Guía del usuario

AWS Clean Rooms

Copyright © 2025 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

AWS Clean Rooms: Guía del usuario

Copyright © 2025 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Las marcas comerciales y la imagen comercial de Amazon no se pueden utilizar en relación con ningún producto o servicio que no sea de Amazon, de ninguna manera que pueda causar confusión entre los clientes y que menosprecie o desacredite a Amazon. Todas las demás marcas registradas que no son propiedad de Amazon son propiedad de sus respectivos propietarios, que pueden o no estar afiliados, conectados o patrocinados por Amazon.

Table of Contents

¿Qué es AWS Clean Rooms? 1
¿Es la primera vez que lo usa AWS Clean Rooms ? 2
¿Cómo funciona AWS Clean Rooms 2
Servicios relacionados 2
AWS servicios 2
Servicios de terceros 4
Acceder AWS Clean Rooms 4
Precios para AWS Clean Rooms 5
Facturación para AWS Clean Rooms 5
Reglas de análisis 6
Tipos de regla de análisis 7
Regla de análisis de agregación 9
Regla de análisis de lista 29
Regla de análisis personalizada 38
Regla de análisis de tablas de asignación de ID 45
AWS Clean Rooms Privacidad diferencial 56
Privacidad diferencial 57
Cómo funciona Differential Privacy in AWS Clean Rooms 57
Política de privacidad diferencial 58
Capacidades de SQL 60
Consejos y ejemplos de consultas SQL 75
Limitaciones 76
AWS Clean Rooms ML 77
Cómo funciona el AWS Clean Rooms aprendizaje automático con AWS los modelos 78
Cómo funciona el aprendizaje AWS Clean Rooms automático con modelos personalizados 79
AWS modelos en Clean Rooms ML 81
Modelos personalizados en Clean Rooms ML 91
Computación criptográfica 99
Consideraciones 101
Tipos de archivo y de datos admitidos 103
Nombres de columnas 109
Tipos de columnas 110
Parámetros 112

Indicadores opcionales 117
Consultas con C3R 120
Directrices 121
Con AWS Clean Rooms figuración 148
Inscríbase en AWS 148
Configure las funciones de servicio para AWS Clean Rooms 148
Creación de un usuario administrador 149
Creación de un rol de IAM para un miembro de la colaboración 150
Cree un rol de servicio para leer datos de Amazon S3 151
Cree un rol de servicio para leer los datos de Amazon Athena 154
Cree un rol de servicio para leer los datos de Snowflake 158
Cree un rol de servicio para leer el código de un bucket de S3 (rol de plantilla de PySpark análisis) 161
Cree un rol de servicio para escribir los resultados de un PySpark trabajo 163
Creación de un rol de servicio para recibir resultados 165
Configure las funciones de servicio para AWS Clean Rooms el aprendizaje automático 169
Configure las funciones de servicio para un modelado similar 170
Configure las funciones de servicio para el modelado personalizado 183
Colaboraciones y suscripciones 198
Seleccionar un tipo de motor de análisis 199
Crear una colaboración 201
Crear una colaboración para consultas 202
Crear una colaboración para consultas y trabajos 212
Crear una colaboración para el modelado de aprendizaje automático 223
Crear una pertenencia y unirse a una colaboración 232
..... 233
Editar colaboraciones 238
Edición del nombre y la descripción de la colaboración 239
Actualice el motor de análisis de colaboración 239
Desactive el almacenamiento de registros 240
Edita la configuración de los registros de colaboración 241
Editar etiquetas de colaboración 242
Edición de etiquetas de pertenencia 243
Editar las etiquetas de tablas asociadas 244
Edición de etiquetas de plantilla de análisis 244
Edición de las etiquetas de la política de privacidad diferencial 245

Eliminar colaboraciones ... 245
Ver colaboraciones ... 246
Invitar a los miembros a una colaboración ... 246
Supervisión de miembros ... 247
Eliminar un miembro de una colaboración ... 248
Abandonar una colaboración ... 248
Tablas de datos ... 250
Formatos de datos ... 251
Formatos de datos compatibles para los trabajos PySpark ... 251
Formatos de datos compatibles para consultas SQL ... 251
Tipos de datos compatibles ... 252
Tipos de compresión de archivos para AWS Clean Rooms ... 254
Cifrado del lado del servidor para AWS Clean Rooms ... 255
Apache Iceberg tablas ... 255
Tipos de datos admitidos para las tablas de Iceberg ... 256
Preparar tablas de datos ... 257
Preparación de tablas de datos en Amazon S3 ... 258
Preparación de tablas de datos en Amazon Athena ... 260
Preparación de tablas de datos en Snowflake ... 262
Preparar tablas de datos cifrados ... 265
Paso 1: completar los requisitos previos ... 265
Paso 2: descargar el cliente de cifrado de C3R ... 266
Paso 3: (Opcional) Visualización de los comandos disponibles en el cliente de cifrado de C3R ... 267
Paso 4: generar un esquema de cifrado para un archivo tabular ... 267
Paso 5: crear una clave secreta compartida ... 275
Paso 6: guardar la clave secreta compartida en la variable de entorno ... 276
Paso 3: cifrar los datos ... 277
Paso 8: verificar el cifrado de datos ... 278
(Opcional) Crear un esquema (usuarios avanzados) ... 279
Descifrar tablas de datos ... 289
Tablas configuradas ... 291
Crear una tabla configurada ... 292
Origen de datos de Amazon S3 ... 292
Fuente de datos de Amazon Athena ... 295
Fuente de datos de Snowflake ... 297

Agregación de una regla de análisis a una tabla configurada 302
Agregación de una regla de análisis de agregación a una tabla (flujo guiado) 303
Agregación de una regla de análisis de lista a una tabla (flujo guiado) 307
Agregación de una regla de análisis personalizada a una tabla (flujo guiado) 309
Agregación de una regla de análisis a una tabla (editor JSON) 313
Pasos a seguir a continuación 315
Asociar una tabla configurada a una colaboración 315
Asociar una tabla configurada desde la página de detalles de la tabla configurada 316
Asociar una tabla configurada desde la página de detalles de la colaboración 320
Pasos a seguir a continuación 323
Agregación de una regla de análisis de colaboración a una tabla configurada 323
Configuración de la política de privacidad diferencial (opcional) 325
Visualización de los registros de uso de la privacidad diferencial 326
Edición de una política de privacidad diferencial 326
Eliminación de una política de privacidad diferencial 327
Visualización de los parámetros de privacidad diferencial calculados 328
Ver tablas y reglas de análisis 329
Editar detalles de tablas configuradas 330
Editar etiquetas de tablas configuradas 330
Editar la regla de análisis de tablas configuradas 331
Eliminar la regla de análisis de tablas configuradas 331
Columnas no permitidas de la tabla configurada 332
Editar asociaciones de tablas configuradas 336
Disociar tablas configuradas 336
AWS Entity Resolution in AWS Clean Rooms 338
Espacios de nombres de ID 339
Creación y asociación de un nuevo espacio de nombres de ID 339
Asociación de un espacio de nombres de ID existente 342
Edición de las asociaciones de espacios de nombres de ID 345
Desasociación de las asociaciones de espacios de nombres de ID 346
Tablas de asignación de ID 347
Creación y llenado de una nueva tabla de asignación de ID 348
Llenado de una tabla de asignación de ID existente 362
Edición de una tabla de asignación de ID 362
Eliminación de una tabla de asignación de ID 363
Plantillas de análisis 364

plantillas de análisis SQL ... 364
Creación de una plantilla de análisis SQL ... 365
Revisión de una plantilla de análisis SQL ... 366
PySpark plantillas de análisis ... 368
Seguridad ... 368
Limitaciones ... 369
Prácticas recomendadas ... 370
Crear un script de usuario ... 371
Crear un entorno virtual (opcional) ... 374
Almacenamiento de un script de usuario y un entorno virtual en S3 ... 375
Crear una plantilla PySpark de análisis ... 377
Revisión de una plantilla PySpark de análisis ... 380
Plantillas de PySpark análisis para la solución ... 383
Solución de problemas con el código ... 383
El trabajo de la plantilla de análisis no se inicia ... 384
El trabajo de la plantilla de análisis se inicia pero se produce un error durante el procesamiento ... 386
La configuración del entorno virtual falla ... 387
Análisis ... 389
Ejecución de consultas SQL ... 389
Consulta de tablas configuradas ... 391
Consulta de tablas de asignación de ID ... 396
Consulta de tablas configuradas mediante una plantilla de análisis SQL ... 398
Consulta con el creador de análisis ... 399
Visualización del impacto de la privacidad diferencial ... 405
Vista de consultas recientes ... 406
Visualización de detalles de consultas ... 406
Ejecutando PySpark trabajos ... 407
Ejecute PySpark el trabajo mediante una plantilla de análisis ... 408
Ver trabajos recientes ... 409
Consultar detalles del trabajo ... 410
Resultados del análisis ... 411
Recibir resultados de consultas ... 412
Recibir los resultados del trabajo ... 413
Edición de los ajustes predeterminados de los resultados de las consultas ... 414
Edición de los valores predeterminados de la configuración de los resultados del trabajo ... 416

Uso del resultado de la consulta en otros Servicios de AWS 417
Modelado de aprendizaje automático para proveedores de datos de formación 418
Importación de datos de entrenamiento 419
Crear un modelo similar 420
Configuración de un modelo similar 421
Asociar un modelo similar configurado 423
Actualización de un modelo similar configurado 423
Modelado de aprendizaje automático para proveedores de datos iniciales 425
Crear un segmento similar 425
Exportación de un segmento similar 427
Modelado personalizado 428
Crear la colaboración 429
Contribuir con datos de entrenamiento 434
Configuración de un algoritmo modelo 438
Asociar el algoritmo del modelo configurado 441
Crear un canal de entrada ML 444
Crear un modelo entrenado 446
Exportación de artefactos del modelo 447
Ejecute la inferencia en un modelo entrenado 449
Pasos a seguir a continuación 451
Solución de problemas 452
Una o más tablas a las que hace referencia la consulta no están accesibles para el rol de servicio asociado. El propietario de la tabla/rol debe conceder al rol de servicio acceso a la tabla. 452
Uno de los conjuntos de datos subyacentes tiene un formato de archivo no compatible. 452
Los resultados de las consultas no son los esperados cuando se utiliza la computación criptográfica para Clean Rooms. 453
Seguridad 454
Protección de los datos 455
Cifrado en reposo 456
Cifrado en tránsito 457
Cifrado de datos subyacentes 457
Política de claves 457
Retención de datos 460
Prácticas recomendadas 461
Mejores prácticas con AWS Clean Rooms 462

Prácticas recomendadas para utilizar reglas de análisis en AWS Clean Rooms 462
Identity and Access Management 464
Público 465
Autenticación con identidades 465
Administración de acceso mediante políticas 469
¿Cómo AWS Clean Rooms funciona con IAM 472
Ejemplos de políticas basadas en identidades 479
AWS políticas gestionadas 482
Solución de problemas 490
Prevención de la sustitución confusa entre servicios 492
Comportamientos de IAM para ML AWS Clean Rooms 493
Comportamientos de IAM para los modelos personalizados de Clean Rooms ML 496
Validación de conformidad 498
Resiliencia 499
Seguridad de la infraestructura 499
Seguridad de la red 500
AWS PrivateLink 500
Consideraciones 501
Creación de un punto de conexión de interfaz 501
Monitorización 503
Inicio de sesión de Analysis AWS Clean Rooms 503
Reciba registros de consultas y trabajos 504
Acciones recomendadas para los registros de consultas y trabajos 506
CloudTrail registros 506
AWS Clean Rooms información en CloudTrail 507
Descripción de las entradas de los archivos de AWS Clean Rooms registro 508
Ejemplos de AWS Clean Rooms CloudTrail eventos 508
AWS CloudFormation recursos 512
AWS Clean Rooms y AWS CloudFormation plantillas 512
Obtenga más información sobre AWS CloudFormation 514
Cuotas 516
AWS Clean Rooms cuotas 516
AWS Clean Rooms límites de parámetros de recursos 524
AWS Clean Rooms Limitación de las cuotas de API 525
AWS Clean Rooms Cuotas de ML 528
Limpia las cuotas de la API Clean Rooms ML 533

Historial de documentos 540
Glosario 550
Regla de análisis de agregación 550
Reglas de análisis 550
Plantilla de análisis 550
AWS Clean Rooms Motor de análisis de SQL 551
Cliente de cifrado de S3 551
Columna de texto sin cifrar 551
Colaboración 551
Creador de la colaboración 552
Tabla configurada 552
Regla de análisis personalizada 553
Descifrado 553
Privacidad diferencial 553
Cifrado 553
Columna de huella digital 553
Método de flujo de trabajo de asignación de ID 553
tabla de asignación de ID 554
Regla de análisis de tablas de asignación de ID 554
Flujo de trabajo de asignación de ID 554
Espacio de nombres de ID 555
Asociación de espacios de nombres de ID 555
Trabajo 555
Regla de análisis de lista 555
Modelo parecido 555
Un segmento similar 555
Miembro 556
Miembro que puede realizar consultas 556
Miembro que puede ejecutar consultas y trabajos 556
Miembro que puede recibir los resultados 556
Miembro que paga los costos de computación de consultas 557
El miembro paga los costos de consulta y cálculo del trabajo 557
Pertenencia 557
Columna sellada 558
Datos iniciales 558
Motor de análisis Spark 558

Consultar 558
..... dlix

¿Qué es AWS Clean Rooms?

AWS Clean Rooms le ayuda a usted y a sus socios a analizar y colaborar en sus conjuntos de datos colectivos para obtener nuevos conocimientos sin revelarse los datos subyacentes entre sí. AWS Clean Rooms es un espacio de trabajo de colaboración seguro, en el que puede crear sus propias salas limpias en cuestión de minutos y analizar sus conjuntos de datos colectivos en tan solo unos pasos. Tú eliges los socios con los que quieres colaborar, seleccionas sus conjuntos de datos y configuras los controles que mejoran la privacidad de esos socios.

Con él AWS Clean Rooms, puede colaborar con miles de empresas que ya los utilizan. AWS La colaboración no requiere mover los datos de otro proveedor de servicios en la nube AWS ni cargarlos en él. Cuando ejecuta consultas o trabajos, AWS Clean Rooms lee los datos de su ubicación original y aplica reglas de análisis integradas para ayudarle a mantener el control sobre esos datos.

AWS Clean Rooms proporciona controles de acceso a los datos y controles de soporte de auditoría integrados que puede configurar. Estos controles incluyen:

- Reglas de análisis para restringir las consultas SQL y proporcionar restricciones de salida.
- Computación criptográfica Clean Rooms para mantener los datos cifrados, incluso mientras se procesan las consultas, a fin de cumplir con las estrictas políticas de manejo de datos.
- Registros de análisis para revisar las consultas y los trabajos AWS Clean Rooms y ayudar a respaldar las auditorías.
- Privacidad diferencial para proteger a los usuarios de los intentos de identificación. AWS Clean Rooms La privacidad diferencial es una capacidad totalmente gestionada que protege la privacidad de sus usuarios con técnicas respaldadas matemáticamente y controles intuitivos que puede aplicar en unos pocos pasos.
- AWS Clean Rooms El aprendizaje automático permite a dos partes identificar a usuarios similares en sus datos sin necesidad de compartir sus datos entre sí. La primera parte crea y configura un modelo similar a partir de sus datos de entrenamiento. A continuación, los datos iniciales se traen a la colaboración para crear un segmento similar que se parece a los datos de entrenamiento.

En el siguiente vídeo se explica más sobre AWS Clean Rooms.

AWS Clean Rooms

¿Es la primera vez que lo usa AWS Clean Rooms ?

Si es la primera vez que lo utiliza AWS Clean Rooms, le recomendamos que comience leyendo las siguientes secciones:

- ¿Cómo funciona AWS Clean Rooms
- Acceder AWS Clean Rooms
- Con AWS Clean Rooms figuración
- AWS Clean Rooms Glosario

¿Cómo funciona AWS Clean Rooms

En AWS Clean Rooms, puedes crear una colaboración y añadir la Cuentas de AWS que quieras invitar, o bien crear una membresía para unirte a una colaboración a la que te hayan invitado. A continuación, vincule los recursos de datos necesarios para su caso de uso: tablas configuradas para datos de eventos, modelos configurados para el modelado de ML o espacios de nombres de ID para la resolución de entidades. Tiene la opción de crear o aprobar plantillas de análisis para acordar de antemano las consultas y los trabajos exactos que desea permitir en una colaboración. Por último, analiza los datos conjuntos mediante la ejecución de consultas o PySpark trabajos de SQL en las tablas configuradas, la resolución de entidades en las tablas de mapeo de ID o el uso de modelos de aprendizaje automático para generar segmentos de audiencia similares.

El siguiente diagrama muestra cómo AWS Clean Rooms funciona.

Servicios relacionados

AWS servicios

Los siguientes Servicios de AWS están relacionados con AWS Clean Rooms:

- Amazon Athena

  Los miembros de Collaboration pueden almacenar los datos que incorporan AWS Clean Rooms como AWS Glue Data Catalog vistas en Amazon Athena. Para obtener más información, consulte los temas siguientes:

  Para obtener más información, consulte los temas siguientes:

Preparación de tablas de datos para consultas en AWS Clean Rooms

Creación de una tabla configurada: fuente de datos de Amazon Athena

¿Qué es Amazon Athena? en la Guía del usuario de Amazon Athena

- AWS CloudFormation

  Cree los siguientes recursos en AWS CloudFormation: colaboraciones, tablas configuradas, asociaciones de tablas configuradas y membresías

  Para obtener más información, consulte Creación de AWS Clean Rooms recursos con AWS CloudFormation.

- AWS CloudTrail

  Úselo AWS Clean Rooms con CloudTrail los registros para mejorar el análisis de la Servicio de AWS actividad.

  Para obtener más información, consulte Registro de llamadas a la AWS Clean Rooms API mediante AWS CloudTrail.

- AWS Entity Resolution

  Úselo AWS Clean Rooms con AWS Entity Resolution para realizar la resolución de entidades.

  Para obtener más información, consulte AWS Entity Resolution in AWS Clean Rooms.

- AWS Glue

  Los miembros de Collaboration pueden crear AWS Glue tablas a partir de sus datos en Amazon S3 para usarlas en ellas AWS Clean Rooms.

  Para obtener más información, consulte los temas siguientes:

  Preparación de tablas de datos para consultas en AWS Clean Rooms

  ¿Qué es AWS Glue? en la Guía para desarrolladores de AWS Glue

- Amazon Simple Storage Service (Amazon S3)

  Los miembros de Collaboration pueden almacenar los datos que introduzcan AWS Clean Rooms en Amazon S3.

Para obtener más información, consulte los temas siguientes:

Preparación de tablas de datos para consultas en AWS Clean Rooms

Creación de una tabla configurada: fuente de datos de Amazon S3

¿Qué es Amazon S3 en la Guía del usuario de Amazon Simple Storage Service

- AWS Secrets Manager

  Los miembros de la colaboración pueden crear secretos para acceder a los datos almacenados en Snowflake y leerlos.

  Para obtener más información, consulte los temas siguientes:

  Cree un rol de servicio para leer los datos de Snowflake

  Preparación de tablas de datos para consultas en AWS Clean Rooms

  ¿Qué es AWS Secrets Manager? en la Guía del usuario de AWS Secrets Manager

Servicios de terceros

El siguiente servicio de terceros está relacionado con: AWS Clean Rooms

- Snowflake

  Los miembros de Collaboration pueden almacenar los datos que introduzcan AWS Clean Rooms en un almacén de Snowflake.

  Para obtener más información, consulte los temas siguientes:

  Preparación de tablas de datos para consultas en AWS Clean Rooms

  Creación de una tabla configurada: fuente de datos de Snowflake

Acceder AWS Clean Rooms

Puede acceder a AWS Clean Rooms través de las siguientes opciones:

- Directamente a través de la AWS Clean Rooms consola en https://console.aws.amazon.com/cleanrooms/.
- Programáticamente a través de la AWS Clean Rooms API. Para obtener más información, consulte la Referencia de la API de AWS Clean Rooms .

Precios para AWS Clean Rooms

Para obtener información acerca de los precios, consulte Precios de AWS Clean Rooms.

> **Note**
>
> En el caso de los miembros de la colaboración que hayan asociado datos almacenados en Snowflake, el proveedor de almacenamiento de datos o el proveedor de nube correspondiente les cobrará tanto la salida como el cálculo de los datos cada vez que se ejecute una consulta que utilice datos almacenados en esas ubicaciones.

Facturación para AWS Clean Rooms

AWS Clean Rooms permite al creador de la colaboración designar qué miembro pagará los costes de cálculo de la consulta o del trabajo en la colaboración.

En la mayoría de los casos, el miembro que puede realizar consultas es también el miembro que paga los costos de computación de las consultas. No obstante, si el miembro que puede realizar consultas y el miembro que paga los costos de computación de las consultas son diferentes, cuando el miembro que puede realizar consultas ejecuta las consultas en su propio recurso de pertenencia, se factura al recurso de pertenencia del miembro que paga los costos de computación de las consultas.

El miembro que paga los costes de cálculo de las consultas no ve ningún evento relacionado con las consultas que se estén ejecutando en su historial de CloudTrail eventos porque el pagador no es quien ejecuta las consultas ni es el propietario del recurso con el que se ejecutan las consultas. Sin embargo, el pagador ve los cargos generados en su recurso de membresía por todas las consultas realizadas por el miembro que puede ejecutar consultas en la colaboración.

Para obtener más información acerca de cómo crear una colaboración y configurar el miembro que paga los costos de computación de las consultas, consulte Crear una colaboración.

Reglas de análisis en AWS Clean Rooms

Como parte de la habilitación de una tabla para utilizarla en AWS Clean Rooms el análisis de la colaboración, el miembro de la colaboración debe configurar una regla de análisis.

Una regla de análisis es un control que mejora la privacidad y que cada propietario de datos define en una tabla configurada. Una regla de análisis determina cómo se puede analizar la tabla configurada.

La regla de análisis es un control de nivel de cuenta en la tabla configurada (un recurso de nivel de cuenta) y que se aplica a cualquier colaboración a la que esté asociada la tabla configurada. Si no hay ninguna regla de análisis configurada, la tabla configurada se puede asociar a colaboraciones, pero no se puede consultar. Las consultas solo pueden hacer referencia a tablas configuradas con el mismo tipo de regla de análisis.

Para configurar una regla de análisis, primero debe seleccionar un tipo de análisis y, a continuación, especificar la regla de análisis. En ambos pasos, debe tener en cuenta el caso de uso que desea habilitar y cómo desea proteger los datos subyacentes.

AWS Clean Rooms aplica los controles más restrictivos en todas las tablas configuradas a las que se hace referencia en una consulta.

En los siguientes ejemplos se ilustran los controles restrictivos.

Example Control restrictivo: restricción de salida

- El colaborador A tiene una restricción de salida de 100 en la columna de identificador.
- El colaborador B tiene una restricción de salida de 150 en la columna de identificador.

  Una consulta de agregación que haga referencia a ambas tablas configuradas requiere al menos 150 valores diferenciados de identificador dentro de una fila de salida para que se muestre en el resultado de la consulta. El resultado de la consulta no indica que los resultados se eliminan debido a la restricción de salida.

Example Control restrictivo: plantilla de análisis no aprobada

- El colaborador A ha permitido una plantilla de análisis con una consulta que hace referencia a tablas configuradas de los colaboradores A y B en su regla de análisis personalizada.
- El colaborador B no ha permitido la plantilla de análisis.

Como el colaborador B no ha permitido la plantilla de análisis, el miembro que puede realizar consultas no puede ejecutar dicha plantilla de análisis.

Tipos de regla de análisis

Existen tres tipos de reglas de análisis: de agregación, de lista y personalizadas. En las siguientes tablas se comparan los tipos de reglas de análisis. Cada tipo tiene dedicada una sección independiente en la que se describe la especificación de la regla de análisis.

> **Note**
>
> Existe un tipo de regla de análisis llamado regla de análisis de tablas de asignación de ID. Sin embargo, esta regla de análisis la administra AWS Clean Rooms y no se puede modificar. Para obtener más información, consulte Regla de análisis de tablas de asignación de ID.

En las siguientes secciones se describen los casos de uso admitidos y los controles para cada tipo de regla de análisis.

Casos de uso admitidos

En las siguientes tablas se muestra un resumen comparativo de los casos de uso admitidos para cada tipo de regla de análisis.

Caso de uso	Agregación	Lista	Personalizada
Análisis admitidos	Consultas que agregan estadísticas mediante las funciones COUNT, SUM y AVG en dimensiones opcionales	Consultas que generan listas de nivel de fila de la superposición entre varias tablas	Cualquier análisis personalizado, siempre que la plantilla de análisis o el creador del análisis se hayan revisado y permitido

Caso de uso	Agregación	Lista	Personalizada
Casos de uso comunes	Análisis de segmentos, medición y atribución	Enriquecimiento, creación de segmentos	Atribución de primer toque, análisis increment ales, detección de audiencias
Constructos SQL	• Instrucciones JOIN: INNER JOIN • Funciones de agregado: COUNT/ COUNT DISTINCT, SUM/SUM DISTINCT y AVG • Funciones escalares: subconjunto limitado	• Instrucciones JOIN: INNER JOIN • Funciones escalares: ninguna	La mayoría de las funciones SQL y constructos SQL están disponibles con el comando SELECT
Subconsultas y expresiones de tabla comunes () CTEs	No	No	Sí
Plantillas de análisis	No	No	Sí

Controles admitidos

En las siguientes tablas se muestra un resumen comparativo del modo en que cada tipo de regla de análisis protege los datos subyacentes.

Controlar	Agregación	Lista	Personalizada
Mecanismo de control	Controle cómo se pueden usar los datos de la tabla en una consulta	Controle cómo se pueden usar los datos de la tabla en una consulta	Controle qué consultas pueden ejecutarse en la tabla

Controlar	Agregación	Lista	Personalizada
	(por ejemplo, permitir COUNT y SUM de la columna hashed_email).	(por ejemplo, permitir el uso de la columna hashed_email solo en combinaciones).	(por ejemplo, permitir solo las consultas definidas en las plantillas de análisis "Consulta personalizada 1").
Técnicas de mejora de la privacidad integrada	• Coincidencia ciega • Agregación obligatoria • Umbral de agregación mín. >= • 2 Estructura de consulta predefinida	• Coincidencia ciega • Solapamiento obligatorio • Estructura de consulta predefinida • Análisis adicionales permitidos	• Privacidad diferencial • Columnas de salida no permitidas
Revisar la consulta antes de que pueda ejecutarse	No	No	Sí, con plantillas de análisis

Para obtener más información sobre las reglas de análisis disponibles en AWS Clean Rooms, consulte los siguientes temas.

- Regla de análisis de agregación
- Regla de análisis de lista
- Regla de análisis personalizada en AWS Clean Rooms

Regla de análisis de agregación

En AWS Clean Rooms, una regla de análisis de agregación genera estadísticas agregadas mediante las funciones COUNT, SUM y/o AVG en dimensiones opcionales. Cuando la regla de análisis de agregación se agrega a una tabla configurada, permite al miembro que puede realizar la consulta ejecutar consultas en la tabla configurada.

La regla de análisis de agregación admite casos de uso tales como la planificación de campañas, el alcance mediático, la medición de la frecuencia y la atribución.

La estructura y sintaxis de consulta admitidas se definen en Estructura y sintaxis de consultas de agregación.

Los parámetros de la regla de análisis, definidos en Regla de análisis de agregación: controles de consulta, incluyen los controles de consulta y los controles de resultados de las consultas. Sus controles de consulta incluyen la posibilidad de imponer como requisito que una tabla configurada se una a al menos una tabla configurada propiedad del miembro que puede realizar la consulta, ya sea de forma directa o transitiva. Este requisito le permite asegurarse de que la consulta se ejecute en la intersección (INNER JOIN) de su tabla y la de ellos.

Estructura y sintaxis de consultas de agregación

Las consultas en tablas que tienen una regla de análisis de agregación deben respetar la siguiente sintaxis.

```
--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}]] [,...]]
```

En la siguiente tabla se explica cada una de las expresiones enumeradas en la sintaxis anterior.

Expression	Definición	Ejemplos
select_aggregate_function_expression	Una lista separada por comas que contiene las siguientes expresiones: • select_aggregation_function_expression • select_aggregate_expression **Note** Debe haber al menos una select_aggregation_function_expression en la select_aggregate_expression .	SELECT SUM(PRICE), user_segment
select_aggregation_function_expression	Una o más funciones de agregación admitidas aplicadas a una o más columnas. Solo se permiten columnas como argumentos de las funciones de agregación. **Note** Debe haber al menos una select_aggregation	AVG(PRICE) COUNT(DISTINCT user_id)

Expression	Definición	Ejemplos
	`_function _expression` en la `select_aggregate_expression` .	

`select_grouping_column_expression`	Expresión que puede contener cualquier expresión que utilice lo siguiente:
	• Nombres de columna de la tabla
	• Funciones escalares admitidas
	• Literales de cadena
	• Literales numéricos
	
	Note
	`select_aggregate_expression` puede asignar un alias a las columnas con o sin el parámetro `AS`. Para obtener más información, consulte Referencia de SQL de AWS Clean Rooms.
		`TRUNC(timestampColumn)`
		
		`UPPER(campaignName)`

Expression	Definición	Ejemplos
table_expression	Una tabla o combinación de tablas que conecta expresion es condicionales de unión con `join_condition` . `join_condition` devuelve un valor booleano. La `table_expression` admite: • Un específico JOIN tipo (INNER JOIN) • La condición de comparación de igualdad dentro de una `join_condition` (=) • Operadores lógicos (`AND`, `OR`).	`FROM consumer_table INNER JOIN provider_table ON consumer_table.identifier1 = provider_table.identifier1 AND consumer_table.identifier2 = provider_table.identifier2`

Expression	Definición	Ejemplos

where_expression	Una expresión condicional que devuelve un valor booleano. Puede constar de lo siguiente:
	• Nombres de columna de la tabla
	• Funciones escalares admitidas
	• Operadores matemáticos
	• Literales de cadena
	• Literales numéricos
	
	Las condiciones de comparación admitidas son (`=, >, <, <=, >=, <>, !=, NOT, IN, NOT IN, LIKE, IS NULL, IS NOT NULL`).
	
	Los operadores lógicos admitidos son (`AND, OR`).
	
	La `where_expression` es opcional.
		`WHERE where_condition`
		
		`WHERE price > 100`
		
		`WHERE TRUNC(timestampColumn) = '1/1/2022'`
		
		`WHERE timestampColumn = timestampColumn2 - 14`

group_by_expression	Lista separada por comas de expresiones que cumplen con los requisitos de `select_grouping_column_expression` .	`GROUP BY TRUNC(timestampColumn), UPPER(campaignName), segment`

Expression	Definición	Ejemplos
having_expression	Una expresión condicion al que devuelve un valor booleano. Tienen una función de agregación admitida aplicada a una sola columna (por ejemplo, `SUM(price)`) y se comparan con un literal numérico. Las condiciones admitidas son (`=, >, <, <=, >=, <>, !=`). Los operadores lógicos admitidos son (`AND, OR`). La `having_expression` es opcional.	`HAVING SUM(SALES) > 500`

Expression	Definición	Ejemplos
order_by_expression	Lista de expresiones separadas por comas que es compatible con los mismos requisitos definidos anteriormente en `select_aggregate_expression` . La `order_by_expression` es opcional. **Note** `order_by_expression` admite los parámetros `ASC` y `DESC`. Para obtener más información, consulte los parámetros ASC DESC en Referencia de SQL de AWS Clean Rooms.	`ORDER BY SUM(SALES), UPPER(campaignName)`

En cuanto a la estructura y sintaxis de las consultas de agregación, tenga en cuenta lo siguiente:

- Comandos SQL distintos de SELECT no se admiten.
- Subconsultas y expresiones de tabla comunes (por ejemplo, WITH) no se admiten.
- Operadores que combinan varias consultas (por ejemplo, UNION) no se admiten.
- TOP, LIMIT, y OFFSET no se admiten los parámetros.

Regla de análisis de agregación: controles de consulta

Los controles de consulta de agregación permiten controlar cómo se utilizan las columnas de la tabla para consultarla. Por ejemplo, puede controlar qué columna se usa para unir, qué columna se puede contar o en qué columna se puede usar WHERE declaraciones.

En las secciones siguientes se explica cada uno de los controles.

Temas

- Controles de agregación
- Controles de combinación
- Controles de dimensión
- Funciones escalares

Controles de agregación

El uso de controles de agregación permite definir qué funciones de agregación se van a permitir y a qué columnas se deben aplicar. Las funciones de agregación se pueden utilizar en el SELECT, HAVING, y ORDER BY expresiones.

Control	Definición	Uso
`aggregateColumns`	Columnas de tablas configuradas que se permite utilizar en las funciones de agregación.	`aggregateColumns` se puede utilizar dentro de una función de agregación en el SELECT, HAVING, y ORDER BY expresiones. Algunas `aggregateColumns` también se pueden categorizar como `joinColumn` (definición disponible más adelante). La `aggregateColumn` dada no se puede categorizar también como `dimension`

Control	Definición	Uso
		`Column` (definición disponible más adelante).
`function`	Las funciones COUNT, SUM y AVG que se permite utilizar además de `aggregateColumns`.	La `function` se puede aplicar a una `aggregateColumns` que esté asociada a ella.

Controles de combinación

Se utiliza una cláusula `JOIN` para combinar filas de dos o más tablas en función de una columna relacionada entre ellas.

Puede utilizar los controles de unión para controlar cómo se puede unir la tabla a otras tablas del`table_expression`. AWS Clean Rooms solo admite INNER JOIN. INNER JOIN las declaraciones solo pueden usar columnas que se hayan clasificado explícitamente como una `joinColumn` en su regla de análisis, sujetas a los controles que usted defina.

La INNER JOIN deben funcionar en una `joinColumn` tabla configurada y en otra tabla configurada `joinColumn` de la colaboración. Usted decide qué columnas de la tabla se pueden usar como `joinColumn`.

Cada condición coincide dentro del ON la cláusula es necesaria para utilizar la condición de comparación de igualdad (=) entre dos columnas.

Varias condiciones de coincidencia dentro de un ON las cláusulas pueden ser:

- Combinación con el operador lógico `AND`
- Separación mediante el operador lógico `OR`

> **Note**
>
> Todos JOIN las condiciones de coincidencia deben coincidir con una fila de cada lado del JOIN. Todos los condicionales conectados por un operador `AND` lógico `OR` o uno lógico también deben cumplir con este requisito.

A continuación se muestra un ejemplo de consulta con un operador lógico AND.

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

A continuación se muestra un ejemplo de consulta con un operador lógico OR.

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

Control	Definición	Uso
`joinColumns`	Las columnas (si las hay) que desea permitir que el miembro que puede realizar la consulta las utilice en el INNER JOIN statement.	También se puede categorizar una `joinColumn` como `aggregateColumn` (consulteControles de agregación). La misma columna no se puede usar como `joinColumn` y `dimensionColumns` a la vez (consulte más adelante). A menos que también se haya clasificado como una`aggregateColumn`, a no se `joinColumn` puede usar en ninguna otra parte de la consulta que no sea la INNER JOIN.
`joinRequired`	Controle si necesita un INNER JOIN con una tabla configura	Si habilita este parámetro, un INNER JOIN es obligatorio. Si

Control	Definición	Uso
	da por el miembro que puede realizar la consulta.	no habilita este parámetro, un INNER JOIN es opcional. Si se habilita este parámetro, el miembro que puede realizar la consulta debe incluir una tabla de su propiedad en el INNER JOIN. Deben JOIN su mesa con la de ellos, ya sea de forma directa o transitiva (es decir, unir su mesa a otra mesa, que a su vez está unida a la suya).

A continuación se muestra un ejemplo de transitividad.

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

> **Note**
>
> El miembro que puede realizar la consulta también puede usar el parámetro `joinRequired`. En ese caso, la consulta debe combinar su tabla con al menos otra tabla.

Controles de dimensión

Los controles de dimensión controlan la columna en la cual se pueden filtrar, agrupar o agregar las columnas de agregación.

Control	Definición	Uso
`dimensionColumns`	Las columnas (si las hay) que permites que utilice el miembro que puede realizar la consulta SELECT, WHERE, GROUP BY, y ORDER BY.	A `dimensionColumn` se puede usar en SELECT (`select_grouping_column_expression`), WHERE, GROUP BY, y ORDER BY. Una misma columna no puede ser a la vez `dimensionColumn`, `joinColumn` y/o `aggregateColumn`.

Funciones escalares

Las funciones escalares controlan qué funciones escalares se pueden usar en las columnas de dimensión.

Control	Definición	Uso
`scalarFunctions`	Las funciones escalares que se pueden utilizar en `dimensionColumns` en la consulta.	Especifica las funciones escalares (si las hay) que se permiten (por ejemplo, CAST) a las que aplicar. `dimensionColumns` Las funciones escalares no se pueden usar además de otras funciones ni dentro de otras funciones. Los argumentos de las funciones escalares pueden ser columnas, literales de cadena o literales numéricos.

Se admiten las siguientes funciones escalares:

- Funciones matemáticas: ABS, CEILING, FLOOR, LOG, LN, ROUND, SQRT
- Funciones de formato de tipos de datos: CAST, CONVERT, TO_CHAR, TO_DATE, TO_NUMBER, TO_TIMESTAMP
- Funciones de cadena: LOWER, UPPER, TRIM, RTRIM, SUBSTRING
  - En el caso de RTRIM, no se permiten conjuntos de caracteres personalizados para recortar.
- Expresiones condicionales: COALESCE
- Funciones de fecha EXTRACT, GETDATE, CURRENT_DATE, DATEADD
- Otras funciones: TRUNC

Para obtener más información, consulte Referencia de SQL de AWS Clean Rooms.

Regla de análisis de agregación: controles de resultados de consulta

Los controles de resultados de consulta de agregación le permiten controlar qué resultados se devuelven especificando una o más condiciones que debe cumplir cada fila de salida. AWS Clean Rooms admite restricciones de agregación en forma de `COUNT (DISTINCT column) >= X`. Este formato requiere que cada fila agregue al menos X valores diferenciados de una selección de la tabla configurada (por ejemplo, un número mínimo de valores `user_id` diferenciados). Este umbral mínimo se aplica automáticamente, incluso si la consulta enviada en sí misma no utiliza la columna especificada. Se aplican de manera conjunta en cada tabla configurada de la consulta desde las tablas configuradas de cada miembro de la colaboración.

Cada tabla configurada debe tener al menos una restricción de agregación en su regla de análisis. Los propietarios de las tablas configuradas pueden añadir varios `columnName` y su `minimum` asociado y estos se aplicarán conjuntamente.

Restricciones de agregación

Las restricciones de agregación controlan qué filas de los resultados de la consulta se devuelven. Para incluirse en los resultados devueltos, una fila debe cumplir con el número mínimo especificado de valores diferenciados en cada columna especificada en la restricción de agregación. Este requisito se aplica incluso si la columna no se menciona explícitamente en la consulta o en otras partes de la regla de análisis.

Control	Definición	Uso
`columnName`	La `aggregateColumn` que se usa en la condición que debe cumplir cada fila de salida.	Puede tratarse de cualquier columna de la tabla configurada.
`minimum`	El número mínimo de valores diferenciados de la `aggregateColumn` asociada que debe tener la fila de salida (por ejemplo, COUNT DISTINCT) para que se devuelva en los resultados de la consulta.	El `minimum` debe tener al menos un valor de 2.

Estructura de la regla de análisis de agregación

El siguiente ejemplo muestra una estructura predefinida para una regla de análisis de agregación.

En el siguiente ejemplo, *`MyTable`* hace referencia a nuestra tabla de datos. Puede reemplazar cada una *`user input placeholder`* con su propia información.

```
{
  "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]
    },
  ]
}
```

Regla de análisis de agregación: ejemplo

El siguiente ejemplo demuestra cómo dos empresas pueden colaborar en el AWS Clean Rooms uso del análisis de agregación.

La empresa A tiene datos de clientes y de ventas. La empresa A está interesada en conocer la actividad de devolución de productos. La empresa B es uno de los minoristas de la empresa A y dispone de datos sobre devoluciones. La empresa B también tiene atributos de segmento sobre los clientes que son útiles para la empresa A (por ejemplo, compra de productos relacionados o uso del servicio de atención al cliente del minorista). La empresa B no quiere proporcionar información sobre atributos ni datos sobre devoluciones de los clientes por fila. La empresa B solo quiere habilitar un conjunto de consultas para que la empresa A obtenga estadísticas agregadas sobre los clientes que se superponen dentro de un umbral de agregación mínimo.

La empresa A y la empresa B deciden colaborar para que la empresa A pueda entender la actividad de devolución de productos y ofrecer mejores productos en la empresa B y en otros canales.

Para crear la colaboración y realizar un análisis de agregación, las empresas hacen lo siguiente:

1. La empresa A crea una colaboración y crea una pertenencia. La colaboración tiene a la empresa B como otro miembro de la colaboración. La empresa A habilita el registro de consultas en la colaboración y habilita el registro de consultas en su cuenta.
2. La empresa B crea una pertenencia en la colaboración. Habilita el registro de consultas en su cuenta.
3. La empresa A crea una tabla configurada de ventas.
4. La empresa A añade la siguiente regla de análisis de agregación a la tabla configurada de ventas.

```
{
  "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`: la empresa A quiere contar el número de clientes únicos que se superponen entre los datos de ventas y los datos de devoluciones. La empresa A también quiere sumar el número de `purchases` fabricados para compararlo con el número de `returns`.

`joinColumns`: la empresa A desea usar `identifier` para cotejar los clientes de los datos de ventas con los clientes de los datos de devoluciones. Esto ayudará a la empresa A a relacionar las devoluciones con las compras correctas. También ayudará a la empresa A a segmentar los clientes superpuestos.

`dimensionColumns`: la empresa A usa `dimensionColumns` para filtrar por un producto específico, comparar las compras y las devoluciones correspondientes a un periodo de tiempo

determinado, asegurarse de que la fecha de devolución sea posterior a la fecha del producto y ayudar a segmentar los clientes superpuestos.

`scalarFunctions`: la empresa A selecciona la función escalar CAST para ayudar a actualizar los formatos de tipo de datos si es necesario, basándose en la tabla configurada que la empresa A ha asociado a la colaboración. También añade funciones escalares para ayudar a dar formato a las columnas si es necesario.

`outputConstraints`: la empresa A establece restricciones de salida mínimas. No necesita restringir los resultados, ya que el analista puede ver los datos de la tabla de ventas por fila.

> **Note**
>
> La empresa A no incluye `joinRequired` en la regla de análisis. Ofrece flexibilidad para que el analista consulte solo la tabla de ventas.

5. La empresa B crea una tabla configurada de devoluciones.
6. La empresa B añade la siguiente regla de análisis de agregación a la tabla configurada de devoluciones.

```
{
  "aggregateColumns": [
    {
      "columnNames": [
        "identifier"
      ],
      "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`: la empresa B permite a la empresa A sumar las `returns` para compararlas con el número de compras. Tienen al menos una columna agregada porque han habilitado una consulta agregada.

`joinColumns`: la empresa B permite a la empresa A combinar por `identifier` para cotejar los clientes de los datos de devoluciones con los clientes de los datos de ventas. Los datos de `identifier` son especialmente sensibles, y tenerlos como `joinColumn` garantiza que los datos no se incluyan nunca en una consulta.

`joinRequired`: la empresa B impone como requisito que las consultas sobre los datos de devoluciones se superpongan con los datos de ventas. No quieren permitir que la empresa A consulte todas las personas de su conjunto de datos. También acordaron esa restricción en su contrato de colaboración.

`dimensionColumns`: la empresa B permite a la empresa A filtrar y agrupar por `state`, `popularpurchases` y `customerserviceuser`, que son atributos únicos que podrían ayudar a realizar el análisis para la empresa A. La empresa B permite a la empresa A usar `returndate` para filtrar la salida por `returndate` que sea posterior a `purchasedate`. Con este filtrado, la salida es más precisa a la hora de evaluar el impacto del cambio de producto.

`scalarFunctions`: la empresa B habilita lo siguiente:

- TRUNC para las fechas
- LOWER y UPPER en caso de que `producttype` se introduzca en un formato distinto en sus datos
- CAST si la empresa A necesita convertir los tipos de datos de las ventas para que sean iguales a los tipos de datos de las devoluciones

La empresa A no habilita otras funciones escalares porque no cree que sean necesarias para las consultas.

`outputConstraints`: la empresa B establece restricciones de salida mínimas en `hashedemail` para reducir en mayor medida la posibilidad de volver a identificar a los clientes. También añade una restricción de salida mínima en `producttype` para reducir en mayor medida la posibilidad de volver a identificar los productos específicos devueltos. Determinados tipos de productos podrían ser más dominantes en función de las dimensiones de la salida (por ejemplo, `state`). Sus restricciones de salida se aplicarán siempre, independientemente de las restricciones de salida que añada la empresa A a sus datos.

7. La empresa A crea una asociación a la tabla de ventas en la colaboración.
8. La empresa B crea una asociación a la tabla de devoluciones en la colaboración.
9. La empresa A realiza consultas, como las del ejemplo siguiente, para entender mejor la cantidad de devoluciones de la empresa B en comparación con el total de compras por ubicación en 2022.

```
SELECT
  companyB.state,
  SUM(companyB.returns),
  COUNT(DISTINCT companyA.hashedemail)
```

```
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. La empresa A y la empresa B revisan los registros de consultas. La empresa B comprueba que la consulta se ajusta a lo acordado en el contrato de colaboración.

Solución de problemas relacionados con reglas de análisis de agregación

Utilice la información que se incluye aquí para diagnosticar y solucionar problemas frecuentes cuando utilice reglas de análisis de agregación.

Problemas

- Mi consulta no ha devuelto ningún resultado

Mi consulta no ha devuelto ningún resultado

Esto puede ocurrir cuando no hay resultados coincidentes o cuando los resultados coincidentes no cumplen uno o más umbrales de agregación mínimos.

Para obtener más información sobre los umbrales de agregación mínimos, consulte Regla de análisis de agregación: ejemplo.

Regla de análisis de lista

En AWS Clean Rooms, una regla de análisis de listas genera listas a nivel de fila de la superposición entre la tabla configurada a la que se agrega y las tablas configuradas del miembro que puede realizar la consulta. El miembro que puede realizar consultas ejecuta consultas que incluyen una regla de análisis de lista.

El tipo de regla de análisis de listas admite casos de uso como el enriquecimiento y la creación de audiencia.

Para obtener más información sobre la estructura y la sintaxis de consulta predefinidas de esta regla de análisis, consulte Estructura predefinida de la regla de análisis de lista.

Los parámetros de la regla de análisis de lista, que se define en Regla de análisis de lista: controles de consulta, tienen controles de consulta. Estos controles de consulta incluyen la posibilidad de seleccionar las columnas que se pueden enumerar en la salida. La consulta debe tener al menos una combinación con una tabla configurada del miembro que puede realizar consultas, ya sea de forma directa o transitiva.

No hay controles de resultados de consulta como los que existen para la regla de análisis de agregación.

Las consultas de lista solo pueden utilizar operadores matemáticos. No pueden usar otras funciones (por ejemplo, de agregación o escalares).

Temas

- Estructura y sintaxis de las consultas de lista
- Regla de análisis de lista: controles de consulta
- Estructura predefinida de la regla de análisis de lista
- Regla de análisis de lista: ejemplo

Estructura y sintaxis de las consultas de lista

Las consultas de las tablas que tienen una regla de análisis de lista deben respetar la siguiente sintaxis.

```
--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]
```

En la siguiente tabla se explica cada una de las expresiones enumeradas en la sintaxis anterior.

Expression	Definición	Ejemplos
select_list_expression	Una lista separada por comas que contiene al menos un nombre de columna de tabla. Es obligatorio un parámetro DISTINCT. **Note** La select_list_expression pueden asignar un alias a las columnas con o sin el parámetro AS. Para obtener más información, consulte Referencia de SQL de AWS Clean Rooms.	SELECT DISTINCT segment
table_expression	Una tabla, o combinación de tablas, con join_condition para conectarla a join_condition . join_condition devuelve un valor booleano. La table_expression admite: • Un tipo JOIN específico (INNER JOIN) • Las condiciones de comparación de igualdad	FROM consumer_table INNER JOIN provider_table ON consumer_table.identifier1 = provider_table.identifier1 AND consumer_table.identifier2 = provider_table.identifier2

Expression	Definición	Ejemplos
	dentro de una `join_condition (=)` • Operadores lógicos (`AND`, `OR`).	

where_expression	Una expresión condicional que devuelve un valor booleano. Puede constar de los siguientes elementos:
	• Nombres de columna de la tabla
	• Operadores matemáticos
	• Literales de cadena
	• Literales numéricos
	
	Las condiciones de comparación admitidas son (`=, >, <, <=, >=, <>, !=, NOT, IN, NOT IN, LIKE, IS NULL, IS NOT NULL`).
	
	Los operadores lógicos admitidos son (`AND, OR`).
	
	La `where_expression` es opcional.
		`WHERE state + '_' + city = 'NY_NYC'`
		
		`WHERE timestampColumn = timestampColumn2 - 14`

limit_expression	Esta expresión debe tener adoptar un entero positivo. La `limit_expression` es opcional.	`LIMIT 100`

En cuanto a la estructura y sintaxis de las consultas de lista, tenga en cuenta lo siguiente:

- No se admiten comandos SQL distintos de SELECT.
- No se admiten subconsultas ni expresiones de tabla comunes (por ejemplo, WITH).
- No se admiten las cláusulas HAVING, GROUP BY ni ORDER BY.
- No se admite el parámetro OFFSET.

Regla de análisis de lista: controles de consulta

Los controles de consulta de lista le permiten controlar cómo se utilizan las columnas de su tabla a la hora de consultar esta última. Por ejemplo, puede controlar qué columna se usa para combinar, o qué columna se puede usar en una instrucción SELECT y en una cláusula WHERE.

En las secciones siguientes se explica cada uno de los controles.

Temas

- Controles de combinación
- Controles de lista

Controles de combinación

Los controles de combinación le permiten controlar cómo se puede combinar su tabla a otras tablas en la table_expression. AWS Clean Rooms solo admite INNER JOIN. En la regla de análisis de lista, se requiere al menos una INNER JOIN, y el miembro que puede realizar consultas debe incluir una tabla de su propiedad en la INNER JOIN. Esto significa que deben combinar su tabla con la suya, ya sea de forma directa o transitiva.

A continuación se muestra un ejemplo de transitividad.

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

Las instrucciones INNER JOIN solo pueden usar columnas que se hayan categorizado explícitamente como `joinColumn` en la regla de análisis.

La instrucción INNER JOIN debe operar en una `joinColumn` de su tabla configurada y en una `joinColumn` de otra tabla configurada de la colaboración. Usted decide qué columnas de la tabla se pueden usar como `joinColumn`.

Cada condición de coincidencia de la cláusula ON debe utilizar la condición de comparación de igualdad (=) entre dos columnas.

Las condiciones de coincidencia múltiples dentro de una cláusula ON pueden ser:

- Combinación con el operador lógico `AND`
- Separación mediante el operador lógico `OR`

> **Note**
>
> Todas las condiciones de coincidencia JOIN deben coincidir con una fila de cada lado de JOIN. Todos los condicionales conectados por un operador lógico `OR` o `AND` también deben cumplir este requisito.

A continuación se muestra un ejemplo de consulta con un operador lógico `AND`.

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

A continuación se muestra un ejemplo de consulta con un operador lógico `OR`.

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

Control	Definición	Uso
`joinColumns`	Las columnas que se desea permitir que el miembro que	Una misma columna no se puede categorizar como `joinColumn` y `listColum`

Control	Definición	Uso
	puede realizar consultas utilice en la instrucción INNER JOIN.	n a la vez (consulte Controles de lista). `joinColumn` no se puede usar en ninguna otra parte de la consulta que no sea INNER JOIN.

Controles de lista

Los controles de lista controlan las columnas que se pueden mostrar en el resultado de la consulta (es decir, que se pueden usar en la instrucción SELECT) o que se pueden usar para filtrar los resultados (es decir, que se pueden usar en la instrucción WHERE).

Control	Definición	Uso
`listColumns`	Las columnas que se permite que el miembro que puede realizar consultas utilice en SELECT y WHERE	Se puede usar una `listColumn` en SELECT y WHERE. Una misma columna no se puede usar como `listColumn` y `joinColumn` a la vez.

Estructura predefinida de la regla de análisis de lista

El siguiente ejemplo incluye una estructura predefinida que muestra cómo completar una regla de análisis de lista.

En el siguiente ejemplo, *`MyTable`* hace referencia a nuestra tabla de datos. Puede reemplazar cada una de ellas *`user input placeholder`* con su propia información.

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

Regla de análisis de lista: ejemplo

El siguiente ejemplo demuestra cómo dos empresas pueden colaborar en el AWS Clean Rooms uso del análisis de listas.

La empresa A dispone de datos de administración de las relaciones con los clientes (CRM). La empresa A desea obtener datos de segmento adicionales sobre sus clientes para conocer mejor a estos últimos y, posiblemente, utilizar los atributos como entrada en otros análisis. La empresa B tiene datos de segmento compuestos por atributos de segmento únicos que creó basándose en sus datos de primera parte. La empresa B desea proporcionar los atributos de segmento únicos a la empresa A solo para aquellos clientes que se superponen entre sus datos y los datos de la empresa A.

Las empresas deciden colaborar para que la empresa A pueda enriquecer los datos superpuestos. La empresa A es el miembro que puede realizar consultas y la empresa B es el colaborador.

Para crear una colaboración y ejecutar en ella un análisis de lista, las empresas hacen lo siguiente:

1. La empresa A crea una colaboración y crea una pertenencia. La colaboración tiene a la empresa B como un miembro más de la colaboración. La empresa A habilita el registro de consultas en la colaboración y habilita el registro de consultas en su cuenta.
2. La empresa B crea una pertenencia en la colaboración. Habilita el registro de consultas en su cuenta.
3. La empresa A crea una tabla configurada de CRM.
4. La empresa A agrega la regla de análisis a la tabla configurada por el cliente, como se muestra en el siguiente ejemplo.

```
{
  "joinColumns": [
    "identifier1",
    "identifier2"
  ],
  "listColumns": [
    "internalid",
    "segment1",
    "segment2",
    "customercategory"
  ]
}
```

`joinColumns`— La empresa A quiere utilizarla `hashedemail` and/or `thirdpartyid` (obtenida de un proveedor de identidad) para hacer coincidir los clientes a partir de los datos de CRM con los clientes a partir de los datos segmentados. Esto ayudará a garantizar que la empresa A relacione los datos enriquecidos con los clientes correctos. Tienen dos JoinColumns para mejorar potencialmente la tasa de coincidencia del análisis.

`listColumns`: la empresa A usa `listColumns` para obtener columnas enriquecidas junto con un `internalid` que utiliza en sus propios sistemas. Añade `segment1`, `segment2` y `customercategory` para, potencialmente, limitar el enriquecimiento a segmentos específicos mediante su uso en filtros.

5. La empresa B crea una tabla configurada de segmento.
6. La empresa B añade la regla de análisis a la tabla configurada por de segmento.

```
{
  "joinColumns": [
    "identifier2"
  ],
  "listColumns": [
    "segment3",
    "segment4"
  ]
}
```

`joinColumns`: la empresa B permite a la empresa A combinarse con `identifier2` para relacionar a los clientes de los datos de segmento con los datos de CRM. La empresa A y la empresa B trabajaron con el proveedor de identidades para obtener un `identifier2` que se ajustara a la colaboración. No añadieron otras `joinColumns` porque creían que `identifier2` proporciona la tasa de coincidencia más alta y precisa y que no se precisaban otros identificadores para las consultas.

`listColumns`: la empresa B permite a la empresa A enriquecer sus datos con los atributos `segment3` y `segment4`, que son atributos únicos que han creado, recopilado y alineado en (con el cliente A) para formar parte del enriquecimiento de datos. Quieren que la empresa A obtenga estos segmentos para solaparlos a nivel de fila, ya que se trata de una colaboración de enriquecimiento de datos.

7. La empresa A crea una asociación a la tabla de CRM en la colaboración.
8. La empresa B crea una asociación a la tabla de segmento en la colaboración.

9. La empresa A ejecuta consultas, como la siguiente, para enriquecer los datos de clientes que se solapan.

```
SELECT companyA.internalid, companyB.segment3, companyB.segment4
INNER JOIN returns companyB
 ON companyA.identifier2 = companyB.identifier2
WHERE companyA.customercategory > 'xxx'
```

10. La empresa A y la empresa B revisan los registros de consultas. La empresa B comprueba que la consulta se ajusta a lo acordado en el contrato de colaboración.

Regla de análisis personalizada en AWS Clean Rooms

En AWS Clean Rooms, una regla de análisis personalizada es un nuevo tipo de regla de análisis que permite ejecutar consultas personalizadas en la tabla configurada. Las consultas SQL personalizadas aún están restringidas a tener solo el SELECT pero puede usar más construcciones de SQL que las consultas de agregación y lista (por ejemplo, funciones de ventana, OUTER JOIN o subconsultas; consulte la Referencia de AWS Clean Rooms SQL para obtener una lista completa). CTEs Las consultas SQL personalizadas no tienen que seguir una estructura de consulta como las consultas de agregación y las consultas de lista.

La regla de análisis personalizada admite casos de uso más avanzados que los que admite una regla de análisis de agregación y de lista, como análisis de atribución personalizado, análisis comparativo, análisis de incrementalidad y detección de audiencias. Esto se suma a un superconjunto de casos de uso compatibles con la regla de análisis de agregación y de lista.

La regla de análisis personalizado también admite privacidad diferencial. La privacidad diferencial es un marco matemáticamente riguroso para la protección de la privacidad de los datos. Para obtener más información, consulte AWS Clean Rooms Privacidad diferencial. Al crear una plantilla de análisis, AWS Clean Rooms Differential Privacy comprueba la plantilla para determinar si es compatible con la estructura de consultas de uso general de Differential Privacy. AWS Clean Rooms Esta validación garantiza que no se cree una plantilla de análisis que no esté permitida con una tabla con protección de privacidad diferencial.

Para configurar la regla de análisis personalizada, los propietarios de los datos pueden optar por permitir que determinadas consultas personalizadas, almacenadas en plantillas de análisis, se ejecuten en sus tablas configuradas. Los propietarios de los datos revisan las plantillas de análisis antes de añadirlas al control de análisis permitido en la regla de análisis personalizada. Las plantillas de análisis están disponibles y visibles solo en la colaboración en la que se crean (incluso si la tabla

está asociada a otras colaboraciones) y solo las puede ejecutar el miembro que pueda realizar consultas en esa colaboración.

Como alternativa, los miembros pueden optar por permitir que otros miembros (proveedores de consultas) creen consultas sin revisión. Los miembros añaden las cuentas de proveedores de consultas que los proveedores de consultas permitidos pueden controlar en la regla de análisis personalizada. Si el proveedor de consultas es el miembro que puede realizar la consulta, este puede ejecutar cualquier consulta directamente en la tabla configurada. Los proveedores de consultas también pueden crear consultas mediante la creación de plantillas de análisis. Todas las consultas que hayan creado los proveedores de consultas pueden ejecutarse automáticamente en la tabla en todas las colaboraciones en las que Cuenta de AWS esté presente y la tabla esté asociada.

Los propietarios de los datos solo pueden permitir que las plantillas de análisis o las cuentas creen consultas, pero no ambas. Si el propietario de los datos lo deja en blanco, el miembro que puede realizar la consulta no puede ejecutar consultas en la tabla configurada.

Temas

- Estructura predefinida de la regla de análisis personalizada
- Ejemplo de regla de análisis personalizada
- Regla de análisis personalizada con privacidad diferencial

Estructura predefinida de la regla de análisis personalizada

El siguiente ejemplo incluye una estructura predefinida que muestra cómo completar una regla de análisis personalizada con la privacidad diferencia activada. El valor `userIdentifier` es la columna que identifica de forma exclusiva a los usuarios, como user_id. Cuando se activan dos o más tablas con la privacidad diferencial activada en una colaboración, AWS Clean Rooms requiere que configure la misma columna que la columna del identificador de usuario en ambas reglas de análisis para mantener una definición coherente de los usuarios en todas las tablas.

```
{
  "allowedAnalyses": ["ANY_QUERY"] | string[],
  "allowedAnalysisProviders": [],
  "differentialPrivacy": {
    "columns": [
      {
        "name": "userIdentifier"
      }
```

```
    ]
  }
}
```

Puede:

- Agregue una plantilla de análisis ARNs para permitir el control de los análisis. En este caso, el control `allowedAnalysisProviders` no está incluido.

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

- Agregue un miembro Cuenta de AWS IDs al `allowedAnalysisProviders` control. En este caso, se añade `ANY_QUERY` al control `allowedAnalyses`.

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

Ejemplo de regla de análisis personalizada

El siguiente ejemplo demuestra cómo dos empresas pueden colaborar en el AWS Clean Rooms uso de la regla de análisis personalizada.

La empresa A tiene datos de clientes y de ventas. La empresa A está interesada en conocer la incrementalidad de las ventas de una campaña publicitaria en el sitio de la empresa B. La empresa B tiene datos de visualizaciones y atributos de segmento que son útiles para la empresa (por ejemplo, el dispositivo utilizado para ver la publicidad).

La empresa A tiene una consulta de incrementalidad específica que quiere ejecutar en la colaboración.

Para crear una colaboración y ejecutar en ella un análisis personalizado, las empresas hacen lo siguiente:

1. La empresa A crea una colaboración y crea una pertenencia. La colaboración tiene a la empresa B como un miembro más de la colaboración. La empresa A habilita el registro de consultas en la colaboración y habilita el registro de consultas en su cuenta.

2. La empresa B crea una pertenencia en la colaboración. Habilita el registro de consultas en su cuenta.
3. La empresa A crea una tabla configurada de CRM.
4. La empresa A añade una regla de análisis personalizada vacía a la tabla de ventas configurada.
5. La empresa A asocia la tabla configurada de ventas a la colaboración.
6. La empresa B crea una tabla configurada de visualizaciones.
7. La empresa B agrega una regla de análisis personalizada vacía a la tabla configurada de visualizaciones.
8. La empresa B asocia la tabla configurada de visualizaciones a la colaboración.
9. La empresa A visualiza la tabla de ventas y la tabla de visualizaciones asociada a la colaboración y crea una plantilla de análisis, añadiendo la consulta de incrementalidad y el parámetro para el mes de la campaña.

```
{
    "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. La empresa A agrega la cuenta (por ejemplo, 444455556666) al control del proveedor de análisis permitido en la regla de análisis personalizada. Utilizan el control de proveedor de análisis permitido porque quieren permitir que las consultas que creen se ejecuten en su tabla configurada de ventas.

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

11. La empresa B ve la plantilla de análisis creada en la colaboración y revisa su contenido, incluidos la cadena de consulta y el parámetro.

12. La empresa B determina que la plantilla de análisis cumple con el caso de uso de incrementalidad y cumple con sus requisitos de privacidad en cuanto a la forma en que se puede consultar su tabla configurada de visualizaciones.

13. La empresa B agrega la plantilla de análisis ARN al control de análisis permitido en la regla de análisis personalizada de la tabla de visualizaciones. Utilizan el control de análisis permitido porque solo quieren permitir que la consulta de incrementalidad se ejecute en su tabla configurada de visualizaciones.

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

14. La empresa A ejecuta la plantilla de análisis y utiliza el valor del parámetro `05-01-2023`.

Regla de análisis personalizada con privacidad diferencial

En AWS Clean Rooms, la regla de análisis personalizado admite la privacidad diferencial. La privacidad diferencial es un marco matemáticamente riguroso para la protección de la privacidad de los datos que le ayuda a proteger sus datos contra los intentos de reidentificación.

La privacidad diferencial permite el análisis agregado, como la planificación de campañas publicitarias, la post-ad-campaign medición, la evaluación comparativa en un consorcio de instituciones financieras y las pruebas A/B para la investigación sanitaria.

La estructura y sintaxis de consulta admitidas se definen en Estructura y sintaxis de las consultas.

Regla de análisis personalizada con privacidad diferencial

> **Note**
>
> AWS Clean Rooms La privacidad diferencial solo está disponible para las colaboraciones que utilizan AWS Clean Rooms SQL como motor de análisis y los datos almacenados en Amazon S3.

Analice el ejemplo de regla de análisis personalizado que se presenta en la sección anterior. En este ejemplo se demuestra cómo puede utilizar la privacidad diferencial para proteger sus datos frente a intentos de reidentificación y, al mismo tiempo, permitir que su socio obtenga información crítica para la empresa a partir de sus datos. Supongamos que la empresa B, que tiene los datos de audiencia, quiere proteger sus datos con una privacidad diferencial. Para completar la configuración de privacidad diferencial, la empresa B realiza los siguientes pasos:

1. La empresa B activa la privacidad diferencial mientras añade una regla de análisis personalizada a la tabla configurada de visualizaciones. La empresa B selecciona `viewershipdata.hashedemail` como columna de identificador de usuario.
2. La empresa B añade una política de privacidad diferencial a la colaboración para que su tabla de datos de audiencia esté disponible para su consulta. La empresa B selecciona la política predeterminada para completar rápidamente la configuración.

La empresa A, que desea comprender la incrementalidad de ventas de una campaña publicitaria en el sitio de la empresa B, utiliza la plantilla de análisis. Como la consulta es compatible con la estructura de consulta de uso general de AWS Clean Rooms Differential Privacy, la consulta se ejecuta correctamente.

Estructura y sintaxis de las consultas

Las consultas que contengan al menos una tabla y que tengan activada la privacidad diferencial deben seguir la siguiente sintaxis.

```
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**
>
> En cuanto a la estructura y sintaxis de las consultas de privacidad diferencial, tenga en cuenta lo siguiente:
>
> - No se admiten las subconsultas.
> - Las expresiones de tabla comunes (CTEs) deberían emitir la columna de identificador de usuario si una tabla o un CTE incluyen datos protegidos por una privacidad diferencial. Los filtros, agrupaciones y agregaciones deben realizarse en el nivel de usuario.
> - Final_select permite las funciones de agregación COUNT DISTINCT, COUNT, SUM, AVG y STDDEV.

Para obtener más información sobre qué palabras clave de SQL se admiten para una privacidad diferencial, consulte Capacidades SQL de AWS Clean Rooms Differential Privacy.

Regla de análisis de tablas de asignación de ID

En AWS Clean Rooms, una regla de análisis de tablas de mapeo de ID no es una regla de análisis independiente. Este tipo de regla de análisis se administra mediante datos de identidad dispares AWS Clean Rooms y se utiliza para unirlos a fin de facilitar las consultas. Se agrega automáticamente a las tablas de asignación de ID y no se puede editar. Hereda los comportamientos de las demás reglas de análisis de la colaboración, siempre que esas reglas de análisis sean homogéneas.

La regla de análisis de la tabla de asignación de ID refuerza la seguridad en una tabla de asignación de ID. Impide que un miembro de la colaboración seleccione o inspeccione directamente la población no superpuesta entre los conjuntos de datos de los dos miembros mediante la tabla de asignación de ID. La regla de análisis de la tabla de asignación de ID se usa para proteger la información confidencial de la tabla de asignación de ID cuando se usa implícitamente en consultas con otras reglas de análisis.

Con la regla de análisis de la tabla de mapeo de ID, se AWS Clean Rooms impone una superposición en ambos lados de la tabla de mapeo de ID en el SQL expandido. Esto le permite realizar las siguientes tareas:

- Utilice la superposición de la tabla de mapeo de ID en JOIN declaraciones.

AWS Clean Rooms permite un INNER, LEFT, or RIGHT únase a la tabla de mapeo de ID si respeta la superposición.

- Utilice las columnas de la tabla de mapeo en JOIN declaraciones.

  No puede usar las columnas de la tabla de mapeo en las siguientes declaraciones: SELECT, WHERE, HAVING, GROUP BY, or ORDER BY (a menos que se modifiquen las protecciones en la asociación del espacio de nombres del ID de origen o en la asociación del espacio de nombres del ID de destino).
- En SQL expandido, también es compatible AWS Clean Rooms OUTER JOIN, implícito JOIN, y CROSS JOIN. Estas uniones no pueden cumplir los requisitos de superposición. En su lugar, se AWS Clean Rooms utiliza `requireOverlap` para especificar en qué columnas se deben unir.

La estructura y la sintaxis de consulta admitidas se definen en Estructura y sintaxis de consultas de la tabla de asignación de ID.

Los parámetros de la regla de análisis, definidos en Controles de consulta de reglas de análisis de tablas de asignación de ID, incluyen los controles de consulta y los controles de resultados de las consultas. Sus controles de consulta incluyen la posibilidad de requerir la superposición de la tabla de mapeo de ID en JOIN declaraciones (es decir,`requireOverlap`).

Temas

- Estructura y sintaxis de consultas de la tabla de asignación de ID
- Controles de consulta de reglas de análisis de tablas de asignación de ID
- Estructura predefinida de la regla de análisis de tablas de asignación de ID
- Regla de análisis de tablas de asignación de ID: ejemplo

Estructura y sintaxis de consultas de la tabla de asignación de ID

Las consultas en tablas que tienen una regla de análisis de tabla de asignación de ID deben respetar la siguiente sintaxis.

```
--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
```

Tablas de colaboración

Las siguientes tablas representan las tablas configuradas que existen en una AWS Clean Rooms colaboración. La columna de ID en las tablas cr_drivers_license y cr_insurance representa una columna que coincide con la tabla de asignación de ID.

cr_drivers_license

id	nombre_del_conductor	estado_de_registro
1	Eduard	TX
2	Dana	MA
3	Gweneth	IL

cr_insurance

id	policyholder_email	número_política
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

tabla de asignación de ID

La siguiente tabla representa una tabla de asignación de ID existente que coincide con las tablas cr_drivers_license y cr_insurance. No todas las entradas serán válidas para ambas tablas de colaboración. IDs

cr_drivers_license_id	cr_insurance_id
1	a
2	null
3	b
null	c

La regla de análisis de la tabla de asignación de ID solo permite que las consultas se ejecuten en el conjunto de datos superpuestos, que tendría el siguiente aspecto:

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

Consultas de ejemplo

Los siguientes ejemplos muestran ubicaciones válidas para las uniones de la tabla de asignación de ID:

```
-- Single ID mapping table
```

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

Consideraciones

Para la estructura y sintaxis de las consultas de tabla de asignación de ID, tenga en cuenta lo siguiente:

- No puede editarlo.
- Se aplica a la tabla de asignación de ID de forma predeterminada.
- Utiliza una asociación de espacio de nombres de ID de origen y destino dentro de la colaboración.
- La tabla de asignación de ID está configurada de forma predeterminada para proporcionar protecciones predeterminadas para la columna que proviene del espacio de nombres de ID. Puede modificar esta configuración para que la columna que proviene del espacio de nombres de ID (`sourceID` o `targetID`) se pueda permitir en cualquier parte de la consulta. Para obtener más información, consulte Espacios de nombres de ID en AWS Clean Rooms.
- La regla de análisis de la tabla de asignación de ID hereda las restricciones de SQL de las demás reglas de análisis de la colaboración.

Controles de consulta de reglas de análisis de tablas de asignación de ID

Con los controles de consulta de tablas de mapeo de ID, controla cómo se utilizan las columnas de la tabla para consultarla. AWS Clean Rooms Por ejemplo, controla qué columnas se utilizan para unir y qué columnas deben superponerse. La regla de análisis de tablas de asignación de ID también incluye una funcionalidad que le permite proyectar `sourceID`, `targetID` o ambas sin necesidad de JOIN.

La tabla siguiente explica cada control.

Control	Definición	Uso
`joinColumns`	Las columnas que el miembro que puede realizar consultas puede usar en la instrucción INNER JOIN.	No puede usar `joinColumns` en ninguna otra parte de la consulta que no sea INNER JOIN. Para obtener más información, consulte Controles de combinación.

Control	Definición	Uso
`dimensionColumns`	Las columnas (si las hay) que el miembro que puede realizar consultas puede utilizar en las instrucciones SELECT y GROUP BY.	A `dimensionColumn` se puede utilizar en SELECT y GROUP BY. Puede aparecer una `dimensionColumn` como `joinKeys`. Solo puede utilizar `dimensionColumns` en la cláusula JOIN si la especifica entre paréntesis.
`queryContraints:RequireOverlap`	Las columnas de la tabla de asignación de ID que se deben unir para que se pueda ejecutar la consulta.	Estas columnas se deben usar para UNIR la tabla de asignación de ID y una tabla de colaboración.

Estructura predefinida de la regla de análisis de tablas de asignación de ID

La estructura predefinida para una regla de análisis de tablas de asignación de ID incluye protecciones predeterminadas que se aplican a `sourceID` y `targetID`. Esto significa que la columna con las protecciones aplicadas se debe usar en las consultas.

Puede configurar la regla de análisis de la tabla de asignación de ID de las siguientes maneras:

- `sourceID` y `targetID` protegidos

  En esta configuración, no se pueden proyectar `sourceID` y `targetID`. `sourceID` y `targetID` se deben usar en una instrucción JOIN cuando se hace referencia a la tabla de asignación de ID.
- Solo `targetID` protegido

  En esta configuración, `targetID` no se puede proyectar. `targetID` se debe usar en una instrucción JOIN cuando se hace referencia a la tabla de asignación de ID. `sourceID` se puede usar en una consulta.
- Solo `sourceID` protegido

En esta configuración, `sourceID` no se puede proyectar. `sourceID` se debe usar en una instrucción JOIN cuando se hace referencia a la tabla de asignación de ID. `targetID` se puede usar en una consulta.

- Ningún `sourceID` o `targetID` protegido

  En esta configuración, la tabla de asignación de ID no está sujeta a ninguna aplicación específica que pueda usarse en una consulta.

El siguiente ejemplo muestra una estructura predefinida para una regla de análisis de tablas de asignación de ID con las protecciones predeterminadas que se aplican a `sourceID` y `targetID`. En este ejemplo, la regla de análisis de la tabla de asignación de ID solo permite una instrucción INNER JOIN en la columna `sourceID` y en la columna `targetID`.

```
{
  "joinColumns": [
    "source_id",
    "target_id"
  ],
  "queryConstraints": [
    {
      "requireOverlap": {
        "columns": [
          "source_id",
          "target_id"
        ]
      }
    }
  ],
  "dimensionColumns": [] // columns that can be used in SELECT and JOIN
}
```

El siguiente ejemplo muestra una estructura predefinida para una regla de análisis de tablas de asignación de ID con las protecciones que se aplican a `targetID`. En este ejemplo, la regla de análisis de la tabla de asignación de ID solo permite una instrucción INNER JOIN en la columna `sourceID`.

```
{
  "joinColumns": [
    "source_id",
```

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

El siguiente ejemplo muestra una estructura predefinida para una regla de análisis de tablas de asignación de ID con las protecciones que se aplican a `sourceID`. En este ejemplo, la regla de análisis de la tabla de asignación de ID solo permite una instrucción INNER JOIN en la columna `targetID`.

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

El siguiente ejemplo muestra una estructura predefinida para una regla de análisis de tablas de asignación de ID sin las protecciones que se aplican a `sourceID` o `targetID`. En este ejemplo, la

regla de análisis de la tabla de asignación de ID permite una instrucción INNER JOIN en la columna `sourceID` y en la columna `targetID`.

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

Regla de análisis de tablas de asignación de ID: ejemplo

En lugar de escribir una larga declaración en cascada que haga referencia a la información de identificación personal (PII), por ejemplo, las empresas pueden utilizar la regla de análisis de la tabla de mapeo de identificadores para utilizar la transcodificación LiveRamp multipartita. En el siguiente ejemplo, se muestra cómo se puede colaborar en el AWS Clean Rooms uso de la regla de análisis de tablas de mapeo de identificadores.

La empresa A es un anunciante que tiene datos de clientes y ventas, que se utilizarán como origen. La empresa A también realiza la transcodificación en nombre de las partes de la colaboración y aporta las LiveRamp credenciales.

La empresa B es un publicador que tiene datos de eventos, que se utilizarán como destino.

> **Note**
>
> La empresa A o la empresa B pueden proporcionar las credenciales de LiveRamp transcodificación y realizar la transcodificación.

Para crear una colaboración que permita el análisis de tablas de asignación de ID en colaboración, las empresas hacen lo siguiente:

1. La empresa A crea una colaboración y crea una pertenencia. Se agrega la empresa B, que también crea una suscripción en la colaboración.
2. La empresa A asocia una fuente de espacio de nombres de ID existente o crea una nueva mediante la consola. AWS Entity Resolution AWS Clean Rooms

   La empresa A crea una tabla configurada con sus datos de ventas y una columna vinculada a `sourceId` en la tabla de asignación de ID.

   El origen del espacio de nombres de ID proporciona datos para la transcodificación.
3. La empresa B asocia un destino de espacio de nombres de ID existente o crea uno nuevo mediante la consola. AWS Entity Resolution AWS Clean Rooms

   La empresa B crea una tabla configurada con sus datos de eventos y una columna vinculada a `targetId` en la tabla de asignación de ID.

   El destino del espacio de nombres de ID no proporciona datos para transcodificar, solo metadatos relacionados con la configuración. LiveRamp
4. La empresa A descubre los dos espacios de nombres de ID asociados a la colaboración y crea y completa una tabla de asignación de ID.
5. La empresa A ejecuta una consulta en los dos conjuntos de datos mediante la unión en la tabla de asignación de ID.

```
--- 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 Privacidad diferencial

> **Note**
>
> Se aplica a: motor de análisis AWS Clean Rooms SQL

AWS Clean Rooms Differential Privacy le ayuda a proteger la privacidad de sus usuarios con una técnica respaldada matemáticamente que se implementa con controles intuitivos en unos pocos clics. Al tratarse de una funcionalidad totalmente gestionada, no es necesario tener experiencia previa en materia de privacidad diferencial para evitar que los usuarios vuelvan a identificarse. AWS Clean Rooms añade automáticamente una cantidad de ruido cuidadosamente calibrada a los resultados de las consultas en tiempo de ejecución para ayudar a proteger sus datos a nivel individual.

AWS Clean Rooms Differential Privacy admite una amplia gama de consultas analíticas y es ideal para una amplia variedad de casos de uso, en los que un pequeño error en los resultados de la consulta no comprometa la utilidad del análisis. Con él, sus socios pueden generar información crítica para la empresa sobre campañas publicitarias, decisiones de inversión, investigaciones clínicas y mucho más, todo ello sin necesidad de ninguna configuración adicional por parte de sus socios.

AWS Clean Rooms Differential Privacy protege contra errores de conversión excesivos o inválidos que utilizan funciones escalares o símbolos de operadores matemáticos de forma malintencionada.

Para obtener más información sobre la privacidad AWS Clean Rooms diferencial, consulte los siguientes temas.

Temas

- Privacidad diferencial
- Cómo funciona Differential Privacy in AWS Clean Rooms
- Política de privacidad diferencial
- Capacidades SQL de AWS Clean Rooms Differential Privacy
- Consejos y ejemplos de consultas sobre privacidad diferencial
- Limitaciones de la AWS Clean Rooms privacidad diferencial

Privacidad diferencial

La privacidad diferencial solo permite obtener información agregada y oculta la contribución de los datos de cualquier persona a esa información. La privacidad diferencial protege los datos de la colaboración para que el miembro pueda obtener resultados al aprender sobre una persona específica. Sin una privacidad diferencial, el miembro que puede recibir los resultados puede intentar deducir los datos de un usuario individual añadiendo o quitando registros sobre una persona y observando la diferencia en los resultados de las consultas.

Cuando se activa la privacidad diferencial, se añade una cantidad específica de ruido a los resultados de la consulta para ocultar la contribución de los usuarios individuales. Si el miembro que puede recibir los resultados intenta observar la diferencia en los resultados de la consulta después de eliminar los registros sobre una persona de su conjunto de datos, la variabilidad del resultado de la consulta ayuda a impedir la identificación de los datos de la persona. AWS Clean Rooms Differential Privacy utiliza el SampCertmuestreador, una implementación comprobada y correcta desarrollada por. AWS

Cómo funciona Differential Privacy in AWS Clean Rooms

El flujo de trabajo para activar la privacidad diferencial AWS Clean Rooms requiere los siguientes pasos adicionales al completar el flujo de trabajo para AWS Clean Rooms:

1. La privacidad diferencial se activa al añadir una regla de análisis personalizada.
2. Puede configurar la política de privacidad diferencial de la colaboración para que las tablas de datos protegidas con privacidad diferencial estén disponibles para su consulta.

Tras completar estos pasos, el miembro que pueda realizar la consulta podrá empezar a ejecutar consultas sobre datos protegidos por la privacidad diferencial. AWS Clean Rooms devuelve resultados que cumplen con la política de privacidad diferencial. AWS Clean Rooms Differential Privacy registra el número estimado de consultas pendientes que puede ejecutar, de forma similar al indicador de combustible de un automóvil que muestra el nivel de combustible actual del automóvil. El número de consultas que puede ejecutar el miembro que puede realizar la consulta está limitado por los parámetros Presupuesto de privacidad y el Ruido añadido por consulta establecidos en Política de privacidad diferencial.

Consideraciones

Cuando utilices la privacidad diferencial en AWS Clean Rooms, ten en cuenta lo siguiente:

- El miembro que puede recibir los resultados no puede usar la privacidad diferencial. Configurará una regla de análisis personalizada con la privacidad diferencial desactivada para las tablas configuradas.
- El miembro que puede realizar consultas no puede unir tablas de dos o más proveedores de datos cuando ambos tienen activada la privacidad diferencial.

Política de privacidad diferencial

La política de privacidad diferencial controla cuántas funciones de agregación puede ejecutar el miembro que puede realizar la consulta en una colaboración. El Presupuesto de privacidad define un recurso común y limitado que se aplica a todas las tablas durante una colaboración. El Ruido agregado por consulta determina la velocidad a la que se agota el presupuesto de privacidad.

Es necesaria una política de privacidad diferencial para que las tablas con protección de privacidad diferencial estén disponibles para su consulta. Se trata de un paso único en una colaboración e incluye dos entradas:

- Presupuesto de privacidad: cuantificado en términos de épsilon, el presupuesto de privacidad controla el nivel de protección de la privacidad. Se trata de un recurso común y limitado que se aplica a todas las tablas protegidas con privacidad diferencial durante la colaboración, ya que el objetivo es conservar la privacidad de los usuarios, cuya información puede estar presente en varias tablas.

  El presupuesto de privacidad se consume cada vez que se ejecuta una consulta en las tablas. Cuando el presupuesto de privacidad se agote por completo, el miembro de la colaboración que puede realizar consultas no podrá ejecutar consultas adicionales hasta que se aumente o actualice. Al establecer un presupuesto de privacidad mayor, el miembro que puede recibir los resultados puede reducir su incertidumbre acerca de las personas que contienen los datos. Elija un presupuesto de privacidad que equilibre sus requisitos de colaboración con sus necesidades de privacidad y tras consultar con los responsables de la toma de decisiones empresariales.

  Si tiene previsto incorporar nuevos datos a la colaboración de forma periódica, puede seleccionar Actualización mensual del presupuesto de privacidad para crear automáticamente un nuevo presupuesto de privacidad cada mes natural. Al elegir esta opción, se mostrarán cantidades arbitrarias de información sobre las filas de datos cuando se consulten repetidamente durante las actualizaciones. Evite elegir esta opción si se van a consultar repetidamente las mismas filas entre las actualizaciones del presupuesto de privacidad.

- El Ruido agregado por consulta se mide en función del número de usuarios cuyas contribuciones desea ocultar. Este valor determina el ritmo al que se agota el presupuesto de privacidad. Un valor de ruido mayor reduce el ritmo al que se agota el presupuesto de privacidad y, por lo tanto, permite ejecutar más consultas sobre los datos. Sin embargo, esto debe equilibrarse con la publicación de información de datos menos precisa. Al establecer este valor, tenga en cuenta la precisión deseada para obtener información sobre la colaboración.

Puede usar la política de privacidad diferencial predeterminada para completar rápidamente la configuración o personalizar su política de privacidad diferencial según su caso de uso. AWS Clean Rooms Differential Privacy proporciona controles intuitivos para configurar la política. AWS Clean Rooms Differential Privacy le permite obtener una vista previa de la utilidad en términos del número de agregaciones posibles en todas las consultas de sus datos y estimar el número de consultas que se pueden ejecutar en una colaboración de datos.

Puede utilizar los ejemplos interactivos para comprender cómo los distintos valores de Presupuesto de privacidad y Ruido añadido por consulta afectarían a los resultados de los distintos tipos de consultas SQL. En general, es necesario equilibrar sus necesidades de privacidad con el número de consultas que desea permitir y la precisión de esas consultas. Un menor Presupuesto de privacidad o un mayor Ruido añadido por consulta puede proteger mejor la privacidad de los usuarios, pero proporciona información menos significativa a sus socios de colaboración.

Si aumenta el Presupuesto de privacidad y mantiene igual el parámetro Ruido añadido por consulta, el miembro que pueda realizar la consulta podrá ejecutar más agregaciones en sus tablas durante la colaboración. Puede aumentar el Presupuesto de privacidad en cualquier momento durante la colaboración. Si reduce el Presupuesto de privacidad y mantiene igual el parámetro Ruido añadido por consulta, el miembro que pueda realizar la consulta podrá ejecutar menos agregaciones. No puede reducir el Presupuesto de privacidad una vez que el miembro que puede realizar la consulta haya empezado a analizar sus datos.

Si aumenta el Ruido añadido por consulta y mantiene igual la entrada de Presupuesto de privacidad, el miembro que pueda realizar la consulta podrá ejecutar más agregaciones en sus tablas durante la colaboración. Si reduce el Ruido añadido por consulta y mantiene igual la entrada en de Presupuesto de privacidad, el miembro que pueda realizar la consulta podrá ejecutar menos agregaciones. Puede aumentar o disminuir el Ruido agregado por consulta en cualquier momento durante la colaboración.

La política de privacidad diferencial se gestiona mediante las acciones de la API de la plantilla de presupuesto de privacidad.

Capacidades SQL de AWS Clean Rooms Differential Privacy

AWS Clean Rooms Differential Privacy utiliza una estructura de consultas de uso general para admitir consultas SQL complejas. Las plantillas de análisis personalizadas se validan según esta estructura para garantizar que se puedan ejecutar en tablas protegidas por privacidad diferencial. En la siguiente tabla se indican las funciones compatibles. Para obtener más información, consulte Estructura y sintaxis de las consultas.

Nombre corto	Constructos SQL	Expresiones de tabla comunes () CTEs	Una cláusula SELECT final
Funciones de agregación	• Función ANY_VALUE • Función APPROXIMATE PERCENTILE_DISC • Función AVG • Funciones COUNT y COUNT DISTINCT • Función LISTAGG • Función MAX • Función MEDIAN • Función MIN • Función PERCENTILE_CONT • Funciones STDDEV_SAMP y STDDEV_POP • Funciones SUM y SUM DISTINCT	Se admiten con la condición de que el CTEs uso de tablas con protección de privacidad diferencial debe generar datos con registros a nivel de usuario. Debe escribir la expresión SELECT en las que CTEs utilicen el `SELECT userIdentifierColumn...' formato.	Agregaciones compatibles: AVG, COUNT, COUNT DISTINCT, STDDEV y SUM.

Nombre corto	Constructos SQL	Expresiones de tabla comunes () CTEs	Una cláusula SELECT final
	• Funciones VAR_SAMP y VAR_POP		
CTEs	Cláusula WITH, subconsulta de cláusula WITH	Se admite con la condición de que el CTEs uso de tablas con protección de privacidad diferencial debe generar datos con registros a nivel de usuario. Debe escribir la expresión SELECT en las que CTEs utilicen el `SELECT userIdentifierColumn...' formato.	N/A
subconsultas	• SELECT • HAVING • JOIN • Condición de JOIN • FROM • WHERE	Puede tener cualquier subconsulta que no haga referencia a las relaciones de privacidad diferenciales en estas construcciones. Puede tener cualquier subconsulta que haga referenci a a relaciones de privacidad diferenciales únicamente en las cláusulas FROM y JOIN.	

Nombre corto	Constructos SQL	Expresiones de tabla comunes () CTEs	Una cláusula SELECT final
Unir cláusulas	• INNER JOIN • LEFT JOIN • RIGHT JOIN • FULL JOIN • Operador [JOIN] OR • CROSS JOIN	Se admite con la condición de que solo se admitan las funciones JOIN que son equi-joins en las columnas del identificador de usuario y son obligatorias cuando se consultan dos o más tablas con la privacidad diferenci al activada. Asegúrese de que las condicion es obligatorias de equi-join sean correctas. Confirme que el propietario de la tabla haya configurado la misma columna de identific ador de usuario en todas las tablas para que la definición de usuario siga siendo coherente en todas las tablas. Las funciones CROSS JOIN no se admiten cuando se combinan dos o más relaciones con la privacidad diferencial activada.	
Operadores de establecimiento	UNION, UNION ALL, INTERSECT , EXCEPT | MINUS (son sinónimos)	Todos son compatibl es	No compatible

Nombre corto	Constructos SQL	Expresiones de tabla comunes () CTEs	Una cláusula SELECT final
Funciones de ventana	Funciones de agregación • Función de ventana AVG • Función de ventana COUNT • Función de ventana CUME_DIST • Función de ventana DENSE_RANK • Función de ventana FIRST_VALUE • Función de ventana LAG • Función de ventana LAST_VALUE • Función de ventana LEAD • Funciones de ventana MAX • Funciones de ventana MEDIAN • Funciones de ventana MIN • Función de ventana DENSE_NTH • Función de ventana RATIO_TO_REPORT	Todos son compatibles con la condición de que la columna de identificación de usuario de la cláusula de partición de la función de ventana sea obligatoria cuando se consulte una relación con la privacidad diferencial activada.	No compatible

Nombre corto	Constructos SQL	Expresiones de tabla comunes () CTEs	Una cláusula SELECT final
	• Las funciones de ventana STDDEV_SAMP y STDDEV_POP (STDDEV_SAMP y STDDEV son sinónimos) • Funciones de ventana SUM • Las funciones de ventana VAR_SAMP y VAR_POP (VAR_SAMP y VARIANCE son sinónimos) Funciones de clasificación • Función de ventana DENSE_RANK • Función de ventana NTILE • Función de ventana PERCENT_RANK • Función de ventana RANK • Función de ventana ROW_NUMBER		

Nombre corto	Constructos SQL	Expresiones de tabla comunes () CTEs	Una cláusula SELECT final
Expresiones condicionales	• Expresión condicional CASE • Expresión COALESCE • Funciones GREATEST y LEAST • Funciones NVL y COALESCE • NVL2 función • Función NULLIF	Todos son compatibles	Todos son compatibles
Condiciones	• Condición de comparación • Condiciones lógicas • Condiciones de coincidencia de patrones • Condiciones de rango BETWEEN • Condición nula	`EXISTS`y `IN` no se pueden usar porque requieren subconsultas. Se admiten todas las demás.	Todos son compatibles

Nombre corto	Constructos SQL	Expresiones de tabla comunes () CTEs	Una cláusula SELECT final
Funciones de fecha y hora	• Funciones de fecha y hora en transacciones • Operador de concatenación • Funciones ADD_MONTHS • Función CONVERT_TIMEZONE • Función CURRENT_DATE • Función DATEADD • Función DATEDIFF • Funciones DATE_PART • Función DATE_TRUNC • Función EXTRACT • Función GETDATE • Funciones TIMEOFDAY • Función TO_TIMESTAMP • Partes de fecha para funciones de fecha o marca temporal	Todas son compatibles	Todos son compatibles

Nombre corto	Constructos SQL	Expresiones de tabla comunes () CTEs	Una cláusula SELECT final
Funciones de cadena	• || Operador (concatenación) • Función BTRIM • Función CHAR_LENGTH • Función CHARACTER _LENGTH • Función CHARINDEX • Función CONCAT • Funciones LEFT y RIGHT • Función LEN • Función LENGTH • Función LOWER • Funciones LPAD y RPAD • Función LTRIM • Funciones POSITION • Función REGEXP_COUNT • Función REGEXP_INSTR • Función REGEXP_RE PLACE • Función REGEXP_SUBSTR	Todos son compatibl es	Todos son compatibl es

Nombre corto	Constructos SQL	Expresiones de tabla comunes () CTEs	Una cláusula SELECT final
	• Función REPEAT • Función REPLACE • Función REPLICATE • Función REVERSE • Función RTRIM • Función SOUNDEX • Función SPLIT_PART • Función STRPOS • Función SUBSTRING • Función TEXTLEN • Función TRANSLATE • Funciones TRIM • Función UPPER		
Funciones de formato de tipo de datos	• Función CAST • TO_CHAR • Función TO_DATE • TO_NUMBER • Cadenas de formatos de fecha y hora • Cadenas de formatos numéricos	Todos son compatibles	Todos son compatibles

Nombre corto	Constructos SQL	Expresiones de tabla comunes () CTEs	Una cláusula SELECT final
Funciones hash	• MD5 función • Función SHA • SHA1 función • SHA2 función • MURMUR3_32_HASH	Todos son compatibles	Todos son compatibles
Símbolos de operadores matemáticos	+, -, *,/,% y @	Todos son compatibles	Todos son compatibles

Nombre corto	Constructos SQL	Expresiones de tabla comunes () CTEs	Una cláusula SELECT final
Funciones matemáticas	• Función ABS • Función ACOS • Función ASIN • Función ATAN • ATAN2 función • Función CBRT • Función CEILING (o CEIL) • Función COS • Función COT • Función DEGREES • Función DEXP • Función LTRIM • DLOG1 función • DLOG1función 0 • Función EXP • Función FLOOR • Función LN • Función LOG • Función MOD • Función PI • Función POWER • Función RADIANS • Función RANDOM • Función ROUND • Función SIGN • Función SIN • Funciones SQRT	Todos son compatibles	Todos son compatibles

Nombre corto	Constructos SQL	Expresiones de tabla comunes () CTEs	Una cláusula SELECT final
	• Función TRUNC		
Funciones de información acerca del tipo SUPER	• Función DECIMAL_PRECISION • Función DECIMAL_SCALE • Función IS_ARRAY • Función IS_BIGINT • Función IS_CHAR • Función IS_DECIMAL • Función IS_FLOAT • Función IS_INTEGER • Función IS_OBJECT • Función IS_SCALAR • Función IS_SMALLINT • Función IS_VARCHAR • Función JSON_TYPEOF	Todos son compatibles	Todos son compatibles

Nombre corto	Constructos SQL	Expresiones de tabla comunes () CTEs	Una cláusula SELECT final
Funciones VARBYTE	• Función FROM_HEX • Función FROM_VARBYTE • Función TO_HEX • Función TO_VARBYTE	Todos son compatibles	Todos son compatibles
JSON	• Función CAN_JSON_PARSE • Función JSON_EXTRACT_ARRAY_ELEMENT_TEXT • Función JSON_EXTRACT_PATH_TEXT • Función JSON_PARSE • Función JSON_SERIALIZE • Funciones JSON_SERALIZE_TO_VARBYTE	Todos son compatibles	Todos son compatibles

Nombre corto	Constructos SQL	Expresiones de tabla comunes () CTEs	Una cláusula SELECT final
Funciones de matriz	• Función array • función array_concat • Función array_flatten • Función get_array_length • Función split_to_array • función de submatriz	No admitido	No admitido
GRUPO AMPLIADO POR	AGRUPAR CONJUNTOS, ENROLLAR, CUBOS	No admitido	No admitido
Operación de clasificación	ORDER BY	Se admite con la condición de que la cláusula ORDER BY solo se admita en la cláusula de partición de una función de ventana cuando se consultan tablas con la privacidad diferencial activada.	Compatible
Límites de filas	LIMIT, OFFSET	No se admite el CTEs uso de tablas con protección de privacidad diferencial	Todas son compatibles

Nombre corto	Constructos SQL	Expresiones de tabla comunes () CTEs	Una cláusula SELECT final
Alias de tablas y columnas		Soportado	Compatible
Funciones matemáticas en funciones agregadas		Soportado	Compatible
Funciones escalares dentro de funciones agregadas		Soportado	Compatible

Alternativas comunes para constructos SQL no admitidos

Categoría	Constructo SQL	Alternativa
Funciones de ventana	• LISTAGG • PERCENTILE_CONT • PERCENTILE_DISC	Puede utilizar la función de agregado equivalente con GROUP BY.
Símbolos de operadores matemáticos	• $column ||/ 2 • $column |/ 2 • $column ^ 2	• CBRT • SQRT • POWER($column, 2)
Funciones escalares	• SYSDATE • $column::integer • convert(type, $column)	• CURRENT_DATE • CAST $column AS integer • CAST $column AS type
Literales	INTERVALO «1 SEGUNDO»	INTERVALO «1» SEGUNDO
Límite de filas	TOP n	LÍMITE n
Join	• USING • NATURAL	La cláusula ON debe contener explícitamente un criterio de unión.

Consejos y ejemplos de consultas sobre privacidad diferencial

AWS Clean Rooms Differential Privacy utiliza una estructura de consulta de uso general para admitir una amplia variedad de construcciones de SQL, como las expresiones de tabla comunes (CTEs) para la preparación de datos y las funciones agregadas de uso común, como, o. `COUNT` `SUM` Con el fin de ocultar la contribución de cualquier posible usuario a los datos y hacer más ruido al agregado de los resultados de las consultas en tiempo de ejecución, AWS Clean Rooms Differential Privacy exige que, en última `SELECT statement` instancia, las funciones de agregación se ejecuten en datos de nivel de usuario.

En el siguiente ejemplo se utilizan dos tablas denominadas `socialco_impressions` y `socialco_users` de un publicador de medios que quiere proteger los datos mediante una privacidad diferencial y, al mismo tiempo, colaborar con una marca deportiva que utiliza datos `athletic_brand_sales`. El publicador multimedia ha configurado la columna `user_id` como columna de identificador de usuario y, al mismo tiempo, ha habilitado una privacidad diferencial en AWS Clean Rooms. El anunciante no necesita una protección de privacidad diferencial y quiere ejecutar una consulta utilizando datos combinados. CTEs Puesto que su CTE utiliza tablas protegidas por la privacidad diferencial, el anunciante incluye la columna de identificador de usuario de esas tablas protegidas en la lista de columnas de CTE y une las tablas protegidas en la columna de identificador de usuario.

```
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
```

Del mismo modo, si desea ejecutar funciones de ventana en tablas de datos con protección de privacidad diferencial, debe incluir la columna de identificador de usuario en la cláusula `PARTITION BY`.

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

Limitaciones de la AWS Clean Rooms privacidad diferencial

AWS Clean Rooms La privacidad diferencial no aborda las siguientes situaciones:

1. AWS Clean Rooms Differential Privacy solo es compatible con tablas respaldadas por Amazon S3. AWS Glue No admite consultas con tablas de Snowflake o Amazon Athena.
2. AWS Clean Rooms Differential Privacy no aborda los ataques cronometrados. Por ejemplo, estos ataques son posibles en situaciones en las que un usuario individual contribuye con un gran número de filas y la adición o eliminación de este usuario cambia considerablemente el tiempo de cálculo de la consulta.
3. La privacidad diferencial de AWS Clean Rooms no garantiza la privacidad diferencial cuando una consulta SQL puede provocar desbordamientos o errores de conversión no válidos en tiempo de ejecución debido al uso de determinadas construcciones de SQL. En la siguiente tabla se muestran algunos constructos de SQL, pero no todos, que pueden producir errores en tiempo de ejecución y que se deben verificar en las plantillas de análisis. Le recomendamos que apruebe plantillas de análisis que minimicen las posibilidades de que se produzcan dichos errores en tiempo de ejecución y que revise periódicamente los registros de consultas para determinar si las consultas se ajustan al acuerdo de colaboración.

   Los siguientes constructos de SQL son vulnerables a los errores de desbordamiento:

   - Funciones de agregación: AVG, LISTAVG, PERCENTILE_COUNT, PERCENTILE_DISC, SUM/ SUM_DISTINCT
   - Funciones de formato de tipo de datos: TO_TIMESTAMP, TO_DATE
   - Funciones de fecha y hora: ADD_MONTHS, DATEADD, DATEDIFF

- Funciones matemáticas: +, -, *, /, POWER
- Funciones de cadena: ||, CONCAT, REPEAT, REPLICATE
- Funciones de ventana: AVG, LISTAGG, PERCENTILE_COUNT, PERCENTILE_DISC, RATIO_TO_REPORT, SUM

La función de formateo de tipos de datos CAST es vulnerable a errores de conversión no válidos.

Puede configurarlo CloudWatch para crear un filtro de métricas para un grupo de registros y, a continuación, crear una CloudWatch alarma en ese filtro de métricas para recibir alertas en caso de que se produzca un posible desbordamiento o un error de conversión. En concreto, debe supervisar los códigos de error `CastError`, `OverflowError` y `ConversionError`. La presencia de estos códigos de error indica un posible ataque de canal lateral, pero podría indicar una consulta SQL errónea.

Para obtener más información, consulte Inicio de sesión de Analysis AWS Clean Rooms.

AWS Clean Rooms ML

AWS Clean Rooms El aprendizaje automático permite a dos o más partes ejecutar modelos de aprendizaje automático en sus datos sin necesidad de compartirlos entre sí. El servicio proporciona controles que mejoran la privacidad y permiten a los propietarios de los datos proteger sus datos y su IP modelo. Puede usar modelos AWS creados o traer su propio modelo personalizado.

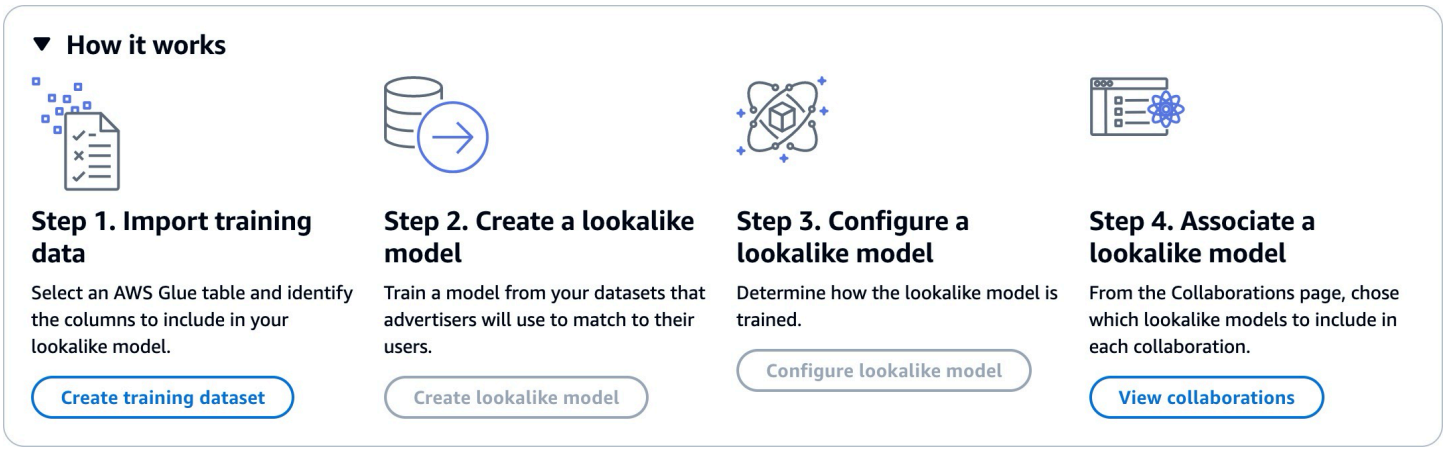
Para obtener una explicación más detallada de cómo funciona esto, consulte Trabajos entre cuentas

Para obtener más información sobre las capacidades de los modelos Clean Rooms ML, consulte los siguientes temas.

Temas

- Cómo funciona el AWS Clean Rooms aprendizaje automático con AWS los modelos
- Cómo funciona el aprendizaje AWS Clean Rooms automático con modelos personalizados
- AWS modelos en Clean Rooms ML
- Modelos personalizados en Clean Rooms ML

Cómo funciona el AWS Clean Rooms aprendizaje automático con AWS los modelos



Trabajar con modelos similares requiere que dos partes, un proveedor de datos de formación y un proveedor de datos iniciales, trabajen secuencialmente AWS Clean Rooms para integrar sus datos en una colaboración. Este es el flujo de trabajo que el proveedor de datos de entrenamiento debe completar primero:

1. Los datos del proveedor de datos de formación deben almacenarse en una tabla de catálogo de AWS Glue datos con las interacciones entre usuarios y elementos. Como mínimo, los datos de entrenamiento deben contener una columna de ID de usuario, una columna de ID de interacción y una columna de marca temporal.
2. El proveedor de datos de entrenamiento registra los datos de entrenamiento con AWS Clean Rooms.
3. El proveedor de datos de entrenamiento crea un modelo similar que se puede compartir con varios proveedores de datos iniciales. El modelo similar es una red neuronal profunda que puede tardar hasta 24 horas en entrenarse. No se reentrena automáticamente y le recomendamos que vuelva a entrenar el modelo una vez por semana.
4. El proveedor de datos de entrenamiento configura el modelo similar, que incluye si desea compartir las métricas de relevancia y la ubicación en Amazon S3 de los segmentos de salida. El proveedor de datos de entrenamiento puede crear varios modelos similares configurados a partir de un único modelo similar.
5. El proveedor de datos de entrenamiento asocia el modelo de audiencia configurado a una colaboración que se comparte con un proveedor de datos iniciales.

Este es el flujo de trabajo que el proveedor de datos iniciales debe completar a continuación:

1. Los datos del proveedor de datos iniciales se pueden almacenar en un bucket de Amazon S3 o pueden provenir de los resultados de una consulta.
2. El proveedor de datos iniciales comienza la colaboración que comparte con el proveedor de datos de entrenamiento.
3. El proveedor de datos iniciales crea un segmento similar en la pestaña de Clean Rooms ML de la página de colaboración.
4. El proveedor de datos iniciales puede evaluar las métricas de relevancia, en caso de que se hayan compartido, y exportar el segmento similar para su uso fuera de AWS Clean Rooms.

Cómo funciona el aprendizaje AWS Clean Rooms automático con modelos personalizados

Con Clean Rooms ML, los miembros de una colaboración pueden utilizar un algoritmo de modelo personalizado dockerizado que se almacena en Amazon ECR para analizar sus datos de forma conjunta. Para ello, el proveedor del modelo debe crear una imagen y almacenarla en Amazon ECR. Siga los pasos de la Guía del usuario de Amazon Elastic Container Registry para crear un repositorio privado que contendrá el modelo de aprendizaje automático personalizado.

Cualquier miembro de una colaboración puede ser el proveedor del modelo, siempre que cuente con los permisos correctos. Todos los miembros de una colaboración pueden aportar datos de entrenamiento, datos de inferencia o ambos al modelo. A los efectos de esta guía, los miembros que aportan datos se denominan proveedores de datos. El miembro que crea la colaboración es el creador de la colaboración y puede ser el proveedor del modelo, uno de los proveedores de datos o ambos.

En el nivel más alto, estos son los pasos que se deben completar para realizar un modelado de aprendizaje automático personalizado:

1. El creador de la colaboración crea una colaboración y asigna a cada miembro las capacidades y la configuración de pago adecuadas. En este paso, el creador de la colaboración debe asignar al miembro la capacidad de recibir los resultados del modelo o los resultados de las inferencias al miembro correspondiente, ya que no se puede actualizar una vez creada la colaboración. Para obtener más información, consulte Crear la colaboración.
2. El proveedor del modelo configura y asocia su modelo de aprendizaje automático en contenedores a la colaboración y se asegura de que se establezcan restricciones de privacidad para los datos exportados. Para obtener más información, consulte Configuración de un algoritmo modelo.

3. Los proveedores de datos aportan sus datos a la colaboración y se aseguran de que se especifiquen sus necesidades de privacidad. Los proveedores de datos deben permitir que el modelo acceda a sus datos. Para obtener más información, consulte Contribuir con datos de entrenamiento y Asociar el algoritmo del modelo configurado.
4. Un miembro de la colaboración crea la configuración de aprendizaje automático, que define a dónde se exportan los artefactos del modelo o los resultados de las inferencias.
5. Un miembro de la colaboración crea un canal de entrada de aprendizaje automático que proporciona información al contenedor de formación o al contenedor de inferencias. El canal de entrada de ML es una consulta que define los datos que se utilizarán en el contexto del algoritmo modelo.
6. Un miembro de la colaboración invoca el entrenamiento del modelo mediante el canal de entrada ML y el algoritmo del modelo configurado. Para obtener más información, consulte Crear un modelo entrenado.
7. (Opcional) El entrenador de modelos invoca el trabajo de exportación del modelo y los artefactos del modelo se envían al receptor de resultados del modelo. Solo los miembros con una configuración de aprendizaje automático válida y con la capacidad de recibir la salida del modelo pueden recibir artefactos del modelo. Para obtener más información, consulte Exportación de artefactos del modelo.
8. (Opcional) Un miembro de la colaboración invoca la inferencia del modelo mediante el canal de entrada ML, el ARN del modelo entrenado y el algoritmo del modelo configurado por inferencia. Los resultados de la inferencia se envían al receptor de salida de la inferencia. Solo los miembros con una configuración de aprendizaje automático válida y con la capacidad de recibir los resultados de la inferencia pueden recibir los resultados de la inferencia.

Estos son los pasos que debe completar el proveedor del modelo:

1. Cree una imagen de docker de Amazon ECR compatible con SageMaker IA. Clean Rooms ML solo admite imágenes de docker compatibles con SageMaker IA.
2. Una vez que haya creado una imagen de docker compatible con SageMaker IA, envíe la imagen a Amazon ECR. Siga las instrucciones de la Guía del usuario de Amazon Elastic Container Registry para crear una imagen de formación sobre contenedores.
3. Configure el algoritmo del modelo para usarlo en Clean Rooms ML.
   a. Proporcione el enlace al repositorio de Amazon ECR y todos los argumentos necesarios para configurar el algoritmo del modelo.

b. Proporcione un rol de acceso al servicio que permita a Clean Rooms ML acceder al repositorio de Amazon ECR.

c. Asocie el algoritmo del modelo configurado a la colaboración. Esto incluye proporcionar una política de privacidad que defina los controles para los registros de contenedores, los registros de errores, CloudWatch las métricas y los límites sobre la cantidad de datos que se pueden exportar de los resultados del contenedor.

Estos son los pasos que debe completar el proveedor de datos para colaborar con un modelo de aprendizaje automático personalizado:

1. Configure una AWS Glue tabla existente con una regla de análisis personalizada. Esto permite que un conjunto específico de consultas preaprobadas o cuentas preaprobadas utilice sus datos.
2. Asocie la tabla configurada a una colaboración y proporcione un rol de acceso al servicio que pueda acceder a sus AWS Glue tablas.
3. Agregue una regla de análisis de colaboración a la tabla que permita a la asociación de algoritmos del modelo configurado acceder a la tabla configurada.
4. Una vez que el modelo y los datos están asociados y configurados en Clean Rooms ML, el miembro con capacidad para ejecutar consultas proporciona una consulta SQL y selecciona el algoritmo del modelo que va a utilizar.

Una vez finalizado el entrenamiento del modelo, ese miembro inicia la exportación de los artefactos del entrenamiento del modelo o de los resultados de las inferencias. Estos artefactos o resultados se envían al miembro para que pueda recibir los resultados del modelo entrenado. El receptor de los resultados debe configurarlos `MachineLearningConfiguration` antes de poder recibir la salida del modelo.

AWS modelos en Clean Rooms ML

AWS Clean Rooms El aprendizaje automático proporciona un método que preserva la privacidad para que dos partes identifiquen a usuarios similares en sus datos sin necesidad de compartir sus datos entre sí. La primera parte aporta los datos de entrenamiento para AWS Clean Rooms poder crear y configurar un modelo similar y asociarlo a una colaboración. A continuación, los datos iniciales se traen a la colaboración para crear un segmento similar que se parece a los datos de entrenamiento.

Para obtener una explicación más detallada de cómo funciona esto, consulte Trabajos entre cuentas

Los siguientes temas proporcionan información sobre cómo crear y configurar un AWS modelo en Clean Rooms ML.

Temas

- AWS Clean Rooms Terminología de ML
- Protecciones de privacidad de ML AWS Clean Rooms
- Requisitos de datos de entrenamiento para Clean Rooms ML
- Requisitos de datos iniciales para Clean Rooms ML
- AWS Clean Rooms Métricas de evaluación del modelo ML

AWS Clean Rooms Terminología de ML

Es importante entender la siguiente terminología al utilizar Clean Rooms ML:

- Proveedor de datos de entrenamiento: la parte que aporta los datos de entrenamiento, crea y configura un modelo similar y, a continuación, lo asocia a una colaboración.
- Proveedor de datos iniciales: la parte que aporta los datos iniciales, genera un segmento similar y exporta su segmento similar.
- Datos de entrenamiento: los datos del proveedor de datos de entrenamiento, que se utilizan para generar un modelo similar. Los datos de entrenamiento se utilizan para medir la similitud en los comportamientos de los usuarios.

  Los datos de entrenamiento deben contener un ID de usuario, un ID de elemento y una columna de marca de tiempo. De forma opcional, los datos de entrenamiento pueden contener otras interacciones, como características numéricas o categóricas. Algunos ejemplos de interacciones son una lista de los vídeos visualizados, los artículos comprados o los artículos leídos.
- Datos iniciales: los datos del proveedor de datos iniciales, que se utilizan para crear un segmento similar. Los datos iniciales se pueden proporcionar directamente o pueden provenir de los resultados de una AWS Clean Rooms consulta. El resultado del segmento similar es un conjunto de usuarios a partir de los datos de entrenamiento que se parece más a los usuarios iniciales.
- Modelo similar: un modelo de machine learning de los datos de entrenamiento que se utiliza para encontrar usuarios similares en otros conjuntos de datos.

  Cuando se utiliza la API, el término modelo de audiencia se utiliza de manera equivalente a modelo similar. Por ejemplo, se utiliza la CreateAudienceModelAPI para crear un modelo similar.

- Segmento similar: es un subconjunto de los datos de entrenamiento que se parecen más a los datos iniciales.

  Al utilizar la API, se crea un segmento similar con la StartAudienceGenerationJobAPI.

Los datos del proveedor de datos de entrenamiento nunca se comparten con el proveedor de datos iniciales, y viceversa. La salida del segmento similar se comparte con el proveedor de datos de entrenamiento, pero nunca con el proveedor de datos iniciales.

Protecciones de privacidad de ML AWS Clean Rooms

Clean Rooms ML está diseñado para reducir el riesgo de ataques de inferencia de suscripción, ya que el proveedor de datos de entrenamiento puede saber quién está en los datos iniciales y el proveedor de datos iniciales puede saber quién está en los datos de entrenamiento. Para prevenir este ataque, se toman varias medidas.

En primer lugar, los proveedores de datos iniciales no observan directamente los resultados de Clean Rooms ML y los proveedores de datos de entrenamiento nunca pueden observar los datos iniciales. Los proveedores de datos iniciales pueden optar por incluir los datos iniciales en el segmento de salida.

A continuación, se crea el modelo similar a partir de una muestra aleatoria de los datos de entrenamiento. Este ejemplo incluye un número significativo de usuarios que no coinciden con la audiencia inicial. Este proceso hace que sea más difícil determinar si un usuario no aparecía en los datos, lo cual es otra forma de deducir la membresía.

Además, se pueden utilizar varios clientes iniciales para cada parámetro del entrenamiento de modelos similares específicos del inicio. Esto limita cuánto puede sobreajustarse el modelo y, por lo tanto, cuánto se puede deducir sobre un usuario. Como resultado, se recomienda que el tamaño mínimo de los datos iniciales sea de 500 usuarios.

Por último, nunca se proporcionan métricas en el nivel de usuario a los proveedores de datos de entrenamiento, lo que elimina otra posibilidad de que se produzca un ataque de inferencia de pertenencia.

Requisitos de datos de entrenamiento para Clean Rooms ML

Para crear correctamente un modelo similar, los datos de entrenamiento deben cumplir los siguientes requisitos:

- Los datos de entrenamiento deben estar en formato Parquet, CSV o JSON.

> **Note**
>
> No se admiten los datos de Parquet comprimidos con Zstandard (ZSTD).

- Tus datos de entrenamiento deben estar catalogados en. AWS Glue Para obtener más información, consulte Introducción al catálogo de datos de AWS Glue en la Guía para AWS Glue desarrolladores. Recomendamos utilizar AWS Glue rastreadores para crear las tablas, ya que el esquema se deduce automáticamente.
- El depósito de Amazon S3 que contiene los datos de entrenamiento y los datos iniciales se encuentra en la misma AWS región que sus otros recursos de aprendizaje automático para salas limpias.
- Los datos de entrenamiento deben contener al menos 100 000 usuarios únicos IDs con al menos dos interacciones entre elementos cada uno.
- Los datos de entrenamiento deben contener al menos 1 millón de registros.
- El esquema especificado en la CreateTrainingDatasetacción debe alinearse con el esquema definido cuando se creó la AWS Glue tabla.
- Los campos obligatorios, tal como se definen en la tabla proporcionada, se definen en la acción CreateTrainingDataset.

Tipo de campo	Tipos de datos compatibles	Obligatorio	Descripción
USER_ID	string, int, bigint	Sí	Un identificador único para cada usuario del conjunto de datos. No debe tener un valor de información de identificación

Tipo de campo	Tipos de datos compatibles	Obligatorio	Descripción
			personal (PII). Es posible que sea un identificador con hash o un ID de cliente.
ITEM_ID	string, int, bigint	Sí	Un identificador único para cada elemento con el que interactúa un usuario.
TIMESTAMP	bigint, int, timestamp	Sí	La hora en que un usuario interactuó con el elemento. Los valores deben estar en formato de tiempo Unix en formato de segundos.

Tipo de campo	Tipos de datos compatibles	Obligatorio	Descripción
CATEGORICAL_FEATURE	string, int, float, bigint, double, boolean, array	No	Captura datos categóricos relacionados con el usuario o el elemento. Esto puede incluir cosas como un tipo de evento (como hacer clic o comprar), los datos demográficos de los usuarios (grupo de edad, sexo: anonimizado), la ubicación del usuario (ciudad, país, anonimizado), la categoría del artículo (como ropa

Tipo de campo	Tipos de datos compatibles	Obligatorio	Descripción
			o artículos electrónicos) o la marca del artículo.

Tipo de campo	Tipos de datos compatibles	Obligatorio	Descripción
NUMERICAL_FEATURE	double, float, int, bigint	No	Captura datos numéricos relacionados con el usuario o el elemento. Esto puede incluir cosas como el historial de compras del usuario (importe total gastado), el precio del artículo, el número de veces que se visita un artículo o las valoraciones de los usuarios sobre los artículos.

- Opcionalmente, puede proporcionar hasta 10 características categóricas o numéricas en total.

Este es un ejemplo de un conjunto de datos de entrenamiento válido en formato CSV

```
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
```

Requisitos de datos iniciales para Clean Rooms ML

Los datos iniciales de un modelo similar pueden provenir directamente de un bucket de Amazon S3 o de los resultados de una consulta SQL.

Los datos iniciales que se proporcionan directamente deben cumplir los siguientes requisitos:

- Los datos iniciales deben estar en formato de líneas JSON con una lista de usuarios IDs.
- El tamaño de la semilla debe estar entre 25 y 500 000 usuarios únicos IDs.
- El número mínimo de usuarios iniciales debe coincidir con el valor mínimo de tamaño inicial coincidente que se especificó al crear el modelo de audiencia configurado.

El siguiente es un ejemplo de un conjunto de datos de entrenamiento válido en formato CSV

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

AWS Clean Rooms Métricas de evaluación del modelo ML

Clean Rooms ML calcula la exhaustividad y la puntuación de relevancia para determinar el rendimiento del modelo. La exhaustividad compara la similitud entre los datos similares y los datos de entrenamiento. La puntuación de relevancia se utiliza para decidir cómo de grande debe ser la audiencia, no si el rendimiento del modelo es bueno.

La exhaustividad es una medida imparcial de la similitud del segmento similar con los datos de entrenamiento. La exhaustividad es el porcentaje de usuarios más similares (de forma

predeterminada, el 20 % más similar) de una muestra de los datos de entrenamiento que se incluyen en la audiencia inicial según el trabajo de generación de audiencia. Los valores oscilan entre 0 y 1, mientras que los valores más altos indican una mejor audiencia. Un valor de exhaustividad aproximadamente igual al porcentaje máximo de intervalo indica que el modelo de audiencia equivale a una selección aleatoria.

Consideramos que esta métrica de evaluación es mejor que la exactitud, la precisión y las puntuaciones de F1 porque Clean Rooms ML no ha etiquetado con precisión a los usuarios que realmente son negativos al crear su modelo.

La puntuación de relevancia del segmento es una medida de similitud con valores que van desde -1 (menos similar) a 1 (más similar). Clean Rooms ML calcula un conjunto de puntuaciones de relevancia para varios tamaños de segmento a fin de ayudarle a determinar el mejor tamaño de segmento para los datos. Las puntuaciones de relevancia disminuyen de forma monótona a medida que aumenta el tamaño del segmento, por lo que, a medida que aumenta el tamaño del segmento, pueden ser menos similares a los datos iniciales. Cuando la puntuación de relevancia del segmento llega a 0, el modelo predice que todos los usuarios del segmento similar provienen de la misma distribución que los datos iniciales. Al aumentar el tamaño de salida, es probable que se incluyan usuarios del segmento similar que no pertenezcan a la misma distribución que los datos iniciales.

Las puntuaciones de relevancia se normalizan en una sola campaña y no se deben utilizar para comparar campañas. Los puntajes de relevancia no deben utilizarse como una fuente única de evidencia para determinar ningún resultado empresarial, ya que se ven afectados por varios factores complejos además de la relevancia, como la calidad del inventario, el tipo de inventario, el momento en que se publica la publicidad, etc.

Las puntuaciones de relevancia no se deben utilizar para juzgar la calidad del inicio, sino para determinar si se puede aumentar o disminuir. Considere los siguientes ejemplos:

- Todas las puntuaciones son positivas: esto indica que hay más usuarios de salida que se consideran similares que los que se incluyen en el segmento similar. Esto es habitual en el caso de los datos iniciales que forman parte de un mercado grande, como el de todos los que han comprado pasta de dientes en el último mes. Le recomendamos que consulte los datos iniciales más pequeños, como los de todas las personas que han comprado pasta dental más de una vez en el último mes.
- Todas las puntuaciones negativas o negativas para su tamaño de segmento similar deseado: esto indica que Clean Rooms ML predice que no hay suficientes usuarios similares en el tamaño de segmento similar deseado. Esto se puede deber a que los datos iniciales son demasiado

específicos o a que el mercado es demasiado pequeño. Recomendamos aplicar menos filtros a los datos iniciales o ampliar el mercado. Por ejemplo, si los datos iniciales originales eran clientes que habían comprado un cochecito y una silla de coche, podría ampliar el mercado a clientes que hayan comprado varios productos para bebés.

Los proveedores de datos de entrenamiento determinan si se exponen las puntuaciones de relevancia y cuáles son los contenedores de bucket en los que se calculan las puntuaciones de relevancia.

Modelos personalizados en Clean Rooms ML

Con Clean Rooms ML, los miembros de una colaboración pueden utilizar un algoritmo de modelo personalizado dockerizado que se almacena en Amazon ECR para analizar sus datos de forma conjunta. Para ello, el proveedor del modelo debe crear una imagen y almacenarla en Amazon ECR. Siga los pasos de la Guía del usuario de Amazon Elastic Container Registry para crear un repositorio privado que contendrá el modelo de aprendizaje automático personalizado.

Cualquier miembro de una colaboración puede ser el proveedor del modelo, siempre que cuente con los permisos correctos. Todos los miembros de una colaboración pueden aportar datos al modelo. A los efectos de esta guía, los miembros que aportan datos se denominan proveedores de datos. El miembro que crea la colaboración es el creador de la colaboración y puede ser el proveedor del modelo, uno de los proveedores de datos o ambos.

En los temas siguientes se describe la información necesaria para crear un modelo de aprendizaje automático personalizado

Temas

- Requisitos previos del modelado ML personalizado
- Pautas de creación de modelos para el contenedor de formación
- Directrices de creación de modelos para el contenedor de inferencias
- Recibir registros y métricas del modelo

Requisitos previos del modelado ML personalizado

Antes de poder realizar un modelado de aprendizaje automático personalizado, debe tener en cuenta lo siguiente:

- Determine si tanto el entrenamiento del modelo como la inferencia sobre el modelo entrenado se realizarán en colaboración.
- Determine la función que desempeñará cada miembro de la colaboración y asígnele las habilidades adecuadas.
  - Asigne la `CAN_QUERY` habilidad al miembro que entrenará el modelo y realizará la inferencia sobre el modelo entrenado.
  - `CAN_RECEIVE_RESULTS`Asígnelo al menos a un miembro de la colaboración.
  - Asigne `CAN_RECEIVE_MODEL_OUTPUT` nuestras `CAN_RECEIVE_INFERENCE_OUTPUT` habilidades al miembro que recibirá las exportaciones de modelos entrenados o el resultado de inferencias, respectivamente. Puedes elegir usar ambas habilidades si así lo requiere tu caso de uso.
- Determine el tamaño máximo de los artefactos del modelo entrenado o de los resultados de inferencia que permitirá exportar.
- Recomendamos que todos los usuarios tengan las `CleanroomsMLFullAccess` políticas `CleanrooomsFullAccess` y políticas asociadas a su función. El uso de modelos de aprendizaje automático personalizados requiere utilizar tanto el AWS Clean Rooms modelo como el AWS Clean Rooms aprendizaje automático SDKs.
- Tenga en cuenta la siguiente información sobre las funciones de IAM.
  - Todos los proveedores de datos deben tener una función de acceso AWS Clean Rooms al servicio que les permita leer los datos de sus AWS Glue catálogos y tablas, y de las ubicaciones subyacentes de Amazon S3. Estas funciones son similares a las que se requieren para las consultas de SQL. Esto le permite utilizar la `CreateConfiguredTableAssociation` acción. Para obtener más información, consulte Cree un rol de servicio para crear una asociación de tablas configurada.
  - Todos los miembros que deseen recibir métricas deben tener una función de acceso al servicio que les permita escribir CloudWatch métricas y registros. Clean Rooms ML utiliza esta función para escribir todas las métricas y registros del modelo para el miembro Cuenta de AWS durante el entrenamiento y la inferencia del modelo. También ofrecemos controles de privacidad para determinar qué miembros tienen acceso a las métricas y los registros. Esto te permite usar la `CreateMLConfiguration` acción. Para obtener más información, consulte, Cree un rol de servicio para el modelado de ML personalizado: Configuración de ML.

    El miembro que reciba los resultados debe proporcionar un rol de acceso al servicio con permisos para escribir en su bucket de Amazon S3. Esta función permite a Clean Rooms ML exportar los resultados (artefactos de modelos entrenados o resultados de inferencias)

a un bucket de Amazon S3. Esto le permite utilizar la `CreateMLConfiguration` acción. Para obtener más información, consulte Cree un rol de servicio para el modelado de ML personalizado: Configuración de ML.

- El proveedor del modelo debe proporcionar un rol de acceso al servicio con permisos para leer su repositorio e imagen de Amazon ECR. Esto le permite utilizar la `CreateConfigureModelAlgorithm` acción. Para obtener más información, consulte Cree un rol de servicio para proporcionar un modelo de aprendizaje automático personalizado.
- El miembro que crea el `MLInputChannel` para generar conjuntos de datos con fines de formación o inferencia debe proporcionar una función de acceso al servicio que permita a Clean Rooms ML ejecutar una consulta SQL en ella. AWS Clean Rooms Esto le permite usar las acciones `CreateTrainedModel` y`StartTrainedModelInferenceJob`. Para obtener más información, consulte Cree un rol de servicio para consultar un conjunto de datos.

- Los autores del modelo deben seguir las Pautas de creación de modelos para el contenedor de formación y Directrices de creación de modelos para el contenedor de inferencias garantizar que las entradas y salidas del modelo estén configuradas según lo esperado AWS Clean Rooms.

Pautas de creación de modelos para el contenedor de formación

En esta sección se detallan las pautas que los proveedores de modelos deben seguir al crear un algoritmo de modelo de aprendizaje automático personalizado para el aprendizaje automático de salas limpias.

- Utilice la imagen base de contenedores adecuada y compatible con la formación en SageMaker IA, tal y como se describe en la Guía para desarrolladores de SageMaker IA. El siguiente código te permite extraer las imágenes de la base de contenedores compatibles de puntos finales de SageMaker IA públicos.

```
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
```

- Cuando cree el modelo localmente, asegúrese de lo siguiente para poder probarlo localmente, en una instancia de desarrollo, en SageMaker AI Training en su Cuenta de AWS empresa y en Clean Rooms ML.

- Recomendamos escribir un guion de formación que acceda a propiedades útiles del entorno de formación a través de diversas variables de entorno. Clean Rooms ML utiliza los siguientes argumentos para invocar el entrenamiento en el código del modelo:`SM_MODEL_DIR`, `SM_OUTPUT_DIRSM_CHANNEL_TRAIN`, y. `FILE_FORMAT` Clean Rooms ML utiliza estos valores predeterminados para entrenar su modelo de aprendizaje automático en su propio entorno de ejecución con los datos de todas las partes.
- Clean Rooms ML hace que sus canales de entrada de formación estén disponibles a través de los `/opt/ml/input/data/`*`channel-name`* directorios del contenedor docker. Cada canal de entrada de ML se mapea en función del correspondiente que `channel_name` se proporciona en la `CreateTrainedModel` solicitud.

```
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"))
```

- Asegúrese de poder generar un conjunto de datos sintético o de prueba basado en el esquema de los colaboradores que se utilizará en el código de su modelo.
- Asegúrese de poder realizar un trabajo de formación en SageMaker IA por su cuenta Cuenta de AWS antes de asociar el algoritmo del modelo a una AWS Clean Rooms colaboración.

  El siguiente código contiene un archivo Docker de muestra que es compatible con las pruebas locales, las pruebas del entorno de formación de SageMaker IA y el aprendizaje automático de salas limpias

```
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"]
```

- Para supervisar mejor los fallos de los contenedores, recomendamos exportar los registros y depurarlos en caso de que se produzcan errores. Como `GetTrainedModel` respuesta, Clean Rooms ML devuelve los primeros 1024 caracteres de este archivo. `StatusDetails`
- Una vez que haya realizado cualquier cambio en el modelo y esté listo para probarlo en el entorno de SageMaker IA, ejecute los siguientes comandos en el orden indicado.

```
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 Doker
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
```

Una vez que haya completado el trabajo de SageMaker IA y esté satisfecho con el algoritmo de su modelo, puede registrar el Registro ECR de Amazon con AWS Clean Rooms ML. Utilice la `CreateConfiguredModelAlgorithm` acción para registrar el algoritmo del modelo y asociarlo a una colaboración. `CreateConfiguredModelAlgorithmAssociation`

Directrices de creación de modelos para el contenedor de inferencias

En esta sección se detallan las pautas que los proveedores de modelos deben seguir al crear un algoritmo de inferencia para Clean Rooms ML.

- Utilice la imagen base de contenedor adecuada compatible con la inferencia de SageMaker IA, tal y como se describe en la SageMaker Guía para desarrolladores de IA. El siguiente código te permite extraer las imágenes de base de contenedores compatibles de puntos finales de IA públicos SageMaker .

```
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
```

- Cuando cree el modelo localmente, asegúrese de lo siguiente para poder probarlo localmente, en una instancia de desarrollo, en SageMaker AI Batch Transform en su Cuenta de AWS entorno y en Clean Rooms ML.
  - Clean Rooms ML hace que los artefactos de su modelo a partir de la inferencia estén disponibles para que los utilice su código de inferencia a través del `/opt/ml/model` directorio del contenedor docker.
  - Clean Rooms ML divide la entrada por línea, utiliza una estrategia por `MultiRecord` lotes y añade un carácter de nueva línea al final de cada registro transformado.
  - Asegúrese de poder generar un conjunto de datos de inferencia sintético o de prueba basado en el esquema de los colaboradores que se utilizará en el código de su modelo.
  - Asegúrese de poder ejecutar un trabajo de transformación por lotes de SageMaker IA por su cuenta Cuenta de AWS antes de asociar el algoritmo del modelo a una AWS Clean Rooms colaboración.

El siguiente código contiene un ejemplo de archivo Docker que es compatible con las pruebas locales, las pruebas de entornos de transformación mediante SageMaker IA y el aprendizaje automático de salas limpias

```
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"]
```

- Una vez que haya realizado cualquier cambio en el modelo y esté listo para probarlo en el entorno de SageMaker IA, ejecute los siguientes comandos en el orden indicado.

```
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
```

Una vez que haya completado el trabajo de SageMaker IA y esté satisfecho con la transformación por lotes, puede registrar el registro Amazon ECR con AWS Clean Rooms ML. Utilice la `CreateConfiguredModelAlgorithm` acción para registrar el algoritmo del modelo y asociarlo a una colaboración. `CreateConfiguredModelAlgorithmAssociation`

Recibir registros y métricas del modelo

Para recibir registros y métricas del entrenamiento o la inferencia de modelos personalizados, los miembros deben haber creado una configuración de aprendizaje automático con un rol válido que proporcione los CloudWatch permisos necesarios (consulte Crear un rol de servicio para el modelado de aprendizaje automático personalizado: configuración de aprendizaje automático).

Métrica del sistema

Las métricas del sistema, tanto para el entrenamiento como para la inferencia, como el uso de la CPU y la memoria, se publican para todos los miembros en colaboración con las configuraciones de aprendizaje automático válidas. Estas métricas se pueden ver a medida que avanza el trabajo mediante CloudWatch las métricas de los espacios de `/aws/cleanroomsml/TrainedModelInferenceJobs` nombres `/aws/cleanroomsml/TrainedModels` o los espacios de nombres, respectivamente.

Registros del modelo

El acceso a los registros del modelo lo proporciona la política de configuración de privacidad de cada algoritmo del modelo configurado. El autor del modelo establece la política de configuración de privacidad al asociar un algoritmo de modelo configurado (mediante la consola o

la `CreateConfiguredModelAlgorithmAssociation` API) a una colaboración. Al establecer la política de configuración de privacidad, se controla qué miembros pueden recibir los registros del modelo.

Además, el autor del modelo puede establecer un patrón de filtro en la política de configuración de privacidad para filtrar los eventos del registro. Todos los registros que envía un contenedor modelo `stdout` o `stderr` que coinciden con el patrón de filtro (si está establecido) se envían a Amazon CloudWatch Logs. Los registros del modelo están disponibles en grupos de CloudWatch registros `/aws/cleanroomsml/TrainedModels` o`/aws/cleanroomsml/TrainedModelInferenceJobs`, respectivamente.

Métricas definidas personalizadas

Al configurar un algoritmo de modelo (mediante la consola o la `CreateConfiguredModelAlgorithm` API), el autor del modelo puede proporcionar nombres de métricas e instrucciones de expresiones regulares específicos para buscarlos en los registros de salida. Se pueden ver a medida que avanza el trabajo mediante CloudWatch Metrics en el `/aws/cleanroomsml/TrainedModels` espacio de nombres. Al asociar un algoritmo de modelo configurado, el autor del modelo puede establecer un nivel de ruido opcional en la configuración de privacidad de las métricas para evitar generar datos sin procesar y, al mismo tiempo, ofrecer visibilidad de las tendencias métricas personalizadas. Si se establece un nivel de ruido, las métricas se publican al final del trabajo y no en tiempo real.

Computación criptográfica para Clean Rooms

Computación criptográfica para Clean Rooms (C3R) es una capacidad AWS Clean Rooms que se puede utilizar además de las reglas de análisis. Con C3R, las organizaciones pueden recopilar datos confidenciales para obtener nuevos conocimientos a partir del análisis de datos al tiempo que limitan criptográficamente la información que puede conocer cualquiera de las partes que interviene en el proceso. C3R lo pueden utilizar dos o más partes que deseen colaborar con sus datos confidenciales, pero que estén obligadas a usar solo datos cifrados en la nube.

El cliente de cifrado C3R es una herramienta de cifrado del lado del cliente que puede utilizar para cifrar sus datos y utilizarlos con ellos. AWS Clean Rooms Cuando utiliza el cliente de cifrado C3R, los datos permanecen protegidos criptográficamente mientras se utilizan en una colaboración. AWS Clean Rooms Al igual que en una AWS Clean Rooms colaboración habitual, los datos de entrada son tablas de bases de datos relacionales y el cálculo se expresa como una consulta SQL. Sin embargo, C3R admite solo un subconjunto limitado de consultas SQL en datos cifrados.

En concreto, C3R es compatible con SQL JOIN y SELECT declaraciones sobre datos protegidos criptográficamente. Cada columna de la tabla de entrada se puede utilizar exactamente en uno de los siguientes tipos de instrucciones SQL:

- Columnas que están protegidas criptográficamente para su uso en JOIN las declaraciones se llaman fingerprint columnas.
- Columnas que están protegidas criptográficamente para su uso en SELECT las declaraciones se llaman sealed columnas.
- Columnas que no están protegidas criptográficamente para su uso en JOIN o SELECT las declaraciones se llaman cleartext columnas.

En algunos casos, GROUP BY las declaraciones se apoyan en fingerprint columnas. Para obtener más información, consulte Fingerprint columns. Actualmente, C3R no admite el uso de otras construcciones de SQL en datos cifrados, como WHERE cláusulas o funciones agregadas como SUM y AVERAGE, incluso si de otro modo estarían permitidas por las reglas de análisis pertinentes.

C3R está diseñado para proteger los datos de celdas específicas de una tabla. Con la configuración predeterminada de C3R, los datos subyacentes que un cliente pone a disposición de terceros a través de una colaboración permanecen cifrados mientras el contenido se utiliza en AWS Clean Rooms. C3R utiliza el cifrado AES-GCM estándar de la industria para todos sealed columnas y una función pseudoaleatoria estándar del sector, conocida como código de autenticación de mensajes (HMAC) basado en hash, para proteger fingerprint columnas.

A pesar de que C3R cifra los datos de las tablas, es posible que aún se pueda inferir la siguiente información:

- Información sobre las propias tablas, incluido el número de columnas, los nombres de columna y el número de filas de la tabla.
- Como ocurre con la mayoría de las formas de cifrado estándar, C3R no intenta ocultar la longitud de los valores cifrados. C3R ofrece la posibilidad de rellenar los valores cifrados para ocultar la longitud exacta del texto sin cifrar. Sin embargo, aun así se podría revelar un límite superior de la longitud del texto sin cifrar de cada columna a un tercero.
- Información de registro, como cuándo se agregó una fila determinada a una tabla cifrada de C3R.

Para obtener más información acerca de CR3, consulte los siguientes temas.

Temas

- Consideraciones al utilizar la computación criptográfica para Clean Rooms
- Tipos de archivos y datos compatibles en la computación criptográfica para Clean Rooms
- Nombres de columnas en computación criptográfica para Clean Rooms
- Tipos de columnas en computación criptográfica para Clean Rooms
- Parámetros de computación criptográfica
- Indicadores opcionales en la computación criptográfica para Clean Rooms
- Consultas con computación criptográfica para Clean Rooms
- Directrices para el cliente de cifrado de C3R

Consideraciones al utilizar la computación criptográfica para Clean Rooms

Computación criptográfica para Clean Rooms (C3R) busca maximizar la protección de datos. Sin embargo, algunos casos de uso pueden beneficiarse de niveles más bajos de protección de datos a cambio de funcionalidades adicionales. Puede hacer estas concesiones específicas modificando C3R a partir de su configuración más segura. Como cliente, debe conocer estas ventajas y desventajas, y determinar si son apropiadas para su caso de uso. Entre las ventajas que debe considerar se incluyen las siguientes:

Temas

- Permitir la mezcla cleartext y datos cifrados en sus tablas
- Permitir la entrada de valores repetidos fingerprint columns
- Disminuir las restricciones sobre cómo fingerprint las columnas se nombran
- Determinar cómo NULL se representan los valores

Para obtener más información sobre cómo establecer parámetros para estos escenarios, consulte Parámetros de computación criptográfica.

Permitir la mezcla cleartext y datos cifrados en sus tablas

Hacer que todos los datos se cifren en el cliente proporciona la máxima protección de datos. Sin embargo, esto limita ciertos tipos de consultas (por ejemplo, las SUM función agregada). El riesgo de permitir cleartext los datos son que es posible que cualquier persona con acceso a las tablas cifradas pueda deducir alguna información sobre los valores cifrados. Esto podría hacerse mediante la realización de un análisis estadístico del cleartext y los datos asociados.

Por ejemplo, supongamos que tiene las columnas de `City` y `State`. La `City` columna es cleartext y la `State` columna está cifrada. Cuando ve el valor `Chicago` en la columna `City`, esto le ayuda a determinar con una alta probabilidad que el `State` es `Illinois`. Por el contrario, si una columna es `City` y la otra columna es`EmailAddress`, cleartext `City`es poco probable que revele algo sobre un cifrado`EmailAddress`.

Para obtener más información acerca del parámetro para este escenario, consulte Permitir cleartext parámetro de columnas.

Permitir la entrada de valores repetidos fingerprint columns

Para un enfoque más seguro, asumimos que cualquier fingerprint la columna contiene exactamente una instancia de una variable. No se puede repetir ningún elemento en un fingerprint columna. El cliente de cifrado C3R los mapea cleartext valores en valores únicos que son indistinguibles de los valores aleatorios. Por lo tanto, es imposible deducir información acerca del cleartext a partir de estos valores aleatorios.

El riesgo de que se repitan valores en un fingerprint columna es que los valores repetidos darán como resultado valores repetidos de aspecto aleatorio. Por lo tanto, cualquier persona que tenga acceso a las tablas cifradas podría, en teoría, realizar un análisis estadístico de fingerprint columnas que podrían revelar información sobre cleartext valores.

Una vez más, supongamos que fingerprint la columna es `State` y cada fila de la tabla corresponde a un hogar estadounidense. Un análisis de frecuencia permitiría inferir qué estado es `California` y cuál es `Wyoming` con una alta probabilidad. Esta inferencia es posible porque `California` tiene muchos más residentes que `Wyoming`. Por el contrario, digamos el fingerprint La columna está en un identificador de hogar y cada hogar apareció en la base de datos entre 1 y 4 veces en una base de datos de millones de entradas. Es poco probable que un análisis de frecuencia revele información útil.

Para obtener más información acerca del parámetro para este escenario, consulte Parámetro Permitir duplicados.

Disminuir las restricciones sobre cómo fingerprint las columnas se nombran

De forma predeterminada, asumimos que cuando dos tablas se unen mediante cifrado fingerprint columnas, esas columnas tienen el mismo nombre en cada tabla. La razón técnica de este resultado es que, de forma predeterminada, obtenemos una clave criptográfica diferente para cifrar cada una fingerprint columna. Esa clave se deriva de una combinación de la clave secreta compartida de la

colaboración y el nombre de la columna. Si intentamos combinar dos columnas con nombres de columna diferentes, obtenemos claves diferentes y no podemos procesar una combinación válida.

Para solucionar este problema, puede desactivar la característica que deriva las claves a partir de cada nombre de columna. A continuación, el cliente de cifrado C3R utiliza una única clave derivada para todos fingerprint columnas. El riesgo reside en que se pueda realizar otro tipo de análisis de frecuencia que pueda revelar información.

Volvamos al ejemplo de `City` y `State`. Si obtenemos los mismos valores aleatorios para cada fingerprint columna (al no incorporar el nombre de la columna). `New York`tiene el mismo valor aleatorio en las `State` columnas `City` y. Nueva York es una de las pocas ciudades de EE. UU. donde el nombre de `City` coincide con el nombre del `State` nombre. Por el contrario, si su conjunto de datos tiene valores completamente diferentes en cada columna, no se filtra información.

Para obtener más información acerca del parámetro para este escenario, consulte Permitir JOIN de columnas con nombres diferentes (parámetro).

Determinar cómo NULL se representan los valores

La opción que tiene a su disposición es si desea procesar criptográficamente (cifrado y HMAC) NULL valores como cualquier otro valor. Si no procesas NULL valores como cualquier otro valor, la información puede revelarse.

Por ejemplo, supongamos que NULL en la `Middle Name` columna del cleartext indica personas sin segundo nombre. Si no cifra esos valores, filtrará qué filas de la tabla cifrada se utilizan para personas sin segundo nombre. Esa información podría ser una señal de identificación para algunas personas de ciertas poblaciones. Pero si lo procesas criptográficamente NULL valores, ciertas consultas SQL actúan de manera diferente. Por ejemplo: GROUP BY las cláusulas no se agruparán fingerprint NULL valores en fingerprint columnas juntas.

Para obtener más información acerca del parámetro para este escenario, consulte Conservar NULL parámetros de valores.

Tipos de archivos y datos compatibles en la computación criptográfica para Clean Rooms

El cliente de cifrado de C3R reconoce los siguientes tipos de archivo:

- Archivos CSV

- Parquet files

Puede utilizar el indicador `--fileFormat` del cliente de cifrado de C3R para especificar un formato de archivo de forma explícita. Cuando se especifica de forma explícita, el formato de archivo no viene determinado por la extensión del archivo.

Temas

- Archivos CSV
- Parquet files
- Cifrar valores que no son de cadena

Archivos CSV

Se presupone que un archivo con la extensión .csv tiene formato CSV y contiene texto codificado en UTF-8. El cliente de cifrado de C3R trata todos los valores como cadenas.

Propiedades admitidas en los archivos .csv

El cliente de cifrado de C3R requiere que los archivos .csv tengan las siguientes propiedades:

- Puede contener o no una fila de encabezado inicial que designe de manera exclusiva cada columna.
- Delimitado por comas. (actualmente no se admiten delimitadores personalizados).
- Texto codificado en UTF-8.

Recorte de espacios en blanco de las entradas .csv

Los espacios en blanco iniciales y finales se recortan de las entradas .csv.

Personalizada NULL codificación para un archivo.csv

Un archivo.csv puede usar custom NULL codificación.

Con el cliente de cifrado C3R, puede especificar codificaciones personalizadas para NULL entradas en los datos de entrada mediante la marca. `--csvInputNULLValue=<csv-input-null>` El cliente de cifrado de C3R puede usar codificaciones personalizadas en el archivo de salida generado para las entradas NULL utilizando el indicador `--csvOutputNULLValue=<csv-output-null>`.

> **Note**
>
> A NULL se considera que una entrada carece de contenido, específicamente en el contexto de un formato tabular más rico, como una tabla SQL. Aunque .csv no admite explícitamente esta caracterización por razones históricas, es una convención común considerar que una entrada vacía que contiene solo espacios en blanco es NULL. Por lo tanto, ese es el comportamiento predeterminado del cliente de cifrado C3R y se puede personalizar según sea necesario.

¿Cómo interpreta C3R las entradas .csv?

En la siguiente tabla se muestran ejemplos de cómo se ordenan las entradas.csv (cleartext a cleartext para mayor claridad) en función de los valores (si los hay) que se proporcionan para los indicadores `--csvInputNULLValue=<csv-input-null>` y`--csvOutputNULLValue=<csv-output-null>`. Los espacios en blanco iniciales y finales fuera de las comillas se recortan antes de que C3R interprete el significado de cualquier valor.

<csv-input-null>	**<csv-output-null>**	Entrada de entrada	Entrada de salida
Ninguno	Ninguna	`,AnyProduct,`	`,AnyProduct,`
Ninguna	Ninguna	`, AnyProduct ,`	`,AnyProduct,`
Ninguna	Ninguna	`,"AnyProduct",`	`,AnyProduct,`
Ninguna	Ninguna	`, "AnyProduct" ,`	`,AnyProduct,`
Ninguna	Ninguna	`,,`	`,,`
Ninguna	Ninguna	`, ,`	`,,`
Ninguna	Ninguna	`,"",`	`,,`
Ninguna	Ninguna	`," ",`	`," ",`
Ninguna	Ninguna	`, " " ,`	`," ",`

<csv-input-null>	**<csv-output-null>**	Entrada de entrada	Entrada de salida
"AnyProduct"	"NULL"	,AnyProduct,	,NULL,
"AnyProduct"	"NULL"	, AnyProduct ,	,NULL,
"AnyProduct"	"NULL"	,"AnyProduct",	,NULL,
"AnyProduct"	"NULL"	, "AnyProdu ct" ,	,NULL,
Ninguna	"NULL"	,,	,NULL,
Ninguna	"NULL"	, ,	,NULL,
Ninguna	"NULL"	,"",	,NULL,
Ninguna	"NULL"	," ",	," ",
Ninguno	"NULL"	, " " ,	," ",
""	"NULL"	,,	,NULL,
""	"NULL"	, ,	,NULL,
""	"NULL"	,"",	,"",
""	"NULL"	," ",	," ",
""	"NULL"	, " " ,	," ",
"\"\""	"NULL"	,,	,,
"\"\""	"NULL"	, ,	,,
"\"\""	"NULL"	,"",	,NULL,
"\"\""	"NULL"	," ",	," ",
"\"\""	"NULL"	, " " ,	," ",

Archivo CSV sin encabezados

No es necesario que el archivo .csv de origen tenga encabezados en la primera fila que asignen un nombre exclusivo a cada columna. Sin embargo, un archivo .csv sin una fila de encabezados requiere un esquema de cifrado posicional. Se requiere el esquema de cifrado posicional en lugar del esquema mapeado típico que se usa para los archivos.csv con una fila de encabezado y Parquet archivos.

Un esquema de cifrado posicional especifica las columnas de salida por posición en lugar de por nombre. Un esquema de cifrado mapeado asigna los nombres de columnas de origen a nombres de columnas de destino. Para obtener más información, incluida una explicación detallada y ejemplos de ambos formatos de esquema, consulte Esquemas de tablas mapeados y posicionales.

Parquet files

Un archivo con un .parquet se supone que la extensión está en Apache Parquet formato.

Compatible Parquet tipos de datos

El cliente de cifrado C3R puede procesar cualquier dato no complejo (es decir, de tipo primitivo) en un Parquet archivo que representa un tipo de datos compatible con. AWS Clean Rooms

Sin embargo, solo se pueden usar columnas de cadena para sealed columnas.

Se admiten los siguientes tipos de datos de Parquet:

- Tipo primitivo `Binary` con las siguientes anotaciones lógicas:
  - Ninguno si `--parquetBinaryAsString` está definido (tipo de datos `STRING`)
  - `Decimal(scale, precision)` (tipo de datos `DECIMAL`)
  - `String` (tipo de datos `STRING`)
- Tipo de datos primitivo `Boolean` sin anotación lógica (tipo de datos `BOOLEAN`)
- Tipo de datos primitivo `Double` sin anotación lógica (tipo de datos `DOUBLE`)
- Tipo primitivo `Fixed_Len_Binary_Array` con anotación lógica `Decimal(scale, precision)` (tipo de datos `DECIMAL`)
- Tipo de datos primitivo `Float` sin anotación lógica (tipo de datos `FLOAT`)
- Tipo primitivo `Int32` con las siguientes anotaciones lógicas:
  - Ninguna (tipo de datos `INT`)

  - `Date` (tipo de datos `DATE`)
  - `Decimal(scale, precision)` (tipo de datos `DECIMAL`)
  - `Int(16, true)` (tipo de datos `SMALLINT`)
  - `Int(32, true)` (tipo de datos `INT`)
- Tipo de datos primitivo `Int64` con las siguientes anotaciones lógicas:
  - Ninguna (tipo de datos `BIGINT`)
  - `Decimal(scale, precision)` (tipo de datos `DECIMAL`)
  - `Int(64, true)` (tipo de datos `BIGINT`)
  - `Timestamp(isUTCAdjusted, TimeUnit.MILLIS)` (tipo de datos `TIMESTAMP`)
  - `Timestamp(isUTCAdjusted, TimeUnit.MICROS)` (tipo de datos `TIMESTAMP`)
  - `Timestamp(isUTCAdjusted, TimeUnit.NANOS)` (tipo de datos `TIMESTAMP`)

Cifrar valores que no son de cadena

Actualmente, solo se admiten valores de cadena sealed columnas.

En el caso de los archivos .csv, el cliente de cifrado de C3R trata todos los valores como texto codificado en UTF-8 y no intenta interpretarlos de forma diferente antes del cifrado.

En el caso de las columnas de huella digital, los tipos se agrupan en clases de equivalencia. Una clase de equivalencia es un conjunto de tipos de datos que se pueden comparar de forma inequívoca para determinar su igualdad mediante un tipo de datos representativo.

Las clases de equivalencia permiten asignar huellas dactilares idénticas al mismo valor semántico independientemente de la representación original. Sin embargo, el mismo valor en dos clases de equivalencia no dará como resultado la misma columna de huella digital.

Por ejemplo, al valor `INTEGRAL 42` se le asignará la misma huella digital independientemente de si originalmente era `SMALLINT`, `INT` o `BIGINT`. Además, el valor `INTEGRAL 0` nunca coincidirá con el valor `BOOLEAN FALSE` (que se representa mediante el valor `0`).

Las columnas de huellas digitales admiten las siguientes clases de equivalencia y los tipos de AWS Clean Rooms datos correspondientes:

Clase de equivalencia	Tipo de datos AWS Clean Rooms admitido
BOOLEAN	BOOLEAN
DATE	DATE
INTEGRAL	BIGINT, INT, SMALLINT
STRING	CHAR, STRING, VARCHAR

Nombres de columnas en computación criptográfica para Clean Rooms

De forma predeterminada, los nombres de las columnas son importantes en la informática criptográfica para Clean Rooms.

Si el valor de Permitir JOIN el parámetro de columnas con nombres diferentes es falso, los nombres de las columnas se utilizan durante el cifrado de fingerprint columnas. Por este motivo, de forma predeterminada, los colaboradores deben coordinarse con antelación y utilizar los mismos nombres de columna de destino para los datos que utilizarán JOIN declaraciones en las consultas. De forma predeterminada, las columnas están cifradas para JOIN con nombres diferentes no se ejecutan correctamente JOIN en cualquier valor.

Si el valor de Permitir JOIN el parámetro de columnas con nombres diferentes es verdadero, JOIN declaraciones entre columnas cifradas como fingerprint las columnas se ejecutan correctamente. Cifrar los datos con este parámetro podría permitir inferir alguna cleartext valores. Por ejemplo, si una fila tiene el mismo valor de código de autenticación de mensajes basado en hash (HMAC) tanto en la columna `City` como en la columna `State`, el valor podría ser `New York`.

Normalización de los nombres de encabezado de columna

El cliente de cifrado de C3R normaliza los nombres de los encabezados de columna. Se eliminan todos los espacios en blanco iniciales y finales, y el nombre de la columna se cambia a minúsculas en el resultado transformado.

La normalización se aplica antes de cualquier otro cómputo, cálculo u otra operación que pueda verse afectada por los nombres de columna. El archivo de salida emitido contiene solo los nombres normalizados.

Tipos de columnas en computación criptográfica para Clean Rooms

En este tema se proporciona información sobre los tipos de columnas en la informática criptográfica para Clean Rooms.

Temas

- Fingerprint columns
- Columnas selladas
- Cleartext columns

Fingerprint columns

Fingerprint las columnas son columnas que están protegidas criptográficamente para su uso en JOIN declaraciones.

Datos de fingerprint las columnas no se pueden descifrar. Solo se pueden descifrar los datos de las columnas selladas.

Fingerprint las columnas solo deben usarse en las siguientes cláusulas y funciones de SQL:

- JOIN (INNER, OUTER, LEFT, RIGHT, or FULL) contra otros fingerprint columnas:
  - Si el valor del `allowJoinsOnColumnsWithDifferentNames` parámetro está establecido en`false`, ambos fingerprint columnas del JOIN también deben tener el mismo nombre.
- `SELECT COUNT()`
- `SELECT COUNT(DISTINCT )`
- `GROUP BY` (usar solo si la colaboración ha definido el valor del parámetro `preserveNulls` como `true`).

Las consultas que infrinjan estas restricciones pueden arrojar resultados incorrectos.

Columnas selladas

Las columnas selladas son columnas que están protegidas criptográficamente para su uso en SELECT declaraciones.

Las columnas selladas solo deben usarse en las siguientes cláusulas y funciones SQL:

- `SELECT`

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

> **Note**
>
> `SELECT COUNT(DISTINCT )` no se admite.

Las consultas que infrinjan estas restricciones pueden arrojar resultados incorrectos.

Relleno de datos para un sealed columna antes del cifrado

Cuando se especifica que una columna debe ser sealed columna, C3R le pregunta qué tipo de relleno elegir. El relleno de datos antes del cifrado es opcional. Sin relleno (tipo blocnone), la longitud de los datos cifrados indica el tamaño del cleartext. En algunas circunstancias, el tamaño del cleartext podría exponer el texto sin formato. Con el relleno (relleno de tipo `fixed` o `max`), todos los valores se rellenan primero hasta alcanzar un tamaño común, y entonces se cifran. Con el relleno, la longitud de los datos cifrados no proporciona información sobre el original cleartext longitud, aparte de indicar un límite superior en su tamaño.

Si desea rellenar una columna y conoce la longitud máxima en bytes de los datos de esa columna, utilice el relleno `fixed`. Utilice un valor `length` que sea al menos tan grande como la mayor longitud en bytes de esa columna.

> **Note**
>
> Si un valor supera la `length` proporcionada, se produce un error y el cifrado falla.

Si desea rellenar una columna y no conoce la longitud máxima en bytes de los datos de esa columna, utilice el relleno `max`. Este modo de relleno rellena todos los datos hasta la longitud del valor más largo, más los bytes de `length` adicionales.

> **Note**
>
> Es posible que desee cifrar los datos por lotes o actualizar sus tablas con nuevos datos de forma periódica. Tenga en cuenta que el relleno `max` rellenará las entradas hasta la longitud (más el byte de `length`) de la entrada de texto sin formato más larga de un lote determinado. Esto significa que la longitud del texto cifrado puede variar de un lote a otro.

Por lo tanto, si conoce la longitud máxima en bytes de una columna, entonces debe utilizar el relleno `fixed` en lugar de `max`.

Cleartext columns

Cleartext las columnas son columnas que no están protegidas criptográficamente para su uso en JOIN o SELECT declaraciones.

Cleartext las columnas se pueden usar en cualquier parte de la consulta SQL.

Parámetros de computación criptográfica

Los parámetros de computación criptográfica están disponibles para colaboraciones que utilizan la computación criptográfica para Clean Rooms (C3R) al crear una colaboración. Puede crear una colaboración mediante la AWS Clean Rooms consola o la operación de la `CreateCollaboration` API. En la consola, puede establecer valores para los parámetros en Parámetros de computación criptográfica después de activar la opción Admitir computación criptográfica. Para obtener más información, consulte los siguientes temas.

Temas

- Permitir cleartext parámetro de columnas
- Parámetro Permitir duplicados
- Permitir JOIN de columnas con nombres diferentes (parámetro)
- Conservar NULL parámetros de valores

Permitir cleartext parámetro de columnas

En la consola, puede configurar el comando Permitir cleartext el parámetro de columnas al crear una colaboración para especificar si cleartext los datos están permitidos en una tabla con datos cifrados.

En la siguiente tabla se describen los valores de la opción Permitir cleartext parámetro de columnas.

Valor del parámetro	Descripción
No	Cleartext no se permiten columnas en la tabla cifrada. Todos los datos están protegidos criptográficamente.

Valor del parámetro	Descripción
Sí	Cleartext se permiten columnas en la tabla cifrada. Cleartext las columnas no están protegidas criptográficamente y se incluyen como cleartext. Deberías tomar nota de cuáles son tus filas. cleartext los datos pueden revelar información sobre los demás datos de la tabla. Para ejecutar SUM o AVG en columnas específicas, las columnas deben estar en cleartext.

Mediante la operación `CreateCollaboration` de la API, para el parámetro `dataEncryptionMetadata`, puede establecer el valor de `allowCleartext` en `true` o `false`. Para obtener más información sobre las operaciones de la API, consulte Referencia de la API de AWS Clean Rooms.

Cleartext las columnas corresponden a columnas clasificadas como cleartext en el esquema específico de la tabla. Los datos de estas columnas no están cifrados y se pueden utilizar de cualquier forma. Cleartext las columnas pueden resultar útiles si los datos no son confidenciales o si se necesita más flexibilidad que una información cifrada sealed columna o fingerprint la columna permite.

Parámetro Permitir duplicados

En la consola, puede configurar el parámetro Permitir duplicados al crear una colaboración para especificar si las columnas están cifradas para JOIN las consultas pueden contener duplicados que no seanNULL valores.

> **Important**
>
> El Permitir duplicados, Permitir JOIN de columnas con diferentes nombres y Preservar NULL los parámetros de valores tienen efectos separados pero relacionados.

En la siguiente tabla se describen los valores del parámetro Permitir duplicados.

Valor del parámetro	Descripción
No	No se permiten valores repetidos en un fingerprint columna. Todos los valores en una sola fingerprint la columna debe ser única.
Sí	Se permiten valores repetidos en un fingerprint columna. Si necesita combinar columnas con valores repetidos, defina este valor en Sí. Cuando se establece en Sí, los patrones de frecuencia aparecen dentro fingerprint las columnas de la tabla C3R o los resultados pueden implicar información adicional sobre la estructura del cleartext datos.

Mediante la operación `CreateCollaboration` de la API, en el parámetro `dataEncryptionMetadata`, puede establecer el valor de `allowDuplicates` en `true` o `false`. Para obtener más información sobre las operaciones de la API, consulte Referencia de la API de AWS Clean Rooms.

De forma predeterminada, si los datos cifrados deben usarse en JOIN En las consultas, el cliente de cifrado C3R requiere que esas columnas no tengan valores duplicados. Con este requisito se pretende aumentar la protección de los datos. Este comportamiento puede contribuir a garantizar que los patrones repetidos en los datos no sean observables. Sin embargo, si desea trabajar con datos cifrados en JOIN consultas y no le preocupan los valores duplicados, el parámetro Permitir duplicados puede deshabilitar esta comprobación conservadora.

Permitir JOIN de columnas con nombres diferentes (parámetro)

En la consola, puede configurar el comando Permitir JOIN de columnas con nombres diferentes al crear una colaboración para especificar si JOIN Se admiten declaraciones entre columnas con nombres diferentes.

Para obtener más información, consulte Normalización de los nombres de encabezado de columna

En la siguiente tabla se describen los valores de la opción Permitir JOIN de columnas con nombres diferentes.

Valor del parámetro	Descripción
No	Uniones de fingerprint no se admiten columnas con nombres diferentes. JOIN las declaraciones solo proporcionan resultado s precisos en las columnas que tienen el mismo nombre. **Important** El valor No proporciona una mayor seguridad de la información, pero requiere que los participantes en la colaboración acuerden previamente los nombres de las columnas. Si dos columnas tienen nombres diferentes cuando se cifran como fingerprint columnas y Permitir JOIN de columnas con nombres diferente s se establece en No, JOIN las declaraciones de esas columnas no producen ningún resultado. Esto se debe a que, después del cifrado, no se comparten valores entre ellas.
Sí	Uniones de fingerprint Se admiten columnas con nombres diferentes. Para mayor flexibilidad, los usuarios pueden establecer este valor en Sí, lo que permite JOIN declaraciones en columnas independientemente del nombre de la columna. Si se establece en Sí, el cliente de cifrado C3R no tiene en cuenta el nombre de la columna al proteger fingerprint columnas. Como resultado, valores comunes en diferentes fingerprint las columnas se pueden observar en la tabla C3R. Por ejemplo, si una fila tiene el mismo cifrado JOIN valor tanto en una `City` columna como en una `State` columna, podría ser razonable inferir que ese valor es`New York`.

Mediante la operación `CreateCollaboration` de la API, para el parámetro `dataEncryptionMetadata`, puede establecer el valor de

`allowJoinsOnColumnsWithDifferentNames` en `true` o `false`. Para obtener más información sobre las operaciones de la API, consulte Referencia de la API de AWS Clean Rooms.

De forma predeterminada, fingerprint el cifrado de columnas se ve afectado por el valor `targetHeader` para esa columna, establecido enPaso 4: generar un esquema de cifrado para un archivo tabular . Por lo tanto, lo mismo cleartext el valor tiene diferentes representaciones cifradas en cada una de ellas fingerprint columna para la que está cifrado.

Este parámetro puede ser útil para evitar la inferencia de cleartext valores en algunos casos. Por ejemplo, ver el mismo valor cifrado en fingerprint columnas `City` y `State` podría usarse para inferir razonablemente que el valor es`New York`. Sin embargo, el uso de este parámetro requiere una coordinación adicional por adelantado, de modo que todas las columnas que se van a combinar en consultas tengan nombres compartidos.

Puede utilizar la opción Permitir JOIN el parámetro de columnas con nombres diferentes para reducir esta restricción. Cuando el valor del parámetro se establece en`Yes`, permite cifrar cualquier columna para JOIN se pueden usar juntas independientemente del nombre.

Conservar NULL parámetros de valores

En la consola, puede configurar el comando Preserve NULL valores el parámetro al crear una colaboración para indicar que no hay ningún valor presente para esa columna.

En la siguiente tabla se describen los valores de Preserve NULL parámetro de valores.

Valor del parámetro	Descripción
No	NULL los valores no se conservan. NULL los valores no aparecen como NULL en una tabla cifrada. NULL los valores aparecen como valores aleatorios únicos en una tabla C3R.
Sí	NULL los valores se conservan. NULL los valores aparecen como NULL en una tabla cifrada. Si necesita una semántica de SQL de NULL valores, puede establecer este valor en Sí. Como resultado, NULL las entradas aparecen como NULL en la tabla C3R, independientemente de si la columna está cifrada o no y del parámetro Permitir duplicados.

Mediante la operación `CreateCollaboration` de la API, para el parámetro `dataEncryptionMetadata`, puede establecer el valor de `preserveNulls` en `true` o `false`. Para obtener más información sobre las operaciones de la API, consulte Referencia de la API de AWS Clean Rooms.

Cuando el Preservar NULL El parámetro de valores está establecido en No para la colaboración:

1. NULL las entradas de `cleartext` las columnas permanecen inalteradas.
2. NULL las entradas de `fingerprint` las columnas cifradas se cifran como valores aleatorios para ocultar su contenido. Unirse a una columna cifrada con NULL entradas en el cleartextla columna no produce ninguna coincidencia con ninguna de las NULL entradas. No se realizan coincidencias porque cada una recibe su propio contenido aleatorio único.
3. NULL las entradas de `sealed` las columnas cifradas están cifradas.

Cuando el valor de la reserva NULL el parámetro de valores está establecido en Sí para la colaboración, NULL las entradas de todas las columnas permanecen como NULL independientemente de si la columna está cifrada.

La reserva NULL El parámetro valores es útil en escenarios como el enriquecimiento de datos, en los que se desea compartir la falta de información expresada como NULL. La reserva NULL El parámetro values también es útil en fingerprint o formato HMAC si tiene NULL valores en la columna que desee JOIN o GROUP BY.

Si el valor de Permitir duplicados y Preservar NULL el parámetro de valores está establecido en No y tiene más de uno NULL entrada en un fingerprint la columna produce un error y detiene el cifrado. Si el valor de cualquiera de los parámetros se establece en Sí, no se produce tal error.

Indicadores opcionales en la computación criptográfica para Clean Rooms

En las siguientes secciones se describen los indicadores opcionales que puede configurar al cifrar datos utilizando el cliente de cifrado de C3R para personalizar y probar archivos tabulares.

Temas

- Indicador --csvInputNULLValue
- Indicador --csvOutputNULLValue
- Indicador --enableStackTraces
- Indicador --dryRun

- Indicador --tempDir

Indicador **--csvInputNULLValue**

Puede usar el `--csvInputNULLValue` indicador para especificar codificaciones personalizadas para NULL entradas en los datos de entrada al cifrar datos mediante el cliente de cifrado C3R.

En la siguiente tabla se resumen el uso y los parámetros de este indicador.

Uso	Parámetros
Opcional. Los usuarios pueden especificar codificaciones personalizadas para NULL entradas en los datos de entrada.	Codificación especificada por el usuario de NULL valores en el archivo CSV de entrada

A NULL la entrada es una entrada que se considera carente de contenido, específicamente en el contexto de un formato tabular más rico, como una tabla SQL. Aunque .csv no admite explícitamente esta caracterización por razones históricas, es una convención común considerar que una entrada vacía que contiene solo espacios en blanco es NULL. Por lo tanto, ese es el comportamiento predeterminado del cliente de cifrado C3R y se puede personalizar según sea necesario.

Indicador **--csvOutputNULLValue**

Puede usar la `--csvOutputNULLValue` marca para especificar codificaciones personalizadas para NULL entradas en los datos de salida al cifrar datos mediante el cliente de cifrado C3R.

En la siguiente tabla se resumen el uso y los parámetros de este indicador.

Uso	Parámetros
Opcional. Los usuarios pueden especificar codificaciones personalizadas en el archivo de salida generado para NULL entradas.	Codificación especificada por el usuario de NULL valores del archivo CSV de salida

A NULL la entrada es una entrada que se considera carente de contenido, específicamente en el contexto de un formato tabular más rico, como una tabla SQL. Aunque .csv no admite explícitamente

esta caracterización por razones históricas, es una convención común considerar que una entrada vacía que contiene solo espacios en blanco es NULL. Por lo tanto, ese es el comportamiento predeterminado del cliente de cifrado C3R y se puede personalizar según sea necesario.

Indicador `--enableStackTraces`

Al cifrar datos utilizando el cliente de cifrado de C3R, utilice el indicador `--enableStackTraces` para proporcionar información contextual adicional para notificar errores cuando C3R detecte un error.

AWS no recopila errores. Si encuentras un error, utiliza el seguimiento de pila para solucionar el error tú mismo o envía el seguimiento de pila a Soporte para obtener ayuda.

En la siguiente tabla se resumen el uso y los parámetros de este indicador.

Uso	Parámetros
Opcional. Se utiliza para proporcionar información contextual adicional para notificar errores cuando el cliente de cifrado de C3R detecta un error.	Ninguno

Indicador `--dryRun`

Los comandos encrypt y decrypt del cliente de cifrado de C3R incluyen un indicador `--dryRun` opcional. El indicador toma todos los argumentos proporcionados por el usuario y comprueba su validez y coherencia.

Puede usar el indicador `--dryRun` para comprobar si su archivo de esquema es válido y coherente con el archivo de entrada correspondiente.

En la siguiente tabla se resumen el uso y los parámetros de este indicador.

Uso	Parámetros
Opcional. Hace que el cliente de cifrado de C3R analice los parámetros y compruebe los	Ninguno

Uso	Parámetros
archivos, pero no realiza ningún cifrado ni descifrado.	

Indicador `--tempDir`

Es posible que desee utilizar un directorio temporal, ya que los archivos cifrados a veces pueden superar en tamaño a los no cifrados, dependiendo de la configuración. Los conjuntos de datos también deben cifrarse por colaboración para funcionar correctamente.

Al cifrar los datos utilizando C3R, utilice el indicador `--tempDir` para especificar la ubicación en la que se pueden crear los archivos temporales mientras se procesa la entrada.

En la siguiente tabla se resumen el uso y los parámetros de este indicador.

Uso	Parámetros
Los usuarios pueden especificar la ubicación en la que se pueden crear los archivos temporales mientras se procesa la entrada.	El valor predeterminado es el directorio temporal del sistema.

Consultas con computación criptográfica para Clean Rooms

En este tema se proporciona información sobre cómo escribir consultas que utilicen tablas de datos cifradas mediante computación criptográfica para Clean Rooms.

Temas

- Consultas que se ramifican NULL
- Asignar una columna de origen a varias columnas de destino
- Utilizar los mismos datos para ambos JOIN y SELECT consultas

Consultas que se ramifican NULL

Para tener una rama de consultas en un NULL declaración significa usar una sintaxis como`IF x IS NULL THEN 0 ELSE 1`.

Las consultas siempre se pueden ramificar NULL declaraciones en cleartext columnas.

Las consultas pueden ramificarse NULL declaraciones en sealed columnas y fingerprint columnas solo cuando el valor del parámetro Conservar valores NULOS (`preserveNulls`) esté establecido en `true`.

Las consultas que infrinjan estas restricciones pueden arrojar resultados incorrectos.

Asignar una columna de origen a varias columnas de destino

Una columna de origen puede asignarse a varias columnas de destino. Por ejemplo, es posible que desee ambos JOIN y SELECT en una columna.

Para obtener más información, consulte Utilizar los mismos datos para ambos JOIN y SELECT consultas.

Utilizar los mismos datos para ambos JOIN y SELECT consultas

Si los datos de una columna no son confidenciales, pueden aparecer en un cleartext columna de destino, lo que permite su uso para cualquier propósito.

Si los datos de una columna son confidenciales y deben usarse para ambos JOIN y SELECT en las consultas, asigne esa columna de origen a dos columnas de destino del archivo de salida. Una columna está cifrada con el `type` como fingerprint columna, y una columna se cifra con la `type` como columna sellada. La generación de esquemas interactivos del cliente de cifrado C3R sugiere sufijos de encabezado de `_fingerprint` y `_sealed`. Estos sufijos de encabezado pueden ser una convención útil para diferenciar dichas columnas rápidamente.

Directrices para el cliente de cifrado de C3R

El cliente de cifrado de C3R es una herramienta que permite a las organizaciones recopilar datos confidenciales para obtener nueva información procesable a partir del análisis de datos. La herramienta limita criptográficamente lo que puede aprender cualquier parte y AWS en el proceso. Si bien esto es de vital importancia, el proceso de protección criptográfica de los datos puede suponer una sobrecarga considerable, tanto en términos de recursos de computación como de almacenamiento. Por lo tanto, es importante entender las ventajas y desventajas de usar cada ajuste y saber cómo optimizar la configuración sin dejar de mantener las garantías criptográficas deseadas. Este tema se centra en las implicaciones para el rendimiento de las diferentes configuraciones del cliente de cifrado de C3R y los esquemas.

Todas las configuraciones de cifrado del cliente de cifrado de C3R ofrecen diferentes garantías criptográficas. De forma predeterminada, los ajustes más seguros son aquellos en el nivel de la colaboración. Al habilitar funciones adicionales al crear una colaboración, se debilitan las garantías de privacidad, ya que se permite realizar actividades como análisis de frecuencia en el texto cifrado. Para obtener más información sobre cómo se utilizan estas configuraciones y cuáles son sus implicaciones, consulte the section called “Computación criptográfica”.

Temas

- Implicaciones en el rendimiento de los tipos de columnas
- Solución de problemas relacionados con el aumento imprevisto de tamaño del texto cifrado

Implicaciones en el rendimiento de los tipos de columnas

C3R utiliza tres tipos de columnas: cleartext, fingerprint, y sealed. Cada uno de estos tipos de columnas proporciona diferentes garantías criptográficas y tiene diferentes usos previstos. En las siguientes secciones se analizan las implicaciones de rendimiento del tipo de columna y el impacto de cada ajuste en el rendimiento.

Temas

- Cleartext columns
- Fingerprint columns
- Sealed columns

Cleartext columns

Cleartext las columnas no se modifican con respecto a su formato original ni se procesan criptográficamente de ninguna manera. Este tipo de columna no se puede configurar y no afecta al rendimiento en términos de almacenamiento o de computación.

Fingerprint columns

Fingerprint las columnas están diseñadas para unir datos en varias tablas. Para ello, el tamaño del texto cifrado resultante debe ser siempre el mismo. Sin embargo, estas columnas se ven afectadas por la configuración del nivel de colaboración. Fingerprint las columnas pueden tener distintos grados de impacto en el tamaño del archivo de salida en función del cleartext contenidas en la entrada.

Temas

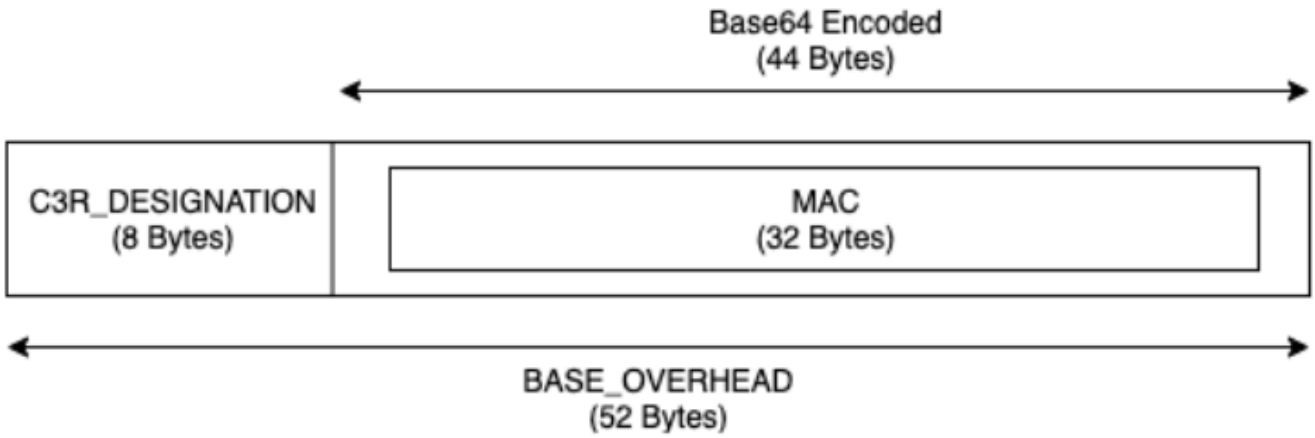
- Superficie base para fingerprint columns
- Configuración de colaboración para fingerprint columns
- Datos de ejemplo para un fingerprint columna
- Solución de problemas fingerprint columns

Superficie base para fingerprint columns

Hay una base superior para fingerprint columnas. Esta sobrecarga es constante y reemplaza el tamaño del cleartext bytes.

Datos en el fingerprint Las columnas se procesan criptográficamente mediante una función de código de autenticación de mensajes (HMAC) basada en hash, que convierte los datos en un código de autenticación de mensajes (MAC) de 32 bytes. A continuación, estos datos se procesan a través de un codificador base64, lo que aumenta el tamaño del byte aproximadamente un 33 por ciento. Va precedido de una designación C3R de 8 bytes para designar el tipo de columna a la que pertenecen los datos y la versión de cliente que los generó. El resultado final es de 52 bytes. Este resultado se multiplica entonces por el número de filas para obtener la sobrecarga de base total (utilice el número total de valores distintos de `null` si `preserveNulls` se establece en true).

En la siguiente imagen se muestra cómo *`BASE_OVERHEAD = C3R_DESIGNATION + (MAC * 1.33)`*



El texto cifrado de salida en la fingerprint las columnas siempre tendrán 52 bytes. Esto puede suponer una disminución de almacenamiento significativa si la entrada cleartext los datos tienen un promedio de más de 52 bytes (por ejemplo, direcciones postales completas). Esto puede suponer un aumento de almacenamiento significativo si la entrada cleartext los datos promedian menos de 52 bytes (por ejemplo, la edad de los clientes).

Configuración de colaboración para fingerprint columns

Ajuste **`preserveNulls`**

Cuando el ajuste de nivel de la colaboración `preserveNulls` es `false` (predeterminado), cada valor `null` se sustituye por 32 bytes únicos y aleatorios y se procesa como si no fuera `null`. El resultado es que cada valor `null` tiene ahora 52 bytes. Esto puede añadir requisitos de almacenamiento importantes para las tablas que contienen datos muy dispersos en comparación con cuando este ajuste es `true` y los valores `null` se transfieren como `null`.

Si no necesita las garantías de privacidad de este ajuste y prefiere retener los valores `null` de sus conjuntos de datos, habilite el ajuste `preserveNulls` en el momento de crear la colaboración. El ajuste `preserveNulls` no se puede cambiar una vez creada la colaboración.

Datos de ejemplo para un fingerprint columna

El siguiente es un ejemplo de conjunto de datos de entrada y salida para un fingerprint columna con ajustes para reproducir. Otros ajustes a nivel de la colaboración, como `allowCleartext` y `allowDuplicates`, no afectan a los resultados y se pueden configurar como `true` o `false` si se intenta reproducirlos localmente.

Secreto compartido de ejemplo: `wJalrXUtnFEMI/K7MDENG/bPxRfiCYEXAMPLEKEY`

ID de colaboración de ejemplo: `a1b2c3d4-5678-90ab-cdef-EXAMPLE11111`

allowJoinsOnColumnsWithDifferentNames: `True` Este ajuste no afecta al rendimiento ni a los requisitos de almacenamiento. Sin embargo, este ajuste hace que la elección del nombre de la columna sea irrelevante al reproducir los valores que se muestran en las siguientes tablas.

Ejemplo 1

Input	`null`
`preserveNulls`	`TRUE`
Salida	`null`
Determinista	`Yes`
Bytes de entrada	0
Bytes de salida	0

Ejemplo 2

Entrada	`null`
`preserveNulls`	`FALSE`
Salida	`01:hmac:3lkFjthvV3IUu6mMvFc1a+XAHwgw/ElmOq4p3Yg25kk=`
Determinista	No
Bytes de entrada	0
Bytes de salida	52

Ejemplo 3

Input	`empty string`
`preserveNulls`	-
Salida	`01:hmac:oKTgi3Gba+eUb3JteSz2EMgXUkF1WgM77UP0Ydw5kPQ=`
Determinista	`Yes`
Bytes de entrada	0
Bytes de salida	52

Ejemplo 4

Input	`abcdefghijklmnopqrstuvwxyz`
`preserveNulls`	-
Salida	`01:hmac:kU/IqwG7FMmzzshr0B9scomE0UJUEE7j9keTctplGww=`
Determinista	`Yes`

Bytes de entrada	26
Bytes de salida	52

Ejemplo 5

Input	`abcdefghijklmnopqrstuvwxyzA BCDEFGHIJKLMNOPQRSTUVWXYZ01 23456789`
`preserveNulls`	-
Salida	`01:hmac:ks3htnQbw2vdhCRFF6J NzW5LMndJaHG57uvE26mBtSs=`
Determinista	`Yes`
Bytes de entrada	62
Bytes de salida	52

Solución de problemas fingerprint columns

¿Por qué está el texto cifrado en mi fingerprint columnas varias veces mayores que el tamaño del cleartext ¿qué contenía?

Texto cifrado en un fingerprint la columna siempre tiene 52 bytes de longitud. Si los datos de entrada eran pequeños (por ejemplo, las edades de los clientes), mostrarán un aumento considerable de tamaño. Esto también puede ocurrir si el ajuste `preserveNulls` se define como `false`.

¿Por qué está el texto cifrado en mi fingerprint columnas varias veces más pequeñas que el tamaño del cleartext ¿qué contenía?

Texto cifrado en un fingerprint la columna siempre tiene 52 bytes de longitud. Si los datos de entrada eran grandes (por ejemplo, las direcciones completas de los clientes), su tamaño se reducirá considerablemente.

¿Cómo sé si necesito las garantías criptográficas que ofrece **`preserveNulls`**?

Lamentablemente, la respuesta es que depende. Como mínimo, se deben revisar los the section called “Parámetros” en cuanto a la forma en que el ajuste `preserveNulls` protege sus datos. No obstante, le recomendamos que consulte los requisitos de manejo de datos de su organización y cualquier contrato aplicable a la colaboración correspondiente.

¿Por qué tengo que incurrir en la sobrecarga de base64?

Para permitir la compatibilidad con los formatos de archivo tabulares como CSV, es necesaria la codificación base64. Aunque algunos formatos de archivo como Parquet puede admitir representaciones binarias de datos, es importante que todos los participantes de una colaboración representen los datos de la misma manera para garantizar que los resultados de las consultas sean correctos.

Sealed columns

Sealed las columnas están diseñadas para ser utilizadas para transferir datos entre los miembros de una colaboración. El texto cifrado de estas columnas no es determinista y tiene un impacto significativo tanto en el rendimiento como en el almacenamiento en función de cómo se configuren las columnas. Estas columnas se pueden configurar de forma individual y, a menudo, son las que más influyen en el rendimiento del cliente de cifrado de C3R y en el tamaño del archivo de salida resultante.

Temas

- Gastos generales básicos para sealed columns
- Configuración de colaboración para sealed columns
- Configuración del esquema sealed columnas: tipos de relleno
- Datos de ejemplo para un sealed columna
- Solución de problemas sealed columns

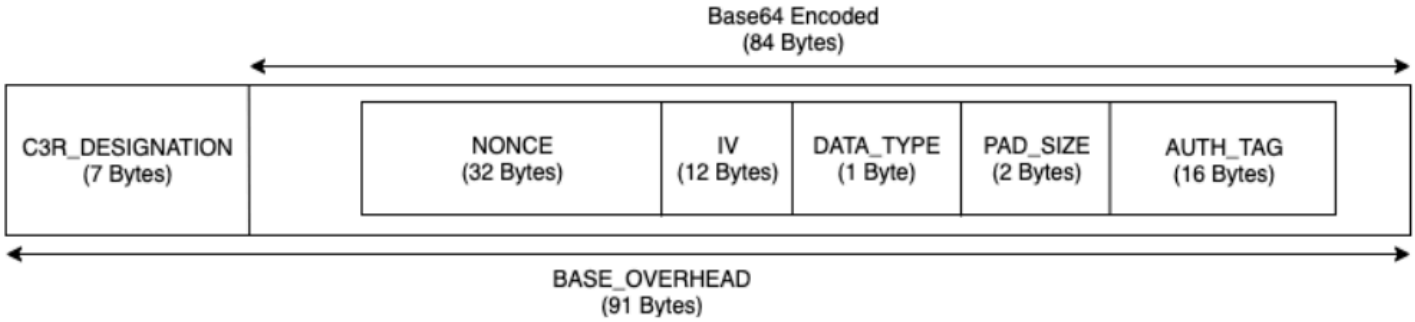
Gastos generales básicos para sealed columns

Hay una base superior para sealed columnas. Esta sobrecarga es constante y, además del tamaño del cleartext y rellenar bytes (si los hay).

Antes de cualquier cifrado, los datos del sealed las columnas van precedidas de un carácter de 1 byte que indica el tipo de datos que contienen. Si se selecciona el relleno, los datos se rellenan y se añaden 2 bytes que indican el tamaño del relleno. Una vez agregados estos bytes, los datos se procesan criptográficamente mediante el AES-GCM y se almacenan con IV (12 bytes), nonce (32

bytes) y Auth Tag (16 bytes). A continuación, estos datos se procesan a través de un codificador base64, lo que aumenta el tamaño del byte aproximadamente un 33 por ciento. Los datos van precedidos de una designación C3R de 7 bytes para indicar el tipo de columna a la que pertenecen los datos y la versión de cliente utilizada para generarlos. El resultado es una sobrecarga de base final de 91 bytes. A continuación, este resultado se puede multiplicar por el número de filas para obtener la sobrecarga de base total (utilice el número total de valores no nulos si `preserveNulls` está definido como true).

En la siguiente imagen se muestra cómo *`BASE_OVERHEAD = C3R_DESIGNATION + ((NONCE + IV + DATA_TYPE + PAD_SIZE + AUTH_TAG) * 1.33)`*



Configuración de colaboración para sealed columns

Ajuste **`preserveNulls`**

Cuando el ajuste en el nivel de la colaboración `preserveNulls` es `false` (predeterminado), cada valor `null` es único, aleatorio de 32 bytes, y se procesa como si no fuera `null`. El resultado es que cada valor `null` tiene ahora 91 bytes (más si se rellena). Esto puede añadir requisitos de almacenamiento importantes para las tablas que contienen datos muy dispersos en comparación con cuando este ajuste es `true` y los valores `null` se transfieren como `null`.

Si no necesita las garantías de privacidad de este ajuste y prefiere retener los valores `null` de sus conjuntos de datos, habilite el ajuste `preserveNulls` en el momento de crear la colaboración. El ajuste `preserveNulls` no se puede cambiar una vez creada la colaboración.

Configuración del esquema sealed columnas: tipos de relleno

Temas

- Tipo de relleno none
- Tipo de relleno fixed
- Tipo de relleno max

Tipo de relleno **none**

Al seleccionar un tipo de almohadilla de `none` no se añade ningún relleno a la cleartext y no añade ninguna sobrecarga adicional a la sobrecarga básica descrita anteriormente. La ausencia de relleno se traduce en el tamaño de salida más eficiente en términos de espacio. Sin embargo, no proporciona las mismas garantías de privacidad que los tipos de relleno `fixed` y `max`. Esto se debe al tamaño del subyacente cleartext es discernible por el tamaño del texto cifrado.

Tipo de relleno **fixed**

Seleccionar un tipo de relleno `fixed` es una medida que salvaguarda la privacidad para ocultar la longitud de los datos contenidos en una columna. Esto se hace rellenando todos los cleartext al proporcionado `pad_length` antes de cifrarlo. Cualquier dato que supere ese tamaño provoca un fallo del cliente de cifrado de C3R.

Dado que el relleno se añade al cleartext antes de cifrarlo, el AES-GCM tiene un mapeo 1 a 1 de cleartext para cifrar bytes de texto. La codificación base64 añadirá un 33 por ciento. La sobrecarga de almacenamiento adicional del relleno se puede calcular restando la longitud media del cleartext del valor de `pad_length` y multiplicándolo por 1,33. El resultado es la sobrecarga promedio de relleno por registro. Este resultado puede entonces multiplicarse por el número de filas para obtener la sobrecarga de relleno total (utilice el número total de valores no `null` si `preserveNulls` está definido como `true`).

*PADDING_OVERHEAD = (PAD_LENGTH - AVG_CLEARTEXT_LENGTH) * 1.33 * ROW_COUNT*

Se recomienda seleccionar la `pad_length` mínima que abarque el mayor valor de columna. Por ejemplo, si el mayor valor es de 50 bytes, basta con una `pad_length` de 50. Un valor superior a ese solo añadirá sobrecarga de almacenamiento adicional.

El relleno fijo no añade ninguna sobrecarga de computación significativa.

Tipo de relleno **max**

Seleccionar un tipo de relleno `max` es una medida que salvaguarda la privacidad para ocultar la longitud de los datos contenidos en una columna. Esto se hace rellenando todos los cleartext al valor más alto de la columna más el adicional `pad_length` antes de cifrarla. Por lo general, el `max` relleno proporciona las mismas garantías que el `fixed` relleno para un único conjunto de datos, al tiempo que permite no conocer el mayor cleartext valor de la columna. Sin embargo, es posible que el relleno `max` no ofrezca las mismas garantías de privacidad que el relleno `fixed` en todas las actualizaciones, ya que el valor más alto de cada conjuntos de datos podría variar.

Se recomienda seleccionar una `pad_length` adicional de 0 al usar el relleno `max`. Esta longitud rellena todos los valores para que tengan el mismo tamaño que el mayor valor de la columna. Un valor superior a ese solo añadirá sobrecarga de almacenamiento adicional.

Si es el más grande cleartext se conoce el valor de una columna determinada, le recomendamos que utilice el tipo `fixed` pad en su lugar. El uso del relleno `fixed` crea coherencia entre los conjuntos de datos actualizados. El uso del relleno `max` hace que cada subconjunto de datos se rellene hasta el mayor valor presente en el subconjunto.

Datos de ejemplo para un sealed columna

El siguiente es un ejemplo de conjunto de datos de entrada y salida para un sealed columna con ajustes para reproducir. Otros ajustes en el nivel de la colaboración, `allowCleartext`, `allowJoinsOnColumnsWithDifferentNames` y `allowDuplicates`, no afectan a los resultados y se pueden definir como `true` o `false` si se intentan reproducir localmente. Si bien estos son los ajustes básicos para reproducir, el sealed la columna no es determinista y los valores cambiarán cada vez. El objetivo es mostrar los bytes de entrada en comparación con los bytes de salida. Los valores `pad_length` de ejemplo se han elegido de forma intencionada. Muestran que el relleno `fixed` da como resultado los mismos valores que el relleno `max` con los ajustes de `pad_length` mínima recomendada o cuando se desea un relleno adicional.

Secreto compartido de ejemplo: `wJalrXUtnFEMI/K7MDENG/bPxRfiCYEXAMPLEKEY`

ID de colaboración de ejemplo: `a1b2c3d4-5678-90ab-cdef-EXAMPLE11111`

Temas

- Tipo de relleno none
- Tipo de relleno fixed (ejemplo 1)
- Tipo de relleno fixed (ejemplo 2)
- Tipo de relleno max (ejemplo 1)
- Tipo de relleno max (ejemplo 2)

Tipo de relleno **none**

Ejemplo 1

Input	`null`
`preserveNulls`	`TRUE`

Salida	`null`
Determinista	`Yes`
Bytes de entrada	0
Bytes de salida	0

Ejemplo 2

Entrada	`null`
`preserveNulls`	`FALSE`
Salida	`01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc4OTBqfRYZ98t5KU6aWfssGSPbNIJfG3iXmu6cbCUrizuV`
Determinista	No
Bytes de entrada	0
Bytes de salida	91

Ejemplo 3

Input	`empty string`
`preserveNulls`	-
Salida	`01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc4OTBqfRYZ98t5KU6aWfstGSPEM6qR8DWC2PB2GMlX41YK`
Determinista	No
Bytes de entrada	0

Bytes de salida	91

Ejemplo 4

Input	`abcdefghijklmnopqrstuvwxyz`
`preserveNulls`	-
Salida	`01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc4OTBqfRYZ98t5KU6aWfsteEE1GKEPiRzyh0h7t6OmWMLTWCvO2ckr6pkx9sGL5VLDQeHzh6DmPpyWNuI=`
Determinista	No
Bytes de entrada	26
Bytes de salida	127

Ejemplo 5

Input	`abcdefghijklmnopqrstuvwxyzABCDEFGHIJKLMNOPQRSTUVWXYZ0123456789`
`preserveNulls`	-
Salida	`01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc4OTBqfRYZ98t5KU6aWfsteEE1GKEPiRzyh0h7t6OmWMLTWCvO2ckr6plwtH/8tRFnn2rF91bcB9G4+n8GiRfJNmqdP4/QOQ3cXb/pbvPcnnohrHIGSX54ua+1/JfcVjc=`
Determinista	No

Bytes de entrada	62
Bytes de salida	175

Tipo de relleno **`fixed`** (ejemplo 1)

En este ejemplo, `pad_length` es 62 y la mayor entrada es de 62 bytes.

Ejemplo 1

Input	`null`
`preserveNulls`	`TRUE`
Salida	`null`
Determinista	`Yes`
Bytes de entrada	0
Bytes de salida	0

Ejemplo 2

Entrada	`null`
`preserveNulls`	`FALSE`
Salida	`01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfssGSNWfMRp7nSb7S MX2s3JKLOhK1+7r75Tk+Mx9jy48 Fcg1yOPvBqRSZ7oqy1V3UKfYTLEZb/ hCz7oaIneVsrcoNpATs0GzbnLkor4L+/ aSuA=`
Determinista	No
Bytes de entrada	0

Bytes de salida	175

Ejemplo 3

Input	`empty string`
`preserveNulls`	-
Salida	`01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc4OTBqfRYZ98t5KU6aWfstGSNWfMRp7nSb7SMX2s3JKLOhK1+7r75Tk+Mx9jy48Fcg1yOPvBqRSZ7oqy1V3UKfYTLEZb/hCz7oaIneVsrcoLB53l07VZpA6OwkuXu29CA=`
Determinista	No
Bytes de entrada	0
Bytes de salida	175

Ejemplo 4

Input	`abcdefghijklmnopqrstuvwxyz`
`preserveNulls`	-
Salida	`01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc4OTBqfRYZ98t5KU6aWfsteEE1GKEPiRzyh0h7t6OmWMLTWCvO2ckr6pkx9jy48Fcg1yOPvBqRSZ7oqy1V3UKfYTLEZb/hCz7oaIneVsrcutBAcO+Mb9tuU2KIHH31AWg=`
Determinista	No

Bytes de entrada	26
Bytes de salida	175

Ejemplo 5

Input	`abcdefghijklmnopqrstuvwxyzA BCDEFGHIJKLMNOPQRSTUVWXYZ01 23456789`
`preserveNulls`	-
Salida	`01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfsteEE1GKEPiRzyh0 h7t6OmWMLTWCvO2ckr6plwtH/8t RFnn2rF91bcB9G4+n8GiRfJNmqdP4/ QOQ3cXb/pbvPcnnohrHIGSX54ua+1/ JfcVjc=`
Determinista	No
Bytes de entrada	62
Bytes de salida	175

Tipo de relleno **`fixed`** (ejemplo 2)

En este ejemplo, `pad_length` es 162 y la mayor entrada es de 62 bytes.

Ejemplo 1

Input	`null`
`preserveNulls`	`TRUE`
Salida	`null`
Determinista	`Yes`

Bytes de entrada	0
Bytes de salida	0

Ejemplo 2

Entrada	`null`
`preserveNulls`	`FALSE`
Salida	`01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfssGSNWfMRp7nSb7S MX2s3JKLOhK1+7r75Tk+Mx9jy48 Fcg1yOPvBqRSZ7oqy1V3UKfYTLE Zb/hCz7oaIneVsrcnkB0xbLWD7z NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvtkSSVc2B80OOGp pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrmwv/xAySX+xcntotL703aBTBb`
Determinista	No
Bytes de entrada	0
Bytes de salida	307

Ejemplo 3

Input	`empty string`
`preserveNulls`	-
Salida	`01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfstGSNWfMRp7nSb7S MX2s3JKLOhK1+7r75Tk+Mx9jy48`

	`Fcg1yOPvBqRSZ7oqy1V3UKfYTLE Zb/hCz7oaIneVsrcnkB0xbLWD7z NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvtkSSVc2B80OOGp pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrmwv84lVaT9Yd+6oQx65/+gdVT`
Determinista	No
Bytes de entrada	0
Bytes de salida	307

Ejemplo 4

Input	`abcdefghijklmnopqrstuvwxyz`
`preserveNulls`	-
Salida	`01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfsteEE1GKEPiRzyh0 h7t6OmWMLTWCvO2ckr6pkx9jy48 Fcg1yOPvBqRSZ7oqy1V3UKfYTLE Zb/hCz7oaIneVsrcnkB0xbLWD7z NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvtkSSVc2B80OOGp pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrmwtX5Hnl+WyfO6ks3QMaRDGSf`
Determinista	No
Bytes de entrada	26
Bytes de salida	307

Ejemplo 5

Input	`abcdefghijklmnopqrstuvwxyzA BCDEFGHIJKLMNOPQRSTUVWXYZ01 23456789`
`preserveNulls`	-
Salida	`01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfsteEE1GKEPiRzyh0 h7t6OmWMLTWCvO2ckr6plwtH/8t RFnn2rF91bcB9G4+n8GiRfJNmqd P4/QOQ3cXb/pbvPcnkB0xbLWD7z NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvtkSSVc2B80OOGp pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrmwjkJXQZOgPdeFX9Yr/8alV5i``
Determinista	No
Bytes de entrada	62
Bytes de salida	307

Tipo de relleno **`max`** (ejemplo 1)

En este ejemplo, `pad_length` es 0 y la mayor entrada es de 62 bytes.

Ejemplo 1

Input	`null`
`preserveNulls`	`TRUE`
Salida	`null`
Determinista	`Yes`

Bytes de entrada	0
Bytes de salida	0

Ejemplo 2

Entrada	`null`
`preserveNulls`	`FALSE`
Salida	`01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc4OTBqfRYZ98t5KU6aWfssGSNWfMRp7nSb7SMX2s3JKLOhK1+7r75Tk+Mx9jy48Fcg1yOPvBqRSZ7oqy1V3UKfYTLEZb/hCz7oaIneVsrcoNpATs0GzbnLkor4L+/aSuA=`
Determinista	No
Bytes de entrada	0
Bytes de salida	175

Ejemplo 3

Input	`empty string`
`preserveNulls`	-
Salida	`01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc4OTBqfRYZ98t5KU6aWfstGSNWfMRp7nSb7SMX2s3JKLOhK1+7r75Tk+Mx9jy48Fcg1yOPvBqRSZ7oqy1V3UKfYTLEZb/hCz7oaIneVsrcoLB53l07VZpA6OwkuXu29CA=`

Determinista	No
Bytes de entrada	0
Bytes de salida	175

Ejemplo 4

Input	`abcdefghijklmnopqrstuvwxyz`
`preserveNulls`	-
Salida	`01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc4OTBqfRYZ98t5KU6aWfsteEE1GKEPiRzyh0h7t6OmWMLTWCvO2ckr6pkx9jy48Fcg1yOPvBqRSZ7oqy1V3UKfYTLEZb/hCz7oaIneVsrcutBAcO+Mb9tuU2KIHH31AWg=`
Determinista	No
Bytes de entrada	26
Bytes de salida	175

Ejemplo 5

Input	`abcdefghijklmnopqrstuvwxyzABCDEFGHIJKLMNOPQRSTUVWXYZ0123456789`
`preserveNulls`	-
Salida	`01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc4OTBqfRYZ98t5KU6aWfsteEE1GKEPiRzyh0h7t6OmWMLTWCvO2ckr6plwtH/8t`

	`RFnn2rF91bcB9G4+n8GiRfJNmqdP4/QOQ3cXb/pbvPcnnohrHIGSX54ua+1/JfcVjc=`
Determinista	No
Bytes de entrada	62
Bytes de salida	175

Tipo de relleno **`max`** (ejemplo 2)

En este ejemplo, `pad_length` es 100 y la mayor entrada es de 62 bytes.

Ejemplo 1

Input	`null`
`preserveNulls`	`TRUE`
Salida	`null`
Determinista	`Yes`
Bytes de entrada	0
Bytes de salida	0

Ejemplo 2

Entrada	`null`
`preserveNulls`	`FALSE`
Salida	`01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc4OTBqfRYZ98t5KU6aWfssGSNWfMRp7nSb7SMX2s3JKLOhK1+7r75Tk+Mx9jy48Fcg1yOPvBqRSZ7oqy1V3UKfYTLEZb/hCz7oaIneVsrcnkB0xbLWD7z`

	`NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvtkSSVc2B80OOGp pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrmwv/xAySX+xcntotL703aBTBb`
Determinista	No
Bytes de entrada	0
Bytes de salida	307

Ejemplo 3

Input	`empty string`
`preserveNulls`	-
Salida	`01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfstGSNWfMRp7nSb7S MX2s3JKLOhK1+7r75Tk+Mx9jy48 Fcg1yOPvBqRSZ7oqy1V3UKfYTLE Zb/hCz7oaIneVsrcnkB0xbLWD7z NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvtkSSVc2B80OOGp pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrmwv84lVaT9Yd+6oQx65/+gdVT`
Determinista	No
Bytes de entrada	0
Bytes de salida	307

Ejemplo 4

Input	`abcdefghijklmnopqrstuvwxyz`
`preserveNulls`	-
Salida	`01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfsteEE1GKEPiRzyh0 h7t6OmWMLTWCvO2ckr6pkx9jy48 Fcg1yOPvBqRSZ7oqy1V3UKfYTLE Zb/hCz7oaIneVsrcnkB0xbLWD7z NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvtkSSVc2B80OOGp pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrmwtX5Hnl+WyfO6ks3QMaRDGSf`
Determinista	No
Bytes de entrada	26
Bytes de salida	307

Ejemplo 5

Input	`abcdefghijklmnopqrstuvwxyzA BCDEFGHIJKLMNOPQRSTUVWXYZ01 23456789`
`preserveNulls`	-
Salida	`01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfsteEE1GKEPiRzyh0 h7t6OmWMLTWCvO2ckr6plwtH/8t RFnn2rF91bcB9G4+n8GiRfJNmqd P4/QOQ3cXb/pbvPcnkB0xbLWD7z`

	`NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvtkSSVc2B80OOGp pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrmwjkJXQZOgPdeFX9Yr/8alV5i`
Determinista	No
Bytes de entrada	62
Bytes de salida	307

Solución de problemas sealed columns

¿Por qué está el texto cifrado en mi sealed columnas varias veces mayores que el tamaño del cleartext ¿qué contenía?

Esto depende de varios factores. Por un lado, el texto cifrado en un Cleartext la columna siempre tiene una longitud mínima de 91 bytes. Si los datos de entrada eran pequeños (por ejemplo, las edades de los clientes), mostrarán un aumento considerable de tamaño. En segundo lugar, si `preserveNulls` se hubiera definido como `false` y los datos de entrada contuvieran un gran número de valores `null`, cada uno de esos valores `null` se habrá convertido en 91 bytes de texto cifrado. Por último, si utiliza el relleno, por definición, los bytes se añaden al cleartext datos antes de cifrarlos.

La mayoría de mis datos están en un sealed la columna es muy pequeña y necesito usar relleno. ¿Puedo simplemente eliminar los valores grandes y procesarlos por separado para ahorrar espacio?

No le recomendamos que elimine los valores grandes y los procese por separado. Al hacerlo, se modifican las garantías de privacidad que ofrece el cliente de cifrado de C3R. Como modelo de amenaza, presuponga que un observador puede ver ambos conjuntos de datos cifrados. Si el observador ve que un subconjunto de datos tiene una columna considerablemente más o menos rellena que otro subconjunto, puede hacer inferencias sobre el tamaño de los datos de cada subconjunto. Por ejemplo, imaginemos que una columna `fullName` se rellena hasta un total de 40 bytes en un archivo y se rellena hasta 800 bytes en otro archivo. Un observador podría deducir que un conjunto de datos contiene el nombre más largo del mundo (747 bytes).

¿Debo proporcionar un relleno adicional al usar el tipo de relleno **`max`**?

No. Al utilizar el relleno `max`, se recomienda establecer en 0 la `pad_length` (también conocida como relleno adicional complementario al mayor valor de la columna).

¿Puedo elegir una **`pad_length`** grande al usar el relleno **`fixed`** para evitar tener que preocuparme de que quepa el valor más grande?

Sí, pero la gran longitud de relleno es ineficiente y utiliza más almacenamiento del necesario. Le recomendamos que compruebe el tamaño del mayor valor y que defina la `pad_length` con ese valor.

¿Cómo sé si necesito las garantías criptográficas que ofrece **`preserveNulls`**?

Lamentablemente, la respuesta es que depende. Como mínimo, se deben revisar los Computación criptográfica para Clean Rooms en cuanto a la forma en que el ajuste `preserveNulls` protege sus datos. No obstante, le recomendamos que consulte los requisitos de manejo de datos de su organización y cualquier contrato aplicable a la colaboración correspondiente.

¿Por qué tengo que incurrir en la sobrecarga de base64?

Para permitir la compatibilidad con los formatos de archivo tabulares, como CSV, es necesaria la codificación base64. Aunque algunos formatos de archivo como Parquet puede admitir representaciones binarias de datos, es importante que todos los participantes de una colaboración representen los datos de la misma manera para garantizar que los resultados de las consultas sean correctos.

Solución de problemas relacionados con el aumento imprevisto de tamaño del texto cifrado

Supongamos que ha cifrado los datos y que el tamaño de los datos resultantes es sorprendentemente grande. Los siguientes pasos pueden ayudarle a identificar dónde se produjo el aumento de tamaño y qué medidas puede adoptar, si las hubiera.

Identificar dónde se produjo el aumento de tamaño

Antes de poder solucionar el problema por el que los datos cifrados son significativamente más grandes que los cleartext datos, primero debe identificar dónde está el aumento de tamaño. Cleartext las columnas se pueden ignorar de forma segura porque no se modifican. Observe los restantes fingerprint y sealed columnas y elige una que parezca significativa.

Identificar el motivo por el que se produjo el aumento de tamaño

A fingerprint columna o una sealed la columna podría contribuir al aumento de tamaño.

Temas

- ¿El aumento de tamaño proviene de un fingerprint ¿columna?
- ¿El aumento de tamaño proviene de un sealed ¿columna?

¿El aumento de tamaño proviene de un fingerprint ¿columna?

Si la columna que más contribuye al aumento del almacenamiento es una fingerprint columna, es probable que esto se deba a que cleartext los datos son pequeños (por ejemplo, la edad del cliente). Cada uno resultante fingerprint el texto cifrado tiene una longitud de 52 bytes. Lamentablemente, no se puede hacer nada al respecto sobre una column-by-column base sólida. Para obtener más información, consulte Superficie base para fingerprint columns para conocer los detalles de esta columna, incluida su incidencia en los requisitos de almacenamiento.

La otra posible causa del aumento de tamaño es fingerprint la columna es la configuración de colaboración,`preserveNulls`. Si la configuración de colaboración `preserveNulls` está deshabilitada (la configuración predeterminada), todos los `null` valores de fingerprint las columnas se convertirán en 52 bytes de texto cifrado. No hay nada que se pueda hacer al respecto en la colaboración actual. El ajuste `preserveNulls` se define en el momento en que se crea la colaboración, y todos los colaboradores deben usar el mismo ajuste para garantizar que los resultados de la consulta sean correctos. Para obtener más información sobre el ajuste `preserveNulls` y sobre cómo su habilitación afecta a las garantías de privacidad de los datos, consulte the section called “Computación criptográfica”.

¿El aumento de tamaño proviene de un sealed ¿columna?

Si la columna que más contribuye al aumento del almacenamiento es una sealed En esta columna, hay algunos detalles que podrían contribuir al aumento de tamaño.

Si el archivo de cleartext los datos son pequeños (por ejemplo, la edad del cliente), y cada uno de ellos resulta sealed el texto cifrado tiene una longitud mínima de 91 bytes. Lamentablemente, no se puede hacer nada al respecto. Para obtener más información, consulte Gastos generales básicos para sealed columns para conocer los detalles de esta columna, incluida su incidencia en los requisitos de almacenamiento.

La segunda causa principal del aumento de almacenamiento en sealed las columnas se rellenan. El relleno añade bytes adicionales al cleartext antes de cifrarlo para ocultar el tamaño de los valores individuales de un conjunto de datos. Se recomienda establecer el relleno en el valor mínimo posible para el conjunto de datos. Como mínimo, se debe definir `pad_length` para el relleno `fixed` de

manera que abarque el mayor valor posible de la columna. Todo ajuste superior a ese no aportará garantías de privacidad adicionales. Por ejemplo, si sabe que el mayor valor posible de una columna puede ser de 50 bytes, le recomendamos que defina la `pad_length` en 50 bytes. Sin embargo, si el sealed la columna está `max` rellenada, le recomendamos que la establezca en `pad_length` 0 bytes. Esto se debe a que el relleno `max` se refiere al relleno adicional que se suma al mayor valor de la columna.

La última causa posible del aumento de tamaño en un sealed la columna es la configuración de colaboración,`preserveNulls`. Si la configuración de colaboración `preserveNulls` está deshabilitada (la configuración predeterminada), todos los `null` valores de sealed las columnas se convertirán en 91 bytes de texto cifrado. No hay nada que se pueda hacer al respecto en la colaboración actual. El ajuste `preserveNulls` se define en el momento en que se crea la colaboración, y todos los colaboradores deben usar el mismo ajuste para garantizar que los resultados de la consulta sean correctos. Para obtener más información sobre los efectos de esta configuración y sobre cómo su activación afecta a las garantías de privacidad de sus datos, consulte the section called “Computación criptográfica”.

Con AWS Clean Rooms figuración

En los temas siguientes se explica cómo configurarlos AWS Clean Rooms.

Temas

- Inscríbase en AWS
- Configure las funciones de servicio para AWS Clean Rooms
- Configure las funciones de servicio para AWS Clean Rooms el aprendizaje automático

Inscríbase en AWS

Antes de poder usar AWS Clean Rooms, o cualquier otro Servicio de AWS, debes registrarte AWS con un Cuenta de AWS.

Si no tienes uno Cuenta de AWS, sigue los siguientes pasos para crearlo.

Para suscribirte a una Cuenta de AWS

1. Abrir https://portal.aws.amazon.com/billing/registro.
2. Siga las instrucciones que se le indiquen.

   Durante el proceso de registro, recibirás una llamada telefónica con un código de verificación que deberás introducir en el teclado del teléfono.
3. Cuando te registras en un Cuenta de AWS, se crea un usuario Cuenta de AWS root. El usuario raíz tendrá acceso a todos los Servicios de AWS y recursos de esa cuenta. Como práctica recomendada de seguridad, asigne acceso administrativo a un usuario administrativo y utilice únicamente el usuario raíz para realizar tareas que requieran acceso de usuario raíz.

Configure las funciones de servicio para AWS Clean Rooms

En las siguientes secciones se describen las funciones necesarias para realizar cada tarea.

Temas

- Creación de un usuario administrador
- Creación de un rol de IAM para un miembro de la colaboración
- Cree un rol de servicio para leer datos de Amazon S3

- Cree un rol de servicio para leer los datos de Amazon Athena
- Cree un rol de servicio para leer los datos de Snowflake
- Cree un rol de servicio para leer el código de un bucket de S3 (rol de plantilla de PySpark análisis)
- Cree un rol de servicio para escribir los resultados de un PySpark trabajo
- Creación de un rol de servicio para recibir resultados

Creación de un usuario administrador

Para AWS Clean Rooms utilizarlos, debe crear un usuario administrador para usted y añadir el usuario administrador a un grupo de administradores.

Para crear un usuario administrador, elija una de las siguientes opciones.

Elegir una forma de administrar el administrador	Para	Haga esto	También puede
En IAM Identity Center (recomendado)	Usar credenciales a corto plazo para acceder a AWS. Esto se ajusta a las prácticas recomendadas de seguridad. Para obtener información sobre las prácticas recomendadas, consulta Prácticas recomendadas de seguridad en IAM en	Siga las instrucciones en Introducción en la Guía del usuario de AWS IAM Identity Center .	Configure el acceso mediante programación configurando el AWS CLI que se utilizará AWS IAM Identity Center en la Guía del AWS Command Line Interface usuario.

Elegir una forma de administrar el administrador	Para	Haga esto	También puede
	la Guía del usuario de IAM.		
En IAM (no recomendado)	Usar credenciales a largo plazo para acceder a AWS.	Siguiendo las instrucciones de Crear un usuario de IAM para acceso de emergencia de la Guía del usuario de IAM.	Configure el acceso programático mediante Administrar las claves de acceso de los usuarios de IAM en la Guía del usuario de IAM.

Creación de un rol de IAM para un miembro de la colaboración

Un miembro es un AWS cliente que participa en una colaboración.

Para crear un rol de IAM para un miembro de la colaboración

1. Siga el procedimiento Creación de un rol para delegar permisos a un usuario de IAM de la Guía del usuario de AWS Identity and Access Management .
2. Para el paso Creación de política, seleccione la pestaña JSON en el Editor de políticas y, a continuación, agregue políticas en función de las capacidades otorgadas al miembro de la colaboración.

   AWS Clean Rooms ofrece las siguientes políticas administradas basadas en casos de uso comunes.

Si desea…	Entonces use...
Ver los recursos y metadatos	AWS política gestionada: `AWSCleanRoomsReadOnlyAccess`

Si desea…	Entonces use...
Consultar	AWS política gestionada: AWSCleanRoomsFullAccess
Consulta y ejecuta trabajos	AWS política gestionada: AWSCleanRoomsFullAccess
Consultar y recibir resultados	AWS política gestionada: AWSCleanRoomsFullAccess
Administración de los recursos de colaboración, pero no consultas	AWS política gestionada: AWSCleanRoomsFullAccessNoQuerying

Para obtener información sobre las diferentes políticas gestionadas que ofrecen AWS Clean Rooms, consulteAWS políticas gestionadas para AWS Clean Rooms,

Cree un rol de servicio para leer datos de Amazon S3

AWS Clean Rooms utiliza un rol de servicio para leer los datos de Amazon S3.

Hay dos formas de crear este rol de servicio.

- Si tiene los permisos de IAM necesarios para crear un rol de servicio, utilice la AWS Clean Rooms consola para crear un rol de servicio.
- Si no tiene `iam:CreateRole` `iam:AttachRolePolicy` permisos o desea crear `iam:CreatePolicy` los roles de IAM manualmente, realice una de las siguientes acciones:
  - Utilice el siguiente procedimiento para crear un rol de servicio mediante políticas de confianza personalizadas.
  - Pídale al administrador que cree el rol de servicio mediante el siguiente procedimiento.

> **Note**
>
> Usted o su administrador de IAM deben seguir este procedimiento únicamente si no tienen los permisos necesarios para crear un rol de servicio mediante la AWS Clean Rooms consola.

Para crear un rol de servicio para leer datos de Amazon S3 mediante políticas de confianza personalizadas

1. Cree un rol mediante políticas de confianza personalizadas. Para obtener más información, consulte el procedimiento Creación de un rol mediante políticas de confianza personalizadas (consola) en la Guía del AWS Identity and Access Management usuario.
2. Utilice la siguiente política de confianza personalizada de acuerdo con el procedimiento Creación de un rol mediante políticas de confianza personalizadas (consola).

> **Note**
>
> Si quiere asegurarse de que el rol se utilice únicamente en el contexto de una determinada membresía de colaboración, puede ampliar la política de confianza. Para obtener más información, consulte Prevención de la sustitución confusa entre servicios.

```
{

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

3. Utilice la siguiente política de permisos de acuerdo con el procedimiento Creación de un rol mediante políticas de confianza personalizadas (consola).

> **Note**
>
> La siguiente política de ejemplo admite los permisos necesarios para leer los metadatos de AWS Glue y los datos de Amazon S3 correspondientes. Sin embargo, es posible que tengas que modificar esta política en función de cómo hayas configurado los datos de Amazon S3. Por ejemplo, si ha configurado una clave de KMS personalizada para

sus datos de Amazon S3, es posible que deba modificar esta política con permisos adicionales AWS Key Management Service (AWS KMS).
Sus AWS Glue recursos y los recursos subyacentes de Amazon S3 deben estar en la Región de AWS misma posición que los de la AWS Clean Rooms colaboración.

```
{

    "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:aws-region:accountId:database/databaseName",
                "arn:aws:glue:aws-region:accountId:table/databaseName/tableName",
                "arn:aws:glue:aws-region:accountId: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":[
                        "s3BucketOwnerAccountId"
                    ]
                }
            }
        },
        {
            "Sid": "NecessaryS3ObjectPermissions",
            "Effect": "Allow",
            "Action": [
                "s3:GetObject"
            ],
            "Resource": [
                "arn:aws:s3:::bucket/prefix/*"
            ],
            "Condition":{
                "StringEquals":{
                    "s3:ResourceAccount":[
                        "s3BucketOwnerAccountId"
                    ]
                }
            }
        }
    ]
}
```

4. Reemplace cada *placeholder* por su propia información.
5. Siga el procedimiento Creación de un rol mediante políticas de confianza personalizadas (consola) para crear el rol.

Cree un rol de servicio para leer los datos de Amazon Athena

AWS Clean Rooms utiliza un rol de servicio para leer los datos de Amazon Athena.

Para crear un rol de servicio para leer datos de Athena mediante políticas de confianza personalizadas

1. Cree un rol mediante políticas de confianza personalizadas. Para obtener más información, consulte el procedimiento Creación de un rol mediante políticas de confianza personalizadas (consola) en la Guía del AWS Identity and Access Management usuario.
2. Utilice la siguiente política de confianza personalizada de acuerdo con el procedimiento Creación de un rol mediante políticas de confianza personalizadas (consola).

> **Note**
>
> Si quiere asegurarse de que el rol se utilice únicamente en el contexto de una determinada membresía de colaboración, puede ampliar la política de confianza. Para obtener más información, consulte Prevención de la sustitución confusa entre servicios.

```
{

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

3. Utilice la siguiente política de permisos de acuerdo con el procedimiento Creación de un rol mediante políticas de confianza personalizadas (consola).

> **Note**
>
> El siguiente ejemplo de política admite los permisos necesarios para leer AWS Glue los metadatos y sus datos de Athena correspondientes. Sin embargo, es posible que tengas que modificar esta política en función de cómo hayas configurado los datos de Amazon S3. Por ejemplo, si ya configuró una clave de KMS personalizada para

sus datos de Amazon S3, es posible que deba modificar esta política con AWS KMS permisos adicionales.
Sus AWS Glue recursos y los recursos subyacentes de Athena deben ser los mismos que los de la Región de AWS AWS Clean Rooms colaboración.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "athena:GetDataCatalog",
                "athena:GetWorkGroup",
                "athena:GetTableMetadata",
                "athena:GetQueryExecution",
                "athena:GetQueryResults",
                "athena:StartQueryExecution"
            ],
            "Resource": [
                "arn:aws:athena:region:accountId:workgroup/workgroup",
                "arn:aws:athena:region:accountId:datacatalog/AwsDataCatalog"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "glue:GetDatabase",
                "glue:GetTable",
                "glue:GetPartitions"
            ],
            "Resource": [
                "arn:aws:glue:region:accountId:catalog",
                "arn:aws:glue:region:accountId:database/database name",
                "arn:aws:glue:region:accountId:table/database name/table name"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "s3:GetObject",
                "s3:GetBucketLocation",
```

```
                "s3:AbortMultipartUpload",
                "s3:ListBucket",
                "s3:PutObject",
                "s3:ListMultipartUploadParts"
            ],
            "Resource": [
                "arn:aws:s3:::bucket",
                "arn:aws:s3:::bucket/*"
            ]
        },
        {
            "Effect": "Allow",
            "Action": "lakeformation:GetDataAccess",
            "Resource": "*"
        },
        {
            "Effect": "Allow",
            "Action": [
                "kms:GenerateDataKey",
                "kms:Decrypt"
            ],
            "Resource": "arn:aws:kms:region:accountId:key/*"
        }
    ]
}
```

4. Reemplace cada *placeholder* por su propia información.
5. Siga el procedimiento Creación de un rol mediante políticas de confianza personalizadas (consola) para crear el rol.

Configurar permisos de Lake Formation

Si consulta recursos protegidos con permisos de Lake Formation, el rol de servicio debe tener permisos de acceso de selección table/view y descripción en la AWS Glue base de datos en la que está almacenada la vista.

Para obtener más información, consulte:

- Utilice Athena para consultar los datos registrados AWS Lake Formation en la Guía del usuario de Amazon Athena

- Cómo incorporar los permisos de Lake Formation en la Guía AWS Lake Formation para desarrolladores

Cree un rol de servicio para leer los datos de Snowflake

AWS Clean Rooms utiliza un rol de servicio para recuperar las credenciales para que Snowflake lea los datos de esta fuente.

Hay dos formas de crear este rol de servicio:

- Si tiene los permisos de IAM necesarios para crear un rol de servicio, utilice la AWS Clean Rooms consola para crear un rol de servicio.
- Si no tiene `iam:CreateRole` `iam:AttachRolePolicy` permisos o desea crear `iam:CreatePolicy` los roles de IAM manualmente, realice una de las siguientes acciones:
  - Utilice el siguiente procedimiento para crear un rol de servicio mediante políticas de confianza personalizadas.
  - Pídale al administrador que cree el rol de servicio mediante el siguiente procedimiento.

> **Note**
>
> Usted o su administrador de IAM deben seguir este procedimiento únicamente si no tienen los permisos necesarios para crear un rol de servicio mediante la AWS Clean Rooms consola.

Para crear un rol de servicio para leer datos de Snowflake mediante políticas de confianza personalizadas

1. Cree un rol mediante políticas de confianza personalizadas. Para obtener más información, consulte el procedimiento Creación de un rol mediante políticas de confianza personalizadas (consola) en la Guía del AWS Identity and Access Management usuario.
2. Utilice la siguiente política de confianza personalizada de acuerdo con el procedimiento Creación de un rol mediante políticas de confianza personalizadas (consola).

> **Note**
>
> Si quiere asegurarse de que el rol se utilice únicamente en el contexto de una determinada membresía de colaboración, puede ampliar la política de confianza. Para obtener más información, consulte Prevención de la sustitución confusa entre servicios.

```
{
    "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:region:accountId:membership/membershipId",

 "arn:aws:cleanrooms:region:queryRunnerAccountId:membership/
queryRunnerMembershipId"
                    ]
                }
            }
        }
    ]
}
```

3. Utilice una de las siguientes políticas de permisos según el procedimiento de creación de un rol mediante políticas de confianza personalizadas (consola).

   Política de permisos para los secretos cifrados con una clave KMS propiedad del cliente

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

```
            "Action": "secretsmanager:GetSecretValue",
            "Resource":
 "arn:aws:secretsmanager:region:secretAccountId:secret:secretIdentifier",
            "Effect": "Allow"
        },
        {
            "Sid": "AllowDecryptViaSecretsManagerForKey",
            "Action": "kms:Decrypt",
            "Resource": "arn:aws:kms:region:keyOwnerAccountId:key/keyIdentifier",
            "Effect": "Allow",
            "Condition": {
                "StringEquals": {
                    "kms:ViaService": "secretsmanager.region.amazonaws.com",
                    "kms:EncryptionContext:SecretARN":
 "arn:aws:secretsmanager:region:secretAccountId:secret:secretIdentifier"
                }
            }
        }
    ]
}
```

Política de permisos para los secretos cifrados con una Clave administrada de AWS

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": "secretsmanager:GetSecretValue",
            "Resource":
 "arn:aws:secretsmanager:region:accountId:secret:secretIdentifier",
            "Effect": "Allow"
        }
    ]
}
```

4. Reemplace cada *placeholder* por su propia información.
5. Siga el procedimiento Creación de un rol mediante políticas de confianza personalizadas (consola) para crear el rol.

Cree un rol de servicio para leer el código de un bucket de S3 (rol de plantilla de PySpark análisis)

AWS Clean Rooms usa un rol de servicio para leer el código del bucket de S3 especificado por un miembro de la colaboración cuando usa una plantilla de PySpark análisis.

Para crear un rol de servicio para leer el código de un bucket de S3

1. Cree un rol mediante políticas de confianza personalizadas. Para obtener más información, consulte el procedimiento Creación de un rol mediante políticas de confianza personalizadas (consola) en la Guía del AWS Identity and Access Management usuario.
2. Utilice la siguiente política de confianza personalizada de acuerdo con el procedimiento Creación de un rol mediante políticas de confianza personalizadas (consola).

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "Service": "cleanrooms.amazonaws.com"
            },
            "Action": "sts:AssumeRole",
            "Condition": {
                "ForAnyValue:ArnEquals": {
                    "aws:SourceArn": [

 "arn:aws:cleanrooms:region:jobRunnerAccountId:membership/jobRunnerMembershipId",

 "arn:aws:cleanrooms:region:analysisTemplateAccountId:membership/
analysisTemplateOwnerMembershipId"
                    ]
                }
            }
        }
    ]
}
```

3. Utilice la siguiente política de permisos de acuerdo con el procedimiento Creación de un rol mediante políticas de confianza personalizadas (consola).

> **Note**
>
> El siguiente ejemplo de política admite los permisos necesarios para leer el código de Amazon S3. No obstante, quizás tenga que modificar esta política en función de cómo haya configurado los datos de S3.
> Sus recursos de Amazon S3 deben estar en los Región de AWS mismos que los de la AWS Clean Rooms colaboración.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "s3:GetObject",
                "s3:GetObjectVersion"
            ],
            "Resource": ["arn:aws:s3:::s3Path"],
            "Condition":{
                "StringEquals":{
                    "s3:ResourceAccount":[
                        "s3BucketOwnerAccountId"
                    ]
                }
            }
        }
    ]
}
```

4. Sustituya cada uno *placeholder* por su propia información:

   - *s3Path*— La ubicación del código en el depósito S3.
   - *s3BucketOwnerAccountId*— El Cuenta de AWS ID del propietario del bucket de S3.
   - *region*: nombre del Región de AWS. Por ejemplo, **us-east-1**.
   - *jobRunnerAccountId*— El Cuenta de AWS ID del miembro que puede ejecutar consultas y trabajos.
   - *jobRunnerMembershipId*— El ID de membresía del miembro que puede consultar y ejecutar trabajos. Puede encontrar el ID de pertenencia en la pestaña Detalles de la

colaboración. Esto garantiza AWS Clean Rooms que solo asuma el rol cuando este miembro ejecute el análisis en esta colaboración.

- *analysisTemplateAccountId*— El Cuenta de AWS ID de la plantilla de análisis.
- *analysisTemplateOwnerMembershipId*— El ID de membresía del miembro propietario de la plantilla de análisis. Puede encontrar el ID de pertenencia en la pestaña Detalles de la colaboración.

5. Siga el procedimiento Creación de un rol mediante políticas de confianza personalizadas (consola) para crear el rol.

Cree un rol de servicio para escribir los resultados de un PySpark trabajo

AWS Clean Rooms usa un rol de servicio para escribir los resultados de un PySpark trabajo en un bucket de S3 específico.

Para crear un rol de servicio para escribir los resultados de un PySpark trabajo

1. Cree un rol mediante políticas de confianza personalizadas. Para obtener más información, consulte el procedimiento Creación de un rol mediante políticas de confianza personalizadas (consola) en la Guía del AWS Identity and Access Management usuario.
2. Utilice la siguiente política de confianza personalizada de acuerdo con el procedimiento Creación de un rol mediante políticas de confianza personalizadas (consola).

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "Service": "cleanrooms.amazonaws.com"
            },
            "Action": "sts:AssumeRole",
            "Condition": {
                "ForAnyValue:ArnEquals": {
                    "aws:SourceArn": [

 "arn:aws:cleanrooms:region:jobRunnerAccountId:membership/jobRunnerMembershipId",

 "arn:aws:cleanrooms:region:rrAccountId:membership/rrMembershipId"
                    ]
```

```
                }
            }
        }
    ]
}
```

3. Utilice la siguiente política de permisos de acuerdo con el procedimiento Creación de un rol mediante políticas de confianza personalizadas (consola).

> **Note**
>
> El siguiente ejemplo de política admite los permisos necesarios para escribir en Amazon S3. Sin embargo, es posible que deba modificar esta política en función de cómo haya configurado S3.
> Sus recursos de Amazon S3 deben estar en los Región de AWS mismos que los de la AWS Clean Rooms colaboración.

```
{
    "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. Sustituya cada uno *placeholder* por su propia información:

   - *region*: nombre del Región de AWS. Por ejemplo, **us-east-1**.
   - *jobRunnerAccountId*— El Cuenta de AWS ID en el que se encuentra el depósito S3.
   - *jobRunnerMembershipId*— El ID de membresía del miembro que puede consultar y ejecutar trabajos. Puede encontrar el ID de pertenencia en la pestaña Detalles de la colaboración. Esto garantiza AWS Clean Rooms que solo asuma el rol cuando este miembro ejecute el análisis en esta colaboración.
   - *rrAccountId*— El Cuenta de AWS ID en el que se encuentra el depósito de S3.
   - *rrMembershipId*— El ID de membresía del miembro que puede recibir los resultados. Puede encontrar el ID de pertenencia en la pestaña Detalles de la colaboración. Esto garantiza AWS Clean Rooms que solo asuma el rol cuando este miembro ejecute el análisis en esta colaboración.
   - *bucket*— El nombre y la ubicación del depósito de S3.
   - *optionalPrefix*— Un prefijo opcional si desea guardar los resultados con un prefijo S3 específico.
   - *s3BucketOwnerAccountId*— El Cuenta de AWS ID del propietario del bucket de S3.

5. Siga el procedimiento Creación de un rol mediante políticas de confianza personalizadas (consola) para crear el rol.

Creación de un rol de servicio para recibir resultados

> **Note**
>
> Si es el único miembro que solo puede recibir resultados (en la consola, sus destrezas como miembro solo son Recibir resultados), siga este procedimiento.

Si es un miembro que puede realizar consultas y recibir resultados (en la consola, Sus destrezas como miembro son Consultar y Recibir resultados), puede omitir este procedimiento.

Para los miembros de la colaboración que solo pueden recibir resultados, AWS Clean Rooms utiliza un rol de servicio para escribir los resultados de los datos consultados en la colaboración en el bucket de S3 especificado.

Hay dos formas de crear este rol de servicio:

- Si tiene los permisos de IAM necesarios para crear un rol de servicio, utilice la AWS Clean Rooms consola para crear un rol de servicio.
- Si no tiene `iam:CreateRole` `iam:AttachRolePolicy` permisos o desea crear `iam:CreatePolicy` los roles de IAM manualmente, realice una de las siguientes acciones:
  - Utilice el siguiente procedimiento para crear un rol de servicio mediante políticas de confianza personalizadas.
  - Pídale al administrador que cree el rol de servicio mediante el siguiente procedimiento.

Note

Usted o su administrador de IAM deben seguir este procedimiento únicamente si no tienen los permisos necesarios para crear un rol de servicio mediante la AWS Clean Rooms consola.

Para crear un rol de servicio para recibir resultados mediante políticas de confianza personalizadas

1. Cree un rol mediante políticas de confianza personalizadas. Para obtener más información, consulte el procedimiento Creación de un rol mediante políticas de confianza personalizadas (consola) en la Guía del AWS Identity and Access Management usuario.
2. Utilice la siguiente política de confianza personalizada de acuerdo con el procedimiento Creación de un rol mediante políticas de confianza personalizadas (consola).

```
{
    "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. Utilice la siguiente política de permisos de acuerdo con el procedimiento Creación de un rol mediante políticas de confianza personalizadas (consola).

> **Note**
>
> La siguiente política de ejemplo admite los permisos necesarios para leer los metadatos de AWS Glue y los datos de Amazon S3 correspondientes. No obstante, quizás tenga que modificar esta política en función de cómo haya configurado los datos de S3.

Sus AWS Glue recursos y los recursos subyacentes de Amazon S3 deben estar en la Región de AWS misma posición que los de la AWS Clean Rooms colaboración.

```
{

    "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. Sustituya cada uno *placeholder* por su propia información:

- *region*: nombre del Región de AWS. Por ejemplo, **us-east-1**.
- *a1b2c3d4-5678-90ab-cdef-EXAMPLEaaaaa*— El ID de membresía del miembro que puede realizar la consulta. Puede encontrar el ID de pertenencia en la pestaña Detalles de la colaboración. Esto garantiza AWS Clean Rooms que solo asuma el rol cuando este miembro ejecute el análisis en esta colaboración.
- *arn:aws:cleanrooms:us-east-1:555555555555:membership/a1b2c3d4-5678-90ab-cdef-EXAMPLEaaaaa*— El ARN de membresía único del miembro que puede realizar la consulta. Puede encontrar el ARN de la pertenencia en la pestaña Detalles de la colaboración. Esto garantiza AWS Clean Rooms que solo asuma el rol cuando este miembro ejecute el análisis en esta colaboración.
- *bucket_name*— El nombre de recurso de Amazon (ARN) del bucket de S3. Puede encontrar el Nombre de recurso de Amazon (ARN) en la pestaña Propiedades del bucket en Amazon S3.
- *accountId*— El Cuenta de AWS ID en el que se encuentra el bucket de S3.

  bucket_name/optional_key_prefix— El nombre del recurso de Amazon (ARN) del destino de los resultados en Amazon S3. Puede encontrar el Nombre de recurso de Amazon (ARN) en la pestaña Propiedades del bucket en Amazon S3.

5. Siga el procedimiento Creación de un rol mediante políticas de confianza personalizadas (consola) para crear el rol.

Configure las funciones de servicio para AWS Clean Rooms el aprendizaje automático

Las funciones necesarias para realizar un modelado similar difieren de las necesarias para utilizar un modelo personalizado. En las siguientes secciones se describen las funciones necesarias para realizar cada tarea.

Temas

- Configure las funciones de servicio para un modelado similar
- Configure las funciones de servicio para el modelado personalizado

Configure las funciones de servicio para un modelado similar

Temas

- Creación de rol de servicio para leer datos de entrenamiento
- Creación de un rol de servicio para escribir un segmento similar
- Creación de rol de servicio para leer datos iniciales

Creación de rol de servicio para leer datos de entrenamiento

AWS Clean Rooms usa un rol de servicio para leer los datos de entrenamiento. Puede crear este rol mediante la consola si dispone de los permisos de IAM necesarios. Si no tiene permisos `CreateRole`, pida al administrador que cree el rol de servicio.

Para crear un rol de servicio para entrenar un conjunto de datos

1. Inicie sesión en la consola de IAM (https://console.aws.amazon.com/iam/) con su cuenta de administrador.
2. En Administración de acceso, seleccione Políticas.
3. Elija Create Policy.
4. En el Editor de políticas, seleccione la pestaña JSON y, a continuación, copie y pegue la siguiente política.

> **Note**
>
> El siguiente ejemplo de política admite los permisos necesarios para leer AWS Glue los metadatos y sus datos correspondientes de Amazon S3. No obstante, quizás tenga que modificar esta política en función de cómo haya configurado los datos de S3. Esta política no incluye una clave de KMS para descifrar los datos.
> Sus AWS Glue recursos y los recursos subyacentes de Amazon S3 deben estar en la Región de AWS misma posición que los de la AWS Clean Rooms colaboración.

```
{
    "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:region:accountId:database/databases",
                "arn:aws:glue:region:accountId:table/databases/tables",
                "arn:aws:glue:region:accountId:catalog",
                "arn:aws:glue:region:accountId:database/default"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "glue:CreateDatabase"
            ],
            "Resource": [
                "arn:aws:glue:region:accountId:database/default"
            ]
        },
        {
            "Effect": "Allow",
            "Action": [
                "s3:ListBucket",
                "s3:GetBucketLocation"
            ],
            "Resource": [
                "arn:aws:s3:::bucket"
            ],
            "Condition":{
                "StringEquals":{
                    "s3:ResourceAccount":[
                        "accountId"
                    ]
                }
            }
        },
        {
```

```
            "Effect": "Allow",
            "Action": [
                "s3:GetObject"
            ],
            "Resource": [
                "arn:aws:s3:::bucketFolders/*"
            ],
            "Condition":{
                "StringEquals":{
                    "s3:ResourceAccount":[
                        "accountId"
                    ]
                }
            }
        }
    ]
}
```

Si es necesario usar una clave KMS para descifrar los datos, añada esta instrucción de AWS KMS a la plantilla anterior:

```
{
            "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. Sustituya cada uno *placeholder* por su propia información:

   - *region*: nombre del Región de AWS. Por ejemplo, **us-east-1**.

- *accountId*— El Cuenta de AWS ID en el que se encuentra el depósito S3.
- *database/databases*, *table/databases/tablescatalog*, y*database/default*: la ubicación de los datos de entrenamiento a los que AWS Clean Rooms hay que acceder.
- *bucket*— El nombre de recurso de Amazon (ARN) del bucket de S3. Puede encontrar el Nombre de recurso de Amazon (ARN) en la pestaña Propiedades del bucket en Amazon S3.
- *bucketFolders*— El nombre de las carpetas específicas del bucket de S3 a AWS Clean Rooms las que se debe acceder.

6. Elija Siguiente.
7. En Revisar y crear, introduzca un Nombre de política y una Descripción y revise el Resumen.
8. Elija Crear política.

   Ha creado una política para AWS Clean Rooms.
9. En Administración de accesos, elija Roles.

   Con Roles, puede crear credenciales a corto plazo, que son las recomendadas aumentar la seguridad. También puede elegir Usuarios para crear credenciales a largo plazo.
10. Elija Crear rol.
11. En el asistente Crear rol, en Tipo de entidad de confianza, elija Política de confianza personalizada.
12. Copie y pegue la siguiente política de confianza personalizada en el editor JSON.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "AllowAssumeRole",
            "Effect": "Allow",
            "Principal": {
                "Service": "cleanrooms-ml.amazonaws.com"
            },
            "Action": "sts:AssumeRole",
            "Condition": {
                "StringEqualsIfExists": {
                    "aws:SourceAccount": ["accountId"]
                },
                "StringLikeIfExists": {
                    "aws:SourceArn": "arn:aws:cleanrooms-
ml:region:accountId:training-dataset/*"
```

```
                }
            }
        }
    ]
}
```

El `SourceAccount` es siempre tuyo Cuenta de AWS. El `SourceArn` puede limitarse a un conjunto de datos de entrenamiento específico, pero solo después de crear ese conjunto de datos. Como aún no conoce el ARN del conjunto de datos de entrenamiento, el comodín se especifica aquí.

*`accountId`*es el ID Cuenta de AWS que contiene los datos de entrenamiento.

13. Elija Siguiente y, en Agregar permisos, introduzca el nombre de la política que acaba de crear. (es posible que tenga que volver a cargar la página).
14. Seleccione la casilla de verificación situada junto al nombre de la política que creó y, a continuación, elija Siguiente.
15. En Nombre, revisar y crear, introduzca el Nombre del rol y la Descripción.

> **Note**
>
> El Nombre del rol debe coincidir con el patrón de los permisos de `passRole` concedidos al miembro que puede realizar consultas y recibir resultados y a roles de miembros.

    a. Revise la sección Seleccionar entidades de confianza y edítela si es necesario.
    b. Revise los permisos en Agregar permisos y edítelos si es necesario.
    c. Revise las Etiquetas y añada etiquetas si es necesario.
    d. Seleccione Crear rol.

Ha creado el rol de servicio para AWS Clean Rooms.

Creación de un rol de servicio para escribir un segmento similar

AWS Clean Rooms usa un rol de servicio para escribir segmentos similares en un segmento. Puede crear este rol mediante la consola si dispone de los permisos de IAM necesarios. Si no tiene permisos `CreateRole`, pida al administrador que cree el rol de servicio.

Creación de un rol de servicio para escribir un segmento similar

1. Inicie sesión en la consola de IAM (https://console.aws.amazon.com/iam/) con su cuenta de administrador.
2. En Administración de acceso, seleccione Políticas.
3. Elija Create Policy.
4. En el Editor de políticas, seleccione la pestaña JSON y, a continuación, copie y pegue la siguiente política.

> **Note**
>
> El siguiente ejemplo de política admite los permisos necesarios para leer AWS Glue los metadatos y sus datos correspondientes de Amazon S3. Sin embargo, es posible que tengas que modificar esta política en función de cómo hayas configurado los datos de Amazon S3. Esta política no incluye una clave de KMS para descifrar los datos.
> Sus AWS Glue recursos y los recursos subyacentes de Amazon S3 deben estar en la Región de AWS misma posición que los de la AWS Clean Rooms colaboración.

```
{
    "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"
                    ]
                }
            }
        }
  ]
}
```

Si es necesario usar una clave de KMS para cifrar los datos, agregue esta instrucción de AWS KMS a la plantilla:

```
{
            "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. Sustituya cada uno *placeholder* por su propia información:

- *buckets*— El nombre de recurso de Amazon (ARN) del bucket de S3. Puede encontrar el Nombre de recurso de Amazon (ARN) en la pestaña Propiedades del bucket en Amazon S3.
- *accountId*— El Cuenta de AWS ID en el que se encuentra el bucket de S3.
- *bucketFolders*— El nombre de las carpetas específicas del depósito de S3 a AWS Clean Rooms las que hay que acceder.
- *region*: nombre del Región de AWS. Por ejemplo, **us-east-1**.
- *keyId*— La clave KMS necesaria para cifrar los datos.

6. Elija Siguiente.
7. En Revisar y crear, introduzca un Nombre de política y una Descripción y revise el Resumen.
8. Elija Crear política.

   Ha creado una política para AWS Clean Rooms.
9. En Administración de accesos, elija Roles.

   Con Roles, puede crear credenciales a corto plazo, que son las recomendadas aumentar la seguridad. También puede elegir Usuarios para crear credenciales a largo plazo.
10. Elija Crear rol.
11. En el asistente Crear rol, en Tipo de entidad de confianza, elija Política de confianza personalizada.
12. Copie y pegue la siguiente política de confianza personalizada en el editor JSON.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "AllowAssumeRole",
            "Effect": "Allow",
            "Principal": {
                "Service": "cleanrooms-ml.amazonaws.com"
            },
            "Action": "sts:AssumeRole",
            "Condition": {
                "StringEqualsIfExists": {
                    "aws:SourceAccount": ["accountId"]

                },
                "StringLikeIfExists": {
```

```
                    "aws:SourceArn": "arn:aws:cleanrooms-
ml:region:accountId:configured-audience-model/*"
                }
            }
        }
    ]
}
```

El `SourceAccount` es siempre tuyo Cuenta de AWS. El `SourceArn` puede limitarse a un conjunto de datos de entrenamiento específico, pero solo después de crear ese conjunto de datos. Como aún no conoce el ARN del conjunto de datos de entrenamiento, el comodín se especifica aquí.

13. Elija Siguiente.
14. Seleccione la casilla de verificación situada junto al nombre de la política que creó y, a continuación, elija Siguiente.
15. En Nombre, revisar y crear, introduzca el Nombre del rol y la Descripción.

> **Note**
>
> El Nombre del rol debe coincidir con el patrón de los permisos de `passRole` concedidos al miembro que puede realizar consultas y recibir resultados y a roles de miembros.

    a. Revise la sección Seleccionar entidades de confianza y edítela si es necesario.
    b. Revise los permisos en Agregar permisos y edítelos si es necesario.
    c. Revise las Etiquetas y añada etiquetas si es necesario.
    d. Seleccione Crear rol.

Ha creado el rol de servicio para. AWS Clean Rooms

Creación de rol de servicio para leer datos iniciales

AWS Clean Rooms utiliza un rol de servicio para leer los datos iniciales. Puede crear este rol mediante la consola si dispone de los permisos de IAM necesarios. Si no tiene permisos `CreateRole`, pida al administrador que cree el rol de servicio.

Para crear un rol de servicio para leer los datos iniciales almacenados en un bucket de S3.

1. Inicie sesión en la consola de IAM (https://console.aws.amazon.com/iam/) con su cuenta de administrador.
2. En Administración de acceso, seleccione Políticas.
3. Elija Create Policy.
4. En el Editor de políticas, seleccione la pestaña JSON y, a continuación, copie y pegue una de las siguientes políticas.

> **Note**
>
> El siguiente ejemplo de política admite los permisos necesarios para leer AWS Glue los metadatos y sus datos correspondientes de Amazon S3. Sin embargo, es posible que tengas que modificar esta política en función de cómo hayas configurado los datos de Amazon S3. Esta política no incluye una clave de KMS para descifrar los datos.
> Sus AWS Glue recursos y los recursos subyacentes de Amazon S3 deben estar en la Región de AWS misma posición que los de la AWS Clean Rooms colaboración.

```
{
    "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**
>
> La siguiente política de ejemplo admite los permisos necesarios para leer los resultados de una consulta SQL y utilizarlos como datos de entrada. No obstante, es posible que tenga que modificar esta política en función de cómo esté estructurada la consulta. Esta política no incluye una clave de KMS para descifrar los datos.

```
{
    "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:region:queryRunnerAccountId:membership/
queryRunnerMembershipId"
            ]
        }
    ]
}
```

Si es necesario usar una clave KMS para descifrar los datos, añada esta instrucción de AWS KMS a la plantilla:

```
{
            "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. Sustituya cada uno *placeholder* por su propia información:

   - *buckets*— El nombre de recurso de Amazon (ARN) del bucket de S3. Puede encontrar el Nombre de recurso de Amazon (ARN) en la pestaña Propiedades del bucket en Amazon S3.
   - *accountId*— El Cuenta de AWS ID en el que se encuentra el bucket de S3.
   - *bucketFolders*— El nombre de las carpetas específicas del depósito de S3 a AWS Clean Rooms las que hay que acceder.

- *region*: nombre del Región de AWS. Por ejemplo, **us-east-1**.
- *queryRunnerAccountId*— El Cuenta de AWS ID de la cuenta que ejecutará las consultas.
- *queryRunnerMembershipId*— El ID de membresía del miembro que puede realizar la consulta. Puede encontrar el ID de pertenencia en la pestaña Detalles de la colaboración. Esto garantiza AWS Clean Rooms que solo asuma el rol cuando este miembro ejecute el análisis en esta colaboración.
- *keyId*— La clave KMS necesaria para cifrar los datos.

6. Elija Siguiente.
7. En Revisar y crear, introduzca un Nombre de política y una Descripción y revise el Resumen.
8. Elija Crear política.

   Ha creado una política para AWS Clean Rooms.
9. En Administración de accesos, elija Roles.

   Con Roles, puede crear credenciales a corto plazo, que son las recomendadas aumentar la seguridad. También puede elegir Usuarios para crear credenciales a largo plazo.
10. Elija Crear rol.
11. En el asistente Crear rol, en Tipo de entidad de confianza, elija Política de confianza personalizada.
12. Copie y pegue la siguiente política de confianza personalizada en el editor JSON.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "AllowAssumeRole",
            "Effect": "Allow",
            "Principal": {
                "Service": "cleanrooms-ml.amazonaws.com"
            },
            "Action": "sts:AssumeRole",
            "Condition": {
                "StringEqualsIfExists": {
                    "aws:SourceAccount": ["accountId"]

                },
                "StringLikeIfExists": {
```

```
                    "aws:SourceArn": "arn:aws:cleanrooms-
ml:region:accountId:audience-generation-job/*"
                }
            }
        }
    ]
}
```

El `SourceAccount` es siempre tuyo Cuenta de AWS. El `SourceArn` puede limitarse a un conjunto de datos de entrenamiento específico, pero solo después de crear ese conjunto de datos. Como aún no conoce el ARN del conjunto de datos de entrenamiento, el comodín se especifica aquí.

13. Elija Siguiente.
14. Seleccione la casilla de verificación situada junto al nombre de la política que creó y, a continuación, elija Siguiente.
15. En Nombre, revisar y crear, introduzca el Nombre del rol y la Descripción.

> **Note**
>
> El Nombre del rol debe coincidir con el patrón de los permisos de `passRole` concedidos al miembro que puede realizar consultas y recibir resultados y a roles de miembros.

    a. Revise la sección Seleccionar entidades de confianza y edítela si es necesario.
    b. Revise los permisos en Agregar permisos y edítelos si es necesario.
    c. Revise las Etiquetas y añada etiquetas si es necesario.
    d. Seleccione Crear rol.

Ha creado el rol de servicio para. AWS Clean Rooms

Configure las funciones de servicio para el modelado personalizado

Temas

- Cree un rol de servicio para el modelado de ML personalizado: Configuración de ML
- Cree un rol de servicio para proporcionar un modelo de aprendizaje automático personalizado
- Cree un rol de servicio para consultar un conjunto de datos

- Cree un rol de servicio para crear una asociación de tablas configurada

Cree un rol de servicio para el modelado de ML personalizado: Configuración de ML

AWS Clean Rooms usa un rol de servicio para controlar quién puede crear una configuración de ML personalizada. Puede crear este rol mediante la consola si dispone de los permisos de IAM necesarios. Si no tiene permisos `CreateRole`, pida al administrador que cree el rol de servicio.

Esta función le permite utilizar la MLConfiguration acción Put.

Para crear un rol de servicio que permita la creación de una configuración de aprendizaje automático personalizada

1. Inicie sesión en la consola de IAM (https://console.aws.amazon.com/iam/) con su cuenta de administrador.
2. En Administración de acceso, seleccione Políticas.
3. Elija Create Policy.
4. En el Editor de políticas, seleccione la pestaña JSON y, a continuación, copie y pegue la siguiente política.

> **Note**
>
> El siguiente ejemplo de política admite los permisos necesarios para acceder a un bucket de S3 y escribirlos y para publicar CloudWatch métricas. Sin embargo, es posible que tengas que modificar esta política en función de cómo hayas configurado los datos de Amazon S3. Esta política no incluye una clave de KMS para descifrar los datos. Sus recursos de Amazon S3 deben estar en los Región de AWS mismos que los de la AWS Clean Rooms colaboración.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "AllowS3ObjectWriteForExport",
            "Effect": "Allow",
            "Action": [
                "s3:PutObject"
```

```
            ],
            "Resource": [
                "arn:aws:s3:::bucket/*"
            ],
            "Condition": {
                "StringEquals": {
                    "s3:ResourceAccount": [
                        "accountId"
                    ]
                }
            }
        },
        {
            "Sid": "AllowS3KMSEncryptForExport",
            "Effect": "Allow",
            "Action": [
                "kms:Encrypt",
                "kms:GenerateDataKey*"
            ],
            "Resource": [
                "arn:aws:kms:region:accountId: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:region:account-id:log-group:/aws/cleanroomsml/*"
            ],
        }
    ]
}
```

5. Sustituya cada uno *placeholder* por su propia información:

   - *bucket*— El nombre de recurso de Amazon (ARN) del bucket de S3. Puede encontrar el Nombre de recurso de Amazon (ARN) en la pestaña Propiedades del bucket en Amazon S3.
   - *region*: nombre del Región de AWS. Por ejemplo, **us-east-1**.
   - *accountId*— El Cuenta de AWS ID en el que se encuentra el bucket de S3.
   - *keyId*— La clave KMS necesaria para cifrar los datos.
6. Elija Siguiente.
7. En Revisar y crear, introduzca un Nombre de política y una Descripción y revise el Resumen.
8. Elija Crear política.

   Ha creado una política para AWS Clean Rooms.
9. En Administración de accesos, elija Roles.

   Con Roles, puede crear credenciales a corto plazo, que son las recomendadas aumentar la seguridad. También puede elegir Usuarios para crear credenciales a largo plazo.
10. Elija Crear rol.
11. En el asistente Crear rol, en Tipo de entidad de confianza, elija Política de confianza personalizada.
12. Copie y pegue la siguiente política de confianza personalizada en el editor JSON.

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

```
            },
            "Action": "sts:AssumeRole",
            "Condition": {
                "StringEquals": {
                    "aws:SourceAccount": "accountId"
                },
                "ArnLike": {
                    "aws:SourceArn":
 "arn:aws:cleanrooms:region:accountId:membership/membershipID"
                }
            }
        }
    ]
}
```

El `SourceAccount` es siempre tuyo Cuenta de AWS. El `SourceArn` puede limitarse a un conjunto de datos de entrenamiento específico, pero solo después de crear ese conjunto de datos. Como aún no conoce el ARN del conjunto de datos de entrenamiento, el comodín se especifica aquí.

13. Elija Siguiente.
14. Seleccione la casilla de verificación situada junto al nombre de la política que creó y, a continuación, elija Siguiente.
15. En Nombre, revisar y crear, introduzca el Nombre del rol y la Descripción.

> **Note**
>
> El Nombre del rol debe coincidir con el patrón de los permisos de `passRole` concedidos al miembro que puede realizar consultas y recibir resultados y a roles de miembros.

a. Revise la sección Seleccionar entidades de confianza y edítela si es necesario.
b. Revise los permisos en Agregar permisos y edítelos si es necesario.
c. Revise las Etiquetas y añada etiquetas si es necesario.
d. Seleccione Crear rol.

Ha creado el rol de servicio para. AWS Clean Rooms

Cree un rol de servicio para proporcionar un modelo de aprendizaje automático personalizado

AWS Clean Rooms utiliza un rol de servicio para controlar quién puede crear un algoritmo de modelo de aprendizaje automático personalizado. Puede crear este rol mediante la consola si dispone de los permisos de IAM necesarios. Si no tiene permisos `CreateRole`, pida al administrador que cree el rol de servicio.

Este rol le permite usar la CreateConfiguredModelAlgorithmacción.

Crear un rol de servicio que permita a un miembro proporcionar un modelo de aprendizaje automático personalizado

1. Inicie sesión en la consola de IAM (https://console.aws.amazon.com/iam/) con su cuenta de administrador.
2. En Administración de acceso, seleccione Políticas.
3. Elija Create Policy.
4. En el Editor de políticas, seleccione la pestaña JSON y, a continuación, copie y pegue la siguiente política.

> **Note**
>
> El siguiente ejemplo de política admite los permisos necesarios para recuperar la imagen del docker que contiene el algoritmo del modelo. Sin embargo, es posible que tengas que modificar esta política en función de cómo hayas configurado los datos de Amazon S3. Esta política no incluye una clave de KMS para descifrar los datos.
> Sus recursos de Amazon S3 deben estar en los Región de AWS mismos que los de la AWS Clean Rooms colaboración.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "AllowECRImageDownloadForTrainingAndInferenceJobs",
            "Effect": "Allow",
            "Action": [
                "ecr:BatchGetImage",
```

```
                "ecr:BatchCheckLayerAvailability",
                "ecr:GetDownloadUrlForLayer"
            ],
            "Resource": "arn:aws:ecr:region:accountID:repository/repoName"
        }
    ]
}
```

5. Sustituya cada uno *placeholder* por su propia información:
   - *region*: nombre del Región de AWS. Por ejemplo, **us-east-1**.
   - *accountId*— El Cuenta de AWS ID en el que se encuentra el depósito S3.
   - *repoName*— El nombre del repositorio que contiene los datos.
6. Elija Siguiente.
7. En Revisar y crear, introduzca un Nombre de política y una Descripción y revise el Resumen.
8. Elija Crear política.

   Ha creado una política para AWS Clean Rooms.
9. En Administración de accesos, elija Roles.

   Con Roles, puede crear credenciales a corto plazo, que son las recomendadas aumentar la seguridad. También puede elegir Usuarios para crear credenciales a largo plazo.
10. Elija Crear rol.
11. En el asistente Crear rol, en Tipo de entidad de confianza, elija Política de confianza personalizada.
12. Copie y pegue la siguiente política de confianza personalizada en el editor JSON.

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

Siempre `SourceAccount` es tuyo. `SourceArn` Puede limitarse a un conjunto de datos de entrenamiento específico, pero solo después de que se haya creado ese conjunto de datos. Cuenta de AWS Como aún no conoce el ARN del conjunto de datos de entrenamiento, el comodín se especifica aquí.

13. Elija Siguiente.
14. Seleccione la casilla de verificación situada junto al nombre de la política que creó y, a continuación, elija Siguiente.
15. En Nombre, revisar y crear, introduzca el Nombre del rol y la Descripción.

> **Note**
>
> El Nombre del rol debe coincidir con el patrón de los permisos de `passRole` concedidos al miembro que puede realizar consultas y recibir resultados y a roles de miembros.

    a. Revise la sección Seleccionar entidades de confianza y edítela si es necesario.
    b. Revise los permisos en Agregar permisos y edítelos si es necesario.
    c. Revise las Etiquetas y añada etiquetas si es necesario.
    d. Seleccione Crear rol.

Ha creado el rol de servicio para. AWS Clean Rooms

Cree un rol de servicio para consultar un conjunto de datos

AWS Clean Rooms usa un rol de servicio para controlar quién puede consultar un conjunto de datos que se utilizará para el modelado de aprendizaje automático personalizado. Puede crear este rol mediante la consola si dispone de los permisos de IAM necesarios. Si no tiene permisos `CreateRole`, pida al administrador que cree el rol de servicio.

Este rol te permite usar la acción Crear MLInput canal.

Para crear un rol de servicio que permita a un miembro consultar un conjunto de datos

1. Inicie sesión en la consola de IAM (https://console.aws.amazon.com/iam/) con su cuenta de administrador.
2. En Administración de acceso, seleccione Políticas.

3. Elija Create Policy.
4. En el Editor de políticas, seleccione la pestaña JSON y, a continuación, copie y pegue la siguiente política.

> **Note**
>
> El siguiente ejemplo de política admite los permisos necesarios para consultar un conjunto de datos que se utilizará para el modelado de aprendizaje automático personalizado. Sin embargo, es posible que tengas que modificar esta política en función de cómo hayas configurado los datos de Amazon S3. Esta política no incluye una clave de KMS para descifrar los datos.
> Sus recursos de Amazon S3 deben estar en los Región de AWS mismos que los de la AWS Clean Rooms colaboración.

```
{
    "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:region:queryRunnerAccountId:membership/
queryRunnerMembershipId"
            ]
        }
    ]
}
```

5. Sustituya cada uno *placeholder* por su propia información:

   - *region*: nombre del Región de AWS. Por ejemplo, **us-east-1**.
   - *queryRunnerAccountId*— El Cuenta de AWS ID de la cuenta que ejecutará las consultas.
   - *queryRunnerMembershipId*— El ID de membresía del miembro que puede realizar la consulta. Puede encontrar el ID de pertenencia en la pestaña Detalles de la colaboración. Esto garantiza AWS Clean Rooms que solo asuma el rol cuando este miembro ejecute el análisis en esta colaboración.
6. Elija Siguiente.
7. En Revisar y crear, introduzca un Nombre de política y una Descripción y revise el Resumen.
8. Elija Crear política.

   Ha creado una política para AWS Clean Rooms.
9. En Administración de accesos, elija Roles.

   Con Roles, puede crear credenciales a corto plazo, que son las recomendadas aumentar la seguridad. También puede elegir Usuarios para crear credenciales a largo plazo.
10. Elija Crear rol.
11. En el asistente Crear rol, en Tipo de entidad de confianza, elija Política de confianza personalizada.
12. Copie y pegue la siguiente política de confianza personalizada en el editor JSON.

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

```
            },
            "Action": "sts:AssumeRole"
        }
    ]
}
```

Siempre `SourceAccount` es tuyo. `SourceArn` Puede limitarse a un conjunto de datos de entrenamiento específico, pero solo después de que se haya creado ese conjunto de datos. Cuenta de AWS Como aún no conoce el ARN del conjunto de datos de entrenamiento, el comodín se especifica aquí.

13. Elija Siguiente.
14. Seleccione la casilla de verificación situada junto al nombre de la política que creó y, a continuación, elija Siguiente.
15. En Nombre, revisar y crear, introduzca el Nombre del rol y la Descripción.

> **Note**
>
> El Nombre del rol debe coincidir con el patrón de los permisos de `passRole` concedidos al miembro que puede realizar consultas y recibir resultados y a roles de miembros.

a. Revise la sección Seleccionar entidades de confianza y edítela si es necesario.
b. Revise los permisos en Agregar permisos y edítelos si es necesario.
c. Revise las Etiquetas y añada etiquetas si es necesario.
d. Seleccione Crear rol.

Ha creado el rol de servicio para. AWS Clean Rooms

Cree un rol de servicio para crear una asociación de tablas configurada

AWS Clean Rooms usa un rol de servicio para controlar quién puede crear una asociación de tablas configurada. Puede crear este rol mediante la consola si dispone de los permisos de IAM necesarios. Si no tiene permisos `CreateRole`, pida al administrador que cree el rol de servicio.

Este rol le permite usar la CreateConfiguredTableAssociation acción.

Para crear un rol de servicio que permita la creación de una asociación de tablas configurada

1. Inicie sesión en la consola de IAM (https://console.aws.amazon.com/iam/) con su cuenta de administrador.
2. En Administración de acceso, seleccione Políticas.
3. Elija Create Policy.
4. En el Editor de políticas, seleccione la pestaña JSON y, a continuación, copie y pegue la siguiente política.

> **Note**
>
> El siguiente ejemplo de política admite la creación de una asociación de tablas configurada. Sin embargo, es posible que tengas que modificar esta política en función de cómo hayas configurado los datos de Amazon S3. Esta política no incluye una clave de KMS para descifrar los datos.
> Sus recursos de Amazon S3 deben estar en los Región de AWS mismos que los de la AWS Clean Rooms colaboración.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "kms:Decrypt",
                "kms:DescribeKey"
            ],
            "Resource": "KMS key used to encrypt the S3 data",
            "Effect": "Allow"
        },
        {
            "Action": [
                "s3:ListBucket",
                "s3:GetBucketLocation"
            ],
            "Resource": "S3 bucket of Glue table",
            "Effect": "Allow"
        },
        {
            "Action": "s3:GetObject",
```

```
            "Resource": "S3 bucket of Glue table/*",
            "Effect": "Allow"
        },
        {
            "Action": [
                "glue:GetDatabase",
                "glue:GetDatabases",
                "glue:GetTable",
                "glue:GetTables",
                "glue:GetPartitions",
                "glue:GetPartition",
                "glue:BatchGetPartition"
            ],
            "Resource": [
                "arn:aws:glue:region:accountID:catalog",
                "arn:aws:glue:region:accountID:database/Glue database name",
                "arn:aws:glue:region:accountID:table/Glue database name/Glue table
 name"
            ],
            "Effect": "Allow"
        },
        {
            "Action": [
                "glue:GetSchema",
                "glue:GetSchemaVersion"
            ],
            "Resource": "*",
            "Effect": "Allow"
        }
    ]
}
```

5. Sustituya cada uno *placeholder* por su propia información:

   - *KMS key used to encrypt the Amazon S3 data*— La clave de KMS que se utilizó para cifrar los datos de Amazon S3. Para descifrar los datos, debe proporcionar la misma clave KMS que se utilizó para cifrar los datos.
   - *Amazon S3 bucket of AWS Glue table*— El nombre del bucket de Amazon S3 que contiene la AWS Glue tabla que contiene sus datos.
   - *region*: nombre del Región de AWS. Por ejemplo, **us-east-1**.
   - *accountId*— El Cuenta de AWS ID de la cuenta propietaria de los datos.

- *AWS Glue database name*— El nombre de la AWS Glue base de datos que contiene los datos.
- *AWS Glue table name*— El nombre de la AWS Glue tabla que contiene los datos.

6. Elija Siguiente.
7. En Revisar y crear, introduzca un Nombre de política y una Descripción y revise el Resumen.
8. Elija Crear política.

   Ha creado una política para AWS Clean Rooms.
9. En Administración de accesos, elija Roles.

   Con Roles, puede crear credenciales a corto plazo, que son las recomendadas aumentar la seguridad. También puede elegir Usuarios para crear credenciales a largo plazo.
10. Elija Crear rol.
11. En el asistente Crear rol, en Tipo de entidad de confianza, elija Política de confianza personalizada.
12. Copie y pegue la siguiente política de confianza personalizada en el editor JSON.

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

   Siempre `SourceAccount` es tuyo. `SourceArn` Puede limitarse a un conjunto de datos de entrenamiento específico, pero solo después de que se haya creado ese conjunto de datos. Cuenta de AWS Como aún no conoce el ARN del conjunto de datos de entrenamiento, el comodín se especifica aquí.
13. Elija Siguiente.
14. Seleccione la casilla de verificación situada junto al nombre de la política que creó y, a continuación, elija Siguiente.

15. En Nombre, revisar y crear, introduzca el Nombre del rol y la Descripción.

> **Note**
>
> El Nombre del rol debe coincidir con el patrón de los permisos de `passRole` concedidos al miembro que puede realizar consultas y recibir resultados y a roles de miembros.

    a. Revise la sección Seleccionar entidades de confianza y edítela si es necesario.
    b. Revise los permisos en Agregar permisos y edítelos si es necesario.
    c. Revise las Etiquetas y añada etiquetas si es necesario.
    d. Seleccione Crear rol.

Ha creado el rol de servicio para. AWS Clean Rooms

Colaboraciones y suscripciones en AWS Clean Rooms

Una colaboración es un límite lógico seguro AWS Clean Rooms en el que los miembros pueden realizar análisis en tablas configuradas.

Cualquier miembro AWS Clean Rooms puede crear una colaboración.

El creador de la colaboración puede designar a un solo miembro para analizar las tablas configuradas y recibir los resultados. Sin embargo, es posible que el creador de la colaboración desee impedir que el miembro que puede ejecutar el análisis tenga acceso a los resultados de la consulta. En ese caso, el creador de la colaboración puede designar a un miembro para que realice la consulta o a un miembro para que ejecute las consultas y los trabajos y a otro miembro para que reciba los resultados.

En la mayoría de los casos, el miembro que puede realizar consultas o el miembro que puede consultar y ejecutar trabajos es también el miembro que paga los costes de procesamiento. No obstante, el creador de la colaboración puede configurar a otro miembro para que se encargue de pagar los costos de computación de consultas.

Para obtener información sobre cómo crear una colaboración mediante la AWS SDKs, consulta la referencia de la AWS Clean Rooms API.

Temas

- Selección de un tipo de motor de análisis en AWS Clean Rooms
- Crear una colaboración
- Crear una pertenencia y unirse a una colaboración
- Editar colaboraciones
- Eliminar colaboraciones
- Ver colaboraciones
- Invitar a los miembros a una colaboración
- Supervisión de miembros
- Eliminar un miembro de una colaboración
- Abandonar una colaboración

Selección de un tipo de motor de análisis en AWS Clean Rooms

> **Note**
>
> AWS Clean Rooms dejará de ser compatible con el antiguo motor de análisis SQL de Clean Rooms el 17 de diciembre de 2025. Antes del 16 de julio de 2025, debe solicitar un aumento del límite a través de AWS Customer Support para crear nuevas colaboraciones basadas en el motor SQL de Clean Rooms. A partir del 17 de julio de 2025, ya no estará disponible la creación de nuevas colaboraciones basadas en el motor SQL de Clean Rooms.

Un motor de análisis es un componente de software que procesa consultas de datos y realiza cálculos analíticos en su interior. AWS Clean Rooms El motor de análisis interpreta los comandos SQL, ejecuta operaciones de procesamiento de datos y devuelve los resultados del análisis. Antes de crear una AWS Clean Rooms colaboración, debe elegir entre dos motores de análisis disponibles en función de sus requisitos técnicos y sus necesidades de procesamiento de datos. Los criterios de selección deben centrarse principalmente en el tamaño del conjunto de datos, la complejidad de las consultas, las funciones compatibles con el motor y la compatibilidad de las fuentes de datos.

En la siguiente tabla se describen los detalles de cada motor de análisis, lo que puede ayudarle a determinar la mejor opción para sus necesidades.

Motor de análisis	¿Cuándo lo usarías?	¿Se admite la regla de análisis de agregación?	¿Se admite la regla de análisis de listas?	¿Se admite una regla de análisis personalizada sin privacidad diferencial?	¿Se admite una regla de análisis personalizada con privacidad diferencial?	¿Soporta la fuente de datos de Amazon S3?	¿Soporta las fuentes de datos de Amazon Athena y Snowflake?
Motor de análisis Spark	• Ejecutan o consultas	Sí	Sí	Sí	No	Sí	Sí

Motor de análisis	¿Cuándo lo usarías?	¿Se admite la regla de análisis de agregación?	¿Se admite la regla de análisis de listas?	¿Se admite una regla de análisis personalizada sin privacidad diferencial?	¿Se admite una regla de análisis personalizada con privacidad diferencial?	¿Soporta la fuente de datos de Amazon S3?	¿Soporta las fuentes de datos de Amazon Athena y Snowflake?
	SQL de Spark • Ejecutando PySpark trabajos • Modelado ML personalizado						
AWS Clean Rooms Motor de análisis SQL	AWS Clean Rooms Ejecutando consultas SQL	Sí	Sí	Sí	Sí	Sí	No

Para obtener información sobre las consultas SQL de Spark, consulta la referencia de SQL de AWS Clean Rooms Spark.

Para obtener información sobre las consultas AWS Clean Rooms SQL, consulta la Referencia de AWS Clean Rooms SQL.

Para obtener información sobre los precios de Spark SQL y AWS Clean Rooms SQL, consulta AWS Clean Rooms los precios.

Una vez que haya determinado qué motor de análisis utilizar en su colaboración, estará listo para seguir los pasos que se indican a continuaciónCrear una colaboración.

Crear una colaboración

Hay tres formas de crear una colaboración en AWS Clean Rooms.

La forma más básica es la colaboración para consultas. Esta colaboración se centra en el análisis de consultas SQL y mantiene una estructura simple con dos funciones principales: un miembro que puede ejecutar consultas y otro que puede recibir los resultados. Esta configuración de colaboración básica funciona bien para tareas sencillas de análisis de datos.

La segunda forma, la colaboración para consultas y trabajos, amplía la funcionalidad al incorporar consultas y PySpark trabajos de SQL y requiere Spark como motor de análisis. Esta configuración de colaboración mantiene la misma estructura de funciones básica, pero amplía los permisos para incluir la ejecución de tareas. Un requisito importante es que el miembro que crea las plantillas de PySpark análisis también debe ser el que reciba los resultados, lo que garantiza una responsabilidad clara en el proceso de análisis.

La tercera forma, la colaboración para el modelado de aprendizaje automático, está diseñada para los flujos de trabajo de aprendizaje automático y requiere Spark como motor de análisis. Esta configuración de colaboración añade dos funciones más: una para los usuarios que necesitan los resultados de modelos entrenados y otra para los usuarios que necesitan los resultados del uso de esos modelos para hacer predicciones. Esta configuración de colaboración ayuda a los miembros de la colaboración a trabajar juntos en proyectos de datos complejos y, al mismo tiempo, mantiene claras las funciones y permisos de todos.

En los siguientes temas se explica cómo crear colaboraciones para consultas, trabajos y modelos de aprendizaje automático.

Temas

- Crear una colaboración para consultas
- Crear una colaboración para consultas y trabajos

- Crear una colaboración para el modelado de aprendizaje automático

Crear una colaboración para consultas

En este procedimiento, usted, como creador de la colaboración, realiza las siguientes tareas:

- Cree una colaboración.
- Invita a uno o más miembros a la colaboración.
- Asigne capacidades a los miembros, como el miembro que puede realizar consultas y el miembro que puede recibir los resultados.

  Si el creador de la colaboración es también el miembro que puede recibir los resultados, especificará el destino y el formato de los resultados. También proporcionan una función de servicio Amazon Resource Name (ARN) para escribir los resultados en el destino de los resultados.
- Configure qué miembro es responsable de pagar los costes informáticos de la colaboración.

Antes de comenzar, asegúrese de que cumple los siguientes requisitos previos:

- Ha determinado el tipo de motor de análisis que desea utilizar.
- Tienes el nombre y el Cuenta de AWS ID de cada miembro al que quieres invitar a la colaboración.
- Tiene permiso para compartir el nombre y la Cuenta de AWS identificación de cada miembro con todos los miembros de la colaboración.

> **Note**
>
> No puedes añadir más miembros después de crear la colaboración.

Para obtener información sobre cómo crear una colaboración mediante la AWS SDKs, consulta la referencia de la AWS Clean Rooms API.

Para crear una colaboración para consultas

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con la Cuenta de AWS que funcionará como creador de la colaboración.
2. En el panel de navegación izquierdo, elija Colaboraciones.

3. En la esquina superior derecha, elija Crear colaboración.
4. En Paso 1: Definir la colaboración, haga lo siguiente:

   a. En Detalles, introduzca el Nombre y la Descripción de la colaboración.

      Esta información estará visible para los miembros de la colaboración que estén invitados a participar en la colaboración. La información de Nombre y Descripción les ayudará a entender a qué se refiere la colaboración.

   b. Elige el motor de análisis que quieras usar.

      Para obtener más información, consulte Selección de un tipo de motor de análisis en AWS Clean Rooms.

      > **Note**
      >
      > AWS Clean Rooms dejará de ser compatible con el antiguo motor de análisis SQL de Clean Rooms el 17 de diciembre de 2025. Antes del 16 de julio de 2025, debe solicitar un aumento del límite a través de AWS Customer Support para crear nuevas colaboraciones basadas en el motor SQL de Clean Rooms. A partir del 17 de julio de 2025, ya no estará disponible la creación de nuevas colaboraciones basadas en el motor SQL de Clean Rooms.

   c. En Miembros:

      i. En Miembro 1: usted, introduzca su Nombre de miembro para mostrar tal y como desee que aparezca en la colaboración.

         > **Note**
         >
         > Su Cuenta de AWS ID se incluye automáticamente como ID de miembro. Cuenta de AWS

      ii. En el caso del miembro 2, introduzca el nombre visible y el Cuenta de AWS ID de miembro del miembro al que desee invitar a la colaboración.

         El Nombre del miembro para mostrar y el ID de Cuenta de AWS del miembro estarán visibles para todos los invitados a participar en la colaboración. Una vez ingresados y guardados los valores de estos campos, no se podrán editar.

> **Note**
>
> Debe informar al miembro de la colaboración de que la información de los campos ID de Cuenta de AWS del miembro y Nombre del miembro para mostrar estará visible para todos los colaboradores activos de la colaboración.

iii. Si desea añadir otro miembro, elija Añadir otro miembro. A continuación, introduzca el nombre de usuario y el Cuenta de AWS ID de miembro de cada miembro que pueda aportar los datos que desee invitar a la colaboración.

d. Si desea habilitar el registro de análisis, seleccione la casilla Habilitar el registro de análisis.

- Seleccione la casilla Registrar a partir de consultas situada en los tipos de registro admitidos.

  Recibirás los registros generados a partir de consultas SQL en tu cuenta de Amazon CloudWatch Logs.

e. (Opcional) Si desea habilitar la capacidad de computación criptográfica, seleccione la casilla Habilitar computación criptográfica.

i. Elija los siguientes parámetros de cobertura criptográfica:

- Permitir plaintext columnas

  Seleccione No si necesita tablas totalmente cifradas.

  Elija Sí si lo desea cleartext columnas permitidas en la tabla cifrada.

  Para ejecutar SUM o AVG en determinadas columnas, las columnas deben estar en cleartext.

- Preservar NULL valores

  Elija No si no quiere conservar NULL valores. NULL los valores no aparecerán como NULL en una tabla cifrada.

  Elija Sí si desea conservar NULL valores. NULL los valores aparecerán como NULL en una tabla cifrada.

ii. Elija los siguientes parámetros de toma de huellas digitales:

- Permitir duplicados

  Seleccione No si no desea que se permitan entradas duplicadas en un fingerprint columna.

  Seleccione Sí si desea que se permitan entradas duplicadas en una fingerprint columna.

- Permitir JOIN de columnas con nombres diferentes

  Elija No si no quiere unirse fingerprint columnas con nombres diferentes.

  Elija Sí si quiere unirse fingerprint columnas con nombres diferentes.

Para obtener más información acerca de los parámetros de computación criptográfica, consulte Parámetros de computación criptográfica.

Para obtener más información sobre cómo cifrar los datos para utilizarlos en AWS Clean Rooms, consultePreparación de tablas de datos cifrados con computación criptográfica para Clean Rooms.

> **Note**
>
> Compruebe estas configuraciones detenidamente antes de completar el siguiente paso. Tras crear la colaboración, solo podrá editar el nombre y la descripción de la colaboración y si los registros se almacenan en Amazon CloudWatch Logs.

f. Si desea habilitar la opción de Etiquetas para el recurso de colaboración, seleccione Añadir nueva etiqueta y, a continuación, introduzca el par de Clave y Valor.

g. Elija Siguiente.

5. En el paso 2: especifique las capacidades de los miembros; en el caso de Análisis mediante consultas y trabajos, en la sección Tipos de análisis compatibles, deje la casilla Consultas seleccionada y tome las medidas recomendadas en función de su objetivo.

Su objetivo	Acción recomendada
Consultar los datos de la colaboración y recibir los resultados	1. Elíjase a usted mismo como miembro que puede Ejecutar las consultas.

Su objetivo	Acción recomendada
	2. Elija uno de los miembros que pueden recibir los resultados de los análisis de la lista desplegable.
Consultar los datos de la colaboración y designar a un miembro distinto para recibir los resultados	1. Elíjase a usted mismo como miembro que puede Ejecutar las consultas. 2. Seleccione el miembro que puede recibir los resultados de los análisis en la lista desplegable.
Recibir los resultados de la consulta en la colaboración y designar a un miembro distinto para consultar los datos	1. Seleccione el miembro que puede Ejecutar las consultas en la lista desplegable. 2. Elija uno de los miembros que pueden recibir los resultados de los análisis de la lista desplegable.
Crear y administrar la colaboración, designar a un miembro distinto para consultar los datos y designar a un miembro distinto para recibir los resultados	1. Seleccione el miembro que puede Ejecutar las consultas en la lista desplegable. 2. Seleccione el miembro que puede recibir los resultados de los análisis en la lista desplegable.

a. Si utiliza Clean Rooms ML, para modelar el aprendizaje automático mediante flujos de trabajo diseñados específicamente,

   i. (Opcional) Seleccione en la lista desplegable el miembro que puede recibir resultados de modelos entrenados.

   ii. (Opcional) Seleccione el miembro que puede recibir los resultados de la inferencia del modelo en la lista desplegable.

b. Vea las habilidades de los miembros en la sección Resolución de ID utilizando. AWS Entity Resolution

c. Elija Siguiente.

6. Para el paso 3: configurar el pago, en el caso del análisis mediante consultas, realice una de las siguientes acciones en función de su objetivo.

Su objetivo	Acción recomendada
Designar al miembro que puede Ejecutar las consultas como miembro que paga los costos de computación de las consultas	1. En el caso de Análisis mediante consultas , elija el miembro que pagará por las consultas para que sea el mismo que el miembro que puede ejecutar consultas. 2. Elija Siguiente.
Designar a otro miembro que pague los costos de computación de las consultas	1. En el caso de Analysis mediante consultas , elija usted mismo como el miembro que pagará por las consultas. 2. Elija Siguiente.

   En el caso del modelado de aprendizaje automático mediante flujos de trabajo diseñados específicamente, el creador del modelo similar configurado es el miembro que pagará por el modelado similar.

   En el caso de la resolución de ID AWS Entity Resolution, el creador de la tabla de mapeo de ID es el miembro que pagará por la tabla de mapeo de ID.

7. Para el paso 4: configurar la membresía, elige una de las siguientes opciones:

   Yes, join by creating membership now

   1. Para la configuración de resultados predeterminada, para la configuración de resultados de consultas, si usted es el miembro que puede recibir los resultados,
      a. Para el destino de los resultados en Amazon S3, introduzca el destino de Amazon S3 o elija Browse S3 para seleccionar un bucket de S3.
      b. Como Formato de los resultados de la consulta, elija CSV o PARQUET.
      c. (Solo Spark) Para los archivos de resultados, elija Múltiple o Único.
      d. (Opcional) Para acceder al servicio, si desea enviar consultas que tarden hasta 24 horas en llegar a su destino de S3, active la casilla Añadir una función de servicio para atender las consultas que tarden hasta 24 horas en completarse.

Las consultas de gran tamaño que tarden hasta 24 horas en completarse se enviarán a su destino de S3.

Si no selecciona la casilla, solo las consultas que se completen en un plazo de 12 horas se enviarán a su ubicación de S3.

e. Especifique los permisos de Acceso a servicios seleccionando Crear y usar un nuevo rol de servicio o Usar un rol de servicio existente.

Si eliges...	Entonces...
Crear y usar un nuevo rol de servicio	• AWS Clean Rooms crea un rol de servicio con la política requerida para esta tabla. • El Nombre del rol de servicio predeterminado es `cleanrooms-result-receiver-<timestamp>` • Debe tener permisos para crear roles y adjuntar políticas.

Si eliges...	Entonces...
Usar un rol de servicio existente	i. Seleccione un Nombre de rol de servicio existente en la lista desplegable. Si tiene permisos de listas de roles, se mostrará la lista de roles. Si no tiene permisos de listas de roles, puede ingresar el nombre de recurso de Amazon (ARN) del rol que desea usar. ii. Consulte el rol de servicio mediante la elección del enlace externo Ver en IAM. Si no hay ningún rol de servicio existente, la opción Usar un rol de servicio existente no estará disponible. De forma predeterminada, AWS Clean Rooms no intenta actualizar la política de roles existente para añadir los permisos necesarios.

Note

- AWS Clean Rooms requiere permisos para realizar consultas de acuerdo con las reglas de análisis. Para obtener más información sobre los permisos para AWS Clean Rooms, consulteAWS políticas gestionadas para AWS Clean Rooms.
- Si el rol no tiene permisos suficientes AWS Clean Rooms, recibirá un mensaje de error que indica que el rol no tiene permisos suficientes para ello AWS Clean Rooms. Debe agregar la política de rol antes de continuar.

- Si no puede modificar la política de roles, recibirá un mensaje de error que indica que no ha AWS Clean Rooms podido encontrar la política para el rol de servicio.

2. En la configuración de Logs, elige una de las siguientes opciones para el almacenamiento de CloudWatch registros en Amazon Logs:

> **Note**
>
> La sección de configuración de registros aparece si decide habilitar el registro de consultas.

a. Selecciona Activar y los registros de consultas relevantes para ti se almacenarán en tu cuenta de Amazon CloudWatch Logs.

   Cada miembro solo puede recibir los registros de las consultas que haya iniciado o que contengan sus datos.

   El miembro que puede recibir los resultados también recibe los registros de todas las consultas que se ejecutan en una colaboración, aunque no se accede a los datos en una consulta.

   En Tipos de registro admitidos, la casilla de verificación Query logs está activada de forma predeterminada.

> **Note**
>
> Tras activar el registro de consultas, el almacenamiento de registros puede tardar unos minutos en configurarse y empezar a recibir registros en Amazon CloudWatch Logs. Durante este breve período, el miembro que puede realizar consultas puede ejecutar consultas que, en realidad, no enviarán registros.

b. Selecciona Desactivar y los registros de consultas relevantes para ti no se almacenarán en tu cuenta de Amazon CloudWatch Logs.

3. Si desea habilitar la opción de Etiquetas para el recurso de pertenencia, seleccione Añadir nueva etiqueta y, a continuación, introduzca el par de Clave y Valor.

4. Si usted es el miembro que paga por Query Compute, indique su aceptación marcando la casilla Acepto pagar los costos de procesamiento de esta colaboración.

> **Note**
>
> Debe seleccionar esta casilla de verificación para continuar.
> Para obtener más información sobre cómo se calculan los precios, consulte Precios para AWS Clean Rooms.

Si es el miembro que paga los costes de cálculo de las consultas, pero no el miembro que puede realizar la consulta, se recomienda configurar un presupuesto AWS Clean Rooms y recibir notificaciones cuando se alcance el presupuesto máximo. AWS Budgets Para obtener más información sobre cómo configurar un presupuesto, consulte Administración de costos con AWS Budgets en la Guía del usuario de AWS Cost Management . Para obtener más información sobre cómo configurar notificaciones, consulte Creación de un tema de Amazon SNS para las notificaciones de presupuesto en la Guía del usuario de AWS Cost Management . Si se ha alcanzado el presupuesto máximo, puede ponerse en contacto con el miembro que puede realizar consultas o abandonar la colaboración. Si abandona la colaboración, no se podrán ejecutar más consultas y, por lo tanto, ya no se le facturarán costos de cómputo de consultas.

5. Elija Siguiente.

Se crean tanto la colaboración como su pertenencia a ella.

Su estado en la colaboración es activo.

No, I will create a membership later

1. Elija Siguiente.

   Se crea solo la colaboración.

   Su estado en la colaboración es inactivo.

8. Para el paso 5: revisar y crear, haga lo siguiente:

   a. Revise las selecciones que realizó en los pasos anteriores y edítelas si es necesario.

   b. Elija una de las opciones.

Si ha elegido...	Entonces seleccione...
Crear una pertenencia en la colaboración (Sí, unirme y crear pertenencia ahora)	Crear colaboración y pertenencia
Crear la colaboración y no crear una pertenencia en este momento (No, crearé una pertenencia más tarde)	Crear colaboración

Una vez que la colaboración se haya creado correctamente, podrá ver la página de detalles de la colaboración en Colaboraciones.

Ya puede hacer lo siguiente:

- Prepare la tabla de datos para analizarla AWS Clean Rooms. (Opcional si desea analizar sus propios datos de eventos o si desea consultar datos de identidad).
- Asociar la tabla configurada a su colaboración (Opcional si quieres analizar los datos de tus propios eventos).
- Agregue una regla de análisis para la tabla configurada. (Opcional si quieres analizar los datos de tus propios eventos).
- Crear una pertenencia y unirse a una colaboración. (Opcional si ya ha creado una suscripción).
- Invita a los miembros a unirse a la colaboración.

Crear una colaboración para consultas y trabajos

En este procedimiento, usted, como creador de la colaboración, realiza las siguientes tareas:

- Cree una colaboración.
- Invita a uno o más miembros a la colaboración.
- Asigne capacidades a los miembros, como el miembro que puede ejecutar consultas y trabajos y el miembro que puede recibir los resultados.

  Si el creador de la colaboración es también el miembro que puede recibir los resultados, especificará el destino y el formato de los resultados. También proporcionan una función

de servicio Amazon Resource Name (ARN) para escribir los resultados en el destino de los resultados.

- Configure qué miembro es responsable de pagar los costos de consulta y procesamiento del trabajo en la colaboración.

Antes de comenzar, asegúrese de que cumple los siguientes requisitos previos:

- Ha determinado el tipo de motor de análisis que desea utilizar.
- Tienes el nombre y el Cuenta de AWS ID de cada miembro al que quieres invitar a la colaboración.
- Tiene permiso para compartir el nombre y la Cuenta de AWS identificación de cada miembro con todos los miembros de la colaboración.

> **Note**
>
> No puedes añadir más miembros después de crear la colaboración.

Para obtener información sobre cómo crear una colaboración mediante la AWS SDKs, consulta la referencia de la AWS Clean Rooms API.

Para crear una colaboración para consultas y trabajos

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con la Cuenta de AWS que funcionará como creador de la colaboración.
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. En la esquina superior derecha, elija Crear colaboración.
4. En Paso 1: Definir la colaboración, haga lo siguiente:

    a. En Detalles, introduzca el Nombre y la Descripción de la colaboración.

    Esta información estará visible para los miembros de la colaboración que estén invitados a participar en la colaboración. La información de Nombre y Descripción les ayudará a entender a qué se refiere la colaboración.

    b. Elige el motor de análisis que quieras usar.

    Para obtener más información, consulte Selección de un tipo de motor de análisis en AWS Clean Rooms.

> **Note**
>
> Si quieres actualizar tu colaboración del motor de análisis de AWS Clean Rooms SQL al motor de análisis de Spark, puedes editar una colaboración existente o volver a crear la colaboración y seleccionar el motor de análisis de Spark.

c. En Miembros:

i. En Miembro 1: usted, introduzca su Nombre de miembro para mostrar tal y como desee que aparezca en la colaboración.

> **Note**
>
> Tu Cuenta de AWS ID se incluye automáticamente como Cuenta de AWS ID de miembro.

ii. En el caso del miembro 2, introduzca el nombre visible y el Cuenta de AWS ID de miembro del miembro al que desee invitar a la colaboración.

El Nombre del miembro para mostrar y el ID de Cuenta de AWS del miembro estarán visibles para todos los invitados a participar en la colaboración. Una vez ingresados y guardados los valores de estos campos, no se podrán editar.

> **Note**
>
> Debe informar al miembro de la colaboración de que la información de los campos ID de Cuenta de AWS del miembro y Nombre del miembro para mostrar estará visible para todos los colaboradores activos de la colaboración.

iii. Si desea añadir otro miembro, elija Añadir otro miembro. A continuación, introduzca el nombre de usuario y el Cuenta de AWS ID de miembro de cada miembro que pueda aportar los datos que desee invitar a la colaboración.

d. Si desea habilitar el registro de análisis, seleccione la casilla Habilitar el registro de análisis y, a continuación, elija los tipos de registro compatibles.

- Si desea recibir los registros generados a partir de consultas SQL, seleccione la casilla Registros a partir de consultas.

- Si desea recibir los registros generados a partir de los trabajos que se utilizan PySpark, seleccione la casilla Registros de los trabajos.

e. (Opcional) Si desea habilitar la capacidad de computación criptográfica, seleccione la casilla Habilitar la computación criptográfica.

i. Elija los siguientes parámetros de cobertura criptográfica:

- Permitir plaintext columnas

  Seleccione No si necesita tablas totalmente cifradas.

  Elija Sí si lo desea cleartext columnas permitidas en la tabla cifrada.

  Para ejecutar SUM o AVG en determinadas columnas, las columnas deben estar en cleartext.

- Preservar NULL valores

  Elija No si no quiere conservar NULL valores. NULL los valores no aparecerán como NULL en una tabla cifrada.

  Elija Sí si desea conservar NULL valores. NULL los valores aparecerán como NULL en una tabla cifrada.

ii. Elija los siguientes parámetros de toma de huellas digitales:

- Permitir duplicados

  Seleccione No si no desea que se permitan entradas duplicadas en un fingerprint columna.

  Seleccione Sí si desea que se permitan entradas duplicadas en una fingerprint columna.

- Permitir JOIN de columnas con nombres diferentes

  Elija No si no quiere unirse fingerprint columnas con nombres diferentes.

  Elija Sí si quiere unirse fingerprint columnas con nombres diferentes.

Para obtener más información acerca de los parámetros de computación criptográfica, consulte Parámetros de computación criptográfica.

Para obtener más información sobre cómo cifrar los datos para utilizarlos en AWS Clean Rooms, consulte[Preparación de tablas de datos cifrados con computación criptográfica para Clean Rooms](#).

> **Note**
>
> Compruebe estas configuraciones detenidamente antes de completar el siguiente paso. Tras crear la colaboración, solo podrá editar el nombre y la descripción de la colaboración y si los registros se almacenan en Amazon CloudWatch Logs.

f. Si desea habilitar la opción de Etiquetas para el recurso de colaboración, seleccione Añadir nueva etiqueta y, a continuación, introduzca el par de Clave y Valor.

g. Elija Siguiente.

5. Para el paso 2: especificar las capacidades de los miembros, haga lo siguiente:

   a. Para el análisis mediante consultas y trabajos, en Tipos de análisis admitidos, seleccione la casilla de verificación Trabajos.

      La casilla de verificación Consultas está seleccionada de forma predeterminada.

      i. Seleccione el miembro que puede ejecutar consultas y trabajos en la lista desplegable.

      ii. Seleccione el miembro que puede recibir los resultados de los análisis en la lista desplegable.

      > **Note**
      >
      > El miembro que crea la plantilla de PySpark análisis también debe ser el miembro que reciba los resultados.

   b. Si utiliza Clean Rooms ML, para modelar el aprendizaje automático mediante flujos de trabajo diseñados específicamente,

      i. (Opcional) Seleccione en la lista desplegable el miembro que puede recibir resultados de modelos entrenados.

      ii. (Opcional) Seleccione el miembro que puede recibir los resultados de la inferencia del modelo en la lista desplegable.

c. Vea las habilidades de los miembros en la sección Resolución de ID utilizando. AWS Entity Resolution

d. Elija Siguiente.

6. Para el paso 3: configurar el pago,

a. Para el análisis mediante consultas y trabajos, elija al miembro que pagará por las consultas y los trabajos.

Puede asignar al miembro que puede ejecutar consultas y trabajos para que sea el miembro que pague las consultas y los trabajos calculen los costes.

Puede asignar a un miembro diferente para que pague los costes de cálculo de las consultas y los trabajos.

b. En el caso del modelado de aprendizaje automático mediante flujos de trabajo diseñados específicamente, el creador del modelo similar configurado es el miembro que pagará por el modelado similar.

c. En el caso de la resolución de ID AWS Entity Resolution, el creador de la tabla de mapeo de ID es el miembro que pagará por la tabla de mapeo de ID.

d. Elija Siguiente.

7. Para el paso 4: configurar la membresía, elige una de las siguientes opciones:

Yes, join by creating membership now

1. Para la configuración de resultados predeterminada, para la configuración de resultados de consultas, si usted es el miembro que puede recibir los resultados,
   a. Seleccione la casilla de verificación Establecer la configuración predeterminada para las consultas. Para el destino de los resultados en Amazon S3, introduzca el destino de Amazon S3 o elija Browse S3 para seleccionar un bucket de S3.
   b. Como Formato de los resultados de la consulta, elija CSV o PARQUET.
   c. (Solo Spark) Para los archivos de resultados, elija Múltiple o Único.
   d. (Opcional) Para acceder al servicio, si desea enviar consultas que tarden hasta 24 horas en llegar a su destino de S3, active la casilla Añadir una función de servicio para atender las consultas que tarden hasta 24 horas en completarse.

      Las consultas de gran tamaño que tarden hasta 24 horas en completarse se enviarán a su destino de S3.

Si no selecciona la casilla, solo las consultas que se completen en un plazo de 12 horas se enviarán a su ubicación de S3.

e. Especifique los permisos de Acceso a servicios seleccionando Crear y usar un nuevo rol de servicio o Usar un rol de servicio existente.

Si eliges...	Entonces...
Crear y usar un nuevo rol de servicio	• AWS Clean Rooms crea un rol de servicio con la política requerida para esta tabla. • El Nombre del rol de servicio predeterminado es `cleanrooms-result-receiver-<timestamp>` • Debe tener permisos para crear roles y adjuntar políticas.

Si eliges...	Entonces...
Usar un rol de servicio existente	i. Seleccione un Nombre de rol de servicio existente en la lista desplegable. Si tiene permisos de listas de roles, se mostrará la lista de roles. Si no tiene permisos de listas de roles, puede ingresar el nombre de recurso de Amazon (ARN) del rol que desea usar. ii. Consulte el rol de servicio mediante la elección del enlace externo Ver en IAM. Si no hay ningún rol de servicio existente, la opción Usar un rol de servicio existente no estará disponible. De forma predeterminada, AWS Clean Rooms no intenta actualizar la política de roles existente para añadir los permisos necesarios.

> **Note**
>
> - AWS Clean Rooms requiere permisos para realizar consultas de acuerdo con las reglas de análisis. Para obtener más información sobre los permisos para AWS Clean Rooms, consulteAWS políticas gestionadas para AWS Clean Rooms.
> - Si el rol no tiene permisos suficientes AWS Clean Rooms, recibirá un mensaje de error que indica que el rol no tiene permisos suficientes para ello AWS Clean Rooms. Debe agregar la política de rol antes de continuar.

- Si no puede modificar la política de roles, recibirá un mensaje de error que indica que no ha AWS Clean Rooms podido encontrar la política para el rol de servicio.

2. Para ver los resultados de Job,

   Example

   Por ejemplo: **`s3://bucket/prefix`**

   a. Seleccione la casilla Establecer la configuración predeterminada para los trabajos y, a continuación, especifique el destino de los resultados en Amazon S3 introduciendo el destino de S3 o elija Browse S3 para seleccionarlos de una lista de buckets de S3 disponibles.

   b. Especifique los permisos de acceso al servicio eligiendo un nombre de rol de servicio existente en la lista desplegable.

3. En la configuración de Logs, elige una de las siguientes opciones para el almacenamiento de CloudWatch registros en Amazon Logs:

   > **Note**
   >
   > La sección de configuración de registros aparece si decide habilitar el registro de consultas.

   a. Selecciona Activar y los registros de consultas relevantes para ti se almacenarán en tu cuenta de Amazon CloudWatch Logs.

   Cada miembro solo puede recibir los registros de las consultas que haya iniciado o que contengan sus datos.

   El miembro que puede recibir los resultados también recibe los registros de todas las consultas que se ejecutan en una colaboración, aunque no se accede a los datos en una consulta.

   En Tipos de registro compatibles, elige entre los tipos de registro que el creador de la colaboración ha decidido admitir:

En Tipos de registro compatibles, las casillas de verificación Registros de consultas y Registros de trabajos están activadas de forma predeterminada.

> **Note**
>
> Tras activar el registro de análisis, el almacenamiento de registros puede tardar unos minutos en configurarse y empezar a recibir registros en Amazon CloudWatch Logs. Durante este breve período, el miembro que puede realizar consultas puede ejecutar consultas que, en realidad, no enviarán registros.

b. Selecciona Desactivar y los registros de consultas relevantes para ti no se almacenarán en tu cuenta de Amazon CloudWatch Logs.

4. Si quieres activar las etiquetas de membresía para el recurso de membresía, selecciona Añadir nueva etiqueta y, a continuación, introduce el par clave y valor.
5. Si usted es el miembro que paga por Query compute, Job compute o ambos, indique su aceptación seleccionando la casilla Acepto pagar los costos de procesamiento en esta colaboración.

> **Note**
>
> Debe seleccionar esta casilla de verificación para continuar.
> Para obtener más información sobre cómo se calculan los precios, consulte Precios para AWS Clean Rooms.

Si es el miembro que paga los costes de cálculo de las consultas, pero no el miembro que puede realizar la consulta, se recomienda configurar un presupuesto AWS Clean Rooms y recibir notificaciones cuando se alcance el presupuesto máximo. AWS Budgets Para obtener más información sobre cómo configurar un presupuesto, consulte Administración de costos con AWS Budgets en la Guía del usuario de AWS Cost Management . Para obtener más información sobre cómo configurar notificaciones, consulte Creación de un tema de Amazon SNS para las notificaciones de presupuesto en la Guía del usuario de AWS Cost Management . Si se ha alcanzado el presupuesto máximo, puede ponerse en contacto con el miembro que puede realizar consultas o abandonar la colaboración. Si abandona la colaboración, no se podrán ejecutar más consultas y, por lo tanto, ya no se le facturarán costos de cómputo de consultas.

6. Elija Siguiente.

   Se crean tanto la colaboración como su pertenencia a ella.

   Su estado en la colaboración es activo.

No, I will create a membership later

1. Elija Siguiente.

   Se crea solo la colaboración.

   Su estado en la colaboración es inactivo.

8. Para el paso 5: revisar y crear, haga lo siguiente:

   a. Revise las selecciones que realizó en los pasos anteriores y edítelas si es necesario.

   b. Elija una de las opciones.

Si ha elegido...	Entonces seleccione...
Crear una pertenencia en la colaboración (Sí, unirme y crear pertenencia ahora)	Crear colaboración y pertenencia
Crear la colaboración y no crear una pertenencia en este momento (No, crearé una pertenencia más tarde)	Crear colaboración

Una vez que la colaboración se haya creado correctamente, podrá ver la página de detalles de la colaboración en Colaboraciones.

Ya puede hacer lo siguiente:

- Prepare la tabla de datos para analizarla AWS Clean Rooms. (Opcional si desea analizar sus propios datos de eventos o si desea consultar datos de identidad).
- Asociar la tabla configurada a su colaboración (Opcional si quieres analizar los datos de tus propios eventos).
- Agregue una regla de análisis para la tabla configurada. (Opcional si quieres analizar los datos de tus propios eventos).

- Crear una pertenencia y unirse a una colaboración. (Opcional si ya ha creado una suscripción).
- Invita a los miembros a unirse a la colaboración.

Crear una colaboración para el modelado de aprendizaje automático

En este procedimiento, usted, como creador de la colaboración, realiza las siguientes tareas:

- Cree una colaboración.
- Invita a uno o más miembros a la colaboración.
- Asigne habilidades a los miembros, como
  - Miembro que puede realizar consultas
  - Miembro que puede recibir los resultados
  - Miembro que puede recibir resultados de modelos entrenados
  - Miembro que puede generar resultados a partir de la inferencia del modelo

  Si el creador de la colaboración es también el miembro que puede recibir los resultados, especificará el destino y el formato de los resultados. También proporcionan una función de servicio Amazon Resource Name (ARN) para escribir los resultados en el destino de los resultados.
- Configure qué miembro es responsable de pagar los costos de cómputo, la capacitación del modelo y los costos de inferencia del modelo en la colaboración.

Antes de comenzar, asegúrese de que cumple los siguientes requisitos previos:

- Ha determinado el tipo de motor de análisis que desea utilizar.
- Tienes el nombre y el Cuenta de AWS ID de cada miembro al que quieres invitar a la colaboración.
- Tiene permiso para compartir el nombre y la Cuenta de AWS identificación de cada miembro con todos los miembros de la colaboración.

> **Note**
>
> No puedes añadir más miembros después de crear la colaboración.

Para obtener información sobre cómo crear una colaboración mediante la AWS SDKs, consulta la referencia de la AWS Clean Rooms API.

Para crear una colaboración para el modelado de aprendizaje automático

1. Inicie sesión en la AWS Clean Rooms consola AWS Management Console y ábrala con la Cuenta de AWS que funcionará como creador de la colaboración.
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. En la esquina superior derecha, elija Crear colaboración.
4. En Paso 1: Definir la colaboración, haga lo siguiente:

    a. En Detalles, introduzca el Nombre y la Descripción de la colaboración.

    Esta información estará visible para los miembros de la colaboración que estén invitados a participar en la colaboración. La información de Nombre y Descripción les ayudará a entender a qué se refiere la colaboración.

    b. Para el motor de análisis, elige Spark.

    c. En Miembros:

        i. En Miembro 1: usted, introduzca su Nombre de miembro para mostrar tal y como desee que aparezca en la colaboración.

        > **Note**
        >
        > Tu Cuenta de AWS ID se incluye automáticamente en el Cuenta de AWS ID de miembro.

        ii. En el caso del miembro 2, introduzca el nombre visible y el Cuenta de AWS ID de miembro del miembro al que desee invitar a la colaboración.

        El Nombre del miembro para mostrar y el ID de Cuenta de AWS del miembro estarán visibles para todos los invitados a participar en la colaboración. Una vez ingresados y guardados los valores de estos campos, no se podrán editar.

> **Note**
>
> Debe informar al miembro de la colaboración de que la información de los campos ID de Cuenta de AWS del miembro y Nombre del miembro para mostrar estará visible para todos los colaboradores activos de la colaboración.

iii. Si desea añadir otro miembro, elija Añadir otro miembro. A continuación, introduzca el nombre de usuario y el Cuenta de AWS ID de miembro de cada miembro que pueda aportar los datos que desee invitar a la colaboración.

d. Si desea habilitar el registro de análisis, seleccione la casilla Habilitar el registro de análisis y, a continuación, en Tipos de registro compatibles, elija Registros a partir de consultas.

e. (Opcional) Si desea habilitar la capacidad de computación criptográfica, active la casilla de verificación Habilitar la computación criptográfica.

i. Elija los siguientes parámetros de cobertura criptográfica:

- Permitir plaintext columnas

  Seleccione No si necesita tablas totalmente cifradas.

  Elija Sí si lo desea cleartext columnas permitidas en la tabla cifrada.

  Para ejecutar SUM o AVG en determinadas columnas, las columnas deben estar en cleartext.

- Preservar NULL valores

  Elija No si no quiere conservar NULL valores. NULL los valores no aparecerán como NULL en una tabla cifrada.

  Elija Sí si desea conservar NULL valores. NULL los valores aparecerán como NULL en una tabla cifrada.

ii. Elija los siguientes parámetros de toma de huellas digitales:

- Permitir duplicados

  Seleccione No si no desea que se permitan entradas duplicadas en un fingerprint columna.

Seleccione Sí si desea que se permitan entradas duplicadas en una fingerprint columna.

- Permitir JOIN de columnas con nombres diferentes

  Elija No si no quiere unirse fingerprint columnas con nombres diferentes.

  Elija Sí si quiere unirse fingerprint columnas con nombres diferentes.

Para obtener más información acerca de los parámetros de computación criptográfica, consulte Parámetros de computación criptográfica.

Para obtener más información sobre cómo cifrar los datos para utilizarlos en AWS Clean Rooms, consultePreparación de tablas de datos cifrados con computación criptográfica para Clean Rooms.

> **Note**
>
> Compruebe estas configuraciones detenidamente antes de completar el siguiente paso. Tras crear la colaboración, solo podrá editar el nombre y la descripción de la colaboración y si los registros se almacenan en Amazon CloudWatch Logs.

f. Si desea habilitar la opción de Etiquetas para el recurso de colaboración, seleccione Añadir nueva etiqueta y, a continuación, introduzca el par de Clave y Valor.

g. Elija Siguiente.

5. Para el paso 2: especifique las capacidades de los miembros,

   a. Para el análisis mediante consultas y trabajos, en la sección Tipos de análisis admitidos, deje seleccionada la casilla Consultas.

   b. En Ejecutar consultas, elija el miembro que iniciará el entrenamiento del modelo

   c. En Recibir los resultados de los análisis, elija uno o más miembros que recibirán los resultados de la consulta.

   d. Para el modelado de aprendizaje automático mediante flujos de trabajo diseñados específicamente,

      i. En Recibir resultados de modelos entrenados, elija el miembro que recibirá los resultados del modelo entrenado, incluidos los artefactos y las métricas del modelo.

ii. En Recibir los resultados de la inferencia del modelo, elija el miembro que recibirá los resultados de la inferencia del modelo.

e. Vea las habilidades de los miembros bajo la resolución de ID utilizando. AWS Entity Resolution

6. Para el paso 3: configurar el pago, en el caso del análisis mediante consultas, realice una de las siguientes acciones en función de su objetivo.

Su objetivo	Acción recomendada
Designar al miembro que puede Ejecutar las consultas como miembro que paga los costos de computación de las consultas	1. Elija el miembro que pagará por las consultas para que sea el mismo miembro que puede ejecutar las consultas. 2. Elija Siguiente.
Designar a otro miembro que pague los costos de computación de las consultas	1. Elige a ti mismo como el miembro que pagará las consultas. 2. Elija Siguiente.

En el caso del modelado de aprendizaje automático mediante flujos de trabajo diseñados específicamente, el creador del modelo similar configurado es el miembro que pagará por el modelado similar.

En el caso de la resolución de ID AWS Entity Resolution, el creador de la tabla de mapeo de ID es el miembro que pagará por la tabla de mapeo de ID.

7. Para el paso 4: configurar la membresía, elige una de las siguientes opciones:

Yes, join by creating membership now

1. Para la configuración de resultados predeterminada, para la configuración de resultados de consultas, si usted es el miembro que puede recibir los resultados,
   a. Para el destino de los resultados en Amazon S3, introduzca el destino de Amazon S3 o elija Browse S3 para seleccionar un bucket de S3.
   b. Como Formato de los resultados de la consulta, elija CSV o PARQUET.
   c. (Solo Spark) Para los archivos de resultados, elija Múltiple o Único.

d. (Opcional) Para acceder al servicio, si desea enviar consultas que tarden hasta 24 horas en llegar a su destino de S3, active la casilla Añadir una función de servicio para atender las consultas que tarden hasta 24 horas en completarse.

   Las consultas de gran tamaño que tarden hasta 24 horas en completarse se enviarán a su destino de S3.

   Si no selecciona la casilla, solo las consultas que se completen en un plazo de 12 horas se enviarán a su ubicación de S3.

e. Especifique los permisos de Acceso a servicios seleccionando Crear y usar un nuevo rol de servicio o Usar un rol de servicio existente.

Si eliges...	Entonces...
Crear y usar un nuevo rol de servicio	• AWS Clean Rooms crea un rol de servicio con la política requerida para esta tabla. • El Nombre del rol de servicio predeterminado es `cleanrooms-result-receiver-<timestamp>` • Debe tener permisos para crear roles y adjuntar políticas.

Si eliges...	Entonces...
Usar un rol de servicio existente	i. Seleccione un Nombre de rol de servicio existente en la lista desplegable. Si tiene permisos de listas de roles, se mostrará la lista de roles. Si no tiene permisos de listas de roles, puede ingresar el nombre de recurso de Amazon (ARN) del rol que desea usar. ii. Consulte el rol de servicio mediante la elección del enlace externo Ver en IAM. Si no hay ningún rol de servicio existente, la opción Usar un rol de servicio existente no estará disponible. De forma predeterminada, AWS Clean Rooms no intenta actualizar la política de roles existente para añadir los permisos necesarios.

Note

- AWS Clean Rooms requiere permisos para realizar consultas de acuerdo con las reglas de análisis. Para obtener más información sobre los permisos para AWS Clean Rooms, consulteAWS políticas gestionadas para AWS Clean Rooms.
- Si el rol no tiene permisos suficientes AWS Clean Rooms, recibirá un mensaje de error que indica que el rol no tiene permisos suficientes para ello AWS Clean Rooms. Debe agregar la política de rol antes de continuar.

- Si no puede modificar la política de roles, recibirá un mensaje de error que indica que no ha AWS Clean Rooms podido encontrar la política para el rol de servicio.

2. Para ver los resultados de Job,

   Example

   Por ejemplo: **`s3://bucket/prefix`**

   a. Seleccione la casilla Establecer la configuración predeterminada para los trabajos y, a continuación, especifique el destino de los resultados en Amazon S3 introduciendo el destino de S3 o elija Browse S3 para seleccionarlos de una lista de buckets de S3 disponibles.

   b. Especifique los permisos de acceso al servicio eligiendo un nombre de rol de servicio existente en la lista desplegable.

3. En la configuración de Logs, elige una de las siguientes opciones para el almacenamiento de CloudWatch registros en Amazon Logs:

   > **Note**
   >
   > La sección de configuración de registros aparece si decide habilitar el registro de consultas.

   a. Selecciona Activar y los registros de consultas relevantes para ti se almacenarán en tu cuenta de Amazon CloudWatch Logs.

   Cada miembro solo puede recibir los registros de las consultas que haya iniciado o que contengan sus datos.

   El miembro que puede recibir los resultados también recibe los registros de todas las consultas que se ejecutan en una colaboración, aunque no se accede a los datos en una consulta.

   En Tipos de registro compatibles, elige entre los tipos de registro que el creador de la colaboración ha decidido admitir:

En Tipos de registro compatibles, la casilla de verificación Consultar registros está activada de forma predeterminada.

> **Note**
>
> Tras activar el registro de análisis, el almacenamiento de registros puede tardar unos minutos en configurarse y empezar a recibir registros en Amazon CloudWatch Logs. Durante este breve período, el miembro que puede realizar consultas puede ejecutar consultas que, en realidad, no enviarán registros.

b. Selecciona Desactivar y los registros de consultas relevantes para ti no se almacenarán en tu cuenta de Amazon CloudWatch Logs.

4. Si desea habilitar la opción de Etiquetas para el recurso de pertenencia, seleccione Añadir nueva etiqueta y, a continuación, introduzca el par de Clave y Valor.
5. Si usted es el miembro que paga por Query Compute, indique su aceptación marcando la casilla Acepto pagar los costos de procesamiento de esta colaboración.

> **Note**
>
> Debe seleccionar esta casilla de verificación para continuar.
> Para obtener más información sobre cómo se calculan los precios, consulte Precios para AWS Clean Rooms.

Si es el miembro que paga los costes de cálculo de las consultas, pero no el miembro que puede realizar la consulta, se recomienda configurar un presupuesto AWS Clean Rooms y recibir notificaciones cuando se alcance el presupuesto máximo. AWS Budgets Para obtener más información sobre cómo configurar un presupuesto, consulte Administración de costos con AWS Budgets en la Guía del usuario de AWS Cost Management . Para obtener más información sobre cómo configurar notificaciones, consulte Creación de un tema de Amazon SNS para las notificaciones de presupuesto en la Guía del usuario de AWS Cost Management . Si se ha alcanzado el presupuesto máximo, puede ponerse en contacto con el miembro que puede realizar consultas o abandonar la colaboración. Si abandona la colaboración, no se podrán ejecutar más consultas y, por lo tanto, ya no se le facturarán costos de cómputo de consultas.

6. Elija Siguiente.

Se crean tanto la colaboración como su pertenencia a ella.

Su estado en la colaboración es activo.

No, I will create a membership later

1. Elija Siguiente.

   Se crea solo la colaboración.

   Su estado en la colaboración es inactivo.

8. Para el paso 5: revisar y crear, haga lo siguiente:

   a. Revise las selecciones que realizó en los pasos anteriores y edítelas si es necesario.

   b. Elija una de las opciones.

Si ha elegido...	Entonces seleccione...
Crear una pertenencia en la colaboración (Sí, unirme y crear pertenencia ahora)	Crear colaboración y pertenencia
Crear la colaboración y no crear una pertenencia en este momento (No, crearé una pertenencia más tarde)	Crear colaboración

Crear una pertenencia y unirse a una colaboración

Una suscripción es un recurso que se crea cuando un miembro se une a una colaboración en AWS Clean Rooms.

Puedes unirte a una colaboración como

- miembro que puede realizar consultas
- miembro que puede realizar consultas y trabajos
- miembro que puede recibir los resultados de una consulta o un trabajo
- miembro que paga los costos de cálculo de la consulta
- miembro que paga por consultas y trabajos

Todos los miembros pueden contribuir con datos.

Para obtener información sobre cómo crear una membresía y unirse a una colaboración mediante la AWS SDKs, consulte la referencia de la AWS Clean Rooms API.

En este procedimiento, el miembro invitado se une a la colaboración creando un recurso de suscripción.

Si el miembro invitado es el miembro que puede recibir los resultados, especificará el destino y el formato de los resultados. También proporcionan un ARN de rol de servicio para escribir en el destino de los resultados.

Si el miembro invitado es el responsable de pagar los costes de cálculo, acepta sus responsabilidades de pago antes de unirse a la colaboración.

Para crear una pertenencia y unirse a una colaboración

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con tu miembro Cuenta de AWS.
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. En la pestaña Disponibles para combinar, en Colaboraciones disponibles para combinar, elija el Nombre de la colaboración.
4. En la página de detalles de la colaboración, en la sección Descripción general, consulta los detalles de la colaboración, incluidos los detalles de tu miembro y una lista de los demás miembros.

   Compruebe que Cuenta de AWS IDs los miembros de cada uno de los miembros de la colaboración son aquellos con los que piensa iniciar la colaboración.
5. Elija Crear pertenencia.
6. En la página Crear membresía, en la sección Descripción general, consulta el nombre de la colaboración, la descripción de la colaboración, el Cuenta de AWS ID del creador de la colaboración, los detalles de tu miembro y el Cuenta de AWS ID del miembro que pagará las consultas.
7. Si el creador de la colaboración ha decidido habilitar el registro de análisis, elija una de las siguientes opciones para el almacenamiento de CloudWatch registros en Amazon Logs:

Si selecciona...	Entonces...
Activar	Los registros relevantes para usted se almacenan en Amazon CloudWatch Logs. Cada miembro solo puede recibir los registros de las consultas que haya iniciado o que contengan sus datos. El miembro que puede recibir los resultado s también recibe los registros de todos los análisis realizados en una colaboración, incluso si no se accede a sus datos en un análisis. En Tipos de registro compatibles, elige entre los tipos de registro que el creador de la colaboración ha decidido admitir: 1. Si desea recibir los registros generados a partir de consultas SQL, seleccione la casilla Registros de consultas. 2. Si desea recibir los registros generados a partir de los trabajos que se utilizan PySpark, seleccione la casilla Registros de los trabajos.
Desactivar	Los registros de consultas relevantes para ti no se almacenan en tu cuenta de Amazon CloudWatch Logs.

> **Note**
>
> Tras activar el registro de análisis, el almacenamiento de registros puede tardar unos minutos en configurarse y empezar a recibir registros en Amazon CloudWatch Logs.

Durante este breve período, el miembro que puede realizar consultas puede ejecutar consultas que, en realidad, no enviarán registros.

8. Si entre las capacidades de sus miembros está la de Recibir resultados, la configuración de resultados predeterminada es la siguiente:

   a. Para los resultados de las consultas, seleccione la casilla Establecer la configuración predeterminada para las consultas y, a continuación, especifique el destino de los resultados en Amazon S3 introduciendo el destino de S3 o elija Browse S3 para seleccionarlos de una lista de buckets de S3 disponibles.

      Example

      Por ejemplo: **`s3://bucket/prefix`**

      i. Para el formato de resultados, elija CSV o PARQUET.

      ii. (Solo Spark) Para los archivos de resultados, elija Múltiple o Único.

      iii. (Opcional) Para acceder al servicio, si desea enviar consultas que tarden hasta 24 horas en llegar a su destino de S3, active la casilla Añadir una función de servicio para atender las consultas que tarden hasta 24 horas en completarse.

      Las consultas grandes que tarden hasta 24 horas en completarse se enviarán a su destino de S3.

      Si no selecciona la casilla de verificación, solo las consultas que se completen en un plazo de 12 horas se enviarán a su ubicación de S3.

      > **Note**
      >
      > Debe seleccionar un rol de servicio existente o tener permisos para crear uno nuevo. Para obtener más información, consulte Creación de un rol de servicio para recibir resultados.

      iv. Especifique los permisos de Acceso a servicios seleccionando Crear y usar un nuevo rol de servicio o Usar un rol de servicio existente.

Create and use a new service role

- AWS Clean Rooms crea un rol de servicio con la política requerida para esta tabla.
- El Nombre del rol de servicio predeterminado es `cleanrooms-result-receiver-<timestamp>`
- Debe tener permisos para crear roles y adjuntar políticas.

Use an existing service role

1. Seleccione un Nombre de rol de servicio existente en la lista desplegable.

   Si tiene permisos de listas de roles, se mostrará la lista de roles.

   Si no tiene permisos de listas de roles, puede ingresar el nombre de recurso de Amazon (ARN) del rol que desea usar.
2. Consulte el rol de servicio mediante la elección del enlace externo Ver en IAM.

   Si no hay ningún rol de servicio existente, la opción Usar un rol de servicio existente no estará disponible.

   De forma predeterminada, AWS Clean Rooms no intenta actualizar la política de roles existente para añadir los permisos necesarios.

> **Note**
>
> - AWS Clean Rooms requiere permisos para realizar consultas de acuerdo con las reglas de análisis. Para obtener más información sobre los permisos para AWS Clean Rooms, consulteAWS políticas gestionadas para AWS Clean Rooms.
> - Si el rol no tiene permisos suficientes AWS Clean Rooms, recibirá un mensaje de error que indica que el rol no tiene permisos suficientes AWS Clean Rooms. Debe agregar la política de rol antes de continuar.

- Si no puede modificar la política de roles, recibirá un mensaje de error que indica que no ha AWS Clean Rooms podido encontrar la política para el rol de servicio.

b. Para los resultados de los trabajos, seleccione la casilla Establecer la configuración predeterminada para los trabajos y, a continuación, especifique el destino de los resultados en Amazon S3 introduciendo el destino de S3 o elija Browse S3 para seleccionarlos de una lista de depósitos de S3 disponibles.

   Example

   Por ejemplo: **`s3://bucket/prefix`**

   - Especifique los permisos de acceso al servicio seleccionando el nombre de un rol de servicio existente en la lista desplegable.

9. Si desea habilitar la opción de Etiquetas para el recurso de pertenencia, seleccione Añadir nueva etiqueta y, a continuación, introduzca el par de Clave y Valor.
10. Si el creador de la colaboración lo ha designado como miembro responsable del pago de las consultas o de las consultas y los trabajos, indique su aceptación marcando la casilla Acepto pagar los costes de procesamiento de esta colaboración.

> **Note**
>
> Debe seleccionar esta casilla de verificación para continuar.
> Para obtener más información sobre cómo se calculan los precios, consulte Precios para AWS Clean Rooms.

Si es el miembro que paga los costes de cálculo de las consultas o el miembro que paga los costes de las consultas y el trabajo, pero no el miembro que puede realizar las consultas AWS Budgets , se recomienda configurar un presupuesto AWS Clean Rooms y recibir notificaciones cuando se alcance el presupuesto máximo. Para obtener más información sobre cómo configurar un presupuesto, consulte Administración de costos con AWS Budgets en la Guía del usuario de AWS Cost Management . Para obtener más información sobre cómo configurar notificaciones, consulte Creación de un tema de Amazon SNS para las notificaciones de presupuesto en la Guía del usuario de AWS Cost Management . Si se ha alcanzado el presupuesto máximo, puedes ponerte en contacto con el miembro que puede gestionar las

consultas y los trabajos o abandonar la colaboración. Si abandona la colaboración, no se podrán ejecutar más consultas y, por lo tanto, ya no se le facturarán costos de cómputo de consultas.

11. Si tiene la certeza de que desea crear una pertenencia y unirse a la colaboración, consulte Crear pertenencia.

Se le concede acceso de lectura a los metadatos de la colaboración. Esto incluye información como el nombre para mostrar y la descripción de la colaboración, además de todos los nombres y Cuenta de AWS IDs de los demás miembros.

Ya puede hacer lo siguiente:

- Preparar su tabla de datos para que se consulte en AWS Clean Rooms (Opcional si desea consultar sus propios datos de eventos o si desea consultar datos de identidad).
- Asocie la tabla configurada a la colaboración: si desea consultar datos de eventos.
- Agregue una regla de análisis para la tabla configurada: si desea consultar datos de eventos.
- Cree y asocie un nuevo espacio de nombres de ID, si desea crear una tabla de mapeo de ID para consultar datos de identidad.

Para obtener información sobre cómo abandonar una colaboración, consulte Abandonar una colaboración.

Editar colaboraciones

Como creador de colaboraciones, puedes editar las distintas partes de una colaboración.

Para obtener información sobre cómo editar una colaboración mediante AWS SDKs, consulte la referencia de la API de salas limpias de AWS.

Temas

- Edición del nombre y la descripción de la colaboración
- Actualice el motor de análisis de colaboración
- Desactive el almacenamiento de registros
- Edita la configuración de los registros de colaboración
- Editar etiquetas de colaboración
- Edición de etiquetas de pertenencia

- Editar las etiquetas de tablas asociadas
- Edición de etiquetas de plantilla de análisis
- Edición de las etiquetas de la política de privacidad diferencial

Edición del nombre y la descripción de la colaboración

Una vez creada una colaboración, solo su nombre y descripción serán editables.

Para editar el nombre y la descripción de la colaboración

1. Inicie sesión en la AWS Clean Rooms consola AWS Management Console y ábrala con la suya Cuenta de AWS (si aún no lo ha hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Elija la colaboración que ha creado.
4. En la página de detalles de la colaboración, elija Acciones y, a continuación, elija Editar colaboración.
5. En la página Editar colaboración, en Detalles, edita el nombre y la descripción de la colaboración.
6. Seleccione Save changes (Guardar cambios).

Actualice el motor de análisis de colaboración

Tras crear la colaboración, puedes cambiar el motor de análisis de AWS Clean Rooms SQL a Spark.

> **Note**
>
> Cambiar el motor de análisis de AWS Clean Rooms SQL a Spark podría interrumpir los flujos de trabajo existentes.

Para actualizar el motor de análisis de colaboración

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.

3. Elija la colaboración que ha creado.
4. En la página de detalles de la colaboración, elija Acciones y, a continuación, elija Editar colaboración.
5. En la página Editar colaboración, para el motor de análisis,

   - Si selecciona AWS Clean Rooms SQL, elija Spark.
   - Si selecciona Spark, elija Enviar un ticket de soporte para enviar un ticket de soporte para cambiar el motor de análisis a AWS Clean Rooms SQL.

   > **Note**
   >
   > AWS Clean Rooms dejará de ser compatible con el antiguo motor de análisis SQL de Clean Rooms el 17 de diciembre de 2025. Antes del 16 de julio de 2025, debe solicitar un aumento del límite a través de AWS Customer Support para crear nuevas colaboraciones basadas en el motor SQL de Clean Rooms. A partir del 16 de julio de 2025, ya no estará disponible la creación de nuevas colaboraciones basadas en el motor SQL de Clean Rooms.

6. Seleccione Save changes (Guardar cambios).

Desactive el almacenamiento de registros

Si has activado el registro de análisis, puedes editar si los registros de análisis se almacenan en tu cuenta de Amazon CloudWatch Logs.

Para desactivar el almacenamiento de registros

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Elige la colaboración que tenga activado el registro de análisis.
4. En la página de detalles de la colaboración, selecciona Acciones y, a continuación, selecciona Desactivar el almacenamiento de registros.

   > **Note**
   >
   > Aparece una advertencia que indica lo siguiente:

- Las consultas nuevas ya no se registrarán en su CloudWatch cuenta.
- Los registros existentes se conservarán de acuerdo con su configuración de retención actual.
- Si reactiva el registro en el futuro, solo se aplicará a las consultas realizadas después de la reactivación.
- Este cambio solo afecta a tus registros; la configuración de registro de los demás miembros del equipo permanece inalterada.

5. Seleccione Desactivar.

Edita la configuración de los registros de colaboración

Si has activado el registro de consultas, puedes editar si los registros de consultas se almacenan en tu cuenta de Amazon CloudWatch Logs.

Para editar la configuración de los registros de colaboración

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Elija la colaboración que ha creado.
4. En la página de detalles de la colaboración, realice una de las siguientes acciones:
   - Elija Acciones y, a continuación, elija Editar la configuración de los registros.
   - En la pestaña Registros, selecciona Editar la configuración de los registros.
5. En el modal Editar configuración de registros, para el almacenamiento de CloudWatch registros en Amazon Logs:
   - Si no quieres que los registros relevantes para ti se almacenen en tu cuenta de Amazon CloudWatch Logs, selecciona Activar.
   - Si quieres que los registros relevantes para ti se almacenen en tu cuenta de Amazon CloudWatch Logs, selecciona Activar.

     Solo puedes recibir los registros de las consultas que hayas iniciado o que contengan tus datos.

El miembro que puede recibir los resultados también recibe los registros de todas las consultas que se ejecutan en una colaboración, aunque no se accede a los datos en una consulta.

1. En Tipos de registro compatibles, elige entre los tipos de registro que el creador de la colaboración ha decidido admitir:
   - Si desea recibir los registros generados a partir de consultas SQL, seleccione la casilla Registros de consultas.
   - Si desea recibir los registros generados a partir de los trabajos que se utilizan PySpark, seleccione la casilla Registros de los trabajos.

6. Seleccione Save changes (Guardar cambios).

> **Note**
>
> Tras activar el registro, el almacenamiento de registros puede tardar unos minutos en configurarse y empezar a recibirlos en Amazon CloudWatch Logs. Durante este breve período, el miembro que puede realizar consultas puede ejecutar consultas que, en realidad, no enviarán registros.

Editar etiquetas de colaboración

Como creador de la colaboración, una vez creada una colaboración, puede administrar las etiquetas del recurso de colaboración.

Para editar las etiquetas de la colaboración

1. Inicia sesión en AWS Management Console y abre la AWS Clean Rooms consola con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Elija la colaboración que ha creado.
4. Seleccione una de las siguientes opciones:

Si usted es...	Entonces...
El creador de la colaboración y un miembro de la colaboración	Elija la pestaña Detalles.
El creador de la colaboración, pero no un miembro de la colaboración	Desplácese hacia abajo por la página hasta la sección Etiquetas.

5. Para ver los Detalles de la colaboración, elija Administrar etiquetas.
6. En la página Administrar etiquetas, puede hacer lo siguiente:
   - Para eliminar una etiqueta, elija Eliminar.
   - Para añadir una etiqueta, elija Añadir nueva etiqueta.
   - Para guardar los cambios, elija Guardar cambios.

Edición de etiquetas de pertenencia

Como creador de colaboraciones, una vez que haya creado una colaboración, puede administrar las etiquetas del recurso de pertenencia.

Para editar las etiquetas de pertenencia

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Elija la colaboración que ha creado.
4. Elija la pestaña Detalles.
5. En Detalles de pertenencia, elija Administrar etiquetas.
6. En la página Administrar etiquetas de pertenencia, puede hacer lo siguiente:
   - Para eliminar una etiqueta, elija Eliminar.
   - Para añadir una etiqueta, elija Añadir nueva etiqueta.
   - Para guardar los cambios, elija Guardar cambios.

Editar las etiquetas de tablas asociadas

Como creador de la colaboración, después de asociar tablas a una colaboración, puede administrar las etiquetas del recurso de tabla asociado.

Para editar las etiquetas de tabla asociadas

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con la tuya Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Elija la colaboración que ha creado.
4. Seleccione la pestaña Tablas.
5. En Tablas asociadas por usted, elija una tabla.
6. En la página de detalles de la tabla configurada, en Etiquetas, elija Administrar etiquetas.

   En la página Administrar etiquetas, puede hacer lo siguiente:

   - Para eliminar una etiqueta, elija Eliminar.
   - Para añadir una etiqueta, elija Añadir nueva etiqueta.
   - Para guardar los cambios, elija Guardar cambios.

Edición de etiquetas de plantilla de análisis

Como creador de la colaboración, una vez creada una colaboración, puede administrar las etiquetas del recurso de plantilla de análisis.

Para editar las etiquetas de pertenencia

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con la tuya Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Elija la colaboración que ha creado.
4. Elija la pestaña Plantillas.
5. En la sección Plantillas de análisis creadas por usted, elija la plantilla de análisis.
6. En la página de detalles de la tabla de plantillas de análisis, desplácese hacia abajo hasta la sección Etiquetas.

7. Elija Administrar etiquetas.
8. En la página Administrar etiquetas, puede hacer lo siguiente:

   - Para eliminar una etiqueta, elija Eliminar.
   - Para añadir una etiqueta, elija Añadir nueva etiqueta.
   - Para guardar los cambios, elija Guardar cambios.

Edición de las etiquetas de la política de privacidad diferencial

Como creador de la colaboración, una vez creada una colaboración, puede administrar las etiquetas del recurso de plantilla de análisis.

Para editar las etiquetas de pertenencia

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con la tuya Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Elija la colaboración que contiene la política de privacidad diferencial que desea editar.
4. Seleccione la pestaña Tablas.
5. En la pestaña Etiquetas, elija Administrar etiquetas.
6. En la página Administrar etiquetas, puede hacer lo siguiente:

   - Para eliminar una etiqueta, elija Eliminar.
   - Para añadir una etiqueta, elija Añadir nueva etiqueta.
   - Para guardar los cambios, elija Guardar cambios.

Eliminar colaboraciones

Como creador de la colaboración, puede eliminar una colaboración que haya creado.

> **Note**
>
> Al eliminar una colaboración, ni usted ni el resto de miembros pueden realizar consultas, recibir resultados o aportar datos. Cada miembro de la colaboración sigue teniendo acceso a sus propios datos como parte de su pertenencia.

Para eliminar una colaboración

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Elija la colaboración que desea eliminar.
4. En Acciones, selecciona Eliminar colaboración.
5. Confirme la eliminación y luego elija Eliminar.

Ver colaboraciones

Como creador de la colaboración, puede ver todas las colaboraciones que ha creado.

Para ver colaboraciones

1. Inicie sesión en AWS Management Console y abra la AWS Clean Rooms consola con su Cuenta de AWS (si aún no lo ha hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. En la página Colaboraciones, en Usadas por última vez, consulte las últimas 5 colaboraciones utilizadas.
4. En la pestaña Con pertenencia activa, consulte la lista de colaboraciones con pertenencia activa.

   Puede ordenar por Nombre, Fecha de creación de la pertenencia y Sus detalles de miembro.

   Puede utilizar la barra de búsqueda para buscar una colaboración.
5. En la pestaña Disponibles para unirse, consulte la lista de Colaboraciones disponibles para unirse.
6. En la pestaña Ya no están disponibles, consulte la lista de colaboraciones eliminadas y Pertenencias a colaboraciones que ya no están disponibles (pertenencias eliminadas).

Invitar a los miembros a una colaboración

Como creador de una colaboración, una vez creada una colaboración, puede enviar un enlace de invitación a los miembros que aparecen en la pestaña Miembros.

Para invitar a los miembros a una colaboración

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Elija la colaboración que ha creado.
4. Seleccione la pestaña Miembros.
5. En la tabla de miembros, selecciona el botón Copiar el enlace de invitación.

   Se copia el enlace de la invitación.
6. Pegue el enlace de la invitación en el método de comunicación seguro que prefiera y envíelo a cada miembro de la colaboración.

Supervisión de miembros

Como creador de una colaboración, una vez creada la colaboración, puede monitorizar el estado de todos los miembros en la pestaña Miembros.

Para comprobar el estado de un miembro

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Elija la colaboración que ha creado.
4. Seleccione la pestaña Miembros.
5. En la tabla de miembros, revise el estado de cada miembro.
6. En la tabla de capacidades de los miembros, revise qué miembros pueden realizar consultas, recibir resultados, aportar datos y realizar otras tareas.
7. En la tabla de configuración de pagos, revise qué miembros pagan por las consultas, las tablas de asignación de ID y los modelos de ML.

Eliminar un miembro de una colaboración

> **Note**
>
> Al eliminar a un miembro, también se eliminan todos sus conjuntos de datos asociados de la colaboración.

Para eliminar un miembro de una colaboración

1. Inicie sesión en AWS Management Console y abra la AWS Clean Rooms consola con su Cuenta de AWS (si aún no lo ha hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Elija la colaboración que ha creado.
4. Seleccione la pestaña Miembros.
5. Seleccione el botón de opción situado junto al miembro que desea eliminar.

   > **Note**
   >
   > El creador de una colaboración no puede elegir su propio ID de cuenta.

6. Elija Eliminar.
7. En el cuadro de diálogo, confirme la decisión de eliminar al miembro escribiendo **`confirm`** en el campo de entrada de texto.

   > **Note**
   >
   > Si elimina al miembro que paga los costos de computación de consultas, no se permite ejecutar más consultas en la colaboración.

Abandonar una colaboración

Como miembro de una colaboración, puede abandonar una colaboración eliminando tu pertenencia. Si es el creador de la colaboración, solo puede abandonarla si elimina la colaboración.

> **Note**
>
> Al eliminar su pertenencia, abandonará la colaboración y no podrá volver a unirse a ella. Si es el miembro que paga los costos de computación de consultas y elimina su pertenencia, no se permite que se ejecuten más consultas.

Para abandonar una colaboración

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. En Con pertenencia activa, elija la colaboración de la que es miembro.
4. Elija Acciones.
5. Seleccione Eliminar pertenencia.
6. En el cuadro de diálogo, confirme la decisión de abandonar la colaboración escribiendo **`confirm`** en el campo de entrada de texto y, a continuación, seleccione Vaciar y eliminar pertenencia.

   En la consola, verá un mensaje que indica que se ha eliminado la pertenencia.

   El creador de la colaboración ve el Estado de miembro como Abandonado.

Prepare tablas de datos en AWS Clean Rooms

> **Note**
>
> Puede preparar las tablas de datos antes o después de unirse a una colaboración. Una vez preparada una tabla, puede reutilizarla en varias colaboraciones siempre que sus necesidades de privacidad para esa tabla sean las mismas.

Como miembro de la colaboración, debe preparar las tablas de datos antes de que el miembro de la colaboración que pueda consultarlas pueda consultarlas. AWS Clean Rooms

Las tablas de datos que se utilizan para las consultas AWS Clean Rooms suelen ser del mismo tipo de tablas de datos que se utilizan para otras aplicaciones. Por ejemplo, los mismos tipos de conjuntos de datos se utilizan con Amazon Athena, Amazon EMR, Amazon Redshift Spectrum y Amazon. QuickSight

Puede consultar los datos en su formato original directamente desde cualquiera de las siguientes fuentes de datos:

- Amazon Simple Storage Service (Amazon S3)
- Amazon Athena
- Snowflake

AWS Clean Rooms accede al conjunto de datos en tiempo de ejecución de la consulta, lo que garantiza que los miembros que pueden realizar consultas accedan siempre a la mayor cantidad de up-to-date datos. Todos los datos que se lean temporalmente en una AWS Clean Rooms colaboración se eliminan una vez finalizada la consulta. Los resultados de la consulta se escriben en su bucket de Amazon S3.

Si su caso de uso implica la consulta de datos de identidad, consulte AWS Entity Resolution in AWS Clean Rooms.

Temas

- Formatos de datos para AWS Clean Rooms
- Apache Iceberg tablas en AWS Clean Rooms

- Preparación de tablas de datos para consultas en AWS Clean Rooms
- Preparación de tablas de datos cifrados con computación criptográfica para Clean Rooms
- Descifrar tablas de datos con el cliente de cifrado de C3R

Formatos de datos para AWS Clean Rooms

Para analizar los datos, los conjuntos de datos deben estar en un formato AWS Clean Rooms compatible.

Temas

- Formatos de datos compatibles para los trabajos PySpark
- Formatos de datos compatibles para consultas SQL
- Tipos de datos compatibles
- Tipos de compresión de archivos para AWS Clean Rooms
- Cifrado del lado del servidor para AWS Clean Rooms

Formatos de datos compatibles para los trabajos PySpark

AWS Clean Rooms admite los siguientes formatos estructurados para ejecutar PySpark trabajos.

- Parquet
- OpenCSV
- JSON

Formatos de datos compatibles para consultas SQL

AWS Clean Rooms admite diferentes formatos estructurados para ejecutar consultas SQL, dependiendo de si eliges el motor de análisis SQL de Spark o el motor de análisis de AWS Clean Rooms SQL.

Spark SQL analytics engine

- Tablas de Apache Iceberg
- Parquet
- OpenCSV

- JSON

AWS Clean Rooms SQL analytics engine

- Tablas de Apache Iceberg
- Parquet
- RCFile
- TextFile
- SequenceFile
- RegexSerde
- OpenCSV
- AVRO
- JSON

> **Note**
>
> Un valor `timestamp` de un archivo de texto debe estar en formato `yyyy-MM-dd HH:mm:ss.SSSSSS`. Por ejemplo: `2017-05-01 11:30:59.000000`.

Recomendamos usar un formato de archivo de almacenamiento en columnas, como Apache Parquet. Con un formato de archivo de almacenamiento en columnas, puede minimizar el movimiento de datos seleccionando solo las columnas que necesite. Para obtener un rendimiento óptimo, los objetos grandes deben dividirse en objetos de 100 MB a 1 GB.

Tipos de datos compatibles

AWS Clean Rooms admite diferentes tipos, en función de si eliges el motor de análisis SQL de Spark o el motor de análisis de AWS Clean Rooms SQL.

Spark SQL analytics engine

- ARRAY
- BIGINT
- BOOLEAN

- BYTE
- CHAR
- DATE
- DECIMAL
- FLOAT
- INTEGER
- INTERVAL
- LONG
- MAP
- REAL
- SHORT
- SMALLINT
- STRUCT
- TIME
- TIMESTAMP_LTZ
- TIMESTAMP_NTZ
- TINYINT
- VARCHAR

Para obtener más información, consulte Tipos de datos en la referencia de SQL.AWS Clean Rooms

AWS Clean Rooms SQL

- ARRAY
- BIGINT
- BOOLEAN
- CHAR
- DATE
- DECIMAL
- DOUBLE PRECISION
- INTEGER

- MAP
- REAL
- SMALLINT
- STRUCT
- SUPER
- TIME
- TIMESTAMP
- TIMESTAMPTZ
- TIMETZ
- VARBYTE
- VARCHAR

Para obtener más información, consulte Tipos de datos en la referencia de AWS Clean Rooms SQL.

Tipos de compresión de archivos para AWS Clean Rooms

Para reducir el espacio de almacenamiento, mejorar el rendimiento y minimizar costos, es muy recomendable comprimir los archivos de datos.

AWS Clean Rooms reconoce los tipos de compresión de archivos en función de la extensión del archivo y admite los tipos y extensiones de compresión que se muestran en la tabla siguiente.

Algoritmo de compresión	Extensión de archivo
GZIP	.gz
Bzip2	.bz2
Snappy	.snappy

Puede aplicar compresión a distintos niveles. En la mayoría de los casos, se comprime un archivo completo o se comprimen bloques individuales dentro de un archivo. La compresión de formatos de columna por archivo no ofrece beneficios en términos de rendimiento.

Cifrado del lado del servidor para AWS Clean Rooms

> **Note**
>
> El cifrado en el servidor no reemplaza a la computación criptográfica en los casos de uso en los que esta última es obligatoria.

AWS Clean Rooms descifra de forma transparente los conjuntos de datos cifrados mediante las siguientes opciones de cifrado:

- SSE-S3: cifrado en el servidor mediante una clave de cifrado AES-256 administrada por Amazon S3
- SSE-KMS: cifrado del lado del servidor con claves administradas por AWS Key Management Service

Para usar SSE-S3, el rol de AWS Clean Rooms servicio utilizado para asociar la tabla configurada a la colaboración debe tener permisos de descifrado por KMS. Para usar SSE-KMS, la política de claves de KMS también debe permitir que la función de servicio descifre. AWS Clean Rooms

AWS Clean Rooms no admite el cifrado del lado del cliente de Amazon S3. Para obtener más información sobre el cifrado en el servidor, consulte Protección de datos con el cifrado del lado del servidor en la Guía del usuario de Amazon Simple Storage Service.

Apache Iceberg tablas en AWS Clean Rooms

Apache Iceberg es un formato de tabla de código abierto para lagos de datos. AWS Clean Rooms puede utilizar las estadísticas almacenadas en Apache Iceberg metadatos para optimizar los planes de consultas y reducir el escaneo de archivos durante el procesamiento de consultas en sala limpia. Para obtener más información, consulte la documentación de Apache Iceberg.

Tenga en cuenta lo siguiente cuando las utilice AWS Clean Rooms con tablas Iceberg:

- Tablas Apache Iceberg para S3: Apache Iceberg las tablas deben definirse en AWS Glue Data Catalog función de la implementación del catálogo Glue de código abierto.
- Tablas Apache Iceberg para Athena: para obtener más información, consulte -iceberg.html https://docs.aws.amazon.com/athena/ latest/ug/querying

- Tablas Apache Iceberg para Snowflake: para obtener más información, consulte user-guide/tables-iceberg https://docs.snowflake.com/en/
- Formato de archivo Parquet: AWS Clean Rooms solo admite tablas Iceberg en el formato de archivo de datos Parquet.
- Compresión GZIP y Snappy: AWS Clean Rooms admite Parquet con GZIP y Snappy compresión.
- Versiones Iceberg: AWS Clean Rooms permite ejecutar consultas en tablas Iceberg de las versiones 1 y 2.
- Particiones: no es necesario añadir particiones manualmente a su Apache Iceberg tablas en AWS Glue. AWS Clean Rooms detecta nuevas particiones en Apache Iceberg tablas automáticamente y no es necesaria ninguna operación manual para actualizar las particiones en la definición de la tabla. Las particiones de Iceberg aparecen como columnas normales en el esquema de tabla de AWS Clean Rooms , y no por separado como una clave de partición en el esquema de la tabla configurada.
- Limitaciones
  - Solo tablas de Iceberg nuevas

    Apache Iceberg tablas convertidas desde Apache Parquet no se admiten tablas.
  - Consultas de viaje en el tiempo

    AWS Clean Rooms no admite consultas sobre viajes en el tiempo con Apache Iceberg tablas.
  - Versión 2 del motor Athena

    Iceberg no se admiten las tablas creadas con la versión 2 del motor Athena.
  - Formatos de archivo

    Avro y no se admiten los formatos de archivo optimizados con filas y columnas (ORC).
  - Compresión

    Zstandard Compresión (Zstd) para Parquet no se admite.

Tipos de datos admitidos para las tablas de Iceberg

AWS Clean Rooms puede consultar Iceberg tablas que contienen los siguientes tipos de datos:

- BOOLEAN
- DATE

- DECIMAL
- DOUBLE
- FLOAT
- INT
- LIST
- LONG
- MAP
- STRING
- STRUCT
- TIMESTAMP WITHOUT TIME ZONE

Para obtener más información sobre los tipos de datos de Iceberg, consulte los esquemas para Iceberg en la documentación de Apache.

Preparación de tablas de datos para consultas en AWS Clean Rooms

Si no es necesario que traiga sus propios datos, puede omitir este procedimiento.

Si su caso de uso implica la consulta de datos de identidad, consulte AWS Entity Resolution in AWS Clean Rooms.

Para obtener más información sobre los formatos de datos que puede utilizar, consulteFormatos de datos para AWS Clean Rooms.

Temas

- Preparación de tablas de datos en Amazon S3
- Preparación de tablas de datos en Amazon Athena
- Preparación de tablas de datos en Snowflake

Preparación de tablas de datos en Amazon S3

Puede analizar las tablas de datos que se han catalogado AWS Glue y almacenado en Amazon S3. Si sus tablas de datos ya están catalogadas AWS Glue, vaya a. Crear una tabla configurada en AWS Clean Rooms

La preparación de las tablas de datos en Amazon S3 implica los siguientes pasos:

Temas

- Paso 1: completar los requisitos previos
- Paso 2: (opcional) preparar sus datos para la computación criptográfica
- Paso 3: cargar la tabla de datos en Amazon S3
- Paso 4: Crear una AWS Glue tabla
- Paso 5: Siguientes pasos

Paso 1: completar los requisitos previos

Para preparar las tablas de datos para usarlas con AWS Clean Rooms ellas, debe cumplir los siguientes requisitos previos:

- Las tablas de datos se guardan como uno de los formatos de datos compatibles para AWS Clean Rooms.
- Las tablas de datos están catalogadas AWS Glue y utilizan los tipos de datos compatibles. AWS Clean Rooms
- Todas las tablas de datos se almacenan en Amazon Simple Storage Service (Amazon S3), en el Región de AWS mismo lugar en el que se creó la colaboración.
- AWS Glue Data Catalog Se encuentra en la misma región en la que se creó la colaboración.
- AWS Glue Data Catalog Está en la misma forma Cuenta de AWS que la membresía.
- El bucket de Amazon S3 no está registrado en AWS Lake Formation.

Paso 2: (opcional) preparar sus datos para la computación criptográfica

(Opcional) Si utiliza la computación criptográfica y su tabla de datos contiene información confidencial que desea cifrar, debe cifrar la tabla de datos mediante el cliente de cifrado de C3R.

Para preparar los datos para la computación criptográfica, siga los procedimientos descritos en Preparación de tablas de datos cifrados con computación criptográfica para Clean Rooms.

Paso 3: cargar la tabla de datos en Amazon S3

> **Note**
>
> Si piensa utilizar tablas de datos cifrados en la colaboración, primero debe cifrar los datos para la computación criptográfica antes de cargar la tabla de datos en Amazon S3. Para obtener más información, consulte Preparación de tablas de datos cifrados con computación criptográfica para Clean Rooms.

Para cargar la tabla de datos en Amazon S3

1. Inicie sesión en la consola de Amazon S3 AWS Management Console y ábrala en https://console.aws.amazon.com/s3/.
2. Elija Buckets y, a continuación, elija un bucket donde desee almacenar la tabla de datos.
3. Elija Cargar y siga las indicaciones de la pantalla.
4. Seleccione la pestaña Objetos para ver el prefijo donde se almacenan sus datos. Anote el nombre de la carpeta.

   Puede seleccionar la carpeta para ver los datos.

Paso 4: Crear una AWS Glue tabla

Si ya tiene una tabla AWS Glue de datos, puede omitir este paso.

En este paso, configuras un rastreador AWS Glue que rastrea todos los archivos del bucket de S3 y crea una AWS Glue tabla. Para obtener más información, consulte Definición de rastreadores en AWS Glue en la Guía para desarrolladores de AWS Glue .

Para obtener más información sobre AWS Glue Data Catalog los tipos de datos compatibles, consulte. Tipos de datos compatibles

> **Note**
>
> AWS Clean Rooms actualmente no es compatible con los buckets S3 registrados en. AWS Lake Formation

El siguiente procedimiento describe cómo crear una AWS Glue tabla. Si desea utilizar un AWS Glue Data Catalog objeto cifrado con una clave AWS Key Management Service (AWS KMS), debe configurar la política de permisos de claves de KMS para permitir el acceso a esa tabla cifrada. Para obtener más información, consulte Configuración del cifrado en AWS Glue, en la Guía del desarrollador de AWS Glue .

Para crear una AWS Glue tabla

1. Siga el procedimiento de trabajo con rastreadores en la AWS Glue consola de la Guía del AWS Glue usuario.
2. Anote el nombre de la AWS Glue base de datos y el nombre de AWS Glue la tabla.

Paso 5: Siguientes pasos

Ahora que ha preparado las tablas de datos en Amazon S3, está preparado para:

- Creación de una tabla configurada
- Creación de un modelo de ML

Las tablas se pueden consultar después de:

- El creador de la colaboración ha configurado una colaboración en AWS Clean Rooms. Para obtener más información, consulte Crear una colaboración.
- El creador de la colaboración le ha enviado el ID de la colaboración como participante en la colaboración.

Preparación de tablas de datos en Amazon Athena

Puede consultar tablas de datos que se hayan creado como vistas AWS Glue Data Catalog (GDC) en Amazon Athena.

Una vista de GDC es una tabla virtual que se crea a partir de una o más tablas subyacentes. AWS Glue Debe crearse con Athena SQL en el catálogo de Athena. `AwsGlueCatalog`

La preparación de las tablas de datos en Amazon Athena implica los siguientes pasos:

Temas

- Paso 1: completar los requisitos previos
- Paso 2: (opcional) preparar sus datos para la computación criptográfica
- Paso 3: Próximos pasos

Paso 1: completar los requisitos previos

Para preparar las tablas de datos para usarlas con AWS Clean Rooms ellas, debe cumplir los siguientes requisitos previos:

- Las tablas de datos se guardan como uno de los formatos de datos compatibles para AWS Clean Rooms.
- Las tablas de datos utilizan los tipos de datos compatibles para AWS Clean Rooms.
- Ha creado una vista GDC en su AWS Glue tabla con Athena SQL en el catálogo de `AwsDataCatalog` Athena.

  La vista aparecerá en:

  - La consola Athena (debajo de`AwsDataCatalog`) como vista: https://console.aws.amazon.com/athena/
  - La AWS Glue consola como AWS Glue mesa: https://console.aws.amazon.com/glue/

  Para obtener más información, consulte Uso de vistas del catálogo de datos en Athena en la Guía del usuario de Amazon Athena.

> **Note**
>
> Necesita los permisos adecuados para crear vistas en Athena y. AWS Glue Además, asegúrese de tener acceso a las tablas subyacentes a las que se hace referencia en su View definition.
> AWS Clean Rooms solo admite el tipo de AWS Glue catálogo para Athena, no los tipos de catálogo Lambda o Hive.

- Sus tablas de datos o vistas de GDC están catalogadas y registradas en AWS Glue ellas. AWS Lake Formation
- Ha creado un depósito de salida independiente en Amazon S3 para recibir los resultados de Athena.
- Ha configurado un rol de servicio para leer los datos de Amazon Athena. Para obtener más información, consulte Cree un rol de servicio para leer los datos de Amazon Athena.
  - El rol de servicio tiene los permisos de acceso Lake Formation Select and Describe en la vista o tabla del GDC.

Paso 2: (opcional) preparar sus datos para la computación criptográfica

(Opcional) Si utiliza la computación criptográfica y su tabla de datos contiene información confidencial que desea cifrar, debe cifrar la tabla de datos mediante el cliente de cifrado de C3R.

Para preparar los datos para la computación criptográfica, siga los procedimientos descritos en Preparación de tablas de datos cifrados con computación criptográfica para Clean Rooms.

Paso 3: Próximos pasos

Ahora que ha preparado las tablas de datos en Amazon Athena, está preparado para:

- Creación de una tabla configurada
- Creación de un modelo de ML

Las tablas se pueden consultar después de:

- El creador de la colaboración ha configurado una colaboración en AWS Clean Rooms. Para obtener más información, consulte Crear una colaboración.
- El creador de la colaboración le ha enviado el ID de la colaboración como participante en la colaboración.

Preparación de tablas de datos en Snowflake

Puede consultar las tablas de datos que se han almacenado en el almacén de datos de Snowflake.

La preparación de las tablas de datos en Snowflake implica los siguientes pasos:

Temas

- Paso 1: completar los requisitos previos
- Paso 2: (opcional) preparar sus datos para la computación criptográfica
- Paso 3: Crea un secreto AWS Secrets Manager
- Paso 4: Siguientes pasos

Paso 1: completar los requisitos previos

Para preparar las tablas de datos para usarlas con ellas AWS Clean Rooms, debe cumplir los siguientes requisitos previos:

- Dispone de Cuenta de AWS los permisos adecuados para leer las tablas de datos. Para obtener más información, consulte Cree un rol de servicio para leer los datos de Snowflake.
- Las tablas de datos se guardan en uno de los formatos de datos admitidos para AWS Clean Rooms.
- Las tablas de datos utilizan los tipos de datos compatibles para AWS Clean Rooms.
- La tabla de datos se almacena en un almacén de Snowflake. Para obtener más información, consulte la documentación de Snowflake.
- Ha configurado un nuevo usuario de Snowflake con privilegios de solo lectura en la tabla de Snowflake que va a asociar a su colaboración.

Paso 2: (opcional) preparar sus datos para la computación criptográfica

(Opcional) Si utiliza la computación criptográfica y su tabla de datos contiene información confidencial que desea cifrar, debe cifrar la tabla de datos mediante el cliente de cifrado de C3R.

Para preparar los datos para la computación criptográfica, siga los procedimientos descritos en Preparación de tablas de datos cifrados con computación criptográfica para Clean Rooms.

Paso 3: Crea un secreto AWS Secrets Manager

Para conectarte a Snowflake desde AWS Clean Rooms, tendrás que crear y almacenar tus credenciales de Snowflake en un AWS Secrets Manager secreto y, a continuación, asociarlo a una tabla de Snowflake. AWS Clean Rooms

> **Note**
>
> Le recomendamos que cree un usuario nuevo que sea exclusivo para. AWS Clean Rooms Ese usuario solo debe tener un rol con permisos de lectura para los datos AWS Clean Rooms a los que desee acceder.

Para crear un AWS Secrets Manager secreto

1. En Snowflake, genere un usuario `snowflakeUser` y una contraseña. `snowflakePassword`
2. Determine con qué almacén de Snowflake interactuará este usuario,. `snowflakeWarehouse` Configúrelo como el formulario `DEFAULT_WAREHOUSE` `snowflakeUser` en Snowflake o recuérdelo para el siguiente paso.
3. En AWS Secrets Manager, cree un secreto con sus credenciales de Snowflake. Para crear un secreto en Secrets Manager, sigue el tutorial disponible en Crear un AWS Secrets Manager secreto en la Guía del AWS Secrets Manager usuario. Tras crear el secreto, conserva el nombre secreto `secretName` para el siguiente paso.

   - Al seleccionar pares clave/valor, crea un par `snowflakeUser` con la clave. `sfUser`
   - Al seleccionar pares clave/valor, cree un par para `snowflakePassword` con la clave. `sfPassword`
   - Al seleccionar pares clave/valor, cree un par para `snowflakeWarehouse` con la clave. `sfWarehouse`

     Esto no es necesario si hay un valor predeterminado en Snowflake. Esto no es necesario si hay un valor predeterminado en Snowflake.
   - Al seleccionar pares clave/valor, cree un par con la clave. `snowflakeRole` `sfrole`

Paso 4: Siguientes pasos

Ahora que ha preparado las tablas de datos en Snowflake, está preparado para:

- Creación de una tabla configurada
- Creación de un modelo de ML

Las tablas se pueden consultar después de:

- El creador de la colaboración ha configurado una colaboración en AWS Clean Rooms. Para obtener más información, consulte Crear una colaboración.
- El creador de la colaboración le ha enviado el ID de la colaboración como participante en la colaboración.

Preparación de tablas de datos cifrados con computación criptográfica para Clean Rooms

Computación criptográfica para Clean Rooms (C3R) es una capacidad en. AWS Clean Rooms Puede utilizar el C3R para limitar criptográficamente lo que puede aprender cualquiera de las partes y en una colaboración. AWS AWS Clean Rooms

Puede cifrar la tabla de datos mediante el cliente de cifrado C3R, una herramienta de cifrado del lado del cliente, antes de cargar la tabla de datos en su fuente de datos: Amazon Simple Storage Service (Amazon S3), Amazon Athena o Snowflake.

Para obtener más información, consulte Computación criptográfica para Clean Rooms.

Para preparar tablas de datos cifrados con C3R, se deben seguir estos pasos:

Pasos

- Paso 1: completar los requisitos previos
- Paso 2: descargar el cliente de cifrado de C3R
- Paso 3: (Opcional) Visualización de los comandos disponibles en el cliente de cifrado de C3R
- Paso 4: generar un esquema de cifrado para un archivo tabular
- Paso 5: crear una clave secreta compartida
- Paso 6: guardar la clave secreta compartida en la variable de entorno
- Paso 3: cifrar los datos
- Paso 8: verificar el cifrado de datos
- (Opcional) Crear un esquema (usuarios avanzados)

Paso 1: completar los requisitos previos

Para preparar las tablas de datos para su uso con C3R, debe cumplir con los siguientes requisitos previos:

- Puede acceder a la informática criptográfica para Clean Rooms repositorio en GitHub:

  https://github.com/aws/c3r

- Ha configurado AWS las credenciales para usar el cliente de cifrado C3R. El cliente de cifrado C3R utiliza estas credenciales para las llamadas a la API de solo lectura con el fin de recuperar los metadatos de la colaboración. AWS Clean Rooms Para obtener más información, consulte Configurar la AWS CLI en la Guía del usuario de la versión 2 de AWS Command Line Interface .
- Tienes Java Runtime Environment (JRE) 11 o una versión posterior instalada en su máquina.
  - El recomendado Java Runtime Environment, Amazon Corretto 11 o superior, se puede descargar de https://aws.amazon.com /corretto.
  - La Java Development Kit (JDK) incluye el correspondiente JRE de la misma versión. Sin embargo, las capacidades adicionales del JDK no son necesarias para ejecutar la computación criptográfica para Clean Rooms cliente de cifrado (C3R).
- Sus archivos de datos tabulares (.csv) o Parquet archivos (.parquet) se guardan localmente.
- Usted u otro miembro de la colaboración tienen la capacidad de crear una clave secreta compartida. Para obtener más información, consulte Paso 5: crear una clave secreta compartida.
- El creador de la colaboración ha creado una colaboración AWS Clean Rooms con la computación criptográfica habilitada para la colaboración. Para obtener más información, consulte Crear una colaboración.
- El creador de la colaboración le ha enviado el ID de la colaboración como participante en la colaboración. El nombre de recurso de Amazon (ARN) de la colaboración se incluye en la invitación enviada, que contiene el ID de la colaboración.

Paso 2: descargar el cliente de cifrado de C3R

Para descargar el cliente de cifrado C3R desde GitHub

1. Vaya a Computación criptográfica para Clean Rooms AWS GitHub repositorio: c3r https://github.com/aws/
2. Seleccione y descargue los archivos.

   El código fuente, las licencias y el material relacionado se pueden clonar o descargar como.zip archivo del GitHub página de inicio del repositorio. (vea el botón Código en la parte superior derecha de la lista de contenido del repositorio).

El último cliente de cifrado C3R firmado Java Executable File (es decir, la aplicación de interfaz de línea de comandos) se encuentra en la página de versiones del GitHub .

El paquete del cliente de cifrado de C3R para Apache Spark (`c3r-cli-spark`) es una versión del c3r-cli que debe enviarse como trabajo a un servidor Apache Spark en ejecución. Para obtener más información, consulte Ejecutar C3R en Apache Spark.

Paso 3: (Opcional) Visualización de los comandos disponibles en el cliente de cifrado de C3R

Utilice este procedimiento para familiarizarse con los comandos disponibles en el cliente de cifrado de C3R.

Para ver los comandos disponibles en el cliente de cifrado de C3R.

1. Desde una interfaz de línea de comandos (CLI), navegue hasta la carpeta que contiene el archivo descargado c3r-cli.jar file.
2. Ejecute el comando siguiente: `java -jar c3r-cli.jar`
3. Vea la lista de comandos y opciones disponibles.

Paso 4: generar un esquema de cifrado para un archivo tabular

Para cifrar datos, se requiere un esquema de cifrado que describa cómo se utilizarán los datos. En esta sección se describe cómo el cliente de cifrado C3R ayuda a generar un esquema de cifrado para un archivo CSV con una fila de encabezado o un Parquet file.

Solo tiene que hacerlo una vez por archivo. Una vez creado el esquema, se puede volver a utilizar para cifrar el mismo archivo (o cualquier archivo con nombres de columna idénticos). Si los nombres de las columnas o el esquema de cifrado deseado cambian, debe actualizar el archivo de esquema. Para obtener más información, consulte (Opcional) Crear un esquema (usuarios avanzados).

> **Important**
>
> Es imprescindible que todas las partes colaboradoras utilicen la misma clave secreta compartida. Las partes colaboradoras también deben coordinar los nombres de las columnas para que coincidan si van a ser JOINEd o comparados de otro

modo para garantizar la igualdad en las consultas. De lo contrario, las consultas SQL podrían producir resultados inesperados o incorrectos. No obstante, esto no será necesario si el creador de la colaboración habilitó la configuración de cifrado `allowJoinsOnColumnsWithDifferentNames` al crear la colaboración. Para obtener más información acerca de los ajustes relevantes para el cifrado, consulte Parámetros de computación criptográfica.

Cuando se ejecuta en modo de esquema, el cliente de cifrado de C3R revisa el archivo de entrada columna por columna y le pregunta si debe tratarse esa columna y de qué manera. Si el archivo contiene un gran número de columnas que no son necesarias para la salida cifrada, la generación del esquema interactivo podría resultar tediosa, ya que habría que saltarse todas las columnas no deseadas. Para evitarlo, puede escribir un esquema manualmente o crear una versión simplificada del archivo de entrada que incluya solo las columnas deseadas. Entonces, el generador de esquemas interactivo podrá ejecutarse en ese archivo reducido. El cliente de cifrado de C3R genera información sobre el archivo de esquema y le pregunta cómo se deben incluir o cifrar las columnas de origen (si es que se deben incluir) en el resultado de destino.

Para cada columna de origen del archivo de entrada, se le preguntará lo siguiente:

1. Cuántas columnas de destino se deben generar
2. Cómo se debe cifrar cada columna de destino (si procede)
3. El nombre de cada columna de destino
4. Cómo deben rellenarse los datos antes del cifrado si la columna se cifra como sealed columna

> **Note**
>
> Al cifrar los datos de una columna que se ha cifrado como sealed columna, debe determinar qué datos deben rellenarse. El cliente de cifrado de C3R sugiere un relleno predeterminado durante la generación del esquema, consistente en rellenar todas las entradas de una columna hasta que alcancen la misma longitud.
> Al determinar la longitud `fixed`, tenga en cuenta que el relleno está en bytes, no en bits.

A continuación se proporciona una tabla de decisión para crear el esquema.

Tabla de decisión del esquema

Decisión	¿Número de columnas de destino de la columna de origen <' '>? name-of-column	Tipo de columna de destino: [c] cleartext, [f] fingerprint, o [s] sealed ?	Nombre del encabezado de la columna de destino <default 'name-of-column'>	Añadir un sufijo <sufijo> al encabezado para indicar cómo se cifró, [y] sí o [n] no <predeterminado 'yes'>	<' name-of-column _sealed'> tipo de relleno: [n] uno, [f] fijo o [m] máximo <default 'max'>
Deje la columna sin cifrar.	1	c	No aplicable	No aplicable	No aplicable
Cifra la columna como fingerprint columna.	1	f	Elija el nombre predeterminado o introduzca un nombre de encabezado nuevo.	Introduzca y para elegir el valor predeterminado (`_fingerprint`) o introduzcan.	No aplicable
Cifra la columna como sealed columna.	1	s	Elija el nombre predeterminado o introduzca un nombre de encabezado nuevo.	Introduzca y para elegir el valor predeterminado (`_sealed`) o introduzcan.	Elija el tipo de relleno. Para obtener más información, consulte (Opcional) Crear un esquema (usuarios avanzados).

Decisión	¿Número de columnas de destino de la columna de origen <' '>? name-of-column	Tipo de columna de destino: [c] cleartext, [f] fingerprint, o [s] sealed ?	Nombre del encabezado de la columna de destino <default 'name-of-column'>	Añadir un sufijo <sufijo> al encabezado para indicar cómo se cifró, [y] sí o [n] no <predeterminado 'yes'>	<' name-of-column _sealed'> tipo de relleno: [n] uno, [f] fijo o [m] máximo <default 'max'>
Cifra la columna como ambas fingerprint y sealed.	2	Introduzca la primera columna de destino: f. Introduzca la segunda columna de destino: s.	Elija los encabezados de destino para cada columna de destino.	Introduzca y para elegir el valor predeterminado o introduzca n. .	Elija el tipo de relleno (para sealed solo columnas). Para obtener más información, consulte (Opcional) Crear un esquema (usuarios avanzados).

A continuación se muestran dos ejemplos de cómo crear esquemas de cifrado. El contenido exacto de la interacción depende del archivo de entrada y de las respuestas que proporcione.

Ejemplos

- Ejemplo: generar un esquema de cifrado para un fingerprint columna y un cleartext columna
- Ejemplo: genere un esquema de cifrado con sealed, fingerprint, y cleartext columns

Ejemplo: generar un esquema de cifrado para un fingerprint columna y un cleartext columna

En este ejemplo, para `ads.csv`, solo hay dos columnas: `username` y `ad_variant`. Para estas columnas, queremos lo siguiente:

- Que la columna `username` se cifre como columna `fingerprint`
- Que la `ad_variant` columna sea una columna `cleartext`

Para generar un esquema de cifrado para un fingerprint columna y un cleartext columna

1. (Opcional) Para garantizar que c3r-cli.jar el archivo y el archivo que se va a cifrar están presentes:

   a. Navegue hasta el directorio deseado y ejecute `ls` (si usa un Mac o Unix/Linux) o `dir` si usa Windows).

   b. Vea la lista de archivos de datos tabulares (por ejemplo, .csv) y elija el archivo que desea cifrar.

   En este ejemplo, `ads.csv` es el archivo que queremos cifrar.

2. Desde la CLI, ejecute el siguiente comando para crear un esquema de forma interactiva.

```
java -jar c3r-cli.jar schema ads.csv --interactive --output=ads.json
```

> **Note**
>
> - Puede ejecutar `java --jar PATH/TO/c3r-cli.jar`. O bien, si ha añadido `PATH/TO/c3r-cli.jar` a su variable de entorno CLASSPATH, también puede ejecutar el nombre de la clase. El cliente de cifrado de C3R buscará en CLASSPATH para encontrarlo (por ejemplo, `java com.amazon.psion.cli.Main`).
> - El indicador `--interactive` selecciona el modo interactivo para desarrollar el esquema. Esto guía al usuario a través de un asistente para crear el esquema. Los usuarios con conocimientos avanzados pueden crear su propio esquema JSON sin necesidad de utilizar el asistente. Para obtener más información, consulte (Opcional) Crear un esquema (usuarios avanzados).
> - El indicador `--output` establece un nombre de salida. Si no incluye el indicador `--output`, el cliente de cifrado de C3R intentará elegir un nombre de

salida predeterminado (por ejemplo, `<input>.out.csv` o, para el esquema, `<input>.json`).

3. En `Number of target columns from source column 'username'?`, escriba **1** y luego presione Intro.
4. En `Target column type: [c]leartext, [f]ingerprint, or [s]ealed?`, escriba **f** y luego presione Intro.
5. En `Target column headername <default 'username'>`, presione Intro.

   Se utiliza el nombre predeterminado '`username`'.
6. En `Add suffix '_fingerprint' to header to indicate how it was encrypted, [y]es or [n]o <default 'yes'>`, escriba **y** y luego presione Intro.

   > **Note**
   >
   > El modo interactivo sugiere añadir sufijos a los encabezados de las columnas cifradas (para `_fingerprint` fingerprint columnas y para `_sealed` sealed columnas). Los sufijos pueden resultar útiles al realizar tareas como cargar datos Servicios de AWS o crear AWS Clean Rooms colaboraciones. Estos sufijos pueden ayudar a indicar qué se puede hacer con los datos cifrados de cada columna. Por ejemplo, las cosas no funcionarán si encriptas una columna como sealed columna (`_sealed`) e intenta JOIN en él o prueba a la inversa.
7. En `Number of target columns from source column 'ad_variant'?`, escriba **1** y luego presione Intro.
8. En `Target column type: [c]leartext, [f]ingerprint, or [s]ealed?`, escriba **c** y luego presione Intro.
9. En `Target column headername <default 'username'>`, presione Intro.

   Se utiliza el nombre predeterminado '`ad_variant`'.

   El esquema se escribe en un nuevo archivo llamado `ads.json`.

   > **Note**
   >
   > Puede ver el esquema abriéndolo en cualquier editor de texto, como Notepad on Windows o TextEdit on macOS.

10. Ahora ya puede cifrar datos.

Ejemplo: genere un esquema de cifrado con sealed, fingerprint, y cleartext columns

En este ejemplo, para `sales.csv`, hay tres columnas: `username`, `purchased` y `product`. Para estas columnas, queremos lo siguiente:

- Que la `product` columna sea una columna `sealed`
- Que la columna `username` se cifre como columna `fingerprint`
- Que la `purchased` columna sea una columna `cleartext`

Para generar un esquema de cifrado con sealed, fingerprint, y cleartext columns

1. (Opcional) Para garantizar la c3r-cli.jar el archivo y el archivo que se va a cifrar están presentes:

   a. Navegue hasta el directorio deseado y ejecute `ls` (si usa un Mac o Unix/Linux) o `dir` si usa Windows).

   b. Vea la lista de archivos de datos tabulares (.csv) y elija el archivo que desea cifrar.

      En este ejemplo, `sales.csv` es el archivo que queremos cifrar.

2. Desde la CLI, ejecute el siguiente comando para crear un esquema de forma interactiva.

```
java -jar c3r-cli.jar schema sales.csv --interactive --
output=sales.json
```

> **Note**
>
> - El indicador `--interactive` selecciona el modo interactivo para desarrollar el esquema. Esto guía al usuario a través de un flujo de trabajo guiado para crear el esquema.
> - Si usted es un usuario avanzado, puede crear su propio esquema JSON sin utilizar el flujo de trabajo guiado. Para obtener más información, consulte (Opcional) Crear un esquema (usuarios avanzados).
> - Para los archivos .csv sin encabezados de columna, consulte el indicador `--noHeaders` del comando de esquema disponible en la CLI.

- El indicador `--output` establece un nombre de salida. Si no incluye el indicador `--output`, el cliente de cifrado de C3R intentará elegir un nombre de salida predeterminado (por ejemplo, `<input>.out` o, para el esquema, `<input>.json`).

3. En `Number of target columns from source column ‘username’?`, escriba **1** y luego presione Intro.
4. En `Target column type: [c]leartext, [f]ingerprint, or [s]ealed?`, escriba **f** y luego presione Intro.
5. En `Target column headername <default 'username'>`, presione Intro.

   Se utiliza el nombre predeterminado ‘`username`’.
6. En `Add suffix '_fingerprint' to header to indicate how it was encrypted, [y]es or [n]o <default 'yes'>`, escriba **y** y luego presione Intro.
7. En `Number of target columns from source column ‘purchased’?`, escriba **1** y luego presione Intro.
8. En `Target column type: [c]leartext, [f]ingerprint, or [s]ealed?`, escriba **c** y luego presione Intro.
9. En `Target column headername <default 'purchased'>`, presione Intro.

   Se utiliza el nombre predeterminado ‘`purchased`’.
10. En `Number of target columns from source column ‘product’?`, escriba **1** y luego presione Intro.
11. En `Target column type: [c]leartext, [f]ingerprint, or [s]ealed?`, escriba **s** y luego presione Intro.
12. En `Target column headername <default 'product'>`, presione Intro.

    Se utiliza el nombre predeterminado ‘`product`’.
13. En `‘product_sealed’ padding type: [n]one, [f]ixed, or [m]ax <default ’max’?>`, presione Intro para elegir el valor predeterminado.
14. Para `Byte-length beyond max length to pad cleartext to in ‘product_sealed’ <default ‘0’>?`, presione Intro para elegir el valor predeterminado.

    El esquema se escribe en un nuevo archivo llamado `sales.json`.
15. Ahora ya puede cifrar datos.

Paso 5: crear una clave secreta compartida

Para cifrar las tablas de datos, los participantes de la colaboración deben acordar una clave secreta compartida y compartirla de forma segura.

La clave secreta compartida debe tener al menos 256 bits (32 bytes). Puede especificar una clave de mayor tamaño, pero no le proporcionará ninguna garantía adicional.

> **Important**
>
> Recuerde que la clave y el ID de la colaboración utilizados para el cifrado y el descifrado deben ser idénticos para todos los participantes de la colaboración.

En las siguientes secciones se proporcionan ejemplos de comandos de consola para generar una clave secreta compartida guardada como `secret.key` en el directorio de trabajo actual del terminal correspondiente.

Temas

- Ejemplo: generación de claves mediante OpenSSL
- Ejemplo: generación de claves en Windows utilización PowerShell

Ejemplo: generación de claves mediante OpenSSL

Para una biblioteca de criptografía común de uso general, ejecute el comando siguiente para crear una clave secreta compartida.

```
openssl rand 32 > secret.key
```

Si estás usando Windows y no tienes OpenSSL instalado, puede generar claves utilizando el ejemplo descrito en Ejemplo: generación de claves en Windows utilización PowerShell.

Ejemplo: generación de claves en Windows utilización PowerShell

En PowerShell, una aplicación de terminal disponible en Windows, ejecute el siguiente comando para crear una clave secreta compartida.

```
$bs = New-Object Byte[](32);
[Security.Cryptography.RandomNumberGenerator]::Create().GetBytes($bs); Set-
Content 'secret.key' -Encoding Byte -Value $bs
```

Paso 6: guardar la clave secreta compartida en la variable de entorno

Una variable de entorno es una forma cómoda y ampliable para que los usuarios proporcionen una clave secreta de varios almacenes de claves, por ejemplo, AWS Secrets Manager y la pasen al cliente de cifrado C3R.

El cliente de cifrado C3R puede utilizar claves almacenadas en Servicios de AWS si se utiliza AWS CLI para almacenar esas claves en la variable de entorno correspondiente. Por ejemplo, el cliente de cifrado C3R puede usar una clave de. AWS Secrets Manager Para obtener más información, consulte Crear y administrar secretos con AWS Secrets Manager en la Guía del usuario de AWS Secrets Manager .

> **Note**
>
> Sin embargo, antes de utilizar una Servicio de AWS como esta AWS Secrets Manager para guardar las claves C3R, compruebe que su caso de uso lo permite. Algunos casos de uso pueden requerir que se oculte la clave. AWS Esto es para garantizar que los datos cifrados y la clave nunca estén en manos del mismo tercero.

Los únicos requisitos para una clave secreta compartida son que la clave secreta compartida sea base64-codificada y almacenada en la variable `C3R_SHARED_SECRET` de entorno.

En las siguientes secciones se describen los comandos de la consola para convertir un `secret.key` archivo en base64 y almacenarlo como una variable de entorno. El archivo `secret.key` podría haberse generado a partir de cualquiera de los comandos enumerados en Paso 5: crear una clave secreta compartida, y este es solo un ejemplo de origen.

Guarde la clave en una variable de entorno en Windows utilización PowerShell

Para convertir a base64 y defina la variable de entorno en Windows utilización PowerShell, ejecute el siguiente comando.

```
$Bytes=[IO.File]::ReadAllBytes((Get-Location).ToString()+'\secret.key');
 $env:C3R_SHARED_SECRET=[Convert]::ToBase64String($Bytes)
```

Guarde la clave en una variable de entorno en Linux o macOS

Para convertir a base64 y defina la variable de entorno en Linux o macOS, ejecute el siguiente comando.

```
export C3R_SHARED_SECRET="$(cat secret.key | base64)"
```

Paso 3: cifrar los datos

Para realizar este paso, debe adquirir el ID de AWS Clean Rooms colaboración y la clave secreta compartida. Para obtener más información, consulte Requisitos previos.

En el siguiente ejemplo, ejecutamos el cifrado en `ads.csv` utilizando el esquema que hemos creado, denominado `ads.json`.

Para cifrar los datos

1. Almacene la clave secreta compartida de la colaboración en Paso 6: guardar la clave secreta compartida en la variable de entorno.
2. Desde la línea de comandos, ingrese el comando siguiente:

   ```
   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. Para *<name of input .csv file>* ello, introduzca el nombre del archivo.csv de entrada.
4. En `schema=`, introduzca el nombre del archivo de esquema de cifrado .json.
5. En `id=`, introduzca el ID de la colaboración.
6. En `output=`, introduzca el nombre del archivo de salida (por ejemplo, `ads-output.csv`).
7. Incluya cualquiera de los indicadores de línea de comandos descritos en Parámetros de computación criptográfica y Indicadores opcionales en la computación criptográfica para Clean Rooms.
8. Ejecute el comando.

En el ejemplo de `ads.csv`, ejecutamos el siguiente comando.

java -jar c3r-cli.jar encrypt ***ads.csv*** --schema=***ads.json*** --id=***123e4567-e89b-42d3-a456-556642440000*** --output=***ads-output.csv***

En el ejemplo de sales.csv, ejecutamos el siguiente comando.

java -jar c3r-cli.jar encrypt ***sales.csv*** --schema=***sales.json*** --id=***123e4567-e89b-42d3-a456-556642440000***

> **Note**
>
> En este ejemplo, no especificamos un nombre de archivo de salida (`--output=sales-output.csv`). Como resultado, se generó el nombre de archivo de salida predeterminado `name-of-file.out.csv`.

Ahora ya puede verificar los datos cifrados.

Paso 8: verificar el cifrado de datos

Para verificar que los datos estaban cifrados

1. Vea el archivo de datos cifrados (por ejemplo, `sales-output.csv`).
2. Compruebe las siguientes columnas:
   a. Columna 1: cifrada (por ejemplo, `username_fingerprint`).

      Para el registro fingerprint En las columnas (HMAC), después de la versión y el prefijo del tipo (por ejemplo,`01:hmac:`), hay 44 caracteres de datos codificados en base64.
   b. Columna 2: sin cifrar (por ejemplo, `purchased`).
   c. Columna 3: cifrada (por ejemplo, `product_sealed`).

      Para cifrados (SELECT) columnas, la longitud de cleartext más cualquier relleno después de la versión y el prefijo de tipo (por ejemplo,`01:enc:`) es directamente proporcional a la longitud del cleartext que estaba encriptado. Es decir, la longitud es el tamaño de la entrada, más aproximadamente un 33 por ciento de sobrecarga debido a la codificación.

Ya puede hacer lo siguiente:

1. Cargar los datos cifrados a S3.
2. Crea una AWS Glue tabla.
3. Crear una tabla configurada en AWS Clean Rooms.

El cliente de cifrado de C3R creará archivos temporales que no contengan datos sin cifrar (a menos que esos datos tampoco estén cifrados en el resultado final). No obstante, es posible que algunos valores cifrados no se rellenen correctamente. Las columnas de huella digital pueden contener

valores duplicados, incluso si el ajuste de colaboración `allowRepeatedFingerprintValue` es `false`. Este problema se debe a que el archivo temporal se escribe antes de comprobar las longitudes de relleno adecuadas y las propiedades de eliminación de duplicados.

Si el cliente de cifrado de C3R falla o se interrumpe durante el cifrado, es posible que se detenga después de escribir el archivo temporal, pero antes de comprobar estas propiedades y eliminar los archivos temporales. Por lo tanto, es posible que estos archivos temporales sigan en el disco. Si este es el caso, el contenido de estos archivos no protege los datos de texto sin formato en los mismos niveles que los de salida. En concreto, estos archivos temporales pueden revelar datos del texto sin formato para los análisis estadísticos, lo que no serviría de nada para el resultado final. El usuario debe eliminar estos archivos (en particular, un SQLite base de datos) para evitar que estos archivos caigan en manos no autorizadas.

(Opcional) Crear un esquema (usuarios avanzados)

La creación manual de esquemas está reserva para los usuarios avanzados.

A continuación se describe el formato del archivo de esquema JSON para los archivos de entrada con o sin encabezados de columna. Los usuarios avanzados pueden escribir o modificar el esquema directamente si lo desean.

> **Note**
>
> El cliente de cifrado de C3R puede ayudarlo a crear un esquema mediante el proceso interactivo descrito en Ejemplo: genere un esquema de cifrado con sealed, fingerprint, y cleartext columns o mediante la creación de una plantilla stub.

Esquemas de tablas mapeados y posicionales

En la siguiente sección se describen dos tipos de esquemas de tabla:

- Esquema de tabla mapeada: este esquema se utiliza para cifrar archivos.csv con una fila de encabezado y Apache Parquet archivos.
- Esquema de tabla posicional: este esquema se usa para cifrar archivos .csv sin fila de encabezado.

El cliente de cifrado de C3R puede cifrar un archivo tabular para una colaboración. Para ello, debe tener un archivo de esquema correspondiente que especifique cómo se debe derivar la salida cifrada a partir de la entrada.

El cliente de cifrado de C3R puede ayudar a generar un esquema para un archivo `INPUT` ejecutando el comando de esquema del cliente de cifrado de C3R en la línea de comandos. Un ejemplo de comando es `java -jar c3r-cli.jar schema --interactive INPUT`.

El esquema especifica la siguiente información:

1. Qué columnas de origen se asignan a qué columnas transformadas en el archivo de salida mediante sus nombres de encabezado (esquemas mapeados) o su posición (esquemas posicionales)
2. ¿Qué columnas de destino van a permanecer cleartext
3. ¿Para qué columnas de destino se van a cifrar SELECT consultas
4. ¿Para qué columnas de destino se van a cifrar JOIN consultas

Esta información está codificada en un archivo de esquema JSON específico de tabla, que consta de un único objeto cuyo campo `headerRow` es un valor booleano. El valor debe ser `true` para Parquet archivos y archivos.csv con una fila de encabezado, `false` etc.

Esquema de la tabla mapeado

El esquema mapeado tiene la siguiente forma.

```
{
  "headerRow": true,
  "columns": [
    {
      "sourceHeader": STRING,
      "targetHeader": STRING,
      "type": TYPE,
      "pad": PAD
    },
    ...
  ]
}
```

Si `headerRow` es `true`, es el siguiente campo del objeto es `columns`, que contiene una matriz de esquemas de columnas que asignan los encabezados de origen a los encabezados de destino (es decir, objetos JSON que describen lo que deben contener las columnas de salida).

- `sourceHeader`: el nombre del encabezado `STRING` de la columna de origen a partir de cual se derivan los datos.

  > **Note**
  >
  > Una misma columna de origen se puede utilizar para varias columnas de destino. Una columna del archivo de entrada que no aparezca enumerada como `sourceHeader` en ninguna parte del esquema no aparecerá en el archivo de salida.

- `targetHeader`: el nombre del encabezado `STRING` de la columna correspondiente del archivo de salida.

  > **Note**
  >
  > Este campo es opcional para los esquemas mapeados. Si se omite este campo, `sourceHeader` se reutiliza para el nombre del encabezado en el archivo de salida. `_fingerprint`O bien `_sealed` se adjunta si la columna de salida es una fingerprint columna o sealed columna, respectivamente.

- `type`: el TYPE de la columna de destino del archivo de salida. Es decir, una de las columnas `cleartext`, `sealed` o `fingerprint` dependiendo de cómo se vaya a utilizar la columna en la colaboración.
- `pad`: un campo de un objeto de esquema de columnas que solo está presente cuando TYPE es `sealed`. Su valor correspondiente de PAD es un objeto que describe cómo deben rellenarse los datos antes del cifrado.

  ```
  {
    "type": PAD_TYPE,
    "length": INT
  }
  ```

  Para especificar el relleno previo al cifrado, se utiliza `type` y `length` de la siguiente manera:

  - PAD_TYPE como none: no se aplicará ningún relleno a los datos de la columna y el campo `length` no es aplicable (es decir, se omite).

- PAD_TYPE como `fixed`: los datos de la columna se rellenan hasta el número de bytes especificado, `length`.
- PAD_TYPE como `max`: los datos de la columna se rellenan hasta el tamaño de la longitud en bytes del valor más largo, más `length` bytes adicionales.

El siguiente es un ejemplo de esquema mapeado, con una columna de cada tipo.

```
{
  "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
      }
    }
  ]
}
```

Un ejemplo más complejo sería el siguiente ejemplo de un archivo .csv con encabezados.

```
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,COO,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
```

En el siguiente ejemplo de esquema mapeado, las columnas `FirstName` y `LastName` son columnas `cleartext`. La columna `State` está cifrada como columna `fingerprint` y como columna `sealed` con un relleno de `none`. Las columnas restantes se omiten.

```
{
  "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"
      }
    }
  ]
```

```
}
```

El siguiente es el archivo .csv que resulta del esquema mapeado.

```
givenname,surname,state_fingerprint,state
John,Doe,01:hmac:UK8s8Cn/WR2JO/To2dTxWD73aDEe2ZUXeSHy3Tv
+1Mk=,01:enc:FQ3n3Ahv9BQQNWQGcugeHzHYzEZE1vapHa2Uu4SRgSAtZ3qObjPA4TcsHt
+BOkMKBcnHWI13BeGG/SBqmj7vKpI=
Paulo,Santos,01:hmac:CHF4eIrtTNgAooU9v4h9Qjc
+txBnMidQTjdjWuaDTTA=,01:enc:KZ5n5GtaXACco65AXk48BQO2durDNR2ULc4YxmMC8NaZZKKJiksU1IwFadAvV4iBQ1
Mateo,Jackson,01:hmac:iIRnjfNBzryusIJ1w35lgNzeY1RQ1bSfq6PDHW8Xrbk=,01:enc:mLKpS5HIOSgphdEsrzhEd
eN9nBO2gAbIygt4OFn4LalYn9Xyj/XUWXlmn8zFe2T4kyDTD8kGOvpQEUGxAUFk=
Diego,Ramirez,01:hmac:UK8s8Cn/WR2JO/To2dTxWD73aDEe2ZUXeSHy3Tv+1Mk=,01:enc:rmZhT98Zm
+IIGw1UTjMIJP4IrW/AAltBLMXcHvnYfRgmWP623VFQ6aUnhsb2MDqEw4G5Uwg5rKKZepUxx5uKbfk=
Jorge,Souza,01:hmac:3BxJdXiFFyZ8HBbYNqqEhBVqhNOd7s2ZiKUe7QiTyo8=,01:enc:vVaqWC1VRbhvkf8gnuR7q0z
Terry,Whitlock01:hmac:UK8s8Cn/WR2JO/To2dTxWD73aDEe2ZUXeSHy3Tv+1Mk=,01:enc:3c9VEWbODO/
xbQjdGuccLvI7oZTBdPU+SyrJIyr2kudfAxbuMQ2uRdU/q7rbgyJjxZS8M2U35ILJf/lDgTyg7cM=
Jane,Doe,01:hmac:UK8s8Cn/WR2JO/To2dTxWD73aDEe2ZUXeSHy3Tv+1Mk=,01:enc:9RWv46YLveykeNZ/
G0NdlYFg+AVdOnu05hHyAYTQkPLHnyX+0/jbzD/g9ZT8GCgVE9aB5bV4ooJIXHGBVMXcjrQ=
```

Esquema de tabla posicional

El esquema posicional tiene la siguiente forma.

```
{
  "headerRow": false,
  "columns": [
    [
      {
        "targetHeader": STRING,
        "type": TYPE,
        "pad": PAD
      },
      {
        "targetHeader": STRING,
        "type": TYPE,
        "pad": PAD
      }
    ],
    [],
    ...
  ]
}
```

Si `headerRow` es `false`, es el siguiente campo del objeto es `columns`, que contiene una matriz de entradas. Cada entrada es en sí misma una matriz de cero o más esquemas de columnas posicionales (sin campo `sourceHeader`), que son objetos JSON que describen lo que debe contener la salida.

- `sourceHeader`: el nombre del encabezado `STRING` de la columna de origen a partir de cual se derivan los datos.

  > **Note**
  >
  > Este campo debe omitirse en los esquemas posicionales. En los esquemas posicionales, la columna de origen se deduce mediante el índice correspondiente de la columna en el archivo de esquema.

- `targetHeader`: el nombre del encabezado `STRING` de la columna correspondiente del archivo de salida.

  > **Note**
  >
  > Este campo es obligatorio para los esquemas posicionales.

- `type`: el TYPE de la columna de destino del archivo de salida. Es decir, una de las columnas `cleartext`, `sealed` o `fingerprint` dependiendo de cómo se vaya a utilizar la columna en la colaboración.
- `pad`: un campo de un objeto de esquema de columnas que solo está presente cuando TYPE es `sealed`. Su valor correspondiente de PAD es un objeto que describe cómo deben rellenarse los datos antes del cifrado.

  ```
  {
    "type": PAD_TYPE,
    "length": INT
  }
  ```

  Para especificar el relleno previo al cifrado, se utiliza `type` y `length` de la siguiente manera:

  - PAD_TYPE como `none`: no se aplicará ningún relleno a los datos de la columna y el campo `length` no es aplicable (es decir, se omite).
  - PAD_TYPE como `fixed`: los datos de la columna se rellenan hasta el número de bytes especificado, `length`.

- PAD_TYPE como `max`: los datos de la columna se rellenan hasta el tamaño de la longitud en bytes del valor más largo, más `length` bytes adicionales.

> **Note**
>
> `fixed` es útil si sabe con antelación cuál es el límite superior del tamaño en bytes de los datos de la columna. Se produce un error si algún dato de esa columna supera la `length` especificada.
> `max` resulta práctico cuando se desconoce el tamaño exacto de los datos de entrada, ya que funciona independientemente del tamaño de los datos. Sin embargo, `max` requiere un tiempo de procesamiento adicional porque cifra los datos dos veces. `max` cifra los datos una vez cuando se leen en el archivo temporal y una vez que se conoce la entrada de datos más larga de la columna.
> Además, la longitud del valor más largo no se guarda entre las invocaciones del cliente. Si tiene previsto cifrar los datos por lotes o cifrar datos nuevos periódicamente, tenga en cuenta que la longitud del texto cifrado resultante puede variar de un lote a otro.

A continuación se muestra un ejemplo de esquema posicional.

```
{
  "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
        }
      }
    ]
  ]
}
```

Como ejemplo complejo, el siguiente es un ejemplo de un archivo .csv sin la primera fila con los encabezados.

```
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,COO, 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
```

El esquema posicional tiene la siguiente forma.

```
{
  "headerRow": false,
  "columns": [
    [
      {
        "targetHeader": "GivenName",
        "type": "cleartext"
      }
    ],
    [
      {
        "targetHeader": "Surname",
        "type": "cleartext"
```

```
      }
    ],
    [],
    [],
    [
      {
        "targetHeader": "State_Join",
        "type": "fingerprint"
      },
      {
        "targetHeader": "State",
        "type": "sealed",
        "pad": {
          "type": "none"
        }
      }
    ],
    [],
    [],
    [],
    []
  ]
}
```

El esquema anterior produce el siguiente archivo de salida con una fila de encabezados que contiene los encabezados de destino especificados.

```
givenname,surname,state_fingerprint,state
Mateo,Jackson,01:hmac:iIRnjfNBzryusIJ1w35lgNzeY1RQ1bSfq6PDHW8Xrbk=,01:enc:ENS6QD3cMVl9vQEGfe9MN
Q8m/Y5SA89dJwKpT5rGPp8e36h6klwDoslpFzGvU0=
Jorge,Souza,01:hmac:3BxJdXiFFyZ8HBbYNqqEhBVqhNOd7s2ZiKUe7QiTyo8=,01:enc:LKo0zirq2+
+XEIIIMNRjAsGMdyWUDwYaum0B+IFP+rUf1BNeZDJjtFe1Z+zbZfXQWwJy52Rt7HqvAb2WIK1oMmk=
Paulo,Santos,01:hmac:CHF4eIrtTNgAooU9v4h9Qjc
+txBnMidQTjdjWuaDTTA=,01:enc:MyQKyWxJ9kvK1xDQQtXlUNwv3F+yrBRr0xrUY/1BGg5KFgOn9pK+MZ7g
+ZNqZEPcPz4lht1u0t/wbTaqzOCLXFQ=
Jane,Doe,01:hmac:UK8s8Cn/WR2JO/To2dTxWD73aDEe2ZUXeSHy3Tv+1Mk=,01:enc:Pd8sbITBfb0/
ttUB4svVsgoYkDfnDvgkvxzeci0Yxq54rLSwccy1o3/B50C3cpkkn56dovCwzgmmPNwrmCmYtb4=
Terry,Whitlock01:hmac:UK8s8Cn/WR2JO/To2dTxWD73aDEe2ZUXeSHy3Tv
+1Mk=,01:enc:Qmtzu3B3GAXKh2KkRYTiEAaMopYedsSdF2e/
ADUiBQ9kv2CxKPzWyYTD3ztmKPMka19dHre5VhUHNpO3O+j1AQ8=
Diego,Ramirez,01:hmac:UK8s8Cn/WR2JO/To2dTxWD73aDEe2ZUXeSHy3Tv+1Mk=,01:enc:ysdg
+GHKdeZrS/geBIooOEPLHG68MsWpx1dh3xjb+fG5rmFmqUcJLNuuYBHhHAlxchM2WVeV1fmHkBX3mvZNvkc=
```

```
John,Doe,01:hmac:UK8s8Cn/WR2JO/To2dTxWD73aDEe2ZUXeSHy3Tv+1Mk=,01:enc:9uX0wZuO7kAPAx
+Hf6uvQownkWqFSKtWS7gQIJSe5aXFquKWCK6yZN0X5Ea2N3bn03Uj1kh0agDWoiP9FRZGJA4=
```

Descifrar tablas de datos con el cliente de cifrado de C3R

Siga este procedimiento para las colaboraciones que utilizan la computación criptográfica para Clean Rooms y el cliente de cifrado C3R para cifrar tablas de datos. Utilice este procedimiento después de haber consultado datos de la colaboración.

La clave secreta compartida y el ID de colaboración son necesarios para este procedimiento.

El miembro que puede recibir los resultados descifra los datos con la misma clave secreta compartida y el mismo ID de colaboración que se utilizaron para cifrar los datos de la colaboración.

> **Note**
>
> AWS Clean Rooms las colaboraciones ya limitan quién puede realizar y ver los resultados de las consultas. Para realizar el descifrado, quien tenga acceso a estos resultados necesita la misma clave secreta compartida y el mismo ID de colaboración que se utilizaron para cifrar los datos.

Para descifrar una tabla de datos cifrada

1. (Opcional) Consulte los comandos disponibles en el cliente de cifrado de C3R.
2. (Opcional) Navegue hasta el directorio deseado y ejecute `ls` (macOS) o `dir` (Windows).
   - Compruebe que el c3r-cli.jar el archivo y el archivo de datos de resultados de la consulta cifrados se encuentran en el directorio deseado.

     > **Note**
     >
     > Si los resultados de la consulta se descargan desde la interfaz de la AWS Clean Rooms consola, es probable que estén en la carpeta Descargas de su cuenta de usuario. (Por ejemplo, la carpeta Descargas del directorio de usuarios de Windows y macOS.) Se recomienda mover el archivo de resultados de la consulta a la misma carpeta que c3r-cli.jar.

3. Guarde la clave secreta compartida en la variable de entorno `C3R_SHARED_SECRET`. Para obtener más información, consulte Paso 6: guardar la clave secreta compartida en la variable de entorno.
4. Desde AWS Command Line Interface (AWS CLI), ejecute el siguiente comando.

   java -jar c3r-cli.jar decrypt ***<name of input .csv file>*** --id=***<collaboration id>*** --output=***<output file name>***
5. Sustituya cada uno *`user input placeholder`* por su propia información:

   a. En `id=`, introduzca el ID de la colaboración.

   b. En `output=`, introduzca el nombre del archivo de salida (por ejemplo, `results-decrypted.csv`).

      Si no especifica un nombre de salida, se mostrará un nombre predeterminado en el terminal.

   c. Vea los datos descifrados en el archivo de salida especificado utilizando el CSV que prefiera o Parquet aplicación de visualización (como Microsoft Excel, un editor de texto u otra aplicación).

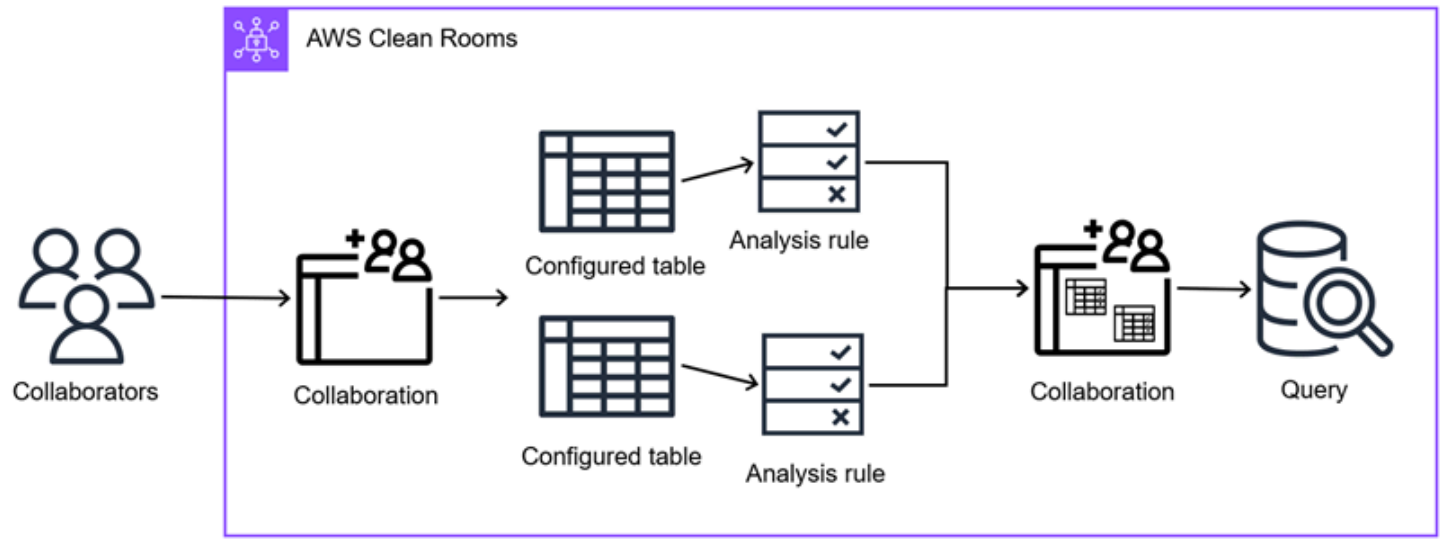
Tablas configuradas en AWS Clean Rooms

Una tabla configurada es una referencia a una tabla existente en una fuente de datos. Contiene una regla de análisis que determina cómo se pueden consultar los datos en AWS Clean Rooms. Las tablas configuradas se pueden asociar a una o más colaboraciones.

Con ella AWS Clean Rooms, puede realizar un análisis de agregación de los datos de los eventos, como el número de compras en comparación con el número de compras. También puede realizar un análisis de listas de datos de eventos, por ejemplo, enriqueciendo los datos de clientes superpuestos pasando de datos de segmentos a datos de CRM. También puede realizar consultas personalizadas y establecer una privacidad diferencial en los datos de los eventos, como los datos de audiencia y los atributos de los segmentos.

En primer lugar, puedes crear una colaboración AWS Clean Rooms y añadir la Cuentas de AWS que quieras invitar, o unirte a una colaboración a la que estés invitado creando una membresía. A continuación, usted y el otro miembro de la colaboración crean tablas configuradas. Agregue una regla de análisis a las tablas configuradas (agregación, lista o personalizada) y asocie las tablas configuradas a la colaboración. Por último, el miembro que puede realizar la consulta ejecuta una consulta en las dos tablas de datos.

En el siguiente diagrama se resume cómo trabajar con los datos de los eventos en AWS Clean Rooms.



Temas

- Crear una tabla configurada en AWS Clean Rooms

- Agregación de una regla de análisis a una tabla configurada
- Asociar una tabla configurada a una colaboración
- Agregación de una regla de análisis de colaboración a una tabla configurada
- Configuración de la política de privacidad diferencial (opcional)
- Ver tablas y reglas de análisis
- Editar detalles de tablas configuradas
- Editar etiquetas de tablas configuradas
- Editar la regla de análisis de tablas configuradas
- Eliminar la regla de análisis de tablas configuradas
- Columnas no permitidas de la tabla configurada
- Editar asociaciones de tablas configuradas
- Disociar tablas configuradas

Crear una tabla configurada en AWS Clean Rooms

Una tabla configurada es una referencia a una tabla existente en una fuente de datos. Contiene una regla de análisis que determina cómo se pueden consultar los datos en AWS Clean Rooms. Las tablas configuradas se pueden asociar a una o más colaboraciones.

Para obtener información sobre cómo crear una tabla configurada mediante la AWS SDKs, consulta la Referencia de la API.AWS Clean Rooms

Temas

- Creación de una tabla configurada: fuente de datos de Amazon S3
- Creación de una tabla configurada: fuente de datos de Amazon Athena
- Creación de una tabla configurada: fuente de datos de Snowflake

Creación de una tabla configurada: fuente de datos de Amazon S3

En este procedimiento, el miembro realiza las siguientes tareas:

- Configura una AWS Glue tabla existente para usarla en. AWS Clean Rooms(este paso se puede realizar antes o después de unirse a una colaboración, a menos que se utilice la computación criptográfica para Clean Rooms).

> **Note**
>
> AWS Clean Rooms admite AWS Glue tablas. Para obtener más información sobre cómo introducir los datos AWS Glue, consulte Paso 3: cargar la tabla de datos en Amazon S3.

- Asigna un nombre a la tabla configurada y elige qué columnas usar en la colaboración.

El siguiente procedimiento asume que:

- El miembro de la colaboración ya ha subido sus tablas de datos a Amazon S3 y ha creado una AWS Glue tabla.

> **Note**
>
> Si utiliza el motor de análisis Spark, el destino de los resultados en Amazon S3 no puede estar dentro del mismo bucket de S3 que cualquier fuente de datos.

- (Opcional) Solo para las tablas de datos cifrados, el miembro de la colaboración ya ha preparado las tablas de datos cifrados con el cliente de cifrado de C3R.

Puede utilizar la generación de estadísticas proporcionada por AWS Glue para calcular las estadísticas a nivel de columna para las tablas. AWS Glue Data Catalog Tras AWS Glue generar las estadísticas para las tablas del catálogo de datos, Amazon Redshift Spectrum las utiliza automáticamente para optimizar el plan de consultas. Para obtener más información sobre cómo calcular las estadísticas a nivel de columna mediante el uso de estadísticas de columnas AWS Glue, consulte Optimización del rendimiento de las consultas mediante estadísticas de columnas en la Guía del AWS Glue usuario. Para obtener más información AWS Glue, consulte la Guía para desarrolladores de AWS Glue.

Para crear una tabla configurada: fuente de datos de Amazon S3

1. Inicie sesión en la consola AWS Management Console y abra la AWS Clean Rooms consola con su Cuenta de AWS (si aún no lo ha hecho).
2. En el panel de navegación izquierdo, elija Tables (Tablas).
3. En la esquina superior derecha, elija Configurar nueva tabla.
4. En Fuente de datos, en Fuentes de AWS datos, elija Amazon S3.

5. En la tabla de Amazon S3:

   a. Elija la base de datos en la lista desplegable.

   b. Seleccione la Tabla que desea configurar en la lista desplegable.

> **Note**
>
> Para comprobar que se trata de la tabla correcta, realice una de las siguientes acciones:
>
> - Seleccione Ver en AWS Glue.
> - Activa Ver esquema desde AWS Glue para ver el esquema.

6. Para ver las columnas y los métodos de análisis permitidos en las colaboraciones,

   a. ¿Para qué columnas quieres permitir en las colaboraciones?

      - Seleccione Todas las columnas para permitir que se consulten todas las columnas de la colaboración.
      - Elija Lista personalizada para poder consultar una o más columnas de la lista desplegable Especificar columnas permitidas en la colaboración.

   b. Para los métodos de análisis permitidos,

      i. Elija Consulta directa para permitir que las consultas SQL se ejecuten directamente en esta tabla

      ii. Elija Trabajo directo para permitir que los PySpark trabajos se ejecuten directamente en esta tabla.

   Example Ejemplo

   Por ejemplo, si desea permitir que los miembros de la colaboración ejecuten consultas SQL directas y PySpark trabajos en todas las columnas, seleccione Todas las columnas, Consulta directa y Trabajo directo.

7. En Detalles de la tabla configurada,

   a. Introduzca un Nombre para la tabla configurada.

      Puede usar el nombre predeterminado o cambiar el nombre de esta tabla.

b. Introduzca una Descripción de la tabla.

La descripción ayuda a diferenciarla de otras tablas configuradas con nombres similares.

8. Si desea habilitar la opción de Etiquetas para el recurso de tabla configurada, seleccione Añadir nueva etiqueta y, a continuación, introduzca el par de Clave y Valor.
9. Seleccione Configurar nuevo cliente.

Ahora que ha creado una tabla configurada, puede hacer lo siguiente:

- Agregación de una regla de análisis a la tabla configurada
- Asociar la tabla configurada a una colaboración

Creación de una tabla configurada: fuente de datos de Amazon Athena

La opción de fuente de datos de Amazon Athena le permite consultar los datos almacenados en Amazon S3, catalogados en el catálogo de AWS Glue datos y con acceso controlado mediante. AWS Lake Formation Se admiten tanto las tablas como las vistas. AWS Glue Data Catalog Los enlaces a los recursos de Lake Formation se pueden usar para compartir tablas y vistas Regiones de AWS a lo largo Cuentas de AWS y ancho de la cuenta del AWS Clean Rooms miembro que los une a una AWS Clean Rooms colaboración.

> **Note**
>
> Solo se pueden consultar los conjuntos de datos basados en Amazon S3 mediante la integración de fuentes de datos de Athena.

En este procedimiento, el miembro realiza las siguientes tareas:

- Configura una tabla o vista existente en el para usarla AWS Glue Data Catalog AWS Clean Rooms
- Asigna un nombre a la tabla configurada y elige qué columnas usar en la colaboración.

El siguiente procedimiento asume que:

- El miembro de la colaboración ya creó la AWS Glue Data Catalog base de datos y la vista de tabla o GDC.

Para crear una tabla configurada: fuente de datos de Athena

1. Inicie sesión AWS Management Console y abra la AWS Clean Rooms consola con su Cuenta de AWS (si aún no lo ha hecho).
2. En el panel de navegación izquierdo, elija Tables (Tablas).
3. En la esquina superior derecha, elija Configurar nueva tabla.
4. En Fuente de datos, en Fuentes AWS de datos, elija Amazon Athena.
5. En la tabla de Amazon Athena:

   a. Elija la base de datos en la lista desplegable.

   b. Seleccione la Tabla que desea configurar en la lista desplegable.

   > **Note**
   >
   > Para comprobar que se trata de la tabla correcta, realice una de las siguientes acciones:
   >
   > - Seleccione Ver en AWS Glue.
   > - Activa Ver esquema desde AWS Glue para ver el esquema.

6. Para las configuraciones de Amazon Athena,

   a. Elija un grupo de trabajo de la lista desplegable.

   b. Para la ubicación de salida de S3, elija una acción recomendada en función de uno de los siguientes escenarios.

Escenario	Acción recomendada
Su grupo de trabajo no tiene una ubicación de salida predeterminada.	Introduzca la ubicación de salida de S3 o elija Browse S3.
Su grupo de trabajo impone su ubicación de salida predeterminada.	La ubicación de salida del S3 se elige automáticamente y no puede cambiarla.
Su grupo de trabajo no impone su ubicación de salida predeterminada.	Introduzca la ubicación de salida de S3 o elija Browse S3.

7. En el caso de las columnas permitidas en las colaboraciones, elige una opción en función de tu objetivo.

Su objetivo	Opción recomendada
Permite el uso de todas las columnas AWS Clean Rooms (sujeto a las reglas de análisis)	Todas las columnas
Permita una o más columnas de la lista desplegable Especificar columnas permitidas	Lista personalizada

8. En Detalles de la tabla configurada,

   a. Introduzca un Nombre para la tabla configurada.

      Puede usar el nombre predeterminado o cambiar el nombre de esta tabla.

   b. Introduzca una Descripción de la tabla.

      La descripción ayuda a diferenciarla de otras tablas configuradas con nombres similares.

   c. Si desea habilitar la opción de Etiquetas para el recurso de tabla configurada, seleccione Añadir nueva etiqueta y, a continuación, introduzca el par de Clave y Valor.

9. Seleccione Configurar nuevo cliente.

Ahora que ha creado una tabla configurada, puede hacer lo siguiente:

- Agregación de una regla de análisis a la tabla configurada
- Asociar la tabla configurada a una colaboración

Creación de una tabla configurada: fuente de datos de Snowflake

En este procedimiento, el miembro realiza las siguientes tareas:

- Configura una tabla Snowflake existente para usarla en. AWS Clean Rooms(este paso se puede realizar antes o después de unirse a una colaboración, a menos que se utilice la computación criptográfica para Clean Rooms).
- Asigna un nombre a la tabla configurada y elige qué columnas usar en la colaboración.

El siguiente procedimiento asume que:

- El miembro de la colaboración ya ha subido sus tablas de datos a Snowflake.
- (Opcional) Solo para las tablas de datos cifrados, el miembro de la colaboración ya ha preparado las tablas de datos cifrados con el cliente de cifrado de C3R.

Para crear una tabla configurada: fuente de datos de Snowflake

1. Inicie sesión en la AWS Clean Rooms consola AWS Management Console y ábrala con su Cuenta de AWS (si aún no lo ha hecho).
2. En el panel de navegación izquierdo, elija Tables (Tablas).
3. En la esquina superior derecha, elija Configurar nueva tabla.
4. En Fuente de datos, en Nubes y fuentes de datos de terceros, selecciona Snowflake.
5. Especifique las credenciales de Snowflake utilizando un ARN secreto existente o almacenando un secreto nuevo para esta tabla.

   Use existing secret ARN

   1. Si tienes un ARN secreto, introdúcelo en el campo ARN secreto.

      Para buscar tu ARN secreto, selecciona Ir a. AWS Secrets Manager
   2. Si ya tiene un secreto de otra tabla, elija Importar el ARN secreto de la tabla existente.

   > **Note**
   >
   > El ARN secreto puede ser multicuenta.

   Store a new secret for this table

   1. Introduzca las siguientes credenciales de Snowflake:
      - Nombre de usuario de Snowflake
      - Contraseña de Snowflake
      - Almacén de Snowflake
      - Papel de Snowflake

2. Para usar el valor predeterminado Clave administrada de AWS, deje desmarcada la casilla Personalizar la configuración de cifrado.
3. Para usar una AWS KMS key, seleccione la casilla Personalizar la configuración de cifrado e introduzca la clave KMS.
4. Introduzca un nombre secreto que le ayude a encontrar sus credenciales más adelante.

6. Para los detalles de la tabla y el esquema de Snowflake, introdúzcalos manualmente o impórtelos automáticamente.

   Enter the details manually

   1. Introduzca el identificador de la cuenta de Snowflake.

      Para obtener más información, consulte los identificadores de cuenta en la documentación de Snowflake.

      El identificador de su cuenta debe tener el formato utilizado para los conductores de Snowflake. Debe sustituir el punto (.) por un guión (-) para que el identificador tenga el formato siguiente. **<orgname>-<account_name>**
   2. Introduzca la base de datos de Snowflake.

      Para obtener más información, consulte la base de datos de Snowflake en la documentación de Snowflake.
   3. Introduzca el nombre del esquema de Snowflake.
   4. Introduzca el nombre de la tabla Snowflake.

      Para obtener más información, consulte Descripción de las estructuras de las tablas de Snowflake en la documentación de Snowflake.
   5. Para el esquema, introduzca el nombre de la columna y elija el tipo de datos en la lista desplegable.
   6. Seleccione Añadir columna para añadir más columnas.
      - Si elige un tipo de datos de objeto, especifique el esquema del objeto.

        Example Ejemplo de esquema de objetos

        ```
        name STRING,
        location OBJECT(
            x INT,
        ```

```
    y INT,
    metadata OBJECT(uuid STRING)
),
history ARRAY(TEXT)
```

- Si elige un tipo de datos de matriz, especifique el esquema de matriz.

  Example Ejemplo de esquema de matriz

  ```
  OBJECT(x INT, y INT)
  ```

- Si elige un tipo de datos de mapa, especifique el esquema de mapa.

  Example Ejemplo de esquema de mapa

  ```
  STRING, OBJECT(x INT, y INT)
  ```

Automatically import the details

1. Exporte la vista de COLUMNAS de Snowflake como un archivo CSV.

   Para obtener más información sobre la vista COLUMNAS de Snowflake, consulte la vista COLUMNAS en la documentación de Snowflake.

2. Seleccione Importar desde un archivo para importar el archivo CSV y especificar cualquier información adicional.

   El nombre de la base de datos, el nombre del esquema, el nombre de la tabla, los nombres de las columnas y los tipos de datos se importan automáticamente.

   - Si elige un tipo de datos de objeto, especifique el esquema del objeto.
   - Si elige un tipo de datos de matriz, especifique el esquema de matriz.
   - Si elige un tipo de datos de mapa, especifique el esquema de mapa.

3. Introduzca el identificador de la cuenta de Snowflake.

   Para obtener más información, consulte los identificadores de cuenta en la documentación de Snowflake.

> **Note**
>
> Solo las tablas de S3 catalogadas se AWS Glue pueden usar para recuperar el esquema de la tabla automáticamente.

7. En el caso de las columnas permitidas en las colaboraciones, elige una opción en función de tu objetivo.

Su objetivo	Opción recomendada
Permite el uso de todas las columnas AWS Clean Rooms (sujeto a las reglas de análisis)	Todas las columnas
Permita una o más columnas de la lista desplegable Especificar columnas permitidas	Lista personalizada

8. En Detalles de la tabla configurada,

   a. Introduzca un Nombre para la tabla configurada.

      Puede usar el nombre predeterminado o cambiar el nombre de esta tabla.

   b. Introduzca una Descripción de la tabla.

      La descripción ayuda a diferenciarla de otras tablas configuradas con nombres similares.

   c. Si desea habilitar la opción de Etiquetas para el recurso de tabla configurada, seleccione Añadir nueva etiqueta y, a continuación, introduzca el par de Clave y Valor.

9. Seleccione Configurar nuevo cliente.

Ahora que ha creado una tabla configurada, puede hacer lo siguiente:

- Agregación de una regla de análisis a la tabla configurada
- Asociar la tabla configurada a una colaboración

Agregación de una regla de análisis a una tabla configurada

En las secciones siguientes se describe cómo agregar una regla de análisis a la tabla configurada. Al definir las reglas de análisis, puede autorizar al miembro que puede realizar consultas a ejecutar consultas que coincidan con una regla de análisis específica admitida en AWS Clean Rooms.

AWS Clean Rooms admite los siguientes tipos de reglas de análisis:

- Regla de análisis de agregación
- Regla de análisis de lista
- Regla de análisis personalizada en AWS Clean Rooms

Puede haber una sola regla de análisis por tabla configurada. Puede configurar la regla de análisis en cualquier momento antes de asociar las tablas configuradas a la colaboración.

> **Important**
>
> Si utiliza la computación criptográfica para Clean Rooms y dispone de tablas de datos cifrados en la colaboración, la regla de análisis que añada a la tabla configurada cifrada debe ser coherente con la forma en que se cifraron los datos. Por ejemplo, si cifró los datos de SELECT (regla de análisis de agregación), no debe añadir la regla de análisis para JOIN (regla de análisis de listas).

Temas

- Agregación de una regla de análisis de agregación a una tabla (flujo guiado)
- Agregación de una regla de análisis de lista a una tabla (flujo guiado)
- Agregación de una regla de análisis personalizada a una tabla (flujo guiado)
- Agregación de una regla de análisis a una tabla (editor JSON)
- Pasos a seguir a continuación

Agregación de una regla de análisis de agregación a una tabla (flujo guiado)

La regla de análisis de agregación permite realizar consultas que agreguen estadísticas sin revelar información a nivel de fila mediante COUNT, SUM, y AVG funciona a lo largo de dimensiones opcionales.

Este procedimiento describe el proceso de añadir una regla de análisis de agregación a la tabla configurada mediante la opción Flujo guiado de la consola de AWS Clean Rooms .

> **Note**
>
> Las tablas configuradas que utilizan fuentes de datos que no son de S3 solo admiten reglas de análisis personalizadas.

Para añadir la regla de análisis de agregación a una tabla (flujo guiado)

1. Inicie sesión AWS Management Console y abra la AWS Clean Rooms consola con su Cuenta de AWS (si aún no lo ha hecho).
2. En el panel de navegación izquierdo, elija Tables (Tablas).
3. Seleccione la tabla configurada.
4. En la página de detalles de la tabla configurada, seleccione Configurar regla de análisis.
5. En el paso 1: elija el tipo de regla de análisis y, en el tipo de regla de análisis, elija la opción Agregación.
6. En Método de creación, seleccione Flujo guiado y, a continuación, seleccione Siguiente.
7. En Paso 2: Especificar controles de consulta, en Funciones de agregación:

    a. Elija una Función de agregación en el menú desplegable:

    - COUNT
    - COUNT DISTINCT
    - SUM
    - SUM DISTINCT
    - AVG

b. Elija qué columnas se pueden usar en la Función de agregación en el menú desplegable Columnas.

c. (Opcional) Seleccione Añadir otra función para añadir otra función de agregación y asociar una o más columnas a esa función.

> **Note**
>
> Se requiere al menos una función de agregación.

d. (Opcional) Seleccione Eliminar para eliminar una función de agregación.

8. En Controles de combinación,

a. Seleccione una opción para Permitir que la tabla se consulte por sí misma:

Si selecciona...	Entonces...
No, solo se puede consultar el solapamiento	La tabla solo se puede consultar si está combinada con una tabla propiedad del miembro que puede realizar consultas.
Sí	La tabla se puede consultar por sí misma o cuando se combina con otras tablas.

b. En Especificar columnas de unión, elija las columnas que desee permitir que se utilicen en el INNER JOIN statement.

Esto es opcional si ha seleccionado Sí en el paso anterior.

c. En Especificar operadores permitidos para la coincidencia, elija qué operadores (si los hay) se pueden usar para hacer coincidir varias columnas de combinación. Si selecciona dos o más JOIN columnas, se requiere uno de estos operadores.

Si selecciona...	Entonces...
AND	Puede incluir `AND` en las condiciones de coincidencia `INNER JOIN` para combinar una columna con otra entre tablas.

Si selecciona...	Entonces...
OR	Puede incluir `OR` en las condiciones de coincidencia `INNER JOIN` para combinar varias coincidencias de columna entre tablas. Este operador lógico es útil para obtener una tasa de coincidencia más alta.

9. (Opcional) Para los controles de dimensiones, en el menú desplegable Especificar columnas de dimensiones, elija qué columnas quiere permitir que se utilicen en la instrucción SELECT y WHERE, GROUP BY, y ORDER BY partes de la consulta.

> **Note**
>
> No se pueden usar columnas de combinación ni de función de agregación como columnas de dimensión.

10. En Funciones escalares, seleccione una opción para ¿Qué funciones escalares desea permitir?

Si selecciona...	Entonces...
Todas ellas son compatibles actualmente con AWS Clean Rooms	Se permiten todas las funciones escalares actualmente admitidas por AWS Clean Rooms. • Puede seleccionar Ver lista para ver la lista completa de Funciones escalares admitidas en AWS Clean Rooms.
Una lista personalizada	Puede personalizar las funciones escalares que desea permitir. • Elija una o más opciones del menú desplegable Especificar las funciones escalares permitidas.
Ninguna	No se permite ninguna función escalar.

Para obtener más información, consulte Funciones escalares.

11. Seleccione Siguiente.
12. En Paso 3: Especificar controles de resultados de consulta, en Restricciones de agregación:
    a. Seleccione cada Nombre de columna en la lista desplegable.
    b. Seleccione la lista desplegable para cada número mínimo de valores distintos que deben cumplirse para que se devuelva cada fila de salida, después del COUNT DISTINCT se le aplica la función.
    c. Seleccione Añadir restricción para añadir más restricciones de agregación.
    d. (Opcional) Seleccione Eliminar para eliminar una restricción de agregación.
13. En análisis adicionales aplicados a los resultados, seleccione una opción en función del objetivo.

Su objetivo	Opción recomendada
Permita solo consultas directas en esta tabla. Impida que se ejecuten análisis adicional es en los resultados de las consultas. La tabla solo se puede utilizar para consultas directas.	No permitido
Permita, pero no exija, tanto las consultas directas como los análisis adicionales en esta tabla.	Permitido
Exija que la tabla solo se pueda utilizar en consultas directas que se procesen con uno de los análisis adicionales necesario s. Las consultas directas de esta tabla se deben procesar más a fondo antes de poder devolverse.	Obligatorio

14. Elija Siguiente.
15. En Paso 4: Revisar y configurar, revise las selecciones que realizó en los pasos anteriores, edítelas si es necesario y, a continuación, seleccione Configurar regla de análisis.

Verá un mensaje de confirmación en el que se indica que ha configurado correctamente una regla de análisis de agregación en la tabla.

Agregación de una regla de análisis de lista a una tabla (flujo guiado)

La regla de análisis de lista permite realizar consultas que generen listas de nivel de fila del solapamiento entre la tabla asociada y una tabla del miembro que puede realizar consultas.

Este procedimiento describe el proceso de agregar la regla de análisis de listas a la tabla configurada mediante la opción de flujo guiado de la AWS Clean Rooms consola.

> **Note**
>
> Las tablas configuradas que utilizan fuentes de datos que no son de S3 solo admiten reglas de análisis personalizadas.

Para añadir una regla de análisis de lista a una tabla (flujo guiado)

1. Inicie sesión AWS Management Console y abra la AWS Clean Rooms consola con su Cuenta de AWS (si aún no lo ha hecho).
2. En el panel de navegación izquierdo, elija Tables (Tablas).
3. Seleccione la tabla configurada.
4. En la página de detalles de la tabla configurada, seleccione Configurar regla de análisis.
5. En el paso 1: elija el tipo de regla de análisis y, en el tipo de regla de análisis, elija la opción Lista.
6. En Método de creación, seleccione Flujo guiado y, a continuación, seleccione Siguiente.
7. En Paso 2: Especificar controles de consulta, en Controles de combinación:

    a. En Especificar columnas de unión, elija las columnas que desee permitir que se utilicen en el INNER JOIN statement.

    b. En Especificar operadores permitidos para la coincidencia, elija qué operadores (si los hay) se pueden usar para hacer coincidir varias columnas de combinación. Si selecciona dos o más JOIN columnas, se requiere uno de estos operadores.

Si selecciona...	Entonces...
AND	Puede incluir `AND` en las condiciones de coincidencia `INNER JOIN` para combinar una columna con otra entre tablas.
OR	Puede incluir `OR` en las condiciones de coincidencia `INNER JOIN` para combinar varias coincidencias de columna entre tablas. Este operador lógico es útil para obtener una tasa de coincidencia más alta.

8. (Opcional) Para los controles de lista, en el menú desplegable Especificar las columnas de la lista, elija qué columnas desea permitir que se usen en el resultado de la consulta (es decir, que se usen en el SELECT declaración), o se utiliza para filtrar los resultados (es decir, WHERE declaración).
9. Elija Siguiente.
10. En el Paso 3: Especificar controles de los resultados de las consultas, para aplicar análisis adicionales a los resultados, seleccione una opción en función del objetivo.

Su objetivo	Opción recomendada
Permita solo consultas directas en esta tabla. Impida que se ejecuten análisis adicional es en los resultados de las consultas. La tabla solo se puede utilizar para consultas directas.	No permitido
Permita, pero no exija, tanto las consultas directas como los análisis adicionales en esta tabla.	Permitido
Exija que la tabla solo se pueda utilizar en consultas directas que se procesen con uno de los análisis adicionales necesario s. Las consultas directas de esta tabla se	Obligatorio

Su objetivo	Opción recomendada
deben procesar más a fondo antes de poder devolverse.	

11. En Paso 4: Revisar y configurar, revise las selecciones que realizó en los pasos anteriores, edítelas si es necesario y, a continuación, seleccione Configurar regla de análisis.

Verá un mensaje de confirmación en el que se indica que ha configurado correctamente una regla de análisis de lista en la tabla.

Agregación de una regla de análisis personalizada a una tabla (flujo guiado)

La regla de análisis personalizada permite realizar consultas o PySpark trabajos de SQL personalizados en una tabla configurada. La regla de análisis personalizada es obligatoria si utiliza:

- Plantillas de análisis para permitir un conjunto específico de consultas o PySpark trabajos de SQL aprobados previamente o un conjunto específico de cuentas que puedan proporcionar consultas que utilicen sus datos.
- AWS Clean Rooms Privacidad diferencial para proteger a los usuarios de los intentos de identificación.
- Fuentes de datos que no son de S3, como Amazon Athena o Snowflake.

Este procedimiento describe el proceso de agregar la regla de análisis personalizada a la tabla configurada mediante la opción de flujo guiado de la consola. AWS Clean Rooms

Para añadir una regla de análisis personalizada a una tabla (flujo guiado)

1. Inicie sesión en la AWS Clean Rooms consola AWS Management Console y ábrala con su Cuenta de AWS (si aún no lo ha hecho).
2. En el panel de navegación izquierdo, elija Tables (Tablas).
3. Seleccione la tabla configurada.
4. En la página de detalles de la tabla configurada, seleccione Configurar regla de análisis.
5. En el paso 1: elija el tipo de regla de análisis y, en el tipo de regla de análisis, elija la opción Personalizada.
6. En Método de creación, seleccione Flujo guiado y, a continuación, seleccione Siguiente.

7. En el paso 2: Especificar los controles de análisis, en el caso de los controles de análisis directos, elija una opción en función de su objetivo.

Su objetivo	Acción recomendada
Revise cada análisis nuevo antes de permitir que se ejecute en esta tabla configurada	1. En Plantillas de análisis que se pueden ejecutar, elija Agregar plantilla de análisis. 2. Elija la plantilla de colaboración y de análisis adecuada en las listas desplegab les. 3. Elija Siguiente.
En esta tabla, permita que colaboradores específicos realicen cualquier análisis del tipo elegido sin necesidad de revisarlos	1. En Tipo de análisis, a. Elija Cualquier consulta para permitir cualquier consulta creada por el usuario Cuenta de AWS que especifique. b. Elija Cualquier consulta para permitir que cualquier trabajo sea creado por la Cuenta de AWS persona que especifiq ue. 2. En Cuentas de AWS Permitido crear cualquier análisis, elija Agregar Cuenta de AWS. 3. Introduzca un Cuenta de AWS ident ificador Cuenta de AWS o selecciónelo en la lista desplegable. 4. (Opcional) Selecciona Añadir otro Cuenta de AWS para añadir otro Cuenta de AWS. 5. Elija Siguiente.

8. En el paso 3: especifique los controles de los resultados del análisis,

   a. Para los controles de resultados de Job, tenga en cuenta que no se admiten controles de resultados adicionales.

b. En los controles de resultados de la consulta, en el caso de las columnas no permitidas en la salida, elija las columnas que desee que se permitan en la salida de la consulta, en función de su objetivo.

Su objetivo	Acción recomendada
Permitir que se devuelvan todas las columnas en los resultados de las consultas	1. Elige Ninguna 2. Continúe con los análisis adicionales aplicados a la salida.
Impedir que se devuelvan determina das columnas en los resultados de las consultas	1. Elija Lista personalizada 2. En Especificar columnas no permitidas, elija las columnas que desee eliminar de los resultados de la consulta.

c. Para aplicar análisis adicionales a la salida, elija si se pueden aplicar análisis adicionales a la salida de la consulta, en función de su objetivo.

Su objetivo	Opción recomendada
• Permita solo consultas directas en esta tabla. • Impida que se ejecuten análisis adicional es en los resultados de las consultas. • La tabla solo se puede utilizar para consultas directas.	No permitido
Permita, pero no exija, tanto las consultas directas como los análisis adicionales en esta tabla.	Permitir

Su objetivo	Opción recomendada
• Exija que la tabla solo se pueda utilizar en consultas directas que se procesen con uno de los análisis adicionales necesarios. • Las consultas directas de esta tabla se deben procesar más a fondo antes de poder devolverse.	Obligatorio

d. Elija Siguiente.

9. (Opcional) En el paso 4: Establecer la privacidad diferencial, determine si desea activar o desactivar la privacidad diferencial.

   La privacidad diferencial es una técnica probada matemáticamente para proteger sus datos de los ataques de reidentificación.

   > **Note**
   >
   > AWS Clean Rooms La privacidad diferencial solo está disponible para las colaboraciones que utilizan AWS Clean Rooms SQL como motor de análisis y los datos almacenados en Amazon S3.

   En el caso de la privacidad diferencial, elija si desea activar o desactivar la privacidad diferencial en función de su objetivo.

Su objetivo	Acción recomendada
• No necesitas protección contra los intentos de reidentificación • La tabla no tiene datos a nivel de usuario	1. Seleccione Desactivar. 2. Elija Siguiente.
• Necesita protección contra los intentos de reidentificación • Su tabla tiene datos a nivel de usuario	1. Elija Turn on. 2. Seleccione la columna del identificador de usuario que contiene el identificador

Su objetivo	Acción recomendada
	único de sus usuarios, como la `user_id` columna cuya privacidad desea proteger. Para activar la privacidad diferencial para dos o más tablas de una colaboración, debe configurar la misma columna que la columna Identificador de usuario en ambas reglas de análisis para mantener una definición coherente de los usuarios en todas las tablas. Si la configuración es incorrecta, el miembro que puede realizar la consulta recibe un mensaje de error que indica que hay dos columnas entre las que elegir para calcular el número de contribuciones de los usuarios (por ejemplo, el número de impresiones de anuncios realizadas por un usuario) al ejecutar la consulta. 3. Elija Siguiente.

10. En Paso 5: Revisar y configurar, revise las selecciones que realizó en los pasos anteriores, edítelas si es necesario y, a continuación, elija Configurar regla de análisis.

Ve un mensaje de confirmación en el que se indica que ha configurado correctamente una regla de análisis personalizada en la tabla.

Agregación de una regla de análisis a una tabla (editor JSON)

El siguiente procedimiento muestra cómo añadir una regla de análisis a una tabla mediante la opción del editor JSON de la AWS Clean Rooms consola.

> **Note**
>
> Las tablas configuradas que utilizan fuentes de datos que no son de S3 solo admiten reglas de análisis personalizadas.

Agregación de una regla de análisis de agregación, de lista o personalizada a una tabla (editor JSON)

1. Inicie sesión AWS Management Console y abra la AWS Clean Rooms consola con su Cuenta de AWS (si aún no lo ha hecho).
2. En el panel de navegación izquierdo, elija Tables (Tablas).
3. Seleccione la tabla configurada.
4. En la página de detalles de la tabla configurada, seleccione Configurar regla de análisis.
5. En el paso 1: elija el tipo de regla de análisis y, en el tipo de regla de análisis, elija la opción Agregación, Lista o Personalizada.
6. En Método de creación, seleccione Editor JSON y, a continuación, seleccione Siguiente.
7. En Paso 2: Especificar controles, puede optar por insertar una estructura de consulta (Insertar plantilla) o insertar un archivo (Importar desde archivo).

Si selecciona...	Entonces...
Insertar plantilla	1. Especifique los parámetros de la regla de análisis seleccionada en Definición de la regla de análisis. 2. Puede pulsar Ctrl + barra espaciadora para habilitar la función de autocompletar. Para obtener más información acerca de los parámetros de la regla de análisis de agregación, consulte Regla de análisis de agregación: controles de consulta. Para obtener más información acerca de los parámetros de la regla de análisis de lista, consulte Regla de análisis de lista: controles de consulta.
Importar desde archivo	1. Seleccione el archivo JSON en su unidad local. 2. Elija Open.

Si selecciona...	Entonces...
	La Definición de la regla de análisis muestra la regla de análisis del archivo cargado.

8. Elija Siguiente.
9. En Paso 3: Revisar y configurar, revise las selecciones que realizó en los pasos anteriores, edítelas si es necesario y, a continuación, seleccione Configurar regla de análisis.

Recibe un mensaje de confirmación en el que se indica que ha configurado correctamente una regla de análisis en la tabla.

Pasos a seguir a continuación

Ahora que ha configurado una regla de análisis para la tabla configurada, ya puede:

- Asociar una tabla configurada a una colaboración
- Consultar las tablas de datos (como miembro que puede realizar consultas)

Asociar una tabla configurada a una colaboración

Una vez que haya creado una tabla configurada y le haya agregado una regla de análisis, puede asociarla a una colaboración y asignar AWS Clean Rooms una función de servicio para acceder a AWS Glue las tablas.

> **Note**
>
> Este rol de servicio tiene permisos relacionados con las tablas. AWS Clean Rooms solo puede asumir el rol de servicio para ejecutar las consultas permitidas en nombre del miembro que puede realizar consultas. Ningún miembro de la colaboración (salvo el propietario de los datos) tiene acceso a las tablas subyacentes de la colaboración. El propietario de los datos puede activar la privacidad diferencial para que sus tablas estén disponibles para que otros miembros las consulten.

Important

Antes de asociar las AWS Glue tablas configuradas a la colaboración, la ubicación de la AWS Glue tabla debe apuntar a una carpeta de Amazon Simple Storage Service (Amazon S3) y no a un único archivo. Puede verificar esta ubicación consultando la tabla de la AWS Glue consola en https://console.aws.amazon.com/glue/.

Note

Si ha configurado el cifrado AWS Glue y creado un rol de servicio, debe conceder acceso a ese rol para que lo utilice AWS KMS keys para descifrar AWS Glue tablas.
Si asoció una tabla configurada respaldada por un conjunto de datos de Amazon S3 AWS KMS cifrado, debe conceder al rol acceso para usar la clave de KMS para descifrar los datos de Amazon S3.
Para obtener más información, consulte Configuración del cifrado en AWS Glue, en la Guía del desarrollador de AWS Glue .

En los temas siguientes se describe cómo asociar una tabla configurada a una colaboración mediante la AWS Clean Rooms consola:

Temas

- Asociar una tabla configurada desde la página de detalles de la tabla configurada
- Asociar una tabla configurada desde la página de detalles de la colaboración
- Pasos a seguir a continuación

Para obtener información sobre cómo asociar las tablas configuradas a la colaboración mediante la AWS SDKs, consulte la referencia de la AWS Clean Rooms API.

Asociar una tabla configurada desde la página de detalles de la tabla configurada

Para asociar AWS Glue tablas a la colaboración desde la página de detalles de la tabla configurada

1. Inicie sesión en la consola AWS Management Console y abra la AWS Clean Rooms consola con su Cuenta de AWS (si aún no lo ha hecho).

2. En el panel de navegación izquierdo, elija Tables (Tablas).
3. Seleccione la tabla configurada.
4. En la página de detalles de la tabla configurada, seleccione Asociar a colaboración.
5. En el cuadro de diálogo Asociar tabla a colaboración, seleccione la Colaboración en la lista desplegable.
6. Seleccione Elegir colaboración.

   En la página Asociar tabla, el nombre de la tabla configurada que ha seleccionado aparece en la sección Elegir tabla configurada.
7. (Opcional) Para Elegir una tabla configurada, haga lo siguiente:

Si desea…	Entonces...
Configurar una nueva tabla	Seleccione Configurar tabla y siga las indicaciones de la página Configurar tabla.
Ver el esquema y la regla de análisis de la tabla configurada	Active Ver esquema y regla de análisis.

8. En Detalles de asociación de tablas,

   a. Introduzca un Nombre para la tabla asociada.

      Puede usar el nombre predeterminado o cambiar el nombre de esta tabla.

   b. (Opcional) Ingrese una Descripción de la tabla.

      La descripción ayuda a escribir las consultas.
9. Especifique los permisos de Acceso a servicios seleccionando Crear y usar un nuevo rol de servicio o Usar un rol de servicio existente.

> **Note**
>
> Si va a asociar una tabla configurada respaldada por Amazon Athena, elija un nombre de rol de servicio existente en la lista desplegable. Asegúrese de que la función de servicio tenga permisos de IAM y, si es necesario, de Lake Formation para el conjunto de datos.

Si selecciona...	Entonces...
Crear y usar un nuevo rol de servicio	• AWS Clean Rooms crea un rol de servicio con la política requerida para esta tabla. • El Nombre del rol de servicio predeterminado es `cleanrooms-<timestamp>` • Debe tener permisos para crear roles y adjuntar políticas. • Si los datos de entrada están cifrados, puede seleccionar Estos datos están cifrados con una clave KMS y, a continuación, introducir una AWS KMS key que se utilizará para descifrar los datos introducidos.

Si selecciona...	Entonces...
Usar un rol de servicio existente	1. Seleccione un Nombre de rol de servicio existente en la lista desplegable. Si tiene permisos de listas de roles, se mostrará la lista de roles. Si no tiene permisos de listas de roles, puede ingresar el nombre de recurso de Amazon (ARN) del rol que desea usar. 2. Consulte el rol de servicio mediante la elección del enlace externo Ver en IAM. Si no hay ningún rol de servicio existente, la opción Usar un rol de servicio existente no estará disponible. De forma predeterminada, AWS Clean Rooms no intenta actualizar la política de roles existente para añadir los permisos necesarios. 3. (Opcional) Seleccione la casilla de verificación Agregar una política preconfig urada con los permisos necesarios para este rol para adjuntar los permisos necesarios al rol. Debe tener permisos para modificar roles y crear políticas.

Note

- AWS Clean Rooms requiere permisos para realizar consultas de acuerdo con las reglas de análisis. Para obtener más información sobre los permisos para AWS Clean Rooms, consulteAWS políticas gestionadas para AWS Clean Rooms.

- Si el rol no tiene permisos suficientes AWS Clean Rooms, recibirá un mensaje de error que indica que el rol no tiene permisos suficientes AWS Clean Rooms. Debe agregar la política de rol antes de continuar.
- Si no puede modificar la política de rol, recibirá un mensaje de error que indica que AWS Clean Rooms no ha podido encontrar la política del rol de servicio.

10. Si desea habilitar Etiquetas para el recurso de asociación de tablas configuradas, seleccione Agregar nueva etiqueta y, a continuación, introduzca el par de Clave y Valor.
11. Seleccione Asociar tabla.

Asociar una tabla configurada desde la página de detalles de la colaboración

Para asociar AWS Glue tablas a la colaboración desde la página de detalles de la colaboración

1. Inicie sesión en la consola AWS Management Console y abra la AWS Clean Rooms consola con su Cuenta de AWS (si aún no lo ha hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Seleccione la colaboración.
4. En la pestaña Tablas, seleccione Asociar tabla.
5. En Elegir tabla configurada, realiza una de las siguientes acciones:

Si desea…	Entonces...
Elegir una tabla configurada existente	Seleccione el Nombre de la tabla configura da que desea asociar a la colaboración en la lista desplegable.
Configurar una nueva tabla	Seleccione Configurar tabla y siga las indicaciones de la página Configurar tabla.
Ver el esquema y la regla de análisis de la tabla configurada	Active Ver esquema y regla de análisis.

6. En Detalles de asociación de tablas,

a. Introduzca un Nombre para la tabla asociada.

   Puede usar el nombre predeterminado o cambiar el nombre de esta tabla.

b. (Opcional) Ingrese una Descripción de la tabla.

   La descripción ayuda a escribir las consultas.

7. Especifique los permisos de Acceso a servicios seleccionando Crear y usar un nuevo rol de servicio o Usar un rol de servicio existente.

> **Note**
>
> Si va a asociar una tabla configurada respaldada por Amazon Athena, elija un nombre de rol de servicio existente en la lista desplegable. Asegúrese de que la función de servicio tenga permisos de IAM y, si es necesario, de Lake Formation para el conjunto de datos.

Si selecciona...	Entonces...
Crear y usar un nuevo rol de servicio	• AWS Clean Rooms crea un rol de servicio con la política requerida para esta tabla. • El Nombre del rol de servicio predeterminado es `cleanrooms-<timestamp>` . • Debe tener permisos para crear roles y adjuntar políticas. • Si los datos de entrada están cifrados, puede seleccionar Estos datos están cifrados con una clave KMS y, a continuación, introducir una AWS KMS key que se utilizará para descifrar los datos introducidos.
Usar un rol de servicio existente	1. Seleccione un Nombre de rol de servicio existente en la lista desplegable.

Si selecciona...	Entonces...
	Si tiene permisos de listas de roles, se mostrará la lista de roles. Si no tiene permisos de listas de roles, puede ingresar el nombre de recurso de Amazon (ARN) del rol que desea usar. 2. Consulte el rol de servicio mediante la elección del enlace externo Ver en IAM. Si no hay ningún rol de servicio existente, la opción Usar un rol de servicio existente no estará disponible. De forma predeterminada, AWS Clean Rooms no intenta actualizar la política de roles existente para añadir los permisos necesarios. 3. (Opcional) Seleccione la casilla de verificación Agregar una política preconfig urada con los permisos necesarios para este rol para adjuntar los permisos necesarios al rol. Debe tener permisos para modificar roles y crear políticas.

> **Note**
>
> - AWS Clean Rooms requiere permisos para realizar consultas de acuerdo con las reglas de análisis. Para obtener más información sobre los permisos para AWS Clean Rooms, consulteAWS políticas gestionadas para AWS Clean Rooms.
> - Si el rol no tiene permisos suficientes AWS Clean Rooms, recibirá un mensaje de error que indica que el rol no tiene permisos suficientes AWS Clean Rooms. Debe agregar la política de rol antes de continuar.

- Si no puede modificar la política de rol, recibirá un mensaje de error que indica que AWS Clean Rooms no ha podido encontrar la política del rol de servicio.

8. Si desea habilitar Etiquetas para el recurso de asociación de tablas configuradas, seleccione Agregar nueva etiqueta y, a continuación, introduzca el par de Clave y Valor.
9. Seleccione Asociar tabla.

Pasos a seguir a continuación

Ahora que ha asociado la tabla de datos configurada a la colaboración, ya puede:

- Agregación de una regla de análisis de colaboración a una tabla configurada
- Editar la colaboración (si es el creador de la colaboración)
- Consultar las tablas de datos (como miembro que puede realizar consultas)

Agregación de una regla de análisis de colaboración a una tabla configurada

La regla de análisis de colaboración le permite especificar los controles que son específicos de esta colaboración. Estos controles funcionan junto con la regla de análisis de tablas configurada para determinar cómo se puede analizar esta tabla en esta colaboración.

Se agrega una regla de análisis de colaboración a una tabla configurada después de haber creado una tabla configurada, haber agregado una regla de análisis y haberla asociado a una colaboración. Debe agregar una regla de análisis de colaboración si la tabla está configurada para admitir el análisis directo o para permitir un análisis adicional.

- Análisis directo: la tabla se puede utilizar en consultas que la analicen directamente. Por ejemplo, en una consulta que genera un análisis de medición agregado o una lista de identificadores para su activación.
- Análisis adicional: la tabla también se puede utilizar como entrada en análisis adicionales, además de en consultas que la analizan directamente. Por ejemplo, la tabla se puede usar en una consulta que sea una semilla para un modelo de ML similar o un canal de entrada de ML para un modelo de ML personalizado.

Agregación de la regla de análisis de agregación a una tabla

1. Inicia sesión en AWS Management Console y abre la AWS Clean Rooms consola con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Seleccione la colaboración.
4. En la pestaña Tablas, en Tablas asociadas por usted, consulte la tabla configurada que ha asociado a la colaboración.

   Si el estado del análisis directo o el estado del análisis adicional tienen el estado Preparado, significa que la tabla está lista para la consulta.
5. Si el estado del análisis directo o el estado del análisis adicional tienen el estado No preparado, seleccione el estado y, a continuación, elija Configurar en el cuadro de diálogo.
6. En la página Configurar regla de análisis de colaboración, expanda Ver regla de análisis de tabla configurada para ver los detalles.
7. Para Análisis adicionales permitidos, elija la opción en función de su objetivo.

Su objetivo	Opción recomendada
Permita los análisis adicionales en la tabla.	Cualquiera
Permita que un miembro específico solo incluya análisis adicionales en la tabla.	Cualquiera realizado por miembros específicos
Permita solo análisis específicos en la tabla.	Lista personalizada

8. Para la entrega de resultados, especifique quién puede recibir los resultados de los miembros autorizados a recibirlos en el menú desplegable de resultados de la consulta.
9. Elija Configurar regla de análisis.

Configuración de la política de privacidad diferencial (opcional)

> **Note**
>
> AWS Clean Rooms La privacidad diferencial solo está disponible para las colaboraciones que utilizan AWS Clean Rooms SQL como motor de análisis y los datos almacenados en Amazon S3.

Este procedimiento describe el proceso de configuración de la política de privacidad diferencial en una colaboración mediante la opción de flujo guiado de la AWS Clean Rooms consola. Este paso solo se ejecuta una vez para todas las tablas con protección de privacidad diferencial.

Para configurar los ajustes de privacidad diferencial (flujo guiado)

1. Inicie sesión en la AWS Clean Rooms consola AWS Management Console y ábrala con su Cuenta de AWS (si aún no lo ha hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Seleccione la colaboración.
4. En la pestaña Tablas de la página de colaboración, elija Configuración de la política de privacidad diferencial.
5. En la página Configuración de la política de privacidad diferencial, elija valores para las siguientes propiedades:

   - Presupuesto de privacidad
   - Actualizar mensualmente el presupuesto de privacidad
   - Ruido agregado por consulta

   Puede utilizar los valores predeterminados o introducir valores personalizados que admitan su caso de uso específico. Tras elegir los valores de Presupuesto de privacidad y Ruido añadido por consulta, puede obtener una vista previa de la utilidad resultante en lo que respecta al número de agregaciones posibles en todas las consultas de sus datos.
6. Elija Configurar.

Verá un mensaje de confirmación en el que se indica que ha configurado correctamente la política de privacidad diferencial para la colaboración.

Ahora que ha configurado la privacidad diferencial, puede hacer lo siguiente:

- Consultar las tablas de datos (como miembro que puede realizar consultas)
- Colaboraciones (si es el creador de la colaboración)

Visualización de los registros de uso de la privacidad diferencial

Como miembro de una colaboración que protege los datos con privacidad diferencial, después de haber creado una colaboración con privacidad diferencial, puede supervisar el uso del presupuesto de privacidad.

Para ver cuántas agregaciones se ejecutaron y qué parte del presupuesto de privacidad se utilizó

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con la tuya Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Seleccione la colaboración.
4. Seleccione la pestaña Tablas.
5. Seleccione Ver registros de uso (texto azul).
6. Consulte los detalles de uso, incluido el presupuesto de privacidad y cuánta utilidad se ha proporcionado.

Edición de una política de privacidad diferencial

Después de configurar la política de privacidad diferencial, puede actualizarla en cualquier momento para que refleje mejor sus necesidades de privacidad.

Para editar la política de privacidad diferencial

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con la tuya Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Seleccione la colaboración.
4. En la pestaña Tablas de la página de colaboración, en Tablas asociadas por usted, elija Editar.
5. En la página Editar privacidad diferencial, elija nuevos valores para las siguientes propiedades:

- Presupuesto de privacidad: mueva la barra del control deslizante para aumentar o disminuir el presupuesto en cualquier momento de una colaboración. No puede reducir el presupuesto después de que el miembro que puede realizar la consulta haya empezado a consultar sus datos. Si se aumenta el presupuesto de privacidad, AWS Clean Rooms seguirá utilizando el presupuesto existente hasta que se agote por completo antes de utilizar el presupuesto de privacidad recién agregado.
- Ruido agregado por consulta: mueva la barra de control deslizante para aumentar o disminuir el ruido agregado por consulta en cualquier momento durante una colaboración.

> **Note**
>
> Puede elegir Ejemplos interactivos para ver cómo los distintos valores del Presupuesto de privacidad y Ruido agregado por consulta afectan al número de funciones agregadas que puede ejecutar.

No se puede cambiar el valor de la Actualización del presupuesto de privacidad. Para cambiar su selección, debe eliminar la política de privacidad diferencial y crear una nueva.

6. Seleccione Save changes (Guardar cambios).

Aparece un mensaje de confirmación en el que se indica que ha editado correctamente la política de privacidad diferencial.

Eliminación de una política de privacidad diferencial

Puede eliminar la política de privacidad diferencial desde la pestaña Tablas de una colaboración.

Eliminación de la política de privacidad diferencial

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con la tuya Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Seleccione la colaboración.
4. En la pestaña Tablas de la página de colaboración, junto a la Política de privacidad diferencial, seleccione Eliminar.

5. Si tiene la certeza de que desea eliminar la política de privacidad diferencial, seleccione Eliminar.

Tras eliminar una política de privacidad diferencial, no podrá acceder a los registros de uso del presupuesto de privacidad de esa política. Las tablas con la privacidad diferencial activada no se pueden consultar si se elimina la política de privacidad diferencial.

Visualización de los parámetros de privacidad diferencial calculados

Los usuarios con experiencia en privacidad diferencial pueden ver los parámetros de privacidad diferencial calculados en la pestaña Consultas de una colaboración.

Visualización de los parámetros de privacidad diferencial calculados

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con la tuya Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Seleccione la colaboración.
4. En la pestaña Consultas, en la sección Resultados, seleccione Ver los parámetros de privacidad diferencial calculados.

En la tabla Parámetros de privacidad diferencial calculados, puede ver los valores de confidencialidad de las funciones agregadas, que se definen como la cantidad máxima en la que el resultado de una función puede cambiar si se añaden, eliminan o modifican los registros de un solo usuario. La lista incluye los siguientes parámetros de privacidad diferencial:

- El límite de contribución del usuario (UCL) es el número máximo de filas que aporta un usuario en una consulta SQL. Por ejemplo, si quieres contar el número total de impresiones coincidentes en una campaña específica, en la que cada usuario puede tener varias impresiones, AWS Clean Rooms Differential Privacy debe limitar el número de impresiones de un solo usuario para garantizar que el cálculo de la privacidad diferencial sea exacto. En otras palabras, si algún usuario tiene más impresiones que el límite, toma AWS Clean Rooms automáticamente una muestra aleatoria uniforme de las impresiones de ese usuario según el valor de UCL calculado y excluye las impresiones restantes de ese usuario al ejecutar la consulta. El valor de la UCL es igual a 1 si se cuenta el número de usuarios únicos. Esto se debe a que la adición, eliminación o modificación de un solo usuario puede cambiar el recuento de usuarios distintos en 1 como máximo.

- El valor mínimo es el límite inferior de una expresión que se utiliza en una función de agregación, como `sum()`. Por ejemplo, si la expresión es una columna denominada `purchase_value`, el valor mínimo es el límite inferior de la columna.
- El valor máximo es el límite superior de una expresión que se utiliza en una función de agregación, como `sum()`. Por ejemplo, si la expresión es una columna denominada `purchase_value`, el valor máximo es el límite superior de la columna.

En la tabla Parámetros de privacidad diferencial calculados, puede utilizar estos parámetros para comprender mejor la cantidad total de ruido en los resultados de las consultas. Por ejemplo, cuando el ruido agregado por consulta es de 30 usuarios y se ejecuta una `COUNT DISTINCT (user_id)` consulta, AWS Clean Rooms Differential Privacy añade un ruido aleatorio que se sitúa entre -30 y 30, con una alta probabilidad porque la sensibilidad de `COUNT DISTINCT` es 1. En el caso de una consulta `COUNT` con la misma configuración, la privacidad diferencial de AWS Clean Rooms añade ruido estadístico que se escala según el límite de contribución del usuario, ya que un solo usuario podría añadir varias filas al resultado de la consulta. En el caso de `SUM` consultas como `SUM (purchase_value)` las que tienen valores positivos en todas las columnas, el ruido total se escala en función del límite de contribución del usuario multiplicado por el valor máximo. AWS Clean Rooms Differential Privacy calcula automáticamente los parámetros de sensibilidad para añadir ruido en tiempo de ejecución de la consulta y agota el presupuesto de privacidad. Es necesario reducir el presupuesto de privacidad porque los parámetros de sensibilidad dependen de los datos.

Ver tablas y reglas de análisis

Para ver las tablas asociadas a la colaboración y a las reglas de análisis

1. Inicia sesión en AWS Management Console y abre la AWS Clean Rooms consola con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Seleccione la colaboración.
4. Seleccione la pestaña Tablas.
5. Seleccione una de las siguientes opciones:

    a. Para ver las tablas asociadas a la colaboración, en Tablas asociadas por usted, elija una tabla (texto azul).
    b. Para ver otras tablas asociadas a la colaboración, en Tablas asociadas por colaboradores, elija una tabla (texto azul).

6. Vea los detalles de la tabla y las reglas de análisis en la página de detalles de la tabla.

Editar detalles de tablas configuradas

Como miembro de la colaboración, puede editar los detalles de la tabla configurada.

Para editar detalles de tablas configuradas

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con la tuya Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Tables (Tablas).
3. Elija la tabla configurada que creó.
4. En la página de detalles de la tabla configurada, desplácese hacia abajo hasta Detalles de tabla configurada.
5. Elija Editar.
6. Actualice el Nombre o la Descripción de la tabla configurada.
7. Elija Guardar cambios.

Editar etiquetas de tablas configuradas

Como miembro de la colaboración, después de crear una tabla configurada, puede administrar las etiquetas del recurso de tabla configurada en la pestaña Tablas configuradas.

Para editar las etiquetas de tablas configuradas

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con la tuya Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Tables (Tablas).
3. Elija la tabla configurada que creó.
4. En la página de detalles de la tabla configurada, desplácese hacia abajo hasta la sección Etiquetas.
5. Elija Administrar etiquetas.
6. En la página Administrar etiquetas, puede hacer lo siguiente:

    - Para eliminar una etiqueta, elija Eliminar.

- Para añadir una etiqueta, elija Añadir nueva etiqueta.
- Para guardar los cambios, elija Guardar cambios.

Editar la regla de análisis de tablas configuradas

Para editar la regla de análisis de tablas configuradas

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con la tuya Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Tables (Tablas).
3. Elija la tabla configurada que creó.
4. En la página de detalles de la tabla configurada, desplácese hacia abajo hasta la sección Regla de análisis de agregación, Regla de análisis de lista o Regla de análisis personalizada (su elección dependerá del tipo de regla de análisis que haya elegido para la tabla configurada).
5. Elija Editar.
6. En la página Editar regla de análisis, puede:
   - Modificar la Definición de la regla de análisis de la siguiente manera:
     - Modificando el editor JSON.
     - Seleccione Importar desde archivo para cargar una nueva definición de la regla de análisis.
   - Para obtener una vista previa de lo que verán los miembros de una colaboración, seleccione una de las siguientes opciones:
     - Vista de tabla
     - JSON
     - Consulta de ejemplo
7. Elija Guardar cambios para guardar los cambios.

Eliminar la regla de análisis de tablas configuradas

> **Warning**
>
> Esta acción no se puede deshacer y afecta a todos los recursos relacionados.

Para eliminar la regla de análisis de tablas configuradas

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con la tuya Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Tables (Tablas).
3. Elija la tabla configurada que creó.
4. En la página de detalles de la tabla configurada, desplácese hacia abajo hasta la sección Regla de análisis de agregación, Regla de análisis de lista o Regla de análisis personalizada (su elección dependerá del tipo de regla de análisis que haya elegido para la tabla configurada).
5. Elija Eliminar.
6. Si tiene la certeza de que desea eliminar la regla de análisis, elija Eliminar.

Columnas no permitidas de la tabla configurada

La configuración de columnas de salida no permitidas es un control de la regla de análisis AWS Clean Rooms personalizada que le permite definir la lista de columnas (si las hay) que no permite que se proyecten en el resultado de la consulta. Las columnas a las que se hace referencia en esta lista se consideran “columnas de salida no permitidas”. Esto significa que cualquier referencia a dicha columna mediante transformación, alias u otros medios puede no estar presente en la SELECCIÓN (proyección) final de la consulta.

Aunque esta funcionalidad prohíbe que las columnas se proyecten directamente en la salida, no impide por completo que los valores subyacentes se deduzcan indirectamente mediante otros mecanismos. Estas columnas se pueden seguir utilizando en una cláusula de proyección (por ejemplo, en una subconsulta o en una expresión de tabla común [CTE]), siempre y cuando no se haga referencia a ellas en la proyección final.

La configuración de columnas de salida no permitidas le ofrece la flexibilidad de aplicar y codificar el control de la tabla, junto con revisiones de nivel de plantilla de análisis en función de los casos de uso y los requisitos de privacidad correspondientes.

Para obtener más información acerca de cómo definir esta configuración, consulte Agregación de una regla de análisis personalizada a una tabla (flujo guiado).

Ejemplos

Los siguientes ejemplos muestran cómo se aplica el control de columnas de salida no permitidas.

- El miembro A colabora con el miembro B.
- El miembro B es quien puede ejecutar consultas.
- El miembro A define a los usuarios de una tabla con las columnas edad, sexo, correo electrónico y nombre. Las columnas de edad y nombre no son columnas de salida permitidas.
- El miembro B define una tabla mascotas con un conjunto similar de columnas edad, género y owner_name. Sin embargo, no establecen ninguna restricción en las columnas de salida, lo que significa que todas las columnas de la tabla se pueden proyectar libremente en la consulta.

Si el miembro B ejecuta la siguiente consulta, se bloquea porque las columnas de salida no permitidas no se pueden proyectar directamente:

```
SELECT
  age
FROM
  users
```

Si el miembro B ejecuta la siguiente consulta, se bloquea porque las columnas de salida no permitidas no se pueden proyectar implícitamente a través de la estrella del proyecto:

```
SELECT
  *
FROM
  users
```

Si el miembro B ejecuta la siguiente consulta, se bloquea porque las transformaciones de las columnas de salida no permitidas no se pueden proyectar:

```
SELECT
  COUNT(age)
FROM
  users
```

Si el miembro B ejecuta la siguiente consulta, se bloquea porque no se puede hacer referencia a las columnas de salida no permitidas en la proyección final con un alias:

```
SELECT
  count_age
FROM
  (SELECT COUNT(age) AS count_age FROM users)
```

Si el miembro B ejecuta la siguiente consulta, se bloquea porque las columnas restringidas transformadas se proyectan en la salida:

```
SELECT
  CONCAT(name, email)
FROM
  users
```

Si el miembro B ejecuta la siguiente consulta, se bloquea porque no se puede hacer referencia a las columnas de salida no permitidas en CTE en la proyección final:

```
WITH cte AS (
  SELECT
    age AS age_alias
  FROM
    users
)
SELECT age_alias FROM cte
```

Si el miembro B ejecuta la siguiente consulta, se bloquea porque no se pueden usar las columnas de salida no permitidas como claves de clasificación o partición en la proyección final:

```
SELECT
  LISTAGG(gender) WITHIN GROUP (ORDER BY age) OVER (PARTITION BY age)
FROM
  users
```

Si el miembro B ejecuta la siguiente consulta, se realiza correctamente porque las columnas que forman parte de las columnas de salida no permitidas se pueden seguir utilizando en otros constructos de la consulta, como en las cláusulas de unión o filtro.

```
SELECT
  u.name,
  p.gender,
  p.age
FROM
  users AS u
JOIN
  pets AS p
ON
  u.name = p.owner_name
```

En el mismo escenario, el miembro B también puede usar la columna de nombres en los usuarios como filtro o clave de clasificación:

```
SELECT
  u.email,
  u.gender
FROM
  users AS u
WHERE
  u.name = 'Mike'
ORDER BY
  u.name
```

Además, las columnas de salida no permitidas por los usuarios se pueden usar en proyecciones intermedias, como subconsultas y CTEs, por ejemplo:

```
WTIH 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)
```

Editar asociaciones de tablas configuradas

Como miembro de una colaboración, puede editar las asociaciones de tablas configuradas que ha creado.

Para editar asociaciones de tablas configuradas

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Seleccione la colaboración.
4. Seleccione la pestaña Tablas.
5. En Tablas asociadas por usted, elija una tabla.
6. En la página de detalles de la tabla, desplácese hacia abajo para ver los Detalles de asociación de tablas.
7. Elija Editar.
8. En la página Editar asociaciones de tablas configuradas, actualice la Descripción o la Información de acceso al servicio.
9. Elija Guardar cambios.

Disociar tablas configuradas

Como miembro de la colaboración, puede disociar una tabla configurada de la colaboración. Esta acción impide que el miembro que puede realizar consultas consulte la tabla.

Para disociar una tabla configurada

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con la tuya Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Seleccione la colaboración.
4. Seleccione la pestaña Tablas.

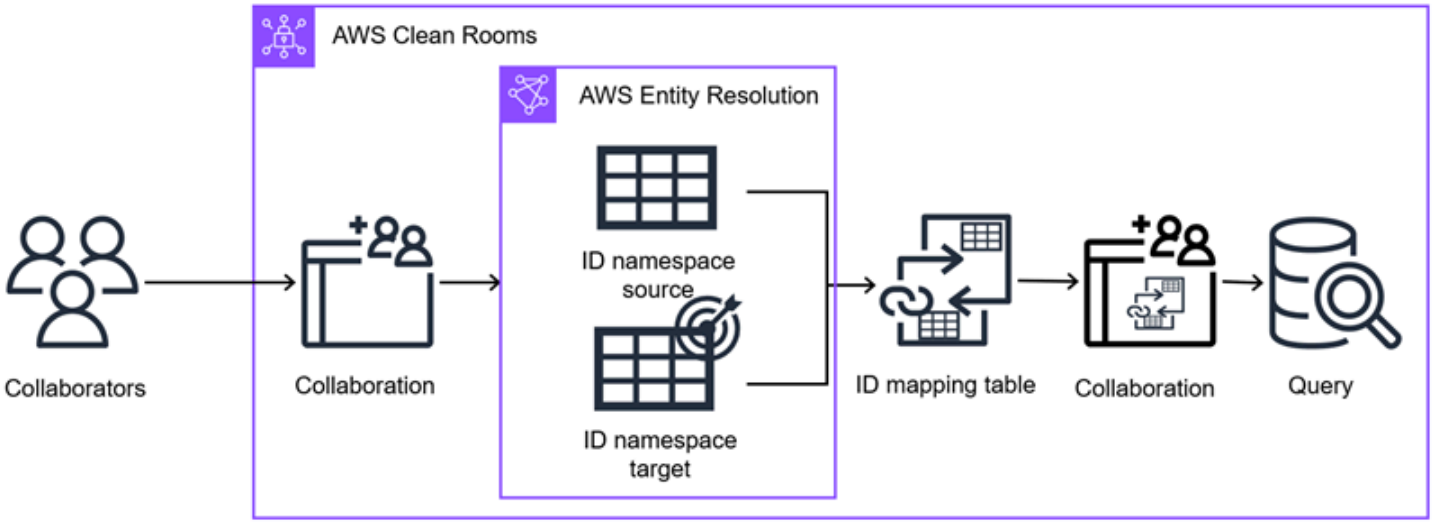
5. En Tablas asociadas por usted, elija el botón de opción situado junto a la tabla que desea disociar.
6. Elija Disociar.
7. En el cuadro de diálogo, confirme la decisión de disociar la tabla configurada, y seleccione Disociar para impedir que el miembro que puede realizar consultas consulte la tabla.

AWS Entity Resolution in AWS Clean Rooms

Con AWS Entity Resolution in AWS Clean Rooms, puede traducir los datos de una fuente a un destino, rellenar una tabla de mapeo de ID con los datos traducidos y consultarlos.

En primer lugar, puedes crear una colaboración en AWS Clean Rooms y añadir la Cuentas de AWS que quieras invitar o unirte a una colaboración a la que estés invitado creando una membresía. A continuación, realice una asignación de ID en dos tablas de datos. Para ello, asocie un origen de espacio de nombres de ID existente o cree uno nuevo en AWS Entity Resolution. El otro miembro de la colaboración asocia un destino de espacio de nombres de ID existente o crea un destino de espacio de nombres de ID nuevo. A continuación, debe crear y rellenar una tabla de asignación de ID a partir de los dos espacios de nombres de ID asociados. Por último, el miembro que puede realizar la consulta ejecuta una consulta en las dos tablas de datos mediante la unión en la tabla de asignación de ID.

En el siguiente diagrama se resume cómo trabajar con ella AWS Entity Resolution . AWS Clean Rooms



> **Note**
>
> El proveedor de servicios de transcodificación compatible actualmente es el siguiente Regiones de AWS: EE.UU. Este (Norte de Virginia), EE.UU. Este (Ohio) y EE.UU. Oeste (Oregón). LiveRamp

Temas

- Espacios de nombres de ID en AWS Clean Rooms
- Tablas de mapeo de ID en AWS Clean Rooms

Espacios de nombres de ID en AWS Clean Rooms

Un espacio de nombres de ID es un encapsulador en torno a la tabla de identidades que le permite proporcionar metadatos que explican el conjunto de datos y cómo usarlo en un flujo de trabajo de asignación de ID. Un flujo de trabajo de asignación de ID es un trabajo de procesamiento de datos que asigna datos de un origen de datos de entrada a un destino de datos de entrada en función del método de asignación de ID especificado. Produce una tabla de asignación de ID.

Hay dos tipos de espacios de nombres de ID: origen y destino. El origen contiene configuraciones para los datos de origen que se procesarán en un flujo de trabajo de asignación de ID. El destino contiene una configuración de los datos de destino que resolverán todos los orígenes. Para definir los datos de entrada que desea resolver en dos Cuentas de AWS, cree una fuente de espacio de nombres de ID y un destino de espacio de nombres de ID para traducir los datos de un conjunto (origen) a otro (destino).

Puede crear un nuevo espacio de nombres de ID o asociar uno existente. Para obtener más información sobre cómo crear un espacio de nombres de ID AWS Entity Resolution, consulte Creación de un espacio de nombres de ID en la Guía del usuario.AWS Entity Resolution

Temas

- Creación y asociación de un nuevo espacio de nombres de ID
- Asociación de un espacio de nombres de ID existente
- Edición de las asociaciones de espacios de nombres de ID
- Desasociación de las asociaciones de espacios de nombres de ID

Creación y asociación de un nuevo espacio de nombres de ID

Cada miembro de la colaboración debe crear y asociar un Origen de espacio de nombres de ID o un Destino de espacio de nombres de ID antes de crear una tabla de asignación de ID para consultar datos de identidad.

Si ya ha creado un espacio de nombres de ID en, vaya a. AWS Entity ResolutionAsociación de un espacio de nombres de ID existente

Creación y asociación de un nuevo espacio de nombres de ID

1. Inicia sesión en AWS Management Console y abre la AWS Clean Rooms consola con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Seleccione la colaboración.
4. En la pestaña Resolución de entidades, elija el Espacio de nombres de ID asociado.
5. En la página del Espacio de nombres de ID asociado, para los datos de resolución de entidades, elija Crear espacio de nombres de ID.

   La AWS Entity Resolution consola aparece en una pestaña nueva.
6. Siga las peticiones en la página Creación de un espacio de nombres de ID en la consola de AWS Entity Resolution .

   a. Para obtener más información, ingrese el nombre y la descripción del espacio de nombres de ID y seleccione el tipo de espacio de nombres de ID (origen o destino).

   b. Para el método del espacio de nombres de ID, elija el método basado en reglas para la coincidencia basada en reglas o los servicios del proveedor para la transcodificación de terceros.

   c. Especifique el tipo de entrada de datos, en función del método de espacio de nombres de ID que haya elegido.

   d. Elija Creación de un espacio de nombres de ID.
7. Vuelve a la AWS Clean Rooms consola.
8. En la página del espacio de nombres de ID asociado, para los datos de resolución de entidades, elija el origen o el destino del espacio de nombres de ID de AWS Entity Resolution que desee asociar a la colaboración en la lista desplegable.
9. Para obtener los detalles de la asociación, siga estos pasos.

   a. Ingrese un nombre para el espacio de nombres de ID asociado.

      Puede usar el nombre predeterminado o cambiar el nombre de este espacio de nombres de ID.

   b. (Opcional) Ingrese una Descripción del espacio de nombres de ID.

      La descripción ayuda a escribir las consultas.

10. Especifique los permisos de acceso de AWS Clean Rooms mediante la selección de una opción y, a continuación, la realización de la acción recomendada.

Opción	Acción recomendada
Permitir AWS Clean Rooms añadir y gestionar la política de permisos	AWS Clean Rooms crea un rol de servicio con la política requerida para esta asociación.
Agregación y administración de los permisos manualmente	Realice una de las siguientes acciones: • Revise la política de recursos y agregue los permisos necesarios a la política. • Use una política existente mediante la elección de Agregar declaración de política. Debe tener permisos para modificar roles y crear políticas. **Note** Si no puede modificar la política de roles, recibirá un mensaje de error que indica que no ha AWS Clean Rooms podido encontrar la política para el rol de servicio.

11. (Opcional) Para las configuraciones de tablas de asignación de ID avanzadas, modifique las protecciones predeterminadas para la columna que proviene del espacio de nombres de ID.

    La tabla de asignación de ID está configurada de forma predeterminada para permitir solo una `INNER JOIN` en la columna `sourceID` y en la columna `targetID`. Puede modificar esta configuración para que la columna que proviene de este espacio de nombres de ID (`sourceID` o `targetID`) se pueda permitir en cualquier parte de la consulta.

Su objetivo	Opción recomendada
Clasificación de la columna como una “columna de unión” y aceptación solo en una cláusula `INNER JOIN`	Sí
Clasifique la columna como una “columna de dimensiones” y colóquela en cualquier parte de la consulta, incluida una cláusula `JOIN`, instrucciones `SELECT`, `WHERE` y `GROUP BY` de la consulta.	No, permita el uso de cualquier parte de la consulta

12. (Opcional) Si desea habilitar la opción de Etiquetas para el recurso de espacio de nombres de ID, elija Agregar nueva etiqueta y, a continuación, ingrese el par de Clave y Valor.
13. Elija Asociar.
14. En la pestaña Resolución de entidades, en la tabla de espacios de nombres de ID asociados, consulte el espacio de nombres de ID asociado y compruebe que el tipo de espacio de nombres de ID es correcto (origen o destino).

Una vez que todos los miembros de la colaboración hayan asociado sus espacios de nombres de ID, puede crear una tabla de asignación de ID y consultar los datos.

Asociación de un espacio de nombres de ID existente

En este procedimiento, cada miembro asocia su origen de espacio de nombres de ID existente o su destino de espacio de nombres de ID en la colaboración.

Asociación de un espacio de nombres de ID existente

1. Inicia sesión en AWS Management Console y abre la AWS Clean Rooms consola con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Seleccione la colaboración.
4. En la pestaña Resolución de entidades, elija el Espacio de nombres de ID asociado.

5. En la página del espacio de nombres de ID asociado, para los datos de resolución de entidades, elija el origen o el destino del espacio de nombres de ID de AWS Entity Resolution que desee asociar a la colaboración en la lista desplegable.
6. Para obtener los detalles de la asociación, siga estos pasos.

   a. Ingrese un nombre para el espacio de nombres de ID asociado.

      Puede usar el nombre predeterminado o cambiar el nombre de este espacio de nombres de ID.

   b. (Opcional) Ingrese una Descripción del espacio de nombres de ID.

      La descripción ayuda a escribir las consultas.
7. Especifique los permisos de acceso de AWS Clean Rooms mediante la selección de una opción y, a continuación, la realización de la acción recomendada.

Opción	Acción recomendada
Permita AWS Clean Rooms agregar y administrar la política de permisos	AWS Clean Rooms crea un rol de servicio con la política requerida para esta asociación.
Agregación y administración de los permisos manualmente	Realice una de las siguientes acciones: • Revise la política de recursos y agregue los permisos necesarios a la política. • Use una política existente mediante la elección de Agregar declaración de política. Debe tener permisos para modificar roles y crear políticas. **Note** Si no puede modificar la política de roles, recibirá un mensaje de error que indica que no ha AWS Clean

Opción	Acción recomendada
	Rooms podido encontrar la política para el rol de servicio.

8. (Opcional) Para las configuraciones de tablas de asignación de ID avanzadas, modifique las protecciones predeterminadas para la columna que proviene del espacio de nombres de ID.

   La tabla de asignación de ID está configurada de forma predeterminada para permitir solo una `INNER JOIN` en la columna `sourceID` y en la columna `targetID`. Puede modificar esta configuración para que la columna que proviene de este espacio de nombres de ID (`sourceID` o `targetID`) se pueda permitir en cualquier parte de la consulta.

Su objetivo	Opción recomendada
Clasifique la columna como una “columna de unión” y acéptela solo en una cláusula `INNER JOIN`.	Sí
Clasifique la columna como una “columna de dimensiones” y colóquela en cualquier parte de la consulta, incluida una cláusula `JOIN`, instrucciones `SELECT`, `WHERE` y `GROUP BY` de la consulta.	No, permita el uso de cualquier parte de la consulta

9. (Opcional) Si desea habilitar la opción de Etiquetas para el recurso de espacio de nombres de ID, elija Agregar nueva etiqueta y, a continuación, ingrese el par de Clave y Valor.
10. Elija Asociar.
11. En la pestaña Resolución de entidades, en la tabla de espacios de nombres de ID asociados, consulte el espacio de nombres de ID asociado y compruebe que el tipo de espacio de nombres de ID es correcto (origen o destino).

Una vez que todos los miembros de la colaboración hayan asociado sus espacios de nombres de ID, puede crear una tabla de asignación de ID y consultar los datos.

Edición de las asociaciones de espacios de nombres de ID

Como miembro de una colaboración, puede editar las asociaciones de espacios de nombres de ID que ha creado.

Edición de una asociación de espacios de nombres de ID

1. Inicia sesión en AWS Management Console y abre la AWS Clean Rooms consola con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Seleccione la colaboración.
4. Elija la pestaña Resolución de entidades.
5. Para los espacios de nombres de ID asociados, elija un espacio de nombres de ID.
6. En la página de detalles del espacio de nombres de ID, desplácese hacia abajo para ver los Detalles de asociación del espacio de nombres de ID.
7. Elija Editar.
8. En la página Editar asociaciones de espacios de nombres de ID, edite cualquiera de las siguientes opciones:

    a. Para ver los detalles de la asociación, actualice el nombre o la descripción.

    b. (Opcional) Para las configuraciones de tablas de asignación de ID avanzadas, modifique las protecciones predeterminadas para la columna que proviene del espacio de nombres de ID.

    La tabla de asignación de ID está configurada de forma predeterminada para permitir solo una `INNER JOIN` en la columna `sourceID` y en la columna `targetID`. Puede modificar esta configuración para que la columna que proviene de este espacio de nombres de ID (`sourceID` o `targetID`) se pueda permitir en cualquier parte de la consulta.

Su objetivo	Opción recomendada
Clasificación de la columna como una “columna de unión” y aceptación solo en una cláusula `INNER JOIN`	Sí
Clasifique la columna como una “columna de dimensiones” y colóquela en cualquier	No, permita el uso de cualquier parte de la consulta

Su objetivo	Opción recomendada
parte de la consulta, incluida una cláusula `JOIN`, instrucciones `SELECT`, `WHERE` y `GROUP BY` de la consulta.	

9. Elija Guardar cambios.

Desasociación de las asociaciones de espacios de nombres de ID

Como miembro de la colaboración, puede desasociar un espacio de nombres de ID de la colaboración. Esta acción impide que el miembro que puede realizar consultas consulte la tabla.

> **Warning**
>
> Al desasociar una asociación de un espacio de nombres de ID de una colaboración, se eliminan todos los datos de las tablas de asignación de ID derivadas, por lo que no se pueden consultar.
> Por ejemplo, si la asociación del espacio de nombres de ID se utilizó como SOURCE en tres tablas de asignación de ID diferentes, todos los datos de estas tablas de asignación de ID se eliminarán cuando desasocie la asociación de espacio de nombres de ID.

Desasociación de una asociación de espacio de nombres de ID

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con la tuya Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Seleccione la colaboración.
4. Elija la pestaña Resolución de entidades.
5. En Nombres de espacios de ID asociados, seleccione el botón de opción situado junto al espacio de nombres de ID que desea desasociar.
6. Elija Desasociar.
7. En el cuadro de diálogo, confirme su decisión de desconectar el espacio de nombres de ID mediante la elección de Desasociar. Esta acción impide que cualquier miembro que pueda realizar consultas acceda a la tabla de asignación de ID.

Si un miembro de la colaboración elimina uno de los espacios de nombres de ID, no podrá volver a rellenar la tabla de asignación de ID si el origen ha abandonado la colaboración.

Aunque la tabla de asignación de ID ya se había rellenado anteriormente, al desasociar el espacio de nombres de ID, ya no se pueden ejecutar consultas en esa tabla.

Tablas de mapeo de ID en AWS Clean Rooms

Una tabla de mapeo de ID es un recurso AWS Clean Rooms que permite el mapeo de identidades multipartitas en una colaboración.

Antes de crear una tabla de asignación de ID, primero debe tener los datos de origen y destino configurados como espacios de nombres de ID.

Tras crear una tabla de asignación de ID, utilice un flujo de trabajo de asignación de ID para traducir el espacio de nombres de ID de origen al espacio de nombres de ID de destino. Puede hacerlo mediante un método basado en reglas o un método de transcodificación de servicios del proveedor.

Un flujo de trabajo de asignación de ID es un trabajo de procesamiento de datos que asigna datos de un origen de datos de entrada a un destino de datos de entrada en función del método de flujo de trabajo de asignación de ID especificado. Este flujo de trabajo rellena una tabla de asignación de ID.

> **Note**
>
> Las tablas de mapeo de ID solo se pueden crear a partir de conjuntos de datos que se almacenan en Amazon S3 y se rastrean en AWS Glue tablas.

Existen dos métodos de flujo de trabajo de asignación de ID: asignación de ID basado en reglas o asignación de ID de servicios de proveedores:

- Asignación de ID basado en reglas: se utilizan reglas de coincidencia para traducir datos propios de un origen a un destino.
- Mapeo de ID de servicios de proveedores: utiliza el servicio de LiveRamp proveedor para traducir datos de terceros de una fuente a un destino.

> **Note**
>
> El proveedor de servicios de transcodificación compatible actualmente es LiveRamp. Cualquier miembro de la colaboración que tenga una suscripción a LiveRamp Through AWS Data Exchange puede crear la tabla de mapeo de ID. Si ya tienes una suscripción a LiveRamp, pero no a través de ella AWS Data Exchange, ponte en contacto con nosotros LiveRamp para obtener una oferta privada. Para obtener más información, consulte Suscripción a un servicio de proveedor en AWS Data Exchange en la Guía del usuario de AWS Entity Resolution .

Temas

- Creación y llenado de una nueva tabla de asignación de ID
- Llenado de una tabla de asignación de ID existente
- Edición de una tabla de asignación de ID
- Eliminación de una tabla de asignación de ID

Creación y llenado de una nueva tabla de asignación de ID

Antes de crear una tabla de asignación de ID, primero debe tener un origen y destino de espacio de nombres de ID asociado. El origen y el destino del espacio de nombres de ID que asocie a la colaboración deben estar configurados para el tipo de asignación de ID que desee realizar (ya sea una asignación de ID basada en reglas o una asignación de ID de servicios de proveedores).

Después de crear una tabla de asignación de ID, tiene dos opciones. Puede rellenarla inmediatamente, lo que ejecuta el flujo de trabajo de asignación de ID. O bien, puede esperar para rellenar la tabla más adelante.

Después de que la tabla de asignación de ID se haya rellenado correctamente, puede ejecutar una consulta de unión de varias tablas en la tabla de asignación de ID para unir `sourceId` con `targetId` y analizar los datos.

Temas

- Creación de una tabla de asignación de ID (basada en reglas)
- Creación de una tabla de asignación de ID (servicios del proveedor)

Creación de una tabla de asignación de ID (basada en reglas)

En este tema se describe el proceso de creación de una tabla de asignación de ID que utilice reglas coincidentes para traducir datos propios de un origen a un destino.

Creación y llenado de una nueva tabla de asignación de ID mediante el método basado en reglas

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Seleccione la colaboración.
4. En la pestaña Resolución de entidades, elija Creación de una tabla de asignación de ID.
5. En la página Creación de una tabla de asignación de ID en la configuración de asignación de ID, lleve a cabo una de las siguientes acciones en función del objetivo.

Su objetivo	Acción recomendada
Creación de un nuevo flujo de trabajo de asignación de ID	1. Deje seleccionada la casilla de verificación Creación de un nuevo flujo de trabajo de asignación de ID. 2. Continúe con el paso 6.
Reutilización de un flujo de trabajo de asignación de ID existente	1. Desactive la casilla de verificación Creación de un nuevo flujo de trabajo de asignación de ID. 2. Seleccione un flujo de trabajo de asignación de ID basado en reglas de la lista desplegable. 3. Vaya al paso 9.

6. En Datos de identidad, lleve a cabo una de las siguientes acciones en función de la situación

Su escenario	Acción recomendada
Solo hay un origen de espacio de nombres de ID y un destino de espacio de nombres de ID en la colaboración	Consulte las asociaciones de los espacios de nombres de ID de origen y destino.

Su escenario	Acción recomendada
Hay varias asociaciones de espacios de nombres de ID en la colaboración	Seleccione las asociaciones del espacio de nombres de ID de origen y destino que desee utilizar de las listas desplegables.

7. En Método, consulte el método de flujo de trabajo de asignación de ID seleccionado: basado en reglas
8. Para los parámetros de la regla, especifique las configuraciones de los controles de la regla, el tipo de comparación y la coincidencia de registros.

   a. Para los controles de la regla, elija si desea que las reglas coincidentes las proporcione el espacio de nombres del ID de destino o de origen.

      Puede ver las reglas activando Mostrar reglas.

      Los controles de reglas deben ser compatibles entre el espacio de nombres del ID de origen y el de destino para poder usarse en un flujo de trabajo de asignación de ID. Por ejemplo, si un espacio de nombres de ID de origen limita las reglas al destino, pero el espacio de nombres de ID de destino limita las reglas al origen, se produce un error.

   b. El tipo de comparación se establece automáticamente en varios campos de entrada.

      Esto se debe a que ambos participantes habían seleccionado esta opción anteriormente.

   c. Especifique el tipo de registro coincidente eligiendo una de las siguientes opciones.

Su objetivo	Opción recomendada
Limite el tipo de registro coincidente para almacenar solo un registro coincidente en el origen por cada registro coinciden te del destino cuando cree el flujo de trabajo de asignación de ID.	De un origen a un destino
Limite el tipo de registro coincidente para almacenar solo registros coincidentes en el origen por cada registro coinciden	Muchos orígenes para un destino

Su objetivo	Opción recomendada
te del destino cuando cree el flujo de trabajo de asignación de ID.	

> **Note**
>
> Las limitaciones especificadas para los espacios de nombres de ID de origen y destino deben ser compatibles.

9. Para obtener detalles de asignación de ID, lleve a cabo las siguientes acciones.

   a. Ingrese un nombre de la tabla de asignación de ID.

      Puede usar el nombre predeterminado o cambiar el nombre de esta tabla de asignación de ID.

   b. (Opcional) Ingrese una Descripción de la tabla de asignación de ID.

      La descripción ayuda a escribir las consultas.

10. Especifique los permisos de acceso de AWS Clean Rooms mediante la elección de una opción y la realización de la acción recomendada.

Opción	Acción recomendada
Permita AWS Clean Rooms agregar y administrar la política de permisos	AWS Clean Rooms crea un rol de servicio con la política requerida para esta asociació n.
Agregación y administración de los permisos manualmente	Realice una de las siguientes acciones: • Revise la política de recursos y agregue los permisos necesarios a la política. • Use una política existente mediante la elección de Agregar declaración de política.

Opción	Acción recomendada
	Debe tener permisos para modificar roles y crear políticas. **Note** Si no puede modificar la política de roles, recibirá un mensaje de error que indica que no ha AWS Clean Rooms podido encontrar la política para el rol de servicio.

11. Especifique los permisos de acceso de AWS Entity Resolution mediante la elección de una opción y la realización de la acción recomendada:

    Esta sección solo está visible si está creando una nueva tabla de asignación de ID.

Opción	Acción recomendada
Crear y usar un nuevo rol de servicio	AWS Clean Rooms crea un rol de servicio con la política requerida para esta tabla. El nombre del rol de servicio predeterminado es `entityresolution-id-mapping-workflow-<timestamp>` Debe tener permisos para crear roles y adjuntar políticas. Si los datos de entrada están cifrados, puede elegir Estos datos están cifrados con una clave de KMS y, a continuación, ingresar una clave de AWS KMS que se utilizará para descifrar los datos ingresados.

Opción	Acción recomendada
Usar un rol de servicio existente	1. Seleccione un Nombre de rol de servicio existente en la lista desplegable. Si tiene permisos de listas de roles, se mostrará la lista de roles. Si no tiene permisos de listas de roles, puede ingresar el nombre de recurso de Amazon (ARN) del rol que desea usar. 2. Consulte el rol de servicio mediante la elección del enlace externo Ver en IAM. Si no hay ningún rol de servicio existente, la opción Usar un rol de servicio existente no estará disponible. De forma predeterminada, AWS Clean Rooms no intenta actualizar la política de roles existente para añadir los permisos necesarios. 3. (Opcional) Seleccione la casilla Agregar una política preconfigurada con los permisos necesarios para este rol para adjuntar los permisos necesarios al rol. Debe tener permisos para modificar roles y crear políticas.

12. (Opcional) Especifique cualquier configuración adicional mediante la selección de una de las siguientes opciones:

    a. Para la tabla de asignación de ID, lleve a cabo una de las siguientes acciones en función del objetivo.

Su objetivo	Acción recomendada
Habilitación de la configuración de cifrado personalizada para la tabla de asignación de ID	Seleccione Personalizar la configuración de cifrado y, a continuación, introduzca la clave. AWS KMS **Note** Esta clave de KMS debe conceder los permisos necesarios para utilizarla dentro de AWS Entity Resolution una política de claves de KMS. `cleanrooms.amazonaws.com` Para obtener más información sobre los permisos necesarios para trabajar con cifrados con un flujo de trabajo de asignación de ID, consulte Crear un rol de trabajo del flujo de trabajo para AWS Entity Resolution en la Guía del usuario de AWS Entity Resolution .
Habilitación de las etiquetas para el recurso de la tabla de asignación de ID	Elija Agregar etiqueta nueva y, a continuación, ingrese el par clave y valor.

b. Para el flujo de trabajo de asignación de ID, lleve a cabo una de las siguientes acciones en función del objetivo.

Esta sección solo está visible si está creando una nueva tabla de asignación de ID.

Su objetivo	Acción recomendada
Modificación del nombre y la descripción del flujo de trabajo de asignación de ID	Desactive la casilla Mantener el mismo nombre y descripción de la tabla de mapeo de ID e introduzca un nombre y una descripción nuevos para el flujo de trabajo de mapeo de ID.
Habilitación de las etiquetas para el recurso del flujo de trabajo de asignación de ID	Elija Agregar etiqueta nueva y, a continuación, ingrese el par clave y valor.

13. Elija una de las siguientes opciones en función del objetivo.

Su objetivo	Opción recomendada
Creación de una tabla de asignación de ID vacía pero no ejecución del flujo de trabajo de asignación de ID	Creación de una tabla de asignación de ID Puede rellenar la tabla de asignación de ID más adelante siguiendo el proceso Llenado de una tabla de asignación de ID existente.
Creación de una tabla de asignación de ID y ejecución del flujo de trabajo de asignación de ID	Creación y llenado de una tabla de asignación de ID Comienza el proceso de flujo de trabajo de asignación de ID. Durante este proceso, la tabla de mapeo de ID se rellena con la traducción IDs. Es posible que el flujo de trabajo de asignación de ID tarde unas horas en procesarse. Después de que la tabla de asignación de ID se haya rellenado correctamente, puede consultar la tabla de asignación de ID para unir `sourceId` con `targetId` y analizar los datos.

Creación de una tabla de asignación de ID (servicios del proveedor)

En este tema se describe el proceso de creación de una tabla de mapeo de ID que utilice un servicio de proveedor (LiveRamp). Los servicios del LiveRamp proveedor traducen un conjunto de Ramp de origen IDs a otro mediante Ramp mantenido o derivadoIDs.

Creación de una nueva tabla de asignación de ID mediante el método de servicios del proveedor

1. Inicie sesión en la consola AWS Management Console y abra la AWS Clean Rooms consola con su Cuenta de AWS (si aún no lo ha hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Seleccione la colaboración.
4. En la pestaña Resolución de entidades, elija Creación de una tabla de asignación de ID.
5. En la página Creación de una tabla de asignación de ID en la configuración de asignación de ID, lleve a cabo una de las siguientes acciones en función del objetivo.

Su objetivo	Acción recomendada
Creación de un nuevo flujo de trabajo de asignación de ID	1. Deje seleccionada la casilla de verificación Creación de un nuevo flujo de trabajo de asignación de ID. 2. Continúe con el paso 6.
Reutilización de un flujo de trabajo de asignación de ID existente	1. Desactive la casilla de verificación Creación de un nuevo flujo de trabajo de asignación de ID. 2. Seleccione un flujo de trabajo de asignación de ID basado en reglas de la lista desplegable. 3. Vaya al paso 9.

6. En Datos de identidad, lleve a cabo una de las siguientes acciones en función de la situación.

Su escenario	Acción recomendada
Solo hay un origen de espacio de nombres de ID y un destino de espacio de nombres de ID en la colaboración	Consulte las asociaciones de los espacios de nombres de ID de origen y destino

Su escenario	Acción recomendada
Hay varias asociaciones de espacios de nombres de ID en la colaboración	Seleccione las asociaciones del espacio de nombres de ID de origen y destino que desee utilizar de las listas desplegables.

7. En Método, compruebe que el método de flujo de trabajo de mapeo de ID seleccionado esté LiveRamp transcodificando.
8. Para LiveRamp las configuraciones, introduzca la siguiente información proporcionada por: LiveRamp

   - LiveRamp Administrador de ID ARN
   - LiveRamp administrador secreto ARN

   Como alternativa, puede elegir Importar desde un flujo de trabajo existente:
9. Para obtener detalles de asignación de ID, lleve a cabo los siguientes pasos.

   a. Ingrese un nombre de la tabla de asignación de ID.

      Puede usar el nombre predeterminado o cambiar el nombre de esta tabla de asignación de ID.

   b. (Opcional) Ingrese una Descripción de la tabla de asignación de ID.

      La descripción ayuda a escribir las consultas.
10. Especifique los permisos de acceso de AWS Clean Rooms mediante la selección de una de las siguientes opciones:

Opción	Acción recomendada
Permitir AWS Clean Rooms agregar y administrar la política de permisos	AWS Clean Rooms crea un rol de servicio con la política requerida para esta asociació n.
Agregación y administración de los permisos manualmente	Realice una de las siguientes acciones: • Revise la política de recursos y agregue los permisos necesarios a la política.

Opción	Acción recomendada
	• Use una política existente mediante la elección de Agregar declaración de política. Debe tener permisos para modificar roles y crear políticas. **Note** Si no puede modificar la política de roles, recibirá un mensaje de error que indica que no ha AWS Clean Rooms podido encontrar la política para el rol de servicio.

11. Especifique los permisos de acceso de AWS Entity Resolution mediante la selección de una opción y la realización de la acción recomendada.

    Esta sección solo está visible si está creando una nueva tabla de asignación de ID.

Opción	Acción recomendada
Crear y usar un nuevo rol de servicio	AWS Clean Rooms crea un rol de servicio con la política requerida para esta tabla. El nombre del rol de servicio predeterminado es `entityresolution-id-mapping-workflow-<timestamp>` Debe tener permisos para crear roles y adjuntar políticas. Si los datos de entrada están cifrados, puede elegir la opción Estos datos están cifrados con una clave de KMS y, a continuación, ingresar una clave de

Opción	Acción recomendada
	AWS KMS que se utilizará para descifrar los datos ingresados.
Usar un rol de servicio existente	1. Seleccione un Nombre de rol de servicio existente en la lista desplegable. Si tiene permisos de listas de roles, se mostrará la lista de roles. Si no tiene permisos de listas de roles, puede ingresar el nombre de recurso de Amazon (ARN) del rol que desea usar. 2. Consulte el rol de servicio mediante la elección de Ver en IAM. Si no hay ningún rol de servicio existente, la opción Usar un rol de servicio existente no estará disponibl e. De forma predeterminada, AWS Clean Rooms no intenta actualizar la política de roles existente para añadir los permisos necesarios. 3. (Opcional) Seleccione la casilla Agregar una política preconfigurada con los permisos necesarios para este rol para agregar los permisos necesarios al rol. Debe tener permisos para modificar roles y crear políticas.

12. (Opcional) Especifique cualquier configuración adicional mediante la selección de una de las siguientes opciones:

    a. Para la tabla de asignación de ID, lleve a cabo una de las siguientes acciones en función del objetivo.

Su objetivo	Acción recomendada
Habilitación de la configuración de cifrado personalizada para la tabla de asignación de ID	Seleccione Personalizar la configuración de cifrado y, a continuación, introduzca la clave. AWS KMS **Note** Esta clave de KMS debe conceder los permisos necesarios para utilizarla dentro de AWS Entity Resolution una política de claves de KMS. `cleanrooms.amazonaws.com` Para obtener más información sobre los permisos necesarios para trabajar con cifrados con un flujo de trabajo de asignación de ID, consulte Crear un rol de trabajo del flujo de trabajo para AWS Entity Resolution en la Guía del usuario de AWS Entity Resolution .
Habilitación de las etiquetas para el recurso de la tabla de asignación de ID	Elija Agregar etiqueta nueva y, a continuación, ingrese el par clave y valor.

b. Para el flujo de trabajo de asignación de ID, lleve a cabo una de las siguientes acciones en función del objetivo.

Esta sección solo está visible si está creando una nueva tabla de asignación de ID.

Su objetivo	Acción recomendada
Modificación del nombre y la descripción del flujo de trabajo de asignación de ID	Desactive la casilla Mantener el mismo nombre y descripción de la tabla de mapeo de ID e introduzca un nombre y una descripción nuevos para el flujo de trabajo de mapeo de ID.
Habilitación de las etiquetas para el recurso del flujo de trabajo de asignación de ID	Elija Agregar etiqueta nueva y, a continuación, ingrese el par clave y valor.

13. Elija una de las siguientes acciones en función del objetivo.

Su objetivo	Acción recomendada
Creación de una tabla de asignación de ID vacía pero no ejecución del flujo de trabajo de asignación de ID	Elija Crear una tabla de asignación de ID. Puede rellenar la tabla de asignación de ID más adelante siguiendo el proceso Llenado de una tabla de asignación de ID existente.
Creación de una tabla de asignación de ID y ejecución del flujo de trabajo de asignación de ID	Elija Creación y llenado de una tabla de asignación de ID. Comienza el proceso de flujo de trabajo de asignación de ID. Durante este proceso, la tabla de mapeo de ID se rellena con datos IDs transcodificados. Es posible que el flujo de trabajo de asignación de ID tarde unas horas en procesarse. Después de que la tabla de asignación de ID se haya rellenado correctamente, puede consultar la tabla de asignación de ID para unir `sourceId` con `targetId` y analizar los datos.

Llenado de una tabla de asignación de ID existente

Cuando se agreguen nuevos datos a un espacio de nombres de ID, utilice este flujo de trabajo.

Llenado de una tabla de asignación de ID existente

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Seleccione la colaboración.
4. En la pestaña Resolución de entidades, en la sección de tablas de asignación de ID, realice una de las siguientes acciones:
   - Elija una tabla de asignación de ID y, a continuación, elija Llenar.
   - Seleccione el botón de opción situado junto a la tabla de asignación de ID y, en la página de detalles de la tabla de asignación de ID, elija Llenar.

Comienza el proceso de flujo de trabajo de asignación de ID. Durante este proceso, la tabla de mapeo de ID se rellena con datos transcodificados IDs. Es posible que el flujo de trabajo de asignación de ID tarde unas horas en procesarse.

Después de que la tabla de asignación de ID se haya rellenado correctamente, puede consultar la tabla de asignación de ID para unir `sourceId` con `targetId`.

Edición de una tabla de asignación de ID

Como miembro de una colaboración, puede editar la tabla de asignación de ID que ha creado.

Edición de una tabla de asignación de ID

1. Inicia sesión en la consola AWS Management Console y abre la AWS Clean Rooms consola con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Seleccione la colaboración.
4. Elija la pestaña Resolución de entidades.
5. Para las tablas de asignación de ID, elija una tabla.

6. En la página de detalles de la tabla de asignación, desplácese hacia abajo para ver los Detalles de la tabla de asignación de ID.
7. Elija Editar.
8. En la página Editar tablas de asignación de ID, actualice la Descripción o la Información de acceso del servicio.
9. Elija Guardar cambios.

Eliminación de una tabla de asignación de ID

Como miembro de una colaboración, puede eliminar una tabla de asignación de ID que haya creado. Esta acción impide que el miembro que puede realizar consultas consulte la tabla.

> **Warning**
>
> Al eliminar una tabla de asignación, se eliminan permanentemente todos los datos rellenados.

Eliminación de una tabla de asignación de ID

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con la tuya Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Seleccione la colaboración.
4. Elija la pestaña Resolución de entidades.
5. Para las tablas de asignación de ID, elija una tabla.
6. En la página de detalles de la tabla de asignación de ID, desplácese hacia abajo para ver las Tablas de asignación de ID.
7. Elija una tabla de asignación de ID y, a continuación, elija Eliminar.
8. Si tiene la certeza de que desea eliminar la tabla de asignación de ID, elija Eliminar.

Plantillas de análisis en AWS Clean Rooms

Las plantillas de análisis funcionan con Regla de análisis personalizada en AWS Clean Rooms. Con una plantilla de análisis, puede definir parámetros que le ayuden a reutilizar la misma consulta. AWS Clean Rooms admite un subconjunto de parametrización con valores literales.

Las plantillas de análisis son específicas de la colaboración. En cada colaboración, los miembros pueden ver solo las consultas de dicha colaboración. Si tiene previsto utilizar la privacidad diferencial en una colaboración, debe asegurarse de que sus plantillas de análisis sean compatibles con la estructura de consultas de uso general de la privacidad diferencial de AWS Clean Rooms .

Puedes crear una plantilla de análisis de dos maneras: con código SQL o con código Python para Spark.

- Las plantillas de análisis SQL están disponibles en colaboraciones que utilizan tanto el motor de análisis Spark como el motor de análisis AWS Clean Rooms SQL.
- PySpark Las plantillas de análisis están disponibles en las colaboraciones que utilizan el motor de análisis de Spark.

Temas

- plantillas de análisis SQL
- PySpark plantillas de análisis
- Plantillas de PySpark análisis para la solución

plantillas de análisis SQL

Las plantillas de análisis SQL le permiten consultar y analizar datos de diferentes conjuntos de datos en el marco de una colaboración. Puede utilizar estas plantillas para realizar varios tipos de análisis, como identificar las superposiciones de audiencia y calcular las métricas agregadas.

Con las plantillas de análisis de SQL, puede:

- Escribir consultas SQL estándar
- Añada parámetros para hacer que sus consultas sean dinámicas
- Controle el acceso a columnas y tablas específicas
- Establezca los requisitos de agregación para los datos confidenciales

Temas

- Creación de una plantilla de análisis SQL
- Revisión de una plantilla de análisis SQL

Creación de una plantilla de análisis SQL

Requisitos previos

Antes de crear una plantilla de análisis SQL, debe tener:

- Una AWS Clean Rooms colaboración activa
- Acceso a al menos una tabla configurada de la colaboración

  Para obtener información sobre la configuración de tablas en AWS Clean Rooms, consulteCrear una tabla configurada en AWS Clean Rooms.
- Permisos para crear plantillas de análisis
- Conocimientos básicos de la sintaxis de consultas SQL

El siguiente procedimiento describe el proceso de creación de una plantilla de análisis SQL mediante la AWS Clean Rooms consola.

Para obtener información sobre cómo crear una plantilla de análisis SQL mediante la AWS SDKs, consulte la referencia de la AWS Clean Rooms API.

Para crear una plantilla de análisis SQL

1. Inicie sesión en la AWS Clean Rooms consola AWS Management Console y ábrala con la Cuenta de AWS que funcionará como creador de la colaboración.
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Seleccione la colaboración.
4. En la pestaña Plantillas, vaya a la sección Plantillas de análisis creadas por usted.
5. Seleccione Crear plantilla de análisis.
6. En la página Crear plantilla de análisis, en Detalles,

   a. Introduzca un nombre para la plantilla de análisis.

   b. (Opcional) Introduzca una Descripción.

c. En Formato, deje seleccionada la opción SQL.

7. En Tablas, consulte las tablas configuradas asociadas a la colaboración.
8. En Definición,

   a. Introduzca la definición de la plantilla de análisis.

   b. Seleccione Importar desde para importar una definición.

   c. (Opcional) Especifique un parámetro en el editor SQL introduciendo dos puntos (`:`) delante del nombre del parámetro.

      Por ejemplo:

      ```
      WHERE table1.date + :date_period > table1.date
      ```

9. Si ha añadido parámetros anteriormente, en Parámetros: opcional, para cada Nombre de parámetro, elija el Tipo y el Valor predeterminado (opcional).
10. Si desea habilitar la opción de Etiquetas para el recurso de tabla configurada, seleccione Añadir nueva etiqueta y, a continuación, introduzca el par de Clave y Valor.
11. Seleccione Crear.
12. Ahora está listo para informar a su colaborador de que puede revisar una plantilla de análisis. (opcional si desea consultar sus propios datos).

Revisión de una plantilla de análisis SQL

Una vez que un miembro de la colaboración haya creado una SQLanalysis plantilla, puede revisarla y aprobarla. Una vez aprobada la plantilla de análisis, se puede utilizar en una consulta en AWS Clean Rooms.

> **Note**
>
> Cuando incorpore su código de análisis a una colaboración, tenga en cuenta lo siguiente:
>
> - AWS Clean Rooms no valida ni garantiza el comportamiento del código de análisis.
>   - Si necesita garantizar un comportamiento determinado, revise directamente el código de su socio colaborador o póngase en contacto con un auditor externo de confianza para que lo revise.
> - En el modelo de seguridad compartida:

- Usted (el cliente) es responsable de la seguridad del código que se ejecuta en el entorno.
- AWS Clean Rooms es responsable de la seguridad del medio ambiente y garantiza que
  - solo se ejecuta el código aprobado
  - solo se puede acceder a las tablas configuradas especificadas
  - el único destino de salida es el depósito S3 del receptor de resultados.

Para revisar una plantilla de análisis SQL mediante la AWS Clean Rooms consola

1. Inicie sesión en la AWS Clean Rooms consola AWS Management Console y ábrala con la Cuenta de AWS que funcionará como creador de la colaboración.
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Seleccione la colaboración.
4. En la pestaña Plantillas, vaya a la sección Plantillas de análisis creadas por usted.
5. Elija la plantilla de análisis cuyo estado Se puede ejecutar sea No precisa su revisión.
6. Elija Revisar.
7. Revise la Descripción general, la definición y los Parámetros de la regla de análisis (si procede).
8. Revise las tablas configuradas que se enumeran en Tablas a las que se hace referencia en la definición.

   El Estado que aparece junto a cada tabla indicará Plantilla no permitida.
9. Elija una tabla.

En el caso de	Entonces elija...
Aprueba la plantilla de análisis	Permitir plantilla en la tabla. Para confirmar su aprobación, seleccione Permitir.
No aprueba la plantilla de análisis	No permitir

Ahora puede consultar la tabla configurada mediante una plantilla de análisis SQL. Para obtener más información, consulte Ejecución de consultas SQL.

PySpark plantillas de análisis

PySpark las plantillas de análisis requieren un script de usuario de Python y un entorno virtual opcional para utilizar bibliotecas personalizadas y de código abierto. Estos archivos se denominan artefactos.

Antes de crear una plantilla de análisis, primero debe crear los artefactos y, a continuación, almacenarlos en un bucket de Amazon S3. AWS Clean Rooms utiliza estos artefactos al ejecutar trabajos de análisis. AWS Clean Rooms solo accede a los artefactos cuando se ejecuta un trabajo.

Antes de ejecutar cualquier código en una plantilla de PySpark análisis, AWS Clean Rooms valida los artefactos de la siguiente manera:

- Comprobar la versión específica del objeto S3 utilizada al crear la plantilla
- Verificar el hash SHA-256 del artefacto
- Falló cualquier trabajo en el que se hayan modificado o eliminado artefactos

> **Note**
>
> El tamaño máximo de todos los artefactos combinados para una plantilla de PySpark análisis determinada AWS Clean Rooms es de 1 GB.

Seguridad para las plantillas PySpark de análisis

Para preservar un entorno informático seguro, AWS Clean Rooms utiliza una arquitectura informática de dos niveles para aislar el código de usuario de las operaciones del sistema. Esta arquitectura se basa en la tecnología Amazon EMR Serverless Fine Grained Access Control, también conocida como Membrane. Para obtener más información, consulte Membrane: controles de acceso a datos seguros y eficaces en Apache Spark en presencia de código imperativo.

Los componentes del entorno informático se dividen en un espacio de usuario y un espacio de sistema independientes. El espacio de usuario ejecuta el PySpark código de la plantilla de PySpark análisis. AWS Clean Rooms utiliza el espacio del sistema para permitir la ejecución del trabajo, lo que incluye el uso de las funciones de servicio proporcionadas por los clientes para leer los datos a fin de ejecutar el trabajo y la implementación de la lista de columnas permitidas. Como resultado de esta arquitectura, el PySpark código de un cliente que afecta al espacio del sistema, que podría incluir una pequeña cantidad de Spark SQL PySpark DataFrames APIs, queda bloqueado.

PySpark limitaciones en AWS Clean Rooms

Cuando los clientes envían una plantilla de PySpark análisis aprobada, AWS Clean Rooms la ejecutan en su propio entorno informático seguro al que ningún cliente puede acceder. El entorno informático implementa una arquitectura informática con un espacio de usuario y un espacio de sistema para preservar un entorno informático seguro. Para obtener más información, consulte Seguridad para las plantillas PySpark de análisis.

Tenga en cuenta las siguientes limitaciones antes PySpark de usarlo AWS Clean Rooms.

Limitaciones

- Solo se admiten las DataFrame salidas
- Una sola sesión de Spark por ejecución de trabajo

Características no admitidas

- Administración de datos
  - Formatos de tablas Iceberg
  - LakeFormation tablas gestionadas
  - Conjuntos de datos distribuidos resilientes (RDD)
  - Streaming de Spark
  - Control de acceso para columnas anidadas
- Funciones y extensiones personalizadas
  - Funciones de tabla definidas por el usuario () UDTFs
  - Colmena UDFs
  - Clases personalizadas en funciones definidas por el usuario
  - Orígenes de datos personalizados
  - Archivos JAR adicionales para:
    - Extensiones Spark
    - Connectors
    - Configuraciones de Metastore
- Supervisión y análisis
  - Registro de chispas

- Interfaz de usuario de Spark
- Comandos de la `ANALYZE TABLE`

> **Important**
>
> Estas limitaciones se han establecido para mantener el aislamiento de seguridad entre los espacios del usuario y del sistema.
> Todas las restricciones se aplican independientemente de la configuración de colaboración.
> Es posible que las actualizaciones futuras añadan compatibilidad con funciones adicionales en función de las evaluaciones de seguridad.

Prácticas recomendadas

Recomendamos las siguientes prácticas recomendadas a la hora de crear plantillas de PySpark análisis.

- Diseñe sus plantillas de análisis teniendo esto PySpark limitaciones en AWS Clean Rooms en cuenta.
- Pruebe primero su código en un entorno de desarrollo.
- Utilice exclusivamente DataFrame las operaciones compatibles.
- Planifique su estructura de salida para que funcione con DataFrame limitaciones.

Recomendamos las siguientes prácticas recomendadas para gestionar los artefactos

- Guarde todos los artefactos de la plantilla de PySpark análisis en un prefijo o depósito de S3 específico.
- Utilice nombres de versión claros para las diferentes versiones de los artefactos.
- Crea nuevas plantillas de análisis cuando necesites actualizar los artefactos.
- Mantenga un inventario de qué plantillas utilizan qué versiones de artefactos.

Para obtener más información sobre cómo escribir el código de Spark, consulta lo siguiente:

- Ejemplos de Apache Spark
- Escribe una aplicación de Spark en la guía de versiones de Amazon EMR

- Tutorial: Cómo escribir un script AWS Glue para Spark en la guía del AWS Glue usuario

En los temas siguientes se explica cómo crear scripts y bibliotecas de usuario de Python antes de crear y revisar la plantilla de análisis.

Temas

- Crear un script de usuario
- Crear un entorno virtual (opcional)
- Almacenamiento de un script de usuario y un entorno virtual en S3
- Crear una plantilla PySpark de análisis
- Revisión de una plantilla PySpark de análisis

Crear un script de usuario

El script de usuario debe contener una función de punto de entrada (en otras palabras, un controlador). Puede asignar un nombre a su archivo de script de usuario con cualquier nombre de archivo de Python válido.

El siguiente procedimiento describe cómo crear un script de usuario para definir la funcionalidad principal del PySpark análisis.

Requisitos previos

- PySpark 1.0 (corresponde a Python 3.9 y Python 3.11 y Spark 3.5.2)
- Los conjuntos de datos de Amazon S3 solo se pueden leer como asociaciones de tablas configuradas en la sesión de Spark que defina.
- Su código no puede llamar directamente a Amazon S3 y AWS Glue
- Su código no puede realizar llamadas de red

Para crear un script de usuario

1. Abra el editor de texto o el entorno de desarrollo integrado (IDE) de su elección.

   Puede usar cualquier editor de texto o IDE (como Visual Studio Code o Notepad++) que admita archivos de Python. PyCharm

2. Cree un nuevo archivo de Python con el nombre que desee (por ejemplo,**`my_analysis.py`**).
3. Defina una función de punto de entrada que acepte un parámetro de objeto de contexto.

```
def entrypoint(context)
```

El parámetro `context` objeto es un diccionario que proporciona acceso a los componentes esenciales de Spark y a las tablas referenciadas. Contiene el acceso a la sesión de Spark para ejecutar las operaciones de Spark y las tablas a las que se hace referencia:

El acceso a las sesiones de Spark está disponible a través de `context['sparkSession']`

Las tablas de referencia están disponibles en `context['referencedTables']`

4. Defina los resultados de la función de punto de entrada:

```
return results
```

`results`Debe devolver un objeto que contenga un diccionario de resultados con nombres de archivos a una salida. DataFrame

> **Note**
>
> AWS Clean Rooms escribe automáticamente los DataFrame objetos en el depósito S3 del receptor de resultados.

5. Ya puede hacer lo siguiente:

   a. Guarde este script de usuario en S3. Para obtener más información, consulte Almacenamiento de un script de usuario y un entorno virtual en S3.

   b. Cree el entorno virtual opcional para admitir cualquier biblioteca adicional que necesite su script de usuario. Para obtener más información, consulte Crear un entorno virtual (opcional).

Example Ejemplo 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
```

La estructura de carpetas de este ejemplo es la siguiente:

```
analysis_results/
#
### output1/ # Basic selected columns
# ### part-00000.parquet
# ### _SUCCESS
#
### output2/ # Joined data
# ### part-00000.parquet
# ### _SUCCESS
#
### analysis_summary/ # Aggregated results
### part-00000.parquet
### _SUCCESS
```

Example Ejemplo 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
```

Crear un entorno virtual (opcional)

Si su script de usuario requiere bibliotecas adicionales, tiene la opción de crear un entorno virtual para almacenar esas bibliotecas. Si no necesita bibliotecas adicionales, puede omitir este paso.

Cuando trabajes con bibliotecas que tienen extensiones nativas, ten en cuenta que PySpark AWS Clean Rooms funciona en Linux con ARM64 arquitectura.

El siguiente procedimiento muestra cómo crear un entorno virtual mediante un comando CLI básico.

Para crear un entorno virtual

1. Abra una terminal o una línea de comandos.
2. Añada el siguiente contenido:

```
# 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. Ahora está listo para almacenar este entorno virtual en S3. Para obtener más información, consulte Almacenamiento de un script de usuario y un entorno virtual en S3.

Para obtener más información sobre cómo trabajar con Docker y Amazon ECR, consulta la Guía de Amazon ECRUser .

Almacenamiento de un script de usuario y un entorno virtual en S3

El siguiente procedimiento explica cómo almacenar un script de usuario y un entorno virtual opcional en Amazon S3. Complete este paso antes de crear una plantilla de PySpark análisis.

> **Important**
>
> No modifique ni elimine artefactos (scripts de usuario o entornos virtuales) después de crear una plantilla de análisis.
> Si lo hace, hará lo siguiente:
>
> - Provoca que todos los futuros trabajos de análisis que utilicen esta plantilla fallen.
> - Requiere la creación de una nueva plantilla de análisis con nuevos artefactos.

- No afecta a los trabajos de análisis completados anteriormente

Requisitos previos

- Y Cuenta de AWS con los permisos adecuados
- Un archivo de script de usuario (comomy_analysis.py)
- (Opcional, si existe) Un paquete (.tar.gzarchivo) de entorno virtual
- Acceso para crear o modificar funciones de IAM

Console

Para almacenar un script de usuario y un entorno virtual en S3 mediante la consola:

1. Inicie sesión en la consola de Amazon S3 AWS Management Console y ábrala en https://console.aws.amazon.com/s3/.
2. Cree un nuevo bucket de S3 o utilice uno existente.
3. Habilite el control de versiones para el bucket.
    a. Seleccione su depósito.
    b. Seleccione Propiedades.
    c. En la sección Control de versiones de buckets, selecciona Editar.
    d. Seleccione Activar y guardar los cambios.
4. Carga tus artefactos y activa el hash SHA-256.
    a. Navega hasta tu bucket.
    b. Seleccione Cargar.
    c. Selecciona Añadir archivos y añade tu archivo de script de usuario.
    d. (Opcional, si existe) Añada su archivo.tar.gz.
    e. Amplíe Propiedades.
    f. En Sumas de verificación, para la función Checksum, seleccione. SHA256
    g. Seleccione Cargar.
5. Ahora está listo para crear una plantilla de PySpark análisis.

CLI

Para almacenar el script de usuario y el entorno virtual en S3 mediante AWS CLI:

1. Ejecuta el siguiente comando:

```
aws s3 cp --checksum-algorithm sha256 pyspark_venv.tar.gz s3://ARTIFACT-BUCKET/
EXAMPLE-PREFIX/
```

2. Ahora está listo para crear una plantilla PySpark de análisis.

> **Note**
>
> Si necesita actualizar el script o el entorno virtual:
>
> 1. Cargue la nueva versión como un objeto independiente.
> 2. Cree una nueva plantilla de análisis con los nuevos artefactos.
> 3. Elimine la plantilla anterior.
> 4. Guarde los artefactos originales en S3 si la plantilla anterior pudiera seguir siendo necesaria.

Crear una plantilla PySpark de análisis

Requisitos previos

Antes de crear una plantilla de PySpark análisis, debe tener:

- Ser miembro de una AWS Clean Rooms colaboración activa
- Acceso a al menos una tabla configurada en la colaboración activa
- Permisos para crear plantillas de análisis
- Un script de usuario de Python y un entorno virtual creados y almacenados en S3
  - El bucket de S3 tiene activado el control de versiones. Para obtener más información, consulte Uso del control de versiones en buckets de S3
  - El bucket de S3 puede calcular las sumas de comprobación del SHA-256 para los artefactos cargados. Para obtener más información, consulte Uso de sumas de comprobación
- Permisos para leer el código de un bucket de S3

Para obtener información sobre cómo crear el rol de servicio necesario, consulteCree un rol de servicio para leer el código de un bucket de S3 (rol de plantilla de PySpark análisis).

El siguiente procedimiento describe el proceso de creación de una plantilla de PySpark análisis mediante la AWS Clean Rooms consola. Se supone que ya ha creado un script de usuario y archivos de entorno virtual y que ha almacenado el script de usuario y los archivos del entorno virtual en un bucket de Amazon S3.

> **Note**
>
> El miembro que crea la plantilla de PySpark análisis también debe ser el miembro que reciba los resultados.

Para obtener información sobre cómo crear una plantilla de PySpark análisis mediante la AWS SDKs, consulte la referencia de la AWS Clean Rooms API.

Para crear una plantilla PySpark de análisis

1. Inicie sesión en la AWS Clean Rooms consola AWS Management Console y ábrala con la Cuenta de AWS que funcionará como creador de la colaboración.
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Seleccione la colaboración.
4. En la pestaña Plantillas, vaya a la sección Plantillas de análisis creadas por usted.
5. Seleccione Crear plantilla de análisis.
6. En la página Crear plantilla de análisis, en Detalles,
    a. Introduzca un nombre para la plantilla de análisis.
    b. (Opcional) Introduzca una Descripción.
    c. En Formato, elija la PySparkopción.
7. En Definición,
    a. Revise los requisitos previos y asegúrese de que se cumplen todos los requisitos previos antes de continuar.
    b. Para el archivo de puntos de entrada, introduzca el bucket de S3 o seleccione Browse S3.

c. (Opcional) Para el archivo de bibliotecas, introduzca el depósito de S3 o seleccione Browse S3.

8. Para las tablas a las que se hace referencia en la definición,

   - Si todas las tablas a las que se hace referencia en la definición se han asociado a la colaboración:
     - Deje seleccionada la casilla de verificación Todas las tablas a las que se hace referencia en la definición se han asociado a la colaboración.
     - En Tablas asociadas a la colaboración, seleccione todas las tablas asociadas a las que se hace referencia en la definición.
   - Si todas las tablas a las que se hace referencia en la definición no se han asociado a la colaboración:
     - Desactive la casilla de verificación Todas las tablas a las que se hace referencia en la definición se han asociado a la colaboración.
     - En Tablas asociadas a la colaboración, seleccione todas las tablas asociadas a las que se hace referencia en la definición.
     - En Tablas que se asociarán más adelante, introduzca un nombre de tabla.
     - Seleccione Mostrar otra tabla para mostrar otra tabla.

9. Especifique los permisos de acceso al servicio seleccionando un nombre de rol de servicio existente en la lista desplegable.

   1. Si tiene permisos de listas de roles, se mostrará la lista de roles.

      Si no tiene permisos de listas de roles, puede ingresar el nombre de recurso de Amazon (ARN) del rol que desea usar.

   2. Consulte el rol de servicio mediante la elección del enlace externo Ver en IAM.

      Si no hay ningún rol de servicio existente, la opción Usar un rol de servicio existente no estará disponible.

      De forma predeterminada, AWS Clean Rooms no intenta actualizar la política de roles existente para añadir los permisos necesarios.

> **Note**
>
> - AWS Clean Rooms requiere permisos para realizar consultas de acuerdo con las reglas de análisis. Para obtener más información sobre los permisos para AWS Clean Rooms, consulteAWS políticas gestionadas para AWS Clean Rooms.
> - Si el rol no tiene permisos suficientes AWS Clean Rooms, recibirá un mensaje de error que indica que el rol no tiene permisos suficientes AWS Clean Rooms. Debe agregar la política de rol antes de continuar.
> - Si no puede modificar la política de rol, recibirá un mensaje de error que indica que AWS Clean Rooms no ha podido encontrar la política del rol de servicio.

10. Si desea habilitar la opción de Etiquetas para el recurso de tabla configurada, seleccione Añadir nueva etiqueta y, a continuación, introduzca el par de Clave y Valor.
11. Seleccione Crear.
12. Ahora puede informar a su miembro de la colaboración de que puede revisar una plantilla de análisis. (opcional si desea consultar sus propios datos).

> **Important**
>
> No modifique ni elimine artefactos (scripts de usuario o entornos virtuales) después de crear una plantilla de análisis.
> Si lo hace, hará lo siguiente:
>
> - Provoca que todos los futuros trabajos de análisis que utilicen esta plantilla fallen.
> - Requiere la creación de una nueva plantilla de análisis con nuevos artefactos.
> - No afecta a los trabajos de análisis completados anteriormente.

Revisión de una plantilla PySpark de análisis

Cuando otro miembro crea una plantilla de análisis en su colaboración, usted debe revisarla y aprobarla antes de poder utilizarla.

El siguiente procedimiento le muestra cómo revisar una plantilla de PySpark análisis, incluidas sus reglas, parámetros y tablas a las que se hace referencia. Como miembro de la colaboración, evaluará si la plantilla se ajusta a sus acuerdos de intercambio de datos y a sus requisitos de seguridad.

Una vez aprobada la plantilla de análisis, se puede utilizar en un trabajo en AWS Clean Rooms.

> **Note**
>
> Cuando incorpore su código de análisis a una colaboración, tenga en cuenta lo siguiente:
>
> - AWS Clean Rooms no valida ni garantiza el comportamiento del código de análisis.
>   - Si necesita garantizar un comportamiento determinado, revise directamente el código de su socio colaborador o póngase en contacto con un auditor externo de confianza para que lo revise.
> - AWS Clean Rooms garantiza que los hashes SHA-256 del código que figuran en la plantilla de PySpark análisis coincidan con el código que se ejecuta en el PySpark entorno de análisis.
> - AWS Clean Rooms no realiza ninguna auditoría ni análisis de seguridad de las bibliotecas adicionales que incorpore al entorno.
> - En el modelo de seguridad compartida:
>   - Usted (el cliente) es responsable de la seguridad del código que se ejecuta en el entorno.
>   - AWS Clean Rooms es responsable de la seguridad del medio ambiente y garantiza que
>   - solo se ejecuta el código aprobado
>   - solo se puede acceder a las tablas configuradas especificadas
>   - el único destino de salida es el depósito S3 del receptor de resultados.

AWS Clean Rooms genera hashes SHA-256 del script de usuario y del entorno virtual para su revisión. Sin embargo, no se puede acceder directamente al script de usuario ni a las bibliotecas en sí. AWS Clean Rooms

Para validar que el script de usuario y las bibliotecas compartidas son los mismos que los referenciados en la plantilla de análisis, puede crear un hash SHA-256 de los archivos compartidos y compararlo con el hash de la plantilla de análisis creado por. AWS Clean Rooms Los hashes del código ejecutado también estarán en los registros de trabajos.

Requisitos previos

- Sistema operativo Linux/Unix o subsistema Windows para Linux (WSL)
- Archivo de script de usuario que desea aplicar un hash
  - Solicite al creador de la plantilla de análisis que comparta el archivo a través de un canal seguro.
- El hash de la plantilla de análisis creado por AWS Clean Rooms

Para revisar una plantilla de PySpark análisis mediante la AWS Clean Rooms consola

1. Inicie sesión en la AWS Clean Rooms consola AWS Management Console y ábrala con la Cuenta de AWS que funcionará como creador de colaboraciones.
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Seleccione la colaboración.
4. En la pestaña Plantillas, vaya a la sección Plantillas de análisis creadas por usted.
5. Elija la plantilla de análisis cuyo estado Se puede ejecutar sea No precisa su revisión.
6. Elija Revisar.
7. Revise la Descripción general, la definición y los Parámetros de la regla de análisis (si procede).
8. Compruebe que el script de usuario compartido y las bibliotecas sean los mismos a los que se hace referencia en la plantilla de análisis.

   a. Cree un hash SHA-256 de los archivos compartidos y compárelo con el hash de la plantilla de análisis creado por. AWS Clean Rooms

      Para generar un hash, diríjase al directorio que contiene el archivo de script de usuario y, a continuación, ejecute el siguiente comando:

      ```
      sha256sum your_script_filename.py
      ```

      Ejemplo de salida:

      ```
      e3b0c44298fc1c149afbf4c8996fb92427ae41e4649b934ca495991b7852b855 my_analysis.py
      ```

   b. Como alternativa, puede utilizar las funciones de suma de comprobación de Amazon S3. Para obtener más información, consulte Comprobar la integridad de los objetos en Amazon S3 en la Guía del usuario de Amazon S3.

   c. Otra alternativa es ver los hashes del código ejecutado en los registros de trabajos.

9. Revise las tablas configuradas que se enumeran en Tablas a las que se hace referencia en la definición.

   El Estado que aparece junto a cada tabla indicará Plantilla no permitida.

10. Elija una tabla.

    a. Para aprobar la plantilla de análisis, elija Permitir la plantilla en la tabla. Para confirmar su aprobación, seleccione Permitir.

    b. Para rechazar la aprobación, seleccione No permitir.

Si ha optado por aprobar la plantilla de análisis, el miembro que puede ejecutar las tareas ahora puede ejecutar una PySpark tarea en una tabla configurada mediante una plantilla de PySpark análisis. Para obtener más información, consulte Ejecutando PySpark trabajos.

Plantillas de PySpark análisis para la solución

Al ejecutar trabajos con plantillas de PySpark análisis, es posible que se produzcan errores durante la inicialización o la ejecución del trabajo. Estos errores suelen estar relacionados con la configuración del script, los permisos de acceso a los datos o la configuración del entorno.

Para obtener más información sobre PySpark las limitaciones, consultePySpark limitaciones en AWS Clean Rooms.

Temas

- Solución de problemas con el código
- El trabajo de la plantilla de análisis no se inicia
- El trabajo de la plantilla de análisis se inicia pero se produce un error durante el procesamiento
- La configuración del entorno virtual falla

Solución de problemas con el código

AWS Clean Rooms restringe los datos confidenciales de los mensajes de error y los registros para proteger los datos subyacentes del cliente. Para ayudarle a desarrollar su código y solucionar sus problemas, le sugerimos que simule AWS Clean Rooms en su propia cuenta y ejecute los trabajos con sus propios datos de prueba.

Puede simular PySpark AWS Clean Rooms en Amazon EMR Serverless con los siguientes pasos. Tendrá pequeñas diferencias con respecto a las salas limpias de AWS, pero PySpark en su mayoría trata sobre cómo se puede ejecutar el código.

Para simular PySpark AWS Clean Rooms en EMR Serverless

1. Cree un conjunto de datos en Amazon S3, catalóguelo en y configure los permisos de Lake Formation. AWS Glue Data Catalog
2. Registre la ubicación S3 en Lake Formation mediante un rol personalizado.
3. Cree una instancia de Amazon EMR Studio si aún no tiene una (se necesita Amazon EMR Studio para usar Amazon EMR Serverless).
4. Cree una aplicación EMR Serverless
   - Seleccione la versión de lanzamiento emr-7.7.0.
   - Seleccione la arquitectura. ARM64
   - Opte por usar configuraciones personalizadas.
   - Deshabilite la capacidad preinicializada.
   - Si planea realizar un trabajo interactivo, seleccione Punto final interactivo > Habilitar punto final para EMR Studio.
   - Seleccione Configuraciones adicionales > Use Lake Formation para un control de acceso detallado.
   - Cree la aplicación.
5. Utilice EMR-S a través de los portátiles de EMR-Studio o de la API. `StartJobRun`

El trabajo de la plantilla de análisis no se inicia

Causas habituales

Los trabajos de la plantilla de análisis pueden fallar inmediatamente al iniciarse debido a tres problemas de configuración principales:

- El nombre del script es incorrecto y no coincide con el formato requerido
- Falta la función de punto de entrada o tiene un formato incorrecto en el script de usuario
- Versión de Python incompatible en el entorno virtual

Resolución

Para resolverlo:

1. Compruebe su secuencia de comandos de usuario:
   - Compruebe que el script de usuario tenga un nombre de archivo de Python válido.

     Los nombres de archivo de Python válidos utilizan letras minúsculas, guiones bajos para separar las palabras y la extensión .py.
2. Compruebe la función de punto de entrada. Si el script de usuario no tiene una función de punto de entrada, añada una.
   a. Abre tu script de usuario.
   b. Añada esta función de punto de entrada:

   ```
   def entrypoint(context):
       # Your analysis code here
   ```

   c. Asegúrese de que el nombre de la función esté escrito exactamente igual. `entrypoint`
   d. Compruebe que la función acepta el `context` parámetro.
3. Compruebe la compatibilidad de las versiones de Python:
   a. Compruebe que su entorno virtual utiliza Python 3.9.
   b. Para comprobar su versión, ejecute: `python --version`
   c. Si es necesario, actualice su entorno virtual:

   ```
   conda create -n analysis-env python=3.9
   conda activate analysis-env
   ```

Prevención

- Utilice el código de inicio de la plantilla de análisis proporcionada, que incluye la estructura de archivos correcta.
- Configure un entorno virtual dedicado con Python 3.9 para todas las plantillas de análisis.
- Pruebe su plantilla de análisis localmente con la herramienta de validación de plantillas antes de enviar los trabajos.

- Implemente comprobaciones de CI/CD para verificar los requisitos de denominación de los scripts y funciones de punto de entrada.

El trabajo de la plantilla de análisis se inicia pero se produce un error durante el procesamiento

Causas habituales

Los trabajos de análisis pueden fallar durante la ejecución por los siguientes motivos de seguridad y formato:

- Intentos de acceso directo no autorizado a AWS servicios como Amazon S3 o AWS Glue
- Devolver el resultado en formatos incorrectos que no coinciden con DataFrame las especificaciones requeridas
- Se bloquearon las llamadas de red debido a restricciones de seguridad en el entorno de ejecución

Resolución

Para resolver

1. Eliminar el acceso directo al AWS servicio:
    a. Busca tu código para importar y llamar directamente a los AWS servicios.
    b. Sustituya el acceso directo a S3 por los métodos de sesión de Spark proporcionados.
    c. Utilice únicamente tablas preconfiguradas a través de la interfaz de colaboración.
2. Formatee los resultados correctamente:
    a. Verifica que todas las salidas sean Spark DataFrames.
    b. Actualiza tu declaración de devolución para que coincida con este formato:

```
return {
    "results": {
        "output1": dataframe1
    }
}
```

    c. Elimine cualquier objeto que no se DataFrame devuelva.

3. Eliminar llamadas de red:

    a. Identifique y elimine cualquier llamada a la API externa.
    b. Elimine cualquier URL, solicitud o biblioteca de red similar.
    c. Elimine cualquier conexión de socket o código de cliente HTTP.

Prevención

- Utilice el lector de códigos suministrado para comprobar si hay AWS importaciones y llamadas de red no autorizadas.
- Pruebe los trabajos en un entorno de desarrollo en el que las restricciones de seguridad coincidan con las de producción.
- Siga el proceso de validación del esquema de salida antes de implementar los trabajos.
- Revise las directrices de seguridad para conocer los patrones de acceso a los servicios aprobados.

La configuración del entorno virtual falla

Causas habituales

Los errores de configuración del entorno virtual suelen producirse debido a:

- La arquitectura de CPU no coincide entre los entornos de desarrollo y ejecución
- Problemas de formato del código Python que impiden la inicialización adecuada del entorno
- Configuración incorrecta de la imagen base en los ajustes del contenedor

Resolución

Para resolverlo

1. Configure la arquitectura correcta:

    a. Compruebe su arquitectura actual con `uname -m`.
    b. Actualice su Dockerfile para especificar: ARM64

```
FROM --platform=linux/arm64 public.ecr.aws/amazonlinux/amazonlinux:2023-minimal
```

c. Reconstruya su contenedor con `docker build --platform=linux/arm64.`

2. Corrija la hendidura de Python:

   a. Ejecuta un formateador de código Python como `black` en tus archivos de código.

   b. Compruebe el uso coherente de los espacios o las pestañas (no de ambos).

   c. Compruebe la hendidura de todos los bloques de códigos:

```
def my_function():
    if condition:
        do_something()
    return result
```

   d. Utilice un IDE con resaltado de hendiduras de Python.

3. Valide la configuración del entorno:

   a. Ejecute `python -m py_compile your_script.py` para comprobar si hay errores de sintaxis.

   b. Pruebe el entorno localmente antes de la implementación.

   c. Compruebe que todas las dependencias estén enumeradas en`requirements.txt`.

Prevención

- Use Visual Studio Code o PyCharm con complementos de formato de Python
- Configure los ganchos de preconfirmación para ejecutar los formateadores de código automáticamente
- Cree y pruebe entornos de forma local con la imagen base proporcionada ARM64
- Implemente la verificación automática del estilo del código en su proceso de CI/CD

Analice los datos en una colaboración

En AWS Clean Rooms, puede analizar los datos mediante la ejecución de consultas o trabajos.

Una consulta es un método para acceder a las tablas configuradas en una colaboración y analizarlas, mediante un conjunto compatible de funciones, clases y variables. El lenguaje de consulta que se admite actualmente AWS Clean Rooms es SQL. Hay tres formas de ejecutar una consulta AWS Clean Rooms: escribir código SQL, utilizar una plantilla de análisis SQL aprobada o utilizar la interfaz de usuario de Analysis Builder.

Un trabajo es un método para acceder a las tablas configuradas y analizarlas en una colaboración mediante un conjunto compatible de funciones, clases y variables. El tipo de trabajo que se admite actualmente AWS Clean Rooms es PySpark. Hay una forma de ejecutar un trabajo AWS Clean Rooms: mediante una plantilla de PySpark análisis aprobada.

En los temas siguientes se describe cómo analizar los datos AWS Clean Rooms mediante la ejecución de consultas o PySpark trabajos de SQL.

Temas

- Ejecución de consultas SQL
- Ejecutando PySpark trabajos

Ejecución de consultas SQL

> **Note**
>
> Solo puede ejecutar consultas si el miembro responsable de pagar los costos de computación de las consultas se ha unido a la colaboración como miembro activo.

Como miembro que puede realizar consultas, puede ejecutar una consulta SQL de las siguientes maneras:

- Creación manual de una consulta SQL mediante el editor de código SQL.
- Uso de una plantilla de análisis SQL aprobada.
- Uso de la interfaz de usuario de Analysis Builder para crear una consulta sin tener que escribir código SQL.

Cuando el miembro que puede realizar la consulta ejecuta una consulta SQL en las tablas de la colaboración, AWS Clean Rooms asume las funciones pertinentes para acceder a las tablas en su nombre. AWS Clean Rooms aplica las reglas de análisis necesarias a la consulta de entrada y a su salida.

Las reglas de análisis y las restricciones de salida se aplican automáticamente. AWS Clean Rooms solo devuelve los resultados que cumplen con las reglas de análisis definidas.

AWS Clean Rooms admite consultas SQL que pueden ser diferentes a las de otros motores de consultas. Para ver las especificaciones, consulte la Referencia de SQL de AWS Clean Rooms. Si desea ejecutar consultas en tablas de datos protegidas con privacidad diferencial, debe asegurarse de que sus consultas sean compatibles con la estructura de consultas de uso general de la privacidad diferencial de AWS Clean Rooms .

> **Note**
>
> Cuando se utiliza la computación criptográfica para Clean Rooms, no todas las operaciones de SQL generan resultados válidos. Por ejemplo, puede realizar un COUNT en una columna cifrada pero realizando un SUM en números cifrados se producen errores. Además, las consultas también pueden arrojar resultados incorrectos. Por ejemplo, consultas que SUM las columnas selladas producen errores. Sin embargo, un GROUP BY la consulta sobre columnas selladas parece tener éxito, pero produce grupos diferentes a los producidos por un GROUP BY consulta sobre texto sin formato.

Las consultas ejecutadas en la colaboración se cobran al miembro que paga los costos de computación de las consultas.

El miembro que puede realizar la consulta puede seleccionar varios miembros que pueden recibir los resultados para recibir los resultados de una sola consulta. Para obtener más información, consulte Consulta de tablas configuradas con el editor de código SQL. Para obtener información general sobre la recepción de los resultados de las consultas, consulteRecepción y uso de los resultados del análisis.

Requisitos previos

Antes de ejecutar una consulta SQL, asegúrese de disponer de lo siguiente:

- Una membresía activa en una AWS Clean Rooms colaboración
- Acceso a al menos una tabla configurada de la colaboración

- Confirmación de que el miembro responsable de los costes informáticos de la consulta es un miembro activo de la colaboración

Para obtener información sobre cómo consultar datos o ver las consultas llamando directamente a la operación de la AWS Clean Rooms `StartProtectedQuery` API o utilizando la AWS SDKs, consulte la referencia de la AWS Clean Rooms API.

Para obtener más información sobre el registro de consultas, consulte Inicio de sesión de Analysis AWS Clean Rooms.

> **Note**
>
> Si ejecuta una consulta en tablas de datos cifrados, los resultados de las columnas cifradas se cifran.

En los siguientes temas se explica cómo consultar datos en una colaboración mediante la consola de AWS Clean Rooms .

Temas

- Consulta de tablas configuradas con el editor de código SQL
- Consulta de tablas de asignación de ID con el editor de código SQL
- Consulta de tablas configuradas mediante una plantilla de análisis SQL
- Consulta con el creador de análisis
- Visualización del impacto de la privacidad diferencial
- Vista de consultas recientes
- Visualización de detalles de consultas

Consulta de tablas configuradas con el editor de código SQL

Como miembro que puede realizar consultas, puede crear una consulta manualmente escribiendo código SQL en el editor de código SQL. El editor de código SQL se encuentra en la sección Análisis de la pestaña Consultas de la AWS Clean Rooms consola.

El editor de código SQL se muestra de forma predeterminada. Si desea usar el creador de análisis para crear consultas, consulteConsulta con el creador de análisis.

> **Important**
>
> Si empieza a escribir una consulta SQL en el editor de código y luego activa la interfaz de usuario del creador de análisis, la consulta no se guarda.

AWS Clean Rooms admite muchos comandos, funciones y condiciones de SQL. Para obtener más información, consulte la Referencia de SQL de AWS Clean Rooms.

> **Tip**
>
> Si se lleva a cabo un mantenimiento programado mientras se está ejecutando una consulta, esta se termina y se revierte. Deberá reiniciar la consulta.

Consulta de tablas configuradas con el editor de código SQL

1. Inicie sesión en la consola AWS Management Console y abra la AWS Clean Rooms consola con su Cuenta de AWS (si aún no lo ha hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Elija la colaboración cuyo estado de Sus capacidades como miembro sea Consultar.
4. En la pestaña Consultas, vaya a la sección Análisis.

   > **Note**
   >
   > La sección Análisis solo se muestra si el miembro que puede recibir los resultados y el miembro responsable de pagar los costos de computación de las consultas se han unido a la colaboración como miembros activos.

5. En la pestaña Consultas, en Tablas, consulte la lista de tablas y su tipo de regla de análisis asociado (regla de análisis de agregación, regla de análisis de lista o regla de análisis personalizada).

   > **Note**
   >
   > Si no ve las tablas que esperaba en la lista, puede deberse a los siguientes motivos:
   >
   > - Las tablas no se han asociado.

- Las tablas no tienen una regla de análisis configurada.

6. (Opcional) Para ver el esquema de la tabla y los controles de las reglas de análisis, amplíe la tabla seleccionando el icono del signo más (+).
7. Cree la consulta escribiéndola en el editor de código SQL.

   Para obtener más información sobre los comandos y funciones de SQL compatibles, consulta la Referencia de AWS Clean Rooms SQL.

   También puede usar las siguientes opciones para crear la consulta.

   Use an example query

   Para usar una consulta de ejemplo

   1. Elija los tres puntos verticales que aparecen junto al nombre de la tabla.
   2. En Insertar en el editor, elija Consulta de ejemplo.

      > **Note**
      >
      > Al insertar una consulta de ejemplo, se anexa a la consulta que ya está en el editor.

      Aparece el ejemplo de consulta. Todas las tablas que aparecen en Tablas se incluyen en la consulta.
   3. Edite los valores de marcadores de posición de la consulta.

   Insert column names or functions

   Para insertar un nombre de columna o una función

   1. Seleccione los tres puntos verticales situados junto a una columna.
   2. En Insertar en el editor, elija Nombre de columna.
   3. Para insertar manualmente una función permitida en una columna, seleccione los tres puntos verticales situados junto a la columna, seleccione Insertar en el editor y, a continuación, seleccione el nombre de la función permitida (por ejemplo INNER JOIN, SUM, SUM DISTINCT, or COUNT).

4. Pulse Ctrl + Espacio para ver los esquemas de la tabla en el editor de código.

> **Note**
>
> Los miembros que pueden realizar consultas pueden ver y usar las columnas de partición de cada asociación de tablas configuradas. Asegúrese de que la columna de partición esté etiquetada como columna de partición en la AWS Glue tabla subyacente a la tabla configurada.

5. Edite los valores de marcadores de posición de la consulta.

8. (Solo para el motor de análisis de Spark) Especifique el tipo de trabajador compatible y el número de trabajadores.

   Usa la siguiente tabla para determinar el tipo y la cantidad de trabajadores que necesitas para tu caso de uso.

> **Note**
>
> Los diferentes tipos y números de trabajadores tienen costos asociados. Para obtener más información sobre los precios, consulta AWS Clean Rooms los precios.

Tipo de empleado	vCPU	Memoria (GB)	Almacenamiento (GB)	Número de procesos de trabajo	Unidades totales de procesamiento de salas limpias (CRPU)
CR.1X (predeterminado)	4	30	100	2	4
				16 (predeterminado)	32
CR.4X	16	120	400	8	64
				32	256

9. En Enviar los resultados a, especifique quién puede recibir los resultados.

> **Note**
>
> Para recibir los resultados, el miembro de la colaboración debe estar configurado como receptor de resultados y debe ser un participante activo en la colaboración (estado: Activo)

10. (El miembro solo puede realizar consultas) La casilla Usar la configuración de resultados predeterminada está seleccionada de forma predeterminada. Manténgala seleccionada si desea conservar la configuración de resultados predeterminada.

    Si desea especificar una configuración de resultados diferente para esta consulta, desactive la casilla Usar la configuración de resultados predeterminada y, a continuación, seleccione lo siguiente.

    a. Formato de resultado (CSV o PARQUET)
    b. Archivos de resultados (individuales o múltiples)
    c. Destino de los resultados en Amazon S3

    Cada miembro que pueda recibir los resultados puede especificar un formato de resultados, archivos de resultados y un destino de resultados diferentes en Amazon S3.

11. Seleccione Ejecutar.

> **Note**
>
> No podrá ejecutar la consulta si el miembro que puede recibir los resultados no ha configurado los ajustes de resultados de consulta.

12. Consulte los resultados.

    Para obtener más información, consulte Recepción y uso de los resultados del análisis.

13. Siga ajustando los parámetros y vuelva a ejecutar la consulta, o pulse el botón + para iniciar una nueva consulta en una pestaña nueva.

> **Note**
>
> AWS Clean Rooms tiene como objetivo proporcionar mensajes de error claros. Si un mensaje de error no contiene detalles suficientes para ayudarle a solucionar el problema, póngase en contacto con el equipo de cuentas. Proporcióneles una descripción de cómo se produjo el error y el mensaje de error (incluidos los identificadores). Para obtener más información, consulte Solución de problemas AWS Clean Rooms.

Consulta de tablas de asignación de ID con el editor de código SQL

El siguiente procedimiento describe cómo ejecutar una consulta de unión de varias tablas en la tabla de asignación de ID para unir `sourceId` con `targetId`.

Antes de consultar la tabla de asignación de ID, la tabla de asignación de ID debe rellenarse correctamente.

Consulta de tablas de asignación de ID con el editor de código SQL

1. Inicia sesión AWS Management Console y abre la AWS Clean Rooms consola con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Elija la colaboración cuyo estado de Sus capacidades como miembro sea Realizar consultas.
4. En la pestaña Consultas, vaya a la sección Análisis.

   > **Note**
   >
   > La sección Análisis solo se muestra si el miembro que puede recibir los resultados y el miembro responsable de pagar los costos de computación de las consultas se han unido a la colaboración como miembros activos.

5. En la pestaña Consultas, en Tablas, consulta la lista de tablas de mapeo de ID (en Administrado por AWS Clean Rooms) y el tipo de regla de análisis asociado (regla de análisis de tablas de mapeo de ID).

> **Note**
>
> Si no ve las tablas de asignación de ID que espera en la lista, es posible que sea porque las tablas de asignación de ID no se han rellenado correctamente. Para obtener más información, consulte Llenado de una tabla de asignación de ID existente.

6. Cree la consulta escribiéndola en el editor de código SQL.

(Opcional) Si desea utilizar una consulta de ejemplo	(Opcional) Si desea insertar un nombre de la tabla
1. Elija los tres puntos verticales que aparecen junto al nombre de la tabla. 2. En Insertar en el editor, elija Instrucción JOIN de ejemplo. **Note** Al insertar un ejemplo de sentencias JOIN, se añade la consulta que ya está en el editor. Aparece la instrucción JOIN de ejemplo. 3. Edite los valores de marcadores de posición de la consulta.	1. Seleccione los tres puntos verticales situados junto a una columna. 2. En Insertar en el editor, elija Nombre de tabla. 3. Edite los valores de marcadores de posición de la consulta.

7. Elija Ejecutar.

> **Note**
>
> No podrá ejecutar la consulta si el miembro que puede recibir los resultados no ha configurado los ajustes de resultados de consulta.

8. Consulte los resultados.

   Para obtener más información, consulte Recepción y uso de los resultados del análisis.

9. Siga ajustando los parámetros y vuelva a ejecutar la consulta, o pulse el botón + para iniciar una nueva consulta en una pestaña nueva.

> **Note**
>
> AWS Clean Rooms tiene como objetivo proporcionar mensajes de error claros. Si un mensaje de error no contiene detalles suficientes para ayudarle a solucionar el problema, póngase en contacto con el equipo de cuentas. Proporcióneles una descripción de cómo se produjo el error y el mensaje de error (incluidos los identificadores). Para obtener más información, consulte Solución de problemas AWS Clean Rooms.

Consulta de tablas configuradas mediante una plantilla de análisis SQL

Este procedimiento muestra cómo utilizar una plantilla de análisis en la AWS Clean Rooms consola para consultar las tablas configuradas con la regla de análisis personalizada.

Para utilizar una plantilla de análisis SQL para consultar tablas configuradas con la regla de análisis personalizada

1. Inicie sesión en la consola AWS Management Console y abra la AWS Clean Rooms consola con su Cuenta de AWS (si aún no lo ha hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Elija la colaboración cuyo estado de Sus capacidades como miembro sea Realizar consultas.
4. En la pestaña Análisis, en la sección Tablas, consulta las tablas y el tipo de regla de análisis asociado (regla de análisis personalizada).

   > **Note**
   >
   > Si no ve las tablas que esperaba en la lista, puede deberse a los siguientes motivos:
   >
   > - Las tablas no se han asociado.
   > - Las tablas no tienen una regla de análisis configurada.

5. En la sección Análisis, seleccione Ejecutar plantillas de análisis y, a continuación, elija la plantilla de análisis de la lista desplegable.
6. Introduzca el valor de los parámetros de la plantilla de análisis que desea utilizar en la consulta.

El valor debe estar en el tipo de datos especificado para el parámetro.

Puede utilizar valores diferentes cada vez que ejecute la plantilla de análisis.

Vacía o NULL no se admiten los valores del parámetro. Uso de parámetros en LIMIT tampoco se admite la cláusula.

7. Seleccione Ejecutar.

> **Note**
>
> No podrá ejecutar la consulta si el miembro que puede recibir los resultados no ha configurado los ajustes de resultados de consulta.

8. Siga ajustando los parámetros y vuelva a ejecutar la consulta, o pulse el botón + para iniciar una nueva consulta en una pestaña nueva.

Consulta con el creador de análisis

Puede usar el creador de análisis para crear consultas sin tener que escribir código SQL. Con el creador de análisis, puede crear una consulta para una colaboración que tenga:

- Una sola tabla que utilice la regla de análisis de agregación sin necesidad de usar JOIN
- Dos tablas (una de cada miembro) que utilicen la regla de análisis de agregación
- Dos tablas (una de cada miembro) que utilicen la regla de análisis de lista
- Dos tablas (una de cada miembro) que utilicen la regla de análisis de agregación y dos tablas (una de cada miembro) que utilicen la regla de análisis de lista

Si desea escribir consultas SQL manualmente, consulte Consulta de tablas configuradas con el editor de código SQL.

El creador de análisis aparece como la opción Interfaz de usuario del creador de análisis en la sección Análisis de la pestaña Consultas de la consola de AWS Clean Rooms .

Important

Si activa Interfaz de usuario del creador de análisis, comienza a crear una consulta en el creador de análisis y, a continuación, desactiva la Interfaz de usuario del creador de análisis, la consulta no se guarda.

Tip

Si se lleva a cabo un mantenimiento programado mientras se está ejecutando una consulta, esta se termina y se revierte. Deberá reiniciar la consulta.

En los siguientes temas se explica cómo utilizar el creador de análisis.

Temas

- Uso del creador de análisis para consultar una sola tabla (agregación)
- Uso del creador de análisis para consultar dos tablas (agregación o lista)

Uso del creador de análisis para consultar una sola tabla (agregación)

Este procedimiento muestra cómo utilizar la interfaz de usuario de Analysis Builder en la AWS Clean Rooms consola para crear una consulta. La consulta es para una colaboración que tiene una sola tabla que usa la regla de análisis de agregación sin JOIN obligatorio.

Para usar el creador de análisis para consultar una sola tabla

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Elija la colaboración cuyo estado de Sus capacidades como miembro sea Consultar.
4. En la pestaña Consultas, en Tablas, consulte la tabla y su tipo de regla de análisis asociada (el tipo de regla de análisis debe ser Regla de análisis de agregación).

   Note

   Si no ve la tabla que esperaba, puede deberse a los siguientes motivos:

- La tabla no se ha asociado.
- La tabla no tiene una regla de análisis configurada.

5. En la sección Análisis, active la Interfaz de usuario del creador de análisis.
6. Cree una consulta.

   Si desea ver todas las métricas de agregación, vaya al paso 9.

   a. En Elegir métricas, revise las métricas agregadas que se han preseleccionado de forma predeterminada y elimine cualquier métrica si es necesario.

   b. (Opcional) En Añadir segmentos: opcional, elija uno o varios parámetros.

   > **Note**
   >
   > Añadir segmentos: opcional solo se muestra si se especifican las dimensiones para la tabla.

   c. (Opcional) En Añadir filtros: opcional, elija Añadir filtro y, a continuación, elija un Parámetro, un operador y un Valor.

   Para añadir más filtros, elija Añadir otro filtro.

   Para eliminar un filtro, elija Eliminar.

   > **Note**
   >
   > ORDER BY no es compatible con las consultas de agregación.
   > Solo el AND los filtros admiten el operador.

   d. (Opcional) En Añadir descripción: opcional, introduzca una descripción que ayude a identificar la consulta en la lista de consultas.
7. Amplíe Vista previa del código SQL.

   a. Consulte el código SQL que se genera a partir del creador de análisis.

   b. Para copiar el código SQL, elija Copiar.

   c. Para editar el código SQL, elija Editar en el editor de código SQL.
8. Elija Ejecutar.

> **Note**
>
> No podrá ejecutar la consulta si el miembro que puede recibir los resultados no ha configurado los ajustes de resultados de consulta.

9. Siga ajustando los parámetros y vuelva a ejecutar la consulta, o pulse el botón + para iniciar una nueva consulta en una pestaña nueva.

> **Note**
>
> AWS Clean Rooms tiene como objetivo proporcionar mensajes de error claros. Si un mensaje de error no contiene detalles suficientes para ayudarle a solucionar el problema, póngase en contacto con el equipo de cuentas. Proporcióneles una descripción de cómo se produjo el error y el mensaje de error (incluidos los identificadores). Para obtener más información, consulte Solución de problemas AWS Clean Rooms.

Uso del creador de análisis para consultar dos tablas (agregación o lista)

Este procedimiento describe cómo utilizar el generador de análisis de la AWS Clean Rooms consola para crear una consulta para una colaboración que tenga:

- Dos tablas (una de cada miembro) que utilicen la regla de análisis de agregación
- Dos tablas (una de cada miembro) que utilicen la regla de análisis de lista
- Dos tablas (una de cada miembro) que utilicen la regla de análisis de agregación y dos tablas (una de cada miembro) que utilicen la regla de análisis de lista

Para usar el creador de análisis para consultar dos tablas

1. Inicie sesión en la AWS Clean Rooms consola AWS Management Console y ábrala con su Cuenta de AWS (si aún no lo ha hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Elija la colaboración cuyo estado de Sus capacidades como miembro sea Consultar.
4. En la pestaña Consultas, en Tablas, consulte las dos tablas y su tipo de regla de análisis asociada (Regla de análisis de agregación o Regla de análisis de lista).

> **Note**
>
> Si no ve las tablas que espera en la lista, puede deberse a los siguientes motivos:
>
> - Las tablas no se han asociado.
> - Las tablas no tienen una regla de análisis configurada.

5. En la sección Análisis, active la Interfaz de usuario del creador de análisis.
6. Cree una consulta.

Si la colaboración contiene dos tablas que utilizan la Regla de análisis de agregación y dos tablas que utilizan la Regla de análisis de lista, elija primero Agregación o Lista y, a continuación, siga las instrucciones según la regla de análisis seleccionada.

Si las dos tablas utilizan la regla de análisis de agregación	Si las dos tablas utilizan la regla de análisis de lista
1. En Elegir métricas, revise las métricas agregadas que se han preseleccionado de forma predeterminada y elimine cualquier métrica si es necesario. 2. En Registros de coincidencia, elija uno o más registros. **Note** Al utilizar el creador de análisis, solo puede hacer coincidir un único par de columnas. 3. (Opcional) En Añadir segmentos: opcional, elija uno o varios parámetros.	1. En Elegir atributos, revise la lista de atributos que se han preseleccionado de forma predeterminada y elimine cualquier métrica si es necesario. 2. En Registros de coincidencia, elija uno o más registros. **Note** Al utilizar el creador de análisis, solo puede hacer coincidir un único par de columnas. 3. (Opcional) En Añadir filtros: opcional, elija Añadir filtro y, a continuación, elija un parámetro, un operador y un valor. Para añadir más filtros, elija Añadir otro filtro.

Si las dos tablas utilizan la regla de análisis de agregación	Si las dos tablas utilizan la regla de análisis de lista

> **Note**
>
> Añadir segmentos: opcional solo se muestra si se especifican las dimensiones para la tabla.

4. (Opcional) En Añadir filtros: opcional, elija Añadir filtro y, a continuación, elija un parámetro, un operador y un valor.

   Para añadir más filtros, elija Añadir otro filtro.

   Para eliminar un filtro, elija Eliminar.

   > **Note**
   >
   > ORDER BY no es compatible con las consultas de agregación.
   > Solo el AND los filtros admiten el operador.

5. (Opcional) En Añadir descripción: opcional, introduzca una descripción que ayude a identificar la consulta en la lista de consultas recientes.

Para eliminar un filtro, elija Eliminar.

> **Note**
>
> LIMIT no es compatible con las consultas de listas.
> Solo el AND los filtros admiten el operador.

4. (Opcional) En Añadir descripción: opcional, introduzca una descripción que ayude a identificar la consulta en la lista de consultas recientes.

7. Amplíe Vista previa del código SQL.

   a. Consulte el código SQL que se genera a partir del creador de análisis.

   b. Para copiar el código SQL, elija Copiar.

   c. Para editar el código SQL, elija Editar en el editor de código SQL.

8. Elija Ejecutar.

> **Note**
>
> No podrá ejecutar la consulta si el miembro que puede recibir los resultados no ha configurado los ajustes de resultados de consulta.

9. Siga ajustando los parámetros y vuelva a ejecutar la consulta, o pulse el botón + para iniciar una nueva consulta en una pestaña nueva.

> **Note**
>
> AWS Clean Rooms tiene como objetivo proporcionar mensajes de error claros. Si un mensaje de error no contiene detalles suficientes para ayudarle a solucionar el problema, póngase en contacto con el equipo de cuentas. Proporcióneles una descripción de cómo se produjo el error y el mensaje de error (incluidos los identificadores). Para obtener más información, consulte Solución de problemas AWS Clean Rooms.

Visualización del impacto de la privacidad diferencial

En general, la escritura y la ejecución de consultas no cambian cuando se activa la privacidad diferencial. Sin embargo, no puede ejecutar una consulta si no queda suficiente presupuesto de privacidad. A medida que ejecuta consultas y consume el presupuesto de privacidad, puede ver aproximadamente cuántas agregaciones puede ejecutar y cómo podrían afectar a las consultas futuras.

Para ver el impacto de la privacidad diferencial en una colaboración

1. Inicia sesión AWS Management Console y abre la AWS Clean Rooms consola con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Elija la colaboración cuyo estado de Sus detalles de miembro sea Ejecutar consultas.
4. En la pestaña Consultas, en Tablas, vea el presupuesto de privacidad restante. Se muestra como el número estimado de funciones de agregación restantes y la utilidad utilizada (representada como porcentaje).

> **Note**
>
> El número estimado de funciones de agregación restantes y el porcentaje de la utilidad utilizada solo muestran para el miembro que puede realizar la consulta.

5. Seleccione Ver impacto para ver cuánto ruido se inyecta en los resultados y aproximadamente cuántas funciones de agregación puede ejecutar.

Vista de consultas recientes

Puedes ver las consultas que se han realizado en los últimos 90 días en la pestaña Análisis.

> **Note**
>
> Si su única capacidad de miembro es Contribuir datos y no es el miembro que paga los costes de cálculo de las consultas, la pestaña Análisis no aparecerá en la consola.

Para ver consultas recientes

1. Inicie sesión en la consola AWS Management Console y abra la AWS Clean Rooms consola con su Cuenta de AWS (si aún no lo ha hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Elija una colaboración.
4. En la pestaña Análisis, en Análisis, selecciona Todas las consultas en el menú desplegable y consulta las consultas que se han realizado en los últimos 90 días.
5. Para ordenar las consultas recientes por Estado, seleccione un estado en la lista desplegable Todos los estados.

   Los estados son: Enviada, Iniciada, Cancelada, Éxito, Error y Agotado tiempo de espera.

Visualización de detalles de consultas

Puede ver los detalles de la consulta como miembro que puede realizar consultas o como miembro que puede recibir los resultados.

Para ver los detalles de la consulta

1. Inicie sesión en AWS Management Console y abra la AWS Clean Rooms consola con su Cuenta de AWS (si aún no lo ha hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Elija una colaboración.
4. En la pestaña Consultas, realice una de las siguientes acciones:
   - Elija el botón de opción correspondiente a la consulta específica que desea ver y, a continuación, elija Ver detalles.
   - Elija el ID de consulta protegida.
5. En la página de Detalles de la consulta,
   - Si es el miembro que puede realizar consultas, vea los Detalles de la consulta, el Texto SQL y los Resultados.

     Ve un mensaje que confirma que los resultados de la consulta se entregaron al miembro que puede recibir los resultados.
   - Si usted es el miembro que puede recibir los resultados, vea los Detalles de la consulta y los Resultados.

Ejecutando PySpark trabajos

Como miembro que puede realizar consultas, puede ejecutar un PySpark trabajo en una tabla configurada mediante una plantilla de PySpark análisis aprobada.

Requisitos previos

Antes de ejecutar un PySpark trabajo, debe tener:

- Ser miembro activo de AWS Clean Rooms Collaboration
- Acceso a al menos una plantilla de análisis de la colaboración
- Acceso a al menos una tabla configurada de la colaboración
- Permisos para escribir los resultados de un PySpark trabajo en un bucket de S3 específico

  Para obtener información sobre cómo crear el rol de servicio requerido, consulteCree un rol de servicio para escribir los resultados de un PySpark trabajo.

- El miembro responsable de pagar los costes informáticos se ha unido a la colaboración como miembro activo

Para obtener información sobre cómo consultar datos o ver las consultas llamando directamente a la operación de la AWS Clean Rooms `StartProtectedJob` API o utilizando la AWS SDKs, consulte la referencia de la AWS Clean Rooms API.

Para obtener información sobre el registro de trabajos, consulteInicio de sesión de Analysis AWS Clean Rooms.

Para obtener información sobre cómo recibir los resultados de los trabajos, consulteRecepción y uso de los resultados del análisis.

En los temas siguientes se explica cómo ejecutar un PySpark trabajo en una tabla configurada en una colaboración mediante la AWS Clean Rooms consola.

Temas

- Ejecutar un PySpark trabajo en una tabla configurada mediante una plantilla PySpark de análisis
- Ver trabajos recientes
- Consultar detalles del trabajo

Ejecutar un PySpark trabajo en una tabla configurada mediante una plantilla PySpark de análisis

Este procedimiento muestra cómo utilizar una plantilla de PySpark análisis en la AWS Clean Rooms consola para analizar las tablas configuradas con la regla de análisis personalizada.

Para ejecutar un PySpark trabajo en una tabla configurada mediante una plantilla de análisis de Pyspark

1. Inicia sesión AWS Management Console y abre la AWS Clean Rooms consola con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Elige la colaboración que tenga el estado de habilidades de miembro de Ejecutar trabajos.
4. En la pestaña Análisis, en la sección Tablas, consulte las tablas y el tipo de regla de análisis asociado (regla de análisis personalizada).

> **Note**
>
> Si no ve las tablas que esperaba en la lista, puede deberse a los siguientes motivos:
>
> - Las tablas no se han asociado.
> - Las tablas no tienen una regla de análisis configurada.

5. En la sección Análisis, seleccione Ejecutar plantillas de análisis y, a continuación, elija la plantilla de PySpark análisis de la lista desplegable.

   Los parámetros de la plantilla de PySpark análisis se rellenarán automáticamente en la definición.

6. Seleccione Ejecutar.

> **Note**
>
> No puede ejecutar el trabajo si el miembro que puede recibir los resultados no ha configurado los ajustes de los resultados del trabajo.

7. Siga ajustando los parámetros y vuelva a ejecutar el trabajo, o pulse el botón + para iniciar un nuevo trabajo en una pestaña nueva.

Ver trabajos recientes

Puede ver los trabajos ejecutados en los últimos 90 días en la pestaña Análisis.

> **Note**
>
> Si su única capacidad de miembro es Contribuir datos y no es el miembro que paga los costes de cálculo del trabajo, la pestaña Análisis no aparecerá en la consola.

Para ver los trabajos recientes

1. Inicia sesión en la consola AWS Management Console y abre la AWS Clean Rooms consola con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.

3. Elija una colaboración.
4. En la pestaña Análisis, en Análisis, selecciona Todos los trabajos en el menú desplegable y consulta los trabajos que se han ejecutado en los últimos 90 días.
5. Para ordenar los trabajos recientes por estado, seleccione un estado de la lista desplegable Todos los estados.

   Los estados son: Enviada, Iniciada, Cancelada, Éxito, Error y Agotado tiempo de espera.

Consultar detalles del trabajo

Puede ver los detalles del trabajo como el miembro que puede ejecutar los trabajos o como el miembro que puede recibir los resultados.

Para ver los detalles del trabajo

1. Inicie sesión en la consola AWS Management Console y abra la AWS Clean Rooms consola con su Cuenta de AWS (si aún no lo ha hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Elija una colaboración.
4. En la pestaña Análisis, en Análisis, selecciona Todos los trabajos en el menú desplegable y, a continuación, realiza una de las siguientes acciones:

   - Elija el botón de opción para el trabajo específico que desee ver y, a continuación, elija Ver detalles.
   - Elige el ID de trabajo protegido.
5. En la página de detalles del trabajo,

   - Si usted es el miembro que puede ejecutar trabajos, consulte los detalles del trabajo, el trabajo y los resultados.

     Aparece un mensaje que confirma que los resultados del trabajo se entregaron al miembro que puede recibirlos.
   - Si usted es el miembro que puede recibir los resultados, consulte los detalles del trabajo y los resultados.

Recepción y uso de los resultados del análisis

El miembro que puede recibir los resultados revisa los resultados de la consulta en la AWS Clean Rooms consola o en el bucket de Amazon S3 que especificó al unirse a la colaboración.

> **Note**
>
> Solo en el caso de las tablas de datos cifrados, el miembro que puede recibir los resultados descifra los resultados de la consulta ejecutando el cliente de cifrado de C3R en modo de descifrado.
> Si utiliza el motor de análisis Spark, el destino de los resultados en Amazon S3 no puede estar dentro del mismo bucket de S3 que cualquier fuente de datos.

En los siguientes temas se explica cómo recibir los resultados de los análisis mediante la AWS Clean Rooms consola.

Temas

- Recibir resultados de consultas
- Recibir los resultados del trabajo
- Edición de los ajustes predeterminados de los resultados de las consultas
- Edición de los valores predeterminados de la configuración de los resultados del trabajo
- Uso del resultado de la consulta en otros Servicios de AWS

Para obtener información sobre cómo consultar datos o ver las consultas llamando a la AWS Clean Rooms API directamente o utilizando la AWS SDKs, consulte la referencia de la AWS Clean Rooms API.

Para obtener más información sobre el registro de consultas, consulte Inicio de sesión de Analysis AWS Clean Rooms.

> **Note**
>
> Si ejecuta una consulta en tablas de datos cifrados, los resultados de las columnas cifradas se cifran.

Recibir resultados de consultas

> **Note**
>
> Si utiliza el motor de análisis Spark, el destino de los resultados en Amazon S3 no puede estar dentro del mismo bucket de S3 que cualquier fuente de datos.

Los resultados de la consulta se encuentran en la sección de valores predeterminados de la configuración de resultados de la pestaña Análisis de la AWS Clean Rooms consola.

Para recibir resultados de consultas

1. Inicie sesión en la AWS Clean Rooms consola AWS Management Console y ábrala con su Cuenta de AWS (si aún no lo ha hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Elija la colaboración cuyo estado de Sus capacidades como miembro sea Recibir resultados.
4. Para recibir los resultados de la consulta directamente AWS Clean Rooms, en la pestaña Análisis, en Análisis, selecciona Todas las consultas en el menú desplegable y, a continuación, en la columna ID de consulta protegida, selecciona la consulta.
5. En la página Detalles de la consulta, en Resultados, realice una de las siguientes acciones:

Si desea...	Entonces elija...
Copiar los resultados.	Copiar
Descargar los resultados.	Download **Note** De forma predeterminada, el nombre del archivo descargado es el `Query id` correspondiente que se mostró cuando al ejecutar la consulta en AWS Clean Rooms.
Consulte los resultados en Amazon S3.	Ver en Amazon S3

Si desea...	Entonces elija...
	Se abre la consola de Amazon S3 en una pestaña aparte.

6. Si utiliza datos cifrados, ahora puede descifrar las tablas de datos.

   Para obtener más información, consulte Descifrar tablas de datos con el cliente de cifrado de C3R.

Recibir los resultados del trabajo

> **Note**
>
> Si utiliza el motor de análisis Spark, el destino de los resultados en Amazon S3 no puede estar dentro del mismo bucket de S3 que cualquier fuente de datos.

Los resultados del trabajo se encuentran en la sección de valores predeterminados de la configuración de resultados de la pestaña Análisis de la AWS Clean Rooms consola.

Para recibir los resultados del trabajo

1. Inicia sesión AWS Management Console y abre la AWS Clean Rooms consola con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Elija la colaboración cuyo estado de Sus capacidades como miembro sea Recibir resultados.
4. Para recibir los resultados del trabajo directamente AWS Clean Rooms, en la pestaña Análisis, en Análisis, selecciona Todos los trabajos en el menú desplegable y, a continuación, en la columna de ID de trabajo protegido, selecciona el trabajo.
5. En la página de detalles del trabajo, en Resultados, copia el ID del trabajo.

   Vuelva a la pestaña Análisis y amplíe los valores predeterminados de la configuración de resultados.

   En Destino de los resultados, seleccione el enlace para ver los resultados en Amazon S3.

   Se abre la consola de Amazon S3 en una pestaña aparte.

En Amazon S3, pegue el ID del trabajo en la barra de búsqueda y pulse Entrar.

Aparece la carpeta que contiene los resultados. Seleccione la carpeta para ver los resultados del trabajo.

Edición de los ajustes predeterminados de los resultados de las consultas

> **Note**
>
> Si utiliza el motor de análisis Spark, el destino de los resultados en Amazon S3 no puede estar dentro del mismo bucket de S3 que cualquier fuente de datos.

Como miembro que puede recibir resultados, puede editar los valores predeterminados de la configuración de los resultados de las consultas en la AWS Clean Rooms consola.

Para editar los ajustes predeterminados de los resultados de las consultas

1. Inicia sesión en la AWS Clean Rooms consola AWS Management Console y ábrela con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Elija la colaboración cuyo estado de Sus capacidades como miembro sea Recibir resultados.
4. En la pestaña Análisis, en los valores predeterminados de la configuración de resultados, selecciona Editar.
5. En la página Editar los valores predeterminados de la configuración de los resultados, modifique cualquiera de las siguientes opciones según sea necesario:

    a. En Resultados de la consulta, modifique el destino de los resultados en Amazon S3, el formato de los resultados o los archivos de resultados.

    b. (Opcional) Para acceder al servicio, si desea enviar consultas que tardan hasta 24 horas en llegar a su destino de S3, seleccione la casilla Añadir una función de servicio para atender las consultas que tarden hasta 24 horas en completarse.

    Las consultas grandes que tarden hasta 24 horas en completarse se enviarán a su destino de S3.

Si no selecciona la casilla de verificación, solo las consultas que se completen en un plazo de 12 horas se enviarán a su ubicación de S3.

- Especifique los permisos de Acceso a servicios seleccionando Crear y usar un nuevo rol de servicio o Usar un rol de servicio existente.

  Create and use a new service role

  - AWS Clean Rooms crea un rol de servicio con la política requerida para esta tabla.
  - El Nombre del rol de servicio predeterminado es `cleanrooms-query-receiver-<timestamp>`
  - Debe tener permisos para crear roles y adjuntar políticas.

  Use an existing service role

  1. Seleccione un Nombre de rol de servicio existente en la lista desplegable.

     Si tiene permisos de listas de roles, se mostrará la lista de roles.

     Si no tiene permisos de listas de roles, puede ingresar el nombre de recurso de Amazon (ARN) del rol que desea usar.
  2. Consulte el rol de servicio mediante la elección del enlace externo Ver en IAM.

     Si no hay ningún rol de servicio existente, la opción Usar un rol de servicio existente no estará disponible.

     De forma predeterminada, AWS Clean Rooms no intenta actualizar la política de roles existente para añadir los permisos necesarios.

  > **Note**
  >
  > - AWS Clean Rooms requiere permisos para realizar consultas de acuerdo con las reglas de análisis. Para obtener más información sobre los permisos para AWS Clean Rooms, consulteAWS políticas gestionadas para AWS Clean Rooms.

- Si el rol no tiene permisos suficientes AWS Clean Rooms, recibirá un mensaje de error que indica que el rol no tiene permisos suficientes AWS Clean Rooms. Debe agregar la política de rol antes de continuar.
- Si no puede modificar la política de roles, recibirá un mensaje de error que indica que no ha AWS Clean Rooms podido encontrar la política para el rol de servicio.

6. Seleccione Save changes (Guardar cambios).
7. Los Ajustes de los resultados de las consultas aparecen en la página de detalles de la colaboración.

Edición de los valores predeterminados de la configuración de los resultados del trabajo

> **Note**
>
> Si utiliza el motor de análisis Spark, el destino de los resultados en Amazon S3 no puede estar dentro del mismo bucket de S3 que cualquier fuente de datos.

Como miembro que puede recibir resultados, puede editar los valores predeterminados de la configuración de los resultados de los trabajos en la AWS Clean Rooms consola.

Para editar los valores predeterminados de la configuración de los resultados del trabajo

1. Inicie sesión en AWS Management Console y abra la AWS Clean Rooms consola con su Cuenta de AWS (si aún no lo ha hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. Elija la colaboración cuyo estado de Sus capacidades como miembro sea Recibir resultados.
4. En la pestaña Análisis, en los valores predeterminados de la configuración de resultados, selecciona Editar.
5. En la página Editar los valores predeterminados de la configuración de los resultados, modifique cualquiera de las siguientes opciones según sea necesario:

   a. En Job results, modifique el destino de los resultados en Amazon S3.

b. En Acceso al servicio, modifique el nombre del rol de servicio existente.

6. Seleccione Save changes (Guardar cambios).
7. La configuración actualizada de los resultados del trabajo aparece en la página de detalles de la colaboración.

Uso del resultado de la consulta en otros Servicios de AWS

El resultado de una consulta SQL se puede utilizar para los datos iniciales para un modelo de Clean Rooms ML. Para obtener más información, consulte ML de AWS Clean Rooms.

El resultado de la consulta AWS Clean Rooms está disponible en la consola (si la consola se utiliza para ejecutar consultas) y se descarga en un bucket de Amazon S3 específico. A partir de ahí, puede utilizar el resultado de la consulta en otros Servicios de AWS, como Amazon QuickSight y Amazon SageMaker AI, en función del uso que esos servicios hagan de los datos de Amazon S3.

Para obtener más información sobre Amazon QuickSight, consulta la QuickSightdocumentación de Amazon.

Para obtener más información sobre Amazon SageMaker AI, consulte la documentación de Amazon SageMaker AI.

Cree AWS Clean Rooms modelos de aprendizaje automático como proveedor de datos de entrenamiento

Un modelo similar es un modelo de los datos de un proveedor de datos de entrenamiento que permite a un proveedor de datos iniciales crear un segmento similar de los datos del proveedor de datos de entrenamiento que se parezca más a sus datos iniciales. Para crear un modelo similar que se pueda utilizar en una colaboración, debe importar los datos de entrenamiento, crear un modelo similar, configurar ese modelo similar y, después, asociarlo a una colaboración.

Trabajar con modelos similares requiere que dos partes, un proveedor de datos de formación y un proveedor de datos iniciales, trabajen secuencialmente AWS Clean Rooms para integrar sus datos en una colaboración. Este es el flujo de trabajo que el proveedor de datos de entrenamiento debe completar primero:

1. Los datos del proveedor de datos de formación deben almacenarse en una tabla de catálogo de AWS Glue datos con las interacciones entre usuarios y elementos. Como mínimo, los datos de entrenamiento deben contener una columna de ID de usuario, una columna de ID de interacción y una columna de marca temporal.
2. El proveedor de datos de entrenamiento registra los datos de entrenamiento con AWS Clean Rooms.
3. El proveedor de datos de entrenamiento crea un modelo similar que se puede compartir con varios proveedores de datos iniciales. El modelo similar es una red neuronal profunda que puede tardar hasta 24 horas en entrenarse. No se reentrena automáticamente y le recomendamos que vuelva a entrenar el modelo una vez por semana.
4. El proveedor de datos de entrenamiento configura el modelo similar, que incluye si desea compartir las métricas de relevancia y la ubicación en Amazon S3 de los segmentos de salida. El proveedor de datos de entrenamiento puede crear varios modelos similares configurados a partir de un único modelo similar.
5. El proveedor de datos de entrenamiento asocia el modelo de audiencia configurado a una colaboración que se comparte con un proveedor de datos iniciales.

Una vez que el proveedor de datos de entrenamiento haya terminado de crear el modelo de aprendizaje automático, el proveedor de datos iniciales puede crear y exportar el segmento similar.

Temas

- Importación de datos de entrenamiento
- Crear un modelo similar
- Configuración de un modelo similar
- Asociar un modelo similar configurado
- Actualización de un modelo similar configurado

Importación de datos de entrenamiento

> **Note**
>
> Solo puede proporcionar un conjunto de datos de entrenamiento para usarlo en un modelo similar a Clean Rooms ML que tenga datos almacenados en Amazon S3. Sin embargo, puede proporcionar los datos iniciales para un modelo similar mediante SQL que agrupe los datos almacenados en cualquier fuente de datos compatible.

Antes de crear un modelo similar, debe especificar la AWS Glue tabla que contiene los datos de entrenamiento. Clean Rooms ML no almacena una copia de estos datos, solo los metadatos que le permiten acceder a los datos.

Para importar los datos de entrenamiento en AWS Clean Rooms

1. Inicia sesión AWS Management Console y abre la AWS Clean Rooms consola con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, selecciona los modelos AWS ML.
3. En la pestaña Conjuntos de datos de entrenamiento, seleccione Crear conjunto de datos de entrenamiento.
4. En la página Crear conjunto de datos de entrenamiento, en los detalles del conjunto de datos de entrenamiento, ingrese un nombre y, de forma opcional, una descripción.
5. Elija el origen de datos de entrenamiento seleccionando la base de datos y la tabla que desee configurar en las listas desplegables.

   > **Note**
   >
   > Para comprobar que se trata de la tabla correcta, realice una de las siguientes acciones:

- Seleccione Ver en AWS Glue.
- Active Ver esquema para ver el esquema.

6. Para Detalles de entrenamiento, elija la Columna de identificador de usuario, la Columna de identificador de elemento y la Columna de marca temporal de las listas desplegables. Los datos de entrenamiento deben contener estas tres columnas. También puede seleccionar cualquier otra columna que quiera incluir en los datos de entrenamiento.

   Los datos de la Columna de marca temporal deben estar en formato de tiempo Unix en formato de segundos.

7. (Opcional) Si tiene columnas adicionales que entrenar, elija el nombre y el tipo de la columna de las listas desplegables.
8. En Acceso al servicio, debe especificar un rol de servicio que pueda acceder a los datos y proporcionar una clave de KMS si los datos están cifrados. Elija Crear y usar un nuevo rol de servicio y Clean Rooms ML creará automáticamente un rol de servicio y agregará la política de permisos necesaria. Elija Usar un rol de servicio existente e ingréselo en el campo Nombre del rol de servicio si tiene un rol de servicio específico que quiera usar.

   Si los datos están cifrados, ingrese la clave de KMS en el campo AWS KMS key o haga clic en Crear una AWS KMS key para generar una nueva clave de KMS.

9. Si desea habilitar la opción de Etiquetas para el conjunto de datos de entrenamiento, seleccione Añadir nueva etiqueta y, a continuación, introduzca el par de Clave y Valor.
10. Elija Crear el conjunto de datos de entrenamiento.

Para ver la acción de API correspondiente, consulte CreateTrainingDataset.

Crear un modelo similar

Una vez que haya creado un conjunto de datos de entrenamiento, estará listo para crear un modelo similar. Puede crear muchos modelos similares a partir de un único conjunto de datos de entrenamiento.

Debe crear una base de datos predeterminada en su función AWS Glue Data Catalog o incluir el `glue:createDatabase` permiso en la función proporcionada.

Para crear un modelo similar en AWS Clean Rooms

1. Inicia sesión en la consola AWS Management Console y abre la AWS Clean Rooms consola con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, selecciona los modelos AWS ML.
3. En la pestaña Modelos similares, seleccione Crear un modelo similar.
4. En la página Crear modelo similar, para Detalles de modelo similar, ingrese un Nombre y una Descripción opcional.

   a. Seleccione el Conjunto de datos de entrenamiento que desea modelar en la lista desplegable.

   > **Note**
   >
   > Para comprobar que se trata del conjunto de datos de entrenamiento correcto, active Mostrar detalles del conjunto de datos de entrenamiento para ver los detalles. Para crear un nuevo conjunto de datos de entrenamiento, elija Crear conjunto de datos de entrenamiento.

   b. (Opcional) Ingrese una ventana de entrenamiento.
5. Si desea activar la configuración de cifrado personalizada para el modelo similar, elija Personalizar la configuración de cifrado y, a continuación, introduzca la clave de KMS.
6. Si desea habilitar la opción de Etiquetas para el modelo similar, seleccione Añadir nueva etiqueta y, a continuación, introduzca el par de Clave y Valor.
7. Seleccione Crear un modelo similar.

> **Note**
>
> El entrenamiento del modelo puede tardar entre varias horas y 2 días.

Para ver la acción de API correspondiente, consulte CreateAudienceModel.

Configuración de un modelo similar

Una vez que haya creado un modelo similar, estará listo para configurarlo para su uso en una colaboración. Puede crear varios modelos similares configurados a partir de un único modelo similar.

Para configurar un modelo similar en AWS Clean Rooms

1. Inicia sesión AWS Management Console y abre la AWS Clean Rooms consola con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, selecciona los modelos AWS ML.
3. En la pestaña Modelos similares configurados, elija Configurar un modelo similar.
4. En la página Configurar modelo similar, para Detalles de modelo similar configurado, ingrese un Nombre y una Descripción opcional.

    a. Seleccione el Modelo similar que desea configurar en la lista desplegable.

    > **Note**
    >
    > Para comprobar que se trata del modelo similar correcto, active Mostrar detalles del modelo similar para ver los detalles.
    > Para crear un nuevo modelo parecido, elija Crear modelo parecido.

    b. Elija el Tamaño mínimo de coincidencia de la lista inicial que desee. Es el número mínimo de usuarios de los datos del proveedor de datos iniciales que se superponen con los usuarios de los datos de entrenamiento. Este valor debe ser mayor que 0.
5. En Métricas para compartir con otros miembros, elija si desea que el proveedor de datos iniciales de su colaboración reciba métricas del modelo, incluidas las puntuaciones de relevancia.
6. Para la ubicación de destino del segmento similar, introduzca el depósito de Amazon S3 al que se exporta el segmento similar. Este bucket debe estar en la misma región que los demás recursos.
7. En Acceso a los servicios, elija el Nombre del rol de servicio existente que se utilizará para acceder a esta tabla.
8. Para la configuración avanzada del tamaño de los contenedores, especifique el tipo de tamaño de la audiencia como un número absoluto o un porcentaje.
9. Si desea habilitar la opción de Etiquetas para el recurso de tabla configurada, seleccione Añadir nueva etiqueta y, a continuación, introduzca el par de Clave y Valor.
10. Elija Configurar un modelo similar.

Para ver la acción de API correspondiente, consulte CreateConfiguredAudienceModel.

Asociar un modelo similar configurado

Después de haber configurado un modelo similar, puede asociarlo a una colaboración.

Para asociar un modelo similar configurado en AWS Clean Rooms

1. Inicia sesión en AWS Management Console y abre la AWS Clean Rooms consola con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. En la pestaña Con pertenencia activa, elija una colaboración.
4. En la pestaña de modelos ML, en Modelos Ready-to-use similares, selecciona Asociar modelo parecido.
5. En la página Asociar modelo parecido configurado, para ver los detalles de la asociación de modelos similares configurados:

   a. Introduzca un Nombre para el modelo de audiencia configurado asociado.

   b. Introduzca una Descripción de la tabla.

      La descripción ayuda a diferenciarlo de otros modelos de audiencia configurados asociados con nombres similares.
6. En Modelo similar configurado, elija un modelo similar configurado de la lista desplegable.
7. Elija Asociar.

Para ver la acción de API correspondiente, consulte. CreateConfiguredAudienceModelAssociation

Actualización de un modelo similar configurado

Una vez que haya asociado un modelo similar configurado, puede actualizarlo para cambiar información como el nombre, las métricas que desea compartir o la ubicación de salida de Amazon S3.

Para actualizar un modelo similar configurado asociado en AWS Clean Rooms

1. Inicie sesión en AWS Management Console y abra la AWS Clean Rooms consola con su Cuenta de AWS (si aún no lo ha hecho).
2. En el panel de navegación izquierdo, selecciona los modelos AWS ML.

3. En la pestaña Modelos similares configurados, en Ready-to-use Modelos similares, elija un modelo similar configurado y seleccione Editar.
4. En la página de edición, para ver los detalles de la asociación de modelos similares configurados:

    a. Actualice un Nombre y una Descripción opcional.
    b. Elija el Modelo similar que desea configurado de la lista desplegable.
    c. Elija el Tamaño mínimo de coincidencia de la lista inicial que desee. Es el número mínimo de usuarios de los datos del proveedor de datos iniciales que se superponen con los usuarios de los datos de entrenamiento. Este valor debe ser mayor que 0.
5. En Métricas para compartir con otros miembros, elija si desea que el proveedor de datos iniciales de su colaboración reciba métricas del modelo, incluidas las puntuaciones de relevancia.
6. En Ubicación de destino del segmento similar, introduzca el bucket de Amazon S3 al que se exporta el segmento similar. Este bucket debe estar en la misma región que los demás recursos.
7. En Acceso a los servicios, elija el Nombre del rol de servicio existente que se utilizará para acceder a esta tabla.
8. Para la configuración avanzada del tamaño de los contenedores, elija cómo desea configurar los tamaños de los contenedores de audiencia.
9. Elija Guardar cambios.

Para ver la acción de API correspondiente, consulte. UpdateConfiguredAudienceModel

Creación AWS Clean Rooms de modelos de aprendizaje automático como proveedor de datos iniciales

Una vez que el proveedor de datos de entrenamiento haya terminado de crear el modelo de aprendizaje automático, el proveedor de datos iniciales podrá crear y exportar el segmento similar. El segmento similar es un subconjunto de los datos de entrenamiento que más se parece a los datos iniciales.

Este es el flujo de trabajo que debe completar el proveedor de datos iniciales:

1. Los datos del proveedor de datos iniciales se pueden almacenar en un bucket de Amazon S3 o pueden provenir de los resultados de una consulta.
2. El proveedor de datos iniciales comienza la colaboración que comparte con el proveedor de datos de entrenamiento.
3. El proveedor de datos iniciales crea un segmento similar en la pestaña de Clean Rooms ML de la página de colaboración.
4. El proveedor de datos iniciales puede evaluar las métricas de relevancia, en caso de que se hayan compartido, y exportar el segmento similar para su uso fuera de AWS Clean Rooms.

Temas

- Crear un segmento similar
- Exportación de un segmento similar

Crear un segmento similar

> **Note**
>
> Solo puede proporcionar un conjunto de datos de entrenamiento para usarlo en un modelo similar a Clean Rooms ML que tenga datos almacenados en Amazon S3. Sin embargo, puede proporcionar los datos iniciales para un modelo similar mediante SQL que analice los datos almacenados en cualquier fuente de datos compatible.

Un segmento similar es un subconjunto de los datos de entrenamiento que se parece más a los datos iniciales.

Para crear un segmento similar en AWS Clean Rooms

1. Inicia sesión en AWS Management Console y abre la AWS Clean Rooms consola con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. En la pestaña Con pertenencia activa, elija una colaboración.
4. En la pestaña Modelos ML, selecciona Crear segmento similar.
5. En la página Crear segmento similar, en Modelo similar configurado asociado, elija el modelo similar configurado asociado para usarlo en este segmento similar.
6. Para Detalles de segmento similar ingrese un Nombre y una Descripción opcional.
7. Para Perfiles iniciales, elija el Método inicial mediante la selección de una opción y, a continuación, realizando la acción recomendada.

Opción	Acción recomendada
Ruta de Amazon S3	1. Seleccione una ubicación de Amazon S3. 2. (Opcional) Elija Incluir perfiles iniciales en la salida.
Consulta SQL	Escriba una consulta SQL y utilice sus resultados como datos iniciales.
Plantilla de análisis	Elija una plantilla de análisis de la lista desplegable y utilice los resultados creados por una plantilla de análisis.

8. Elija el tipo de trabajador y el número de trabajadores que se utilizarán al crear este canal de datos.
9. En Acceso a los servicios, elija el Nombre del rol de servicio existente que se utilizará para acceder a esta tabla.
10. Si desea habilitar la opción de Etiquetas para el conjunto de datos de entrenamiento, seleccione Añadir nueva etiqueta y, a continuación, introduzca el par de Clave y Valor.

11. Elija Crear segmento similar.

Para ver la acción de API correspondiente, consulte StartAudienceGenerationJob.

Exportación de un segmento similar

Después de crear un segmento similar, puede exportar esos datos a un bucket de Amazon S3.

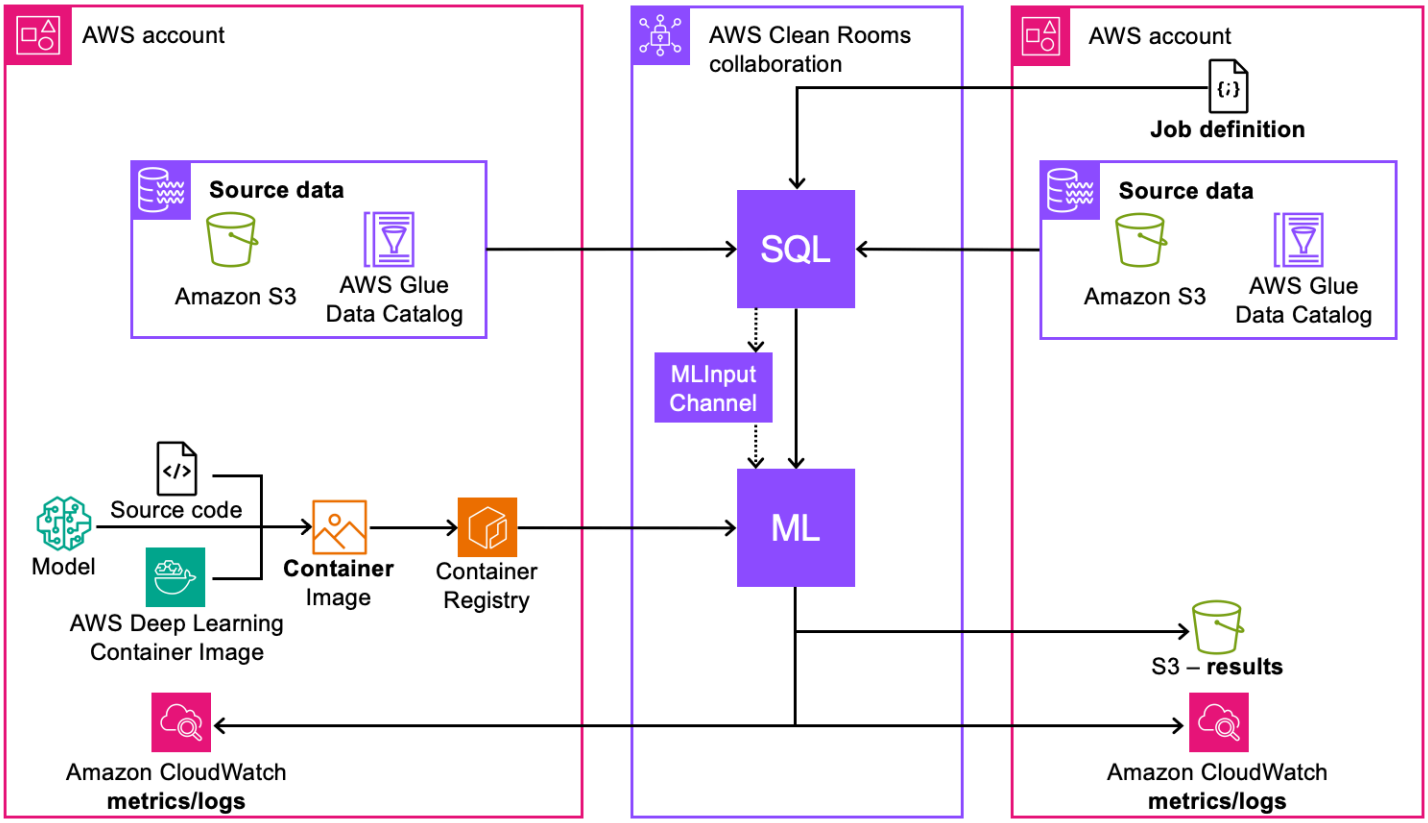
Para exportar un segmento similar a AWS Clean Rooms

1. Inicia sesión en AWS Management Console y abre la AWS Clean Rooms consola con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. En la pestaña Con pertenencia activa, elija una colaboración.
4. En la pestaña Modelos ML, selecciona un segmento similar y selecciona Exportar.
5. En Exportar modelo similar, para Exportar detalles de modelo similar, introduzca un Nombre y una Descripción opcional.
6. En Tamaño del segmento, elija el tamaño que desee para el segmento exportado.
7. Seleccione Exportar.

Para ver la acción de API correspondiente, consulte StartAudienceExportJob.

AWS Clean Rooms Modelado personalizado de ML

Desde un punto de vista técnico, el siguiente diagrama describe cómo funciona el modelado de ML personalizado en AWS Clean Rooms ML.



1. Package sus modelos (de formación o inferencia) en una imagen de contenedor y publíquelos en Amazon ECR.
2. Cree los recursos de AWS Clean Rooms aprendizaje automático y de Clean Rooms necesarios para realizar el entrenamiento de modelos.
3. Asocie el algoritmo modelo a la colaboración.
4. Lea los datos de las cuentas del proveedor de datos para generar el canal de entrada de aprendizaje automático que se utiliza para el entrenamiento o la inferencia.
5. Ejecute el trabajo de formación en aprendizaje automático con la información de los pasos #1 y #4.
6. (Opcional) Exporte los artefactos del modelo entrenado al receptor de resultados.
7. (Opcional) Ejecute el trabajo de inferencia de aprendizaje automático con la información de los pasos #1, #4 y #5.

Antes de empezar, consulte Requisitos previos del modelado ML personalizado y Pautas de creación de modelos para el contenedor de formación para obtener más información.

Temas

- Crear la colaboración
- Contribuir con datos de entrenamiento
- Configuración de un algoritmo modelo
- Asociar el algoritmo del modelo configurado
- Crear un canal de entrada ML
- Crear un modelo entrenado
- Exportación de artefactos del modelo
- Ejecute la inferencia en un modelo entrenado
- Pasos a seguir a continuación

Crear la colaboración

El creador de la colaboración es responsable de crear la colaboración, invitar a los miembros y asignar sus funciones:

Console

1. Cree una colaboración e invite a uno o más miembros a unirse a ella
2. Asigne las siguientes habilidades a los miembros para el análisis mediante consultas:
   - Ejecute consultas: se asignan al miembro que iniciará el entrenamiento del modelo.
   - Reciba los resultados de las consultas, que se asignan a los miembros que recibirán los resultados de las consultas.

   Asigne las siguientes habilidades a los miembros para el modelado de aprendizaje automático mediante flujos de trabajo diseñados específicamente:
   - Reciba los resultados de los modelos entrenados, que se asignarán al miembro que recibirá los resultados del modelo entrenado, incluidos los artefactos y las métricas del modelo.
   - Reciba los resultados de la inferencia del modelo, que se asignan al miembro que recibirá los resultados de la inferencia del modelo.

Si el creador de la colaboración también es el receptor de los resultados, también debe especificar el destino y el formato de los resultados de la consulta durante la creación de la colaboración.

3. Especifique los miembros que pagarán los costes de cálculo de consultas, formación de modelos e inferencia de modelos. Cada uno de estos costos se puede asignar a los mismos miembros o a miembros diferentes. Si un miembro invitado es el miembro responsable de pagar los costes de pago, debe aceptar sus responsabilidades de pago antes de unirse a la colaboración.
4. A continuación, el creador de la colaboración debe configurar la configuración del aprendizaje automático. La configuración de ML proporciona una función para que Clean Rooms ML publique métricas en un Cuenta de AWS. Si el creador de la colaboración también recibe artefactos de modelos entrenados, puede especificar el bucket de Amazon S3 utilizado para recibir los resultados.

   En la sección de configuraciones de ML, especifique el destino de salida del modelo en Amazon S3 y la función de acceso al servicio necesaria para acceder a esta ubicación.

API

1. Cree una colaboración e invite a uno o más miembros a unirse a ella
2. Asigne las siguientes funciones a los miembros de la colaboración:
   - `CAN_QUERY`- asignado al miembro que iniciará el entrenamiento y la inferencia del modelo.
   - `CAN_RECEIVE_MODEL_OUTPUT`- se asigna a los miembros que recibirán los resultados del modelo entrenado.
   - `CAN_RECEIVE_INFERENCE_OUTPUT`- se asigna a los miembros que recibirán los resultados de la inferencia del modelo.

   Si el creador de la colaboración es también el receptor de los resultados, también debe especificar el destino y el formato de los resultados de la consulta durante la creación de la colaboración. También proporcionará un nombre de recurso de Amazon (ARN) del rol de servicio para escribir los resultados en el destino de resultados de las consultas.
3. Especifique los miembros que pagarán los costes de cálculo de consultas, formación de modelos e inferencia de modelos. Cada uno de estos costos se puede asignar a los mismos miembros o a miembros diferentes. Si un miembro invitado es el miembro responsable de

pagar los costes de pago, debe aceptar sus responsabilidades de pago antes de unirse a la colaboración.

4. El siguiente código crea una colaboración, invita a un miembro que puede realizar consultas y recibir resultados, y especifica al creador de la colaboración como receptor de los artefactos modelo.

```
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(f"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. A continuación, el creador de la colaboración debe configurar la configuración del aprendizaje automático. La configuración de ML proporciona una función para que Clean Rooms ML publique métricas y registros en un Cuenta de AWS. Si el creador de la colaboración también recibe resultados (artefactos del modelo o resultados de inferencias), puede especificar el bucket de Amazon S3 utilizado para recibir los resultados.

```
import boto3
acr_ml_client= boto3.client('cleanroomsml')

acr_ml_client.put_ml_configuration(
        membershipIdentifier=membership_id,
        defaultOutputLocation={
           'roleArn':'arn:aws:iam::account:role/roleName',
           'destination':{
                's3Destination':{
                    's3Uri':"s3://bucketName/prefix"
                }
            }
        }
    )
```

Una vez que el creador de la colaboración haya terminado sus tareas, los miembros invitados deberán completar las suyas.

Console

1. Si el miembro invitado es el miembro que puede recibir los resultados, especificará el destino y el formato de los resultados de las consultas. También proporcionan un rol de servicio (ARN) que permite al servicio escribir en el destino de los resultados de la consulta.

   Si el miembro invitado es el miembro responsable de pagar, incluidos los costos de procesamiento de consultas, capacitación e inferencia de modelos, debe aceptar sus responsabilidades de pago antes de unirse a la colaboración.
2. El miembro invitado configura la configuración de ML, que proporciona a Clean Rooms ML la función de publicar las métricas del modelo en un Cuenta de AWS. Si también es el miembro que recibe los artefactos de modelos entrenados, deben proporcionar un depósito de Amazon S3 donde se almacenan los artefactos de modelos entrenados.

API

1. Si el miembro invitado es el miembro que puede recibir los resultados, especificará el destino y el formato de los resultados de las consultas. También proporcionan un rol de servicio (ARN) que permite al servicio escribir en el destino de los resultados de la consulta.

   Si el miembro invitado es el miembro responsable de pagar, incluidos los costos de procesamiento de consultas, capacitación e inferencia de modelos, debe aceptar sus responsabilidades de pago antes de unirse a la colaboración.

   Si el miembro invitado es el responsable de pagar la formación de modelos y la inferencia de modelos para el modelado personalizado, debe aceptar sus responsabilidades de pago antes de unirse a la colaboración.

   ```
   import boto3
   acr_client= boto3.client('cleanrooms')

   acr_client.create_membership(
       membershipIdentifier='membership_id',
       queryLogStatus='ENABLED'
   )
   ```

2. El miembro invitado configura la configuración de ML, que proporciona a Clean Rooms ML la función de publicar las métricas del modelo en un Cuenta de AWS. Si también es el miembro que recibe los artefactos de modelos entrenados, deben proporcionar un depósito de Amazon S3 donde se almacenan los artefactos de modelos entrenados.

```
import boto3
acr_ml_client= boto3.client('cleanroomsml')

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"
                }
            }
        }
    )
```

Contribuir con datos de entrenamiento

Una vez que el creador de la colaboración haya creado la colaboración y se hayan unido los miembros invitados, estará listo para contribuir con los datos de formación a la colaboración. Cualquier miembro puede aportar datos de formación y, para ello, debe seguir estos pasos:

Console

Para aportar datos de formación en AWS Clean Rooms

1. Inicia sesión AWS Management Console y abre la AWS Clean Rooms consola con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Tables (Tablas).
3. En la página Tablas, selecciona Configurar tabla nueva.
4. En Configurar nueva tabla, en Fuente de datos, elija Amazon S3.

   Para Amazon S3, elija una base de datos de la lista desplegable. A continuación, seleccione la tabla de la base de datos.

5. En Columnas permitidas en colaboraciones, seleccione Todas las columnas o Lista personalizada.
6. En los detalles de la tabla configurada, proporcione el nombre y una descripción opcional para esta tabla.
7. Si desea informar sobre las métricas del modelo, introduzca el nombre de las métricas y la instrucción Regex que buscará la métrica en los registros de salida.
8. Seleccione Configurar nuevo cliente.
9. En la página de detalles de la tabla, elija Configurar regla de análisis para configurar una regla de análisis personalizada para esta tabla. Una regla de análisis personalizada limita el acceso a los datos. Puede permitir un conjunto específico de consultas preautorizadas en sus datos o permitir que un conjunto específico de cuentas consulte sus datos.
10. Para el tipo de regla de análisis, elija Personalizado y para Método de creación, elija Flujo guiado.
11. Elija Siguiente.
12. Para Privacidad diferencial, selecciona Desactivar.
13. Elija Siguiente.
14. En el caso de los análisis para consultas directas, elija entre Revisar cada nuevo análisis antes de permitir que se ejecute en esta tabla o Permitir que cualquier consulta creada por colaboradores específicos se ejecute sin revisión en esta tabla.
15. Elija Siguiente.
16. En el campo Columnas no permitidas en la salida, especifique si desea excluir alguna columna de la salida. Si selecciona Ninguna, no se excluirá ninguna columna de la salida. Si elige Lista personalizada, puede especificar determinadas columnas que se eliminarán de la salida.
17. Para los análisis adicionales aplicados a la salida, especifique si desea permitir, denegar o requerir un análisis adicional antes de que se generen los resultados.
18. Elija Siguiente.
19. Revise la información de la página Revisar y configurar y, a continuación, seleccione Configurar regla de análisis.
20. En la página de detalles de la tabla, elija Asociar a la colaboración.
21. En la ventana Asociar tabla, seleccione la colaboración a la que desee asociar esta tabla y elija Elegir colaboración.

22. En la página Asociar la tabla, revise la información de los detalles de la asociación de tablas, el acceso a los servicios y las etiquetas. Si es correcto, selecciona Asociar tabla.
23. En las tablas asociadas a su tabla, seleccione el botón de radio situado junto a la tabla que acaba de asociar. En el menú Acciones, seleccione Configurar en el grupo de reglas de análisis de colaboración.
24. En los análisis adicionales permitidos, elija si algún miembro de la colaboración o miembro específico de la colaboración puede realizar análisis adicionales.

    En la opción Entrega de resultados, elija qué miembros pueden recibir los resultados de las consultas.
25. Elija Configurar regla de análisis.

API

1. Configure una AWS Glue tabla existente para AWS Clean Rooms utilizarla proporcionando la tabla y las columnas que se pueden utilizar.

```
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. Configure una regla de análisis personalizada que limite el acceso a sus datos. Puede permitir un conjunto específico de consultas preautorizadas en sus datos o permitir que un conjunto específico de cuentas consulte sus datos.

```
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"
            }
        }
    }
)
```

En este ejemplo, una cuenta específica puede ejecutar cualquier consulta sobre los datos y es necesario realizar un análisis adicional.

3. Asocie una tabla configurada a la colaboración y proporcione una función de acceso al servicio a las AWS Glue tablas.

```
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**
>
> Este rol de servicio tiene permisos relacionados con las tablas. La función de servicio solo se puede asumir si se ejecutan las consultas permitidas en nombre del miembro que puede realizar la consulta. AWS Clean Rooms Ningún miembro de la colaboración (salvo el propietario de los datos) tiene acceso a las tablas subyacentes de la colaboración. El propietario de los datos puede desactivar la privacidad diferencial para que sus tablas estén disponibles para que otros miembros las consulten.

4. Por último, añada una regla de análisis a la asociación de tablas configurada.

```
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']
            }
        }
    }
)
```

Configuración de un algoritmo modelo

Tras crear un repositorio privado en Amazon ECR, debe configurar el algoritmo del modelo. La configuración de un algoritmo modelo permite que esté disponible para asociarlo a una colaboración.

Console

Para configurar un algoritmo de modelo de aprendizaje automático personalizado en AWS Clean Rooms

1. Inicie sesión en AWS Management Console y abra la AWS Clean Rooms consola con su Cuenta de AWS (si aún no lo ha hecho).
2. En el panel de navegación izquierdo, selecciona modelos de aprendizaje automático personalizados.
3. En la página de modelos ML personalizados, elija Configurar algoritmo de modelo.
4. En Configurar el algoritmo del modelo, en los detalles del algoritmo del modelo, introduzca un nombre y, si lo desea, una descripción.
5. Si desea realizar un entrenamiento de modelos, para ver los detalles del contenedor ECR de imágenes de entrenamiento,

a. Seleccione la casilla de verificación Especificar el URI de la imagen de entrenamiento.

b. Seleccione el repositorio que contiene el modelo de entrenamiento, el contenedor de inferencias o ambos en la lista desplegable.

c. Seleccione la imagen.

d. (Opcional) Introduzca el valor de los puntos de entrada para acceder a la imagen de entrenamiento.

e. (Opcional) Introduzca el valor de los argumentos.

6. Si desea elaborar un informe sobre las métricas del modelo, en el caso de las métricas de entrenamiento, introduzca el nombre de las métricas y la instrucción Regex que buscará la métrica en los registros de salida.
7. Si desea realizar una inferencia de modelos, para ver los detalles del contenedor ECR de la imagen de inferencia,

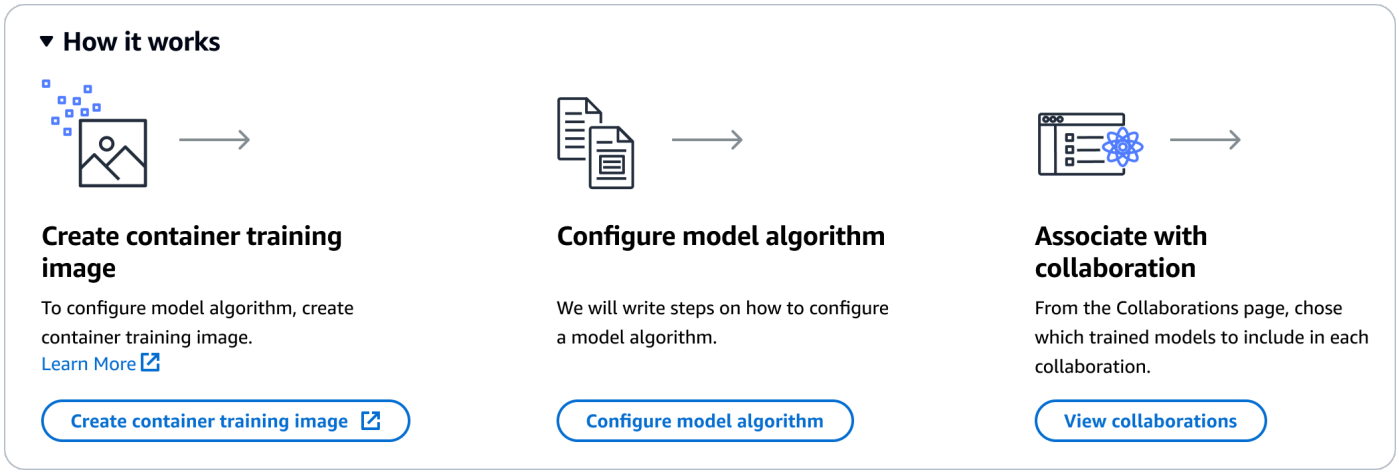
a. Seleccione la casilla de verificación Especificar el URI de la imagen de inferencia.

b. Seleccione el repositorio en la lista desplegable.

c. Seleccione la imagen.

8. En Acceso a los servicios, elija el Nombre del rol de servicio existente que se utilizará para acceder a esta tabla.
9. Para el cifrado, elija Personalizar la configuración de cifrado para especificar su propia clave de KMS y la información relacionada. De lo contrario, Clean Rooms ML gestionará el cifrado
10. Si desea activar las etiquetas, seleccione Añadir nueva etiqueta y, a continuación, introduzca el par clave y valor.
11. Elija Configurar el algoritmo del modelo.

API



1. Crea una imagen de docker compatible con SageMaker IA. Clean Rooms ML solo admite imágenes de docker compatibles con SageMaker IA.
2. Después de crear una imagen de docker compatible con SageMaker IA, utilice Amazon ECR para crear una imagen de entrenamiento. Siga las instrucciones de la Guía del usuario de Amazon Elastic Container Registry para crear una imagen de formación sobre contenedores.
3. Configure el algoritmo del modelo para usarlo en Clean Rooms ML. Debe proporcionar la siguiente información:
   - El enlace al repositorio de Amazon ECR y argumentos adicionales para entrenar el modelo y ejecutar la inferencia. Clean Rooms ML permite ejecutar trabajos de transformación por lotes en un contenedor de inferencias.
   - Un rol de acceso al servicio que permite a Clean Rooms ML acceder al repositorio.
   - (Opcional) Un contenedor de inferencias. Aunque puede proporcionarlo en un algoritmo modelo configurado por separado, le recomendamos que lo haga en este paso para que tanto el contenedor de entrenamiento como el de inferencia se administren como parte del mismo recurso.

```
import boto3
acr_ml_client= boto3.client('cleanroomsml')

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'
)
```

Asociar el algoritmo del modelo configurado

Una vez que haya configurado el algoritmo del modelo, estará listo para asociar el algoritmo del modelo a una colaboración. Al asociar un algoritmo modelo, el algoritmo modelo estará disponible para todos los miembros de la colaboración.

Console

Para asociar un algoritmo de modelo ML personalizado en AWS Clean Rooms

1. Inicie sesión en AWS Management Console y abra la AWS Clean Rooms consola con su Cuenta de AWS (si aún no lo ha hecho).
2. En el panel de navegación izquierdo, selecciona los modelos de aprendizaje automático personalizados.
3. En la página de modelos de aprendizaje automático personalizados, elija el algoritmo de modelo configurado que desee asociar a una colaboración y haga clic en Asociar a la colaboración.
4. En la ventana Asociar el algoritmo del modelo configurado, elija la colaboración a la que desee asociarse.
5. Seleccione Elegir colaboración.

API

Asocie el algoritmo del modelo configurado a la colaboración. También proporciona una política de privacidad que define quién tiene acceso a los diferentes registros, permite a los clientes definir las expresiones regulares y cuántos datos se pueden exportar de los resultados del modelo de entrenamiento o de las inferencias.

> **Note**
>
> Las asociaciones de algoritmos de modelos configurados son inmutables.

```
import boto3
acr_ml_client= boto3.client('cleanroomsml')

acr_ml_client.create_configured_model_algorithm_association(
    name='configured_model_algorithm_association_name',
    description='purpose of the association',
    membershipIdentifier='membership_id',
    configuredModelAlgorithmArn= 'arn:aws:cleanrooms-ml:region:account:membership/
membershipIdentifier/configured-model-algorithm/identifier',
    privacyConfiguration = {
        "policies": {
            "trainedModels": {
                "containerLogs": [
                    {
                        "allowedAccountIds": ['member_account_id'],

                    },
                    {
                        "allowedAccountIds": ['member_account_id'],
                        "filterPattern": "INFO"
                    }
                ],
                "containerMetrics": {
                    "noiseLevel": 'noise value'
                }
            },
            "trainedModelInferenceJobs": {
                "containerLogs": [
                    {
                        "allowedAccountIds": ['member_account_id']
                    }
                ]
            },
            trainedModelExports: {
                maxSize: {
                    unit: GB,
                    value: 5
```

```
                },
                filesToExport: [
                    "MODEL",    // final model artifacts that container should write
 to /opt/ml/model directory
                    "OUTPUT"   // other artifacts that container should write to /
opt/ml/output/data directory
                ]
            }
        }
    }
)
```

Una vez que el algoritmo modelo configurado esté asociado a la colaboración, los proveedores de datos de formación deben añadir una regla de análisis de colaboración a su tabla. Esta regla permite que la asociación de algoritmos del modelo configurado acceda a su tabla configurada. Todos los proveedores de datos de formación que contribuyan deben ejecutar el siguiente código:

```
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**
>
> Como las asociaciones de algoritmos de modelos configurados son inmutables, recomendamos que los proveedores de datos de formación que deseen permitir su uso en listas de modelos utilicen caracteres comodín `allowedAdditionalAnalyses` durante las primeras iteraciones de la configuración del modelo personalizado. Esto

permite a los proveedores de modelos iterar su código sin necesidad de que otros proveedores de formación vuelvan a asociarlo antes de entrenar su código de modelo actualizado con datos.

Crear un canal de entrada ML

Un canal de entrada ML es un flujo de datos que se crea a partir de una consulta de datos específica. Los miembros con la capacidad de consultar datos pueden preparar sus datos para el entrenamiento y la inferencia mediante la creación de un canal de entrada de aprendizaje automático. La creación de un canal de entrada de aprendizaje automático permite que los datos se utilicen en diferentes modelos de entrenamiento dentro de la misma colaboración. Debe crear canales de entrada de aprendizaje automático independientes para el entrenamiento y la inferencia.

Para crear un canal de entrada de ML, debe especificar la consulta SQL que se utiliza para consultar los datos de entrada y crear el canal de entrada de ML. Los resultados de esta consulta nunca se comparten con ningún miembro y permanecen dentro de los límites de Clean Rooms ML. El nombre de recurso de Amazon (ARN) de referencia se utiliza en los siguientes pasos para entrenar un modelo o ejecutar una inferencia.

Console

Para crear un canal de entrada de aprendizaje automático en AWS Clean Rooms

1. Inicia sesión en AWS Management Console y abre la AWS Clean Rooms consola con tu Cuenta de AWS (si aún no lo has hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. En la página de colaboraciones, elige la colaboración en la que quieres crear un canal de entrada de aprendizaje automático.
4. Cuando se abra la colaboración, selecciona la pestaña de modelos de aprendizaje automático y, a continuación, selecciona Crear canal de entrada de aprendizaje automático.
5. En Crear canal de entrada ML, en los detalles del canal de entrada ML, introduzca un nombre, una descripción opcional y el algoritmo del modelo asociado que desee utilizar.
6. En Dataset, elija Plantilla de análisis para usar los resultados de una plantilla de análisis como conjunto de datos de entrenamiento o consulta SQL para usar los resultados de una consulta SQL como conjunto de datos de entrenamiento. Si eligió la plantilla de análisis,

especifique la plantilla de análisis que desee. Si eligió una consulta SQL, introduzca la consulta en el campo de consulta SQL.

7. Elija el tipo de trabajador y el número de trabajadores que se utilizarán al crear este canal de datos.
8. En Retención de datos en días, especifique cuánto tiempo se conservarán los datos.
9. Para acceder al servicio, elija el nombre del rol de servicio existente que se usará para acceder a esta tabla o elija Crear y usar un nuevo rol de servicio.
10. Para el cifrado, elija Personalizar la configuración de cifrado para especificar su propia clave de KMS y la información relacionada. De lo contrario, Clean Rooms ML gestionará el cifrado.
11. Elija Crear canal de entrada ML.

API

Para crear un canal de entrada ML, ejecute el siguiente código:

```
import boto3
acr_client = boto3.client('cleanroomsml')

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"
                }
            }
        },
        "roleArn": "arn:aws:iam::111122223333:role/ezcrc-ctm-role"
    }
)
channel_arn = resp['ML Input Channel ARN']
```

Crear un modelo entrenado

Una vez que haya asociado el algoritmo del modelo configurado a una colaboración y, a continuación, haya creado y configurado un canal de entrada de aprendizaje automático, estará listo para crear un modelo entrenado. Los miembros de una colaboración utilizan un modelo entrenado para analizar sus datos de forma conjunta.

Console

Para crear un modelo entrenado en AWS Clean Rooms

1. Inicie sesión en la consola AWS Management Console y abra la AWS Clean Rooms consola con su Cuenta de AWS (si aún no lo ha hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. En la página Colaboraciones, elige la colaboración en la que quieres crear un modelo entrenado.
4. Cuando se abra la colaboración, selecciona la pestaña de modelos de aprendizaje automático y, a continuación, selecciona Crear modelo entrenado.
5. En Crear un modelo entrenado, en los detalles del modelo personalizado entrenado, introduzca un nombre y, si lo desea, una descripción.
6. En el conjunto de datos de entrenamiento, elija el canal de entrada ML para este modelo entrenado.
7. En el caso de los hiperparámetros, especifique los parámetros específicos del algoritmo y sus valores previstos. Los hiperparámetros son específicos del modelo que se está entrenando y se utilizan para ajustar el entrenamiento del modelo.
8. En el caso de las variables de entorno, especifique cualquier variable específica del algoritmo y sus valores previstos. Las variables de entorno se configuran en el contenedor de Docker.
9. Para acceder al servicio, elija el nombre del rol de servicio existente que se usará para acceder a esta tabla o elija Crear y usar un nuevo rol de servicio.
10. En la configuración de EC2 recursos, especifique la información sobre los recursos informáticos que se utilizan para el entrenamiento del modelo. Debe especificar el tipo de instancia y el tamaño del volumen que se utilizan.
11. Elija Crear modelo entrenado.

API

El miembro con la capacidad de entrenar un modelo comienza a entrenar seleccionando el canal de entrada ML y el algoritmo del modelo:

```
import boto3
acr_ml_client= boto3.client('cleanroomsml')

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"
        }
    ]
)
```

Exportación de artefactos del modelo

Esta tarea es opcional y debe completarse cuando haya asignado la capacidad de `CAN_RECEIVE_MODEL_OUTPUT` miembro a un miembro de la colaboración.

Una vez finalizada la formación del modelo, el miembro que lo entrenó puede iniciar la exportación de los artefactos del modelo. El miembro que entrenó el modelo elige quién recibirá los artefactos del modelo, siempre que dicho miembro tenga la capacidad de recibir los resultados y una configuración de aprendizaje automático válida.

Console

Para configurar un algoritmo de modelo de aprendizaje automático personalizado en AWS Clean Rooms

1. Inicie sesión en AWS Management Console y abra la AWS Clean Rooms consola con su Cuenta de AWS (si aún no lo ha hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. En la página Colaboraciones, elige la colaboración que contiene el modelo personalizado que deseas exportar.
4. Cuando se abra la colaboración, seleccione la pestaña Modelos de aprendizaje automático y, a continuación, elija su modelo de la tabla de modelos entrenados personalizados
5. En la página de detalles del modelo entrenado personalizado, haga clic en Exportar la salida del modelo.
6. En Exportar la salida del modelo, en Exportar detalles de salida del modelo, introduzca un nombre y, si lo desea, una descripción.

   Elija qué miembro recibirá los artefactos del modelo en la lista desplegable del modelo exportado a los miembros de la colaboración.
7. Seleccione Exportar.

   Los resultados se exportan a la siguiente ruta en la ubicación de Amazon S3 que se especificó en la configuración de ML:`yourSpecifiedS3Path/collaborationIdentifier/trainedModelName/callerAccountId/jobName`. Solo se exportan los archivos que se van a exportar, hasta el tamaño máximo de archivo especificado, que haya seleccionado al asociar el algoritmo del modelo configurado.

API

Para iniciar la exportación del modelo, ejecute el código siguiente:

```
import boto3
acr_ml_client= boto3.client('cleanroomsml')

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'
)
```

Los resultados se exportan a la siguiente ruta en la ubicación de Amazon S3 que se especificó en la configuración de ML:`yourSpecifiedS3Path/collaborationIdentifier/trainedModelName/callerAccountId/jobName`. Solo se exporta el algoritmo del modelo configurado, hasta el `maxSize` especificado, que haya seleccionado al asociar. `filesToExport`

Ejecute la inferencia en un modelo entrenado

Los miembros con la capacidad de ejecutar consultas también pueden iniciar el trabajo de inferencia una vez finalizado el trabajo de formación. Escogen el conjunto de datos de inferencias con el que quieren ejecutar la inferencia y hacen referencia a los resultados del modelo entrenado con los que desean ejecutar el contenedor de inferencias.

Al miembro que recibirá el resultado de la inferencia se le debe conceder la habilidad de miembro. `CAN_RECEIVE_INFERENCE_OUTPUT`

Console

Para crear un trabajo de inferencia de modelos en AWS Clean Rooms

1. Inicie sesión en AWS Management Console y abra la AWS Clean Rooms consola con su Cuenta de AWS (si aún no lo ha hecho).
2. En el panel de navegación izquierdo, elija Colaboraciones.
3. En la página Colaboraciones, elige la colaboración que contiene el modelo personalizado sobre el que quieres crear un trabajo de inferencia.
4. Cuando se abra la colaboración, seleccione la pestaña Modelos de aprendizaje automático y, a continuación, elija su modelo de la tabla de modelos entrenados personalizados.
5. En la página de detalles del modelo entrenado personalizado, haga clic en Iniciar trabajo de inferencia.
6. En Iniciar trabajo de inferencia, en Detalles del trabajo de inferencia, introduzca un nombre y, si lo desea, una descripción.

Introduzca la información siguiente:

- Algoritmo del modelo asociado: el algoritmo del modelo asociado que se utiliza durante el trabajo de inferencia.
- Detalles del canal de entrada ML: el canal de entrada ML que proporcionará los datos para este trabajo de inferencia.
- Recursos de transformación: la instancia de procesamiento que se utiliza para realizar la función de transformación del trabajo de inferencia.
- Configuración de salida: quién recibirá el resultado del trabajo de inferencia y el tipo MIME del resultado.
- Cifrado: elija Personalizar la configuración de cifrado para especificar su propia clave de KMS y la información relacionada. De lo contrario, Clean Rooms ML gestionará el cifrado.
- Transformar los detalles del trabajo: la carga útil máxima del trabajo de inferencia, en MB.
- Variables de entorno: cualquier variable de entorno necesaria para acceder a la imagen del contenedor de trabajos de inferencia.

7. Seleccione Iniciar trabajo de inferencia.

   Los resultados se exportan a la siguiente ruta en la ubicación de Amazon S3 que se especificó en la configuración de ML:`yourSpecifiedS3Path/collaborationIdentifier/trainedModelName/callerAccountId/jobName`.

API

Para iniciar el trabajo de inferencia, ejecute el siguiente código:

```
import boto3
acr_ml_client= boto3.client('cleanroomsml')

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'
            }
        ]
    }
)
```

Los resultados se exportan a la siguiente ruta en la ubicación de Amazon S3 que se especificó en la configuración de ML:`yourSpecifiedS3Path/collaborationIdentifier/trainedModelName/callerAccountId/jobName`.

Pasos a seguir a continuación

Una vez que haya creado un modelo personalizado, estará listo para:

- Cree una colaboración y una membresía en AWS Clean Rooms

Solución de problemas AWS Clean Rooms

En esta sección se describen algunos problemas comunes que pueden surgir al usarlos AWS Clean Rooms y cómo solucionarlos.

Problemas

- Una o más tablas a las que hace referencia la consulta no están accesibles para el rol de servicio asociado. El propietario de la tabla/rol debe conceder al rol de servicio acceso a la tabla.
- Uno de los conjuntos de datos subyacentes tiene un formato de archivo no compatible.
- Los resultados de las consultas no son los esperados cuando se utiliza la computación criptográfica para Clean Rooms.

Una o más tablas a las que hace referencia la consulta no están accesibles para el rol de servicio asociado. El propietario de la tabla/rol debe conceder al rol de servicio acceso a la tabla.

- Asegúrese de que los permisos del rol de servicio estén configurados como corresponda. Para obtener más información, consulte Con AWS Clean Rooms figuración.

Uno de los conjuntos de datos subyacentes tiene un formato de archivo no compatible.

- Asegúrese de que su conjunto de datos esté en uno de los formatos de archivo compatibles:
  - Parquet
  - RCFile
  - TextFile
  - SequenceFile
  - RegexSerde
  - OpenCSV
  - AVRO
  - JSON

Para obtener más información, consulte Formatos de datos para AWS Clean Rooms.

Los resultados de las consultas no son los esperados cuando se utiliza la computación criptográfica para Clean Rooms.

Si utiliza la computación criptográfica para Clean Rooms (C3R), compruebe que la consulta utiliza correctamente las columnas cifradas:

- La sealed las columnas solo se utilizan en SELECT cláusulas.
- La fingerprint las columnas solo se utilizan en JOIN cláusulas (y GROUP BY cláusulas bajo ciertas condiciones).
- Que solo eres JOINing fingerprint columnas con el mismo nombre si la configuración de colaboración lo requiere.

Para obtener más información, consulte the section called “Computación criptográfica” y the section called “Tipos de columnas”.

Seguridad en AWS Clean Rooms

La seguridad en la nube AWS es la máxima prioridad. Como AWS cliente, usted se beneficia de una arquitectura de centro de datos y red diseñada para cumplir con los requisitos de las organizaciones más sensibles a la seguridad.

La seguridad es una responsabilidad compartida entre usted AWS y usted. El modelo de responsabilidad compartida la describe como seguridad de la nube y seguridad en la nube:

- Seguridad de la nube: AWS es responsable de proteger la infraestructura que ejecuta AWS los servicios en la AWS nube. AWS también le proporciona servicios que puede utilizar de forma segura. Los auditores externos prueban y verifican periódicamente la eficacia de nuestra seguridad como parte de los AWS programas de de . Para obtener más información sobre los programas de conformidad aplicables AWS Clean Rooms, consulte Servicios de AWS dentro del alcance por programa de conformidad .
- Seguridad en la nube: su responsabilidad viene determinada por el AWS servicio que utilice. También es responsable de otros factores, incluida la confidencialidad de los datos, los requisitos de la empresa y la legislación y los reglamentos aplicables

Esta documentación le ayuda a comprender cómo aplicar el modelo de responsabilidad compartida cuando se utiliza AWS Clean Rooms. Le muestra cómo configurarlo para AWS Clean Rooms cumplir sus objetivos de seguridad y conformidad. También aprenderá a utilizar otros AWS servicios que le ayudan a supervisar y proteger sus AWS Clean Rooms recursos.

Contenido

- Protección de datos en AWS Clean Rooms
- Retención de datos en AWS Clean Rooms
- Mejores prácticas para la colaboración de datos en AWS Clean Rooms
- Identity and Access Management para AWS Clean Rooms
- Validación de conformidad para AWS Clean Rooms
- Resiliencia en AWS Clean Rooms
- Seguridad de la infraestructura en AWS Clean Rooms
- Acceda al aprendizaje AWS Clean Rooms automático mediante AWS Clean Rooms un punto final de interfaz ()AWS PrivateLink

Protección de datos en AWS Clean Rooms

El modelo de responsabilidad AWS compartida modelo se aplica a la protección de datos en AWS Clean Rooms. Como se describe en este modelo, AWS es responsable de proteger la infraestructura global que ejecuta todos los Nube de AWS. Eres responsable de mantener el control sobre el contenido alojado en esta infraestructura. También eres responsable de las tareas de administración y configuración de seguridad para los Servicios de AWS que utiliza. Para obtener más información sobre la privacidad de los datos, consulta las Preguntas frecuentes sobre la privacidad de datos. Para obtener información sobre la protección de datos en Europa, consulta la publicación de blog sobre el Modelo de responsabilidad compartida de AWS y GDPR en el Blog de seguridad de AWS .

Con fines de protección de datos, le recomendamos que proteja Cuenta de AWS las credenciales y configure los usuarios individuales con AWS IAM Identity Center o AWS Identity and Access Management (IAM). De esta manera, solo se otorgan a cada usuario los permisos necesarios para cumplir sus obligaciones laborales. También recomendamos proteger sus datos de la siguiente manera:

- Utiliza la autenticación multifactor (MFA) en cada cuenta.
- Se utiliza SSL/TLS para comunicarse con AWS los recursos. Se recomienda el uso de TLS 1.2 y recomendamos TLS 1.3.
- Configure la API y el registro de actividad de los usuarios con AWS CloudTrail. Para obtener información sobre el uso de CloudTrail senderos para capturar AWS actividades, consulte Cómo trabajar con CloudTrail senderos en la Guía del AWS CloudTrail usuario.
- Utilice soluciones de AWS cifrado, junto con todos los controles de seguridad predeterminados Servicios de AWS.
- Utiliza servicios de seguridad administrados avanzados, como Amazon Macie, que lo ayuden a detectar y proteger los datos confidenciales almacenados en Amazon S3.
- Si necesita módulos criptográficos validados por FIPS 140-3 para acceder a AWS través de una interfaz de línea de comandos o una API, utilice un punto final FIPS. Para obtener más información sobre los puntos de conexión de FIPS disponibles, consulta Estándar de procesamiento de la información federal (FIPS) 140-3.

Se recomienda encarecidamente no introducir nunca información confidencial o sensible, como por ejemplo, direcciones de correo electrónico de clientes, en etiquetas o campos de formato libre, tales como el campo Nombre. Esto incluye cuando trabaja con AWS Clean Rooms o Servicios de AWS utiliza la consola, la API o. AWS CLI AWS SDKs Cualquier dato que ingrese en etiquetas o campos

de texto de formato libre utilizados para nombres se puede emplear para los registros de facturación o diagnóstico. Si proporciona una URL a un servidor externo, recomendamos encarecidamente que no incluya la información de las credenciales en la URL para validar la solicitud para ese servidor.

Cifrado en reposo

AWS Clean Rooms siempre cifra todos los metadatos del servicio en reposo sin necesidad de ninguna configuración adicional. Este cifrado es automático cuando se utiliza AWS Clean Rooms.

Clean Rooms ML cifra todos los datos almacenados en el servicio en reposo. AWS KMS Si decide proporcionar su propia clave KMS, el contenido de sus modelos similares y de los trabajos de generación de segmentos similares se cifra en reposo con su clave KMS.

Cuando se utilizan modelos de aprendizaje automático AWS Clean Rooms personalizados, el servicio cifra todos los datos almacenados en reposo. AWS KMS AWS Clean Rooms admite el uso de claves simétricas administradas por el cliente que usted crea, posee y administra para cifrar los datos en reposo. Si no se especifican las claves administradas por el cliente, Claves propiedad de AWS se utilizan de forma predeterminada.

AWS Clean Rooms utiliza concesiones y políticas clave para acceder a las claves gestionadas por el cliente. Puede revocar el acceso a la concesión o eliminar el acceso del servicio a la clave administrada por el cliente en cualquier momento. Si lo hace, AWS Clean Rooms no podrá acceder a ninguno de los datos cifrados por la clave gestionada por el cliente, lo que afectará a las operaciones que dependen de esos datos. Por ejemplo, si intenta crear un modelo entrenado a partir de un canal de entrada de aprendizaje automático cifrado al que no AWS Clean Rooms se puede acceder, la operación devolverá un `ValidationException` error.

> **Note**
>
> Puede utilizar las opciones de cifrado de Amazon S3 para proteger sus datos en reposo. Para obtener más información, consulte Especificación del cifrado de Amazon S3 en la Guía del usuario de Amazon S3.

Cuando se utiliza una tabla de mapeo de ID interna AWS Clean Rooms, el servicio cifra todos los datos almacenados en reposo. AWS KMS Si decides proporcionar tu propia clave de KMS, el contenido de la tabla de mapeo de ID se cifra en reposo con tu clave de KMS mediante AWS Entity Resolution. Para obtener más información sobre los permisos necesarios para trabajar con cifrados

con un flujo de trabajo de asignación de ID, consulte Crear un rol de trabajo del flujo de trabajo para AWS Entity Resolution en la Guía del usuario de AWS Entity Resolution .

Cifrado en tránsito

AWS Clean Rooms utiliza Transport Layer Security (TLS) para el cifrado en tránsito. La comunicación siempre AWS Clean Rooms se realiza a través de HTTPS, por lo que sus datos siempre se cifran en tránsito, independientemente de si están almacenados en Amazon S3, Amazon Athena o Snowflake. Esto incluye todos los datos en tránsito cuando se utiliza Clean Rooms ML.

Cifrado de datos subyacentes

Para obtener más información sobre cómo cifrar sus datos subyacentes, consulte Computación criptográfica para Clean Rooms.

Política de claves

Las políticas de clave controlan el acceso a la clave administrada por el cliente. Cada clave administrada por el cliente debe tener exactamente una política de clave, que contiene instrucciones que determinan quién puede usar la clave y cómo puede utilizarla. Cuando crea la clave administrada por el cliente, puede especificar una política de clave. Para obtener más información, consulte Administrar el acceso a las claves administradas por el cliente en la AWS Key Management Service Guía para desarrolladores.

Para utilizar la clave gestionada por el cliente con los modelos de aprendizaje automático AWS Clean Rooms personalizados, la política de claves debe permitir las siguientes operaciones de API:

- `kms:DescribeKey`— Proporciona los detalles de la clave gestionada por el cliente AWS Clean Rooms para permitir su validación.
- `kms:Decrypt`— Proporciona acceso AWS Clean Rooms para descifrar los datos cifrados y utilizarlos en tareas relacionadas.
- `kms:CreateGrant`- Clean Rooms ML cifra las imágenes de formación e inferencia inactivas en Amazon ECR mediante la creación de subvenciones para Amazon ECR. Para obtener más información, consulte Encryption at Rest in Amazon ECR. Clean Rooms ML también utiliza Amazon SageMaker AI para realizar trabajos de formación e inferencia, y crea subvenciones para que SageMaker AI cifre los volúmenes de Amazon EBS adjuntos a las instancias, así como los datos de salida en Amazon S3. Para obtener más información, consulte Proteger los datos en reposo mediante el cifrado en Amazon SageMaker AI.

- `kms:GenerateDataKey`- Clean Rooms ML cifra los datos en reposo almacenados en Amazon S3 mediante el cifrado del lado del servidor con. AWS KMS keys Para obtener más información, consulte Uso del cifrado del lado del servidor con AWS KMS keys (SSE-KMS) en Amazon S3.

Los siguientes son ejemplos de declaraciones de políticas que puede añadir a los AWS Clean Rooms siguientes recursos:

Canal de entrada ML

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
        "Sid": "Allow access to principals authorized to use Clean Rooms ML",
        "Effect": "Allow",
        "Principal": {
            "AWS": "arn:aws:iam::444455556666:role/ExampleRole"
        },
        "Action": [
            "kms:DescribeKey",
            "kms:GenerateDataKey",
            "kms:Decrypt"
        ],
        "Resource": "*",
        "Condition": {
            "StringEquals": {
                "kms:ViaService": "cleanrooms-ml.region.amazonaws.com"
            }
        }
    },
    {
        "Sid": "Allow access to Clean Rooms ML service principal",
        "Effect": "Allow",
        "Principal": {
            "Service": "cleanrooms-ml.amazonaws.com"
        },
        "Action": [
            "kms:DescribeKey",
            "kms:GenerateDataKey",
            "kms:Decrypt"
        ],
        "Resource": "*"
```

```
    }
  ]
}
```

Trabajo de modelo entrenado o trabajo de inferencia de modelo entrenado

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
        "Sid": "Allow access to principals authorized to use Clean Rooms ML",
        "Effect": "Allow",
        "Principal": { "AWS": "arn:aws:iam::444455556666:role/ExampleRole" },
        "Action": [
            "kms:GenerateDataKey",
            "kms:DescribeKey",
            "kms:CreateGrant",
            "kms:Decrypt"
        ],
        "Resource": "*",
        "Condition": {
            "StringEquals": {
                "kms:ViaService": "cleanrooms-ml.region.amazonaws.com"
            }
            "ForAllValues:StringEquals": {
                "kms:GrantOperations": [
                    "Decrypt",
                        "Encrypt",
                        "GenerateDataKeyWithoutPlaintext",
                        "ReEncryptFrom",
                        "ReEncryptTo",
                        "CreateGrant",
                        "DescribeKey",
                        "RetireGrant",
                        "GenerateDataKey"
                ]
              },
            "BoolIfExists": {
              "kms:GrantIsForAWSResource": true
            }
        }
    },
    {
```

```
        "Sid": "Allow access to Clean Rooms ML service principal",
        "Effect": "Allow",
        "Principal": {
            "Service": "cleanrooms-ml.amazonaws.com"
        },
        "Action": [
            "kms:GenerateDataKey",
            "kms:DescribeKey",
            "kms:CreateGrant",
            "kms:Decrypt"
        ],
        "Resource": "*",
        "Condition": {
            "ForAllValues:StringEquals": {
                "kms:GrantOperations": [
                        "Decrypt",
                        "Encrypt",
                        "GenerateDataKeyWithoutPlaintext",
                        "ReEncryptFrom",
                        "ReEncryptTo",
                        "CreateGrant",
                        "DescribeKey",
                        "RetireGrant",
                        "GenerateDataKey"
                ]
              }
        }
    }
  ]
}
```

Clean Rooms ML no permite especificar el contexto de cifrado del servicio o el contexto de origen en las políticas clave administradas por el cliente. Los clientes pueden ver el contexto de cifrado utilizado internamente por el servicio CloudTrail.

Retención de datos en AWS Clean Rooms

Todos los datos que se lean temporalmente en una AWS Clean Rooms colaboración se eliminan una vez finalizada la consulta.

Al crear un modelo similar, Clean Rooms ML lee los datos de entrenamiento, los transforma en un formato adecuado para nuestro modelo de ML y almacena los parámetros del modelo

entrenado dentro de Clean Rooms ML. Clean Rooms ML no conserva una copia de sus datos de entrenamiento. AWS Clean Rooms Las consultas SQL no conservan ninguno de sus datos una vez ejecutada la consulta. A continuación, Clean Rooms ML utiliza el modelo entrenado para resumir el comportamiento de todos los usuarios. Clean Rooms ML almacena un conjunto de datos en el nivel de usuario para cada usuario de los datos durante el tiempo que el modelo similar esté activo.

Al iniciar un trabajo de generación de segmentos similares, Clean Rooms ML lee los datos iniciales, lee los resúmenes de comportamiento del modelo similar asociado y crea un segmento similar que se almacena en el servicio. AWS Clean Rooms Clean Rooms ML no retiene ninguna copia de los datos iniciales. Clean Rooms ML almacena la salida del trabajo en el nivel de usuario mientras el trabajo esté activo.

Si los datos iniciales provienen de una consulta SQL, el resultado de esa consulta solo se almacena en el servicio mientras dure el trabajo. Los resultados de la consulta están cifrados en reposo y en tránsito.

Utilice la API para eliminar los datos del trabajo de generación de segmentos o modelos similares. Clean Rooms ML elimina de forma asíncrona todos los datos asociados al modelo o trabajo. Una vez finalizado este proceso, Clean Rooms ML elimina los metadatos del modelo o trabajo y ya no están visibles en la API. Clean Rooms ML retiene los datos eliminados durante 3 días para evitar la recuperación ante desastres. Una vez que el trabajo o el modelo ya no estén visibles en la API y hayan transcurrido 3 días, todos los datos asociados al modelo o al trabajo se eliminarán permanentemente.

Mejores prácticas para la colaboración de datos en AWS Clean Rooms

En este tema se describen las prácticas recomendadas para realizar colaboraciones de datos en AWS Clean Rooms.

AWS Clean Rooms sigue el modelo de responsabilidad AWS compartida. AWS Clean Rooms ofrece reglas de análisis que puede configurar para reforzar su capacidad de proteger los datos confidenciales en una colaboración. Las reglas de análisis que configure AWS Clean Rooms impondrán las restricciones (controles de consulta y controles de salida de consultas) que haya configurado. Usted es responsable de determinar las restricciones y configurar las reglas de análisis como corresponda.

Las colaboraciones de datos pueden implicar algo más que su uso AWS Clean Rooms. Para ayudarlo a maximizar los beneficios de las colaboraciones de datos, le recomendamos que lleve a cabo las siguientes prácticas recomendadas con el uso de las reglas de análisis AWS Clean Rooms y, específicamente, con ellas.

Temas

- Mejores prácticas con AWS Clean Rooms
- Prácticas recomendadas para utilizar reglas de análisis en AWS Clean Rooms

Mejores prácticas con AWS Clean Rooms

Usted es responsable de evaluar el riesgo de cada colaboración de datos y compararlo con sus requisitos de privacidad (por ejemplo, con sus programas y políticas de cumplimiento externos e internos). Le recomendamos que tome medidas adicionales al usar AWS Clean Rooms. Estas acciones pueden ayudarle a administrar mejor los riesgos y a protegerse de los intentos de reidentificación de los datos por parte de terceros (por ejemplo, ataques diferenciales o ataques de canal lateral).

Por ejemplo, considere la posibilidad de realizar comprobaciones de diligencia debida en sus otros colaboradores y de formalizar acuerdos legales con ellos antes de iniciar una colaboración. Para monitorizar el uso de sus datos, considere también la posibilidad de adoptar otros mecanismos de auditoría cuando utilice AWS Clean Rooms.

Prácticas recomendadas para utilizar reglas de análisis en AWS Clean Rooms

Las reglas de análisis AWS Clean Rooms permiten restringir las consultas que se pueden ejecutar configurando los controles de consulta en una tabla configurada. Por ejemplo, puede definir un control de consulta sobre cómo se puede combinar una tabla configurada y qué columnas se pueden seleccionar. También puede restringir el resultado de la consulta configurando controles de resultados de consulta, como umbrales de agregación en las filas de salida. El servicio rechaza cualquier consulta y elimina las filas que no cumplen las reglas de análisis establecidas por los miembros en sus tablas configuradas de la consulta.

Es aconsejable que siga estas 10 prácticas recomendadas a la hora de usar reglas de análisis en su tabla configurada:

- Cree tablas configuradas distintas para distintos casos de uso de consultas (por ejemplo, planificación de audiencias o atribución). Puede crear varias tablas configuradas con la misma tabla de AWS Glue subyacente.
- Especifique las columnas de la regla de análisis (por ejemplo, columnas de dimensión, columnas de lista, columnas de combinación) que sean necesarias para las consultas de una colaboración. Esto puede contribuir a mitigar el riesgo de ataques diferenciales o de que otros miembros apliquen técnicas de ingeniería inversa a sus datos. Use la característica de lista de columnas permitidas para anotar otras columnas que quizás desee habilitar para consulta en el futuro. Para personalizar las columnas que se pueden usar para una colaboración determinada, cree tablas configuradas adicionales con la misma AWS Glue tabla subyacente.
- Especifique las funciones de la regla de análisis que son necesarias para el análisis en la colaboración. Esto puede contribuir a mitigar el riesgo derivado de errores de función poco frecuentes que puedan presentar información sobre un punto de datos concreto. Para personalizar las funciones que se pueden utilizar en una determinada colaboración, cree tablas configuradas adicionales con la misma tabla de AWS Glue subyacente.
- Añada restricciones de agregación a aquellas columnas cuyos valores a nivel de fila sean confidenciales. Esto incluye las columnas de la tabla configurada que también existen en las tablas y reglas de análisis de otros miembros de la colaboración como una restricción de agregación. También incluye las columnas de su tabla configurada que no se pueden consultar; es decir, las columnas que figuran en su tabla configurada, pero no en la regla de análisis. Las restricciones de agregación pueden contribuir a mitigar el riesgo de correlacionar los resultados de las consultas con datos externos a la colaboración.
- Cree colaboraciones y reglas de análisis de prueba para probar las restricciones creadas con las reglas de análisis especificadas.
- Revise las tablas configuradas de los colaboradores y las reglas de análisis de los miembros en las tablas configuradas para comprobar que coinciden con las condiciones acordadas para la colaboración. Esto puede contribuir a mitigar el riesgo de que otros miembros manipulen sus propios datos para ejecutar consultas no acordadas.
- Revise la consulta de ejemplo proporcionada (solo en la consola) que se habilita en la tabla configurada después de configurar la regla de análisis.

> **Note**
>
> Además de la consulta de ejemplo proporcionada, es posible realizar otras consultas en función de la regla de análisis y de otras tablas y reglas de análisis de los miembros de la colaboración.

- Puede agregar o actualizar una regla de análisis para una tabla configurada en una colaboración. Cuando lo haga, revise todas las colaboraciones a las que está asociada la tabla configurada y el efecto resultante. Esto le ayuda a asegurarse de que ninguna colaboración utilice reglas de análisis obsoletas.
- Revise las consultas ejecutadas en la colaboración para comprobar que coincidan con los casos de uso o con las consultas acordados para la colaboración (las consultas están disponibles en los registros de consultas cuando la característica Registro de consultas está activada). Esto puede contribuir a mitigar el riesgo de que los miembros ejecuten análisis no acordados y de que se produzcan posibles ataques (por ejemplo, ataques de canal lateral).
- Revise las columnas de la tabla configurada que se utilizan en las reglas de análisis de los miembros de la colaboración y en las consultas para asegurarse de que coinciden con lo acordado en la colaboración (las consultas están disponibles en los registros de consultas cuando la característica está activada). Esto puede contribuir a mitigar el riesgo de que otros miembros manipulen sus propios datos para ejecutar consultas no acordadas.

Identity and Access Management para AWS Clean Rooms

AWS Identity and Access Management (IAM) es una herramienta Servicio de AWS que ayuda al administrador a controlar de forma segura el acceso a AWS los recursos. Los administradores de IAM controlan quién puede autenticarse (iniciar sesión) y quién puede autorizarse (tener permisos) para usar los recursos. AWS Clean Rooms La IAM es una Servicio de AWS opción que puede utilizar sin coste adicional.

Temas

- Público
- Autenticación con identidades
- Administración de acceso mediante políticas
- ¿Cómo AWS Clean Rooms funciona con IAM

- Ejemplos de políticas basadas en la identidad para AWS Clean Rooms
- AWS políticas gestionadas para AWS Clean Rooms
- Solución de problemas AWS Clean Rooms de identidad y acceso
- Prevención de la sustitución confusa entre servicios
- Comportamientos de IAM para ML AWS Clean Rooms
- Comportamientos de IAM para los modelos personalizados de Clean Rooms ML

Público

La forma de usar AWS Identity and Access Management (IAM) varía según el trabajo en el que se realice. AWS Clean Rooms

Usuario del servicio: si utiliza el AWS Clean Rooms servicio para realizar su trabajo, el administrador le proporcionará las credenciales y los permisos que necesita. A medida que vaya utilizando más AWS Clean Rooms funciones para realizar su trabajo, es posible que necesite permisos adicionales. Entender cómo se administra el acceso puede ayudarle a solicitar los permisos correctos al administrador. Si no puede acceder a una característica en AWS Clean Rooms, consulte Solución de problemas AWS Clean Rooms de identidad y acceso.

Administrador de servicios: si estás a cargo de AWS Clean Rooms los recursos de tu empresa, probablemente tengas acceso total a ellos AWS Clean Rooms. Su trabajo consiste en determinar a qué AWS Clean Rooms funciones y recursos deben acceder los usuarios del servicio. Luego, debe enviar solicitudes a su gestionador de IAM para cambiar los permisos de los usuarios de su servicio. Revise la información de esta página para conocer los conceptos básicos de IAM. Para obtener más información sobre cómo su empresa puede utilizar la IAM AWS Clean Rooms, consulte¿Cómo AWS Clean Rooms funciona con IAM.

Administrador de IAM: si es un administrador de IAM, es posible que quiera conocer más detalles sobre cómo escribir políticas para administrar el acceso a AWS Clean Rooms. Para ver ejemplos de políticas AWS Clean Rooms basadas en la identidad que puede utilizar en IAM, consulte. Ejemplos de políticas basadas en la identidad para AWS Clean Rooms

Autenticación con identidades

La autenticación es la forma de iniciar sesión AWS con sus credenciales de identidad. Debe estar autenticado (con quien haya iniciado sesión AWS) como usuario de IAM o asumiendo una función de IAM. Usuario raíz de la cuenta de AWS

Puede iniciar sesión AWS como una identidad federada mediante las credenciales proporcionadas a través de una fuente de identidad. AWS IAM Identity Center Los usuarios (Centro de identidades de IAM) o la autenticación de inicio de sesión único de su empresa son ejemplos de identidades federadas. Al iniciar sesión como una identidad federada, su gestionador habrá configurado previamente la federación de identidades mediante roles de IAM. Cuando accede AWS mediante la federación, asume indirectamente un rol.

Según el tipo de usuario que sea, puede iniciar sesión en el portal AWS Management Console o en el de AWS acceso. Para obtener más información sobre cómo iniciar sesión AWS, consulte Cómo iniciar sesión Cuenta de AWS en su Guía del AWS Sign-In usuario.

Si accede AWS mediante programación, AWS proporciona un kit de desarrollo de software (SDK) y una interfaz de línea de comandos (CLI) para firmar criptográficamente sus solicitudes con sus credenciales. Si no utilizas AWS herramientas, debes firmar las solicitudes tú mismo. Para obtener más información sobre el uso del método recomendado para firmar las solicitudes usted mismo, consulte Proceso de firma de Signature Version 4 en la Referencia general de AWS.

Independientemente del método de autenticación que use, es posible que deba proporcionar información de seguridad adicional. Por ejemplo, le AWS recomienda que utilice la autenticación multifactor (MFA) para aumentar la seguridad de su cuenta. Para obtener más información, consulte Autenticación multifactor en la Guía del usuario de AWS IAM Identity Center y Uso de la autenticación multifactor (MFA) en AWSen la Guía del usuario de IAM.

Cuenta de AWS usuario root

Al crear una Cuenta de AWS, comienza con una identidad de inicio de sesión que tiene acceso completo a todos Servicios de AWS los recursos de la cuenta. Esta identidad recibe el nombre de usuario raíz de la Cuenta de AWS y se accede a ella iniciando sesión con el email y la contraseña que utilizó para crear la cuenta. Recomendamos que no utilice el usuario raíz para las tareas cotidianas. Proteja las credenciales del usuario raíz y utilícelas sólo para las tareas que el usuario raíz pueda realizar. Para obtener la lista completa de tareas que requieren que inicie sesión como usuario raíz, consulte Credenciales de Usuario raíz de la cuenta de AWS e identidades de IAM en la Referencia general de AWS.

Identidad federada

Como práctica recomendada, exija a los usuarios humanos, incluidos los que requieren acceso de administrador, que utilicen la federación con un proveedor de identidades para acceder Servicios de AWS mediante credenciales temporales.

Una identidad federada es un usuario del directorio de usuarios de su empresa, un proveedor de identidades web AWS Directory Service, el directorio del Centro de Identidad o cualquier usuario al que acceda Servicios de AWS mediante las credenciales proporcionadas a través de una fuente de identidad. Cuando las identidades federadas acceden Cuentas de AWS, asumen funciones y las funciones proporcionan credenciales temporales.

Para una administración de acceso centralizada, le recomendamos que utiliza AWS IAM Identity Center. Puede crear usuarios y grupos en el Centro de identidades de IAM, o puede conectarse y sincronizarse con un conjunto de usuarios y grupos de su propia fuente de identidad para usarlos en todas sus Cuentas de AWS aplicaciones. Para obtener más información, consulta ¿Qué es el Centro de identidades de IAM? en la Guía del usuario de AWS IAM Identity Center .

Usuarios y grupos de IAM

Un usuario de IAM es una identidad propia Cuenta de AWS que tiene permisos específicos para una sola persona o aplicación. Siempre que sea posible, recomendamos emplear credenciales temporales, en lugar de crear usuarios de IAM que tengan credenciales de larga duración como contraseñas y claves de acceso. No obstante, si tiene casos de uso específicos que requieran credenciales de larga duración con usuarios de IAM, recomendamos rotar las claves de acceso. Para más información, consulte Rotar las claves de acceso periódicamente para casos de uso que requieran credenciales de larga duración en la Guía del usuario de IAM.

Un grupo de IAM es una identidad que especifica un conjunto de usuarios de IAM. No puedes iniciar sesión como grupo. Puedes usar los grupos para especificar permisos para varios usuarios a la vez. Los grupos facilitan la administración de los permisos para grandes conjuntos de usuarios. Por ejemplo, puede asignar un nombre a un grupo IAMAdminsy concederle permisos para administrar los recursos de IAM.

Los usuarios son diferentes de los roles. Un usuario se asocia exclusivamente a una persona o aplicación, pero la intención es que cualquier usuario pueda asumir un rol que necesite. Los usuarios tienen credenciales de larga duración permanentes; no obstante, los roles proporcionan credenciales temporales. Para obtener más información, consulte Casos de uso para usuarios de IAM en la Guía del usuario de IAM.

Roles de IAM

Un rol de IAM es una identidad dentro de usted Cuenta de AWS que tiene permisos específicos. Es similar a un usuario de IAM, pero no está asociado a una persona determinada. Para asumir temporalmente un rol de IAM en el AWS Management Console, puede cambiar de un rol de usuario

a uno de IAM (consola). Puedes asumir un rol llamando a una operación de AWS API AWS CLI o usando una URL personalizada. Para más información sobre los métodos para el uso de roles, consulta Métodos para asumir un rol en la Guía del usuario de IAM.

Los roles de IAM con credenciales temporales son útiles en las siguientes situaciones:

- Acceso de usuario federado: para asignar permisos a una identidad federada, puede crear un rol y definir sus permisos. Cuando se autentica una identidad federada, se asocia la identidad al rol y se le conceden los permisos define el rol. Para obtener información acerca de roles de federación, consulte Crear un rol para un proveedor de identidad de terceros (federación) en la Guía de usuario de IAM. Si utiliza el IAM Identity Center, debe configurar un conjunto de permisos. IAM Identity Center correlaciona el conjunto de permisos con un rol en IAM para controlar a qué pueden acceder las identidades después de autenticarse. Para obtener información acerca de los conjuntos de permisos, consulta Conjuntos de permisos en la Guía del usuario de AWS IAM Identity Center .
- Permisos de usuario de IAM temporales: un usuario de IAM puede asumir un rol de IAM para recibir temporalmente permisos distintos que le permitan realizar una tarea concreta.
- Acceso entre cuentas: puede utilizar un rol de IAM para permitir que alguien (una entidad principal de confianza) de otra cuenta acceda a los recursos de la cuenta. Los roles son la forma principal de conceder acceso entre cuentas. Sin embargo, con algunas Servicios de AWS, puedes adjuntar una política directamente a un recurso (en lugar de usar un rol como proxy). Para obtener información acerca de la diferencia entre los roles y las políticas basadas en recursos para el acceso entre cuentas, consulta Acceso a recursos entre cuentas en IAM en la Guía del usuario de IAM.
- Acceso entre servicios: algunos Servicios de AWS utilizan funciones en otros Servicios de AWS. Por ejemplo, cuando realizas una llamada en un servicio, es habitual que ese servicio ejecute aplicaciones en Amazon EC2 o almacene objetos en Amazon S3. Es posible que un servicio haga esto usando los permisos de la entidad principal, usando un rol de servicio o usando un rol vinculado al servicio.
  - Sesiones de acceso directo (FAS): cuando utilizas un usuario o un rol de IAM para realizar acciones en AWS ellas, se te considera principal. Cuando utiliza algunos servicios, es posible que realice una acción que desencadene otra acción en un servicio diferente. El FAS utiliza los permisos del principal que llama Servicio de AWS y los solicita Servicio de AWS para realizar solicitudes a los servicios descendentes. Las solicitudes de FAS solo se realizan cuando un servicio recibe una solicitud que requiere interacciones con otros Servicios de AWS recursos para completarse. En este caso, debe tener permisos para realizar ambas acciones. Para

obtener información sobre las políticas a la hora de realizar solicitudes de FAS, consulte Reenviar sesiones de acceso.

- Rol de servicio: un rol de servicio es un rol de IAM que adopta un servicio para realizar acciones en su nombre. Un administrador de IAM puede crear, modificar y eliminar un rol de servicio desde IAM. Para obtener más información, consulte Creación de un rol para delegar permisos a un Servicio de AWS en la Guía del usuario de IAM.
- Función vinculada al servicio: una función vinculada a un servicio es un tipo de función de servicio que está vinculada a un. Servicio de AWS El servicio puede asumir el rol para realizar una acción en su nombre. Los roles vinculados al servicio aparecen en usted Cuenta de AWS y son propiedad del servicio. Un administrador de IAM puede ver, pero no editar, los permisos de los roles vinculados a servicios.

- Aplicaciones que se ejecutan en Amazon EC2: puedes usar un rol de IAM para administrar las credenciales temporales de las aplicaciones que se ejecutan en una EC2 instancia y realizan AWS CLI solicitudes a la AWS API. Esto es preferible a almacenar las claves de acceso en la EC2 instancia. Para asignar un AWS rol a una EC2 instancia y ponerlo a disposición de todas sus aplicaciones, debe crear un perfil de instancia adjunto a la instancia. Un perfil de instancia contiene el rol y permite que los programas que se ejecutan en la EC2 instancia obtengan credenciales temporales. Para obtener más información, consulte Usar un rol de IAM para conceder permisos a las aplicaciones que se ejecutan en EC2 instancias de Amazon en la Guía del usuario de IAM.

Administración de acceso mediante políticas

El acceso se controla AWS creando políticas y adjuntándolas a AWS identidades o recursos. Una política es un objeto AWS que, cuando se asocia a una identidad o un recurso, define sus permisos. AWS evalúa estas políticas cuando un director (usuario, usuario raíz o sesión de rol) realiza una solicitud. Los permisos en las políticas determinan si la solicitud se permite o se deniega. La mayoría de las políticas se almacenan AWS como documentos JSON. Para obtener más información sobre la estructura y el contenido de los documentos de política JSON, consulte Información general de políticas JSON en la Guía del usuario de IAM.

Los administradores pueden usar las políticas de AWS JSON para especificar quién tiene acceso a qué. Es decir, qué entidad principal puede realizar acciones en qué recursos y bajo qué condiciones.

Cada entidad de IAM (usuario o rol) comienza sin permisos. De forma predeterminada, los usuarios no pueden hacer nada, ni siquiera cambiar sus propias contraseñas. Para conceder permiso a un usuario para hacer algo, el administrador debe asociar una política de permisos a un usuario. O bien

el administrador puede agregar al usuario a un grupo que tenga los permisos necesarios. Cuando el administrador concede permisos a un grupo, todos los usuarios de ese grupo obtienen los permisos.

Las políticas de IAM definen permisos para una acción independientemente del método que se utilice para realizar la operación. Por ejemplo, suponga que dispone de una política que permite la acción `iam:GetRole`. Un usuario con esa política puede obtener información sobre el rol de la API AWS Management Console AWS CLI, la o la AWS API.

Políticas basadas en identidades

Las políticas basadas en identidad son documentos de políticas de permisos JSON que puede asociar a una identidad, como un usuario de IAM, un grupo de usuarios o un rol. Estas políticas controlan qué acciones pueden realizar los usuarios y los roles, en qué recursos y en qué condiciones. Para obtener más información sobre cómo crear una política basada en identidad, consulte Creación de políticas de IAM en la Guía del usuario de IAM.

Las políticas basadas en identidades pueden clasificarse además como políticas insertadas o políticas administradas. Las políticas insertadas se integran directamente en un único usuario, grupo o rol. Las políticas administradas son políticas independientes que puedes adjuntar a varios usuarios, grupos y roles de su Cuenta de AWS. Las políticas administradas incluyen políticas AWS administradas y políticas administradas por el cliente. Para más información sobre cómo elegir una política administrada o una política insertada, consulte Elegir entre políticas administradas y políticas insertadas en la Guía del usuario de IAM.

Políticas basadas en recursos

Las políticas basadas en recursos son documentos de política JSON que se asocian a un recurso. Los ejemplos de políticas basadas en recursos son las políticas de confianza de roles de IAM y las políticas de bucket de Amazon S3. En los servicios que admiten políticas basadas en recursos, los administradores de servicios pueden utilizarlos para controlar el acceso a un recurso específico. Para el recurso al que se asocia la política, la política define qué acciones puede realizar una entidad principal especificada en ese recurso y en qué condiciones. Debe especificar una entidad principal en una política en función de recursos. Los principales pueden incluir cuentas, usuarios, roles, usuarios federados o. Servicios de AWS

Las políticas basadas en recursos son políticas insertadas que se encuentran en ese servicio. No puedes usar políticas AWS gestionadas de IAM en una política basada en recursos.

Otros tipos de políticas

AWS admite tipos de políticas adicionales y menos comunes. Estos tipos de políticas pueden establecer el máximo de permisos que los tipos de políticas más frecuentes le conceden.

- Límites de permisos: un límite de permisos es una característica avanzada que le permite establecer los permisos máximos que una política basada en identidad puede conceder a una entidad de IAM (usuario o rol de IAM). Puede establecer un límite de permisos para una identidad. Los permisos resultantes son la intersección de las políticas basadas en identidad de la entidad y los límites de sus permisos. Las políticas basadas en recursos que especifiquen el usuario o rol en el campo `Principal` no estarán restringidas por el límite de permisos. Una denegación explícita en cualquiera de estas políticas anulará el permiso. Para obtener más información sobre los límites de los permisos, consulte Límites de permisos para las entidades de IAM en la Guía del usuario de IAM.
- Políticas de control de servicios (SCPs): SCPs son políticas de JSON que especifican los permisos máximos para una organización o unidad organizativa (OU). AWS Organizations AWS Organizations es un servicio para agrupar y administrar de forma centralizada varios de los Cuentas de AWS que son propiedad de su empresa. Si habilitas todas las funciones de una organización, puedes aplicar políticas de control de servicios (SCPs) a una o a todas tus cuentas. El SCP limita los permisos de las entidades en las cuentas de los miembros, incluidas las de cada una Usuario raíz de la cuenta de AWS. Para obtener más información sobre Organizations SCPs, consulte Cómo SCPs trabajar en la Guía del AWS Organizations usuario.
- Políticas de sesión: las políticas de sesión son políticas avanzadas que se pasan como parámetro cuando se crea una sesión temporal mediante programación para un rol o un usuario federado. Los permisos de la sesión resultantes son la intersección de las políticas basadas en identidades del rol y las políticas de la sesión. Los permisos también pueden proceder de una política en función de recursos. Una denegación explícita en cualquiera de estas políticas anulará el permiso. Para más información, consulte Políticas de sesión en la Guía del usuario de IAM.

Varios tipos de políticas

Cuando se aplican varios tipos de políticas a una solicitud, los permisos resultantes son más complicados de entender. Para saber cómo se AWS determina si se debe permitir una solicitud cuando se trata de varios tipos de políticas, consulte la lógica de evaluación de políticas en la Guía del usuario de IAM.

¿Cómo AWS Clean Rooms funciona con IAM

Antes de utilizar IAM para gestionar el acceso AWS Clean Rooms, infórmese sobre las funciones de IAM disponibles para su uso. AWS Clean Rooms

Funciones de IAM que puede utilizar con AWS Clean Rooms

Característica de IAM	AWS Clean Rooms soporte
Políticas basadas en identidades	Sí
Políticas basadas en recursos	Parcial
Acciones de políticas	Sí
Recursos de políticas	Sí
Claves de condición de política (específicas del servicio)	Parcial
ACLs	No
ABAC (etiquetas en políticas)	Sí
Credenciales temporales	Sí
Sesiones de acceso directo (FAS)	Sí
Roles de servicio	Sí
Roles vinculados al servicio	No

Para obtener una visión general de cómo AWS Clean Rooms y otras funciones Servicios de AWS funcionan con la mayoría de las funciones de IAM, consulte Servicios de AWS la Guía del usuario de IAM sobre cómo funcionan con IAM.

Políticas basadas en la identidad para AWS Clean Rooms

Compatibilidad con las políticas basadas en identidad: sí

Las políticas basadas en identidad son documentos de políticas de permisos JSON que puede asociar a una identidad, como un usuario de IAM, un grupo de usuarios o un rol. Estas políticas controlan qué acciones pueden realizar los usuarios y los roles, en qué recursos y en qué condiciones. Para obtener más información sobre cómo crear una política basada en identidad, consulte Creación de políticas de IAM en la Guía del usuario de IAM.

Con las políticas basadas en identidades de IAM, puede especificar las acciones y los recursos permitidos o denegados, así como las condiciones en las que se permiten o deniegan las acciones. No es posible especificar la entidad principal en una política basada en identidad porque se aplica al usuario o rol al que está asociada. Para obtener más información sobre los elementos que puede utilizar en una política de JSON, consulte Referencia de los elementos de las políticas de JSON de IAM en la Guía del usuario de IAM.

Ejemplos de políticas basadas en la identidad para AWS Clean Rooms

Para ver ejemplos de políticas AWS Clean Rooms basadas en la identidad, consulte. Ejemplos de políticas basadas en la identidad para AWS Clean Rooms

Políticas basadas en recursos dentro de AWS Clean Rooms

Compatibilidad con las políticas basadas en recursos: parcial

Las políticas basadas en recursos son documentos de política JSON que se asocian a un recurso. Los ejemplos de políticas basadas en recursos son las políticas de confianza de roles de IAM y las políticas de bucket de Amazon S3. En los servicios que admiten políticas basadas en recursos, los administradores de servicios pueden utilizarlos para controlar el acceso a un recurso específico. Para el recurso al que se asocia la política, la política define qué acciones puede realizar una entidad principal especificada en ese recurso y en qué condiciones. Debe especificar una entidad principal en una política en función de recursos. Los principales pueden incluir cuentas, usuarios, roles, usuarios federados o. Servicios de AWS

Para habilitar el acceso entre cuentas, puede especificar toda una cuenta o entidades de IAM de otra cuenta como la entidad principal de una política en función de recursos. Añadir a una política en función de recursos una entidad principal entre cuentas es solo una parte del establecimiento de una relación de confianza. Cuando el principal y el recurso son diferentes Cuentas de AWS, el administrador de IAM de la cuenta de confianza también debe conceder a la entidad principal (usuario o rol) permiso para acceder al recurso. Para conceder el permiso, adjunte la entidad a una política basada en identidad. Sin embargo, si la política basada en recursos concede acceso a una

entidad principal de la misma cuenta, no es necesaria una política basada en identidad adicional. Para obtener más información, consulte Cross account resource access in IAM en la Guía del usuario de IAM.

El AWS Clean Rooms servicio solo admite un tipo de política basada en recursos denominada política de recursos gestionados de modelo similar configurada, que se adjunta a un modelo similar configurado. Esta política define qué entidades principales pueden realizar acciones en el modelo similar configurado.

Para obtener información sobre cómo asociar una política basada en recursos a un modelo similar configurado, consulte Comportamientos de IAM para ML AWS Clean Rooms.

Acciones políticas para AWS Clean Rooms

Compatibilidad con las acciones de políticas: sí

Los administradores pueden usar las políticas de AWS JSON para especificar quién tiene acceso a qué. Es decir, qué entidad principal puede realizar acciones en qué recursos y en qué condiciones.

El elemento `Action` de una política JSON describe las acciones que puede utilizar para conceder o denegar el acceso en una política. Las acciones políticas suelen tener el mismo nombre que la operación de AWS API asociada. Hay algunas excepciones, como acciones de solo permiso que no tienen una operación de API coincidente. También hay algunas operaciones que requieren varias acciones en una política. Estas acciones adicionales se denominan acciones dependientes.

Incluya acciones en una política para conceder permisos y así llevar a cabo la operación asociada.

Para ver una lista de AWS Clean Rooms acciones, consulta las acciones definidas AWS Clean Rooms en la Referencia de autorización del servicio.

Las acciones políticas AWS Clean Rooms utilizan el siguiente prefijo antes de la acción.

```
cleanrooms
```

Para especificar varias acciones en una única instrucción, sepárelas con comas.

```
"Action": [
      "cleanrooms:action1",
      "cleanrooms:action2"
```

```
        ]
```

Para ver ejemplos de políticas AWS Clean Rooms basadas en la identidad, consulte. Ejemplos de políticas basadas en la identidad para AWS Clean Rooms

Recursos de políticas para AWS Clean Rooms

Compatibilidad con los recursos de políticas: sí

Los administradores pueden usar las políticas de AWS JSON para especificar quién tiene acceso a qué. Es decir, qué entidad principal puedes realizar acciones en qué recursos y en qué condiciones.

El elemento `Resource` de la política JSON especifica el objeto u objetos a los que se aplica la acción. Las instrucciones deben contener un elemento `Resource` o `NotResource`. Como práctica recomendada, especifique un recurso utilizando el Nombre de recurso de Amazon (ARN). Puedes hacerlo para acciones que admitan un tipo de recurso específico, conocido como permisos de nivel de recurso.

Para las acciones que no admiten permisos de nivel de recurso, como las operaciones de descripción, utiliza un carácter comodín (*) para indicar que la instrucción se aplica a todos los recursos.

```
"Resource": "*"
```

Para ver una lista de los tipos de AWS Clean Rooms recursos y sus tipos ARNs, consulte los recursos definidos AWS Clean Rooms en la Referencia de autorización de servicios. Para obtener información sobre las acciones con las que puede especificar el ARN de cada recurso, consulte Acciones definidas por AWS Clean Rooms.

Para ver ejemplos de políticas AWS Clean Rooms basadas en la identidad, consulte. Ejemplos de políticas basadas en la identidad para AWS Clean Rooms

Claves de condición de la política para AWS Clean Rooms

Compatibilidad con claves de condición de políticas específicas del servicio: parcial

Los administradores pueden usar las políticas de AWS JSON para especificar quién tiene acceso a qué. Es decir, qué entidad principal puedes realizar acciones en qué recursos y en qué condiciones.

El elemento `Condition` (o bloque de `Condition`) permite especificar condiciones en las que entra en vigor una instrucción. El elemento `Condition` es opcional. Puedes crear expresiones condicionales que utilizan operadores de condición, tales como igual o menor que, para que la condición de la política coincida con los valores de la solicitud.

Si especifica varios elementos de `Condition` en una instrucción o varias claves en un único elemento de `Condition`, AWS las evalúa mediante una operación `AND` lógica. Si especifica varios valores para una única clave de condición, AWS evalúa la condición mediante una `OR` operación lógica. Se deben cumplir todas las condiciones antes de que se concedan los permisos de la instrucción.

También puedes utilizar variables de marcador de posición al especificar condiciones. Por ejemplo, puedes conceder un permiso de usuario de IAM para acceder a un recurso solo si está etiquetado con su nombre de usuario de IAM. Para más información, consulta Elementos de la política de IAM: variables y etiquetas en la Guía del usuario de IAM.

AWS admite claves de condición globales y claves de condición específicas del servicio. Para ver todas las claves de condición AWS globales, consulte las claves de contexto de condición AWS globales en la Guía del usuario de IAM.

Para saber cómo AWS Clean Rooms ML utiliza las claves de condición de las políticas, consulte Comportamientos de IAM para ML AWS Clean Rooms.

ACLs in AWS Clean Rooms

Soporta ACLs: No

Las listas de control de acceso (ACLs) controlan qué directores (miembros de la cuenta, usuarios o roles) tienen permisos para acceder a un recurso. ACLs son similares a las políticas basadas en recursos, aunque no utilizan el formato de documento de políticas JSON.

ABAC con AWS Clean Rooms

Admite ABAC (etiquetas en las políticas): sí

El control de acceso basado en atributos (ABAC) es una estrategia de autorización que define permisos en función de atributos. En AWS, estos atributos se denominan etiquetas. Puede adjuntar

etiquetas a las entidades de IAM (usuarios o roles) y a muchos AWS recursos. El etiquetado de entidades y recursos es el primer paso de ABAC. A continuación, designa las políticas de ABAC para permitir operaciones cuando la etiqueta de la entidad principal coincida con la etiqueta del recurso al que se intenta acceder.

ABAC es útil en entornos que crecen con rapidez y ayuda en situaciones en las que la administración de las políticas resulta engorrosa.

Para controlar el acceso en función de etiquetas, debe proporcionar información de las etiquetas en el elemento de condición de una política utilizando las claves de condición `aws:ResourceTag/key-name`, `aws:RequestTag/key-name` o `aws:TagKeys`.

Si un servicio admite las tres claves de condición para cada tipo de recurso, el valor es Sí para el servicio. Si un servicio admite las tres claves de condición solo para algunos tipos de recursos, el valor es Parcial.

Para obtener más información sobre ABAC, consulte Definición de permisos con la autorización de ABAC en la Guía del usuario de IAM. Para ver un tutorial con los pasos para configurar ABAC, consulte Uso del control de acceso basado en atributos (ABAC) en la Guía del usuario de IAM.

Utilizar credenciales temporales con AWS Clean Rooms

Compatibilidad con credenciales temporales: sí

Algunas Servicios de AWS no funcionan cuando inicias sesión con credenciales temporales. Para obtener información adicional, incluidas las que Servicios de AWS funcionan con credenciales temporales, consulta Cómo Servicios de AWS funcionan con IAM en la Guía del usuario de IAM.

Utiliza credenciales temporales si inicia sesión en ellas AWS Management Console mediante cualquier método excepto un nombre de usuario y una contraseña. Por ejemplo, cuando accedes AWS mediante el enlace de inicio de sesión único (SSO) de tu empresa, ese proceso crea automáticamente credenciales temporales. También crea credenciales temporales de forma automática cuando inicia sesión en la consola como usuario y luego cambia de rol. Para obtener más información sobre el cambio de roles, consulte Cambio de un usuario a un rol de IAM (consola) en la Guía del usuario de IAM.

Puedes crear credenciales temporales manualmente mediante la AWS CLI API o. AWS A continuación, puede utilizar esas credenciales temporales para acceder AWS. AWS recomienda generar credenciales temporales de forma dinámica en lugar de utilizar claves de acceso a largo plazo. Para obtener más información, consulte Credenciales de seguridad temporales en IAM.

Sesiones de acceso directo para AWS Clean Rooms

Admite sesiones de acceso directo (FAS): sí

Cuando utiliza un usuario o un rol de IAM para realizar acciones en AWSél, se le considera director. Cuando utiliza algunos servicios, es posible que realice una acción que desencadene otra acción en un servicio diferente. FAS utiliza los permisos del principal que llama y los que solicita Servicio de AWS para realizar solicitudes a los servicios descendentes. Servicio de AWS Las solicitudes de FAS solo se realizan cuando un servicio recibe una solicitud que requiere interacciones con otros Servicios de AWS recursos para completarse. En este caso, debe tener permisos para realizar ambas acciones. Para obtener información sobre las políticas a la hora de realizar solicitudes de FAS, consulte Reenviar sesiones de acceso.

Roles de servicio para AWS Clean Rooms

Compatibilidad con roles de servicio: sí

Un rol de servicio es un rol de IAM que asume un servicio para realizar acciones en su nombre. Un administrador de IAM puede crear, modificar y eliminar un rol de servicio desde IAM. Para obtener más información, consulte Creación de un rol para delegar permisos a un Servicio de AWS en la Guía del usuario de IAM.

> **Warning**
>
> Si se cambian los permisos de un rol de servicio, es posible que se interrumpa AWS Clean Rooms la funcionalidad. Edite las funciones de servicio solo cuando se AWS Clean Rooms proporcionen instrucciones para hacerlo.

Funciones vinculadas al servicio para AWS Clean Rooms

Compatibilidad con roles vinculados al servicio: no

Un rol vinculado a un servicio es un tipo de rol de servicio que está vinculado a un. Servicio de AWS El servicio puede asumir el rol para realizar una acción en su nombre. Los roles vinculados al servicio aparecen en usted Cuenta de AWS y son propiedad del servicio. Un administrador de IAM puedes ver, pero no editar, los permisos de los roles vinculados a servicios.

Para más información sobre cómo crear o administrar roles vinculados a servicios, consulta Servicios de AWS que funcionan con IAM. Busque un servicio en la tabla que incluya `Yes` en la columna Rol

vinculado a un servicio. Seleccione el vínculo Sí para ver la documentación acerca del rol vinculado a servicios para ese servicio.

Ejemplos de políticas basadas en la identidad para AWS Clean Rooms

De forma predeterminada, los usuarios y roles no tienen permiso para crear, ver ni modificar recursos de AWS Clean Rooms . Tampoco pueden realizar tareas mediante la AWS Management Console, AWS Command Line Interface (AWS CLI) o AWS la API. Un administrador de IAM puede crear políticas de IAM para conceder permisos a los usuarios para realizar acciones en los recursos que necesitan. A continuación, el administrador puedes añadir las políticas de IAM a roles y los usuarios puedes asumirlos.

Para obtener información acerca de cómo crear una política basada en identidades de IAM mediante el uso de estos documentos de políticas JSON de ejemplo, consulte Creación de políticas de IAM (consola) en la Guía del usuario de IAM.

Para obtener más información sobre las acciones y los tipos de recursos definidos AWS Clean Rooms, incluido el formato ARNs de cada uno de los tipos de recursos, consulte las claves de condición, recursos y acciones de AWS Clean Rooms la Referencia de autorización de servicios.

Temas

- Prácticas recomendadas sobre las políticas
- Mediante la consola de AWS Clean Rooms
- Cómo permitir a los usuarios consultar sus propios permisos

Prácticas recomendadas sobre las políticas

Las políticas basadas en la identidad determinan si alguien puede crear AWS Clean Rooms recursos de tu cuenta, acceder a ellos o eliminarlos. Estas acciones pueden generar costos adicionales para su Cuenta de AWS. Siga estas directrices y recomendaciones al crear o editar políticas basadas en identidades:

- Comience con las políticas AWS administradas y avance hacia los permisos con privilegios mínimos: para empezar a conceder permisos a sus usuarios y cargas de trabajo, utilice las políticas AWS administradas que otorgan permisos para muchos casos de uso comunes. Están disponibles en su. Cuenta de AWS Le recomendamos que reduzca aún más los permisos definiendo políticas administradas por el AWS cliente que sean específicas para sus casos de uso.

Con el fin de obtener más información, consulta las políticas administradas por AWS o las políticas administradas por AWS para funciones de tarea en la Guía de usuario de IAM.

- Aplique permisos de privilegio mínimo: cuando establezca permisos con políticas de IAM, conceda solo los permisos necesarios para realizar una tarea. Para ello, debe definir las acciones que se pueden llevar a cabo en determinados recursos en condiciones específicas, también conocidos como permisos de privilegios mínimos. Con el fin de obtener más información sobre el uso de IAM para aplicar permisos, consulta Políticas y permisos en IAM en la Guía del usuario de IAM.
- Utilice condiciones en las políticas de IAM para restringir aún más el acceso: puede agregar una condición a sus políticas para limitar el acceso a las acciones y los recursos. Por ejemplo, puede escribir una condición de políticas para especificar que todas las solicitudes deben enviarse utilizando SSL. También puedes usar condiciones para conceder el acceso a las acciones del servicio si se utilizan a través de una acción específica Servicio de AWS, por ejemplo AWS CloudFormation. Para obtener más información, consulta Elementos de la política de JSON de IAM: Condición en la Guía del usuario de IAM.
- Utiliza el analizador de acceso de IAM para validar las políticas de IAM con el fin de garantizar la seguridad y funcionalidad de los permisos: el analizador de acceso de IAM valida políticas nuevas y existentes para que respeten el lenguaje (JSON) de las políticas de IAM y las prácticas recomendadas de IAM. El analizador de acceso de IAM proporciona más de 100 verificaciones de políticas y recomendaciones procesables para ayudar a crear políticas seguras y funcionales. Para más información, consulte Validación de políticas con el Analizador de acceso de IAM en la Guía del usuario de IAM.
- Requerir autenticación multifactor (MFA): si tiene un escenario que requiere usuarios de IAM o un usuario raíz en Cuenta de AWS su cuenta, active la MFA para mayor seguridad. Para exigir la MFA cuando se invoquen las operaciones de la API, añada condiciones de MFA a sus políticas. Para más información, consulte Acceso seguro a la API con MFA en la Guía del usuario de IAM.

Para obtener más información sobre las prácticas recomendadas de IAM, consulte Prácticas recomendadas de seguridad en IAM en la Guía del usuario de IAM.

Mediante la consola de AWS Clean Rooms

Para acceder a la AWS Clean Rooms consola, debe tener un conjunto mínimo de permisos. Estos permisos deben permitirle enumerar y ver detalles sobre los AWS Clean Rooms recursos de su cuenta Cuenta de AWS. Si crea una política basada en identidades que sea más restrictiva que el mínimo de permisos necesarios, la consola no funcionará del modo esperado para las entidades (usuarios o roles) que tengan esa política.

No es necesario que concedas permisos mínimos de consola a los usuarios que solo realicen llamadas a la API AWS CLI o a la AWS API. En su lugar, permite el acceso únicamente a las acciones que coincidan con la operación de API que intentan realizar.

Para garantizar que los usuarios y los roles puedan seguir utilizando la AWS Clean Rooms consola, adjunte también la política *ReadOnly* AWS gestionada AWS Clean Rooms *FullAccess* o la política gestionada a las entidades. Para obtener más información, consulte Adición de permisos a un usuario en la Guía del usuario de IAM:

Cómo permitir a los usuarios consultar sus propios permisos

En este ejemplo, se muestra cómo podría crear una política que permita a los usuarios de IAM ver las políticas gestionadas e insertadas que se asocian a la identidad de sus usuarios. Esta política incluye permisos para completar esta acción en la consola o mediante programación mediante la API AWS CLI o AWS .

```
{
    "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 políticas gestionadas para AWS Clean Rooms

Una política AWS administrada es una política independiente creada y administrada por AWS. AWS Las políticas administradas están diseñadas para proporcionar permisos para muchos casos de uso comunes, de modo que pueda empezar a asignar permisos a usuarios, grupos y funciones.

Ten en cuenta que es posible que las políticas AWS administradas no otorguen permisos con privilegios mínimos para tus casos de uso específicos, ya que están disponibles para que los usen todos los AWS clientes. Se recomienda definir políticas administradas por el cliente específicas para sus casos de uso a fin de reducir aún más los permisos.

No puedes cambiar los permisos definidos en AWS las políticas administradas. Si AWS actualiza los permisos definidos en una política AWS administrada, la actualización afecta a todas las identidades principales (usuarios, grupos y roles) a las que está asociada la política. AWS es más probable que actualice una política AWS administrada cuando Servicio de AWS se lance una nueva o cuando estén disponibles nuevas operaciones de API para los servicios existentes.

Para obtener más información, consulte Políticas administradas de AWS en la Guía del usuario de IAM.

AWS política gestionada: **`AWSCleanRoomsReadOnlyAccess`**

Puede adjuntar `AWSCleanRoomsReadOnlyAccess` a sus entidades principales de IAM.

Esta política otorga permisos de solo lectura a recursos y metadatos de una colaboración de `AWSCleanRoomsReadOnlyAccess`.

Detalles de los permisos

Esta política incluye los permisos siguientes:

- `CleanRoomsRead`: concede a las entidades principales acceso de solo lectura al servicio.
- `ConsoleDisplayTables`— Permite a los directores acceder de solo lectura a los AWS Glue metadatos necesarios para mostrar los datos sobre AWS Glue las tablas subyacentes en la consola.
- `ConsoleLogSummaryQueryLogs`: concede a las entidades principales permiso para ver los registros de consultas.
- `ConsoleLogSummaryObtainLogs`: concede a las entidades principales permiso para recuperar los resultados de registro.

Para obtener una lista en JSON de los detalles de la política, consulte la Guía AWSCleanRoomsReadOnlyAccessde referencia de políticas AWS administradas.

AWS política gestionada: **`AWSCleanRoomsFullAccess`**

Puede adjuntar `AWSCleanRoomsFullAccess` a sus entidades principales de IAM.

Esta política otorga permisos administrativos que permiten el acceso total (lectura, escritura y actualización) a los recursos y metadatos de una AWS Clean Rooms colaboración. Esta política incluye el acceso para realizar consultas.

Detalles de los permisos

Esta política incluye los permisos siguientes:

- `CleanRoomsAccess`— Otorga acceso total a todas las acciones de todos los recursos para AWS Clean Rooms.
- `PassServiceRole`: concede acceso para transferir un rol de servicio únicamente al servicio (condición `PassedToService`) cuyo nombre contenga "cleanrooms".
- `ListRolesToPickServiceRole`— Permite a los directores enumerar todas sus funciones para poder elegir una función de servicio cuando la utilicen AWS Clean Rooms.
- `GetRoleAndListRolePoliciesToInspectServiceRole`: permite a las entidades principales ver el rol de servicio y la política correspondiente en IAM.
- `ListPoliciesToInspectServiceRolePolicy`: permite a las entidades principales ver el rol de servicio y la política correspondiente en IAM.
- `GetPolicyToInspectServiceRolePolicy`: permite a las entidades principales ver el rol de servicio y la política correspondiente en IAM.

- `ConsoleDisplayTables`— Permite a los directores acceder de solo lectura a los AWS Glue metadatos necesarios para mostrar los datos sobre AWS Glue las tablas subyacentes de la consola.
- `ConsolePickQueryResultsBucketListAll`: permite a las entidades principales elegir, de una lista de todos los buckets de S3 disponibles, un bucket de Amazon S3 en el que se escriban los resultados de sus consultas.
- `SetQueryResultsBucket`: permite a las entidades principales elegir un bucket de S3 en el que se escriban los resultados de sus consultas.
- `ConsoleDisplayQueryResults`: permite a las entidades principales mostrar al cliente los resultados de las consultas, leídos desde el bucket de S3.
- `WriteQueryResults`: permite a las entidades principales escribir los resultados de las consultas en un bucket de S3 propiedad del cliente.
- `EstablishLogDeliveries`— Permite a los directores enviar registros de consultas al grupo de registros de Amazon CloudWatch Logs de un cliente.
- `SetupLogGroupsDescribe`— Permite a los directores utilizar el proceso de creación de grupos de CloudWatch registros de Amazon Logs.
- `SetupLogGroupsCreate`— Permite a los directores crear un grupo de CloudWatch registros de Amazon Logs.
- `SetupLogGroupsResourcePolicy`— Permite a los directores configurar una política de recursos en el grupo de CloudWatch registros de Amazon Logs.
- `ConsoleLogSummaryQueryLogs`: concede a las entidades principales permiso para ver los registros de consultas.
- `ConsoleLogSummaryObtainLogs`: concede a las entidades principales permiso para recuperar los resultados de registro.

Para ver una lista en JSON de los detalles de la política, consulte AWSCleanRoomsFullAccessla Guía de referencia de políticas AWS administradas.

AWS política gestionada: **`AWSCleanRoomsFullAccessNoQuerying`**

Puede adjuntar `AWSCleanRoomsFullAccessNoQuerying` a sus IAM principals.

Esta política otorga permisos administrativos que permiten el acceso total (lectura, escritura y actualización) a los recursos y metadatos de una AWS Clean Rooms colaboración. Esta política no incluye el acceso para realizar consultas.

Detalles de los permisos

Esta política incluye los permisos siguientes:

- `CleanRoomsAccess`— Otorga acceso total a todas las acciones de todos los recursos AWS Clean Rooms, excepto para las consultas en las colaboraciones.
- `CleanRoomsNoQuerying`: deniega explícitamente `StartProtectedQuery` y `UpdateProtectedQuery` para impedir las consultas.
- `PassServiceRole`: concede acceso para transferir un rol de servicio únicamente al servicio (condición `PassedToService`) cuyo nombre contenga "cleanrooms".
- `ListRolesToPickServiceRole`— Permite a los directores enumerar todas sus funciones para poder elegir una función de servicio cuando la utilicen. AWS Clean Rooms
- `GetRoleAndListRolePoliciesToInspectServiceRole`: permite a las entidades principales ver el rol de servicio y la política correspondiente en IAM.
- `ListPoliciesToInspectServiceRolePolicy`: permite a las entidades principales ver el rol de servicio y la política correspondiente en IAM.
- `GetPolicyToInspectServiceRolePolicy`: permite a las entidades principales ver el rol de servicio y la política correspondiente en IAM.
- `ConsoleDisplayTables`— Permite a los directores acceder de solo lectura a los AWS Glue metadatos necesarios para mostrar los datos sobre AWS Glue las tablas subyacentes de la consola.
- `EstablishLogDeliveries`— Permite a los directores enviar registros de consultas al grupo de registros de Amazon CloudWatch Logs de un cliente.
- `SetupLogGroupsDescribe`— Permite a los directores utilizar el proceso de creación de grupos de CloudWatch registros de Amazon Logs.
- `SetupLogGroupsCreate`— Permite a los directores crear un grupo de CloudWatch registros de Amazon Logs.
- `SetupLogGroupsResourcePolicy`— Permite a los directores configurar una política de recursos en el grupo de CloudWatch registros de Amazon Logs.
- `ConsoleLogSummaryQueryLogs`: concede a las entidades principales permiso para ver los registros de consultas.
- `ConsoleLogSummaryObtainLogs`: concede a las entidades principales permiso para recuperar los resultados de registro.

- `cleanrooms`: administre las colaboraciones, las plantillas de análisis, las tablas configuradas, las suscripciones y los recursos asociados dentro del servicio de AWS Clean Rooms . Realice diversas operaciones, como crear, actualizar, eliminar, mostrar y recuperar información sobre estos recursos.
- `iam`— Transfiere al servicio las funciones de AWS Clean Rooms servicio cuyos nombres contengan `cleanrooms` «». Enumere las funciones y políticas e inspeccione las funciones y políticas de servicio relacionadas con el AWS Clean Rooms servicio.
- `glue`— Recupere información sobre bases de datos, tablas, particiones y esquemas de AWS Glue. Esto es necesario para que el AWS Clean Rooms servicio muestre las fuentes de datos subyacentes e interactúe con ellas.
- `logs`— Gestione las entregas de registros, los grupos de registros y las políticas de recursos para CloudWatch los registros. Consulte y recupere los registros relacionados con el AWS Clean Rooms servicio. Estos permisos son necesarios para la supervisión, la auditoría y la solución de problemas dentro del servicio.

La política también deniega explícitamente las acciones `cleanrooms:StartProtectedQuery` y `cleanrooms:UpdateProtectedQuery` para evitar que los usuarios ejecuten o actualicen directamente las consultas protegidas, lo que debe hacerse a través de los mecanismos controlados de AWS Clean Rooms .

Para obtener una lista en JSON de los detalles de la política, consulte AWSCleanRoomsFullAccessNoQueryingla Guía de referencia de políticas AWS gestionadas.

AWS política gestionada: **`AWSCleanRoomsMLReadOnlyAccess`**

Puede adjuntar `AWSCleanRoomsMLReadOnlyAccess` a sus entidades principales de IAM.

Esta política otorga permisos de solo lectura a recursos y metadatos de una colaboración de `AWSCleanRoomsMLReadOnlyAccess`.

Esta política incluye los permisos siguientes:

- `CleanRoomsConsoleNavigation`— Otorga acceso para ver las pantallas de la AWS Clean Rooms consola.
- `CleanRoomsMLRead`: permite a las entidades principales el acceso de solo lectura al servicio de Clean Rooms ML.
- `PassCleanRoomsResources`— Otorga acceso para pasar AWS Clean Rooms recursos específicos.

Para ver una lista en JSON de los detalles de la política, consulta AWSCleanHabitaciones MLRead OnlyAccess en la Guía de referencia sobre políticas AWS gestionadas.

AWS política gestionada: **AWSCleanRoomsMLFullAccess**

Puede adjuntar `AWSCleanRoomsMLFullAcces` a sus entidades principales de IAM. Esta política concede permisos administrativos que permiten el acceso total (lectura, escritura y actualización) a los recursos y metadatos necesarios de Clean Rooms ML.

Detalles de los permisos

Esta política incluye los permisos siguientes:

- `CleanRoomsMLFullAccess`: concede acceso a todas las acciones de Clean Rooms ML.
- `PassServiceRole`: concede acceso para transferir un rol de servicio únicamente al servicio (condición `PassedToService`) cuyo nombre contenga "cleanrooms-ml".
- `CleanRoomsConsoleNavigation`— Otorga acceso para ver las pantallas de la AWS Clean Rooms consola.
- `CollaborationMembershipCheck`— Al iniciar un trabajo de generación de audiencia (segmento similar) dentro de una colaboración, el servicio Clean Rooms ML llama `ListMembers` para comprobar que la colaboración es válida, que la persona que llama es un miembro activo y que el propietario del modelo de audiencia configurado es un miembro activo. Este permiso siempre es obligatorio; el SID de navegación de la consola solo es necesario para los usuarios de la consola.
- `PassCleanRoomsResources`— Otorga acceso para transferir recursos específicos AWS Clean Rooms .
- `AssociateModels`: permite a las entidades principales asociar un modelo de Clean Rooms ML a la colaboración.
- `TagAssociations`: permite a las entidades principales añadir etiquetas a la asociación entre un modelo similar y una colaboración.
- `ListRolesToPickServiceRole`— Permite a los directores enumerar todas sus funciones para poder elegir una función de servicio cuando la utilicen AWS Clean Rooms.
- `GetRoleAndListRolePoliciesToInspectServiceRole`: permite a las entidades principales ver el rol de servicio y la política correspondiente en IAM.
- `ListPoliciesToInspectServiceRolePolicy`: permite a las entidades principales ver el rol de servicio y la política correspondiente en IAM.

- `GetPolicyToInspectServiceRolePolicy`: permite a las entidades principales ver el rol de servicio y la política correspondiente en IAM.
- `ConsoleDisplayTables`— Permite a los directores acceder de solo lectura a los AWS Glue metadatos necesarios para mostrar los datos sobre AWS Glue las tablas subyacentes de la consola.
- `ConsolePickOutputBucket`: permite a las entidades principales seleccionar buckets de Amazon S3 para las salidas configuradas del modelo de audiencia.
- `ConsolePickS3Location`: permite a las entidades principales seleccionar la ubicación dentro de un bucket para los resultados configurados del modelo de audiencia.
- `ConsoleDescribeECRRepositories`— Permite a los directores describir los repositorios e imágenes de Amazon ECR.

Para ver una lista en JSON de los detalles de la política, consulte AWSCleanRooms MLFull Access en la Guía de referencia de políticas AWS gestionadas.

AWS Clean Rooms actualizaciones de las políticas AWS gestionadas

Consulte los detalles sobre las actualizaciones de las políticas AWS administradas AWS Clean Rooms desde que este servicio comenzó a rastrear estos cambios. Para recibir alertas automáticas sobre los cambios en esta página, suscríbase a la fuente RSS de la página del historial del AWS Clean Rooms documento.

Cambio	Descripción	Fecha
AWSCleanRoomsMLReadOnlyAccess: actualización de una política actual AWSCleanRoomsMLFullAccess: actualización de una política actual	Se agregó PassCleanRoomsResources a AWSCleanRoomsMLReadOnlyAccess. Se agregaron PassCleanRoomsResources y ConsoleDescribeECRRepositories a AWSCleanRoomsMLFullAccess.	10 de enero de 2025
AWSCleanRoomsFullAccessNoQuerying: actualización de una política actual	Se agregó cleanrooms:BatchGetSchemaAnalysisRule a CleanRoomsAccess.	13 de mayo de 2024
AWSCleanRoomsFullAccess: actualización de una política actual	Se ha actualizado el ID de la declaración en AWSCleanRoomsFullAccess	21 de marzo de 2024

Cambio	Descripción	Fecha
	de ConsolePickQueryResultsBucket a SetQueryResultsBucket en esta política para representar mejor los permisos, ya que los permisos son necesarios para configurar el bucket de resultados de las consultas con y sin la consola.	
AWSCleanRoomsMLReadOnlyAccess: política nueva AWSCleanRoomsMLFullAccess: política nueva	Se ha añadido AWSCleanRoomsMLReadOnlyAccess y AWSCleanRoomsMLFullAccess para admitir AWS Clean Rooms ML.	29 de noviembre de 2023
AWSCleanRoomsFullAccessNoQuerying: actualización de una política actual	Se ha agregado cleanrooms:CreateAnalysisTemplate, cleanrooms:GetAnalysisTemplate, cleanrooms:UpdateAnalysisTemplate, cleanrooms:DeleteAnalysisTemplate, cleanrooms:ListAnalysisTemplates, cleanrooms:GetCollaborationAnalysisTemplate, cleanrooms:BatchGetCollaborationAnalysisTemplate y cleanrooms:ListCollaborationAnalysisTemplates a CleanRoomsAccess para habilitar la nueva característica de plantillas de análisis.	31 de julio de 2023
AWSCleanRoomsFullAccessNoQuerying: actualización de una política actual	Se ha añadido cleanrooms:ListTagsForResource, cleanrooms:UntagResource y cleanrooms:TagResource a CleanRoomsAccess para habilitar el etiquetado de recursos.	21 de marzo de 2023
AWS Clean Rooms comenzó a rastrear los cambios	AWS Clean Rooms comenzó a rastrear los cambios de sus políticas AWS gestionadas.	12 de enero de 2023

Solución de problemas AWS Clean Rooms de identidad y acceso

Utilice la siguiente información como ayuda para diagnosticar y solucionar los problemas habituales que pueden surgir al trabajar con un AWS Clean Rooms IAM.

Temas

- No estoy autorizado a realizar ninguna acción en AWS Clean Rooms
- No estoy autorizado a realizar tareas como: PassRole
- Quiero permitir que personas ajenas a mí accedan Cuenta de AWS a mis AWS Clean Rooms recursos

No estoy autorizado a realizar ninguna acción en AWS Clean Rooms

Si recibe un error que indica que no tiene autorización para realizar una acción, las políticas se deben actualizar para permitirle realizar la acción.

En el siguiente ejemplo, el error se produce cuando el usuario de IAM `mateojackson` intenta utilizar la consola para consultar los detalles acerca de un recurso ficticio *`my-example-widget`*, pero no tiene los permisos ficticios `cleanrooms:`*`GetWidget`*.

```
User: arn:aws:iam::123456789012:user/mateojackson is not authorized to perform:
 cleanrooms:GetWidget on resource: my-example-widget
```

En este caso, la política de Mateo se debe actualizar para permitirle acceder al recurso *`my-example-widget`* mediante la acción `cleanrooms:`*`GetWidget`*.

Si necesita ayuda, póngase en contacto con su AWS administrador. El gestionador es la persona que le proporcionó las credenciales de inicio de sesión.

No estoy autorizado a realizar tareas como: PassRole

Si recibe un error que indica que no tiene autorización para realizar la acción `iam:PassRole`, las políticas deben actualizarse a fin de permitirle pasar un rol a AWS Clean Rooms.

Algunas Servicios de AWS permiten transferir una función existente a ese servicio en lugar de crear una nueva función de servicio o una función vinculada a un servicio. Para ello, debe tener permisos para transferir el rol al servicio.

En el siguiente ejemplo, el error se produce cuando un usuario de IAM denominado `marymajor` intenta utilizar la consola para realizar una acción en AWS Clean Rooms. Sin embargo, la acción requiere que el servicio cuente con permisos que otorguen un rol de servicio. Mary no tiene permisos para transferir el rol al servicio.

```
User: arn:aws:iam::123456789012:user/marymajor is not authorized to perform:
 iam:PassRole
```

En este caso, las políticas de Mary se deben actualizar para permitirle realizar la acción `iam:PassRole`.

Si necesita ayuda, póngase en contacto con su administrador. AWS El gestionador es la persona que le proporcionó las credenciales de inicio de sesión.

Quiero permitir que personas ajenas a mí accedan Cuenta de AWS a mis AWS Clean Rooms recursos

Puede crear un rol que los usuarios de otras cuentas o las personas externas a la organización puedan utilizar para acceder a sus recursos. Puede especificar una persona de confianza para que asuma el rol.

Para obtener más información, consulte lo siguiente:

- Para saber si AWS Clean Rooms es compatible con estas funciones, consulte¿Cómo AWS Clean Rooms funciona con IAM.
- Para obtener información sobre cómo proporcionar acceso a los recursos de su Cuentas de AWS propiedad, consulte Proporcionar acceso a un usuario de IAM en otro usuario de su propiedad Cuenta de AWS en la Guía del usuario de IAM.
- Para obtener información sobre cómo proporcionar acceso a tus recursos a terceros Cuentas de AWS, consulta Cómo proporcionar acceso a recursos que Cuentas de AWS son propiedad de terceros en la Guía del usuario de IAM.
- Para obtener información sobre cómo proporcionar acceso mediante una federación de identidades, consulte Proporcionar acceso a usuarios autenticados externamente (identidad federada) en la Guía del usuario de IAM.
- Para obtener información sobre la diferencia entre los roles y las políticas basadas en recursos para el acceso entre cuentas, consulte Cómo los roles de IAM difieren de las políticas basadas en recursos en la Guía del usuario de IAM.

Prevención de la sustitución confusa entre servicios

El problema de la sustitución confusa es un problema de seguridad en el que una entidad que no tiene permiso para realizar una acción puede obligar a una entidad con más privilegios a realizar la acción. En AWS, la suplantación de identidad entre servicios puede provocar el confuso problema de un diputado. La suplantación entre servicios puedes producirse cuando un servicio (el servicio que lleva a cabo las llamadas) llama a otro servicio (el servicio al que se llama). El servicio que lleva a cabo las llamadas se puedes manipular para utilizar sus permisos a fin de actuar en función de los recursos de otro cliente de una manera en la que no debe tener permiso para acceder. Para evitarlo, AWS proporciona herramientas que lo ayudan a proteger sus datos para todos los servicios con entidades principales de servicio a las que se les ha dado acceso a los recursos de su cuenta.

Se recomienda utilizar las claves de contexto de condición global `aws:SourceArn` en las políticas de recursos para limitar los permisos que AWS Clean Rooms concede a otro servicio para el recurso. Utiliza `aws:SourceArn` si desea que solo se asocie un recurso al acceso entre servicios.

La forma más eficaz de protegerse contra el problema de la sustitución confusa es utilizar la clave de contexto de condición global de `aws:SourceArn` con el ARN completo del recurso. En AWSClean Rooms, también tienes que compararla con la clave de condición. `sts:ExternalId`

El valor de `aws:SourceArn` debe definirse con el ARN de pertenencia del rol asumido.

El siguiente ejemplo muestra cómo usar la clave de contexto de condición global `aws:SourceArn` en AWS Clean Rooms para evitar el problema del suplente confuso.

> **Note**
>
> La política de ejemplo se aplica a la política de confianza del rol de servicio que utiliza AWS Clean Rooms para acceder a los datos de los clientes.
> El valor de *membershipID* es su ID de AWSClean Rooms miembro en la colaboración.

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

```
            },
            "Action": "sts:AssumeRole",
            "Condition": {
                "StringLike": {
                    "sts:ExternalId": "arn:aws:*:aws-region:*:dbuser:*/membershipID*"
                }
            }
        },
        {
            "Sid": "AllowIfSourceArnMatches",
            "Effect": "Allow",
            "Principal": {
                "Service": "cleanrooms.amazonaws.com"
            },
            "Action": "sts:AssumeRole",
            "Condition": {
                "ForAnyValue:ArnEquals": {
                    "aws:SourceArn": "arn:aws:cleanrooms:aws-
region:123456789012:membership/membershipID"
                }
            }
        }
    ]
}
```

Comportamientos de IAM para ML AWS Clean Rooms

Trabajos entre cuentas

Clean Rooms ML permite que otra Cuenta de AWS persona acceda de forma segura Cuenta de AWS a determinados recursos creados por una persona en su cuenta. Cuando un cliente de Cuenta de AWS A recurre `StartAudienceGenerationJob` a un `ConfiguredAudienceModel` recurso propiedad de Cuenta de AWS B, Clean Rooms ML crea dos ARNs para esa tarea. Un ARN en Cuenta de AWS A y otro en B. Cuenta de AWS ARNs Son idénticos excepto por sus. Cuenta de AWS

Clean Rooms ML crea dos ARNs para el trabajo a fin de garantizar que ambas cuentas puedan aplicar sus propias políticas de IAM a los trabajos. Por ejemplo, ambas cuentas pueden usar el control de acceso basado en etiquetas y aplicar las políticas de su AWS organización. El trabajo procesa los datos de ambas cuentas, de modo que ambas cuentas pueden eliminar el trabajo y sus datos asociados. Ninguna de las dos cuentas puede impedir que la otra elimine el trabajo.

Solo hay una ejecución de trabajo y ambas cuentas pueden ver la tarea cuando llaman a `ListAudienceGenerationJobs`. Ambas cuentas pueden llamar al `GetDelete`, y `Export` APIs estar trabajando usando el ARN con su propia Cuenta de AWS identificación.

Ninguno de los dos Cuenta de AWS puede acceder al trabajo cuando se utiliza un ARN con el otro Cuenta de AWS ID.

El nombre del trabajo debe ser único dentro de una Cuenta de AWS. El nombre en Cuenta de AWS B es*$accountA-$name*. El nombre elegido por Cuenta de AWS A lleva el prefijo Cuenta de AWS A cuando el trabajo se visualiza en B. Cuenta de AWS

Para que una cuenta cruzada `StartAudienceGenerationJob` tenga éxito, Cuenta de AWS B debe permitir esa acción tanto en el nuevo trabajo de B como `ConfiguredAudienceModel` en el de Cuenta de AWS Cuenta de AWS B mediante una política de recursos similar a la del siguiente ejemplo:

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "Clean-Rooms-<CAMA ID>",
            "Effect": "Allow",
            "Principal": {
                "AWS": [
                    "accountA"
                ]
            },
            "Action": [
                "cleanrooms-ml:StartAudienceGenerationJob"
            ],
            "Resource": [
                "arn:aws:cleanrooms-ml:us-west-1:AccountB:configured-audience-
model/id",
                "arn:aws:cleanrooms-ml:us-west-1:AccountB:audience-generation-job/*"
            ],
            // optional - always set by AWS Clean Rooms
"Condition":{"StringEquals":{"cleanrooms-ml:CollaborationId":"UUID"}}
        }
    ]
}
```

Si utilizas la API de AWS Clean Rooms ML para crear un modelo similar `manageResourcePolicies` configurado con el valor true, AWS Clean Rooms crea esta política automáticamente.

Además, la política de identidad de la persona que llama en Cuenta de AWS A necesita `StartAudienceGenerationJob` permiso. `arn:aws:cleanrooms-ml:us-west-1:AccountA:audience-generation-job/*` Por lo tanto, hay tres recursos de acción de IAM`StartAudienceGenerationJob`: el trabajo Cuenta de AWS A, el trabajo Cuenta de AWS B y el Cuenta de AWS trabajo B. `ConfiguredAudienceModel`

> **Warning**
>
> El Cuenta de AWS que inició el trabajo recibe un evento de registro de AWS CloudTrail auditoría sobre el trabajo. La Cuenta de AWS propietaria de `ConfiguredAudienceModel` no recibe ningún evento de registro de auditoría de AWS CloudTrail .

Etiquetado de trabajos

Al establecer el parámetro `childResourceTagOnCreatePolicy=FROM_PARENT_RESOURCE` de `CreateConfiguredAudienceModel`, todos los trabajos de generación de segmentos similares de su cuenta que se creen a partir de ese modelo similar configurado tendrán de forma predeterminada las mismas etiquetas que el modelo similar configurado. El modelo similar configurado es el principal y el trabajo de generación de segmentos similares es el secundario.

Si está creando un trabajo en su propia cuenta, las etiquetas de solicitud del trabajo anulan las etiquetas principales. Los trabajos creados por otras cuentas nunca crean etiquetas en su cuenta. Si establece `childResourceTagOnCreatePolicy=FROM_PARENT_RESOURCE` y otra cuenta crea un trabajo, hay dos copias del trabajo. La copia de su cuenta tiene las etiquetas del recurso principal y la copia de la cuenta del remitente del trabajo tiene las etiquetas de la solicitud.

Validación de colaboradores

Al conceder permisos a otros miembros de una AWS Clean Rooms colaboración, la política de recursos debe incluir la clave de condición`cleanrooms-ml:CollaborationId`. Esto obliga a incluir el `collaborationId` parámetro en la StartAudienceGenerationJobsolicitud. Cuando el parámetro `collaborationId` se incluye en la solicitud, Clean Rooms ML valida que la colaboración existe, que el remitente del trabajo es un miembro activo de la colaboración y que el propietario del modelo similar configurado es un miembro activo de la colaboración.

Cuando AWS Clean Rooms administre la política de recursos del modelo similar configurada (el `manageResourcePolicies` parámetro está `TRUE` en la CreateConfiguredAudienceModelAssociation solicitud), esta clave de condición se establecerá en la política de recursos. Por lo tanto, debe especificar la entrada`collaborationId`. StartAudienceGenerationJob

Acceso entre cuentas

Solo se puede llamar a `StartAudienceGenerationJob` entre cuentas. El resto de Clean Rooms ML solo se APIs pueden utilizar con los recursos de su propia cuenta. Esto garantiza que sus datos de entrenamiento, la configuración del modelo similar y otra información permanezcan privados.

Clean Rooms ML nunca revela Amazon S3 ni AWS Glue las ubicaciones de las cuentas. La ubicación de los datos de entrenamiento, la ubicación de salida del modelo similar configurado y la ubicación inicial del trabajo de generación de segmentos similar nunca están visibles entre cuentas. A menos que el registro de consultas esté habilitado en la colaboración, las cuentas no pueden ver si los datos iniciales provienen de una consulta SQL o la consulta en sí. Si `Get` un trabajo de generación de audiencia enviado por otra cuenta, el servicio no mostrará la ubicación inicial.

Comportamientos de IAM para los modelos personalizados de Clean Rooms ML

Trabajos entre cuentas

Clean Rooms ML permite que otra Cuenta de AWS persona acceda de forma segura a determinados recursos asociados Cuenta de AWS a una colaboración creada por una persona desde su cuenta. Un cliente de Cuenta de AWS A con la capacidad de realizar consultas como miembro puede llamar `CreateTrainedModel` o utilizar `StartTrainedModelInferenceJob` un `ConfiguredModelAlgorithmAssociation` recurso propiedad de otro miembro de la colaboración, siempre que lo permita la `ConfiguredModelAlgorithmAssociation` regla de análisis personalizada creada con ella`CreateConfiguredTableAnalysisRule`. `CreateMLInputChannel`

Además, cualquier miembro activo de una colaboración puede eliminar los datos asociados a un modelo entrenado o a un canal de entrada de aprendizaje automático mediante el comando `DeleteTrainedModelOutput` and `DeleteMLInputChannelData` APIs.

Acceso entre cuentas

Clean Rooms ML permite a los usuarios recuperar metadatos sobre los recursos creados por otras cuentas a través de `GetCollaboration` y `ListCollaboration` APIs. Clean Rooms ML no revela la clave del KMS ARNs, las etiquetas, las variables de entorno ni los hiperparámetros (para la `TrainedModel` acción) a otras cuentas.

Acceso a membresías y colaboraciones

Al acceder a los recursos de membresía y colaboración en el contexto de los modelos personalizados de Clean Rooms ML, la política de identidad del usuario necesita permisos para las acciones `cleanrooms:PassMembershipcleanrooms:PassCollaboration`, o ambos. Todos los APIs que aceptan `membershipId` necesitan el `cleanrooms:PassMembership` permiso y todos los APIs que aceptan `collaborationId` necesitan el `cleanrooms:PassCollaboration` permiso. Se proporciona un ejemplo de política de identidad para un rol al que se puede llamar `createTrainedModel` en el contexto de un ID de membresía y al que se puede llamar `GetCollaborationTrainedModel` en el contexto de un ID de colaboración.

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "AllowCleanroomsMLActions",
            "Effect": "Allow",
            "Action": [
                "cleanrooms-ml:PassMembership",
                "cleanrooms-ml:PassCollaboration",
            ],
            "Resource": ["*"]
        },
        {
            "Sid": "AllowMembership",
            "Effect": "Allow",
            "Action": [
                "cleanrooms-ml:PassMembership",
            ],
            "Resource": ["arn:aws:cleanrooms:region:account:membership/memberId"]
        },
        {
            "Sid": "AllowCollaboration",
            "Effect": "Allow",
```

```
            "Action": [
                "cleanrooms-ml:PassCollaboration",
            ],
            "Resource":
 ["arn:aws:cleanrooms:region:account:collaboration/collaborationId"]
        }
    ]
}
```

Validación de conformidad para AWS Clean Rooms

Para saber si uno Servicio de AWS está dentro del ámbito de aplicación de programas de cumplimiento específicos, consulte Servicios de AWS Alcance por programa de cumplimiento Servicios de AWS de cumplimiento y elija el programa de cumplimiento que le interese. Para obtener información general, consulte Programas de AWS cumplimiento > Programas AWS .

Puede descargar informes de auditoría de terceros utilizando AWS Artifact. Para obtener más información, consulte Descarga de informes en AWS Artifact .

Su responsabilidad de cumplimiento al Servicios de AWS utilizarlos viene determinada por la confidencialidad de sus datos, los objetivos de cumplimiento de su empresa y las leyes y reglamentos aplicables. AWS proporciona los siguientes recursos para ayudar con el cumplimiento:

- Cumplimiento de seguridad y gobernanza: en estas guías se explican las consideraciones de arquitectura y se proporcionan pasos para implementar las características de seguridad y cumplimiento.
- Referencia de servicios válidos de HIPAA: muestra una lista con los servicios válidos de HIPAA. No todos Servicios de AWS cumplen con los requisitos de la HIPAA.
- AWS Recursos de de cumplimiento: esta colección de libros de trabajo y guías puede aplicarse a su industria y ubicación.
- AWS Guías de cumplimiento para clientes: comprenda el modelo de responsabilidad compartida desde el punto de vista del cumplimiento. Las guías resumen las mejores prácticas para garantizar la seguridad Servicios de AWS y orientan los controles de seguridad en varios marcos (incluidos el Instituto Nacional de Estándares y Tecnología (NIST), el Consejo de Normas de Seguridad del Sector de Tarjetas de Pago (PCI) y la Organización Internacional de Normalización (ISO)).
- Evaluación de los recursos con reglas en la guía para AWS Config desarrolladores: el AWS Config servicio evalúa en qué medida las configuraciones de los recursos cumplen con las prácticas internas, las directrices del sector y las normas.

- AWS Security Hub— Esto Servicio de AWS proporciona una visión completa del estado de su seguridad interior AWS. Security Hub utiliza controles de seguridad para evaluar sus recursos de AWS y comprobar su cumplimiento con los estándares y las prácticas recomendadas del sector de la seguridad. Para obtener una lista de los servicios y controles compatibles, consulte la Referencia de controles de Security Hub.
- Amazon GuardDuty: Servicio de AWS detecta posibles amenazas para sus cargas de trabajo Cuentas de AWS, contenedores y datos mediante la supervisión de su entorno para detectar actividades sospechosas y maliciosas. GuardDuty puede ayudarlo a cumplir con varios requisitos de conformidad, como el PCI DSS, al cumplir con los requisitos de detección de intrusiones exigidos por ciertos marcos de cumplimiento.
- AWS Audit Manager— Esto le Servicio de AWS ayuda a auditar continuamente su AWS uso para simplificar la gestión del riesgo y el cumplimiento de las normativas y los estándares del sector.

Resiliencia en AWS Clean Rooms

La infraestructura AWS global se basa en AWS regiones y zonas de disponibilidad. Las regiones proporcionan varias zonas de disponibilidad físicamente independientes y aisladas que se encuentran conectadas mediante redes con un alto nivel de rendimiento y redundancia, además de baja demora. Con las zonas de disponibilidad, puede diseñar y utilizar aplicaciones y bases de datos que realizan una conmutación por error automática entre las zonas sin interrupciones. Las zonas de disponibilidad tienen una mayor disponibilidad, tolerancia a errores y escalabilidad que las infraestructuras tradicionales de uno o varios centros de datos.

Para obtener más información sobre AWS las regiones y las zonas de disponibilidad, consulte Infraestructura AWS global.

Seguridad de la infraestructura en AWS Clean Rooms

Como servicio gestionado, AWS Clean Rooms está protegido por la seguridad de la red AWS global. Para obtener información sobre los servicios AWS de seguridad y cómo se AWS protege la infraestructura, consulte Seguridad AWS en la nube. Para diseñar su AWS entorno utilizando las mejores prácticas de seguridad de la infraestructura, consulte Protección de infraestructuras en un marco de buena AWS arquitectura basado en el pilar de la seguridad.

Utiliza las llamadas a la API AWS publicadas para acceder a AWS Clean Rooms través de la red. Los clientes deben admitir lo siguiente:

- Seguridad de la capa de transporte (TLS). Exigimos TLS 1.2 y recomendamos TLS 1.3.
- Conjuntos de cifrado con confidencialidad directa total (PFS) como DHE (Ephemeral Diffie-Hellman) o ECDHE (Elliptic Curve Ephemeral Diffie-Hellman). La mayoría de los sistemas modernos como Java 7 y posteriores son compatibles con estos modos.

Además, las solicitudes deben estar firmadas mediante un ID de clave de acceso y una clave de acceso secreta que esté asociada a una entidad principal de IAM. También puedes utilizar AWS Security Token Service (AWS STS) para generar credenciales de seguridad temporales para firmar solicitudes.

Seguridad de la red

Cuando se AWS Clean Rooms lee desde el bucket de S3 durante la ejecución de la consulta, el tráfico entre Amazon S3 AWS Clean Rooms y Amazon S3 se enruta de forma segura a través de la red AWS privada. El tráfico en vuelo se firma mediante el protocolo Amazon Signature versión 4 (SIGv4) y se cifra mediante HTTPS. Este tráfico se autoriza en función del rol de servicio de IAM que se haya configurado para la tabla configurada.

Puede conectarse mediante programación a AWS Clean Rooms través de un punto final. Para obtener una lista de puntos de conexión de servicio, consulte Puntos de conexión y cuotas de AWS Clean Rooms en la Referencia general de AWS.

Todos los puntos de conexión de servicio son únicamente HTTPS. Puede utilizar los puntos de conexión de Amazon Virtual Private Cloud (VPC) en caso de que desee conectarse desde AWS Clean Rooms su VPC y no desee tener conectividad a Internet. Para obtener más información, consulte Acceder a AWS los servicios AWS PrivateLink en la Guía.AWS PrivateLink

Puede asignar políticas de IAM a sus directores de IAM que utilicen las claves aws: SourceVpce context para restringir a su director de IAM a fin de que solo pueda realizar llamadas a través de AWS Clean Rooms un punto final de VPC y no a través de Internet.

Acceda al aprendizaje AWS Clean Rooms automático mediante AWS Clean Rooms un punto final de interfaz ()AWS PrivateLink

Puede usarlo AWS PrivateLink para crear una conexión privada entre su nube privada virtual (VPC) y ML AWS Clean Rooms . AWS Clean Rooms Puede acceder a nuestro AWS Clean Rooms AWS Clean Rooms ML como si estuviera en su VPC, sin el uso de una puerta de enlace a Internet, un

dispositivo NAT, una conexión VPN o AWS Direct Connect una conexión. Las instancias de la VPC no necesitan direcciones IP públicas para acceder a AWS Clean Rooms.

Esta conexión privada se establece mediante la creación de un punto de conexión de interfaz alimentado por AWS PrivateLink. Creamos una interfaz de red de punto de conexión en cada subred habilitada para el punto de conexión de interfaz. Se trata de interfaces de red administradas por el solicitante que sirven como punto de entrada para el tráfico destinado a AWS Clean Rooms.

Para obtener más información, consulte Acceso a los Servicios de AWS a través de AWS PrivateLink en la Guía de AWS PrivateLink .

Consideraciones sobre AWS Clean Rooms

Antes de configurar un punto final de interfaz para AWS Clean Rooms, consulte las consideraciones de la AWS PrivateLink guía.

AWS Clean Rooms y AWS Clean Rooms ML permiten realizar llamadas a todas sus acciones de API a través del punto final de la interfaz.

Las políticas de puntos de conexión de VPC no son compatibles con el ML AWS Clean Rooms . AWS Clean Rooms De forma predeterminada, se permite el acceso total AWS Clean Rooms y el AWS Clean Rooms aprendizaje automático a través del punto final de la interfaz. Como alternativa, puede asociar un grupo de seguridad a las interfaces de red de los puntos finales para controlar el tráfico hacia AWS Clean Rooms o el aprendizaje AWS Clean Rooms automático a través del punto final de la interfaz.

Cree un punto final de interfaz para AWS Clean Rooms

Puede crear un punto de enlace de interfaz para AWS Clean Rooms el AWS Clean Rooms aprendizaje automático mediante la consola de Amazon VPC o el AWS Command Line Interface ()AWS CLI. Para obtener más información, consulte Creación de un punto de conexión de interfaz en la Guía de AWS PrivateLink .

Cree un punto final de interfaz para AWS Clean Rooms usar el siguiente nombre de servicio.

```
com.amazonaws.region.cleanrooms
```

Cree un punto final de interfaz para AWS Clean Rooms ML con el siguiente nombre de servicio.

```
com.amazonaws.region.cleanrooms-ml
```

Si habilita DNS privado para el punto de conexión de interfaz, puede realizar solicitudes a la API para AWS Clean Rooms usando su nombre de DNS predeterminado para la región. Por ejemplo, `cleanrooms-ml.us-east-1.amazonaws.com`.

Monitorización AWS Clean Rooms

La supervisión es una parte importante del mantenimiento de la confiabilidad, la disponibilidad y el rendimiento de AWS Clean Rooms AWS las demás soluciones. AWS proporciona las siguientes herramientas de monitoreo para observar AWS Clean Rooms, informar cuando algo anda mal y tomar medidas automáticas cuando sea apropiado:

- Amazon CloudWatch Logs le permite supervisar, almacenar y acceder a sus archivos de registro desde EC2 instancias de Amazon y otras fuentes. AWS CloudTrail Amazon CloudWatch Logs puede supervisar la información de los archivos de registro y notificarle cuando se alcancen determinados umbrales. También se pueden archivar los datos del registro en un almacenamiento de larga duración. Para obtener más información, consulta la Guía del usuario CloudWatch de Amazon Logs.

  Clean Rooms ML permite realizar trabajos entre cuentas para determinadas acciones de la API. La Cuenta de AWS persona que inició el trabajo recibe el evento del registro de AWS CloudTrail auditoría del trabajo. Para obtener más información, consulte Comportamientos de IAM para ML AWS Clean Rooms
- AWS CloudTrailcaptura las llamadas a la API y los eventos relacionados realizados por usted o en su nombre Cuenta de AWS y entrega los archivos de registro a un bucket de Amazon S3 que especifique. Puede identificar qué usuarios y cuentas llamaron AWS, la dirección IP de origen desde la que se realizaron las llamadas y cuándo se produjeron. Para obtener más información, consulte la Guía del usuario de AWS CloudTrail.

Inicio de sesión de Analysis AWS Clean Rooms

El registro de análisis es una función de AWS Clean Rooms. Al crear una colaboración y activar el registro de análisis, los miembros pueden almacenar los registros relevantes de las consultas o los registros de los trabajos en Amazon CloudWatch Logs.

Con los registros de consultas y los registros de trabajos, los miembros pueden determinar si las consultas cumplen con las reglas de análisis y se ajustan al acuerdo de colaboración. Además, los registros de consultas respaldan las auditorías.

Cuando la opción de registro de análisis está activada en la AWS Clean Rooms consola, los registros de consultas incluyen lo siguiente:

- `analysisRule`: la regla de análisis de la tabla configurada.
- `analysisTemplateArn`: la plantilla de análisis que se ejecutó (aparece en función de la regla de análisis).
- `collaborationId`: el identificador único de la colaboración en la que se ejecutó la consulta.
- `configuredTableID`: el identificador único de la tabla configurada a la que hace referencia la consulta.
- `directQueryAnalysisRulePolicy.custom.allowedAnalysis`: la plantilla de análisis que se puede ejecutar en una tabla configurada (aparece según la regla de análisis).
- `directQueryAnalysisRulePolicy.v1.custom.allowedAnalysisProviders`: los proveedores de consultas que tienen permitido crear consultas (aparece según la regla de análisis).
- `errorCode`: el código de error que se produce cuando una consulta no se ejecuta correctamente.
- `errorMessage`: el mensaje de error que se produce cuando una consulta no se ejecuta correctamente.
- `eventID`: el identificador único de ejecución de la consulta. Desde el 31 de agosto de 2023, el identificador único coincidirá con el `protectedQueryID`.
- `eventTimestamp`: el momento de ejecución de la consulta.
- `parameters.parametervalue`: los valores de los parámetros (aparecen en función del texto de la consulta).
- `queryText`: la definición SQL de la ejecución de la consulta. Si hay parámetros, se etiquetan como `:parametervalue`.
- `queryValidationErrors`: los errores de consulta durante la validación de la consulta.
- `schemaName`: el nombre de la asociación de tablas configuradas a la que se hace referencia en la consulta.
- `status`: el estado de la ejecución de la consulta.

Reciba registros de consultas y trabajos

No necesita realizar ninguna acción aparte de AWS Clean Rooms configurar los registros de consultas y trabajos. AWS Clean Rooms crea grupos de registro para las colaboraciones después de que cada miembro de la colaboración cree una membresía.

Los miembros que puedan realizar consultas, los miembros que puedan ejecutar consultas y trabajos, los miembros que puedan recibir resultados y los miembros cuyas tablas de configuración estén referenciadas en la consulta recibirán un registro de consultas o un registro de trabajos.

El miembro que puede realizar consultas y el miembro que puede recibir los resultados recibirán registros de consultas para cada tabla configurada a la que se haga referencia en la consulta. Si no son el propietario de la tabla configurada, no podrán ver el ID de la tabla configurada (`configuredTableID`).

El miembro que pueda ejecutar consultas y trabajos y el miembro que pueda recibir los resultados recibirán los registros de trabajos de cada tabla configurada a la que se haga referencia en el trabajo. Si no son el propietario de la tabla configurada, no podrán ver el ID de la tabla configurada (`configuredTableID`).

Si un miembro tiene varias asociaciones de tablas configuradas a las que se hace referencia en la consulta, recibirá un registro de consultas por cada tabla configurada.

Si un miembro tiene varias asociaciones de tablas configuradas a las que se hace referencia en el trabajo, recibirá un registro de trabajos para cada tabla configurada.

Se crean registros para las consultas que contienen SQL admitido y no admitido en AWS Clean Rooms. Para obtener más información, consulte Referencia de SQL de AWS Clean Rooms.

Los registros también se crean cuando las consultas o los trabajos hacen referencia a tablas configuradas que no están asociadas a la colaboración.

Los registros no se crean si el código SQL no es correcto AWS Clean Rooms.

Los registros de consultas y trabajos indican el estado de una consulta, pero no indican si el resultado de la consulta se entregó. Confirman que el miembro que puede realizar la consulta ha enviado una consulta o un trabajo. Los registros de consultas también confirman que la consulta contiene SQL compatible AWS Clean Rooms y hace referencia a las tablas configuradas asociadas a la colaboración.

Example

Por ejemplo, no se genera un registro si la consulta se canceló después de AWS Clean Rooms validar su conformidad con las reglas de análisis y durante el procesamiento de la consulta.

Si elimina el grupo de registros, debe volver a crearlo manualmente con el mismo nombre de grupo de registros (ID de colaboración de la colaboración). O bien, puede activar y desactivar el registro en su pertenencia.

Para obtener más información sobre cómo activar el registro de análisis, consulteCrear una colaboración.

Para obtener más información sobre Amazon CloudWatch Logs, consulta la Guía del usuario de Amazon CloudWatch Logs.

Acciones recomendadas para los registros de consultas y trabajos

Es aconsejable que los miembros tomen las siguientes medidas de forma periódica:

- Para comprobar que las consultas y los trabajos coinciden con los casos de uso o las consultas acordados para la colaboración, revise las consultas y los trabajos que se ejecutan en la colaboración.

  Para obtener más información sobre cómo ver consultas recientes, consulte Visualización de consultas recientes.

  Para obtener más información sobre cómo ver los trabajos recientes, consulteVer trabajos recientes.
- Revisar las columnas de la tabla configurada que se utilizan en las reglas de análisis de los miembros de la colaboración y en las consultas para asegurarse de que las columnas de la tabla configurada se ajustan a lo acordado para la colaboración.

  Para obtener más información sobre cómo ver las columnas configuradas, consulte Ver tablas y reglas de análisis.

Registro de llamadas a la AWS Clean Rooms API mediante AWS CloudTrail

AWS Clean Rooms está integrado con AWS CloudTrail un servicio que proporciona un registro de las acciones realizadas por un usuario, un rol o un Servicio de AWS usuario AWS Clean Rooms. CloudTrail captura todas las llamadas a la API AWS Clean Rooms como eventos. Las llamadas capturadas incluyen llamadas desde la AWS Clean Rooms consola y llamadas en código a las operaciones de la AWS Clean Rooms API. Si crea una ruta, puede habilitar la entrega continua de CloudTrail eventos a un bucket de Amazon S3, incluidos los eventos para AWS Clean Rooms. Si no configura una ruta, podrá ver los eventos más recientes en la CloudTrail consola, en el historial de eventos. Con la información recopilada por usted CloudTrail, puede determinar a AWS Clean Rooms qué dirección IP se realizó la solicitud, quién la realizó, cuándo se realizó y detalles adicionales.

Para obtener más información CloudTrail, consulte la Guía AWS CloudTrail del usuario.

AWS Clean Rooms información en CloudTrail

CloudTrail está habilitada en tu cuenta Cuenta de AWS al crear la cuenta. Cuando se produce una actividad en AWS Clean Rooms, esa actividad se registra en un CloudTrail evento junto con otros Servicio de AWS eventos del historial de eventos. Puede ver, buscar y descargar eventos recientes en su Cuenta de AWS. Para obtener más información, consulte Visualización de eventos con el historial de CloudTrail eventos.

Para tener un registro continuo de tus eventos Cuenta de AWS, incluidos los eventos para AWS Clean Rooms ti, crea una ruta. Un rastro permite CloudTrail entregar archivos de registro a un bucket de Amazon S3. De forma predeterminada, cuando se crea un registro de seguimiento en la consola, el registro de seguimiento se aplica a todas las Regiones de AWS. La ruta registra los eventos de todas las regiones de la AWS partición y envía los archivos de registro al bucket de Amazon S3 que especifique. Además, puede configurar otros Servicios de AWS para que analicen más a fondo los datos de eventos recopilados en los CloudTrail registros y actúen en función de ellos. Para más información, consulte los siguientes temas:

- Introducción a la creación de registros de seguimiento
- CloudTrail servicios e integraciones compatibles
- Configuración de las notificaciones de Amazon SNS para CloudTrail
- Recibir archivos de CloudTrail registro de varias regiones
- Recibir archivos de CloudTrail registro de varias cuentas

Todas AWS Clean Rooms las acciones se registran CloudTrail y se documentan en la Referencia de la AWS Clean Rooms API.

Cada entrada de registro o evento contiene información sobre quién generó la solicitud. La información de identidad del usuario le ayuda a determinar lo siguiente:

- Si la solicitud se realizó con las credenciales del usuario raíz o del usuario de IAM.
- Si la solicitud se realizó con credenciales de seguridad temporales de un rol o fue un usuario federado.
- Si la solicitud la realizó otro Servicio de AWS.

Para obtener más información, consulte el elemento userIdentity de CloudTrail .

Descripción de las entradas de los archivos de AWS Clean Rooms registro

Un rastro es una configuración que permite la entrega de eventos como archivos de registro a un bucket de Amazon S3 que usted especifique. CloudTrail Los archivos de registro contienen una o más entradas de registro. Un evento representa una solicitud única de cualquier fuente e incluye información sobre la acción solicitada, la fecha y la hora de la acción, los parámetros de la solicitud, etc. CloudTrail Los archivos de registro no son un registro ordenado de las llamadas a la API pública, por lo que no aparecen en ningún orden específico.

Ejemplos de AWS Clean Rooms CloudTrail eventos

Los siguientes ejemplos muestran CloudTrail eventos para:

Temas

- StartProtectedQuery (exitoso)
- StartProtectedQuery (fallido)

StartProtectedQuery (exitoso)


```
{
    "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 (fallido)


```
{
    "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"
}
```

Creación de AWS Clean Rooms recursos con AWS CloudFormation

AWS Clean Rooms está integrado con AWS CloudFormation un servicio que le ayuda a modelar y configurar sus AWS recursos. Como resultado de esta integración, puede dedicar menos tiempo a crear y administrar sus recursos e infraestructura. Crea una plantilla que describe todos los AWS recursos que desea y AWS CloudFormation aprovisiona y configura esos recursos por usted. Algunos ejemplos de recursos son las colaboraciones, las tablas configuradas, las asociaciones de tablas configuradas y las pertenencias.

Cuando la utilice AWS CloudFormation, podrá reutilizar la plantilla para configurar los AWS Clean Rooms recursos de forma coherente y repetida. Describa sus recursos una vez y, a continuación, aprovisione los mismos recursos una y otra vez en varios Cuentas de AWS y Regiones de AWS.

AWS Clean Rooms y AWS CloudFormation plantillas

Para aprovisionar y configurar recursos AWS Clean Rooms y servicios relacionados, debe conocer AWS CloudFormation las plantillas. Las plantillas son archivos de texto con formato JSON o YAML. Estas plantillas describen los recursos que desea aprovisionar en sus AWS CloudFormation pilas. Si no estás familiarizado con JSON o YAML, puedes usar AWS CloudFormation Designer para ayudarte a empezar con AWS CloudFormation las plantillas. Para obtener más información, consulte ¿Qué es Designer de AWS CloudFormation ? en la Guía del usuario de AWS CloudFormation .

AWS Clean Rooms permite crear colaboraciones, tablas configuradas, asociaciones de tablas configuradas y membresías en ellas. AWS CloudFormation Para obtener más información, incluidos ejemplos de plantillas JSON y YAML para colaboraciones, tablas configuradas, asociaciones de tablas configuradas y pertenencias, consulte la Referencia de tipos de recursos de AWS Clean Rooms en la Guía del usuario de AWS CloudFormation .

Están disponibles las siguientes plantillas:

- Plantilla de análisis

  Especifique una plantilla de AWS Clean Rooms análisis, que incluya el nombre, la descripción, el formato, la fuente, los parámetros y las etiquetas.

  Para obtener más información, consulte los temas siguientes:

AWS::CleanRooms::AnalysisTemplate en la Guía del usuario de AWS Clean Rooms .

CreateAnalysisTemplate en la Referencia de la API de AWS Clean Rooms

- Colaboración

  Especifique una AWS Clean Rooms colaboración, incluidos el nombre, la descripción, el tipo, los parámetros y las etiquetas.

  Para obtener más información, consulte los temas siguientes:

  AWS::CleanRooms::Collaboration en la Guía del usuario de AWS CloudFormation .

  CreateCollaboration en la Referencia de la API de AWS Clean Rooms

- Tabla configurada

  Especifique una tabla configurada AWS Clean Rooms, incluidas las columnas permitidas, el método de análisis, la descripción, el nombre, la referencia de la tabla, el presupuesto de privacidad y las etiquetas. Las tablas configuradas representan una referencia a una tabla existente en la AWS Glue Data Catalog que se ha configurado para su uso en AWS Clean Rooms. Una tabla configurada contiene una regla de análisis que determina cómo se pueden utilizar los datos.

  Para obtener más información, consulte los temas siguientes:

  AWS::CleanRooms::ConfiguredTable en la Guía del usuario de AWS CloudFormation .

  CreateConfiguredTable en la Referencia de la API de AWS Clean Rooms

- Asociación de tablas configuradas

  Especifique una asociación de tabla configurada en AWS Clean Rooms, incluidos el ID, la descripción, el ID de membresía, el nombre, el rol, el nombre de recurso de Amazon (ARN) y las etiquetas. Una asociación de tablas configuradas vincula una tabla configurada a una colaboración.

  Para obtener más información, consulte los temas siguientes:

  AWS::CleanRooms::ConfiguredTableAssociation en la Guía del usuario de AWS CloudFormation .

  CreateConfiguredTableAssociation en la Referencia de la API de AWS Clean Rooms

- Pertenencia

  Especifique la pertenencia para un identificador de colaboración específico y únase a la colaboración en AWS Clean Rooms.

  Para obtener más información, consulte los temas siguientes:

  AWS::CleanRooms::Membership en la Guía del usuario de AWS CloudFormation .

  CreateMembership en la Referencia de la API de AWS Clean Rooms
- Plantilla de presupuesto de privacidad

  Especifica una plantilla AWS Clean Rooms de presupuesto de privacidad, que incluya un presupuesto de privacidad, el ruido añadido por consulta y una actualización mensual del presupuesto de privacidad.

  Para obtener más información, consulte los temas siguientes:

  AWS::CleanRooms::PrivacyBudgetTemplate en la Guía del usuario de AWS CloudFormation .

  CreatePrivacyBudgetTemplate en la Referencia de la API de AWS Clean Rooms
- Creación de un conjunto de datos de entrenamiento

  Especifique un conjunto de datos de entrenamiento para un modelo Clean Rooms ML a partir de una AWS Glue tabla.

  Para obtener más información, consulte los temas siguientes:

  AWS::CleanRoomsML::TrainingDataset en la Guía del usuario de AWS CloudFormation .

  CreateTrainingDataset en la referencia de la API de Clean Rooms ML

Obtenga más información sobre AWS CloudFormation

Para obtener más información AWS CloudFormation, consulte los siguientes recursos:

- AWS CloudFormation
- AWS CloudFormation Guía del usuario
- AWS CloudFormation Referencia de la API

- Guía del usuario de la interfaz de la línea de comandos de AWS CloudFormation

Cuotas para AWS Clean Rooms

Tienes Cuenta de AWS cuotas predeterminadas, antes denominadas límites, para cada una de ellas Servicio de AWS. A menos que se indique lo contrario, cada cuota es específica de un Región de AWS. Puede solicitar aumentos para algunas cuotas, pero otras no se pueden aumentar.

Para ver las cuotas AWS Clean Rooms, abra la consola Service Quotas. En el panel de navegación, elija Servicios de AWS y seleccione AWS Clean Rooms.

Para solicitar un aumento de cuota, consulte Solicitud de aumento de cuota en la Guía del usuario de Service Quotas. Si la cuota no está todavía disponible en Service Quotas, utilice el Formulario de aumento del límite de servicio.

Temas

- AWS Clean Rooms cuotas
- AWS Clean Rooms Cuotas de ML

AWS Clean Rooms cuotas

Cuenta de AWS Tiene las siguientes cuotas relacionadas con. AWS Clean Rooms

Nombre	Valor predeterminado	Ajusta	Descripción
Tamaño de regla de análisis	Cada región admitida: 100 kilobytes	No	Tamaño máximo de JSON para una regla de análisis
Plantillas de análisis por membresía	Cada región admitida: 25	No	Número máximo de plantillas de análisis por membresía
Colaboraciones creadas por cuenta	Cada región admitida: 10	Sí	Número máximo de colaboraciones creadas por cuenta

Nombre	Valor predeterminado	Ajusta	Descripción
Lista de columnas permitidas por tabla configurada	Cada región admitida: 100	No	Número máximo de columnas que se pueden incluir en la lista por tabla configurada
Uso de vCPU en PySpark tareas simultáneas por cuenta	us-east-1:1.024 Cada una de las demás regiones compatibles: 512	Sí	Uso máximo total de vCPU de todos los trabajos que se ejecutan PySpark simultáneamente por cuenta
Trabajos simultáneos PySpark por cuenta	us-east-1: 5 Cada una de las demás regiones compatibles: 2	Sí	Número máximo de PySpark trabajos en ejecución simultánea por cuenta
Consultas SQL simultáneas por cuenta	us-east-1: 5 Cada una de las demás regiones compatibles: 2	Sí	Número máximo de consultas SQL que se ejecutan simultáneamente por cuenta
Uso simultáneo de vCPU de consultas SQL por cuenta	Cada región admitida: 512	Sí	Uso máximo total de vCPU de todas las consultas SQL que se ejecutan simultáneamente por cuenta
Trabajo continuo simultáneo por membresía	Cada región admitida: 1	No	Número máximo de trabajos continuos simultáneos por membresía

Nombre	Valor predeterminado	Ajusta	Descripción
Consultas simultáneas en curso por membresía	Cada región admitida: 5	No	Número máximo de consultas simultáneas en curso por membresía
Asociaciones de modelos de audiencia configuradas por membresía	Cada región admitida: 5	No	Número máximo de asociaciones de modelos de audiencia configuradas por membresía
Tablas configuradas por cuenta	Cada región admitida: 60	No	Número máximo de tablas configuradas creadas por cuenta
Tablas configuradas por consulta protegida	Cada región admitida: 15	No	Número máximo de tablas configuradas en una consulta protegida
Tablas de mapeo de ID por membresía	Cada región admitida: 5	Sí	Número máximo de tablas de mapeo de ID por membresía
Asociaciones de espacios de nombres de ID por suscripción	Cada región admitida: 10	Sí	Número máximo de asociaciones de espacios de nombres de ID por membresía
Miembros invitados por colaboración	Cada región admitida: 5	Sí	Número máximo de miembros invitados por colaboración
Membresías por cuenta	Cada región admitida: 100	Sí	Número máximo de membresías por cuenta
Longitud del texto de la consulta	Cada región admitida: 16 kilobytes	No	Longitud máxima del texto para una instrucción de consulta SQL

Nombre	Valor predeterminado	Ajusta	Descripción
Tasa de solicitudes BatchGetSchema	Cada región admitida: 5	Sí	Número máximo de llamadas a la BatchGetSchema API por segundo
Tasa de CreateCollaboration solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la CreateCollaboration API por segundo
Tasa de CreateConfiguredTable solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la CreateConfiguredTable API por segundo
Tasa de CreateConfiguredTableAnalysisRule solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la CreateConfiguredTableAnalysisRule API por segundo
Tasa de CreateConfiguredTableAssociation solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la CreateConfiguredTableAssociation API por segundo
Tasa de CreateMembership solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la CreateMembership API por segundo
Tasa de DeleteCollaboration solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la DeleteCollaboration API por segundo

Nombre	Valor predeterminado	Ajusta	Descripción
Tasa de DeleteConfiguredTable solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la DeleteConfiguredTable API por segundo
Tasa de DeleteConfiguredTableAnalysisRule solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la DeleteConfiguredTableAnalysisRule API por segundo
Tasa de DeleteConfiguredTableAssociation solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la DeleteConfiguredTableAssociation API por segundo
Tasa de DeleteMember solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la DeleteMember API por segundo
Tasa de DeleteMembership solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la DeleteMembership API por segundo
Tasa de GetCollaboration solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la GetCollaboration API por segundo
Tasa de GetConfiguredTable solicitudes	Cada región admitida: 20	Sí	Número máximo de llamadas a la GetConfiguredTable API por segundo

Nombre	Valor predeterminado	Ajusta	Descripción
Tasa de GetConfiguredTableAnalysisRule solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la GetConfiguredTableAnalysisRule API por segundo
Tasa de GetConfiguredTableAssociation solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la GetConfiguredTableAssociation API por segundo
Tasa de GetMembership solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la GetMembership API por segundo
Tasa de GetProtectedJob solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la GetProtectedJob API por segundo
Tasa de GetProtectedQuery solicitudes	Cada región admitida: 20	Sí	Número máximo de llamadas a la GetProtectedQuery API por segundo
Tasa de GetSchema solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la GetSchema API por segundo
Tasa de GetSchemaAnalysisRule solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la GetSchemaAnalysisRule API por segundo

Nombre	Valor predeterminado	Ajusta	Descripción
Tasa de ListCollaborations solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la ListCollaborations API por segundo
Tasa de ListConfiguredTableAssociations solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la ListConfiguredTableAssociations API por segundo
Tasa de ListConfiguredTables solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la ListConfiguredTables API por segundo
Tasa de ListMembers solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la ListMembers API por segundo
Tasa de ListMemberships solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la ListMemberships API por segundo
Tasa de ListProtectedJobs solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la ListProtectedJobs API por segundo
Tasa de ListProtectedQueries solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la ListProtectedQueries API por segundo

Nombre	Valor predeterminado	Ajusta	Descripción
Tasa de ListSchemas solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la ListSchemas API por segundo
Tasa de StartProtectedJob solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la StartProtectedJob API por segundo
Tasa de StartProtectedQuery solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la StartProtectedQuery API por segundo
Tasa de UpdateCollaboration solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la UpdateCollaboration API por segundo
Tasa de UpdateConfiguredTable solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la UpdateConfiguredTable API por segundo
Tasa de UpdateConfiguredTableAnalysisRule solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la UpdateConfiguredTableAnalysisRule API por segundo
Tasa de UpdateConfiguredTableAssociation solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la UpdateConfiguredTableAssociation API por segundo

Nombre	Valor predeterminado	Ajusta	Descripción
Tasa de UpdateProtectedJob solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la UpdateProtectedJob API por segundo
Tasa de UpdateProtectedQuery solicitudes	Cada región admitida: 5	Sí	Número máximo de llamadas a la UpdateProtectedQuery API por segundo
Asociaciones de tabla por membresía	Cada región admitida: 25	No	Número máximo de asociaciones de tabla por membresía

AWS Clean Rooms límites de parámetros de recursos

Recurso	Predeterminado/a	Descripción
Longitud del texto de consulta	90 KB	Longitud máxima del texto de una instrucción de consulta SQL
Longitud del texto de la consulta (utilizando la privacidad diferencial)	8 KB	Longitud máxima del texto para una sentencia de consulta SQL con privacidad diferencial
Tiempo de ejecución de consultas	12 horas	Tiempo máximo que se ejecuta una consulta antes de que se agote el tiempo de espera

AWS Clean Rooms Limitación de las cuotas de API

Cuenta de AWS Tiene las siguientes cuotas de transacciones de API por segundo (TPS) por cuenta y punto final para los siguientes recursos:

- AnalysisTemplate
- ConfiguredAudienceModelAssociation
- PrivacyBudgetTempate
- CollaborationConfiguredAudienceModelAssociation

Recurso	Límite de frecuencia	Descripción
Tasa de solicitudes `BatchGetCollaborationAnalysisTemplate`	5 TPS	Número máximo de llamadas a la API `BatchGetCollaborationAnalysisTemplate` por segundo
Tasa de solicitudes `CreateAnalysisTemplate`	5 TPS	Número máximo de llamadas a la API `CreateAnalysisTemplate` por segundo
Tasa de solicitudes `CreateConfiguredAudienceModelAssociation`	5 TPS	Número máximo de llamadas `CreateConfiguredAudienceModelAssociation` por segundo
Tasa de solicitudes `CreatePrivacyBudgetTempate`	5 TPS	Número máximo de llamadas `CreatePrivacyBudgetTemplate` por segundo
Tasa de solicitudes `DeleteAnalysisTemplate`	5 TPS	Número máximo de llamadas `DeleteAnalysisTemplate` por segundo
Tasa de solicitudes `DeleteConfiguredAu`	5 TPS	Número máximo de llamadas `DeleteConfiguredAu`

Recurso	Límite de frecuencia	Descripción
dienceModelAssocia tion		dienceModelAssocia tion por segundo
Tasa de solicitudes DeletePrivacyBudge tTemplate	5 TPS	Número máximo de llamadas DeletePrivacyBudge tTemplate por segundo
Tasa de solicitudes GetAnalysisTemplate	5 TPS	Número máximo de llamadas GetAnalysisTemplate por segundo
Tasa de solicitudes GetCollaborationCo nfiguredAudienceMo delAssociation	5 TPS	Número máximo de llamadas GetCollaborationCo nfiguredAudienceMo delAssociation por segundo
Tasa de solicitudes GetCollaborationPr ivacyBudgetTemplate	5 TPS	Número máximo de llamadas GetCollaborationPr ivacyBudgetTemplate por segundo
Tasa de solicitudes GetConfiguredAudie nceModelAssociation	5 TPS	Número máximo de llamadas GetConfiguredAudie nceModelAssociation por segundo
Tasa de solicitudes GetPrivacyBudgetTe mplate	5 TPS	Número máximo de llamadas GetPrivacyBudgetTe mplate por segundo
Tasa de solicitudes ListAnalysisTempla tes	5 TPS	Número máximo de llamadas ListAnalysisTempla tes por segundo

Recurso	Límite de frecuencia	Descripción
Tasa de solicitudes `ListCollaborationConfiguredAudienceModelAssociations`	5 TPS	Número máximo de llamadas `ListCollaborationConfiguredAudienceModelAssociations` por segundo
Tasa de solicitudes `ListCollaborationPrivacyBudgets`	5 TPS	Número máximo de llamadas `ListCollaborationPrivacyBudgets` por segundo
Tasa de solicitudes `ListCollaborationPrivacyBudgetTemplates`	5 TPS	Número máximo de llamadas `ListCollaborationPrivacyBudgetTemplates` por segundo
Tasa de solicitudes `ListConfiguredAudienceModelAssociations`	5 TPS	Número máximo de llamadas `ListConfiguredAudienceModelAssociations` por segundo
Tasa de solicitudes `ListPrivacyBudgets`	5 TPS	Número máximo de llamadas `ListPrivacyBudgets` por segundo
Tasa de solicitudes `ListPrivacyBudgetTemplates`	5 TPS	Número máximo de llamadas `ListPrivacyBudgetTemplates` por segundo
Tasa de solicitudes `UpdateAnalysisTemplate`	5 TPS	Número máximo de llamadas `UpdateAnalysisTemplate` por segundo

Recurso	Límite de frecuencia	Descripción
Tasa de solicitudes `UpdateConfiguredAudienceModelAssociation`	5 TPS	Número máximo de llamadas `UpdateConfiguredAudienceModelAssociation` por segundo
Tasa de solicitudes `UpdatePrivacyBudgetTemplate`	5 TPS	Número máximo de llamadas `UpdatePrivacyBudgetTemplate` por segundo

AWS Clean Rooms Cuotas de ML

Cuenta de AWS Tiene las siguientes cuotas relacionadas con Clean Rooms ML.

Nombre	Valor predeterminado	Ajusta	Descripción
Trabajos de exportación de audiencia activa por trabajo de generación de audiencia	Cada región admitida: 25	No	El número máximo de trabajos de exportación de audiencia activa para un trabajo de generación de audiencia
Asociaciones activas de algoritmos de modelos configurados por membresía	Cada región admitida: 1000	Sí	El número máximo de asociaciones de algoritmos de modelo configuradas activas por membresía
Algoritmos de modelo configurados activos por membresía	Cada región admitida: 1000	Sí	El número máximo de algoritmos de modelo configurados activos por membresía
Canales de entrada de modelos personalizados activos por membresía	Cada región admitida: 100	Sí	El número máximo de canales de entrada de

Nombre	Valor predeterminado	Ajusta	Descripción
			modelos personalizados activos por membresía
Los trabajos de exportación de audiencia pendientes o en curso por cliente	Cada región admitida: 20	No	El número máximo de trabajos de exportación de audiencia pendientes o en curso por cliente
Los trabajos de generación de audiencia pendientes o en curso por cliente	Cada región admitida: 10	Sí	El número máximo de trabajos de generación de audiencia pendientes o en curso por cliente
Los modelos de audiencia pendientes o en curso por cliente	Cada región admitida: 2	Sí	El número máximo de trabajos de entrenamiento del modelo de audiencia pendientes o en curso por cliente
Trabajos de inferencia de modelos personalizados pendientes o en curso por cuenta	Cada región admitida: 10	Sí	El número máximo de trabajos de inferencia de modelos personalizados pendientes o en curso por cuenta
Trabajos de inferencia de modelos personalizados pendientes o en curso por membresía	Cada región admitida: 5	Sí	El número máximo de trabajos de inferencia de modelos personalizados pendientes o en curso por membresía

Nombre	Valor predeterminado	Ajusta	Descripción
Trabajos de formación de modelos personalizados pendientes o en curso por cuenta	Cada región admitida: 10	Sí	El número máximo de trabajos de formación de modelos personalizados pendientes o en curso por cuenta
Trabajos de formación de modelos personalizados pendientes o en curso por membresía	Cada región admitida: 5	Sí	El número máximo de trabajos de formación de modelos personalizados pendientes o en curso por membresía

Cuotas de Clean Rooms ML

Recurso	Predeterminado/a	Descripción
Duración máxima de un trabajo de inferencia de modelado personalizado	25 horas	
Número máximo de interacciones	20 mil millones	Número máximo de interacciones permitidas en los datos de entrenamiento. Las entradas más grandes se muestrean.
Número mínimo de interacciones	1 millón	
Número máximo de usuarios distintos para el entrenamiento de modelos similares	100 millones	Si se incluyen más, solo se utilizan los primeros 100 millones, ordenados por número de interacciones.

Recurso	Predeterminado/a	Descripción
Número mínimo de usuarios distintos para el entrenamiento de modelos similares	100 000	
Número mínimo de usuarios para el trabajo de segmentos similares (audiencia) de exportación	10 000	
Número máximo de elementos distintos que se utilizan para el entrenamiento de modelos.	1 millón	Puede incluir hasta 50 millones de artículos, pero solo se utiliza el 1 millón más popular.
Número máximo de columnas de características en el conjunto de datos de entrenamiento.	10	
Número mínimo de elementos distintos por usuario	2	AWS Clean Rooms El aprendizaje automático requiere que cada fila o usuario tenga dos o más elementos, incluidos los elementos repetidos.
Tamaño máximo de la audiencia inicial	500.000	
Tamaño mínimo de la audiencia inicial	500	El proveedor de datos de entrenamiento puede establecer este valor tan bajo como 25.
APIs	por cliente	

Recurso	Predeterminado/a	Descripción
Número total de conjuntos de datos de entrenamiento activos.	500	
Número total de modelos similares activos (modelos de audiencia)	500	
Número total de modelos similares configurados y activos (modelos de audiencia)	10 000	
Número total de trabajos de generación de segmentos similares (audiencia) completados	Sin límite	
Número total de trabajos de segmentos similares de exportación (audiencia) completados	Sin límite	
Duración máxima de un trabajo de generación de modelos similares (modelo de audiencia)	1 día (24 horas)	

Recurso	Predeterminado/a	Descripción
Duración máxima de un trabajo de generación de segmentos similares (audiencia)	10 horas	Tras proporcionar un origen, Clean Rooms ML tarda un máximo de 10 horas en generar un segmento similar. Si utiliza una consulta de SQL como dato inicial, la consulta puede tardar hasta 12 horas en ejecutarse, además de las 10 horas que se tardan en generar el segmento similar.
Porcentaje mínimo para el tamaño de la papelera de un segmento (audiencia)	1%	
Porcentaje máximo para el tamaño de la papelera de un segmento (audiencia)	20%	
Tamaño absoluto mínimo para el tamaño de la papelera de un segmento (audiencia)	El 1 % del número de usuarios distintos	
Tamaño absoluto máximo para el tamaño de la papelera de un segmento (audiencia)	El 20 % del número de usuarios distintos	

Limpia las cuotas de la API Clean Rooms ML

Cuenta de AWS Tiene las siguientes cuotas de transacciones de API por segundo (TPS) por cuenta y punto final.

Recurso	Límite de frecuencia	Descripción
Tasa de solicitudes de `CreateAudienceModel`	Tasa de 1 TPS, ráfaga de 3 TPS	Número máximo de llamadas a la API `CreateAudienceModel` por segundo
Tasa de solicitudes `CreateConfiguredAudienceModel`	10 TPS	Número máximo de llamadas a la API `CreateConfiguredAudienceModel` por segundo
Tasa de solicitudes `CreateTrainingDataset`	10 TPS	Número máximo de llamadas a la API `CreateTrainingDataset` por segundo
Tasa de solicitudes `DeleteAudienceGenerationJob`	Tasa de 2 TPS, ráfaga de 10 TPS	Número máximo de llamadas a la API `DeleteAudienceGenerationJob` por segundo
Tasa de solicitudes `DeleteAudienceModel`	Tasa de 2 TPS, ráfaga de 10 TPS	Número máximo de llamadas a la API `DeleteAudienceModel` por segundo
Tasa de solicitudes `DeleteConfiguredAudienceModel`	10 TPS	Número máximo de llamadas a la API `DeleteConfiguredAudienceModel` por segundo
Tasa de solicitudes `DeleteConfiguredAudienceModelPolicy`	25 TPS	Número máximo de llamadas a la API `DeleteConfiguredAudienceModelPolicy` por segundo
Tasa de solicitudes `DeleteTrainingDataset`	10 TPS	Número máximo de llamadas a la API `DeleteTra`

Recurso	Límite de frecuencia	Descripción
		iningDataset por segundo
Tasa de solicitudes GetAudienceGenerationJob	50 TPS	Número máximo de llamadas a la API GetAudienceGenerationJob por segundo
Tasa de solicitudes GetAudienceModel	50 TPS	Número máximo de llamadas a la API GetAudienceModel por segundo
Tasa de solicitudes GetConfiguredAudienceModel	50 TPS	Número máximo de llamadas a la API GetConfiguredAudienceModel por segundo
Tasa de solicitudes GetConfiguredAudienceModelPolicy	50 TPS	Número máximo de llamadas a la API GetConfiguredAudienceModelPolicy por segundo
Tasa de solicitudes GetTrainingDataset	50 TPS	Número máximo de llamadas a la API GetTrainingDataset por segundo
Tasa de solicitudes ListAudienceExportJobs	50 TPS	Número máximo de llamadas a la API ListAudienceExportJobs por segundo
Tasa de solicitudes ListAudienceGenerationJobs	50 TPS	Número máximo de llamadas a la API ListAudienceGenerationJobs por segundo

Recurso	Límite de frecuencia	Descripción
Tasa de solicitudes `ListAudienceModels`	50 TPS	Número máximo de llamadas a la API `ListAudienceModels` por segundo
Tasa de solicitudes `ListConfiguredAudienceModels`	50 TPS	Número máximo de llamadas a la API `ListConfiguredAudienceModels` por segundo
Tasa de solicitudes `ListTagsForResource`	50 TPS	Número máximo de llamadas a la API `ListTagsForResource` por segundo
Tasa de solicitudes `ListTrainingDatasets`	50 TPS	Número máximo de llamadas a la API `ListTrainingDatasets` por segundo
Tasa de solicitudes `PutConfiguredAudienceModelPolicy`	25 TPS	Número máximo de llamadas a la API `PutConfiguredAudienceModelPolicy` por segundo
Tasa de solicitudes `StartAudienceExportJob`	Tasa de 1 TPS, ráfaga de 3 TPS	Número máximo de llamadas a la API `StartAudienceExportJob` por segundo
Tasa de solicitudes `StartAudienceGenerationJob`	Tasa de 1 TPS, ráfaga de 5 TPS	Número máximo de llamadas a la API `StartAudienceGenerationJob` por segundo
Tasa de solicitudes `TagResource`	10 TPS	Número máximo de llamadas a la API `TagResource` por segundo

Recurso	Límite de frecuencia	Descripción
Tasa de solicitudes `UntagResource`	50 TPS	Número máximo de llamadas a la API `UntagResource` por segundo
Tasa de solicitudes `UpdateConfiguredAudienceModel`	10 TPS	Número máximo de llamadas a la API `UpdateConfiguredAudienceModel` por segundo
Tasa de solicitudes `CreateConfiguredModelAlgorithm`	10 TPS	Número máximo de llamadas a la `CreateConfiguredModelAlgorithm` API por segundo.
Tasa de solicitudes `CreateConfiguredModelAlgorithmAssociation`	10 TPS	Número máximo de llamadas a la `CreateConfiguredModelAlgorithmAssociaton` API por segundo.
Tasa de solicitudes `PutMLConfiguration`	10 TPS	Número máximo de llamadas a la `PutMLConfiguration` API por segundo.
Tasa de solicitudes `CreateTrainedModel`	Tasa de 1 TPS, ráfaga de 3 TPS	Número máximo de llamadas a la `CreateTrainedModel` API por segundo.
Tasa de solicitudes `StartTrainedModelExportJob`	10 TPS	Número máximo de llamadas a la `StartTrainedModelExportJob` API por segundo.
Tasa de solicitudes `StartTrainedModelInferenceJob`	Tasa de 1 TPS, tasa de 3 TPS	Número máximo de llamadas a la `StartTrainedModelInferenceJob` API por segundo.

Recurso	Límite de frecuencia	Descripción
Tasa de `GetConfiguredModelAlgorithm` solicitud	50 TPS	Número máximo de llamadas a la `GetConfiguredModelAlgorithm` API por segundo.
Tasa de `GetConfiguredModelAlgorithmAssociation` solicitud	50 TPS	Número máximo de llamadas a la `GetConfiguredModelAlgorithmAssociaton` API por segundo.
Tasa de solicitudes `GetTrainedModel`	50 TPS	Número máximo de llamadas a la `GetTrainedModel` API por segundo.
Tasa de solicitudes `GetMLConfiguration`	50 TPS	Número máximo de llamadas a la `GetMLConfiguration` API por segundo.
Tasa de solicitudes `GetTrainedModelInferenceJob`	50 TPS	Número máximo de llamadas a la `GetTrainedModelInferenceJob` API por segundo.
Tasa de solicitudes `ListConfiguredModelAlgorithm`	50 TPS	Número máximo de llamadas a la `ListConfiguredModelAlgorithm` API por segundo.
Tasa de solicitudes `ListConfiguredModelAlgorithmAssociations`	50 TPS	Número máximo de llamadas a la `ListConfiguredModelAlgorithmAssociations` API por segundo.
Tasa de solicitudes `ListTrainedModels`	50 TPS	Número máximo de llamadas a la `ListTrainedModels` API por segundo.

Recurso	Límite de frecuencia	Descripción
Tasa de solicitudes `ListCollaborationTrainedModelExportJobs`	50 TPS	Número máximo de llamadas a la `ListCollaborationTrainedModelExportJobs` API por segundo.
Tasa de solicitudes `ListCollaborationTrainedModelInferenceJobs`	50 TPS	Número máximo de llamadas a la `ListCollaborationTrainedModelInferenceJobs` API por segundo.
Tasa de solicitudes `DeleteConfiguredModelAlgorithm`	Tasa de 2 TPS, ráfaga de 10 TPS	Número máximo de llamadas a la `DeleteConfiguredModelAlgorithm` API por segundo.
Tasa de solicitudes `DeleteConfiguredModelAlgorithmAssociation`	Tasa de 2 TPS, ráfaga de 10 TPS	Número máximo de solicitudes de `DeleteConfiguredModelAlgorithmAssociaton` API por segundo.
Tasa de solicitudes `DeleteMLConfiguration`	Tasa de 2 TPS, ráfaga de 10 TPS	Número máximo de solicitudes de `DeleteMLConfiguration` API por segundo.
Tasa de solicitudes `DeleteTrainedModelOutput`	Tasa de 2 TPS, ráfaga de 10 TPS	Número máximo de solicitudes de `DeleteTrainedModelOutput` API por segundo.

Historial de documentos de la Guía AWS Clean Rooms del usuario

En la siguiente tabla se describen las versiones de la documentación de AWS Clean Rooms.

Para obtener notificaciones sobre las actualizaciones de esta documentación, puede suscribirse a la fuente RSS. Para suscribirse a las actualizaciones RSS, debe tener un complemento de RSS habilitado para el navegador que esté utilizando.

Cambio	Descripción	Fecha
Plazo de finalización del soporte para el motor de análisis SQL de Clean Rooms	A partir de ahora y hasta el 16 de julio de 2025, las nuevas colaboraciones que utilicen el antiguo motor de análisis SQL de Clean Rooms requieren un aumento del límite. A partir del 17 de julio de 2025, no podrás crear nuevas colaboraciones con este motor. AWS Clean Rooms finaliza la compatibi lidad con el motor de análisis SQL de Clean Rooms el 17 de diciembre de 2025.	22 de mayo de 2025
Actualización de las plantillas PySpark de análisis	Los clientes ahora pueden asignar un nombre a un archivo de script de usuario con cualquier nombre de archivo Python válido.	15 de mayo de 2025
Support para múltiples receptores de resultados	Los clientes ahora pueden seleccionar varios miembros para recibir los resultados de una sola consulta.	30 de abril de 2025

Support para migrar colaboraciones a Spark SQL	AWS Clean Rooms SQL ahora admite reglas de agregación y análisis de listas, además de reglas de análisis personalizadas. Además, los clientes pueden actualizar una colaboración existente para usar el motor de análisis de Spark que impulsa Spark SQL.	2 de abril de 2025
Support for PySpark jobs	Los clientes ahora pueden analizar los datos mediante la ejecución de tareas mediante plantillas de PySpark análisis aprobadas.	18 de marzo de 2025
Actualización de políticas existentes	Se ha agregado el permiso nuevo siguiente a la política administrada `AWSCleanRoomsMLReadOnlyAccess` : `PassCleanRoomsResources` . Se han agregado los siguientes permisos nuevos a la política `AWSCleanRoomsMLFullAccess` administrada: `PassCleanRoomsResources` y `ConsoleDescribeECRRepositories` .	10 de enero de 2025
Support para varios trabajadores de cómputo	Los clientes ahora pueden especificar qué tipo de trabajadores informáticos y cuántos aprovisionar al crear un segmento similar.	17 de diciembre de 2024

Support para múltiples fuentes de datos y nubes	Los clientes ahora pueden usar múltiples fuentes de datos y nubes, como Amazon Athena y Snowflake, para colaborar con los conjuntos de datos de sus socios.	1 de diciembre de 2024
El modelado personalizado de Clean Rooms ML ya está disponible	Los clientes ahora pueden usar sus propios modelos de aprendizaje automático personalizados en colaboración.	7 de noviembre de 2024
Nuevo motor de análisis	Los clientes que tienen conjuntos de datos de gran tamaño ahora pueden ejecutar consultas complejas mediante las funciones de SQL compatibles con el motor de análisis SQL de Spark.	29 de octubre de 2024
Protección de privacidad mejorada, creación de audiencias parecidas, elección de varios receptores de resultados	Puede proteger los datos y, al mismo tiempo, permitir consultas de activación complejas mediante los análisis adicionales y la regla de análisis de colaboración. Puede crear modelos de audiencia similares a partir de consultas SQL o plantillas de análisis. Puede elegir varios miembros para recibir los resultados.	24 de julio de 2024

Resolución de entidades en AWS Clean Rooms	Con AWS Entity Resolution in AWS Clean Rooms, puede crear una tabla de mapeo de ID entre dos espacios de nombres de ID para consultar datos de eventos en espacios de identidad dispares.	23 de julio de 2024
Actualización de una política existente	Se ha agregado el permiso nuevo siguiente a la política administrada `AWSCleanRoomsFullAccessNoQuerying` : `cleanrooms:BatchGetSchemaAnalysisRule` .	13 de mayo de 2024
AWS Clean Rooms El aprendizaje automático ya está totalmente disponible	AWS Clean Rooms El aprendizaje automático proporciona un método que mejora la privacidad para que dos partes identifiquen a usuarios similares en sus datos sin necesidad de compartir sus datos entre sí.	3 de abril de 2024
Actualización de una política existente	El ID de declaración de la política administrada por `AWSCleanRoomsFullAccess` se ha actualizado de ConsolePickQueryResultsBucket a SetQueryResultsBucket para representar mejor los permisos desde los permisos.	21 de marzo de 2024

Nuevas políticas gestionadas para el aprendizaje automático AWS Clean Rooms	Se han añadido dos nuevas políticas administradas: `AWSCleanRoomsMLReadOnlyAccess` y `AWSCleanRoomsMLFullAccess` .	29 de noviembre de 2023
AWS Clean Rooms ML (versión preliminar)	AWS Clean Rooms El aprendizaje automático proporciona un método que mejora la privacidad para que dos partes identifiquen a usuarios similares en sus datos sin necesidad de compartir sus datos entre sí.	29 de noviembre de 2023
AWS Clean Rooms Privacidad diferencial (versión preliminar)	Los clientes ahora pueden usar AWS Clean Rooms Differential Privacy para ayudar a proteger la privacidad de sus usuarios.	29 de noviembre de 2023
Configuración de pagos	El creador de la colaboración ahora puede designar al miembro que puede ejecutar consultas o a otro miembro de la colaboración para que se le facturen los costos de computación de las consultas.	14 de noviembre de 2023
Tiempo de ejecución de consulta: actualización	El tiempo máximo que se ejecuta una consulta antes de que se agote el tiempo de espera se ha actualizado de 4 a 12 horas.	6 de octubre de 2023

AWS CloudFormation recursos: actualización	AWS Clean Rooms ha agregado los siguiente s recursos nuevos: `AWS::CleanRooms::M embership Protected QueryOutputConfigu ration AWS::Clea nRooms::Membership ProtectedQueryResu ltConfiguration , yAWS::CleanRooms::M embership Protected QueryS3OutputConfi guration .`	7 de septiembre de 2023
AWS CloudFormation recursos: actualizar	AWS Clean Rooms ha añadido los siguientes recursos nuevos: `AWS::Clea nRooms::AnalysisTe mplate yAWS::Clea nRooms::Configured Table AnalysisR uleCustom .`	31 de agosto de 2023
Capacidades independientes de los miembros	El creador de la colaborac ión ahora puede designar a un miembro que realice las consultas y a otro miembro que reciba los resultados. Esto permite al creador de la colaboración asegurarse de que el miembro que puede realizar las consultas no tenga acceso a los resultados de la consulta.	30 de agosto de 2023

AWS Clean Rooms Glosario	Actualización exclusiva de la documentación para añadir un glosario de términos. AWS Clean Rooms	30 de agosto de 2023
Compatibilidad para tablas de Apache Iceberg (vista previa)	AWS Clean Rooms ahora admite Apache Iceberg tablas (vista previa).	25 de agosto de 2023
Actualización de cuotas	La sección Cuotas se ha actualizado para reflejar la nueva cuota predeterminada de pertenencias por cuenta.	9 de agosto de 2023
Actualización de la política existente	Se han añadido los siguiente s nuevos permisos a la política administrada de `AWSCleanRoomsFullAccessNoQuerying` : `cleanrooms:CreateAnalysisTemplate` , `cleanrooms:GetAnalysisTemplate` , `cleanrooms:UpdateAnalysisTemplate` , `cleanrooms:DeleteAnalysisTemplate` , `cleanrooms:ListAnalysisTemplates` , `cleanrooms:GetCollaborationAnalysisTemplate` , `cleanrooms:BatchGetCollaborationAnalysisTemplate` y `cleanrooms:ListCollaborationAnalysisTemplates` .	31 de julio de 2023

Plantillas de análisis y regla de análisis personalizada	AWS Clean Rooms ahora admite plantillas de análisis y la regla de análisis personalizada. Las plantillas de análisis permiten a los colaboradores crear o importar su propia consulta SQL personalizada para utilizarla en la colaboración. Con la regla de análisis personalizada, el propietario de la tabla puede aprobar consultas SQL personalizadas en sus tablas configuradas.	31 de julio de 2023
Las reglas de análisis admiten la condición lógica `OR`	AWS Clean Rooms las reglas de análisis ahora admiten la condición `OR` lógica de la `JOIN` cláusula.	29 de junio de 2023
CloudFormation integration	AWS Clean Rooms ahora se integra con AWS CloudFormation.	15 de junio de 2023
Creador de análisis	Los miembros que pueden realizar consultas y recibir resultados ahora tienen la posibilidad de ejecutar consultas en algunas tablas sin escribir código SQL utilizando la interfaz de usuario del creador de análisis.	15 de junio de 2023
Funciones SQL	Actualización solo de la documentación para aclarar las funciones SQL admitidas.	5 de mayo de 2023

Solución de problemas	Actualización solo de la documentación para añadir una sección de solución de problemas para los problemas más frecuentes.	27 de abril de 2023
Tipos de datos compatibles para AWS Clean Rooms	Actualización exclusiva de la documentación para añadir una nueva sección que muestre los tipos de AWS Glue Data Catalog datos compatibles.	26 de abril de 2023
Ejemplos de eventos AWS CloudTrail	Actualización solo de la documentación para agregar ejemplos de CloudTrail eventos StartProtectedQuery (exitosos) y StartProtectedQuery (fallidos).	20 de abril de 2023
Actualización de la política existente	Se han añadido los siguientes nuevos permisos a la política administrada `AWSCleanRoomsFullAccessNoQuerying` : `cleanrooms:ListTagsForResource` , `cleanrooms:UntagResource` y `cleanrooms:TagResource` . Para obtener más información, consulte Políticas administradas de AWS.	21 de marzo de 2023
Disponibilidad general	AWS Clean Rooms ahora está disponible de forma general.	21 de marzo de 2023

Versión de prueba	Versión preliminar de la Guía AWS Clean Rooms del usuario	12 de enero de 2023

AWS Clean Rooms Glosario

Consulte este glosario para familiarizarse con la terminología utilizada en AWS Clean Rooms.

Regla de análisis de agregación

La restricción de consultas que permite realizar consultas que agreguen análisis mediante COUNT, SUM, or AVG funciona a lo largo de dimensiones opcionales. Estas consultas no revelarán información de nivel de fila.

Admite casos de uso como la planificación de campañas, el alcance mediático, la frecuencia y la medición de conversiones.

Otros tipos de reglas de análisis son las personalizadas y las de lista.

Reglas de análisis

Las restricciones de consulta que autorizan un tipo específico de consulta.

El tipo de regla de análisis determina qué tipo de análisis se puede ejecutar en la tabla configurada. Cada tipo tiene una estructura de consulta predefinida. Usted controla cómo se pueden utilizar las columnas de la tabla en la estructura mediante los controles de consulta.

Los tipos de reglas de análisis son de agregación, de lista y personalizadas.

Plantilla de análisis

Una consulta preaprobada y específica de la colaboración que se puede reutilizar.

Formatos compatibles: código SQL o código Python para Spark.

Si utiliza SQL, la plantilla de análisis puede contener parámetros siempre que aparezca un valor literal en una consulta SQL. Para obtener más información sobre los tipos de parámetros admitidos, consulte Tipos de datos en la Referencia de SQL de AWS Clean Rooms .

Las plantillas de análisis solo funcionan con la regla de análisis personalizada.

AWS Clean Rooms Motor de análisis de SQL

Un sistema de procesamiento de consultas integrado AWS Clean Rooms que permite a los usuarios consultar los datos almacenados en Amazon S3 mediante las funciones de SQL compatibles con AWS Clean Rooms. Es compatible con varios formatos de datos y ofrece funciones para ejecutar consultas SQL en conjuntos de datos colaborativos y, al mismo tiempo, mantener la privacidad y el control de los datos, e incluye características como la privacidad diferencial. Este motor está diseñado para cada caso de AWS Clean Rooms uso y ofrece un equilibrio entre la funcionalidad de SQL, las características de privacidad de los datos y la integración con otras AWS Clean Rooms capacidades, lo que lo hace adecuado para los usuarios que no necesitan las capacidades avanzadas o la escala del motor de análisis SQL de Spark.

Cuando creas una colaboración mediante la CreateCollaborationAPI, el valor del motor de análisis de AWS Clean Rooms SQL es fundamental`CLEAN_ROOMS_SQL`.

Cliente de cifrado de S3

La computación criptográfica para Clean Rooms cliente de cifrado (C3R).

C3R, que se utiliza para cifrar y descifrar datos, es un SDK de cifrado del cliente con una interfaz de línea de comandos.

Columna de texto sin cifrar

Una columna que no está protegida criptográficamente para ninguno de los JOIN o SELECT Construcción SQL.

Las columnas de texto sin cifrar se pueden usar en cualquier parte de la consulta SQL.

Colaboración

Límite lógico seguro AWS Clean Rooms en el que los miembros pueden realizar consultas SQL en tablas configuradas.

Las colaboraciones las crea el creador de la colaboración.

Solo los miembros que hayan sido invitados a la colaboración pueden unirse a ella.

Una colaboración solo puede tener un miembro que pueda consultar datos o un miembro que pueda ejecutar consultas y trabajos.

Una colaboración solo puede tener un miembro que pueda recibir los resultados.

Una colaboración solo puede tener un miembro que pague los costos de procesamiento de consultas o un miembro que pague los costos de procesamiento de consultas y trabajos.

Todos los miembros pueden ver la lista de participantes invitados a la colaboración antes de unirse a ella.

Creador de la colaboración

El miembro que crea una colaboración.

Solo hay un creador por colaboración.

El creador de la colaboración es el único que puede eliminar miembros de la colaboración o eliminar la colaboración.

Tabla configurada

Cada tabla configurada representa una referencia a una tabla existente en la AWS Glue Data Catalog que se ha configurado para su uso en AWS Clean Rooms. Una tabla configurada contiene una regla de análisis que determina cómo se pueden utilizar los datos.

Actualmente, AWS Clean Rooms admite la asociación de datos almacenados en Amazon Simple Storage Service (Amazon S3) que están catalogados mediante. AWS Glue

Para obtener más información al respecto AWS Glue, consulte la Guía para AWS Glue desarrolladores.

Las tablas configuradas se pueden asociar a una o más colaboraciones.

> **Note**
>
> AWS Clean Rooms actualmente no es compatible con las ubicaciones de bucket de Amazon S3 en las que esté registrado AWS Lake Formation.

Regla de análisis personalizada

La restricción de consultas que permite un conjunto específico de consultas preaprobadas (plantillas de análisis) o permite que un conjunto específico de cuentas puedan proporcionar consultas o trabajos que utilicen sus datos.

Admite casos de uso como la atribución de primer toque, los análisis incrementales y los análisis de detección de audiencia.

Admite la privacidad diferencial.

Otros tipos de reglas de análisis son de agregación y lista.

Descifrado

El proceso de transformar los datos cifrados para devolverles su forma original. El descifrado solo se puede realizar si se tiene el acceso a la clave secreta.

Privacidad diferencial

Una técnica matemáticamente rigurosa que protege los datos de la colaboración para que el miembro pueda obtener resultados al aprender sobre una persona específica.

Cifrado

Proceso de codificación de datos en un formato aparentemente aleatorio utilizando un valor secreto denominado clave. Es imposible determinar el texto sin formato original sin tener acceso a la clave.

Columna de huella digital

Una columna que está protegida criptográficamente para un JOIN Construcción SQL.

Método de flujo de trabajo de asignación de ID

Cómo desea que se realice la asignación de ID.

Existen dos métodos de flujo de trabajo de asignación de ID:

- Asignación de ID basado en reglas: el método por el que se utilizan reglas de coincidencia para traducir datos propios de un origen a un destino en un flujo de trabajo de asignación de ID.
- Asignación de ID de servicios de proveedores: el método por el que se utiliza un servicio de proveedores para traducir datos codificados de terceros de un origen a un destino en un flujo de trabajo de asignación de ID.

  AWS Clean Rooms actualmente admite LiveRamp como método de flujo de trabajo de mapeo de ID basado en servicios de proveedores. Debe tener una suscripción para LiveRamp poder AWS Data Exchange utilizar este método. Para obtener más información, consulte Suscripción a un servicio de proveedor en AWS Data Exchange en la Guía del usuario de AWS Entity Resolution .

tabla de asignación de ID

Recurso AWS Clean Rooms que permite establecer reglas de coincidencia propias o transcodificar identidades multipartitas en una colaboración.

Una tabla de asignación de ID es una referencia a una tabla existente en AWS Glue Data Catalog. Contiene una regla de análisis de tabla de asignación de ID que determina cómo se pueden consultar los datos en AWS Clean Rooms. Las tablas de asignación de ID se pueden asociar a una o más colaboraciones.

Regla de análisis de tablas de asignación de ID

Un tipo de regla de análisis se administra mediante AWS Clean Rooms y se usa para unir datos de identidad dispares a fin de facilitar las consultas. Se agrega automáticamente a las tablas de asignación de ID y no se puede editar. Hereda los comportamientos de las demás reglas de análisis de la colaboración, siempre que esas reglas de análisis sean homogéneas.

Flujo de trabajo de asignación de ID

Un trabajo de procesamiento de datos que asigna datos de un origen a un destino en función del método de flujo de trabajo de asignación de ID especificado. Produce una tabla de asignación de ID.

Espacio de nombres de ID

Un recurso AWS Clean Rooms que contiene metadatos que explican los conjuntos de datos de varios conjuntos de datos Cuentas de AWS y cómo utilizarlos en un flujo de trabajo de mapeo de identidades.

Asociación de espacios de nombres de ID

Una asociación de un recurso de espacio de nombres de ID que le ayuda a detectar las entradas del flujo de trabajo de asignación de ID.

Trabajo

Método para acceder a las tablas configuradas y analizarlas en una colaboración mediante un conjunto compatible de funciones, clases y variables.

AWS Clean Rooms actualmente admite el tipo de PySpark trabajo.

AWS Clean Rooms actualmente admite la ejecución de trabajos mediante una plantilla de PySpark análisis.

Regla de análisis de lista

La restricción de consultas que permite realizar consultas que generen un análisis de atributos de nivel de fila del solapamiento entre esta tabla y las tablas del miembro que puede realizar consultas.

Admite casos de uso como el enriquecimiento y la creación o supresión de audiencias.

Otros tipos de reglas de análisis son de agregación y personalizadas.

Modelo parecido

Un modelo de los datos de un proveedor de datos de entrenamiento que permite a un proveedor de datos iniciales crear un segmento similar de los datos del proveedor de datos de entrenamiento que se parezca más a sus datos iniciales.

Un segmento similar

Un subconjunto de los datos de entrenamiento que más se parece a los datos iniciales.

Miembro

Un AWS cliente que participa en una colaboración.

Un miembro se identifica mediante su Cuenta de AWS.

Todos los miembros pueden contribuir con datos.

Miembro que puede realizar consultas

El miembro que puede consultar los datos de la colaboración.

Solo hay un miembro que puede realizar consultas por colaboración y ese miembro es inmutable.

Un usuario administrativo puede usar los permisos AWS Identity and Access Management (de IAM) para controlar cuáles de sus directores de IAM (como los usuarios o las funciones) pueden consultar los datos de la colaboración. Para obtener más información, consulte Cree un rol de servicio para leer datos de Amazon S3.

Miembro que puede ejecutar consultas y trabajos

El miembro que puede ejecutar consultas y tareas con los datos de la colaboración.

Solo hay un miembro que puede ejecutar consultas y trabajos por colaboración, y ese miembro es inmutable.

Un usuario administrativo puede usar los permisos AWS Identity and Access Management (de IAM) para controlar cuáles de sus directores de IAM (como usuarios o roles) pueden ejecutar consultas y trabajos en la colaboración. Para obtener más información, consulte Cree un rol de servicio para leer datos de Amazon S3.

Miembro que puede recibir los resultados

Un miembro que puede recibir los resultados de las consultas. Un miembro que puede recibir los resultados especifica la configuración de los resultados de la consulta para el destino de Amazon S3 y el formato de los resultados de la consulta (CSV o Parquet).

En el caso de los análisis con el motor de análisis Spark, el miembro que puede recibir los resultados también especifica si los archivos deben generarse como un solo archivo o como varios archivos.

Puede haber más de un miembro que pueda recibir los resultados en una colaboración.

Miembro que paga los costos de computación de consultas

El miembro responsable de pagar los costos de computación de consultas.

Solo hay un miembro responsable de pagar los costos de computación de consultas por colaboración, y ese miembro es inmutable.

Si el creador de la colaboración no ha designado a nadie como miembro que paga los costos de computación de consultas, el miembro que puede realizar consultas es el pagador predeterminado.

El miembro que paga los costos de computación de las consultas recibe una factura por las consultas que se han ejecutado en la colaboración.

El miembro paga los costos de consulta y cálculo del trabajo

El miembro responsable de pagar los costos de consulta y cálculo del trabajo.

Solo hay un miembro que es responsable de pagar los costes de consulta y cálculo del trabajo por colaboración, y ese miembro es inmutable.

Si el creador de la colaboración no ha especificado a nadie como miembro que paga los costes de consulta y cálculo del trabajo, el miembro que puede realizar la consulta es el pagador predeterminado.

El miembro que pague los costes de consulta y cálculo del trabajo recibirá una factura por las consultas que se hayan realizado en la colaboración.

Pertenencia

Un recurso que se crea cuando un miembro se une a una colaboración.

Todos los recursos que el miembro asocia a una colaboración forman parte de la pertenencia o están asociados a ella.

Solo el miembro propietario de la pertenencia puede añadir, eliminar o editar los recursos de esa pertenencia.

Columna sellada

Una columna que está protegida criptográficamente para un SELECT Construcción SQL.

Datos iniciales

Los datos del proveedor de datos iniciales, que se utilizan para crear un segmento similar. Los datos iniciales se pueden proporcionar directamente o pueden provenir de los resultados de una consulta de AWS Clean Rooms . El resultado del segmento similar es un conjunto de usuarios a partir de los datos de entrenamiento que se parece más a los usuarios iniciales.

Motor de análisis Spark

Una opción de análisis AWS Clean Rooms que permite a los clientes ejecutar consultas complejas en grandes conjuntos de datos almacenados en Amazon S3, Amazon Athena o Snowflake mediante las funciones SQL de Apache Spark. Sirve como alternativa al motor de análisis de AWS Clean Rooms SQL y también admite el análisis interno. PySpark AWS Clean Rooms

Cuando creas una colaboración mediante la CreateCollaborationAPI, el valor del motor de análisis de Spark lo es`SPARK`.

Consultar

Método para acceder a las tablas configuradas de una colaboración y analizarlas, mediante un conjunto compatible de funciones, clases y variables.

AWS Clean Rooms actualmente es compatible con el lenguaje de consultas SQL.

AWS Clean Rooms actualmente admite la ejecución de consultas SQL directas o la ejecución de consultas mediante una plantilla de análisis SQL.

Las traducciones son generadas a través de traducción automática. En caso de conflicto entre la traducción y la version original de inglés, prevalecerá la version en inglés.