kms:CreateGrant",
        "kms:Decrypt"
      ],
      "Resource": "*",
      "Condition": {
        "StringEquals": {
          "kms:ViaService": "cleanrooms-ml.us-east-1.amazonaws.com"
        }
      }
    }
  ]
}
```



```

    },
    "ForAllValues:StringEquals": {
        "kms:GrantOperations": [
            "Decrypt",
            "Encrypt",
            "GenerateDataKeyWithoutPlaintext",
            "ReEncryptFrom",
            "ReEncryptTo",
            "CreateGrant",
            "DescribeKey",
            "RetireGrant",
            "GenerateDataKey"
        ]
    },
    "BoolIfExists": {
        "kms:GrantIsForAWSResource": true
    }
},
{
    "Sid": "Allow describe key for principals authorized to use AWS Clean Rooms
ML",
    "Effect": "Allow",
    "Principal": {
        "AWS": "arn:aws:iam::444455556666:role/ExampleRole"
    },
    "Action": [
        "kms:DescribeKey"
    ],
    "Resource": "*",
    "Condition": {
        "StringEquals": {
            "kms:ViaService": "cleanrooms-ml.us-east-1.amazonaws.com"
        }
    }
},
{
    "Sid": "Allow grant operations for AWS Clean Rooms ML service principal",
    "Effect": "Allow",
    "Principal": {
        "Service": "cleanrooms-ml.amazonaws.com"
    },
    "Action": [
        "kms:GenerateDataKey",

```

```

        "kms:CreateGrant",
        "kms:Decrypt"
    ],
    "Resource": "*",
    "Condition": {
        "ForAllValues:StringEquals": {
            "kms:GrantOperations": [
                "Decrypt",
                "Encrypt",
                "GenerateDataKeyWithoutPlaintext",
                "ReEncryptFrom",
                "ReEncryptTo",
                "CreateGrant",
                "DescribeKey",
                "RetireGrant",
                "GenerateDataKey"
            ]
        }
    }
},
{
    "Sid": "Allow describe key for AWS Clean Rooms ML service principal",
    "Effect": "Allow",
    "Principal": {
        "Service": "cleanrooms-ml.amazonaws.com"
    },
    "Action": [
        "kms:DescribeKey"
    ],
    "Resource": "*"
}
]
}

```

ML-Eingangskanal ohne synthetische Daten

JSON

```

{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "Allow access to principals authorized to use AWS Clean Rooms ML",

```

```

    "Effect": "Allow",
    "Principal": {
      "AWS": "arn:aws:iam::666666666666:role/ExampleRole"
    },
    "Action": [
      "kms:DescribeKey",
      "kms:GenerateDataKey",
      "kms:Decrypt"
    ],
    "Resource": "*",
    "Condition": {
      "StringEquals": {
        "kms:ViaService": "cleanrooms-ml.us-east-1.amazonaws.com"
      }
    }
  },
  {
    "Sid": "Allow access to AWS Clean Rooms ML service principal",
    "Effect": "Allow",
    "Principal": {
      "Service": "cleanrooms-ml.amazonaws.com"
    },
    "Action": [
      "kms:DescribeKey",
      "kms:GenerateDataKey",
      "kms:Decrypt"
    ],
    "Resource": "*"
  }
]
}

```

Trainierter Modelljob oder Job mit trainierter Modellinferenz

JSON

```

{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "Allow grant operations for principals authorized to use AWS  
Clean Rooms ML",

```

```

    "Effect": "Allow",
    "Principal": {
      "AWS": "arn:aws:iam::666666666666:role/ExampleRole"
    },
    "Action": [
      "kms:GenerateDataKey",
      "kms:CreateGrant",
      "kms:Decrypt"
    ],
    "Resource": "*",
    "Condition": {
      "StringEquals": {
        "kms:ViaService": "cleanrooms-ml.us-east-1.amazonaws.com"
      },
      "ForAllValues:StringEquals": {
        "kms:GrantOperations": [
          "Decrypt",
          "Encrypt",
          "GenerateDataKeyWithoutPlaintext",
          "ReEncryptFrom",
          "ReEncryptTo",
          "CreateGrant",
          "DescribeKey",
          "RetireGrant",
          "GenerateDataKey"
        ]
      },
      "BoolIfExists": {
        "kms:GrantIsForAWSResource": true
      }
    }
  },
  {
    "Sid": "Allow describe key for principals authorized to use AWS Clean Rooms ML",
    "Effect": "Allow",
    "Principal": {
      "AWS": "arn:aws:iam::444455556666:role/ExampleRole"
    },
    "Action": [
      "kms:DescribeKey"
    ],
    "Resource": "*",
    "Condition": {

```

```

        "StringEquals": {
            "kms:ViaService": "cleanrooms-ml.us-east-1.amazonaws.com"
        }
    },
    {
        "Sid": "Allow grant operations for AWS Clean Rooms ML service principal",
        "Effect": "Allow",
        "Principal": {
            "Service": "cleanrooms-ml.amazonaws.com"
        },
        "Action": [
            "kms:GenerateDataKey",
            "kms:CreateGrant",
            "kms:Decrypt"
        ],
        "Resource": "*",
        "Condition": {
            "ForAllValues:StringEquals": {
                "kms:GrantOperations": [
                    "Decrypt",
                    "Encrypt",
                    "GenerateDataKeyWithoutPlaintext",
                    "ReEncryptFrom",
                    "ReEncryptTo",
                    "CreateGrant",
                    "DescribeKey",
                    "RetireGrant",
                    "GenerateDataKey"
                ]
            }
        }
    },
    {
        "Sid": "Allow describe key for AWS Clean Rooms ML service principal",
        "Effect": "Allow",
        "Principal": {
            "Service": "cleanrooms-ml.amazonaws.com"
        },
        "Action": [
            "kms:DescribeKey"
        ],
        "Resource": "*"
    }
}

```

```
}  
]  
}
```

Clean Rooms ML unterstützt nicht die Angabe des Dienstverschlüsselungskontextes oder des Quellkontextes in vom Kunden verwalteten Schlüsselrichtlinien. Der vom Service intern verwendete Verschlüsselungskontext ist für Kunden in sichtbar CloudTrail.

Verwenden von serviceverknüpften Rollen für AWS Clean Rooms

AWS Clean Rooms verwendet AWS Identity and Access Management (IAM) [serviceverknüpfte](#) Rollen. Eine serviceverknüpfte Rolle ist ein einzigartiger Typ von IAM-Rolle, mit der direkt verknüpft ist. AWS Clean Rooms mit Diensten verknüpfte Rollen sind vordefiniert AWS Clean Rooms und enthalten alle Berechtigungen, die der Dienst benötigt, um andere AWS Dienste in Ihrem Namen aufzurufen.

Eine dienstbezogene Rolle AWS Clean Rooms erleichtert die Einrichtung, da Sie die erforderlichen Berechtigungen nicht manuell hinzufügen müssen. AWS Clean Rooms definiert die Berechtigungen ihrer dienstbezogenen Rollen und AWS Clean Rooms kann, sofern nicht anders definiert, nur ihre Rollen übernehmen. Die definierten Berechtigungen umfassen die Vertrauens- und Berechtigungsrichtlinie. Diese Berechtigungsrichtlinie kann keinen anderen IAM-Entitäten zugewiesen werden.

Sie können eine serviceverknüpfte Rolle erst löschen, nachdem ihre verwandten Ressourcen gelöscht wurden. Dadurch werden Ihre AWS Clean Rooms Ressourcen geschützt, da Sie nicht versehentlich die Zugriffsberechtigung für die Ressourcen entziehen können.

Informationen zu anderen Diensten, die dienstverknüpfte Rollen unterstützen, finden Sie unter [AWS Dienste, die mit IAM funktionieren](#). Suchen Sie in der Spalte Dienstverknüpfte Rollen nach den Diensten, für die Ja steht. Wählen Sie über einen Link Ja aus, um die Dokumentation zu einer serviceverknüpften Rolle für diesen Service anzuzeigen.

Berechtigungen für dienstverknüpfte Rollen AWS Clean Rooms

AWS Clean Rooms verwendet die serviceverknüpfte Rolle mit dem Namen „AWSServiceRoleForAWSCleanRooms“, um CloudWatch Metriken im Zusammenhang mit Clean Rooms in Ihrem Konto zu veröffentlichen. AWS

Die `AWSServiceRoleForAWSCleanRooms`-Rolle vertraut darauf, dass die folgenden Dienste die Rolle übernehmen:

- `cleanrooms.amazonaws.com`

Die genannte Richtlinie für Rollenberechtigungen `AWSCleanRoomsServiceRolePolicy` AWS Clean Rooms ermöglicht es, die folgenden Aktionen für die angegebenen Ressourcen durchzuführen:

- Aktion: `cloudwatch:PutMetricData` für all AWS resources, restricted to the AWS Clean Rooms namespace

Sie müssen Berechtigungen konfigurieren, damit eine Benutzer, Gruppen oder Rollen eine serviceverknüpfte Rolle erstellen, bearbeiten oder löschen können. Weitere Informationen finden Sie unter [serviceverknüpfte Rollenberechtigung](#) im IAM-Benutzerhandbuch.

Erstellen einer dienstbezogenen Rolle für AWS Clean Rooms

Sie können die IAM-Konsole verwenden, um eine serviceverknüpfte Rolle mit dem Anwendungsfall `AWSServiceRoleForAWSCleanRooms` zu erstellen. Erstellen Sie in der AWS CLI oder der AWS API eine dienstverknüpfte Rolle mit dem `cleanrooms.amazonaws.com` Dienstenamen. Weitere Informationen finden Sie unter [Erstellen einer serviceverknüpften Rolle](#) im IAM-Benutzerhandbuch. Wenn Sie diese serviceverknüpfte Rolle löschen, können Sie mit demselben Verfahren die Rolle erneut erstellen.

Bearbeiten einer dienstbezogenen Rolle für AWS Clean Rooms

AWS Clean Rooms erlaubt es Ihnen nicht, die serviceverknüpfte `AWSServiceRoleForAWSCleanRooms`-Rolle zu bearbeiten. Da möglicherweise verschiedene Entitäten auf die Rolle verweisen, kann der Rollename nach dem Erstellen einer serviceverknüpften Rolle nicht mehr geändert werden. Sie können jedoch die Beschreibung der Rolle mit IAM bearbeiten. Weitere Informationen finden Sie unter [Bearbeiten einer serviceverknüpften Rolle](#) im IAM-Benutzerhandbuch.

Löschen einer serviceverknüpften Rolle für AWS Clean Rooms

Wenn Sie ein Feature oder einen Dienst, die bzw. der eine serviceverknüpften Rolle erfordert, nicht mehr benötigen, sollten Sie diese Rolle löschen. Auf diese Weise haben Sie keine ungenutzte juristische Stelle, die nicht aktiv überwacht oder verwaltet wird. Sie müssen jedoch die Ressourcen für Ihre serviceverknüpften Rolle zunächst bereinigen, bevor Sie sie manuell löschen können.

Um die `AWSServiceRoleForAWSCleanRooms` zu löschen, müssen Sie zunächst alle [Kollaborationen](#) und [Mitgliedschaften](#) in Ihrem löschen. AWS-Konto

 Note

Wenn der AWS Clean Rooms Dienst die Rolle verwendet, wenn Sie versuchen, die Ressourcen zu löschen, schlägt das Löschen möglicherweise fehl. Wenn dies passiert, warten Sie einige Minuten und versuchen Sie es erneut.

So löschen Sie die serviceverknüpfte Rolle mit IAM

Verwenden Sie die IAM-Konsole, die oder die AWS API AWS CLI, um die mit dem Dienst verknüpfte `AWSServiceRoleForAWSCleanRooms`-Rolle zu löschen. Weitere Informationen finden Sie unter [Löschen einer serviceverknüpften Rolle](#) im IAM-Benutzerhandbuch.

Unterstützte Regionen für serviceverknüpfte Rollen AWS Clean Rooms

AWS Clean Rooms unterstützt die Verwendung von dienstbezogenen Rollen in allen Regionen, in denen der Dienst verfügbar ist. Weitere Informationen finden Sie unter [AWS -Regionen und Endpunkte](#).

Aufbewahrung von Daten in AWS Clean Rooms

Alle Daten, die vorübergehend in eine AWS Clean Rooms Kollaboration eingelesen werden, werden nach Abschluss der Abfrage gelöscht.

Wenn Sie ein Lookalike-Modell erstellen, liest Clean Rooms ML Ihre Trainingsdaten, wandelt sie in ein für unser ML-Modell geeignetes Format um und speichert die trainierten Modellparameter in Clean Rooms ML. Clean Rooms ML speichert keine Kopie Ihrer Trainingsdaten. AWS Clean Rooms In SQL-Abfragen werden keine Ihrer Daten gespeichert, nachdem die Abfrage ausgeführt wurde. Clean Rooms ML verwendet dann das trainierte Modell, um das Verhalten all Ihrer Benutzer zusammenzufassen. Clean Rooms ML speichert für jeden Benutzer in Ihren Daten einen Datensatz auf Benutzerebene, solange Ihr Lookalike-Modell aktiv ist.

Wenn Sie einen Job zur Generierung von Lookalike-Segmenten starten, liest Clean Rooms ML die Ausgangsdaten, liest die Verhaltenszusammenfassungen aus dem zugehörigen Lookalike-Modell und erstellt ein Lookalike-Segment, das im Service gespeichert wird. AWS Clean Rooms Clean

Rooms ML speichert keine Kopie Ihrer Ausgangsdaten. Clean Rooms ML speichert die Ausgabe des Jobs auf Benutzerebene, solange der Job aktiv ist.

Wenn Ihre Ausgangsdaten aus einer SQL-Abfrage stammen, wird die Ausgabe dieser Abfrage nur für die Dauer des Jobs im Dienst gespeichert. Die Ergebnisse der Abfrage werden im Ruhezustand und bei der Übertragung verschlüsselt.

Wenn Sie die Auftragsdaten Ihres Lookalike-Modells oder der Generierung von Lookalike-Segmenten entfernen möchten, verwenden Sie die API, um sie zu löschen. Clean Rooms ML löscht asynchron alle mit dem Modell oder Job verknüpften Daten. Sobald dieser Vorgang abgeschlossen ist, löscht Clean Rooms ML die Metadaten für das Modell oder den Job und sie sind in der API nicht mehr sichtbar. Clean Rooms ML bewahrt gelöschte Daten 3 Tage lang auf, um eine Notfallwiederherstellung zu verhindern. Sobald der Job oder das Modell in der API nicht mehr sichtbar ist und 3 Tage vergangen sind, wurden alle mit dem Modell oder Job verknüpften Daten dauerhaft gelöscht.

Bewährte Methoden für die Zusammenarbeit bei Daten in AWS Clean Rooms

In diesem Thema werden die bewährten Methoden für die Durchführung von Datenkooperationen in beschrieben. AWS Clean Rooms

AWS Clean Rooms folgt dem [Modell der AWS gemeinsamen Verantwortung](#). AWS Clean Rooms bietet [Analyseregeln](#), die Sie konfigurieren können, um Ihre Fähigkeit zu verbessern, vertrauliche Daten in einer Zusammenarbeit zu schützen. Die Analyseregeln, in denen Sie konfigurieren, setzen die von AWS Clean Rooms Ihnen konfigurierten Einschränkungen (Abfragesteuerelemente und Abfrageausgabesteuerungen) durch. Sie sind dafür verantwortlich, die Einschränkungen festzulegen und die Analyseregeln entsprechend zu konfigurieren.

Datenkooperationen können mehr als nur Ihre Nutzung von AWS Clean Rooms beinhalten. Damit Sie den größtmöglichen Nutzen aus Datenkooperationen ziehen können, empfehlen wir Ihnen, bei der Verwendung von Analyseregeln AWS Clean Rooms und insbesondere bei der Verwendung von Analyseregeln die folgenden bewährten Methoden anzuwenden.

Themen

- [Bewährte Methoden mit AWS Clean Rooms](#)
- [Bewährte Methoden für die Verwendung von Analyseregeln in AWS Clean Rooms](#)

Bewährte Methoden mit AWS Clean Rooms

Sie sind dafür verantwortlich, das Risiko jeder Datenzusammenarbeit zu bewerten und es mit Ihren Datenschutzanforderungen wie externen und internen Compliance-Programmen und -Richtlinien zu vergleichen. Wir empfehlen Ihnen, bei der Verwendung von zusätzliche Maßnahmen zu ergreifen AWS Clean Rooms. Diese Maßnahmen können dazu beitragen, Risiken besser zu managen und vor Versuchen Dritter zu schützen, Ihre Daten neu zu identifizieren (z. B. differenzierende Angriffe oder Side-Channel-Angriffe).

Erwägen Sie beispielsweise, bei Ihren anderen Mitarbeitern eine Due-Diligence-Prüfung durchzuführen und rechtliche Vereinbarungen mit ihnen zu treffen, bevor Sie eine Zusammenarbeit eingehen. Um die Verwendung Ihrer Daten zu überwachen, sollten Sie auch die Einführung anderer Prüfmechanismen in Betracht ziehen. AWS Clean Rooms

Bewährte Methoden für die Verwendung von Analyseregeln in AWS Clean Rooms

Mit den Analyseregeln in AWS Clean Rooms können Sie die Abfragen einschränken, die ausgeführt werden können, indem Sie die Abfragesteuerelemente für eine konfigurierte Tabelle festlegen. Sie können beispielsweise eine Abfragesteuerung dafür einrichten, wie eine konfigurierte Tabelle verknüpft und welche Spalten ausgewählt werden können. Sie können die Abfrageausgabe auch einschränken, indem Sie Steuerelemente für Abfrageergebnisse festlegen, z. B. Aggregationsschwellenwerte für Ausgabezeilen. Der Dienst lehnt jede Abfrage ab und entfernt Zeilen, die nicht den Analyseregeln entsprechen, die von Mitgliedern in ihren konfigurierten Tabellen in der Abfrage festgelegt wurden.

Wir empfehlen die folgenden 10 bewährten Methoden für die Verwendung von Analyseregeln in Ihrer konfigurierten Tabelle:

- Erstellen Sie separate konfigurierte Tabellen für separate Anwendungsfälle für Abfragen (z. B. Zielgruppenplanung oder Zuordnung). Sie können mehrere konfigurierte Tabellen mit derselben zugrunde liegenden AWS Glue Tabelle erstellen.
- Geben Sie in der Analyseregeln Spalten an (z. B. Dimensionsspalten, Listenspalten, Verbindungsspalten), die für Abfragen in einer Kollaboration erforderlich sind. Dies kann dazu beitragen, das Risiko zu verringern, dass Angriffe differenziert werden oder dass andere Mitglieder Ihre Daten zurückentwickeln können. Verwenden Sie die Funktion Allowlist-Spalten, um andere Spalten zu notieren, die Sie möglicherweise in future abfragbar machen möchten. Um die Spalten

anzupassen, die für eine bestimmte Zusammenarbeit verwendet werden können, erstellen Sie zusätzliche konfigurierte Tabellen mit derselben Basistabelle. AWS Glue

- Geben Sie in der Analyseregeln die Funktionen an, die für die Analyse in der Kollaboration erforderlich sind. Dies kann dazu beitragen, das Risiko zu verringern, das durch seltene Funktionsfehler entsteht, die Informationen zu einem einzelnen Datenpunkt enthalten können. Um die Funktionen anzupassen, die für eine bestimmte Zusammenarbeit verwendet werden können, erstellen Sie zusätzliche konfigurierte Tabellen mit derselben zugrunde liegenden AWS Glue Tabelle.
- Fügen Sie Aggregationseinschränkungen für alle Spalten hinzu, deren Werte auf Zeilenebene sensibel sind. Dies schließt Spalten in Ihrer konfigurierten Tabelle ein, die auch in den Tabellen und Analyseregeln anderer Kollaborationsmitglieder als Aggregationseinschränkung vorhanden sind. Dazu gehören auch Spalten in Ihrer konfigurierten Tabelle, die nicht abfragbar sind, d. h. Spalten, die sich in Ihrer konfigurierten Tabelle befinden, aber nicht in der Analyseregeln enthalten sind. Aggregationsbeschränkungen können dazu beitragen, das Risiko zu verringern, das durch die Korrelation von Abfrageergebnissen mit Daten außerhalb der Zusammenarbeit entsteht.
- Erstellen Sie Testkollaborationen und Analyseregeln, um Einschränkungen zu testen, die mit bestimmten Analyseregeln erstellt wurden.
- Überprüfen Sie die von den Mitarbeitern konfigurierten Tabellen und die Analyseregeln der Mitglieder in den konfigurierten Tabellen, um sicherzustellen, dass sie den für die Zusammenarbeit vereinbarten Regeln entsprechen. Dies kann dazu beitragen, das Risiko zu verringern, dass andere Mitglieder ihre eigenen Daten manipulieren, um Abfragen auszuführen, die nicht vereinbart wurden.
- Sehen Sie sich die bereitgestellte Beispielabfrage (nur Konsole) an, die in Ihrer konfigurierten Tabelle aktiviert ist, nachdem Sie die Analyseregeln eingerichtet haben.

 Note

Zusätzlich zu der bereitgestellten Beispielabfrage sind weitere Abfragen möglich, die auf der Analyseregeln und anderen Tabellen und Analyseregeln für Kollaborationsmitglieder basieren.

- Sie können eine Analyseregeln für eine konfigurierte Tabelle in einer Kollaboration hinzufügen oder aktualisieren. Wenn Sie dies tun, überprüfen Sie alle Kollaborationen, denen die konfigurierte Tabelle zugeordnet ist, und die sich daraus ergebenden Auswirkungen. Auf diese Weise können Sie sicherstellen, dass keine Kollaborationen veraltete Analyseregeln verwenden.

- Überprüfen Sie die in der Kollaboration ausgeführten Abfragen, um sicherzustellen, dass die Abfragen den Anwendungsfällen oder Abfragen entsprechen, die für die Zusammenarbeit vereinbart wurden. (Die Abfragen sind in den Abfrageprotokollen verfügbar, wenn die Funktion zur Abfrageprotokollierung aktiviert ist.) Dies kann dazu beitragen, das Risiko zu verringern, dass Mitglieder Analysen durchführen, die nicht vereinbart wurden, und potenzielle Angriffe wie Seitenkanalangriffe.
- Überprüfen Sie die konfigurierten Tabellenspalten, die in den Analyseregeln der Kollaborationsmitglieder und in Abfragen verwendet werden, um sicherzustellen, dass sie den in der Zusammenarbeit vereinbarten Werten entsprechen. (Die Abfragen sind in den Abfrageprotokollen verfügbar, wenn diese Funktion aktiviert ist.) Dies kann dazu beitragen, das Risiko zu verringern, dass andere Mitglieder ihre eigenen Daten manipulieren, um Abfragen durchzuführen, über die keine Einigung erzielt wurde.

Identity and Access Management für AWS Clean Rooms

AWS Identity and Access Management (IAM) hilft einem Administrator AWS-Service, den Zugriff auf AWS Ressourcen sicher zu kontrollieren. IAM-Administratoren kontrollieren, wer authentifiziert (angemeldet) und autorisiert werden kann (über Berechtigungen verfügt), um Ressourcen zu verwenden. AWS Clean Rooms IAM ist ein Programm AWS-Service, das Sie ohne zusätzliche Kosten nutzen können.

Themen

- [Zielgruppe](#)
- [Authentifizierung mit Identitäten](#)
- [Verwalten des Zugriffs mit Richtlinien](#)
- [Wie AWS Clean Rooms funktioniert mit IAM](#)
- [Beispiele für identitätsbasierte Richtlinien für AWS Clean Rooms](#)
- [AWS verwaltete Richtlinien für AWS Clean Rooms](#)
- [Fehlerbehebung bei AWS Clean Rooms Identität und Zugriff](#)
- [Serviceübergreifende Confused-Deputy-Prävention](#)
- [IAM-Verhalten für AWS Clean Rooms ML](#)
- [IAM-Verhalten für benutzerdefinierte Clean Rooms ML-Modelle](#)

Zielgruppe

Wie Sie AWS Identity and Access Management (IAM) verwenden, hängt von Ihrer Rolle ab:

- Servicebenutzer – Fordern Sie von Ihrem Administrator Berechtigungen an, wenn Sie nicht auf Features zugreifen können (siehe [Fehlerbehebung bei AWS Clean Rooms Identität und Zugriff](#)).
- Serviceadministrator – Bestimmen Sie den Benutzerzugriff und stellen Sie Berechtigungsanfragen (siehe [Wie AWS Clean Rooms funktioniert mit IAM](#)).
- IAM-Administrator – Schreiben Sie Richtlinien zur Zugriffsverwaltung (siehe [Beispiele für identitätsbasierte Richtlinien für AWS Clean Rooms](#)).

Authentifizierung mit Identitäten

Authentifizierung ist die Art und Weise, wie Sie sich AWS mit Ihren Identitätsdaten anmelden. Sie müssen als IAM-Benutzer authentifiziert (angemeldet AWS) sein oder eine IAM-Rolle annehmen. Root-Benutzer des AWS-Kontos

Sie können sich AWS als föderierte Identität anmelden, indem Sie Anmeldeinformationen verwenden, die über eine Identitätsquelle bereitgestellt wurden. AWS IAM Identity Center (IAM Identity Center) - Benutzer oder die Single Sign-On-Authentifizierung Ihres Unternehmens sind Beispiele für föderierte Identitäten. Wenn Sie sich als Verbundidentität anmelden, hat der Administrator vorher mithilfe von IAM-Rollen einen Identitätsverbund eingerichtet. Wenn Sie über einen Verbund darauf zugreifen AWS , übernehmen Sie indirekt eine Rolle.

Je nachdem, welcher Benutzertyp Sie sind, können Sie sich beim AWS-Managementkonsole oder beim AWS Zugangsportal anmelden. Weitere Informationen zur Anmeldung finden Sie AWS unter [So melden Sie sich bei Ihrem an AWS-Konto](#) im AWS-Anmeldung Benutzerhandbuch.

Wenn Sie AWS programmgesteuert zugreifen, AWS stellt es ein Software Development Kit (SDK) und eine Befehlszeilenschnittstelle (CLI) bereit, um Ihre Anfragen mit Ihren Anmeldeinformationen kryptografisch zu signieren. Wenn Sie keine AWS Tools verwenden, müssen Sie Anfragen selbst signieren. Weitere Informationen über die empfohlene Methode zur eigenständigen Signierung von Anfragen finden Sie unter [Signierprozess mit Signaturversion 4](#) in der Allgemeine AWS-Referenz.

Unabhängig von der verwendeten Authentifizierungsmethode müssen Sie möglicherweise zusätzliche Sicherheitsinformationen bereitstellen. AWS Empfiehlt beispielsweise, die Multi-Faktor-Authentifizierung (MFA) zu verwenden, um die Sicherheit Ihres Kontos zu erhöhen. Weitere

Informationen finden Sie unter [Multi-Faktor-Authentifizierung](#) im AWS IAM Identity Center - Benutzerhandbuch und [Verwenden der Multi-Faktor-Authentifizierung \(MFA\) in AWS](#) im IAM-Benutzerhandbuch.

AWS-Konto Root-Benutzer

Wenn Sie ein AWS-Konto erstellen, beginnen Sie mit einer Anmeldeidentität, die vollständigen Zugriff auf alle AWS-Services Ressourcen im Konto hat. Diese Identität wird als AWS-Konto -Root-Benutzer bezeichnet. Für den Zugriff auf den Root-Benutzer müssen Sie sich mit der E-Mail-Adresse und dem Passwort anmelden, die zur Erstellung des Kontos verwendet wurden. Wir raten ausdrücklich davon ab, den Root-Benutzer für Alltagsaufgaben zu verwenden, auch nicht für administrative Aufgaben. Schützen Sie Ihre Root-Benutzer-Anmeldeinformationen und verwenden Sie diese, um die Aufgaben auszuführen, die nur der Root-Benutzer ausführen kann. Eine vollständige Liste der Aufgaben, für die Sie sich als Root-Benutzer anmelden müssen, finden Sie unter [Root-Benutzer des AWS-Kontos Anmeldeinformationen und IAM-Identitäten](#) in der Allgemeinen AWS-Referenz.

Verbundidentität

Es hat sich bewährt, dass menschliche Benutzer für den Zugriff mit temporären Anmeldeinformationen einen Verbund mit einem Identitätsanbieter AWS-Services verwenden müssen.

Eine föderierte Identität ist ein Benutzer aus Ihrem Unternehmensverzeichnis, Ihrem Directory Service Web-Identitätsanbieter oder der AWS-Services mithilfe von Anmeldeinformationen aus einer Identitätsquelle zugreift. Verbundene Identitäten übernehmen Rollen, die temporäre Anmeldeinformationen bereitstellen.

Für die zentrale Zugriffsverwaltung empfehlen wir AWS IAM Identity Center. Weitere Informationen finden Sie unter [Was ist IAM Identity Center?](#) im AWS IAM Identity Center -Benutzerhandbuch.

IAM-Benutzer und -Gruppen

Ein [IAM-Benutzer](#) ist eine Identität mit bestimmten Berechtigungen für eine einzelne Person oder Anwendung. Wir empfehlen die Verwendung temporärer Anmeldeinformationen anstelle von IAM-Benutzern mit langfristigen Anmeldeinformationen. Weitere Informationen finden Sie im IAM-Benutzerhandbuch unter [Erfordern, dass menschliche Benutzer den Verbund mit einem Identitätsanbieter verwenden müssen, um AWS mithilfe temporärer Anmeldeinformationen darauf zugreifen zu können](#).

Eine [IAM-Gruppe](#) spezifiziert eine Sammlung von IAM-Benutzern und erleichtert die Verwaltung von Berechtigungen für große Gruppen von Benutzern. Weitere Informationen finden Sie unter [Anwendungsfälle für IAM-Benutzer](#) im IAM-Benutzerhandbuch.

IAM-Rollen

Eine [IAM-Rolle](#) ist eine Identität mit spezifischen Berechtigungen, die temporäre Anmeldeinformationen bereitstellt. Sie können eine Rolle übernehmen, indem Sie [von einer Benutzer- zu einer IAM-Rolle \(Konsole\) wechseln](#) AWS CLI oder einen AWS API-Vorgang aufrufen. Weitere Informationen finden Sie unter [Methoden, um eine Rolle zu übernehmen](#) im IAM-Benutzerhandbuch.

IAM-Rollen sind nützlich für den Verbundbenutzer-Zugriff, temporäre IAM-Benutzerberechtigungen, kontoübergreifenden Zugriff, serviceübergreifenden Zugriff und Anwendungen, die auf Amazon EC2 laufen. Weitere Informationen finden Sie unter [Kontoübergreifender Ressourcenzugriff in IAM](#) im IAM-Benutzerhandbuch.

Verwalten des Zugriffs mit Richtlinien

Sie kontrollieren den Zugriff, AWS indem Sie Richtlinien erstellen und diese an AWS Identitäten oder Ressourcen anhängen. Eine Richtlinie ist ein Objekt, AWS das, wenn es einer Identität oder Ressource zugeordnet ist, deren Berechtigungen definiert. AWS wertet diese Richtlinien aus, wenn ein Prinzipal (Benutzer, Root-Benutzer oder Rollensitzung) eine Anfrage stellt. Die Berechtigungen in den Richtlinien legen fest, ob eine Anforderung zugelassen oder abgelehnt wird. Die meisten Richtlinien werden AWS als JSON-Dokumente gespeichert. Weitere Informationen zu Struktur und Inhalten von JSON-Richtliniendokumenten finden Sie unter [Übersicht über JSON-Richtlinien](#) im IAM-Benutzerhandbuch.

Administratoren können mithilfe von AWS JSON-Richtlinien angeben, wer auf was Zugriff hat. Das heißt, welcher Prinzipal kann Aktionen für welche Ressourcen und unter welchen Bedingungen ausführen.

Eine IAM-Entität (Benutzer oder Rolle) besitzt zunächst keine Berechtigungen. Standardmäßig können Benutzer nichts tun, nicht einmal ihr eigenes Passwort ändern. Um einem Benutzer die Berechtigung für eine Aktion zu erteilen, muss ein Administrator einem Benutzer eine Berechtigungsrichtlinie zuweisen. Alternativ kann der Administrator den Benutzer zu einer Gruppe hinzufügen, die über die gewünschten Berechtigungen verfügt. Wenn ein Administrator einer Gruppe Berechtigungen erteilt, erhalten alle Benutzer in dieser Gruppe diese Berechtigungen.

IAM-Richtlinien definieren Berechtigungen für eine Aktion unabhängig von der Methode, die Sie zur Ausführung der Aktion verwenden. Angenommen, es gibt eine Richtlinie, die Berechtigungen für die `iam:GetRole`-Aktion erteilt. Ein Benutzer mit dieser Richtlinie kann Rolleninformationen von der AWS-Managementkonsole, der AWS CLI, der oder der AWS API abrufen.

Identitätsbasierte Richtlinien

Identitätsbasierte Richtlinien sind JSON-Berechtigungsrichtliniendokumente, die Sie einer Identität anfügen können, wie z. B. IAM-Benutzern, -Benutzergruppen oder -Rollen. Diese Richtlinien steuern, welche Aktionen die Benutzer und Rollen für welche Ressourcen und unter welchen Bedingungen ausführen können. Informationen zum Erstellen identitätsbasierter Richtlinien finden Sie unter [Definieren benutzerdefinierter IAM-Berechtigungen mit vom Kunden verwalteten Richtlinien](#) im IAM-Benutzerhandbuch.

Identitätsbasierte Richtlinien können weiter als Inline-Richtlinien oder verwaltete Richtlinien kategorisiert werden. Inline-Richtlinien sind direkt in einen einzelnen Benutzer, eine einzelne Gruppe oder eine einzelne Rolle eingebettet. Verwaltete Richtlinien sind eigenständige Richtlinien, die Sie mehreren Benutzern, Gruppen und Rollen in Ihrem AWS-Konto anfügen können. Zu den verwalteten Richtlinien gehören AWS verwaltete Richtlinien und vom Kunden verwaltete Richtlinien. Informationen dazu, wie Sie zwischen einer verwalteten Richtlinie und einer eingebundenen Richtlinie wählen, finden Sie unter [Auswahl zwischen verwalteten und eingebundenen Richtlinien](#) im IAM-Benutzerhandbuch.

Ressourcenbasierte Richtlinien

Ressourcenbasierte Richtlinien sind JSON-Richtliniendokumente, die Sie an eine Ressource anfügen. Beispiele für ressourcenbasierte Richtlinien sind IAM-Rollen-Vertrauensrichtlinien und Amazon-S3-Bucket-Richtlinien. In Services, die ressourcenbasierte Richtlinien unterstützen, können Service-Administratoren sie verwenden, um den Zugriff auf eine bestimmte Ressource zu steuern. Für die Ressource, an welche die Richtlinie angehängt ist, legt die Richtlinie fest, welche Aktionen ein bestimmter Prinzipal unter welchen Bedingungen für diese Ressource ausführen kann. Sie müssen in einer ressourcenbasierten Richtlinie [einen Prinzipal angeben](#). Zu den Prinzipalen können Konten, Benutzer, Rollen, Verbundbenutzer oder gehören. AWS-Services

Ressourcenbasierte Richtlinien sind Richtlinien innerhalb dieses Diensts. Sie können AWS verwaltete Richtlinien von IAM nicht in einer ressourcenbasierten Richtlinie verwenden.

Weitere Richtlinientypen

AWS unterstützt zusätzliche, weniger verbreitete Richtlinientypen. Mit diesen Richtlinientypen können Sie die maximalen Berechtigungen festlegen, die Ihnen durch die gängigeren Richtlinientypen gewährt werden.

- **Berechtigungsgrenzen** – Eine Berechtigungsgrenze ist ein erweitertes Feature, mit der Sie die maximalen Berechtigungen festlegen können, die eine identitätsbasierte Richtlinie einer IAM-Entität (IAM-Benutzer oder -Rolle) erteilen kann. Sie können eine Berechtigungsgrenze für eine Entität festlegen. Die resultierenden Berechtigungen sind eine Schnittmenge der identitätsbasierten Richtlinien der Entität und ihrer Berechtigungsgrenzen. Ressourcenbasierte Richtlinien, die den Benutzer oder die Rolle im Feld `Principal` angeben, werden nicht durch Berechtigungsgrenzen eingeschränkt. Eine explizite Zugriffsverweigerung in einer dieser Richtlinien setzt eine Zugriffserlaubnis außer Kraft. Weitere Informationen über Berechtigungsgrenzen finden Sie unter [Berechtigungsgrenzen für IAM-Entitäten](#) im IAM-Benutzerhandbuch.
- **Dienststeuerungsrichtlinien (SCPs)** — SCPs sind JSON-Richtlinien, die die maximalen Berechtigungen für eine Organisation oder Organisationseinheit (OU) in festlegen AWS Organizations. AWS Organizations ist ein Dienst zur Gruppierung und zentralen Verwaltung mehrerer Objekte AWS-Konten, die Ihrem Unternehmen gehören. Wenn Sie alle Funktionen in einer Organisation aktivieren, können Sie Richtlinien zur Servicesteuerung (SCPs) auf einige oder alle Ihre Konten anwenden. Das SCP schränkt die Berechtigungen für Entitäten in Mitgliedskonten ein, einschließlich der einzelnen Root-Benutzer des AWS-Kontos Entitäten. Weitere Informationen zu Organizations und SCPs finden Sie unter [So SCPs arbeiten](#) Sie im AWS Organizations Benutzerhandbuch.
- **Sitzungsrichtlinien** – Sitzungsrichtlinien sind erweiterte Richtlinien, die Sie als Parameter übergeben, wenn Sie eine temporäre Sitzung für eine Rolle oder einen verbundenen Benutzer programmgesteuert erstellen. Die resultierenden Sitzungsberechtigungen sind eine Schnittmenge der auf der Identität des Benutzers oder der Rolle basierenden Richtlinien und der Sitzungsrichtlinien. Berechtigungen können auch aus einer ressourcenbasierten Richtlinie stammen. Eine explizite Zugriffsverweigerung in einer dieser Richtlinien setzt eine Zugriffserlaubnis außer Kraft. Weitere Informationen finden Sie unter [Sitzungsrichtlinien](#) im IAM-Benutzerhandbuch.

Mehrere Richtlinientypen

Wenn für eine Anfrage mehrere Arten von Richtlinien gelten, sind die sich daraus ergebenden Berechtigungen schwieriger zu verstehen. Informationen darüber, wie AWS bestimmt wird, ob eine

Anfrage zulässig ist, wenn mehrere Richtlinientypen betroffen sind, finden Sie unter [Bewertungslogik für Richtlinien](#) im IAM-Benutzerhandbuch.

Wie AWS Clean Rooms funktioniert mit IAM

Bevor Sie IAM verwenden, um den Zugriff auf zu verwalten AWS Clean Rooms, sollten Sie sich darüber informieren, welche IAM-Funktionen zur Verfügung stehen. AWS Clean Rooms

IAM-Funktionen, die Sie mit verwenden können AWS Clean Rooms

IAM-Feature	AWS Clean Rooms Unterstützung
Identitätsbasierte Richtlinien	Ja
Ressourcenbasierte Richtlinien	Teilweise
Richtlinienaktionen	Ja
Richtlinienressourcen	Ja
Richtlinienbedingungsschlüssel (servicespezifisch)	Teilweise
ACLs	Nein
ABAC (Tags in Richtlinien)	Ja
Temporäre Anmeldeinformationen	Ja
Forward Access Sessions (FAS)	Ja
Servicerollen	Ja
Service-verknüpfte Rollen	Nein

Einen allgemeinen Überblick darüber, wie die meisten IAM-Funktionen AWS-Services funktionieren AWS Clean Rooms und wie sie [funktionieren AWS-Services](#), finden Sie im IAM-Benutzerhandbuch.

Identitätsbasierte Richtlinien für AWS Clean Rooms

Unterstützt Richtlinien auf Identitätsbasis: Ja

Identitätsbasierte Richtlinien sind JSON-Berechtigungsrichtliniendokumente, die Sie einer Identität anfügen können, wie z. B. IAM-Benutzern, -Benutzergruppen oder -Rollen. Diese Richtlinien steuern, welche Aktionen die Benutzer und Rollen für welche Ressourcen und unter welchen Bedingungen ausführen können. Informationen zum Erstellen identitätsbasierter Richtlinien finden Sie unter [Definieren benutzerdefinierter IAM-Berechtigungen mit vom Kunden verwalteten Richtlinien](#) im IAM-Benutzerhandbuch.

Mit identitätsbasierten IAM-Richtlinien können Sie angeben, welche Aktionen und Ressourcen zugelassen oder abgelehnt werden. Darüber hinaus können Sie die Bedingungen festlegen, unter denen Aktionen zugelassen oder abgelehnt werden. Informationen zu sämtlichen Elementen, die Sie in einer JSON-Richtlinie verwenden, finden Sie in der [IAM-Referenz für JSON-Richtlinienelemente](#) im IAM-Benutzerhandbuch.

Beispiele für identitätsbasierte Richtlinien für AWS Clean Rooms

Beispiele für AWS Clean Rooms identitätsbasierte Richtlinien finden Sie unter [Beispiele für identitätsbasierte Richtlinien für AWS Clean Rooms](#)

Ressourcenbasierte Richtlinien finden Sie in AWS Clean Rooms

Unterstützt ressourcenbasierte Richtlinien: Teilweise

Ressourcenbasierte Richtlinien sind JSON-Richtliniendokumente, die Sie an eine Ressource anfügen. Beispiele für ressourcenbasierte Richtlinien sind IAM-Rollen-Vertrauensrichtlinien und Amazon-S3-Bucket-Richtlinien. In Services, die ressourcenbasierte Richtlinien unterstützen, können Service-Administratoren sie verwenden, um den Zugriff auf eine bestimmte Ressource zu steuern. Für die Ressource, an welche die Richtlinie angehängt ist, legt die Richtlinie fest, welche Aktionen ein bestimmter Prinzipal unter welchen Bedingungen für diese Ressource ausführen kann. Sie müssen in einer ressourcenbasierten Richtlinie [einen Prinzipal angeben](#). Zu den Prinzipalen können Konten, Benutzer, Rollen, Verbundbenutzer oder gehören. AWS-Services

Um kontoübergreifenden Zugriff zu ermöglichen, können Sie ein gesamtes Konto oder IAM-Entitäten in einem anderen Konto als Prinzipal in einer ressourcenbasierten Richtlinie angeben. Weitere Informationen finden Sie unter [Kontoübergreifender Ressourcenzugriff in IAM](#) im IAM-Benutzerhandbuch.

Der AWS Clean Rooms Dienst unterstützt nur eine Art von ressourcenbasierter Richtlinie, die als konfiguriertes Lookalike-Modell bezeichnet wird (verwaltete Ressourcenrichtlinie), die an ein konfiguriertes Lookalike-Modell angehängt ist. Diese Richtlinie definiert, welche Principals Aktionen auf dem konfigurierten Lookalike-Modell ausführen können.

Informationen zum Anhängen einer ressourcenbasierten Richtlinie an ein konfiguriertes Lookalike-Modell finden Sie unter. [IAM-Verhalten für AWS Clean Rooms ML](#)

Politische Maßnahmen für AWS Clean Rooms

Unterstützt Richtlinienaktionen: Ja

Administratoren können mithilfe von AWS JSON-Richtlinien angeben, wer Zugriff auf was hat. Das heißt, welcher Prinzipal Aktionen für welche Ressourcen und unter welchen Bedingungen ausführen kann.

Das Element Action einer JSON-Richtlinie beschreibt die Aktionen, mit denen Sie den Zugriff in einer Richtlinie zulassen oder verweigern können. Nehmen Sie Aktionen in eine Richtlinie auf, um Berechtigungen zur Ausführung des zugehörigen Vorgangs zu erteilen.

Eine Liste der AWS Clean Rooms Aktionen finden Sie unter [Aktionen definiert von AWS Clean Rooms](#) in der Serviceautorisierungsreferenz.

Bei Richtlinienaktionen wird vor der Aktion das folgende Präfix AWS Clean Rooms verwendet.

```
cleanrooms
```

Um mehrere Aktionen in einer einzigen Anweisung anzugeben, trennen Sie sie mit Kommata:

```
"Action": [  
    "cleanrooms:action1",  
    "cleanrooms:action2"  
]
```

Beispiele für AWS Clean Rooms identitätsbasierte Richtlinien finden Sie unter. [Beispiele für identitätsbasierte Richtlinien für AWS Clean Rooms](#)

Politische Ressourcen für AWS Clean Rooms

Unterstützt Richtlinienressourcen: Ja

Administratoren können mithilfe von AWS JSON-Richtlinien angeben, wer Zugriff auf was hat. Das heißt, welcher Prinzipal Aktionen für welche Ressourcen und unter welchen Bedingungen ausführen kann.

Das JSON-Richtlinienelement `Resource` gibt die Objekte an, auf welche die Aktion angewendet wird. Als Best Practice geben Sie eine Ressource mit dem zugehörigen [Amazon-Ressourcennamen \(ARN\)](#) an. Verwenden Sie für Aktionen, die keine Berechtigungen auf Ressourcenebene unterstützen, einen Platzhalter (*), um anzugeben, dass die Anweisung für alle Ressourcen gilt.

```
"Resource": "*" 
```

Eine Liste der AWS Clean Rooms Ressourcentypen und ihrer ARNs Eigenschaften finden Sie unter [Resources defined by AWS Clean Rooms](#) in der Service Authorization Reference. Informationen zu den Aktionen, mit denen Sie den ARN einzelner Ressourcen angeben können, finden Sie unter [Von AWS Clean Rooms definierte Aktionen](#).

Beispiele für AWS Clean Rooms identitätsbasierte Richtlinien finden Sie unter. [Beispiele für identitätsbasierte Richtlinien für AWS Clean Rooms](#)

Bedingungsschlüssel für Richtlinien für AWS Clean Rooms

Unterstützt dienstspezifische Richtlinien-Bedingungsschlüssel: Teilweise

Administratoren können mithilfe von AWS JSON-Richtlinien angeben, wer Zugriff auf was hat. Das heißt, welcher Prinzipal Aktionen für welche Ressourcen und unter welchen Bedingungen ausführen kann.

Das Element `Condition` gibt an, wann Anweisungen auf der Grundlage definierter Kriterien ausgeführt werden. Sie können bedingte Ausdrücke erstellen, die [Bedingungsoperatoren](#) verwenden, z. B. ist gleich oder kleiner als, damit die Bedingung in der Richtlinie mit Werten in der Anforderung übereinstimmt. Eine Übersicht aller AWS globalen Bedingungsschlüssel finden Sie unter [Kontextschlüssel für AWS globale Bedingungen](#) im IAM-Benutzerhandbuch.

Informationen darüber, wie AWS Clean Rooms ML Richtlinien-Bedingungsschlüssel verwendet, finden Sie unter [IAM-Verhalten für AWS Clean Rooms ML](#).

ACLs in AWS Clean Rooms

Unterstützt ACLs: Nein

Zugriffskontrolllisten (ACLs) steuern, welche Principals (Kontomitglieder, Benutzer oder Rollen) über Zugriffsberechtigungen für eine Ressource verfügen. ACLs ähneln ressourcenbasierten Richtlinien, verwenden jedoch nicht das JSON-Richtliniendokumentformat.

ABAC mit AWS Clean Rooms

Unterstützt ABAC (Tags in Richtlinien): Ja

Die attributbasierte Zugriffskontrolle (ABAC) ist eine Autorisierungsstrategie, bei der Berechtigungen basierend auf Attributen, auch als Tags bezeichnet, definiert werden. Sie können Tags an IAM-Entitäten und AWS -Ressourcen anhängen und dann ABAC-Richtlinien entwerfen, die Operationen zulassen, wenn das Tag des Prinzipals mit dem Tag auf der Ressource übereinstimmt.

Um den Zugriff auf der Grundlage von Tags zu steuern, geben Sie im Bedingungelement einer [Richtlinie Tag-Informationen](#) an, indem Sie die Schlüssel `aws:ResourceTag/key-name`, `aws:RequestTag/key-name`, oder Bedingung `aws:TagKeys` verwenden.

Wenn ein Service alle drei Bedingungsschlüssel für jeden Ressourcentyp unterstützt, lautet der Wert für den Service Ja. Wenn ein Service alle drei Bedingungsschlüssel für nur einige Ressourcentypen unterstützt, lautet der Wert Teilweise.

Weitere Informationen zu ABAC finden Sie unter [Definieren von Berechtigungen mit ABAC-Autorisierung](#) im IAM-Benutzerhandbuch. Um ein Tutorial mit Schritten zur Einstellung von ABAC anzuzeigen, siehe [Attributbasierte Zugriffskontrolle \(ABAC\)](#) verwenden im IAM-Benutzerhandbuch.

Verwenden temporärer Anmeldeinformationen mit AWS Clean Rooms

Unterstützt temporäre Anmeldeinformationen: Ja

Temporäre Anmeldeinformationen ermöglichen kurzfristigen Zugriff auf AWS Ressourcen und werden automatisch erstellt, wenn Sie einen Verbund verwenden oder die Rollen wechseln. AWS empfiehlt, temporäre Anmeldeinformationen dynamisch zu generieren, anstatt langfristige Zugriffsschlüssel zu verwenden. Weitere Informationen finden Sie unter [Temporäre Anmeldeinformationen in IAM und AWS-Services , die mit IAM funktionieren](#) im IAM-Benutzerhandbuch.

Zugriffssitzungen weiterleiten für AWS Clean Rooms

Unterstützt Forward Access Sessions (FAS): Ja

Forward-Access-Sitzungen (FAS) verwenden die Berechtigungen des Prinzipals, der einen aufruft AWS-Service, in Kombination mit der Anforderung, Anfragen AWS-Service an nachgelagerte Dienste zu stellen. Einzelheiten zu den Richtlinien für FAS-Anforderungen finden Sie unter [Zugriffssitzungen weiterleiten](#).

Servicerollen für AWS Clean Rooms

Unterstützt Servicerollen: Ja

Eine Servicerolle ist eine [IAM-Rolle](#), die ein Service annimmt, um Aktionen in Ihrem Namen auszuführen. Ein IAM-Administrator kann eine Servicerolle innerhalb von IAM erstellen, ändern und löschen. Weitere Informationen finden Sie unter [Erstellen einer Rolle zum Delegieren von Berechtigungen an einen AWS-Service](#) im IAM-Benutzerhandbuch.

Warning

Durch das Ändern der Berechtigungen für eine Servicerolle kann die AWS Clean Rooms Funktionalität beeinträchtigt werden. Bearbeiten Sie Servicerollen nur, AWS Clean Rooms wenn Sie dazu eine Anleitung erhalten.

Dienstbezogene Rollen für AWS Clean Rooms

Unterstützt serviceverknüpfte Rollen: Ja

Eine dienstbezogene Rolle ist eine Art von Servicerolle, die mit einer verknüpft ist. AWS-Service Der Service kann die Rolle übernehmen, um eine Aktion in Ihrem Namen auszuführen. Dienstbezogene Rollen werden in Ihrem Dienst angezeigt AWS-Konto und gehören dem Dienst. Ein IAM-Administrator kann die Berechtigungen für Service-verknüpfte Rollen anzeigen, aber nicht bearbeiten.

Details zum Erstellen oder Verwalten von serviceverknüpften Rollen finden Sie unter [AWS -Services, die mit IAM funktionieren](#). Suchen Sie in der Tabelle nach einem Service mit einem Yes in der Spalte Service-linked role (Serviceverknüpfte Rolle). Wählen Sie den Link Yes (Ja) aus, um die Dokumentation für die serviceverknüpfte Rolle für diesen Service anzuzeigen.

Beispiele für identitätsbasierte Richtlinien für AWS Clean Rooms

Benutzer und Rollen haben standardmäßig nicht die Berechtigung, AWS Clean Rooms -Ressourcen zu erstellen oder zu ändern. Ein IAM-Administrator muss IAM-Richtlinien erstellen, die Benutzern die Berechtigung erteilen, Aktionen für die Ressourcen auszuführen, die sie benötigen.

Informationen dazu, wie Sie unter Verwendung dieser beispielhaften JSON-Richtliniendokumente eine identitätsbasierte IAM-Richtlinie erstellen, finden Sie unter [Erstellen von IAM-Richtlinien \(Konsole\)](#) im IAM-Benutzerhandbuch.

Einzelheiten zu Aktionen und Ressourcentypen, die von definiert wurden AWS Clean Rooms, einschließlich des Formats von ARNs für die einzelnen Ressourcentypen, finden Sie unter [Aktionen, Ressourcen und Bedingungsschlüssel für AWS Clean Rooms](#) in der Service Authorization Reference.

Themen

- [Best Practices für Richtlinien](#)
- [Verwenden der Konsole AWS Clean Rooms](#)
- [Gewähren der Berechtigung zur Anzeige der eigenen Berechtigungen für Benutzer](#)

Best Practices für Richtlinien

Identitätsbasierte Richtlinien legen fest, ob jemand AWS Clean Rooms Ressourcen in Ihrem Konto erstellen, darauf zugreifen oder sie löschen kann. Dies kann zusätzliche Kosten für Ihr verursachen AWS-Konto. Beachten Sie beim Erstellen oder Bearbeiten identitätsbasierter Richtlinien die folgenden Richtlinien und Empfehlungen:

- Erste Schritte mit AWS verwalteten Richtlinien und Umstellung auf Berechtigungen mit den geringsten Rechten — Verwenden Sie die AWS verwalteten Richtlinien, die Berechtigungen für viele gängige Anwendungsfälle gewähren, um damit zu beginnen, Ihren Benutzern und Workloads Berechtigungen zu gewähren. Sie sind in Ihrem verfügbar. AWS-Konto Wir empfehlen Ihnen, die Berechtigungen weiter zu reduzieren, indem Sie vom AWS Kunden verwaltete Richtlinien definieren, die speziell auf Ihre Anwendungsfälle zugeschnitten sind. Weitere Informationen finden Sie unter [Von AWS verwaltete Richtlinien](#) oder [Von AWS verwaltete Richtlinien für Auftragsfunktionen](#) im IAM-Benutzerhandbuch.
- Anwendung von Berechtigungen mit den geringsten Rechten – Wenn Sie mit IAM-Richtlinien Berechtigungen festlegen, gewähren Sie nur die Berechtigungen, die für die Durchführung einer Aufgabe erforderlich sind. Sie tun dies, indem Sie die Aktionen definieren, die für bestimmte Ressourcen unter bestimmten Bedingungen durchgeführt werden können, auch bekannt als die geringsten Berechtigungen. Weitere Informationen zur Verwendung von IAM zum Anwenden von Berechtigungen finden Sie unter [Richtlinien und Berechtigungen in IAM](#) im IAM-Benutzerhandbuch.
- Verwenden von Bedingungen in IAM-Richtlinien zur weiteren Einschränkung des Zugriffs – Sie können Ihren Richtlinien eine Bedingung hinzufügen, um den Zugriff auf Aktionen und

Ressourcen zu beschränken. Sie können beispielsweise eine Richtlinienbedingung schreiben, um festzulegen, dass alle Anforderungen mithilfe von SSL gesendet werden müssen. Sie können auch Bedingungen verwenden, um Zugriff auf Serviceaktionen zu gewähren, wenn diese für einen bestimmten Zweck verwendet werden AWS-Service, z. CloudFormation B. Weitere Informationen finden Sie unter [IAM-JSON-Richtlinienelemente: Bedingung](#) im IAM-Benutzerhandbuch.

- Verwenden von IAM Access Analyzer zur Validierung Ihrer IAM-Richtlinien, um sichere und funktionale Berechtigungen zu gewährleisten – IAM Access Analyzer validiert neue und vorhandene Richtlinien, damit die Richtlinien der IAM-Richtliniensprache (JSON) und den bewährten IAM-Methoden entsprechen. IAM Access Analyzer stellt mehr als 100 Richtlinienprüfungen und umsetzbare Empfehlungen zur Verfügung, damit Sie sichere und funktionale Richtlinien erstellen können. Weitere Informationen finden Sie unter [Richtlinienvvalidierung mit IAM Access Analyzer](#) im IAM-Benutzerhandbuch.
- Multi-Faktor-Authentifizierung (MFA) erforderlich — Wenn Sie ein Szenario haben, das IAM-Benutzer oder einen Root-Benutzer in Ihrem System erfordert AWS-Konto, aktivieren Sie MFA für zusätzliche Sicherheit. Um MFA beim Aufrufen von API-Vorgängen anzufordern, fügen Sie Ihren Richtlinien MFA-Bedingungen hinzu. Weitere Informationen finden Sie unter [Sicherer API-Zugriff mit MFA](#) im IAM-Benutzerhandbuch.

Weitere Informationen zu bewährten Methoden in IAM finden Sie unter [Best Practices für die Sicherheit in IAM](#) im IAM-Benutzerhandbuch.

Verwenden der Konsole AWS Clean Rooms

Um auf die AWS Clean Rooms Konsole zugreifen zu können, benötigen Sie ein Mindestmaß an Berechtigungen. Diese Berechtigungen müssen es Ihnen ermöglichen, Details zu den AWS Clean Rooms Ressourcen in Ihrem aufzulisten und anzuzeigen AWS-Konto. Wenn Sie eine identitätsbasierte Richtlinie erstellen, die strenger ist als die mindestens erforderlichen Berechtigungen, funktioniert die Konsole nicht wie vorgesehen für Entitäten (Benutzer oder Rollen) mit dieser Richtlinie.

Sie müssen Benutzern, die nur die API AWS CLI oder die AWS API aufrufen, keine Mindestberechtigungen für die Konsole gewähren. Stattdessen sollten Sie nur Zugriff auf die Aktionen zulassen, die der API-Operation entsprechen, die die Benutzer ausführen möchten.

Um sicherzustellen, dass Benutzer und Rollen die AWS Clean Rooms Konsole weiterhin verwenden können, fügen Sie den Entitäten auch die AWS Clean Rooms *FullAccess* oder die *ReadOnly* AWS

verwaltete Richtlinie hinzu. Weitere Informationen finden Sie unter [Hinzufügen von Berechtigungen zu einem Benutzer](#) im IAM-Benutzerhandbuch.

Gewähren der Berechtigung zur Anzeige der eigenen Berechtigungen für Benutzer

In diesem Beispiel wird gezeigt, wie Sie eine Richtlinie erstellen, die IAM-Benutzern die Berechtigung zum Anzeigen der eingebundenen Richtlinien und verwalteten Richtlinien gewährt, die ihrer Benutzeridentität angefügt sind. Diese Richtlinie umfasst Berechtigungen zum Ausführen dieser Aktion auf der Konsole oder programmgesteuert mithilfe der AWS CLI AWS OR-API.

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "ViewOwnUserInfo",
      "Effect": "Allow",
      "Action": [
        "iam:GetUserPolicy",
        "iam:ListGroupsForUser",
        "iam:ListAttachedUserPolicies",
        "iam:ListUserPolicies",
        "iam:GetUser"
      ],
      "Resource": ["arn:aws:iam::*:user/${aws:username}"]
    },
    {
      "Sid": "NavigateInConsole",
      "Effect": "Allow",
      "Action": [
        "iam:GetGroupPolicy",
        "iam:GetPolicyVersion",
        "iam:GetPolicy",
        "iam:ListAttachedGroupPolicies",
        "iam:ListGroupPolicies",
        "iam:ListPolicyVersions",
        "iam:ListPolicies",
        "iam:ListUsers"
      ],
      "Resource": "*"
    }
  ]
}
```

AWS verwaltete Richtlinien für AWS Clean Rooms

Eine AWS verwaltete Richtlinie ist eine eigenständige Richtlinie, die von erstellt und verwaltet wird AWS. AWS Verwaltete Richtlinien sind so konzipiert, dass sie Berechtigungen für viele gängige Anwendungsfälle bereitstellen, sodass Sie damit beginnen können, Benutzern, Gruppen und Rollen Berechtigungen zuzuweisen.

Beachten Sie, dass AWS verwaltete Richtlinien für Ihre speziellen Anwendungsfälle möglicherweise keine Berechtigungen mit den geringsten Rechten gewähren, da sie allen AWS Kunden zur Verfügung stehen. Wir empfehlen Ihnen, die Berechtigungen weiter zu reduzieren, indem Sie [vom Kunden verwaltete Richtlinien](#) definieren, die speziell auf Ihre Anwendungsfälle zugeschnitten sind.

Sie können die in AWS verwalteten Richtlinien definierten Berechtigungen nicht ändern. Wenn die in einer AWS verwalteten Richtlinie definierten Berechtigungen AWS aktualisiert werden, wirkt sich das Update auf alle Prinzidentitäten (Benutzer, Gruppen und Rollen) aus, denen die Richtlinie zugeordnet ist. AWS aktualisiert eine AWS verwaltete Richtlinie höchstwahrscheinlich, wenn eine neue Richtlinie eingeführt AWS-Service wird oder neue API-Operationen für bestehende Dienste verfügbar werden.

Weitere Informationen finden Sie unter [Von AWS verwaltete Richtlinien](#) im IAM-Benutzerhandbuch.

AWS verwaltete Richtlinie: **AWSCleanRoomsReadOnlyAccess**

Sie können eine Verbindung **AWSCleanRoomsReadOnlyAccess** zu Ihren IAM-Prinzipalen herstellen.

Diese Richtlinie gewährt nur Leseberechtigungen für Ressourcen und Metadaten in einer Kollaboration. **AWSCleanRoomsReadOnlyAccess**

Details zu Berechtigungen

Diese Richtlinie umfasst die folgenden Berechtigungen:

- **CleanRoomsRead**— Ermöglicht Prinzipalen nur Lesezugriff auf den Dienst.
- **ConsoleDisplayTables**— Ermöglicht Prinzipalen den schreibgeschützten Zugriff auf die AWS Glue Metadaten, die für die Anzeige von Daten zu den zugrunde liegenden Tabellen auf der Konsole erforderlich sind. AWS Glue
- **ConsoleLogSummaryQueryLogs**— Ermöglicht es den Prinzipalen, die Abfrageprotokolle zu sehen.

- `ConsoleLogSummaryObtainLogs`— Ermöglicht Prinzipalen das Abrufen der Protokollergebnisse.

Eine JSON-Liste der Richtliniendetails finden Sie [AWSCleanRoomsReadOnlyAccessim](#) Referenzhandbuch für AWS verwaltete Richtlinien.

AWS verwaltete Richtlinie: **AWSCleanRoomsFullAccess**

Sie können eine Verbindung `AWSCleanRoomsFullAccess` zu Ihren IAM-Prinzipalen herstellen.

Diese Richtlinie gewährt Administratorberechtigungen, die vollen Zugriff (Lesen, Schreiben und Aktualisieren) auf Ressourcen und Metadaten in einer AWS Clean Rooms Kollaboration ermöglichen. Diese Richtlinie beinhaltet den Zugriff zur Durchführung von Abfragen.

Details zu Berechtigungen

Diese Richtlinie umfasst die folgenden Berechtigungen:

- `CleanRoomsAccess`— Gewährt vollen Zugriff auf alle Aktionen auf allen Ressourcen für AWS Clean Rooms.
- `PassServiceRole`— Gewährt Zugriff zur Übergabe einer Servicerolle nur an den Dienst (`PassedToServiceBedingung`), dessen Name `cleanrooms ""` enthält.
- `ListRolesToPickServiceRole`— Ermöglicht es Prinzipalen, alle ihre Rollen aufzulisten, um bei der Verwendung AWS Clean Rooms eine Servicerolle auszuwählen.
- `GetRoleAndListRolePoliciesToInspectServiceRole`— Ermöglicht Prinzipalen, die Servicerolle und die entsprechende Richtlinie in IAM zu sehen.
- `ListPoliciesToInspectServiceRolePolicy`— Ermöglicht Prinzipalen, die Servicerolle und die entsprechende Richtlinie in IAM zu sehen.
- `GetPolicyToInspectServiceRolePolicy`— Ermöglicht Prinzipalen, die Servicerolle und die entsprechende Richtlinie in IAM zu sehen.
- `ConsoleDisplayTables`— Ermöglicht Prinzipalen den schreibgeschützten Zugriff auf die AWS Glue Metadaten, die für die Anzeige von Daten zu den zugrunde liegenden AWS Glue Tabellen auf der Konsole erforderlich sind.
- `ConsolePickQueryResultsBucketListAll`— Ermöglicht Prinzipalen, einen Amazon S3 S3-Bucket aus einer Liste aller verfügbaren S3-Buckets auszuwählen, in die ihre Abfrageergebnisse geschrieben werden.

- **SetQueryResultsBucket**— Ermöglicht Prinzipalen, einen S3-Bucket auszuwählen, in den ihre Abfrageergebnisse geschrieben werden.
- **ConsoleDisplayQueryResults**— Ermöglicht es den Prinzipalen, dem Kunden die Abfrageergebnisse anzuzeigen, die aus dem S3-Bucket gelesen wurden.
- **WriteQueryResults**— Ermöglicht Prinzipalen, die Abfrageergebnisse in einen kundeneigenen S3-Bucket zu schreiben.
- **EstablishLogDeliveries**— Ermöglicht Principals, Abfrageprotokolle an die Amazon CloudWatch Logs-Protokollgruppe eines Kunden zu senden.
- **SetupLogGroupsDescribe**— Ermöglicht Prinzipalen, den Prozess zur Erstellung von Amazon CloudWatch Logs-Protokollgruppen zu verwenden.
- **SetupLogGroupsCreate**— Ermöglicht Prinzipalen, eine Amazon CloudWatch Logs-Protokollgruppe zu erstellen.
- **SetupLogGroupsResourcePolicy**— Ermöglicht Prinzipalen, eine Ressourcenrichtlinie für die Amazon CloudWatch Logs-Protokollgruppe einzurichten.
- **ConsoleLogSummaryQueryLogs**— Ermöglicht es den Prinzipalen, die Abfrageprotokolle einzusehen.
- **ConsoleLogSummaryObtainLogs**— Ermöglicht Prinzipalen das Abrufen der Protokollergebnisse.

Eine JSON-Liste der Richtliniendetails finden Sie [AWSCleanRoomsFullAccess](#) im Referenzhandbuch für AWS verwaltete Richtlinien.

AWS verwaltete Richtlinie: **AWSCleanRoomsFullAccessNoQuerying**

Sie können es `AWSCleanRoomsFullAccessNoQuerying` an Ihre anhängen IAM principals.

Diese Richtlinie gewährt Administratorberechtigungen, die vollen Zugriff (Lesen, Schreiben und Aktualisieren) auf Ressourcen und Metadaten in einer AWS Clean Rooms Kollaboration ermöglichen. Diese Richtlinie schließt den Zugriff zur Durchführung von Abfragen aus.

Details zu Berechtigungen

Diese Richtlinie umfasst die folgenden Berechtigungen:

- **CleanRoomsAccess**— Gewährt vollen Zugriff auf alle Aktionen auf allen Ressourcen für AWS Clean Rooms, mit Ausnahme von Abfragen in Kollaborationen.

- **CleanRoomsNoQuerying**— Verweigert ausdrücklich das Abfragen `StartProtectedQuery` und verhindert `UpdateProtectedQuery` es.
- **PassServiceRole**— Gewährt Zugriff auf die Übergabe einer Servicerolle nur an den Dienst (`PassedToService` Bedingung), dessen Name "cleanrooms" enthält.
- **ListRolesToPickServiceRole**— Ermöglicht es Prinzipalen, alle ihre Rollen aufzulisten, um bei der Verwendung AWS Clean Rooms eine Servicerolle auszuwählen.
- **GetRoleAndListRolePoliciesToInspectServiceRole**— Ermöglicht Prinzipalen, die Servicerolle und die entsprechende Richtlinie in IAM zu sehen.
- **ListPoliciesToInspectServiceRolePolicy**— Ermöglicht Prinzipalen, die Servicerolle und die entsprechende Richtlinie in IAM zu sehen.
- **GetPolicyToInspectServiceRolePolicy**— Ermöglicht Prinzipalen, die Servicerolle und die entsprechende Richtlinie in IAM zu sehen.
- **ConsoleDisplayTables**— Ermöglicht Prinzipalen den schreibgeschützten Zugriff auf die AWS Glue Metadaten, die für die Anzeige von Daten zu den zugrunde liegenden AWS Glue Tabellen auf der Konsole erforderlich sind.
- **EstablishLogDeliveries**— Ermöglicht Principals, Abfrageprotokolle an die Amazon CloudWatch Logs-Protokollgruppe eines Kunden zu senden.
- **SetupLogGroupsDescribe**— Ermöglicht Prinzipalen, den Prozess zur Erstellung von Amazon CloudWatch Logs-Protokollgruppen zu verwenden.
- **SetupLogGroupsCreate**— Ermöglicht Prinzipalen, eine Amazon CloudWatch Logs-Protokollgruppe zu erstellen.
- **SetupLogGroupsResourcePolicy**— Ermöglicht Prinzipalen, eine Ressourcenrichtlinie für die Amazon CloudWatch Logs-Protokollgruppe einzurichten.
- **ConsoleLogSummaryQueryLogs**— Ermöglicht es den Prinzipalen, die Abfrageprotokolle einzusehen.
- **ConsoleLogSummaryObtainLogs**— Ermöglicht Prinzipalen das Abrufen der Protokollergebnisse.
- **cleanrooms**— Verwaltet Kollaborationen, Analysevorlagen, konfigurierte Tabellen, Mitgliedschaften und zugehörige Ressourcen innerhalb des Service. AWS Clean Rooms Führen Sie verschiedene Operationen durch, z. B. das Erstellen, Aktualisieren, Löschen, Auflisten und Abrufen von Informationen zu diesen Ressourcen.
- **iam**— Übergibt Dienstrollen, deren Namen "cleanrooms" enthalten, an den AWS Clean Rooms Dienst. Listen Sie Rollen und Richtlinien auf und überprüfen Sie die Dienstrollen und Richtlinien, die sich auf den AWS Clean Rooms Dienst beziehen.

- **glue**— Rufen Sie Informationen zu Datenbanken, Tabellen, Partitionen und Schemas von ab AWS Glue. Dies ist erforderlich, damit der AWS Clean Rooms Dienst die zugrunde liegenden Datenquellen anzeigen und mit ihnen interagieren kann.
- **logs**— Verwalten Sie Protokollzustellungen, Protokollgruppen und Ressourcenrichtlinien für CloudWatch Protokolle. Abfragen und Abrufen von Protokollen, die sich auf den AWS Clean Rooms Dienst beziehen. Diese Berechtigungen sind für Überwachungs-, Prüf- und Fehlerbehebungs Zwecke innerhalb des Dienstes erforderlich.

Die Richtlinie lehnt die Aktionen auch ausdrücklich ab `cleanrooms:StartProtectedQuery` und verhindert `cleanrooms:UpdateProtectedQuery`, dass Benutzer geschützte Abfragen direkt ausführen oder aktualisieren, was über die AWS Clean Rooms kontrollierten Mechanismen geschehen sollte.

Eine JSON-Liste der Richtliniendetails finden Sie [AWSCleanRoomsFullAccessNoQuerying](#) im Referenzhandbuch für AWS verwaltete Richtlinien.

AWS verwaltete Richtlinie: **AWSCleanRoomsMLReadOnlyAccess**

Sie können eine Verbindung `AWSCleanRoomsMLReadOnlyAccess` zu Ihren IAM-Prinzipalen herstellen.

Diese Richtlinie gewährt nur Leseberechtigungen für Ressourcen und Metadaten in einer Kollaboration. `AWSCleanRoomsMLReadOnlyAccess`

Diese Richtlinie umfasst die folgenden Berechtigungen:

- **CleanRoomsConsoleNavigation**— Gewährt Zugriff auf die Bildschirme der AWS Clean Rooms Konsole.
- **CleanRoomsMLRead**— Ermöglicht Prinzipalen nur Lesezugriff auf den Clean Rooms ML-Dienst.
- **PassCleanRoomsResources**— Gewährt Zugriff zur Weitergabe bestimmter Ressourcen. AWS Clean Rooms

Eine JSON-Liste der Richtliniendetails finden Sie unter [AWSCleanRäume MLRead OnlyAccess](#) im Referenzhandbuch für AWS verwaltete Richtlinien.

AWS verwaltete Richtlinie: **AWSCleanRoomsMLFullAccess**

Sie können eine Verbindung **AWSCleanRoomsMLFullAccess** zu Ihren IAM-Prinzipalen herstellen. Diese Richtlinie gewährt Administratorberechtigungen, die vollen Zugriff (Lesen, Schreiben und Aktualisieren) auf Ressourcen und Metadaten ermöglichen, die von Clean Rooms ML benötigt werden.

Details zu Berechtigungen

Diese Richtlinie umfasst die folgenden Berechtigungen:

- **CleanRoomsMLFullAccess**— Gewährt Zugriff auf alle Clean Rooms ML-Aktionen.
- **PassServiceRole**— Gewährt Zugriff auf die Weitergabe einer Servicерolle nur an den Dienst (die **PassedToService** Bedingung), dessen Name `cleanrooms-ml` enthält.
- **CleanRoomsConsoleNavigation**— Gewährt Zugriff auf die Bildschirme der AWS Clean Rooms Konsole.
- **CollaborationMembershipCheck**— Wenn Sie innerhalb einer Kollaboration einen Job zur Zielgruppengenerierung (Lookalike-Segment) starten, ruft der Clean Rooms ML-Service an, **ListMembers** um zu überprüfen, ob die Kollaboration gültig ist, der Anrufer ein aktives Mitglied und der Besitzer des konfigurierten Zielgruppenmodells ein aktives Mitglied ist. Diese Berechtigung ist immer erforderlich. Die SID für die Konsolennavigation ist nur für Konsolenbenutzer erforderlich.
- **PassCleanRoomsResources**— Gewährt Zugriff auf die Weitergabe bestimmter AWS Clean Rooms Ressourcen.
- **AssociateModels**— Ermöglicht Prinzipalen, Ihrer Zusammenarbeit ein Clean Rooms-ML-Modell zuzuordnen.
- **TagAssociations**— Ermöglicht es Prinzipalen, der Verknüpfung zwischen einem Lookalike-Modell und einer Kollaboration Tags hinzuzufügen.
- **ListRolesToPickServiceRole**— Ermöglicht es Prinzipalen, alle ihre Rollen aufzulisten, um bei der Verwendung eine Servicерolle auszuwählen. AWS Clean Rooms
- **GetRoleAndListRolePoliciesToInspectServiceRole**— Ermöglicht Prinzipalen, die Servicерolle und die entsprechende Richtlinie in IAM zu sehen.
- **ListPoliciesToInspectServiceRolePolicy**— Ermöglicht Prinzipalen, die Servicерolle und die entsprechende Richtlinie in IAM zu sehen.
- **GetPolicyToInspectServiceRolePolicy**— Ermöglicht Prinzipalen, die Servicерolle und die entsprechende Richtlinie in IAM zu sehen.

- **ConsoleDisplayTables**— Ermöglicht Prinzipalen den schreibgeschützten Zugriff auf die AWS Glue Metadaten, die für die Anzeige von Daten zu den zugrunde liegenden AWS Glue Tabellen auf der Konsole erforderlich sind.
- **ConsolePickOutputBucket**— Ermöglicht Prinzipalen die Auswahl von Amazon S3 S3-Buckets für konfigurierte Zielgruppenmodellausgaben.
- **ConsolePickS3Location**— Ermöglicht Prinzipalen die Auswahl des Speicherorts innerhalb eines Buckets für konfigurierte Zielgruppenmodell-Ausgaben.
- **ConsoleDescribeECRRepositories**— Ermöglicht Prinzipalen die Beschreibung von Amazon ECR-Repositorys und -Images.

Eine JSON-Liste der Richtliniendetails finden Sie unter [AWSCleanRooms MLFull Access](#) im AWS Managed Policy Reference Guide.

AWS Clean Rooms Aktualisierungen der AWS verwalteten Richtlinien

Hier finden Sie Informationen zu Aktualisierungen AWS verwalteter Richtlinien AWS Clean Rooms seit Beginn der Nachverfolgung dieser Änderungen durch diesen Dienst. Abonnieren Sie den RSS-Feed auf der Seite AWS Clean Rooms Dokumentenverlauf, um automatische Benachrichtigungen über Änderungen an dieser Seite zu erhalten.

Änderungen	Beschreibung	Date
AWSCleanRoomsFullAccessNoQuering — Aktualisierung der bestehenden Richtlinie	Reinräume: UpdateConfiguredTableAllowedColumns und Reinräume: UpdateConfiguredTableReference zu hinzugefügt. CleanRoomsAccess	29. Juli 2025
AWSCleanRoomsMLReadOnlyAccess – Aktualisierung auf eine bestehende Richtlinie	Hinzufügung von PassCleanRoomsResources zu AWSCleanRoomsMLReadOnlyAccess. PassCleanRoomsResources und ConsoleDescribeECRRepositories wurden zu AWSCleanRoomsMLFullAccess hinzugefügt.	10. Januar 2025

Änderungen	Beschreibung	Date
AWSCleanRoomsFullAccessNoQuering – Aktualisierung auf eine bestehende Richtlinie	Hinzufügung von cleanrooms:BatchGetSchemaAnalysisRule zu CleanRoomsAccess.	13. Mai 2024
AWSCleanRoomsFullAccess – Aktualisierung auf eine bestehende Richtlinie	Die Statement ID in der Zeile AWSCleanRoomsFullAccess von ConsolePickQueryResultsBucket bis wurde SetQueryResultsBucket in dieser Richtlinie aktualisiert, um die Berechtigungen besser darzustellen, da die Berechtigungen für die Einstellung des Abfrageergebnis-Buckets sowohl mit als auch ohne Konsole benötigt werden.	21. März 2024
AWSCleanRoomsMLReadOnlyAccess – Neue Richtlinie AWSCleanRoomsMLFullAccess – Neue Richtlinie	Hinzugefügt AWSCleanRoomsMLReadOnlyAccess und AWSCleanRoomsMLFullAccess zur Unterstützung von AWS Clean Rooms ML.	29. November 2023
AWSCleanRoomsFullAccessNoQuering – Aktualisierung auf eine bestehende Richtlinie	Es wurden cleanrooms:CreateAnalysisTemplatecleanrooms:GetAnalysisTemplate,,cleanrooms:UpdateAnalysisTemplate, cleanrooms:DeleteAnalysisTemplate, cleanrooms:ListAnalysisTemplatescleanrooms:GetCollaborationAnalysisTemplate, und cleanrooms:ListCollaborationAnalysisTemplates zu hinzugefügtcleanrooms:BatchGetCollaborationAnalysisTemplate, CleanRoomsAccess um die neue Funktion für Analysevorlagen zu aktivieren.	31. Juli 2023

Änderungen	Beschreibung	Date
AWSCleanRoomsFullAccessNoQuering – Aktualisierung auf eine bestehende Richtlinie	cleanrooms:ListTagsForResource, und cleanrooms:TagResource zu hinzugefügt cleanrooms:UntagResource, CleanRoomsAccess um das Ressourcen-Tagging zu aktivieren.	21. März 2023
AWS Clean Rooms hat begonnen, Änderungen zu verfolgen	AWS Clean Rooms hat begonnen, Änderungen für die AWS verwalteten Richtlinien zu verfolgen.	12. Januar 2023

Fehlerbehebung bei AWS Clean Rooms Identität und Zugriff

Verwenden Sie die folgenden Informationen, um häufig auftretende Probleme zu diagnostizieren und zu beheben, die bei der Arbeit mit AWS Clean Rooms und IAM auftreten können.

Themen

- [Ich bin nicht berechtigt, eine Aktion durchzuführen in AWS Clean Rooms](#)
- [Ich bin nicht berechtigt, iam auszuführen: PassRole](#)
- [Ich möchte Personen außerhalb von mir den Zugriff AWS-Konto auf meine AWS Clean Rooms Ressourcen ermöglichen](#)

Ich bin nicht berechtigt, eine Aktion durchzuführen in AWS Clean Rooms

Wenn Sie eine Fehlermeldung erhalten, dass Sie nicht zur Durchführung einer Aktion berechtigt sind, müssen Ihre Richtlinien aktualisiert werden, damit Sie die Aktion durchführen können.

Der folgende Beispielfehler tritt auf, wenn der mateojackson IAM-Benutzer versucht, die Konsole zum Anzeigen von Details zu einer fiktiven *my-example-widget*-Ressource zu verwenden, jedoch nicht über `cleanrooms:GetWidget`-Berechtigungen verfügt.

```
User: arn:aws:iam::123456789012:user/mateojackson is not authorized to perform:
cleanrooms:GetWidget on resource: my-example-widget
```

In diesem Fall muss die Mateo-Richtlinie aktualisiert werden, damit er mit der `cleanrooms:GetWidget`-Aktion auf die *my-example-widget*-Ressource zugreifen kann.

Wenn Sie Hilfe benötigen, wenden Sie sich an Ihren AWS Administrator. Ihr Administrator hat Ihnen Ihre Anmeldeinformationen zur Verfügung gestellt.

Ich bin nicht berechtigt, iam auszuführen: PassRole

Wenn Sie die Fehlermeldung erhalten, dass Sie nicht zum Durchführen der `iam:PassRole`-Aktion autorisiert sind, müssen Ihre Richtlinien aktualisiert werden, um eine Rolle an AWS Clean Rooms übergeben zu können.

Einige AWS-Services ermöglichen es Ihnen, eine bestehende Rolle an diesen Dienst zu übergeben, anstatt eine neue Servicerolle oder eine dienstverknüpfte Rolle zu erstellen. Hierzu benötigen Sie Berechtigungen für die Übergabe der Rolle an den Dienst.

Der folgende Beispielfehler tritt auf, wenn ein IAM-Benutzer mit dem Namen `marymajor` versucht, die Konsole zu verwenden, um eine Aktion in AWS Clean Rooms auszuführen. Die Aktion erfordert jedoch, dass der Service über Berechtigungen verfügt, die durch eine Servicerolle gewährt werden. Mary besitzt keine Berechtigungen für die Übergabe der Rolle an den Dienst.

```
User: arn:aws:iam::123456789012:user/marymajor is not authorized to perform:
iam:PassRole
```

In diesem Fall müssen die Richtlinien von Mary aktualisiert werden, um die Aktion `iam:PassRole` ausführen zu können.

Wenn Sie Hilfe benötigen, wenden Sie sich an Ihren AWS Administrator. Ihr Administrator hat Ihnen Ihre Anmeldeinformationen zur Verfügung gestellt.

Ich möchte Personen außerhalb von mir den Zugriff AWS-Konto auf meine AWS Clean Rooms Ressourcen ermöglichen

Sie können eine Rolle erstellen, mit der Benutzer in anderen Konten oder Personen außerhalb Ihrer Organisation auf Ihre Ressourcen zugreifen können. Sie können festlegen, wem die Übernahme der Rolle anvertraut wird.

Weitere Informationen dazu finden Sie hier:

- Informationen darüber, ob diese Funktionen AWS Clean Rooms unterstützt werden, finden Sie unter [Wie AWS Clean Rooms funktioniert mit IAM](#).

- Informationen dazu, wie Sie Zugriff auf Ihre Ressourcen gewähren können, AWS-Konten die Ihnen gehören, finden Sie im IAM-Benutzerhandbuch unter [Gewähren des Zugriffs auf einen IAM-Benutzer in einem anderen AWS-Konto , den Sie besitzen](#).
- Informationen dazu, wie Sie Dritten Zugriff auf Ihre Ressourcen gewähren können AWS-Konten, finden Sie [AWS-Konten im IAM-Benutzerhandbuch unter Gewähren des Zugriffs für Dritte](#).
- Informationen dazu, wie Sie über einen Identitätsverbund Zugriff gewähren, finden Sie unter [Gewähren von Zugriff für extern authentifizierte Benutzer \(Identitätsverbund\)](#) im IAM-Benutzerhandbuch.
- Informationen zum Unterschied zwischen der Verwendung von Rollen und ressourcenbasierten Richtlinien für den kontoübergreifenden Zugriff finden Sie unter [So unterscheiden sich IAM-Rollen von ressourcenbasierten Richtlinien](#) im IAM-Benutzerhandbuch.

Serviceübergreifende Confused-Deputy-Prävention

Das Problem des verwirrten Stellvertreters ist ein Sicherheitsproblem, bei dem eine Entität, die keine Berechtigung zur Durchführung einer Aktion hat, eine privilegiere Entität zur Durchführung der Aktion zwingen kann. In AWS kann ein dienstübergreifendes Identitätswechsels zum Problem des verwirrten Stellvertreters führen. Ein dienstübergreifender Identitätswechsel kann auftreten, wenn ein Dienst (der Anruf-Dienst) einen anderen Dienst anruft (den aufgerufenen Dienst). Der aufrufende Service kann manipuliert werden, um seine Berechtigungen zu verwenden, um Aktionen auf die Ressourcen eines anderen Kunden auszuführen, für die er sonst keine Zugriffsberechtigung haben sollte. Um dies zu verhindern, bietet AWS Tools, mit denen Sie Ihre Daten für alle Services mit Serviceprinzipalen schützen können, die Zugriff auf Ressourcen in Ihrem Konto erhalten haben.

Wir empfehlen, die [aws:SourceArn](#) globalen Bedingungskontextschlüssel in Ressourcenrichtlinien zu verwenden, um die Berechtigungen einzuschränken, die der AWS Clean Rooms Ressource einen anderen Dienst gewähren. Verwenden Sie `aws:SourceArn`, wenn Sie nur eine Ressource mit dem betriebsübergreifenden Zugriff verknüpfen möchten.

Der effektivste Weg, um sich vor dem Confused-Deputy-Problem zu schützen, ist die Verwendung des globalen Bedingungskontext-Schlüssels `aws:SourceArn` mit dem vollständigen ARN der Ressource. AWS Clean Rooms müssen Sie auch mit dem `sts:ExternalId` Bedingungsschlüssel vergleichen.

Der Wert von `aws:SourceArn` muss auf den ARN der Mitgliedschaft der übernommenen Rolle gesetzt werden.

Das folgende Beispiel zeigt, wie Sie den Kontextschlüssel für `aws:SourceArn` globale Bedingungen verwenden können, AWS Clean Rooms um das Problem des verwirrten Stellvertreters zu vermeiden.

 Note

Die Beispielrichtlinie bezieht sich auf die Vertrauensrichtlinie der Servicerolle, die für den Zugriff auf Daten und Metadaten für eine konfigurierte Tabelle AWS Clean Rooms verwendet wird.

Der Wert für `<query-runner-membership-id>` muss auf die Mitglieds-ID des Abfragerunners gesetzt werden.

Alle Mitglieder der Kollaboration können die konfigurierten Tabellenmetadaten anzeigen, sodass jeder Mitglieds-ARN in der Mitgliedschaftsliste enthalten sein muss ARNs.

 Note

Wenn eine Servicerolle über die AWS Clean Rooms Konsole erstellt wird, werden standardmäßig alle aktuellen Mitglieder der Kollaboration in den Status Confused Deputy aufgenommen.

Wenn Sie einer Kollaboration, der bereits konfigurierte Tabellen zugeordnet sind, neue Mitglieder hinzufügen, stellen Sie sicher, dass Sie den Status Confused Deputy Ihrer Servicerolle mit dem Mitglieds-ARN des neuen Mitglieds aktualisieren.

Wenn Sie nach dem Hinzufügen eines neuen Mitglieds den Status „Verwirrter Stellvertreter“ Ihrer Servicerolle nicht aktualisieren, kann das neue Mitglied nicht auf die Informationen zugreifen AWS Clean Rooms , die mit dieser Rolle abgerufen wurden.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowIfExternalIdMatches",
      "Effect": "Allow",
      "Principal": {
        "Service": "cleanrooms.amazonaws.com"
      },
    },
  ],
}
```

```

        "Action": "sts:AssumeRole",
        "Condition": {
            "StringLike": {
                "sts:ExternalId": "arn:aws:*:us-east-1::dbuser:*/<query-  

runner-membership-id>*"
            }
        },
    },
    {
        "Sid": "AllowIfSourceArnMatches",
        "Effect": "Allow",
        "Principal": {
            "Service": "cleanrooms.amazonaws.com"
        },
        "Action": "sts:AssumeRole",
        "Condition": {
            "ForAnyValue:ArnEquals": {
                "aws:SourceArn": [
                    "arn:aws:cleanrooms:us-east-1:111122223333:membership/<member-1-membership-id>",
                    "arn:aws:cleanrooms:us-east-1:444455556666:membership/<member-2-membership-id>"
                ]
            }
        }
    }
]
}

```

IAM-Verhalten für AWS Clean Rooms ML

Kontoübergreifende Jobs

Mit Clean Rooms ML können bestimmte Ressourcen, die von einer Person erstellt wurden AWS-Konto , von einer anderen AWS-Konto Person sicher in ihrem Konto abgerufen werden. Wenn ein Kunde in AWS-Konto A eine ConfiguredAudienceModel Ressource StartAudienceGenerationJob anruft, die AWS-Konto B gehört, erstellt Clean Rooms ML zwei Ressourcen ARNs für den Job. Ein ARN in AWS-Konto A und ein weiterer in AWS-Konto B. Sie ARNs sind bis auf ihre identisch AWS-Konto.

Clean Rooms ML erstellt zwei ARNs für den Job, um sicherzustellen, dass beide Konten ihre eigenen IAM-Richtlinien auf die Jobs anwenden können. Beispielsweise können beide Konten eine tagbasierte Zugriffskontrolle verwenden und die Richtlinien ihrer AWS Organisation anwenden. Der Job verarbeitet Daten von beiden Konten, sodass beide Konten den Job und die zugehörigen Daten löschen können. Keines der Konten kann das andere Konto daran hindern, den Job zu löschen.

Es gibt nur eine Auftragsausführung und beide Konten können den Job sehen, wenn sie aufrufen `ListAudienceGenerationJobs`. Beide Konten können `Get`, und `Export` APIs on the Job aufrufen `Delete`, indem sie den ARN mit ihrer eigenen AWS-Konto ID verwenden.

Keiner AWS-Konto kann auf den Job zugreifen, wenn er einen ARN mit der anderen AWS-Konto ID verwendet.

Der Name des Jobs muss innerhalb eines eindeutig sein AWS-Konto. Der Name in AWS-Konto B ist `$accountA-$name`. Dem von AWS-Konto A ausgewählten Namen wird A vorangestellt, wenn der Job in AWS-Konto B angezeigt wird. AWS-Konto

Damit ein Cross-Account `StartAudienceGenerationJob` erfolgreich ist, muss AWS-Konto B diese Aktion sowohl für den neuen Job in B als auch für den Job in AWS-Konto B zulassen. Dabei `ConfiguredAudienceModel` muss eine Ressourcenrichtlinie verwendet werden, die dem folgenden Beispiel ähnelt: AWS-Konto

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "Clean-Rooms-CAMA-ID",
      "Effect": "Allow",
      "Principal": {
        "AWS": [
          "111122223333"
        ]
      },
      "Action": [
        "cleanrooms-ml:StartAudienceGenerationJob"
      ],
      "Resource": [
        "arn:aws:cleanrooms-ml:us-east-1:444455556666:configured-audience-model/id",

```



```

        "arn:aws:cleanrooms-ml:us-east-1:444455556666:audience-
generation-job/*"
      ],
      "Condition":{"StringEquals":{"cleanrooms-ml:CollaborationId":"UUID"}}
    }
  ]
}

```

Note

Diese AWS Clean Rooms ML-Ressourcenrichtlinie bezieht sich auf zwei unterschiedliche Methoden, AWS-Konto IDs um die kontenübergreifende Zielgruppengenerierung zu unterstützen:

- 111122223333 — Dies ist das Konto, das den Prinzipal (Benutzer, Rolle oder Dienst) enthält, der autorisiert ist, Jobs zur Zielgruppengenerierung zu starten. Dieses Konto initiiert den ML-Verarbeitungsworkflow.
- 444455556666 — Dies ist das Konto, dem die AWS Clean Rooms ML-Ressourcen gehören (das konfigurierte Zielgruppenmodell und die Jobs zur Zielgruppengenerierung). Dieses Konto hostet die ML-Modelle und verwaltet die Auftragsausführung.

Zusätzliche Hinweise zur Konfiguration:

- Statement ID (Sid): CAMA-ID Ersetzen Sie es durch Ihre tatsächliche AWS Clean Rooms Audience Model Application (CAMA) -ID, damit die Grundsatzerklärung leicht identifizierbar ist.
- Ressource IDs: **id** Ersetzen Sie diese durch die tatsächliche ID Ihres konfigurierten Zielgruppenmodells und **UUID** durch Ihre spezifische Kollaborations-ID.
- Bedingung: Die `cleanrooms-ml:CollaborationId` Bedingung stellt sicher, dass Jobs zur Zielgruppengenerierung nur im Kontext der angegebenen AWS Clean Rooms Zusammenarbeit gestartet werden können, wodurch eine zusätzliche Sicherheitsgrenze geschaffen wird.

Diese kontenübergreifende Konfiguration ermöglicht Szenarien, in denen eine Organisation die ML-Modelle und die Infrastruktur verwaltet und es autorisierten Partnern

gleichzeitig ermöglicht, im Rahmen ihrer Kooperationsvereinbarung Prozesse zur Zielgruppengenerierung einzuleiten.

Wenn Sie die [AWS Clean Rooms ML-API](#) verwenden, um ein konfiguriertes Lookalike-Modell mit dem Wert `manageResourcePolicies true` zu erstellen, AWS Clean Rooms erstellt diese Richtlinie für Sie.

Darüber hinaus muss für die Identitätsrichtlinie des Anrufers in AWS-Konto A eine `StartAudienceGenerationJob` Genehmigung aktiviert werden. `arn:aws:cleanrooms-ml:us-west-1:AccountA:audience-generation-job/*` Es gibt also drei IAM-Ressourcen für Aktionen `StartAudienceGenerationJob`: den AWS-Konto A-Job, den AWS-Konto B-Job und den AWS-Konto B-Job. `ConfiguredAudienceModel`

Warning

Derjenige AWS-Konto, der den Job gestartet hat, erhält ein AWS CloudTrail Audit-Log-Ereignis über den Job. AWS-Konto Derjenige, dem der gehört, empfängt `ConfiguredAudienceModel` kein AWS CloudTrail Überwachungsprotokollereignis.

Jobs taggen

Wenn Sie den `childResourceTagOnCreatePolicy=FROM_PARENT_RESOURCE` Parameter von `festlegenCreateConfiguredAudienceModel`, haben alle Jobs zur Generierung von Lookalike-Segmenten in Ihrem Konto, die anhand dieses konfigurierten Lookalike-Modells erstellt wurden, standardmäßig dieselben Tags wie das konfigurierte Lookalike-Modell. Das konfigurierte Lookalike-Modell ist das übergeordnete Modell und der Job zur Generierung von Lookalike-Segmenten ist das untergeordnete Modell.

Wenn Sie einen Job in Ihrem eigenen Konto erstellen, haben die Anforderungs-Tags des Jobs Vorrang vor den übergeordneten Tags. Jobs, die von anderen Konten erstellt wurden, erzeugen niemals Stichwörter in Ihrem Konto. Wenn Sie einen Job einrichten `childResourceTagOnCreatePolicy=FROM_PARENT_RESOURCE` und ein anderer Account erstellt, gibt es zwei Kopien des Jobs. Die Kopie in Ihrem Konto enthält die übergeordneten Ressourcen-Tags und die Kopie im Konto des Jobeinreichers enthält Stichwörter aus der Anfrage.

Mitarbeiter werden validiert

Wenn Sie anderen Mitgliedern einer AWS Clean Rooms Kollaboration Berechtigungen gewähren, sollte die Ressourcenrichtlinie den Bedingungsschlüssel enthalten. `cleanrooms-ml:CollaborationId` Dadurch wird erzwungen, dass der `collaborationId` Parameter in der [StartAudienceGenerationJob](#)Anfrage enthalten ist. Wenn der `collaborationId` Parameter in der Anfrage enthalten ist, überprüft Clean Rooms ML, ob die Kollaboration existiert, dass der Jobeinreicher ein aktives Mitglied der Kollaboration ist und der Besitzer des konfigurierten Lookalike-Modells ein aktives Mitglied der Kollaboration ist.

Wenn Ihre konfigurierte Ressourcenrichtlinie für das Lookalike-Modell AWS Clean Rooms verwaltet wird (der `manageResourcePolicies` Parameter ist [CreateConfiguredAudienceModelAssociation](#) [angefordert](#)), wird dieser Bedingungsschlüssel `TRUE` in der Ressourcenrichtlinie festgelegt. Daher müssen Sie den `collaborationId` in [StartAudienceGenerationJob](#) angeben.

Kontoübergreifender Zugriff

`StartAudienceGenerationJob` Kann nur kontenübergreifend aufgerufen werden. Alle anderen Clean Rooms ML APIs können nur mit Ressourcen in Ihrem eigenen Konto verwendet werden. Dadurch wird sichergestellt, dass Ihre Trainingsdaten, die Konfiguration eines Lookalike-Modells und andere Informationen vertraulich bleiben.

Clean Rooms ML gibt niemals Amazon S3 oder AWS Glue Standorte für mehrere Konten preis. Der Speicherort der Trainingsdaten, der konfigurierte Speicherort für die Ausgabe eines Lookalike-Modells und der Standort der Jobstartdaten für die Lookalike-Segmentgenerierung sind nicht für alle Konten sichtbar. Sofern die Abfrageprotokollierung in der Kollaboration nicht aktiviert ist, ist es nicht kontenübergreifend sichtbar, ob die Ausgangsdaten aus einer SQL-Abfrage stammen, und die Abfrage selbst. Wenn Sie Get einen Job zur Zielgruppengenerierung haben, der von einem anderen Account eingereicht wurde, zeigt der Service den Ausgangsort nicht an.

IAM-Verhalten für benutzerdefinierte Clean Rooms ML-Modelle

Kontoübergreifende Jobs

Clean Rooms ML ermöglicht den sicheren Zugriff auf bestimmte Ressourcen, die mit einer von einer Person erstellten Zusammenarbeit verknüpft sind AWS-Konto , von einer anderen AWS-Konto Person in ihrem Konto sicher abgerufen werden. Ein Client in AWS-Konto A, der über die Fähigkeit eines Mitglieds verfügt, Abfragen auszuführen `CreateTrainedModelCreateMLInputChannel`, kann eine `ConfiguredModelAlgorithmAssociation` Ressource aufrufen, die

einem anderen Mitglied der Kollaboration gehört, sofern dies durch die benutzerdefinierte Analyseregeln, die mit erstellt wurde, zulässig `ConfiguredModelAlgorithmAssociation` `isCreateConfiguredTableAnalysisRule`. `StartTrainedModelInferenceJob`

Darüber hinaus kann jedes aktive Mitglied einer Kollaboration Daten löschen, die einem trainierten Modell oder einem ML-Eingangskanal zugeordnet sind, über den Befehl `DeleteTrainedModelOutput` und `DeleteMLInputChannelData` APIs.

Kontoübergreifender Zugriff

Clean Rooms ML ermöglicht es Benutzern, Metadaten zu Ressourcen, die von anderen Konten erstellt wurden, über `GetCollaboration` und abzurufen `ListCollaboration` APIs. Clean Rooms ML gibt keine KMS-Schlüssel ARNs, -Tags, Umgebungsvariablen oder Hyperparameter (für die `TrainedModel` Aktion) an andere Konten weiter.

Zugriff auf Mitgliedschaft und Zusammenarbeit

Beim Zugriff auf Mitgliedschafts- und Kollaborationsressourcen im Kontext von benutzerdefinierten Clean Rooms ML-Modellen benötigt die Identitätsrichtlinie eines Benutzers Berechtigungen für die Aktionen `cleanrooms:PassMembership` oder beides. `cleanrooms:PassCollaboration` Alle APIs, die zustimmen, `membershipId` benötigen die `cleanrooms:PassMembership` Erlaubnis, und alle APIs, die zustimmen, `collaborationId` benötigen die `cleanrooms:PassCollaboration` Erlaubnis. Es wird ein Beispiel für eine Identitätsrichtlinie für eine Rolle bereitgestellt, die `createTrainedModel` im Kontext einer Mitglieds-ID aufrufen kann, die `GetCollaborationTrainedModel` im Kontext einer Kollaborations-ID aufgerufen werden kann.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowCleanroomsMLActions",
      "Effect": "Allow",
      "Action": [
        "cleanrooms:PassCollaboration",
        "cleanrooms:PassMembership"
      ],
    },
  ],
}
```

```

        "Resource": [
            "*"
        ]
    },
    {
        "Sid": "AllowMembershipAccess",
        "Effect": "Allow",
        "Action": [
            "cleanrooms:GetMembership"
        ],
        "Resource": [
            "arn:aws:cleanrooms:us-east-1:111122223333:membership/memberId"
        ]
    },
    {
        "Sid": "AllowCollaborationAccess",
        "Effect": "Allow",
        "Action": [
            "cleanrooms:GetCollaboration"
        ],
        "Resource": [
            "arn:aws:cleanrooms:us-east-1:111122223333:collaboration/collaborationId"
        ]
    }
]
}

```

Konformitätsprüfung für AWS Clean Rooms

Informationen darüber, ob AWS-Service ein [AWS-Services in den Geltungsbereich bestimmter Compliance-Programme fällt, finden Sie unter Umfang nach Compliance-Programm AWS-Services unter](#) . Wählen Sie dort das Compliance-Programm aus, an dem Sie interessiert sind. Allgemeine Informationen finden Sie unter [AWS Compliance-Programme AWS](#) .

Sie können Prüfberichte von Drittanbietern unter herunterladen AWS Artifact. Weitere Informationen finden Sie unter [Berichte herunterladen unter](#) .

Ihre Verantwortung für die Einhaltung der Vorschriften bei der Nutzung AWS-Services hängt von der Vertraulichkeit Ihrer Daten, den Compliance-Zielen Ihres Unternehmens und den geltenden Gesetzen

und Vorschriften ab. Weitere Informationen zu Ihrer Verantwortung für die Einhaltung der Vorschriften bei der Nutzung AWS-Services finden Sie in der [AWS Sicherheitsdokumentation](#).

Resilienz in AWS Clean Rooms

Die AWS globale Infrastruktur basiert auf AWS Regionen und Availability Zones. Regionen stellen mehrere physisch getrennte und isolierte Availability Zones bereit, die über hoch redundante Netzwerke mit niedriger Latenz und hohen Durchsätzen verbunden sind. Mithilfe von Availability Zones können Sie Anwendungen und Datenbanken erstellen und ausführen, die automatisch Failover zwischen Zonen ausführen, ohne dass es zu Unterbrechungen kommt. Availability Zones sind besser verfügbar, fehlertoleranter und skalierbarer als herkömmliche Infrastrukturen mit einem oder mehreren Rechenzentren.

Weitere Informationen zu AWS Regionen und Availability Zones finden Sie unter [AWS Globale Infrastruktur](#).

Sicherheit der Infrastruktur in AWS Clean Rooms

Als verwalteter Dienst AWS Clean Rooms ist er durch AWS globale Netzwerksicherheit geschützt. Informationen zu AWS Sicherheitsdiensten und zum AWS Schutz der Infrastruktur finden Sie unter [AWS Cloud-Sicherheit](#). Informationen zum Entwerfen Ihrer AWS Umgebung unter Verwendung der bewährten Methoden für die Infrastruktursicherheit finden Sie unter [Infrastructure Protection](#) in Security Pillar AWS Well-Architected Framework.

Sie verwenden AWS veröffentlichte API-Aufrufe für den Zugriff AWS Clean Rooms über das Netzwerk. Kunden müssen Folgendes unterstützen:

- Transport Layer Security (TLS). Wir benötigen TLS 1.2 und empfehlen TLS 1.3.
- Verschlüsselungs-Suiten mit Perfect Forward Secrecy (PFS) wie DHE (Ephemeral Diffie-Hellman) oder ECDHE (Elliptic Curve Ephemeral Diffie-Hellman). Die meisten modernen Systeme wie Java 7 und höher unterstützen diese Modi.

Netzwerksicherheit

Wenn während der Abfrageausführung aus Ihrem S3-Bucket AWS Clean Rooms gelesen wird, wird der Datenverkehr zwischen Amazon S3 AWS Clean Rooms und Amazon S3 sicher durch das AWS private Netzwerk geleitet. Der Flugverkehr wird mit dem Amazon Signature Version 4-Protokoll

(SIGv4) signiert und mit HTTPS verschlüsselt. Dieser Datenverkehr wird auf der Grundlage der IAM-Servicerolle autorisiert, die Sie für Ihre konfigurierte Tabelle eingerichtet haben.

Sie können programmgesteuert eine Verbindung AWS Clean Rooms über einen Endpunkt herstellen. Eine Liste der Dienstendpunkte finden Sie unter [AWS Clean Rooms Endpunkte und Kontingente](#) in der Allgemeinen AWS-Referenz.

Alle Dienstendpunkte sind nur für HTTPS verfügbar. Sie können Amazon Virtual Private Cloud (VPC)-Endpunkte verwenden, falls Sie AWS Clean Rooms von Ihrer VPC aus eine Verbindung herstellen möchten und keine Internetverbindung haben möchten. Weitere Informationen finden Sie unter [Access AWS services through AWS PrivateLink im Handbuch](#). AWS PrivateLink

Sie können Ihren IAM-Prinzipalen IAM-Richtlinien zuweisen, die die [SourceVpce aws:-Kontextschlüssel](#) verwenden, um Ihren IAM-Prinzipal darauf zu beschränken, Anrufe nur AWS Clean Rooms über einen VPC-Endpunkt und nicht über das Internet tätigen zu können.

Zugriff AWS Clean Rooms oder AWS Clean Rooms ML über einen Schnittstellen-Endpunkt ()AWS PrivateLink

Sie können AWS PrivateLink damit eine private Verbindung zwischen Ihrer Virtual Private Cloud (VPC) AWS Clean Rooms und/oder AWS Clean Rooms ML herstellen. Sie können auf AWS Clean Rooms oder AWS Clean Rooms ML zugreifen, als ob es in Ihrer VPC wäre, ohne ein Internet-Gateway, ein NAT-Gerät, eine VPN-Verbindung oder Direct Connect eine Verbindung verwenden zu müssen. Instances in Ihrer VPC benötigen für den Zugriff AWS Clean Rooms keine öffentlichen IP-Adressen.

Sie stellen diese private Verbindung her, indem Sie einen Schnittstellen-Endpunkt erstellen, der von AWS PrivateLink unterstützt wird. Wir erstellen eine Endpunkt-Netzwerkschnittstelle in jedem Subnetz, das Sie für den Schnittstellen-Endpunkt aktivieren. Hierbei handelt es sich um vom Anforderer verwaltete Netzwerkschnittstellen, die als Eingangspunkt für den Datenverkehr dienen, der für AWS Clean Rooms bestimmt ist.

Weitere Informationen finden Sie unter [Zugriff auf AWS-Services über AWS PrivateLink](#) im AWS PrivateLink -Leitfaden.

Überlegungen zu AWS Clean Rooms

Bevor Sie einen Schnittstellenendpunkt für einrichten AWS Clean Rooms, lesen Sie die [Überlegungen](#) im AWS PrivateLink Handbuch.

AWS Clean Rooms und AWS Clean Rooms ML unterstützen das Aufrufen all ihrer API-Aktionen über den Schnittstellenendpunkt.

VPC-Endpunktrichtlinien werden für AWS Clean Rooms oder AWS Clean Rooms ML nicht unterstützt. Standardmäßig ist Vollzugriff auf AWS Clean Rooms und AWS Clean Rooms ML über den Schnittstellenendpunkt zulässig. Alternativ können Sie den Endpunkt-Netzwerkschnittstellen eine Sicherheitsgruppe zuordnen, um den Datenverkehr zum AWS Clean Rooms oder AWS Clean Rooms ML über den Schnittstellenendpunkt zu steuern.

Erstellen Sie einen Schnittstellenendpunkt für AWS Clean Rooms

Sie können einen Schnittstellenendpunkt für AWS Clean Rooms oder AWS Clean Rooms ML entweder mit der Amazon VPC-Konsole oder mit AWS Command Line Interface (AWS CLI) erstellen. Weitere Informationen finden Sie unter [Erstellen eines Schnittstellenendpunkts](#) im AWS PrivateLink - Leitfaden.

Erstellen Sie einen Schnittstellenendpunkt für die AWS Clean Rooms Verwendung des folgenden Servicenamens.

```
com.amazonaws.region.cleanrooms
```

Erstellen Sie einen Schnittstellenendpunkt für AWS Clean Rooms ML mit dem folgenden Dienstnamen.

```
com.amazonaws.region.cleanrooms-ml
```

Wenn Sie privates DNS für den Schnittstellenendpunkt aktivieren, können Sie API-Anfragen an die AWS Clean Rooms Verwendung des standardmäßigen regionalen DNS-Namens stellen. Beispiel, `cleanrooms-ml.us-east-1.amazonaws.com`.

Überwachung AWS Clean Rooms

Die Überwachung ist ein wichtiger Bestandteil der Aufrechterhaltung der Zuverlässigkeit, Verfügbarkeit und Leistung Ihrer AWS Clean Rooms anderen AWS Lösungen. AWS bietet die folgenden Überwachungstools, mit denen Sie beobachten AWS Clean Rooms, melden können, wenn etwas nicht stimmt, und gegebenenfalls automatische Maßnahmen ergreifen können:

- Mit Amazon CloudWatch Logs können Sie Ihre Protokolldateien von Amazon EC2 EC2-Instances und anderen Quellen überwachen AWS CloudTrail, speichern und darauf zugreifen. Amazon CloudWatch Logs kann Informationen in den Protokolldateien überwachen und Sie benachrichtigen, wenn bestimmte Schwellenwerte erreicht werden. Sie können Ihre Protokolldaten auch in einem sehr robusten Speicher archivieren. Weitere Informationen finden Sie im [Amazon CloudWatch Logs-Benutzerhandbuch](#).

Clean Rooms ML ermöglicht kontenübergreifende Jobs für bestimmte API-Aktionen. Derjenige AWS-Konto, der den Job gestartet hat, erhält das AWS CloudTrail Audit-Log-Ereignis für den Job. Weitere Informationen finden Sie unter [IAM-Verhalten für AWS Clean Rooms ML](#).

- AWS CloudTrailerfasst API-Aufrufe und zugehörige Ereignisse, die von Ihnen oder in Ihrem Namen getätigt wurden, AWS-Konto und übermittelt die Protokolldateien an einen von Ihnen angegebenen Amazon S3 S3-Bucket. Sie können feststellen, welche Benutzer und Konten angerufen wurden AWS, von welcher Quell-IP-Adresse aus die Anrufe getätigt wurden und wann die Aufrufe erfolgten. Weitere Informationen finden Sie im [AWS CloudTrail -Benutzerhandbuch](#).
- Amazon EventBridge ist ein serverloser Event-Bus-Service, der es einfach macht, Ihre Anwendungen mit Daten aus einer Vielzahl von Quellen zu verbinden. EventBridge liefert einen Stream von Echtzeitdaten aus Ihren eigenen Anwendungen, Software-as-a-Service (SaaS-) Anwendungen und AWS Diensten und leitet diese Daten an Ziele wie Lambda weiter. Auf diese Weise können Sie Ereignisse überwachen, die in Services auftreten, und ereignisgesteuerte Architekturen erstellen. Weitere Informationen finden Sie im [EventBridge Amazon-Benutzerhandbuch](#) und in der [Amazon EventBridge Events-Referenz](#).
- AWS Fakturierung und Kostenmanagementist ein Service, der Funktionen bietet, mit denen Sie Ihre Rechnungen bezahlen und Ihre Kosten optimieren können. Amazon Web Services berechnet Ihr Konto nutzungsabhängig und stellt so sicher, dass Sie nur für das bezahlen, was Sie tatsächlich nutzen. AWS ermöglicht es Ihnen auch, [Kostenzuweisungs-Tags auf](#) Ihre Ressourcen anzuwenden, um Ihre AWS-Kosten zu verfolgen und zu kategorisieren. In können Sie AWS Clean Rooms beispielsweise Tags auf Kollaborationsmitgliedschaften anwenden, um die

Kosten pro Zusammenarbeit nachzuverfolgen. Weitere Informationen finden Sie im [AWS Billing - Benutzerhandbuch](#).

Analyse Einloggen AWS Clean Rooms

Die Analyseprotokollierung ist eine Funktion in AWS Clean Rooms. Wenn Sie [eine Kollaboration erstellen](#) und die Analysis-Protokollierung aktivieren, können Mitglieder relevante Protokolle von Abfragen oder Protokolle von Jobs in Amazon CloudWatch Logs speichern.

Anhand von Abfrageprotokollen und Jobprotokollen können Mitglieder feststellen, ob die Anfragen den Analyseregeln entsprechen und mit der Kooperationsvereinbarung übereinstimmen. Darüber hinaus unterstützen Abfrageprotokolle Audits.

Wenn die Option Analyse-Protokollierung in der AWS Clean Rooms Konsole aktiviert ist, enthalten die Abfrageprotokolle Folgendes:

- `analysisRule`— Die Analyseregeln für die konfigurierte Tabelle.
- `analysisTemplateArn`— Die Analysevorlage, die ausgeführt wurde (wird je nach Analyseregeln angezeigt).
- `collaborationId`— Die eindeutige Kennung für die Zusammenarbeit, in der die Abfrage ausgeführt wurde.
- `configuredTableID`— Die eindeutige Kennung für die konfigurierte Tabelle, auf die in der Abfrage verwiesen wird.
- `directQueryAnalysisRulePolicy.custom.allowedAnalysis`— Die Analysevorlage, die für die konfigurierte Tabelle ausgeführt werden darf (wird je nach Analyseregeln angezeigt).
- `directQueryAnalysisRulePolicy.v1.custom.allowedAnalysisProviders`— Die Abfrageanbieter, die eine Abfrage erstellen dürfen (wird je nach Analyseregeln angezeigt).
- `errorCode`— Der Fehlercode, wenn eine Abfrage nicht ordnungsgemäß ausgeführt werden konnte.
- `errorMessage`— Die Fehlermeldung, wenn eine Abfrage nicht ordnungsgemäß ausgeführt werden konnte.
- `eventID`— Die eindeutige Kennung für den Abfragelauf. Nach dem 31. August 2023 ist der eindeutige Bezeichner derselbe wie `derprotectedQueryID`.
- `eventTimestamp`— Die Laufzeit der Abfrage.
- `parameters.parameterValue`— Die Parameterwerte (erscheint je nach Abfragetext).

- `queryText`— Die SQL-Definition von Query Run. Wenn es Parameter gibt, werden sie als `:parametervalue` gekennzeichnet.
- `queryValidationErrors`— Die Abfragefehler bei der Abfragevalidierung.
- `schemaName`— Der Name der konfigurierten Tabellenverknüpfung, auf die in der Abfrage verwiesen wird.
- `status`— Der Ausführungsstatus der Abfrage.

Empfangen von Abfrage- und Jobprotokollen

Sie müssen außer der Einrichtung von AWS Clean Rooms Abfrageprotokollen und Jobprotokollen keine weiteren Aktionen ausführen. AWS Clean Rooms erstellt Protokollgruppen für Kollaborationen, nachdem jedes Kollaborationsmitglied [eine Mitgliedschaft erstellt hat](#).

Mitglieder, die Abfragen durchführen können, Mitglieder, die Abfragen und Jobs ausführen können, Mitglieder, die Ergebnisse empfangen können, und Mitglieder, auf deren Konfigurationstabellen in der Abfrage verwiesen wird, erhalten ein Abfrageprotokoll oder ein Jobprotokoll.

Das Mitglied, das Abfragen durchführen kann, und das Mitglied, das Ergebnisse empfangen kann, erhalten Abfrageprotokolle für jede konfigurierte Tabelle, auf die in der Abfrage verwiesen wird. Wenn sie nicht Eigentümer der konfigurierten Tabelle sind, können sie die konfigurierte Tabellen-ID (`configuredTableID`) nicht einsehen.

Das Mitglied, das Abfragen und Jobs ausführen kann, und das Mitglied, das Ergebnisse empfangen kann, erhalten Jobprotokolle für jede konfigurierte Tabelle, auf die im Job verwiesen wird. Wenn sie nicht Eigentümer der konfigurierten Tabelle sind, können sie die konfigurierte Tabellen-ID (`configuredTableID`) nicht einsehen.

Wenn ein Mitglied über mehrere konfigurierte Tabellenzuordnungen verfügt, auf die in der Abfrage verwiesen wird, erhält es für jede konfigurierte Tabelle ein Abfrageprotokoll.

Wenn ein Mitglied über mehrere konfigurierte Tabellenzuordnungen verfügt, auf die im Job verwiesen wird, erhält es für jede konfigurierte Tabelle ein Jobprotokoll.

Protokolle werden für Abfragen erstellt, die nicht unterstütztes und unterstütztes SQL in AWS Clean Rooms enthalten. Weitere Informationen finden Sie in der [AWS Clean Rooms SQL-Referenz](#).

Protokolle werden auch erstellt, wenn Abfragen oder Jobs auf konfigurierte Tabellen verweisen, die nicht mit der Kollaboration verknüpft sind.

Protokolle können Informationen über falsches SQL enthalten.

Abfrage- und Auftragsprotokolle geben den Status einer Abfrage an, geben jedoch nicht an, ob die Abfrageausgabe zugestellt wurde. Sie bestätigen, dass eine Anfrage oder ein Auftrag von dem Mitglied eingereicht wurde, das eine Anfrage abfragen kann. Abfrageprotokolle bestätigen auch, dass die Abfrage unterstütztes SQL in enthält AWS Clean Rooms und auf konfigurierte Tabellen verweist, die der Kollaboration zugeordnet sind.

Example

Beispielsweise wird kein Protokoll erstellt, wenn die Abfrage abgebrochen wurde, nachdem AWS Clean Rooms überprüft wurde, ob sie den Analyseregeln entspricht, und während der Abfrageverarbeitung.

Wenn Sie die Protokollgruppe löschen, müssen Sie die Protokollgruppe manuell mit demselben Protokollgruppennamen (Kollaborations-ID der Kollaboration) neu erstellen. Oder Sie können die Protokollierung in Ihrer Mitgliedschaft ein- und ausschalten.

Weitere Informationen zum Aktivieren der Analyseprotokollierung finden Sie unter [Eine Zusammenarbeit erstellen](#).

Weitere Informationen zu Amazon CloudWatch Logs finden Sie im [Amazon CloudWatch Logs-Benutzerhandbuch](#).

Empfohlene Aktionen für Abfrage- und Jobprotokolle

Wir empfehlen Mitgliedern, regelmäßig die folgenden Maßnahmen zu ergreifen:

- Um sicherzustellen, dass die Abfragen und Jobs den Anwendungsfällen oder Abfragen entsprechen, die für die Zusammenarbeit vereinbart wurden, überprüfen Sie die Abfragen und Jobs, die in der Kollaboration ausgeführt werden.

Weitere Informationen zum Anzeigen der letzten Abfragen finden Sie unter [Aktuelle Abfragen anzeigen](#).

Weitere Informationen zum Anzeigen der letzten Jobs finden Sie unter [Aktuelle Jobs anzeigen](#).

- Überprüfen Sie die konfigurierten Tabellenspalten, die in den Analyseregeln der Kollaborationsmitglieder und in Abfragen verwendet werden, um zu überprüfen, ob die konfigurierten Tabellenspalten mit dem übereinstimmen, was für die Kollaboration vereinbart wurde.

Weitere Informationen zum Anzeigen der konfigurierten Spalten finden Sie unter [Tabellen und Analyseregeln anzeigen](#).

Detailliertes Monitoring mit CloudWatch in AWS Clean Rooms

Die detaillierte Überwachung ist eine optionale Funktion von AWS Clean Rooms. Wenn du [eine Kollaboration erstellst](#) und die detaillierte Überwachung aktivierst, können Mitglieder, die Abfragen ausführen können, und konfigurierte Zahler auswählen, ob sie detaillierte Messdaten zur Beobachtbarkeit in ihrem Konto erhalten möchten. CloudWatch

Dank der detaillierten Überwachung können Mitglieder die Abfrageleistung überwachen, die Ressourcenauslastung verfolgen und Dashboards für betriebliche Einblicke erstellen. Diese Kennzahlen helfen bei der Kapazitätsplanung, Kostenanalyse und Behebung von Leistungsproblemen.

Note

Die Verfügbarkeit von Metriken hängt von den konfigurierten Tabellendatenquellen ab:

- Bei Analysen, die auf konfigurierte Tabellen mit Athena als Datenquelle verweisen, geben die QueryFileScanned Metriken BytesRead und immer 0 zurück.
- Bei Analysen, die auf konfigurierte Tabellen mit Snowflake als Datenquelle verweisen, geben die RecordsRead Metriken BytesReadQueryFileScanned, und immer 0 zurück.
- Bei Analysen, die auf mehrere konfigurierte Tabellen mit unterschiedlichen Datenquellen verweisen (z. B. Snowflake- und S3-gestützte AWS Glue Tabellen) und Metriken nur für S3-gestützte BytesRead Tabellen QueryFileScanned zurückgeben. AWS Glue RecordsRead gibt Metriken für S3-gestützte AWS Glue Tabellen und Athena-Tabellen zurück.

Arten von Metriken

AWS Clean Rooms veröffentlicht zwei Kategorien von Metriken für: CloudWatch

Aggregierte detaillierte Monitoring-Metriken

Diese Kennzahlen bieten einen Überblick über alle Analysen hinweg auf aggregierter Ebene:

AnalysisRuntime

Zeigt die Gesamtdauer einer Analyse in Millisekunden an.

CRPUConsumed

Die für eine bestimmte Analyse verbrauchten CRPU (Clean Room Processing Units). [Weitere Informationen zur CRPU-Berechnung finden Sie unter AWS Clean Rooms Preisgestaltung.](#)

ConcurrentVCPUs

Gleichzeitige Nutzung der vCPU zu einem bestimmten Zeitpunkt (bezogen auf Dienstkontingente).

AnalysesCount

Anzahl der Analysen, die über einen bestimmten Zeitraum für eine bestimmte Dimension ausgeführt wurden.

ConcurrentQueries

Anzahl gleichzeitiger Abfragen zu einem bestimmten Zeitpunkt (bezogen auf Servicekontingente).

Detaillierte Überwachungsmetriken auf Abfrageebene

Diese Metriken bieten detaillierte Einblicke in die Ausführung einzelner Abfragen:

BytesRead

Gesamtzahl der Daten (in Byte), die für eine Analyse aus Quelltabellen gelesen wurden.

BytesWritten

Gesamtzahl der für eine Analyse geschriebenen Daten (in Byte).

RecordsRead

Gesamtzahl der Datensätze, die für eine Analyse aus Quelltabellen gelesen wurden.

RecordsWritten

Gesamtzahl der im Analyseergebnis zurückgegebenen Datensätze.

QueryFileScanned

Gesamtzahl der Datendateien, die für eine Analyse gescannt wurden.

Memory utilization

Für eine Analyse verwendete Speicherressource.

vCPU utilization

Für eine Analyse verwendete vCPU-Ressource.

Disk utilization

Für eine Analyse verwendete Speicherressource.

Metrikdimensionen

Die Metriken sind nach den folgenden Dimensionen gegliedert:

- Status der Analyse: Abgeschlossen
- Analysetyp: SparkSQL
- Mitgliedschaft: Die Mitglieds-ID
- Abfrage: Individuelle Abfrage-ID (für Metriken auf Abfrageebene)
- Konto/Region: Ihr Konto und Ihre Region AWS-Konto

Alle Metriken werden im AWS/CleanRooms Namespace in veröffentlicht. CloudWatch

Wer kann auf Metriken zugreifen

Detaillierte Monitoring-Metriken sind verfügbar für:

- Query Runner: Mitglieder, die Abfragen in der Kollaboration ausführen können
- Konfigurierte Kostenträger: Mitglieder, die dazu bestimmt sind, die Kosten für die Ausführung von Abfragen zu tragen

Mitglieder, die nur Daten beisteuern, haben keinen Zugriff auf detaillierte Monitoring-Metriken.

Preisgestaltung

AWS Clean Roomserhebt keine Gebühren für die Veröffentlichung von Metriken auf CloudWatch. Für die Speicherung von Kennzahlen, die Nutzung des Dashboards und Alarmer fallen jedoch

CloudWatch Standardgebühren an. Weitere Informationen finden Sie unter [CloudWatch Amazon-Preise](#).

Protokollieren von AWS Clean Rooms API-Aufrufen mit AWS CloudTrail

AWS Clean Rooms ist in einen Dienst integriert AWS CloudTrail, der eine Aufzeichnung der von einem Benutzer, einer Rolle oder einem AWS-Service Mitglied ausgeführten Aktionen bereitstellt AWS Clean Rooms. CloudTrail erfasst alle API-Aufrufe AWS Clean Rooms als Ereignisse. Zu den erfassten Aufrufen gehören Aufrufe von der AWS Clean Rooms Konsole und Codeaufrufen für die AWS Clean Rooms API-Operationen. Wenn Sie einen Trail erstellen, können Sie die kontinuierliche Bereitstellung von CloudTrail Ereignissen an einen Amazon S3 S3-Bucket aktivieren, einschließlich Ereignissen für AWS Clean Rooms. Wenn Sie keinen Trail konfigurieren, können Sie die neuesten Ereignisse trotzdem in der CloudTrail Konsole im Ereignisverlauf anzeigen. Anhand der von gesammelten Informationen können Sie die Anfrage ermitteln CloudTrail, an die die Anfrage gestellt wurde AWS Clean Rooms, die IP-Adresse, von der aus die Anfrage gestellt wurde, wer die Anfrage gestellt hat, wann sie gestellt wurde, und weitere Details.

Weitere Informationen CloudTrail dazu finden Sie im [AWS CloudTrail Benutzerhandbuch](#).

AWS Clean Rooms Informationen in CloudTrail

CloudTrail ist auf Ihrem aktiviert AWS-Konto , wenn Sie das Konto erstellen. Wenn eine Aktivität in stattfindet AWS Clean Rooms, wird diese Aktivität zusammen mit anderen AWS-Service Ereignissen in der CloudTrail Ereignishistorie in einem Ereignis aufgezeichnet. Sie können in Ihrem AWS-Konto die neusten Ereignisse anzeigen, suchen und herunterladen. Weitere Informationen finden Sie unter [Ereignisse im CloudTrail Ereignisverlauf anzeigen](#).

Für eine fortlaufende Aufzeichnung der Ereignisse in Ihrem AWS-Konto, einschließlich der Ereignisse für AWS Clean Rooms, erstellen Sie einen Trail. Ein Trail ermöglicht CloudTrail die Übermittlung von Protokolldateien an einen Amazon S3 S3-Bucket. Wenn Sie einen Trail in der Konsole anlegen, gilt dieser für alle AWS-Regionen-Regionen. Der Trail protokolliert Ereignisse aus allen Regionen der AWS Partition und übermittelt die Protokolldateien an den von Ihnen angegebenen Amazon S3 S3-Bucket. Darüber hinaus können Sie andere konfigurieren, AWS-Services um die in den CloudTrail Protokollen gesammelten Ereignisdaten weiter zu analysieren und darauf zu reagieren. Weitere Informationen finden Sie hier:

- [Übersicht zum Erstellen eines Trails](#)

- [CloudTrail unterstützte Dienste und Integrationen](#)
- [Konfiguration von Amazon SNS SNS-Benachrichtigungen für CloudTrail](#)
- [Empfangen von CloudTrail Protokolldateien aus mehreren Regionen](#)
- [Empfangen von CloudTrail Protokolldateien von mehreren Konten](#)

Alle AWS Clean Rooms Aktionen werden von der [AWS Clean Rooms API-Referenz](#) protokolliert CloudTrail und sind in dieser dokumentiert.

Jeder Ereignis- oder Protokolleintrag enthält Informationen zu dem Benutzer, der die Anforderung generiert hat. Die Identitätsinformationen unterstützen Sie bei der Ermittlung der folgenden Punkte:

- Ob die Anfrage mit Anmeldeinformationen des Stammbenutzers oder des IAM-Benutzers gestellt wurde.
- Gibt an, ob die Anforderung mit temporären Sicherheitsanmeldeinformationen für eine Rolle oder einen Verbundbenutzer gesendet wurde.
- Ob die Anforderung aus einem anderen AWS-Service gesendet wurde.

Weitere Informationen finden Sie unter [CloudTrail -Element userIdentity](#).

AWS Clean Rooms Logdateieinträge verstehen

Ein Trail ist eine Konfiguration, die die Übertragung von Ereignissen als Protokolldateien an einen von Ihnen angegebenen Amazon S3 S3-Bucket ermöglicht. CloudTrail Protokolldateien enthalten einen oder mehrere Protokolleinträge. Ein Ereignis stellt eine einzelne Anforderung aus einer beliebigen Quelle dar und enthält Informationen über die angeforderte Aktion, Datum und Uhrzeit der Aktion, Anforderungsparameter usw. CloudTrail Protokolldateien sind kein geordneter Stack-Trace der öffentlichen API-Aufrufe, sodass sie nicht in einer bestimmten Reihenfolge angezeigt werden.

Beispiele für AWS Clean Rooms CloudTrail Ereignisse

Die folgenden Beispiele zeigen CloudTrail Ereignisse für:

Themen

- [StartProtectedQuery \(erfolgreich\)](#)
- [StartProtectedQuery\(gescheitert\)](#)

StartProtectedQuery (erfolgreich)

```
{
  "eventVersion": "1.08",
  "userIdentity": {
    "type": "AssumedRole",
    "principalId": "EXAMPLE_PRINCIPAL_ID",
    "arn": "arn:aws:sts::123456789012:assumed-role/query-runner/jdoe",
    "accountId": "123456789012",
    "accessKeyId": "EXAMPLE_KEY_ID",
    "sessionContext": {
      "sessionIssuer": {
        "type": "Role",
        "principalId": "EXAMPLE_PRINCIPAL_ID",
        "arn": "arn:aws:iam::123456789012:role/query-runner",
        "accountId": "123456789012",
        "userName": "query-runner"
      },
      "webIdFederationData": {},
      "attributes": {
        "creationDate": "2023-04-07T19:34:32Z",
        "mfaAuthenticated": "false"
      }
    }
  },
  "eventTime": "2023-04-07T19:53:32Z",
  "eventSource": "cleanrooms.amazonaws.com",
  "eventName": "StartProtectedQuery",
  "awsRegion": "us-east-2",
  "sourceIPAddress": "203.0.113.1",
  "userAgent": "aws-internal/3",
  "requestParameters": {
    "resultConfiguration": {
      "outputConfiguration": {
        "s3": {
          "resultFormat": "CSV",
          "bucket": "cleanrooms-queryresults-jdoe-test",
          "keyPrefix": "test"
        }
      }
    }
  },
  "sqlParameters": "****",
  "membershipIdentifier": "a1b2c3d4-5678-90ab-cdef-EXAMPLE11111",
}
```

```

        "type": "SQL"
    },
    "responseElements": {
        "Access-Control-Expose-Headers": "x-amzn-RequestId,x-amzn-ErrorType,x-amzn-ErrorMessage,Date",
        "protectedQuery": {
            "createTime": 1680897212.279,
            "id": "f5988bf1-771a-4141-82a8-26fcc4e41c9f",
            "membershipArn": "arn:aws:cleanrooms:us-east-2:123456789012:membership/a1b2c3d4-5678-90ab-cdef-EXAMPLE11111",
            "membershipId": "a1b2c3d4-5678-90ab-cdef-EXAMPLE11111",
            "resultConfiguration": {
                "outputConfiguration": {
                    "s3": {
                        "bucket": "cleanrooms-queryresults-jdoe-test",
                        "keyPrefix": "test",
                        "resultFormat": "CSV"
                    }
                }
            },
            "sqlParameters": "****",
            "status": "SUBMITTED"
        }
    },
    "requestID": "7464211b-2277-4b55-9723-fb4f259aefd2",
    "eventID": "f7610f5e-74b9-420f-ae43-206571ebcbf7",
    "readOnly": false,
    "eventType": "AwsApiCall",
    "managementEvent": true,
    "recipientAccountId": "123456789012",
    "eventCategory": "Management"
}

```

StartProtectedQuery(gescheitert)

```

{
    "eventVersion": "1.08",
    "userIdentity": {
        "type": "AssumedRole",
        "principalId": "EXAMPLE_PRINCIPAL_ID",
        "arn": "arn:aws:sts::123456789012:assumed-role/query-runner/jdoe",
        "accountId": "123456789012",
        "accessKeyId": "EXAMPLE_KEY_ID",

```

```

    "sessionContext": {
      "sessionIssuer": {
        "type": "Role",
        "principalId": "EXAMPLE_PRINCIPAL_ID",
        "arn": "arn:aws:iam::123456789012:role/query-runner",
        "accountId": "123456789012",
        "userName": "query-runner"
      },
      "webIdFederationData": {},
      "attributes": {
        "creationDate": "2023-04-07T19:34:32Z",
        "mfaAuthenticated": "false"
      }
    },
    "eventTime": "2023-04-07T19:47:27Z",
    "eventSource": "cleanrooms.amazonaws.com",
    "eventName": "StartProtectedQuery",
    "awsRegion": "us-east-2",
    "sourceIPAddress": "203.0.113.1",
    "userAgent": "aws-internal/3",
    "errorCode": "ValidationException",
    "requestParameters": {
      "resultConfiguration": {
        "outputConfiguration": {
          "s3": {
            "resultFormat": "CSV",
            "bucket": "cleanrooms-queryresults-jdoe-test",
            "keyPrefix": "test"
          }
        }
      },
      "sqlParameters": "****",
      "membershipIdentifier": "a1b2c3d4-5678-90ab-cdef-EXAMPLE11111",
      "type": "SQL"
    },
    "responseElements": {
      "Access-Control-Expose-Headers": "x-amzn-RequestId,x-amzn-ErrorType,x-amzn-ErrorMessage,Date",
      "message": "Column(s) [identifier] is not allowed in select"
    },
    "requestID": "e29f9f74-8299-4a83-9d18-5ddce7302f07",
    "eventID": "c8ee3498-8e4e-44b5-87e4-ab9477e56eb5",
    "readOnly": false,

```

```
"eventType": "AwsApiCall",  
"managementEvent": true,  
"recipientAccountId": "123456789012",  
"eventCategory": "Management"  
}
```

Integration AWS Clean Rooms in ereignisgesteuerte Anwendungen mithilfe von Amazon EventBridge

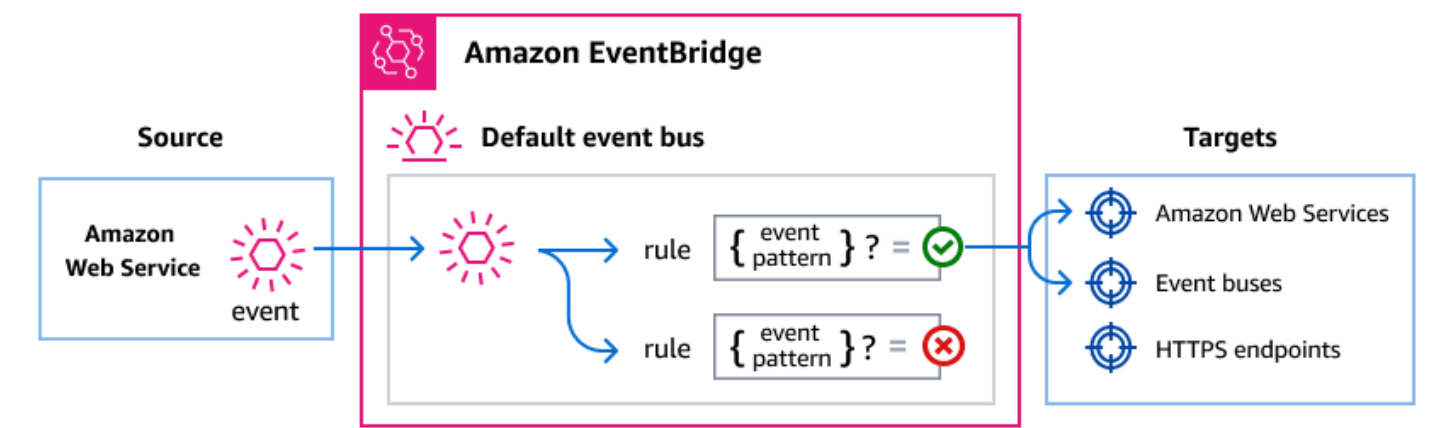
Sie können sie AWS Clean Rooms in ereignisgesteuerte Anwendungen (EDAs) integrieren, die Ereignisse verwenden, um zwischen Anwendungskomponenten AWS Clean Rooms zu kommunizieren und nachgelagerte Prozesse einzuleiten. Sie tun dies, indem Sie Amazon EventBridge verwenden, um Ereignisse von AWS Clean Rooms zu anderen Softwarekomponenten weiterzuleiten. Amazon EventBridge ist ein serverloser Service, der Ereignisse verwendet, um Anwendungskomponenten miteinander zu verbinden, sodass Sie AWS Dienste einfacher integrieren können, z. B. AWS Clean Rooms in ereignisgesteuerte Architekturen ohne zusätzlichen Code und zusätzliche Operationen.

Bei der ereignisgesteuerten Architektur werden lose gekoppelte Softwaresysteme entwickelt, die zusammenarbeiten, indem sie Ereignisse senden und darauf reagieren. In diesem Modell steht ein Ereignis für eine Änderung in einer Ressource oder Umgebung.

So EventBridge funktioniert das mit AWS Clean Rooms:

AWS Clean Rooms Generiert wie bei vielen AWS Diensten Ereignisse und sendet sie an den EventBridge Standard-Event-Bus. Ein Eventbus ist ein Router, der Ereignisse empfängt und sie an die von Ihnen angegebenen Ziele weiterleitet. Ziele können andere AWS Dienste, benutzerdefinierte Anwendungen und SaaS-Partneranwendungen umfassen.

EventBridge leitet Ereignisse gemäß den Regeln weiter, die Sie im Event-Bus erstellen. Für jede Regel geben Sie einen Filter oder ein Ereignismuster an, um nur die gewünschten Ereignisse auszuwählen. Jedes Mal, wenn ein Ereignis an den Event-Bus gesendet wird, wird es mit den einzelnen Regeln EventBridge verglichen. Wenn das Ereignis der Regel entspricht EventBridge, wird das Ereignis an die angegebenen Ziele weitergeleitet.



Nehmen wir zum Beispiel an, Sie möchten jedes Mal wissen, wenn in Ihrem Konto eine neue AWS Clean Rooms Kollaboration erstellt wird. Sie könnten eine Regel für den Standard-Event-Bus erstellen. In der Regel würden Sie ein Ereignismuster erstellen, das Ereignisse aus AWS Clean Rooms mit dem Namen angibt **Collaboration Created**. Jedes Mal, wenn ein Ereignis EventBridge empfangen wurde, das diesen Eigenschaften entspricht, wurde das Ereignis an den angegebenen Workflow weitergeleitet.

AWS Clean Rooms Ereignisse

AWS Dienste können Ereignisse direkt an den EventBridge Standard-Event-Bus senden. AWS CloudTrail Sendet außerdem Ereignisse, die von zahlreichen AWS Diensten stammen EventBridge , an. Zu diesen Ereignissen können API-Aufrufe, Konsolenanmeldungen und -aktionen, Serviceereignisse und CloudTrail Insights gehören. Weitere Informationen finden Sie AWS CloudTrail im EventBridge Benutzerhandbuch unter [AWS Serviceereignisse, die über bereitgestellt](#) werden.

Eine vollständige Liste der AWS Clean Rooms Ereignisse, an die gesendet wurden EventBridge, finden Sie unter dem AWS Clean Rooms Thema in der [EventBridgeEreignisreferenz](#).

Ereignisdetailtyp	Description
Analysevorlage erstellt	Der Besitzer der Analysevorlage und alle aktiven Mitglieder der Kollaboration werden benachrichtigt, wenn eine Analysevorlage erstellt wird.

Ereignisdetailtyp	Description
<u>Die Analysevorlage wurde aktualisiert</u>	Der Besitzer der Analysevorlage und alle aktiven Mitglieder der Kollaboration, die Einblick in die Aktualisierung haben, werden benachrichtigt, wenn eine Analysevorlage aktualisiert wird.
<u>Die Analysevorlage wurde gelöscht</u>	Der Besitzer der Analysevorlage und alle aktiven Mitglieder der Kollaboration werden benachrichtigt, wenn eine Analysevorlage gelöscht wird.
<u>Die Zusammenarbeit wurde erstellt</u>	Der Eigentümer der Kollaboration wird benachrichtigt, wenn eine Kollaboration erstellt wird.
<u>Die Zusammenarbeit wurde aktualisiert</u>	Der Eigentümer der Kollaboration und alle aktiven Mitglieder der Kollaboration, die Einblick in die Aktualisierung haben, werden benachrichtigt, wenn eine Kollaboration aktualisiert wird.
<u>Änderungsanfrage für die Kollaboration wurde erstellt</u>	Der Eigentümer der Kollaboration und alle aktiven Mitglieder der Kollaboration werden benachrichtigt, wenn eine Änderungsanfrage für die Kollaboration erstellt wird.
<u>Änderungsantrag zur Kollaboration genehmigt</u>	Der Eigentümer der Kollaboration und alle aktiven Mitglieder der Kollaboration werden benachrichtigt, wenn eine Kollaborationsänderungsanfrage genehmigt wurde.
<u>Anfrage zur Änderung der Kollaboration storniert</u>	Der Eigentümer der Kollaboration und alle aktiven Mitglieder der Kollaboration werden benachrichtigt, wenn eine Kollaborationsänderungsanfrage storniert wird.
<u>Anfrage zur Änderung der Kollaboration bestätigt</u>	Der Eigentümer der Kollaboration und alle aktiven Mitglieder der Kollaboration werden benachrichtigt, wenn eine Kollaborationsänderungsanfrage bestätigt wird.
<u>Die konfigurierte Tabellenzuweisung wurde erstellt</u>	Der Besitzer der konfigurierten Tabellenzuordnung und alle aktiven Mitglieder der Kollaboration werden benachrichtigt, wenn eine konfigurierte Tabellenzuordnung erstellt wird.

Ereignisdetailtyp	Description
<u>Die konfigurierte Tabellenzuordnung wurde aktualisiert</u>	Der Besitzer der konfigurierten Tabellenzuordnung und alle aktiven Mitglieder der Kollaboration, die Einblick in das Update haben, werden benachrichtigt, wenn eine konfigurierte Tabellenzuordnung aktualisiert wird.
<u>Die konfigurierte Tabellenzuordnung wurde gelöscht</u>	Der Besitzer der konfigurierten Tabellenzuordnung und alle aktiven Mitglieder der Kollaboration werden benachrichtigt, wenn eine konfigurierte Tabellenzuordnung gelöscht wird.
<u>Die Analyseregeln für konfigurierte Tabellenzuordnungen wurde erstellt</u>	Der Besitzer der Regel zur Analyse der konfigurierten Tabellenverknüpfung und alle aktiven Mitglieder der Kollaboration werden benachrichtigt, wenn eine Regel zur Analyse der konfigurierten Tabellenverknüpfung erstellt wird.
<u>Die Regel zur Analyse der konfigurierten Tabellenverknüpfung wurde aktualisiert</u>	Der Besitzer der Regel zur Analyse der konfigurierten Tabellenverknüpfung und alle aktiven Mitglieder der Kollaboration, die Einblick in das Update haben, werden benachrichtigt, wenn eine Regel zur Analyse der konfigurierten Tabellenverknüpfung aktualisiert wird.
<u>Die Regel zur Analyse der konfigurierten Tabellenverknüpfung wurde gelöscht</u>	Der Besitzer der konfigurierten Tabellenzuordnungs-Analyse-Regel und alle aktiven Mitglieder der Kollaboration werden benachrichtigt, wenn eine Regel zur Analyse der konfigurierten Tabellenverknüpfung gelöscht wird.
<u>Die ID-Zuordnungstabelle wurde erstellt</u>	Der Eigentümer der ID-Zuordnungstabelle und alle aktiven Mitglieder der Kollaboration werden benachrichtigt, wenn eine ID-Zuordnungstabelle erstellt wird.
<u>Die ID-Zuordnungstabelle wurde aktualisiert</u>	Der Eigentümer der ID-Zuordnungstabelle und alle aktiven Mitglieder der Kollaboration, die Einblick in die Aktualisierung haben, werden benachrichtigt, wenn eine ID-Zuordnungstabelle aktualisiert wird.

Ereignisdetailtyp	Description
<u>Die ID-Zuordnungstabelle wurde gelöscht</u>	Der Eigentümer der ID-Zuordnungstabelle und alle aktiven Mitglieder der Kollaboration werden benachrichtigt, wenn eine ID-Zuordnungstabelle gelöscht wird.
<u>ID-Namespace-Zuordnung wurde erstellt</u>	Der Eigentümer der ID-Namespace-Assoziation und alle aktiven Mitglieder der Kollaboration werden benachrichtigt, wenn eine ID-Namespace-Zuordnung erstellt wird.
<u>Die ID-Namespace-Zuordnung wurde aktualisiert</u>	Der Eigentümer der ID-Namespace-Assoziation und alle aktiven Mitglieder der Kollaboration, die Einblick in das Update haben, werden benachrichtigt, wenn eine ID-Namespace-Zuordnung aktualisiert wird.
<u>Die ID-Namespace-Zuordnung wurde gelöscht</u>	Der Eigentümer der ID-Namespace-Assoziation und alle aktiven Mitglieder der Kollaboration werden benachrichtigt, wenn eine ID-Namespace-Zuordnung gelöscht wird.
<u>Zur Kollaboration eingeladen</u>	Das eingeladene Mitglied wird benachrichtigt, wenn es zu einer Kollaboration eingeladen wird.
<u>Die Mitgliedschaft wurde erstellt</u>	Der Inhaber der Mitgliedschaft und alle aktiven Mitglieder der Kollaboration werden benachrichtigt, wenn eine Mitgliedschaft erstellt wird.
<u>Die Mitgliedschaft wurde aktualisiert</u>	Der Eigentümer der Mitgliedschaft wird benachrichtigt, wenn eine Mitgliedschaft aktualisiert wird, es sei denn, die Mitgliedschaft wurde aus der Kollaboration entfernt. In diesem Fall werden alle aktiven Mitglieder der Kollaboration benachrichtigt.
<u>Die Mitgliedschaft wurde gelöscht</u>	Der Inhaber der Mitgliedschaft und alle aktiven Mitglieder der Kollaboration werden benachrichtigt, wenn eine Mitgliedschaft gelöscht wird.
<u>Geschützter Job eingereicht</u>	Der Job Runner, der Job Payer und der Ergebnisempfänger eines geschützten Jobs werden benachrichtigt, wenn der geschützte Job eingereicht wird.

Ereignisdetailtyp	Description
<u>Geschützter Job gestartet</u>	Der Job Runner, der Job Payer und der Ergebnisempfänger eines geschützten Jobs werden benachrichtigt, wenn der geschützte Job gestartet wird.
<u>Geschützte Kündigung von Job</u>	Der Job Runner, der Job Payer und der Ergebnisempfänger eines geschützten Jobs werden benachrichtigt, wenn der geschützte Job storniert wird.
<u>Geschützter Job storniert</u>	Der Job Runner, der Job Payer und der Ergebnisempfänger eines geschützten Jobs werden benachrichtigt, wenn der geschützte Job storniert wird.
<u>Geschützter Job erfolgreich</u>	Der Job Runner, der Job Payer und der Ergebnisempfänger eines geschützten Jobs werden benachrichtigt, wenn der geschützte Job erfolgreich war.
<u>Geschützter Job ist fehlgeschlagen</u>	Der Job Runner, der Job Payer und der Ergebnisempfänger eines geschützten Jobs werden benachrichtigt, wenn der geschützte Job fehlgeschlagen ist.
<u>Geschützte Abfrage gesendet</u>	Der Query Runner, der Query Payer und die Ergebnisempfänger einer geschützten Abfrage werden benachrichtigt, wenn die geschützte Abfrage gesendet wird.
<u>Geschützte Abfrage wurde gestartet</u>	Der Query Runner, der Query Payer und die Ergebnisempfänger einer geschützten Abfrage werden benachrichtigt, wenn die geschützte Abfrage gestartet wird.
<u>Abbrechen geschützter Abfragen</u>	Der Query Runner, der Query Payer und die Ergebnisempfänger einer geschützten Abfrage werden benachrichtigt, wenn die geschützte Abfrage abgebrochen wird.
<u>Geschützte Abfrage wurde abgebrochen</u>	Der Query Runner, der Query Payer und die Ergebnisempfänger einer geschützten Abfrage werden benachrichtigt, wenn die geschützte Abfrage storniert wird.

Ereignisdetailtyp	Description
Geschützte Abfrage war erfolgreich	Der Query Runner, der Query Payer und die Ergebnisempfänger einer geschützten Abfrage werden benachrichtigt, wenn die geschützte Abfrage erfolgreich war.
Geschützte Abfrage ist fehlgeschlagen	Der Query Runner, der Query Payer und die Ergebnisempfänger einer geschützten Abfrage werden benachrichtigt, wenn die geschützte Abfrage fehlgeschlagen ist.
Bei geschützter Abfrage wurde das Zeitlimit überschritten	Der Query Runner, der Query Payer und die Ergebnisempfänger einer geschützten Abfrage werden benachrichtigt, wenn das Timeout für die geschützte Abfrage abgelaufen ist.

Ereignisse AWS Clean Rooms weiterleiten mit EventBridge

Um AWS Clean Rooms Ereignisse an ein Ziel EventBridge weiterleiten zu können, müssen Sie eine Regel erstellen. Jede Regel enthält ein Ereignismuster, das EventBridge mit jedem Ereignis übereinstimmt, das auf dem Event-Bus empfangen wird. Wenn die Ereignisdaten mit dem angegebenen Ereignismuster EventBridge übereinstimmen, wird das Ereignis an die Ziele der Regel weitergeleitet.

Umfassende Anweisungen zum Erstellen von Event-Bus-Regeln finden Sie im EventBridge Benutzerhandbuch unter [Erstellen von Regeln, die auf Ereignisse reagieren](#).

Erstellen von Ereignismustern, die AWS Clean Rooms Ereignissen entsprechen

Jedes Ereignismuster ist ein JSON-Objekt, das Folgendes enthält:

- (Optional): Ein `source` Attribut, das den Dienst identifiziert, der das Ereignis sendet. Für AWS Clean Rooms Ereignisse ist die Quelle `aws.cleanrooms`.
- (Optional): Ein `detail-type` Attribut, das ein Array mit den passenden Ereignisnamen enthält.
- (Optional): Ein `detail`-Attribut, das alle anderen Ereignisdaten für den Abgleich enthält.

Das folgende Ereignismuster entspricht beispielsweise allen Id Namespace Association Updated-Ereignissen, aus AWS Clean Rooms denen die Kollaboration gelöscht wurde:

```
{
```

```
"source": ["aws.cleanrooms"],
"detail-type": ["Id Namespace Association Updated"],
"detail": {
  "status": ["COLLABORATION_DELETED"]
}
}
```

Weitere Informationen zum Schreiben von Ereignismustern finden Sie unter [Ereignismuster](#) im EventBridge Benutzerhandbuch.

AWS Clean Rooms Referenz mit Einzelheiten zu Ereignissen

Alle Ereignisse von AWS Diensten haben einen gemeinsamen Satz von Feldern, die Metadaten über das Ereignis enthalten, z. B. den AWS Dienst, der die Quelle des Ereignisses ist, den Zeitpunkt, zu dem das Ereignis generiert wurde, das Konto und die Region, in der das Ereignis stattgefunden hat, und andere. Definitionen dieser allgemeinen Felder finden Sie unter [Event-Struktur](#) in der Amazon EventBridge Events-Referenz.

Darüber hinaus weist jedes Ereignis ein `detail`-Feld auf, das spezifische Daten für das betreffende Ereignis enthält. In der folgenden Referenz werden die Detailfelder für die verschiedenen AWS Clean Rooms Ereignisse definiert.

Bei der EventBridge Auswahl und Verwaltung von AWS Clean Rooms Ereignissen ist es hilfreich, Folgendes zu beachten:

- Das `source` Feld für alle Ereignisse von AWS Clean Rooms ist auf `aws.cleanrooms`.
- Das Feld `detail-type` gibt den Ereignistyp an.

Beispiel, `Collaboration Created`.

- Das Feld `detail` enthält die Daten, die für das betreffende Ereignis spezifisch sind.

Informationen zur Erstellung von Ereignismustern, die es Regeln ermöglichen, AWS Clean Rooms Ereignissen zuzuordnen, finden Sie unter [Ereignismuster](#) im EventBridge Amazon-Benutzerhandbuch.

Analysevorlage: Ereignis erstellt

Im Folgenden finden Sie die Detailfelder für das `Analysis Template Created` Ereignis.

```
{
```

```
"source": ["aws.cleanrooms"],  
"detail-type": ["Analysis Template Created"]  
}
```

Analysevorlage: Aktualisiertes Ereignis

Nachfolgend finden Sie die Detailfelder für das `Analysis Template Updated` Ereignis.

```
{  
  "source": ["aws.cleanrooms"],  
  "detail-type": ["Analysis Template Updated"]  
}
```

Analysevorlage: Ereignis gelöscht

Im Folgenden finden Sie die Detailfelder für das `Analysis Template Deleted` Ereignis.

```
{  
  "source": ["aws.cleanrooms"],  
  "detail-type": ["Analysis Template Deleted"]  
}
```

Zusammenarbeit: Ereignis erstellt

Im Folgenden finden Sie die Detailfelder für die `Collaboration Created` Veranstaltung.

```
{  
  "source": ["aws.cleanrooms"],  
  "detail-type": ["Collaboration Created"]  
}
```

Zusammenarbeit Aktualisierte Veranstaltung

Nachfolgend finden Sie die Detailfelder für die `Collaboration Updated` Veranstaltung.

```
{  
  "source": ["aws.cleanrooms"],  
  "detail-type": ["Collaboration Updated"]  
}
```

Anfrage zur Zusammenarbeit, Änderung, Ereignis erstellt

Im Folgenden finden Sie die Detailfelder für das Collaboration Change Request Created Ereignis.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Collaboration Change Request Created"]
}
```

Anfrage zur Zusammenarbeit, Änderungsantrag genehmigt, Ereignis genehmigt

Nachfolgend finden Sie die Detailfelder für die Collaboration Change Request Approved Veranstaltung.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Collaboration Change Request Approved"]
}
```

Anfrage zur Änderung der Zusammenarbeit, Veranstaltung abgesagt

Nachfolgend finden Sie die Detailfelder für die Collaboration Change Request Cancelled Veranstaltung.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Collaboration Change Request Cancelled"]
}
```

Anfrage zur Änderung der Zusammenarbeit: Zugesagtes Ereignis

Im Folgenden finden Sie die Detailfelder für die Collaboration Change Request Committed Veranstaltung.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Collaboration Change Request Committed"]
}
```

Ereignis „Konfigurierte Tabellenzuweisung wurde erstellt“

Im Folgenden finden Sie die Detailfelder für das Configured Table Association Created Ereignis.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Configured Table Association Created"]
}
```

Ereignis „Konfigurierte Tabellenzuweisung“ aktualisiert

Im Folgenden finden Sie die Detailfelder für das Configured Table Association Updated Ereignis.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Configured Table Association Updated"]
}
```

Ereignis „Konfigurierte Tabellenzuordnung gelöscht“

Im Folgenden finden Sie die Detailfelder für das Configured Table Association Deleted Ereignis.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Configured Table Association Deleted"]
}
```

Ereignis „Konfigurierte Tabellenzuordnungsanalyse, Regel erstellt“

Im Folgenden finden Sie die Detailfelder für das Configured Table Association Analysis Rule Created Ereignis.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Configured Table Association Analysis Rule Created"]
}
```

Analyseregel für konfigurierte Tabellenzuordnungen (aktualisiertes Ereignis)

Im Folgenden finden Sie die Detailfelder für das Configured Table Association Analysis Rule Updated Ereignis.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Configured Table Association Analysis Rule Updated"]
}
```

Analyseregel für konfigurierte Tabellenzuordnungen (Ereignis gelöscht)

Im Folgenden finden Sie die Detailfelder für das Configured Table Association Analysis Rule Deleted Ereignis.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Configured Table Association Analysis Rule Deleted"]
}
```

Ereignis „ID-Zuordnungstabelle erstellt“

Im Folgenden finden Sie die Detailfelder für das Id Mapping Table Created Ereignis.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Id Mapping Table Created"]
}
```

ID-Zuordnungstabelle Das Ereignis wurde aktualisiert

Nachfolgend finden Sie die Detailfelder für das Id Mapping Table Updated Ereignis.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Id Mapping Table Updated"]
}
```

ID-Zuordnungstabelle Gelöschtes Ereignis

Im Folgenden finden Sie die Detailfelder für das Id Mapping Table Deleted Ereignis.


```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Id Mapping Table Deleted"]
}
```

Ereignis, das von der ID Namespace Association erstellt wurde

Im Folgenden finden Sie die Detailfelder für das Id Namespace Association Created Ereignis.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Id Namespace Association Created"]
}
```

Ereignis „Id Namespace Association“ wurde aktualisiert

Nachfolgend finden Sie die Detailfelder für das Id Namespace Association Updated Ereignis.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Id Namespace Association Updated"]
}
```

Note

Wenn eine Kollaboration gelöscht wird, wird das Ereignis im `detail.status` Feld „ID Namespace Association Updated“ erfasst.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Id Namespace Association Updated"],
  "detail": {
    "status": ["COLLABORATION_DELETED"]
  }
}
```

Ereignis „ID Namespace Association Deleted“

Im Folgenden finden Sie die Detailfelder für das Id Namespace Association Deleted Ereignis.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Id Namespace Association Deleted"]
}
```

Zu einer Collaboration-Veranstaltung eingeladen

Nachfolgend finden Sie die Detailfelder für die Invited To Collaboration Veranstaltung.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Invited To Collaboration"]
}
```

Mitgliedschaft: Veranstaltung erstellt

Nachfolgend finden Sie die Detailfelder für die Membership Created Veranstaltung.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Membership Created"]
}
```

Mitgliedschaft Aktualisierte Veranstaltung

Nachfolgend finden Sie die Detailfelder für die Membership Updated Veranstaltung.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Membership Updated"]
}
```

Mitgliedschaft Gelöschtes Ereignis

Nachfolgend finden Sie die Detailfelder für die Membership Deleted Veranstaltung.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Membership Deleted"]
}
```

Geschützter Job Eingereichtes Ereignis

Nachfolgend finden Sie die Detailfelder für die Protected Job Submitted Veranstaltung.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Protected Job Submitted"]
}
```

Ereignis „Geschützter Job gestartet“

Im Folgenden finden Sie die Detailfelder für das Protected Job Started Ereignis.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Protected Job Started"]
}
```

Geschütztes Ereignis zur Kündigung von Job

Nachfolgend finden Sie die Detailfelder für die Protected Job Cancelling Veranstaltung.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Protected Job Cancelling"]
}
```

Geschützter Job Veranstaltung abgesagt

Nachfolgend finden Sie die Detailfelder für die Protected Job Cancelled Veranstaltung.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Protected Job Cancelled"]
}
```

Geschützter Job Erfolgreiches Ereignis

Im Folgenden finden Sie die Detailfelder für das Protected Job Succeeded Ereignis.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Protected Job Succeeded"]
}
```

Ereignis „Geschützter Job fehlgeschlagen“

Im Folgenden finden Sie die Detailfelder für das Protected Job Failed Ereignis.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Protected Job Failed"]
}
```

Geschützte Anfrage: Ereignis eingereicht.

Im Folgenden finden Sie die Detailfelder für das Protected Query Submitted Ereignis.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Protected Query Submitted"]
}
```

Ereignis „Geschützte Abfrage gestartet“

Im Folgenden finden Sie die Detailfelder für das Protected Query Started Ereignis.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Protected Query Started"]
}
```

Geschützte Abfrage, Ereignis wird storniert

Im Folgenden finden Sie die Detailfelder für das Protected Query Cancelling Ereignis.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Protected Query Cancelling"]
}
```

Geschützte Abfrage Abgebrochenes Ereignis

Nachfolgend finden Sie die Detailfelder für das Protected Query Cancelled Ereignis.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Protected Query Cancelled"]
}
```

Geschützte Abfrage Erfolgreiches Ereignis

Im Folgenden finden Sie die Detailfelder für das Protected Query Succeeded Ereignis.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Protected Query Succeeded"]
}
```

Ereignis „Geschützte Abfrage fehlgeschlagen“

Im Folgenden finden Sie die Detailfelder für das Protected Query Failed Ereignis.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Protected Query Failed"]
}
```

Ereignis „Geschütztes Abfrage-Timeout“

Im Folgenden finden Sie die Detailfelder für das Protected Query Timed Out Ereignis.

```
{
  "source": ["aws.cleanrooms"],
  "detail-type": ["Protected Query Timed Out"]
}
```

Kostenzuordnungs-Tags

AWS Clean Rooms unterstützt die Verwendung von Cost Allocation Tags zur Erfassung Ihrer AWS Kosten. Sie aktivieren diese Tags auf dem AWS Fakturierung und Kostenmanagement Dashboard.

AWS verwendet die Tags, um Ihre Kosten zu kategorisieren und Ihnen einen monatlichen Kostenverteilungsbericht zu senden. Benutzerdefinierte Kostenzuweisungs-Tags können auf AWS Clean Rooms Ressourcen angewendet werden, um Ihnen zu helfen, die Kosten in Ihrer gesamten Zusammenarbeit nachzuverfolgen und zuzuordnen.

Taggbare Ressourcen für die Kostenzuweisung

In der folgenden Tabelle sind die AWS Clean Rooms Ressourcen aufgeführt, für die Gebühren anfallen und die mit benutzerdefinierten Kostenzuordnungs-Tags für die Nachverfolgung und Organisation von Kosten gekennzeichnet werden können.

Abgerechnete Ressource	Verschlagwortet mit Ressource
SQL-Abfrage	Mitgliedschaften
PySpark Job	Mitgliedschaften
ML-Ausbildungsjob	Mitgliedschaften
ML-Inferenzjob	Mitgliedschaften
Generierung synthetischer Daten	Mitgliedschaften
Training mit Lookalike-Modellen	Lookalike-Modell
Export von Lookalike-Segmenten	Ähnliches Segment

AWS Clean Rooms Ressourcen erstellen mit AWS CloudFormation

AWS Clean Rooms ist in einen Service integriert AWS CloudFormation, der Sie bei der Modellierung und Einrichtung Ihrer AWS Ressourcen unterstützt. Dank dieser Integration müssen Sie weniger Zeit für die Erstellung und Verwaltung Ihrer Ressourcen und Infrastruktur aufwenden. Sie erstellen eine Vorlage, die alle gewünschten AWS Ressourcen beschreibt und diese Ressourcen für Sie CloudFormation bereitstellt und konfiguriert. Zu den Ressourcen gehören beispielsweise Kollaborationen, konfigurierte Tabellen, konfigurierte Tabellenzuordnungen und Mitgliedschaften.

Wenn Sie Ihre Vorlage verwenden CloudFormation, können Sie sie wiederverwenden, um Ihre AWS Clean Rooms Ressourcen einheitlich und wiederholt einzurichten. Beschreiben Sie Ihre Ressourcen einmal und stellen Sie dann dieselben Ressourcen immer wieder mehrfach AWS-Konten bereit AWS-Regionen.

AWS Clean Rooms und CloudFormation Vorlagen

Um Ressourcen für und zugehörige Dienste bereitzustellen AWS Clean Rooms und zu konfigurieren, müssen Sie [CloudFormation Vorlagen](#) verstehen. Vorlagen sind formatierte Textdateien in JSON oder YAML. Diese Vorlagen beschreiben die Ressourcen, die Sie in Ihren CloudFormation Stacks bereitstellen möchten. Wenn Sie mit JSON oder YAML nicht vertraut sind, können Sie CloudFormation Designer verwenden, um Ihnen die ersten Schritte mit Vorlagen zu erleichtern. CloudFormation Weitere Informationen finden Sie unter [Was ist CloudFormation -Designer?](#) im AWS CloudFormation -Benutzerhandbuch.

AWS Clean Rooms unterstützt das Erstellen von Kollaborationen, konfigurierten Tabellen, konfigurierten Tabellenzuordnungen und Mitgliedschaften in. CloudFormationWeitere Informationen, darunter Beispiele für JSON- und YAML-Vorlagen für Kollaborationen, konfigurierte Tabellen, konfigurierte Tabellenzuordnungen und Mitgliedschaften, finden Sie in den Referenzen zu den [ML-Ressourcentypen AWS Clean Roomsund AWS Clean Rooms ML](#) im AWS CloudFormation Benutzerhandbuch.

Erfahren Sie mehr über CloudFormation

Weitere Informationen CloudFormation finden Sie in den folgenden Ressourcen:

- [AWS CloudFormation](#)
- [AWS CloudFormation Benutzerhandbuch](#)
- [CloudFormation API Reference](#)
- [AWS CloudFormation -Benutzerhandbuch für die Befehlszeilenschnittstelle](#)

Kontingente für AWS Clean Rooms

Ihr AWS-Konto verfügt über Standardkontingente, die früher als Limits bezeichnet wurden, für jedes Kontingent AWS-Service. Sofern nicht anders angegeben, ist jedes Kontingent spezifisch für ein AWS-Region. Sie können Erhöhungen für einige Kontingente beantragen und andere Kontingente können nicht erhöht werden.

Um die Kontingente für anzuzeigen AWS Clean Rooms, öffnen Sie die [Konsole Service Quotas](#). Wählen Sie im Navigationsbereich AWS-Services aus und wählen Sie AWS Clean Rooms.

Informationen zum Beantragen einer Kontingenterhöhung finden Sie unter [Beantragen einer Kontingenterhöhung](#) im Service-Quotas-Benutzerhandbuch. Wenn das Kontingent noch nicht unter Servicekontingente verfügbar ist, verwenden Sie das [Formular zur Erhöhung des Servicelimits](#).

Themen

- [AWS Clean Rooms Kontingente](#)
- [ML-Kontingente für AWS Clean Rooms](#)

AWS Clean Rooms Kontingente

Ihr AWS-Konto hat die folgenden Kontingente im Zusammenhang mit. AWS Clean Rooms

Name	Standard	Anpas	Description
Größe der Analyseregeln	Jede unterstützte Region: 100 Kilobyte	Nein	Maximale JSON-Größe für eine Analyseregeln
Analysevorlagen pro Mitgliedschaft	Jede unterstützte Region: 25	Ja	Maximale Anzahl von Analysevorlagen pro Mitgliedschaft
Pro Konto erstellte Zusammenarbeiten	Jede unterstützte Region: 10	Yes (Ja)	Maximale Anzahl von Kollaborationen, die pro Konto erstellt wurden

Name	Standard	Anpas	Description
Spalten pro konfigurierter Tabellenzulassungsliste	Jede unterstützte Region: 100	Yes (Ja)	Maximale Anzahl von Spalten, die pro konfigurierter Tabelle zugelassen werden können
Gleichzeitige PySpark Job-vCPU-Nutzung pro Konto	us-east-1: 1.024 Jede der anderen unterstützten Regionen: 512	Ja	Maximale Gesamt-vCPU-Auslastung aller gleichzeitig ausgeführten PySpark Jobs pro Konto
Gleichzeitige Jobs pro Konto PySpark	us-east-1: 5 Jede der anderen unterstützten Regionen: 2	Ja	Maximale Anzahl gleichzeitig ausgeführter PySpark Jobs pro Konto
Gleichzeitige SQL-Abfragen pro Konto	us-east-1: 5 Jede der anderen unterstützten Regionen: 2	Ja	Maximale Anzahl gleichzeitig ausgeführter SQL-Abfragen pro Konto
Gleichzeitige SQL-Abfrage vCPU-Nutzung pro Konto	Jede unterstützte Region: 512	Ja	Maximale Gesamt-vCPU-Auslastung aller gleichzeitig laufenden SQL-Abfragen pro Konto
Gleichzeitiger laufender Job pro Mitgliedschaft	Jede unterstützte Region: 1	Nein	Maximale Anzahl gleichzeitiger laufender Jobs pro Mitgliedschaft
Gleichzeitig laufende Anfragen pro Mitgliedschaft	Jede unterstützte Region: 5	Nein	Maximale Anzahl gleichzeitiger laufender Anfragen pro Mitgliedschaft

Name	Standard	Anpas	Description
Konfigurierte Zuordnungen im Zielgruppenmodell pro Mitgliedschaft	Jede unterstützte Region: 5	Yes (Ja)	Maximale Anzahl konfigurierter Zuordnungen für ein Zielgruppenmodell pro Mitgliedschaft
Konfigurierte Tabellen pro Konto	Jede unterstützte Region: 250	Ja	Maximale Anzahl von konfigurierten Tabellen, die pro Konto erstellt wurden
Konfigurierte Tabellen pro geschützter Abfrage	Jede unterstützte Region: 15	Ja	Maximale Anzahl konfigurierter Tabellen in einer geschützten Abfrage
ID-Zuordnungstabellen pro Mitgliedschaft	Jede unterstützte Region: 5	Yes (Ja)	Maximale Anzahl von ID-Zuordnungstabellen pro Mitgliedschaft
ID-Namespace-Zuordnungen pro Mitgliedschaft	Jede unterstützte Region: 10	Yes (Ja)	Maximale Anzahl von ID-Namespace-Zuordnungen pro Mitgliedschaft
Pro Zusammenarbeit eingeladene Mitglieder	Jede unterstützte Region: 5	Yes (Ja)	Maximale Anzahl eingeladener Mitglieder pro Kollaboration
Mitgliedschaften pro Konto	Jede unterstützte Region: 100	Yes (Ja)	Maximale Anzahl von Mitgliedschaften pro Konto

Name	Standard	Anpas	Description
Vorlagen für Datenschutzbudgets für Zugriffsbudgets pro Zusammenarbeit	Jede unterstützte Region: 25	Ja	Maximale Anzahl von Vorlagen für Datenschutzbudgets für Zugriffsbudgets pro Zusammenarbeit
Textlänge abfragen (unter Verwendung von Differential Privacy)	Jede unterstützte Region: 8 Kilobyte	Nein	Maximale Textlänge für eine SQL-Abfrageanweisung unter Verwendung von Differential Privacy
Textlänge in der Clean Rooms SQL Analytics Engine abfragen	Jede unterstützte Region: 90 Kilobyte	Nein	Maximale Textlänge für eine SQL-Abfrageanweisung in der Clean Rooms SQL Analytics Engine
Textlänge in der Spark-Analyse-Engine abfragen	Jede unterstützte Region: 500 Kilobyte	Nein	Maximale Textlänge für eine SQL-Abfrageanweisung in der Spark-Analytics-Engine
Rate der BatchGetCollaborationAnalysisTemplate Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von BatchGetCollaborationAnalysisTemplate API-Aufrufen pro Sekunde
Rate der BatchGetSchema Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von BatchGetSchema API-Aufrufen pro Sekunde
Rate der BatchGetSchemaAnalysisRule Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von BatchGetSchemaAnalysisRule API-Aufrufen pro Sekunde

Name	Standard	Anpas	Description
Rate der CreateAnalysisTemplate Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von CreateAnalysisTemplate API-Aufrufen pro Sekunde
Rate der CreateCollaboration Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von CreateCollaboration API-Aufrufen pro Sekunde
Rate der CreateCollaborationChangeRequest Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von CreateCollaborationChangeRequest API-Aufrufen pro Sekunde
Rate der CreateConfiguredAudienceModelAssociation Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von CreateConfiguredAudienceModelAssociation API-Aufrufen pro Sekunde
Rate der CreateConfiguredTable Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von CreateConfiguredTable API-Aufrufen pro Sekunde
Rate der CreateConfiguredTableAnalysisRule Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von CreateConfiguredTableAnalysisRule API-Aufrufen pro Sekunde
Rate der CreateConfiguredTableAssociation Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von CreateConfiguredTableAssociation API-Aufrufen pro Sekunde

Name	Standard	Anpas	Description
Rate der CreateConfiguredTableAssociationAnalysisRule Anfragen	Jede unterstützte Region: 5 pro Sekunde	<u>Ja</u>	Maximale Anzahl von CreateConfiguredTableAssociationAnalysisRule API-Aufrufen pro Sekunde
Rate der CreateIdMappingTable Anfragen	Jede unterstützte Region: 5 pro Sekunde	<u>Ja</u>	Maximale Anzahl von CreateIdMappingTable API-Aufrufen pro Sekunde
Rate der CreateIdNamespaceAssociation Anfragen	Jede unterstützte Region: 5 pro Sekunde	<u>Ja</u>	Maximale Anzahl von CreateIdNamespaceAssociation API-Aufrufen pro Sekunde
Rate der CreateMembership Anfragen	Jede unterstützte Region: 5 pro Sekunde	<u>Ja</u>	Maximale Anzahl von CreateMembership API-Aufrufen pro Sekunde
Rate der CreatePrivacyBudgetTemplate Anfragen	Jede unterstützte Region: 5 pro Sekunde	<u>Ja</u>	Maximale Anzahl von CreatePrivacyBudgetTemplate API-Aufrufen pro Sekunde
Rate der DeleteAnalysisTemplate Anfragen	Jede unterstützte Region: 5 pro Sekunde	<u>Ja</u>	Maximale Anzahl von DeleteAnalysisTemplate API-Aufrufen pro Sekunde
Rate der DeleteCollaboration Anfragen	Jede unterstützte Region: 5 pro Sekunde	<u>Ja</u>	Maximale Anzahl von DeleteCollaboration API-Aufrufen pro Sekunde

Name	Standard	Anpas	Description
Rate der DeleteConfiguredAudienceModelAssociation Anfragen	Jede unterstützte Region: 5 pro Sekunde	<u>Ja</u>	Maximale Anzahl von DeleteConfiguredAudienceModelAssociation API-Aufrufen pro Sekunde
Rate der DeleteConfiguredTable Anfragen	Jede unterstützte Region: 5 pro Sekunde	<u>Ja</u>	Maximale Anzahl von DeleteConfiguredTable API-Aufrufen pro Sekunde
Rate der DeleteConfiguredTableAnalysisRule Anfragen	Jede unterstützte Region: 5 pro Sekunde	<u>Ja</u>	Maximale Anzahl von DeleteConfiguredTableAnalysisRule API-Aufrufen pro Sekunde
Rate der DeleteConfiguredTableAssociation Anfragen	Jede unterstützte Region: 5 pro Sekunde	<u>Ja</u>	Maximale Anzahl von DeleteConfiguredTableAssociation API-Aufrufen pro Sekunde
Rate der DeleteConfiguredTableAssociationAnalysisRule Anfragen	Jede unterstützte Region: 5 pro Sekunde	<u>Ja</u>	Maximale Anzahl von DeleteConfiguredTableAssociationAnalysisRule API-Aufrufen pro Sekunde
Rate der DeleteIdMappingTable Anfragen	Jede unterstützte Region: 5 pro Sekunde	<u>Ja</u>	Maximale Anzahl von DeleteIdMappingTable API-Aufrufen pro Sekunde
Rate der DeleteIdNamespaceAssociations Anfragen	Jede unterstützte Region: 5 pro Sekunde	<u>Ja</u>	Maximale Anzahl von DeleteIdNamespaceAssociation API-Aufrufen pro Sekunde

Name	Standard	Anpas	Description
Rate der DeleteMember Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von DeleteMember API-Aufrufen pro Sekunde
Rate der DeleteMembership Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von DeleteMembership API-Aufrufen pro Sekunde
Rate der DeletePrivacyBudgetTemplate Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von DeletePrivacyBudgetTemplate API-Aufrufen pro Sekunde
Rate der GetAnalysisTemplate Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von GetAnalysisTemplate API-Aufrufen pro Sekunde
Rate der GetCollaboration Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von GetCollaboration API-Aufrufen pro Sekunde
Rate der GetCollaborationAnalysisTemplate Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von GetCollaborationAnalysisTemplate API-Aufrufen pro Sekunde
Rate der GetCollaborationChangeRequest Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von GetCollaborationChangeRequest API-Aufrufen pro Sekunde

Name	Standard	Anpas	Description
Rate der GetCollaborationConfiguredAudienceModelAssociation Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von GetCollaborationConfiguredAudienceModelAssociation API-Aufrufen pro Sekunde
Rate der GetCollaborationIdNamespaceAssociation Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von GetCollaborationIdNamespaceAssociation API-Aufrufen pro Sekunde
Rate der GetCollaborationPrivacyBudgetTemplate Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von GetCollaborationPrivacyBudgetTemplate API-Aufrufen pro Sekunde
Rate der GetConfiguredAudienceModelAssociation Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von GetConfiguredAudienceModelAssociation API-Aufrufen pro Sekunde
Rate der GetConfiguredTable Anfragen	Jede unterstützte Region: 20 pro Sekunde	Ja	Maximale Anzahl von GetConfiguredTable API-Aufrufen pro Sekunde
Rate der GetConfiguredTableAnalysisRule Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von GetConfiguredTableAnalysisRule API-Aufrufen pro Sekunde

Name	Standard	Anpas	Description
Rate der GetConfiguredTableAssociation Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von GetConfiguredTableAssociation API-Aufrufen pro Sekunde
Rate der GetConfiguredTableAssociationAnalysisRule Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von GetConfiguredTableAssociationAnalysisRule API-Aufrufen pro Sekunde
Rate der GetIdMappingTable Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von GetIdMappingTable API-Aufrufen pro Sekunde
Rate der GetIdNamespaceAssociation Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von GetIdNamespaceAssociation API-Aufrufen pro Sekunde
Rate der GetMembership Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von GetMembership API-Aufrufen pro Sekunde
Rate der GetPrivacyBudgetTemplate Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von GetPrivacyBudgetTemplate API-Aufrufen pro Sekunde
Rate der GetProtectedJob Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von GetProtectedJob API-Aufrufen pro Sekunde
Rate der GetProtectedQuery Anfragen	Jede unterstützte Region: 20 pro Sekunde	Ja	Maximale Anzahl von GetProtectedQuery API-Aufrufen pro Sekunde

Name	Standard	Anpas	Description
Rate der GetSchema Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von GetSchema API-Aufrufen pro Sekunde
Rate der GetSchemaAnalysisRule Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von GetSchemaAnalysisRule API-Aufrufen pro Sekunde
Rate der ListAnalysisTemplates Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von ListAnalysisTemplates API-Aufrufen pro Sekunde
Rate der ListCollaborationAnalysisTemplates Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von ListCollaborationAnalysisTemplates API-Aufrufen pro Sekunde
Rate der ListCollaborationChangeRequests Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von ListCollaborationChangeRequests API-Aufrufen pro Sekunde
Rate der ListCollaborationConfiguredAudienceModelAssociations Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von ListCollaborationConfiguredAudienceModelAssociations API-Aufrufen pro Sekunde
Rate der ListCollaborationIdNamespaceAssociations Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von ListCollaborationIdNamespaceAssociations API-Aufrufen pro Sekunde

Name	Standard	Anpas	Description
Rate der ListCollaborationPrivacyBudgetTemplates Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von ListCollaborationPrivacyBudgetTemplates API-Aufrufen pro Sekunde
Rate der ListCollaborationPrivacyBudgets Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von ListCollaborationPrivacyBudgets API-Aufrufen pro Sekunde
Rate der ListCollaborations Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von ListCollaborations API-Aufrufen pro Sekunde
Rate der ListConfiguredAudienceModelAssociations Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von ListConfiguredAudienceModelAssociations API-Aufrufen pro Sekunde
Rate der ListConfiguredTableAssociations Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von ListConfiguredTableAssociations API-Aufrufen pro Sekunde
Rate der ListConfiguredTables Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von ListConfiguredTables API-Aufrufen pro Sekunde
Rate der ListIdMappingTables Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von ListIdMappingTables API-Aufrufen pro Sekunde

Name	Standard	Anpas	Description
Rate der ListIdNamespaceAssociations Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von ListIdNamespaceAssociations API-Aufrufen pro Sekunde
Rate der ListMembers Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von ListMembers API-Aufrufen pro Sekunde
Rate der ListMemberships Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von ListMemberships API-Aufrufen pro Sekunde
Rate der ListPrivacyBudgetTemplates Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von ListPrivacyBudgetTemplates API-Aufrufen pro Sekunde
Rate der ListPrivacyBudgets Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von ListPrivacyBudgets API-Aufrufen pro Sekunde
Rate der ListProtectedJobs Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von ListProtectedJobs API-Aufrufen pro Sekunde
Rate der ListProtectedQueries Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von ListProtectedQueries API-Aufrufen pro Sekunde
Rate der ListSchemas Anfragen	Jede unterstützte Region: 5	Yes (Ja)	Maximale Anzahl von ListSchemas API-Aufrufen pro Sekunde

Name	Standard	Anpas	Description
Rate der ListTagsForResource Anfragen	Jede unterstützte Region: 5 pro Sekunde	<u>Ja</u>	Maximale Anzahl von ListTagsForResource API-Aufrufen pro Sekunde
Rate der PopulateIdMappingTable Anfragen	Jede unterstützte Region: 5 pro Sekunde	<u>Ja</u>	Maximale Anzahl von PopulateIdMappingTable API-Aufrufen pro Sekunde
Rate der PreviewPrivacyImpact Anfragen	Jede unterstützte Region: 5 pro Sekunde	<u>Ja</u>	Maximale Anzahl von PreviewPrivacyImpact API-Aufrufen pro Sekunde
Rate der StartProtectedJob Anfragen	Jede unterstützte Region: 5 pro Sekunde	<u>Ja</u>	Maximale Anzahl von StartProtectedJob API-Aufrufen pro Sekunde
Rate der StartProtectedQuery Anfragen	Jede unterstützte Region: 5 pro Sekunde	<u>Ja</u>	Maximale Anzahl von StartProtectedQuery API-Aufrufen pro Sekunde
Rate der TagResource Anfragen	Jede unterstützte Region: 5 pro Sekunde	<u>Ja</u>	Maximale Anzahl von TagResource API-Aufrufen pro Sekunde
Rate der UntagResource Anfragen	Jede unterstützte Region: 5 pro Sekunde	<u>Ja</u>	Maximale Anzahl von UntagResource API-Aufrufen pro Sekunde
Rate der UpdateAnalysisTemplate Anfragen	Jede unterstützte Region: 5 pro Sekunde	<u>Ja</u>	Maximale Anzahl von UpdateAnalysisTemplate API-Aufrufen pro Sekunde

Name	Standard	Anpas	Description
Rate der UpdateCollaboration Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von UpdateCollaboration API-Aufrufen pro Sekunde
Rate der UpdateCollaborationChangeRequest Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von UpdateCollaborationChangeRequest API-Aufrufen pro Sekunde
Rate der UpdateConfiguredAudienceModelAssociation Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von UpdateConfiguredAudienceModelAssociation API-Aufrufen pro Sekunde
Rate der UpdateConfiguredTable Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von UpdateConfiguredTable API-Aufrufen pro Sekunde
Rate der UpdateConfiguredTableAnalysisRule Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von UpdateConfiguredTableAnalysisRule API-Aufrufen pro Sekunde
Rate der UpdateConfiguredTableAssociation Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von UpdateConfiguredTableAssociation API-Aufrufen pro Sekunde
Rate der UpdateConfiguredTableAssociationAnalysisRule Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von UpdateConfiguredTableAssociationAnalysisRule API-Aufrufen pro Sekunde

Name	Standard	Anpas	Description
Rate der UpdateIdMappingTable Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von UpdateIdMappingTable API-Aufrufen pro Sekunde
Rate der UpdateIdNamespaceAssociation Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von UpdateIdNamespaceAssociation API-Aufrufen pro Sekunde
Rate der UpdatePrivacyBudgetTemplate Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von UpdatePrivacyBudgetTemplate API-Aufrufen pro Sekunde
Rate der UpdateProtectedJob Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von UpdateProtectedJob API-Aufrufen pro Sekunde
Rate der UpdateProtectedQuery Anfragen	Jede unterstützte Region: 5 pro Sekunde	Ja	Maximale Anzahl von UpdateProtectedQuery API-Aufrufen pro Sekunde
Tabellenzuordnungen pro Mitgliedschaft	Jede unterstützte Region: 25	Ja	Maximale Anzahl von Tabellenzuordnungen pro Mitgliedschaft

AWS Clean Rooms Grenzwerte für Ressourcenparameter

Ressource	Standard	Description
Länge des Abfragetexts	90 KB	Maximale Textlänge für eine SQL-Abfrageanweisung

Ressource	Standard	Description
Länge des Abfragetextes (unter Verwendung von Differential Privacy)	90 KB	Maximale Textlänge für eine SQL-Abfrageanweisung unter Verwendung von Differential Privacy
Laufzeit der Abfrage	24 Stunden	Maximale Dauer, für die eine Abfrage vor dem Timeout ausgeführt wird

ML-Kontingente für AWS Clean Rooms

Ihr AWS-Konto hat die Standardkontingente, früher als Limits bezeichnet, für AWS Clean Rooms ML.

Gehen Sie wie folgt vor, um Service-Kontingente für AWS Clean Rooms ML anzuzeigen:

- Folgen Sie den Schritten unter [Service-Kontingente anzeigen](#) im Service Quotas User Guide und wählen Sie AWS Clean Rooms ML als Service aus.
- Weitere Informationen zu den [ML-Servicekontingenten von AWS Clean Rooms](#) finden Sie in der Allgemeine Amazon Web Services-Referenz.

Um die Leistung des Service aufrechtzuerhalten und eine angemessene Nutzung von AWS Clean Rooms ML sicherzustellen, können die einem Konto zugewiesenen Standardkontingente je nach regionalen Faktoren, Zahlungsverlauf, betrügerischer Nutzung und and/or Genehmigung einer Anfrage zur Erhöhung des Kontingents aktualisiert werden.

Name	Standard	Anpas	Description
Exportaufträge für aktive Zielgruppen pro Auftrag zur Zielgruppengenerierung	Jede unterstützte Region: 25	Nein	Die maximale Anzahl von aktiven Zielgrupp en-Exportaufträgen für einen Job zur Zielgrupp engenerierung
Aktive konfigurierte Modellalgorithmusz uordnungen pro Mitgliedschaft	Jede unterstützte Region: 1 000	Ja	Die maximale Anzahl von aktiven konfiguri

Name	Standard	Anpas	Description
			erten Modellalgorithmusz uordnungen pro Mitglieds chaft
Aktive konfigurierte Modellalgorithmen pro Mitgliedschaft	Jede unterstützte Region: 1 000	Ja	Die maximale Anzahl von aktiven konfigurierten Modellalgorithmen pro Mitgliedschaft
Aktive Eingangskanäle für benutzerd efinierte Modelle pro Mitgliedschaft	Jede unterstützte Region: 100	Yes (Ja)	Die maximale Anzahl aktiver Eingangskanäle für benutzerdefinierte Modelle pro Mitglieds chaft
Maximale Anzahl aktiver Trainings instanzen pro Konto	Jede unterstützte Region: 10	Yes (Ja)	Die maximale Anzahl aktiver Trainings instanzen, die pro Konto für die Erstellung trainiert er Modelle verwendet werden können.
Maximale Anzahl von mit active/pe nding/in Fortschritt trainierten Modellver sionen pro trainiertem Modell	Jede unterstützte Region: 100	Yes (Ja)	Die maximale Anzahl von mit active/pending/ in Progress trainierten Modellversionen, die pro trainiertem Modell erstellt werden können.
Maximale Anzahl von Eingabespalten zur Generierung synthetischer Daten	Jede unterstützte Region: 1 000	Nein	Die maximale Anzahl von Eingabespalten zur Generierung synthetis cher Daten

Name	Standard	Anpas	Description
Maximale Anzahl von Eingabezeilen zum Generieren synthetischer Daten	Jede unterstützte Region: 2.500.000	Nein	Die maximale Anzahl von Eingabezeilen zum Generieren synthetischer Daten
Maximale Anzahl von ml.c4.2xlarge-Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c4.2xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.c4.4xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c4.4xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.c4.8xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c4.8xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.c4.xlarge-Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c4.xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.

Name	Standard	Anpas	Description
Maximale Anzahl von ml.c5.18xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c5.18xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.c5.2xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c5.2xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.c5.4xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c5.4xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.c5.9xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c5.9xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.c5.xlarge-Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c5.xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.

Name	Standard	Anpas	Description
Maximale Anzahl von ml.c5n.18xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c5n.18xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.c5n.2xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c5n.2xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.c5n.4xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c5n.4xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.c5n.9xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c5n.9xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.c5n.xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c5n.xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.

Name	Standard	Anpas	Description
Maximale Anzahl von ml.c6i.12xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c6i.12xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.c6i.16xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c6i.16xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.c6i.24xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c6i.24xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.c6i.2xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c6i.2xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.c6i.32xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c6i.32xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.

Name	Standard	Anpas	Description
Maximale Anzahl von ml.c6i.4xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c6i.4xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.c6i.8xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c6i.8xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.c6i.xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c6i.xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.c7i.12xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c7i.12xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.c7i.16xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c7i.16xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.

Name	Standard	Anpas	Description
Maximale Anzahl von ml.c7i.24xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c7i.24xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.c7i.2xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c7i.2xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.c7i.48xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c7i.48xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.c7i.4xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c7i.4xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.c7i.8xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c7i.8xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.

Name	Standard	Anpas	Description
Maximale Anzahl von ml.c7i.large Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c7i.large-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.c7i.xlarge-Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.c7i.xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.g4dn.12xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g4dn.12xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.g4dn.16xlarge-Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g4dn.16xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.

Name	Standard	Anpas	Description
Maximale Anzahl von ml.g4dn.2xlarge-Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g4dn.2xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.g4dn.4xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g4dn.4xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.g4dn.8xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g4dn.8xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.g4dn.xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g4dn.xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.g5.12xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g5.12xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.

Name	Standard	Anpas	Description
Maximale Anzahl von ml.g5.16xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g5.16xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.g5.24xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g5.24xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.g5.2xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g5.2xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.g5.48xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g5.48xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.g5.4xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g5.4xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.

Name	Standard	Anpas	Description
Maximale Anzahl von ml.g5.8xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g5.8xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.g5.xlarge-Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g5.xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.g6.12xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g6.12xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.g6.16xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g6.16xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.g6.24xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g6.24xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.

Name	Standard	Anpas	Description
Maximale Anzahl von ml.g6.2xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g6.2xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.g6.48xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g6.48xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.g6.4xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g6.4xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.g6.8xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g6.8xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.g6.xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g6.xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.

Name	Standard	Anpas	Description
Maximale Anzahl von ml.g6e.12xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g6e.12xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.g6e.16xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g6e.16xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.g6e.24xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g6e.24xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.g6e.2xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g6e.2xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.g6e.48xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g6e.48xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.

Name	Standard	Anpas	Description
Maximale Anzahl von ml.g6e.4xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g6e.4xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.g6e.8xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g6e.8xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.g6e.xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.g6e.xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.m4.10xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.m4.10xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.m4.16xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.m4.16xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.

Name	Standard	Anpas	Description
Maximale Anzahl von ml.m4.2xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.m4.2xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.m4.4xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.m4.4xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.m4.xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.m4.xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.m5.12xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 3	Ja	Die maximale Anzahl von ml.m5.12xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.m5.24xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.m5.24xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.

Name	Standard	Anpas	Description
Maximale Anzahl von ml.m5.2xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 3	Ja	Die maximale Anzahl von ml.m5.2xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.m5.4xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 3	Ja	Die maximale Anzahl von ml.m5.4xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.m5.large-Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.m5.large-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.m5.xlarge-Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.m5.xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.

Name	Standard	Anpas	Description
Maximale Anzahl von ml.m6i.12xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.m6i.12xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.m6i.16xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.m6i.16xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.m6i.24xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.m6i.24xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.m6i.2xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.m6i.2xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.

Name	Standard	Anpas	Description
Maximale Anzahl von ml.m6i.32xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.m6i.32xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.m6i.4xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.m6i.4xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.m6i.8xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.m6i.8xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.m6i.large Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.m6i.large-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.

Name	Standard	Anpas	Description
Maximale Anzahl von ml.m6i.xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.m6i.xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.m7i.12xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.m7i.12xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.m7i.16xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.m7i.16xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.m7i.24xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.m7i.24xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.

Name	Standard	Anpas	Description
Maximale Anzahl von ml.m7i.2xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	<u>Ja</u>	Die maximale Anzahl von ml.m7i.2xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.m7i.48xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	<u>Ja</u>	Die maximale Anzahl von ml.m7i.48xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.m7i.4xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	<u>Ja</u>	Die maximale Anzahl von ml.m7i.4xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.m7i.8xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	<u>Ja</u>	Die maximale Anzahl von ml.m7i.8xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.

Name	Standard	Anpas	Description
Maximale Anzahl von ml.m7i.large Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.m7i.large-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.m7i.xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.m7i.xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.p2.16xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.p2.16xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.p2.8xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.p2.8xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.

Name	Standard	Anpas	Description
Maximale Anzahl von ml.p2.xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.p2.xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.p4d.24xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.p4d.24xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.p4de.24xlarge-Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.p4de.24xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.p5.48xlarge-Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.p5.48xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.

Name	Standard	Anpas	Description
Maximale Anzahl von ml.p5en.48xlarge-Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.p5en.48xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.r5.12xlarge-Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.r5.12xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.r5.16xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.r5.16xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.r5.24xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.r5.24xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.r5.2xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.r5.2xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.

Name	Standard	Anpas	Description
Maximale Anzahl von ml.r5.4xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.r5.4xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.r5.8xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.r5.8xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.r5.large-Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.r5.large-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.r5.xlarge-Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.r5.xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.

Name	Standard	Anpas	Description
Maximale Anzahl von ml.r5d.12xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	<u>Ja</u>	Die maximale Anzahl von ml.r5d.12xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.r5d.16xlarge-Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	<u>Ja</u>	Die maximale Anzahl von ml.r5d.16xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.r5d.24xlarge-Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	<u>Ja</u>	Die maximale Anzahl von ml.r5d.24xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.r5d.2xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	<u>Ja</u>	Die maximale Anzahl von ml.r5d.2xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.

Name	Standard	Anpas	Description
Maximale Anzahl von ml.r5d.4xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	<u>Ja</u>	Die maximale Anzahl von ml.r5d.4xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.r5d.8xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	<u>Ja</u>	Die maximale Anzahl von ml.r5d.8xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.r5d.large Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	<u>Ja</u>	Die maximale Anzahl von ml.r5d.large-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.r5d.xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	<u>Ja</u>	Die maximale Anzahl von ml.r5d.xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.

Name	Standard	Anpas	Description
Maximale Anzahl von ml.r7i.12xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.r7i.12xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.r7i.16xlarge-Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.r7i.16xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.r7i.24xlarge-Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.r7i.24xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.r7i.2xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.r7i.2xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.

Name	Standard	Anpas	Description
Maximale Anzahl von ml.r7i.48xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.r7i.48xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.r7i.4xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.r7i.4xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.r7i.8xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.r7i.8xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.r7i.large Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.r7i.large-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.r7i.xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.r7i.xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.

Name	Standard	Anpas	Description
Maximale Anzahl von ml.t3.2xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.t3.2xlarge Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.t3.large-Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.t3.large-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.t3.medium-Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.t3.medium-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.t3.xlarge-Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.t3.xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.

Name	Standard	Anpas	Description
Maximale Anzahl von ml.trn1.2xlarge Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.trn1.2xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.trn1.32xlarge-Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.trn1.32xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Anzahl von ml.trn1n.32xlarge-Trainingsinstanzen pro Konto	Jede unterstützte Region: 0	Ja	Die maximale Anzahl von ml.trn1n.32xlarge-Trainingsinstanzen, die pro Konto für die Erstellung trainierter Modelle verwendet werden können.
Maximale Einzelwerte in der Spalte mit kategorialen Vorhersagen	Jede unterstützte Region: 100	Nein	Die maximalen Einzelwerte in der Spalte mit der kategorialen Vorhersage
Mindestanzahl von Eingabespalten, die zur Generierung synthetischer Daten erforderlich sind	Jede unterstützte Region: 5	Nein	Die Mindestanzahl von Eingabespalten, die zur Generierung synthetischer Daten erforderlich ist

Name	Standard	Anpas	Description
Mindestanzahl von Eingabezeilen, die zur Generierung synthetischer Daten erforderlich sind	Jede unterstützte Region: 1 500	Nein	Die Mindestanzahl von Eingabezeilen, die zur Generierung synthetischer Daten erforderlich ist
Ausstehende oder in Bearbeitung befindliche Zielgruppenexportaufträge pro Kunde	Jede unterstützte Region: 20	Nein	Die maximale Anzahl von ausstehenden/laufenden Zielgruppen-Exportaufträgen pro Kunde
Ausstehende/laufende Jobs zur Zielgruppengenerierung pro Kunde	Jede unterstützte Region: 10	Yes (Ja)	Die maximale Anzahl ausstehender oder laufender Jobs zur Zielgruppengenerierung pro Kunde
Ausstehende oder in Bearbeitung befindliche Zielgruppenmodelle pro Kunde	Jede unterstützte Region: 2	Ja	Die maximale Anzahl von ausstehenden/laufenden Schulungsaufträgen für Zielgruppenmodelle pro Kunde
Ausstehend/in Bearbeitung befindliche Inferenzjobs mit benutzerdefiniertem Modell pro Konto	Jede unterstützte Region: 10	Yes (Ja)	Die maximale Anzahl ausstehend/in Bearbeitung befindlicher Inferenzjobs für benutzerdefinierte Modelle pro Konto
Ausstehend/in Bearbeitung befindliche Inferenzaufträge für benutzerdefinierte Modelle pro Mitgliedschaft	Jede unterstützte Region: 5	Yes (Ja)	Die maximale Anzahl ausstehend/in Bearbeitung befindlicher Inferenzjobs für benutzerdefinierte Modelle pro Mitgliedschaft

Name	Standard	Anpas	Description
Ausstehend/in Bearbeitung befindliche Schulungsaufträge für benutzerdefinierte Modelle pro Konto	Jede unterstützte Region: 10	Yes (Ja)	Die maximale Anzahl ausstehend/in Bearbeitung befindlicher Schulungsjobs für benutzerdefinierte Modelle pro Konto
Ausstehende/laufende Trainingsjobs mit benutzerdefiniertem Modell pro Mitgliedschaft	Jede unterstützte Region: 5	Yes (Ja)	Die maximale Anzahl ausstehender oder in Bearbeitung befindlicher Schulungsjobs mit benutzerdefiniertem Modell pro Mitgliedschaft
Ausstehende/laufende Jobs zur synthetischen MLIC-Generierung pro Konto	Jede unterstützte Region: 2	Ja	Die maximale Anzahl von ausstehenden/laufenden Aufträgen zur synthetischen MLIC-Generierung pro Konto

AWS Clean Rooms ML hat zusätzliche Kontingente, die unten aufgeführt sind

ML-Kontingente für Clean Rooms

Ressource	Standard	Description
Maximale Dauer eines benutzerdefinierten Modellierungs-Inferenzjobs	25 Stunden	
Maximale Anzahl von Interaktionen	20 Milliarden	Maximale Anzahl von Interaktionen, die in Trainingsdaten zulässig sind. Größere Eingaben werden abgetastet.
Minimale Anzahl von Interaktionen	1 Mio.	

Ressource	Standard	Description
Maximale Anzahl verschiedener Benutzer für das Training mit ähnlichen Modellen	100 Mio.	Wenn mehr berücksichtigt werden, werden nur die besten 100 Millionen verwendet, geordnet nach der Anzahl der Interaktionen.
Mindestanzahl verschiedener Benutzer für das Training mit einem ähnlichen Modell	100 000	
Mindestanzahl von Benutzern für den Export eines Jobs im Lookalike-Segment (Zielgruppe)	10.000	
Maximale Anzahl verschiedener Elemente, die für das Modelltraining verwendet werden.	1 Mio.	Sie können bis zu 50 Millionen Elemente hinzufügen, es werden jedoch nur die beliebtesten 1 Million verwendet.
Maximale Anzahl von Feature-Spalten im Trainingsdatensatz	10	
Mindestanzahl unterschiedlicher Elemente pro Benutzer	2	AWS Clean Rooms ML erfordert, dass jede Zeile oder jeder Benutzer zwei oder mehr Elemente enthält, einschließlich sich wiederholender Elemente.
Maximale Größe der Stammzielgruppe	500 000	

Ressource	Standard	Description
Mindestgröße des Startpublikums	500	Der Anbieter von Trainingsdaten kann diesen Wert auf einen niedrigen Wert von 25 festlegen.
APIs	pro Kunde	
Gesamtzahl der aktiven Trainingsdatensätze	500	
Gesamtzahl der aktiven Lookalike-Modelle (Zielgruppenmodelle)	500	
Gesamtzahl der aktiven konfigurierten Lookalike-Modelle (Zielgruppenmodelle)	10.000	
Gesamtzahl der abgeschlossenen Jobs zur Generierung von Lookalike-Segmenten (Audience)	Kein Limit	
Gesamtzahl der abgeschlossenen Aufträge für den Export von Lookalike-Segmenten (Audience)	Kein Limit	
Maximale Dauer eines Jobs zur Generierung eines Lookalike-Modells (Zielgruppenmodell)	1 Tag (24 Stunden)	

Ressource	Standard	Description
Maximale Dauer eines Jobs zur Generierung eines Lookalike-Segments (Audience)	10 Stunden	Nachdem Sie einen Seed bereitgestellt haben, benötigt Clean Rooms ML maximal 10 Stunden, um ein ähnliches Segment zu generieren. Wenn Sie eine SQL-Abfrage als Ausgangsdaten verwenden, kann die Ausführung der Abfrage zusätzlich zu den 10 Stunden zur Generierung des Lookalike-Segments bis zu 12 Stunden dauern.
Mindestprozentsatz für einen Bereich mit Segmentgröße (Zielgruppengröße)	1%	
Maximaler Prozentsatz für einen Bereich mit Segmentgröße (Zielgruppe)	20 %	
Absolute Mindestgröße für einen Bereich mit Segmentgröße (Zielgruppengröße)	1% der Anzahl der einzelnen Benutzer	
Maximale absolute Größe für ein Segment (Zielgruppengröße)	20% der Anzahl der einzelnen Benutzer	

Drosselungsquoten für die ML-API von Clean Rooms

Ihr Konto AWS-Konto verfügt über die folgenden Kontingente für API-Transaktionen pro Sekunde (TPS) pro Konto und Endpunkt.

Ressource	Ratenlimit	Description
Rate der Anfragen CreateAudienceModel	1 TPS-Rate, 3 TPS-Burst	Maximale Anzahl von CreateAudienceModel API-Aufrufen pro Sekunde
Rate der CreateCon figuredAudienceMod el Anfragen	10 TPS	Maximale Anzahl von CreateConfiguredAu dienceModel API-Auf rufen pro Sekunde
Rate der CreateTra iningDataset Anfragen	10 TPS	Maximale Anzahl von CreateTrainingData set API-Aufrufen pro Sekunde
Rate der DeleteAud ienceGenerationJob Anfragen	Geschwindigkeit von 2 TPS, Burst von 10 TPS	Maximale Anzahl von DeleteAudienceGene rationJob API-Aufrufen pro Sekunde
Rate der DeleteAud ienceModel Anfragen	Geschwindigkeit von 2 TPS, Burst von 10 TPS	Maximale Anzahl von DeleteAudienceModel API-Aufrufen pro Sekunde
Rate der DeleteCon figuredAudienceMod el Anfragen	10 TPS	Maximale Anzahl von DeleteConfiguredAu dienceModel API-Auf rufen pro Sekunde
Rate der DeleteCon figuredAudienceMod elPolicy Anfragen	25 TPS	Maximale Anzahl von DeleteConfiguredAu dienceModelPolicy API- Aufrufen pro Sekunde
Rate der DeleteTra iningDataset Anfragen	10 TPS	Maximale Anzahl von DeleteTrainingData

Ressource	Ratenlimit	Description
		set API-Aufrufen pro Sekunde
Rate der GetAudienceGenerationJob Anfragen	50 TPS	Maximale Anzahl von GetAudienceGenerationJob API-Aufrufen pro Sekunde
Rate der GetAudienceModel Anfragen	50 TPS	Maximale Anzahl von GetAudienceModel API-Aufrufen pro Sekunde
Rate der GetConfiguredAudienceModel Anfragen	50 TPS	Maximale Anzahl von GetConfiguredAudienceModel API-Aufrufen pro Sekunde
Rate der GetConfiguredAudienceModelPolicy Anfragen	50 TPS	Maximale Anzahl von GetConfiguredAudienceModelPolicy API-Aufrufen pro Sekunde
Rate der GetTrainingDataset Anfragen	50 TPS	Maximale Anzahl von GetTrainingDataset API-Aufrufen pro Sekunde
Rate der ListAudienceExportJobs Anfragen	50 TPS	Maximale Anzahl von ListAudienceExportJobs API-Aufrufen pro Sekunde
Rate der ListAudienceGenerationJobs Anfragen	50 TPS	Maximale Anzahl von ListAudienceGenerationJobs API-Aufrufen pro Sekunde

Ressource	Ratenlimit	Description
Rate der ListAudienceModels Anfragen	50 TPS	Maximale Anzahl von ListAudienceModels API-Aufrufen pro Sekunde
Rate der ListConfiguredAudienceModels Anfragen	50 TPS	Maximale Anzahl von ListConfiguredAudienceModels API-Aufrufen pro Sekunde
Rate der ListTagsForResource Anfragen	50 TPS	Maximale Anzahl von ListTagsForResource API-Aufrufen pro Sekunde
Rate der ListTrainingDatasets Anfragen	50 TPS	Maximale Anzahl von ListTrainingDatasets API-Aufrufen pro Sekunde
Rate der PutConfiguredAudienceModelPolicy Anfragen	25 TPS	Maximale Anzahl von PutConfiguredAudienceModelPolicy API-Aufrufen pro Sekunde
Rate der StartAudienceExportJob Anfragen	1 TPS-Rate, 3 TPS-Burst	Maximale Anzahl von StartAudienceExportJob API-Aufrufen pro Sekunde
Rate der StartAudienceGenerationJob Anfragen	1 TPS-Rate, 5 TPS-Burst	Maximale Anzahl von StartAudienceGenerationJob API-Aufrufen pro Sekunde
Rate der TagResource Anfragen	10 TPS	Maximale Anzahl von TagResource API-Aufrufen pro Sekunde

Ressource	Ratenlimit	Description
Rate der UntagResource Anfragen	50 TPS	Maximale Anzahl von UntagResource API-Aufrufen pro Sekunde
Rate der UpdateConfiguredAudienceModel Anfragen	10 TPS	Maximale Anzahl von UpdateConfiguredAudienceModel API-Aufrufen pro Sekunde
Rate der CreateConfiguredModelAlgorithm Anfragen	10 TPS	Maximale Anzahl von CreateConfiguredModelAlgorithm API-Aufrufen pro Sekunde.
Rate der CreateConfiguredModelAlgorithmAssociation Anfragen	10 TPS	Maximale Anzahl von CreateConfiguredModelAlgorithmAssociation API-Aufrufen pro Sekunde.
Rate der PutMLConfiguration Anfragen	10 TPS	Maximale Anzahl von PutMLConfiguration API-Aufrufen pro Sekunde.
Rate der CreateTrainedModel Anfragen	1 TPS-Rate, 3 TPS-Burst	Maximale Anzahl von CreateTrainedModel API-Aufrufen pro Sekunde.
Rate der StartTrainedModelExportJob Anfragen	10 TPS	Maximale Anzahl von StartTrainedModelExportJob API-Aufrufen pro Sekunde.

Ressource	Ratenlimit	Description
Rate der StartTrainedModelInferenceJob Anfragen	Rate 1 TPS, Rate 3 TPS	Maximale Anzahl von StartTrainedModelInferenceJob API-Aufrufen pro Sekunde.
Rate der GetConfiguredModelAlgorithm Anfrage	50 TPS	Maximale Anzahl von GetConfiguredModelAlgorithm API-Aufrufen pro Sekunde.
Rate der GetConfiguredModelAlgorithm Association Anfrage	50 TPS	Maximale Anzahl von GetConfiguredModelAlgorithmAssociation API-Aufrufen pro Sekunde.
Rate der GetTrainedModel Anfragen	50 TPS	Maximale Anzahl von GetTrainedModel API-Aufrufen pro Sekunde.
Rate der GetMLConfiguration Anfragen	50 TPS	Maximale Anzahl von GetMLConfiguration API-Aufrufen pro Sekunde.
Rate der GetTrainedModelInferenceJob Anfragen	50 TPS	Maximale Anzahl von GetTrainedModelInferenceJob API-Aufrufen pro Sekunde.
Rate der ListConfiguredModelAlgorithm Anfragen	50 TPS	Maximale Anzahl von ListConfiguredModelAlgorithm API-Aufrufen pro Sekunde.

Ressource	Ratenlimit	Description
Rate der ListConfiguredModelAlgorithmAssociations Anfragen	50 TPS	Maximale Anzahl von ListConfiguredModelAlgorithmAssociations API-Aufrufen pro Sekunde.
Rate der ListTrainedModels Anfragen	50 TPS	Maximale Anzahl von ListTrainedModels API-Aufrufen pro Sekunde.
Rate der ListCollaborationTrainedModelExportJobs Anfragen	50 TPS	Maximale Anzahl von ListCollaborationTrainedModelExportJobs API-Aufrufen pro Sekunde.
Rate der ListCollaborationTrainedModelInferenceJobs Anfragen	50 TPS	Maximale Anzahl von ListCollaborationTrainedModelInferenceJobs API-Aufrufen pro Sekunde.
Rate der DeleteConfiguredModelAlgorithm Anfragen	Geschwindigkeit von 2 TPS, Burst von 10 TPS	Maximale Anzahl von DeleteConfiguredModelAlgorithm API-Aufrufen pro Sekunde.
Rate der DeleteConfiguredModelAlgorithmAssociation Anfragen	Geschwindigkeit von 2 TPS, Burst von 10 TPS	Maximale Anzahl von DeleteConfiguredModelAlgorithmAssociation API-Anfragen pro Sekunde.

Ressource	Ratenlimit	Description
Rate der DeleteMLC onfiguration Anfragen	Geschwindigkeit von 2 TPS, Burst von 10 TPS	Maximale Anzahl von DeleteMLConfigurat ion API-Anfragen pro Sekunde.
Rate der DeleteTra inedModelOutput Anfragen	Geschwindigkeit von 2 TPS, Burst von 10 TPS	Maximale Anzahl von DeleteTrainedModel Output API-Anfragen pro Sekunde.

Dokumentenverlauf für das AWS Clean Rooms Benutzerhandbuch

In der folgenden Tabelle werden die Dokumentationsversionen für beschriebene AWS Clean Rooms.

Um Benachrichtigungen über Aktualisierungen dieser Dokumentation zu erhalten, können Sie den RSS-Feed abonnieren. Um RSS-Updates zu abonnieren, müssen Sie ein RSS-Plugin für den von Ihnen verwendeten Browser aktiviert haben.

Änderung	Beschreibung	Datum
Die Dokumentation für den Abschnitt „Cost Allocation Tagging“ wurde aktualisiert	Es wurde eine Dokumentation für markierbare Ressourcen im Zusammenhang mit der Kostenzuweisung hinzugefügt AWS Fakturierung und Kostenmanagement und ein Abschnitt erstellt.	20. Februar 2026
Support für die Cleanrooms-ml Spark Compute-Konfiguration	Kunden können jetzt das Laufzeitverhalten von Spark anpassen und die SQL-Leistung optimieren, indem sie unterstützte Spark-Eigenschaften für SQL-Abfragen konfigurieren. Diese Funktion ist jetzt für die Generierung von Zielgruppen, die Erstellung von Jobs und die Erstellung von MLInput Kanälen verfügbar.	5. Januar 2026
Support für detailliertes Monitoring mit CloudWatch	Kunden können sich jetzt für eine detaillierte Überwachung mit CloudWatch Kennzahlen für die betriebliche Überwachung	2. Januar 2026

ng der Zusammenarbeit entscheiden, einschließlich der Abfrageleistung und der Ressourcennutzung.

[Support von EventBridge Benachrichtigungsereignissen für Collaboration Change Requests](#)

Es wurden neue EventBridge Ereignisse für Lebenszyklusereignisse von Kollaborationsänderungsanfragen hinzugefügt.

19. Dezember 2025

[Support für Änderungsanfragen zur Verwaltung der Mitgliedschaft und Einstellungen für die Zusammenarbeit](#)

Ersteller von Kollaborationen können jetzt Änderungsanfragen einreichen, um neue Mitglieder hinzuzufügen, die Fähigkeiten vorhandener Mitglieder zu aktualisieren und die Einstellungen für die automatische Genehmigung zu ändern. Alle Mitglieder einer Kollaboration müssen Änderungsanträge genehmigen, bevor die Änderungen wirksam werden.

18. Dezember 2025

[Zeitplan für das Ende des Supports für die AWS Clean Rooms SQL Analytics Engine](#)

Der Support für die AWS Clean Rooms SQL Analytics Engine wird am 17. Dezember 2025 eingestellt. Kunden können jetzt nur noch die Spark SQL Analytics Engine verwenden.

17. Dezember 2025

[Support für zusätzliche EventBridge Benachrichtigungseignisse](#)

Neue EventBridge Ereignisse für Analysis Template, Configured Table Association, Configured Table Association Analysis Rule, ID Mapping Table, Id Namespace Association, Protected Job und Protected Query wurden hinzugefügt.

15. Dezember 2025

[Support für die Generierung synthetischer Datensätze mit verbessertem Datenschutz zum Trainieren benutzerdefinierter Modelle für maschinelles Lernen](#)

Kunden können jetzt synthetische Datensätze mit verbessertem Datenschutz generieren, um benutzerdefinierte Modelle für maschinelles Lernen zu trainieren und gleichzeitig sensible Daten zu schützen.

30. November 2025

[Support für konfigurierbare Spark-Eigenschaften in SQL-Abfragen](#)

Kunden können jetzt das Laufzeitverhalten von Spark anpassen, indem sie unterstützte Spark-Eigenschaften für SQL-Abfragen konfigurieren.

30. Oktober 2025

[Support für die regionenübergreifende Zusammenarbeit](#)

AWS Clean Rooms unterstützt jetzt die regionsübergreifende Zusammenarbeit für AWS Glue und Athena-Datentabellen, sodass Kunden in einer Datenbank gespeicherte Daten verarbeiten AWS-Regionen und Ergebnisse an Regionen liefern können, die von allen Kollaborationsmitgliedern genehmigt wurden.

3. Oktober 2025

[Support von Budgets für den Datenzugriff](#)

Kunden können jetzt Budgets für den Datenzugriff festlegen und verwalten, um die Tabellennutzung zu kontrollieren.

1. Oktober 2025

[ID-Zuordnungstabelle auffüllen — Update](#)

Kunden können jetzt die inkrementelle Verarbeitung in einem regelbasierten ID-Mapping-Workflow verwenden, um große Datensätze effizienter zu verarbeiten. Kunden können auch Datensätze aus einem ID-Mapping-Workflow löschen, um Datenverwaltungsvorschriften einzuhalten.

23. September 2025

[Support für variable Workload-Größen für PySpark Jobs](#)

Kunden können jetzt ihre Rechengröße wählen und anpassen, um PySpark Jobs effektiver ausführen zu können.

4. September 2025

[Support für das Hinzufügen neuer Mitglieder zu einer bestehenden Kollaboration](#)

Kunden können jetzt neue Mitglieder zu bestehenden Kollaborationen hinzufügen. Diese Funktion umfasst Optionen zum Konfigurieren automatischer Genehmigungseinstellungen und zum Blockieren von Mitgliedschaftsanfragen.

3. September 2025

[Unterstützung für die Konfiguration von Fehlerzusammenfassungen und die Schwärzung von benutzerdefinierten Modellprotokollen wurde hinzugefügt](#)

Kunden können jetzt Fehlerzusammenfassungen und Schwärzungen für benutzerdefinierte Modellprotokolle in der konfigurierten Modellalgorithmus-Zuordnung konfigurieren.

3. September 2025

[Support für die Snowflake-Authentifizierungsoption](#)

Kunden können jetzt die Schlüsselpaar-Authentifizierung mit einem privaten Snowflake-PEM-Schlüssel verwenden, wenn sie ihre Snowflake-Anmeldeinformationen in for connections speichern. AWS Secrets Manager AWS Clean Rooms

29. August 2025

[Fehlerkonfiguration für PySpark Analysevorlagen hinzugefügt](#)

Kunden können jetzt grundlegende oder detaillierte Stufen für Fehlermeldungen in PySpark Analysevorlagen konfigurieren, um die Sichtbarkeit von Fehlern und die Funktionen zur Fehlerbehebung zu verwalten.

19. August 2025

[Support für EventBridge](#)

Kunden können jetzt Amazon EventBridge verwenden, um Ereignisse von AWS Clean Rooms zu anderen Softwarekomponenten weiterzuleiten.

31. Juli 2025

[Aktualisierung auf eine bestehende Richtlinie](#)

Die folgenden neuen Berechtigungen wurden zur `AWSCleanRoomsFullAccessNoQuering` verwalteten Richtlinie hinzugefügt: `cleanrooms:UpdateConfiguredTableAllowedColumns` und `cleanrooms:UpdateConfiguredTableReference` .

29. Juli 2025

[Verbesserte Bearbeitungsmöglichkeiten für konfigurierte Tabellen](#)

Kunden können jetzt zulässige Spalten und Tabellenreferenzfelder für konfigurierte Tabellen ändern.

29. Juli 2025

[Der ML-Eingangskanal unterstützt den Parquet-Datentyp](#)

Kunden können jetzt den Parquet-Datentyp als Datenformat wählen, das der ML-Eingangskanal verwenden kann.

17. Juli 2025

[Support für differenziellen Datenschutz in der Spark SQL Analytics-Engine](#)

Kunden können jetzt Differential Privacy mit der Spark SQL Analytics Engine verwenden . Darüber hinaus beträgt die maximale Textlänge für eine SQL-Abfrageanweisung, die Differential Privacy verwendet, jetzt 90 KB.

16. Juli 2025

[Zeitplan für das Ende des Supports für die AWS Clean Rooms SQL-Analyse-Engine](#)

Die Support für neue Kollaborationen mit der Clean Rooms SQL Analytics Engine endet am 16. Juli 2025. Ab dem 17. Juli 2025 können Sie mit der alten Clean Rooms SQL-Analyse-Engine keine neuen Kollaborationen mehr erstellen. AWS Clean Rooms beendet die Unterstützung für die Clean Rooms SQL-Analyse-Engine am 17. Dezember 2025.

16. Juli 2025

[Support für inkrementelle und verteilte Schulungen](#)

Kunden können jetzt inkrementelle Schulungen und verteilte Schulungen für ihre individuellen Modelle nutzen.

1. Juli 2025

[Zeitplan für das Ende des Supports für die Clean Rooms SQL Analytics Engine](#)

Ab sofort bis zum 16. Juli 2025 ist für neue Kollaborationen, die die alte Clean Rooms SQL-Analyse-Engine verwenden, eine Erhöhung des Limits erforderlich. Ab dem 17. Juli 2025 können Sie mit dieser Engine keine neuen Kollaborationen mehr erstellen. AWS Clean Rooms beendet die Unterstützung für die Clean Rooms SQL Analytics Engine am 17. Dezember 2025.

22. Mai 2025

<u>Aktualisierung der PySpark Analysevorlagen</u>	Kunden können jetzt eine Benutzerskriptdatei mit einem beliebigen gültigen Python-Datennamen benennen.	15. Mai 2025
<u>Support für mehrere Ergebnisse mpfänger</u>	Kunden können jetzt mehrere Mitglieder auswählen, um Abfrageergebnisse aus einer einzigen Abfrage zu erhalten.	30. April 2025
<u>Support für die Migration von Kollaborationen zu Spark SQL</u>	AWS Clean Rooms SQL unterstützt jetzt zusätzlich zu benutzerdefinierten Analyseregeln auch Aggregations- und Listenanalyseregeln. Darüber hinaus können Kunden eine bestehende Zusammenarbeit aktualisieren, um die Spark-Analyse-Engine zu verwenden, die Spark SQL unterstützt.	2. April 2025
<u>Support für PySpark Jobs</u>	Kunden können jetzt Daten analysieren, indem sie Jobs mit genehmigten PySpark Analysevorlagen ausführen.	18. März 2025

[Aktualisierung der bestehenden Richtlinien](#)

Die folgende neue Berechtigung wurde der AWSCleanRoomsMLReadOnlyAccess verwalteten Richtlinie hinzugefügt: PassCleanRoomsResources . Die folgenden neuen Berechtigungen wurden der AWSCleanRoomsMLFullAccess verwalteten Richtlinie hinzugefügt: PassCleanRoomsResources undConsoleDescribeECRRepositories .

10. Januar 2025

[Support für mehrere Compute-Worker](#)

Kunden können jetzt bei der Erstellung eines Lookalike-Segments angeben, welche Art von Rechenarbeitern und wie viele bereitgestellt werden sollen.

17. Dezember 2024

[Support für mehrere Datenquellen und Clouds](#)

Kunden können jetzt mehrere Datenquellen und Clouds wie Amazon Athena und Snowflake verwenden, um mit den Datensätzen ihrer Partner zusammenzuarbeiten.

01. Dezember 2024

[Clean Rooms ML Custom Modeling ist jetzt verfügbar](#)

Kunden können jetzt ihre eigenen benutzerdefinierten ML-Modelle in einer Zusammenarbeit verwenden.

7. November 2024

<u>Neue Analyse-Engine</u>	Kunden mit großen Datensätzen können jetzt komplexe Abfragen mithilfe von SQL-Funktionen ausführen, die von der Spark SQL Analytics-Engine unterstützt werden.	29. Oktober 2024
<u>Verbesserter Datenschutz, Aufbau ähnlicher Zielgruppen, Auswahl mehrerer Ergebnisse mpfänger</u>	Mithilfe zusätzlicher Analysen und der Regel für die Kollaborationsanalyse können Sie Ihre Daten schützen und gleichzeitig komplexe Aktivierungsabfragen ermöglichen. Sie können anhand von SQL-Abfragen oder Analysevorgängen Lookalike-Zielgruppenmodelle erstellen. Sie können mehrere Mitglieder auswählen, um Ergebnisse zu erhalten.	24. Juli 2024
<u>Auflösung der Entität in AWS Clean Rooms</u>	Mit AWS Entity Resolution in AWS Clean Rooms können Sie eine ID-Zuordnungstabelle zwischen zwei ID-Namespace erstellen, um Ereignisdaten aus unterschiedlichen Identitätsräumen abzufragen.	23. Juli 2024
<u>Aktualisierung auf eine bestehende Richtlinie</u>	Die folgende neue Berechtigung wurde der verwalteten Richtlinie hinzugefügt: <code>AWSCleanRoomsFullAccessNoQuerying</code> <code>cleanrooms:BatchGetSchemaAnalysisRule</code>	13. Mai 2024

[AWS Clean Rooms ML ist jetzt vollständig verfügbar](#)

AWS Clean Rooms ML bietet eine Methode zur Verbesserung des Datenschutzes, mit der zwei Parteien ähnliche Benutzer in ihren Daten identifizieren können, ohne ihre Daten miteinander teilen zu müssen.

3. April 2024

[Aktualisierung auf eine bestehende Richtlinie](#)

Die Statement ID in der `AWSCleanRoomsFullAccess` verwalteten Richtlinie wurde von `pickQueryResultsBucket` bis `setQueryResultsBucket` aktualisiert, um die seit den Berechtigungen bestehenden Berechtigungen besser darzustellen.

21. März 2024

[Neue verwaltete Richtlinien für AWS Clean Rooms ML](#)

Zwei neue verwaltete Richtlinien wurden hinzugefügt: `AWSCleanRoomsMLReadOnlyAccess` und `AWSCleanRoomsMLFullAccess`.

29. November 2023

[AWS Clean Rooms ML \(Vorschau\)](#)

AWS Clean Rooms ML bietet eine Methode zur Verbesserung des Datenschutzes, mit der zwei Parteien ähnliche Benutzer in ihren Daten identifizieren können, ohne ihre Daten miteinander teilen zu müssen.

29. November 2023

[AWS Clean Rooms Differenzierter Datenschutz \(Vorschau\)](#)

Kunden können jetzt AWS Clean Rooms Differential Privacy verwenden, um die Privatsphäre ihrer Benutzer zu schützen.

29. November 2023

[Konfiguration der Zahlung](#)

Der Kollaborationsersteller kann nun entweder das Mitglied, das Abfragen ausführen kann, oder ein anderes Mitglied der Kollaboration so konfigurieren, dass ihm die Rechenkosten für Abfragen in Rechnung gestellt werden.

14. November 2023

[Laufzeit der Abfrage — Aktualisierung](#)

Die maximale Dauer der Ausführung einer Abfrage vor dem Timeout wurde von 4 Stunden auf 12 Stunden aktualisiert.

06. Oktober 2023

[CloudFormation Ressourcen — aktualisieren](#)

AWS Clean Rooms hat die folgenden neuen Ressourcen hinzugefügt: `AWS::CleanRooms::Membership Protected QueryOutputConfiguration` `AWS::CleanRooms::Membership ProtectedQueryResultConfiguration` , und `AWS::CleanRooms::Membership Protected QueryS3OutputConfiguration` .

07. September 2023

[CloudFormation Ressourcen
— aktualisieren](#)

AWS Clean Rooms hat die folgenden neuen Ressourcen hinzugefügt: `AWS::CleanRooms::AnalysisTemplate` und `AWS::CleanRooms::ConfiguredTableAnalysisRuleCustom`.

31. August 2023

[Separate Fähigkeiten der Mitglieder](#)

Der Ersteller der Kollaboration kann jetzt ein Mitglied als das Mitglied bestimmen, das Abfragen durchführen kann, und ein anderes Mitglied als das Mitglied, das Ergebnisse erhalten kann. Dadurch kann der Kollaborationsersteller sicherstellen, dass das Mitglied, das Abfragen durchführen kann, keinen Zugriff auf die Abfrageergebnisse hat.

30. August 2023

[AWS Clean Rooms Glossar](#)

Aktualisierung nur für die Dokumentation, um ein Glossar mit Begriffen hinzuzufügen. AWS Clean Rooms

30. August 2023

[Support für Apache Iceberg Tabellen \(Vorschau\)](#)

AWS Clean Rooms unterstützt jetzt Apache Iceberg Tabellen (Vorschau).

25. August 2023

Aktualisierung der KontingenteDer Abschnitt Kontingente

9. August 2023

wurde aktualisiert, um das neue Standardkontingent für Mitgliedschaften pro Konto widerzuspiegeln.

Aktualisierung auf eine bestehende Richtlinie

Die folgenden neuen Berechtigungen wurden der AWSCleanR

31. Juli 2023

oomsFullAccessNoQu
erying verwalteten Richtlini
e hinzugefügt:cleanroom
s:CreateAnalysisTe
mplate ,cleanroom
s:GetAnalysisTempl
ate ,cleanroom
s:UpdateAnalysisTe
mplate , cleanroom
s>DeleteAnalysisTe
mplate ,cleanroom
s:ListAna
lysisTemplates ,
cleanrooms:GetColl
aborationAnalysisT
emplate cleanroom
s:BatchGetCollabor
ationAnalysisTempl
ate ,und cleanroom
s:ListCollaboratio
nAnalysisTemplates .

Analysevorlagen und benutzerdefinierte Analyseregeln	AWS Clean Rooms unterstützt jetzt Analysevorlagen und die benutzerdefinierte Analyseregeln. Analysevorlagen ermöglichen es Mitarbeitern, ihre eigene benutzerdefinierte SQL-Abfrage zu erstellen oder zu importieren, um sie in der Zusammenarbeit zu verwenden. Mit der benutzerdefinierten Analyseregeln kann der Tabellenbesitzer benutzerdefinierte SQL-Abfragen für seine konfigurierten Tabellen genehmigen.	31. Juli 2023
Analyseregeln unterstützen die OR logische Bedingung	AWS Clean Rooms Analyseregeln unterstützen jetzt die OR logische Bedingung in der JOIN Klausel.	29. Juni 2023
CloudFormation Integration	AWS Clean Rooms integriert sich jetzt in CloudFormation.	15. Juni 2023
Analyse-Builder	Mitglieder, die Ergebnisse abfragen und empfangen können, können nun mithilfe der Benutzeroberfläche von Analysis Builder Abfragen für einige Tabellen ausführen, ohne SQL-Code schreiben zu müssen.	15. Juni 2023
SQL-Funktionen	Nur zur Dokumentation verwendetes Update zur Erläuterung der unterstützten SQL-Funktionen.	5. Mai 2023

Fehlersuche	Update nur für die Dokumentation, um einen Abschnitt zur Fehlerbehebung für häufig auftretende Probleme hinzuzufügen.	27. April 2023
Unterstützte Datentypen für AWS Clean Rooms	Update nur für die Dokumentation, um einen neuen Abschnitt hinzuzufügen, der die unterstützten AWS Glue Data Catalog Datentypen auflistet.	26. April 2023
Beispiele für AWS CloudTrail Ereignisse	Update nur für die Dokumentation, um Beispiele für CloudTrail Ereignisse für StartProtectedQuery (erfolgreich) und StartProtectedQuery (fehlgeschlagen) hinzuzufügen.	20. April 2023
Aktualisierung auf eine bestehende Richtlinie	Die folgenden neuen Berechtigungen wurden der AWSCleanRoomsFullAccessNoQuerying verwalteten Richtlinie hinzugefügt: <code>cleanrooms:ListTagsForResource</code> <code>cleanrooms:UntagResource</code> , und <code>cleanrooms:TagResource</code> Weitere Informationen finden Sie unter AWS Verwaltete Richtlinien .	21. März 2023
Allgemeine Verfügbarkeit	AWS Clean Rooms ist jetzt allgemein verfügbar.	21. März 2023

[Vorschauversion](#)

Vorschauversion des AWS
Clean Rooms Benutzerh
andbuchs

12. Januar 2023

AWS Clean Rooms Glossar

Konsultieren Sie dieses Glossar, um sich mit der verwendeten Terminologie vertraut zu machen.
AWS Clean Rooms

Regel für die Aggregationsanalyse

Die Abfrageeinschränkung, die Abfragen ermöglicht, bei denen Analysen mithilfe von COUNT, oder aggregiert werden SUM, und AVG zwar entlang optionaler Dimensionen. Diese Abfragen geben keine Informationen auf Zeilenebene preis.

Unterstützt Anwendungsfälle wie Kampagnenplanung, Messung der Medienreichweite, Häufigkeit und Konversionsmessung.

Andere Arten von Analyseregeln sind [benutzerdefinierte](#) Regeln und [Listenregeln](#).

Analyseregeln

Die Abfrageeinschränkungen, die einen bestimmten Abfragetyp autorisieren.

Der Analyseregeltyp bestimmt, welche Art von Analyse für die konfigurierte Tabelle ausgeführt werden kann. Jeder Typ hat eine vordefinierte Abfragestruktur. Über die Abfragesteuerelemente steuern Sie, wie Ihre Tabellenspalten in der Struktur verwendet werden können.

Die Arten von Analyseregeln sind [Aggregation](#), [Liste](#) und [Benutzerdefiniert](#).

Analysevorlage

Eine für die Zusammenarbeit spezifische, vorab genehmigte Abfrage, die wiederverwendet werden kann.

Unterstützte Formate: SQL-Code oder Python-Code für Spark.

Wenn Sie SQL verwenden, kann die Analysevorlage überall dort Parameter enthalten, wo ein Literalwert normalerweise in einer SQL-Abfrage vorkommen könnte. Weitere Informationen zu unterstützten Parametertypen finden Sie unter [Datentypen](#) in der AWS Clean Rooms SQL-Referenz.

Analysevorlagen funktionieren nur mit der [benutzerdefinierten Analyseregeln](#).

C3R-Verschlüsselungsclient

Der Verschlüsselungsclient von Cryptographic Computing for Clean Rooms (C3R).

C3R ist ein clientseitiges Verschlüsselungs-SDK mit einer Befehlszeilenschnittstelle, das zum Verschlüsseln und Entschlüsseln von Daten verwendet wird.

Spalte mit klarem Text

Eine Spalte, die weder für ein noch für ein JOIN SQL-Konstrukt kryptografisch geschützt ist. SELECT Klartextspalten können in jedem Teil der SQL-Abfrage verwendet werden.

Zusammenarbeit

Eine sichere logische Grenze, innerhalb AWS Clean Rooms derer Mitglieder SQL-Abfragen an konfigurierten Tabellen ausführen können.

Kollaborationen werden vom [Ersteller der Kollaboration](#) erstellt.

Nur Mitglieder, die zu der Kollaboration eingeladen wurden, können der Kollaboration beitreten.

Eine Kollaboration kann nur ein [Mitglied haben, das Daten abfragen kann](#), oder ein [Mitglied, das Abfragen und Jobs ausführen kann](#).

Eine Kollaboration kann nur ein [Mitglied haben, das Ergebnisse erhalten kann](#).

In einer Kollaboration kann nur ein [Mitglied für die Rechenkosten für Abfragen](#) oder ein [Mitglied für die Berechnung von Abfragen und Aufträgen zahlen](#).

Alle Mitglieder können die Liste der eingeladenen Teilnehmer der Kollaboration sehen, bevor sie der Kollaboration beitreten.

Ersteller der Kollaboration

Das Mitglied, das eine Kollaboration erstellt.

Pro Kollaboration gibt es nur einen Kollaborationsersteller.

Nur der Ersteller der Kollaboration kann Mitglieder aus der Kollaboration entfernen oder die Kollaboration löschen.

Konfigurierte Tabelle

Jede konfigurierte Tabelle stellt einen Verweis auf eine vorhandene Tabelle in der dar AWS Glue Data Catalog , die für die Verwendung in konfiguriert wurde AWS Clean Rooms. Eine konfigurierte Tabelle enthält eine Analyseregeln, die bestimmt, wie die Daten verwendet werden können.

AWS Clean Rooms Unterstützt derzeit das Zuordnen von Daten, die in Amazon Simple Storage Service (Amazon S3) gespeichert sind und über katalogisiert wurden. AWS Glue

Weitere Informationen AWS Glue dazu finden Sie im [AWS Glue Entwicklerhandbuch](#).

Konfigurierte Tabellen können einer oder mehreren Kollaborationen zugeordnet werden.

Note

AWS Clean Rooms unterstützt derzeit keine Amazon S3 S3-Bucket-Standorte, bei denen registriert ist AWS Lake Formation.

Benutzerdefinierte Analyseregeln

Die Abfrageeinschränkung, die einen bestimmten Satz vorab genehmigter Abfragen ([Analysevorlagen](#)) oder eine bestimmte Gruppe von Konten zulässt, die Abfragen oder Jobs bereitstellen können, die Ihre Daten verwenden.

Unterstützt Anwendungsfälle wie First-Touch-Attribution, inkrementelle Analysen und Analysen zur Zielgruppenfindung.

Unterstützt differenziellen Datenschutz.

Andere Arten von Analyseregeln sind [Aggregation](#) und [Liste](#).

Entschlüsselung

Der Prozess der Rücktransformation verschlüsselter Daten in ihre ursprüngliche Form. Die Entschlüsselung kann nur durchgeführt werden, wenn Sie Zugriff auf den geheimen Schlüssel haben.

Differenzielle Privatsphäre

Eine mathematisch strenge Technik, die die Kollaborationsdaten vor Mitgliedern schützt, die Ergebnisse erhalten können, wenn sie mehr über eine bestimmte Person erfahren.

Verschlüsselung

Der Prozess, bei dem Daten mithilfe eines geheimen Werts, eines sogenannten Schlüssels, in eine Form kodiert werden, die zufällig erscheint. Ohne Zugriff auf den Schlüssel ist es unmöglich, den ursprünglichen Klartext zu ermitteln.

Spalte „Fingerabdruck“

Eine Spalte, die für ein JOIN SQL-Konstrukt kryptografisch geschützt ist.

Workflow-Methode für die ID-Zuordnung

Wie die ID-Zuordnung durchgeführt werden soll.

Es gibt zwei Workflow-Methoden für die ID-Zuordnung:

- Regelbasierte ID-Zuordnung — Die Methode, mit der Sie Abgleichsregeln verwenden, um First-Party-Daten in einem ID-Mapping-Workflow von einer Quelle in ein Ziel zu übersetzen.
- ID-Zuordnung von Providerdiensten — Die Methode, mit der Sie einen Provider-Service verwenden, um in einem ID-Mapping-Workflow von Drittanbietern codierte Daten von einer Quelle in ein Ziel zu übersetzen.

AWS Clean Rooms unterstützt derzeit die LiveRamp auf Providerdiensten basierende Workflow-Methode für die ID-Zuordnung. Sie müssen über ein Abonnement für LiveRamp Through verfügen, um diese AWS Data Exchange Methode verwenden zu können. Weitere Informationen finden [Sie unter Abonnieren eines Anbieterdienstes AWS Data Exchange](#) im AWS Entity Resolution Benutzerhandbuch.

Tabelle mit ID-Zuordnung

Eine Ressource AWS Clean Rooms , die entweder Regeln für den Abgleich von Erstanbietern oder die Transcodierung von Identitäten mehrerer Parteien in einer Zusammenarbeit ermöglicht.

Eine ID-Zuordnungstabelle ist ein Verweis auf eine vorhandene Tabelle in der AWS Glue Data Catalog. Sie enthält eine [Analyseregel für die ID-Zuordnungstabelle](#), die bestimmt, wie die Daten abgefragt werden können. AWS Clean Rooms ID-Zuordnungstabellen können einer oder mehreren Kollaborationen zugeordnet werden.

Regel zur Analyse von ID-Zuordnungstabellen

Eine Art von Analyseregel, die von AWS Clean Rooms verwaltet und verwendet wird, um unterschiedliche Identitätsdaten zusammenzuführen, um Abfragen zu erleichtern. Sie wird automatisch zu den [ID-Zuordnungstabellen](#) hinzugefügt und kann nicht bearbeitet werden. Es erbt das Verhalten der anderen Analyseregeln in der Zusammenarbeit — sofern diese Analyseregeln homogen sind.

Arbeitsablauf bei der ID-Zuordnung

Ein Datenverarbeitungsjob, der Daten von einer Quelle einem Ziel auf der Grundlage der angegebenen [Workflow-Methode für die ID-Zuordnung zuordnet](#). Es erzeugt eine [ID-Zuordnungstabelle](#).

ID-Namespaces

Eine Ressource AWS Clean Rooms, die Metadaten enthält, die mehrere Datensätze AWS-Konten und die Verwendung dieser Datensätze in einem [ID-Mapping-Workflow](#) erläutern.

Zuordnung des ID-Namespaces

Eine Zuordnung einer ID-Namespaces-Ressource, die Ihnen hilft, Eingaben in ihren [ID-Mapping-Workflow](#) zu ermitteln.

Aufgabe

Eine Methode für den Zugriff auf und die Analyse konfigurierter Tabellen in einer Kollaboration mithilfe eines unterstützten Satzes von Funktionen, Klassen und Variablen.

AWS Clean Rooms unterstützt derzeit den PySpark Jobtyp.

AWS Clean Rooms unterstützt derzeit das Ausführen von Jobs mithilfe einer PySpark Analysevorlage.

Analyseregel auflisten

Die Abfrageeinschränkung, die Abfragen ermöglicht, die eine Attributanalyse der Überschneidung zwischen dieser Tabelle und den Tabellen des Mitglieds, das Abfragen durchführen kann, auf Zeilenebene ausgeben.

Unterstützt Anwendungsfälle wie Bereicherung und Zielgruppenbildung oder -unterbindung.

[Andere Arten von Analyseregeln sind Aggregationsregeln und benutzerdefinierte Regeln.](#)

Lookalike-Modell

Ein Modell der Daten eines Trainingsdatenanbieters, das es einem Anbieter von Ausgangsdaten ermöglicht, ein ähnliches [Segment der Daten](#) eines Trainingsdatenanbieters zu erstellen, das seinen Ausgangsdaten am ähnlichsten ist.

Ähnliches Segment

Eine Teilmenge der Trainingsdaten, die den Ausgangsdaten am ähnlichsten ist.

Mitglied

Ein AWS Kunde, der an einer [Zusammenarbeit](#) teilnimmt.

Ein Mitglied wird anhand seines identifiziert AWS-Konto.

Alle Mitglieder können Daten beitragen.

Mitglied, das Abfragen durchführen kann

Das Mitglied, das Daten in der [Kollaboration](#) abfragen kann.

Es gibt nur ein Mitglied, das Abfragen pro Kollaboration durchführen kann, und dieses Mitglied ist unveränderlich.

Ein Administratorbenutzer kann mithilfe von AWS Identity and Access Management (IAM-) Berechtigungen steuern, welche seiner IAM-Prinzipale (z. B. Benutzer oder Rollen) Daten in der

Kollaboration abfragen können. Weitere Informationen finden Sie unter [Erstellen Sie eine Servicerolle zum Lesen von Daten aus Amazon S3](#).

Mitglied, das Abfragen und Jobs ausführen kann

Das Mitglied, das Abfragen und Jobs für die Daten in der [Kollaboration](#) ausführen kann.

Es gibt nur ein Mitglied, das Abfragen und Jobs pro Kollaboration ausführen kann, und dieses Mitglied ist unveränderlich.

Ein Administratorbenutzer kann mithilfe von AWS Identity and Access Management (IAM-) Berechtigungen steuern, welche seiner IAM-Prinzipale (z. B. Benutzer oder Rollen) Abfragen und Jobs in der Kollaboration ausführen können. Weitere Informationen finden Sie unter [Erstellen Sie eine Servicerolle zum Lesen von Daten aus Amazon S3](#).

Mitglied, das Ergebnisse erhalten kann

Ein Mitglied, das Abfrageergebnisse erhalten kann. Ein Mitglied, das Ergebnisse empfangen kann, legt die Einstellungen für die Abfrageergebnisse für das Amazon S3 S3-Ziel und das Abfrageergebnisformat (CSV oder Parquet) fest.

Bei Analysen mit der Spark-Analyse-Engine gibt ein Mitglied, das Ergebnisse empfangen kann, auch an, ob Dateien als einzelne Datei oder als mehrere Dateien ausgegeben werden sollen.

In einer Zusammenarbeit kann es mehr als ein Mitglied geben, das Ergebnisse erhalten kann.

Das Mitglied zahlt die Kosten für die Berechnung von Abfragen

Das Mitglied, das für die Bezahlung der Kosten für die Query Compute verantwortlich ist.

Es gibt nur ein Mitglied, das für die Bezahlung der Abfrageberechnungskosten pro Zusammenarbeit verantwortlich ist, und dieses Mitglied ist unveränderlich.

Wenn der Ersteller der Kollaboration niemanden als das Mitglied angegeben hat, das die Kosten für die Abfrageverarbeitung bezahlt, ist das [Mitglied, das Abfragen durchführen kann](#), der Standardzahler.

Das Mitglied, das die Kosten für die Query-Compute bezahlt, erhält eine Rechnung für die Abfragen, die im Rahmen der Kollaboration ausgeführt wurden.

Das Mitglied zahlt für die Kosten der Abfrage- und Auftragsverarbeitung

Das Mitglied, das für die Bezahlung der Kosten für Abfragen und Auftragsberechnungen verantwortlich ist.

Es gibt nur ein Mitglied, das für die Zahlung der Kosten für die Abfrage- und Auftragsverarbeitung pro Zusammenarbeit verantwortlich ist, und dieses Mitglied ist unveränderlich.

Wenn der Ersteller der Kollaboration niemanden als das Mitglied angegeben hat, das die Kosten für die Abfrage und die Auftragsberechnung bezahlt, ist das [Mitglied, das Abfragen durchführen kann](#), der Standardzahler.

Das Mitglied, das die Kosten für die Abfrage- und Job-Compute bezahlt, erhält eine Rechnung für die Abfragen, die im Rahmen der Kollaboration ausgeführt wurden.

Mitgliedschaften

Eine Ressource, die erstellt wird, wenn ein [Mitglied](#) einer [Kollaboration](#) beitrifft.

Alle Ressourcen, die das Mitglied einer Kollaboration zuordnet, sind Teil der Mitgliedschaft oder mit der Mitgliedschaft verknüpft.

Nur das Mitglied, dem die Mitgliedschaft gehört, kann Ressourcen in dieser Mitgliedschaft hinzufügen, entfernen oder bearbeiten.

Versiegelte Spalte

Eine Spalte, die für ein SELECT SQL-Konstrukt kryptografisch geschützt ist.

Ausgangsdaten

Die Daten des Seed-Datenanbieters, die zur Erstellung eines [Lookalike-Segments](#) verwendet werden. Die Ausgangsdaten können direkt bereitgestellt werden oder aus den Ergebnissen einer AWS Clean Rooms Abfrage stammen. Bei der Ausgabe eines Lookalike-Segments handelt es sich um eine Gruppe von Benutzern aus den Trainingsdaten, die den Ausgangsbenutzern am ähnlichsten sind.

Spark-Analyse-Engine

Eine Analyseoption AWS Clean Rooms , mit der Kunden mithilfe von Apache Spark SQL-Funktionen komplexe Abfragen für große Datensätze ausführen können, die in Amazon S3, Amazon Athena oder Snowflake gespeichert sind. Es unterstützt auch Analysen in. PySpark AWS Clean Rooms

Wenn Sie mithilfe der [CreateCollaborationAPI](#) eine Kollaboration erstellen, ist SPARK der Wert der Spark-Analyse-Engine:

Query

Eine Methode, um auf konfigurierte Tabellen in einer Kollaboration zuzugreifen und diese zu analysieren, wobei ein unterstützter Satz von Funktionen, Klassen und Variablen verwendet wird.

AWS Clean Rooms unterstützt derzeit die SQL-Abfragesprache.

AWS Clean Rooms unterstützt derzeit das Ausführen von direkten SQL-Abfragen oder das Ausführen von Abfragen mithilfe einer SQL-Analysevorlage.

Die vorliegende Übersetzung wurde maschinell erstellt. Im Falle eines Konflikts oder eines Widerspruchs zwischen dieser übersetzten Fassung und der englischen Fassung (einschließlich infolge von Verzögerungen bei der Übersetzung) ist die englische Fassung maßgeblich.