사용자 가이드

AWS Clean Rooms

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

AWS Clean Rooms: 사용자 가이드

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

Amazon의 상표 및 트레이드 드레스는 Amazon 외 제품 또는 서비스와 함께, Amazon 브랜드 이미지를 떨어뜨리거나 고객에게 혼동을 일으킬 수 있는 방식으로 사용할 수 없습니다. Amazon이 소유하지 않은 기타 모든 상표는 Amazon과 제휴 관계이거나 관련이 있거나 후원 관계와 관계없이 해당 소유자의 자산입니다.

Table of Contents

AWS Clean Rooms란 무엇인가요? 1
처음 AWS Clean Rooms 사용하시나요? 1
AWS Clean Rooms 작동 방식 2
관련 서비스 2
AWS 서비스 2
타사 서비스 4
액세스 AWS Clean Rooms 4
에 대한 요금 AWS Clean Rooms 4
에 대한 결제 AWS Clean Rooms 5
분석 엔진 5
분석 엔진 정보 6
추가 리소스 6
분석 규칙 7
분석 규칙 유형 8
집계 분석 규칙 10
목록 분석 규칙 27
사용자 지정 분석 규칙 35
ID 매핑 테이블 분석 규칙 42
AWS Clean Rooms 차등 프라이버시 51
차등 프라이버시 52
의 차등 프라이버시 AWS Clean Rooms 작동 방식 52
차등 프라이버시 정책 53
SQL 기능 54
SQL 쿼리 팁 및 예제 113
제한 사항 114
AWS Clean Rooms ML 116
AWS Clean Rooms ML이 AWS 모델과 작동하는 방식 117
AWS Clean Rooms ML이 사용자 지정 모델과 작동하는 방식 118
AWS Clean Rooms ML의 모델 120
Clean Rooms ML의 사용자 지정 모델 127
암호화 컴퓨팅 135
고려 사항 136
지원되는 파일 및 데이터 유형 138
열 이름 143

열 유형 144
파라미터 146
선택적 플래그 150
C3R을 사용한 쿼리 153
지침 154
설 AWS Clean Rooms정 179
에 가입 AWS 179
에 대한 서비스 역할 설정 AWS Clean Rooms 179
관리자 사용자 생성하기 180
공동 작업 구성원의 IAM 역할 생성 181
Amazon S3에서 데이터를 읽을 서비스 역할 생성 181
Amazon Athena에서 데이터를 읽을 서비스 역할 생성 185
서비스 역할을 생성하여 Snowflake에서 데이터 읽기 188
서비스 역할을 생성하여 S3 버킷에서 코드 읽기(PySpark 분석 템플릿 역할) 191
서비스 역할을 생성하여 PySpark 작업의 결과 작성 193
결과를 수신할 서비스 역할 생성 196
AWS Clean Rooms ML에 대한 서비스 역할 설정 199
유사 모델링을 위한 서비스 역할 설정 200
사용자 지정 모델링을 위한 서비스 역할 설정 213
공동 작업 및 멤버십 228
공동 작업 생성 228
쿼리에 대한 공동 작업 생성 229
쿼리 및 작업에 대한 공동 작업 생성 238
ML 모델링을 위한 공동 작업 생성 246
멤버십 생성 및 공동 작업 참여 255
..... 255
공동 작업 편집 260
공동 작업 이름 및 설명 편집 260
공동 작업 분석 엔진 업데이트 261
로그 스토리지 끄기 262
공동 작업 로그 설정 편집 262
공동 작업 태그 편집 263
멤버십 태그 편집 264
관련 테이블 태그 편집 265
분석 템플릿 태그 편집 265
차등 프라이버시 정책 태그 편집 266

공동 작업 삭제 ... 266
공동 작업 보기 ... 267
공동 작업에 멤버 초대 ... 267
구성원 모니터링 ... 268
공동 작업에서 구성원 제거 ... 268
공동 작업 탈퇴 ... 269
데이터 테이블 ... 271
데이터 형식 ... 272
PySpark 작업에 지원되는 데이터 형식 ... 272
SQL 쿼리에 지원되는 데이터 형식 ... 272
지원되는 데이터 유형 ... 273
에 대한 파일 압축 유형 AWS Clean Rooms ... 275
에 대한 서버 측 암호화 AWS Clean Rooms ... 275
Apache Iceberg 테이블 ... 276
Iceberg 테이블에 대해 지원되는 데이터 형식 ... 277
데이터 테이블 준비 ... 277
Amazon S3에서 데이터 테이블 준비 ... 278
Amazon Athena에서 데이터 테이블 준비 ... 280
Snowflake에서 데이터 테이블 준비 ... 282
암호화된 데이터 테이블 준비 ... 285
1단계: 필수 구성 요소 완성 ... 285
2단계: C3R 암호화 클라이언트 다운로드 ... 286
3단계: (선택 사항) C3R 암호화 클라이언트에서 사용 가능한 명령 보기 ... 286
4단계: 표 형식 파일의 암호화 스키마 생성 ... 287
5단계: 공유 암호 키 생성 ... 293
6단계: 환경 변수에 공유 암호 키 저장 ... 294
7단계: 데이터 암호화 ... 295
8단계: 데이터 암호화 확인 ... 296
(선택 사항) 스키마 생성(고급 사용자) ... 297
데이터 테이블 암호 해독 ... 307
구성된 테이블 ... 309
구성된 테이블 생성 ... 310
Amazon S3 데이터 소스 ... 310
Amazon Athena 데이터 소스 ... 313
Snowflake 데이터 소스 ... 315
구성된 테이블에 분석 규칙 추가 ... 319

테이블에 집계 분석 규칙 추가(안내식 흐름) 320
테이블에 목록 분석 규칙 추가(안내식 흐름) 323
테이블에 사용자 지정 분석 규칙 추가(안내식 흐름) 325
테이블에 분석 규칙 추가(JSON 편집기) 329
다음 단계 330
구성된 테이블을 공동 작업에 연결 330
구성된 테이블 세부 정보 페이지에서 구성된 테이블을 연결 332
공동 작업 세부 정보 페이지에서 구성된 테이블을 연결 335
다음 단계 338
구성된 테이블에 공동 작업 분석 규칙 추가 338
차등 프라이버시 정책 구성(선택 사항) 339
차등 프라이버시 사용량 로그 보기 340
차등 프라이버시 정책 편집 341
차등 프라이버시 정책 삭제 342
계산된 차등 프라이버시 파라미터 보기 342
테이블 및 분석 규칙 보기 343
구성된 테이블 세부 정보 편집 344
구성된 테이블 태그 편집 344
구성된 테이블 분석 규칙 편집 345
구성된 테이블 분석 규칙 삭제 345
구성된 테이블의 허용되지 않는 열 346
구성된 테이블 연결 편집 349
구성된 테이블 분리 350
AWS Entity Resolution in AWS Clean Rooms 351
ID 네임스페이스 352
새 ID 네임스페이스 생성 및 연결 352
기존 ID 네임스페이스 연결 355
ID 네임스페이스 연결 편집 357
ID 네임스페이스 연결 해제 358
ID 매핑 테이블 358
새 ID 매핑 테이블 생성 및 채우기 360
기존 ID 매핑 테이블 채우기 371
ID 매핑 테이블 편집 371
ID 매핑 테이블 삭제 372
분석 템플릿 373
SQL 분석 템플릿 373

SQL 분석 템플릿 생성 ... 374
SQL 분석 템플릿 검토 ... 375
PySpark 분석 템플릿 ... 376
보안 ... 377
제한 사항 ... 377
모범 사례 ... 379
사용자 스크립트 생성 ... 380
가상 환경 생성(선택 사항) ... 383
S3에 사용자 스크립트 및 가상 환경 저장 ... 384
PySpark 분석 템플릿 생성 ... 386
PySpark 분석 템플릿 검토 ... 388
PySpark 분석 템플릿 문제 해결 ... 391
코드 문제 해결 ... 391
분석 템플릿 작업이 시작되지 않음 ... 392
분석 템플릿 작업이 시작되지만 처리 중에 실패함 ... 393
가상 환경 설정 실패 ... 395
분석 ... 397
SQL 쿼리 실행 ... 397
구성된 테이블 쿼리 ... 399
ID 매핑 테이블 쿼리 ... 403
SQL 분석 템플릿을 사용하여 구성된 테이블 쿼리 ... 405
분석 빌더로 쿼리 ... 406
차등 프라이버시의 영향 확인 ... 412
최근 쿼리 보기 ... 412
쿼리 세부 정보 보기 ... 413
PySpark 작업 실행 ... 414
분석 템플릿을 사용하여 PySpark 작업 실행 ... 415
최근 작업 보기 ... 415
작업 세부 정보 보기 ... 416
분석 결과 ... 418
쿼리 결과 수신 ... 419
작업 결과 수신 ... 420
쿼리 결과 설정의 기본값 편집 ... 421
작업 결과 설정의 기본값 편집 ... 423
다른에서 쿼리 출력 사용 AWS 서비스 ... 423
훈련 데이터 공급자를 위한 ML 모델링 ... 425

학습 데이터 가져오기 ... 426
유사 모델 생성 ... 427
유사 모델 구성 ... 428
구성된 유사 모델 연결 ... 429
구성된 유사 모델 업데이트 ... 430
시드 데이터 공급자를 위한 ML 모델링 ... 431
유사 세그먼트 생성 ... 431
유사 세그먼트 내보내기 ... 432
사용자 지정 모델링 ... 434
공동 작업 생성 및 참여 ... 435
기계 학습을 위한 공동 작업 생성 ... 436
공동 작업 참여 ... 439
훈련 데이터 기여 ... 440
모델 알고리즘 구성 ... 445
구성된 모델 알고리즘 연결 ... 448
ML 입력 채널 생성 ... 451
훈련된 모델 생성 ... 454
증분 훈련 사용 ... 457
분산 훈련 사용 ... 461
모델 아티팩트 내보내기 ... 464
훈련된 모델에서 추론 실행 ... 466
문제 해결 ... 469
쿼리에서 참조하는 하나 이상의 테이블은 관련 서비스 역할로 액세스할 수 없습니다. 테이블/역할 소유자는 서비스 역할에 테이블에 대한 액세스 권한을 부여해야 합니다. ... 469
기본 데이터 세트 중 하나에 지원되지 않는 파일 형식이 있습니다. ... 469
Clean Rooms에 대한 암호화 컴퓨팅을 사용하는 경우 쿼리 결과가 예상과 다릅니다. ... 470
보안 ... 471
데이터 보호 ... 472
저장된 데이터 암호화 ... 472
전송 중 암호화 ... 473
기본 데이터 암호화 ... 473
키 정책 ... 474
데이터 보존 ... 477
모범 사례 ... 478
의 모범 사례 AWS Clean Rooms ... 478
AWS Clean Rooms에서 분석 규칙을 사용하는 모범 사례 ... 478

ID 및 액세스 관리 480
대상 480
ID를 통한 인증 481
정책을 사용하여 액세스 관리 484
AWS Clean Rooms 에서 IAM을 사용하는 방법 486
자격 증명 기반 정책 예제 492
AWS 관리형 정책 495
문제 해결 502
교차 서비스 혼동된 대리자 예방 504
AWS Clean Rooms ML에 대한 IAM 동작 506
Clean Rooms ML Custom 모델의 IAM 동작 508
규정 준수 확인 510
복원성 511
인프라 보안 511
네트워크 보안 512
AWS PrivateLink 512
고려 사항 512
인터페이스 엔드포인트 생성 513
모니터링 514
의 분석 로깅 AWS Clean Rooms 514
쿼리 및 작업 로그 수신 515
쿼리 및 작업 로그에 대한 권장 작업 516
CloudTrail 로그 517
AWS Clean Rooms CloudTrail의 정보 517
AWS Clean Rooms 로그 파일 항목 이해 518
Example AWS Clean Rooms CloudTrail 이벤트 518
AWS CloudFormation 리소스 522
AWS Clean Rooms 및 AWS CloudFormation 템플릿 522
에 대해 자세히 알아보기 AWS CloudFormation 522
할당량 523
AWS Clean Rooms 할당량 523
AWS Clean Rooms 리소스 파라미터 제한 530
AWS Clean Rooms API 제한 할당량 530
AWS Clean Rooms ML 할당량 533
Clean Rooms ML API 제한 할당량 552
문서 기록 558

용어집 567
집계 분석 규칙 567
분석 규칙 567
분석 템플릿 567
AWS Clean Rooms SQL 분석 엔진 567
C3R 암호화 클라이언트 568
일반 텍스트 열 568
공동 작업 568
공동 작업 생성자 568
구성된 테이블 569
사용자 지정 분석 규칙 569
해독 569
차등 프라이버시 570
암호화 570
핑거프린트 컬럼 570
ID 매핑 워크플로 방법 570
ID 매핑 테이블 570
ID 매핑 테이블 분석 규칙 571
ID 매핑 워크플로 571
ID 네임스페이스 571
ID 네임스페이스 연결 571
작업 571
목록 분석 규칙 571
유사 모델 572
유사 세그먼트 572
Member 572
쿼리할 수 있는 회원 572
쿼리 및 작업을 실행할 수 있는 구성원 572
결과를 받을 수 있는 구성원 573
구성원은 쿼리 컴퓨팅 비용을 지불합니다. 573
쿼리 및 작업 컴퓨팅 비용을 지불하는 구성원 573
멤버십 573
봉인 열 574
시드 데이터 574
Spark 분석 엔진 574
Query 574

..... dlxxv

AWS Clean Rooms란 무엇인가요?

AWS Clean Rooms 를 사용하면 사용자와 파트너가 집합 데이터 세트를 분석하고 공동 작업하여 기본 데이터를 서로 공개하지 않고도 새로운 인사이트를 얻을 수 있습니다. AWS Clean Rooms 는 몇 분 만에 자체 클린룸을 생성하고 몇 단계만으로 집합 데이터 세트를 분석하는 안전한 공동 작업 영역입니다. 협업할 파트너를 선택하고, 데이터 세트를 선택하고, 해당 파트너에 대한 개인 정보 보호 강화 제어를 구성합니다.

AWS Clean Rooms를 사용하면 이미 사용 중인 수천 개의 회사와 협업할 수 있습니다 AWS. 공동 작업에서는 데이터를 다른 클라우드 서비스 공급자로 이동 AWS 하거나 로드할 필요가 없습니다. 쿼리 또는 작업을 실행할 때는 해당 데이터의 원래 위치에서 데이터를 AWS Clean Rooms 읽고 기본 제공 분석 규칙을 적용하여 해당 데이터에 대한 제어를 유지합니다.

AWS Clean Rooms 는 구성할 수 있는 기본 제공 데이터 액세스 제어 및 감사 지원 제어를 제공합니다. 이러한 제어에는 다음이 포함됩니다.

- SQL 쿼리를 제한하고 출력 제약 조건을 제공하는 분석 규칙입니다.
- 용 암호화 컴퓨팅Clean Rooms은 쿼리가 처리되더라도 엄격한 데이터 처리 정책을 준수하기 위해 데이터를 암호화합니다.
- 에서 쿼리 및 작업을 검토하고 감사를 AWS Clean Rooms 지원하는 분석 로그입니다.
- 사용자 식별 시도로부터 보호하기 위한 차등 프라이버시입니다. AWS Clean Rooms 차등 프라이버시는 몇 가지 단계에서 적용할 수 있는 수학적으로 지원되는 기술과 직관적인 제어를 통해 사용자의 프라이버시를 보호하는 완전 관리형 기능입니다.
- AWS Clean Rooms ML을 사용하면 두 당사자가 데이터를 서로 공유할 필요 없이 데이터에서 유사한 사용자를 식별할 수 있습니다. 첫 번째 당사자는 훈련 데이터에서 유사 모델을 만들고 구성합니다. 그런 다음 시드 데이터를 공동 작업에 가져와서 훈련 데이터와 유사한 유사 세그먼트를 만듭니다.

다음 동영상에서는에 대해 자세히 설명합니다 AWS Clean Rooms.

AWS Clean Rooms

처음 AWS Clean Rooms 사용하시나요?

를 처음 사용하는 경우 먼저 다음 섹션을 읽는 것이 AWS Clean Rooms좋습니다.

- AWS Clean Rooms 작동 방식

- 액세스 AWS Clean Rooms
- 설 AWS Clean Rooms정
- AWS Clean Rooms 용어집

AWS Clean Rooms 작동 방식

에서 공동 작업을 AWS Clean Rooms생성하고 초대하려는 AWS 계정 를 추가하거나 초대받은 공동 작업에 참여할 멤버십을 생성합니다. 그런 다음 이벤트 데이터에 대해 구성된 테이블, ML 모델링에 대해 구성된 모델 또는 엔터티 확인을 위한 ID 네임스페이스 등 사용 사례에 필요한 데이터 리소스를 연결합니다. 분석 템플릿을 생성하거나 승인하여 공동 작업에서 허용하려는 정확한 쿼리 및 작업에 미리 동의할 수 있습니다. 마지막으로 구성된 테이블에서 SQL 쿼리 또는 PySpark 작업을 실행하거나, ID 매핑 테이블에서 개체 확인을 수행하거나, ML 모델링을 사용하여 유사 대상 세그먼트를 생성하여 관절 데이터를 분석합니다.

다음 다이어그램은 AWS Clean Rooms 가 작동하는 방식을 보여줍니다.

관련 서비스

AWS 서비스

다음은 다음과 AWS 서비스 관련이 있습니다 AWS Clean Rooms.

- Amazon Athena

  협업 구성원은 Amazon Athena에서 AWS Glue Data Catalog 뷰 AWS Clean Rooms 로 가져오는 데이터를 저장할 수 있습니다. 자세한 정보는 다음의 주제를 참조하세요.

  자세한 정보는 다음의 주제를 참조하세요.

  에서 쿼리를 위한 데이터 테이블 준비 AWS Clean Rooms

  구성된 테이블 생성 - Amazon Athena 데이터 소스

  Amazon Athena 사용 설명서의 Amazon Athena란 무엇인가요? Amazon Athena
- AWS CloudFormation

  에서 공동 AWS CloudFormation작업, 구성된 테이블, 구성된 테이블 연결 및 멤버십 리소스를 생성합니다.

자세한 내용은 를 사용하여 AWS Clean Rooms 리소스 생성 AWS CloudFormation 단원을 참조하십시오.

- AWS CloudTrail

  CloudTrail 로그 AWS Clean Rooms 와 함께를 사용하여 AWS 서비스 활동 분석을 개선합니다.

  자세한 내용은 를 사용하여 AWS Clean Rooms API 호출 로깅 AWS CloudTrail 단원을 참조하십시오.

- AWS Entity Resolution

  와 AWS Clean Rooms 함께 AWS Entity Resolution 를 사용하여 개체 확인을 수행합니다.

  자세한 내용은 AWS Entity Resolution in AWS Clean Rooms 단원을 참조하십시오.

- AWS Glue

  공동 작업 구성원은에서 사용할 Amazon S3의 데이터에서 AWS Glue 테이블을 생성할 수 있습니다 AWS Clean Rooms.

  자세한 정보는 다음의 주제를 참조하세요.

  에서 쿼리를 위한 데이터 테이블 준비 AWS Clean Rooms

  AWS Glue 개발자 가이드의 AWS Glue(이)란 무엇인가요?

- Amazon Simple Storage Service(S3)

  공동 작업 구성원은 Amazon S3 AWS Clean Rooms 에서 가져온 데이터를 저장할 수 있습니다.

  자세한 정보는 다음의 주제를 참조하세요.

  에서 쿼리를 위한 데이터 테이블 준비 AWS Clean Rooms

  구성된 테이블 생성 - Amazon S3 데이터 소스

  Amazon Simple Storage Service 사용 설명서의 Amazon S3란 무엇입니까?

- AWS Secrets Manager

  공동 작업 구성원은 보안 암호를 생성하여 Snowflake에 저장된 데이터에 액세스하고 읽을 수 있습니다.

  자세한 정보는 다음의 주제를 참조하세요.

서비스 역할을 생성하여 Snowflake에서 데이터 읽기

에서 쿼리를 위한 데이터 테이블 준비 AWS Clean Rooms

AWS Secrets Manager 사용 설명서의 AWS Secrets Manager란 무엇입니까?

타사 서비스

다음 타사 서비스는 AWS Clean Rooms다음과 관련이 있습니다.

- Snowflake

  공동 작업 구성원은 Snowflake 웨어하우스 AWS Clean Rooms 에 가져온 데이터를 저장할 수 있습니다.

  자세한 정보는 다음의 주제를 참조하세요.

  에서 쿼리를 위한 데이터 테이블 준비 AWS Clean Rooms

  구성된 테이블 생성 - Snowflake 데이터 소스

액세스 AWS Clean Rooms

다음 옵션을 AWS Clean Rooms 통해에 액세스할 수 있습니다.

- https://console.aws.amazon.com/cleanrooms/ AWS Clean Rooms 콘솔을 통해 직접.
- AWS Clean Rooms API를 통해 프로그래밍 방식으로. 자세한 내용은 AWS Clean Rooms API 참조를 참조하세요.

에 대한 요금 AWS Clean Rooms

요금 정보는 AWS Clean Rooms 요금을 참조하세요.

> **Note**
>
> Snowflake에 저장된 데이터를 연결한 공동 작업 구성원의 경우 해당 위치에 저장된 데이터를 사용하는 쿼리가 실행될 때마다 해당 데이터 웨어하우스 공급자 또는 클라우드 공급자가 데이터 송신 및 컴퓨팅 모두에 대해 요금을 청구합니다.

에 대한 결제 AWS Clean Rooms

AWS Clean Rooms 는 공동 작업 생성자에게 공동 작업에서 쿼리 또는 작업 컴퓨팅 비용을 지불할 구성원을 지정할 수 있는 기능을 제공합니다.

대부분의 경우 쿼리를 할 수 있는 구성원과 쿼리 컴퓨팅 비용을 지불하는 구성원은 동일합니다. 하지만 쿼리를 할 수 있는 구성원과 쿼리 계산 비용을 지불하는 구성원이 다르면 쿼리를 할 수 있는 구성원이 자신의 구성원 리소스에 대해 쿼리를 실행하면 쿼리 계산 비용을 지불하는 구성원의 구성원 리소스에 요금이 청구됩니다.

쿼리 컴퓨팅 비용을 지불하는 회원은 CloudTrail Event 기록에서 쿼리가 실행되는 이벤트를 볼 수 없습니다. 지불자가 쿼리를 실행하는 사람도 아니고 쿼리가 실행되는 리소스의 소유자도 아니기 때문입니다. 그러나 지급인은 공동 작업에서 쿼리를 실행할 수 있는 구성원이 실행하는 모든 쿼리에 대해 멤버십 리소스에 생성된 요금을 볼 수 있습니다.

컬래버레이션을 생성하고 구성원이 쿼리 컴퓨팅 비용을 지불하도록 구성하는 방법에 대한 자세한 내용은 공동 작업 생성 섹션을 참조하세요.

의 분석 엔진 AWS Clean Rooms

> **Important**
>
> AWS Clean Rooms 는 2025년 12월 17일에 레거시 Clean Rooms SQL 분석 엔진에 대한 지원을 종료합니다. 2025년 7월 16일 이전에 AWS 는 고객 지원을 통해 한도 증가를 요청하여 새로운 Clean Rooms SQL 엔진 기반 공동 작업을 생성해야 합니다. 2025년 7월 17일부터 새로운 Clean Rooms SQL 엔진 기반 공동 작업을 더 이상 생성할 수 없습니다.

분석 엔진 정보

분석 엔진은 데이터 쿼리를 처리하고 내부에서 분석 계산을 수행하는 소프트웨어 구성 요소입니다 AWS Clean Rooms. 분석 엔진은 SQL 명령을 해석하고, 데이터 처리 작업을 실행하고, 분석 결과를 반환합니다.

다음 표에서는 Spark 분석 엔진(권장)을 레거시 Clean Rooms SQL 분석 엔진과 비교합니다.

분석 엔진	언제 사용하나요?	집계 분석 규칙이 지원되나요?	지원되는 분석 규칙 나열	차등 프라이버시가 지원되지 않는 사용자 지정 분석 규칙	차등 프라이버시가 지원되는 사용자 지정 분석 규칙	Amazon S3 데이터 소스가 지원되나요?	Amazon Athena 및 Snowflake 데이터 소스가 지원되나요?
Spark 분석 엔진	• Spark SQL 쿼리 실행 • PySpark 작업 실행 • 사용자 지정 ML 모델링	예	예	예	예	예	예
AWS Clean Rooms SQL 분석 엔진	AWS Clean Rooms SQL 쿼리 실행	예	예	예	예	예	아니요

추가 리소스

Spark SQL 쿼리에 대한 자세한 내용은 AWS Clean Rooms Spark SQL 참조를 참조하세요.

AWS Clean Rooms SQL 쿼리에 대한 자세한 내용은 AWS Clean Rooms SQL 참조를 참조하세요.

Spark SQL 및 AWS Clean Rooms SQL에 대한 요금 정보는 AWS Clean Rooms 요금을 참조하세요.

의 분석 규칙 AWS Clean Rooms

공동 작업 분석에에서 테이블을 사용할 수 있도록 하기 AWS Clean Rooms 위해 공동 작업 구성원은 분석 규칙을 구성해야 합니다.

분석 규칙은 각 데이터 소유자가 구성된 테이블에 설정하는 개인 정보 보호 강화 컨트롤입니다. 분석 규칙은 구성된 테이블을 분석할 수 있는 방법을 결정합니다.

분석 규칙은 구성된 테이블(계정 수준 리소스)의 계정 수준 컨트롤이며, 구성된 테이블이 연결된 모든 공동 작업에 적용됩니다. 구성된 분석 규칙이 없는 경우 구성된 테이블을 공동 작업에 연결할 수는 있지만 쿼리할 수는 없습니다. 쿼리는 분석 규칙 유형이 동일한 구성된 테이블만 참조할 수 있습니다.

분석 규칙을 구성하려면 먼저 분석 유형을 선택한 다음 분석 규칙을 지정합니다. 두 단계 모두에서 활성화하려는 사용 사례와 기본 데이터를 보호하는 방법을 고려해야 합니다.

AWS Clean Rooms 는 쿼리에서 참조되는 구성된 모든 테이블에 대해 보다 제한적인 제어를 적용합니다.

다음 예는 제한적 제어를 설명합니다.

Example 제한적 제어: 출력 제약조건

- 협업자 A의 식별자 열에 대한 출력 제한은 100입니다.
- 협업자 B의 식별자 열에 대한 출력 제한은 150입니다.

  구성된 두 테이블을 모두 참조하는 집계 쿼리의 경우 쿼리 출력에 표시되려면 출력 행 내에 최소 150개의 고유한 식별자 값이 있어야 합니다. 쿼리 출력은 출력 제한으로 인해 결과가 제거되었음을 나타내지 않습니다.

Example 제한적 제어: 분석 템플릿이 승인되지 않았습니다

- 협업자 A는 사용자 지정 분석 규칙에서 협업자 A와 협업자 B의 구성된 테이블을 참조하는 쿼리가 포함된 분석 템플릿을 허용했습니다.
- 협업자 B는 분석 템플릿을 허용하지 않았습니다.

  협업자 B가 분석 템플릿을 허용하지 않았기 때문에 쿼리할 수 있는 구성원은 해당 분석 템플릿을 실행할 수 없습니다.

분석 규칙 유형

분석 규칙에는 집계, 목록, 사용자 지정의 세 가지 유형이 있습니다. 다음 표는 분석 규칙 유형을 비교합니다. 각 유형에는 분석 규칙 지정을 설명하는 별도의 섹션이 있습니다.

> **Note**
>
> ID 매핑 테이블 분석 규칙이라는 분석 규칙 유형이 있습니다. 그러나이 분석 규칙은에서 관리 AWS Clean Rooms 하며 수정할 수 없습니다. 자세한 내용은 ID 매핑 테이블 분석 규칙 단원을 참조하십시오.

다음 단원에서는 분석 규칙 유형별 지원되는 사용 사례 및 컨트롤에 대해 설명합니다.

지원되는 사용 사례

다음 표에는 각 분석 규칙 유형별로 지원되는 사용 사례를 비교한 요약이 나와 있습니다.

사용 사례	집계	목록	사용자 지정
지원되는 분석	필요에 따라 측정기준과 함께 COUNT, SUM, AVG 함수를 사용하여 통계를 집계하는 쿼리	여러 테이블 간의 겹침에 대한 행 수준 목록을 출력하는 쿼리	모든 사용자 지정 분석(분석 템플릿 또는 분석 생성자가 검토되고 허용된 경우에 한함)
일반 사용 사례	세그먼트 분석, 측정, 속성	강화, 세그먼트 구축	퍼스트 터치 어트리뷰션, 증분 분석, 오디언스 디스커버리
SQL 구조	• JOIN 문: INNER JOIN • 집계 함수: COUNT/COUNT DISTINCT, SUM/SUM DISTINCT, AVG • 스칼라 함수: 제한된 서브셋	• JOIN 문: INNER JOIN • 스칼라 함수: 없음	SELECT 명령으로 대부분의 SQL 함수 및 SQL 구문을 사용할 수 있습니다

사용 사례	집계	목록	사용자 지정
하위 쿼리 및 공통 테이블 표현식(CTE)	아니요	아니요	예
분석 템플릿	아니요	아니요	예

지원되는 컨트롤

다음 표는 각 분석 규칙 유형이 기본 데이터를 보호하는 방법을 비교 요약한 것입니다.

컨트롤	집계	목록	사용자 지정
제어 메커니즘	테이블의 데이터를 쿼리에 사용할 수 있는 방법을 제어합니다 (예를 들어, hashed_email 열의 개수 및 합계를 허용합니다.)	테이블의 데이터를 쿼리에 사용할 수 있는 방법을 제어합니다 (예를 들어, hashed_email 열은 조인에만 사용할 수 있습니다.)	테이블에서 실행할 수 있는 쿼리를 제어합니다 (예: 분석 템플릿 "사용자 지정 쿼리 1"에 정의된 쿼리만 허용)
기본 제공되는 개인 정보 보호 강화 기술	• 블라인드 매치 • 집계 필요 • 최소 집계 임계값 >= • 2) 사전 정의된 쿼리 구조	• 블라인드 매치 • 오버랩 필요 • 사전 정의된 쿼리 구조 • 추가 분석이 허용됨	• 차등 프라이버시 • 허용되지 않는 출력 열
쿼리를 실행하기 전에 검토하세요	아니요	아니요	예, 분석 템플릿 사용

에서 사용할 수 있는 분석 규칙에 대한 자세한 내용은 다음 주제를 AWS Clean Rooms참조하세요.

- 집계 분석 규칙
- 목록 분석 규칙

- 의 사용자 지정 분석 규칙 AWS Clean Rooms

집계 분석 규칙

에서 집계 분석 규칙 AWS Clean Rooms은 선택적 차원과 함께 COUNT, SUM 및/또는 AVG 함수를 사용하여 집계 통계를 생성합니다. 구성된 테이블에 집계 분석 규칙을 추가하면 쿼리를 할 수 있는 구성원이 구성된 테이블에서 쿼리를 실행할 수 있습니다.

집계 분석 규칙은 캠페인 계획, 미디어 도달 범위, 빈도 측정, 속성 등의 사용 사례를 지원합니다.

지원되는 쿼리 구조 및 구문은 집계 쿼리 구조 및 구문에 정의되어 있습니다.

집계 분석 규칙 - 쿼리 제어에 정의된 분석 규칙의 매개 변수에는 쿼리 컨트롤 및 쿼리 결과 컨트롤이 포함됩니다. 쿼리 컨트롤에는 쿼리할 수 있는 구성원이 소유한 하나 이상의 구성된 테이블에 구성된 테이블을 직접 또는 전이적으로 조인하도록 요구하는 기능이 포함됩니다. 이 요구 사항을 통해 쿼리가 테이블과 해당 테이블의 교차점(INNERJOIN)에서 실행되도록 할 수 있습니다.

집계 쿼리 구조 및 구문

집계 분석 규칙이 있는 테이블에 대한 쿼리는 다음 구문을 준수해야 합니다.

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

다음 표에서는 위 구문에 나열된 각 표현식에 대해 설명합니다.

표현식	정의	예시
select_aggregate_function_expression	다음 표현식을 포함하는 쉼표로 구분된 쉼표로 구분된 목록입니다. • select_aggregation_function_expression • select_aggregate_expression **Note** select_aggregate_expression 에 select_aggregation_function_expression 이 하나 이상 있어야 합니다.	SELECT SUM(PRICE), user_segment
select_aggregation_function_expression	지원되는 집계 함수가 하나 이상의 열에 하나 이상 적용되었습니다. 열만 집계 함수의 인수로 사용할 수 있습니다. **Note** select_aggregate_e	AVG(PRICE) COUNT(DISTINCT user_id)

표현식	정의	예시
	`xpression` 에 `select_ag gregation _function _expression` 이 하나 이상 있어야 합니다.	

`select_grouping_co lumn_expression`	다음을 사용하는 모든 식을 포함할 수 있는 표현식:
	• 테이블 열 이름
	• 지원되는 스칼라 함수
	• 리터럴 문자열입니다
	• 수치 리터럴
	
	Note
	`select_ag gregate_e xpression` 은 AS 매개 변수를 사용하거나 사용하지 않고 열에 별칭을 지정할 수 있습니다. 자세한 내용은 AWS Clean Rooms SQL 참조 섹션을 참조하세요.
		`TRUNC(timestampCol umn)`
		
		`UPPER(campaignName)`

표현식	정의	예시
table_expression	테이블 또는 테이블의 조인으로, 조인 조건식을 join_condition 과 연결합니다. join_condition 은 BOOLEAN을 반환합니다. table_expression 은 다음을 지원합니다. • 특정 JOIN 유형(INNERJOIN) • join_condition (=) 내의 동등 비교 조건 • 논리 연산자(AND, OR).	FROM consumer_table INNER JOIN provider_table ON consumer_table.identifier1 = provider_table.identifier1 AND consumer_table.identifier2 = provider_table.identifier2

표현식	정의	예시
where_expression	부울을 반환하는 조건식. 다음과 같이 구성될 수 있습니다. • 테이블 열 이름 • 지원되는 스칼라 함수 • 수학적 연산 • 문자열 리터럴 • 숫자 리터럴 지원되는 비교 조건은 (=, >, <, <=, >=, <>, !=, NOT, IN, NOT IN, LIKE, IS NULL, IS NOT NULL)입니다. 지원되는 논리 연산자는 (AND, OR)입니다. where_expression 은 선택 사항입니다.	WHERE where_condition WHERE price > 100 WHERE TRUNC(timestampColumn) = '1/1/2022' WHERE timestampColumn = timestampColumn2 - 14
group_by_expression	select_grouping_column_expression 의 요구 사항과 일치하는 표혁식 목록으로, 쉼표로 구분됩니다.	GROUP BY TRUNC(timestampColumn), UPPER(campaignName), segment

표현식	정의	예시
having_expression	조건은 부울 결과가 있는 표현식입니다. 지원되는 집계 함수가 단일 열(예:SUM(price))에 적용되며 숫자 리터럴과 비교됩니다. 지원되는 조건은 (=, >, <, <=, >=, <>, !=)입니다. 지원되는 논리 연산자는 (AND, OR)입니다. having_expression 은 선택 사항입니다.	HAVING SUM(SALES) > 500
order_by_expression	앞에서 select_aggregate_expression 에서 정의한 것과 동일한 요구 사항과 호환되는 표현식 목록으로 쉼표로 구분됩니다. order_by_expression 은 선택 사항입니다. **Note** order_by_expression 은 ASC와 DESC 매개변수를 허가합니다. 자세한 내용은 AWS Clean Rooms SQL 참조의 ASC DESC 매개변수를 참조하세요.	ORDER BY SUM(SALES), UPPER(campaignName)

집계 쿼리 구조 및 구문의 경우 다음 사항에 유의해야 합니다.

- SELECT 이외의 SQL 명령은 지원되지 않습니다.
- 하위 쿼리 및 공통 테이블 표현식(예:WITH)은 지원되지 않습니다.
- 여러 쿼리를 결합하는 연산자(예:UNION)는 지원되지 않습니다.
- TOP, LIMIT, 및 OFFSET 파라미터는 지원되지 않습니다.

집계 분석 규칙 - 쿼리 제어

집계 쿼리 컨트롤을 사용하면 테이블의 열을 사용하여 테이블을 쿼리하는 방법을 제어할 수 있습니다. 예를 들어 조인에 사용할 열, 계산할 수 있는 열 또는 WHERE 명령문에 사용할 수 있는 열을 제어할 수 있습니다.

다음 단원에서는 각 컨트롤에 대해 설명합니다.

주제

- 집계 제어
- 조인 컨트롤
- 치수 제어
- 스칼라 함수

집계 제어

집계 제어를 사용하면 허용할 집계 함수와 해당 집계 함수를 적용해야 하는 열을 정의할 수 있습니다. 집계 함수는 SELECT, HAVING, 및 ORDER BY 표현식에서 사용할 수 있습니다.

컨트롤	정의	사용법
`aggregateColumns`	집계 함수 내에서 사용할 수 있도록 구성된 테이블 열의 열	`aggregateColumns` 은 SELECT, HAVING, 및 ORDER BY 표현식의 집계 함수 내에서 사용할 수 있습니다. 일부 `aggregateColumns` 은 `joinColumn` (나중에 정의됨)으로 분류할 수도 있습니다.

컨트롤	정의	사용법
		주어진 aggregate Column 을 dimension Column (나중에 정의됨)으로도 분류할 수 없습니다.
function	카운트, 합계 및 평균 함수는 aggregateColumns 위에서 사용할 수 있습니다.	function은 관련된 aggregateColumns 에 적용할 수 있습니다.

조인 컨트롤

JOIN 절은 두 개 이상의 테이블에서 관련 열을 기반으로 두 개 이상의 테이블에서 행을 결합하는 데 사용됩니다.

조인 컨트롤을 사용하여 테이블을의 다른 테이블에 조인하는 방법을 제어할 수 있습니다table_expression. AWS Clean Rooms 만 지원합니다INNERJOIN. INNER JOIN 문은 사용자가 정의한 컨트롤에 따라 분석 규칙joinColumn에서 로 명시적으로 분류된 열만 사용할 수 있습니다.

INNER JOIN은 사용자가 구성한 테이블의 joinColumn과 공동 작업의 다른 구성된 테이블의 joinColumn에서 작동해야 합니다. 테이블에서 joinColumn으로 사용할 수 있는 열을 결정합니다.

ON 절의 각 일치 조건은 두 열 간의 등식 비교 조건(=)을 사용해야 합니다.

ON 절 내의 여러 일치 조건은 다음과 같을 수 있습니다.

- AND 논리 연산자를 사용하여 조합합니다
- OR 논리 연산자를 사용하여 구분합니다

> **Note**
>
> 모든 JOIN 일치 조건은 JOIN의 각 변의 한 행과 일치해야 합니다. OR 또는 AND 논리 연산자로 연결된 모든 조건문도 이 요구 사항을 준수해야 합니다.

다음은 AND 논리 연산자를 사용한 쿼리의 예입니다.

```
SELECT some_col, other_col
```

```
FROM table1
    JOIN table2
    ON table1.id = table2.id AND table1.name = table2.name
```

다음은 OR 논리 연산자를 사용하는 쿼리의 예입니다.

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

컨트롤	정의	사용법
joinColumns	쿼리를 할 수 있는 구성원이 INNER JOIN 명령문에서 사용할 수 있도록 허용하려는 열(있는 경우)	특정 joinColumn 을 aggregateColumn 으로 분류할 수도 있습니다(집계 제어 참조). 동일한 열을 joinColumn 및 dimensionColumns 으로 모두 사용할 수는 없습니다(나중에 참조). aggregateColumn 으로 분류되지 않는 한, joinColumn 은 INNER JOIN가 아닌 쿼리의 다른 부분에서는 사용할 수 없습니다.
joinRequired	쿼리할 수 있는 멤버에게 구성된 테이블이 있는 INNER JOIN이 필요한지 여부를 제어합니다.	이 파라미터는 이 지정된 경우에만 INNER JOIN이 필요합니다. 이 매개 변수를 활성화하지 않는 경우 INNER JOIN은 선택 사항입니다. 이 매개 변수를 활성화하면 쿼리할 수 있는 구성원은 자신이 소유한 테이블을 INNER JOIN에 포함해야 합니다. 그들은 직

컨트롤	정의	사용법
		접적으로든 간접적으로든(즉, 그들의 테이블을 다른 테이블에 조인하거나 그것이 자신의 테이블과 조인된 경우) 자신의 테이블과 JOIN해야 합니다

다음은 전이성의 예입니다.

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

> **Note**
>
> 쿼리할 수 있는 멤버도 `joinRequired` 파라미터를 사용할 수 있습니다. 이 경우 쿼리는 테이블을 하나 이상의 다른 테이블과 조인해야 합니다.

치수 제어

차원 컨트롤은 집계 열을 필터링, 그룹화 또는 집계할 수 있는 열을 제어합니다.

컨트롤	정의	사용법
`dimensionColumns`	쿼리할 수 있는 구성원이 SELECT,WHERE, GROUP BY 및 ORDER BY에서 사용할 수 있도록 허용하는 열(있는 경우).	`dimensionColumn` 은 SELECT(`select_grouping_column_expression`), WHERE, GROUPBY, 및 ORDER BY에서 사용할 수 있습니다. 동일한 열에 `dimensionColumn` , `joinColumn` , 및/

컨트롤	정의	사용법
		또는aggregateColumn 이 모두 있을 수는 없습니다.

스칼라 함수

스칼라 함수는 차원 열에 사용할 수 있는 스칼라 함수를 제어합니다.

컨트롤	정의	사용법
scalarFunctions	쿼리의 dimension Columns 에서 사용할 수 있는 스칼라 함수.	dimensionColumns 에 적용할 수 있는 스칼라 함수(있는 경우)를 지정합니다(예:CAST). 스칼라 함수는 다른 함수 위나 다른 함수 내에서 사용할 수 없습니다. 스칼라 함수의 인수는 열, 문자열 리터럴 또는 숫자형 리터럴일 수 있습니다.

지원되는 스칼라 함수는 다음과 같습니다.

- 수학 함수 — AABS, CEILING, FLOOR, LOG, LN, ROUND, SQRT
- 데이터 형식 지정 함수 – CAST, CONVERT, TO_CHAR, TO_DATE, TO_NUMBER, TO_TIMESTAMP
- 문자열 함수 - LOWER, UPPER, TRIM, RTRIM, SUBSTRING
  - RTRIM의 경우 트리밍할 사용자 정의 문자 세트를 사용할 수 없습니다.
- 조건부 표현식 – COALESCE
- 날짜 함수 - EXTRACT, GETDATE, CURRENT_DATE, DATEADD
- 기타 함수 — TRUNC

자세한 내용은 AWS Clean Rooms SQL 참조를 참조하세요.

집계 분석 규칙 - 쿼리 결과 제어

집계 쿼리 결과 컨트롤을 사용하면 각 출력 행이 충족해야 반환되는 조건을 하나 이상 지정하여 반환되는 결과를 제어할 수 있습니다. AWS Clean Rooms 은 `COUNT (DISTINCT column) >= X`와 같은 형식의 집계 제약 조건을 지원합니다. 이 양식을 사용하려면 각 행이 구성된 테이블에서 선택한 고유한 값을 X개 이상 집계해야 합니다(예: 고유 `user_id` 값의 최소 수). 제출된 쿼리 자체에서 지정된 열을 사용하지 않는 경우에도 이 최소 임계값은 자동으로 적용됩니다. 이 규칙들은 집합적으로 적용되며 쿼리에서 각 구성된 테이블에 걸쳐 적용되며, 공동 작업 참여자 각각의 구성된 테이블을 포함합니다.

구성된 각 테이블의 분석 규칙에는 하나 이상의 집계 제약조건이 있어야 합니다. 구성된 테이블 소유자는 여러 개의 `columnName`과 연결된 `minimum`을 추가할 수 있으며, 이는 일괄적으로 적용됩니다.

집계 제약 조건

집계 제약 조건은 쿼리 결과에서 반환되는 행을 제어합니다. 행이 반환되려면 집계 제약 조건에 지정된 각 열의 고유 값의 지정된 최소 개수를 충족해야 합니다. 이 요구 사항은 쿼리나 분석 규칙의 다른 부분에서 열이 명시적으로 언급되지 않은 경우에도 적용됩니다.

컨트롤	정의	사용법
`columnName`	각 출력 행이 충족해야 하는 조건에 사용되는 `aggregateColumn` 입니다.	구성된 테이블의 모든 열이 될 수 있습니다.
`minimum`	쿼리 결과에서 반환되기 위해 출력 행에 있어야 하는 연관된 `aggregateColumn` 의 최소 고유값 수(예: COUNT DISTINCT)입니다.	`minimum`은 최소 2의 값이어야 합니다.

집계 분석 규칙 구조

다음 예제에서는 집계 분석 규칙의 사전 정의된 구조를 보여줍니다.

다음 예제에서는 데이터 *`MyTable`*은 데이터 테이블을 나타냅니다. *`user input placeholder`*를 사용자의 정보로 바꿉니다.

```
{
```

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

집계 분석 규칙 - 예제

다음 예제에서는 두 회사가 집계 분석을 AWS Clean Rooms 사용하여에서 협업하는 방법을 보여줍니다.

회사 A에는 고객 및 판매 데이터가 있습니다. 회사 A는 제품 반품 활동을 이해하는 데 관심이 있습니다. 회사 B는 회사 A의 소매업체 중 하나이며 반품 데이터를 보유하고 있습니다. 또한 B사는 A에 유용한 고객 세그먼트 속성을 가지고 있습니다(예: 관련 제품 구매, 소매업체의 고객 서비스 이용). 회사 B는 행 수준의 고객 반품 데이터 및 속성 정보를 제공하고 싶지 않습니다. 회사 B는 회사 A가 최소 집계 임계값으로 중복되는 고객에 대한 집계 통계를 얻을 수 있도록 일련의 쿼리를 활성화하려는 것뿐입니다.

A사와 B사는 A사가 제품 반품 활동을 이해하고 B사 및 기타 채널에서 더 나은 제품을 제공할 수 있도록 협력하기로 결정했습니다.

공동 작업을 구축하고 집계 분석을 실행하기 위해 회사는 다음과 같은 작업을 수행합니다.

1. 회사 A는 공동 작업을 만들고 멤버십을 생성합니다. 공동 작업에는 회사 B가 협업의 또 다른 구성원으로 포함됩니다. 회사 A는 공동 작업에서 쿼리 로깅을 활성화하고 해당 계정에서 쿼리 로깅을 활성화합니다.
2. 회사 B는 공동 작업을 통해 멤버십을 생성합니다. 이를 통해 해당 계정에서 쿼리 로그인이 가능합니다.
3. 회사 A는 판매 구성 테이블을 생성합니다.
4. 회사 A는 판매 구성 테이블에 다음과 같은 집계 분석 규칙을 추가합니다.

```
{
  "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`— 회사 A는 판매 데이터와 반품 데이터가 겹치는 순 고객 수를 세려고 합니다. 또한 A 회사는 `purchases` 제조 건수를 합산하여 `returns`의 수와 비교하려고 합니다.

`joinColumns`— 회사 A는 판매 데이터의 고객과 반품 데이터의 고객을 매칭하는 데 `identifier`를 사용하려고 합니다. 이렇게 하면 회사 A가 올바른 구매에 대한 반품을 매칭하는 데 도움이 됩니다. 또한 A사가 중복되는 고객을 세분화하는 데도 도움이 됩니다.

`dimensionColumns`— 회사 A는 특정 제품별로 필터링하고, 일정 기간 동안의 구매 및 반품을 비교하고, 반품 날짜가 제품 날짜 이후인지 확인하고, 중복되는 고객을 분류하기 위해 `dimensionColumns`을 사용합니다.

`scalarFunctions`— 회사 A는 회사 A가 공동 작업에 연결하는 구성 테이블을 기반으로 필요한 경우 데이터 유형 형식을 업데이트하는 데 도움이 되는 `CAST` 스칼라 함수를 선택합니다. 또한 필요한 경우 열 서식을 지정하는 데 도움이 되는 스칼라 함수도 추가되었습니다.

`outputConstraints`— 회사 A는 최소 출력 제약 조건을 설정합니다. 분석가가 판매 테이블에서 행 수준 데이터를 볼 수 있으므로 결과를 제한할 필요가 없습니다.

> **Note**
>
> 회사 A는 분석 규칙에 `joinRequired`을 포함하지 않습니다. 이를 통해 분석가는 유연하게 판매 테이블만 쿼리할 수 있습니다.

5. 회사 B는 반품 구성 테이블을 생성합니다.
6. 회사 B는 반품 구성 테이블에 다음과 같은 집계 분석 규칙을 추가합니다.

```
{
  "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— 회사 B를 사용하면 회사 A의 합계를 returns 구매 건수와 비교할 수 있습니다. 집계 쿼리를 활성화하기 때문에 집계 열이 하나 이상 있습니다.

joinColumns— 회사 B를 사용하면 회사 A가 identifier에 참여하여 반품 데이터에 있는 고객과 판매 데이터를 바탕으로 고객을 매칭할 수 있습니다. identifier 데이터는 특히 민감한 데이터이므로 이를 joinColumn으로 설정하면 쿼리에서 데이터가 출력되지 않습니다.

joinRequired— 회사 B에서는 반품 데이터에 대한 쿼리가 판매 데이터와 중복되도록 요구합니다. 회사 A가 데이터 세트에 있는 모든 개인을 쿼리할 수 있도록 하고 싶지는 않습니다. 또한 공동 작업 계약에서도 이러한 제한에 동의했습니다.

dimensionColumns— 회사 B는 회사 A에 대한 분석에 도움이 될 수 있는 고유한 속성인 state, popularpurchases, 및 customerserviceuser를 기준으로 필터링 및 그룹화할 수 있도록 합니다. 회사 B는 회사 B를 통해 회사 A가 returndate를 사용하여 purchasedate 이후에 발생하는 returndate의 출력을 필터링할 수 있도록 합니다. 이 필터링을 사용하면 제품 변경의 영향을 평가하는 데 더 정확한 결과를 얻을 수 있습니다.

scalarFunctions— 회사 B는 다음을 가능하게 합니다.

- 날짜용 TRUNC
- 데이터에 producttype이 다른 형식으로 입력된 경우 LOWER 및 UPPER를 사용합니다.
- CAST 회사 A가 매출의 데이터 유형을 수익의 데이터 유형과 동일하게 변환해야 하는 경우

회사 A는 다른 스칼라 함수가 쿼리에 필요하지 않다고 생각하기 때문에 다른 스칼라 함수를 활성화하지 않습니다.

outputConstraints— 회사 B는 고객을 재식별하는 능력을 줄이기 위해 hashedemail에 최소 출력 제약 조건을 설정합니다. 또한 반품된 특정 제품을 재식별하는 능력을 줄이기 위해 producttype에 최소 생산량 제한을 추가합니다. 생산량 규모에 따라 특정 제품 유형이 더 우세할 수 있습니다 (예: state). 회사 A가 데이터에 추가한 출력 제약 조건과 관계없이 해당 출력 제약은 항상 적용됩니다.

7. 회사 A는 공동 작업과 관련된 판매 테이블을 생성합니다.
8. 회사 B는 공동 작업에 대한 반품 테이블 연결을 생성합니다.
9. 회사 A는 다음 예와 같은 쿼리를 실행하여 2022년 위치별 총 구매액과 비교하여 회사 B의 반품 수량을 더 잘 파악합니다.

```
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;
```

10회사 A와 회사 B는 쿼리 로그를 검토합니다. 회사 B는 쿼리가 공동 작업 계약에서 합의한 내용과 일치하는지 확인합니다.

집계 분석 규칙 문제 해결

여기에 있는 정보를 사용하여 집계 분석 규칙으로 작업할 때 흔히 발생하는 문제를 진단하고 해결하는 데 도움을 받을 수 있습니다.

문제

- 쿼리에서 결과가 반환되지 않았습니다.

쿼리에서 결과가 반환되지 않았습니다.

일치하는 결과가 없거나 일치하는 결과가 하나 이상의 최소 집계 임계값을 충족하지 않을 때 이런 일이 발생할 수 있습니다.

최소 집계 임계값에 대한 자세한 내용은 집계 분석 규칙 - 예제 섹션을 참조하세요.

목록 분석 규칙

에서 AWS Clean Rooms목록 분석 규칙은 추가된 구성된 테이블과 쿼리할 수 있는 구성원의 구성된 테이블 간의 중첩 행 수준 목록을 출력합니다. 쿼리할 수 있는 구성원은 목록 분석 규칙이 포함된 쿼리를 실행합니다.

목록 분석 규칙 유형은 강화 및 대상 구축과 같은 사용 사례를 지원합니다.

이 분석 규칙의 사전 정의된 쿼리 구조 및 구문에 대한 자세한 내용은 목록 분석 규칙 사전 정의된 구조을 참조하세요.

목록 분석 규칙 - 쿼리 제어에 정의된 목록 분석 규칙의 매개 변수에는 쿼리 컨트롤이 있습니다. 쿼리 컨트롤에는 출력에 나열할 수 있는 열을 선택하는 기능이 포함됩니다. 쿼리에는 직접 또는 전이적으로 쿼리할 수 있는 구성원의 구성된 테이블을 사용한 조인이 하나 이상 있어야 합니다.

집계 분석 규칙과 같은 쿼리 결과 컨트롤은 없습니다.

목록 쿼리는 수학 연산자만 사용할 수 있습니다. 다른 함수(예: 집계 또는 스칼라)는 사용할 수 없습니다.

주제

- 쿼리 구조 및 구문 목록
- 목록 분석 규칙 - 쿼리 제어
- 목록 분석 규칙 사전 정의된 구조
- 목록 분석 규칙 - 예제

쿼리 구조 및 구문 목록

목록 분석 규칙이 있는 테이블에 대한 쿼리는 다음 구문을 준수해야 합니다.

```
--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]
```

다음 표에서는 위 구문에 나열된 각 표현식에 대해 설명합니다.

표현식	정의	예시
select_list_expression	테이블 열 이름을 하나 이상 포함하는 쉼표로 구분된 목록입니다. DISTINCT 파라미터가 필요합니다.	SELECT DISTINCT segment

표현식	정의	예시
	Note `select_list_expression` 은 AS 매개변수를 포함하거나 포함하지 않고 열에 별칭을 지정할 수 있습니다. 자세한 내용은 AWS Clean Rooms SQL 참조 섹션을 참조하세요.	
table_expression	`join_condition` 를 `join_condition` 에 연결하는 테이블 또는 테이블의 조인. `join_condition` 은 BOOLEAN을 반환합니다. `table_expression` 은 다음을 지원합니다. • 특정 조인 유형(INNER조인) • `join_condition` (=) 내의 동등 비교 조건 • 논리 연산자(AND, OR).	`FROM consumer_table INNER JOIN provider_table ON consumer_table.identifier1 = provider_table.identifier1 AND consumer_table.identifier2 = provider_table.identifier2`

표현식	정의	예시

where_expression	부울을 반환하는 조건식입니다. 다음과 같이 구성될 수 있습니다:
	• 테이블 열 이름
	• 수학적 연산
	• 문자열 리터럴
	• 숫자 리터럴
	
	지원되는 비교 조건은 (=, >, <, <=, >=, <>, !=, NOT, IN, NOT IN, LIKE, IS NULL, IS NOT NULL)입니다.
	
	지원되는 논리 연산자는 (AND, OR)입니다.
	
	where_expression 은 선택 사항입니다.
		`WHERE state + '_' + city = 'NY_NYC'`
		
		`WHERE timestampColumn = timestampColumn2 - 14`

limit_expression	이 표현식은 양의 정수를 사용해야 합니다. limit_expression 은 선택 사항입니다.	`LIMIT 100`

목록 쿼리 구조 및 구문에 대해서는 다음 사항에 유의해야 합니다.

- SELECT 이외의 SQL 명령은 지원되지 않습니다.
- 하위 쿼리 및 공통 테이블 표현식(예:WITH)은 지원되지 않습니다
- HAVING GROUP BY, 및 BY ORDER 절은 지원되지 않습니다
- OFFSET 파라미터는 지원되지 않습니다

목록 분석 규칙 - 쿼리 제어

목록 쿼리 컨트롤을 사용하면 테이블의 열을 사용하여 테이블을 쿼리하는 방법을 제어할 수 있습니다. 예를 들어, 조인에 사용되는 열 또는 SELECT 명령문 및 WHERE 절에서 사용할 수 있는 열을 제어할 수 있습니다.

다음 단원에서는 각 컨트롤에 대해 설명합니다.

주제

- 조인 컨트롤
- 목록 컨트롤

조인 컨트롤

조인 컨트롤을 사용하면 테이블이 table_expression의 다른 테이블에 조인되는 방식을 제어할 수 있습니다. AWS Clean Rooms 은 INNER JOIN만 지원합니다. 목록 분석 규칙에는 최소 하나 이상의 INNER JOIN이 필요하며 쿼리가 가능한 구성원은 자신이 소유한 테이블을 INNER JOIN에 포함해야 합니다. 즉, 직접 또는 전이적으로 테이블을 자신의 테이블과 조인해야 합니다.

다음은 전이성의 예입니다.

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

INNER JOIN 문은 분석 규칙에서 명시적으로 `joinColumn`으로 분류된 열만 사용할 수 있습니다.

INNER JOIN은 구성된 테이블의 `joinColumn`과 공동 작업의 다른 구성된 테이블의 `joinColumn`에 대해 작동해야 합니다. 테이블에서 `joinColumn`으로 사용할 수 있는 열을 결정합니다.

ON 절의 각 일치 조건은 두 열 간의 등식 비교 조건(=)을 사용해야 합니다.

ON 조항 내의 여러 일치 조건은 다음과 같을 수 있습니다.

- `AND` 논리 연산자를 사용하여 조합합니다
- `OR` 논리 연산자를 사용하여 구분합니다

> **Note**
>
> 모든 JOIN 일치 조건은 JOIN의 각 변의 한 행과 일치해야 합니다. OR 또는 AND 논리 연산자로 연결된 모든 조건문도 이 요구 사항을 준수해야 합니다.

다음은 AND 논리 연산자를 사용한 쿼리의 예입니다.

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

다음은 OR 논리 연산자를 사용하는 쿼리의 예입니다.

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

컨트롤	정의	사용법
joinColumns	쿼리를 할 수 있는 구성원이 INNER JOIN 문에서 사용할 수 있도록 허용하려는 열입니다.	동일한 열을 joinColumn 과 listColumn (목록 컨트롤 참조) 모두로 분류할 수는 없습니다. joinColumn 은 INNER JOIN 이외의 쿼리 다른 부분에서는 사용할 수 없습니다.

목록 컨트롤

목록 컨트롤은 쿼리 출력에 나열될 수 있는 열(즉, SELECT 문에 사용) 또는 결과를 필터링하는 데 사용할 수 있는 열(즉, WHERE 문에서 사용)을 제어합니다.

컨트롤	정의	사용법
`listColumns`	쿼리할 수 있는 구성원이 SELECT와 WHERE에서 사용하도록 허용한 열	`listColumn` 은 SELECT 및 WHERE에서 사용할 수 있습니다. 동일한 열을 `listColumn` 과 `joinColumn` 둘 다로 사용할 수는 없습니다.

목록 분석 규칙 사전 정의된 구조

다음 예제에는 목록 분석 규칙을 완성하는 방법을 보여주는 사전 정의된 구조가 포함되어 있습니다.

다음 예제에서는 *`MyTable`*은 데이터 테이블을 의미합니다. 각 *`user input placeholder`*를 사용자의 정보로 바꿀 수 있습니다.

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

목록 분석 규칙 - 예제

다음 예제에서는 두 회사가 목록 분석을 AWS Clean Rooms 사용하여에서 협업하는 방법을 보여줍니다.

회사 A에는 고객 관계 관리(CRM) 데이터가 있습니다. 회사 A는 고객에 대해 더 자세히 알아보고 잠재적으로 속성을 다른 분석에 입력 자료로 사용할 수 있도록 고객에 대한 추가 세그먼트 데이터를 확보하고자 합니다. 회사 B는 자사 데이터를 기반으로 생성한 고유한 세그먼트 속성으로 구성된 세그먼트 데이터를 보유하고 있습니다. 회사 B는 고객 데이터와 회사 A 데이터 간에 중복되는 고객에 대해서만 회사 A에 고유한 세그먼트 속성을 제공하려고 합니다.

두 회사는 A사가 중복되는 데이터를 보강할 수 있도록 협력하기로 결정합니다. 회사 A는 쿼리를 할 수 있는 구성원이고 회사 B는 기여자입니다.

공동 작업을 생성하고 공동 작업으로 목록 분석을 실행하기 위해 회사는 다음과 같은 작업을 수행합니다.

1. 회사 A는 공동 작업을 만들고 멤버십을 생성합니다. 공동 작업에는 회사 B가 공동 작업의 또 다른 구성원으로 참여합니다. 회사 A는 공동 작업에서 쿼리 로깅을 활성화하고 해당 계정에서 쿼리 로깅을 활성화합니다.
2. 회사 B는 공동 작업 멤버십을 생성합니다. 이를 통해 해당 계정에서 쿼리 로깅이 가능합니다.
3. 회사 A는 CRM으로 구성된 테이블을 생성합니다
4. 회사 A는 다음 예에 나와 있는 대로 고객 구성 테이블에 분석 규칙을 추가합니다.

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

`joinColumns`— 회사 A는 `hashedemail` 및/또는 `thirdpartyid`(ID 공급업체로부터 획득)를 사용하여 CRM 데이터의 고객과 세그먼트 데이터의 고객을 매칭하려고 합니다. 이를 통해 A사는 풍부한 데이터를 적합한 고객에게 매칭할 수 있습니다. 두 개의 JoinColumn이 있어 분석의 일치율을 잠재적으로 개선할 수 있습니다.

`listColumns`— 회사 A는 자체 시스템에서 사용하는 `internalid` 외에도 풍부한 열을 얻기 위해 `listColumns`을 사용합니다. `segment1`, `segment2`, 및 `customercategory`를을 추가하여 필터에 사용함으로써 잠재적으로 특정 세그먼트로 보강을 제한할 수 있습니다.

5. 회사 B는 세그먼트 구성 테이블을 만듭니다.
6. 회사 B는 세그먼트 구성 테이블에 분석 규칙을 추가합니다.

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

```
}
```

`joinColumns`— 회사 B는 회사 A가 `identifier2`에서 조인하여 세그먼트 데이터에서 CRM 데이터로 고객을 일치시킬 수 있도록 지원합니다. 회사 A와 회사 B는 ID 공급업체와 협력하여 이번 공동 작업에 적합한 `identifier2`를 확보했습니다. `identifier2`가 가장 높고 정확한 일치율을 제공하며 다른 식별자는 쿼리에 필요하지 않다고 판단했기 때문에 다른 `joinColumns`을 추가하지 않았습니다.

`listColumns`— 회사 B는 회사 A가 데이터 보강의 일부로(고객 A와 함께) 생성, 수집, 정렬한 고유 속성인 `segment3` 및 `segment4` 속성을 사용하여 데이터를 보강할 수 있도록 지원합니다. 그들은 A 회사가 행 수준에서 중복되는 부분에 대해 이러한 세그먼트를 확보하기를 원합니다. 이는 데이터 강화 공동 작업이기 때문입니다.

7. 회사 A는 공동 작업에 CRM 테이블 연결을 생성합니다.
8. 회사 B는 공동 작업에 대한 세그먼트 테이블 연결을 생성합니다.
9. 회사 A는 다음과 같은 쿼리를 실행하여 중복되는 고객 데이터를 보강합니다.

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

10. 회사 A와 회사 B는 쿼리 로그를 검토합니다. 회사 B는 쿼리가 공동 작업 계약에서 합의한 내용과 일치하는지 확인합니다.

의 사용자 지정 분석 규칙 AWS Clean Rooms

에서 AWS Clean Rooms사용자 지정 분석 규칙은 구성된 테이블에서 사용자 지정 쿼리를 실행할 수 있는 새로운 유형의 분석 규칙입니다. 사용자 지정 SQL 쿼리는 여전히 SELECT 명령만 사용하는 것으로 제한되지만 집계 및 목록 쿼리보다 더 많은 SQL 구성을 사용할 수 있습니다(예: 창 함수, OUTER JOIN, CTE 또는 하위 쿼리가 있고, 전체 목록은 AWS Clean Rooms SQL 참조를 참조하세요). 사용자 지정 SQL 쿼리는 집계 및 목록 쿼리와 같은 쿼리 구조를 따를 필요가 없습니다.

사용자 지정 분석 규칙은 사용자 지정 속성 분석, 벤치마킹, 증분 분석, 대상 발견과 같은 집계 및 목록 분석 규칙이 지원하는 것보다 더 고급 사용 사례를 지원합니다. 이는 집계 및 목록 분석 규칙이 지원하는 사용 사례의 일부에 추가됩니다.

사용자 지정 분석 규칙은 차등 프라이버시 기능도 지원합니다. 차등 프라이버시는 데이터 프라이버시를 보호하기 위해 수학적으로 엄격한 프레임워크입니다. 자세한 내용은 AWS Clean Rooms 차등 프라

이버시 단원을 참조하십시오. 분석 템플릿을 생성할 때 AWS Clean Rooms 차등 프라이버시는 템플릿을 확인하여 AWS Clean Rooms 차등 프라이버시에 대한 범용 쿼리 구조와 호환되는지 확인합니다. 이 검증을 통해 차등 프라이버시 보호 테이블에서 허용되지 않는 분석 템플릿을 생성하지 않을 수 있습니다.

사용자 지정 분석 규칙을 구성하기 위해 데이터 소유자는 분석 템플릿에 저장된 특정 사용자 지정 쿼리가 구성된 테이블에서 실행되도록 허용하도록 선택할 수 있습니다. 데이터 소유자는 사용자 지정 분석 규칙에서 허용된 분석 컨트롤에 추가하기 전에 분석 템플릿을 검토합니다. 분석 템플릿은 테이블이 다른 공동 작업과 연결되어 있더라도 해당 템플릿을 만든 공동 작업에서만 사용할 수 있고 볼 수 있으며 해당 공동 작업에서 쿼리할 수 있는 구성원만 실행할 수 있습니다.

또는 구성원이 다른 구성원(쿼리 제공자)이 검토 없이 쿼리를 생성하도록 허용할 수도 있습니다. 구성원은 사용자 지정 분석 규칙에서 허용된 쿼리 제공자가 제어하는 쿼리 제공자의 계정을 추가합니다. 쿼리 제공자가 쿼리를 수행할 수 있는 구성원인 경우 구성원은 구성된 테이블에서 직접 모든 쿼리를 실행할 수 있습니다. 쿼리 제공자는 분석 템플릿을 생성하여 쿼리를 생성할 수도 있습니다. 쿼리 공급자가 생성한 모든 쿼리는 AWS 계정 가 존재하고 테이블이 연결된 모든 공동 작업의 테이블에서 자동으로 실행될 수 있습니다.

데이터 소유자는 분석 템플릿 또는 계정만 쿼리를 생성하도록 허용할 수 있으며 둘 다 생성할 수는 없습니다. 데이터 소유자가 테이블을 비워 두면 쿼리를 할 수 있는 구성원은 구성된 테이블에서 쿼리를 실행할 수 없습니다.

주제

- 사용자 지정 분석 규칙, 사전 정의된 구조
- 사용자 지정 분석 규칙 예제
- 차등 프라이버시가 사용되는 사용자 지정 분석 규칙

사용자 지정 분석 규칙, 사전 정의된 구조

다음 예제에는 차등 프라이버시 기능을 켜고 사용자 지정 분석 규칙을 완성하는 방법을 보여 주는 사전 정의된 구조가 포함되어 있습니다. `userIdentifier` 값은 사용자를 고유하게 식별하는 열(예: user_id)입니다. 공동 작업에서 차등 프라이버시를 사용하는 테이블이 2개 이상 있는 경우 AWS Clean Rooms 은 2개 이상의 분석 규칙의 사용자 식별자 열과 동일한 열을 구성하여 테이블 간에 사용자에 대한 정의를 일관되게 유지해야 합니다.

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

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

다음 작업 중 하나를 수행할 수 있습니다.

- 허용된 분석 제어에 분석 템플릿 ARN을 추가합니다. 이 경우 `allowedAnalysisProviders` 컨트롤은 포함되지 않습니다.

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

- `allowedAnalysisProviders` 제어에 AWS 계정 IDs 추가합니다. 이 경우에는 `ANY_QUERY`를 `allowedAnalyses` 컨트롤에 추가합니다.

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

사용자 지정 분석 규칙 예제

다음 예제에서는 두 회사가 사용자 지정 분석 규칙을 AWS Clean Rooms 사용하여에서 협업하는 방법을 보여줍니다.

회사 A에는 고객 및 판매 데이터가 있습니다. 회사 A는 회사 B 사이트에서 광고 캠페인의 매출 증대를 파악하는 데 관심이 있습니다. 회사 B에는 회사에 유용한 시청률 데이터 및 세그먼트 속성(예: 광고를 볼 때 사용한 장치)이 있습니다.

회사 A에는 공동 작업에서 실행하려는 특정 증분 쿼리가 있습니다.

공동 작업을 생성하고 공동 작업을 통해 사용자 지정 분석을 실행하기 위해 회사는 다음을 수행합니다.

1. 회사 A는 공동 작업을 만들고 멤버십을 생성합니다. 공동 작업에는 회사 B가 공동 작업의 또 다른 구성원으로 참여합니다. 회사 A는 공동 작업에서 쿼리 로깅을 활성화하고 해당 계정에서 쿼리 로깅을 활성화합니다.
2. 회사 B는 공동 작업 멤버십을 생성합니다. 이를 통해 해당 계정에서 쿼리 로깅이 가능합니다.
3. 회사 A는 CRM으로 구성된 테이블을 생성합니다.
4. 회사 A는 판매 구성 테이블에 빈 사용자 지정 분석 규칙을 추가합니다.
5. 회사 A는 판매 구성 테이블을 공동 작업에 연결합니다.
6. 회사 B는 시청률 구성 테이블을 만듭니다.
7. 회사 B는 시청률 구성 테이블에 빈 사용자 지정 분석 규칙을 추가합니다.
8. 회사 B는 시청률 구성 테이블을 공동 작업과 연결합니다.
9. 회사 A는 공동 작업과 관련된 판매 테이블 및 시청률 테이블을 보고 캠페인 월의 증분 쿼리 및 매개 변수를 추가하여 분석 템플릿을 만듭니다.

```
{
    "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.회사 A는 사용자 지정 분석 규칙에서 허용된 분석 제공자 컨트롤에 해당 계정(예: 444455556666)을 추가합니다. 이들은 자신이 만든 모든 쿼리가 판매 구성 테이블에서 실행되도록 허용하기를 원하기 때문에 허용된 분석 공급자 컨트롤을 사용합니다.

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

11.회사 B는 공동 작업에서 생성된 분석 템플릿을 보고 쿼리 문자열 및 파라미터를 포함한 내용을 검토합니다.

12.회사 B는 분석 템플릿이 증분 사용 사례를 충족하고 시청률 구성 테이블을 쿼리할 수 있는 방법에 대한 프라이버시 요구 사항을 충족한다고 판단합니다.

13.회사 B는 시청률 테이블의 사용자 지정 분석 규칙에서 허용된 분석 컨트롤에 분석 템플릿 ARN을 추가합니다. 시청률이 구성된 테이블에서 증분 쿼리만 실행되도록 허용하려고 하기 때문에 허용된 분석 제어를 사용합니다.

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

14.회사 A는 분석 템플릿을 실행하고 매개변수 값 `05-01-2023`을 사용합니다.

차등 프라이버시가 사용되는 사용자 지정 분석 규칙

에서 사용자 지정 분석 규칙은 차등 프라이버시 AWS Clean Rooms를 지원합니다. 차등 프라이버시는 재식별 시도로부터 데이터를 보호하는 데 도움이 되는 수학적으로 엄격한 데이터 프라이버시 보호 프레임워크입니다.

차등 프라이버시는 광고 캠페인 계획, 광고 캠페인 후 측정, 금융 기관 컨소시엄의 벤치마킹, 의료 연구를 위한 A/B 테스트와 같은 집계 분석을 지원합니다.

지원되는 쿼리 구조 및 구문은 쿼리 구조 및 구문에 정의되어 있습니다.

차등 프라이버시가 사용되는 사용자 지정 분석 규칙 예제

> **Note**
>
> AWS Clean Rooms 차등 프라이버시는 AWS Clean Rooms SQL을 분석 엔진으로 사용하고 Amazon S3에 저장된 데이터를 사용하는 공동 작업에만 사용할 수 있습니다.

이전 섹션에 제시된 사용자 지정 분석 규칙 예제를 생각해 봅니다. 이 예제는 차등 프라이버시를 사용하여 재식별 시도로부터 데이터를 보호하는 동시에 파트너가 데이터에서 비즈니스에 중요한 인사이트를 얻을 수 있도록 하는 방법을 보여 줍니다. 시청률 데이터를 보유한 회사 B가 차등 프라이버시를 통해 데이터를 보호하고자 한다고 가정해 보겠습니다. 차등 프라이버시 설정을 완료하기 위해 회사 B는 다음 단계를 완료합니다.

1. 회사 B는 시청률 구성 테이블에 사용자 지정 분석 규칙을 추가하면서 차등 프라이버시를 사용 설정합니다. 회사 B는 사용자 식별자 열로 `viewershipdata.hashedemail`을 선택합니다.
2. 회사 B는 시청률 데이터 테이블을 쿼리에 사용할 수 있도록 공동 작업에 차등 프라이버시 정책을 추가합니다. 회사 B는 설정을 빠르게 완료하기 위해 기본 정책을 선택합니다.

회사 B의 사이트에서 광고 캠페인의 매출 증대를 파악하려는 회사 A가 분석 템플릿을 실행합니다. 쿼리는 AWS Clean Rooms 차등 프라이버시의 범용 쿼리 구조와 호환되므로 쿼리가 성공적으로 실행됩니다.

쿼리 구조 및 구문

차등 프라이버시 기능이 켜진 테이블이 적어도 하나 이상 포함된 쿼리는 다음 구문을 따라야 합니다.

```
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

차등 프라이버시 쿼리 구조 및 구문에 대해서는 다음 사항에 유의해야 합니다.

- 하위 쿼리는 지원되지 않습니다.
- 테이블 또는 CTE에 차등 프라이버시로 보호되는 데이터가 포함된 경우 일반 테이블 표현식(CTE)에서 사용자 식별자 열을 내보내야 합니다. 필터, 그룹화 및 집계는 사용자 수준에서 수행해야 합니다.
- Final_select는 COUNT DISTINCT, COUNT, SUM, AVG, STDDEV 집계 함수를 허용합니다.

차등 프라이버시에서 지원되는 SQL 키워드에 대한 자세한 내용은 AWS Clean Rooms 차등 프라이버시의 SQL 기능 섹션을 참조하세요.

ID 매핑 테이블 분석 규칙

에서 AWS Clean RoomsID 매핑 테이블 분석 규칙은 독립 실행형 분석 규칙이 아닙니다. 이러한 유형의 분석 규칙은에서 관리 AWS Clean Rooms 하며 쿼리를 용이하게 하기 위해 서로 다른 자격 증명 데이터를 조인하는 데 사용됩니다. ID 매핑 테이블에 자동으로 추가되며 편집할 수 없습니다. 또한 공동 작업 내 다른 분석 규칙의 동작을 상속합니다. 단, 해당 분석 규칙들이 동질적이어야 합니다.

ID 매핑 테이블 분석 규칙은 ID 매핑 테이블에 보안을 적용합니다. 이 규칙은 공동 작업 구성원이 ID 매핑 테이블을 사용하여 두 구성원의 데이터 세트 간에 중첩되지 않는 모집단을 직접 선택하거나 검사하는 것을 제한합니다. ID 매핑 테이블 분석 규칙은 다른 분석 규칙과 함께 쿼리에 암시적으로 사용될 때 ID 매핑 테이블의 민감한 데이터를 보호하는 데 사용됩니다.

ID 매핑 테이블 분석 규칙을 사용하면는 확장된 SQL에서 ID 매핑 테이블의 양쪽에 중첩을 AWS Clean Rooms 적용합니다. 이를 통해 다음 작업을 수행할 수 있습니다:

- JOIN 문에서 ID 매핑 테이블의 중첩을 사용할 수 있습니다.

  AWS Clean Rooms 는 중첩을 준수하는 경우 ID 매핑 테이블에서 INNERLEFT, 또는 RIGHT 조인을 허용합니다. 민감한 매핑 정보를 보호하려면 ID 매핑 테이블이 항상 JOIN 작업의 "inner" 측에 있어야 합니다. 예를 들어 다음 JOIN 작업이 유효합니다.

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

  다음 JOIN 작업은 유효하지 않습니다.

  - id_mapping_table LEFT JOIN table
  - table RIGHT JOIN id_mapping_table

  이렇게 하면 데이터 세트에 일치하는 항목이 없는 매핑 레코드가 노출되는 것을 방지할 수 있습니다. 이러한 작업을 허용하면 다른 공동 작업 구성원의 데이터 매핑에 대한 민감한 정보가 잠재적으로 공개될 수 있습니다.
- JOIN 문에서 매핑 테이블 열을 사용할 수 있습니다.

  그러나 SELECT, WHERE, HAVING, GROUP BY 또는 ORDER BY 문에서는 매핑 테이블 열을 사용할 수 없습니다. 단, 소스 ID 네임스페이스 연결 또는 대상 ID 네임스페이스 연결에서 보호가 수정된 경우는 예외입니다.

- 확장된 SQL에서는 OUTER JOIN, 암시적 및 JOINCROSS도 AWS Clean Rooms 지원합니다 JOIN. 이러한 조인은 중첩 요구 사항을 충족할 수 없습니다. 대신 `requireOverlap`를 AWS Clean Rooms 사용하여 조인해야 하는 열을 지정합니다.

지원되는 쿼리 구조 및 구문은 ID 매핑 테이블 쿼리 구조 및 구문에 정의되어 있습니다.

ID 매핑 테이블 분석 규칙 쿼리 컨트롤에 정의된 분석 규칙의 파라미터에는 쿼리 컨트롤 및 쿼리 결과 컨트롤이 포함됩니다. 쿼리 컨트롤에는 JOIN 문에서 ID 매핑 테이블의 중첩을 요구하는 기능(즉, `requireOverlap`)이 포함됩니다.

주제

- ID 매핑 테이블 쿼리 구조 및 구문
- ID 매핑 테이블 분석 규칙 쿼리 컨트롤
- ID 매핑 테이블 분석 규칙의 사전 정의된 구조
- ID 매핑 테이블 분석 규칙 - 예

ID 매핑 테이블 쿼리 구조 및 구문

ID 매핑 테이블 분석 규칙이 있는 테이블에 대한 쿼리는 다음 구문을 준수해야 합니다.

```
--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
```

공동 작업 테이블

다음 표는 AWS Clean Rooms 공동 작업에 있는 구성된 테이블을 나타냅니다. cr_drivers_license 및 cr_insurance 테이블의 id 열은 ID 매핑 테이블과 일치하는 열을 나타냅니다.

cr_drivers_license

id	driver_name	state_of_registration
1	Eduard	TX
2	Dana	MA
3	Gweneth	IL

cr_insurance

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

ID 매핑 테이블

다음 표는 cr_drivers_license 및 cr_insurance 테이블과 일치하는 기존 ID 매핑 테이블을 나타냅니다. 두 공동 작업 테이블에 대한 ID가 없는 항목도 있습니다.

cr_drivers_license_id	cr_insurance_id
1	a
2	null
3	b
null	c

ID 매핑 테이블 분석 규칙은 다음과 같이 중첩 데이터 세트에 대해서만 쿼리를 실행할 수 있도록 허용합니다.

cr_drivers_license_id	cr_insurance_id	driver_name	state_of_registration	policyholder_email	policy_number
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

쿼리 예제

다음 예제에서는 ID 매핑 테이블 조인의 유효한 위치를 보여줍니다.

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

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

고려 사항

ID 매핑 테이블 쿼리 구조 및 구문에 대해서는 다음 사항에 유의해야 합니다.

- 편집할 수 없습니다.
- 기본적으로 ID 매핑 테이블에 적용됩니다.
- 공동 작업 내에서 소스 및 대상 ID 네임스페이스 연결을 사용합니다.
- ID 매핑 테이블은 ID 네임스페이스에서 가져온 열에 대한 기본 보호를 제공하도록 기본적으로 구성됩니다. ID 네임스페이스(`sourceID` 또는 `targetID`)에서 가져온 열이 쿼리의 어느 곳에서든 허용될 수 있도록 이 구성을 수정할 수 있습니다. 자세한 내용은 의 ID 네임스페이스 AWS Clean Rooms 단원을 참조하십시오.
- ID 매핑 테이블 분석 규칙은 공동 작업 내 다른 분석 규칙의 SQL 제한을 상속합니다.

ID 매핑 테이블 분석 규칙 쿼리 컨트롤

ID 매핑 테이블 쿼리 제어를 사용하면가 테이블의 열을 사용하여 테이블을 쿼리하는 방법을 AWS Clean Rooms 제어합니다. 예를 들어 조인에 사용되는 열과 중첩이 필요한 열을 제어합니다. ID 매핑 테이블 분석 규칙에는 JOIN 없이도 `sourceID`, `targetID` 또는 둘 다를 표시할 수 있는 기능도 포함되어 있습니다.

다음 표에서는 각 컨트롤에 대해 설명합니다.

컨트롤	정의	사용법
joinColumns	쿼리를 할 수 있는 구성원이 INNER JOIN 문에서 사용할 수 있는 열입니다.	INNER JOIN 이외의 다른 쿼리 부분에서는 joinColumns 를 사용할 수 없습니다. 자세한 내용은 조인 컨트롤 단원을 참조하십시오.
dimensionColumns	쿼리할 수 있는 구성원이 SELECT 및 GROUP BY 문에서 사용할 수 있는 열입니다(있는 경우).	dimensionColumn 은 SELECT 및 GROUP BY에서 사용할 수 있습니다. dimensionColumn 은 joinKeys로 표시될 수 있습니다. 괄호 안에 지정한 경우에만 JOIN 절에서 dimension Columns 를 사용할 수 있습니다.
queryContraints:RequireOverlap	쿼리를 실행할 수 있도록 조인해야 하는 ID 매핑 테이블의 열입니다.	이러한 열은 ID 매핑 테이블과 공동 작업 테이블을 조인하는 데 사용해야 합니다.

ID 매핑 테이블 분석 규칙의 사전 정의된 구조

ID 매핑 테이블 분석 규칙의 사전 정의된 구조에는 sourceID 및 targetID에 적용되는 기본 보호 기능이 포함되어 있습니다. 즉, 보호 기능이 적용된 열을 쿼리에 사용해야 합니다.

ID 매핑 테이블 분석 규칙은 다음과 같은 방법으로 구성할 수 있습니다.

- sourceID 및 모두 targetID 보호됨

  이 구성에서는 sourceID 및 targetID를 모두 표시할 수 없습니다. ID 매핑 테이블을 참조할 경우에는 sourceID 및 targetID를 JOIN에 사용해야 합니다.
- targetID만 보호됨

이 구성에서는 targetID를 표시할 수 없습니다. ID 매핑 테이블을 참조할 경우에는 targetID를 JOIN에 사용해야 합니다. sourceID는 쿼리에 사용할 수 있습니다.

- sourceID만 보호됨

  이 구성에서는 sourceID를 표시할 수 없습니다. ID 매핑 테이블을 참조할 경우에는 sourceID를 JOIN에 사용해야 합니다. targetID는 쿼리에 사용할 수 있습니다.

- sourceID 또는 targetID 모두 보호되지 않음

  이 구성에서는 ID 매핑 테이블에 대해 쿼리에 사용할 수 있는 특정 제한이 적용되지 않습니다.

다음 예시는 sourceID와 targetID에 기본 보호가 적용된 ID 매핑 테이블 분석 규칙의 사전 정의된 구조를 보여줍니다. 이 예시에서 ID 매핑 테이블 분석 규칙은 sourceID 열과 targetID 열 모두에 대한 INNER JOIN만 허용합니다.

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

다음 예시는 targetID에 보호 기능이 적용된 ID 매핑 테이블 분석 규칙의 미리 정의된 구조를 보여줍니다. 이 예시에서 ID 매핑 테이블 분석 규칙은 sourceID 열에 대한 INNER JOIN만 허용합니다.

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

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

다음 예시는 sourceID에 보호 기능이 적용된 ID 매핑 테이블 분석 규칙의 미리 정의된 구조를 보여줍니다. 이 예시에서 ID 매핑 테이블 분석 규칙은 targetID 열에 대한 INNER JOIN만 허용합니다.

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

다음 예시는 sourceID 또는 targetID에 보호 기능이 적용되지 않은 ID 매핑 테이블 분석 규칙의 미리 정의된 구조를 보여줍니다. 이 예시에서 ID 매핑 테이블 분석 규칙은 sourceID 열과 targetID 열 모두에 대한 INNER JOIN을 허용합니다.

```
{
```

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

ID 매핑 테이블 분석 규칙 - 예

예를 들어 개인 식별 정보(PII)를 참조하는 긴 워터폴 문을 작성하는 대신, 기업은 ID 매핑 테이블 분석 규칙을 통해 다자간 LiveRamp 트랜스코딩을 활용할 수 있습니다. 다음 예제에서는 ID 매핑 테이블 분석 규칙을 AWS Clean Rooms 사용하여에서 협업하는 방법을 보여줍니다.

A사는 소스로 사용될 고객 및 판매 데이터를 보유한 광고주입니다. 또한 A사는 공동 작업의 당사자들을 대신하여 트랜스코딩을 수행하고 LiveRamp 자격 증명을 제공합니다.

B사는 대상으로 사용될 이벤트 데이터를 보유한 게시자입니다.

> **Note**
>
> A사 또는 B사는 LiveRamp 트랜스코딩 자격 증명을 제공하고 트랜스코딩을 수행할 수 있습니다.

공동 작업을 생성하고 공동 작업을 통해 ID 매핑 테이블 분석을 수행하기 위해 두 회사는 다음과 같은 작업을 수행합니다.

1. A사는 공동 작업을 만들고 멤버십을 생성합니다. B사를 추가하고 이 회사의 공동 작업 멤버십도 생성합니다.
2. 회사 A는 기존 ID 네임스페이스 소스를 연결하거나 AWS Clean Rooms 콘솔을 AWS Entity Resolution 사용하여에서 새 ID 네임스페이스 소스를 생성합니다.

A사는 판매 데이터로 구성된 테이블을 생성하고, 이 테이블에 ID 매핑 테이블의 `sourceId`에 연결된 열을 추가합니다.

ID 네임스페이스 소스는 트랜스코딩할 데이터를 제공합니다.

3. 회사 B는 AWS Clean Rooms 콘솔을 AWS Entity Resolution 사용하여 기존 ID 네임스페이스 대상을 연결하거나에서 새 대상을 생성합니다.

   B사는 이벤트 데이터로 구성된 테이블을 생성하고, 이 테이블에 ID 매핑 테이블의 `targetId`에 연결된 열을 추가합니다.

   ID 네임스페이스 대상은 트랜스코딩할 데이터를 제공하지 않고 LiveRamp 구성과 관련된 메타데이터만 제공합니다.
4. A사는 공동 작업에 연결된 두 개의 ID 네임스페이스를 검색하고 ID 매핑 테이블을 생성하여 채웁니다.
5. A사는 ID 매핑 테이블을 기준으로 두 데이터세트를 조인하여 쿼리를 실행합니다.

```
--- 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 차등 프라이버시

AWS Clean Rooms 차등 프라이버시를 사용하면 몇 번의 클릭만으로 직관적인 제어로 구현되는 수학 기반 기술을 통해 사용자의 개인 정보를 보호할 수 있습니다. 완전 관리형 기능으로서 사용자의 재식별을 방지하는 데 도움이 되는 사전 차등 프라이버시 환경은 필요하지 않습니다.는 개별 수준 데이터를 보호하기 위해 런타임 시 쿼리 결과에 신중하게 보정된 양의 노이즈를 AWS Clean Rooms 자동으로 추가합니다.

AWS Clean Rooms 차등 프라이버시는 광범위한 분석 쿼리를 지원하며 쿼리 결과의 소량의 오류가 분석의 유용성을 손상시키지 않는 다양한 사용 사례에 적합합니다. 이를 통해 파트너는 파트너 측면에서 추가 설정을 하지 않아도 광고 캠페인, 투자 결정, 임상 연구 등에 관한 비즈니스에 중요한 인사이트를 얻을 수 있습니다.

AWS Clean Rooms 차등 프라이버시는 스칼라 함수 또는 수학 연산자 기호를 악의적으로 사용하는 오버플로 또는 잘못된 캐스팅 오류로부터 보호합니다.

AWS Clean Rooms 차등 프라이버시에 대한 자세한 내용은 다음 주제를 참조하세요.

주제

- 차등 프라이버시
- 의 차등 프라이버시 AWS Clean Rooms 작동 방식
- 차등 프라이버시 정책
- AWS Clean Rooms 차등 프라이버시의 SQL 기능
- 차등 프라이버시 쿼리 팁 및 예제
- AWS Clean Rooms 차등 프라이버시의 제한 사항

차등 프라이버시

차등 프라이버시를 사용하면 집계된 인사이트만 사용할 수 있고 해당 인사이트에 있는 개인 데이터의 기여도는 난독화됩니다. 차등 프라이버시는 특정 개인에 대해 학습한 결과를 수신할 수 있는 구성원에서 얻은 공동 작업 데이터를 보호합니다. 차등 프라이버시를 사용하지 않고 결과를 수신할 수 있는 구성원은 개인에 대한 기록을 추가하거나 제거하고 쿼리 결과의 차이점을 관찰하여 개별 사용자 데이터를 추론할 수 있습니다.

차등 프라이버시 기능이 켜져 있으면 쿼리 결과에 특정 양의 노이즈가 추가되어 개별 사용자의 기여도가 난독화됩니다. 결과를 받을 수 있는 구성원이 데이터 세트에서 개별에 대한 레코드를 제거한 후 쿼리 결과의 차이를 관찰하려고 하면 쿼리 결과의 변동성이 개별 데이터의 식별을 방지하는 데 도움이 됩니다. AWS Clean Rooms 차등 프라이버시는에서 개발한 검증된 올바른 샘플러 구현인 SampCert 샘플러를 사용합니다 AWS.

의 차등 프라이버시 AWS Clean Rooms 작동 방식

에서 차등 프라이버시를 활성화하는 워크플로에는 워크플로를 완료할 때 AWS Clean Rooms 다음과 같은 추가 단계가 AWS Clean Rooms 필요합니다.

1. 사용자 지정 분석 규칙을 추가할 때 차등 프라이버시 기능을 켭니다.
2. 쿼리에 사용할 수 있는 차등 프라이버시 기능으로 데이터 테이블을 보호하도록 공동 작업의 차등 프라이버시 정책을 구성합니다.

이 단계를 완료하면 쿼리할 수 있는 구성원이 차등 프라이버시 보호 데이터에 대한 쿼리 실행을 시작할 수 있습니다. 차등 프라이버시 정책을 준수하는 결과를 AWS Clean Rooms 반환합니다. AWS Clean Rooms 차등 프라이버시는 차량의 현재 연료 수준을 보여주는 자동차의 가스 게이지와 유사하게 실행할 수 있는 나머지 쿼리의 예상 수를 추적합니다. 쿼리할 수 있는 구성원이 실행할 수 있는 쿼리 수는 차등 프라이버시 정책에 설정된 프라이버시 예산 및 쿼리당 추가된 노이즈 파라미터에 따라 제한됩니다.

고려 사항

에서 차등 프라이버시를 사용할 때는 다음 사항을 AWS Clean Rooms고려하세요.

- 결과를 수신할 수 있는 구성원은 차등 프라이버시 기능을 사용할 수 없습니다. 구성한 테이블에서 차등 프라이버시 기능이 꺼진 상태로 사용자 지정 분석 규칙을 구성할 수 있습니다.
- 둘 이상의 데이터 공급자 모두에 차등 프라이버시 기능이 켜져 있으면 쿼리를 수행할 수 있는 구성원은 둘 이상의 데이터 공급자의 테이블을 조인할 수 없습니다.

차등 프라이버시 정책

차등 프라이버시 정책은 쿼리할 수 있는 구성원이 공동 작업에서 실행할 수 있는 집계 함수의 수를 제어합니다. 프라이버시 예산은 공동 작업의 모든 테이블에 적용되는 한정된 공통 리소스를 정의합니다. 쿼리당 추가되는 노이즈는 프라이버시 예산이 고갈되는 비율을 결정합니다.

차등 프라이버시 테이블을 쿼리에 사용할 수 있게 하려면 차등 프라이버시 정책이 필요합니다. 이는 공동 작업에서 한 번만 실행되는 단계이며 다음 2가지를 입력해야 합니다.

- 프라이버시 예산 - 엡실론의 측면에서 수치화하면 프라이버시 예산에서 프라이버시 수준을 제어할 수 있습니다. 여러 테이블에 정보가 있을 수 있는 사용자의 개인 정보를 보호하는 것이 목표이므로 공동 작업 시 차등 프라이버시로 보호되는 모든 테이블에 적용되는 일반적이고 유한한 리소스입니다.

  프라이버시 예산은 테이블에서 쿼리가 실행될 때마다 소비됩니다. 프라이버시 예산이 모두 소진되면 쿼리할 수 있는 공동 작업 구성원은 예산이 증대되거나 새로 고침될 때까지 추가 쿼리를 실행할 수 없습니다. 프라이버시 예산을 더 많이 설정하면 결과를 수신할 수 있는 구성원이 데이터 내 개인에 대한 불확실성을 줄일 수 있습니다. 비즈니스 의사 결정권자와 상의한 후 프라이버시 요구 사항과 공동 작업 요구 사항 간의 균형을 맞출 수 있는 프라이버시 예산을 선택합니다.

  공동 작업에 새 데이터를 정기적으로 가져오려는 경우 매월 프라이버시 예산 새로 고침을 선택하여 매월 새 프라이버시 예산을 자동으로 생성할 수 있습니다. 이 옵션을 선택하면 새로 고침 과정에서 반복해서 쿼리할 때 데이터 행에 대해 임의의 양의 정보가 공개될 수 있습니다. 프라이버시 예산 새로 고침 사이에 동일한 행이 반복적으로 쿼리되는 경우에는 이 옵션을 선택하지 않습니다.

- 쿼리당 추가되는 노이즈는 기여도를 난독화하고 싶은 사용자의 수를 기준으로 측정됩니다. 이 값에 따라 프라이버시 예산이 고갈되는 비율이 결정됩니다. 노이즈 값이 클수록 프라이버시 예산이 고갈되는 비율이 감소하므로 데이터에 대해 더 많은 쿼리를 실행할 수 있습니다. 하지만 정확도가 떨어지는 데이터 인사이트를 릴리스하는 것과 균형을 맞춰야 합니다. 이 값을 설정할 때는 공동 작업 인사이트에 필요한 정확도를 고려합니다.

기본 차등 프라이버시 정책을 사용하여 설정을 빠르게 완료하거나 사용 사례에 따라 차등 프라이버시 정책을 사용자 지정할 수 있습니다. AWS Clean Rooms 차등 프라이버시는 정책을 구성하는 직관적인 제어를 제공합니다. AWS Clean Rooms 차등 프라이버시를 사용하면 데이터의 모든 쿼리에서 가능한 집계 수 측면에서 유틸리티를 미리 보고 데이터 공동 작업에서 실행할 수 있는 쿼리 수를 추정할 수 있습니다.

대화형 예제를 사용하여 다양한 프라이버시 예산 값 및 쿼리당 추가되는 노이즈가 다양한 유형의 SQL 쿼리 결과에 어떤 영향을 미치는지 이해할 수 있습니다. 일반적으로 허용하려는 쿼리 수 및 해당 쿼리의 정확성과 프라이버시 요구 사항의 균형을 맞춰야 합니다. 프라이버시 예산을 줄이거나 쿼리당 추가되는 노이즈를 더 늘리면 사용자 프라이버시를 더 잘 보호할 수 있지만 공동 작업 파트너에게 의미 있는 인사이트를 제공하지 못할 수 있습니다.

쿼리당 추가되는 노이즈 파라미터를 동일하게 유지하면서 프라이버시 예산을 늘리면 쿼리할 수 있는 구성원이 공동 작업에서 테이블에 대해 더 많은 집계를 실행할 수 있습니다. 공동 작업 중에 언제든지 프라이버시 예산을 늘릴 수 있습니다. 쿼리당 추가되는 노이즈 파라미터를 동일하게 유지하면서 프라이버시 예산을 줄이면 쿼리할 수 있는 구성원이 실행할 수 있는 집계가 줄어듭니다. 쿼리할 수 있는 구성원이 데이터 분석을 시작하면 프라이버시 예산을 줄일 수 없습니다.

프라이버시 예산 입력값을 동일하게 유지하면서 쿼리당 추가되는 노이즈를 늘리면 쿼리할 수 있는 구성원이 공동 작업에서 테이블에 대해 더 많은 집계를 실행할 수 있습니다. 프라이버시 예산 입력값을 동일하게 유지하면서 쿼리당 추가되는 노이즈를 줄이면 쿼리할 수 있는 구성원이 실행할 수 있는 집계가 줄어듭니다. 공동 작업 기간 동안 언제든지 쿼리당 추가되는 노이즈를 늘리거나 줄일 수 있습니다.

차등 프라이버시 정책은 프라이버시 예산 템플릿 API 작업에서 관리됩니다.

AWS Clean Rooms 차등 프라이버시의 SQL 기능

AWS Clean Rooms 차등 프라이버시는 범용 쿼리 구조를 사용하여 복잡한 SQL 쿼리를 지원합니다. 사용자 지정 분석 템플릿은 차등 프라이버시로 보호되는 테이블에서 실행할 수 있도록 이 구조에 대해 검증됩니다. 다음 테이블에는 지원되는 함수가 나와 있습니다. 자세한 정보는 쿼리 구조 및 구문 섹션을 참조하세요.

SQL
SELECT
(CTE)
AWS Clean Rooms
SQL

ANY_VALUE
APPROXIMATE PERCENTILE_DISC
AVG
COUNT
COUNT DISTINCT

SQL 구문 SELECT 쿼리에 공통 테이블 표현식(CTE)이 포함된 AWS Clean Rooms SQL 구문

- LISTAGG 함수
- MEDIAN 함수
- PERCENTILE_CONT 함수
- PERCENTILE_DISC 함수
- STDDEV_SAMP 및 STDDEV_POP 함수

SQL
문
SELECT
쿼
리
결
과
에
서
제
공
(CTE)
되
는
AWS
Clean
Rooms
SQL
구
문

함
수
STDDEV_PO
P
- SUM
함
수
SUM
- DISTINCT
함
수
SUM
DISTINCT
- VAR_SAMP
함
수
VAR_POP
- VAR_SAMP
함
수
VAR_POP
함
수

SELECT
(CTE)
AWS Clean Rooms
SQL

WITH clause subquery

CTEs using different ial privacy protected tables must result in data with user-

SELECT
(CTE)
AWS
Clean
Rooms
SQL

leve
l
records.
You
should
write
the
SELECT
expressio
n
in
those
CTEs
using
`SELECT
userIdent
ifierColu

SQL
SELECT
(CTE)
AWS
Clean
Rooms
SQL

mn...'
format.

SQL
문
SELECT
쿼
리
에
서
공
통
테
이
블
식
(CTE)
또
는
AWS
Clean
Rooms
SQL
구
문

subquery
• SELECT
• HAVING
• JOIN
• JOIN
조
건
• FROM
• WHERE

SQL
구문
SELECT
쿼리
예제
제한
(CTE)
및
AWS
Clean
Rooms
SQL
구
문

have
any
subquery
that
reference
s
different
ial
privacy
relations
in
a
FROM
and
JOIN
clause
only.

SQL
문
SELECT
쉘
엔
결
예
쉘
제
(C
되
는
AWS
Clean
Rooms
SQL
구
문

- JOIN
- 자
- 식
- 별
- 자
- 열
- 의미
- 동
- 등
- 조
- 인
- 인
- JOIN
- 함
- 수
- 만
- 조

INNER
JOIN
LEFT
JOIN
RIGHT
JOIN
FULL
JOIN
[JOIN]
OR
연
쪽
에
티
CROSS
JOIN
인
RIGHT
JOIN

Spark SQL 플랜 SELECT 쿼리 엔진 결과 예제 테이블 제한(CTE) 되는 AWS Clean Rooms SQL 구문

- 전체 외부 조인(FULL OUTER JOIN)
- 로지스틱 회귀 분석[JOIN]
- 모델 OR 연산
- 배열 함수
- 교차 조인(CROSS JOIN)

프라이버시가 설정된 상

SQL
문
SELECT
쉘
엔
결
예
체
제
(CTE)
되
는
AWS
Clean
Rooms
SQL
구
문

태
에
서
둘
이
상
의
테
이
블
을
쿼
리
할
때
는
필
수
상

SQL
문 SELECT
쌜
엔
결
예
체
제
(CTE)
되
는
AWS
Clean
Rooms
SQL
구
문

태
로
지
원
됩
니
다.
필
수
동
등
조
인
조
건
이
올
바
른

SQL
문
SELECT
쉘
엔
결
예
쉘
제
(CTE)
되
는
AWS
Clean
Rooms
SQL
구
문

지
확
인
합
니
다.
테
이
블
소
유
자
가
모
든
테
이
블
에

SQL
문
SELECT
쉘
엔
결
예
체
제
(CTE)
되
는
AWS
Clean
Rooms
SQL
구
문

서
동
일
한
사
용
자
식
별
자
열
을
구
성
하
면
사
용
자

SQL 문 SELECT 쿼리 엔진 결과 예제 테이블 제공 (CTE) 되는 AWS Clean Rooms SQL 구문

정의가 테이블 간에 일관되게 유지됩니다.

SQL
SELECT
(CTE)
AWS Clean Rooms
SQL
차등 프라이버시가 켜진 상태에서 둘 이상의 관

SQL
문
SELECT
쉘
엔
결
예
쉘
제
(CTE)
되
는
AWS
Clean
Rooms
SQL
구
문

계
를
결
합
할
때
CROSS
JOIN
함
수
는
지
원
되
지
않
습

SQL
문
SELECT
쿼
리
결
과
세
트
(CTE)
또
는
AWS
Clean
Rooms
SQL
구
문
니
다.

SUNION,
UNION
ALL,
INTERSECT
,
EXCEPT
|
MINUS
(these
are
synonyms)

SQL 구문 SELECT 쿼리 엔진 결과 예제 쿼리 제한(CTE) 또는 AWS Clean Rooms SQL 구문

Window functions supported with the condition that the user identifier column in the window function's partition

- AVG 창 함수
- COUNT 창 함수
- CUME_DIST 창 함수

SQL 구문 및 SELECT 쿼리 예제 쿼리 예제 (CTE) 또는 AWS Clean Rooms SQL 구문

- classifier DENSE_RANK
- first_value FIRST_VALUE
- lag LAG
- last_value LAST_VALUE

SQL
구
문
SELECT
목
록
절
예
제
(CTE)
및
AWS
Clean
Rooms
SQL
구
문

함
수

- LEAD
  창
  함
  수
- MAX
  창
  함
  수
- MEDIAN
  창
  함
  수
- MIN
  창
  함
  수

SQL
문
SELECT
절
에
대
한
제
한
(CTE)
및
AWS
Clean
Rooms
SQL
구
문

- NTH_VALUE
창
함
수
- STDDEV_SA
MP
창
함
수
- STDDEV_PO
P
창
함
수
- STDDEV_SA
MP
(STDDEV
_S
AMP
)
창
함
수
- STDDEV_PO
P
(STDDEV

SQL
SELECT
(CTE)
AWS
Clean
Rooms
SQL
SAMP
STDDEV
SUM
VAR_SAMP
SUM
VAR_POP
VAR_SAMP
VAR_POP
VARIANCE
VAR_SA

SQL
SELECT
(CTE)
AWS Clean Rooms
SQL
VARIANCE

- DENSE_RANK
- DENSE_RANK
- PERCENT_RANK

SQL
문
SELECT
섹
엔
결
예
섹
제
(CTE)
되
는
AWS
Clean
Rooms
SQL
구
문

- 창
  함
  참
  수
- RANK
- 창
  참
  창
  함
- ROW_NUMBE
  R
- 창
  함
  참
  수
- ROW_NUMBE
  R
  창
  함
  수

SQL
구
문
SELECT
쿼
리
예
제
(CTE)
및
AWS
Clean
Rooms
SQL
구
문

ANSI
SQL
표
준
에
서
지
원
하
는
식

- COALESCE
  표
  현
  식
- GREATEST
  및
  LEAST
  함
  수
- NVL
  및 COALESCE
  함
  수

SQL
구
문
SELECT
절
예
제
(CTE)
또
는
AWS
Clean
Rooms
SQL
구
문

- NVL2
  함
  수
- NULLIF
  함
  수

SQL
SELECT
공통
테이블
표현식
(CTE)
및
재귀
공통
테이블
표현식
(CTE)
AWS
Clean
Rooms
SQL
구
문

Additional
and
not
supported
cannot
be
used
because
they
require
subqueries.
All
others
are
supported.

- BETWEEN
- 패턴
- 조건
- 비교
- 논리
- 집합
- 서브쿼리
- 조건
- 기타

SQL
SELECT
(CTE)
AWS
Clean
Rooms
SQL

- NULL

SQL
구
문
SELECT
절
예
제
재
귀
공
통
테
이
블
식
(CTE)
AWS
Clean
Rooms
SQL
구
문

BAE-
table
fsupported

-
-

SQL
구
문
SELECT
절
예
제
(CTE)
또
는
AWS
Clean
Rooms
SQL
구
문

- ADD_MONTHS
  함
  수
- CONVERT_TIMEZONE
  함
  수
- CURRENT_DATE
  함
  수
- DATEADD
  함
  수
- DATEDIFF
  함
  수

SQL
구
문
SELECT
절
에
서
의
제
한
(CTE)
되
는
AWS
Clean
Rooms
SQL
구
문

- DATE_PART
  함
  수
- DATE_TRUNC
  함
  수
- EXTRACT
  함
  수
- GETDATE
  함
  수
- TIMEOFDAY
- 함
  수
- TO_TIMESTAMP

SQL
SELECT
(CTE)
AWS Clean Rooms
SQL

- 함수
  - 날짜 부분

SQL
문
SELECT
쿼
리
예
제
재
귀
공
통
테
이
블
식
(CTE)
및
AWS
Clean
Rooms
SQL
구
문

All
supported
operators
and
functions

- 연산자
- BTRIM 함수
- CHAR_LENGTH 함수
- CHARACTER_LENGTH 함수

SQL
문
SELECT
섹
엔
결
예
첨
제
문(CTE)
되
는
SQL
Clean
Rooms
SQL
구
문

- CONCAT
  함
  수
- LEFT
  및
  RIGHT
  함
  수
- LEN
  함
  수
- LENGTH
  함
  수
- LOWER
  함
  수

SQL
구
문
SELECT
쿼
리
결
과
세
트
제
한
(CTE)
및
AWS
Clean
Rooms
SQL
구
문

- LPAD
  및
  RPAD
  함
  수
- ltrim
  함
  수
- POSITION
  함
  수
- REGEXP_CO
  UNT
  함
  수
- REGEXP_IN
  STR

SQL
SELECT
CTE)
AWS
Clean
Rooms
SQL

- REGEXP_REPLACE
- REGEXP_SUBSTR
- REPEAT
- REPLACE

SQL
SELECT
(CTE)
AWS Clean Rooms
SQL

- REVERSE
- REVERSE
- SPLIT_PART
- SOUNDEX
- SUBSTRING
- SPLIT_PART
- TRANSLATE

SQL
문
SELECT
쿼
리
결
과
세
트
공
통
테
이
블
식
(CTE)
및
재
귀
CTE
AWS
Clean
Rooms
SQL
구
문

- STRPOS
  함
  수
- SUBSTRING
  함
  수
- TEXTLEN
  함
  수
- TRANSLATE
  함
  수
- TRIM
  함
  수

SQL
문
SELECT
쿼
리
결
과
에
서
제
외
(CTE)
되
는
AWS
Clean
Rooms
SQL
구
문

- UPPER
  함
  수

SQL
SELECT
CTE)
AWS
Clean
Rooms
SQL

CAST
type
supported
functions

- TO_CHAR
- TO_DATE
- TO_NUMBER
-

SQL
SELECT
(CTE)
AWS
Clean
Rooms
SQL

- 숫자 형식 문자열

SQL
구
문
SELECT
쿼
리
결
과
예
제
쿼
리
제
한
(CTE)
및
하
위
AWS
쿼
Clean
리
Rooms
SQL
구
문

AWS_ENCRY
functions
supported
AWS_DECRY

- SHA
- ENCODE
- DECODE
- SHA1
- MD5
- SHA2
- SHA1
- MURMUR3_3
- SHA2
- XX_HASH64

SQL 구문 SELECT 쿼리 예제 결합 예제 쿼리 제한 공통 테이블 표현식(CTE) 및 재귀 CTE AWS Clean Rooms SQL 구문

Mathematical functions supported in AWS Clean Rooms SQL and @

SQL
구
문
SELECT
구
문
테
이
블
조
인
구
문
및
공
통
테
이
블
식
(CTE)
에
는
AWS
Clean
Rooms
SQL
구
문
이
나
supported
functions
에
서
지
원
하
는
함
수
가
포
함
되
어
서
는
안
됩
니
다
.

- ACOS
  함
  수
- ASIN
  함
  수
- ATAN
  함
  수
- ATAN2
  함
  수

SQL
문
SELECT
섹
엔
결
예
섹
제
식(CTE)
되
는
AWS
Clean
Rooms
SQL
구
문

- CBRT
  함
  수
- CEILING(또
  는
  CEIL)
  함
  수
- COS
  함
  수
- COT
  함
  수
- DEGREES
  함
  수

SQL
구
문
SELECT
쿼
리
예
제
(CTE)
및
하
위
쿼
리
AWS
Clean
Rooms
SQL
구
문

- DATEXP
  함
  수
- EXP
  함
  수
- FLOOR
  함
  수
- LN
  함
  수
- LOG
  함
  수
- MOD
  함
  수

SQL
SELECT
CTE)
AWS
Clean
Rooms
SQL

- PI
- POWER
- RADIANS
- RANDOM
- ROUND
- SIGN

SQL
문
SELECT
쿼
리
결
과
세
트
제
한
(CTE)
되
는
AWS
Clean
Rooms
SQL
구
문

- RANDOM
  함
  수
- ROUND
  함
  수
- SIGN
  함
  수
- SIN
  함
  수
- SQRT
  함
  수

SQL
구
문
SELECT
문
에
대
한
공
통
테
이
블
식
(CTE)
및
AWS
Clean
Rooms
SQL
구
문

- TRUNC
  함
  수
- VARBYTE_
  UNBASE64
  함
  수
- FROM_VARB
  YTE
  함
  수
- HLL_SKETC
  H_AGG,
  함
  수
- HLL_SKETC
  H_ESTIMAT
  E
  함
  수
- HLL_UNION
  _AGG
  함
  수

SELECT 문에서 공통 테이블 표현식(CTE)으로 사용되는 AWS Clean Rooms SQL 구문

- GET_JSON_OBJECT 함수
- JSON_EXTRACT_ARRAY_ELEMENT_TEXT 함수
- JSON_EXTRACT_PATH_TEXT 함수
- JSON_PARSE

SELECT
(CTE)
SQL
Clean
Rooms
SQL

함
수

- JSON_SERI
  ALIZE
  함
  수
- JSON_SERA
  LIZE_TO_V
  ARBYTE
  함
  수

SQL
구
문
SELECT
문
에
대
한
참
조
(CTE)
를
사
용
하
여
AWS
Clean
Rooms
SQL
구
문

ARRAY_CON
CAT
함
수
ARRAY_DIS
TINCT
함
수
ARRAY_EXC
EPT
함
수
ARRAY_INT
ERSECT
함
수
ARRAY_JOI
N
함
수
ARRAY_REM
OVE
함
수
ARRAY_SOR
T
함
수
ARRAY_UNI
ON
함
수

- array_con
- array_fla
- get_array
_length
- split_to_
array

SQL
문
SELECT
쿼
리
결
과
에
서
제
외
(CTE)
되
는
AWS
Clean
Rooms
SQL
구
문

- 함
  수
- 부
  분
  배
  열
  함
  수

GROUPING
SETS,
ROLLUP,
CUBE

SELECT
SQL
Clean
Rooms
SQL
구
문

ORDER BY
the
condition
that
an
ORDER
BY
clause
is
only
supported
in
a
window
function'
s
partition

SQL
구
문
SELECT
명
령
문
(CTE)
및
AWS
Clean
Rooms
SQL
구
문

clause
when
querying
tables
with
different
ial
privacy
turned
on.

SQL
SELECT
(CTE)
AWS Clean Rooms
SQL

supported
CTEs
using
different
ial
privacy
protected
tables

Supported
and
column
aliasing

Spark SQL 쿼리 SELECT 섹션 엔진 절 예제 섹션 제한(CTE) 되는 AWS Clean Rooms SQL 구문

Supported functions on aggregate functions

Supported functions within aggregate functions

지원되지 않는 SQL 구성에 대한 공통 대안

범주	SQL 구성	대안
윈도 함수	• LISTAGG • PERCENTILE_CONT • PERCENTILE_DISC	You can use the equivalent aggregate function with GROUP BY.
Mathematical operator symbols	• $column ||/ 2 • $column |/ 2 • $column ^ 2	• CBRT • SQRT • POWER($column, 2)
Scalar functions	• SYSDATE • $column::정수 • convert(유형, $column)	• CURRENT_DATE • CAST $column AS 정수 • CAST $column AS 유형
Literals	INTERVAL ‘1 SECOND'	INTERVAL '1' SECOND
Row limiting	TOP n	LIMIT n
Join	• USING • NATURAL	ON clause should explicitly contain a join criterion.

차등 프라이버시 쿼리 팁 및 예제

AWS Clean Rooms 차등 프라이버시는 범용 쿼리 구조를 사용하여 데이터 준비를 위한 공통 테이블 표현식(CTEs) 및 COUNT또는와 같이 일반적으로 사용되는 집계 함수와 같은 다양한 SQL 구문을 지원합니다SUM. 런타임 시 집계 쿼리 결과에 노이즈를 추가하여 데이터에서 가능한 사용자의 기여도를 난독화하려면 AWS Clean Rooms 차등 프라이버시를 사용하려면 최종의 집계 함수가 사용자 수준 데이터에서 실행SELECT statement되어야 합니다.

다음 예제에서는 athletic_brand_sales 데이터가 포함된 스포츠 브랜드와 공동 작업하면서 차등 프라이버시를 사용하여 데이터를 보호하려는 미디어 게시자의 socialco_impressions 및 socialco_users라는 테이블 2개를 사용합니다. 미디어 게시자는 이 user_id 열을 사용자 식별자 열로 구성하면서 차등 프라이버시를 AWS Clean Rooms에서 사용 설정했습니다. 광고주는 차등 프라이버시 보호가 필요하지 않으므로 결합된 데이터에서 CTE를 사용하여 쿼리를 실행하려고 합니다.

CTE는 차등 프라이버시 보호 테이블을 사용하기 때문에 광고주는 보호 테이블의 사용자 식별자 열을 CTE 열 목록에 포함시키고 사용자 식별자 열에 보호 테이블을 조인합니다.

```
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
```

마찬가지로 차등 프라이버시 보호 데이터 테이블에서 창 함수를 실행하려면 `PARTITION BY` 절에 사용자 식별자 열을 포함해야 합니다.

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

AWS Clean Rooms 차등 프라이버시의 제한 사항

AWS Clean Rooms 차등 프라이버시는 다음 상황을 해결하지 않습니다.

1. AWS Clean Rooms 차등 프라이버시는 Amazon S3 지원 AWS Glue 테이블을 사용하는 쿼리만 지원합니다. Snowflake 또는 Amazon Athena 테이블을 사용한 쿼리는 지원하지 않습니다.

2. AWS Clean Rooms 차등 프라이버시는 타이밍 공격을 해결하지 않습니다. 예를 들어 개별 사용자가 많은 행을 제공하고 이 사용자를 추가하거나 제거하면 쿼리 계산 시간이 크게 달라지는 시나리오에서 이러한 공격이 발생할 수 있습니다.
3. AWS Clean Rooms 차등 프라이버시는 SQL 쿼리로 인해 특정 SQL 구문의 사용으로 인해 런타임에 오버플로 또는 잘못된 캐스트 오류가 발생할 수 있는 경우 차등 프라이버시를 보장하지 않습니다.

   다음 표는 런타임 오류를 생성할 수 있으므로 분석 템플릿에서 확인해야 하는 일부 SQL 구문(전부는 아님) 목록입니다. 이러한 런타임 오류의 가능성을 최소화하는 분석 템플릿을 승인하고 쿼리 로그를 주기적으로 검토하여 쿼리가 공동 작업 계약에 부합하는지 확인하는 것이 좋습니다.

   다음 SQL 구문은 오버플로 오류에 취약합니다.

범주	Spark SQL 분석 엔진의 오버플로 오류에 취약한 SQL 구문	SQL 분석 엔진의 오버플로 오류에 취약한 AWS Clean Rooms SQL 구문
Aggregate functions	• AVG • SUM/SUM_DISTINCT	• AVG • LISTAVG • PERCENTILE_COUNT • PERCENTILE_DISC • SUM/SUM_DISTINCT
Data type formatting functions	• TO_TIMESTAMP • TO_DATE	• TO_TIMESTAMP • TO_DATE
Date and time functions	• ADD_MONTHS • DATEADD • DATEDIFF	• ADD_MONTHS • DATEADD • DATEDIFF
Math functions	• +, -, *, / • POWER	• +, -, *, / • POWER
String functions	• || • CONCAT • 반복	• || • CONCAT • 반복 • REPLICATE

범주	Spark SQL 분석 엔진의 오버플로 오류에 취약한 SQL 구문	SQL 분석 엔진의 오버플로 오류에 취약한 AWS Clean Rooms SQL 구문
Window functions	• AVG • SUM	• AVG • LISTAVG • PERCENTILE_COUNT • PERCENTILE_DISC • RATIO_TO_REPORT • SUM

4. CAST 데이터 유형 서식 설정 함수는 잘못된 캐스트 오류에 취약합니다.

   로그 그룹에 대한 지표 필터를 생성하도록 CloudWatch를 구성한 다음 해당 지표 필터에 대해 CloudWatch 경보를 생성하여 잠재적인 오버플로 또는 캐스팅 오류가 발생할 경우 알림을 받을 수 있습니다.

   특히 오류 코드 `CastError`, `OverflowError`, `ConversionError`를 모니터링해야 합니다. 이러한 오류 코드는 잠재적인 부채널 공격을 나타내지만 잘못된 SQL 쿼리를 나타낼 수도 있습니다.

   자세한 내용은 의 분석 로깅 AWS Clean Rooms 단원을 참조하십시오.

AWS Clean Rooms ML

AWS Clean Rooms ML을 사용하면 두 명 이상의 당사자가 데이터를 서로 공유할 필요 없이 데이터에 대해 기계 학습 모델을 실행할 수 있습니다. 이 서비스는 데이터 소유자가 데이터와 모델 IP를 안전하게 보호할 수 있도록 개인 정보 보호 강화 제어를 제공합니다. AWS 작성 모델을 사용하거나 자체 사용자 지정 모델을 가져올 수 있습니다.

작동하는 방식에 대한 자세한 설명은 교차 계정 작업 섹션을 참조하세요.

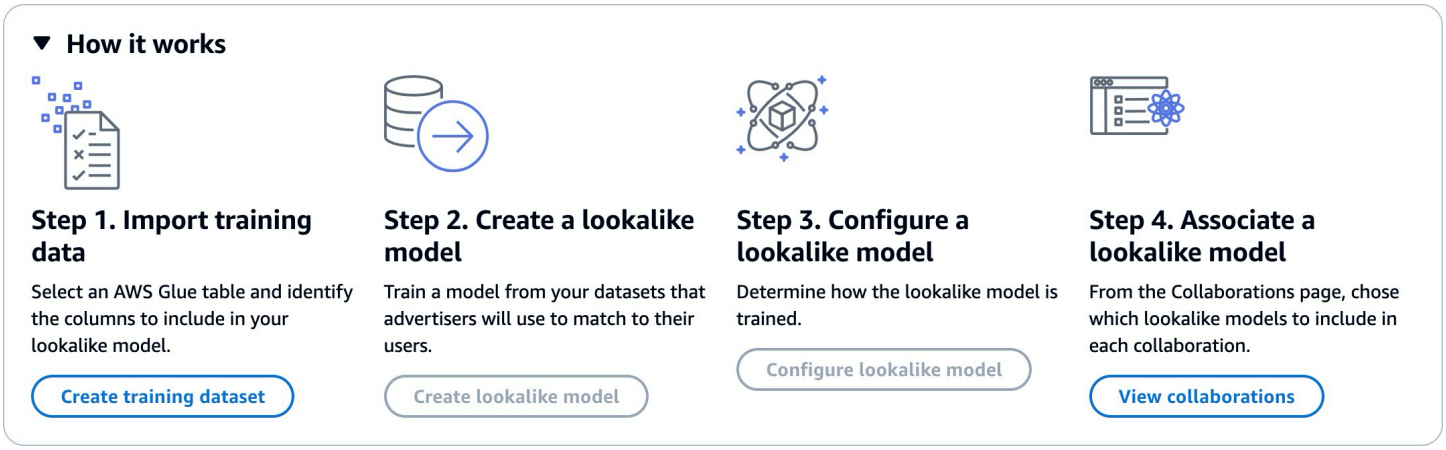
Clean Rooms ML 모델의 기능에 대한 자세한 내용은 다음 주제를 참조하세요.

주제

- AWS Clean Rooms ML이 AWS 모델과 작동하는 방식
- AWS Clean Rooms ML이 사용자 지정 모델과 작동하는 방식
- AWS Clean Rooms ML의 모델

- Clean Rooms ML의 사용자 지정 모델

AWS Clean Rooms ML이 AWS 모델과 작동하는 방식



유사 모델을 사용하려면 훈련 데이터 공급자와 시드 데이터 공급자라는 두 당사자가 순차적으로에서 작업 AWS Clean Rooms 하여 데이터를 공동 작업으로 가져와야 합니다. 훈련 데이터 공급자가 먼저 완료해야 하는 워크플로는 다음과 같습니다.

1. 훈련 데이터 공급자의 데이터는 사용자-항목 상호 작용의 AWS Glue 데이터 카탈로그 테이블에 저장되어야 합니다. 훈련 데이터에는 최소한 사용자 ID 열, 상호 작용 ID 열 및 타임스탬프 열이 포함되어야 합니다.
2. 훈련 데이터 공급자는 훈련 데이터를에 등록합니다 AWS Clean Rooms.
3. 훈련 데이터 공급자는 여러 시드 데이터 공급자와 공유할 수 있는 유사 모델을 생성합니다. 유사 모델은 신경망이며, 훈련하는 데 최대 24시간이 걸릴 수 있습니다. 자동으로 재훈련되지 않으므로 매주 모델을 재훈련하는 것이 좋습니다.
4. 훈련 데이터 공급자는 관련성 지표 공유 여부 및 출력 세그먼트의 Amazon S3 위치를 포함하여 유사 모델을 구성합니다. 훈련 데이터 공급자는 단일 유사 모델에서 구성된 유사 모델을 여러 개 생성할 수 있습니다.
5. 훈련 데이터 공급자는 구성된 대상 모델을 시드 데이터 공급자와 공유하는 공동 작업에 연결합니다.

시드 데이터 공급자가 다음으로 완료해야 하는 워크플로는 다음과 같습니다.

1. 시드 데이터 공급자의 데이터는 Amazon S3 버킷에 저장하거나 쿼리 결과에서 가져올 수 있습니다.
2. 시드 데이터 공급자는 훈련 데이터 공급자와 공유하는 공동 작업을 엽니다.
3. 시드 데이터 공급자는 공동 작업 페이지의 Clean Rooms ML 탭에서 유사 세그먼트를 만듭니다.

4. 시드 데이터 공급자는 관련성 지표가 공유된 경우 이를 평가하고 유사 세그먼트를 내보내 AWS Clean Rooms외부에서 사용할 수 있습니다.

AWS Clean Rooms ML이 사용자 지정 모델과 작동하는 방식

Clean Rooms ML을 사용하면 공동 작업 구성원이 Amazon ECR에 저장된 도킹된 사용자 지정 모델 알고리즘을 사용하여 데이터를 공동 분석할 수 있습니다. 이렇게 하려면 모델 공급자가 이미지를 생성하고 Amazon ECR에 저장해야 합니다. Amazon Elastic Container Registry 사용 설명서의 단계에 따라 사용자 지정 ML 모델을 포함할 프라이빗 리포지토리를 생성합니다.

올바른 권한이 있는 경우 공동 작업의 모든 구성원이 모델 공급자가 될 수 있습니다. 공동 작업의 모든 구성원은 훈련 데이터, 추론 데이터 또는 둘 다를 모델에 제공할 수 있습니다. 이 안내서의 목적상 데이터를 제공하는 구성원을 데이터 공급자라고 합니다. 공동 작업을 생성하는 구성원은 공동 작업 생성자이며,이 구성원은 모델 공급자, 데이터 공급자 중 하나 또는 둘 다일 수 있습니다.

다음은 가장 높은 수준에서 사용자 지정 ML 모델링을 수행하기 위해 완료해야 하는 단계입니다.

1. 공동 작업 생성자는 공동 작업을 생성하고 각 구성원에게 적절한 구성원 기능 및 결제 구성을 할당합니다. 공동 작업 생성자는이 단계에서 모델 출력을 수신하거나 추론 결과를 수신할 수 있는 구성원 기능을 적절한 구성원에게 할당해야 합니다. 공동 작업이 생성된 후에는 업데이트할 수 없기 때문입니다. 자세한 내용은 AWS Clean Rooms ML에서 공동 작업 생성 및 참여 단원을 참조하십시오.
2. 모델 공급자는 컨테이너화된 ML 모델을 구성하여 공동 작업에 연결하고 내보낸 데이터에 대해 개인 정보 보호 제약 조건이 설정되도록 합니다. 자세한 내용은 AWS Clean Rooms ML에서 모델 알고리즘 구성 단원을 참조하십시오.
3. 데이터 공급자는 데이터를 공동 작업에 기여하고 개인 정보 보호 요구 사항이 지정되도록 합니다. 데이터 공급자는 모델이 데이터에 액세스할 수 있도록 허용해야 합니다. 자세한 내용은 AWS Clean Rooms ML에서 훈련 데이터 기여 및 AWS Clean Rooms ML에서 구성된 모델 알고리즘 연결 섹션을 참조하세요.
4. 공동 작업 구성원은 모델 아티팩트 또는 추론 결과를 내보낼 위치를 정의하는 ML 구성을 생성합니다.
5. 공동 작업 구성원은 훈련 컨테이너 또는 추론 컨테이너에 입력을 제공하는 ML 입력 채널을 생성합니다. ML 입력 채널은 모델 알고리즘의 컨텍스트에서 사용할 데이터를 정의하는 쿼리입니다.
6. 공동 작업 구성원은 ML 입력 채널과 구성된 모델 알고리즘을 사용하여 모델 훈련을 호출합니다. 자세한 내용은 AWS Clean Rooms ML에서 훈련된 모델 생성 단원을 참조하십시오.
7. (선택 사항) 모델 트레이너는 모델 내보내기 작업을 호출하고 모델 아티팩트는 모델 결과 수신기로 전송됩니다. 유효한 ML 구성을 가진 멤버와 모델 출력을 수신할 수 있는 멤버만 모델 아티팩트를 수

신할 수 있습니다. 자세한 내용은 AWS Clean Rooms ML에서 모델 아티팩트 내보내기 단원을 참조하십시오.

8. (선택 사항) 공동 작업 구성원은 ML 입력 채널, 훈련된 모델 ARN 및 추론 구성 모델 알고리즘을 사용하여 모델 추론을 호출합니다. 추론 결과는 추론 출력 수신기로 전송됩니다. 유효한 ML 구성을 가진 멤버와 추론 출력을 수신할 수 있는 멤버만 추론 결과를 수신할 수 있습니다.

다음은 모델 공급자가 완료해야 하는 단계입니다.

1. SageMaker AI 호환 Amazon ECR Docker 이미지를 생성합니다. Clean Rooms ML은 SageMaker AI 호환 도커 이미지만 지원합니다.
2. SageMaker AI 호환 도커 이미지를 생성한 후 이미지를 Amazon ECR에 푸시합니다. Amazon Elastic Container Registry 사용 설명서의 지침에 따라 컨테이너 훈련 이미지를 생성합니다.
3. Clean Rooms ML에서 사용할 모델 알고리즘을 구성합니다.
   a. Amazon ECR 리포지토리 링크와 모델 알고리즘을 구성하는 데 필요한 인수를 제공합니다.
   b. Clean Rooms ML이 Amazon ECR 리포지토리에 액세스할 수 있는 서비스 액세스 역할을 제공합니다.
   c. 구성된 모델 알고리즘을 공동 작업과 연결합니다. 여기에는 컨테이너 로그, 실패 로그, CloudWatch 지표 및 컨테이너 결과에서 내보낼 수 있는 데이터의 양에 대한 제한에 대한 제어를 정의하는 개인 정보 보호 정책 제공이 포함됩니다.

다음은 데이터 공급자가 사용자 지정 ML 모델과 협업하기 위해 완료해야 하는 단계입니다.

1. 사용자 지정 분석 규칙을 사용하여 기존 AWS Glue 테이블을 구성합니다. 이렇게 하면 사전 승인된 특정 쿼리 세트 또는 사전 승인된 계정이 데이터를 사용할 수 있습니다.
2. 구성된 테이블을 공동 작업과 연결하고 AWS Glue 테이블에 액세스할 수 있는 서비스 액세스 역할을 제공합니다.
3. 구성된 모델 알고리즘 연결이 구성된 테이블에 액세스할 수 있도록 하는 공동 작업 분석 규칙을 테이블에 추가합니다.
4. Clean Rooms ML에서 모델과 데이터를 연결하고 구성한 후 쿼리를 실행할 수 있는 구성원은 SQL 쿼리를 제공하고 사용할 모델 알고리즘을 선택합니다.

모델 훈련이 완료되면 해당 구성원은 모델 훈련 아티팩트 또는 추론 결과의 내보내기를 시작합니다. 이러한 아티팩트 또는 결과는 훈련된 모델 출력을 수신할 수 있는 기능을 가진 멤버에게 전송됩니다. 결과 수신기는 모델 출력을 수신하기 `MachineLearningConfiguration` 전에를 구성해야 합니다.

AWS Clean Rooms ML의 모델

AWS Clean Rooms ML은 두 당사자가 데이터를 서로 공유할 필요 없이 데이터에서 유사한 사용자를 식별할 수 있는 개인 정보 보호 방법을 제공합니다. 첫 번째 당사자는 유사 모델을 생성 및 구성하고 이를 공동 작업과 연결할 수 AWS Clean Rooms 있도록 훈련 데이터를에 가져옵니다. 그런 다음 시드 데이터를 공동 작업에 가져와서 훈련 데이터와 유사한 유사 세그먼트를 만듭니다.

작동하는 방식에 대한 자세한 설명은 교차 계정 작업 섹션을 참조하세요.

다음 주제에서는 Clean Rooms ML에서 AWS 모델을 생성하고 구성하는 방법에 대한 정보를 제공합니다.

주제

- AWS Clean Rooms ML 용어
- AWS Clean Rooms ML의 개인 정보 보호
- Clean Rooms ML의 훈련 데이터 요구 사항
- Clean Rooms ML의 시드 데이터 요구 사항
- AWS Clean Rooms ML 모델 평가 지표

AWS Clean Rooms ML 용어

Clean Rooms ML을 사용할 때는 다음 용어를 이해하는 것이 중요합니다.

- 훈련 데이터 공급자 – 훈련 데이터를 제공하고 유사 모델을 생성 및 구성한 다음 해당 유사 모델을 공동 작업에 연결하는 역할을 합니다.
- 시드 데이터 공급자 - 시드 데이터를 제공하고 유사 세그먼트를 생성하여 유사 세그먼트를 내보내는 역할을 합니다.
- 훈련 데이터 - 유사 모델을 생성하는 데 사용되는 훈련 데이터 공급자의 데이터입니다. 훈련 데이터는 사용자 동작의 유사성을 측정하는 데 사용됩니다.

  훈련 데이터에는 사용자 ID, 항목 ID 및 타임스탬프 열이 포함되어야 합니다. 필요에 따라 훈련 데이터에 다른 상호작용을 수치적 특징 또는 범주형 특징으로 포함할 수 있습니다. 상호작용의 예로는 시청한 동영상, 구매한 항목, 읽은 기사 목록 등이 있습니다.
- 시드 데이터 - 유사 세그먼트를 만드는 데 사용되는 시드 데이터 공급자의 데이터입니다. 시드 데이터는 직접 제공하거나 AWS Clean Rooms 쿼리 결과에서 가져올 수 있습니다. 유사 세그먼트는 시드 사용자와 가장 유사한 훈련 데이터의 사용자 집합입니다.

- 유사 모델 - 다른 데이터 세트에서 유사한 사용자를 찾는 데 사용되는 훈련 데이터의 기계 학습 모델입니다.

  API를 사용할 때 대상 모델이라는 용어는 유사 모델과 동일하게 사용됩니다. 예를 들어 CreateAudienceModel API를 사용하여 유사 모델을 만들 수 있습니다.
- 유사 세그먼트 - 시드 데이터와 가장 유사한 훈련 데이터의 하위 집합입니다.

  API를 사용할 때 StartAudienceGenerationJob API를 사용하여 유사 세그먼트를 생성합니다.

훈련 데이터 공급자의 데이터는 시드 데이터 공급자와 공유되지 않으며 시드 데이터 공급자의 데이터도 훈련 데이터 공급자와 공유되지 않습니다. 유사 세그먼트 출력은 훈련 데이터 공급자와 공유되지만 시드 데이터 공급자와는 공유되지 않습니다.

AWS Clean Rooms ML의 개인 정보 보호

Clean Rooms ML은 훈련 데이터 공급자가 시드 데이터에 있는 사용자를 알고 시드 데이터 공급자가 훈련 데이터에 있는 사용자를 알 수 있는 멤버십 추론 공격의 위험을 줄이도록 설계되었습니다. 이 공격을 방지하기 위해 취할 수 있는 몇 가지 단계가 있습니다.

첫째, 시드 데이터 공급자는 Clean Rooms ML 결과를 직접 관찰하지 않으며 훈련 데이터 공급자는 시드 데이터를 절대 관찰할 수 없습니다. 시드 데이터 공급자는 출력 세그먼트에 시드 데이터를 포함하도록 선택할 수 있습니다.

다음으로, 훈련 데이터의 랜덤 샘플에서 유사 모델을 만듭니다. 이 샘플에는 시드 대상과 일치하지 않는 상당수의 사용자가 포함되어 있습니다. 이 프로세스를 통해 사용자가 데이터에 없는지 확인하기가 더 어려워지며, 이는 멤버십 추론의 또 다른 방법입니다.

또한 시드별 유사 모델 훈련의 모든 파라미터에 여러 시드 고객을 사용할 수 있습니다. 이로 인해 모델이 오버피팅할 수 있는 양과 사용자에 대해 추론할 수 있는 양이 제한됩니다. 따라서 시드 데이터의 최소 크기는 사용자 500명으로 설정하는 것이 좋습니다.

마지막으로, 사용자 수준 지표는 훈련 데이터 공급자에게 절대 제공되지 않으므로 멤버십 추론 공격의 또 다른 수단이 없어집니다.

Clean Rooms ML의 훈련 데이터 요구 사항

유사 모델을 성공적으로 생성하려면 훈련 데이터가 다음 요구 사항을 충족해야 합니다.

- 훈련 데이터는 Parquet, CSV 또는 JSON 형식이어야 합니다.

> **Note**
>
> Zstandard(ZSTD) 압축 Parquet 데이터는 지원되지 않습니다.

- 훈련 데이터는 카탈로그로 작성해야 합니다 AWS Glue. 자세한 내용은 AWS Glue 개발자 안내서의 AWS Glue 데이터 카탈로그 시작하기를 참조하세요. 스키마가 자동으로 추론되므로 AWS Glue 크롤러를 사용하여 테이블을 생성하는 것이 좋습니다.
- 훈련 데이터 및 시드 데이터가 포함된 Amazon S3 버킷은 다른 Clean Rooms ML 리소스와 동일한 AWS 리전에 있습니다.
- 훈련 데이터에는 항목 상호 작용이 각각 두 개 이상 있는 고유한 사용자 ID가 100,000개 이상 포함되어야 합니다.
- 학습 데이터에는 최소 1백만 개의 레코드가 포함되어야 합니다.
- CreateTrainingDataset 작업에 지정된 스키마는 AWS Glue 테이블이 생성될 때 정의된 스키마와 일치해야 합니다.
- 제공된 표에 정의된 필수 필드는 CreateTrainingDataset 작업에 정의되어 있습니다.

필드 유형	지원되는 데이터 유형	필수	설명
USER_ID	string, int, bigint	예	데이터 세트의 각 사용자에 대한 고유 식별자입니다. 개인 식별 정보(PII) 값이 아닌 값이어야 합니다. 해시 식별자 또는 고객 ID일 수 있습니다.
ITEM_ID	string, int, bigint	예	사용자가 상호 작용하는

필드 유형	지원되는 데이터 유형	필수	설명
			각 항목의 고유 식별자입니다.
TIMESTAMP	bigint, int, timestamp	예	사용자가 항목과 상호 작용한 시간입니다. 값은 Unix Epoch 시간(초) 형식이어야 합니다.

필드 유형	지원되는 데이터 유형	필수	설명
CATEGORICAL_FEATURE	string, int, float, bigint, double, boolean, array	아니요	사용자 또는 항목과 관련된 범주형 데이터를 캡처합니다. 여기에는 이벤트 유형(클릭 또는 구매 등), 사용자 인구 통계(연령 그룹, 성별 - 익명화), 사용자 위치(도시, 국가 - 익명화), 항목 범주(의류 또는 전자 제품 등) 또는 항목 브랜드가 포함될 수 있습니다.

필드 유형	지원되는 데이터 유형	필수	설명
NUMERICAL_FEATURE	double, float, int, bigint	아니요	사용자 또는 항목과 관련된 숫자 데이터를 캡처합니다. 여기에는 사용자 구매 내역(사용된 총 금액), 항목 가격, 항목을 방문한 횟수 또는 항목에 대한 사용자 평점이 포함될 수 있습니다.

- 선택적으로 범주형 또는 숫자형 기능을 최대 10개 제공할 수 있습니다.

다음은 CSV 형식의 유효한 훈련 데이터 세트의 예입니다.

```
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
```

Clean Rooms ML의 시드 데이터 요구 사항

유사 모델의 시드 데이터는 Amazon S3 버킷에서 직접 가져오거나 SQL 쿼리의 결과에서 가져올 수 있습니다.

직접 제공되는 시드 데이터는 다음 요구 사항을 충족해야 합니다.

- 시드 데이터는 사용자 IDs.
- 시드 크기는 25~500,000개의 고유 사용자 ID여야 합니다.
- 최소 시드 사용자 수는 구성된 대상 모델을 생성할 때 지정된 최소 일치 시드 크기 값과 일치해야 합니다.

다음은 CSV 형식의 유효한 훈련 데이터 세트의 예입니다.

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

AWS Clean Rooms ML 모델 평가 지표

Clean Rooms ML은 리콜 및 관련성 점수를 계산하여 모델의 성능을 결정합니다. 리콜은 유사 데이터와 훈련 데이터 간의 유사성을 비교합니다. 관련성 점수는 모델 성능이 좋은지 여부가 아니라 대상의 규모를 결정하는 데 사용됩니다.

리콜은 유사 세그먼트가 훈련 데이터와 얼마나 유사한지를 편향 없이 측정한 것입니다. 리콜은 대상 생성 작업을 통해 시드 대상에 포함된 훈련 데이터 샘플에서 가장 유사한 사용자(기본적으로 가장 유사한 20%)의 백분율입니다. 값 범위는 0~1이며, 값이 클수록 더 나은 대상을 나타냅니다. 최대 빈 백분율과 거의 동일한 리콜 값은 대상 모델이 무작위 선택과 동일함을 나타냅니다.

Clean Rooms ML은 모델을 구축할 때 실제 부정적 사용자를 정확하게 분류하지 않기 때문에 정확도, 정밀도, F1 점수보다 이 평가 지표가 더 나은 평가 지표로 간주됩니다.

세그먼트 수준 관련성 점수는 -1(가장 유사하지 않음)에서 1(가장 유사함) 사이의 값을 갖는 유사성 척도입니다. Clean Rooms ML은 다양한 세그먼트 크기에 대한 관련성 점수 집합을 계산하여 데이터에 가장 적합한 세그먼트 크기를 결정하는 데 도움을 줍니다. 관련성 점수는 세그먼트 크기가 증가함에 따라 단조롭게 감소하므로 세그먼트 크기가 증가함에 따라 시드 데이터와 덜 유사할 수 있습니다. 세그먼트 수준 관련성 점수가 0에 도달하면 모델은 유사 세그먼트의 모든 사용자가 시드 데이터와 동일한 분포에 속한다고 예측합니다. 출력 크기를 늘리면 유사 세그먼트에 시드 데이터와 동일한 분포에 속하지 않는 사용자가 포함될 가능성이 높습니다.

관련성 점수는 단일 캠페인 내에서 정규화되므로 여러 캠페인을 비교하는 데 사용해서는 안 됩니다. 관련성 점수는 인벤토리 품질, 인벤토리 유형, 광고 시기 등과 같은 관련성 외에도 여러 복잡한 요인의 영향을 받기 때문에 비즈니스 성과에 대한 단일 소스 증거로 사용해서는 안 됩니다.

관련성 점수는 시드의 품질을 판단하는 데 사용할 것이 아니라 높이거나 낮출 수 있는지를 판단하는 데 사용해야 합니다. 다음 예제를 살펴보세요.

- 전부 플러스인 점수 - 이는 유사 세그먼트에 포함된 것보다 유사한 것으로 예측된 출력 사용자가 더 많다는 것을 나타냅니다. 이는 지난 한 달 동안 치약을 구매한 모든 사람과 같이 규모가 큰 시장의 일부 시드 데이터에서 흔히 볼 수 있습니다. 지난 한 달 동안 치약을 두 번 이상 구매한 모든 사람과 같이 소규모 시드 데이터를 살펴보는 것이 좋습니다.
- 원하는 유사 세그먼트 크기에서 전부 마이너스 점수 - 이는 Clean Rooms ML이 원하는 유사 세그먼트 크기에서 유사한 사용자가 충분하지 않을 것으로 예측한다는 것을 나타냅니다. 이는 시드 데이터가 너무 구체적이거나 시장 규모가 너무 작기 때문일 수 있습니다. 시드 데이터에 적용할 필터 수를 줄이거나 시장을 확대하는 것이 좋습니다. 예를 들어 원래 시드 데이터가 유모차와 카시트를 구매한 고객이었다면 유아용품을 여러 개 구매한 고객으로 시장을 확대할 수 있습니다.

훈련 데이터 공급자는 관련성 점수의 노출 여부와 관련성 점수를 계산하는 버킷 빈을 결정합니다.

Clean Rooms ML의 사용자 지정 모델

Clean Rooms ML을 사용하면 공동 작업 구성원이 Amazon ECR에 저장된 도킹된 사용자 지정 모델 알고리즘을 사용하여 데이터를 공동 분석할 수 있습니다. 이렇게 하려면 모델 공급자가 이미지를 생성하고 Amazon ECR에 저장해야 합니다. Amazon Elastic Container Registry 사용 설명서의 단계에 따라 사용자 지정 ML 모델을 포함할 프라이빗 리포지토리를 생성합니다.

올바른 권한이 있는 경우 공동 작업의 모든 구성원이 모델 공급자가 될 수 있습니다. 공동 작업의 모든 구성원은 모델에 데이터를 제공할 수 있습니다. 이 안내서의 목적상 데이터를 제공하는 구성원을 데이터 공급자라고 합니다. 공동 작업을 생성하는 구성원은 공동 작업 생성자이며,이 구성원은 모델 공급자, 데이터 공급자 중 하나 또는 둘 다일 수 있습니다.

다음 주제에서는 사용자 지정 ML 모델을 생성하는 데 필요한 정보를 설명합니다.

주제

- 사용자 지정 ML 모델링 사전 조건
- 훈련 컨테이너에 대한 모델 작성 지침
- 추론 컨테이너에 대한 모델 작성 지침

- 모델 로그 및 지표 수신

사용자 지정 ML 모델링 사전 조건

사용자 지정 ML 모델링을 수행하려면 먼저 다음 사항을 고려해야 합니다.

- 훈련된 모델에 대한 모델 훈련과 추론을 공동 작업에서 수행할지 여부를 결정합니다.
- 각 공동 작업 구성원이 수행할 역할을 결정하고 적절한 능력을 할당합니다.
  - 모델을 훈련하고 훈련된 모델에서 추론을 실행할 구성원에게 CAN_QUERY 기능을 할당합니다.
  - 공동 작업의 구성원 한 명 이상CAN_RECEIVE_RESULTS에게를 할당합니다.
  - 훈련된 모델 내보내기 CAN_RECEIVE_MODEL_OUTPUT 또는 추론 출력을 각각 받을 구성원에게 또는 CAN_RECEIVE_INFERENCE_OUTPUT 기능을 할당합니다. 사용 사례에 필요한 경우 두 기능을 모두 사용하도록 선택할 수 있습니다.
- 내보내도록 허용할 훈련된 모델 아티팩트 또는 추론 결과의 최대 크기를 결정합니다.
- 모든 사용자에게 역할에 연결된 CleanrooomsFullAccess 및 CleanroomsMLFullAccess 정책이 있는 것이 좋습니다. 사용자 지정 ML 모델을 사용하려면 AWS Clean Rooms 및 AWS Clean Rooms ML SDKs.
- IAM 역할에 대한 다음 정보를 고려합니다.
  - 모든 데이터 공급자는가 AWS Glue 카탈로그 및 테이블과 기본 Amazon S3 위치에서 데이터를 AWS Clean Rooms 읽을 수 있도록 허용하는 서비스 액세스 역할이 있어야 합니다. 이러한 역할은 SQL 쿼리에 필요한 역할과 유사합니다. 이렇게 하면 CreateConfiguredTableAssociation 작업을 사용할 수 있습니다. 자세한 내용은 서비스 역할을 생성하여 구성된 테이블 연결 생성 단원을 참조하십시오.
  - 지표를 수신하려는 모든 멤버는 CloudWatch 지표 및 로그를 작성할 수 있는 서비스 액세스 역할이 있어야 합니다. 이 역할은 Clean Rooms ML에서 모델 훈련 및 추론 AWS 계정 중에 모든 모델 지표와 로그를 멤버의에 기록하는 데 사용됩니다. 또한 지표 및 로그에 액세스할 수 있는 구성원을 결정하기 위한 개인 정보 보호 제어 기능도 제공합니다. 이렇게 하면 CreateMLConfiguration 작업을 사용할 수 있습니다. 자세한 내용은 단원을 참조하십시오사용자 지정 ML 모델링을 위한 서비스 역할 생성 - ML 구성.

    결과를 받는 구성원은 서비스 액세스 역할에 Amazon S3 버킷에 쓸 수 있는 권한을 제공해야 합니다. 이 역할을 통해 Clean Rooms ML은 결과(학습된 모델 아티팩트 또는 추론 결과)를 Amazon S3 버킷으로 내보낼 수 있습니다. 이렇게 하면 CreateMLConfiguration 작업을 사용할 수 있습니다. 자세한 내용은 사용자 지정 ML 모델링을 위한 서비스 역할 생성 - ML 구성 단원을 참조하십시오.

- 모델 공급자는 서비스 액세스 역할에 Amazon ECR 리포지토리 및 이미지를 읽을 수 있는 권한을 제공해야 합니다. 이렇게 하면 `CreateConfigureModelAlgorithm` 작업을 사용할 수 있습니다. 자세한 내용은 서비스 역할을 생성하여 사용자 지정 ML 모델 제공 단원을 참조하십시오.
- 훈련 또는 추론`MLInputChannel`용 데이터 세트를 생성하기 위해를 생성하는 구성원은 Clean Rooms ML이 SQL 쿼리를 실행할 수 있도록 허용하는 서비스 액세스 역할을 제공해야 합니다 AWS Clean Rooms. 이렇게 하면 `CreateTrainedModel` 및 `StartTrainedModelInferenceJob` 작업을 사용할 수 있습니다. 자세한 내용은 데이터 세트를 쿼리할 서비스 역할 생성 단원을 참조하십시오.

- 모델 작성자는 훈련 컨테이너에 대한 모델 작성 지침 모델 입력 및 출력이 예상대로 구성되도록 추론 컨테이너에 대한 모델 작성 지침 및를 따라야 합니다 AWS Clean Rooms.

훈련 컨테이너에 대한 모델 작성 지침

이 섹션에서는 Clean Rooms ML에 대한 사용자 지정 ML 모델 알고리즘을 생성할 때 모델 공급자가 따라야 하는 지침을 자세히 설명합니다.

- SageMaker AI 개발자 안내서에 설명된 대로 적절한 SageMaker AI 훈련 지원 컨테이너 기본 이미지를 사용합니다. 다음 코드를 사용하면 퍼블릭 SageMaker AI 엔드포인트에서 지원되는 컨테이너 기본 이미지를 가져올 수 있습니다.

```
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
```

- 모델을 로컬에서 작성할 때 개발 인스턴스,의 SageMaker AI 훈련 및 AWS 계정 Clean Rooms ML에서 로컬에서 모델을 테스트할 수 있도록 다음을 확인하세요.
  - 다양한 환경 변수를 통해 훈련 환경에 대한 유용한 속성에 액세스하는 훈련 스크립트를 작성하는 것이 좋습니다. Clean Rooms ML은 `SM_MODEL_DIR`, `SM_OUTPUT_DIR`, `SM_CHANNEL_TRAIN`및 인수를 사용하여 모델 코드에 대한 훈련을 호출합니다`FILE_FORMAT`. 이러한 기본값은 Clean Rooms ML에서 모든 당사자의 데이터를 사용하여 자체 실행 환경에서 ML 모델을 훈련하는 데 사용됩니다.
  - Clean Rooms ML을 사용하면 Docker 컨테이너의 `/opt/ml/input/data/channel-name` 디렉터리를 통해 훈련 입력 채널을 사용할 수 있습니다. 각 ML 입력 채널은 `CreateTrainedModel` 요청에 `channel_name` 제공된 해당를 기반으로 매핑됩니다.

```
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"))
```

- 모델 코드에 사용할 공동 작업자의 스키마를 기반으로 합성 또는 테스트 데이터 세트를 생성할 수 있는지 확인합니다.
- 모델 알고리즘 AWS Clean Rooms 을 공동 작업에 연결하기 AWS 계정 전에 SageMaker AI 훈련 작업을 직접 실행할 수 있는지 확인합니다.

  다음 코드에는 로컬 테스트, SageMaker AI 훈련 환경 테스트 및 Clean Rooms ML과 호환되는 샘플 Docker 파일이 포함되어 있습니다.

```
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"]
```

- 컨테이너 장애를 가장 잘 모니터링하려면 오류로 인해 로그를 내보내고 디버깅하는 것이 좋습니다. 이에 GetTrainedModel 대한 응답으로 Clean Rooms ML은 아래의이 파일에서 처음 1024자를 반환합니다StatusDetails.
- 모델 변경을 완료하고 SageMaker AI 환경에서 테스트할 준비가 되면 제공된 순서대로 다음 명령을 실행합니다.

```
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
```

SageMaker AI 작업이 완료되고 모델 알고리즘에 만족하면 AWS Clean Rooms ML에 Amazon ECR 레지스트리를 등록할 수 있습니다. CreateConfiguredModelAlgorithm 작업을 사용하여 모델 알고리즘을 등록하고 CreateConfiguredModelAlgorithmAssociation를 사용하여 이를 공동 작업에 연결합니다.

추론 컨테이너에 대한 모델 작성 지침

이 섹션에서는 Clean Rooms ML에 대한 추론 알고리즘을 생성할 때 모델 공급자가 따라야 하는 지침을 자세히 설명합니다.

- SageMaker AI 개발자 안내서에 설명된 대로 적절한 SageMaker AI 추론 지원 컨테이너 기본 이미지를 사용합니다. 다음 코드를 사용하면 퍼블릭 SageMaker AI 엔드포인트에서 지원되는 컨테이너 기본 이미지를 가져올 수 있습니다.

```
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
```

- 모델을 로컬에서 작성할 때 개발 인스턴스,의 SageMaker AI 배치 변환 및 AWS 계정 Clean Rooms ML에서 로컬에서 모델을 테스트할 수 있도록 다음을 확인하세요.
  - Clean Rooms ML을 사용하면 추론의 모델 아티팩트를 Docker 컨테이너의 `/opt/ml/model` 디렉터리를 통해 추론 코드에서 사용할 수 있습니다.
  - Clean Rooms ML은 입력을 줄별로 분할하고 배치 `MultiRecord` 전략을 사용하며 변환된 모든 레코드의 끝에 줄 바꿈 문자를 추가합니다.
  - 모델 코드에 사용할 공동 작업자의 스키마를 기반으로 합성 또는 테스트 추론 데이터 세트를 생성할 수 있는지 확인합니다.
  - 모델 알고리즘 AWS Clean Rooms 을 공동 작업에 연결하기 AWS 계정 전에 SageMaker AI 배치 변환 작업을 직접 실행할 수 있는지 확인합니다.

    다음 코드에는 로컬 테스트, SageMaker AI 변환 환경 테스트 및 Clean Rooms ML과 호환되는 샘플 Docker 파일이 포함되어 있습니다.

```
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"]
```

- 모델 변경을 완료하고 SageMaker AI 환경에서 테스트할 준비가 되면 제공된 순서대로 다음 명령을 실행합니다.

```
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
```

SageMaker AI 작업이 완료되고 배치 변환에 만족하면 AWS Clean Rooms ML에 Amazon ECR 레지스트리를 등록할 수 있습니다. `CreateConfiguredModelAlgorithm` 작업을 사용하여 모델 알고

리즘을 등록하고 CreateConfiguredModelAlgorithmAssociation를 사용하여 이를 공동 작업에 연결합니다.

모델 로그 및 지표 수신

사용자 지정 모델 훈련 또는 추론에서 로그 및 지표를 수신하려면 멤버가 필요한 CloudWatch 권한을 제공하는 유효한 역할로 ML 구성을 생성해야 합니다(사용자 지정 ML 모델링을 위한 서비스 역할 생성 - ML 구성 참조).

시스템 지표

CPU 및 메모리 사용률과 같은 훈련 및 추론 모두에 대한 시스템 지표는 유효한 ML 구성을 사용하여 공동 작업의 모든 구성원에게 게시됩니다. 이러한 지표는 각각 /aws/cleanroomsml/TrainedModels 또는 /aws/cleanroomsml/TrainedModelInferenceJobs 네임스페이스의 CloudWatch 지표를 통해 작업이 진행됨에 따라 볼 수 있습니다.

모델 로그

모델 로그에 대한 액세스는 구성된 각 모델 알고리즘의 개인 정보 보호 구성 정책에 의해 제공됩니다. 모델 작성자는 구성된 모델 알고리즘(콘솔 또는 CreateConfiguredModelAlgorithmAssociation API를 통해)을 공동 작업에 연결할 때 개인 정보 보호 구성 정책을 설정합니다. 개인 정보 보호 구성 정책을 설정하면 모델 로그를 수신할 수 있는 멤버가 제어됩니다.

또한 모델 작성자는 개인 정보 보호 구성 정책에서 필터 패턴을 설정하여 로그 이벤트를 필터링할 수 있습니다. 모델 컨테이너가 stdout 또는 로 보내고 필터 패턴(설정된 경우)stderr과 일치하는 모든 로그는 Amazon CloudWatch Logs로 전송됩니다. 모델 로그는 /aws/cleanroomsml/TrainedModelInferenceJobs각각 CloudWatch 로그 그룹 /aws/cleanroomsml/TrainedModels 또는에서 사용할 수 있습니다.

사용자 지정 정의 지표

모델 알고리즘을 구성할 때(콘솔 또는 CreateConfiguredModelAlgorithm API를 통해) 모델 작성자는 출력 로그에서 검색할 특정 지표 이름과 정규식 문을 제공할 수 있습니다. 이는 /aws/cleanroomsml/TrainedModels 네임스페이스의 CloudWatch 지표를 통해 작업이 진행됨에 따라 볼 수 있습니다. 구성된 모델 알고리즘을 연결할 때 모델 작성자는 지표 프라이버시 구성에서 선택적 노이즈 수준을 설정하여 사용자 지정 지표 추세에 대한 가시성을 제공하면서 원시 데이터를 출력하지 않도록 할 수 있습니다. 노이즈 수준이 설정된 경우 지표는 실시간으로가 아닌 작업 종료 시 게시됩니다.

Clean Rooms에 대한 암호화 컴퓨팅

Clean Rooms (C3R)에 대한 암호화 컴퓨팅은 분석 규칙 외에도 사용할 수 AWS Clean Rooms 있는의 기능입니다. C3R을 사용하면 조직은 민감한 데이터를 한데 모아 데이터 분석을 통해 새로운 통찰력을 도출하는 동시에 프로세스에서 모든 당사자가 학습할 수 있는 내용을 암호학적으로 제한할 수 있습니다. 민감한 데이터로 공동 작업하고 싶지만 클라우드에서는 암호화된 데이터만 사용해야 하는 두 명 이상의 당사자가 C3R을 사용할 수 있습니다.

C3R 암호화 클라이언트는 사용할 데이터를 암호화하는 데 사용할 수 있는 클라이언트 측 암호화 도구입니다 AWS Clean Rooms. C3R 암호화 클라이언트를 사용하는 경우 데이터는 AWS Clean Rooms 공동 작업에 사용되는 동안 암호화 방식으로 보호됩니다. 정기적인 AWS Clean Rooms 공동 작업과 마찬가지로 입력 데이터는 관계형 데이터베이스 테이블이며 계산은 SQL 쿼리로 표현됩니다. 하지만 C3R은 암호화된 데이터에 대한 제한된 SQL 쿼리 하위 집합만 지원합니다.

특히 C3R은 암호로 보호된 데이터에 대한 SQL JOIN 및 SELECT 명령문을 지원합니다. 입력 테이블의 각 열은 다음 SQL 문 유형 중 정확히 하나로 사용할 수 있습니다.

- JOIN 명령문에 사용할 수 있도록 암호로 보호되는 열을 fingerprint 열이라고 합니다.
- SELECT명령문에 사용할 수 있도록 암호로 보호되는 열을 sealed 열이라고 합니다.
- JOIN 또는 SELECT 문에서 사용할 수 있도록 암호로 보호되지 않은 열을 cleartext 열이라고 합니다.

경우에 따라 fingerprint 열에서 GROUP BY 명령문이 지원됩니다. 자세한 내용은 Fingerprint 열 단원을 참조하십시오. 현재 C3R은 관련 분석 규칙에서 허용하는 경우라도 암호화된 데이터에 WHERE 절이나 SUM, AVERAGE와(과) 같은 집계 함수 같은 다른 SQL 구성을 사용하는 것을 지원하지 않습니다.

C3R은 테이블의 개별 셀에 있는 데이터를 보호하도록 설계되었습니다. C3R의 기본 구성을 사용하면 고객이 공동 작업을 통해 제3자에게 제공하는 기본 데이터는 콘텐츠를 AWS Clean Rooms내에서 사용하는 동안 암호화된 상태로 유지됩니다. C3R은 모든 sealed 열에 업계 표준 AES-GCM 암호화를 사용하고 fingerprint 열 보호를 위해 해시 기반 메시지 인증 코드(HMAC)로 알려진 업계 표준 의사 랜덤 함수를 사용합니다.

C3R은 테이블의 데이터를 암호화하지만 다음 정보는 여전히 유추할 수 있습니다.

- 테이블의 열 수, 열 이름, 행 수 등 테이블 자체에 대한 정보.
- 대부분의 표준 암호화 형식과 마찬가지로 C3R은 암호화된 값의 길이를 숨기려고 하지 않습니다. C3R은 암호화된 값을 패딩하여 일반 텍스트의 정확한 길이를 숨길 수 있는 기능을 제공합니다. 그러나 각 열의 일반 텍스트 길이의 상한선이 여전히 다른 당사자에게 공개될 수 있습니다.

- 로깅 수준 정보(예: 암호화된 C3R 테이블에 특정 행이 추가된 시점).

C3R에 대한 자세한 내용은 다음 주제를 참조하세요.

주제

- Clean Rooms에 대한 암호화 컴퓨팅 사용 시 고려 사항
- Clean Rooms용 암호화 컴퓨팅에서 지원되는 파일 및 데이터 유형
- Clean Rooms에 대한 암호화 컴퓨팅의 열 이름
- Clean Rooms용 암호화 컴퓨팅의 열 유형
- 암호화 컴퓨팅 파라미터
- Clean Rooms용 암호화 컴퓨팅의 선택적 플래그
- Clean Rooms에 대한 암호화 컴퓨팅을 사용한 쿼리
- C3R 암호화 클라이언트에 대한 지침

Clean Rooms에 대한 암호화 컴퓨팅 사용 시 고려 사항

Clean Rooms에 대한 암호화 컴퓨팅(C3R)은 데이터 보호를 극대화하기 위한 것입니다. 그러나 일부 사용 사례에서는 추가 기능을 제공하는 대신 낮은 수준의 데이터 보호를 통해 혜택을 받을 수 있습니다. 가장 안전한 구성에서 C3R을 수정하여 이러한 특정 절충안을 만들 수 있습니다. 고객은 이러한 장단점을 인지하고 사용 사례에 적합한지 판단해야 합니다. 고려해야 할 중요한 요소는 다음과 같습니다.

주제

- 테이블에 혼합 cleartext 및 암호화된 데이터 허용
- fingerprint 열에 반복되는 값 허용
- fingerprint 열 이름 지정 방법에 대한 제한 완화
- NULL 값 표현 방식 결정

이들 파라미터를 설정하는 방법에 대한 자세한 내용은 암호화 컴퓨팅 파라미터 섹션을 참조하세요.

테이블에 혼합 cleartext 및 암호화된 데이터 허용

모든 데이터를 클라이언트 측에서 암호화하면 데이터 보호를 극대화할 수 있습니다. 하지만 이렇게 하면 특정 종류의 쿼리(예: SUM 집계 함수)가 제한됩니다. cleartext 데이터를 허용할 경우 암호화된 테이블에 액세스할 수 있는 모든 사용자가 암호화된 값에 대한 일부 정보를 유추할 수 있다는 위험이 있습니다. 이는 cleartext 및 관련 데이터에 대한 통계 분석을 수행하여 수행할 수 있습니다.

예를 들어, City 및 State의 열이 있다고 가정해 보겠습니다. City 열은 cleartext이고 State 열은 암호화되어 있습니다. City 열에 Chicago라는 값이 표시되면 State가 Illinois일 확률이 높다는 것을 알 수 있습니다. 반대로, 한 열이 City이고 다른 열이 EmailAddress인 경우 cleartext City는 암호화된 EmailAddress에 대해 아무 것도 드러내지 않을 것입니다.

이 창의 파라미터에 대한 자세한 내용은 cleartext 열 허용 매개 변수 섹션을 참조하세요.

fingerprint 열에 반복되는 값 허용

가장 안전한 접근 방식을 위해 모든 fingerprint 열에 정확히 하나의 변수 인스턴스가 포함되어 있다고 가정합니다. 하나의 fingerprint 열에서 어떤 항목도 반복할 수 없습니다. C3R 암호화 클라이언트는 이러한 cleartext 값을 임의 값과 구별할 수 없는 고유한 값으로 매핑합니다. 따라서 이러한 임의 값으로는 cleartext에 대한 정보를 유추할 수 없습니다.

하나의 fingerprint 열에 값이 반복되면 값이 반복되어 무작위로 보이는 값이 반복될 위험이 있습니다. 따라서 암호화된 테이블에 액세스할 수 있는 사람은 이론적으로 fingerprint 열에 대한 통계 분석을 수행하여 cleartext 값에 대한 정보를 확인할 수 있습니다.

다시 말하지만, fingerprint 열이 State이고 테이블의 모든 행이 미국 가정에 해당한다고 가정해 보겠습니다. 빈도 분석을 수행하면 어떤 상태가 California이고 어떤 상태가 Wyoming인지 높은 확률로 추론할 수 있습니다. 이러한 추론이 가능한 이유는 California의 거주자 수가 Wyoming보다 많기 때문입니다. 반대로, fingerprint 열이 세대 식별자에 관한 것이고 수백만 개의 항목이 있는 데이터베이스에서 각 가구가 데이터베이스에 1~4번 나타났다고 가정해 보겠습니다. 빈도 분석을 통해 유용한 정보가 나올 가능성은 거의 없습니다.

이 시나리오의 파라미터에 대한 자세한 내용은 중복 매개 변수 허용 섹션을 참조하세요.

fingerprint 열 이름 지정 방법에 대한 제한 완화

기본적으로 암호화된 fingerprint 열을 사용하여 두 테이블을 조인하면 각 테이블에서 해당 열의 이름이 같다고 가정합니다. 이 결과가 나오는 기술적인 이유는 기본적으로 각 fingerprint 열을 암호화하기 위해 서로 다른 암호화 키를 생성하기 때문입니다. 이 키는 공동 작업을 위한 공유 비밀 키와 열 이름의 조합에서 파생됩니다. 열 이름이 다른 두 열을 조인하려고 하면 다른 키가 파생되고 유효한 조인을 계산할 수 없습니다.

이 문제를 해결하려면 각 열 이름에서 키를 추출하는 기능을 끌 수 있습니다. 그러면 C3R 암호화 클라이언트는 모든 fingerprint 열에 단일 파생 키를 사용합니다. 위험은 정보를 드러낼 수 있는 또 다른 종류의 주파수 분석을 수행할 수 있다는 것입니다.

City와 State 예제를 다시 사용해봅시다. 열 이름을 포함하지 않고 각 fingerprint 열에 대해 동일한 임의 값을 도출하는 경우 New York은 City및 State 열에 동일한 임의 값을 갖습니다. 뉴욕은 미국에

서 City 이름과 같은 State 이름을 가진 몇 안 되는 도시 중 하나입니다. 반대로 데이터 세트의 각 열에 있는 값이 완전히 다른 경우에는 정보가 유출되지 않습니다.

이 시나리오의 파라미터에 대한 자세한 내용은 이름이 다른 열의 JOIN 허용 매개변수 섹션을 참조하세요.

NULL 값 표현 방식 결정

사용할 수 있는 옵션은 다른 값과 마찬가지로 NULL 값을 암호화(암호화 및 HMAC)로 처리할지 여부입니다. 다른 값처럼 NULL 값을 처리하지 않으면 정보가 노출될 수 있습니다.

예를 들어, cleartext의 Middle Name 열에 있는 NULL이 중간 이름이 없는 사람을 나타낸다고 가정해 보겠습니다. 이러한 값을 암호화하지 않으면 암호화된 테이블에서 중간 이름이 없는 사람들이 사용한 행이 유출됩니다. 이 정보는 일부 집단의 일부 사람들을 식별하는 신호가 될 수 있습니다. 그러나 NULL 값을 암호적으로 처리하는 경우 특정 SQL 쿼리는 다르게 작동합니다. 예를 들어 GROUP BY 절은 fingerprint 열의 fingerprint NULL 값을 함께 그룹화하지 않습니다.

이 시나리오의 파라미터에 대한 자세한 내용은 NULL 값 보존(매개변수) 섹션을 참조하세요.

Clean Rooms용 암호화 컴퓨팅에서 지원되는 파일 및 데이터 유형

C3R 암호화 클라이언트는 다음 파일 유형을 인식합니다.

- CSV 파일
- Parquet files

C3R 암호화 클라이언트의 --fileFormat 플래그를 사용하여 파일 형식을 명시적으로 지정할 수 있습니다. 명시적으로 지정한 경우 파일 형식은 파일 확장자에 의해 결정되지 않습니다.

주제

- CSV 파일
- Parquet files
- 문자열이 아닌 값 암호화

CSV 파일

확장명이.csv인 파일은 CSV 형식이고 UTF-8 인코딩 텍스트를 포함하는 것으로 간주됩니다. C3R 암호화 클라이언트는 모든 값을 문자열로 취급합니다.

.csv 파일에서 지원되는 속성

C3R 암호화 클라이언트를 사용하려면 .csv 파일에 다음과 같은 속성이 있어야 합니다.

- 각 열의 이름을 고유하게 지정하는 초기 헤더 행을 포함할 수도 있고 포함하지 않을 수도 있습니다.
- 쉼표로 구분. (현재 사용자 지정 구분 기호는 지원되지 않습니다.)
- UTF-8 인코딩 텍스트.

.csv 항목에서 스페이스 제거

선행 공백과 후행 스페이스 모두.csv 항목에서 제거됩니다.

.csv 파일의 사용자 지정 NULL 인코딩

.csv 파일은 사용자 지정 NULL 인코딩을 사용할 수 있습니다.

C3R 암호화 클라이언트에서는 `--csvInputNULLValue=<csv-input-null>` 플래그를 사용하여 입력 데이터의 NULL 항목에 대한 사용자 지정 인코딩을 지정할 수 있습니다. C3R 암호화 클라이언트는 `--csvOutputNULLValue=<csv-output-null>` 플래그를 사용하여 NULL 항목에 대해 생성된 출력 파일에서 사용자 지정 인코딩을 사용할 수 있습니다.

> **Note**
>
> NULL 항목은 내용이 부족한 것으로 간주됩니다. 특히 SQL 테이블과 같은 풍부한 테이블 형식의 컨텍스트에서는 더욱 그렇습니다. .csv는 역사적 이유로 이 특성화를 명시적으로 지원하지 않지만, 스페이스만 포함된 빈 항목은 NULL로 간주하는 것이 일반적인 관례입니다 따라서 이는 C3R 암호화 클라이언트의 기본 동작이며 필요에 따라 사용자 지정할 수 있습니다.

C3R에서.csv 항목을 해석하는 방법

다음 표는 `--csvInputNULLValue=<csv-input-null>` 및 `--csvOutputNULLValue=<csv-output-null>` 플래그에 제공된 값(있는 경우)에 따라 .csv 항목이 어떻게 마샬링되는지(명확성을 위해 cleartext에서 cleartext로 보여주는 예입니다. 따옴표 밖의 선행 및 후행 스페이스는 C3R이 값의 의미를 해석하기 전에 잘립니다.

<csv-input-null>	**<csv-output-null>**	항목	출력 항목
None	None	, ## ##,	, ## ##,
None	None	, ## ##,	, ## ##,
None	None	,"## ##",	, ## ##,
None	None	, "## ##" ,	, ## ##,
None	None	,,	,,
None	None	, ,	,,
None	None	,"",	,,
None	None	," ",	," ",
None	None	, " " ,	," ",
"## ##"	"NULL"	, ## ##,	,NULL,
"## ##"	"NULL"	, ## ##,	,NULL,
"## ##"	"NULL"	,"## ##",	,NULL,
"## ##"	"NULL"	, "## ##" ,	,NULL,
None	"NULL"	,,	,NULL,
None	"NULL"	, ,	,NULL,
None	"NULL"	,"",	,NULL,
None	"NULL"	," ",	," ",
None	"NULL"	, " " ,	," ",
""	"NULL"	,,	,NULL,
""	"NULL"	, ,	,NULL,

<csv-input-null>	<csv-output-null>	항목	출력 항목
""	"NULL"	,"",	,"",
""	"NULL"	," ",	," ",
""	"NULL"	, " " ,	," ",
"\"\""	"NULL"	,,	,,
"\"\""	"NULL"	, ,	,,
"\"\""	"NULL"	,"",	,NULL,
"\"\""	"NULL"	," ",	," ",
"\"\""	"NULL"	, " " ,	," ",

헤더가 없는 CSV 파일

소스.csv 파일의 첫 번째 행에는 각 열의 이름을 고유하게 지정하는 헤더가 없어도 됩니다. 하지만 헤더 행이 없는.csv 파일에는 위치 암호화 스키마가 필요합니다. 헤더 행이 있는.csv 파일과 Parquet 파일 모두에 사용되는 일반적인 매핑 스키마 대신 위치 암호화 스키마가 필요합니다.

위치 암호화 스키마는 이름 대신 위치를 기준으로 출력 열을 지정합니다. 매핑된 암호화 스키마는 원본 열 이름을 대상 열 이름에 매핑합니다. 두 스키마 형식에 대한 자세한 설명과 예를 포함한 자세한 내용은 매핑된 테이블 스키마와 위치 테이블 스키마을 참조하세요.

Parquet files

.parquet 확장자가 있는 파일은 Apache Parquet와 같은 형식으로 간주됩니다.

지원되는 Parquet 데이터 형식

C3R 암호화 클라이언트는 AWS Clean Rooms에서 지원하는 데이터 유형을 나타내는 Parquet 파일에서 복잡하지 않은 데이터(즉, 기본 유형)를 모두 처리할 수 있습니다.

하지만 sealed 열에는 문자열 열만 사용할 수 있습니다.

지원되는 Parquet 데이터 유형은 다음과 같습니다:

- 다음과 같은 논리적 주석이 포함된 Binary 기본 유형:
  - --parquetBinaryAsString이 설정된 경우, 없음(STRING데이터 유형)
  - Decimal(scale, precision)(DECIMAL 데이터 유형)
  - String(STRING 데이터 유형)
- 논리적 주석이 없는 Boolean 기본 데이터 유형(BOOLEAN 데이터 유형)
- 논리적 주석이 없는 Double 기본 데이터 유형(DOUBLE 데이터 유형)
- Decimal(scale, precision) 논리적 주석이 있는 Fixed_Len_Binary_Array 기본 유형(DECIMAL 데이터 유형)
- 논리적 주석이 없는 Float 기본 데이터 유형(FLOAT 데이터 유형)
- 다음과 같은 논리적 주석이 포함된 Int32 기본 유형:
  - 없음(INT 데이터 유형)
  - Date(DATE 데이터 유형)
  - Decimal(scale, precision)(DECIMAL 데이터 유형)
  - Int(16, true)(SMALLINT 데이터 유형)
  - Int(32, true)(INT 데이터 유형)
- 다음과 같은 논리적 주석이 포함된 Int64 기본 데이터 유형:
  - 없음(BIGINT 데이터 유형)
  - Decimal(scale, precision)(DECIMAL 데이터 유형)
  - Int(64, true)(BIGINT 데이터 유형)
  - Timestamp(isUTCAdjusted, TimeUnit.MILLIS)(TIMESTAMP 데이터 유형)
  - Timestamp(isUTCAdjusted, TimeUnit.MICROS)(TIMESTAMP 데이터 유형)
  - Timestamp(isUTCAdjusted, TimeUnit.NANOS)(TIMESTAMP 데이터 유형)

문자열이 아닌 값 암호화

현재 sealed 열에는 문자열 값만 지원됩니다.

.csv 파일의 경우 C3R 암호화 클라이언트는 모든 값을 UTF-8 인코딩 텍스트로 취급하며 암호화 전에 이러한 값을 다르게 해석하려고 시도하지 않습니다.

핑거프린트 열의 경우 유형은 동등한 클래스로 그룹화됩니다. 등가 클래스는 대표적인 데이터 유형을 통해 동등성을 명확하게 비교할 수 있는 데이터 유형 집합입니다.

등가 클래스를 사용하면 원래 표현과 관계없이 동일한 지문을 동일한 시맨틱 값에 할당할 수 있습니다. 하지만 두 등가 클래스의 값이 같더라도 핑거프린트 열이 같아지지는 않습니다.

예를 들어, INTEGRAL 값 42는 원래 SMALLINT, INT, 또는 BIGINT였는지 여부에 관계없이 동일한 지문을 할당받게 됩니다. 또한 INTEGRAL 값 0은 BOOLEAN 값 FALSE (값 0으로 표시됨)와 절대 일치하지 않습니다.

지문 열에서 지원되는 등가 클래스 및 해당 AWS Clean Rooms 데이터 형식은 다음과 같습니다.

등가 클래스	지원되는 AWS Clean Rooms 데이터 형식
BOOLEAN	BOOLEAN
DATE	DATE
INTEGRAL	BIGINT, INT, SMALLINT
STRING	CHAR, STRING, VARCHAR

Clean Rooms에 대한 암호화 컴퓨팅의 열 이름

Clean Rooms에 대한 암호화 컴퓨팅의 경우 기본적으로 열 이름이 중요합니다.

이름이 다른 열 JOIN 허용 매개 변수의 값이 거짓인 경우 열 암호화 시 fingerprint 열 이름이 사용됩니다. 따라서 기본적으로 협업자는 사전에 조율하여 쿼리에서 JOIN 문을 사용할 데이터에 대해 동일한 대상 열 이름을 사용해야 합니다. 기본적으로 열을 다른 JOIN 이름으로 암호화하면 어떤 값에서도 제대로 JOIN하지 않습니다.

이름이 다른 열의 JOIN 허용 매개 변수의 값이 참인 경우 fingerprint 열로 암호화된 열 전체에서 JOIN 명령문이 성공합니다. 이 매개 변수로 데이터를 암호화하면 cleartext 값을 어느 정도 추론할 수 있습니다. 예를 들어 행의 City 열과 State 열 모두에 동일한 해시 기반 메시지 인증 코드(HMAC)값이 있는 경우 값은 New York과 같을 수 있습니다.

열 헤더 이름의 정규화

열 헤더 이름은 C3R 암호화 클라이언트에 의해 정규화됩니다. 변환된 출력에서는 선행 및 후행 스페이스가 모두 제거되고 열 이름은 소문자로 바뀝니다.

정규화는 열 이름의 영향을 받을 수 있는 다른 모든 계산, 연산 또는 기타 작업보다 먼저 적용됩니다. 내보낸 출력 파일에는 정규화된 이름만 포함됩니다.

Clean Rooms용 암호화 컴퓨팅의 열 유형

이 항목에서는 Clean Rooms용 암호화 컴퓨팅의 열 유형에 대한 정보를 제공합니다.

주제

- Fingerprint 열
- 밀폐형 열
- Cleartext 열

Fingerprint 열

Fingerprint 열은 JOIN 명령문에 사용할 수 있도록 암호로 보호되는 열입니다.

fingerprint 열의 데이터는 해독할 수 없습니다. 봉인된 열의 데이터만 해독할 수 있습니다.

Fingerprint 열은 다음 SQL 절 및 함수에서만 사용해야 합니다.

- JOIN (INNER, OUTER, LEFT, RIGHT, or FULL)을 다른 fingerprint 열과 대조합니다.
  - `allowJoinsOnColumnsWithDifferentNames` 매개 변수 값이 `false`로 설정된 경우 JOIN의 두 fingerprint 열 모두 이름이 같아야 합니다.
- `SELECT COUNT()`
- `SELECT COUNT(DISTINCT )`
- `GROUP BY`(공동 작업에서 `preserveNulls` 파라미터 값을 `true`(으)로 설정한 경우에만 사용)

이러한 제약조건을 위반하는 쿼리는 잘못된 결과를 초래할 수 있습니다.

밀폐형 열

봉인된 열은 SELECT 명령문에 사용할 수 있도록 암호로 보호되는 열입니다.

봉인된 열은 다음 SQL 절 및 함수에서만 사용해야 합니다.

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

> **Note**
>
> `SELECT COUNT(DISTINCT )`는 지원되지 않습니다.

이러한 제약 조건을 위반하는 쿼리는 잘못된 결과를 초래할 수 있습니다.

암호화 전에 sealed 열의 데이터 패딩

하나의 열을 sealed 열로 지정하면 C3R은 어떤 종류의 패딩을 선택할지 묻습니다. 암호화 전에 데이터 패딩은 선택 사항입니다. 패딩이 없는 경우(패드 유형none), 암호화된 데이터의 길이는 cleartext의 크기를 나타냅니다. 경우에 따라 cleartext의 크기로 인해 일반 텍스트가 노출될 수 있습니다. 패딩(패드 유형이 `fixed` 또는`max`인 경우)을 사용하면 모든 값이 먼저 공통 크기로 채워진 다음 암호화됩니다. 패딩을 사용하면 암호화된 데이터의 길이에 따라 데이터 크기에 상한이 주어지는 것 외에는 원래 cleartext 길이에 대한 정보가 제공되지 않습니다.

열에 패딩을 적용하고 해당 열에 있는 데이터의 최대 바이트 길이를 알고 있는 경우 `fixed` 패딩을 사용합니다. 적어도 `length` 열에 있는 가장 긴 값의 바이트 길이만큼 큰 값을 사용합니다.

> **Note**
>
> 값이 제공된 `length`보다 길면 오류가 발생하고 암호화에 실패합니다.

열에 패딩을 적용하고 해당 열에 있는 데이터의 최대 바이트 길이를 알 수 없는 경우 `max` 패딩을 사용합니다. 이 패딩 모드는 모든 데이터를 가장 긴 값에 추가 `length` 바이트를 더한 길이로 채웁니다.

> **Note**
>
> 데이터를 일괄적으로 암호화하거나 정기적으로 새 데이터로 테이블을 업데이트하는 것이 좋습니다. `max`패딩은 주어진 배치에서 가장 긴 일반 텍스트 항목의 길이 (`length`바이트를 더한 값) 로 항목을 채우게 된다는 점에 유의하세요. 즉, 사이퍼텍스트 길이는 배치마다 다를 수 있습니다. 따라서 열의 최대 바이트 길이를 알고 있는 경우에는 `max` 대신 `fixed`를 사용해야 합니다.

Cleartext 열

Cleartext 열은 JOIN 또는 SELECT 문에서 사용할 때 암호로 보호되지 않는 열입니다.

Cleartext 열은 SQL 쿼리의 어느 부분에서나 사용할 수 있습니다.

암호화 컴퓨팅 파라미터

공동 작업을 생성할 때 Clean Rooms(C3R)에 대한 암호화 컴퓨팅을 사용하는 공동 작업에 암호화 컴퓨팅 파라미터를 사용할 수 있습니다. AWS Clean Rooms 콘솔 또는 `CreateCollaboration` API 작업을 사용하여 공동 작업을 생성할 수 있습니다. 콘솔에서 암호화 컴퓨팅 지원 옵션을 켠 후 암호화 컴퓨팅 매개 변수의 매개 변수 값을 설정할 수 있습니다. 자세한 내용은 다음 항목을 참조하세요.

주제

- cleartext 열 허용 매개 변수
- 중복 매개 변수 허용
- 이름이 다른 열의 JOIN 허용 매개변수
- NULL 값 보존(매개변수)

cleartext 열 허용 매개 변수

콘솔에서 공동 작업을 생성할 때 cleartext 열 허용 매개변수를 설정하여 암호화된 cleartext 데이터가 포함된 테이블에 데이터를 허용할지 여부를 지정할 수 있습니다.

다음 표에는 cleartext 열 허용 매개 변수의 값이 설명되어 있습니다.

파라미터값	설명
아니요	암호화된 테이블에는 Cleartext 열을 사용할 수 없습니다. 모든 데이터는 암호로 보호됩니다.
예	암호화된 테이블에는 Cleartext 열을 사용할 수 있습니다. Cleartext 열은 암호로 보호되지 않으며 cleartext로 포함됩니다. 행의 cleartext 데이터가 테이블에 있는 다른 데이터에 대해 무엇을 알려줄 수 있는지 내용을 기록해 두어야 합니다. 특정 열에서 SUM 또는 AVG를 실행하거나 해당 열이 cleartext 안에 있어야 합니다.

`CreateCollaboration` API 작업을 사용하여 `dataEncryptionMetadata` 매개변수의 경우 `allowCleartext`의 값을 `true` 또는 `false`로 설정할 수 있습니다. 이러한 API 작업을 사용하는 방법에 대한 자세한 내용은 AWS Clean Rooms API 참조를 참조하세요.

Cleartext 열은 테이블별 스키마에서 cleartext로 분류된 열에 해당합니다. 이러한 열의 데이터는 암호화되지 않으며 어떤 방식으로든 사용할 수 있습니다. Cleartext 열은 데이터가 민감하지 않거나 암호화된 sealed 열 또는 fingerprint 열에서 허용하는 것보다 더 많은 유연성이 필요한 경우에 유용할 수 있습니다.

중복 매개 변수 허용

콘솔에서 공동 작업을 생성할 때 중복 허용 매개변수를 설정하여 JOIN 쿼리용으로 암호화된 열에 NULL 값이 아닌 중복이 포함될 수 있는지 여부를 지정할 수 있습니다.

> **Important**
>
> 중복 허용, 이름이 다른 열의 JOIN 허용 및 NULL 값 보존 매개 변수는 별개이지만 서로 연관된 효과가 있습니다.

다음 표에는 중복 허용 매개 변수의 값이 설명되어 있습니다.

파라미터값	설명
아니요	반복되는 값은 fingerprint 열에 허용되지 않습니다. 단일 fingerprint 열의 모든 값은 고유해야 합니다.
예	한 fingerprint 열에 반복되는 값을 사용할 수 있습니다. 반복되는 값이 있는 열을 결합해야 하는 경우 이 값을 예로 설정하세요. 예로 설정하면 C3R 테이블 또는 결과의 fingerprint 열 내에 나타나는 빈도 패턴은 cleartext 데이터 구조에 대한 몇 가지 추가 정보를 암시할 수 있습니다.

`CreateCollaboration` API 작업을 사용하여 `dataEncryptionMetadata` 파라미터의 `allowDuplicates` 값을 `true` 또는 `false`로 설정할 수 있습니다. 이러한 API 작업을 사용하는 방법에 대한 자세한 내용은 AWS Clean Rooms API 참조를 참조하세요.

기본적으로 JOIN 쿼리에 암호화된 데이터를 사용해야 하는 경우 C3R 암호화 클라이언트는 해당 열에 중복 값이 없어야 합니다. 이 요구 사항은 데이터 보호를 강화하기 위한 노력입니다. 이 동작은 데이터에서 반복되는 패턴을 관찰할 수 없도록 하는 데 도움이 될 수 있습니다. 하지만 JOIN 쿼리에서 암호화된 데이터로 작업하고 싶고 값이 중복될 염려가 없는 경우에는 중복 허용 매개 변수를 사용하여 이러한 보수적 검사를 비활성화할 수 있습니다.

이름이 다른 열의 JOIN 허용 매개변수

콘솔에서 공동 작업을 생성할 때 이름이 다른 열의 JOIN 허용 매개 변수를 설정하여 이름이 다른 열 간의 JOIN 명령문이 지원되는지 여부를 지정할 수 있습니다.

자세한 내용은 열 헤더 이름의 정규화 단원을 참조하세요.

다음 표에는 이름이 다른 열의 JOIN 허용 매개 변수의 값이 설명되어 있습니다.

파라미터값	설명
아니요	이름이 다른 fingerprint 열의 조인은 지원되지 않습니다. JOIN 명령문은 이름이 같은 열에 대해서만 정확한 결과를 제공합니다. **Important** 아니오 값을 사용하면 정보 보안이 강화되지만 협동 참여자가 열 이름에 대해 사전에 동의해야 합니다. fingerprint 열로 암호화할 때 두 열의 이름이 다르고 이름이 다른 열 JOIN 허용 이 아니오로 설정된 경우 해당 열에 대한 JOIN 명령문은 결과를 생성하지 않습니다. 이는 암호화 후 값이 두 항목 간에 공유되지 않기 때문입니다.
예	이름이 다른 fingerprint 열의 조인이 지원됩니다. 유연성을 높이기 위해 사용자는 이 값을 예로 설정할 수 있습니다. 이렇게 하면 열 이름에 관계없이 열에 JOIN 명령문을 입력할 수 있습니다. 예로 설정하면 C3R 암호화 클라이언트는 fingerprint 열을 보호할 때 열의 이름을 고려하지 않습니다. 그 결과, C3R 테이블에서 서로 다른 fingerprint 열의 공통 값을 관찰할 수 있습니다.

파라미터값	설명
	예를 들어 행의 City 열과 State 열 모두에 동일한 암호화된 JOIN 값이 있는 경우 해당 값이 암호화된 New York 값이라고 추론하는 것이 합리적일 수 있습니다.

CreateCollaboration API 작업을 사용하여 dataEncryptionMetadata 매개 변수의 경우, allowJoinsOnColumnsWithDifferentNames 값을 true 또는 false로 설정할 수 있습니다. 이러한 API 작업을 사용하는 방법에 대한 자세한 내용은 AWS Clean Rooms API 참조를 참조하세요.

기본적으로 fingerprint 열 암호화는 4단계: 표 형식 파일의 암호화 스키마 생성 에 설정된 해당 열의 targetHeader에 영향을 받습니다. 따라서 동일한 cleartext 값이라도 암호화 대상 fingerprint 열마다 암호화된 표현이 서로 다릅니다.

이 매개 변수는 경우에 따라 cleartext 값 추론을 방지하는 데 유용할 수 있습니다. 예를 들어, fingerprint 열 City와 State에서 동일한 암호화된 값을 확인하면 해당 값이 New York라고 합리적으로 추론할 수 있습니다. 하지만 이 매개 변수를 사용하려면 쿼리에서 조인할 모든 열이 공유 이름을 갖도록 사전에 추가 조정이 필요합니다.

이름이 다른 열의 JOIN 허용 매개 변수를 사용하여 이러한 제한을 완화할 수 있습니다. 매개 변수 값을Yes로 설정하면 이름에 관계없이 JOIN에 대해 암호화된 모든 열을 함께 사용할 수 있습니다.

NULL 값 보존(매개변수)

콘솔에서 공동 작업을 생성할 때 NULL 값 보존 매개변수를 설정하여 해당 열에 값이 없음을 표시할 수 있습니다.

다음 표에서는 NULL값 보존 매개 변수의 값을 설명합니다.

파라미터값	설명
아니요	NULL 값은 보존되지 않습니다. NULL 값은 암호화된 테이블에서 NULL처럼 표시되지 않습니다. NULL 값은 C3R 테이블에서 고유한 임의 값으로 나타납니다.
예	NULL 값은 보존됩니다. NULL 값은 암호화된 테이블에서 NULL처럼 표시됩니다. NULL 값의 SQL 시맨틱이 필요한 경우 이 값을 예로 설정할 수 있습니다. 따라서 열의 암호화 여부와 중복

파라미터값	설명
	허용의 매개 변수 설정에 관계없이 NULL 항목이 C3R 테이블에서 NULL같이 나타납니다.

`CreateCollaboration` API 작업을 사용하여 `dataEncryptionMetadata` 매개변수의 `preserveNulls` 값을 `true` 또는 `false`로 설정할 수 있습니다. 이러한 API 작업을 사용하는 방법에 대한 자세한 내용은 AWS Clean Rooms API 참조를 참조하세요.

공동 작업을 위해 NULL 값 보존 매개변수를 아니요로 설정한 경우:

1. `cleartext` 열의 NULL 항목은 변경되지 않습니다.
2. 암호화된 `fingerprint` 열의 NULL 항목은 내용을 숨기기 위해 임의 값으로 암호화됩니다. 암호화된 열을 cleartext 열의 NULL 항목과 결합해도 해당 항목과 일치하는 NULL 항목이 하나도 생성되지 않습니다. 각각 고유한 무작위 콘텐츠를 받기 때문에 매칭이 이루어지지 않습니다.
3. 암호화된 `sealed` 열의 NULL 항목은 암호화됩니다.

공동 작업에 대해 NULL값 보존 매개 변수의 값을 예로 설정하면 열의 암호화 여부에 관계없이 모든 열의 NULL 항목이 NULL로 유지됩니다.

NULL값 보존 매개변수는 NULL과 같이 표현된 정보가 부족한 부분을 공유하려는 데이터 보강과 같은 시나리오에서 유용합니다. NULL값 보존 매개변수는 JOIN 또는 GROUP BY로 변환하려는 열에 NULL 값이 있는 경우 fingerprint 또는 HMAC 형식에서도 유용합니다.

중복 허용 및 값 NULL보존 값 매개 변수가 아니요로 설정된 경우 한 fingerprint 열에 NULL 항목이 두 개 이상 있으면 오류가 발생하고 암호화가 중지됩니다. 두 매개 변수 중 하나의 값이 예로 설정된 경우 해당 오류는 발생하지 않습니다.

Clean Rooms용 암호화 컴퓨팅의 선택적 플래그

다음 섹션에서는 표 형식 파일 사용자 지정 및 테스트를 위해 C3R 암호화 클라이언트를 사용하여 데이터를 암호화할 때 설정할 수 있는 선택적 플래그에 대해 설명합니다.

주제

- --csvInputNULLValue 플래그
- --csvOutputNULLValue 플래그
- --enableStackTraces 플래그

- --dryRun 플래그
- --tempDir 플래그

`--csvInputNULLValue` 플래그

C3R 암호화 클라이언트를 사용하여 데이터를 암호화할 때 `--csvInputNULLValue` 플래그를 사용하여 입력 데이터 NULL 항목의 사용자 지정 인코딩을 지정할 수 있습니다.

다음 표에는 이 플래그의 사용 및 매개변수가 요약되어 있습니다.

사용법	파라미터
선택 사항. 사용자는 입력 데이터의 NULL 항목에 대해 사용자 지정 인코딩을 지정할 수 있습니다.	입력 CSV 파일의 사용자 지정 NULL 값 인코딩

NULL 항목은 특히 SQL 테이블과 같은 풍부한 표 형식의 컨텍스트에서 내용이 부족한 것으로 간주되는 항목입니다. .csv는 역사적 이유로 이 특성화를 명시적으로 지원하지 않지만, 스페이스만 포함된 빈 항목을 NULL로 간주하는 것이 일반적입니다. 따라서 이는 C3R 암호화 클라이언트의 기본 동작이며 필요에 따라 사용자 지정할 수 있습니다.

`--csvOutputNULLValue` 플래그

C3R 암호화 클라이언트를 사용하여 데이터를 암호화할 때 `--csvOutputNULLValue` 플래그를 사용하여 출력 데이터의 NULL 항목에 대한 사용자 지정 인코딩을 지정할 수 있습니다.

다음 표에는 이 플래그의 사용 및 매개변수가 요약되어 있습니다.

사용법	파라미터
선택 사항. 사용자는 생성된 출력 파일에서 NULL 항목에 대한 사용자 지정 인코딩을 지정할 수 있습니다.	출력 CSV 파일의 사용자 지정 NULL 값 인코딩

NULL 항목은 특히 SQL 테이블과 같은 풍부한 표 형식의 컨텍스트에서 내용이 부족한 것으로 간주되는 항목입니다. .csv는 역사적 이유로 이 특성화를 명시적으로 지원하지 않지만, 스페이스만 포함된 빈

항목을 NULL로 간주하는 것이 일반적입니다. 따라서 이는 C3R 암호화 클라이언트의 기본 동작이며 필요에 따라 사용자 지정할 수 있습니다.

`--enableStackTraces` 플래그

C3R 암호화 클라이언트를 사용하여 데이터를 암호화하는 경우 `--enableStackTraces` 플래그를 사용하여 C3R에서 오류가 발생할 경우 오류 보고를 위한 추가 컨텍스트 정보를 제공합니다.

AWS 는 오류를 수집하지 않습니다. 오류가 발생하면 스택 트레이스를 사용하여 오류를 직접 해결하거나 스택 트레이스를에 보내 지원을 지원 받으세요.

다음 표에는 이 플래그의 사용 및 매개변수가 요약되어 있습니다.

사용법	파라미터
선택 사항. C3R 암호화 클라이언트에서 오류가 발생할 경우 오류 보고를 위한 추가 컨텍스트 정보를 제공하는 데 사용됩니다.	없음

`--dryRun` 플래그

C3R 암호화 클라이언트 암호화 및 암호 해독 명령에는 선택 사항인 `--dryRun` 플래그가 포함되어 있습니다. 플래그는 사용자가 제공한 모든 인수를 가져와 유효성과 일관성을 검사합니다.

`--dryRun` 플래그를 사용하여 스키마 파일이 유효하고 해당 입력 파일과 일치하는지 확인할 수 있습니다.

다음 표에는 이 플래그의 사용 및 매개변수가 요약되어 있습니다.

사용법	파라미터
선택 사항. C3R 암호화 클라이언트가 매개변수를 분석하고 파일을 검사하지만 암호화나 암호 해독은 수행하지 않도록 합니다.	없음

`--tempDir` 플래그

설정에 따라 암호화된 파일이 암호화되지 않은 파일보다 클 수 있으므로 임시 디렉터리를 사용하는 것이 좋습니다. 또한 공동 작업별로 데이터 세트를 암호화해야 제대로 작동할 수 있습니다.

C3R을 사용하여 데이터를 암호화하는 경우 `--tempDir` 플래그를 사용하여 입력을 처리하는 동안 임시 파일을 생성할 수 있는 위치를 지정합니다.

다음 표에는 이 플래그의 사용 및 매개변수가 요약되어 있습니다.

사용법	파라미터
사용자는 입력을 처리하는 동안 임시 파일을 생성할 수 있는 위치를 지정할 수 있습니다.	기본값은 시스템 임시 디렉터리입니다.

Clean Rooms에 대한 암호화 컴퓨팅을 사용한 쿼리

이 항목에서는 Clean Rooms에 대한 암호화 컴퓨팅을 사용하여 암호화된 데이터 테이블을 사용하는 쿼리 작성에 대한 정보를 제공합니다.

주제

- NULL에서 분기된 쿼리
- 하나의 소스 열을 여러 대상 열에 매핑
- JOIN 및 SELECT 쿼리 모두에 동일한 데이터 사용

NULL에서 분기된 쿼리

NULL 명령문에 쿼리 브랜치가 있다는 것은 `IF x IS NULL THEN 0 ELSE 1`와 같은 구문을 사용한다는 의미입니다.

쿼리는 항상 cleartext 열의 NULL 명령문에서 분기할 수 있습니다.

NULL 값 보존 매개 변수(`preserveNulls`)의 값이 `true`로 설정된 경우에만 sealed 열과 fingerprint 열의 NULL 명령문에 대해 쿼리를 분기할 수 있습니다.

이러한 제약조건을 위반하는 쿼리는 잘못된 결과를 초래할 수 있습니다.

하나의 소스 열을 여러 대상 열에 매핑

하나의 소스 열을 여러 대상 열에 매핑할 수 있습니다. 예를 들어 하나의 열에 JOIN 및 SELECT 둘다를 모두 매핑하고 싶을 수 있습니다.

자세한 내용은 JOIN 및 SELECT 쿼리 모두에 동일한 데이터 사용 단원을 참조하십시오.

JOIN 및 SELECT 쿼리 모두에 동일한 데이터 사용

열의 데이터가 민감하지 않은 경우 cleartext 대상 열에 표시되므로 어떤 용도로든 사용할 수 있습니다.

열의 데이터가 민감하고 JOIN 및 SELECT 쿼리 모두에 사용해야 하는 경우 해당 소스 열을 출력 파일의 두 대상 열에 매핑합니다. 한 열은 `type`을 fingerprint 열로 암호화하고, 한 열은 `type`을 봉인된 열로 암호화합니다. C3R 암호화 클라이언트의 대화형 스키마 생성은 헤더 접미사 `_fingerprint` 및 `_sealed`를 제안합니다. 이러한 헤더 접미사는 이러한 열을 빠르게 구분하는 데 유용한 규칙이 될 수 있습니다.

C3R 암호화 클라이언트에 대한 지침

C3R 암호화 클라이언트는 조직이 민감한 데이터를 한데 모아 데이터 분석에서 새로운 인사이트를 도출할 수 있도록 지원하는 도구입니다. 도구는 당사자 및 프로세스 AWS 에서 학습할 수 있는 내용을 암호화 방식으로 제한합니다. 이는 매우 중요하지만, 데이터를 암호화하여 보호하는 과정은 컴퓨팅 및 스토리지 리소스 측면에서 상당한 오버헤드를 유발할 수 있습니다. 따라서 각 설정 사용의 장단점을 이해하고 원하는 암호화 보장을 유지하면서 설정을 최적화하는 방법을 이해하는 것이 중요합니다. 이 항목에서는 C3R 암호화 클라이언트 및 스키마의 다양한 설정이 성능에 미치는 영향을 중점적으로 다룹니다.

모든 C3R 암호화 클라이언트 암호화 설정은 서로 다른 암호화 보장을 제공합니다. 공동 작업 수준 설정은 기본적으로 가장 안전합니다. 공동 작업을 생성하는 동안 추가 기능을 활성화하면 개인정보 보호가 약화되어 빈도 분석과 같은 활동이 사이퍼텍스트에서 수행될 수 있습니다. 이러한 설정이 사용되는 방식 및 그 영향에 대한 자세한 내용은 the section called “암호화 컴퓨팅”을 참조하세요.

주제

- 열 유형에 대한 성능 영향
- 예상하지 못한 사이퍼텍스트 크기 증가 문제 해결

열 유형에 대한 성능 영향

C3R은 세 가지 열 유형(cleartext, fingerprint 및 sealed)을 사용합니다. 각 열 유형은 서로 다른 암호화 보장을 제공하며 용도도 다릅니다. 다음 섹션에서는 열 유형이 성능에 미치는 영향과 각 설정이 성능에 미치는 영향에 대해 설명합니다.

주제

- Cleartext 열
- Fingerprint 열
- Sealed 열

Cleartext 열

Cleartext 열은 원래 형식에서 변경되지 않으며 어떤 방식으로도 암호화 처리되지 않습니다. 이 열 유형은 구성할 수 없으며 스토리지 또는 컴퓨팅 성능에 영향을 주지 않습니다.

Fingerprint 열

Fingerprint 열은 여러 테이블의 데이터를 조인하는 데 사용됩니다. 이를 위해서는 결과 사이퍼텍스트 크기가 항상 같아야 합니다. 그러나 이러한 열은 공동 작업 수준 설정의 영향을 받습니다. Fingerprint 열은 입력에 포함된 cleartext 열에 따라 출력 파일 크기에 다양한 정도의 영향을 미칠 수 있습니다.

주제

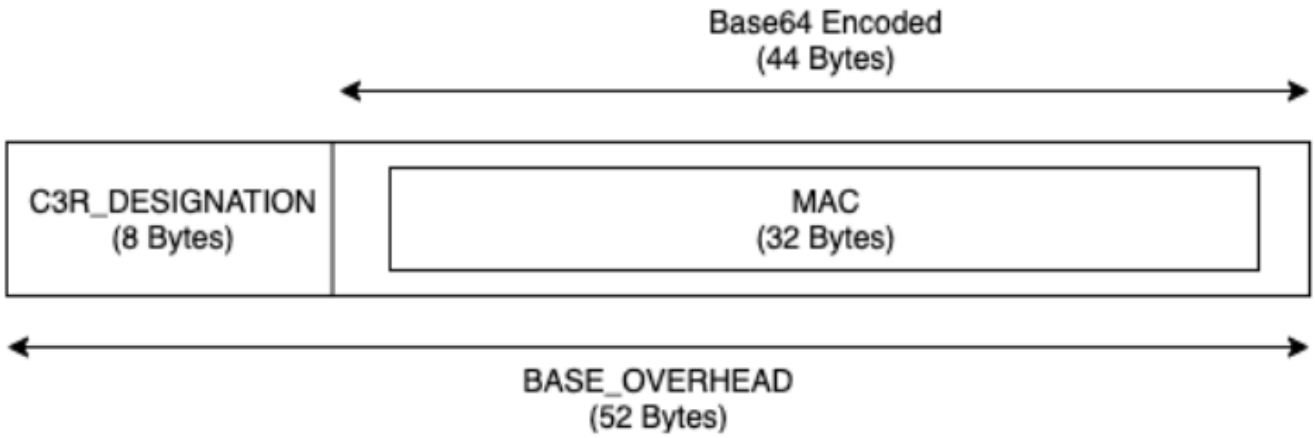
- fingerprint 열의 기본 오버헤드
- fingerprint 열에 대한 공동 작업 설정
- fingerprint 열의 예제 데이터
- 문제 해결 fingerprint 열

fingerprint 열의 기본 오버헤드

fingerprint 열에 대한 기본 오버헤드가 있습니다. 이 오버헤드는 일정하며 cleartext 바이트 크기를 대체합니다.

fingerprint 열의 데이터는 해시 기반 메시지 인증 코드(HMAC) 함수를 통해 암호화 처리되며, HMAC(해시 기반 메시지 인증 코드) 함수는 데이터를 32바이트 메시지 인증 코드(MAC)로 변환합니다. 그런 다음 이 데이터는 base64 인코더를 통해 처리되어 바이트 크기에 약 33% 가 추가됩니다. 데이터가 속하는 열의 유형과 해당 열을 생성한 클라이언트 버전을 지정하는 8바이트 C3R 지정이 앞에 추가됩니다. 최종 결과는 52바이트입니다. 그런 다음 이 결과에 행 수를 곱하여 총 기본 오버헤드를 구합니다 (`preserveNulls`이 참으로 설정된 경우 `null`이 아닌 총 값의 수 사용).

다음 이미지는 *BASE_OVERHEAD = C3R_DESIGNATION + (MAC * 1.33)* 방법을 보여줍니다



fingerprint 열의 출력 사이퍼텍스트은 항상 52바이트입니다. 입력 cleartext 데이터가 평균 52바이트를 초과하는 경우(예: 전체 주소) 스토리지가 크게 감소할 수 있습니다. 입력 cleartext 데이터가 평균 52바이트 미만인 경우(예: 고객 사용 기간) 스토리지가 크게 증가할 수 있습니다.

fingerprint 열에 대한 공동 작업 설정

`preserveNulls` 설정

공동 작업 수준 설정 `preserveNulls`이(가) `false`(기본값)인 경우, 각 `null` 값은 고유한 임의의 32바이트로 대체되어 `null`이(가) 아닌 것처럼 처리됩니다. 그 결과 이제 각 `null` 값은 52바이트가 되었습니다. 이렇게 하면 이 설정이 `true`이고 `null` 값이 `null`로 전달될 때와 비교하여 매우 희소한 데이터를 포함하는 테이블에 대한 스토리지 요구 사항이 크게 증가할 수 있습니다.

이 설정의 개인정보 보호가 필요하지 않고 데이터 세트 내에 `null` 값을 보관하고 싶다면 공동 작업을 만들 때 `preserveNulls` 설정을 활성화하세요. 공동 작업을 생성한 후에는 `preserveNulls` 설정을 변경할 수 없습니다.

fingerprint 열의 예제 데이터

다음은 재현할 설정이 있는 fingerprint 열의 입력 및 출력 데이터 세트 예제입니다. `allowCleartext`와(과) `allowDuplicates`같은 다른 공동 작업 수준 설정은 결과에 영향을 주지 않으므로 로컬에서 재현하려는 경우, `true` 또는 `false`(으)로 설정할 수 있습니다.

공유 비밀의 예: `wJalrXUtnFEMI/K7MDENG/bPxRfiCYEXAMPLEKEY`

공동 작업 ID 예시: `a1b2c3d4-5678-90ab-cdef-EXAMPLE11111`

allowJoinsOnColumnsWithDifferentNames: `True` 이 설정은 성능이나 스토리지 요구 사항에 영향을 주지 않습니다. 하지만 이 설정을 사용하면 다음 표에 표시된 값을 재현할 때 열 이름을 선택할 필요가 없습니다.

예시 1

입력	null
preserveNulls	TRUE
출력	null
DETERMINISTIC	Yes
입력 바이트	0
출력 바이트	0

예시 2

입력	null
preserveNulls	FALSE
출력	01:hmac:3lkFjthvV3IUu6mMvFc1a +XAHwgw/ElmOq4p3Yg25kk=
DETERMINISTIC	No
입력 바이트	0
출력 바이트	52

예 3

입력	empty string
preserveNulls	-
출력	01:hmac:oKTgi3Gba+eUb3JteSz 2EMgXUkF1WgM77UP0Ydw5kPQ=
DETERMINISTIC	Yes

입력 바이트	0
출력 바이트	52

예 4

입력	abcdefghijklmnopqrstuvwxyz
preserveNulls	-
출력	01:hmac:kU/IqwG7FMmzzshr0B9scomE0UJUEE7j9keTctplGww=
DETERMINISTIC	Yes
입력 바이트	26
출력 바이트	52

예 5

입력	abcdefghijklmnopqrstuvwxyzABCDEFGHIJKLMNOPQRSTUVWXYZ0123456789
preserveNulls	-
출력	01:hmac:ks3htnQbw2vdhCRFF6JNzW5LMndJaHG57uvE26mBtSs=
DETERMINISTIC	Yes
입력 바이트	62
출력 바이트	52

문제 해결 fingerprint 열

fingerprint 열의 사이퍼텍스트 텍스트가 그 안에 들어간 cleartext 열의 크기보다 몇 배나 큰 이유는 무엇인가요?

fingerprint 열의 사이퍼텍스트 길이는 항상 52바이트입니다. 입력 데이터가 작을 경우(예: 고객 연령) 크기가 크게 증가합니다. `preserveNulls` 설정이 `false`(으)로 설정된 경우에도 이런 일이 발생할 수 있습니다.

fingerprint 열의 사이퍼텍스트가 그 안에 들어간 cleartext 열의 크기보다 몇 배나 작은 이유는 무엇인가요?

fingerprint 열의 사이퍼텍스트 길이는 항상 52바이트입니다. 입력 데이터가 큰 경우(예: 고객의 전체 주소) 크기가 크게 줄어듭니다.

`preserveNulls`에서 제공하는 암호화 보증이 필요한지 어떻게 알 수 있나요?

안타깝게도 답은 상황에 따라 다르다는 것입니다. 최소한 `preserveNulls` 설정이 데이터를 어떻게 보호하는지 the section called “파라미터”을 검토해야 합니다. 하지만 조직의 데이터 처리 요구 사항 및 각 공동 작업에 적용되는 모든 계약을 참조하는 것이 좋습니다.

base64의 오버헤드가 발생해야 하는 이유는 무엇입니까?

CSV와 같은 표 형식 파일 형식과의 호환성을 허용하려면 base64 인코딩이 필요합니다. Parquet와 같은 일부 파일 형식은 데이터의 바이너리 표현을 지원할 수 있지만 올바른 쿼리 결과를 얻으려면 공동 작업의 모든 참여자가 동일한 방식으로 데이터를 나타내는 것이 중요합니다.

Sealed 열

Sealed 열은 공동 작업 구성원 간에 데이터를 전송할 때 사용됩니다. 이러한 열의 사이퍼텍스트는 결정적이지 않으며 열 구성 방식에 따라 성능과 스토리지 모두에 상당한 영향을 미칩니다. 이러한 열은 개별적으로 구성할 수 있으며 대개 C3R 암호화 클라이언트의 성능과 결과 출력 파일 크기에 가장 큰 영향을 미칩니다.

주제

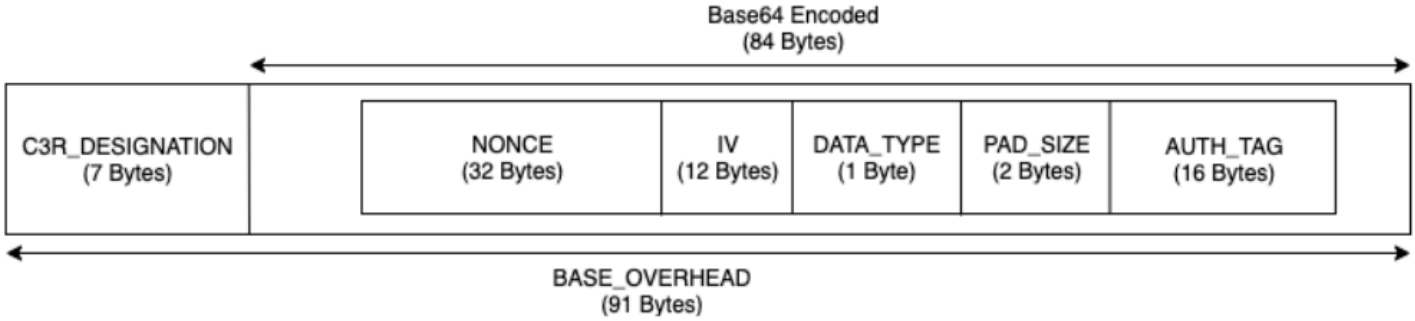
- sealed 열의 기본 오버헤드
- sealed 열에 대한 공동 작업 설정
- 스키마 설정 sealed 열: 패딩 유형
- sealed 열에 대한 예제 데이터
- 문제 해결 sealed 열

sealed 열의 기본 오버헤드

sealed 열에는 기본 오버헤드가 있습니다. 이 오버헤드는 일정하며, cleartext 및 패딩(있는 경우) 바이트의 크기와 함께 발생합니다.

암호화하기 전에는 sealed 열의 데이터 앞에 포함되는 데이터 유형을 지정하는 1바이트 문자가 추가됩니다. 패딩을 선택하면 데이터가 채워지고 패드 크기를 나타내는 2바이트가 추가됩니다. 이러한 바이트가 추가되면 AES-GCM을 사용하여 데이터를 암호화 처리하여 IV (12바이트), nonce (32바이트) 및 (16바이트) 로 저장합니다. Auth Tag 그런 다음 이 데이터는 base64 인코더를 통해 처리되어 바이트 크기에 약 33% 가 추가됩니다. 데이터 앞에 7바이트 C3R 지정이 추가되어 데이터가 속하는 열 유형과 데이터를 생성하는 데 사용된 클라이언트 버전을 지정합니다. 결과적으로 최종 베이스 오버헤드는 91바이트입니다. 그런 다음 이 결과에 행 수를 곱하여 총 기본 오버헤드를 구할 수 있습니다 (`preserveNulls`이 참으로 설정된 경우 null이 아닌 총 값 수 사용).

다음 이미지는 *BASE_OVERHEAD = C3R_DESIGNATION + ((NONCE + IV + DATA_TYPE + PAD_SIZE + AUTH_TAG) * 1.33)* 방법을 보여줍니다



sealed 열에 대한 공동 작업 설정

`preserveNulls` 설정

공동 작업 수준 설정 `preserveNulls`이(가) `false`(기본값)인 경우 각 `null` 값은 고유한 32바이트의 무작위 값이며, `null`이(가) 아닌 것처럼 처리됩니다. 그 결과 이제 각 `null` 값은 91바이트가 됩니다 (패딩된 경우 더 많음). 이렇게 하면 이 설정이 `true`이고 `null` 값이 `null`로 전달될 때와 비교하여 매우 희박한 데이터가 포함된 테이블의 경우 상당한 스토리지 요구 사항이 추가될 수 있습니다.

이 설정의 개인정보 보호가 필요하지 않고 데이터 세트 내에 `null` 값을 보관하고 싶다면 공동 작업을 만들 때 `preserveNulls` 설정을 활성화하세요. 공동 작업을 생성한 후에는 `preserveNulls` 설정을 변경할 수 없습니다.

스키마 설정 sealed 열: 패딩 유형

주제

- 패드 유형: none
- 패드 유형: fixed
- max의 패드 유형

패드 유형: **none**

none 패드 유형을 선택하면 패딩이 cleartext에 추가되지 않으며 앞서 설명한 기본 오버헤드에 추가 오버헤드가 추가되지 않습니다. 패딩이 없으면 출력 크기가 가장 공간 효율적입니다. 그러나 `fixed` 및 `max` 패딩 유형과 동일한 개인 정보 보호 보장을 제공하지는 않습니다. 이는 사이퍼텍스트의 크기로 기본 cleartext 파일의 크기를 식별할 수 있기 때문입니다.

패드 유형: **`fixed`**

`fixed` 패드 유형을 선택하면 개인 정보 보호를 위해 열에 포함된 데이터의 길이를 숨길 수 있습니다. 이는 암호화하기 전에 모든 cleartext를 제공된 `pad_length`에 패딩하는 방식으로 수행됩니다. 데이터가 이 크기를 초과하면 C3R 암호화 클라이언트가 실패합니다.

암호화되기 전에 cleartext에 패딩이 추가된다는 점을 감안할 때, AES-GCM은 cleartext를 사이퍼텍스트 바이트에 1대 1로 매핑합니다. base64 인코딩은 33% 를 추가할 것입니다. 패딩의 추가 스토리지 오버헤드는 `pad_length`의 값에서 cleartext의 평균 길이를 뺀 다음 1.33을 곱하여 계산할 수 있습니다. 결과는 레코드당 평균 패딩 오버헤드입니다. 그런 다음 이 결과에 행 수를 곱하여 총 패딩 오버헤드를 구할 수 있습니다(`preserveNulls`이 `true`로 설정된 경우 `null`이 아닌 총 값의 수 사용).

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

열의 가장 큰 값을 포함하는 최소 `pad_length`를 선택하는 것이 좋습니다. 예를 들어, 가장 큰 값이 50바이트인 경우 50의 `pad_length`면 충분합니다. 이보다 큰 값은 스토리지 오버헤드만 추가될 뿐입니다.

고정 패딩은 큰 컴퓨팅 오버헤드를 추가하지 않습니다.

`max`의 패드 유형

`max`의 패드 유형을 선택하면 개인 정보 보호를 위해 열에 포함된 데이터의 길이를 숨길 수 있습니다. 암호화하기 전에 모든 cleartext를 열의 가장 큰 값에 더한 추가 `pad_length`로 모두 채워 넣으면 됩니다. 일반적으로 `max` 패딩은 단일 데이터 세트의 `fixed` 패딩과 동일한 보장을 제공하지만 열의 가장 큰 cleartext을(를) 알 수 없도록 합니다. 그러나 개별 데이터 세트에서 가장 큰 값이 다를 수 있으므로 `max` 패딩은 업데이트 전반의 `fixed` 패딩과 동일한 개인정보 보호 보장을 제공하지 않을 수 있습니다.

max 패딩을 사용할 때는 pad_length를 0으로 추가 선택하는 것이 좋습니다. 이 길이는 모든 값을 열의 가장 큰 값과 같은 크기로 채웁니다. 이보다 큰 값은 스토리지 오버헤드만 추가될 뿐입니다.

해당 열의 최대 cleartext 값이 알려진 경우 fixed 패드 유형을 대신 사용하는 것이 좋습니다. fixed 패딩을 사용하면 업데이트된 데이터 세트 간에 일관성을 유지할 수 있습니다. max 패딩을 사용하면 각 데이터 하위 집합이 하위 집합에 있었던 가장 큰 값으로 채워집니다.

sealed 열에 대한 예제 데이터

다음은 재현할 설정이 있는 sealed 열의 입력 및 출력 데이터 세트 예제입니다. allowCleartext, allowJoinsOnColumnsWithDifferentNames, 및 allowDuplicates같은 기타 공동 작업 수준 설정은 결과에 영향을 주지 않으며 로컬에서 재현하려는 경우 true 또는 false(으)로 설정할 수 있습니다. 이러한 설정이 재현을 위한 기본 설정이긴 하지만 sealed 열은 결정적이지 않으므로 값이 매번 변경됩니다. 목표는 입력 바이트와 출력 바이트를 비교하여 표시하는 것입니다. 예제 pad_length 값은 의도적으로 선택되었습니다. fixed 패딩을 사용하면 권장되는 최소 pad_length 설정 또는 추가 패딩이 필요한 경우 max 패딩과 동일한 값을 얻을 수 있습니다.

공유 비밀의 예: wJalrXUtnFEMI/K7MDENG/bPxRfiCYEXAMPLEKEY

공동 작업 ID 예시: a1b2c3d4-5678-90ab-cdef-EXAMPLE11111

주제

- none의 패드 유형
- fixed의 패드 유형(예 1)
- fixed의 패드 유형(예 2)
- max의 패드 유형(예시 1)
- 패드 유형 max (예 2)

none의 패드 유형

예시 1

입력	null
preserveNulls	TRUE
출력	null

DETERMINISTIC	Yes
입력 바이트	0
출력 바이트	0

예시 2

입력	null
preserveNulls	FALSE
출력	01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc4OTBqfRYZ98t5KU6aWfssGSPbNIJfG3iXmu6cbCUrizuV
DETERMINISTIC	No
입력 바이트	0
출력 바이트	91

예 3

입력	empty string
preserveNulls	-
출력	01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc4OTBqfRYZ98t5KU6aWfstGSPEM6qR8DWC2PB2GMlX41YK
DETERMINISTIC	No
입력 바이트	0
출력 바이트	91

예 4

입력	abcdefghijklmnopqrstuvwxyz
preserveNulls	-
출력	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfsteEE1GKEPiRzyh0 h7t6OmWMLTWCvO2ckr6pkx9sGL5 VLDQeHzh6DmPpyWNuI=
DETERMINISTIC	No
입력 바이트	26
출력 바이트	127

예 5

입력	abcdefghijklmnopqrstuvwxyzA BCDEFGHIJKLMNOPQRSTUVWXYZ01 23456789
preserveNulls	-
출력	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfsteEE1GKEPiRzyh0 h7t6OmWMLTWCvO2ckr6plwtH/8t RFnn2rF91bcB9G4+n8GiRfJNmqdP4/ QOQ3cXb/pbvPcnnohrHIGSX54ua+1/ JfcVjc=
DETERMINISTIC	No
입력 바이트	62
출력 바이트	175

fixed의 패드 유형(예 1)

이 예제에서 pad_length는 62이고 최대 입력은 62바이트입니다.

예시 1

입력	null
preserveNulls	TRUE
출력	null
DETERMINISTIC	Yes
입력 바이트	0
출력 바이트	0

예시 2

입력	null
preserveNulls	FALSE
출력	01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc4OTBqfRYZ98t5KU6aWfssGSNWfMRp7nSb7SMX2s3JKLOhK1+7r75Tk+Mx9jy48Fcg1yOPvBqRSZ7oqy1V3UKfYTLEZb/hCz7oaIneVsrcoNpATs0GzbnLkor4L+/aSuA=
DETERMINISTIC	No
입력 바이트	0
출력 바이트	175

예 3

입력	empty string
preserveNulls	-
출력	01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc4OTBqfRYZ98t5KU6aWfstGSNWfMRp7nSb7SMX2s3JKLOhK1+7r75Tk+Mx9jy48Fcg1yOPvBqRSZ7oqy1V3UKfYTLEZb/hCz7oaIneVsrcoLB53l07VZpA6OwkuXu29CA=
DETERMINISTIC	No
입력 바이트	0
출력 바이트	175

예 4

입력	abcdefghijklmnopqrstuvwxyz
preserveNulls	-
출력	01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc4OTBqfRYZ98t5KU6aWfsteEE1GKEPiRzyh0h7t6OmWMLTWCvO2ckr6pkx9jy48Fcg1yOPvBqRSZ7oqy1V3UKfYTLEZb/hCz7oaIneVsrcutBAcO+Mb9tuU2KIHH31AWg=
DETERMINISTIC	No
입력 바이트	26
출력 바이트	175

예 5

입력	abcdefghijklmnopqrstuvwxyzA BCDEFGHIJKLMNOPQRSTUVWXYZ01 23456789
preserveNulls	-
출력	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfsteEE1GKEPiRzyh0 h7t6OmWMLTWCvO2ckr6plwtH/8t RFnn2rF91bcB9G4+n8GiRfJNmqdP4/ QOQ3cXb/pbvPcnnohrHIGSX54ua+1/ JfcVjc=
DETERMINISTIC	No
입력 바이트	62
출력 바이트	175

fixed의 패드 유형(예 2)

이 예제에서 pad_length는 162이고 최대 입력은 62바이트입니다.

예시 1

입력	null
preserveNulls	TRUE
출력	null
DETERMINISTIC	Yes
입력 바이트	0
출력 바이트	0

예시 2

입력	null
preserveNulls	FALSE
출력	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfssGSNWfMRp7nSb7S MX2s3JKLOhK1+7r75Tk+Mx9jy48 Fcg1yOPvBqRSZ7oqy1V3UKfYTLE Zb/hCz7oaIneVsrcnkB0xbLWD7z NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvtkSSVc2B80OOGp pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrmwv/xAySX+xcntotL703aBTBb
DETERMINISTIC	No
입력 바이트	0
출력 바이트	307

예 3

입력	empty string
preserveNulls	-
출력	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfstGSNWfMRp7nSb7S MX2s3JKLOhK1+7r75Tk+Mx9jy48 Fcg1yOPvBqRSZ7oqy1V3UKfYTLE Zb/hCz7oaIneVsrcnkB0xbLWD7z NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvtkSSVc2B80OOGp

	pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrmwv84lVaT9Yd+6oQx65/+gdVT
DETERMINISTIC	No
입력 바이트	0
출력 바이트	307

예 4

입력	abcdefghijklmnopqrstuvwxyz
preserveNulls	-
출력	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfsteEE1GKEPiRzyh0 h7t6OmWMLTWCvO2ckr6pkx9jy48 Fcg1yOPvBqRSZ7oqy1V3UKfYTLE Zb/hCz7oaIneVsrcnkB0xbLWD7z NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvtkSSVc2B80OOGp pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrmwtX5Hnl+WyfO6ks3QMaRDGSf
DETERMINISTIC	No
입력 바이트	26
출력 바이트	307

예 5

입력	abcdefghijklmnopqrstuvwxyzA BCDEFGHIJKLMNOPQRSTUVWXYZ01 23456789
preserveNulls	-
출력	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfsteEE1GKEPiRzyh0 h7t6OmWMLTWCvO2ckr6plwtH/8t RFnn2rF91bcB9G4+n8GiRfJNmqd P4/QOQ3cXb/pbvPcnkB0xbLWD7z NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvtkSSVc2B80OOGp pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrmwjkJXQZOgPdeFX9Yr/8alV5i
DETERMINISTIC	No
입력 바이트	62
출력 바이트	307

max의 패드 유형(예시 1)

이 예제에서 pad_length는 0이고 최대 입력은 62바이트입니다.

예시 1

입력	null
preserveNulls	TRUE
출력	null
DETERMINISTIC	Yes

입력 바이트	0
출력 바이트	0

예시 2

입력	null
preserveNulls	FALSE
출력	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfssGSNWfMRp7nSb7S MX2s3JKLOhK1+7r75Tk+Mx9jy48 Fcg1yOPvBqRSZ7oqy1V3UKfYTLEZb/ hCz7oaIneVsrcoNpATs0GzbnLkor4L+/ aSuA=
DETERMINISTIC	No
입력 바이트	0
출력 바이트	175

예 3

입력	empty string
preserveNulls	-
출력	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfstGSNWfMRp7nSb7S MX2s3JKLOhK1+7r75Tk+Mx9jy48 Fcg1yOPvBqRSZ7oqy1V3UKfYTLE Zb/hCz7oaIneVsrcoLB53l07VZp A6OwkuXu29CA=

DETERMINISTIC	No
입력 바이트	0
출력 바이트	175

예 4

입력	abcdefghijklmnopqrstuvwxyz
preserveNulls	-
출력	01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc4OTBqfRYZ98t5KU6aWfsteEE1GKEPiRzyh0h7t6OmWMLTWCvO2ckr6pkx9jy48Fcg1yOPvBqRSZ7oqy1V3UKfYTLEZb/hCz7oaIneVsrcutBAcO+Mb9tuU2KIHH31AWg=
DETERMINISTIC	No
입력 바이트	26
출력 바이트	175

예 5

입력	abcdefghijklmnopqrstuvwxyzABCDEFGHIJKLMNOPQRSTUVWXYZ0123456789
preserveNulls	-
출력	01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc4OTBqfRYZ98t5KU6aWfsteEE1GKEPiRzyh0h7t6OmWMLTWCvO2ckr6plwtH/8t

	RFnn2rF91bcB9G4+n8GiRfJNmqdP4/QOQ3cXb/pbvPcnnohrHIGSX54ua+1/JfcVjc=
DETERMINISTIC	No
입력 바이트	62
출력 바이트	175

패드 유형 **max** (예 2)

이 예제에서 pad_length 는 100이고 최대 입력은 62바이트입니다.

예시 1

입력	null
preserveNulls	TRUE
출력	null
DETERMINISTIC	Yes
입력 바이트	0
출력 바이트	0

예시 2

입력	null
preserveNulls	FALSE
출력	01:enc:bm9uY2UwMTIzNDU2Nzg5MG5vbmNlMDEyMzQ1Njc4OTBqfRYZ98t5KU6aWfssGSNWfMRp7nSb7SMX2s3JKLOhK1+7r75Tk+Mx9jy48Fcg1yOPvBqRSZ7oqy1V3UKfYTLEZb/hCz7oaIneVsrcnkB0xbLWD7z

	NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvtkSSVc2B80OOGp pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrmwv/xAySX+xcntotL703aBTBb
DETERMINISTIC	No
입력 바이트	0
출력 바이트	307

예 3

입력	empty string
preserveNulls	-
출력	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfstGSNWfMRp7nSb7S MX2s3JKLOhK1+7r75Tk+Mx9jy48 Fcg1yOPvBqRSZ7oqy1V3UKfYTLE Zb/hCz7oaIneVsrcnkB0xbLWD7z NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvtkSSVc2B80OOGp pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrmwv84lVaT9Yd+6oQx65/+gdVT
DETERMINISTIC	No
입력 바이트	0
출력 바이트	307

예 4

입력	abcdefghijklmnopqrstuvwxyz
preserveNulls	-
출력	01:enc:bm9uY2UwMTIzNDU2Nzg5 MG5vbmNlMDEyMzQ1Njc4OTBqfRY Z98t5KU6aWfsteEE1GKEPiRzyh0 h7t6OmWMLTWCvO2ckr6pkx9jy48 Fcg1yOPvBqRSZ7oqy1V3UKfYTLE Zb/hCz7oaIneVsrcnkB0xbLWD7z NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvtkSSVc2B80OOGp pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrmwtX5Hnl+WyfO6ks3QMaRDGSf
DETERMINISTIC	No
입력 바이트	26
출력 바이트	307

예 5

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

	NdAqQGR0rXoSESdW0I0vpNoGcBf v4cJbG0A3h1DvtkSSVc2B80OOGp pzdDqhrUVN5wFNyn8vgfPMqDaeJk5bn +8o4WtG/ClipNcjDXvXVtK4vfCohcCA6 uwrmwjkJXQZOgPdeFX9Yr/8alV5i
DETERMINISTIC	No
입력 바이트	62
출력 바이트	307

문제 해결 sealed 열

sealed 열의 사이퍼텍스트가 그 안에 들어간 cleartext 열의 크기보다 몇 배나 큰 이유는 무엇인가요?

이는 여러 요인에 따라 달라집니다. 예를 들어, Cleartext 열의 사이퍼텍스트 길이는 항상 91바이트 이상입니다. 입력 데이터가 작을 경우(예: 고객 연령) 크기가 크게 증가합니다. 둘째, `preserveNulls`이 `false`로 설정되어 있고 입력 데이터에 `null` 값이 많이 포함되어 있는 경우, 각 `null` 값은 91바이트의 사이퍼텍스트로 변환됩니다. 마지막으로, 패딩을 사용하면 정의상 데이터가 암호화되기 전에 cleartext 데이터에 바이트가 추가됩니다.

sealed 열에 있는 대부분의 데이터는 매우 작아서 패딩을 사용해야 합니다. 스페이스를 절약하기 위해 큰 값을 제거하고 개별적으로 처리해도 될까요?

큰 값을 삭제하고 별도로 처리하지 않는 것이 좋습니다. 이렇게 하면 C3R 암호화 클라이언트가 제공하는 개인 정보 보호 보장이 변경됩니다. 위협 모델로서 관찰자가 암호화된 데이터 세트를 모두 볼 수 있다고 가정해 봅시다. 관찰자는 한 데이터 하위 집합의 열이 다른 하위 집합보다 훨씬 많거나 적게 채워진 것을 발견하면 각 하위 집합의 데이터 크기를 추론할 수 있습니다. 예를 들어 한 파일에서는 `fullName` 열이 총 40바이트로 채워지고 다른 파일에서는 800바이트로 채워져 있다고 가정해 보겠습니다. 관찰자는 한 데이터 세트에 세계에서 가장 긴 이름 (747바이트) 이 포함되어 있다고 가정할 수 있습니다.

`max` 패딩 유형을 사용할 때 추가 패딩을 제공해야 하나요?

`max` 패딩을 사용하는 경우 열에서 가장 큰 값을 초과하는 추가 패딩이라고도 하는 `pad_length`를 0으로 설정하는 것이 좋습니다.

가장 큰 값이 맞을지 걱정하지 않도록 **`fixed`** 패딩을 사용할 때 큰 **`pad_length`**를 선택해도 되나요?

네, 하지만 패드 길이가 길면 비효율적이며 필요 이상으로 많은 저장 공간을 사용합니다. 가장 큰 값이 얼마나 큰지 확인하고 `pad_length`를 해당 값으로 설정하는 것이 좋습니다.

`preserveNulls`에서 제공하는 암호화 보증이 필요한지 어떻게 알 수 있나요?

안타깝게도 답은 상황에 따라 다르다는 것입니다. 최소한 `preserveNulls` 설정이 데이터를 어떻게 보호하는지 Clean Rooms에 대한 암호화 컴퓨팅을 검토해야 합니다. 하지만 조직의 데이터 처리 요구 사항 및 각 공동 작업에 적용되는 모든 계약을 참조하는 것이 좋습니다.

base64의 오버헤드가 발생해야 하는 이유는 무엇입니까?

CSV와 같은 표 형식 파일 형식과의 호환성을 허용하려면 base64 인코딩이 필요합니다. Parquet같은 일부 파일 형식은 데이터의 바이너리 표현을 지원할 수 있지만 올바른 쿼리 결과를 얻으려면 공동 작업의 모든 참여자가 동일한 방식으로 데이터를 나타내는 것이 중요합니다.

예상하지 못한 사이퍼텍스트 크기 증가 문제 해결

데이터를 암호화했는데 결과 데이터의 크기가 놀라울 정도로 크다고 가정해 보겠습니다. 다음 단계는 크기 증가가 발생한 위치와 취할 수 있는 조치(있는 경우)를 식별하는 데 도움이 될 수 있습니다.

크기 증가가 발생한 위치 식별

암호화된 데이터가 cleartext 데이터보다 훨씬 큰 이유를 해결하려면 먼저 크기가 어디서 증가했는지 확인해야 합니다. Cleartext 열은 변경되지 않으므로 무시해도 됩니다. 나머지 fingerprint 열과 sealed 열을 살펴보고 중요해 보이는 열을 선택하세요.

크기 증가가 발생한 원인 파악

fingerprint 열 또는 sealed 열이 크기 증가에 영향을 줄 수 있습니다.

주제

- fingerprint 열에서 크기가 커졌나요?
- sealed 열을 통해 크기가 커졌나요?

fingerprint 열에서 크기가 커졌나요?

스토리지 증가에 가장 큰 영향을 미치는 fingerprint 열인 경우, cleartext 데이터가 적기 때문일 수 있습니다(예: 고객 연령). 결과로 생성되는 각 fingerprint 사이퍼텍스트의 길이는 52바이트입니다. 안타깝게도 이 문제에 대해 열별로 수행할 수 있는 작업은 없습니다. 이 열이 스토리지 요구 사항에 미치는 영향을 포함하여 이 열에 대한 자세한 내용은 fingerprint 열의 기본 오버헤드을(를) 참조하세요.

fingerprint 열의 크기가 증가하는 다른 가능한 원인은 공동 작업 설정인 `preserveNulls`입니다. `preserveNulls`에 대한 공동 작업 설정을 사용하지 않도록 설정하면(기본 설정) fingerprint 열의 모든 `null` 값이 52바이트의 사이퍼텍스트가 됩니다. 현재 공동 작업에서는 이에 대해 수행할 수 있는 작업이 없습니다. `preserveNulls` 설정은 공동 작업을 생성할 때 설정되며 올바른 쿼리 결과를 얻으려면 모든 협업자가 동일한 설정을 사용해야 합니다. `preserveNulls` 설정 및 설정 활성화가 데이터의 개인 정보 보장에 미치는 영향에 대한 자세한 내용은 the section called “암호화 컴퓨팅”을 참조하세요.

sealed 열을 통해 크기가 커졌나요?

저장용량 증가에 가장 큰 영향을 미치는 sealed 열이 열인 경우 크기 증가에 기여할 수 있는 몇 가지 세부 정보가 있습니다.

cleartext 데이터가 작은 경우(예: 고객 연령) 결과로 생성되는 각 sealed 사이퍼텍스트의 길이는 91바이트 이상입니다. 안타깝게도 이 문제에 대해서는 아무 것도 할 수 없습니다. 스토리지 요구 사항에 미치는 영향을 포함하여 이 열에 대한 자세한 내용은 sealed 열의 기본 오버헤드에 대한 세부 정보를 참조하세요.

sealed 열 스토리지 증가의 두 번째 주요 원인은 패딩입니다. 패딩은 데이터 세트에서 개별 값의 크기를 숨기기 위해 암호화하기 전에 cleartext에 추가 바이트를 추가합니다. 패딩을 데이터 세트의 가능한 최소값으로 설정하는 것이 좋습니다. 최소한 열에서 가능한 가장 큰 값을 포함하도록 `fixed`에 대한 `pad_length` 패딩을 설정해야 합니다. 이보다 높게 설정해도 개인 정보 보장이 추가되지는 않습니다. 예를 들어 열의 가능한 최대 값이 50바이트라는 것을 알고 있는 경우 `pad_length`를 50바이트로 설정하는 것이 좋습니다. 그러나 sealed 열이 `max` 패딩을 사용하는 경우에는 `pad_length`를 0바이트로 설정하는 것이 좋습니다. 이는 `max` 패딩이 열의 가장 큰 값을 초과하는 추가 패딩을 의미하기 때문입니다.

sealed 열의 크기가 커질 수 있는 마지막 원인은 공동 작업 설정인 `preserveNulls`입니다. `preserveNulls`에 대한 공동 작업 설정을 사용하지 않도록 설정하면(기본 설정) sealed 열의 모든 `null` 값이 91바이트의 사이퍼텍스트가 됩니다. 현재 공동 작업에서는 이에 대해 수행할 수 있는 작업이 없습니다. `preserveNulls` 설정은 공동 작업이 생성될 때 설정되며 올바른 쿼리 결과를 얻으려면 모든 협업자가 동일한 설정을 사용해야 합니다. 이 설정의 영향 및 활성화가 데이터의 개인 정보 보호 보장에 미치는 영향에 대한 자세한 내용은 the section called “암호화 컴퓨팅”을 참조하세요.

설 AWS Clean Rooms정

다음 주제에서는 설정 방법을 설명합니다 AWS Clean Rooms.

주제

- 에 가입 AWS
- 에 대한 서비스 역할 설정 AWS Clean Rooms
- AWS Clean Rooms ML에 대한 서비스 역할 설정

에 가입 AWS

AWS Clean Rooms또는를 사용하려면 먼저 AWS 에 가입 AWS 서비스해야 합니다 AWS 계정.

이 없는 경우 다음 단계를 AWS 계정완료하여 생성합니다.

에 가입하려면 AWS 계정

1. https://portal.aws.amazon.com/billing/signup을 엽니다.
2. 온라인 지시 사항을 따릅니다.

   가입 절차 중에 전화 키패드에 입력할 확인 코드가 포함된 전화를 받게 됩니다.
3. 에 가입하면 AWS 계정AWS 계정 루트 사용자가 생성됩니다. 루트 사용자에게는 계정의 모든 AWS 서비스 및 리소스에 액세스할 권한이 있습니다. 보안 모범 사례는 관리 사용자에게 관리자 액세스 권한을 할당하고, 루트 사용자만 루트 사용자 액세스 권한이 필요한 태스크를 수행하는 것입니다.

에 대한 서비스 역할 설정 AWS Clean Rooms

다음 섹션에서는 각 작업을 수행하는 데 필요한 역할을 설명합니다.

주제

- 관리자 사용자 생성하기
- 공동 작업 구성원의 IAM 역할 생성
- Amazon S3에서 데이터를 읽을 서비스 역할 생성
- Amazon Athena에서 데이터를 읽을 서비스 역할 생성

- 서비스 역할을 생성하여 Snowflake에서 데이터 읽기
- 서비스 역할을 생성하여 S3 버킷에서 코드 읽기(PySpark 분석 템플릿 역할)
- 서비스 역할을 생성하여 PySpark 작업의 결과 작성
- 결과를 수신할 서비스 역할 생성

관리자 사용자 생성하기

를 사용하려면 관리자 사용자를 직접 생성하고 관리자 그룹에 관리자 사용자를 추가 AWS Clean Rooms해야 합니다.

다음 옵션 중 하나를 선택하여 관리 사용자를 생성합니다.

관리자를 관리하는 방법 한 가지 선택	목적	By	다른 방법
IAM Identity Center에서 (권장)	단기 보안 인증 정보를 사용하여 AWS에 액세스합니다. 이는 보안 모범 사례와 일치합니다. 모범 사례에 대한 자세한 내용은 IAM 사용 설명서의 IAM의 보안 모범 사례를 참조하세요.	AWS IAM Identity Center 사용 설명서의 시작하기 지침을 따르세요.	AWS Command Line Interface 사용 설명서에서 사용하도록 AWS CLI 를 구성 AWS IAM Identity Center하여 프로그래밍 방식 액세스를 구성합니다.
IAM에서 (권장되지 않음)	장기 보안 인증 정보를 사용하여 AWS에 액세스합니다.	IAM 사용 설명서의 비상 액세스를 위한 IAM 사용자 생성에 나와 있는 지침을 따르세요.	IAM 사용 설명서에 나온 IAM 사용자의 액세스 키 관리 단계를 수행하여 프로그래밍 방식의 액세스를 구성합니다.

공동 작업 구성원의 IAM 역할 생성

멤버는 공동 작업의 참가자인 AWS 고객입니다.

공동 작업 구성원의 IAM 역할을 생성하는 방법

1. AWS Identity and Access Management 사용 설명서에 설명된 IAM 사용자에게 권한을 위임하는 역할 생성 절차를 수행합니다.
2. 정책 생성 단계의 정책 편집기에서 JSON 탭을 선택한 다음 공동 작업 구성원에게 부여된 권한에 따라 정책을 추가합니다.

   AWS Clean Rooms 는 일반적인 사용 사례를 기반으로 다음과 같은 관리형 정책을 제공합니다.

다음을 수행하려는 경우 ...	그럼... 을(를) 사용하세요.
리소스 및 메타데이터 보기	AWS 관리형 정책: AWSCleanRoomsReadOnlyAccess
Query	AWS 관리형 정책: AWSCleanRoomsFullAccess
쿼리 및 실행 작업	AWS 관리형 정책: AWSCleanRoomsFullAccess
결과 쿼리 및 수신	AWS 관리형 정책: AWSCleanRoomsFullAccess
공동 작업 리소스는 관리하지만 쿼리하지 않음	AWS 관리형 정책: AWSCleanRoomsFullAccessNoQuerying

   에서 제공하는 다양한 관리형 정책에 대한 자세한 내용은 단원 AWS Clean Rooms을 참조하십시오AWS 에 대한 관리형 정책 AWS Clean Rooms.

Amazon S3에서 데이터를 읽을 서비스 역할 생성

AWS Clean Rooms 는 서비스 역할을 사용하여 Amazon S3에서 데이터를 읽습니다.

이 서비스 역할을 생성하는 방법에는 두 가지가 있습니다.

- 서비스 역할을 생성하는 데 필요한 IAM 권한이 있는 경우 AWS Clean Rooms 콘솔을 사용하여 서비스 역할을 생성합니다.
- `iam:CreateRole`, `iam:CreatePolicy` 및 `iam:AttachRolePolicy` 권한이 없거나 IAM 역할을 수동으로 생성하려는 경우 다음 중 하나를 수행합니다.
  - 다음 절차에 따라 사용자 지정 신뢰 정책을 사용하여 서비스 역할을 생성합니다.
  - 다음 절차에 따라 서비스 역할을 생성하도록 관리자에게 요청합니다.

> **Note**
>
> 사용자 또는 IAM 관리자는 콘솔을 사용하여 AWS Clean Rooms 서비스 역할을 생성하는 데 필요한 권한이 없는 경우에만 이 절차를 따라야 합니다.

사용자 지정 신뢰 정책을 사용하여 Amazon S3에서 데이터를 읽을 서비스 역할을 생성하려면

1. 사용자 지정 신뢰 정책을 사용하여 역할을 생성합니다. 자세한 내용은 AWS Identity and Access Management 사용 설명서의 사용자 지정 신뢰 정책을 사용하여 역할 생성(콘솔) 절차를 참조하세요.
2. 사용자 지정 신뢰 정책을 사용하여 역할 생성(콘솔) 절차에 따라 다음 사용자 지정 신뢰 정책을 사용합니다.

   > **Note**
   >
   > 역할이 특정 공동 작업 멤버십의 컨텍스트에서만 사용되도록 하려면 신뢰 정책의 범위를 더 좁힐 수 있습니다. 자세한 내용은 교차 서비스 혼동된 대리자 예방 단원을 참조하십시오.

   JSON

```
{

    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "RoleTrustPolicyForCleanRoomsService",
            "Effect": "Allow",
```

```
            "Principal": {
                "Service": ".amazonaws.com"
            },
            "Action": "sts:AssumeRole"
        }
    ]
}
```

3. 사용자 지정 신뢰 정책을 사용하여 역할 생성(콘솔) 절차에 따라 다음 권한 정책을 사용합니다.

> **Note**
>
> 다음 예제 정책은 AWS Glue 메타데이터와 해당 Amazon S3 데이터를 읽는 데 필요한 권한을 지원합니다. 그러나 Amazon S3 데이터를 설정한 방법에 따라이 정책을 수정해야 할 수 있습니다. 예를 들어 Amazon S3 데이터에 대한 사용자 지정 KMS 키를 설정한 경우 추가 AWS Key Management Service (AWS KMS) 권한으로이 정책을 수정해야 할 수 있습니다.
>
> AWS Glue 리소스와 기본 Amazon S3 리소스는 AWS Clean Rooms 공동 작업 AWS 리전과 동일한에 있어야 합니다.

JSON

```
{

    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "NecessaryGluePermissions",
            "Effect": "Allow",
            "Action": [
                "glue:GetDatabase",
                "glue:GetDatabases",
                "glue:GetTable",
                "glue:GetTables",
                "glue:GetPartition",
                "glue:GetPartitions",
                "glue:BatchGetPartition"
            ],
            "Resource": [
                "arn:aws:glue: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. 각 ## ###를 자신의 정보로 바꿉니다.
5. 계속해서 사용자 지정 신뢰 정책을 사용하여 역할 생성(콘솔) 절차에 따라 역할을 생성합니다.

Amazon Athena에서 데이터를 읽을 서비스 역할 생성

AWS Clean Rooms 는 서비스 역할을 사용하여 Amazon Athena에서 데이터를 읽습니다.

사용자 지정 신뢰 정책을 사용하여 Athena에서 데이터를 읽을 서비스 역할을 생성하려면

1. 사용자 지정 신뢰 정책을 사용하여 역할을 생성합니다. 자세한 내용은 AWS Identity and Access Management 사용 설명서의 사용자 지정 신뢰 정책을 사용하여 역할 생성(콘솔) 절차를 참조하세요.
2. 사용자 지정 신뢰 정책을 사용하여 역할 생성(콘솔) 절차에 따라 다음 사용자 지정 신뢰 정책을 사용합니다.

> **Note**
>
> 역할이 특정 공동 작업 멤버십의 컨텍스트에서만 사용되도록 하려면 신뢰 정책의 범위를 더 좁힐 수 있습니다. 자세한 내용은 교차 서비스 혼동된 대리자 예방 단원을 참조하십시오.

JSON

```
{

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

```
            "Effect": "Allow",
            "Principal": {
                "Service": ".amazonaws.com"
            },
            "Action": "sts:AssumeRole"
        }
    ]
}
```

3. 사용자 지정 신뢰 정책을 사용하여 역할 생성(콘솔) 절차에 따라 다음 권한 정책을 사용합니다.

> **Note**
>
> 다음 예제 정책은 AWS Glue 메타데이터 및 해당 Athena 데이터를 읽는 데 필요한 권한을 지원합니다. 그러나 Amazon S3 데이터를 설정한 방법에 따라이 정책을 수정해야 할 수 있습니다. 예를 들어 Amazon S3 데이터에 대한 사용자 지정 KMS 키를 이미 설정한 경우 추가 AWS KMS 권한으로이 정책을 수정해야 할 수 있습니다.
> AWS Glue 리소스와 기본 Athena 리소스는 AWS Clean Rooms 공동 작업 AWS 리전 과 동일한에 있어야 합니다.

JSON

```
{
    "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. 각 ## ###를 자신의 정보로 바꿉니다.
5. 계속해서 사용자 지정 신뢰 정책을 사용하여 역할 생성(콘솔) 절차에 따라 역할을 생성합니다.

Lake Formation 권한 설정

Lake Formation 권한으로 보호되는 리소스를 쿼리하는 경우 서비스 역할에는 테이블/보기에 대한 선택 및 설명 액세스 권한과 보기가 저장된 AWS Glue 데이터베이스에 대한 설명 액세스 권한이 있어야 합니다.

자세한 내용은 다음을 참조하세요.

- Athena를 사용하여 Amazon Athena 사용 설명서의에 등록된 데이터 쿼리 AWS Lake Formation Amazon Athena
- AWS Lake Formation 개발자 안내서의 Lake Formation에 온보딩 권한

서비스 역할을 생성하여 Snowflake에서 데이터 읽기

AWS Clean Rooms 는 서비스 역할을 사용하여 Snowflake가이 소스에서 데이터를 읽을 수 있도록 자격 증명을 검색합니다.

이 서비스 역할을 만드는 방법은 다음 두 가지입니다.

- 서비스 역할을 생성하는 데 필요한 IAM 권한이 있는 경우 AWS Clean Rooms 콘솔을 사용하여 서비스 역할을 생성합니다.
- `iam:CreateRole`, `iam:CreatePolicy` 및 `iam:AttachRolePolicy` 권한이 없거나 IAM 역할을 수동으로 생성하려는 경우 다음 중 하나를 수행합니다.
  - 다음 절차에 따라 사용자 지정 신뢰 정책을 사용하여 서비스 역할을 생성합니다.
  - 다음 절차에 따라 서비스 역할을 생성하도록 관리자에게 요청합니다.

> **Note**
>
> 사용자 또는 IAM 관리자는 콘솔을 사용하여 AWS Clean Rooms 서비스 역할을 생성하는 데 필요한 권한이 없는 경우에만이 절차를 따라야 합니다.

사용자 지정 신뢰 정책을 사용하여 Snowflake에서 데이터를 읽을 서비스 역할을 생성하려면

1. 사용자 지정 신뢰 정책을 사용하여 역할을 생성합니다. 자세한 내용은 AWS Identity and Access Management 사용 설명서의 사용자 지정 신뢰 정책을 사용하여 역할 생성(콘솔) 절차를 참조하세요.
2. 사용자 지정 신뢰 정책을 사용하여 역할 생성(콘솔) 절차에 따라 다음 사용자 지정 신뢰 정책을 사용합니다.

> **Note**
>
> 역할이 특정 공동 작업 멤버십의 컨텍스트에서만 사용되도록 하려면 신뢰 정책의 범위를 더 좁힐 수 있습니다. 자세한 내용은 교차 서비스 혼동된 대리자 예방 단원을 참조하십시오.

JSON

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

 "arn:aws:cleanrooms:region:accountId:membership/membershipId",

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

```
}
```

3. 사용자 지정 신뢰 정책을 사용하여 역할 생성(콘솔) 절차에 따라 다음 권한 정책 중 하나를 사용합니다.

   고객 소유 KMS 키로 암호화된 보안 암호에 대한 권한 정책

   JSON

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

   로 암호화된 보안 암호에 대한 권한 정책 AWS 관리형 키

   JSON

```
{
```

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

4. 각 ## ###를 자신의 정보로 바꿉니다.
5. 계속해서 사용자 지정 신뢰 정책을 사용하여 역할 생성(콘솔) 절차에 따라 역할을 생성합니다.

서비스 역할을 생성하여 S3 버킷에서 코드 읽기(PySpark 분석 템플릿 역할)

AWS Clean Rooms 는 PySpark 분석 템플릿을 사용할 때 서비스 역할을 사용하여 공동 작업 구성원의 지정된 S3 버킷에서 코드를 읽습니다.

S3 버킷에서 코드를 읽을 서비스 역할을 생성하려면

1. 사용자 지정 신뢰 정책을 사용하여 역할을 생성합니다. 자세한 내용은 AWS Identity and Access Management 사용 설명서의 사용자 지정 신뢰 정책을 사용하여 역할 생성(콘솔) 절차를 참조하세요.
2. 사용자 지정 신뢰 정책을 사용하여 역할 생성(콘솔) 절차에 따라 다음 사용자 지정 신뢰 정책을 사용합니다.

   JSON

```
{
    "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. 사용자 지정 신뢰 정책을 사용하여 역할 생성(콘솔) 절차에 따라 다음 권한 정책을 사용합니다.

> **Note**
>
> 다음 예제 정책은 Amazon S3에서 코드를 읽는 데 필요한 권한을 지원합니다. 하지만 S3 데이터를 설정한 방법에 따라 이 정책을 수정해야 할 수도 있습니다.
> Amazon S3 리소스는 공동 작업 AWS 리전 과 AWS Clean Rooms 동일한에 있어야 합니다.

JSON

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

```
                }
            }
        }
    ]
}
```

4. 각 *## ###*를 사용자의 정보로 바꿉니다.

   - *s3Path* - 코드의 S3 버킷 위치입니다.
   - *s3BucketOwnerAccountId* - S3 버킷 소유자의 AWS 계정 ID입니다.
   - *##* – AWS 리전의 이름. 예를 들어 **us-east-1**입니다.
   - *jobRunnerAccountId* - 쿼리 및 작업을 실행할 수 있는 구성원의 AWS 계정 ID입니다.
   - *jobRunnerMembershipId* - 작업을 쿼리하고 실행할 수 있는 멤버의 멤버십 ID입니다. 멤버십 ID는 공동 작업의 세부 정보 탭에서 찾을 수 있습니다. 이렇게 하면이 구성원 AWS Clean Rooms 이이 공동 작업에서 분석을 실행할 때만가 역할을 수임할 수 있습니다.
   - *analysisTemplateAccountId* - 분석 템플릿의 AWS 계정 ID입니다.
   - *analysisTemplateOwnerMembershipId* - 분석 템플릿을 소유한 구성원의 멤버십 ID입니다. 멤버십 ID는 공동 작업의 세부 정보 탭에서 찾을 수 있습니다.
5. 계속해서 사용자 지정 신뢰 정책을 사용하여 역할 생성(콘솔) 절차에 따라 역할을 생성합니다.

서비스 역할을 생성하여 PySpark 작업의 결과 작성

AWS Clean Rooms 는 서비스 역할을 사용하여 PySpark 작업의 결과를 지정된 S3 버킷에 기록합니다.

PySpark 작업의 결과를 작성할 서비스 역할을 생성하려면

1. 사용자 지정 신뢰 정책을 사용하여 역할을 생성합니다. 자세한 내용은 AWS Identity and Access Management 사용 설명서의 사용자 지정 신뢰 정책을 사용하여 역할 생성(콘솔) 절차를 참조하세요.
2. 사용자 지정 신뢰 정책을 사용하여 역할 생성(콘솔) 절차에 따라 다음 사용자 지정 신뢰 정책을 사용합니다.

   JSON

```
{
    "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. 사용자 지정 신뢰 정책을 사용하여 역할 생성(콘솔) 절차에 따라 다음 권한 정책을 사용합니다.

> **Note**
>
> 다음 예제 정책은 Amazon S3에 쓰는 데 필요한 권한을 지원합니다. 그러나 S3 설정 방법에 따라이 정책을 수정해야 할 수 있습니다.
> Amazon S3 리소스는 공동 작업 AWS 리전 과 AWS Clean Rooms 동일한에 있어야 합니다.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "s3:PutObject"
            ],
```

```
            "Resource": "arn:aws:s3:::bucket/optionalPrefix/*",
            "Condition":{
                "StringEquals":{
                    "s3:ResourceAccount":[
                        "s3BucketOwnerAccountId"
                    ]
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": [
                "s3:GetBucketLocation",
                "s3:ListBucket"
            ],
            "Resource": "arn:aws:s3:::bucket",
            "Condition":{
                "StringEquals":{
                    "s3:ResourceAccount":[
                        "s3BucketOwnerAccountId"
                    ]
                }
            }
        }
    ]
}
```

4. 각 *## ###*를 자신의 정보로 바꿉니다.

   - *##* – AWS 리전의 이름. 예를 들어 **us-east-1**입니다.
   - *jobRunnerAccountId* - S3 버킷이 위치한 AWS 계정 ID입니다.
   - *jobRunnerMembershipId* - 작업을 쿼리하고 실행할 수 있는 멤버의 멤버십 ID입니다. 멤버십 ID는 공동 작업의 세부 정보 탭에서 찾을 수 있습니다. 이렇게 하면이 구성원 AWS Clean Rooms 이이 공동 작업에서 분석을 실행할 때만가 역할을 수임할 수 있습니다.
   - *rrAccountId* AWS 계정 - S3 버킷이 위치한 ID입니다.
   - *rrMembershipId* - 결과를 받을 수 있는 멤버의 멤버십 ID입니다. 멤버십 ID는 공동 작업의 세부 정보 탭에서 찾을 수 있습니다. 이렇게 하면이 구성원 AWS Clean Rooms 이이 공동 작업에서 분석을 실행할 때만가 역할을 수임할 수 있습니다.
   - *##* - S3 버킷의 이름과 위치입니다.
   - *optionalPrefix* - 결과를 특정 S3 접두사에 저장하려는 경우의 선택적 접두사입니다.

- *s3BucketOwnerAccountId* - S3 버킷 소유자의 AWS 계정 ID입니다.

5. 계속해서 사용자 지정 신뢰 정책을 사용하여 역할 생성(콘솔) 절차에 따라 역할을 생성합니다.

결과를 수신할 서비스 역할 생성

> **Note**
>
> 결과 수신만 가능한 구성원인 경우(콘솔에서 사용자의 멤버 자격이 결과 받기로만 표시됨) 다음 절차를 따르세요.
> 쿼리와 결과 수신이 모두 가능한 구성원인 경우(콘솔에서 사용자의 멤버 자격이 쿼리와 결과 받기 모두 해당) 이 절차를 건너뛸 수 있습니다.

결과만 수신할 수 있는 공동 작업 구성원의 경우는 서비스 역할을 AWS Clean Rooms 사용하여 공동 작업에서 쿼리된 데이터의 결과를 지정된 S3 버킷에 기록합니다.

이 서비스 역할을 만드는 방법은 다음 두 가지입니다.

- 서비스 역할을 생성하는 데 필요한 IAM 권한이 있는 경우 AWS Clean Rooms 콘솔을 사용하여 서비스 역할을 생성합니다.
- `iam:CreateRole`, `iam:CreatePolicy` 및 `iam:AttachRolePolicy` 권한이 없거나 IAM 역할을 수동으로 생성하려는 경우 다음 중 하나를 수행합니다.
  - 다음 절차에 따라 사용자 지정 신뢰 정책을 사용하여 서비스 역할을 생성합니다.
  - 다음 절차에 따라 서비스 역할을 생성하도록 관리자에게 요청합니다.

> **Note**
>
> 사용자 또는 IAM 관리자는 콘솔을 사용하여 AWS Clean Rooms 서비스 역할을 생성하는 데 필요한 권한이 없는 경우에만이 절차를 따라야 합니다.

사용자 지정 신뢰 정책을 사용하여 결과를 수신할 서비스 역할을 생성하려면

1. 사용자 지정 신뢰 정책을 사용하여 역할을 생성합니다. 자세한 내용은 AWS Identity and Access Management 사용 설명서의 사용자 지정 신뢰 정책을 사용하여 역할 생성(콘솔) 절차를 참조하세요.

2. 사용자 지정 신뢰 정책을 사용하여 역할 생성(콘솔) 절차에 따라 다음 사용자 지정 신뢰 정책을 사용합니다.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "AllowIfExternalIdMatches",
            "Effect": "Allow",
            "Principal": {
                "Service": "cleanrooms.amazonaws.com"
            },
            "Action": "sts:AssumeRole",
            "Condition": {
                "ArnLike": {
                    "sts:ExternalId":
 "arn:aws:*:region:*:dbuser:*/a1b2c3d4-5678-90ab-cdef-EXAMPLEaaaaa*"
                }
            }
        },
        {
            "Sid": "AllowIfSourceArnMatches",
            "Effect": "Allow",
            "Principal": {
                "Service": "cleanrooms.amazonaws.com"
            },
            "Action": "sts:AssumeRole",
            "Condition": {
                "ForAnyValue:ArnEquals": {
                    "aws:SourceArn": [
                        "arn:aws:cleanrooms:us-
east-1:555555555555:membership/a1b2c3d4-5678-90ab-cdef-EXAMPLEaaaaa"
                    ]

                }
            }
        }
    ]
}
```

3. 사용자 지정 신뢰 정책을 사용하여 역할 생성(콘솔) 절차에 따라 다음 권한 정책을 사용합니다.

> **Note**
>
> 다음 예제 정책은 AWS Glue 메타데이터와 해당 Amazon S3 데이터를 읽는 데 필요한 권한을 지원합니다. 하지만 S3 데이터를 설정한 방법에 따라 이 정책을 수정해야 할 수도 있습니다.
> AWS Glue 리소스와 기본 Amazon S3 리소스는 AWS Clean Rooms 공동 작업 AWS 리전과 동일한에 있어야 합니다.

JSON

```
{

    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "s3:GetBucketLocation",
                "s3:ListBucket"
            ],
            "Resource": [
                "arn:aws:s3:::bucket_name"
            ],
            "Condition": {
                "StringEquals": {
                    "aws:ResourceAccount":"accountId"
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": [
                "s3:PutObject"
            ],
            "Resource": [
                "arn:aws:s3:::bucket_name/optional_key_prefix/*"
            ],
            "Condition": {
```

```
                "StringEquals": {
                    "aws:ResourceAccount":"accountId"
                }
            }
        }
    ]
}
```

4. 각 *## ###*를 자신의 정보로 바꿉니다.

   - *##* – AWS 리전의 이름. 예를 들어 **us-east-1**입니다.
   - *a1b2c3d4-5678-90ab-cdef-EXAMPLEaaaaa* — 쿼리할 수 있는 구성원의 멤버십 ID입니다. 멤버십 ID는 공동 작업의 세부 정보 탭에서 찾을 수 있습니다. 이렇게 하면이 구성원 AWS Clean Rooms 이이 공동 작업에서 분석을 실행할 때만가 역할을 수임할 수 있습니다.
   - *arn:aws:cleanrooms:us-east-1:555555555555:membership/a1b2c3d4-5678-90ab-cdef-EXAMPLEaaaaa* — 쿼리할 수 있는 구성원의 단일 멤버십 ARN입니다. 멤버십 ARN은 공동 작업의 세부 정보 탭에서 찾을 수 있습니다. 이렇게 하면이 구성원 AWS Clean Rooms 이이 공동 작업에서 분석을 실행할 때만가 역할을 수임할 수 있습니다.
   - *bucket_name* – S3 버킷의 Amazon 리소스 이름(ARN). Amazon 리소스 이름(ARN)은 Amazon S3에 있는 버킷의 속성 탭에서 찾을 수 있습니다.
   - *accountId* - S3 버킷이 위치한 AWS 계정 ID입니다.

     bucket_name/optional_key_prefix – Amazon S3의 결과 대상의 Amazon 리소스 이름(ARN)입니다. Amazon S3 Amazon 리소스 이름(ARN)은 Amazon S3에 있는 버킷의 속성 탭에서 찾을 수 있습니다.

5. 계속해서 사용자 지정 신뢰 정책을 사용하여 역할 생성(콘솔) 절차에 따라 역할을 생성합니다.

AWS Clean Rooms ML에 대한 서비스 역할 설정

유사 모델링을 수행하는 데 필요한 역할은 사용자 지정 모델을 사용하는 데 필요한 역할과 다릅니다. 다음 섹션에서는 각 작업을 수행하는 데 필요한 역할을 설명합니다.

주제

- 유사 모델링을 위한 서비스 역할 설정
- 사용자 지정 모델링을 위한 서비스 역할 설정

유사 모델링을 위한 서비스 역할 설정

주제

- 서비스 역할 생성하여 훈련 데이터 읽기
- 서비스 역할을 생성하여 유사 세그먼트를 작성
- 서비스 역할 생성하여 시드 데이터 읽기

서비스 역할 생성하여 훈련 데이터 읽기

AWS Clean Rooms 는 서비스 역할을 사용하여 훈련 데이터를 읽습니다. 필수 IAM 권한이 있는 경우 콘솔을 사용하여 이 역할을 생성할 수 있습니다. `CreateRole` 권한이 없는 경우 관리자에게 서비스 역할을 생성해 달라고 요청하세요.

서비스 역할을 생성하여 데이터 세트를 훈련하려면

1. 관리자 계정으로 https://console.aws.amazon.com/iam/의 IAM 콘솔에 로그인합니다.
2. 액세스 관리(Access management)에서 정책(Policies)을 선택합니다.
3. 정책 생성을 선택합니다.
4. 정책 편집기에서 JSON 탭을 선택한 다음 다음 정책을 복사하여 붙여넣습니다.

> **Note**
>
> 다음 예제 정책은 AWS Glue 메타데이터와 해당 Amazon S3 데이터를 읽는 데 필요한 권한을 지원합니다. 하지만 S3 데이터를 설정한 방법에 따라 이 정책을 수정해야 할 수도 있습니다. 이 정책에는 데이터를 복호화하기 위한 KMS 키가 포함되어 있지 않습니다. AWS Glue 리소스와 기본 Amazon S3 리소스는 AWS Clean Rooms 공동 작업 AWS 리전과 동일한에 있어야 합니다.

JSON

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

```
                "glue:GetDatabase",
                "glue:GetDatabases",
                "glue:GetTable",
                "glue:GetTables",
                "glue:GetPartitions",
                "glue:GetPartition",
                "glue:BatchGetPartition",
                "glue:GetUserDefinedFunctions"
            ],
            "Resource": [
                "arn:aws:glue: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"
                    ]
                }
            }
        }
    ]
}
```

KMS 키를 사용하여 데이터를 복호화해야 하는 경우 이전 템플릿에 이 AWS KMS 문을 추가하세요.

```
{
            "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. 각 ## ###를 자신의 정보로 바꿉니다.

   - ## – AWS 리전의 이름. 예를 들어 **us-east-1**입니다.

- *accountId* - S3 버킷이 위치한 AWS 계정 ID입니다.
- *######/######*, *table/databases/tables*블, *####* 및 *######/###* -에서 액세스 AWS Clean Rooms 해야 하는 훈련 데이터의 위치입니다.
- *##* - S3 버킷의 Amazon 리소스 이름(ARN)입니다. Amazon 리소스 이름(ARN)은 Amazon S3에 있는 버킷의 속성 탭에서 찾을 수 있습니다.
- *bucketFolders* -가 액세스 AWS Clean Rooms 해야 하는 S3 버킷의 특정 폴더 이름입니다.

6. 다음을 선택합니다.
7. 검토 및 생성에서 정책 이름 및 설명을 입력하고 요약을 검토합니다.
8. 정책 생성을 선택합니다.

   에 대한 정책을 생성했습니다 AWS Clean Rooms.
9. 액세스 관리에서 역할을 선택합니다.

   역할을 사용하면 단기 보안 인증을 만들 수 있으며, 보안 강화를 위해 이 방법을 사용하는 것이 좋습니다. 사용자를 선택하여 장기 보안 인증을 생성할 수도 있습니다.
10. 역할 생성을 선택합니다.
11. 역할 생성 마법사의 신뢰할 수 있는 엔터티 유형에서 사용자 지정 신뢰 정책을 선택합니다.
12. 다음 사용자 지정 신뢰 정책을 복사하여 JSON 편집기에 붙여넣습니다.

    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/*"
                }
            }
        }
    ]
}
```

는 항상 사용자의 SourceAccount입니다 AWS 계정. 특정 훈련 데이터 세트로 SourceArn을 제한할 수 있지만 해당 데이터 세트가 생성된 후에만 가능합니다. 훈련 데이터 세트 ARN을 아직 모르기 때문에 와일드카드가 여기에 지정됩니다.

*accountId*는 훈련 데이터가 AWS 계정 포함된의 ID입니다.

13. 다음을 선택하고 권한 추가에서 방금 생성한 정책의 이름을 입력합니다. (페이지를 새로 고쳐야 할 수 있습니다.)
14. 생성한 정책 이름 옆에 있는 확인란을 선택하고 다음을 선택합니다.
15. 이름 지정, 생성의 경우 역할의 이름과 설명을 입력합니다.

> **Note**
>
> 역할 이름은 결과 및 구성원 역할을 쿼리하고 받을 수 있는 구성원에게 부여된 passRole 권한의 패턴과 일치해야 합니다.

a. 검토: 신뢰할 수 있는 엔티티를 선택하고 필요한 경우 편집합니다.
b. 권한 추가에서 권한을 검토하고 필요한 경우 편집합니다.
c. 태그를 검토하고 필요한 경우 태그를 추가합니다.
d. 역할 생성을 선택합니다.

에 대한 서비스 역할을 생성했습니다 AWS Clean Rooms.

서비스 역할을 생성하여 유사 세그먼트를 작성

AWS Clean Rooms 는 서비스 역할을 사용하여 유사 세그먼트를 버킷에 씁니다. 필수 IAM 권한이 있는 경우 콘솔을 사용하여 이 역할을 생성할 수 있습니다. CreateRole 권한이 없는 경우 관리자에게 서비스 역할을 생성해 달라고 요청하세요.

서비스 역할을 생성하여 유사 세그먼트를 작성하려면

1. 관리자 계정으로 https://console.aws.amazon.com/iam/의 IAM 콘솔에 로그인합니다.
2. 액세스 관리(Access management)에서 정책(Policies)을 선택합니다.
3. 정책 생성을 선택합니다.
4. 정책 편집기에서 JSON 탭을 선택한 다음 다음 정책을 복사하여 붙여넣습니다.

> **Note**
>
> 다음 예제 정책은 AWS Glue 메타데이터와 해당 Amazon S3 데이터를 읽는 데 필요한 권한을 지원합니다. 그러나 Amazon S3 데이터를 설정한 방법에 따라이 정책을 수정해야 할 수 있습니다. 이 정책에는 데이터를 복호화하기 위한 KMS 키가 포함되어 있지 않습니다. AWS Glue 리소스와 기본 Amazon S3 리소스는 AWS Clean Rooms 공동 작업 AWS 리전과 동일한에 있어야 합니다.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
    {
            "Effect": "Allow",
            "Action": [
                "s3:ListBucket",
                "s3:GetBucketLocation"
            ],
            "Resource": [
                "arn:aws:s3:::buckets"
            ],
            "Condition":{
                "StringEquals":{
                    "s3:ResourceAccount":[
                        "accountId"
                    ]
                }
            }
        },
        {
            "Effect": "Allow",
```

```
            "Action": [
                "s3:PutObject"
            ],
            "Resource": [
                "arn:aws:s3:::bucketFolders/*"
            ],
            "Condition":{
                "StringEquals":{
                    "s3:ResourceAccount":[
                        "accountId"
                    ]
                }
            }
        }
  ]
}
```

KMS 키를 사용하여 데이터를 암호화해야 하는 경우 템플릿에 이 AWS KMS 문을 추가하세요.

```
{
            "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. 각 ## ###를 사용자의 정보로 바꿉니다.

- *##* - S3 버킷의 Amazon 리소스 이름(ARN)입니다. Amazon 리소스 이름(ARN)은 Amazon S3에 있는 버킷의 속성 탭에서 찾을 수 있습니다.
- *accountId* - S3 버킷이 위치한 AWS 계정 ID입니다.
- *bucketFolders* -가 액세스 AWS Clean Rooms 해야 하는 S3 버킷의 특정 폴더 이름입니다.
- *##* – AWS 리전의 이름. 예를 들어 **us-east-1**입니다.
- *keyId* - 데이터를 암호화하는 데 필요한 KMS 키입니다.

6. 다음을 선택합니다.
7. 검토 및 생성에서 정책 이름 및 설명을 입력하고 요약을 검토합니다.
8. 정책 생성을 선택합니다.

   에 대한 정책을 생성했습니다 AWS Clean Rooms.
9. 액세스 관리에서 역할을 선택합니다.

   역할을 사용하면 단기 보안 인증을 만들 수 있으며, 보안 강화를 위해 이 방법을 사용하는 것이 좋습니다. 사용자를 선택하여 장기 보안 인증을 생성할 수도 있습니다.
10. 역할 생성을 선택합니다.
11. 역할 생성 마법사의 신뢰할 수 있는 엔터티 유형에서 사용자 지정 신뢰 정책을 선택합니다.
12. 다음 사용자 지정 신뢰 정책을 복사하여 JSON 편집기에 붙여넣습니다.

    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/*"
                }
            }
        }
    ]
}
```

는 항상 사용자의 SourceAccount입니다 AWS 계정. 특정 훈련 데이터 세트로 SourceArn을 제한할 수 있지만 해당 데이터 세트가 생성된 후에만 가능합니다. 훈련 데이터 세트 ARN을 아직 모르기 때문에 와일드카드가 여기에 지정됩니다.

13. 다음을 선택합니다.
14. 생성한 정책 이름 옆에 있는 확인란을 선택하고 다음을 선택합니다.
15. 이름 지정, 생성의 경우 역할의 이름과 설명을 입력합니다.

> **Note**
>
> 역할 이름은 결과 및 구성원 역할을 쿼리하고 받을 수 있는 구성원에게 부여된 passRole 권한의 패턴과 일치해야 합니다.

a. 검토: 신뢰할 수 있는 엔티티를 선택하고 필요한 경우 편집합니다.
b. 권한 추가에서 권한을 검토하고 필요한 경우 편집합니다.
c. 태그를 검토하고 필요한 경우 태그를 추가합니다.
d. 역할 생성을 선택합니다.

에 대한 서비스 역할을 생성했습니다 AWS Clean Rooms.

서비스 역할 생성하여 시드 데이터 읽기

AWS Clean Rooms 는 서비스 역할을 사용하여 시드 데이터를 읽습니다. 필수 IAM 권한이 있는 경우 콘솔을 사용하여 이 역할을 생성할 수 있습니다. CreateRole 권한이 없는 경우 관리자에게 서비스 역할을 생성해 달라고 요청하세요.

서비스 역할을 생성하여 S3 버킷에 저장된 시드 데이터를 읽습니다.

1. 관리자 계정으로 https://console.aws.amazon.com/iam/의 IAM 콘솔에 로그인합니다.
2. 액세스 관리(Access management)에서 정책(Policies)을 선택합니다.
3. 정책 생성을 선택합니다.
4. 정책 편집기에서 JSON 탭을 선택한 후 다음 정책 중 하나를 복사하여 붙여넣습니다.

> **Note**
>
> 다음 예제 정책은 AWS Glue 메타데이터와 해당 Amazon S3 데이터를 읽는 데 필요한 권한을 지원합니다. 그러나 Amazon S3 데이터를 설정한 방법에 따라이 정책을 수정해야 할 수 있습니다. 이 정책에는 데이터를 복호화하기 위한 KMS 키가 포함되어 있지 않습니다. AWS Glue 리소스와 기본 Amazon S3 리소스는 AWS Clean Rooms 공동 작업 AWS 리전과 동일한에 있어야 합니다.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
    {
            "Effect": "Allow",
            "Action": [
                "s3:ListBucket",
            ],
            "Resource": [
                "arn:aws:s3:::buckets"
            ],
            "Condition":{
                "StringEquals":{
                    "s3:ResourceAccount":[
                        "accountId"
                    ]
                }
            }
        },
        {
            "Effect": "Allow",
            "Action": [
```

```
                "s3:GetObject"
            ],
            "Resource": [
                "arn:aws:s3:::bucketFolders/*"
            ],
            "Condition":{
                "StringEquals":{
                    "s3:ResourceAccount":[
                        "accountId"
                    ]
                }
            }
        }
  ]
}
```

> **Note**
>
> 다음 예제 정책은 SQL 쿼리의 결과를 읽고 이를 입력 데이터로 사용하는 데 필요한 권한을 지원합니다. 그러나 쿼리 구조에 따라 이 정책을 수정해야 할 수 있습니다. 이 정책에는 데이터를 복호화하기 위한 KMS 키가 포함되어 있지 않습니다.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "AllowCleanRoomsStartQuery",
            "Effect": "Allow",
            "Action": [
                "cleanrooms:GetCollaborationAnalysisTemplate",
                "cleanrooms:GetSchema",
                "cleanrooms:StartProtectedQuery"
            ],
            "Resource": "*"
        },
        {
            "Sid": "AllowCleanRoomsGetAndUpdateQuery",
```

```
            "Effect": "Allow",
            "Action": [
                "cleanrooms:GetProtectedQuery",
                "cleanrooms:UpdateProtectedQuery"
            ],
            "Resource": [

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

KMS 키를 사용하여 데이터를 복호화해야 하는 경우 템플릿에 이 AWS KMS 문을 추가하세요.

```
{
            "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. 각 *## ###*를 사용자의 정보로 바꿉니다.

   - *##* - S3 버킷의 Amazon 리소스 이름(ARN)입니다. Amazon 리소스 이름(ARN)은 Amazon S3에 있는 버킷의 속성 탭에서 찾을 수 있습니다.
   - *accountId* - S3 버킷이 위치한 AWS 계정 ID입니다.
   - *bucketFolders* -가 액세스 AWS Clean Rooms 해야 하는 S3 버킷의 특정 폴더 이름입니다.

- *##* – AWS 리전의 이름. 예를 들어 **us-east-1**입니다.
- *queryRunnerAccountId* - 쿼리를 실행할 계정의 AWS 계정 ID입니다.
- *queryRunnerMembershipId* - 쿼리할 수 있는 멤버의 멤버십 ID입니다. 멤버십 ID는 공동 작업의 세부 정보 탭에서 찾을 수 있습니다. 이렇게 하면이 구성원 AWS Clean Rooms 이이 공동 작업에서 분석을 실행할 때만가 역할을 수임할 수 있습니다.
- *keyId* - 데이터를 암호화하는 데 필요한 KMS 키입니다.

6. 다음을 선택합니다.
7. 검토 및 생성에서 정책 이름 및 설명을 입력하고 요약을 검토합니다.
8. 정책 생성을 선택합니다.

   에 대한 정책을 생성했습니다 AWS Clean Rooms.
9. 액세스 관리에서 역할을 선택합니다.

   역할을 사용하면 단기 보안 인증을 만들 수 있으며, 보안 강화를 위해 이 방법을 사용하는 것이 좋습니다. 사용자를 선택하여 장기 보안 인증을 생성할 수도 있습니다.
10. 역할 생성을 선택합니다.
11. 역할 생성 마법사의 신뢰할 수 있는 엔터티 유형에서 사용자 지정 신뢰 정책을 선택합니다.
12. 다음 사용자 지정 신뢰 정책을 복사하여 JSON 편집기에 붙여넣습니다.

    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/*"
                }
            }
        }
    ]
}
```

는 항상 사용자의 SourceAccount입니다 AWS 계정. 특정 훈련 데이터 세트로 SourceArn을 제한할 수 있지만 해당 데이터 세트가 생성된 후에만 가능합니다. 훈련 데이터 세트 ARN을 아직 모르기 때문에 와일드카드가 여기에 지정됩니다.

13. 다음을 선택합니다.
14. 생성한 정책 이름 옆에 있는 확인란을 선택하고 다음을 선택합니다.
15. 이름 지정, 생성의 경우 역할의 이름과 설명을 입력합니다.

> **Note**
>
> 역할 이름은 결과 및 구성원 역할을 쿼리하고 받을 수 있는 구성원에게 부여된 passRole 권한의 패턴과 일치해야 합니다.

a. 검토: 신뢰할 수 있는 엔티티를 선택하고 필요한 경우 편집합니다.
b. 권한 추가에서 권한을 검토하고 필요한 경우 편집합니다.
c. 태그를 검토하고 필요한 경우 태그를 추가합니다.
d. 역할 생성을 선택합니다.

에 대한 서비스 역할을 생성했습니다 AWS Clean Rooms.

사용자 지정 모델링을 위한 서비스 역할 설정

주제

- 사용자 지정 ML 모델링을 위한 서비스 역할 생성 - ML 구성
- 서비스 역할을 생성하여 사용자 지정 ML 모델 제공
- 데이터 세트를 쿼리할 서비스 역할 생성

- 서비스 역할을 생성하여 구성된 테이블 연결 생성

사용자 지정 ML 모델링을 위한 서비스 역할 생성 - ML 구성

AWS Clean Rooms 는 서비스 역할을 사용하여 사용자 지정 ML 구성을 생성할 수 있는 사용자를 제어합니다. 필수 IAM 권한이 있는 경우 콘솔을 사용하여 이 역할을 생성할 수 있습니다. `CreateRole` 권한이 없는 경우 관리자에게 서비스 역할을 생성해 달라고 요청하세요.

이 역할을 사용하면 PutMLConfiguration 작업을 사용할 수 있습니다.

사용자 지정 ML 구성 생성을 허용하는 서비스 역할을 생성하려면

1. 관리자 계정으로 https://console.aws.amazon.com/iam/의 IAM 콘솔에 로그인합니다.
2. 액세스 관리(Access management)에서 정책(Policies)을 선택합니다.
3. 정책 생성을 선택합니다.
4. 정책 편집기에서 JSON 탭을 선택한 다음 다음 정책을 복사하여 붙여넣습니다.

> **Note**
>
> 다음 예제 정책은 S3 버킷에 액세스하고 데이터를 쓰고 CloudWatch 지표를 게시하는 데 필요한 권한을 지원합니다. 그러나 Amazon S3 데이터를 설정한 방법에 따라이 정책을 수정해야 할 수 있습니다. 이 정책에는 데이터를 복호화하기 위한 KMS 키가 포함되어 있지 않습니다.
>
> Amazon S3 리소스는 AWS Clean Rooms 공동 작업 AWS 리전 과 동일한에 있어야 합니다.

JSON

```
{
    "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. 각 *## ###*를 사용자의 정보로 바꿉니다.

   - *##* - S3 버킷의 Amazon 리소스 이름(ARN)입니다. Amazon 리소스 이름(ARN)은 Amazon S3에 있는 버킷의 속성 탭에서 찾을 수 있습니다.
   - *##* – AWS 리전의 이름. 예를 들어 **us-east-1**입니다.
   - *accountId* — S3 버킷이 위치한 AWS 계정 ID입니다.
   - *keyId* - 데이터를 암호화하는 데 필요한 KMS 키입니다.
6. 다음을 선택합니다.
7. 검토 및 생성에서 정책 이름 및 설명을 입력하고 요약을 검토합니다.
8. 정책 생성을 선택합니다.

   에 대한 정책을 생성했습니다 AWS Clean Rooms.
9. 액세스 관리에서 역할을 선택합니다.

   역할을 사용하면 단기 보안 인증을 만들 수 있으며, 보안 강화를 위해 이 방법을 사용하는 것이 좋습니다. 사용자를 선택하여 장기 보안 인증을 생성할 수도 있습니다.
10. 역할 생성을 선택합니다.
11. 역할 생성 마법사의 신뢰할 수 있는 엔터티 유형에서 사용자 지정 신뢰 정책을 선택합니다.
12. 다음 사용자 지정 신뢰 정책을 복사하여 JSON 편집기에 붙여넣습니다.

    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"
                }
            }
        }
    ]
}
```

는 항상 사용자의 SourceAccount입니다 AWS 계정. 특정 훈련 데이터 세트로 SourceArn을 제한할 수 있지만 해당 데이터 세트가 생성된 후에만 가능합니다. 훈련 데이터 세트 ARN을 아직 모르기 때문에 와일드카드가 여기에 지정됩니다.

13. 다음을 선택합니다.
14. 생성한 정책 이름 옆에 있는 확인란을 선택하고 다음을 선택합니다.
15. 이름 지정, 생성의 경우 역할의 이름과 설명을 입력합니다.

> **Note**
>
> 역할 이름은 결과 및 구성원 역할을 쿼리하고 받을 수 있는 구성원에게 부여된 passRole 권한의 패턴과 일치해야 합니다.

a. 검토: 신뢰할 수 있는 엔티티를 선택하고 필요한 경우 편집합니다.
b. 권한 추가에서 권한을 검토하고 필요한 경우 편집합니다.
c. 태그를 검토하고 필요한 경우 태그를 추가합니다.
d. 역할 생성을 선택합니다.

에 대한 서비스 역할을 생성했습니다 AWS Clean Rooms.

서비스 역할을 생성하여 사용자 지정 ML 모델 제공

AWS Clean Rooms 는 서비스 역할을 사용하여 사용자 지정 ML 모델 알고리즘을 생성할 수 있는 사용자를 제어합니다. 필수 IAM 권한이 있는 경우 콘솔을 사용하여 이 역할을 생성할 수 있습니다. `CreateRole` 권한이 없는 경우 관리자에게 서비스 역할을 생성해 달라고 요청하세요.

이 역할을 사용하면 CreateConfiguredModelAlgorithm 작업을 사용할 수 있습니다.

멤버가 사용자 지정 ML 모델을 제공하도록 허용하는 서비스 역할을 생성하려면

1. 관리자 계정으로 https://console.aws.amazon.com/iam/의 IAM 콘솔에 로그인합니다.
2. 액세스 관리(Access management)에서 정책(Policies)을 선택합니다.
3. 정책 생성을 선택합니다.
4. 정책 편집기에서 JSON 탭을 선택한 다음 다음 정책을 복사하여 붙여넣습니다.

> **Note**
>
> 다음 예제 정책은 모델 알고리즘이 포함된 Docker 이미지를 검색하는 데 필요한 권한을 지원합니다. 그러나 Amazon S3 데이터를 설정한 방법에 따라이 정책을 수정해야 할 수 있습니다. 이 정책에는 데이터를 복호화하기 위한 KMS 키가 포함되어 있지 않습니다. Amazon S3 리소스는 AWS Clean Rooms 공동 작업 AWS 리전 과 동일한에 있어야 합니다.

JSON

```
{
    "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. 각 *## ###*를 자신의 정보로 바꿉니다.

   - *##* – AWS 리전의 이름. 예를 들어 **us-east-1**입니다.
   - *accountId* — S3 버킷이 위치한 AWS 계정 ID입니다.
   - *repoName* - 데이터가 포함된 리포지토리의 이름입니다.
6. 다음을 선택합니다.
7. 검토 및 생성에서 정책 이름 및 설명을 입력하고 요약을 검토합니다.
8. 정책 생성을 선택합니다.

   에 대한 정책을 생성했습니다 AWS Clean Rooms.
9. 액세스 관리에서 역할을 선택합니다.

   역할을 사용하면 단기 보안 인증을 만들 수 있으며, 보안 강화를 위해 이 방법을 사용하는 것이 좋습니다. 사용자를 선택하여 장기 보안 인증을 생성할 수도 있습니다.
10. 역할 생성을 선택합니다.
11. 역할 생성 마법사의 신뢰할 수 있는 엔터티 유형에서 사용자 지정 신뢰 정책을 선택합니다.
12. 다음 사용자 지정 신뢰 정책을 복사하여 JSON 편집기에 붙여넣습니다.

    JSON

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

`SourceAccount`는 항상 사용자의 입니다 AWS 계정 .는 특정 훈련 데이터 세트로 제한될 `SourceArn` 수 있지만 해당 데이터 세트가 생성된 후에만 가능합니다. 훈련 데이터 세트 ARN을 아직 모르기 때문에 와일드카드가 여기에 지정됩니다.

13. 다음을 선택합니다.
14. 생성한 정책 이름 옆에 있는 확인란을 선택하고 다음을 선택합니다.
15. 이름 지정, 생성의 경우 역할의 이름과 설명을 입력합니다.

> **Note**
>
> 역할 이름은 결과 및 구성원 역할을 쿼리하고 받을 수 있는 구성원에게 부여된 `passRole` 권한의 패턴과 일치해야 합니다.

a. 검토: 신뢰할 수 있는 엔티티를 선택하고 필요한 경우 편집합니다.
b. 권한 추가에서 권한을 검토하고 필요한 경우 편집합니다.
c. 태그를 검토하고 필요한 경우 태그를 추가합니다.
d. 역할 생성을 선택합니다.

에 대한 서비스 역할을 생성했습니다 AWS Clean Rooms.

데이터 세트를 쿼리할 서비스 역할 생성

AWS Clean Rooms 는 서비스 역할을 사용하여 사용자 지정 ML 모델링에 사용할 데이터 세트를 쿼리할 수 있는 사용자를 제어합니다. 필수 IAM 권한이 있는 경우 콘솔을 사용하여 이 역할을 생성할 수 있습니다. `CreateRole` 권한이 없는 경우 관리자에게 서비스 역할을 생성해 달라고 요청하세요.

이 역할을 사용하면 CreateMLInputChannel 작업을 사용할 수 있습니다.

멤버가 데이터 세트를 쿼리할 수 있도록 서비스 역할을 생성하려면

1. 관리자 계정으로 https://console.aws.amazon.com/iam/의 IAM 콘솔에 로그인합니다.
2. 액세스 관리(Access management)에서 정책(Policies)을 선택합니다.
3. 정책 생성을 선택합니다.
4. 정책 편집기에서 JSON 탭을 선택한 다음 다음 정책을 복사하여 붙여넣습니다.

> **Note**
>
> 다음 예제 정책은 사용자 지정 ML 모델링에 사용할 데이터 세트를 쿼리하는 데 필요한 권한을 지원합니다. 그러나 Amazon S3 데이터를 설정한 방법에 따라이 정책을 수정해야 할 수 있습니다. 이 정책에는 데이터를 복호화하기 위한 KMS 키가 포함되어 있지 않습니다. Amazon S3 리소스는 AWS Clean Rooms 공동 작업 AWS 리전 과 동일한에 있어야 합니다.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "AllowCleanRoomsStartQueryForMLInputChannel",
            "Effect": "Allow",
            "Action": "cleanrooms:StartProtectedQuery",
            "Resource": "*"
        },
        {
            "Sid":
 "AllowCleanroomsGetSchemaAndGetAnalysisTemplateForMLInputChannel",
            "Effect": "Allow",
            "Action": [
                "cleanrooms:GetSchema",
                "cleanrooms:GetCollaborationAnalysisTemplate"
            ],
            "Resource": "*"
        },
        {
            "Sid": "AllowCleanRoomsGetAndUpdateQueryForMLInputChannel",
            "Effect": "Allow",
            "Action": [
                "cleanrooms:GetProtectedQuery",
                "cleanrooms:UpdateProtectedQuery"
            ],
            "Resource": [
```

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

5. 각 *## ###*를 자신의 정보로 바꿉니다.

   - *##* – AWS 리전의 이름. 예를 들어 **us-east-1**입니다.
   - *queryRunnerAccountId* - 쿼리를 실행할 계정의 AWS 계정 ID입니다.
   - *queryRunnerMembershipId* - 쿼리할 수 있는 멤버의 멤버십 ID입니다. 멤버십 ID는 공동 작업의 세부 정보 탭에서 찾을 수 있습니다. 이렇게 하면이 구성원 AWS Clean Rooms 이이 공동 작업에서 분석을 실행할 때만가 역할을 수임할 수 있습니다.
6. 다음을 선택합니다.
7. 검토 및 생성에서 정책 이름 및 설명을 입력하고 요약을 검토합니다.
8. 정책 생성을 선택합니다.

   에 대한 정책을 생성했습니다 AWS Clean Rooms.
9. 액세스 관리에서 역할을 선택합니다.

   역할을 사용하면 단기 보안 인증을 만들 수 있으며, 보안 강화를 위해 이 방법을 사용하는 것이 좋습니다. 사용자를 선택하여 장기 보안 인증을 생성할 수도 있습니다.
10. 역할 생성을 선택합니다.
11. 역할 생성 마법사의 신뢰할 수 있는 엔터티 유형에서 사용자 지정 신뢰 정책을 선택합니다.
12. 다음 사용자 지정 신뢰 정책을 복사하여 JSON 편집기에 붙여넣습니다.

    JSON

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

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

SourceAccount는 항상 사용자의 입니다 AWS 계정 .는 특정 훈련 데이터 세트로 제한될 SourceArn 수 있지만 해당 데이터 세트가 생성된 후에만 가능합니다. 훈련 데이터 세트 ARN을 아직 모르기 때문에 와일드카드가 여기에 지정됩니다.

13. 다음을 선택합니다.
14. 생성한 정책 이름 옆에 있는 확인란을 선택하고 다음을 선택합니다.
15. 이름 지정, 생성의 경우 역할의 이름과 설명을 입력합니다.

> **Note**
>
> 역할 이름은 결과 및 구성원 역할을 쿼리하고 받을 수 있는 구성원에게 부여된 passRole 권한의 패턴과 일치해야 합니다.

a. 검토: 신뢰할 수 있는 엔티티를 선택하고 필요한 경우 편집합니다.
b. 권한 추가에서 권한을 검토하고 필요한 경우 편집합니다.
c. 태그를 검토하고 필요한 경우 태그를 추가합니다.
d. 역할 생성을 선택합니다.

에 대한 서비스 역할을 생성했습니다 AWS Clean Rooms.

서비스 역할을 생성하여 구성된 테이블 연결 생성

AWS Clean Rooms 는 서비스 역할을 사용하여 구성된 테이블 연결을 생성할 수 있는 사용자를 제어합니다. 필수 IAM 권한이 있는 경우 콘솔을 사용하여 이 역할을 생성할 수 있습니다. CreateRole 권한이 없는 경우 관리자에게 서비스 역할을 생성해 달라고 요청하세요.

이 역할을 사용하면 CreateConfiguredTableAssociation 작업을 사용할 수 있습니다.

구성된 테이블 연결을 생성할 수 있도록 서비스 역할을 생성하려면

1. 관리자 계정으로 https://console.aws.amazon.com/iam/의 IAM 콘솔에 로그인합니다.
2. 액세스 관리(Access management)에서 정책(Policies)을 선택합니다.

3. 정책 생성을 선택합니다.
4. 정책 편집기에서 JSON 탭을 선택한 다음 다음 정책을 복사하여 붙여넣습니다.

> **Note**
>
> 다음 예제 정책은 구성된 테이블 연결 생성을 지원합니다. 그러나 Amazon S3 데이터를 설정한 방법에 따라이 정책을 수정해야 할 수 있습니다. 이 정책에는 데이터를 복호화하기 위한 KMS 키가 포함되어 있지 않습니다.
> Amazon S3 리소스는 AWS Clean Rooms 공동 작업 AWS 리전 과 동일한에 있어야 합니다.

JSON

```
{
    "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. 각 *## ###*를 사용자의 정보로 바꿉니다.

   - *Amazon S3 #### ##### # #### KMS #* - Amazon S3 데이터를 암호화하는 데 사용된 KMS 키입니다. 데이터를 복호화하려면 데이터를 암호화하는 데 사용된 것과 동일한 KMS 키를 제공해야 합니다.
   - *AWS Glue #### Amazon S3 ##* - 데이터가 포함된 AWS Glue 테이블이 포함된 Amazon S3 버킷의 이름입니다.
   - *##* – AWS 리전의 이름. 예를 들어 **us-east-1**입니다.
   - *accountId* - 데이터를 소유한 계정의 AWS 계정 ID입니다.
   - *AWS Glue ###### ##* - 데이터가 포함된 AWS Glue 데이터베이스의 이름입니다.
   - *AWS Glue ### ##* - 데이터가 포함된 AWS Glue 테이블의 이름입니다.

6. 다음을 선택합니다.
7. 검토 및 생성에서 정책 이름 및 설명을 입력하고 요약을 검토합니다.

8. 정책 생성을 선택합니다.

   에 대한 정책을 생성했습니다 AWS Clean Rooms.

9. 액세스 관리에서 역할을 선택합니다.

   역할을 사용하면 단기 보안 인증을 만들 수 있으며, 보안 강화를 위해 이 방법을 사용하는 것이 좋습니다. 사용자를 선택하여 장기 보안 인증을 생성할 수도 있습니다.

10. 역할 생성을 선택합니다.
11. 역할 생성 마법사의 신뢰할 수 있는 엔터티 유형에서 사용자 지정 신뢰 정책을 선택합니다.
12. 다음 사용자 지정 신뢰 정책을 복사하여 JSON 편집기에 붙여넣습니다.

    JSON

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

    SourceAccount는 항상 사용자의 입니다 AWS 계정 .는 특정 훈련 데이터 세트로 제한될 SourceArn 수 있지만 해당 데이터 세트가 생성된 후에만 가능합니다. 훈련 데이터 세트 ARN을 아직 모르기 때문에 와일드카드가 여기에 지정됩니다.

13. 다음을 선택합니다.
14. 생성한 정책 이름 옆에 있는 확인란을 선택하고 다음을 선택합니다.
15. 이름 지정, 생성의 경우 역할의 이름과 설명을 입력합니다.

> **Note**
>
> 역할 이름은 결과 및 구성원 역할을 쿼리하고 받을 수 있는 구성원에게 부여된 `passRole` 권한의 패턴과 일치해야 합니다.

a. 검토: 신뢰할 수 있는 엔티티를 선택하고 필요한 경우 편집합니다.
b. 권한 추가에서 권한을 검토하고 필요한 경우 편집합니다.
c. 태그를 검토하고 필요한 경우 태그를 추가합니다.
d. 역할 생성을 선택합니다.

에 대한 서비스 역할을 생성했습니다 AWS Clean Rooms.

AWS Clean Rooms의 공동 작업 및 멤버십

공동 작업은 구성원이 구성된 테이블에 대한 분석을 수행할 수 있는 안전한 논리적 경계 AWS Clean Rooms 입니다.

의 모든 멤버는 공동 작업을 생성할 AWS Clean Rooms 수 있습니다.

공동 작업 생성자는 구성된 테이블을 분석하고 결과를 수신할 단일 멤버를 지정할 수 있습니다. 그러나 공동 작업 생성자는 분석을 실행할 수 있는 구성원이 쿼리 결과에 액세스하지 못하도록 하려고 할 수 있습니다. 이 경우 공동 작업 생성자는 쿼리할 수 있는 구성원 또는 쿼리 및 작업을 실행할 수 있는 구성원과 결과를 받을 수 있는 다른 구성원을 지정할 수 있습니다.

대부분의 경우 쿼리할 수 있는 구성원 또는 쿼리하고 작업을 실행할 수 있는 구성원도 컴퓨팅 비용을 지불하는 구성원입니다. 하지만 공동 작업 생성자는 쿼리 컴퓨팅 비용을 지불하도록 다른 구성원을 구성할 수 있습니다.

AWS SDKs. AWS Clean Rooms

주제

- 공동 작업 생성
- 멤버십 생성 및 공동 작업 참여
- 공동 작업 편집
- 공동 작업 삭제
- 공동 작업 보기
- 공동 작업에 멤버 초대
- 구성원 모니터링
- 공동 작업에서 구성원 제거
- 공동 작업 탈퇴

공동 작업 생성

에서 공동 작업을 생성하는 세 가지 방법이 있습니다 AWS Clean Rooms.

가장 기본적인 형식은 쿼리에 대한 공동 작업입니다. 이 공동 작업은 SQL 쿼리 분석에 중점을 두고 두 가지 주요 역할, 즉 쿼리를 실행할 수 있는 구성원과 결과를 받을 수 있는 다른 역할로 구성된 간단한 구조를 유지합니다. 이 기본 공동 작업 설정은 간단한 데이터 분석 작업에 적합합니다.

두 번째 형식인 쿼리 및 작업에 대한 공동 작업은 SQL 쿼리와 PySpark 작업을 모두 통합하여 기능을 확장하며 Spark가 분석 엔진으로 필요합니다. 이 공동 작업 설정은 동일한 기본 역할 구조를 유지하지만 작업 실행을 포함하도록 권한을 확장합니다. PySpark 분석 템플릿을 생성하는 구성원도 결과를 받는 구성원이어야 하므로 분석 프로세스에서 명확한 책임을 져야 합니다.

세 번째 형식인 ML 모델링을 위한 협업은 기계 학습 워크플로를 위해 구축되었으며 분석 엔진으로 Spark가 필요합니다. 이 공동 작업 설정에는 두 가지 역할이 추가됩니다. 하나는 훈련된 모델의 결과가 필요한 사용자용이고 다른 하나는 해당 모델을 사용하여 예측한 결과가 필요한 사용자용입니다. 이 공동 작업 설정은 공동 작업 구성원이 복잡한 데이터 프로젝트에서 함께 작업하는 동시에 모든 사람의 역할과 권한을 명확하게 유지하는 데 도움이 됩니다.

다음 주제에서는 쿼리, 작업 및 ML 모델링에 대한 공동 작업을 생성하는 방법을 설명합니다.

주제

- 쿼리에 대한 공동 작업 생성
- 쿼리 및 작업에 대한 공동 작업 생성
- ML 모델링을 위한 공동 작업 생성

쿼리에 대한 공동 작업 생성

이 절차에서는 공동 작업 생성자가 다음 작업을 수행합니다.

- 공동 작업을 생성합니다.
- 공동 작업에 한 명 이상의 구성원을 초대합니다.
- 쿼리할 수 있는 구성원 및 결과를 받을 수 있는 구성원과 같은 구성원에게 기능을 할당합니다.

  공동 작업 생성자가 결과를 받을 수 있는 구성원이기도 한 경우 결과 대상과 형식을 지정합니다. 또한 서비스 역할 Amazon 리소스 이름(ARN)을 제공하여 결과를 결과 대상에 기록합니다.
- 공동 작업에서 컴퓨팅 비용을 지불할 책임이 있는 구성원을 구성합니다.

시작하기 전에 다음 사전 조건을 완료했는지 확인합니다.

-
- 공동 작업에 초대하려는 각 멤버의 이름과 AWS 계정 ID가 있습니다.
- 각 멤버의 이름과 AWS 계정 ID를 공동 작업의 모든 멤버와 공유할 수 있는 권한이 있습니다.

> **Note**
>
> 공동 작업을 생성한 후에는 멤버를 더 추가할 수 없습니다.

AWS SDKs. AWS Clean Rooms

쿼리에 대한 공동 작업을 생성하려면

1. 에 로그인 AWS Management Console 하고 공동 작업 생성자 역할을 AWS 계정 할를 사용하여 AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 오른쪽 상단 모서리에서 공동 작업 생성을 선택합니다.
4. 1단계: 공동 작업 정의의 경우 다음을 수행합니다:

   a. 세부 정보에는 공동 작업의 이름 및 설명을 입력합니다.

      이 정보는 공동 작업에 참여하도록 초대받은 공동 작업 구성원에게 표시됩니다. 이름과 설명은 공동 작업이 무엇을 의미하는지 이해하는 데 도움이 됩니다.

   b. 분석 엔진에서 Spark를 선택합니다.

      > **Note**
      >
      > AWS Clean Rooms 는 2025년 12월 17일에 레거시 Clean Rooms SQL 분석 엔진에 대한 지원을 종료합니다. 2025년 7월 16일 이전에 AWS 는 고객 지원을 통해 한도 증가를 요청하여 새로운 Clean Rooms SQL 엔진 기반 공동 작업을 생성해야 합니다. 2025년 7월 17일부터 새로운 Clean Rooms SQL 엔진 기반 공동 작업을 더 이상 생성할 수 없습니다.
      > AWS Clean Rooms SQL 분석 엔진에서 Spark 분석 엔진으로 공동 작업을 업데이트하려면 기존 공동 작업을 편집하거나 공동 작업을 다시 생성하고 Spark 분석 엔진을 선택합니다.

   c. 구성원의 경우:

      i. 구성원 1: 본인, 공동 작업에 표시할 구성원 표시명을 원하는 대로 입력합니다.

> **Note**
>
> AWS 계정 ID는 멤버 AWS 계정 ID에 자동으로 포함됩니다.

ii. 멤버 2에 공동 작업에 초대하려는 멤버의 멤버 표시 이름과 멤버 AWS 계정 ID를 입력합니다.

구성원 표시 이름과 구성원 AWS 계정 ID는 공동 작업에 초대된 모든 사람이 볼 수 있습니다. 이러한 필드의 값을 입력하고 저장한 후에는 편집할 수 없습니다.

> **Note**
>
> 공동 작업에 초대되거나 활동 중인 모든 협업자가 구성원 AWS 계정 ID와 구성원 표시명을 볼 수 있다는 사실을 공동 작업 구성원에게 알려야 합니다.

iii. 다른 구성원을 추가하려면 다른 구성원 추가를 선택합니다. 그런 다음 공동 작업에 초대하려는 데이터를 제공할 수 있는 각 멤버의 멤버 표시 이름과 멤버 AWS 계정 ID를 입력합니다.

d. 분석 로깅을 활성화하려면 분석 로깅 활성화 확인란을 선택합니다.

- 지원되는 로그 유형에서 쿼리의 로그 확인란을 선택합니다.

  Amazon CloudWatch Logs 계정의 SQL 쿼리에서 생성된 로그를 받게 됩니다.

e. (선택 사항) 암호화 컴퓨팅 기능을 활성화하려면 암호화 컴퓨팅 활성화 확인란을 선택합니다.

i. 다음 암호화 적용 범위 파라미터를 선택합니다.

- plaintext 열 허용

  완전히 암호화된 테이블이 필요한 경우 아니요를 선택합니다.

  암호화된 테이블에 cleartext 열을 허용하려면 예를 선택합니다.

  특정 열에서 SUM 또는 AVG을(를) 실행하려면 해당 열이 실행하려면 cleartext 안에 있어야 합니다.

- NULL 값 보존

  NULL 값을 보존하지 않으려면 아니오를 선택합니다. NULL 값은 암호화된 테이블에서 NULL처럼 표시되지 않습니다.

NULL 값을 보존하려면 예를 선택합니다. NULL 값은 암호화된 테이블에서와 NULL같이 표시됩니다.

ii. 다음 지문 파라미터를 선택합니다.

- 중복 허용

  fingerprint 열에 중복 항목을 허용하지 않으려면 아니오를 선택합니다.

  fingerprint 열에 중복 항목을 허용하려면 예를 선택합니다.

- 이름이 다른 열의 JOIN 허용

  이름이 다른 fingerprint 열에 가입하지 않으려면 아니오를 선택합니다.

  이름이 다른 fingerprint 열에 참여하려면 예를 선택합니다.

암호화 컴퓨팅 매개변수에 대한 자세한 내용은 암호화 컴퓨팅 파라미터을 참조하세요.

에서 사용할 데이터를 암호화하는 방법에 대한 자세한 내용은 섹션을 AWS Clean Rooms참조하세요Clean Rooms에 대한 암호화 컴퓨팅으로 암호화된 데이터 테이블 준비.

> **Note**
>
> 다음 단계를 완료하기 전에 이러한 구성을 주의 깊게 확인하세요. 공동 작업을 생성한 후에는 공동 작업 이름, 설명 및 로그가 Amazon CloudWatch Logs에 저장되는지 여부만 편집할 수 있습니다.

f. 공동 작업 리소스에 태그를 활성화하려면 새 태그 추가를 선택한 다음 키와 값 쌍을 입력합니다.

g. 다음을 선택합니다.

5. 2단계: 멤버 기능 지정, 쿼리 및 작업을 사용한 분석의 경우 지원되는 분석 유형에서 쿼리 확인란을 선택한 상태로 두고 목표에 따라 권장 조치를 취합니다.

목표	권장 조치
공동 작업의 데이터를 쿼리하고 결과를 받으세요	1. 쿼리를 실행할 수 있는 구성원으로 자신을 선택하세요.

목표	권장 조치
	2. 드롭다운 목록에서 분석 결과를 받을 수 있는 멤버로 자신을 선택합니다.
공동 작업의 데이터를 쿼리하고 결과를 받을 다른 구성원을 배정합니다	1. 쿼리를 실행할 수 있는 구성원으로 자신을 선택하세요. 2. 드롭다운 목록에서 분석 결과를 받을 수 있는 멤버를 선택합니다.
공동 작업에서 쿼리 결과를 수신하고 데이터를 쿼리할 다른 구성원을 지정합니다	1. 드롭다운 목록에서 쿼리를 실행할 수 있는 구성원을 선택합니다. 2. 드롭다운 목록에서 분석 결과를 받을 수 있는 멤버로 자신을 선택합니다.
공동 작업을 생성 및 관리하고, 데이터를 쿼리할 다른 구성원을 지정하고, 결과를 받을 다른 구성원을 배정합니다	1. 드롭다운 목록에서 쿼리를 실행할 수 있는 구성원을 선택합니다. 2. 드롭다운 목록에서 분석 결과를 받을 수 있는 멤버를 선택합니다.

a. Clean Rooms ML을 사용하는 경우 특별히 구축된 워크플로를 사용한 ML 모델링의 경우

   i. (선택 사항) 드롭다운 목록에서 훈련된 모델에서 출력을 수신할 수 있는 멤버를 선택합니다.

   ii. (선택 사항) 드롭다운 목록에서 모델 추론에서 출력을 수신할 수 있는 멤버를 선택합니다.

b. 를 사용하여 ID 확인 AWS Entity Resolution에서 멤버 기능을 봅니다.

c. 다음을 선택합니다.

6. 3단계: 결제 구성의 경우 쿼리를 사용한 분석에서 목표에 따라 다음 작업 중 하나를 수행합니다.

목표	권장 조치
쿼리를 실행할 수 있는 구성원을 쿼리 계산 비용을 지불하는 구성원으로 배정합니다	1. 쿼리를 사용한 분석에서 쿼리 비용을 지불할 멤버를 쿼리를 실행할 수 있는 멤버와 동일하게 선택합니다. 2. 다음을 선택합니다.

목표	권장 조치
쿼리 컴퓨팅 비용을 지불할 다른 구성원을 배정하세요	1. 쿼리를 사용한 분석에서 쿼리 비용을 지불할 멤버로 자신을 선택합니다. 2. 다음을 선택합니다.

특별히 구축된 워크플로를 사용한 ML 모델링의 경우 구성된 유사 모델의 생성자는 유사 모델링 비용을 지불할 구성원입니다.

를 사용한 ID 확인 AWS Entity Resolution의 경우 ID 매핑 테이블의 생성자는 ID 매핑 테이블 비용을 지불할 멤버입니다.

7. 4단계: 멤버십 구성에서 다음 옵션 중 하나를 선택합니다.

   Yes, join by creating membership now

   1. 결과 설정 기본값의 경우 쿼리 결과 설정의 경우 결과를 받을 수 있는 멤버인 경우
      a. Amazon S3의 결과 대상에 Amazon S3 대상을 입력하거나 S3 찾아보기를 선택하여 S3 버킷을 선택합니다.
      b. 쿼리 결과 형식의 경우 CSV 또는 PARQUET 중 하나를 선택합니다.
      c. (Spark만 해당) 결과 파일에서 다중 또는 단일을 선택합니다.
      d. (선택 사항) 서비스 액세스에서 S3 대상에 최대 24시간이 걸리는 쿼리를 전송하려면 완료에 최대 24시간이 걸리는 쿼리를 지원하는 서비스 역할 추가 확인란을 선택합니다.

         완료하는 데 최대 24시간이 걸리는 대규모 쿼리는 S3 대상으로 전송됩니다.

         확인란을 선택하지 않으면 12시간 이내에 완료된 쿼리만 S3 위치로 전송됩니다.
      e. 새 서비스 역할 생성 및 사용 또는 기존 서비스 역할 사용을 선택하여 서비스 액세스 권한을 지정합니다.

를 선택하는 경우...	THEN ...
새 서비스 역할 생성 및 사용	• AWS Clean Rooms 는이 테이블에 필요한 정책을 사용하여 서비스 역할을 생성합니다.

를 선택하는 경우...	THEN ...
	• 기본 서비스 역할 이름은 `cleanrooms-result-receiver-<timestamp>` 입니다 • 역할을 생성하고 정책을 연결할 수 있는 권한이 있어야 합니다.
기존 서비스 역할 사용	i. 드롭다운 목록에서 기존 서비스 역할 이름을 선택합니다. 역할을 나열할 권한이 있는 경우 역할 목록이 표시됩니다. 역할을 나열할 수 있는 권한이 없는 경우 사용하려는 역할의 Amazon 리소스 이름(ARN)을 입력할 수 있습니다. ii. IAM에서 보기 외부 링크를 선택하여 서비스 역할을 확인합니다. 기존 서비스 역할이 없는 경우 기존 서비스 역할 사용 옵션을 사용할 수 없습니다. 기본적으로 AWS Clean Rooms 는 필요한 권한을 추가하기 위해 기존 역할 정책을 업데이트하려고 시도하지 않습니다.

Note

- AWS Clean Rooms 에는 분석 규칙에 따라 쿼리할 수 있는 권한이 필요합니다. 권한에 대한 자세한 내용은 섹션을 AWS Clean Rooms참조하세요AWS 에 대한 관리형 정책 AWS Clean Rooms.
- 역할에 충분한 권한이 없는 경우 역할에 충분한 권한이 없다는 오류 메시지가 AWS Clean Rooms표시됩니다 AWS Clean Rooms. 계속하기 전에 역할 정책을 추가해야 합니다.

- 역할 정책을 수정할 수 없는 경우에서 서비스 역할에 대한 정책을 찾을 AWS Clean Rooms 수 없다는 오류 메시지가 표시됩니다.

2. 로그 설정에서 Amazon CloudWatch Logs의 로그 스토리지에 대한 다음 옵션 중 하나를 선택합니다.

> **Note**
>
> 쿼리 로깅을 활성화하도록 선택한 경우 로그 설정 섹션이 나타납니다.

a. 켜기를 선택하면 관련 쿼리 로그가 Amazon CloudWatch Logs 계정에 저장됩니다.

각 구성원은 자신이 시작했거나 자신의 데이터가 포함된 쿼리에 대한 로그만 받을 수 있습니다.

결과를 받을 수 있는 구성원은 공동 작업 작업에서 실행되는 모든 쿼리에 대한 로그도 받습니다. 쿼리에서 해당 데이터에 액세스하지 않았더라도 마찬가지입니다.

지원되는 로그 유형에서 로그 쿼리 확인란은 기본적으로 켜져 있습니다.

> **Note**
>
> 쿼리 로깅을 사용 설정한 후 로그 저장소가 설정되고 Amazon CloudWatch Logs에서 로그 수신을 시작하는 데 몇 분 정도 걸릴 수 있습니다. 이 짧은 기간 동안 쿼리를 할 수 있는 구성원이 실제로 로그를 전송하지 않는 쿼리를 실행할 수도 있습니다.

b. 끄기를 선택하면 관련 쿼리 로그가 Amazon CloudWatch Logs 계정에 저장되지 않습니다.

3. 멤버십 리소스의 태그를 활성화하려면 새 태그 추가를 선택한 다음 키와 값 쌍을 입력합니다.

4. 쿼리 컴퓨팅 비용을 지불하는 멤버인 경우 이 공동 작업에서 컴퓨팅 비용을 지불하는 데 동의합니다 확인란을 선택하여 수락을 표시합니다.

> **Note**
>
> 계속하려면이 확인란을 선택해야 합니다.

체크섬 계산에 대한 자세한 내용은 에 대한 요금 AWS Clean Rooms 섹션을 참조하세요.

쿼리 컴퓨팅 비용을 지불하는 구성원이지만 쿼리할 수 있는 구성원이 아닌 경우 AWS Budgets 를 사용하여에 대한 예산을 구성 AWS Clean Rooms 하고 최대 예산에 도달하면 알림을 받는 것이 좋습니다. 예산 설정에 대한 자세한 내용은AWS Cost Management 사용 설명서의 AWS Budgets(으)로 비용 관리를 참조하세요. 알림 설정에 대한 자세한 내용은 AWS Cost Management 사용 설명서의 예산 알림을 위한 Amazon SNS 주제 생성을 참조하세요. 최대 예산에 도달한 경우 쿼리를 실행하거나 공동 작업을 중단할 수 있는 구성원에게 문의할 수 있습니다. 공동 작업을 종료하면 더 이상 쿼리를 실행할 수 없으므로 더 이상 쿼리 컴퓨팅 비용이 청구되지 않습니다.

5. 다음을 선택합니다.

   공동 작업과 멤버십이 모두 생성됩니다.

   공동 작업 상태는 활성 상태입니다.

No, I will create a membership later

1. 다음을 선택합니다.

   공동 작업만 생성됩니다.

   공동 작업 상태는 비활성 상태입니다.

8. 5단계: 검토 및 생성에서 다음을 수행합니다.

   a. 이전 단계에서 선택한 내용을 검토하고 필요한 경우 편집하세요.

   b. 다음 옵션 중 하나를 선택합니다.

수행할 작업...	그런 다음을 선택합니다…
공동 작업으로 멤버십 생성합니다(예, 지금 멤버십을 만들어 가입하세요)	공동 작업 및 멤버십 생성
공동 작업을 만들고 지금은 멤버십을 만들지 마세요(아니요, 나중에 멤버십을 만들겠습니다)	공동 작업 생성

공동 작업이 성공적으로 생성되면 공동 작업 아래에서 공동 작업 세부 정보 페이지를 볼 수 있습니다.

이제 다음에 대한 준비가 되었습니다.

- 에서 분석할 데이터 테이블을 준비합니다 AWS Clean Rooms. (자체 이벤트 데이터를 분석하거나 자격 증명 데이터를 쿼리하려는 경우 선택 사항)
- 구성된 테이블을 컬에 연결합니다. (자체 이벤트 데이터를 분석하려는 경우 선택 사항)
- 구성된 테이블에 대한 분석 규칙을 추가합니다. (자체 이벤트 데이터를 분석하려는 경우 선택 사항)
- 멤버십을 만들고 공동 작업에 참여합니다. (이미 멤버십을 생성한 경우 선택 사항)
- 멤버를 공동 작업에 초대합니다.

쿼리 및 작업에 대한 공동 작업 생성

이 절차에서는 공동 작업 생성자가 다음 작업을 수행합니다.

- 공동 작업을 생성합니다.
- 공동 작업에 한 명 이상의 멤버를 초대합니다.
- 쿼리 및 작업을 실행할 수 있는 구성원과 결과를 받을 수 있는 구성원과 같은 구성원에게 기능을 할당합니다.

  공동 작업 생성자가 결과를 받을 수 있는 구성원이기도 한 경우 결과 대상과 형식을 지정합니다. 또한 서비스 역할 Amazon 리소스 이름(ARN)을 제공하여 결과를 결과 대상에 기록합니다.
- 공동 작업에서 쿼리 및 작업 컴퓨팅 비용을 지불할 책임이 있는 구성원을 구성합니다.

시작하기 전에 다음 사전 조건을 완료했는지 확인합니다.

- 공동 작업에 초대하려는 각 멤버의 이름과 AWS 계정 ID가 있습니다.
- 각 멤버의 이름과 AWS 계정 ID를 공동 작업의 모든 멤버와 공유할 수 있는 권한이 있습니다.

> **Note**
>
> 공동 작업을 생성한 후에는 구성원을 더 추가할 수 없습니다.

AWS SDKs. AWS Clean Rooms

쿼리 및 작업에 대한 공동 작업을 생성하려면

1. 에 로그인 AWS Management Console 하고 공동 작업 생성자 역할을 AWS 계정 할를 사용하여 AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 오른쪽 상단 모서리에서 공동 작업 생성을 선택합니다.
4. 1단계: 공동 작업 정의의 경우 다음을 수행합니다:

   a. 세부 정보에는 공동 작업의 이름 및 설명을 입력합니다.

      이 정보는 공동 작업에 참여하도록 초대받은 공동 작업 구성원에게 표시됩니다. 이름과 설명은 공동 작업이 무엇을 의미하는지 이해하는 데 도움이 됩니다.

   b. 분석 엔진에서 Spark를 선택합니다.

      > **Note**
      >
      > AWS Clean Rooms 는 2025년 12월 17일에 레거시 Clean Rooms SQL 분석 엔진에 대한 지원을 종료합니다. 2025년 7월 16일 이전에는 AWS 고객 지원을 통해 한도 증가를 요청하여 새로운 Clean Rooms SQL 엔진 기반 공동 작업을 생성해야 합니다. 2025년 7월 17일부터 새로운 Clean Rooms SQL 엔진 기반 공동 작업을 더 이상 생성할 수 없습니다.
      > AWS Clean Rooms SQL 분석 엔진에서 Spark 분석 엔진으로 공동 작업을 업데이트하려면 기존 공동 작업을 편집하거나 공동 작업을 다시 생성하고 Spark 분석 엔진을 선택합니다.

   c. 구성원의 경우:

      i. 구성원 1: 본인, 공동 작업에 표시할 구성원 표시명을 원하는 대로 입력합니다.

         > **Note**
         >
         > ID AWS 계정 는 멤버 AWS 계정 ID에 자동으로 포함됩니다.

      ii. 멤버 2에 공동 작업에 초대하려는 멤버의 멤버 표시 이름과 멤버 AWS 계정 ID를 입력합니다.

         구성원 표시 이름과 구성원 AWS 계정 ID는 공동 작업에 초대된 모든 사람이 볼 수 있습니다. 이러한 필드의 값을 입력하고 저장한 후에는 편집할 수 없습니다.

> **Note**
>
> 공동 작업에 초대되거나 활동 중인 모든 협업자가 구성원 AWS 계정 ID와 구성원 표시명을 볼 수 있다는 사실을 공동 작업 구성원에게 알려야 합니다.

iii. 다른 구성원을 추가하려면 다른 구성원 추가를 선택합니다. 그런 다음 공동 작업에 초대하려는 데이터를 제공할 수 있는 각 멤버의 멤버 표시 이름과 멤버 AWS 계정 ID를 입력합니다.

d. 분석 로깅을 활성화하려면 분석 로깅 활성화 확인란을 선택한 다음 지원되는 로그 유형을 선택합니다.

- SQL 쿼리에서 생성된 로그를 수신하려면 쿼리에서 로그 확인란을 선택합니다.
- PySpark를 사용하여 작업에서 생성된 로그를 수신하려면 작업에서 로그 확인란을 선택합니다.

e. (선택 사항) 암호화 컴퓨팅 기능을 활성화하려면 암호화 컴퓨팅 활성화 확인란을 선택합니다.

i. 다음 암호화 적용 범위 파라미터를 선택합니다.

- plaintext 열 허용

  완전히 암호화된 테이블이 필요한 경우 아니요를 선택합니다.

  암호화된 테이블에 cleartext 열을 허용하려면 예를 선택합니다.

  특정 열에서 SUM 또는 AVG을(를) 실행하려면 해당 열이 실행하려면 cleartext 안에 있어야 합니다.

- NULL 값 보존

  NULL 값을 보존하지 않으려면 아니오를 선택합니다. NULL 값은 암호화된 테이블에서 NULL처럼 표시되지 않습니다.

  NULL 값을 보존하려면 예를 선택합니다. NULL 값은 암호화된 테이블에서와 NULL같이 표시됩니다.

ii. 다음 지문 파라미터를 선택합니다.

- 중복 허용

  fingerprint 열에 중복 항목을 허용하지 않으려면 아니오를 선택합니다.

fingerprint 열에 중복 항목을 허용하려면 예를 선택합니다.

- 이름이 다른 열의 JOIN 허용

  이름이 다른 fingerprint 열에 가입하지 않으려면 아니오를 선택합니다.

  이름이 다른 fingerprint 열에 참여하려면 예를 선택합니다.

암호화 컴퓨팅 매개변수에 대한 자세한 내용은 암호화 컴퓨팅 파라미터을 참조하세요.

에서 사용할 데이터를 암호화하는 방법에 대한 자세한 내용은 섹션을 AWS Clean Rooms참조하세요Clean Rooms에 대한 암호화 컴퓨팅으로 암호화된 데이터 테이블 준비.

> **Note**
>
> 다음 단계를 완료하기 전에 이러한 구성을 주의 깊게 확인하세요. 공동 작업을 생성한 후에는 공동 작업 이름, 설명 및 로그가 Amazon CloudWatch Logs에 저장되는지 여부만 편집할 수 있습니다.

f. 공동 작업 리소스에 태그를 활성화하려면 새 태그 추가를 선택한 다음 키와 값 쌍을 입력합니다.

g. 다음을 선택합니다.

5. 2단계: 멤버 기능 지정에서 다음을 수행합니다.

   a. 쿼리 및 작업을 사용한 분석의 지원되는 분석 유형에서 작업 확인란을 선택합니다.

      쿼리 확인란은 기본적으로 선택되어 있습니다.

      i. 드롭다운 목록에서 쿼리 및 작업을 실행할 수 있는 구성원을 선택합니다.

      ii. 드롭다운 목록에서 분석 결과를 받을 수 있는 멤버를 선택합니다.

      > **Note**
      >
      > PySpark 분석 템플릿을 생성하는 구성원도 결과를 수신하는 구성원이어야 합니다.

   b. Clean Rooms ML을 사용하는 경우 특별히 구축된 워크플로를 사용한 ML 모델링의 경우

i. (선택 사항) 드롭다운 목록에서 훈련된 모델에서 출력을 수신할 수 있는 멤버를 선택합니다.

ii. (선택 사항) 드롭다운 목록에서 모델 추론에서 출력을 수신할 수 있는 멤버를 선택합니다.

c. 를 사용하여 ID 확인 AWS Entity Resolution에서 멤버 기능을 봅니다.

d. 다음을 선택합니다.

6. 3단계: 결제 구성의 경우

a. 쿼리 및 작업을 사용한 분석에서 쿼리 및 작업에 대한 비용을 지불할 멤버를 선택합니다.

쿼리 및 작업을 실행할 수 있는 구성원을 쿼리 및 작업 컴퓨팅 비용을 지불하는 구성원으로 할당할 수 있습니다.

쿼리 및 작업 컴퓨팅 비용을 지불할 다른 멤버를 할당할 수 있습니다.

b. 특별히 구축된 워크플로를 사용한 ML 모델링의 경우 구성된 유사 모델의 생성자는 유사 모델링 비용을 지불할 구성원입니다.

c. 를 사용한 ID 확인 AWS Entity Resolution의 경우 ID 매핑 테이블의 생성자는 ID 매핑 테이블 비용을 지불할 멤버입니다.

d. 다음을 선택합니다.

7. 4단계: 멤버십 구성에서 다음 옵션 중 하나를 선택합니다.

Yes, join by creating membership now

1. 결과 설정 기본값의 경우 쿼리 결과 설정의 경우 결과를 받을 수 있는 멤버인 경우

a. 쿼리에 대한 기본 설정 설정 확인란을 선택합니다. Amazon S3의 결과 대상에 Amazon S3 대상을 입력하거나 S3 찾아보기를 선택하여 S3 버킷을 선택합니다.

b. 쿼리 결과 형식의 경우 CSV 또는 PARQUET 중 하나를 선택합니다.

c. (Spark만 해당) 결과 파일에서 다중 또는 단일을 선택합니다.

d. (선택 사항) 서비스 액세스에서 S3 대상에 최대 24시간이 걸리는 쿼리를 전송하려면 완료하는 데 최대 24시간이 걸리는 쿼리를 지원하는 서비스 역할 추가 확인란을 선택합니다.

완료하는 데 최대 24시간이 걸리는 대규모 쿼리는 S3 대상으로 전송됩니다.

확인란을 선택하지 않으면 12시간 이내에 완료된 쿼리만 S3 위치로 전송됩니다.

e. 새 서비스 역할 생성 및 사용 또는 기존 서비스 역할 사용을 선택하여 서비스 액세스 권한을 지정합니다.

를 선택하는 경우...	THEN ...
새 서비스 역할 생성 및 사용	• AWS Clean Rooms 는이 테이블에 필요한 정책을 사용하여 서비스 역할을 생성합니다. • 기본 서비스 역할 이름은 `cleanrooms-result-receiver-<timestamp>` 입니다 • 역할을 생성하고 정책을 연결할 수 있는 권한이 있어야 합니다.
기존 서비스 역할 사용	i. 드롭다운 목록에서 기존 서비스 역할 이름을 선택합니다. 역할을 나열할 권한이 있는 경우 역할 목록이 표시됩니다. 역할을 나열할 수 있는 권한이 없는 경우 사용하려는 역할의 Amazon 리소스 이름(ARN)을 입력할 수 있습니다. ii. IAM에서 보기 외부 링크를 선택하여 서비스 역할을 확인합니다. 기존 서비스 역할이 없는 경우 기존 서비스 역할 사용 옵션을 사용할 수 없습니다. 기본적으로 AWS Clean Rooms 는 필요한 권한을 추가하기 위해 기존 역할 정책을 업데이트하려고 시도하지 않습니다.

> **Note**
>
> - AWS Clean Rooms 에는 분석 규칙에 따라 쿼리할 수 있는 권한이 필요합니다. 권한에 대한 자세한 내용은 섹션을 AWS Clean Rooms참조하세요AWS 에 대한 관리형 정책 AWS Clean Rooms.
> - 역할에 충분한 권한이 없는 경우 역할에 충분한 권한이 없다는 오류 메시지가 AWS Clean Rooms표시됩니다 AWS Clean Rooms. 계속하기 전에 역할 정책을 추가해야 합니다.
> - 역할 정책을 수정할 수 없는 경우에서 서비스 역할에 대한 정책을 찾을 AWS Clean Rooms 수 없다는 오류 메시지가 표시됩니다.

2. 작업 결과의 경우

   Example

   예: **`s3://bucket/prefix`**

   a. 작업에 대한 기본 설정 설정 확인란을 선택한 다음 S3 대상을 입력하여 Amazon S3에서 결과 대상을 지정하거나 S3 찾아보기를 선택하여 사용 가능한 S3 버킷 목록에서 선택합니다. S3

   b. 드롭다운 목록에서 기존 서비스 역할 이름을 선택하여 서비스 액세스 권한을 지정합니다.

3. 로그 설정에서 Amazon CloudWatch Logs의 로그 스토리지에 대한 다음 옵션 중 하나를 선택합니다.

   > **Note**
   >
   > 쿼리 로깅을 활성화하도록 선택한 경우 로그 설정 섹션이 나타납니다.

   a. 켜기를 선택하면 관련 쿼리 로그가 Amazon CloudWatch Logs 계정에 저장됩니다.

   각 구성원은 자신이 시작했거나 자신의 데이터가 포함된 쿼리에 대한 로그만 받을 수 있습니다.

   결과를 받을 수 있는 구성원은 공동 작업 작업에서 실행되는 모든 쿼리에 대한 로그도 받습니다. 쿼리에서 해당 데이터에 액세스하지 않았더라도 마찬가지입니다.

지원되는 로그 유형에서 공동 작업 생성자가 지원하도록 선택한 로그 유형 중에서 선택합니다.

지원되는 로그 유형에서 쿼리 로그 및 작업 로그 확인란은 기본적으로 켜져 있습니다.

> **Note**
>
> 분석 로깅을 켠 후 로그 스토리지를 설정하고 Amazon CloudWatch Logs에서 로그 수신을 시작하는 데 몇 분 정도 걸릴 수 있습니다. 이 짧은 기간 동안 쿼리를 할 수 있는 구성원이 실제로 로그를 전송하지 않는 쿼리를 실행할 수도 있습니다.

b. 끄기를 선택하면 관련 쿼리 로그가 Amazon CloudWatch Logs 계정에 저장되지 않습니다.

4. 멤버십 리소스에 대해 멤버십 태그를 활성화하려면 새 태그 추가를 선택한 다음 키 및 값 페어를 입력합니다.

5. 쿼리 컴퓨팅, 작업 컴퓨팅 또는 둘 다에 대한 비용을 지불하는 멤버인 경우 이 공동 작업에서 컴퓨팅 비용 지불에 동의합니다 확인란을 선택하여 수락을 표시합니다.

> **Note**
>
> 계속하려면이 확인란을 선택해야 합니다.
> 체크섬 계산에 대한 자세한 내용은 에 대한 요금 AWS Clean Rooms 섹션을 참조하세요.

쿼리 컴퓨팅 비용을 지불하는 구성원이지만 쿼리할 수 있는 구성원이 아닌 경우 AWS Budgets 를 사용하여에 대한 예산을 구성 AWS Clean Rooms 하고 최대 예산에 도달하면 알림을 받는 것이 좋습니다. 예산 설정에 대한 자세한 내용은AWS Cost Management 사용 설명서의 AWS Budgets(으)로 비용 관리를 참조하세요. 알림 설정에 대한 자세한 내용은 AWS Cost Management 사용 설명서의 예산 알림을 위한 Amazon SNS 주제 생성을 참조하세요. 최대 예산에 도달한 경우 쿼리를 실행하거나 공동 작업을 중단할 수 있는 구성원에게 문의할 수 있습니다. 공동 작업을 종료하면 더 이상 쿼리를 실행할 수 없으므로 더 이상 쿼리 컴퓨팅 비용이 청구되지 않습니다.

6. 다음을 선택합니다.

공동 작업과 멤버십이 모두 생성됩니다.

공동 작업 상태는 활성 상태입니다.

No, I will create a membership later

1. 다음을 선택합니다.

   공동 작업만 생성됩니다.

   공동 작업 상태는 비활성 상태입니다.

8. 5단계: 검토 및 생성에서 다음을 수행합니다.

   a. 이전 단계에서 선택한 내용을 검토하고 필요한 경우 편집하세요.

   b. 다음 옵션 중 하나를 선택합니다.

수행할 작업...	그런 다음을 선택합니다…
공동 작업으로 멤버십 생성합니다(예, 지금 멤버십을 만들어 가입하세요)	공동 작업 및 멤버십 생성
공동 작업을 만들고 지금은 멤버십을 만들지 마세요(아니요, 나중에 멤버십을 만들겠습니다)	공동 작업 생성

공동 작업이 성공적으로 생성되면 공동 작업 아래에서 공동 작업 세부 정보 페이지를 볼 수 있습니다.

이제 다음에 대한 준비가 되었습니다.

- 에서 분석할 데이터 테이블을 준비합니다 AWS Clean Rooms. (자체 이벤트 데이터를 분석하거나 자격 증명 데이터를 쿼리하려는 경우 선택 사항)
- 구성된 테이블을 컬에 연결합니다. (자체 이벤트 데이터를 분석하려는 경우 선택 사항)
- 구성된 테이블에 대한 분석 규칙을 추가합니다. (자체 이벤트 데이터를 분석하려는 경우 선택 사항)
- 멤버십을 만들고 공동 작업에 참여합니다. (이미 멤버십을 생성한 경우 선택 사항)
- 멤버를 공동 작업에 초대합니다.

ML 모델링을 위한 공동 작업 생성

이 절차에서는 공동 작업 생성자가 다음 작업을 수행합니다.

- 공동 작업을 생성합니다.
- 공동 작업에 한 명 이상의 멤버를 초대합니다.
- 와 같은 기능을 멤버에게 할당
  - 쿼리할 수 있는 멤버
  - 결과를 받을 수 있는 멤버
  - 훈련된 모델로부터 출력을 받을 수 있는 구성원
  - 모델 추론에서 출력할 수 있는 멤버

  공동 작업 생성자가 결과를 받을 수 있는 구성원이기도 한 경우 결과 대상과 형식을 지정합니다. 또한 서비스 역할 Amazon 리소스 이름(ARN)을 제공하여 결과를 결과 대상에 기록합니다.
- 공동 작업에서 컴퓨팅 비용, 모델 훈련 및 모델 추론 비용을 지불할 책임이 있는 구성원을 구성합니다.

시작하기 전에 다음 사전 조건을 완료했는지 확인합니다.

- 공동 작업에 초대하려는 각 멤버의 이름과 AWS 계정 ID가 있습니다.
- 각 멤버의 이름과 AWS 계정 ID를 공동 작업의 모든 멤버와 공유할 수 있는 권한이 있습니다.

> **Note**
>
> 공동 작업을 생성한 후에는 구성원을 더 추가할 수 없습니다.

AWS SDKs. AWS Clean Rooms

ML 모델링을 위한 공동 작업을 생성하려면

1. 에 로그인 AWS Management Console 하고 공동 작업 생성자 역할을 AWS 계정 할를 사용하여 AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 오른쪽 상단 모서리에서 공동 작업 생성을 선택합니다.
4. 1단계: 공동 작업 정의의 경우 다음을 수행합니다:

   a. 세부 정보에는 공동 작업의 이름 및 설명을 입력합니다.

이 정보는 공동 작업에 참여하도록 초대받은 공동 작업 구성원에게 표시됩니다. 이름과 설명은 공동 작업이 무엇을 의미하는지 이해하는 데 도움이 됩니다.

b. 분석 엔진에서 Spark를 선택합니다.

c. 구성원의 경우:

i. 구성원 1: 본인, 공동 작업에 표시할 구성원 표시명을 원하는 대로 입력합니다.

> **Note**
>
> ID AWS 계정 는 멤버 AWS 계정 ID에 자동으로 포함됩니다.

ii. 멤버 2에 공동 작업에 초대하려는 멤버의 멤버 표시 이름과 멤버 AWS 계정 ID를 입력합니다.

구성원 표시 이름과 구성원 AWS 계정 ID는 공동 작업에 초대된 모든 사람이 볼 수 있습니다. 이러한 필드의 값을 입력하고 저장한 후에는 편집할 수 없습니다.

> **Note**
>
> 공동 작업에 초대되거나 활동 중인 모든 협업자가 구성원 AWS 계정 ID와 구성원 표시명을 볼 수 있다는 사실을 공동 작업 구성원에게 알려야 합니다.

iii. 다른 구성원을 추가하려면 다른 구성원 추가를 선택합니다. 그런 다음 공동 작업에 초대하려는 데이터를 제공할 수 있는 각 멤버의 멤버 표시 이름과 멤버 AWS 계정 ID를 입력합니다.

d. 분석 로깅을 활성화하려면 분석 로깅 활성화 확인란을 선택한 다음 지원되는 로그 유형에서 쿼리의 로그를 선택합니다.

e. (선택 사항) 암호화 컴퓨팅 기능을 활성화하려면 암호화 컴퓨팅 활성화 확인란을 선택합니다.

i. 다음 암호화 적용 범위 파라미터를 선택합니다.

- plaintext 열 허용

  완전히 암호화된 테이블이 필요한 경우 아니요를 선택합니다.

  암호화된 테이블에 cleartext 열을 허용하려면 예를 선택합니다.

특정 열에서 SUM 또는 AVG을(를) 실행하려면 해당 열이 실행하려면 cleartext 안에 있어야 합니다.

- NULL 값 보존

  NULL 값을 보존하지 않으려면 아니오를 선택합니다. NULL 값은 암호화된 테이블에서 NULL처럼 표시되지 않습니다.

  NULL 값을 보존하려면 예를 선택합니다. NULL 값은 암호화된 테이블에서와 NULL같이 표시됩니다.

ii. 다음 지문 파라미터를 선택합니다.

- 중복 허용

  fingerprint 열에 중복 항목을 허용하지 않으려면 아니오를 선택합니다.

  fingerprint 열에 중복 항목을 허용하려면 예를 선택합니다.

- 이름이 다른 열의 JOIN 허용

  이름이 다른 fingerprint 열에 가입하지 않으려면 아니오를 선택합니다.

  이름이 다른 fingerprint 열에 참여하려면 예를 선택합니다.

암호화 컴퓨팅 매개변수에 대한 자세한 내용은 암호화 컴퓨팅 파라미터을 참조하세요.

에서 사용할 데이터를 암호화하는 방법에 대한 자세한 내용은 섹션을 AWS Clean Rooms참조하세요Clean Rooms에 대한 암호화 컴퓨팅으로 암호화된 데이터 테이블 준비.

> **Note**
>
> 다음 단계를 완료하기 전에 이러한 구성을 주의 깊게 확인하세요. 공동 작업을 생성한 후에는 공동 작업 이름, 설명 및 로그가 Amazon CloudWatch Logs에 저장되는지 여부만 편집할 수 있습니다.

f. 공동 작업 리소스에 태그를 활성화하려면 새 태그 추가를 선택한 다음 키와 값 쌍을 입력합니다.

g. 다음을 선택합니다.

5. 2단계: 멤버 기능 지정의 경우

a. 쿼리 및 작업을 사용한 분석의 경우 지원되는 분석 유형에서 쿼리 확인란을 선택한 상태로 둡니다.

b. 쿼리 실행에서 모델 훈련을 시작할 멤버를 선택합니다.

c. 분석에서 결과 수신에서 쿼리 결과를 받을 멤버를 하나 이상 선택합니다.

d. 특별히 구축된 워크플로를 사용한 ML 모델링의 경우

   i. 훈련된 모델에서 출력 수신에서 모델 아티팩트 및 지표를 포함하여 훈련된 모델 결과를 받을 멤버를 선택합니다.

   ii. 모델 추론에서 출력 수신에서 모델 추론 결과를 받을 멤버를 선택합니다.

e. 를 사용하여 ID 확인 AWS Entity Resolution에서 멤버 기능을 봅니다.

6. 3단계: 결제 구성의 경우 쿼리를 사용한 분석에서 목표에 따라 다음 작업 중 하나를 수행합니다.

목표	권장 조치
쿼리를 실행할 수 있는 구성원을 쿼리 계산 비용을 지불하는 구성원으로 배정합니다	1. 쿼리를 실행할 수 있는 멤버와 동일하게 쿼리 비용을 지불할 멤버를 선택합니다. 2. 다음을 선택합니다.
쿼리 컴퓨팅 비용을 지불할 다른 구성원을 배정하세요	1. 쿼리 비용을 지불할 멤버로 자신을 선택합니다. 2. 다음을 선택합니다.

특별히 구축된 워크플로를 사용한 ML 모델링의 경우 구성된 유사 모델의 생성자는 유사 모델링 비용을 지불할 구성원입니다.

를 사용한 ID 확인 AWS Entity Resolution의 경우 ID 매핑 테이블의 생성자는 ID 매핑 테이블 비용을 지불할 멤버입니다.

7. 4단계: 멤버십 구성에서 다음 옵션 중 하나를 선택합니다.

   Yes, join by creating membership now

   1. 결과 설정 기본값의 경우 쿼리 결과 설정의 경우 결과를 받을 수 있는 멤버인 경우

      a. Amazon S3의 결과 대상에 Amazon S3 대상을 입력하거나 S3 찾아보기를 선택하여 S3 버킷을 선택합니다.

b. 쿼리 결과 형식의 경우 CSV 또는 PARQUET 중 하나를 선택합니다.

c. (Spark만 해당) 결과 파일에서 다중 또는 단일을 선택합니다.

d. (선택 사항) 서비스 액세스에서 S3 대상에 최대 24시간이 걸리는 쿼리를 전송하려면 완료하는 데 최대 24시간이 걸리는 쿼리를 지원하는 서비스 역할 추가 확인란을 선택합니다.

   완료하는 데 최대 24시간이 걸리는 대규모 쿼리는 S3 대상으로 전송됩니다.

   확인란을 선택하지 않으면 12시간 이내에 완료된 쿼리만 S3 위치로 전송됩니다.

e. 새 서비스 역할 생성 및 사용 또는 기존 서비스 역할 사용을 선택하여 서비스 액세스 권한을 지정합니다.

를 선택하는 경우...	THEN ...
새 서비스 역할 생성 및 사용	• AWS Clean Rooms 는이 테이블에 필요한 정책을 사용하여 서비스 역할을 생성합니다. • 기본 서비스 역할 이름은 `cleanrooms-result-receiver-<timestamp>` 입니다 • 역할을 생성하고 정책을 연결할 수 있는 권한이 있어야 합니다.

를 선택하는 경우...	THEN ...
기존 서비스 역할 사용	i. 드롭다운 목록에서 기존 서비스 역할 이름을 선택합니다. 역할을 나열할 권한이 있는 경우 역할 목록이 표시됩니다. 역할을 나열할 수 있는 권한이 없는 경우 사용하려는 역할의 Amazon 리소스 이름(ARN)을 입력할 수 있습니다. ii. IAM에서 보기 외부 링크를 선택하여 서비스 역할을 확인합니다. 기존 서비스 역할이 없는 경우 기존 서비스 역할 사용 옵션을 사용할 수 없습니다. 기본적으로 AWS Clean Rooms 는 필요한 권한을 추가하기 위해 기존 역할 정책을 업데이트하려고 시도하지 않습니다.

> **Note**
>
> - AWS Clean Rooms 에는 분석 규칙에 따라 쿼리할 수 있는 권한이 필요합니다. 권한에 대한 자세한 내용은 섹션을 AWS Clean Rooms참조하세요AWS 에 대한 관리형 정책 AWS Clean Rooms.
> - 역할에 충분한 권한이 없는 경우 역할에 충분한 권한이 없다는 오류 메시지가 AWS Clean Rooms표시됩니다 AWS Clean Rooms. 계속하기 전에 역할 정책을 추가해야 합니다.
> - 역할 정책을 수정할 수 없는 경우에서 서비스 역할에 대한 정책을 찾을 AWS Clean Rooms 수 없다는 오류 메시지가 표시됩니다.

2. 작업 결과의 경우

Example

예: **`s3://bucket/prefix`**

a. 작업에 대한 기본 설정 설정 확인란을 선택한 다음 S3 대상을 입력하여 Amazon S3에서 결과 대상을 지정하거나 S3 찾아보기를 선택하여 사용 가능한 S3 버킷 목록에서 선택합니다. S3

b. 드롭다운 목록에서 기존 서비스 역할 이름을 선택하여 서비스 액세스 권한을 지정합니다.

3. 로그 설정에서 Amazon CloudWatch Logs의 로그 스토리지에 대한 다음 옵션 중 하나를 선택합니다.

> **Note**
>
> 쿼리 로깅을 활성화하도록 선택한 경우 로그 설정 섹션이 나타납니다.

a. 켜기를 선택하면 관련 쿼리 로그가 Amazon CloudWatch Logs 계정에 저장됩니다.

각 구성원은 자신이 시작했거나 자신의 데이터가 포함된 쿼리에 대한 로그만 받을 수 있습니다.

결과를 받을 수 있는 구성원은 공동 작업 작업에서 실행되는 모든 쿼리에 대한 로그도 받습니다. 쿼리에서 해당 데이터에 액세스하지 않았더라도 마찬가지입니다.

지원되는 로그 유형에서 공동 작업 생성자가 지원하도록 선택한 로그 유형 중에서 선택합니다.

지원되는 로그 유형에서 로그 쿼리 확인란은 기본적으로 켜져 있습니다.

> **Note**
>
> 분석 로깅을 켠 후 로그 스토리지를 설정하고 Amazon CloudWatch Logs에서 로그 수신을 시작하는 데 몇 분 정도 걸릴 수 있습니다. 이 짧은 기간 동안 쿼리를 할 수 있는 구성원이 실제로 로그를 전송하지 않는 쿼리를 실행할 수도 있습니다.

b. 끄기를 선택하면 관련 쿼리 로그가 Amazon CloudWatch Logs 계정에 저장되지 않습니다.

4. 멤버십 리소스의 태그를 활성화하려면 새 태그 추가를 선택한 다음 키와 값 쌍을 입력합니다.
5. 쿼리 컴퓨팅 비용을 지불하는 멤버인 경우 이 공동 작업에서 컴퓨팅 비용을 지불하는 데 동의합니다 확인란을 선택하여 수락을 표시합니다.

> **Note**
>
> 계속하려면이 확인란을 선택해야 합니다.
> 체크섬 계산에 대한 자세한 내용은 에 대한 요금 AWS Clean Rooms 섹션을 참조하세요.

쿼리 컴퓨팅 비용을 지불하는 구성원이지만 쿼리할 수 있는 구성원이 아닌 경우 AWS Budgets 를 사용하여에 대한 예산을 구성 AWS Clean Rooms 하고 최대 예산에 도달하면 알림을 받는 것이 좋습니다. 예산 설정에 대한 자세한 내용은AWS Cost Management 사용 설명서의 AWS Budgets(으)로 비용 관리를 참조하세요. 알림 설정에 대한 자세한 내용은 AWS Cost Management 사용 설명서의 예산 알림을 위한 Amazon SNS 주제 생성을 참조하세요. 최대 예산에 도달한 경우 쿼리를 실행하거나 공동 작업을 중단할 수 있는 구성원에게 문의할 수 있습니다. 공동 작업을 종료하면 더 이상 쿼리를 실행할 수 없으므로 더 이상 쿼리 컴퓨팅 비용이 청구되지 않습니다.

6. 다음을 선택합니다.

공동 작업과 멤버십이 모두 생성됩니다.

공동 작업 상태는 활성 상태입니다.

No, I will create a membership later

1. 다음을 선택합니다.

   공동 작업만 생성됩니다.

   공동 작업 상태는 비활성 상태입니다.

8. 5단계: 검토 및 생성에서 다음을 수행합니다.

   a. 이전 단계에서 선택한 내용을 검토하고 필요한 경우 편집하세요.

   b. 다음 옵션 중 하나를 선택합니다.

수행할 작업...	그런 다음을 선택합니다...
공동 작업으로 멤버십 생성합니다(예, 지금 멤버십을 만들어 가입하세요)	공동 작업 및 멤버십 생성
공동 작업을 만들고 지금은 멤버십을 만들지 마세요(아니요, 나중에 멤버십을 만들겠습니다)	공동 작업 생성

멤버십 생성 및 공동 작업 참여

멤버십은 구성원이 AWS Clean Rooms에서 공동 작업에 참여할 때 생성되는 리소스입니다.

로 공동 작업에 참여할 수 있습니다.

- 쿼리할 수 있는 멤버
- 쿼리 및 작업을 실행할 수 있는 멤버
- 쿼리 또는 작업의 결과를 수신할 수 있는 멤버
- 쿼리 컴퓨팅 비용을 지불하는 멤버
- 쿼리 및 작업 비용을 지불하는 멤버

모든 구성원이 데이터를 제공할 수 있습니다.

AWS SDK를 사용하여 멤버십을 생성하고 공동 작업에 참여하는 방법에 대한 자세한 내용은 AWS Clean Rooms API 참조를 참조하세요.

이 절차에서 초대된 구성원은 멤버십 리소스를 생성하여 공동 작업에 참여합니다.

초대된 멤버가 결과를 받을 수 있는 멤버인 경우 결과 대상과 형식을 지정합니다. 또한 결과 대상에 쓸 서비스 역할 ARN을 제공합니다.

초대된 구성원이 컴퓨팅 비용을 지불할 책임이 있는 구성원인 경우 공동 작업에 참여하기 전에 지불 책임을 수락합니다.

멤버십을 만들고 공동 작업에 참여하려면

1. 에 로그인 AWS Management Console 하고 멤버 로 AWS Clean Rooms 콘솔을 엽니다 AWS 계정.
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 참여 가능 탭에서 참여할 수 있는 공동 작업의 경우 공동 작업의 이름을 선택합니다.
4. 공동 작업 세부 정보 페이지의 개요 섹션에서 멤버 세부 정보 및 다른 멤버 목록을 포함한 공동 작업 세부 정보를 봅니다.

   공동 작업의 각 구성원에 대한 AWS 계정 IDs가 공동 작업에 참여하려는 ID인지 확인합니다.
5. 멤버십 생성을 선택합니다.
6. 멤버십 생성 페이지의 개요에서 공동 작업 이름, 공동 작업 설명, 공동 작업 생성자의 AWS 계정 ID, 멤버 세부 정보 및 쿼리 비용을 지불할 멤버의 AWS 계정 ID를 확인합니다.
7. 공동 작업 생성자가 분석 로깅을 활성화하도록 선택한 경우 Amazon CloudWatch Logs의 로그 스토리지에 대해 다음 옵션 중 하나를 선택합니다.

사용자 선택 항목...	THEN ...
켜기	사용자와 관련된 로그는 Amazon CloudWatch Logs에 저장됩니다. 각 구성원은 자신이 시작했거나 자신의 데이터가 포함된 쿼리에 대한 로그만 받을 수 있습니다. 결과를 받을 수 있는 구성원은 분석에서 데이터에 액세스하지 않더라도 공동 작업에서 실행되는 모든 분석에 대한 로그도 수신합니다. 지원되는 로그 유형에서 공동 작업 생성자가 지원하도록 선택한 로그 유형 중에서 선택합니다. 1. SQL 쿼리에서 생성된 로그를 수신하려면 쿼리에서 로그 확인란을 선택합니다.

사용자 선택 항목...	THEN ...
	2. PySpark를 사용하여 작업에서 생성된 로그를 수신하려면 작업에서 로그 확인란을 선택합니다.
끄기	사용자와 관련된 쿼리 로그는 Amazon CloudWatch Logs 계정에 저장되지 않습니다.

> **Note**
>
> 분석 로깅을 켠 후 로그 스토리지를 설정하고 Amazon CloudWatch Logs에서 로그 수신을 시작하는 데 몇 분 정도 걸릴 수 있습니다. 이 짧은 기간 동안 쿼리를 할 수 있는 구성원이 실제로 로그를 전송하지 않는 쿼리를 실행할 수도 있습니다.

8. 멤버 기능에 결과 수신이 포함된 경우 결과 설정의 기본값은 다음과 같습니다.

   a. 쿼리 결과에서 쿼리에 대한 기본 설정 설정 확인란을 선택한 다음 S3 대상을 입력하여 Amazon S3에서 결과 대상을 지정하거나 S3 찾아보기를 선택하여 사용 가능한 S3 버킷 목록에서 선택합니다. S3

      Example

      예: **`s3://bucket/prefix`**

      i. 결과 형식에서 CSV 또는 PARQUET를 선택합니다.

      ii. (Spark만 해당) 결과 파일에서 다중 또는 단일을 선택합니다.

      iii. (선택 사항) 서비스 액세스에서 S3 대상에 최대 24시간이 걸리는 쿼리를 전송하려면 완료에 최대 24시간이 걸리는 쿼리를 지원하는 서비스 역할 추가 확인란을 선택합니다.

      완료하는 데 최대 24시간이 걸리는 대규모 쿼리는 S3 대상으로 전송됩니다.

      확인란을 선택하지 않으면 12시간 이내에 완료된 쿼리만 S3 위치로 전송됩니다.

      > **Note**
      >
      > 기존 서비스 역할을 선택하거나 새 서비스 역할을 생성할 수 있는 권한이 있어야 합니다. 자세한 내용은 결과를 수신할 서비스 역할 생성 단원을 참조하십시오.

iv. 새 서비스 역할 생성 및 사용 또는 기존 서비스 역할 사용을 선택하여 서비스 액세스 권한을 지정합니다.

Create and use a new service role

- AWS Clean Rooms 는이 테이블에 필요한 정책을 사용하여 서비스 역할을 생성합니다.
- 기본 서비스 역할 이름은 `cleanrooms-result-receiver-<timestamp>`입니다
- 역할을 생성하고 정책을 연결할 수 있는 권한이 있어야 합니다.

Use an existing service role

1. 드롭다운 목록에서 기존 서비스 역할 이름을 선택합니다.

   역할을 나열할 권한이 있는 경우 역할 목록이 표시됩니다.

   역할을 나열할 수 있는 권한이 없는 경우 사용하려는 역할의 Amazon 리소스 이름(ARN)을 입력할 수 있습니다.
2. IAM에서 보기 외부 링크를 선택하여 서비스 역할을 확인합니다.

   기존 서비스 역할이 없는 경우 기존 서비스 역할 사용 옵션을 사용할 수 없습니다.

   기본적으로 AWS Clean Rooms 는 필요한 권한을 추가하기 위해 기존 역할 정책을 업데이트하려고 시도하지 않습니다.

> **Note**
>
> - AWS Clean Rooms 에는 분석 규칙에 따라 쿼리할 수 있는 권한이 필요합니다. 권한에 대한 자세한 내용은 섹션을 AWS Clean Rooms참조하세요AWS 에 대한 관리형 정책 AWS Clean Rooms.
> - 역할에 충분한 권한이 없는 경우 역할에 충분한 권한이 없다는 오류 메시지가 AWS Clean Rooms표시됩니다 AWS Clean Rooms. 계속하기 전에 역할 정책을 추가해야 합니다.
> - 역할 정책을 수정할 수 없는 경우에서 서비스 역할에 대한 정책을 찾을 AWS Clean Rooms 수 없다는 오류 메시지가 표시됩니다.

b. 작업 결과에서 작업에 대한 기본 설정 설정 확인란을 선택한 다음 S3 대상을 입력하여 Amazon S3에서 결과 대상을 지정하거나 S3 찾아보기를 선택하여 사용 가능한 S3 버킷 목록에서 선택합니다. S3

Example

예: **`s3://bucket/prefix`**

- 드롭다운 목록에서 기존 서비스 역할 이름을 선택하여 서비스 액세스 권한을 지정합니다.

9. 멤버십 리소스의 태그를 활성화하려면 새 태그 추가를 선택한 다음 키와 값 쌍을 입력합니다.
10. 공동 작업 생성자가 쿼리에 대해 지불하거나 쿼리 및 작업에 대해 지불하도록 사용자를 멤버로 지정한 경우 이 공동 작업에서 컴퓨팅 비용 지불에 동의합니다 확인란을 선택하여 수락을 표시합니다.

> **Note**
>
> 계속하려면이 확인란을 선택해야 합니다.
> 체크섬 계산에 대한 자세한 내용은 에 대한 요금 AWS Clean Rooms 섹션을 참조하세요.

쿼리 컴퓨팅 비용을 지불하는 멤버이거나 쿼리할 수 있는 멤버가 아닌 쿼리 및 작업 컴퓨팅 비용을 지불하는 멤버인 경우 AWS Budgets 를 사용하여에 대한 예산을 구성 AWS Clean Rooms 하고 최대 예산에 도달하면 알림을 받는 것이 좋습니다. 예산 설정에 대한 자세한 내용은AWS Cost Management 사용 설명서의 AWS Budgets(으)로 비용 관리를 참조하세요. 알림 설정에 대한 자세한 내용은AWS Cost Management 사용 설명서의 예산 알림을 위한 Amazon SNS 주제 생성을 참조하세요. 최대 예산에 도달한 경우 쿼리 및 작업을 실행하거나 공동 작업을 나갈 수 있는 구성원에게 문의할 수 있습니다. 공동 작업을 종료하면 더 이상 쿼리를 실행할 수 없으므로 더 이상 쿼리 컴퓨팅 비용이 청구되지 않습니다.

11. 멤버십을 만들고 공동 작업에 참여하려는 것이 확실하다면 멤버십 생성을 선택하세요.

공동 작업 메타데이터에 대한 읽기 권한이 부여됩니다. 여기에는 다른 멤버의 모든 이름 및 AWS 계정 IDs 외에도 공동 작업의 표시 이름 및 설명과 같은 정보가 포함됩니다.

이제 다음에 대한 준비가 되었습니다.

- AWS Clean Rooms에서 쿼리할 데이터 테이블을 준비하세요. (자체 이벤트 데이터를 쿼리하거나 ID 데이터를 쿼리하려는 경우 선택 사항)
- 구성된 테이블을 공동 작업에 연결 - 이벤트 데이터를 쿼리하려는 경우
- 구성된 테이블에 대한 분석 규칙 추가 - 이벤트 데이터를 쿼리하려는 경우
- 새 ID 네임스페이스 생성 및 연결 - ID 매핑 테이블을 생성하여 ID 데이터를 쿼리하려는 경우.

공동 작업에서 탈퇴하는 방법에 대한 자세한 내용은 공동 작업 탈퇴 섹션을 참조하세요.

공동 작업 편집

공동 작업 생성자는 공동 작업의 다양한 부분을 편집할 수 있습니다.

AWS SDKs를 사용하여 공동 작업을 편집하는 방법에 대한 자세한 내용은 AWS Clean Rooms API 참조를 참조하세요.

주제

- 공동 작업 이름 및 설명 편집
- 공동 작업 분석 엔진 업데이트
- 로그 스토리지 끄기
- 공동 작업 로그 설정 편집
- 공동 작업 태그 편집
- 멤버십 태그 편집
- 관련 테이블 태그 편집
- 분석 템플릿 태그 편집
- 차등 프라이버시 정책 태그 편집

공동 작업 이름 및 설명 편집

공동 작업을 생성한 후에는 공동 작업 이름 및 설명만 편집할 수 있습니다.

공동 작업 이름 및 설명을 편집하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.

3. 생성한 공동 작업을 선택합니다.
4. 공동 작업 세부정보 페이지에서 작업을 선택한 다음 공동 작업 편집을 선택합니다.
5. 공동 작업 편집 페이지의 세부 정보에서 공동 작업의 이름 및 설명을 편집합니다.
6. 변경 사항 저장을 선택합니다.

공동 작업 분석 엔진 업데이트

공동 작업을 생성한 후 분석 엔진을 AWS Clean Rooms SQL에서 Spark로 변경할 수 있습니다.

> **Note**
>
> 분석 엔진을 AWS Clean Rooms SQL에서 Spark로 변경하면 기존 워크플로가 중단될 수 있습니다.

공동 작업 분석 엔진을 업데이트하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 생성한 공동 작업을 선택합니다.
4. 공동 작업 세부정보 페이지에서 작업을 선택한 다음 공동 작업 편집을 선택합니다.
5. 공동 작업 편집 페이지의 분석 엔진에서

   - AWS Clean Rooms SQL을 선택한 경우 Spark를 선택합니다.
   - Spark를 선택한 경우 지원 티켓 제출을 선택하여 지원 티켓을 제출하여 분석 엔진을 AWS Clean Rooms SQL로 변경합니다.

   > **Note**
   >
   > AWS Clean Rooms 는 2025년 12월 17일에 레거시 Clean Rooms SQL 분석 엔진에 대한 지원을 종료합니다. 2025년 7월 16일 이전에 AWS 는 고객 지원을 통해 한도 증가를 요청하여 새로운 Clean Rooms SQL 엔진 기반 공동 작업을 생성해야 합니다. 2025년 7월 16일 이후에는 새로운 Clean Rooms SQL 엔진 기반 공동 작업을 더 이상 생성할 수 없습니다.

6. 변경 사항 저장을 선택합니다.

로그 스토리지 끄기

분석 로깅을 활성화한 경우 분석 로그가 Amazon CloudWatch Logs 계정에 저장되는지 여부를 편집할 수 있습니다.

로그 스토리지를 끄려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 분석 로깅이 켜져 있는 공동 작업을 선택합니다.
4. 공동 작업 세부 정보 페이지에서 작업을 선택한 다음 로그 스토리지 끄기를 선택합니다.

> **Note**
>
> 다음을 나타내는 경고가 나타납니다.
>
> - 새 쿼리는 더 이상 CloudWatch 계정에 기록되지 않습니다.
> - 기존 로그는 현재 보존 설정에 따라 보존됩니다.
> - 나중에 로깅을 다시 활성화하면 다시 활성화한 후 수행된 쿼리에만 적용됩니다.
> - 이 변경 사항은 로그에만 영향을 미치며, 다른 팀원의 로깅 설정은 변경되지 않습니다.

5. 끄기를 선택합니다.

공동 작업 로그 설정 편집

쿼리 로깅을 활성화한 경우 Amazon CloudWatch Logs 계정에 쿼리 로그를 저장할지 여부를 편집할 수 있습니다.

공동 작업 로그 설정을 편집하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 생성한 공동 작업을 선택합니다.

4. 공동 작업 세부 정보 페이지에서 다음 중 하나를 수행합니다.
   - 작업을 선택한 다음 로그 설정 편집을 선택합니다.
   - 로그 탭에서 로그 설정 편집을 선택합니다.
5. 로그 설정 편집 모달에서 Amazon CloudWatch Logs의 로그 스토리지에 대해:
   - 사용자와 관련된 로그를 Amazon CloudWatch Logs 계정에 저장하지 않으려면 끄기를 선택합니다.
   - 사용자와 관련된 로그를 Amazon CloudWatch Logs 계정에 저장하려면 켜기를 선택합니다.

     시작했거나 데이터가 포함된 쿼리에 대한 로그만 수신할 수 있습니다.

     결과를 받을 수 있는 구성원은 공동 작업 작업에서 실행되는 모든 쿼리에 대한 로그도 받습니다. 쿼리에서 해당 데이터에 액세스하지 않았더라도 마찬가지입니다.

     1. 지원되는 로그 유형에서 공동 작업 생성자가 지원하도록 선택한 로그 유형 중에서 선택합니다.
        - SQL 쿼리에서 생성된 로그를 수신하려면 쿼리에서 로그 확인란을 선택합니다.
        - PySpark를 사용하여 작업에서 생성된 로그를 수신하려면 작업에서 로그 확인란을 선택합니다.
6. 변경 사항 저장을 선택합니다.

> **Note**
>
> 로깅을 켠 후 로그 스토리지를 설정하고 Amazon CloudWatch Logs에서 로그 수신을 시작하는 데 몇 분 정도 걸릴 수 있습니다. 이 짧은 기간 동안 쿼리를 할 수 있는 구성원이 실제로 로그를 전송하지 않는 쿼리를 실행할 수도 있습니다.

공동 작업 태그 편집

공동 작업 생성자는 공동 작업을 생성한 후 공동 작업 리소스에서 태그를 관리할 수 있습니다.

공동 작업 태그를 편집하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.

3. 생성한 공동 작업을 선택합니다.
4. 다음 중 하나를 선택합니다.

상황...	THEN ...
공동 작업 생성자 및 공동 작업 구성원	세부 정보 탭을 선택하십시오.
공동 작업 생성자이지만 공동 작업의 구성원은 아닙니다	페이지의 태그 섹션으로 스크롤을 내립니다.

5. 공동 작업 세부 정보를 보려면 태그 관리를 선택합니다.
6. 태그 관리 페이지에서 다음 작업을 수행할 수 있습니다.
   - 태그를 제거하려면 제거를 선택합니다.
   - 태그를 추가하려면 태그 추가(Add new tag)를 선택합니다
   - 변경 사항을 저장하려면 변경 사항 저장을 선택합니다.

멤버십 태그 편집

공동 작업 생성자는 공동 작업을 생성한 후 멤버십 리소스에서 태그를 관리할 수 있습니다.

멤버십 태그를 편집하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 생성한 공동 작업을 선택합니다.
4. 세부 정보 탭을 선택하십시오.
5. 멤버십 세부 정보에서 태그 관리를 선택합니다.
6. 태그 멤버십 관리 페이지에서 다음 작업을 수행할 수 있습니다.
   - 태그를 제거하려면 제거를 선택합니다.
   - 태그를 추가하려면 태그 추가를 선택합니다.
   - 변경 사항을 저장하려면 변경 사항 저장을 선택합니다.

관련 테이블 태그 편집

공동 작업 생성자는 테이블을 공동 작업에 연결한 후 연결된 테이블 리소스의 태그를 관리할 수 있습니다.

관련 테이블 태그를 편집하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 생성한 공동 작업을 선택합니다.
4. 테이블 탭을 선택합니다.
5. 귀하와 연결된 테이블의 경우 테이블을 선택하세요.
6. 구성된 테이블 세부 정보 페이지에서 태그의 경우, 태그 관리를 선택합니다.

   태그 관리 페이지에서 다음 작업을 수행할 수 있습니다.

   - 태그를 제거하려면 제거를 선택합니다.
   - 태그를 추가하려면 태그 추가를 선택합니다.
   - 변경 사항을 저장하려면 변경 사항 저장을 선택합니다.

분석 템플릿 태그 편집

공동 작업 생성자는 공동 작업을 생성한 후 분석 템플릿 리소스에서 태그를 관리할 수 있습니다.

멤버십 태그를 편집하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 생성한 공동 작업을 선택합니다.
4. 템플릿 탭을 선택합니다.
5. 섹션에서 생성한 분석 템플릿에서 분석 템플릿을 선택합니다.
6. 분석 템플릿 테이블 세부 정보 페이지에서 태그 섹션까지 아래로 스크롤합니다.
7. 태그 관리를 선택합니다.

8. 태그 관리 페이지에서 다음 작업을 수행할 수 있습니다.

   - 태그를 제거하려면 제거를 선택합니다.
   - 태그를 추가하려면 태그 추가를 선택합니다.
   - 변경 사항을 저장하려면 변경 사항 저장을 선택합니다.

차등 프라이버시 정책 태그 편집

공동 작업 생성자는 공동 작업을 생성한 후 분석 템플릿 리소스에서 태그를 관리할 수 있습니다.

멤버십 태그를 편집하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 편집할 차등 프라이버시 정책이 포함된 공동 작업을 선택합니다.
4. 테이블 탭을 선택합니다.
5. 테이블 탭에서 태그 관리를 선택합니다.
6. 태그 관리 페이지에서 다음 작업을 수행할 수 있습니다.

   - 태그를 제거하려면 제거를 선택합니다.
   - 태그를 추가하려면 태그 추가를 선택합니다.
   - 변경 사항을 저장하려면 변경 사항 저장을 선택합니다.

공동 작업 삭제

공동 작업 생성자는 자신이 만든 공동 작업을 삭제할 수 있습니다.

> **Note**
>
> 공동 작업을 삭제하면 나와 모든 구성원은 쿼리를 실행하거나 결과를 받거나 데이터를 제공할 수 없습니다. 각 공동 작업 구성원은 멤버십의 일환으로 자신의 데이터에 계속 접근할 수 있습니다.

공동 작업을 삭제하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 삭제하려는 공동 작업을 선택합니다.
4. 작업에서 공동 작업 삭제를 선택합니다.
5. 삭제를 확인한 다음 삭제를 선택합니다.

공동 작업 보기

공동 작업 생성자는 자신이 만든 모든 협업을 볼 수 있습니다.

공동 작업을 보려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업 페이지의 마지막 사용에서 최근에 사용한 5개의 공동 작업을 확인합니다.
4. 활성 멤버십 사용 탭에서 활성 멤버십을 사용한 공동 작업 목록을 볼 수 있습니다.

   이름, 멤버십 생성 날짜, 회원 세부 정보별로 정렬할 수 있습니다.

   검색 창을 사용하여 공동 작업을 검색할 수 있습니다.
5. 참여 가능 탭에서 참여할 수 있는 공동 작업 목록을 확인하세요.
6. 더 이상 사용할 수 없음 탭에서 삭제된 공동 작업 목록과 더 이상 사용할 수 없는 공동 작업의 멤버십(제거된 멤버십)을 볼 수 있습니다.

공동 작업에 멤버 초대

공동 작업 생성자는 공동 작업을 생성한 후 멤버 탭에 나열된 멤버에게 초대 링크를 보낼 수 있습니다.

공동 작업에 멤버를 초대하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).

2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 생성한 공동 작업을 선택합니다.
4. 구성원 탭을 선택합니다.
5. 멤버 테이블에서 초대 복사 링크 버튼을 선택합니다.

   초대 링크가 복사됩니다.
6. 선택한 보안 통신 방법에 초대 링크를 붙여넣고 공동 작업의 각 구성원에게 보냅니다.

구성원 모니터링

공동 작업 생성자는 공동 작업을 생성한 후 구성원 탭에서 모든 구성원의 상태를 모니터링할 수 있습니다.

구성원의 상태를 확인하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 수행하지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 생성한 공동 작업을 선택합니다.
4. 구성원 탭을 선택합니다.
5. 구성원 테이블에서 각 구성원의 상태를 검토합니다.
6. 멤버 자격 테이블에서 쿼리, 결과 수신, 데이터 제공 및 기타 작업을 수행할 수 있는 구성원을 검토합니다.
7. 결제 구성 테이블에서 쿼리, ID 매핑 테이블 및 ML 모델링에 대해 비용을 지불하는 구성원을 검토합니다.

공동 작업에서 구성원 제거

> **Note**
>
> 구성원을 제거하면 관련 데이터 세트도 모두 공동 작업에서 제거됩니다.

공동 작업에서 구성원을 제거하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 생성한 공동 작업을 선택합니다.
4. 구성원 탭을 선택합니다.
5. 제거할 멤버 옆에 있는 옵션 버튼을 선택합니다.

> **Note**
>
> 공동 작업 생성자는 자신의 계정 ID를 선택할 수 없습니다.

6. 제거를 선택합니다.
7. 대화 상자에서 텍스트 입력 필드에 **`confirm`**을 입력하여 구성원을 제거할지 여부를 확인합니다.

> **Note**
>
> 쿼리 컴퓨팅 비용을 지불하는 구성원을 제거하면 공동 작업에서 더 이상 쿼리를 실행할 수 없습니다.

공동 작업 탈퇴

공동 작업 구성원은 멤버십을 삭제하여 공동 작업에서 탈퇴할 수 있습니다. 공동 작업 생성자인 경우 공동 작업을 삭제해야만 공동 작업을 탈퇴할 수 있습니다.

> **Note**
>
> 멤버십을 삭제하면 공동 작업에서 탈퇴하고 다시 가입할 수 없습니다. 쿼리 컴퓨팅 비용을 지불하는 회원이고 멤버십을 삭제하면 더 이상 쿼리를 실행할 수 없습니다.

공동 작업에서 탈퇴하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 수행하지 않은 경우).

2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 활성 멤버십 포함에서 구성원으로 속해 있는 공동 작업을 선택합니다.
4. 작업을 선택합니다.
5. 멤버십 삭제를 선택합니다.
6. 대화 상자에서 텍스트 입력 필드를 **confirm**을 입력하여 공동 작업에서 탈퇴할지 여부를 확인한 다음 비우기 및 멤버십 삭제를 선택합니다.

   콘솔에 멤버십이 삭제되었다는 메시지가 표시됩니다.

   공동 작업 생성자는 구성원 상태를 왼쪽으로 볼 수 있습니다.

에서 데이터 테이블 준비 AWS Clean Rooms

> **Note**
>
> 공동 작업에 참여하기 전이나 후에 데이터 테이블을 준비할 수 있습니다. 테이블이 준비된 후에는 해당 테이블에 대한 프라이버시 요구 사항이 동일하다면 여러 공동 작업에서 테이블을 재사용할 수 있습니다.

공동 작업의 구성원은 쿼리할 수 있는 공동 작업 구성원 AWS Clean Rooms 이 데이터 테이블을 쿼리하기 전에 먼저 준비해야 합니다.

에서 쿼리에 사용하는 데이터 테이블 AWS Clean Rooms 은 일반적으로 다른 애플리케이션에 사용하는 것과 동일한 유형의 데이터 테이블입니다. 예를 들어, 동일한 유형의 데이터 세트가 Amazon Athena, Amazon EMR, Amazon Redshift Spectrum 및 Amazon QuickSight에 사용됩니다.

다음 데이터 소스에서 직접 원래 형식으로 데이터를 쿼리할 수 있습니다.

- Amazon Simple Storage Service(S3)
- Amazon Athena
- Snowflake

AWS Clean Rooms 는 쿼리 실행 시 데이터 세트에 액세스하여 쿼리할 수 있는 멤버가 항상 up-to-date 데이터에 액세스할 수 있도록 합니다. AWS Clean Rooms 공동 작업에 일시적으로 읽는 모든 데이터는 쿼리가 완료된 후 삭제됩니다. 쿼리 결과는 Amazon S3 버킷에 기록됩니다.

사용 사례에 ID 데이터 쿼리가 포함된 경우 AWS Entity Resolution in AWS Clean Rooms 단원을 참조하세요.

주제

- 의 데이터 형식 AWS Clean Rooms
- Apache Iceberg의 테이블 AWS Clean Rooms
- 에서 쿼리를 위한 데이터 테이블 준비 AWS Clean Rooms
- Clean Rooms에 대한 암호화 컴퓨팅으로 암호화된 데이터 테이블 준비
- C3R 암호화 클라이언트를 사용한 데이터 테이블 복호화

의 데이터 형식 AWS Clean Rooms

데이터를 분석하려면 데이터 세트가에서 AWS Clean Rooms 지원하는 형식이어야 합니다.

주제

- PySpark 작업에 지원되는 데이터 형식
- SQL 쿼리에 지원되는 데이터 형식
- 지원되는 데이터 유형
- 에 대한 파일 압축 유형 AWS Clean Rooms
- 에 대한 서버 측 암호화 AWS Clean Rooms

PySpark 작업에 지원되는 데이터 형식

AWS Clean Rooms 는 PySpark 작업을 실행하기 위해 다음과 같은 구조화된 형식을 지원합니다.

- Parquet
- OpenCSV
- JSON

SQL 쿼리에 지원되는 데이터 형식

AWS Clean Rooms 는 Spark SQL 분석 엔진 또는 SQL 분석 엔진을 선택하는지 여부에 따라 AWS Clean Rooms SQL 쿼리를 실행하기 위한 다양한 구조화된 형식을 지원합니다.

Spark SQL analytics engine

- Apache Iceberg 테이블
- Parquet
- OpenCSV
- JSON

AWS Clean Rooms SQL analytics engine

- Apache Iceberg 테이블
- Parquet

- RCFile
- TextFile
- SequenceFile
- RegexSerde
- OpenCSV
- AVRO
- JSON

> **Note**
>
> 텍스트 파일의 `timestamp` 값은 `yyyy-MM-dd HH:mm:ss.SSSSSS` 형식이어야 합니다. 예: `2017-05-01 11:30:59.000000`.

Apache Parquet 같은 컬럼 형식 스토리지 파일을 사용하는 것이 좋습니다. 열 기반 스토리지 파일 형식을 사용하면 필요한 열만 선택하여 데이터 이동을 최소화할 수 있습니다. 최적의 성능을 위해 대형 오브젝트는 100mb~1gb 오브젝트로 분할해야 합니다.

지원되는 데이터 유형

AWS Clean Rooms 는 Spark SQL 분석 엔진 또는 AWS Clean Rooms SQL 분석 엔진을 선택하는지 여부에 따라 다양한 유형을 지원합니다.

Spark SQL analytics engine

- ARRAY
- BIGINT
- BOOLEAN
- BYTE
- CHAR
- 날짜
- DECIMAL
- FLOAT
- INTEGER
- INTERVAL

- LONG
- MAP
- REAL
- SHORT
- SMALLINT
- STRUCT
- TIME
- TIMESTAMP_LTZ
- TIMESTAMP_NTZ
- TINYINT
- VARCHAR

자세한 내용은 AWS Clean Rooms SQL 참조의 데이터 유형을 참조하세요.

AWS Clean Rooms SQL

- ARRAY
- BIGINT
- BOOLEAN
- CHAR
- 날짜
- DECIMAL
- DOUBLE PRECISION
- INTEGER
- MAP
- REAL
- SMALLINT
- STRUCT
- SUPER
- TIME
- TIMESTAMP
- TIMESTAMPTZ

- TIMETZ
- VARBYTE
- VARCHAR

자세한 내용은 AWS Clean Rooms SQL 참조의 데이터 유형을 참조하세요.

에 대한 파일 압축 유형 AWS Clean Rooms

스토리지 스페이스를 줄이고 성능을 높이며 비용을 최소화하려면 데이터 세트를 압축하는 것이 좋습니다.

AWS Clean Rooms 는 파일 확장명을 기반으로 파일 압축 유형을 인식하고 다음 표에 표시된 압축 유형 및 확장명을 지원합니다.

압축 알고리즘	파일 확장명
GZIP	.gz
Bzip2	.bz2
Snappy	.snappy

여러 레벨에서 압축을 적용할 수 있습니다. 일반적으로 전체 파일을 압축하거나 파일 내의 개별 블록을 압축합니다. 파일 수준에서 열 형식을 압축해도 성능상의 이점이 없습니다.

에 대한 서버 측 암호화 AWS Clean Rooms

> **Note**
>
> 서버측 암호화는 암호화 컴퓨팅을 필요로 하는 사용 사례에서 암호화 컴퓨팅을 대체하지 않습니다.

AWS Clean Rooms 는 다음 암호화 옵션을 사용하여 암호화된 데이터 세트를 투명하게 해독합니다.

- SSE-S3 - Amazon S3에서 관리하는 AES-256 암호화 키를 사용하는 서버 측 암호화
- SSE-KMS -에서 관리하는 키를 사용한 서버 측 암호화 AWS Key Management Service

SSE-S3를 사용하려면 구성된 테이블을 공동 작업에 연결하는 데 사용되는 AWS Clean Rooms 서비스 역할에 KMS 암호 해독 권한이 있어야 합니다. SSE-KMS를 사용하려면 KMS 키 정책에서 AWS Clean Rooms 서비스 역할의 암호 해독도 허용해야 합니다.

AWS Clean Rooms 는 Amazon S3 클라이언트 측 암호화를 지원하지 않습니다. 서버 측 암호화에 대한 자세한 내용은 Amazon Simple Storage Service 사용 설명서에서 서버 측 암호화를 사용하여 데이터 보호를 참조하세요.

Apache Iceberg의 테이블 AWS Clean Rooms

Apache Iceberg는 데이터 lakes에 대한 오픈 소스 테이블 형식입니다.는 Apache Iceberg 메타데이터에 저장된 통계를 사용하여 쿼리 계획을 최적화하고 클린룸 쿼리 처리 중에 파일 스캔을 줄일 AWS Clean Rooms 수 있습니다. 자세한 내용은 Apache Iceberg를 참조하세요.

Iceberg 테이블과 AWS Clean Rooms 함께를 사용할 때는 다음 사항을 고려하세요.

- S3용 Apache Iceberg 테이블 - Apache Iceberg 테이블은 오픈 소스 글루 카탈로그 구현을 AWS Glue Data Catalog 기반으로에 정의되어야 합니다.
- Athena용 Apache Iceberg 테이블 - 자세한 내용은 https://docs.aws.amazon.com/athena/latest/ug/querying-iceberg.html 참조하세요.
- Snowflake용 Apache Iceberg 테이블 - 자세한 내용은 https://docs.snowflake.com/en/user-guide/tables-iceberg 참조하세요.
- Parquet 파일 형식 -는 Parquet 데이터 파일 형식의 Iceberg 테이블 AWS Clean Rooms 만 지원합니다.
- GZIP 및 Snappy 압축 - GZIP 및 Snappy 압축을 통해 Parquet를 AWS Clean Rooms 지원합니다.
- Iceberg 버전 - 버전 1 및 버전 2 Iceberg 테이블에 대한 쿼리 실행을 AWS Clean Rooms 지원합니다.
- 파티션 -의 Apache Iceberg 테이블에 파티션을 수동으로 추가할 필요가 없습니다 AWS Glue.는 Apache Iceberg 테이블의 새 파티션을 자동으로 AWS Clean Rooms 감지하므로 테이블 정의의 파티션을 업데이트하는 데 수동 작업이 필요하지 않습니다. Iceberg 파티션은 구성된 AWS Clean Rooms 테이블 스키마에서 파티션 키로 별도로 표시되지 않고 테이블 스키마에서 일반 열로 나타납니다.
- 제한 사항
  - 새 아이스버그 테이블에만 해당

    Apache Parquet 테이블에서 변환된 Apache Iceberg 테이블은 지원되지 않습니다.
  - 시간 이동 쿼리

AWS Clean Rooms 는 Apache Iceberg 테이블을 사용한 시간 이동 쿼리를 지원하지 않습니다.

- Athena 엔진 버전 2

  Athena 엔진 버전 2로 생성된 Iceberg 테이블은 지원되지 않습니다.

- 파일 형식

  Avro 및 Optimized Row Columnar(ORC) 파일 형식은 지원되지 않습니다.

- 압축

  Parquet에 대한 Zstandard(Zstd) 압축은 지원되지 않습니다.

Iceberg 테이블에 대해 지원되는 데이터 형식

AWS Clean Rooms 는 다음 데이터 형식이 포함된 Iceberg 테이블을 쿼리할 수 있습니다.

- BOOLEAN
- DATE
- DECIMAL
- DOUBLE
- FLOAT
- INT
- LIST
- LONG
- MAP
- STRING
- STRUCT
- TIMESTAMP WITHOUT TIME ZONE

Iceberg 데이터 형식에 대한 자세한 내용은 Apache Iceberg 설명서에서 Iceberg용 스키마를 참조하세요.

에서 쿼리를 위한 데이터 테이블 준비 AWS Clean Rooms

사용 사례에서 자체 데이터를 가져오지 않아도 되는 경우에는 이 절차를 건너뛸 수 있습니다.

사용 사례에 ID 데이터 쿼리가 포함된 경우 AWS Entity Resolution in AWS Clean Rooms 단원을 참조하세요.

사용할 수 있는 데이터 형식에 대한 자세한 내용은 섹션을 참조하세요의 데이터 형식 AWS Clean Rooms.

주제

- Amazon S3에서 데이터 테이블 준비
- Amazon Athena에서 데이터 테이블 준비
- Snowflake에서 데이터 테이블 준비

Amazon S3에서 데이터 테이블 준비

에서 카탈로그화 AWS Glue 되고 Amazon S3에 저장된 데이터 테이블을 분석할 수 있습니다. 데이터 테이블이 이미에 카탈로그화되어 있는 경우 로 AWS Glue건너뜁니다AWS Clean Rooms에서 구성된 테이블 생성하기.

Amazon S3에서 데이터 테이블을 준비하려면 다음 단계를 수행해야 합니다.

주제

- 1단계: 필수 구성 요소 완성
- 2단계: (선택 사항) 암호화 컴퓨팅용 데이터 준비
- 3단계: 데이터 테이블을 Amazon S3에 업로드
- 4단계: AWS Glue 테이블 생성
- 5단계: 다음 절차

1단계: 필수 구성 요소 완성

에서 사용할 데이터 테이블을 준비하려면 다음 사전 조건을 완료해야 AWS Clean Rooms합니다.

- 데이터 테이블은 지원되는 데이터 형식 AWS Clean Rooms 중 하나로 저장됩니다.
- 데이터 테이블은에서 카탈로그화 AWS Glue 되며 지원되는 데이터 형식을 AWS Clean Rooms 사용합니다.
- 모든 데이터 테이블은 공동 작업이 생성된 동일한의 Amazon Simple Storage Service(Amazon S3) AWS 리전 에 저장됩니다.
- AWS Glue Data Catalog 는 공동 작업이 생성된 리전과 동일한 리전에 있습니다.

- AWS Glue Data Catalog 는 멤버십 AWS 계정 과 동일한에 있습니다.
- Amazon S3 버킷이에 등록되지 않았습니다 AWS Lake Formation.

2단계: (선택 사항) 암호화 컴퓨팅용 데이터 준비

(선택 사항) 암호화 컴퓨팅을 사용 중이고 데이터 테이블에 암호화하려는 민감한 정보가 포함되어 있는 경우 C3R 암호화 클라이언트를 사용하여 데이터 테이블을 암호화해야 합니다.

암호화 컴퓨팅에 사용할 데이터를 준비하려면 Clean Rooms에 대한 암호화 컴퓨팅으로 암호화된 데이터 테이블 준비의 절차를 따릅니다.

3단계: 데이터 테이블을 Amazon S3에 업로드

> **Note**
>
> 공동 작업에서 암호화된 데이터 테이블을 사용하려는 경우, 데이터 테이블을 Amazon S3에 업로드하기 전에 먼저 암호화 컴퓨팅을 위해 데이터를 암호화해야 합니다. 자세한 내용은 Clean Rooms에 대한 암호화 컴퓨팅으로 암호화된 데이터 테이블 준비 단원을 참조하십시오.

Amazon S3에 데이터 테이블을 업로드하려면

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/s3/://https://https://https://://https://://https://://https://://Amazon S3https://://://https://://://https://https://https
2. 버킷을 선택한 다음 데이터 테이블을 저장할 버킷을 선택합니다.
3. 업로드를 선택한 다음 안내를 따릅니다.
4. 개체 탭을 선택하여 데이터가 저장되는 접두사를 확인합니다. 폴더의 이름을 메모해 둡니다.

   데이터를 확인할 폴더를 선택할 수 있습니다.

4단계: AWS Glue 테이블 생성

AWS Glue 데이터 테이블이 이미 있는 경우이 단계를 건너뛸 수 있습니다.

이 단계에서는 S3 버킷의 모든 파일을 크롤링하고 AWS Glue 테이블을 생성하는에서 AWS Glue 크롤러를 설정합니다. 자세한 내용은 AWS Glue 개발자 안내서의 AWS Glue에서 크롤러 정의를 참조하세요.

지원되는 AWS Glue Data Catalog 데이터 형식에 대한 자세한 내용은 섹션을 참조하세요지원되는 데이터 유형.

> **Note**
>
> AWS Clean Rooms 는 현재에 등록된 S3 버킷을 지원하지 않습니다 AWS Lake Formation.

다음 절차에서는 AWS Glue 테이블을 생성하는 방법을 설명합니다. 암호화된 AWS Glue Data Catalog 객체를 AWS Key Management Service (AWS KMS) 키와 함께 사용하려면 암호화된 테이블에 대한 액세스를 허용하도록 KMS 키 권한 정책을 구성해야 합니다. 자세한 내용은AWS Glue 개발자 설명서의 AWS Glue에서의 암호화 설정을 참조하세요.

AWS Glue 테이블을 생성하려면

1. AWS Glue 사용 설명서의 AWS Glue 콘솔에서 크롤러 작업 절차를 따릅니다.
2. AWS Glue 데이터베이스 이름과 AWS Glue 테이블 이름을 기록해 둡니다.

5단계: 다음 절차

이제 Amazon S3에서 데이터 테이블을 준비했으므로 다음을 수행할 준비가 되었습니다.

- 구성된 테이블 생성
- ML 모델 생성

테이블은 다음 후에 쿼리할 수 있습니다.

- 공동 작업 생성자가 AWS Clean Rooms에서 공동 작업을 설정했습니다. 자세한 내용은 공동 작업 생성 단원을 참조하십시오.
- 공동 작업 생성자가 공동 작업 참여자인 사용자에게 공동 작업 ID를 전송했습니다.

Amazon Athena에서 데이터 테이블 준비

Amazon Athena에서 AWS Glue Data Catalog (GDC) 뷰로 생성된 데이터 테이블을 쿼리할 수 있습니다.

GDC View는 하나 이상의 기본 테이블에서 생성된 가상 AWS Glue 테이블입니다. Athena `AwsGlueCatalog` 카탈로그에서 Athena SQL을 사용하여 생성해야 합니다.

Amazon Athena에서 데이터 테이블을 준비하려면 다음 단계를 수행해야 합니다.

주제

- 1단계: 필수 구성 요소 완성
- 2단계: (선택 사항) 암호화 컴퓨팅용 데이터 준비
- 3단계: 다음 단계

1단계: 필수 구성 요소 완성

에서 사용할 데이터 테이블을 준비하려면 다음 사전 조건을 완료해야 AWS Clean Rooms합니다.

- 데이터 테이블은 지원되는 데이터 형식 AWS Clean Rooms 중 하나로 저장됩니다.
- 데이터 테이블은 지원되는 데이터 형식을 AWS Clean Rooms 사용합니다.
- Athena `AwsDataCatalog` 카탈로그에서 Athena SQL을 사용하여 AWS Glue 테이블에 대한GDC 뷰를 생성했습니다.

  뷰는 다음과 같이 표시됩니다.

  - Athena 콘솔(아래`AwsDataCatalog`)을 보기로 표시: https://console.aws.amazon.com/athena/://
  - AWS Glue 테이블로서의 AWS Glue 콘솔: https://console.aws.amazon.com/glue/://

  자세한 내용은 Amazon Athena Athena 사용 설명서의 Athena에서 데이터 카탈로그 보기 사용을 참조하세요.

  > **Note**
  >
  > Athena 및에서 뷰를 생성하려면 적절한 권한이 필요합니다 AWS Glue. 또한 보기 정의에서 참조되는 기본 테이블에 액세스할 수 있는지 확인합니다.
  > AWS Clean Rooms 는 Lambda 또는 Hive AWS Glue 카탈로그 유형이 아닌 Athena에 대한 카탈로그 유형만 지원합니다.

- 데이터 테이블 또는GDC 보기는에 카탈로그화 AWS Glue 되어 있으며에 등록됩니다 AWS Lake Formation.
- Athena 결과를 수신하기 위해 Amazon S3에 별도의 출력 버킷을 생성했습니다.
- Amazon Athena에서 데이터를 읽도록 서비스 역할을 설정했습니다. 자세한 내용은 Amazon Athena에서 데이터를 읽을 서비스 역할 생성 단원을 참조하십시오.

- 서비스 역할에는GDC 보기 또는 테이블에 대한 Lake Formation Select 및 Describe 액세스 권한이 있습니다.

2단계: (선택 사항) 암호화 컴퓨팅용 데이터 준비

(선택 사항) 암호화 컴퓨팅을 사용 중이고 데이터 테이블에 암호화하려는 민감한 정보가 포함되어 있는 경우 C3R 암호화 클라이언트를 사용하여 데이터 테이블을 암호화해야 합니다.

암호화 컴퓨팅에 사용할 데이터를 준비하려면 Clean Rooms에 대한 암호화 컴퓨팅으로 암호화된 데이터 테이블 준비의 절차를 따릅니다.

3단계: 다음 단계

Amazon Athena에서 데이터 테이블을 준비했으므로 이제 다음을 수행할 준비가 되었습니다.

- 구성된 테이블 생성
- ML 모델 생성

테이블은 다음 후에 쿼리할 수 있습니다.

- 공동 작업 생성자가 AWS Clean Rooms에서 공동 작업을 설정했습니다. 자세한 내용은 공동 작업 생성 단원을 참조하십시오.
- 공동 작업 생성자가 공동 작업 참여자인 사용자에게 공동 작업 ID를 전송했습니다.

Snowflake에서 데이터 테이블 준비

Snowflake 데이터 웨어하우스에 저장된 데이터 테이블을 쿼리할 수 있습니다.

Snowflake에서 데이터 테이블을 준비하려면 다음 단계를 수행해야 합니다.

주제

- 1단계: 필수 구성 요소 완성
- 2단계: (선택 사항) 암호화 컴퓨팅용 데이터 준비
- 3단계: 보안 암호 AWS Secrets Manager 생성
- 4단계: 다음 단계

1단계: 필수 구성 요소 완성

에서 사용할 데이터 테이블을 준비하려면 다음 사전 조건을 완료해야 AWS Clean Rooms합니다.

- 데이터 테이블을 읽을 수 AWS 계정 있는 적절한 권한이 부여된가 있습니다. 자세한 내용은 서비스 역할을 생성하여 Snowflake에서 데이터 읽기 단원을 참조하십시오.
- 데이터 테이블은 지원되는 데이터 형식 AWS Clean Rooms 중 하나로 저장됩니다.
- 데이터 테이블은 지원되는 데이터 형식을 AWS Clean Rooms 사용합니다.
- 데이터 테이블은 Snowflake 웨어하우스에 저장됩니다. 자세한 내용은 Snowflake 설명서를 참조하세요.
- 공동 작업과 연결할 Snowflake 테이블에 대한 읽기 전용 권한을 가진 새 Snowflake 사용자를 설정했습니다.

2단계: (선택 사항) 암호화 컴퓨팅용 데이터 준비

(선택 사항) 암호화 컴퓨팅을 사용 중이고 데이터 테이블에 암호화하려는 민감한 정보가 포함되어 있는 경우 C3R 암호화 클라이언트를 사용하여 데이터 테이블을 암호화해야 합니다.

암호화 컴퓨팅에 사용할 데이터를 준비하려면 Clean Rooms에 대한 암호화 컴퓨팅으로 암호화된 데이터 테이블 준비의 절차를 따릅니다.

3단계: 보안 암호 AWS Secrets Manager 생성

에서 Snowflake에 연결하려면 Snowflake 자격 증명을 생성하여 AWS Secrets Manager 보안 암호에 저장한 다음 해당 보안 암호를의 Snowflake 테이블과 연결 AWS Clean Rooms해야 합니다 AWS Clean Rooms.

> **Note**
>
> 전용인 새 사용자를 생성하는 것이 좋습니다 AWS Clean Rooms. 해당 사용자에게는 액세스 AWS Clean Rooms 하려는 데이터에 대한 읽기 권한이 있는 역할만 있어야 합니다.

AWS Secrets Manager 보안 암호를 생성하려면

1. Snowflake에서 사용자 `snowflakeUser` 및 암호를 생성합니다`snowflakePassword`.

2. 이 사용자가 상호 작용할 Snowflake 웨어하우스를 결정합니다snowflakeWarehouse. SnowflakesnowflakeUser에서에 DEFAULT_WAREHOUSE 대한 로 설정하거나 다음 단계에서 기억합니다.
3. AWS Secrets Manager에서 Snowflake 보안 인증을 사용하여 보안 암호를 생성합니다. Secrets Manager에서 보안 암호를 생성하려면 AWS Secrets Manager 사용 설명서의 AWS Secrets Manager 보안 암호 생성에서 제공되는 자습서를 따르세요. 보안 암호를 생성한 후 다음 단계를 secretName 위해 보안 암호 이름을 유지합니다.
   - 키/값 페어를 선택할 때 키를 snowflakeUser 사용하여에 대한 페어를 생성합니다sfUser.
   - 키/값 페어를 선택할 때 키를 snowflakePassword 사용하여에 대한 페어를 생성합니다sfPassword.
   - 키/값 페어를 선택할 때 키를 snowflakeWarehouse 사용하여에 대한 페어를 생성합니다sfWarehouse.

     Snowflake에 기본값이 설정된 경우에는 필요하지 않습니다. Snowflake에 기본값이 설정된 경우에는 필요하지 않습니다.
   - 키/값 페어를 선택할 때 키 snowflakeRole를 사용하여에 대한 페어를 생성합니다sfrole.

4단계: 다음 단계

이제 Snowflake에서 데이터 테이블을 준비했으므로 다음을 수행할 준비가 되었습니다.

- 구성된 테이블 생성
- ML 모델 생성

테이블은 다음 후에 쿼리할 수 있습니다.

- 공동 작업 생성자가 AWS Clean Rooms에서 공동 작업을 설정했습니다. 자세한 내용은 공동 작업 생성 단원을 참조하십시오.
- 공동 작업 생성자가 공동 작업 참여자인 사용자에게 공동 작업 ID를 전송했습니다.

Clean Rooms에 대한 암호화 컴퓨팅으로 암호화된 데이터 테이블 준비

Clean Rooms (C3R)에 대한 암호화 컴퓨팅은의 기능입니다 AWS Clean Rooms. C3R을 사용하여 당사자와 공동 작업 AWS 에서 학습할 수 있는 내용을 암호화 방식으로 제한할 수 있습니다 AWS Clean Rooms .

Amazon Simple Storage Service(Amazon S33), Amazon Athena 또는 Snowflake 데이터 소스에 데이터 테이블을 업로드하기 전에 클라이언트 측 암호화 도구인 C3R 암호화 클라이언트를 사용하여 데이터 테이블을 암호화할 수 있습니다. Amazon Athena

자세한 내용은 Clean Rooms에 대한 암호화 컴퓨팅 단원을 참조하십시오.

C3R로 암호화된 데이터 테이블을 준비하려면 다음 단계를 거쳐야 합니다.

단계

- 1단계: 필수 구성 요소 완성
- 2단계: C3R 암호화 클라이언트 다운로드
- 3단계: (선택 사항) C3R 암호화 클라이언트에서 사용 가능한 명령 보기
- 4단계: 표 형식 파일의 암호화 스키마 생성
- 5단계: 공유 암호 키 생성
- 6단계: 환경 변수에 공유 암호 키 저장
- 7단계: 데이터 암호화
- 8단계: 데이터 암호화 확인
- (선택 사항) 스키마 생성(고급 사용자)

1단계: 필수 구성 요소 완성

C3R에 사용할 수 있는 데이터 테이블을 준비하려면 다음과 같은 사전 조건을 완료해야 합니다.

- GitHub의 Clean Rooms 리포지토리를 위한 암호화 컴퓨팅에 액세스할 수 있습니다.

  https://github.com/aws/c3r
- C3R 암호화 클라이언트를 사용하도록 AWS 자격 증명을 설정했습니다. 이러한 자격 증명은 C3R 암호화 클라이언트가에 대한 읽기 전용 API 호출을 통해 공동 작업 메타데이터를 검색하는 AWS

Clean Rooms 데 사용됩니다. 자세한 내용은AWS Command Line Interface 버전 2 사용 설명서AWS CLI의 구성을 참조하세요.

- 컴퓨터에 Java Runtime Environment(JRE) 11 이상이 설치되어 있습니다.
  - 권장 Java Runtime Environment 제품인 Amazon Corretto 11 이상은 https://aws.amazon.com/corretto에서 다운로드할 수 있습니다.
  - Java Development Kit(JDK)에는 동일한 JRE 버전의 해당 버전이 포함되어 있습니다. 그러나 JDK의 추가 기능은 Clean Rooms에 대한 암호화 클라이언트(C3R)를 실행하는 데 필요하지 않습니다.
- 표 형식 데이터 파일(.csv) 또는 Parquet 파일(.parquet)은 로컬에 저장됩니다.
- 본인 또는 공동 작업에 참여한 다른 구성원은 공유 비밀 키를 생성할 수 있습니다. 자세한 내용은 5단계: 공유 암호 키 생성 단원을 참조하십시오.
- 공동 작업 생성자가에서 공동 작업에 대해 활성화된 암호화 컴퓨팅 AWS Clean Rooms 을 사용하여 공동 작업을 생성했습니다. 자세한 내용은 공동 작업 생성 단원을 참조하십시오.
- 공동 작업 생성자가 공동 작업 참여자인 귀하에게 공동 작업 ID를 전송했습니다. 전송된 초대에는 공동 작업 Amazon 리소스 이름(ARN)이 포함되며, 초대장에는 공동 작업 ID가 포함됩니다.

2단계: C3R 암호화 클라이언트 다운로드

GitHub에서 C3R 암호화 클라이언트를 다운로드하려는 경우

1. Clean Rooms AWS GitHub 리포지토리용 암호화 컴퓨팅: https://github.com/aws/c3r 이동합니다.
2. 파일을 선택하고 다운로드합니다.

   소스 코드, 라이선스 및 관련 자료는 GitHub 리포지토리 랜딩 페이지의 파일로부터 .zip으로 복제하거나 다운로드할 수 있습니다. (리포지토리 콘텐츠 목록 오른쪽 상단의 코드 버튼 참조).

   가장 최근에 서명된 C3R 암호화 클라이언트 Java Executable File(즉, 명령줄 인터페이스 응용 프로그램)은 GitHub 리포지토리의 릴리스 페이지에 있습니다.

   Apache Spark용 C3R 암호화 클라이언트 패키지(`c3r-cli-spark`)는 실행 중인 Apache Spark 서버에 작업으로 제출해야 하는 c3r-cli의 버전입니다. 자세한 내용은 Apache Spark에서 C3R 실행을 참조하세요.

3단계: (선택 사항) C3R 암호화 클라이언트에서 사용 가능한 명령 보기

이 절차를 사용하여 C3R 암호화 클라이언트에서 사용 가능한 명령을 익히세요.

C3R 암호화 클라이언트에서 사용 가능한 모든 명령을 보려는 경우

1. 명령줄 인터페이스(CLI) 에서 다운로드한 c3r-cli.jar 파일이 있는 폴더로 이동합니다.
2. `java -jar c3r-cli.jar` 명령을 실행합니다.
3. 사용 가능한 명령 및 옵션을 확인합니다.

4단계: 표 형식 파일의 암호화 스키마 생성

데이터를 암호화하려면 데이터 사용 방법을 설명하는 암호화 스키마가 필요합니다. 이 섹션에서는 C3R 암호화 클라이언트가 헤더 행 또는 Parquet 파일이 있는 CSV 파일의 암호화 스키마를 생성하는 데 어떤 도움을 주는지 설명합니다.

요청은 파일당 한 번만 하면 됩니다. 스키마가 존재하면 이를 다시 사용하여 동일한 파일(또는 열 이름이 동일한 파일)을 암호화할 수 있습니다. 열 이름이나 원하는 암호화 스키마가 변경되면 스키마 파일을 업데이트해야 합니다. 자세한 내용은 (선택 사항) 스키마 생성(고급 사용자) 단원을 참조하십시오.

> **Important**
>
> 모든 공동 작업 당사자가 동일한 공유 비밀 키를 사용하는 것이 가장 중요합니다. 또한 공동 작업 당사자는 쿼리에서 동일성을 유지하기 위해 열 이름이 JOIN될지 또는 다른 방식으로 비교될지 일치하도록 열 이름을 조정해야 합니다. 그렇지 않으면 SQL 쿼리에서 예상치 못한 결과가 나오거나 잘못된 결과가 나올 수 있습니다. 하지만 공동 작업 생성자가 공동 작업 생성 중에 `allowJoinsOnColumnsWithDifferentNames` 암호화 설정을 활성화한 경우에는 이 방법이 필요하지 않습니다. 암호화 관련 설정에 대한 자세한 내용은 암호화 컴퓨팅 파라미터 섹션을 참조하세요.

스키마 모드에서 실행할 경우 C3R 암호화 클라이언트는 입력 파일을 열별로 검토하여 해당 열을 처리할지 여부와 처리 방법을 묻는 메시지를 표시합니다. 파일에 암호화된 출력에 필요하지 않은 열이 많이 포함된 경우 원하지 않는 각 열을 건너뛰어야 하므로 대화형 스키마 생성이 번거로울 수 있습니다. 이를 방지하려면 스키마를 수동으로 작성하거나 원하는 열만 포함하는 입력 파일의 단순화된 버전을 만들 수 있습니다. 그러면 축소된 파일에서 대화형 스키마 생성기를 실행할 수 있습니다. C3R 암호화 클라이언트는 스키마 파일에 대한 정보를 출력하고 대상 출력에 소스 열을 포함하거나 암호화하는 방법(있는 경우)을 묻습니다.

입력 파일의 각 소스 열에 대해 를 묻는 메시지가 표시됩니다.

1. 생성해야 하는 대상 열의 수

2. 각 대상 열을 암호화하는 방법(있는 경우)

3. 각 대상 열의 이름

4. 열을 sealed 열로 암호화하는 경우 암호화 전에 데이터를 패딩하는 방법

> **Note**
>
> sealed 열로 암호화된 열의 데이터를 암호화할 때는 패딩이 필요한 데이터를 결정해야 합니다. C3R 암호화 클라이언트는 스키마 생성 중에 열의 모든 항목을 동일한 길이로 채우는 기본 패딩을 제안합니다.
> `fixed`의 길이를 결정할 때는 패딩이 비트가 아니라 바이트 단위라는 점에 유의하세요.

다음은 스키마 생성에 대한 의사 결정 테이블입니다.

스키마 의사결정표

결정	소스 열<'열 이름'>의 대상 열 수?	대상 열 유형: [c] cleartext, [f] fingerprint, 또는 [s] sealed?	대상 열 헤더 이름 <default 'name-of-column'>	<suffix>암호화된 방식을 나타내는 접미사를 헤더에 추가합니다 ([y] 예 또는 [n] 아니요 <기본 '예'>	<'열 이름_봉인'> 패딩 유형: [n] 하나, [f] 고정 또는 최대 [m] <기본 '최대'>
열을 암호화하지 않은 상태로 두세요.	1	c	해당 사항 없음	해당 사항 없음	해당 사항 없음
열을 fingerprint 열로 암호화합니다.	1	f	기본값을 선택하거나 새 헤더 이름을 입력합니다.	y를 입력하여 기본값 (`_fingerprint`)을 선택하거나 n을 입력합니다.	해당 사항 없음

결정	소스 열<'열 이름'>의 대상 열 수?	대상 열 유형: [c] cleartext, [f] fingerprint, 또는 [s] sealed?	대상 열 헤더 이름 <default 'name-of-column'>	<suffix>암호화된 방식을 나타내는 접미사를 헤더에 추가합니다 ([y] 예 또는 [n] 아니요 <기본 '예'>	<'열 이름_봉인'> 패딩 유형: [n] 하나, [f] 고정 또는 최대 [m] <기본 '최대'>
열을 sealed 열로 암호화합니다.	1	s	기본값을 선택하거나 새 헤더 이름을 입력합니다.	y를 입력하여 기본값 (_sealed)을 선택하거나 n을 입력합니다.	패딩 유형을 선택합니다. 자세한 내용은 (선택 사항) 스키마 생성(고급 사용자) 단원을 참조하십시오.
열을 fingerprint 및 sealed로 모두 암호화합니다.	2	첫 번째 대상 열 입력: f. 두 번째 대상 열 입력: s.	각 대상 열의 대상 헤더를 선택합니다.	y를 입력하여 기본값을 선택하거나 n.을 입력합니다	패딩 유형을 선택합니다 (sealed 열만 해당). 자세한 내용은 (선택 사항) 스키마 생성(고급 사용자) 단원을 참조하십시오.

다음은 암호화 스키마 생성 방법에 대한 두 가지 예제입니다. 상호 작용의 정확한 내용은 입력 파일과 제공하는 응답에 따라 달라집니다.

예시

- 예: fingerprint 열과 cleartext 열에 대한 암호화 스키마 생성
- 예:sealed, fingerprint, cleartext 열을 사용하여 암호화 스키마 생성

예: fingerprint 열과 cleartext 열에 대한 암호화 스키마 생성

이 예제에서 ads.csv의 경우 열은 username 및 ad_variant 두 개뿐입니다. 이 열에는 다음이 필요합니다.

- username 열을 fingerprint 열로 암호화하려면
- ad_variant 열을 cleartext 열로 만들려면

fingerprint 열 및 cleartext 열에 대한 암호화 스키마를 생성하려는 경우

1. (선택 사항) 암호화할 c3r-cli.jar 파일 및 파일이 있는지 확인하려면:

   a. 원하는 디렉토리로 이동하여 ls(Mac 또는 Unix/Linux를 사용하는 경우) 또는 dir(Windows을 사용하는 경우)을 실행합니다.

   b. 테이블 형식 데이터 파일(예: .csv) 목록을 보고 암호화할 파일을 선택합니다.

      이 예에서는 ads.csv가 암호화하려는 파일입니다.

2. CLI에서 다음 명령을 실행하여 대화식으로 스키마를 생성합니다.

   ```
   java -jar c3r-cli.jar schema ads.csv --interactive --output=ads.json
   ```

   > **Note**
   >
   > - java --jar PATH/TO/c3r-cli.jar를 실행할 수 있습니다. 또는 CLASSPATH 환경 변수에 PATH/TO/c3r-cli.jar를 추가한 경우 클래스 이름을 실행할 수도 있습니다. C3R 암호화 클라이언트는 CLASSPATH를 검색하여 찾습니다 (예:java com.amazon.psion.cli.Main).
   > - --interactive 플래그는 스키마 개발을 위한 대화형 모드를 선택합니다. 이를 통해 스키마 생성을 위한 마법사가 사용자를 안내합니다. 고급 기술을 갖춘 사용자는 마법사를 사용하지 않고도 자신만의 스키마 JSON을 만들 수 있습니다. 자세한 내용은 (선택 사항) 스키마 생성(고급 사용자) 단원을 참조하십시오.
   > - --output 플래그는 출력 이름을 설정합니다. --output 플래그를 포함하지 않으면 C3R 암호화 클라이언트는 기본 출력 이름(예: <input>.out.csv 또는 스키마의 경우, <input>.json)을 선택하려고 합니다.

3. Number of target columns from source column 'username'?의 경우 **1**을 입력한 다음 Enter 키를 누릅니다.

4. `Target column type: [c]leartext, [f]ingerprint, or [s]ealed?`의 경우 **f**를 입력한 다음 Enter 키를 누릅니다.
5. `Target column headername <default 'username'>`의 경우 Enter 키를 누릅니다.

   기본 사용자 이름은 '`username`'입니다.
6. `Add suffix '_fingerprint' to header to indicate how it was encrypted, [y]es or [n]o <default 'yes'>`의 경우 **y**를 입력한 다음 Enter 키를 누릅니다.

   > **Note**
   >
   > 대화형 모드에서는 암호화된 열 헤더(fingerprint 열의 경우 `_fingerprint`, sealed 열의 경우 `_sealed`)에 추가할 접미사를 제안합니다. 접미사는에 데이터를 업로드 AWS 서비스 하거나 AWS Clean Rooms 공동 작업을 생성하는 등의 작업을 수행할 때 유용할 수 있습니다. 이러한 접미사는 각 열의 암호화된 데이터로 수행할 수 있는 작업을 나타내는 데 도움이 됩니다. 예를 들어 sealed 열을 열(`_sealed`)로 암호화한 후 그 위에 JOIN을 시도하거나 그 반대의 경우를 시도하면 작동하지 않습니다.
7. `Number of target columns from source column 'ad_variant'?`의 경우 **1**을 입력한 다음 Enter 키를 누릅니다.
8. `Target column type: [c]leartext, [f]ingerprint, or [s]ealed?`의 경우 **c**를 입력한 다음 Enter 키를 누릅니다.
9. `Target column headername <default 'username'>`의 경우 Enter 키를 누릅니다.

   기본 사용자 이름은 '`ad_variant`'입니다.

   `ads.json`이라는 새 파일에 스키마가 기록됩니다.

   > **Note**
   >
   > 스키마는 텍스트 편집기에서 열면 볼 수 있습니다(예: Windows에서 Notepad 또는 macOS에서 TextEdit).
10. 이제 데이터를 암호화할 준비가 되었습니다.

예:sealed, fingerprint, cleartext 열을 사용하여 암호화 스키마 생성

이 예제에서는 `sales.csv`의 경우, `username` , `purchased`, 및 `product` 세 개의 열이 있습니다. 이 열에는 다음이 필요합니다.

- `product` 열을 `sealed` 열로 만들려면
- `username` 열을 `fingerprint` 열로 암호화하려면
- `purchased` 열을 `cleartext` 열로 만들려면

sealed, fingerprint, cleartext 열이 있는 암호화 스키마를 생성하려는 경우

1. (선택 사항) 암호화할 c3r-cli.jar 파일 및 파일이 있는지 확인하려는 경우:

   a. 원하는 디렉토리로 이동하여 `ls`(Mac 또는 Unix/Linux를 사용하는 경우) 또는 `dir`(Windows를 사용하는 경우)을 실행합니다.

   b. 표 형식 데이터 파일(.csv) 목록을 보고 암호화할 파일을 선택합니다.

   이 예에서는 `sales.csv`가 암호화하려는 파일입니다.

2. CLI에서 다음 명령을 실행하여 대화식으로 스키마를 생성합니다.

```
java -jar c3r-cli.jar schema sales.csv --interactive --
output=sales.json
```

> **Note**
>
> - `--interactive` 플래그는 스키마 개발을 위한 대화형 모드를 선택합니다. 이를 통해 사용자는 스키마 생성을 위한 안내가 있는 워크플로를 안내합니다.
> - 고급 사용자인 경우 안내식 워크플로를 사용하지 않고도 자체 스키마 JSON을 만들 수 있습니다. 자세한 내용은 (선택 사항) 스키마 생성(고급 사용자) 단원을 참조하십시오.
> - 열 헤더가 없는.csv 파일의 경우 CLI에서 사용할 수 있는 스키마 명령의 `--noHeaders` 플래그를 참조하세요.
> - `--output` 플래그는 출력 이름을 설정합니다. `--output` 플래그를 포함하지 않으면 C3R 암호화 클라이언트는 기본 출력 이름(예: `<input>.out` 또는 스키마의 경우, `<input>.json`)을 선택하려고 합니다.

3. `Number of target columns from source column 'username'?`의 경우 **1**을 입력한 다음 Enter 키를 누릅니다.

4. `Target column type: [c]leartext, [f]ingerprint, or [s]ealed?`의 경우 **f**를 입력한 다음 Enter 키를 누릅니다.
5. `Target column headername <default 'username'>`의 경우 Enter 키를 누릅니다.

   기본 사용자 이름은 '`username`'입니다.
6. `Add suffix '_fingerprint' to header to indicate how it was encrypted, [y]es or [n]o <default 'yes'>`의 경우 **y**을 입력한 다음 Enter 키를 누릅니다.
7. `Number of target columns from source column 'purchased'?`의 경우 **1**을 입력한 다음 Enter 키를 누릅니다.
8. `Target column type: [c]leartext, [f]ingerprint, or [s]ealed?`의 경우 **c**를 입력한 다음 Enter 키를 누릅니다.
9. `Target column headername <default 'purchased'>`의 경우 Enter 키를 누릅니다.

   기본 사용자 이름은 '`purchased`'입니다.
10. `Number of target columns from source column 'product'?`의 경우 **1**을 입력한 다음 Enter 키를 누릅니다.
11. `Target column type: [c]leartext, [f]ingerprint, or [s]ealed?`의 경우 **s**를 입력한 다음 Enter 키를 누릅니다.
12. `Target column headername <default 'product'>`의 경우 Enter 키를 누릅니다.

    기본 사용자 이름은 '`product`'입니다.
13. `'product_sealed' padding type: [n]one, [f]ixed, or [m]ax <default 'max'?>`의 경우 Enter 키를 눌러 기본값을 선택합니다.
14. `Byte-length beyond max length to pad cleartext to in 'product_sealed' <default '0'>?`의 경우 Enter 키를 눌러 기본값을 선택합니다.

    `sales.json`이라는 새 파일에 스키마가 기록됩니다.
15. 이제 데이터를 암호화할 준비가 되었습니다.

5단계: 공유 암호 키 생성

데이터 테이블을 암호화하려면 공동 작업 참여자가 공유 비밀 키에 동의하고 안전하게 공유해야 합니다.

공유 암호 키는 256비트(32바이트) 이상이어야 합니다. 더 큰 키를 지정할 수 있지만 추가 보안을 제공하지는 않습니다.

> **Important**
>
> 암호화와 암호 해독에 사용되는 키와 공동 작업 ID는 모든 공동 작업 참여자에 대해 동일해야 한다는 점을 기억하세요.

다음 섹션에서는 각 터미널의 현재 작업 디렉터리에 `secret.key`로 저장된 공유 비밀 키를 생성하는 콘솔 명령의 예를 제공합니다.

주제

- 예: OpenSSL을 사용한 키 생성
- 예: PowerShell을 사용하여 Windows에서 키 생성

예: OpenSSL을 사용한 키 생성

일반적인 범용 암호화 라이브러리의 경우 다음 명령을 실행하여 공유 비밀 키를 생성합니다.

```
openssl rand 32 > secret.key
```

Windows를 사용 중이고 아직 OpenSSL을 설치하지 않은 경우 PowerShell을 사용하여 Windows에서 키 생성하기에 설명된 예시를 사용하여 키를 생성할 수 있습니다.

예: PowerShell을 사용하여 Windows에서 키 생성

Windows에서 사용할 수 있는 터미널 애플리케이션인 PowerShell의 경우 다음 명령을 실행하여 공유 비밀 키를 생성합니다.

```
$bs = New-Object Byte[](32);
[Security.Cryptography.RandomNumberGenerator]::Create().GetBytes($bs); Set-
Content 'secret.key' -Encoding Byte -Value $bs
```

6단계: 환경 변수에 공유 암호 키 저장

환경 변수는 사용자가와 같은 다양한 키 스토어의 보안 키를 제공하고 C3R 암호화 클라이언트에 전달할 수 있는 편리 AWS Secrets Manager 하고 확장 가능한 방법입니다.

를 사용하여 관련 환경 변수에 키를 AWS CLI 저장하는 AWS 서비스 경우 C3R 암호화 클라이언트는에 저장된 키를 사용할 수 있습니다. 예를 들어 C3R 암호화 클라이언트는의 키를 사용할 수 있습니다 AWS Secrets Manager. 자세한 내용은 AWS Secrets Manager 사용 설명서에서 AWS Secrets Manager으로 기본 보안 암호 생성을 참조하세요.

> **Note**
>
> 그러나와 AWS 서비스 같은를 사용하여 C3R 키를 보관 AWS Secrets Manager 하기 전에 사용 사례에서 허용하는지 확인합니다. 특정 사용 사례에서는 키를 보류해야 할 수 있습니다 AWS. 이는 암호화된 데이터와 키를 동일한 제3자가 보유하지 않도록 하기 위한 것입니다.

공유 비밀 키의 유일한 요구 사항은 공유 비밀 키가 base64-인코딩되어 환경 변수 `C3R_SHARED_SECRET`에 저장되어 있어야 한다는 것입니다.

다음 섹션에서는 `secret.key` 파일을 base64로 변환하고 환경 변수로 저장하는 콘솔 명령에 대해 설명합니다. `secret.key` 파일은 5단계: 공유 암호 키 생성에 나열된 명령 중 하나에서 생성될 수 있으며 예제 소스일 뿐입니다.

PowerShell을 사용하여 Windows의 환경 변수에 키를 저장합니다

base64로 변환하고 PowerShell을 사용하여 Windows에서 환경 변수를 설정하려면 다음 명령을 실행합니다.

```
$Bytes=[IO.File]::ReadAllBytes((Get-Location).ToString()+'\secret.key');
 $env:C3R_SHARED_SECRET=[Convert]::ToBase64String($Bytes)
```

Linux 또는 macOS의 환경 변수에 키를 저장합니다

base64로 변환하고 Linux 또는 macOS에서 환경 변수를 설정하려면 다음 명령을 실행합니다.

```
export C3R_SHARED_SECRET="$(cat secret.key | base64)"
```

7단계: 데이터 암호화

이 단계를 수행하려면 AWS Clean Rooms 공동 작업 ID와 공유 보안 키를 획득해야 합니다. 자세한 내용은 사전 조건을 참조하세요.

다음 예제에서는 `ads.json`이라고 생성한 스키마를 사용하여 `ads.csv`에서 암호화를 실행합니다.

데이터를 암호화하려는 경우

1. 6단계: 환경 변수에 공유 암호 키 저장에서 공동 작업을 위한 공유 비밀 키를 저장합니다.

2. 명령줄에서 다음 명령을 입력합니다.

   ```
   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. *<name of input .csv file>*의 경우 입력.csv 파일의 이름을 입력합니다.
4. schema=의 경우 .json 암호화 스키마 파일의 이름을 입력합니다.
5. id=의 경우 공동 작업 ID를 입력합니다.
6. output=의 경우 출력 파일의 이름(예: ads-output.csv)을 입력합니다.
7. 암호화 컴퓨팅 파라미터 및 Clean Rooms용 암호화 컴퓨팅의 선택적 플래그 에 설명된 명령줄 플래그를 모두 포함합니다.
8. 명령을 실행합니다.

ads.csv의 예제에서는 다음 명령을 실행합니다.

java -jar c3r-cli.jar encrypt ***ads.csv*** --schema=***ads.json*** --id=***123e4567-e89b-42d3-a456-556642440000*** --output=***ads-output.csv***

sales.csv의 예제에서는 다음 명령을 실행합니다.

java -jar c3r-cli.jar encrypt ***sales.csv*** --schema=***sales.json*** --id=***123e4567-e89b-42d3-a456-556642440000***

> **Note**
>
> 이 예제에서는 출력 파일 이름(--output=*sales-output.csv*)을 지정하지 않습니다. 그 결과 기본 출력 파일 이름 name-of-file.out.csv가 생성되었습니다.

이제 암호화된 데이터를 확인할 준비가 되었습니다.

8단계: 데이터 암호화 확인

데이터가 암호화되었는지 확인하려면

1. 암호화된 데이터 파일(예:sales-output.csv)을 확인합니다.
2. 다음 열을 확인합니다.

a. 열 1 — 암호화됨(예:`username_fingerprint`).

   fingerprint 열(HMAC)의 경우 버전 및 유형 접두사(예:`01:hmac:`) 뒤에 44자의 base64로 인코딩된 데이터가 있습니다.

b. 열 2 — 암호화되지 않음(예: `purchased`).

c. 열 3 — 암호화됨(예: `product_sealed`).

   암호화된(SELECT) 열의 경우 cleartext의 길이와 버전 및 유형 접두사 뒤에 오는 패딩(예:`01:enc:`)의 길이는 암호화된 cleartext의 길이에 정비례합니다. 즉, 길이는 입력 크기에 인코딩으로 인한 약 33% 의 오버헤드를 더한 값입니다.

이제 다음에 대한 준비가 되었습니다.

1. 암호화된 데이터를 S3에 업로드합니다.
2. AWS Glue 테이블을 생성합니다.
3. AWS Clean Rooms에서 구성된 테이블을 생성합니다.

C3R 암호화 클라이언트는 암호화되지 않은 데이터를 포함하지 않는 임시 파일을 생성합니다(최종 출력에서 해당 데이터도 암호화되지 않는 경우 제외). 하지만 일부 암호화된 값은 제대로 채워지지 않을 수 있습니다. 공동 작업 설정 `allowRepeatedFingerprintValue`가 `false`인 경우에도 지문 열에 중복된 값이 포함될 수 있습니다. 이 문제는 적절한 패딩 길이와 중복 제거 속성을 확인하기 전에 임시 파일이 작성되기 때문에 발생합니다.

C3R 암호화 클라이언트가 실패하거나 암호화 중에 중단되는 경우 임시 파일을 작성한 후 이러한 속성을 확인하고 임시 파일을 삭제하기 전에 중단될 수 있습니다. 따라서 이러한 임시 파일은 여전히 디스크에 있을 수 있습니다. 이 경우 이러한 파일의 내용은 출력과 동일한 수준으로 일반 텍스트 데이터를 보호하지 못합니다. 특히 이러한 임시 파일은 통계 분석에 일반 텍스트 데이터를 공개할 수 있지만 최종 출력에는 적합하지 않을 수 있습니다. 사용자는 이러한 파일(특히 SQLite 데이터베이스)을 삭제하여 파일이 무단으로 유출되지 않도록 해야 합니다.

(선택 사항) 스키마 생성(고급 사용자)

스키마 생성 기능은 고급 사용자를 위한 것입니다.

다음은 열 헤더가 있거나 없는 입력 파일의 JSON 스키마 파일 형식에 대한 설명입니다. 고급 사용자는 원하는 경우 스키마를 직접 작성하거나 수정할 수 있습니다.

> **Note**
>
> C3R 암호화 클라이언트는 예:sealed, fingerprint, cleartext 열을 사용하여 암호화 스키마 생성에서 설명하는 대화형 프로세스 또는 스텁 템플릿 생성을 통해 스키마를 만드는 데 도움을 줄 수 있습니다.

매핑된 테이블 스키마와 위치 테이블 스키마

다음 섹션에서는 두 종류의 테이블 스키마에 대해 설명합니다.

- 매핑된 테이블 스키마 - 이 스키마는 헤더 행과 Apache Parquet 파일이 있는.csv 파일을 암호화하는 데 사용됩니다.
- 위치 테이블 스키마 - 이 스키마는 헤더 행이 없는.csv 파일을 암호화하는 데 사용됩니다.

C3R 암호화 클라이언트는 공동 작업을 위해 표 형식 파일을 암호화할 수 있습니다. 이렇게 하려면 입력에서 암호화된 출력을 도출하는 방법을 지정하는 해당 스키마 파일이 있어야 합니다.

C3R 암호화 클라이언트는 명령줄에서 C3R 암호화 클라이언트 스키마 명령을 실행하여 `INPUT` 파일에 대한 스키마를 생성하는 데 도움을 줄 수 있습니다. 명령의 예는 `java -jar c3r-cli.jar schema --interactive INPUT`입니다.

이 스키마는 다음 정보를 지정합니다:

1. 헤더 이름(매핑된 스키마) 또는 위치(위치 스키마)를 통해 출력 파일의 변환된 열에 매핑되는 원본 열
2. 어떤 대상 열을 cleartext로 유지해야 할까요
3. SELECT 쿼리를 위해 어떤 대상 열을 암호화해야 할까요
4. JOIN 쿼리를 위해 어떤 대상 열을 암호화해야 할까요

이 정보는 테이블별 JSON 스키마 파일에 인코딩되며, 이 파일은 `headerRow` 필드가 부울 값인 단일 객체로 구성됩니다. 헤더 행이 있는 Parquet 파일 및 .csv 파일의 경우 값은 `true`여야 하며, 그렇지 않은 경우 `false`여야 합니다.

매핑된 테이블 스키마

매핑된 스키마의 모양은 다음과 같습니다.

```
{
  "headerRow": true,
  "columns": [
    {
      "sourceHeader": STRING,
      "targetHeader": STRING,
      "type": TYPE,
      "pad": PAD
    },
    ...
  ]
}
```

headerRow가 true인 경우, 객체의 다음 필드는 columns이고, 이 필드는 소스 헤더를 대상 헤더에 매핑하는 열 스키마 배열(즉, 출력 열에 포함되어야 하는 내용을 설명하는 JSON 객체)을 포함합니다.

- sourceHeader— 데이터가 파생된 소스 열의 STRING 헤더 이름.

> **Note**
>
> 동일한 소스 열을 여러 대상 열에 사용할 수 있습니다.
> 스키마 어디에도 sourceHeader로 나열되지 않은 입력 파일의 열은 출력 파일에 나타나지 않습니다.

- targetHeader – 출력 파일에 있는 해당 열의 STRING 헤더 이름.

> **Note**
>
> 이 필드는 매핑된 스키마의 경우 선택 사항입니다. 이 필드를 생략하면 출력의 헤더 이름으로 sourceHeader가 다시 사용됩니다. 출력 열이 fingerprint 열 또는 sealed 열인 경우 _fingerprint 또는 _sealed가 각각 추가됩니다.

- type – 출력 파일에 있는 대상 열의 TYPE. 즉, 공동 작업에서 열이 사용되는 방식에 따라 cleartext, sealed 또는 fingerprint 중 하나를 선택합니다.
- pad— TYPE이 sealed인 경우에만 존재하는 열 스키마 개체의 필드입니다.. PAD의 해당 값은 데이터를 암호화하기 전에 데이터를 어떻게 패딩해야 하는지를 설명하는 객체입니다.

```
{
  "type": PAD_TYPE,
```

```
  "length": INT
}
```

사전 암호화 패딩을 지정하기 위해 type 및 length가 사용되며 다음과 같이 사용됩니다.

- PAD_TYPE as none — 열의 데이터에 패딩이 적용되지 않으며 length 필드를 적용할 수 없습니다(즉, 생략).
- PAD_TYPE as fixed — 열의 데이터가 지정된 length 바이트만큼 채워집니다.
- PAD_TYPE as max — 가장 긴 값의 바이트 길이에 추가 length 바이트를 더한 크기로 열의 데이터가 채워집니다.

다음은 각 유형의 열이 있는 매핑된 스키마의 예시입니다.

```
{
  "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
```

```
      }
    }
  ]
}
```

좀 더 복잡한 예로, 다음은 헤더가 있는.csv 파일 예제입니다.

```
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
```

다음 매핑된 스키마 예제에서 열 `FirstName` 및 `LastName`은 열은 `cleartext` 열입니다. `State` 열은 `fingerprint` 열로 암호화되고 none의 패딩이 있는 `sealed` 열로 암호화됩니다. 나머지 열은 생략됩니다.

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

다음은 매핑된 스키마의 결과인.csv 파일입니다.

```
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=
```

위치 테이블 스키마

위치 스키마의 모양은 다음과 같습니다.

```
{
  "headerRow": false,
  "columns": [
    [
      {
        "targetHeader": STRING,
        "type": TYPE,
        "pad": PAD
      },
      {
        "targetHeader": STRING,
        "type": TYPE,
        "pad": PAD
```

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

headerRow가 false인 경우, 개체의 다음 필드는 columns이고 이는 항목 배열을 포함하는 필드입니다. 각 항목 자체는 0개 이상의 위치 열 스키마(sourceHeader 필드 없음)로 구성된 배열입니다. 이 스키마는 출력에 포함되어야 하는 내용을 설명하는 JSON 객체입니다.

- sourceHeader— 데이터가 파생되는 소스 열의 STRING 헤더 이름.

> **Note**
>
> 위치 스키마에서는 이 필드를 생략해야 합니다. 위치 스키마에서 소스 열은 스키마 파일에 있는 열의 해당 인덱스에 의해 유추됩니다.

- targetHeader— 출력 파일에 있는 해당 열의 STRING 헤더 이름.

> **Note**
>
> 이 필드는 위치 스키마에 필요합니다.

- type— 출력 파일에 있는 대상 열의 TYPE. 즉, 공동 작업에서 열이 사용되는 방식에 따라 cleartext, sealed 또는 fingerprint 중 하나를 선택합니다.
- pad— TYPE이 sealed인 경우에만 존재하는 열 스키마 개체의 필드입니다.. PAD의 해당 값은 데이터를 암호화하기 전에 데이터를 어떻게 패딩해야 하는지를 설명하는 객체입니다.

```
{
  "type": PAD_TYPE,
  "length": INT
}
```

사전 암호화 패딩을 지정하기 위해 type 및 length가 사용되며 다음과 같이 사용됩니다.

- PAD_TYPE as none — 열의 데이터에 패딩이 적용되지 않으며 length 필드를 적용할 수 없습니다(즉, 생략).
- PAD_TYPE as fixed — 열의 데이터가 지정된 length 바이트만큼 채워집니다.

- PAD_TYPE as max — 가장 긴 값의 바이트 길이에 추가 length 바이트를 더한 크기로 열의 데이터가 채워집니다.

> **Note**
>
> fixed 열 데이터의 바이트 크기 상한을 미리 알고 있는 경우에 유용합니다. 해당 열의 데이터가 지정된 length 값보다 길면 오류가 발생합니다.
> max는 입력 데이터의 정확한 크기를 알 수 없는 경우 데이터 크기와 상관없이 작동하므로 편리합니다. 하지만 max는 데이터를 두 번 암호화하기 때문에 추가 처리 시간이 필요합니다. max는 임시 파일에 읽어올 때 데이터를 한 번 암호화하고 열에 가장 긴 데이터 입력을 알고 나면 한 번 암호화합니다.
> 또한 가장 긴 값의 길이는 클라이언트 호출 사이에 저장되지 않습니다. 데이터를 일괄적으로 암호화하거나 새 데이터를 주기적으로 암호화하려는 경우 결과 사이퍼텍스트 -길이는 배치마다 다를 수 있다는 점에 유의하세요.

다음은 위치 스키마의 예입니다.

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

복잡한 예로, 다음은.csv 파일의 첫 번째 행에 헤더가 없는 경우의 예시.csv 파일입니다.

```
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
```

위치 스키마의 형식은 다음과 같습니다.

```
{
  "headerRow": false,
  "columns": [
    [
      {
        "targetHeader": "GivenName",
        "type": "cleartext"
      }
    ],
    [
      {
        "targetHeader": "Surname",
        "type": "cleartext"
      }
    ],
    [],
```

```
    [],
    [
      {
        "targetHeader": "State_Join",
        "type": "fingerprint"
      },
      {
        "targetHeader": "State",
        "type": "sealed",
        "pad": {
          "type": "none"
        }
      }
    ],
    [],
    [],
    [],
    []
  ]
}
```

위의 스키마는 지정된 대상 헤더를 포함하는 헤더 행과 함께 다음과 같은 출력 파일을 생성합니다.

```
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=
```

C3R 암호화 클라이언트를 사용한 데이터 테이블 복호화

Clean Rooms에 대한 암호화 컴퓨팅 및 C3R 암호화 클라이언트를 사용하여 데이터 테이블을 암호화하는 공동 작업의 경우 이 절차를 따릅니다. 공동 작업에서 데이터를 쿼리한 후에 이 절차를 사용합니다.

이 절차를 수행하려면 공유 비밀 키와 공동 작업 ID가 필요합니다.

결과를 받을 수 있는 구성원은 공동 작업 데이터를 암호화하는 데 사용한 것과 동일한 공유 비밀 키와 공동 작업 ID를 사용하여 데이터를 해독합니다.

> **Note**
>
> AWS Clean Rooms 공동 작업은 쿼리 결과를 수행하고 볼 수 있는 사용자를 이미 제한하고 있습니다. 암호 해독을 수행하려면 이러한 결과에 액세스할 수 있는 사람은 데이터를 암호화하는 데 사용된 것과 동일한 공유 비밀 키와 공동 작업 ID가 필요합니다.

암호화된 데이터 테이블을 해독하려는 경우

1. (선택 사항) C3R 암호화 클라이언트에서 사용 가능한 명령을 확인하세요.
2. (선택 사항) 원하는 디렉토리로 이동하여 `ls`(macOS) 또는 `dir`(Windows)를 실행합니다.
   - c3r-cli.jar 파일 및 암호화된 쿼리 결과 데이터 파일이 원하는 디렉터리에 있는지 확인합니다.

     > **Note**
     >
     > AWS Clean Rooms 콘솔 인터페이스에서 쿼리 결과를 다운로드하는 경우 사용자 계정의 다운로드 폴더에 있을 수 있습니다. (예를 들어, 사용자 디렉토리의 다운로드 폴더는 Windows 및 macOS에 있습니다.) 쿼리 결과 파일을 c3r-cli.jar와 동일한 폴더로 옮기는 것이 좋습니다.
3. 공유 비밀 키를 `C3R_SHARED_SECRET` 환경 변수에 저장합니다. 자세한 내용은 6단계: 환경 변수에 공유 암호 키 저장 단원을 참조하십시오.
4. AWS Command Line Interface (AWS CLI)에서 다음 명령을 실행합니다.

   java -jar c3r-cli.jar decrypt ***<name of input .csv file>*** --id=***<collaboration id>*** --output=***<output file name>***
5. *`user input placeholder`*를 사용자의 정보로 바꿉니다.

a. `id=`의 경우 공동 작업 ID를 입력합니다.

b. `output=`의 경우 출력 파일의 이름(예:`results-decrypted.csv`)을 입력합니다.

   출력 이름을 지정하지 않으면 터미널에 기본 이름이 표시됩니다.

c. 선호하는 CSV 또는 Parquet 보기 응용 프로그램(예: Microsoft Excel, 텍스트 편집기 또는 기타 응용 프로그램)을 사용하여 지정된 출력 파일의 해독된 데이터를 볼 수 있습니다.

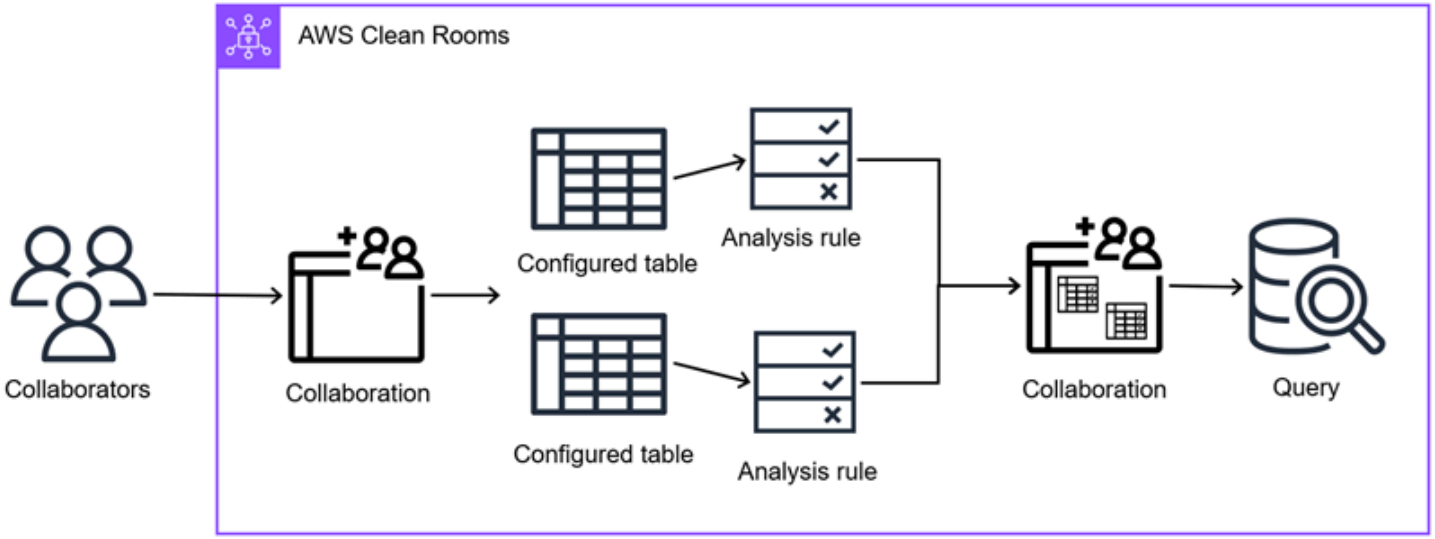
에서 구성된 테이블 AWS Clean Rooms

구성된 테이블은 데이터 소스의 기존 테이블에 대한 참조입니다. 여기에는 AWS Clean Rooms에서 데이터를 쿼리할 수 있는 방법을 결정하는 분석 규칙이 포함되어 있습니다. 구성된 테이블을 하나 이상의 공동 작업에 연결할 수 있습니다.

를 사용하면 구매 수와 구매 수의 비교와 같은 이벤트 데이터에 대한 집계 분석을 수행할 AWS Clean Rooms수 있습니다. 세그먼트 데이터에서 CRM 데이터로 중복되는 고객 데이터를 보강하는 등 이벤트 데이터에 대한 목록 분석을 수행할 수도 있습니다. 사용자 지정 쿼리를 수행하고 시청률 데이터 및 세그먼트 속성과 같은 이벤트 데이터에 대한 차등 프라이버시를 설정할 수도 있습니다.

먼저에서 공동 작업을 생성하고 초대 AWS 계정 할를 AWS Clean Rooms 추가하거나 멤버십을 생성하여 초대할 공동 작업에 참여합니다. 그런 다음 공동 작업의 다른 구성원과 함께 구성된 테이블을 생성합니다. 두 사람 모두 구성된 테이블에 분석 규칙(집계, 목록 또는 사용자 지정)을 추가하고 구성된 테이블을 공동 작업에 연결합니다. 마지막으로 쿼리 권한이 있는 구성원이 두 데이터 테이블에 대한 쿼리를 실행합니다.

다음 다이어그램은에서 이벤트 데이터를 사용하는 방법을 요약합니다 AWS Clean Rooms.



주제

- AWS Clean Rooms에서 구성된 테이블 생성하기
- 구성된 테이블에 분석 규칙 추가
- 구성된 테이블을 공동 작업에 연결
- 구성된 테이블에 공동 작업 분석 규칙 추가
- 차등 프라이버시 정책 구성(선택 사항)

- 테이블 및 분석 규칙 보기
- 구성된 테이블 세부 정보 편집
- 구성된 테이블 태그 편집
- 구성된 테이블 분석 규칙 편집
- 구성된 테이블 분석 규칙 삭제
- 구성된 테이블의 허용되지 않는 열
- 구성된 테이블 연결 편집
- 구성된 테이블 분리

AWS Clean Rooms에서 구성된 테이블 생성하기

구성된 테이블은 데이터 소스의 기존 테이블에 대한 참조입니다. 여기에는 AWS Clean Rooms에서 데이터를 쿼리할 수 있는 방법을 결정하는 분석 규칙이 포함되어 있습니다. 구성된 테이블을 하나 이상의 공동 작업에 연결할 수 있습니다.

AWS SDKs. AWS Clean Rooms

주제

- 구성된 테이블 생성 - Amazon S3 데이터 소스
- 구성된 테이블 생성 - Amazon Athena 데이터 소스
- 구성된 테이블 생성 - Snowflake 데이터 소스

구성된 테이블 생성 - Amazon S3 데이터 소스

이 절차에서 구성원은 다음 작업을 수행합니다.

- 에서 사용할 기존 AWS Glue 테이블을 구성합니다 AWS Clean Rooms. (이 단계는 Clean Rooms에 대한 암호화 컴퓨팅을 사용하지 않는 한 공동 작업에 참여하기 전이나 후에 수행할 수 있습니다.)

> **Note**
>
> AWS Clean Rooms 는 AWS Glue 테이블을 지원합니다. 데이터 가져오기에 대한 자세한 내용은 섹션을 AWS Glue참조하세요3단계: 데이터 테이블을 Amazon S3에 업로드.

- 구성된 테이블의 이름을 지정하고 공동 작업에 사용할 열을 선택합니다.

다음 절차에서는 다음과 같이 가정합니다.

- 공동 작업 구성원이 이미 Amazon S3에 데이터 테이블을 업로드하고 AWS Glue 테이블을 생성했습니다.

> **Note**
>
> Spark 분석 엔진을 사용하는 경우 Amazon S3의 결과 대상은 데이터 소스와 동일한 S3 버킷 내에 있을 수 없습니다.

- (선택 사항) 암호화된 데이터 테이블의 경우에만 공동 작업 구성원은 이미 C3R 암호화 클라이언트를 사용하여 암호화된 데이터 테이블을 준비했습니다.

에서 제공하는 통계 생성을 사용하여 AWS Glue Data Catalog 테이블에 대한 열 수준 통계를 계산 AWS Glue 할 수 있습니다. 가 데이터 카탈로그의 테이블에 대한 통계를 AWS Glue 생성한 후 Amazon Redshift Spectrum은 해당 통계를 자동으로 사용하여 쿼리 계획을 최적화합니다. 를 사용하여 열 수준 통계를 계산하는 방법에 대한 자세한 내용은 AWS Glue 사용 설명서의 열 통계를 사용하여 쿼리 성능 최적화를 AWS Glue참조하세요. 에 대한 자세한 내용은 AWS Glue 개발자 안내서를 AWS Glue참조하십시오.

구성된 테이블을 생성하려면 - Amazon S3 데이터 소스

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/cleanrooms AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 테이블을 선택합니다.
3. 오른쪽 상단 모서리에서 새 테이블 구성을 선택합니다.
4. 데이터 소스의 AWS 데이터 소스에서 Amazon S3를 선택합니다.
5. Amazon S3 테이블에서:
   a. 드롭다운 목록에서 데이터베이스를 선택합니다.
   b. 드롭다운 목록에서 구성하고 싶은 테이블을 선택합니다.

> **Note**
>
> 테이블이 올바른지 확인하려면 다음 중 하나를 수행합니다.
>
> - 에서 보기를 AWS Glue 선택합니다.

- 에서 스키마 보기를 AWS Glue 켜면 스키마를 볼 수 있습니다.

> **Important**
>
> 데이터가 CSV 형식인 AWS Glue 테이블의 경우 Glue 스키마의 열 이름과 순서가 CSV 데이터와 정확히 일치해야 합니다. 정렬되지 않으면 구성된 테이블에 허용되는 열 목록이 제대로 적용되지 않을 수 있습니다.

6. 공동 작업에 허용되는 열 및 분석 방법의 경우

   a. 공동 작업에서 허용할 열은 무엇입니까?

      - 공동 작업에서 모든 열을 쿼리하도록 허용하려면 모든 열을 선택합니다.
      - 사용자 지정 목록을 선택하여 허용된 열 지정 드롭다운 목록에서 하나 이상의 열을 공동 작업에서 쿼리할 수 있도록 허용합니다.

   b. 허용된 분석 방법의 경우

      i. 이 테이블에서 SQL 쿼리를 직접 실행하려면 직접 쿼리를 선택합니다.

      ii. 이 테이블에서 PySpark 작업을 직접 실행하려면 직접 작업을 선택합니다.

   Example 예제

   예를 들어 공동 작업 구성원이 모든 열에서 직접 SQL 쿼리와 PySpark 작업을 모두 실행하도록 허용하려면 모든 열, 직접 쿼리 및 직접 작업을 선택합니다.

7. 구성된 테이블 세부 정보의 경우,

   a. 구성된 테이블의 이름을 입력합니다.

      기본 이름을 사용하거나 이 테이블의 이름을 바꿀 수 있습니다.

   b. 테이블에 대한 설명을 입력합니다.

      설명은 비슷한 이름을 가진 다른 구성된 테이블을 구분하는 데 도움이 됩니다.

8. 구성된 테이블 리소스에 대해 태그를 활성화하려면 새 태그 추가를 선택한 다음 키와 값 쌍을 입력합니다.

9. 새 클라이언트 구성를 선택합니다.

이제 구성된 테이블을 만들었으므로 다음을 수행할 준비가 되었습니다.

- 구성된 테이블에 분석 규칙 구성
- 구성된 테이블을 공동 작업에 연결합니다

구성된 테이블 생성 - Amazon Athena 데이터 소스

Amazon Athena 데이터 소스 옵션을 사용하면 데이터 AWS Glue 카탈로그에 카탈로그화된 Amazon S3에 저장된 데이터를 쿼리하고를 통해 액세스를 제어할 수 있습니다 AWS Lake Formation. 테이블과 AWS Glue Data Catalog 뷰가 모두 지원됩니다. Lake Formation 리소스 링크를 사용하여 AWS Clean Rooms 공동 작업에 조인하는 AWS 계정 AWS 리전 AWS Clean Rooms 멤버 계정 간에 테이블과 뷰를 공유할 수 있습니다.

> **Note**
>
> Athena 데이터 소스 통합을 통해 Amazon S3 기반 데이터 세트만 쿼리할 수 있습니다.

이 절차에서 구성원은 다음 작업을 수행합니다.

- 를 사용하기 AWS Glue Data Catalog 위해에서 기존 테이블 또는 뷰를 구성합니다. AWS Clean Rooms
- 구성된 테이블의 이름을 지정하고 공동 작업에 사용할 열을 선택합니다.

다음 절차에서는 다음과 같이 가정합니다.

- 공동 작업 구성원이 AWS Glue Data Catalog 데이터베이스 및 테이블 또는GDC 뷰를 이미 생성했습니다.

구성된 테이블을 생성하려면 - Athena 데이터 소스

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/cleanrooms AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 테이블을 선택합니다.
3. 오른쪽 상단 모서리에서 새 테이블 구성을 선택합니다.
4. 데이터 소스의 AWS 데이터 소스에서 Amazon Athena를 선택합니다.

5. Amazon Athena 테이블에서:

   a. 드롭다운 목록에서 데이터베이스를 선택합니다.

   b. 드롭다운 목록에서 구성하고 싶은 테이블을 선택합니다.

> **Note**
>
> 테이블이 올바른지 확인하려면 다음 중 하나를 수행합니다.
>
> - 에서 보기를 AWS Glue 선택합니다.
> - 에서 스키마 보기를 AWS Glue 켜면 스키마를 볼 수 있습니다.

6. Amazon Athena 구성의 경우

   a. 드롭다운 목록에서 작업 그룹을 선택합니다.

   b. S3 출력 위치의 경우 다음 시나리오 중 하나를 기반으로 권장 작업을 선택합니다.

시나리오	권장 조치
작업 그룹에 기본 출력 위치가 없습니다.	S3 출력 위치를 입력하거나 S3 찾아보기를 선택합니다.
작업 그룹은 기본 출력 위치를 적용합니다.	S3 출력 위치가 자동으로 선택되며 변경할 수 없습니다.
Workgroup은 기본 출력 위치를 적용하지 않습니다.	S3 출력 위치를 입력하거나 S3 찾아보기를 선택합니다.

7. 공동 작업에 허용된 열에서 목표에 따라 옵션을 선택합니다.

목표	권장 옵션
에서 모든 열 사용 허용 AWS Clean Rooms (분석 규칙에 따라 다름)	모든 컬럼
허용된 열 지정 드롭다운 목록에서 하나 이상의 열 허용	사용자 지정 목록

8. 구성된 테이블 세부 정보의 경우,

   a. 구성된 테이블의 이름을 입력합니다.

      기본 이름을 사용하거나 이 테이블의 이름을 바꿀 수 있습니다.

   b. 테이블에 대한 설명을 입력합니다.

      설명은 비슷한 이름을 가진 다른 구성된 테이블을 구분하는 데 도움이 됩니다.

   c. 구성된 테이블 리소스에 대해 태그를 활성화하려면 새 태그 추가를 선택한 다음 키와 값 쌍을 입력합니다.

9. 새 클라이언트 구성를 선택합니다.

이제 구성된 테이블을 만들었으므로 다음을 수행할 준비가 되었습니다.

- 구성된 테이블에 분석 규칙 구성
- 구성된 테이블을 공동 작업에 연결합니다

구성된 테이블 생성 - Snowflake 데이터 소스

이 절차에서 구성원은 다음 작업을 수행합니다.

- 에서 사용할 기존 Snowflake 테이블을 구성합니다 AWS Clean Rooms. (이 단계는 Clean Rooms에 대한 암호화 컴퓨팅을 사용하지 않는 한 공동 작업에 참여하기 전이나 후에 수행할 수 있습니다.)
- 구성된 테이블의 이름을 지정하고 공동 작업에 사용할 열을 선택합니다.

다음 절차에서는 다음과 같이 가정합니다.

- 공동 작업 구성원이 이미 Snowflake에 데이터 테이블을 업로드했습니다.
- (선택 사항) 암호화된 데이터 테이블의 경우에만 공동 작업 구성원은 이미 C3R 암호화 클라이언트를 사용하여 암호화된 데이터 테이블을 준비했습니다.

구성된 테이블을 생성하려면 - Snowflake 데이터 소스

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/cleanrooms AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 테이블을 선택합니다.

3. 오른쪽 상단 모서리에서 새 테이블 구성을 선택합니다.
4. 데이터 소스의 타사 클라우드 및 데이터 소스에서 Snowflake를 선택합니다.
5. 기존 보안 암호 ARN을 사용하거나이 테이블에 대한 새 보안 암호를 저장하여 Snowflake 보안 인증 정보를 지정합니다.

   Use existing secret ARN

   1. 보안 암호 ARN이 있는 경우 보안 암호 ARN 필드에 입력합니다.

      이동을 선택하여 보안 암호 ARN을 조회할 AWS Secrets Manager 수 있습니다.
   2. 다른 테이블의 기존 보안 암호가 있는 경우 기존 테이블에서 보안 암호 ARN 가져오기를 선택합니다.

   > **Note**
   >
   > 보안 암호 ARN은 교차 계정일 수 있습니다.

   Store a new secret for this table

   1. 다음 Snowflake 자격 증명을 입력합니다.
      - Snowflake 사용자 이름
      - Snowflake 암호
      - Snowflake 웨어하우스
      - Snowflake 역할
   2. 기본값을 사용하려면 암호화 설정 사용자 지정 확인란을 선택하지 않은 AWS 관리형 키상태로 둡니다.
   3. 를 사용하려면 암호화 설정 AWS KMS key사용자 지정 확인란을 선택하고 KMS 키를 입력합니다.
   4. 나중에 자격 증명을 찾는 데 도움이 되는 보안 암호 이름을 입력합니다.
6. Snowflake 테이블 및 스키마 세부 정보에 세부 정보를 수동으로 입력하거나 자동으로 세부 정보를 가져옵니다.

   Enter the details manually

   1. Snowflake 계정 식별자를 입력합니다.

자세한 내용은 Snowflake 설명서의 계정 식별자를 참조하세요.

계정 식별자는 Snowflake 드라이버에 사용되는 형식이어야 합니다. 식별자의 형식이 되도록 마침표(.)를 하이픈(-)으로 바꿔야 합니다**<orgname>-<account_name>**.

2. Snowflake 데이터베이스를 입력합니다.

   자세한 내용은 Snowflake 설명서의 Snowflake 데이터베이스를 참조하세요.

3. Snowflake 스키마 이름을 입력합니다.
4. Snowflake 테이블 이름을 입력합니다.

   자세한 내용은 Snowflake 설명서의 Snowflake 테이블 구조 이해를 참조하세요.

5. 스키마에 열 이름을 입력하고 드롭다운 목록에서 데이터 유형을 선택합니다.
6. 열 추가를 선택하여 열을 더 추가합니다.
   - 객체 데이터 유형을 선택하는 경우 객체 스키마를 지정합니다.

     Example 객체 스키마 예제

     ```
     name STRING,
     location OBJECT(
         x INT,
         y INT,
         metadata OBJECT(uuid STRING)
     ),
     history ARRAY(TEXT)
     ```

   - 배열 데이터 유형을 선택하는 경우 배열 스키마를 지정합니다.

     Example 배열 스키마 예제

     ```
     OBJECT(x INT, y INT)
     ```

   - 맵 데이터 유형을 선택하는 경우 맵 스키마를 지정합니다.

     Example 맵 스키마 예제

     ```
     STRING, OBJECT(x INT, y INT)
     ```

Automatically import the details

1. Snowflake에서 COLUMNS 뷰를 CSV 파일로 내보냅니다.

   Snowflake COLUMNS 뷰에 대한 자세한 내용은 Snowflake 설명서의 COLUMNS 뷰를 참조하세요.
2. 파일에서 가져오기를 선택하여 CSV 파일을 가져오고 추가 정보를 지정합니다.

   데이터베이스 이름, 스키마 이름, 테이블 이름, 열 이름 및 데이터 유형을 자동으로 가져옵니다.
   - 객체 데이터 유형을 선택하는 경우 객체 스키마를 지정합니다.
   - 배열 데이터 유형을 선택하는 경우 배열 스키마를 지정합니다.
   - 맵 데이터 유형을 선택하는 경우 맵 스키마를 지정합니다.
3. Snowflake 계정 식별자를 입력합니다.

   자세한 내용은 Snowflake 설명서의 계정 식별자를 참조하세요.

> **Note**
>
> 에 카탈로그화된 S3 테이블만 테이블 스키마를 자동으로 검색하는 데 사용할 AWS Glue 수 있습니다.

7. 공동 작업에 허용된 열에서 목표에 따라 옵션을 선택합니다.

목표	권장 옵션
에서 모든 열 사용 허용 AWS Clean Rooms (분석 규칙에 따라 다름)	모든 컬럼
허용된 열 지정 드롭다운 목록에서 하나 이상의 열 허용	사용자 지정 목록

8. 구성된 테이블 세부 정보의 경우,

   a. 구성된 테이블의 이름을 입력합니다.

      기본 이름을 사용하거나 이 테이블의 이름을 바꿀 수 있습니다.

b. 테이블에 대한 설명을 입력합니다.

   설명은 비슷한 이름을 가진 다른 구성된 테이블을 구분하는 데 도움이 됩니다.

c. 구성된 테이블 리소스에 대해 태그를 활성화하려면 새 태그 추가를 선택한 다음 키와 값 쌍을 입력합니다.

9. 새 클라이언트 구성를 선택합니다.

이제 구성된 테이블을 만들었으므로 다음을 수행할 준비가 되었습니다.

- 구성된 테이블에 분석 규칙 구성
- 구성된 테이블을 공동 작업에 연결합니다

구성된 테이블에 분석 규칙 추가

다음 단원에서는 구성된 테이블에 분석 규칙을 추가하는 방법에 대해 설명합니다. 분석 규칙을 정의하면 AWS Clean Rooms에서 지원하는 특정 분석 규칙과 일치하는 쿼리를 실행하도록 쿼리할 수 있는 구성원을 승인할 수 있습니다.

AWS Clean Rooms 는 다음과 같은 유형의 분석 규칙을 지원합니다.

- 집계 분석 규칙
- 목록 분석 규칙
- 의 사용자 지정 분석 규칙 AWS Clean Rooms

구성된 테이블당 하나의 분석 규칙만 있을 수 있습니다. 구성된 테이블을 공동 작업과 연결하기 전에 언제든지 분석 규칙을 구성할 수 있습니다.

> **Important**
>
> Clean Rooms에 대한 암호화 컴퓨팅을 사용하고 공동 작업에서 데이터 테이블을 암호화한 경우 암호화된 구성 테이블에 추가하는 분석 규칙은 데이터가 암호화된 방식과 일치해야 합니다. 예를 들어 SELECT(집계 분석 규칙)에 대한 데이터를 암호화한 경우 JOIN(목록 분석 규칙)에 대한 분석 규칙을 추가해서는 안 됩니다.

주제

- 테이블에 집계 분석 규칙 추가(안내식 흐름)
- 테이블에 목록 분석 규칙 추가(안내식 흐름)
- 테이블에 사용자 지정 분석 규칙 추가(안내식 흐름)
- 테이블에 분석 규칙 추가(JSON 편집기)
- 다음 단계

테이블에 집계 분석 규칙 추가(안내식 흐름)

집계 분석 규칙을 사용하면 선택적 차원에 따라 COUNT, SUM, AVG 함수를 사용하여 행 수준 정보를 표시하지 않고 통계를 집계하는 쿼리를 사용할 수 있습니다.

이 절차에서는 AWS Clean Rooms 콘솔의 안내식 흐름 옵션을 사용하여 구성된 테이블에 집계 분석 규칙을 추가하는 프로세스를 설명합니다.

> **Note**
>
> non-S3 데이터 소스를 사용하여 구성된 테이블은 사용자 지정 분석 규칙만 지원합니다.

테이블에 집계 분석 규칙을 추가하려면(안내식 흐름)

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/cleanrooms AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 테이블을 선택합니다.
3. 구성된 테이블을 선택합니다.
4. 구성된 테이블 세부 정보 페이지에서 분석 규칙 구성을 선택합니다.
5. 1단계: 분석 규칙 유형 선택의 분석 규칙 유형에서 집계 옵션을 선택합니다.
6. 생성 방법에서 유도 흐름을 선택하고 다음을 선택합니다.
7. 2단계: 쿼리 컨트롤 지정에서 집계 함수의 경우:

    a. 드롭다운에서 집계 함수를 선택합니다.

    - COUNT
    - COUNT DISTINCT
    - SUM

- SUM DISTINCT
- AVG

b. 열 드롭다운에서 집계 함수에 사용할 수 있는 열을 선택합니다.

c. (선택 사항) 다른 함수 추가를 선택하여 다른 집계 함수를 추가하고 하나 이상의 열을 해당 함수에 연결합니다.

> **Note**
>
> 집계 함수가 최소 1개 이상 필요합니다.

d. (선택 사항) 집계 함수를 제거하려면 제거를 선택합니다.

8. 조인 컨트롤의 경우,

a. 테이블을 단독으로 쿼리하도록 허용하는 옵션 하나를 선택합니다.

사용자 선택 항목...	THEN ...
아니요. 오버랩만 쿼리할 수 있습니다	쿼리가 가능한 구성원이 소유한 테이블에 조인된 경우에만 테이블을 쿼리할 수 있습니다.
예	테이블은 단독으로 쿼리하거나 다른 테이블에 조인할 때 쿼리할 수 있습니다.

b. 지정된 조인 열에서는 INNER JOIN 명령문에서 사용할 수 있도록 허용하려는 열을 선택하세요.

이전 단계에서 예를 선택한 경우 이 옵션은 선택 사항입니다.

c. 일치에 사용할 수 있는 연산자 지정에서 여러 조인 열에서 일치시키는 데 사용할 수 있는 연산자(있는 경우)를 선택합니다. 두 개 이상의 JOIN 열을 선택하는 경우 이러한 연산자 중 하나가 필요합니다.

사용자 선택 항목...	THEN ...
AND	테이블 간의 한 열을 다른 열에 조인하는 것을 `INNER JOIN` 일치 조건에 `AND`를 포함할 수 있습니다.

사용자 선택 항목...	THEN ...
또는	`INNER JOIN` 일치 조건에 `OR`을 포함시켜 테이블 간에 여러 열 일치 항목을 결합할 수 있습니다. 이 논리 연산자는 일치율을 높이는 데 유용합니다.

9. (선택 사항) 차원 컨트롤의 경우 차원 열 지정 드롭다운에서 SELECT 문에 사용할 수 있도록 허용할 열과 쿼리의 WHERE, GROUP BY, 및 ORDER BY 일부를 선택합니다.

> **Note**
>
> 집계 함수 또는 조인 열은 차원 열로 사용할 수 없습니다.

10. 스칼라 함수의 경우 어떤 스칼라 함수를 허용하시겠습니까?에 대한 옵션 하나를 선택합니다

사용자 선택 항목...	THEN ...
현재에서 지원되는 모든 AWS Clean Rooms	AWS Clean Rooms에서 현재 지원하는 모든 스칼라 함수를 허용합니다. • 목록 보기를 선택하여 AWS Clean Rooms에서 지원되는 스칼라 함수의 전체 목록을 볼 수 있습니다.
사용자 지정 목록	허용할 스칼라 함수를 사용자 지정할 수 있습니다. • 허용된 스칼라 함수 지정 드롭다운에서 하나 이상의 옵션을 선택합니다.
없음	스칼라 함수는 허용하지 않는 것이 좋습니다.

자세한 내용은 스칼라 함수 단원을 참조하십시오.

11. 다음을 선택합니다.
12. 3단계: 쿼리 결과 제어 지정에서 집계 제약 조건에 대해 다음을 수행하세요.

    a. 각 열 이름의 드롭다운 목록을 선택합니다.

b. 각 출력 행에 COUNT DISTINCT 함수를 적용한 후 반환될 각 출력 행에 대해 충족되어야 하는 고유 값의 최소 수에 대한 드롭다운 목록을 선택합니다.

c. 제약 조건 추가를 선택하여 집계 제약 조건을 더 추가합니다.

d. (선택 사항) 집계 제약 조건을 제거하려면 제거를 선택합니다.

13. 출력에 적용된 추가 분석에서 목표에 따라 옵션을 선택합니다.

목표	권장 옵션
이 테이블에 대한 직접 쿼리만 허용합니다. 쿼리 결과에 대한 추가 분석이 실행되지 않도록 합니다. 이 테이블은 직접 쿼리에만 사용할 수 있습니다.	허용되지 않음
이 테이블에 대한 직접 쿼리와 추가 분석을 모두 허용하지만 반드시 필요한 것은 아닙니다.	허용
필요한 추가 분석 중 하나로 처리되는 직접 쿼리에서만 테이블을 사용할 수 있어야 합니다. 이 테이블에 대한 직접 쿼리는 반환하기 전에 추가로 처리해야 합니다.	필수

14. 다음을 선택합니다.

15. 4단계: 검토 및 구성에서 이전 단계에서 선택한 내용을 검토하고 필요한 경우 편집한 다음 분석 규칙 구성을 선택합니다.

테이블에 집계 분석 규칙을 성공적으로 구성했다는 확인 메시지가 표시됩니다.

테이블에 목록 분석 규칙 추가(안내식 흐름)

목록 분석 규칙을 사용하면 관련 테이블과 쿼리 가능한 구성원의 테이블 간에 겹치는 행 수준 목록을 출력하는 쿼리를 수행할 수 있습니다.

이 절차에서는 AWS Clean Rooms 콘솔의 안내 흐름 옵션을 사용하여 구성된 테이블에 목록 분석 규칙을 추가하는 프로세스를 설명합니다.

> **Note**
>
> non-S3 데이터 소스를 사용하여 구성된 테이블은 사용자 지정 분석 규칙만 지원합니다.

테이블에 목록 분석 규칙 추가하기(안내식 흐름)

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/cleanrooms AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 테이블을 선택합니다.
3. 구성된 테이블을 선택합니다.
4. 구성된 테이블 세부 정보 페이지에서 분석 규칙 구성을 선택합니다.
5. 1단계: 분석 규칙 유형 선택의 분석 규칙 유형에서 목록 옵션을 선택합니다.
6. 생성 방법에서 흐름 안내를 선택하고 다음을 선택합니다.
7. 2단계: 쿼리 컨트롤 지정에서 조인 컨트롤의 경우:

   a. 조인 열 지정에서 INNER JOIN 명령문에 사용하도록 허용하려는 열을 선택합니다.

   b. 일치 허용 연산자 지정에서 다중 조인 열의 일치 연산자 지정에 사용할 수 있는 연산자(있는 경우)를 선택합니다. 두 개 이상의 JOIN 열을 선택하는 경우 이러한 연산자 중 하나가 필요합니다.

사용자 선택 항목...	THEN ...
AND	테이블 간의 한 열을 다른 열에 조인하는 것을 `INNER JOIN` 일치 조건에 `AND`를 포함할 수 있습니다.
또는	`INNER JOIN` 일치 조건에 `OR`을 포함시켜 테이블 간에 여러 열 일치 항목을 결합할 수 있습니다. 이 논리 연산자는 일치율을 높이는 데 유용합니다.

8. (선택 사항) 목록 컨트롤의 경우 목록 열 지정 드롭다운에서 쿼리 출력에 사용(즉, SELECT 명령문에 사용)하거나 결과를 필터링하는 데 사용할 열(즉, WHERE 명령문)을 선택합니다.
9. 다음을 선택합니다.

10. 3단계: 쿼리 결과 컨트롤 지정 아래의 출력에 적용된 추가 분석에서 목표에 따라 옵션을 선택합니다.

목표	권장 옵션
이 테이블에 대한 직접 쿼리만 허용합니다. 쿼리 결과에 대한 추가 분석이 실행되지 않도록 합니다. 이 테이블은 직접 쿼리에만 사용할 수 있습니다.	허용되지 않음
이 테이블에 대한 직접 쿼리와 추가 분석을 모두 허용하지만 반드시 필요한 것은 아닙니다.	허용
필요한 추가 분석 중 하나로 처리되는 직접 쿼리에서만 테이블을 사용할 수 있어야 합니다. 이 테이블에 대한 직접 쿼리는 반환하기 전에 추가로 처리해야 합니다.	필수

11. 4단계: 검토 및 구성에서 이전 단계에서 선택한 내용을 검토하고 필요한 경우 편집한 다음 분석 규칙 구성을 선택합니다.

테이블에 대한 목록 분석 규칙을 성공적으로 구성했다는 확인 메시지가 표시됩니다.

테이블에 사용자 지정 분석 규칙 추가(안내식 흐름)

사용자 지정 분석 규칙은 구성된 테이블에서 사용자 지정 SQL 쿼리 또는 PySpark 작업을 활성화합니다. 다음을 사용하는 경우 사용자 지정 분석 규칙이 필요합니다.

- 사전 승인된 특정 SQL 쿼리 또는 PySpark 작업 세트 또는 데이터를 사용하는 쿼리를 제공할 수 있는 특정 계정 세트를 허용하는 분석 템플릿입니다.
- 사용자 식별 시도로부터 보호하기 위한 AWS Clean Rooms 차등 프라이버시.
- Amazon Athena 또는 Snowflake와 같은 Non-S3 데이터 소스입니다.

이 절차에서는 AWS Clean Rooms 콘솔의 안내 흐름 옵션을 사용하여 구성된 테이블에 사용자 지정 분석 규칙을 추가하는 프로세스를 설명합니다.

테이블에 사용자 지정 분석 규칙 추가하기(안내식 흐름)

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/cleanrooms AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 테이블을 선택합니다.
3. 구성된 테이블을 선택합니다.
4. 구성된 테이블 세부 정보 페이지에서 분석 규칙 구성을 선택합니다.
5. 1단계: 분석 규칙 유형 선택의 분석 규칙 유형에서 사용자 지정 옵션을 선택합니다.
6. 생성 방법에서 흐름 안내를 선택하고 다음을 선택합니다.
7. 2단계: 분석 컨트롤 지정에서 직접 분석 컨트롤의 경우 목표에 따라 옵션을 선택합니다.

목표	권장 조치
이 구성된 테이블에서 실행이 허용되기 전에 각 새 분석을 검토합니다.	1. 실행이 허용된 분석 템플릿에서 분석 템플릿 추가를 선택합니다. 2. 드롭다운 목록에서 적절한 협업 및 분석 템플릿을 선택합니다. 3. 다음을 선택합니다.
특정 공동 작업자가이 테이블에서 검토하지 않고 선택한 유형의 분석을 실행하도록 허용	1. 분석 유형에서 a. AWS 계정 지정한에서 생성된 모든 쿼리를 허용하려면 모든 쿼리를 선택합니다. b. AWS 계정 모든 쿼리를 선택하여 지정한에서 생성한 모든 작업을 허용합니다. 2. AWS 계정 분석을 생성할 수 있도록 허용에서 추가 AWS 계정를 선택합니다. 3. 를 입력 AWS 계정 하거나 드롭다운 목록에서 AWS 계정 ID를 선택합니다. 4. (선택 사항) 다른 추가 AWS 계정를 선택하여 다른를 추가합니다 AWS 계정. 5. 다음을 선택합니다.

8. 3단계: 분석 결과 컨트롤 지정에서

a. 작업 결과 제어의 경우 추가 결과 제어는 지원되지 않습니다.

b. 쿼리 결과 제어의 출력에 허용되지 않는 열에서 목표에 따라 쿼리 출력에 허용할 열을 선택합니다.

목표	권장 조치
쿼리 출력에 모든 열이 반환되도록 허용	1. 없음을 선택합니다. 2. 출력에 적용된 추가 분석으로 진행합니다.
쿼리 출력에서 특정 열이 반환되지 않도록 허용	1. 사용자 지정 목록 선택 2. 허용되지 않는 열 지정에서 쿼리 출력에서 제거할 열을 선택합니다.

c. 출력에 적용된 추가 분석의 경우 목표에 따라 쿼리 출력에 추가 분석을 적용할 수 있는지 여부를 선택합니다.

목표	권장 옵션
• 이 테이블에 대한 직접 쿼리만 허용합니다. • 쿼리 결과에 대한 추가 분석이 실행되지 않도록 합니다. • 이 테이블은 직접 쿼리에만 사용할 수 있습니다.	허용되지 않음
이 테이블에 대한 직접 쿼리와 추가 분석을 모두 허용하지만 반드시 필요한 것은 아닙니다.	허용
• 필요한 추가 분석 중 하나로 처리되는 직접 쿼리에서만 테이블을 사용할 수 있어야 합니다. • 이 테이블에 대한 직접 쿼리는 반환하기 전에 추가로 처리해야 합니다.	필수

d. 다음을 선택합니다.

9. (선택 사항) 4단계: 차등 프라이버시 설정에서 차등 프라이버시를 켜거나 끌지 여부를 결정합니다.

차등 프라이버시는 재식별 공격으로부터 데이터를 보호하는 수학적으로 입증된 기술입니다.

> **Note**
>
> AWS Clean Rooms 차등 프라이버시는 Amazon S3에 저장된 데이터를 사용하는 공동 작업에만 사용할 수 있습니다.

차등 프라이버시의 경우 목표에 따라 차등 프라이버시를 켤지 아니면 끌지 선택합니다.

목표	권장 조치
• 재식별 시도에 대한 보호가 필요하지 않습니다. • 테이블에 사용자 수준 데이터가 없습니다.	1. 끄기를 선택합니다. 2. 다음을 선택합니다.
• 재식별 시도로부터 보호해야 합니다. • 테이블에 사용자 수준 데이터가 있음	1. 켜기를 선택합니다. 2. 개인 정보를 보호하려는 열과 같은 사용자의 고유 식별자가 포함된 사용자 식별자 `user_id` 열을 선택합니다. 공동 작업에서 2개 이상의 테이블에 대해 차등 프라이버시를 켜려면 두 분석 규칙에서 사용자 식별자 열과 동일한 열을 구성하여 테이블 간에 사용자에 대한 정의를 일관되게 유지해야 합니다. 구성이 잘못된 경우 쿼리를 수행할 수 있는 구성원에게는 쿼리를 실행하는 동안 사용자 기여도 숫자(예: 사용자의 광고 노출 수)를 계산하기 위해 2개의 열 중에서 선택할 수 있다는 오류 메시지가 표시됩니다. 3. 다음을 선택합니다.

10. 5단계: 검토 및 구성에서 이전 단계에서 선택한 내용을 검토하고 필요한 경우 편집한 다음 분석 규칙 구성을 선택합니다.

테이블에 대한 사용자 지정 분석 규칙을 성공적으로 구성했다는 확인 메시지가 표시됩니다.

테이블에 분석 규칙 추가(JSON 편집기)

다음 절차에서는 AWS Clean Rooms 콘솔의 JSON 편집기 옵션을 사용하여 테이블에 분석 규칙을 추가하는 방법을 보여줍니다.

> **Note**
>
> non-S3 데이터 소스를 사용하여 구성된 테이블은 사용자 지정 분석 규칙만 지원합니다.

테이블에 집계, 목록 또는 사용자 지정 분석 규칙을 추가하려면(JSON 편집기)

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/cleanrooms AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 테이블을 선택합니다.
3. 구성된 테이블을 선택합니다.
4. 구성된 테이블 세부 정보 페이지에서 분석 규칙 구성을 선택합니다.
5. 1단계: 분석 규칙 유형 선택의 분석 규칙 유형에서 집계, 목록 또는 사용자 지정 옵션을 선택합니다.
6. 생성 방법에서 JSON 편집기를 선택하고 다음을 선택합니다.
7. 2단계: 컨트롤 지정에서 쿼리 구조를 삽입(템플릿 삽입)하거나 파일을 삽입(파일에서 가져오기)하도록 선택할 수 있습니다.

사용자 선택 항목...	THEN ...
템플릿 삽입	1. 분석 규칙 정의에서 선택한 분석 규칙의 매개변수를 지정합니다. 2. Ctrl + 스페이스바를 눌러 자동 완성을 활성화할 수 있습니다.

사용자 선택 항목...	THEN ...
	집계 분석 규칙 파라미터에 대한 자세한 내용은 집계 분석 규칙 - 쿼리 제어 섹션을 참조하세요. 리스트 분석 규칙 파라미터에 대한 자세한 내용은 목록 분석 규칙 - 쿼리 제어 섹션을 참조하세요.
Import From File	1. 로컬 드라이브에서 JSON 파일을 선택합니다. 2. Open을 선택합니다. 분석 규칙 정의에는 업로드된 파일의 분석 규칙이 표시됩니다.

8. 다음을 선택합니다.
9. 3단계: 검토 및 구성에서 이전 단계에서 선택한 내용을 검토하고 필요한 경우 편집한 다음 분석 규칙 구성을 선택합니다.

테이블에 대한 분석 규칙을 성공적으로 구성했다는 확인 메시지가 나타납니다.

다음 단계

구성된 테이블에 분석 규칙을 구성했으므로 이제 다음을 수행할 준비가 되었습니다.

- 구성된 테이블을 공동 작업에 연결
- (쿼리할 수 있는 구성원으로)데이터 테이블을 쿼리

구성된 테이블을 공동 작업에 연결

구성된 테이블을 생성하고 여기에 분석 규칙을 추가한 후 이를 공동 작업에 연결하고 AWS Glue 테이블에 액세스할 수 AWS Clean Rooms 있는 서비스 역할을 부여할 수 있습니다.

> **Note**
>
> 이 서비스 역할에는 테이블에 대한 권한이 있습니다. 서비스 역할은 쿼리할 수 있는 회원을 대신하여 허용된 쿼리를 실행할 수 있는 AWS Clean Rooms 만 맡을 수 있습니다. 공동 작업 구성원(데이터 소유자 제외)은 공동 작업의 기본 테이블에 액세스할 수 없습니다. 데이터 소유자는 차등 프라이버시 기능을 켜서 다른 구성원이 자신의 테이블을 쿼리할 수 있도록 할 수 있습니다.

> **Important**
>
> 구성된 AWS Glue 테이블을 공동 작업에 연결하기 전에 AWS Glue 테이블 위치가 단일 파일이 아닌 Amazon Simple Storage Service(Amazon S3) 폴더를 가리켜야 합니다. https://console.aws.amazon.com/glue/ AWS Glue 콘솔에서 테이블을 확인하여이 위치를 확인할 수 있습니다.

> **Note**
>
> 에서 암호화를 구성 AWS Glue 하고 서비스 역할을 생성한 경우 해당 역할에 AWS Glue 테이블을 복호화하는 AWS KMS keys 데 사용할 수 있는 액세스 권한을 부여해야 합니다.
> AWS KMS암호화된 Amazon S3 데이터 세트가 지원하는 구성된 테이블을 연결한 경우 역할에 KMS 키를 사용하여 Amazon S3 데이터를 해독할 수 있는 액세스 권한을 부여해야 합니다.
> 자세한 내용은 AWS Glue 개발자 안내서의 AWS Glue에서 암호화 설정을 참조하세요.

다음 주제에서는 AWS Clean Rooms 콘솔을 사용하여 구성된 테이블을 공동 작업에 연결하는 방법을 설명합니다.

주제

- 구성된 테이블 세부 정보 페이지에서 구성된 테이블을 연결
- 공동 작업 세부 정보 페이지에서 구성된 테이블을 연결
- 다음 단계

AWS SDK를 사용하여 구성된 테이블을 공동 작업에 연결하는 방법에 대한 자세한 내용은 AWS Clean Rooms API 참조를 참조하세요.

구성된 테이블 세부 정보 페이지에서 구성된 테이블을 연결

구성된 AWS Glue 테이블 세부 정보 페이지에서 테이블을 공동 작업에 연결하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 테이블을 선택합니다.
3. 구성된 테이블을 선택합니다.
4. 구성된 테이블 세부 정보 페이지에서 공동 작업에 연결을 선택합니다.
5. 공동 작업에 테이블 연결 대화 상자의 경우 드롭다운 목록에서 공동 작업을 선택합니다.
6. 공동 작업 선택을 선택합니다.

   테이블 연결 페이지에서 선택한 구성 테이블의 이름이 구성된 테이블 선택 섹션 아래에 표시됩니다.
7. (선택 사항) 구성된 테이블 선택에서 다음을 수행합니다.

다음을 수행하려는 경우	THEN ...
새 테이블 생성	테이블 구성을 선택하고 테이블 구성 페이지의 지시를 따릅니다.
구성된 테이블의 체계 및 분석 규칙 보기	체계 및 분석 규칙 보기를 켭니다.

8. 테이블 연결 세부 정보의 경우,
   a. 관련 테이블의 이름을 입력합니다.

      기본 이름을 사용하거나 이 테이블의 이름을 바꿀 수 있습니다.
   b. (선택 사항) 작업에 대한 설명을 입력합니다.

      설명은 쿼리 작성에 도움이 됩니다.
9. 새 서비스 역할 생성 및 사용 또는 기존 서비스 역할 사용을 선택하여 서비스 액세스 권한을 지정합니다.

> **Note**
>
> Amazon Athena에서 지원하는 구성된 테이블을 연결하는 경우 드롭다운 목록에서 기존 서비스 역할 이름을 선택합니다. 서비스 역할에 데이터 세트에 대한 IAM 및 필요한 경우 Lake Formation 권한이 있는지 확인합니다.

사용자 선택 항목...	THEN ...
새 서비스 역할 생성 및 사용	• AWS Clean Rooms 는이 테이블에 필요한 정책을 사용하여 서비스 역할을 생성합니다. • 기본 서비스 역할 이름은 `cleanrooms-<timestamp>` 입니다 • 역할을 생성하고 정책을 연결할 수 있는 권한이 있어야 합니다. • 입력 데이터가 암호화된 경우 이 데이터는 KMS 키로 암호화됨을 선택한 다음 데이터 입력을 복호화 AWS KMS key 하는 데 사용할를 입력할 수 있습니다.

사용자 선택 항목...	THEN ...
기존 서비스 역할 사용	1. 드롭다운 목록에서 기존 서비스 역할 이름을 선택합니다. 역할을 나열할 권한이 있는 경우 역할 목록이 표시됩니다. 역할을 나열할 수 있는 권한이 없는 경우 사용하려는 역할의 Amazon 리소스 이름(ARN)을 입력할 수 있습니다. 2. IAM에서 보기 외부 링크를 선택하여 서비스 역할을 확인합니다. 기존 서비스 역할이 없는 경우 기존 서비스 역할 사용 옵션을 사용할 수 없습니다. 기본적으로 AWS Clean Rooms 는 필요한 권한을 추가하기 위해 기존 역할 정책을 업데이트하려고 시도하지 않습니다. 3. (선택 사항) 필요한 권한이 포함된 사전 구성된 정책을 이 역할에 추가 확인란을 선택하여 필요한 권한을 역할에 추가합니다. 역할을 수정하고 정책을 생성할 권한이 있어야 합니다.

Note

- AWS Clean Rooms 에는 분석 규칙에 따라 쿼리할 수 있는 권한이 필요합니다. 권한에 대한 자세한 내용은 섹션을 AWS Clean Rooms참조하세요AWS 에 대한 관리형 정책 AWS Clean Rooms.
- 역할에에 대한 충분한 권한이 없는 경우 역할에에 대한 충분한 권한이 없다는 오류 메시지가 AWS Clean Rooms표시됩니다 AWS Clean Rooms. 계속하기 전에 역할 정책을 추가해야 합니다.

- 역할 정책을 수정할 수 없는 경우 AWS Clean Rooms 에서 서비스 역할에 대한 정책을 찾을 수 없다는 오류 메시지가 나타납니다.

10. 구성된 테이블 연결 리소스에 대해 태그를 활성화하려면 새 태그 추가를 선택한 다음 키 및 값 쌍을 입력합니다.
11. 연결 테이블을 선택합니다.

공동 작업 세부 정보 페이지에서 구성된 테이블을 연결

공동 작업 세부 정보 페이지에서 AWS Glue 테이블을 공동 작업에 연결하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업을 선택합니다.
4. 테이블 탭에서 테이블 연결을 선택합니다.
5. 구성된 테이블 선택에서 다음 중 하나를 수행합니다.

다음을 수행하려는 경우	THEN ...
기존 구성된 테이블 선택	드롭다운 목록에서 공동 작업과 연결할 구성된 테이블 이름을 선택합니다.
새 테이블 생성	테이블 구성을 선택하고 테이블 구성 페이지의 지시를 따릅니다.
구성된 테이블의 체계 및 분석 규칙 보기	체계 및 분석 규칙 보기를 켭니다.

6. 테이블 연결 세부 정보의 경우,

   a. 관련 테이블의 이름을 입력합니다.

      기본 이름을 사용하거나 이 테이블의 이름을 바꿀 수 있습니다.

   b. (선택 사항) 작업에 대한 설명을 입력합니다.

      설명은 쿼리 작성에 도움이 됩니다.

7. 새 서비스 역할 생성 및 사용 또는 기존 서비스 역할 사용을 선택하여 서비스 액세스 권한을 지정합니다.

> **Note**
>
> Amazon Athena에서 지원하는 구성된 테이블을 연결하는 경우 드롭다운 목록에서 기존 서비스 역할 이름을 선택합니다. 서비스 역할에 데이터 세트에 대한 IAM 및 필요한 경우 Lake Formation 권한이 있는지 확인합니다.

사용자 선택 항목...	THEN ...
새 서비스 역할 생성 및 사용	• AWS Clean Rooms 는이 테이블에 필요한 정책을 사용하여 서비스 역할을 생성합니다. • 기본 서비스 역할 이름은 `cleanrooms-<timestamp>` 입니다. • 역할을 생성하고 정책을 연결할 수 있는 권한이 있어야 합니다. • 입력 데이터가 암호화된 경우 이 데이터는 KMS 키로 암호화됨을 선택한 다음 데이터 입력을 복호화 AWS KMS key 하는 데 사용할를 입력할 수 있습니다.

사용자 선택 항목...	THEN ...
기존 서비스 역할 사용	1. 드롭다운 목록에서 기존 서비스 역할 이름을 선택합니다. 역할을 나열할 권한이 있는 경우 역할 목록이 표시됩니다. 역할을 나열할 수 있는 권한이 없는 경우 사용하려는 역할의 Amazon 리소스 이름(ARN)을 입력할 수 있습니다. 2. IAM에서 보기 외부 링크를 선택하여 서비스 역할을 확인합니다. 기존 서비스 역할이 없는 경우 기존 서비스 역할 사용 옵션을 사용할 수 없습니다. 기본적으로 AWS Clean Rooms 는 필요한 권한을 추가하기 위해 기존 역할 정책을 업데이트하려고 시도하지 않습니다. 3. (선택 사항) 필요한 권한이 포함된 사전 구성된 정책을 이 역할에 추가 확인란을 선택하여 필요한 권한을 역할에 추가합니다. 역할을 수정하고 정책을 생성할 권한이 있어야 합니다.

Note

- AWS Clean Rooms 에는 분석 규칙에 따라 쿼리할 수 있는 권한이 필요합니다. 권한에 대한 자세한 내용은 섹션을 AWS Clean Rooms참조하세요AWS 에 대한 관리형 정책 AWS Clean Rooms.
- 역할에에 대한 충분한 권한이 없는 경우 역할에에 대한 충분한 권한이 없다는 오류 메시지가 AWS Clean Rooms표시됩니다 AWS Clean Rooms. 계속하기 전에 역할 정책을 추가해야 합니다.

- 역할 정책을 수정할 수 없는 경우 AWS Clean Rooms 에서 서비스 역할에 대한 정책을 찾을 수 없다는 오류 메시지가 나타납니다.

8. 구성된 테이블 연결 리소스에 대해 태그를 활성화하려면 새 태그 추가를 선택한 다음 키 및 값 쌍을 입력합니다.
9. 연결 테이블을 선택합니다.

다음 단계

구성된 데이터 테이블을 공동 작업에 연결했으니 이제 다음을 수행할 준비가 되었습니다.

- 구성된 테이블에 공동 작업 분석 규칙 추가
- 공동 작업 생성자인 경우 공동 작업 편집
- 데이터 테이블 쿼리 (쿼리할 수 있는 멤버로서)

구성된 테이블에 공동 작업 분석 규칙 추가

공동 작업 분석 규칙을 사용하면 이 공동 작업과 관련된 컨트롤을 지정할 수 있습니다. 이러한 컨트롤은 구성된 테이블 분석 규칙과 함께 작동하여 이 공동 작업 내에서 이 테이블을 분석할 수 있는 방법을 결정합니다.

공동 작업 분석 규칙은 구성된 테이블을 생성하고 분석 규칙을 추가하고 이를 공동 작업에 연결한 후 구성된 테이블에 추가합니다. 테이블이 직접 분석을 지원하거나 추가 분석을 허용하도록 구성된 경우 공동 작업 분석 규칙을 추가해야 합니다.

- 직접 분석 - 테이블을 직접 분석하는 쿼리에서 테이블을 사용할 수 있습니다. 예를 들어, 집계 측정 분석 또는 활성화를 위한 식별자 목록을 출력하는 쿼리에서 사용할 수 있습니다.
- 추가 분석 - 테이블을 직접 분석하는 쿼리 외에도 추가 분석에 대한 입력으로 사용할 수도 있습니다. 예를 들어 테이블은 유사 ML 모델의 시드인 쿼리 또는 사용자 지정 ML 모델의 ML 입력 채널에 사용할 수 있습니다.

테이블에 공동 작업 분석 규칙을 추가하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 수행하지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.

3. 공동 작업을 선택합니다.
4. 테이블 탭의 내가 연결한 테이블에서 공동 작업에 연결한 구성된 테이블을 확인합니다.

   직접 분석 상태 또는 추가 분석 상태가 준비 완료 상태인 경우 테이블을 쿼리할 준비가 된 것입니다.
5. 직접 분석 상태 또는 추가 분석 상태가 준비되지 않음 상태인 경우 상태를 선택한 다음 대화 상자에서 구성을 선택합니다.
6. 공동 작업 분석 규칙 구성 페이지에서 구성된 테이블 분석 규칙 보기를 확장하여 세부 정보를 확인합니다.
7. 허용된 추가 분석에서 목표에 따라 옵션을 선택합니다.

목표	권장 옵션
테이블에 대한 추가 분석을 허용합니다.	모두
특정 멤버의 테이블에 대한 추가 분석만 허용합니다.	특정 멤버가 수행하는 모든 항목
테이블에 대한 특정 분석만 허용합니다.	사용자 지정 목록

8. 결과 전달에서 쿼리 출력에 대한 결과를 수신할 수 있는 멤버 드롭다운에서 결과를 수신할 수 있는 사용자를 지정합니다.
9. 분석 규칙 구성을 선택합니다.

차등 프라이버시 정책 구성(선택 사항)

> **Note**
>
> AWS Clean Rooms 차등 프라이버시는 AWS Clean Rooms SQL을 분석 엔진으로 사용하고 Amazon S3에 저장된 데이터를 사용하는 공동 작업에만 사용할 수 있습니다.

이 절차에서는 AWS Clean Rooms 콘솔의 안내 흐름 옵션을 사용하여 공동 작업에서 차등 프라이버시 정책을 구성하는 프로세스를 설명합니다. 차등 프라이버시 보호 기능을 갖춘 모든 테이블에서 이 단계를 한 번 거쳐야 합니다.

차등 프라이버시 설정을 구성하려면(안내식 흐름)

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/cleanrooms AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업을 선택합니다.
4. 공동 작업 페이지의 테이블 탭에서 차등 프라이버시 정책 구성을 선택합니다.
5. 차등 프라이버시 정책 구성 페이지에서 다음 속성의 값을 선택합니다.

   - 프라이버시 예산
   - 매월 프라이버시 예산 새로 고침
   - 쿼리당 추가된 노이즈

   기본값을 사용하거나 특정 사용 사례를 지원하는 사용자 지정 값을 입력할 수 있습니다. 프라이버시 예산 및 쿼리당 추가된 노이즈 값을 선택한 후 데이터에 대한 모든 쿼리에서 가능한 집계 수를 기준으로 결과 유틸리티를 미리 볼 수 있습니다.
6. 구성을 선택합니다.

공동 작업을 위한 차등 프라이버시 정책을 성공적으로 구성했다는 확인 메시지가 표시됩니다.

차등 프라이버시를 구성했으므로 이제 다음을 수행할 준비가 되었습니다.

- 데이터 테이블 쿼리 (쿼리할 수 있는 멤버로서)
- 공동 작업(공동 작업 생성자인 경우)

차등 프라이버시 사용량 로그 보기

차등 프라이버시를 사용해 데이터를 보호하는 공동 작업 구성원은 차등 프라이버시 기능을 사용하는 공동 작업을 만든 후 프라이버시 예산 사용량을 모니터링할 수 있습니다.

실행된 집계 수와 프라이버시 예산 중 사용된 프라이버시 예산을 보려면

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/cleanrooms AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.

3. 공동 작업을 선택합니다.
4. 테이블 탭을 선택합니다.
5. 사용량 로그 보기(파란색 텍스트)를 선택합니다.
6. 프라이버시 예산 및 제공된 유틸리티 금액을 포함한 사용량 세부 정보를 확인하세요.

차등 프라이버시 정책 편집

차등 프라이버시 정책을 구성한 후 언제든지 프라이버시 요구 사항이 더 잘 반영되도록 업데이트할 수 있습니다.

차등 프라이버시 정책을 편집하려면

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/cleanrooms AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업을 선택합니다.
4. 공동 작업 페이지의 테이블 탭에 있는 내가 연결한 테이블에서 편집을 선택합니다.
5. 차등 프라이버시 정책 편집 페이지에서 다음 속성의 값을 새로 선택합니다.

   - 프라이버시 예산 - 공동 작업 중 언제든지 슬라이더 막대를 움직여 예산을 늘리거나 줄일 수 있습니다. 쿼리할 수 있는 구성원이 데이터 쿼리를 시작하면 예산을 줄일 수 없습니다. 개인 정보 보호 예산이 증가하면 AWS Clean Rooms 는 새로 추가된 개인 정보 보호 예산을 활용하기 전에 기존 예산이 완전히 소비될 때까지 기존 예산을 계속 사용합니다.
   - 쿼리당 추가된 노이즈 - 공동 작업 중 언제든지 슬라이더 막대를 움직여 쿼리당 추가되는 노이즈를 늘리거나 줄일 수 있습니다.

   > **Note**
   >
   > 대화형 예제를 선택하여 프라이버시 예산 및 쿼리당 추가되는 노이즈 값의 차이가 실행 가능한 집계 함수의 수에 어떤 영향을 미치는지 살펴볼 수 있습니다.

   프라이버시 예산 새로 고침의 값은 변경할 수 없습니다. 선택 항목을 변경하려면 차등 프라이버시를 삭제하고 새 정책을 생성해야 합니다.
6. 변경 사항 저장을 선택합니다.

차등 프라이버시 정책을 성공적으로 편집했다는 확인 메시지가 표시됩니다.

차등 프라이버시 정책 삭제

공동 작업의 테이블 탭에서 차등 프라이버시 정책을 삭제할 수 있습니다.

차등 프라이버시 정책을 삭제하려면

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/cleanrooms AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업을 선택합니다.
4. 공동 작업 페이지의 테이블 탭에서 차등 프라이버시 정책 옆의 삭제를 선택합니다.
5. 차등 프라이버시 정책을 삭제하려면 삭제를 선택합니다.

차등 프라이버시 정책을 삭제하면 해당 정책에서 프라이버시 예산 사용 로그에 액세스할 수 없습니다. 차등 프라이버시 정책이 삭제되면 차등 프라이버시 기능이 켜진 테이블은 쿼리할 수 없습니다.

계산된 차등 프라이버시 파라미터 보기

차등 프라이버시에 대한 전문 지식이 있는 사용자의 경우 공동 작업의 쿼리 탭에서 계산된 차등 프라이버시 파라미터를 볼 수 있습니다.

계산된 차등 프라이버시 파라미터를 보려면

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/cleanrooms AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업을 선택합니다.
4. 쿼리 탭의 결과 섹션에서 계산된 차등 프라이버시 파라미터 보기를 선택합니다.

계산된 차등 프라이버시 파라미터 표에서 집계 함수의 민감도 값을 볼 수 있습니다. 이 값은 단일 사용자 레코드가 추가, 제거 또는 수정될 경우 함수 결과가 변경될 수 있는 최대량으로 정의됩니다. 목록에는 다음 차등 프라이버시 파라미터가 포함됩니다.

- 사용자 기여 한도(UCL)는 SQL 쿼리에서 사용자가 기여한 최대 행 수입니다. 예를 들어 각 사용자가 여러 개의 노출을 가질 수 있는 지정된 캠페인에서 일치하는 총 노출 수를 계산하려는 경우 AWS

Clean Rooms 차등 프라이버시 계산이 정확하려면 차등 프라이버시가 단일 사용자의 노출 수를 제한해야 합니다. 즉, 사용자에게 바인딩보다 많은 노출이 있는 경우는 계산된 UCL 값에 따라 해당 사용자의 노출에 대한 균일한 무작위 샘플을 AWS Clean Rooms 자동으로 가져와 쿼리를 실행하는 동안 해당 사용자의 나머지 노출을 제외합니다. 고유 사용자 수를 세는 경우 UCL 값은 1이 됩니다. 한 명의 사용자를 추가, 제거 또는 수정하면 개별 사용자 수가 최대 1명까지 변경될 수 있기 때문입니다.

- 최소값은 다음과 같은 집계 함수 내에서 사용되는 표현식의 하한입니다(예: `sum( )`). 예를 들어 표현식이 `purchase_value`로 알려진 열인 경우 최솟값은 열의 하한입니다.
- 최대값은 다음과 같은 집계 함수 내에서 사용되는 표현식의 상한입니다(예: `sum( )`). 예를 들어 표현식이 `purchase_value`로 알려진 열인 경우 최댓값은 열의 상한입니다.

계산된 차등 프라이버시 파라미터 표에서 이러한 파라미터를 사용하여 쿼리 결과의 총 노이즈 양을 더 잘 이해할 수 있습니다. 예를 들어 쿼리당 추가된 구성된 노이즈가 사용자 30명이고 `COUNT DISTINCT (user_id)` 쿼리가 실행되는 경우 차 AWS Clean Rooms 등 프라이버시는의 민감도가 1이기 때문에 확률이 높은 -30~30 사이의 임의 노이즈`COUNT DISTINCT`를 추가합니다. 동일한 구성이 포함된 쿼리의 경우 AWS Clean Rooms 차등 프라이버시를 사용하면 한 명의 사용자가 쿼리 결과에 여러 행을 입력할 수 있으므로 사용자 기여도 한도에 따라 규모가 조정되는 통계적 노이즈가 추가됩니다. 모든 열 값이 양수인 `SUM` `SUM (purchase_value)` 쿼리의 경우 총 노이즈는 사용자 기여 한도에 최대값을 곱하여 조정됩니다. AWS Clean Rooms 차등 프라이버시는 자동으로 민감도 파라미터를 계산하여 쿼리 런타임에 노이즈를 추가하고 프라이버시 예산을 고갈합니다. 민감도 파라미터는 데이터에 따라 달라지므로 프라이버시 예산이 사용되어야 합니다.

테이블 및 분석 규칙 보기

공동 작업 및 분석 규칙과 관련된 테이블을 보려면

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/cleanrooms AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업을 선택합니다.
4. 테이블 탭을 선택합니다.
5. 다음 중 하나를 선택합니다.

    a. 공동 작업과 관련된 테이블을 보려면 귀하와 연결된 테이블의 경우 테이블(파란색 텍스트)을 선택합니다.

b. 공동 작업과 관련된 다른 테이블을 보려면 공동 작업자가 연계한 테이블의 경우 테이블(파란색 텍스트)을 선택합니다.

6. 테이블 세부정보 페이지에서 테이블 세부 정보 및 분석 규칙을 확인하세요.

구성된 테이블 세부 정보 편집

공동 작업 구성원은 구성된 테이블 세부 정보를 편집할 수 있습니다.

구성된 테이블 세부 정보를 편집하려면

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/cleanrooms AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 테이블을 선택합니다.
3. 생성한 구성된 테이블을 선택합니다.
4. 구성된 테이블 세부 정보 페이지에서 구성된 테이블 세부 정보까지 아래로 스크롤합니다.
5. 편집을 선택합니다.
6. 구성된 테이블의 이름 또는 설명을 업데이트합니다.
7. 변경 사항 저장을 선택합니다.

구성된 테이블 태그 편집

공동 작업 구성원은 구성된 테이블을 만든 후 구성된 테이블 탭에서 구성된 테이블 리소스의 태그를 관리할 수 있습니다.

구성된 테이블 태그를 편집하려면

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/cleanrooms AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 테이블을 선택합니다.
3. 생성한 구성된 테이블을 선택합니다.
4. 구성된 테이블 세부 정보 페이지에서 태그 섹션으로 스크롤합니다.
5. 태그 관리를 선택합니다.
6. 태그 관리 페이지에서 다음 작업을 수행할 수 있습니다.

- 태그를 제거하려면 제거를 선택합니다.
- 태그를 추가하려면 태그 추가를 선택합니다.
- 변경 사항을 저장하려면 변경 사항 저장을 선택합니다.

구성된 테이블 분석 규칙 편집

구성된 테이블 분석 규칙을 편집하려면

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/cleanrooms AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 테이블을 선택합니다.
3. 생성한 구성된 테이블을 선택합니다.
4. 구성된 테이블 세부 정보 페이지에서 아래로 스크롤하여 집계 분석 규칙, 목록 분석 규칙 또는 사용자 지정 분석 규칙 섹션으로 이동합니다. (선택은 구성된 테이블에 대해 선택한 분석 규칙 유형에 따라 달라집니다.)
5. 편집을 선택합니다.
6. 분석 규칙 편집 페이지에서 다음을 수행할 수 있습니다.
   - 다음과 같이 분석 규칙 정의를 수정합니다.
     - JSON 편집기를 수정합니다.
     - 파일에서 가져오기를 선택하여 새 분석 규칙 정의를 업로드합니다.
   - 다음 옵션 중에서 선택하여 구성원이 공동 작업에서 보게 될 내용을 미리 볼 수 있습니다.
     - 테이블 보기
     - JSON
     - 쿼리 예
7. 변경 사항을 저장하려면 변경 사항 저장을 선택합니다.

구성된 테이블 분석 규칙 삭제

> **Warning**
>
> 이 작업은 취소할 수 없으며 모든 관련 리소스에 영향을 줍니다.

구성된 테이블 분석 규칙을 삭제하려면

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/cleanrooms AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 테이블을 선택합니다.
3. 생성한 구성된 테이블을 선택합니다.
4. 구성된 테이블 세부 정보 페이지에서 아래로 스크롤하여 집계 분석 규칙, 목록 분석 규칙 또는 사용자 지정 분석 규칙 섹션으로 이동합니다. (선택은 구성된 테이블에 대해 선택한 분석 규칙 유형에 따라 달라집니다.)
5. 삭제를 선택합니다.
6. 분석 규칙을 삭제하려는 것이 확실하다면 삭제를 선택합니다.

구성된 테이블의 허용되지 않는 열

허용되지 않는 출력 열 구성은 쿼리 결과에가 프로젝션되도록 허용하지 않는 열 목록(있는 경우)을 정의할 수 있는 AWS Clean Rooms 사용자 지정 분석 규칙의 제어입니다. 이 목록에 참조된 열은 '허용되지 않는 출력 열'로 간주됩니다. 즉, 이러한 열을 변환, 별칭 지정 또는 기타 방법으로 참조한 경우 쿼리의 최종 SELECT(프로젝션)에 표시되지 않을 수 있습니다.

이 기능은 열이 출력에 직접 표시되는 것은 차단하지만, 다른 메커니즘을 통해 기본값을 간접적으로 유추하는 것을 완전히 차단하지는 못합니다. 이러한 열은 최종 프로젝션에서 참조되지 않는 한 하위 쿼리나 공통 테이블 식(CTE)과 같은 프로젝션 절에서 계속 사용할 수 있습니다.

허용되지 않는 출력 열 구성은 사용 사례와 해당 개인 정보 보호 요구 사항에 따라 분석 템플릿 수준의 검토와 결합하여 테이블에 대한 제어를 적용하고 코드화할 있는 유연성을 제공합니다.

이 구성을 설정하는 방법에 대한 자세한 내용은 테이블에 사용자 지정 분석 규칙 추가(안내식 흐름) 단원을 참조하세요.

예시

다음 예시는 허용되지 않는 출력 열 제어가 적용되는 방식을 보여줍니다.

- 구성원 A는 구성원 B와 공동 작업을 하고 있습니다.
- 구성원 B는 쿼리를 실행할 수 있는 구성원입니다.
- 구성원 A는 age, gender, email, name 열이 포함된 users 테이블을 정의합니다. age 및 name 열은 허용되지 않는 출력 열입니다.

- 구성원 B는 age, gender, owner_name이라는 유사한 열 집합이 포함된 pets 테이블을 정의합니다. 출력 열에 대한 제약 조건을 설정하지 않았으므로, 테이블의 모든 열이 쿼리에 자유롭게 표시될 수 있습니다.

구성원 B가 다음 쿼리를 실행할 경우 허용되지 않는 출력 열을 직접 표시할 수 없으므로 차단됩니다.

```
SELECT
  age
FROM
  users
```

구성원 B가 다음 쿼리를 실행할 경우 허용되지 않는 출력 열을 project star를 통해 암시적으로 표시할 수 없으므로 차단됩니다.

```
SELECT
  *
FROM
  users
```

구성원 B가 다음 쿼리를 실행할 경우 허용되지 않는 출력 열의 변환을 표시할 수 없으므로 차단됩니다.

```
SELECT
  COUNT(age)
FROM
  users
```

구성원 B가 다음 쿼리를 실행할 경우 별칭을 사용하여 허용되지 않는 출력 열을 최종 프로젝션에서 참조할 수 없으므로 차단됩니다.

```
SELECT
  count_age
FROM
  (SELECT COUNT(age) AS count_age FROM users)
```

구성원 B가 다음 쿼리를 실행할 경우 변환된 제한 열이 출력에 투영되므로 차단됩니다.

```
SELECT
  CONCAT(name, email)
FROM
  users
```

구성원 B가 다음 쿼리를 실행할 경우 CTE에 정의된 허용되지 않는 출력 열을 최종 프로젝션에서 참조할 수 없으므로 차단됩니다.

```
WITH cte AS (
  SELECT
    age AS age_alias
  FROM
    users
)
SELECT age_alias FROM cte
```

구성원 B가 다음 쿼리를 실행할 경우 허용되지 않는 출력 열을 최종 프로젝션에서 정렬 또는 파티션 키로 사용할 수 없으므로 차단됩니다.

```
SELECT
  LISTAGG(gender) WITHIN GROUP (ORDER BY age) OVER (PARTITION BY age)
FROM
  users
```

구성원 B가 다음 쿼리를 실행할 경우 허용되지 않는 출력 열의 일부인 열은 조인 또는 필터 절과 같은 쿼리의 다른 구문에서 계속 사용할 수 있으므로 쿼리가 성공합니다.

```
SELECT
  u.name,
  p.gender,
  p.age
FROM
  users AS u
JOIN
  pets AS p
ON
  u.name = p.owner_name
```

동일한 시나리오에서 구성원 B는 users의 name 열을 필터 또는 정렬 키로 사용할 수도 있습니다.

```
SELECT
  u.email,
  u.gender
FROM
  users AS u
WHERE
  u.name = 'Mike'
ORDER BY
  u.name
```

또한 users의 허용되지 않는 출력 열은 다음과 같은 하위 쿼리 및 CTE와 같은 중간 프로젝션에 사용할 수 있습니다.

```
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)
```

구성된 테이블 연결 편집

공동 작업 구성원은 자신이 만든 구성된 테이블 연결을 편집할 수 있습니다.

구성된 테이블 연결을 편집하려는 경우

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/cleanrooms AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업을 선택합니다.

4. 테이블 탭을 선택합니다.
5. 귀하에게 연결된 테이블의 경우, 테이블을 선택하세요.
6. 테이블 세부정보 페이지에서 아래로 스크롤하여 테이블 연결 세부 정보를 확인합니다.
7. 편집을 선택합니다.
8. 구성된 테이블 연결 편집 페이지에서 설명 또는 서비스 액세스 정보를 업데이트합니다.
9. 변경 사항 저장을 선택합니다.

구성된 테이블 분리

공동 작업 구성원은 구성된 테이블을 공동 작업에서 분리할 수 있습니다. 이 작업을 수행하면 쿼리할 수 있는 구성원이 테이블을 쿼리할 수 없습니다.

구성된 테이블의 연결을 해제하려면

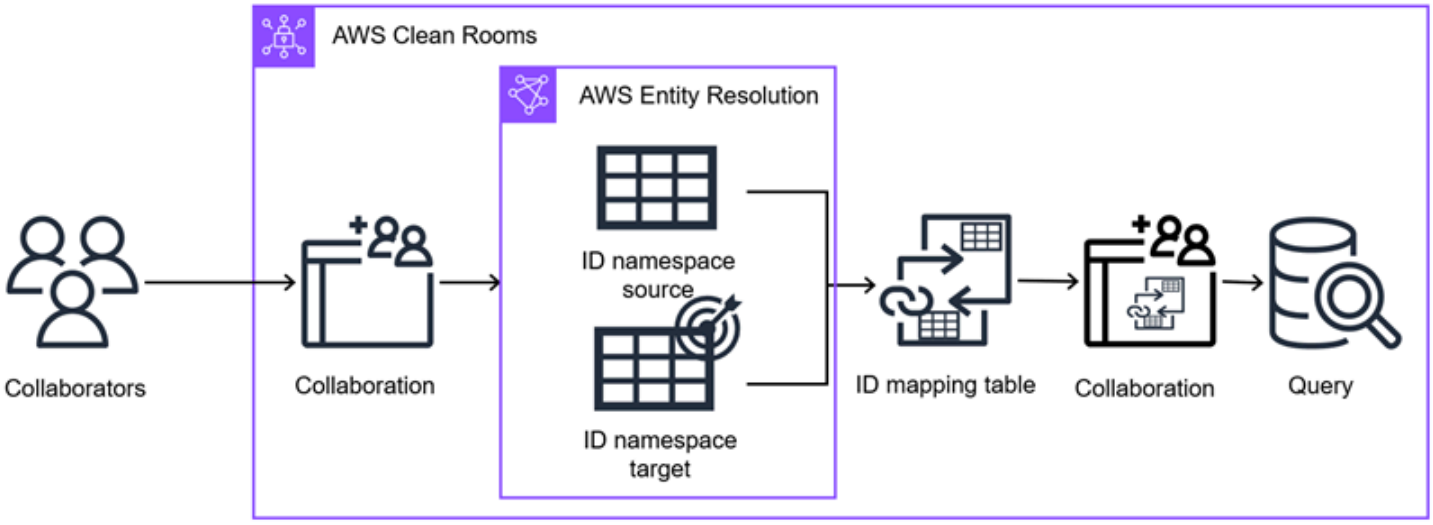
1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/cleanrooms AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업을 선택합니다.
4. 테이블 탭을 선택합니다.
5. 귀하에게 생성한 테이블에 대해 연결을 해제하려는 테이블 옆에 있는 옵션 버튼을 선택합니다.
6. 연결 해제를 선택합니다.
7. 대화 상자에서 연결 해제를 선택하여 구성된 테이블의 연결을 해제할지 여부를 확인하고 쿼리할 수 있는 구성원이 테이블을 쿼리하지 못하도록 합니다.

AWS Entity Resolution in AWS Clean Rooms

AWS Entity Resolution 를 입력 AWS Clean Rooms하면 소스의 데이터를 대상으로 변환하고, ID 매핑 테이블을 변환된 데이터로 채우고, 데이터를 쿼리할 수 있습니다.

먼저에서 공동 작업을 생성하고 초대 AWS 계정 할를 AWS Clean Rooms 추가하거나 멤버십을 생성하여 초대할 공동 작업에 참여합니다. 그런 다음 두 데이터 테이블에 대해 ID 매핑을 수행합니다. 이를 위해 기존 ID 네임스페이스 소스를 연결하거나 AWS Entity Resolution에서 소스를 새로 생성합니다. 공동 작업의 다른 구성원은 기존 ID 네임스페이스 대상을 연결하거나 ID 네임스페이스 대상을 새로 생성합니다. 그런 다음 연결된 두 ID 네임스페이스에서 ID 매핑 테이블을 생성하고 채웁니다. 마지막으로 쿼리 권한이 있는 구성원이 ID 매핑 테이블에 참여하여 두 데이터 테이블에 대한 쿼리를 실행합니다.

다음 다이어그램은 AWS Entity Resolution 에서 작업하는 방법을 요약합니다 AWS Clean Rooms.



> **Note**
>
> 현재 지원되는 트랜스코딩 서비스 공급자는 LiveRamp이며, AWS 리전미국 동부(버지니아 북부), 미국 동부(오하이오) 및 미국 서부(오레곤)에서 사용할 수 있습니다.

주제

- 의 ID 네임스페이스 AWS Clean Rooms
- 의 ID 매핑 테이블 AWS Clean Rooms

의 ID 네임스페이스 AWS Clean Rooms

ID 네임스페이스는 ID 테이블을 감싸는 래퍼로, 데이터세트와 ID 매핑 워크플로에서 이를 사용하는 방법을 설명하는 메타데이터를 제공할 수 있습니다. ID 매핑 워크플로는 지정된 ID 매핑 방법에 따라 입력 데이터 소스의 데이터를 입력 데이터 대상에 매핑하는 데이터 처리 작업입니다. ID 매핑 테이블을 생성합니다.

ID 네임스페이스는 소스 및 대상이라는 두 가지 유형으로 나뉩니다. 소스에는 ID 매핑 워크플로에서 처리될 소스 데이터에 대한 구성이 포함되어 있습니다. 대상에는 모든 소스가 확인될 대상 데이터의 구성이 포함되어 있습니다. 둘 다 확인하려는 입력 데이터를 정의하려면 ID 네임스페이스 소스와 ID 네임스페이스 대상을 AWS 계정생성하여 한 세트(소스)에서 다른 세트(대상)로 데이터를 변환합니다.

새 ID 네임스페이스를 생성하거나 기존 네임스페이스를 연결할 수 있습니다. 에서 ID 네임스페이스를 생성하는 방법에 대한 자세한 내용은 AWS Entity Resolution 사용 설명서의 ID 네임스페이스 생성을 AWS Entity Resolution참조하세요.

주제

- 새 ID 네임스페이스 생성 및 연결
- 기존 ID 네임스페이스 연결
- ID 네임스페이스 연결 편집
- ID 네임스페이스 연결 해제

새 ID 네임스페이스 생성 및 연결

공동 작업의 각 구성원은 ID 데이터를 쿼리할 ID 매핑 테이블을 생성하기 전에 ID 네임스페이스 소스 또는 ID 네임스페이스 대상을 생성하고 연결해야 합니다.

에서 ID 네임스페이스를 이미 생성한 경우 로 AWS Entity Resolution건너뜁니다기존 ID 네임스페이스 연결.

새 ID 네임스페이스를 생성하고 연결하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 수행하지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업을 선택합니다.
4. 엔터티 확인 탭에서 ID 네임스페이스 연결을 선택합니다.

5. ID 네임스페이스 연결 페이지의 엔터티 확인 데이터에서 ID 네임스페이스 생성을 선택합니다.

   AWS Entity Resolution 콘솔이 새 탭에 나타납니다.

6. AWS Entity Resolution 콘솔의 ID 네임스페이스 생성 페이지에 표시되는 메시지를 따릅니다.

   a. 세부 정보에 ID 네임스페이스 이름, 설명을 입력하고 ID 네임스페이스 유형(소스 또는 대상)을 선택합니다.

   b. ID 네임스페이스 방법에서 규칙 기반 매칭을 사용하려면 규칙 기반 방법을, 서드 파티 트랜스코딩을 사용하려면 공급자 서비스를 선택합니다.

   c. 선택한 ID 네임스페이스 방법에 따라 데이터 입력 유형을 지정합니다.

   d. ID 네임스페이스 생성을 선택합니다.

7. AWS Clean Rooms 콘솔로 돌아갑니다.

8. ID 네임스페이스 연결 페이지에서 엔터티 확인 데이터에 대해 드롭다운 목록에서 공동 작업과 연결할 AWS Entity Resolution ID 네임스페이스 소스 또는 대상을 선택합니다.

9. 연결 세부 정보에서 다음 단계를 수행합니다.

   a. 연결된 ID 네임스페이스의 이름을 입력합니다.

      기본 이름을 사용하거나 이 ID 네임스페이스의 이름을 바꿀 수 있습니다.

   b. (선택 사항) ID 네임스페이스에 대한 설명을 입력합니다.

      설명은 쿼리 작성에 도움이 됩니다.

10. 옵션을 선택한 다음 권장 조치를 수행하여 AWS Clean Rooms 액세스 권한을 지정합니다.

옵션	권장 조치
AWS Clean Rooms 가 권한 정책을 추가하고 관리하도록 허용	AWS Clean Rooms 는이 연결에 필요한 정책을 사용하여 서비스 역할을 생성합니다.
수동으로 권한 추가 및 관리	다음 중 하나를 수행합니다. • 리소스 정책을 검토하고 필요한 권한을 정책에 추가합니다. • 정책 설명 추가를 선택하여 기존 정책을 사용합니다.

옵션	권장 조치
	역할을 수정하고 정책을 생성할 권한이 있어야 합니다. **Note** 역할 정책을 수정할 수 없는 경우에서 서비스 역할에 대한 정책을 찾을 AWS Clean Rooms 수 없다는 오류 메시지가 표시됩니다.

11. (선택 사항) 고급 ID 매핑 테이블 구성에 대해 ID 네임스페이스에서 가져온 열의 기본 보호를 수정합니다.

    ID 매핑 테이블은 기본적으로 `sourceID` 열과 `targetID` 열 모두에서 `INNER JOIN`만 허용하도록 구성됩니다. 이 ID 네임스페이스(`sourceID` 또는 `targetID`)에서 가져온 열이 쿼리의 어느 곳에서든 허용될 수 있도록 이 구성을 수정할 수 있습니다.

목표	권장 옵션
열을 '조인 열'로 분류하고 `INNER JOIN` 절에서만 허용합니다.	예
열을 '차원 열'로 분류하고 쿼리의 `JOIN` 절, `SELECT`, `WHERE`, `GROUP BY` 문을 포함하여 쿼리의 모든 위치에서 허용합니다.	아니요. 쿼리의 모든 위치에서 허용합니다.

12. (선택 사항) ID 네임스페이스 리소스의 태그를 활성화하려면 새 태그 추가를 선택한 다음 키와 값 쌍을 입력합니다.
13. 연결을 선택합니다.
14. 엔티티 확인 탭의 연결된 ID 네임스페이스 테이블에서 연결된 ID 네임스페이스를 확인하고 ID 네임스페이스 유형이 올바른지 확인합니다(소스 또는 대상).

모든 공동 작업 구성원의 ID 네임스페이스가 연결되면 ID 매핑 테이블을 생성하고 데이터를 쿼리할 수 있습니다.

기존 ID 네임스페이스 연결

이 절차에서는 공동 작업에서 각 구성원의 기존 ID 네임스페이스 소스 또는 ID 네임스페이스 대상을 연결합니다.

기존 ID 네임스페이스를 연결하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 수행하지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업을 선택합니다.
4. 엔터티 확인 탭에서 ID 네임스페이스 연결을 선택합니다.
5. ID 네임스페이스 연결 페이지에서 엔터티 확인 데이터에 대해 드롭다운 목록에서 공동 작업과 연결할 AWS Entity Resolution ID 네임스페이스 소스 또는 대상을 선택합니다.
6. 연결 세부 정보에서 다음 단계를 수행합니다.

   a. 연결된 ID 네임스페이스의 이름을 입력합니다.

      기본 이름을 사용하거나 이 ID 네임스페이스의 이름을 바꿀 수 있습니다.

   b. (선택 사항) ID 네임스페이스에 대한 설명을 입력합니다.

      설명은 쿼리 작성에 도움이 됩니다.
7. 옵션을 선택한 다음 권장 조치를 수행하여 AWS Clean Rooms 액세스 권한을 지정합니다.

옵션	권장 조치
AWS Clean Rooms 가 권한 정책을 추가하고 관리하도록 허용	AWS Clean Rooms 는이 연결에 필요한 정책을 사용하여 서비스 역할을 생성합니다.
수동으로 권한 추가 및 관리	다음 중 하나를 수행합니다. • 리소스 정책을 검토하고 필요한 권한을 정책에 추가합니다. • 정책 설명 추가를 선택하여 기존 정책을 사용합니다.

옵션	권장 조치
	역할을 수정하고 정책을 생성할 권한이 있어야 합니다. **Note** 역할 정책을 수정할 수 없는 경우에서 서비스 역할에 대한 정책을 찾을 AWS Clean Rooms 수 없다는 오류 메시지가 표시됩니다.

8. (선택 사항) 고급 ID 매핑 테이블 구성에 대해 ID 네임스페이스에서 가져온 열의 기본 보호를 수정합니다.

   ID 매핑 테이블은 기본적으로 `sourceID` 열과 `targetID` 열 모두에서 `INNER JOIN`만 허용하도록 구성됩니다. 이 ID 네임스페이스(`sourceID` 또는 `targetID`)에서 가져온 열이 쿼리의 어느 곳에서든 허용될 수 있도록 이 구성을 수정할 수 있습니다.

목표	권장 옵션
열을 '조인 열'로 분류하고 `INNER JOIN` 절에서만 허용합니다.	예
열을 '차원 열'로 분류하고 쿼리의 `JOIN` 절, `SELECT`, `WHERE`, `GROUP BY` 문을 포함하여 쿼리의 모든 위치에서 허용합니다.	아니요. 쿼리의 모든 위치에서 허용합니다.

9. (선택 사항) ID 네임스페이스 리소스의 태그를 활성화하려면 새 태그 추가를 선택한 다음 키와 값 쌍을 입력합니다.
10. 연결을 선택합니다.
11. 엔티티 확인 탭의 연결된 ID 네임스페이스 테이블에서 연결된 ID 네임스페이스를 확인하고 ID 네임스페이스 유형이 올바른지 확인합니다(소스 또는 대상).

모든 공동 작업 구성원의 ID 네임스페이스가 연결되면 ID 매핑 테이블을 생성하고 데이터를 쿼리할 수 있습니다.

ID 네임스페이스 연결 편집

공동 작업 구성원은 자신이 생성한 ID 네임스페이스 연결을 편집할 수 있습니다.

ID 네임스페이스 연결을 편집하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 수행하지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업을 선택합니다.
4. 엔터티 확인 탭을 선택합니다.
5. 연결된 ID 네임스페이스에서 ID 네임스페이스를 선택합니다.
6. ID 네임스페이스 세부 정보 페이지에서 아래로 스크롤하여 ID 네임스페이스 연결 세부 정보를 확인합니다.
7. 편집을 선택합니다.
8. ID 네임스페이스 연결 편집 페이지에서 다음 중 하나를 편집합니다.

    a. 연결 세부 정보 에서 이름 또는 설명을 업데이트합니다.

    b. (선택 사항) 고급 ID 매핑 테이블 구성에 대해 ID 네임스페이스에서 가져온 열의 기본 보호를 수정합니다.

    ID 매핑 테이블은 기본적으로 `sourceID` 열과 `targetID` 열 모두에서 `INNER JOIN`만 허용하도록 구성됩니다. 이 ID 네임스페이스(`sourceID` 또는 `targetID`)에서 가져온 열이 쿼리의 어느 곳에서든 허용될 수 있도록 이 구성을 수정할 수 있습니다.

목표	권장 옵션
열을 '조인 열'로 분류하고 `INNER JOIN` 절에서만 허용합니다.	예
열을 '차원 열'로 분류하고 쿼리의 `JOIN` 절, `SELECT`, `WHERE`, `GROUP BY` 문을 포함하여 쿼리의 모든 위치에서 허용합니다.	아니요. 쿼리의 모든 위치에서 허용합니다.

9. 변경 사항 저장을 선택합니다.

ID 네임스페이스 연결 해제

공동 작업 구성원은 공동 작업에서 ID 네임스페이스의 연결을 해제할 수 있습니다. 이 작업을 수행하면 쿼리할 수 있는 구성원이 테이블을 쿼리할 수 없습니다.

> **Warning**
>
> 공동 작업에서 ID 네임스페이스 연결을 해제하면 파생된 ID 매핑 테이블에서 모든 데이터가 삭제되어 쿼리할 수 없게 됩니다.
> 예를 들어 ID 네임스페이스 연결이 세 개의 다른 ID 매핑 테이블에서 소스로 사용된 경우 ID 네임스페이스 연결을 해제하면 해당 ID 매핑 테이블의 모든 데이터가 삭제됩니다.

ID 네임스페이스 연결을 해제하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 수행하지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업을 선택합니다.
4. 엔터티 확인 탭을 선택합니다.
5. 연결된 ID 네임스페이스에서 연결을 해제하려는 ID 네임스페이스 옆의 옵션 버튼을 선택합니다.
6. 연결 해제를 선택합니다.
7. 대화 상자에서 연결 해제를 선택하여 ID 네임스페이스 연결 해제 결정을 확인합니다. 이 작업을 수행하면 쿼리할 수 있는 모든 구성원이 ID 매핑 테이블에 액세스할 수 없게 됩니다.

   공동 작업의 구성원이 ID 네임스페이스 중 하나를 제거할 경우 소스에서 공동 작업이 종료되었다면 ID 매핑 테이블을 다시 채울 수 없습니다.

   ID 매핑 테이블이 이전에 채워졌더라도 ID 네임스페이스의 연결을 해제하면 해당 테이블에서 더 이상 쿼리를 실행할 수 없습니다.

의 ID 매핑 테이블 AWS Clean Rooms

ID 매핑 테이블은 공동 작업에서 다자간 자격 증명 매핑을 활성화 AWS Clean Rooms 하는의 리소스입니다.

ID 매핑 테이블을 생성하려면 먼저 소스 데이터와 대상 데이터가 모두 ID 네임스페이스로 구성되어 있어야 합니다.

ID 매핑 테이블을 생성한 후에는 ID 매핑 워크플로를 사용하여 소스 ID 네임스페이스를 대상 ID 네임스페이스로 변환합니다. 이 작업은 규칙 기반 방법 또는 공급자 서비스 트랜스코딩 방법을 사용하여 수행할 수 있습니다.

ID 매핑 워크플로는 지정된 ID 매핑 워크플로 방법에 따라 입력 데이터 소스의 데이터를 입력 데이터 대상에 매핑하는 데이터 처리 작업입니다. 이 워크플로를 통해 ID 매핑 테이블이 채워집니다.

> **Note**
>
> ID 매핑 테이블은 Amazon S3에 저장되고 AWS Glue 테이블로 크롤링된 데이터 세트에서만 생성할 수 있습니다.

ID 매핑 워크플로 방법에는 규칙 기반 ID 매핑 또는 공급자 서비스 ID 매핑의 두 가지가 있습니다.

- 규칙 기반 ID 매핑 - 일치하는 규칙을 사용하여 퍼스트 파티 데이터를 소스에서 대상으로 변환합니다.
- 공급자 서비스 ID 매핑 - LiveRamp 공급자 서비스를 사용하여 서드 파티 데이터를 소스에서 대상으로 변환합니다.

> **Note**
>
> 현재 지원되는 트랜스코딩 서비스 공급자는 LiveRamp입니다. LiveRamp를 통해 구독하는 공동 작업의 모든 구성원은 ID 매핑 테이블을 생성할 AWS Data Exchange 수 있습니다. 이미 LiveRamp 구독이 있지만를 통해 구독하지 않은 경우 LiveRamp에 AWS Data Exchange 문의하여 비공개 제안을 받으세요. 자세한 내용은 AWS Entity Resolution 사용 설명서의 AWS Data Exchange에서 공급자 서비스 구독을 참조하세요.

주제

- 새 ID 매핑 테이블 생성 및 채우기
- 기존 ID 매핑 테이블 채우기
- ID 매핑 테이블 편집
- ID 매핑 테이블 삭제

새 ID 매핑 테이블 생성 및 채우기

ID 매핑 테이블을 생성하려면 먼저 연결된 ID 네임스페이스 소스와 대상이 있어야 합니다. 공동 작업에 연결할 ID 네임스페이스 소스 및 대상은 수행하려는 ID 매핑 유형(규칙 기반 ID 매핑 또는 공급자 서비스 ID 매핑)에 맞게 구성해야 합니다.

ID 매핑 테이블을 생성한 후에는 두 가지 옵션이 제공됩니다. 즉시 테이블을 채우고 ID 매핑 워크플로를 실행할 수 있습니다. 또는 나중에 테이블을 채울 때까지 기다릴 수도 있습니다.

ID 매핑 테이블이 성공적으로 채워지면 ID 매핑 테이블에 대해 다중 테이블 조인 쿼리를 실행하여 `sourceId`를 `targetId`와 조인하고 데이터를 분석할 수 있습니다.

주제

- ID 매핑 테이블 생성(규칙 기반)
- ID 매핑 테이블 생성(공급자 서비스)

ID 매핑 테이블 생성(규칙 기반)

이 주제는 일치하는 규칙을 사용하여 퍼스트 파티 데이터를 소스에서 대상으로 변환하는 ID 매핑 테이블 생성 프로세스에 대해 설명합니다.

규칙 기반 방법을 사용하여 새 ID 매핑 테이블을 생성하고 채우려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업을 선택합니다.
4. 엔터티 확인 탭에서 ID 매핑 테이블 생성을 선택합니다.
5. ID 매핑 테이블 생성 페이지의 ID 매핑 설정 아래에서 목표에 따라 다음 조치 중 하나를 수행합니다.

목표	권장 조치
새 ID 매핑 워크플로 생성	1. 새 ID 매핑 워크플로 생성 확인란을 선택한 상태로 둡니다. 2. 6단계로 이동합니다.

목표	권장 조치
기존 ID 매핑 워크플로 재사용	1. 새 ID 매핑 워크플로 생성 확인란의 선택을 취소합니다. 2. 드롭다운 목록에서 규칙 기반 ID 매핑 워크플로를 선택합니다. 3. 9단계로 건너뜁니다.

6. ID 데이터에서 시나리오에 따라 다음 조치 중 하나를 수행합니다.

시나리오	권장 조치
공동 작업에 ID 네임스페이스 소스와 ID 네임스페이스 대상이 하나만 있습니다.	소스 및 대상 ID 네임스페이스 연결을 확인합니다.
공동 작업에 여러 ID 네임스페이스 연결이 있습니다.	드롭다운 목록에서 사용하려는 소스 및 대상 ID 네임스페이스 연결을 선택합니다.

7. 방법에서 선택한 ID 매핑 워크플로 방법이 규칙 기반으로 표시되는지 확인합니다.
8. 규칙 파라미터에서 규칙 제어, 비교 유형, 레코드 매칭 구성을 지정합니다.

   a. 규칙 제어에서 매칭 규칙을 대상 또는 소스 ID 네임스페이스에서 제공할지 여부를 선택합니다.

      규칙 표시를 켜면 규칙을 볼 수 있습니다.

      규칙 제어는 소스 및 대상 ID 네임스페이스 간에 호환되어야 ID 매핑 워크플로에서 사용할 수 있습니다. 예를 들어 소스 ID 네임스페이스는 규칙을 대상으로 제한하지만 대상 ID 네임스페이스는 규칙을 소스로 제한하는 경우 오류가 발생합니다.

   b. 비교 유형은 다중 입력 필드로 자동 설정됩니다.

      두 참가자 모두 이전에 이 옵션을 선택했기 때문입니다.

   c. 다음 옵션 중 하나를 선택하여 레코드 매칭 유형을 지정합니다.

목표	권장 옵션
ID 매핑 워크플로를 생성할 때 대상의 각 매칭 레코드당 소스의 매칭 레코드 하나	하나의 소스를 하나의 대상에

목표	권장 옵션
씩만 저장하도록 레코드 매칭 유형을 제한합니다.	
ID 매핑 워크플로를 생성할 때 대상의 각 매칭 레코드당 소스의 모든 매칭 레코드를 저장하도록 레코드 매칭 유형을 제한합니다.	여러 소스를 하나의 대상에

> **Note**
>
> 소스 및 대상 ID 네임스페이스에 대해 지정된 제한은 호환되어야 합니다.

9. ID 매핑 세부 정보에서 다음 작업을 수행합니다.

   a. ID 매핑 테이블 이름을 입력합니다.

      기본 이름을 사용하거나 이 ID 매핑 테이블의 이름을 바꿀 수 있습니다.

   b. (선택 사항) ID 매핑 테이블에 대한 설명을 입력합니다.

      설명은 쿼리 작성에 도움이 됩니다.

10. 옵션을 선택한 다음 권장 조치를 수행하여 AWS Clean Rooms 액세스 권한을 지정합니다.

옵션	권장 조치
AWS Clean Rooms 가 권한 정책을 추가하고 관리하도록 허용	AWS Clean Rooms 는이 연결에 필요한 정책을 사용하여 서비스 역할을 생성합니다.
수동으로 권한 추가 및 관리	다음 중 하나를 수행합니다. • 리소스 정책을 검토하고 필요한 권한을 정책에 추가합니다. • 정책 설명 추가를 선택하여 기존 정책을 사용합니다.

옵션	권장 조치
	역할을 수정하고 정책을 생성할 권한이 있어야 합니다. **Note** 역할 정책을 수정할 수 없는 경우에서 서비스 역할에 대한 정책을 찾을 AWS Clean Rooms 수 없다는 오류 메시지가 표시됩니다.

11. 옵션을 선택한 다음 권장 조치를 수행하여 AWS Entity Resolution 액세스 권한을 지정합니다.

    이 단원은 새 ID 매핑 테이블을 생성하는 경우에만 표시됩니다.

옵션	권장 조치
새 서비스 역할 생성 및 사용	AWS Clean Rooms 는이 테이블에 필요한 정책을 사용하여 서비스 역할을 생성합니다. 기본 서비스 역할 이름은 `entityresolution-id-mapping-workflow-<timestamp>` 입니다 역할을 생성하고 정책을 연결할 수 있는 권한이 있어야 합니다. 입력 데이터가 암호화된 경우 이 데이터는 KMS 키로 암호화됨을 선택한 다음 데이터 입력을 복호화하는 데 사용할 AWS KMS 키를 입력할 수 있습니다.
기존 서비스 역할 사용	1. 드롭다운 목록에서 기존 서비스 역할 이름을 선택합니다.

옵션	권장 조치
	역할을 나열할 권한이 있는 경우 역할 목록이 표시됩니다. 역할을 나열할 수 있는 권한이 없는 경우 사용하려는 역할의 Amazon 리소스 이름(ARN)을 입력할 수 있습니다. 2. IAM에서 보기 외부 링크를 선택하여 서비스 역할을 확인합니다. 기존 서비스 역할이 없는 경우 기존 서비스 역할 사용 옵션을 사용할 수 없습니다. 기본적으로 AWS Clean Rooms 는 필요한 권한을 추가하기 위해 기존 역할 정책을 업데이트하려고 시도하지 않습니다. 3. (선택 사항) 이 역할에 필요한 권한이 있는 사전 구성된 정책 추가 확인란을 선택하여 역할에 필요한 권한을 연결합니다. 역할을 수정하고 정책을 생성할 권한이 있어야 합니다.

12. (선택 사항) 다음 중 하나를 선택하여 추가 설정을 지정합니다.

    a. ID 매핑 테이블의 경우 목표에 따라 다음 작업 중 하나를 수행합니다.

목표	권장 조치
ID 매핑 테이블에 대한 사용자 지정 암호화 설정 활성화	암호화 설정 사용자 지정을 선택한 다음 AWS KMS 키를 입력합니다. **Note** 이 KMS 키는 KMS 키 정책을 `cleanrooms.amazonaws.com` 사용하여 내에서를 사용하는 AWS Entity Resolution 데 필요한 권한

목표	권장 조치
	을 부여해야 합니다. ID 매핑 워크플로로 암호화 작업을 하는 데 필요한 권한에 대한 자세한 내용은 AWS Entity Resolution 사용 설명서의 AWS Entity Resolution용 워크플로 작업 역할 생성을 참조하세요.
ID 매핑 테이블 리소스에 대한 태그 활성화	새 태그 추가를 선택한 다음 키와 값 쌍을 입력합니다.

b. ID 매핑 워크플로의 경우 목표에 따라 다음 작업 중 하나를 수행합니다.

이 단원은 새 ID 매핑 테이블을 생성하는 경우에만 표시됩니다.

목표	권장 조치
ID 매핑 워크플로의 이름 및 설명 수정	동일한 ID 매핑 테이블 이름 및 설명 유지 확인란의 선택을 취소하고 새 ID 매핑 워크플로 이름 및 설명을 입력합니다.
ID 매핑 워크플로 리소스에 대한 태그 활성화	새 태그 추가를 선택한 다음 키와 값 쌍을 입력합니다.

13. 목표에 따라 다음 옵션 중 하나를 선택합니다.

목표	권장 옵션
빈 ID 매핑 테이블을 만들되 ID 매핑 워크플로는 실행하지 않음	ID 매핑 테이블 생성 기존 ID 매핑 테이블 채우기 프로세스에 따라 나중에 ID 매핑 테이블을 채울 수 있습니다.
ID 매핑 테이블 생성 및 ID 매핑 워크플로 실행	ID 매핑 테이블 생성 및 채우기 ID 매핑 워크플로 프로세스가 시작됩니다. 이 프로세스 중에 ID 매핑 테이블은 변환된 ID로

목표	권장 옵션
	채워집니다. ID 매핑 워크플로를 처리하는 데 몇 시간이 걸릴 수 있습니다. ID 매핑 테이블이 성공적으로 채워지면 ID 매핑 테이블을 쿼리하여 `sourceId`를 `targetId`와 조인하고 데이터를 분석할 수 있습니다.

ID 매핑 테이블 생성(공급자 서비스)

이 주제는 공급자 서비스(LiveRamp)를 사용하는 ID 매핑 테이블을 생성하는 프로세스에 대해 설명합니다. LiveRamp 공급자 서비스는 유지 관리되거나 파생된 RampID를 사용하여 일련의 소스 RampID를 다른 RampID로 변환합니다.

공급자 서비스 방법을 사용하여 새 ID 매핑 테이블을 생성하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업을 선택합니다.
4. 엔터티 확인 탭에서 ID 매핑 테이블 생성을 선택합니다.
5. ID 매핑 테이블 생성 페이지의 ID 매핑 설정 아래에서 목표에 따라 다음 조치 중 하나를 수행합니다.

목표	권장 조치
새 ID 매핑 워크플로 생성	1. 새 ID 매핑 워크플로 생성 확인란을 선택한 상태로 둡니다. 2. 6단계로 이동합니다.
기존 ID 매핑 워크플로 재사용	1. 새 ID 매핑 워크플로 생성 확인란의 선택을 취소합니다. 2. 드롭다운 목록에서 규칙 기반 ID 매핑 워크플로를 선택합니다.

목표	권장 조치
	3. 9단계로 건너뜁니다.

6. ID 데이터에서 시나리오에 따라 다음 조치 중 하나를 수행합니다.

시나리오	권장 조치
공동 작업에 ID 네임스페이스 소스와 ID 네임스페이스 대상이 하나만 있습니다.	소스 및 대상 ID 네임스페이스 연결을 확인합니다.
공동 작업에 여러 ID 네임스페이스 연결이 있습니다.	드롭다운 목록에서 사용하려는 소스 및 대상 ID 네임스페이스 연결을 선택합니다.

7. 방법에서 선택한 ID 매핑 워크플로 방법이 LiveRamp 트랜스코딩인지 확인합니다.
8. LiveRamp 구성에 LiveRamp에서 제공하는 다음 정보를 입력합니다.

   - LiveRamp ID 관리자 ARN
   - LiveRamp 암호 관리자 ARN

   또는 기존 워크플로에서 가져오기를 선택할 수 있습니다.
9. ID 매핑 세부 정보에서 다음 단계를 수행합니다.

   a. ID 매핑 테이블 이름을 입력합니다.

      기본 이름을 사용하거나 이 ID 매핑 테이블의 이름을 바꿀 수 있습니다.

   b. (선택 사항) ID 매핑 테이블에 대한 설명을 입력합니다.

      설명은 쿼리 작성에 도움이 됩니다.
10. 다음 중 하나를 선택하여 AWS Clean Rooms 액세스 권한을 지정합니다.

옵션	권장 조치
AWS Clean Rooms 가 권한 정책을 추가하고 관리하도록 허용	AWS Clean Rooms 는이 연결에 필요한 정책을 사용하여 서비스 역할을 생성합니다.
수동으로 권한 추가 및 관리	다음 중 하나를 수행합니다.

옵션	권장 조치
	• 리소스 정책을 검토하고 필요한 권한을 정책에 추가합니다. • 정책 설명 추가를 선택하여 기존 정책을 사용합니다. 역할을 수정하고 정책을 생성할 권한이 있어야 합니다. **Note** 역할 정책을 수정할 수 없는 경우에서 서비스 역할에 대한 정책을 찾을 AWS Clean Rooms 수 없다는 오류 메시지가 표시됩니다.

11. 옵션을 선택한 다음 권장 조치를 수행하여 AWS Entity Resolution 액세스 권한을 지정합니다.

    이 단원은 새 ID 매핑 테이블을 생성하는 경우에만 표시됩니다.

옵션	권장 조치
새 서비스 역할 생성 및 사용	AWS Clean Rooms 는이 테이블에 필요한 정책을 사용하여 서비스 역할을 생성합니다. 기본 서비스 역할 이름은 `entityresolution-id-mapping-workflow-<timestamp>` 입니다 역할을 생성하고 정책을 연결할 수 있는 권한이 있어야 합니다. 입력 데이터가 암호화된 경우 이 데이터는 KMS 키로 암호화됨 옵션을 선택한 다음 데이터 입력을 복호화하는 데 사용할 AWS KMS 키를 입력할 수 있습니다.

옵션	권장 조치
기존 서비스 역할 사용	1. 드롭다운 목록에서 기존 서비스 역할 이름을 선택합니다. 역할을 나열할 권한이 있는 경우 역할 목록이 표시됩니다. 역할을 나열할 수 있는 권한이 없는 경우 사용하려는 역할의 Amazon 리소스 이름(ARN)을 입력할 수 있습니다. 2. IAM에서 보기를 선택하여 서비스 역할을 확인합니다. 기존 서비스 역할이 없는 경우 기존 서비스 역할 사용 옵션을 사용할 수 없습니다. 기본적으로 AWS Clean Rooms 는 필요한 권한을 추가하기 위해 기존 역할 정책을 업데이트하려고 시도하지 않습니다. 3. (선택 사항) 이 역할에 필요한 권한이 있는 사전 구성된 정책 추가 확인란을 선택하여 역할에 필요한 권한을 추가합니다. 역할을 수정하고 정책을 생성할 권한이 있어야 합니다.

12. (선택 사항) 다음 중 하나를 선택하여 추가 설정을 지정합니다.

    a. ID 매핑 테이블의 경우 목표에 따라 다음 작업 중 하나를 수행합니다.

목표	권장 조치
ID 매핑 테이블에 대한 사용자 지정 암호화 설정 활성화	암호화 설정 사용자 지정을 선택한 다음 AWS KMS 키를 입력합니다. **Note** 이 KMS 키는 KMS 키 정책을 `cleanrooms.amazonaws.com`

목표	권장 조치
	사용하여 내에서를 사용하는 AWS Entity Resolution 데 필요한 권한을 부여해야 합니다. ID 매핑 워크플로로 암호화 작업을 하는 데 필요한 권한에 대한 자세한 내용은 AWS Entity Resolution 사용 설명서의 AWS Entity Resolution용 워크플로 작업 역할 생성을 참조하세요.
ID 매핑 테이블 리소스에 대한 태그 활성화	새 태그 추가를 선택한 다음 키와 값 쌍을 입력합니다.

b. ID 매핑 워크플로의 경우 목표에 따라 다음 작업 중 하나를 수행합니다.

이 단원은 새 ID 매핑 테이블을 생성하는 경우에만 표시됩니다.

목표	권장 조치
ID 매핑 워크플로의 이름 및 설명 수정	동일한 ID 매핑 테이블 이름 및 설명 유지 확인란의 선택을 취소하고 새 ID 매핑 워크플로 이름 및 설명을 입력합니다.
ID 매핑 워크플로 리소스에 대한 태그 활성화	새 태그 추가를 선택한 다음 키와 값 쌍을 입력합니다.

13. 목표에 따라 다음 조치 중 하나를 선택합니다.

목표	권장 조치
빈 ID 매핑 테이블을 만들되 ID 매핑 워크플로는 실행하지 않음	ID 매핑 테이블 생성을 선택합니다. 기존 ID 매핑 테이블 채우기 프로세스에 따라 나중에 ID 매핑 테이블을 채울 수 있습니다.
ID 매핑 테이블 생성 및 ID 매핑 워크플로 실행	ID 매핑 테이블 생성 및 채우기를 선택합니다.

목표	권장 조치
	ID 매핑 워크플로 프로세스가 시작됩니다. 이 프로세스 중에 ID 매핑 테이블은 트랜스코딩된 ID로 채워집니다. ID 매핑 워크플로를 처리하는 데 몇 시간이 걸릴 수 있습니다. ID 매핑 테이블이 성공적으로 채워지면 ID 매핑 테이블을 쿼리하여 `sourceId`를 `targetId`와 조인하고 데이터를 분석할 수 있습니다.

기존 ID 매핑 테이블 채우기

새 데이터가 ID 네임스페이스에 추가되면 이 워크플로를 사용합니다.

기존 ID 매핑 테이블을 채우려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 수행하지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업을 선택합니다.
4. 엔터티 확인 탭의 ID 매핑 테이블 섹션에서 다음 중 하나를 수행합니다.
   - ID 매핑 테이블을 선택한 다음 채우기를 선택합니다.
   - ID 매핑 테이블 옆의 옵션 버튼을 선택하고 ID 매핑 테이블 세부 정보 페이지에서 채우기를 선택합니다.

ID 매핑 워크플로 프로세스가 시작됩니다. 이 프로세스 중에 ID 매핑 테이블은 트랜스코딩된 ID로 채워집니다. ID 매핑 워크플로를 처리하는 데 몇 시간이 걸릴 수 있습니다.

ID 매핑 테이블이 성공적으로 채워지면 ID 매핑 테이블을 쿼리하여 `sourceId`를 `targetId`와 조인할 수 있습니다.

ID 매핑 테이블 편집

공동 작업 구성원은 자신이 생성한 ID 매핑 테이블을 편집할 수 있습니다.

ID 매핑 테이블을 편집하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 수행하지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업을 선택합니다.
4. 엔터티 확인 탭을 선택합니다.
5. ID 매핑 테이블에서 테이블을 선택합니다.
6. ID 매핑 테이블 세부 정보 페이지에서 아래로 스크롤하여 ID 매핑 테이블 세부 정보를 확인합니다.
7. 편집을 선택합니다.
8. ID 매핑 테이블 편집 페이지에서 설명 또는 서비스 액세스 정보를 업데이트합니다.
9. Save changes(변경 사항 저장)를 선택합니다.

ID 매핑 테이블 삭제

공동 작업 구성원은 자신이 생성한 ID 매핑 테이블을 삭제할 수 있습니다. 이 작업을 수행하면 쿼리할 수 있는 구성원이 테이블을 쿼리할 수 없습니다.

> **Warning**
> 매핑 테이블을 삭제하면 채워진 데이터가 영구적으로 제거됩니다.

ID 매핑 테이블을 삭제하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 수행하지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업을 선택합니다.
4. 엔터티 확인 탭을 선택합니다.
5. ID 매핑 테이블에서 테이블을 선택합니다.
6. ID 매핑 테이블 세부 정보 페이지에서 아래로 스크롤하여 ID 매핑 테이블을 확인합니다.
7. ID 매핑 테이블을 선택한 다음 삭제를 선택합니다.
8. ID 매핑 테이블을 삭제하려면 삭제를 선택합니다.

의 분석 템플릿 AWS Clean Rooms

분석 템플릿은 의 사용자 지정 분석 규칙 AWS Clean Rooms에서 사용할 수 있습니다. 분석 템플릿을 사용하면 동일한 쿼리를 재사용하는 데 도움이 되는 파라미터를 정의할 수 있습니다.는 리터럴 값으로 파라미터화의 하위 집합을 AWS Clean Rooms 지원합니다.

분석 템플릿은 공동 작업별로 다릅니다. 각 협업에 대해 구성원은 해당 공동 작업의 쿼리만 볼 수 있습니다. 공동 작업에서 차등 프라이버시를 사용하려는 경우 분석 템플릿이 AWS Clean Rooms 차등 프라이버시의 범용 쿼리 구조와 호환되는지 확인해야 합니다.

SQL 코드를 사용하거나 Spark용 Python 코드를 사용하는 두 가지 방법으로 분석 템플릿을 생성할 수 있습니다.

- SQL 분석 템플릿은 Spark 분석 엔진과 AWS Clean Rooms SQL 분석 엔진을 모두 사용하는 공동 작업에서 사용할 수 있습니다.
- PySpark 분석 템플릿은 Spark 분석 엔진을 사용하는 공동 작업에서 사용할 수 있습니다.

주제

- SQL 분석 템플릿
- PySpark 분석 템플릿
- PySpark 분석 템플릿 문제 해결

SQL 분석 템플릿

SQL 분석 템플릿을 사용하면 공동 작업 내에서 다양한 데이터 세트의 데이터를 쿼리하고 분석할 수 있습니다. 이러한 템플릿을 사용하여 대상 중복을 식별하고 집계된 지표를 계산하는 등 다양한 유형의 분석을 수행할 수 있습니다.

SQL 분석 템플릿을 사용하면 다음을 수행할 수 있습니다.

- 표준 SQL 쿼리 쓰기
- 파라미터를 추가하여 쿼리를 동적으로 만들기
- 특정 열 및 테이블에 대한 액세스 제어
- 민감한 데이터에 대한 집계 요구 사항 설정

주제

- SQL 분석 템플릿 생성
- SQL 분석 템플릿 검토

SQL 분석 템플릿 생성

사전 조건

SQL 분석 템플릿을 생성하기 전에 다음이 있어야 합니다.

- 활성 AWS Clean Rooms 공동 작업
- 공동 작업에서 하나 이상의 구성된 테이블에 대한 액세스

  에서 테이블을 구성하는 방법에 대한 자세한 내용은 단원을 AWS Clean Rooms참조하십시오AWS Clean Rooms에서 구성된 테이블 생성하기.
- 분석 템플릿을 생성할 수 있는 권한
- SQL 쿼리 구문에 대한 기본 지식

다음 절차에서는 AWS Clean Rooms 콘솔을 사용하여 SQL 분석 템플릿을 생성하는 프로세스를 설명합니다.

AWS SDKs. AWS Clean Rooms

SQL 분석 템플릿을 생성하려면

1. 에 로그인 AWS Management Console 하고 공동 작업 생성자 역할을 AWS 계정 할를 사용하여 AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업을 선택합니다.
4. 템플릿 탭에서 직접 생성한 분석 템플릿 섹션으로 이동합니다.
5. 분석 템플릿 생성을 선택합니다.
6. 분석 템플릿 생성 페이지의 세부 정보에서

   a. 분석 템플릿의 이름을 입력합니다.

   b. (선택 사항) 설명을 입력합니다.

   c. 형식에서 SQL 옵션을 선택한 상태로 둡니다.

7. 테이블의 경우 공동 작업과 관련된 구성된 테이블을 확인하세요.
8. 정의의 경우,

   a. 분석 템플릿의 정의를 입력합니다.
   b. 정의를 가져오려면 가져오기 대상을 선택합니다.
   c. (선택 사항) SQL 편집기에서 매개 변수 이름 앞에 콜론(:)을 입력하여 매개 변수를 지정합니다.

      예시:

      ```
      WHERE table1.date + :date_period > table1.date
      ```
9. 이전에 매개 변수를 추가한 경우 매개 변수 — 선택 사항에서 각 매개 변수 이름에 대해 유형 및 기본값(선택 사항)을 선택합니다.
10. 구성된 테이블 리소스에 대해 태그를 활성화하려면 새 태그 추가를 선택한 다음 키와 값 쌍을 입력합니다.
11. 생성(Create)을 선택합니다.
12. 이제 공동 작업 구성원에게 분석 템플릿을 검토할 수 있음을 알릴 준비가 되었습니다. (자체 데이터를 쿼리하려는 경우 선택 사항)

SQL 분석 템플릿 검토

공동 작업 구성원이 SQLanalysis 템플릿을 생성한 후 이를 검토하고 승인할 수 있습니다. 분석 템플릿이 승인되고 나면의 쿼리에 사용할 수 있습니다 AWS Clean Rooms.

> **Note**
>
> 분석 코드를 공동 작업으로 가져올 때 다음 사항에 유의하세요.
>
> - AWS Clean Rooms 는 분석 코드의 동작을 검증하거나 보장하지 않습니다.
>   - 특정 동작을 확인해야 하는 경우 공동 작업 파트너의 코드를 직접 검토하거나 신뢰할 수 있는 타사 감사자와 협력하여 검토합니다.
> - 공유 보안 모델에서:
>   - 사용자(고객)는 환경에서 실행되는 코드의 보안에 대한 책임이 있습니다.
>   - AWS Clean Rooms 는 환경의 보안을 담당하여 다음을 보장합니다.
>   - 승인된 코드만 실행됩니다.
>   - 지정된 구성된 테이블만 액세스할 수 있습니다.

- 유일한 출력 대상은 결과 수신자의 S3 버킷입니다.

AWS Clean Rooms 콘솔을 사용하여 SQL 분석 템플릿을 검토하려면

1. 에 로그인 AWS Management Console 하고 공동 작업 생성자로 AWS 계정 작동하는를 사용하여 AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업을 선택합니다.
4. 템플릿 탭에서 다른 구성원이 만든 분석 템플릿 섹션으로 이동합니다.
5. 실행 가능 상태가 아니요 - 직접 검토가 필요함인 분석 템플릿을 선택합니다.
6. 검토를 선택합니다.
7. 분석 규칙 개요, 정의 및 매개변수(있는 경우)를 검토합니다.
8. 정의에 참조된 테이블 아래에 나열된 구성된 테이블을 검토합니다.

   각 테이블 옆의 상태는 템플릿 허용 불가로 표시됩니다.
9. 테이블을 선택합니다.

만약	그런 다음을 선택합니다...
분석 템플릿을 승인합니다	테이블에 템플릿을 허용하세요. 허용을 선택하여 승인을 확인합니다.
분석 템플릿을 승인하지 마세요	허용 안 함

이제 SQL 분석 템플릿을 사용하여 구성된 테이블을 쿼리할 준비가 되었습니다. 자세한 내용은 SQL 쿼리 실행 단원을 참조하십시오.

PySpark 분석 템플릿

사용자 지정 및 오픈 소스 라이브러리를 사용하려면 PySpark 분석 템플릿에 Python 사용자 스크립트와 선택적 가상 환경이 필요합니다. 이러한 파일을 아티팩트라고 합니다.

분석 템플릿을 생성하기 전에 먼저 아티팩트를 생성한 다음 Amazon S3 버킷에 아티팩트를 저장합니다.는 분석 작업을 실행할 때 이러한 아티팩트를 AWS Clean Rooms 사용합니다.는 작업을 실행할 때 아티팩트 AWS Clean Rooms 에만 액세스합니다.

PySpark 분석 템플릿에서 코드를 실행하기 전에는 다음을 통해 아티팩트를 AWS Clean Rooms 검증합니다.

- 템플릿을 생성할 때 사용되는 특정 S3 객체 버전 확인
- 아티팩트의 SHA-256 해시 확인
- 아티팩트가 수정 또는 제거된 작업 실패

> **Note**
>
> 에서 지정된 PySpark 분석 템플릿에 대한 모든 결합된 아티팩트의 최대 크기는 1GB AWS Clean Rooms 입니다.

PySpark용 보안 분석 템플릿

보안 컴퓨팅 환경을 유지하기 위해는 2계층 컴퓨팅 아키텍처를 AWS Clean Rooms 사용하여 시스템 작업에서 사용자 코드를 격리합니다. 이 아키텍처는 Membrane이라고도 하는 Amazon EMR Serverless Fine Grained Access Control 기술을 기반으로 합니다. 자세한 내용은 Membrane – 필수 코드가 있는 경우 Apache Spark의 안전하고 성능이 뛰어난 데이터 액세스 제어를 참조하세요.

컴퓨팅 환경 구성 요소는 별도의 사용자 공간과 시스템 공간으로 나뉩니다. 사용자 공간은 PySpark 분석 템플릿에서 PySpark 코드를 실행합니다.는 시스템 공간을 AWS Clean Rooms 사용하여 고객이 제공한 서비스 역할을 사용하여 데이터를 읽고 작업을 실행하고 열 허용 목록을 구현하는 등 작업을 실행할 수 있습니다. 이 아키텍처의 결과로 적은 수의 Spark SQL 및 PySpark DataFrames APIs를 포함할 수 있는 시스템 공간에 영향을 미치는 고객의 PySpark 코드가 차단됩니다.

의 PySpark 제한 사항 AWS Clean Rooms

고객이 승인된 PySpark 분석 템플릿을 제출하면는 고객이 액세스할 수 없는 자체 보안 컴퓨팅 환경에서 템플릿을 AWS Clean Rooms 실행합니다. 컴퓨팅 환경은 사용자 공간과 시스템 공간이 있는 컴퓨팅 아키텍처를 구현하여 안전한 컴퓨팅 환경을 유지합니다. 자세한 내용은 PySpark용 보안 분석 템플릿 단원을 참조하십시오.

PySpark를 사용하기 전에 다음 제한 사항을 고려하세요 AWS Clean Rooms.

제한 사항

- DataFrame 출력만 지원됩니다.

- 작업 실행당 단일 Spark 세션

지원되지 않는 기능

- 데이터 관리
  - Iceberg 테이블 형식
  - LakeFormation 관리형 테이블
  - 복원력 있는 분산 데이터세트(RDD)
  - Spark 스트리밍
  - 중첩된 열에 대한 액세스 제어
- 사용자 지정 함수 및 확장
  - 사용자 정의 테이블 함수(UDTFs)
  - HiveUDFs
  - 사용자 정의 함수의 사용자 지정 클래스
  - 사용자 지정 데이터 소스
  - 다음을 위한 추가 JAR 파일:
    - Spark 확장
    - 커넥터
    - 메타스토어 구성
- 모니터링 및 분석
  - Spark 로깅
  - Spark UI
  - `ANALYZE TABLE` 명령

Important

이러한 제한은 사용자와 시스템 공간 간의 보안 격리를 유지하기 위해 적용됩니다.
공동 작업 구성에 관계없이 모든 제한이 적용됩니다.
향후 업데이트는 보안 평가에 따라 추가 기능에 대한 지원을 추가할 수 있습니다.

모범 사례

PySpark 분석 템플릿을 생성할 때 다음 모범 사례를 따르는 것이 좋습니다.

- 를 의 PySpark 제한 사항 AWS Clean Rooms 염두에 두고 분석 템플릿을 설계합니다.
- 먼저 개발 환경에서 코드를 테스트합니다.
- 지원되는 DataFrame 작업만 사용합니다.
- DataFrame 제한 사항에 맞게 출력 구조를 계획합니다.

아티팩트 관리를 위한 다음 모범 사례를 권장합니다.

- 모든 PySpark 분석 템플릿 아티팩트를 전용 S3 버킷 또는 접두사에 보관합니다.
- 다양한 아티팩트 버전에 대해 명확한 버전 이름 지정을 사용합니다.
- 아티팩트 업데이트가 필요한 경우 새 분석 템플릿을 생성합니다.
- 어떤 템플릿이 어떤 아티팩트 버전을 사용하는지 인벤토리를 유지 관리합니다.

Spark 코드 작성 방법에 대한 자세한 내용은 다음을 참조하세요.

- Apache Spark 예제
- Amazon EMR 릴리스 가이드의 Spark 애플리케이션 작성
- 자습서: 사용 설명서의 AWS Glue for Spark 스크립트 작성 AWS Glue

다음 주제에서는 분석 템플릿을 생성하고 검토하기 전에 Python 사용자 스크립트 및 라이브러리를 생성하는 방법을 설명합니다.

주제

- 사용자 스크립트 생성
- 가상 환경 생성(선택 사항)
- S3에 사용자 스크립트 및 가상 환경 저장
- PySpark 분석 템플릿 생성
- PySpark 분석 템플릿 검토

사용자 스크립트 생성

사용자 스크립트에는 진입점 함수(즉, 핸들러)가 포함되어야 합니다. 사용자 스크립트 파일의 이름을 유효한 Python 파일 이름으로 지정할 수 있습니다.

다음 절차에서는 사용자 스크립트를 생성하여 PySpark 분석의 핵심 기능을 정의하는 방법을 설명합니다.

사전 조건

- PySpark 1.0(Python 3.9 및 Python 3.11 및 Spark 3.5.2에 해당)
- Amazon S3의 데이터세트는 정의한 Spark 세션에서 구성된 테이블 연결로만 읽을 수 있습니다.
- 코드가 Amazon S3 및를 직접 호출할 수 없습니다. AWS Glue
- 코드가 네트워크 호출을 수행할 수 없음

사용자 스크립트를 생성하려면

1. 원하는 텍스트 편집기 또는 통합 개발 환경(IDE)을 엽니다.

   Python 파일을 지원하는 모든 텍스트 편집기 또는 IDE(예: Visual Studio Code, PyCharm 또는 Notepad++)를 사용할 수 있습니다.
2. 선택한 이름으로 새 Python 파일을 생성합니다(예: **`my_analysis.py`**).
3. 컨텍스트 객체 파라미터를 수락하는 진입점 함수를 정의합니다.

   ```
   def entrypoint(context)
   ```

   `context` 객체 파라미터는 필수 Spark 구성 요소 및 참조 테이블에 대한 액세스를 제공하는 사전입니다. 여기에는 Spark 작업 및 참조된 테이블을 실행하기 위한 Spark 세션 액세스가 포함됩니다.

   Spark 세션 액세스는를 통해 사용할 수 있습니다. `context['sparkSession']`

   참조된 테이블은를 통해 사용할 수 있습니다. `context['referencedTables']`
4. 진입점 함수의 결과를 정의합니다.

   ```
   return results
   ```

   는 파일 이름의 결과 사전이 포함된 객체를 출력 DataFrame에 반환해야 `results` 합니다.

> **Note**
>
> AWS Clean Rooms 는 결과 수신기의 S3 버킷에 DataFrame 객체를 자동으로 기록합니다.

5. 이제 다음에 대한 준비가 되었습니다.

   a. 이 사용자 스크립트를 S3에 저장합니다. 자세한 내용은 S3에 사용자 스크립트 및 가상 환경 저장 단원을 참조하십시오.

   b. 사용자 스크립트에 필요한 추가 라이브러리를 지원하도록 선택적 가상 환경을 생성합니다. 자세한 내용은 가상 환경 생성(선택 사항) 단원을 참조하십시오.

Example 예시 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
```

이 예제의 폴더 구조는 다음과 같습니다.

```
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 예시 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
```

가상 환경 생성(선택 사항)

사용자 스크립트에 필요한 추가 라이브러리가 있는 경우 해당 라이브러리를 저장할 가상 환경을 생성할 수 있습니다. 추가 라이브러리가 필요하지 않은 경우이 단계를 건너뛸 수 있습니다.

네이티브 확장이 있는 라이브러리로 작업할 때는의 PySpark가 ARM64 아키텍처를 사용하는 Linux에서 AWS Clean Rooms 작동한다는 점에 유의하세요.

다음 절차에서는 기본 CLI 명령을 사용하여 가상 환경을 생성하는 방법을 보여줍니다.

가상 환경을 생성하려면

1. 터미널 또는 명령 프롬프트를 엽니다.
2. 다음 콘텐츠를 추가합니다.

```
# 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. 이제이 가상 환경을 S3에 저장할 준비가 되었습니다. 자세한 내용은 S3에 사용자 스크립트 및 가상 환경 저장 단원을 참조하십시오.

Docker 및 Amazon ECR 작업에 대한 자세한 내용은 Amazon ECRUser 설명서를 참조하세요.

S3에 사용자 스크립트 및 가상 환경 저장

다음 절차에서는 Amazon S3에 사용자 스크립트와 선택적 가상 환경을 저장하는 방법을 설명합니다. PySpark 분석 템플릿을 생성하기 전에이 단계를 완료합니다.

> **Important**
>
> 분석 템플릿을 생성한 후에는 아티팩트(사용자 스크립트 또는 가상 환경)를 수정하거나 제거하지 마십시오.
> 이렇게 하면 다음이 수행됩니다.
>
> - 이 템플릿을 사용하는 향후 모든 분석 작업이 실패합니다.
> - 새 아티팩트가 포함된 새 분석 템플릿을 생성해야 합니다.
> - 이전에 완료된 분석 작업에 영향을 주지 않음

사전 조건

- 적절한 권한이 AWS 계정 있는
- 사용자 스크립트 파일(예: `my_analysis.py`)
- (선택 사항, 존재하는 경우) 가상 환경 패키지(`.tar.gz` 파일)
- IAM 역할을 생성하거나 수정할 수 있는 액세스 권한

Console

콘솔을 사용하여 S3에 사용자 스크립트와 가상 환경을 저장하려면:

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/s3/ Amazon S3 콘솔을 엽니다.
2. 새 S3 버킷을 생성하거나 기존 버킷을 사용합니다.
3. 버킷에 대한 버전 관리를 활성화합니다.

a. 버킷을 선택합니다.
b. 속성을 선택합니다.
c. 버킷 버전 관리 섹션에서 편집을 선택합니다.
d. 변경 사항 활성화 및 저장을 선택합니다.

4. 아티팩트를 업로드하고 SHA-256 해시를 활성화합니다.

a. 버킷으로 이동합니다.
b. 업로드를 선택합니다.
c. 파일 추가를 선택하고 사용자 스크립트 파일을 추가합니다.
d. (선택 사항, 있는 경우) .tar.gz 파일을 추가합니다.
e. 속성을 확장합니다.
f. 체크섬의 체크섬 함수에서 SHA256을 선택합니다.
g. 업로드를 선택합니다.

5. 이제 PySpark 분석 템플릿을 생성할 준비가 되었습니다.

CLI

AWS CLI다음을 사용하여 사용자 스크립트와 가상 환경을 S3에 저장하려면

1. 다음 명령 실행:

```
aws s3 cp --checksum-algorithm sha256 pyspark_venv.tar.gz s3://ARTIFACT-BUCKET/
EXAMPLE-PREFIX/
```

2. 이제 PySpark 분석 템플릿을 생성할 준비가 되었습니다.

> **Note**
>
> 스크립트 또는 가상 환경을 업데이트해야 하는 경우:
>
> 1. 새 버전을 별도의 객체로 업로드합니다.
> 2. 새 아티팩트를 사용하여 새 분석 템플릿을 생성합니다.
> 3. 이전 템플릿을 사용 중지합니다.
> 4. 이전 템플릿이 여전히 필요할 수 있는 경우 S3에 원래 아티팩트를 유지합니다.

PySpark 분석 템플릿 생성

사전 조건

PySpark 분석 템플릿을 생성하기 전에 다음이 있어야 합니다.

- 활성 AWS Clean Rooms 공동 작업의 멤버십
- 활성 공동 작업에서 구성된 테이블 하나 이상에 대한 액세스
- 분석 템플릿을 생성할 수 있는 권한
- Python 사용자 스크립트 및 S3에 생성 및 저장된 가상 환경
  - S3 버킷에 버전 관리가 활성화되어 있습니다. 자세한 내용은 S3 버킷에서 버전 관리 사용을 참조하세요.
  - S3 버킷은 업로드된 아티팩트에 대한 SHA-256 체크섬을 계산할 수 있습니다. 자세한 내용은 체크섬 사용을 참조하세요.
- S3 버킷에서 코드를 읽을 수 있는 권한

  필요한 서비스 역할 생성에 대한 자세한 내용은 섹션을 참조하세요서비스 역할을 생성하여 S3 버킷에서 코드 읽기(PySpark 분석 템플릿 역할).

다음 절차에서는 AWS Clean Rooms 콘솔을 사용하여 PySpark 분석 템플릿을 생성하는 프로세스를 설명합니다. 사용자 스크립트와 가상 환경 파일을 이미 생성하고 사용자 스크립트와 가상 환경 파일을 Amazon S3 버킷에 저장했다고 가정합니다.

> **Note**
>
> PySpark 분석 템플릿을 생성하는 구성원도 결과를 수신하는 구성원이어야 합니다.

SDK를 사용하여 PySpark 분석 템플릿을 생성하는 방법에 대한 자세한 내용은 AWS Clean Rooms API 참조를 참조하세요. AWS SDKs

PySpark 분석 템플릿을 생성하려면

1. 에 로그인 AWS Management Console 하고 공동 작업 생성자로 AWS 계정 작동하는를 사용하여 AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업을 선택합니다.

4. 템플릿 탭에서 직접 생성한 분석 템플릿 섹션으로 이동합니다.
5. 분석 템플릿 생성을 선택합니다.
6. 분석 템플릿 생성 페이지의 세부 정보에서

   a. 분석 템플릿의 이름을 입력합니다.

   b. (선택 사항) 설명을 입력합니다.

   c. 형식에서 PySpark 옵션을 선택합니다.
7. 정의의 경우,

   a. 계속하기 전에 사전 조건을 검토하고 각 사전 조건이 충족되었는지 확인합니다.

   b. 진입점 파일에 S3 버킷을 입력하거나 S3 찾아보기를 선택합니다.

   c. (선택 사항) 라이브러리 파일에 S3 버킷을 입력하거나 S3 찾아보기를 선택합니다.
8. 정의에서 참조되는 테이블의 경우

   - 정의에서 참조되는 모든 테이블이 공동 작업에 연결된 경우:
     - 정의에서 참조된 모든 테이블이 공동 작업 확인란에 연결된 상태로 둡니다.
     - 공동 작업과 연결된 테이블에서 정의에서 참조되는 모든 연결된 테이블을 선택합니다.
   - 정의에서 참조되는 모든 테이블이 공동 작업에 연결되지 않은 경우:
     - 정의에서 참조된 모든 테이블이 공동 작업과 연결됨 확인란의 선택을 취소합니다.
     - 공동 작업과 연결된 테이블에서 정의에서 참조되는 모든 연결된 테이블을 선택합니다.
     - 나중에 연결할 테이블에 테이블 이름을 입력합니다.
     - 다른 테이블 나열을 선택하여 다른 테이블을 나열합니다.
9. 드롭다운 목록에서 기존 서비스 역할 이름을 선택하여 서비스 액세스 권한을 지정합니다.

   1. 역할을 나열할 권한이 있는 경우 역할 목록이 표시됩니다.

      역할을 나열할 수 있는 권한이 없는 경우 사용하려는 역할의 Amazon 리소스 이름(ARN)을 입력할 수 있습니다.

   2. IAM에서 보기 외부 링크를 선택하여 서비스 역할을 확인합니다.

      기존 서비스 역할이 없는 경우 기존 서비스 역할 사용 옵션을 사용할 수 없습니다.

      기본적으로 AWS Clean Rooms 는 필요한 권한을 추가하기 위해 기존 역할 정책을 업데이트하려고 시도하지 않습니다.

> **Note**
>
> - AWS Clean Rooms 에는 분석 규칙에 따라 쿼리할 수 있는 권한이 필요합니다. 권한에 대한 자세한 내용은 섹션을 AWS Clean Rooms참조하세요AWS 에 대한 관리형 정책 AWS Clean Rooms.
> - 역할에에 대한 충분한 권한이 없는 경우 역할에에 대한 충분한 권한이 없다는 오류 메시지가 AWS Clean Rooms표시됩니다 AWS Clean Rooms. 계속하기 전에 역할 정책을 추가해야 합니다.
> - 역할 정책을 수정할 수 없는 경우 AWS Clean Rooms 에서 서비스 역할에 대한 정책을 찾을 수 없다는 오류 메시지가 나타납니다.

10. 구성된 테이블 리소스에 대해 태그를 활성화하려면 새 태그 추가를 선택한 다음 키와 값 쌍을 입력합니다.
11. 생성(Create)을 선택합니다.
12. 이제 공동 작업 구성원에게 분석 템플릿을 검토할 수 있음을 알릴 준비가 되었습니다. (자체 데이터를 쿼리하려는 경우 선택 사항)

> **Important**
>
> 분석 템플릿을 생성한 후에는 아티팩트(사용자 스크립트 또는 가상 환경)를 수정하거나 제거하지 마십시오.
> 이렇게 하면 다음이 수행됩니다.
>
> - 이 템플릿을 사용하는 향후 모든 분석 작업이 실패합니다.
> - 새 아티팩트가 포함된 새 분석 템플릿을 생성해야 합니다.
> - 이전에 완료된 분석 작업에는 영향을 주지 않습니다.

PySpark 분석 템플릿 검토

다른 구성원이 공동 작업에서 분석 템플릿을 생성할 때 이를 검토하고 승인해야 사용할 수 있습니다.

다음 절차에서는 규칙, 파라미터 및 참조 테이블을 포함하여 PySpark 분석 템플릿을 검토하는 방법을 보여줍니다. 공동 작업 구성원은 템플릿이 데이터 공유 계약 및 보안 요구 사항과 일치하는지 평가합니다.

분석 템플릿이 승인되고 나면의 작업에 사용할 수 있습니다 AWS Clean Rooms.

> **Note**
>
> 분석 코드를 공동 작업으로 가져올 때 다음 사항에 유의하세요.
>
> - AWS Clean Rooms 는 분석 코드의 동작을 검증하거나 보장하지 않습니다.
>   - 특정 동작을 확인해야 하는 경우 공동 작업 파트너의 코드를 직접 검토하거나 신뢰할 수 있는 타사 감사자와 협력하여 검토합니다.
> - AWS Clean Rooms 는 PySpark 분석 템플릿에 나열된 코드의 SHA-256 해시가 PySpark 분석 환경에서 실행되는 코드와 일치하도록 보장합니다.
> - AWS Clean Rooms 는 환경에 가져오는 추가 라이브러리에 대한 감사 또는 보안 분석을 수행하지 않습니다.
> - 공유 보안 모델에서:
>   - 사용자(고객)는 환경에서 실행되는 코드의 보안에 대한 책임이 있습니다.
>   - AWS Clean Rooms 는 환경의 보안을 담당하여 다음을 보장합니다.
>   - 승인된 코드만 실행됩니다.
>   - 지정된 구성된 테이블만 액세스할 수 있습니다.
>   - 유일한 출력 대상은 결과 수신자의 S3 버킷입니다.

AWS Clean Rooms 는 검토를 위해 사용자 스크립트 및 가상 환경의 SHA-256 해시를 생성합니다. 그러나 실제 사용자 스크립트 및 라이브러리는 내에서 직접 액세스할 수 없습니다 AWS Clean Rooms.

공유된 사용자 스크립트 및 라이브러리가 분석 템플릿에서 참조된 것과 동일한지 확인하려면 공유된 파일의 SHA-256 해시를 생성하여에서 생성한 분석 템플릿 해시와 비교할 수 있습니다 AWS Clean Rooms. 코드 실행의 해시도 작업 로그에 포함됩니다.

사전 조건

- Linux/Unix 운영 체제 또는 Linux용 Windows 하위 시스템(WSL)
- 해시하려는 사용자 스크립트 파일
  - 분석 템플릿 생성자가 보안 채널을 통해 파일을 공유하도록 요청합니다.

- 에서 생성한 분석 템플릿 해시 AWS Clean Rooms

AWS Clean Rooms 콘솔을 사용하여 PySpark 분석 템플릿을 검토하려면

1. 에 로그인 AWS Management Console 하고 공동 작업 생성자 역할을 AWS 계정 할를 사용하여 AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업을 선택합니다.
4. 템플릿 탭에서 다른 구성원이 만든 분석 템플릿 섹션으로 이동합니다.
5. 실행 가능 상태가 아니요 - 직접 검토가 필요함인 분석 템플릿을 선택합니다.
6. 검토를 선택합니다.
7. 분석 규칙 개요, 정의 및 매개변수(있는 경우)를 검토합니다.
8. 공유 사용자 스크립트 및 라이브러리가 분석 템플릿에서 참조된 것과 동일한지 확인합니다.

   a. 공유된 파일의 SHA-256 해시를 생성하고에서 생성한 분석 템플릿 해시와 비교합니다 AWS Clean Rooms.

      사용자 스크립트 파일이 포함된 디렉터리로 이동한 다음 다음 명령을 실행하여 해시를 생성할 수 있습니다.

      ```
      sha256sum your_script_filename.py
      ```

      출력 예시:

      ```
      e3b0c44298fc1c149afbf4c8996fb92427ae41e4649b934ca495991b7852b855 my_analysis.py
      ```

   b. 또는 Amazon S3 체크섬 기능을 사용할 수 있습니다. 자세한 내용은 Amazon S3 사용 설명서의 Amazon S3에서 객체 무결성 확인을 참조하세요. Amazon S3

   c. 또 다른 대안은 작업 로그에서 실행된 코드의 해시를 보는 것입니다.
9. 정의에 참조된 테이블 아래에 나열된 구성된 테이블을 검토합니다.

   각 테이블 옆의 상태는 템플릿 허용 불가로 표시됩니다.
10. 테이블을 선택합니다.

    a. 분석 템플릿을 승인하려면 테이블에서 템플릿 허용을 선택합니다. 허용을 선택하여 승인을 확인합니다.

b. 승인을 거부하려면 허용 안 함을 선택합니다.

분석 템플릿을 승인하기로 선택한 경우 작업을 실행할 수 있는 구성원은 이제 PySpark 분석 템플릿을 사용하여 구성된 테이블에서 PySpark 작업을 실행할 수 있습니다. 자세한 내용은 PySpark 작업 실행 단원을 참조하십시오.

PySpark 분석 템플릿 문제 해결

PySpark 분석 템플릿을 사용하여 작업을 실행할 때 작업 초기화 또는 실행 중에 오류가 발생할 수 있습니다. 이러한 실패는 일반적으로 스크립트 구성, 데이터 액세스 권한 또는 환경 설정과 관련이 있습니다.

PySpark 제한 사항에 대한 자세한 내용은 섹션을 참조하세요의 PySpark 제한 사항 AWS Clean Rooms.

주제

- 코드 문제 해결
- 분석 템플릿 작업이 시작되지 않음
- 분석 템플릿 작업이 시작되지만 처리 중에 실패함
- 가상 환경 설정 실패

코드 문제 해결

AWS Clean Rooms 는 오류 메시지 및 로그에서 민감한 데이터를 제한하여 고객의 기본 데이터를 보호합니다. 코드를 개발하고 문제를 해결하는 데 도움이 되도록 자체 계정 AWS Clean Rooms 에서 시뮬레이션하고 자체 테스트 데이터를 사용하여 작업을 실행하는 것이 좋습니다.

다음 단계를 통해 Amazon EMR Serverless AWS Clean Rooms 의에서 PySpark를 시뮬레이션할 수 있습니다. AWS Clean Rooms의 PySpark와 약간의 차이점이 있지만 주로 코드 실행 방법을 다룹니다.

EMR Serverless AWS Clean Rooms 에서의 PySpark를 시뮬레이션하려면

1. Amazon S3에서 데이터 세트를 생성하고,에서 데이터 세트를 카탈로그화하고 AWS Glue Data Catalog, Lake Formation 권한을 설정합니다.
2. 사용자 지정 역할을 사용하여 Lake Formation에 S3 위치를 등록합니다.
3. Amazon EMR Studio 인스턴스가 아직 없는 경우 생성합니다(Amazon EMR Serverless를 사용하려면 Amazon EMR Studio 필요).

4. EMR Serverless 앱 생성
    - 릴리스 버전 emr-7.7.0을 선택합니다.
    - ARM64 아키텍처를 선택합니다.
    - 사용자 지정 설정 사용을 선택합니다.
    - 사전 초기화된 용량을 비활성화합니다.
    - 대화형 작업을 수행하려는 경우 대화형 엔드포인트 > EMR 스튜디오에 엔드포인트 활성화를 선택합니다.
    - 추가 구성 > 세분화된 액세스 제어를 위해 Lake Formation 사용을 선택합니다.
    - 애플리케이션을 생성합니다.
5. EMR-Studio 노트북 또는 `StartJobRun` API를 통해 EMR-S를 사용합니다.

분석 템플릿 작업이 시작되지 않음

일반적인 원인

분석 템플릿 작업은 세 가지 주요 구성 문제로 인해 시작 시 즉시 실패할 수 있습니다.

- 필요한 형식과 일치하지 않는 잘못된 스크립트 이름 지정
- 사용자 스크립트의 진입점 함수가 누락되었거나 형식이 잘못 지정됨
- 가상 환경의 호환되지 않는 Python 버전

해결 방법

이 문제를 해결하려면:

1. 사용자 스크립트를 확인합니다.
    - 사용자 스크립트에 유효한 Python 파일 이름이 있는지 확인합니다.

      유효한 Python 파일 이름은 소문자, 밑줄을 사용하여 단어를 구분하고 .py 확장자를 사용합니다.
2. 진입점 함수를 확인합니다. 사용자 스크립트에 진입점 함수가 없는 경우 진입점 함수를 추가합니다.
    a. 사용자 스크립트를 엽니다.

b. 이 진입점 함수를 추가합니다.

```
def entrypoint(context):
    # Your analysis code here
```

c. 함수 이름의 철자가와 정확히 일치하는지 확인합니다`entrypoint`.

d. 함수가 `context` 파라미터를 수락하는지 확인합니다.

3. Python 버전 호환성 확인:

a. 가상 환경에서 Python 3.9를 사용하는지 확인합니다.

b. 버전을 확인하려면 다음을 실행합니다. `python --version`

c. 필요한 경우 가상 환경을 업데이트합니다.

```
conda create -n analysis-env python=3.9
conda activate analysis-env
```

예방책

- 올바른 파일 구조가 포함된 제공된 분석 템플릿 스타터 코드를 사용합니다.
- 모든 분석 템플릿에 대해 Python 3.9를 사용하여 전용 가상 환경을 설정합니다.
- 작업을 제출하기 전에 템플릿 검증 도구를 사용하여 분석 템플릿을 로컬에서 테스트합니다.
- CI/CD 검사를 구현하여 스크립트 이름 지정 및 진입점 함수 요구 사항을 확인합니다.

분석 템플릿 작업이 시작되지만 처리 중에 실패함

일반적인 원인

다음과 같은 보안 및 형식 지정 이유로 실행 중에 분석 작업이 실패할 수 있습니다.

- Amazon S3 또는와 같은 AWS 서비스에 대한 무단 직접 액세스 시도 AWS Glue
- 필수 DataFrame 사양과 일치하지 않는 잘못된 형식의 출력 반환
- 실행 환경의 보안 제한으로 인한 네트워크 호출 차단

해결 방법

해결하려면

1. 직접 AWS 서비스 액세스 제거:
   a. 코드를 검색하여 직접 AWS 서비스 가져오기 및 호출을 찾습니다.
   b. 직접 S3 액세스를 제공된 Spark 세션 메서드로 바꿉니다.
   c. 공동 작업 인터페이스를 통해 미리 구성된 테이블만 사용합니다.
2. 올바른 출력 형식 지정:
   a. 모든 출력이 Spark DataFrames인지 확인합니다.
   b. 다음 형식과 일치하도록 반환 문을 업데이트합니다.

```
return {
    "results": {
        "output1": dataframe1
    }
}
```

   c. DataFrame이 아닌 반환 객체를 제거합니다.
3. 네트워크 호출 제거:
   a. 외부 API 호출을 식별하고 제거합니다.
   b. urllib, 요청 또는 유사한 네트워크 라이브러리를 제거합니다.
   c. 소켓 연결 또는 HTTP 클라이언트 코드를 제거합니다.

예방책

- 제공된 코드 린터를 사용하여 무단 AWS 가져오기 및 네트워크 호출을 확인합니다.
- 보안 제한이 프로덕션과 일치하는 개발 환경에서 작업을 테스트합니다.
- 작업을 배포하기 전에 출력 스키마 검증 프로세스를 따릅니다.
- 승인된 서비스 액세스 패턴에 대한 보안 지침을 검토합니다.

가상 환경 설정 실패

일반적인 원인

가상 환경 구성 실패는 일반적으로 다음과 같은 이유로 발생합니다.

- 개발 환경과 실행 환경 간의 CPU 아키텍처 불일치
- 적절한 환경 초기화를 방해하는 Python 코드 형식 지정 문제
- 컨테이너 설정의 기본 이미지 구성이 잘못되었습니다.

해결 방법

해결하려면

1. 올바른 아키텍처를 구성합니다.
   a. 를 사용하여 현재 아키텍처 확인 `uname -m`.
   b. Dockerfile을 업데이트하여 ARM64를 지정합니다.

   ```
   FROM --platform=linux/arm64 public.ecr.aws/amazonlinux/amazonlinux:2023-minimal
   ```

   c. 를 사용하여 컨테이너 재구축 `docker build --platform=linux/arm64`.
2. Python 들여쓰기 수정:
   a. 코드 파일`black`에서와 같은 Python 코드 형식 지정자를 실행합니다.
   b. 공백 또는 탭(둘 다 아님)의 일관된 사용을 확인합니다.
   c. 모든 코드 블록의 들여쓰기를 확인합니다.

   ```
   def my_function():
       if condition:
           do_something()
       return result
   ```

   d. Python 들여쓰기 강조 표시와 함께 IDE를 사용합니다.
3. 환경 구성 검증:
   a. 를 실행`python -m py_compile your_script.py`하여 구문 오류를 확인합니다.
   b. 배포 전에 로컬에서 환경을 테스트합니다.

c. 모든 종속성이에 나열되어 있는지 확인합니다`requirements.txt`.

예방책

- Python 형식 플러그인과 함께 Visual Studio 코드 또는 PyCharm 사용
- 코드 형식자를 자동으로 실행하도록 사전 커밋 후크 구성
- 제공된 ARM64 기본 이미지를 사용하여 로컬에서 환경 빌드 및 테스트
- CI/CD 파이프라인에서 자동 코드 스타일 검사 구현

공동 작업에서 데이터 분석

에서는 쿼리 또는 작업을 실행하여 데이터를 분석할 AWS Clean Rooms수 있습니다.

쿼리는 지원되는 함수, 클래스 및 변수 세트를 사용하여 공동 작업에서 구성된 테이블에 액세스하고 분석하는 방법입니다. 에서 현재 지원되는 쿼리 언어는 SQL AWS Clean Rooms 입니다. 에서 쿼리를 실행하는 방법에는 SQL 코드 작성, 승인된 SQL 분석 템플릿 사용 또는 분석 빌더 UI 사용 AWS Clean Rooms의 세 가지가 있습니다.

작업은 지원되는 함수, 클래스 및 변수 세트를 사용하여 공동 작업에서 구성된 테이블에 액세스하고 분석하는 방법입니다. 에서 현재 지원되는 작업 유형은 PySpark AWS Clean Rooms 입니다. 에서 작업을 실행하는 한 가지 방법은 승인된 PySpark 분석 템플릿을 AWS Clean Rooms사용하는 것입니다.

다음 주제에서는 SQL 쿼리 또는 PySpark 작업을 실행 AWS Clean Rooms 하여의 데이터를 분석하는 방법을 설명합니다.

주제

- SQL 쿼리 실행
- PySpark 작업 실행

SQL 쿼리 실행

> **Note**
>
> 쿼리 컴퓨팅 비용을 지불해야 하는 구성원이 활성 구성원으로 공동 작업에 참여한 경우에만 쿼리를 실행할 수 있습니다.

쿼리할 수 있는 멤버는 다음을 통해 SQL 쿼리를 실행할 수 있습니다.

- SQL 코드 편집기를 사용하여 SQL 쿼리를 수동으로 빌드합니다.
- 승인된 SQL 분석 템플릿 사용.
- 분석 빌더 UI를 사용하여 SQL 코드를 작성할 필요 없이 쿼리를 빌드합니다.

쿼리할 수 있는 구성원이 공동 작업의 테이블에서 SQL 쿼리를 실행하면는 사용자를 대신하여 테이블에 액세스하기 위한 관련 역할을 AWS Clean Rooms 수임합니다.는 필요에 따라 분석 규칙을 입력 쿼리 및 해당 출력에 AWS Clean Rooms 적용합니다.

분석 규칙 및 출력 제약 조건은 자동으로 적용됩니다.는 정의된 분석 규칙을 준수하는 결과 AWS Clean Rooms 만 반환합니다.

AWS Clean Rooms 는 다른 쿼리 엔진과 다를 수 있는 SQL 쿼리를 지원합니다. 사양은 AWS Clean Rooms SQL 참조를 참조하세요. 차등 프라이버시로 보호되는 데이터 테이블에서 쿼리를 실행하려면 쿼리가 AWS Clean Rooms 차등 프라이버시의 범용 쿼리 구조와 호환되는지 확인해야 합니다.

> **Note**
>
> Clean Rooms에 암호화 컴퓨팅을 사용하는 경우 일부 SQL 작업에서는 올바른 결과가 생성되지 않습니다. 예를 들어, 암호화된 열에서 COUNT를 수행할 수 있지만 암호화된 번호에서 SUM을 수행하면 오류가 발생합니다. 또한 쿼리로 인해 잘못된 결과가 나올 수도 있습니다. 예를 들어 SUM이 열을 봉인하는 쿼리는 오류를 발생시킵니다. 하지만 봉인된 열에 대한 GROUP BY 쿼리는 성공한 것 같지만 일반 텍스트에 대한 GROUP BY 쿼리로 생성된 그룹과는 다른 그룹을 생성합니다.

쿼리 컴퓨팅 비용을 지불하는 멤버에게는 공동 작업에서 실행한 쿼리에 대한 요금이 부과됩니다.

쿼리할 수 있는 구성원은 결과를 받을 수 있는 여러 구성원을 선택하여 단일 쿼리에서 결과를 받을 수 있습니다. 자세한 내용은 SQL 코드 편집기를 사용하여 구성된 테이블 쿼리 단원을 참조하십시오. 쿼리 결과 수신에 대한 일반적인 내용은 섹션을 참조하세요분석 결과 수신 및 사용.

사전 조건

SQL 쿼리를 실행하기 전에 다음이 있는지 확인합니다.

- AWS Clean Rooms 공동 작업 중인 활성 멤버십
- 공동 작업에서 하나 이상의 구성된 테이블에 대한 액세스
- 쿼리 컴퓨팅 비용을 담당하는 구성원이 활성 공동 작업 구성원인지 확인

`StartProtectedQuery` API 작업을 직접 호출하거나SDK를 AWS Clean Rooms 사용하여 데이터를 쿼리하거나 쿼리를 보는 방법에 대한 자세한 내용은 AWS Clean Rooms API 참조를 참조하세요. AWS SDKs

쿼리 로깅에 대한 자세한 내용은 의 분석 로깅 AWS Clean Rooms 섹션을 참조하세요.

> **Note**
>
> 암호화된 데이터 테이블에서 쿼리를 실행하면 암호화된열의 결과가 암호화됩니다.

다음 항목에서는 AWS Clean Rooms 콘솔을 사용하여 공동 작업에서 데이터를 쿼리하는 방법을 설명합니다.

주제

- SQL 코드 편집기를 사용하여 구성된 테이블 쿼리
- SQL 코드 편집기를 사용하여 ID 매핑 테이블 쿼리
- SQL 분석 템플릿을 사용하여 구성된 테이블 쿼리
- 분석 빌더로 쿼리
- 차등 프라이버시의 영향 확인
- 최근 쿼리 보기
- 쿼리 세부 정보 보기

SQL 코드 편집기를 사용하여 구성된 테이블 쿼리

쿼리를 할 수 있는 구성원은 SQL 코드 편집기에서 SQL 코드를 작성하여 수동으로 쿼리를 작성할 수 있습니다. SQL 코드 편집기는 AWS Clean Rooms 콘솔의 쿼리 탭에 있는 분석 섹션에 있습니다.

SQL 코드 편집기가 기본적으로 표시됩니다. 분석 빌더를 사용하여 쿼리를 작성하려면 분석 빌더로 쿼리을 참조하세요.

> **Important**
>
> 코드 편집기에서 SQL 쿼리 작성을 시작한 다음 분석 빌더 UI를 켜면 쿼리가 저장되지 않습니다.

AWS Clean Rooms 는 많은 SQL 명령, 함수 및 조건을 지원합니다. 자세한 내용은 AWS Clean Rooms SQL 참조의 섹션을 참조하세요.

> **Tip**
>
> 쿼리 실행 도중 예정된 유지 관리가 실행되면 쿼리는 종료 후 롤백됩니다. 쿼리를 다시 시작해야 합니다.

SQL 코드 편집기를 사용하여 구성된 테이블을 쿼리하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 구성원 권한 상태가 쿼리인 공동 작업을 선택합니다.
4. 쿼리 탭에서 분석 섹션으로 이동합니다.

   > **Note**
   >
   > 분석 섹션은 결과를 받을 수 있는 구성원과 쿼리 컴퓨팅 비용을 지불해야 하는 구성원이 활성 구성원으로 공동 작업에 참여한 경우에만 표시됩니다.

5. 쿼리 탭의 테이블에서 테이블 목록 및 관련 분석 규칙 유형(집계 분석 규칙, 목록 분석 규칙 또는 사용자 지정 분석 규칙)을 볼 수 있습니다.

   > **Note**
   >
   > 목록에 예상한 테이블이 보이지 않는 경우 다음과 같은 이유가 있을 수 있습니다.
   >
   > - 테이블이 연결되지 않았습니다.
   > - 테이블에는 분석 규칙이 구성되어 있지 않습니다.

6. (선택 사항) 테이블의 스키마와 분석 규칙 컨트롤을 보려면 더하기 기호 아이콘(+)을 선택하여 테이블을 확장합니다.
7. SQL 코드 편집기에 쿼리를 입력하여 쿼리를 작성합니다.

   지원되는 SQL 명령 및 함수에 대한 자세한 내용은 AWS Clean Rooms SQL 참조를 참조하세요.

   다음 옵션을 사용하여 쿼리를 빌드할 수도 있습니다.

Use an example query

예제 쿼리를 사용하려면

1. 테이블 옆의 세로로 된 세 점을 선택합니다.
2. 편집기에 삽입에서 예제 쿼리를 선택합니다.

> **Note**
> 예제 쿼리를 삽입하면 편집기에 이미 있는 쿼리에 추가됩니다.

쿼리 예제가 나타납니다. 테이블 아래에 나열된 모든 테이블이 쿼리에 포함됩니다.

3. 쿼리에서 자리 표시자 값을 편집합니다.

Insert column names or functions

열 이름 또는 함수를 삽입하려면

1. 열 옆에 있는 세 개의 수직 점을 선택합니다.
2. 편집기에 삽입에서 열 이름을 선택합니다.
3. 열에 허용되는 함수를 수동으로 삽입하려면 열 옆에 있는 세 개의 세로 점을 선택하고 편집기에 삽입을 선택한 다음 허용된 함수의 이름(예: INNER JOIN, SUMSUM DISTINCT, 또는 COUNT)을 선택합니다.
4. Ctrl + Space를 눌러 코드 편집기에서 테이블 스키마를 볼 수 있습니다.

> **Note**
> 쿼리가 가능한 구성원은 구성된 각 테이블 연결의 파티션 열을 보고 사용할 수 있습니다. 구성된 AWS Glue 테이블의 기반이 되는 테이블에서 파티션 열에 파티션 열로 레이블이 지정되어 있는지 확인합니다.

5. 쿼리에서 자리 표시자 값을 편집합니다.

8. (Spark 분석 엔진만 해당) 지원되는 작업자 유형과 작업자 수를 지정합니다.

다음 표를 사용하여 사용 사례에 필요한 유형과 수 또는 작업자를 결정합니다.

> **Note**
>
> 작업자 유형과 작업자 수가 다르면 관련 비용이 발생합니다. 요금에 대한 자세한 내용은 AWS Clean Rooms 요금을 참조하세요.

작업자 유형	vCPU	메모리(GB)	스토리지(GB)	작업자 수	총 클린룸 처리 단위(CRPU)
CR.1X(기본값)	4	30	100	2	4
				16(기본값)	32
CR.4X	16	120	400	8	64
				32	256

9. 결과 전송 대상에서 결과를 받을 수 있는 사람을 지정합니다.

> **Note**
>
> 결과를 수신하려면 공동 작업 구성원을 결과 수신자로 구성해야 하며 공동 작업의 활성 참가자여야 합니다(상태: 활성).

10. (쿼리만 가능한 구성원) 기본 결과 설정 사용 확인란이 기본적으로 선택되어 있습니다. 기본 결과 설정을 유지하려면이 옵션을 선택한 상태로 유지합니다.

    이 쿼리에 대해 다른 결과 설정을 지정하려면 기본 결과 설정 사용 확인란의 선택을 취소한 후 다음을 선택합니다.

    a. 결과 형식(CSV 또는 PARQUET)

    b. 결과 파일(단일 또는 다중)

    c. Amazon S3의 결과 대상

결과를 받을 수 있는 각 멤버는 Amazon S3에서 다른 결과 형식, 결과 파일 및 결과 대상을 지정할 수 있습니다. Amazon S3

11. Run(실행)을 선택합니다.

> **Note**
>
> 결과를 받을 수 있는 구성원이 쿼리 결과 설정을 구성하지 않은 경우 쿼리를 실행할 수 없습니다.

12. 결과를 확인합니다.

    자세한 내용은 분석 결과 수신 및 사용 단원을 참조하십시오.

13. 계속해서 매개변수를 조정하고 쿼리를 다시 실행하거나 + 버튼을 선택하여 새 탭에서 새 쿼리를 시작합니다.

> **Note**
>
> AWS Clean Rooms 는 명확한 오류 메시징을 제공하는 것을 목표로 합니다. 오류 메시지에 문제 해결에 도움이 되는 세부 정보가 충분하지 않은 경우 계정 팀에 문의하세요. 오류 발생 경위에 대한 설명과 오류 메시지(식별자 포함)를 제공하세요. 자세한 내용은 문제 해결 AWS Clean Rooms 단원을 참조하십시오.

SQL 코드 편집기를 사용하여 ID 매핑 테이블 쿼리

다음 절차에서는 ID 매핑 테이블에 대해 다중 테이블 조인 쿼리를 실행하여 `sourceId`를 `targetId`와 조인하는 방법에 대해 설명합니다.

ID 매핑 테이블을 쿼리하기 전에 ID 매핑 테이블을 성공적으로 채워야 합니다.

SQL 코드 편집기를 사용하여 ID 매핑 테이블을 쿼리하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 사용자의 멤버 자격 상태가 쿼리 실행인 공동 작업을 선택합니다.

4. 쿼리 탭에서 분석 섹션으로 이동합니다.

> **Note**
>
> 분석 섹션은 결과를 받을 수 있는 구성원과 쿼리 컴퓨팅 비용을 지불해야 하는 구성원이 활성 구성원으로 공동 작업에 참여한 경우에만 표시됩니다.

5. 쿼리 탭의 테이블에서 ID 매핑 테이블 목록(관리자 AWS Clean Rooms 아래) 및 관련 분석 규칙 유형(ID 매핑 테이블 분석 규칙)을 확인합니다.

> **Note**
>
> 예상한 ID 매핑 테이블이 목록에 표시되지 않는다면 ID 매핑 테이블이 성공적으로 채워지지 않았기 때문일 수 있습니다. 자세한 내용은 기존 ID 매핑 테이블 채우기 단원을 참조하십시오.

6. SQL 코드 편집기에 쿼리를 입력하여 쿼리를 작성합니다.

(선택 사항) 예제 쿼리를 사용하려는 경우	(선택 사항) 테이블 이름을 삽입하려는 경우
1. 테이블 옆의 세로로 된 세 점을 선택합니다. 2. 편집기에 삽입에서 JOIN 명령문 예시를 선택합니다. **Note** 예제 JOIN 문을 삽입하면 편집기에 이미 쿼리가 추가됩니다. JOIN 명령문 예시가 나타납니다. 3. 쿼리에서 자리 표시자 값을 편집합니다.	1. 열 옆에 있는 세 개의 수직 점을 선택합니다. 2. 편집기에 삽입에서 테이블 이름을 선택합니다. 3. 쿼리에서 자리 표시자 값을 편집합니다.

7. Run(실행)을 선택합니다.

> **Note**
>
> 결과를 받을 수 있는 구성원이 쿼리 결과 설정을 구성하지 않은 경우 쿼리를 실행할 수 없습니다.

8. 결과를 확인합니다.

   자세한 내용은 분석 결과 수신 및 사용 단원을 참조하십시오.

9. 계속해서 매개변수를 조정하고 쿼리를 다시 실행하거나 + 버튼을 선택하여 새 탭에서 새 쿼리를 시작합니다.

> **Note**
>
> AWS Clean Rooms 는 명확한 오류 메시징을 제공하는 것을 목표로 합니다. 오류 메시지에 문제 해결에 도움이 되는 세부 정보가 충분하지 않은 경우 계정 팀에 문의하세요. 오류 발생 경위에 대한 설명과 오류 메시지(식별자 포함)를 제공하세요. 자세한 내용은 문제 해결 AWS Clean Rooms 단원을 참조하십시오.

SQL 분석 템플릿을 사용하여 구성된 테이블 쿼리

이 절차에서는 AWS Clean Rooms 콘솔에서 분석 템플릿을 사용하여 사용자 지정 분석 규칙을 사용하여 구성된 테이블을 쿼리하는 방법을 보여줍니다.

SQL 분석 템플릿을 사용하여 사용자 지정 분석 규칙을 사용하여 구성된 테이블을 쿼리하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 사용자의 멤버 자격 상태가 쿼리 실행인 공동 작업을 선택합니다.
4. 분석 탭의 테이블 섹션에서 테이블과 관련 분석 규칙 유형(사용자 지정 분석 규칙)을 확인합니다.

> **Note**
>
> 목록에 예상한 테이블이 보이지 않는 경우 다음과 같은 이유가 있을 수 있습니다.
>
> - 테이블이 연결되지 않았습니다.

- 테이블에는 분석 규칙이 구성되어 있지 않습니다.

5. 분석 섹션의 분석 템플릿 실행을 선택한 다음 드롭다운 목록에서 분석 템플릿을 선택합니다.
6. 쿼리에 사용할 분석 템플릿의 매개변수 값을 입력합니다.

   값은 매개변수의 지정된 데이터 유형이어야 합니다.

   분석 템플릿을 실행할 때마다 다른 값을 사용할 수 있습니다.

   파라미터의 빈 값 또는 NULL 값은 지원되지 않습니다. LIMIT 절에서 파라미터를 사용하는 것도 지원되지 않습니다.
7. Run(실행)을 선택합니다.

> **Note**
>
> 결과를 받을 수 있는 구성원이 쿼리 결과 설정을 구성하지 않은 경우 쿼리를 실행할 수 없습니다.

8. 계속해서 매개변수를 조정하고 쿼리를 다시 실행하거나 + 버튼을 선택하여 새 탭에서 새 쿼리를 시작합니다.

분석 빌더로 쿼리

SQL 코드를 작성할 필요 없이 분석 빌더를 사용하여 쿼리를 작성할 수 있습니다. 분석 빌더를 사용하면 다음과 같은 기능을 갖춘 공동 작업을 위한 쿼리를 작성할 수 있습니다.

- 집계 분석 규칙을 사용하는 단일 테이블(JOIN 필요 없음)
- 집계 분석 규칙을 모두 사용하는 두 테이블(각 멤버에서 하나씩)
- 목록 분석 규칙을 모두 사용하는 두 테이블(각 멤버에서 하나씩)
- 집계 분석 규칙을 모두 사용하는 두 개의 테이블(각 구성원에서 하나씩)과 목록 분석 규칙을 모두 사용하는 두 개의 테이블(각 구성원에서 하나씩)

SQL 쿼리를 수동으로 작성하려면 SQL 코드 편집기를 사용하여 구성된 테이블 쿼리을 참조하세요.

분석 빌더는 AWS Clean Rooms 콘솔의 쿼리 탭에 있는 분석 섹션에 분석 빌더 UI 옵션으로 표시됩니다.

> **Important**
>
> 분석 빌더 UI를 켜고 분석 빌더에서 쿼리 작성을 시작한 다음 분석 빌더 UI를 끄면 쿼리가 저장되지 않습니다.

> **Tip**
>
> 쿼리 실행 도중 예정된 유지 관리가 실행되면 쿼리는 종료 후 롤백됩니다. 쿼리를 다시 시작해야 합니다.

다음 항목에서는 분석 체계를 사용하는 방법을 설명합니다.

주제

- 분석 빌더를 사용하여 단일 테이블 (집계) 을 쿼리할 수 있습니다
- 분석 빌더를 사용하여 두 테이블(집계 또는 목록)을 쿼리할 수 있습니다

분석 빌더를 사용하여 단일 테이블 (집계) 을 쿼리할 수 있습니다

이 절차에서는 AWS Clean Rooms 콘솔에서 분석 빌더 UI를 사용하여 쿼리를 빌드하는 방법을 보여줍니다. 이 쿼리는 별도 필요 없이 JOIN집계 분석 규칙을 사용하는 단일 테이블이 있는 공동 작업용입니다.

분석 빌더를 사용하여 단일 테이블을 쿼리하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 구성원 권한 상태가 쿼리인 공동 작업을 선택합니다.
4. 쿼리 탭의 테이블에서 테이블 및 관련 분석 규칙 유형을 확인합니다. (분석 규칙 유형은 집계 분석 규칙이어야 합니다.)

> **Note**
>
> 예상한 테이블이 보이지 않는 경우 다음과 같은 이유가 있을 수 있습니다.

- 테이블이 연결되지 않았습니다.
- 테이블에는 분석 규칙이 구성되어 있지 않습니다.

5. 분석 섹션에서 분석 빌더 UI를 켜세요.
6. 쿼리를 작성하세요.

   모든 집계 지표를 보려면 9단계로 건너뛰세요.

   a. 지표 선택에서는 기본적으로 사전 선택된 집계 지표를 검토하고 필요한 경우 지표를 제거합니다.

   b. (선택 사항) 세그먼트 추가 — 선택 사항의 경우 하나 이상의 매개 변수를 선택합니다.

   > **Note**
   >
   > 세그먼트 추가 - 선택 사항은 테이블에 치수가 지정된 경우에만 표시됩니다.

   c. (선택 사항) 필터 추가 - 선택 사항의 경우 필터 추가를 선택한 다음 매개 변수, 연산자 및 값을 선택합니다.

   다른 필터 그룹을 추가하려는 경우, 필터 그룹 추가를 선택합니다.

   필터를 제거하려면 제거를 선택합니다.

   > **Note**
   >
   > ORDER BY는 집계 쿼리를 지원하지 않습니다.
   > AND 연산자만 필터에서 지원됩니다.

   d. (선택 사항) 설명 추가 - 선택 사항에는 쿼리 목록에서 쿼리를 식별하는 데 도움이 되는 설명을 입력합니다.

7. SQL 코드 미리 보기를 확장합니다.

   a. 분석 빌더에서 생성된 SQL 코드를 확인합니다.

   b. SQL 코드를 복사하려면 복사를 선택합니다.

   c. SQL 코드를 편집하려면 SQL 코드 편집기에서 편집을 선택합니다.

8. Run(실행)을 선택합니다.

> **Note**
>
> 결과를 받을 수 있는 구성원이 쿼리 결과 설정을 구성하지 않은 경우 쿼리를 실행할 수 없습니다.

9. 계속해서 매개변수를 조정하고 쿼리를 다시 실행하거나 + 버튼을 선택하여 새 탭에서 새 쿼리를 시작합니다.

> **Note**
>
> AWS Clean Rooms 는 명확한 오류 메시징을 제공하는 것을 목표로 합니다. 오류 메시지에 문제 해결에 도움이 되는 세부 정보가 충분하지 않은 경우 계정 팀에 문의하세요. 오류 발생 경위에 대한 설명과 오류 메시지(식별자 포함)를 제공하세요. 자세한 내용은 문제 해결 AWS Clean Rooms 단원을 참조하십시오.

분석 빌더를 사용하여 두 테이블(집계 또는 목록)을 쿼리할 수 있습니다

이 절차에서는 AWS Clean Rooms 콘솔에서 분석 빌더를 사용하여 다음과 같은 공동 작업에 대한 쿼리를 작성하는 방법을 설명합니다.

- 집계 분석 규칙을 모두 사용하는 두 테이블(각 구성원에서 하나씩)
- 목록 분석 규칙을 모두 사용하는 두 테이블(각 멤버에서 하나씩)
- 집계 분석 규칙을 모두 사용하는 두 개의 테이블(각 구성원에서 하나씩)과 목록 분석 규칙을 모두 사용하는 두 개의 테이블(각 구성원에서 하나씩)

분석 빌더를 사용하여 두 테이블을 쿼리하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 구성원 권한 상태가 쿼리인 공동 작업을 선택합니다.
4. 쿼리 탭의 테이블에서 두 테이블과 관련 분석 규칙 유형(집계 분석 규칙 또는 목록 분석 규칙)을 확인합니다.

> **Note**
>
> 목록에 예상한 테이블이 보이지 않는 경우 다음과 같은 이유가 있을 수 있습니다.
>
> - 테이블이 연결되지 않았습니다.
> - 테이블에는 분석 규칙이 구성되어 있지 않습니다.

5. 분석 섹션에서 분석 빌더 UI를 켜세요.
6. 쿼리를 작성하세요.

공동 작업에 집계 분석 규칙을 사용하는 두 개의 테이블과 목록 분석 규칙을 사용하는 두 개의 테이블이 포함된 경우 먼저 집계 또는 목록을 선택한 다음 선택한 분석 규칙을 기반으로 프롬프트를 따릅니다.

두 테이블에서 집계 분석 규칙을 사용하는 경우	두 테이블이 목록 분석 규칙을 사용하는 경우
1. 지표 선택에서는 기본적으로 미리 선택된 집계 지표를 검토하고 필요한 경우 지표를 제거합니다. 2. 매치 레코드의 경우 하나 이상의 레코드를 선택합니다. **Note** 분석 빌더를 사용하는 경우 한 쌍의 열에서만 매칭할 수 있습니다. 3. (선택 사항) 세그먼트 추가 — 선택 사항의 경우 하나 이상의 매개 변수를 선택합니다. **Note** 세그먼트 추가 - 선택 사항은 테이블에 치수가 지정된 경우에만 표시됩니다.	1. 속성 선택의 경우 기본적으로 미리 선택된 목록 속성을 검토하고 필요한 경우 지표를 제거합니다. 2. 매치 레코드의 경우 하나 이상의 레코드를 선택합니다. **Note** 분석 빌더를 사용하는 경우 한 쌍의 열에서만 매칭할 수 있습니다. 3. (선택 사항) 필터 추가 - 선택 사항의 경우 필터 추가를 선택한 다음 매개 변수, 연산자 및 값을 선택합니다. 다른 필터 그룹을 추가하려는 경우, 필터 그룹 추가를 선택합니다. 필터를 제거하려면 제거를 선택합니다.

두 테이블에서 집계 분석 규칙을 사용하는 경우	두 테이블이 목록 분석 규칙을 사용하는 경우
4. (선택 사항) 필터 추가 - 선택 사항의 경우 필터 추가를 선택한 다음 매개 변수, 연산자 및 값을 선택합니다. 다른 필터 그룹을 추가하려는 경우, 필터 그룹 추가를 선택합니다. 필터를 제거하려면 제거를 선택합니다. **Note** ORDER BY는 집계 쿼리에는 지원되지 않습니다. AND 연산자만 필터에서 지원됩니다. 5. (선택 사항) 설명 추가 - 선택 사항에는 최근 쿼리 목록에서 쿼리를 식별하는 데 도움이 되는 설명을 입력합니다.	**Note** 목록 쿼리에는 LIMIT이 지원되지 않습니다. AND 연산자만 필터에서 지원됩니다. 4. (선택 사항) 설명 추가 - 선택 사항에는 최근 쿼리 목록에서 쿼리를 식별하는 데 도움이 되는 설명을 입력합니다.

7. SQL 코드 미리 보기를 확장합니다.
   a. 분석 빌더에서 생성된 SQL 코드를 확인합니다.
   b. SQL 코드를 복사하려면 복사를 선택합니다.
   c. SQL 코드를 편집하려면 SQL 코드 편집기에서 편집을 선택합니다.
8. Run(실행)을 선택합니다.

Note
결과를 받을 수 있는 구성원이 쿼리 결과 설정을 구성하지 않은 경우 쿼리를 실행할 수 없습니다.

9. 계속해서 매개변수를 조정하고 쿼리를 다시 실행하거나 + 버튼을 선택하여 새 탭에서 새 쿼리를 시작합니다.

> **Note**
>
> AWS Clean Rooms 는 명확한 오류 메시징을 제공하는 것을 목표로 합니다. 오류 메시지에 문제 해결에 도움이 되는 세부 정보가 충분하지 않은 경우 계정 팀에 문의하세요. 오류 발생 경위에 대한 설명과 오류 메시지(식별자 포함)를 제공하세요. 자세한 내용은 문제 해결 AWS Clean Rooms 단원을 참조하십시오.

차등 프라이버시의 영향 확인

일반적으로 차등 프라이버시가 켜져 있어도 쿼리 작성 및 실행은 달라지지 않습니다. 하지만 남은 프라이버시 예산이 충분하지 않으면 쿼리를 실행할 수 없습니다. 쿼리를 실행하고 프라이버시 예산을 사용함에 따라 실행할 수 있는 집계 수와 향후 쿼리에 미치는 영향을 대략적으로 확인할 수 있습니다.

공동 작업에서 차등 프라이버시가 미치는 영향을 보려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 구성원 세부 정보 상태가 쿼리 실행인 공동 작업을 선택합니다.
4. 쿼리 탭의 테이블에서 남은 프라이버시 예산을 확인합니다. 남은 집계 함수의 예상 수와 사용된 유틸리티(백분율로 표시)로 표시됩니다.

   > **Note**
   >
   > 남은 집계 함수의 예상 수와 사용된 유틸리티의 백분율은 쿼리가 가능한 구성원에 대해서만 표시됩니다.

5. 결과에 주입되는 노이즈의 양과 실행할 수 있는 집계 함수의 대략적인 수를 보려면 영향 보기를 선택합니다.

최근 쿼리 보기

분석 탭에서 지난 90일 동안 실행된 쿼리를 볼 수 있습니다.

> **Note**
>
> 유일한 멤버 기능이 데이터 기여이고 쿼리 컴퓨팅 비용을 지불하는 멤버가 아닌 경우 분석 탭이 콘솔에 표시되지 않습니다.

최근 쿼리 보려는 경우

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업을 선택합니다.
4. 분석 탭의 분석에서 드롭다운에서 모든 쿼리를 선택하고 지난 90일 동안 실행된 쿼리를 확인합니다.
5. 최근 쿼리를 상태별로 정렬하려면 모든 상태 드롭다운 목록에서 상태를 선택합니다.

   상태는 제출됨, 시작됨, 취소됨, 성공, 실패, 시간 초과입니다.

쿼리 세부 정보 보기

쿼리를 실행할 수 있는 구성원 또는 결과를 받을 수 있는 구성원으로 쿼리 세부 정보를 볼 수 있습니다.

쿼리 세부 정보를 보려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업을 선택합니다.
4. 쿼리 탭에서 다음 중 한 가지를 수행합니다.
   - 보려는 특정 쿼리에 대한 옵션 버튼을 선택한 다음 세부정보 보기를 선택합니다.
   - 보호된 쿼리 ID를 선택합니다.
5. 쿼리 세부정보 페이지에서,
   - 쿼리를 실행할 수 있는 구성원인 경우 쿼리 세부 정보, SQL 텍스트 및 결과를 확인하세요.

     결과를 받을 수 있는 구성원에게 쿼리 결과가 전달되었음을 확인하는 메시지가 표시됩니다.

- 결과를 받을 수 있는 회원인 경우 쿼리 세부 정보 및 결과를 확인하세요.

PySpark 작업 실행

쿼리할 수 있는 구성원은 승인된 PySpark 분석 템플릿을 사용하여 구성된 테이블에서 PySpark 작업을 실행할 수 있습니다.

사전 조건

PySpark 작업을 실행하기 전에 다음이 있어야 합니다.

- AWS Clean Rooms 공동 작업 중인 활성 멤버십
- 공동 작업에서 하나 이상의 분석 템플릿에 대한 액세스
- 공동 작업에서 하나 이상의 구성된 테이블에 대한 액세스
- 지정된 S3 버킷에 PySpark 작업의 결과를 쓸 수 있는 권한

  필요한 서비스 역할 생성에 대한 자세한 내용은 섹션을 참조하세요서비스 역할을 생성하여 PySpark 작업의 결과 작성.
- 컴퓨팅 비용을 지불할 책임이 있는 구성원이 활성 구성원으로 공동 작업에 참여했습니다.

`StartProtectedJob` API 작업을 직접 호출하거나SDK를 AWS Clean Rooms 사용하여 데이터를 쿼리하거나 쿼리를 보는 방법에 대한 자세한 내용은 AWS Clean Rooms API 참조를 참조하세요. AWS SDKs

작업 로깅에 대한 자세한 내용은 섹션을 참조하세요의 분석 로깅 AWS Clean Rooms.

작업 결과 수신에 대한 자세한 내용은 섹션을 참조하세요분석 결과 수신 및 사용.

다음 주제에서는 AWS Clean Rooms 콘솔을 사용하여 공동 작업의 구성된 테이블에서 PySpark 작업을 실행하는 방법을 설명합니다.

주제

- PySpark 분석 템플릿을 사용하여 구성된 테이블에서 PySpark 작업 실행
- 최근 작업 보기
- 작업 세부 정보 보기

PySpark 분석 템플릿을 사용하여 구성된 테이블에서 PySpark 작업 실행

이 절차에서는 콘솔에서 AWS Clean Rooms PySpark 분석 템플릿을 사용하여 사용자 지정 분석 규칙을 사용하여 구성된 테이블을 분석하는 방법을 보여줍니다.

Pyspark 분석 템플릿을 사용하여 구성된 테이블에서 PySpark 작업을 실행하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 멤버 기능 상태가 작업 실행인 공동 작업을 선택합니다.
4. 분석 탭의 테이블 섹션에서 테이블과 관련 분석 규칙 유형(사용자 지정 분석 규칙)을 확인합니다.

   > **Note**
   >
   > 목록에 예상한 테이블이 보이지 않는 경우 다음과 같은 이유가 있을 수 있습니다.
   >
   > - 테이블이 연결되지 않았습니다.
   > - 테이블에는 분석 규칙이 구성되어 있지 않습니다.

5. 분석 섹션의 분석 템플릿 실행을 선택한 다음 드롭다운 목록에서 PySpark 분석 템플릿을 선택합니다.

   PySpark 분석 템플릿의 파라미터는 정의에 자동으로 채워집니다.

6. Run(실행)을 선택합니다.

   > **Note**
   >
   > 결과를 받을 수 있는 구성원이 작업 결과 설정을 구성하지 않은 경우 작업을 실행할 수 없습니다.

7. 파라미터를 계속 조정하고 작업을 다시 실행하거나 + 버튼을 선택하여 새 탭에서 새 작업을 시작합니다.

최근 작업 보기

분석 탭에서 지난 90일 동안 실행된 작업을 볼 수 있습니다.

> **Note**
>
> 유일한 멤버 기능이 데이터 기여이고 멤버가 작업 컴퓨팅 비용을 지불하지 않는 경우 분석 탭이 콘솔에 표시되지 않습니다.

최근 작업을 보려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업을 선택합니다.
4. 분석 탭의 분석에서 드롭다운에서 모든 작업을 선택하고 지난 90일 동안 실행된 작업을 확인합니다.
5. 최근 작업을 상태별로 정렬하려면 모든 상태 드롭다운 목록에서 상태를 선택합니다.

   상태는 제출됨, 시작됨, 취소됨, 성공, 실패, 시간 초과입니다.

작업 세부 정보 보기

작업을 실행할 수 있는 멤버 또는 결과를 받을 수 있는 멤버로 작업 세부 정보를 볼 수 있습니다.

작업의 세부 정보를 보려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업을 선택합니다.
4. 분석 탭의 분석에서 드롭다운에서 모든 작업을 선택한 후 다음 중 하나를 수행합니다.
   - 보려는 특정 작업에 대한 옵션 버튼을 선택한 다음 세부 정보 보기를 선택합니다.
   - 보호된 작업 ID를 선택합니다.
5. 작업 세부 정보 페이지에서
   - 작업을 실행할 수 있는 구성원인 경우 작업 세부 정보, 작업 및 결과를 확인합니다.

     결과를 받을 수 있는 구성원에게 작업 결과가 전달되었음을 확인하는 메시지가 표시됩니다.

- 결과를 받을 수 있는 구성원인 경우 작업 세부 정보 및 결과를 확인합니다.

분석 결과 수신 및 사용

결과를 받을 수 있는 구성원은 AWS Clean Rooms 콘솔 또는 공동 작업에 참여할 때 지정한 Amazon S3 버킷에서 쿼리 결과를 검토합니다.

> **Note**
>
> 암호화된 데이터 테이블의 경우에만 결과를 수신할 수 있는 구성원이 C3R 암호화 클라이언트를 암호 해독 모드로 실행하여 쿼리 결과를 복호화합니다.
> Spark 분석 엔진을 사용하는 경우 Amazon S3의 결과 대상은 데이터 소스와 동일한 S3 버킷 내에 있을 수 없습니다.

다음 주제에서는 AWS Clean Rooms 콘솔을 사용하여 분석 결과를 수신하는 방법을 설명합니다.

주제

- 쿼리 결과 수신
- 작업 결과 수신
- 쿼리 결과 설정의 기본값 편집
- 작업 결과 설정의 기본값 편집
- 다른에서 쿼리 출력 사용 AWS 서비스

AWS Clean Rooms API를 직접 호출하거나 AWS SDKs. AWS Clean Rooms

쿼리 로깅에 대한 자세한 내용은 의 분석 로깅 AWS Clean Rooms 섹션을 참조하세요.

> **Note**
>
> 암호화된 데이터 테이블에서 쿼리를 실행하면 암호화된열의 결과가 암호화됩니다.

쿼리 결과 수신

> **Note**
>
> Spark 분석 엔진을 사용하는 경우 Amazon S3의 결과 대상은 데이터 소스와 동일한 S3 버킷 내에 있을 수 없습니다.

쿼리 결과는 AWS Clean Rooms 콘솔의 분석 탭에 있는 결과 설정 기본값 섹션에 있습니다.

쿼리 결과를 수신하려는 경우

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 구성원 능력 상태가 결과 받기인 공동 작업을 선택합니다.
4. 쿼리 결과를 직접 수신하려면 분석 탭 AWS Clean Rooms의 분석에서 드롭다운의 모든 쿼리를 선택한 다음 보호된 쿼리 ID 열에서 쿼리를 선택합니다.
5. 쿼리 세부정보 페이지의 결과에서 다음 중 하나를 수행합니다:

다음을 수행하려는 경우...	그런 다음을 선택합니다...
결과를 복사합니다.	복사
결과를 다운로드합니다.	다운로드 **Note** 기본적으로 다운로드된 파일의 이름은 AWS Clean Rooms에서 쿼리를 실행할 때 표시된 해당 `Query id`입니다.
Amazon S3에서 결과를 볼 수 있습니다.	아마존 S3에서 보기 Amazon S3 콘솔이 별도의 탭에서 열립니다.

6. 암호화된 데이터를 사용하는 경우 이제 데이터 테이블을 해독할 수 있습니다.

   자세한 내용은 C3R 암호화 클라이언트를 사용한 데이터 테이블 복호화 단원을 참조하십시오.

작업 결과 수신

> **Note**
>
> Spark 분석 엔진을 사용하는 경우 Amazon S3의 결과 대상은 데이터 소스와 동일한 S3 버킷 내에 있을 수 없습니다.

작업 결과는 AWS Clean Rooms 콘솔의 분석 탭에 있는 결과 설정 기본값 섹션에 있습니다.

작업 결과를 수신하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 구성원 능력 상태가 결과 받기인 공동 작업을 선택합니다.
4. 작업 결과를 직접 수신하려면 분석 탭 AWS Clean Rooms의 분석에서 드롭다운의 모든 작업을 선택한 다음 보호된 작업 ID 열에서 작업을 선택합니다.
5. 작업 세부 정보 페이지의 결과에서 작업 ID를 복사합니다.

   분석 탭으로 돌아가서 결과 설정 기본값을 확장합니다.

   결과 대상에서 링크를 선택하여 Amazon S3에서 결과를 봅니다.

   Amazon S3 콘솔이 별도의 탭에서 열립니다.

   Amazon S3의 검색 창에 작업 ID를 붙여넣고 Enter 키를 누릅니다.

   결과가 포함된 폴더가 나타납니다. 폴더를 선택하여 작업 결과를 봅니다.

쿼리 결과 설정의 기본값 편집

> **Note**
>
> Spark 분석 엔진을 사용하는 경우 Amazon S3의 결과 대상은 데이터 소스와 동일한 S3 버킷 내에 있을 수 없습니다.

결과를 받을 수 있는 구성원은 AWS Clean Rooms 콘솔에서 쿼리 결과 설정의 기본값을 편집할 수 있습니다.

쿼리 결과 설정의 기본값을 편집하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 구성원 능력 상태가 결과 받기인 공동 작업을 선택합니다.
4. 분석 탭의 결과 설정 기본값에서 편집을 선택합니다.
5. 결과 설정 기본값 편집 페이지에서 필요에 따라 다음 중 하나를 수정합니다.

   a. 쿼리 결과에서 Amazon S3의 결과 대상, 결과 형식 또는 결과 파일을 수정합니다.

   b. (선택 사항) 서비스 액세스에서 S3 대상에 최대 24시간이 걸리는 쿼리를 전송하려면 완료에 최대 24시간이 걸리는 쿼리를 지원하는 서비스 역할 추가 확인란을 선택합니다.

      완료하는 데 최대 24시간이 걸리는 대규모 쿼리는 S3 대상으로 전송됩니다.

      확인란을 선택하지 않으면 12시간 이내에 완료된 쿼리만 S3 위치로 전송됩니다.

      - 새 서비스 역할 생성 및 사용 또는 기존 서비스 역할 사용을 선택하여 서비스 액세스 권한을 지정합니다.

        Create and use a new service role

        - AWS Clean Rooms 는이 테이블에 필요한 정책을 사용하여 서비스 역할을 생성합니다.
        - 기본 서비스 역할 이름은 `cleanrooms-query-receiver-<timestamp>`입니다
        - 역할을 생성하고 정책을 연결할 수 있는 권한이 있어야 합니다.

Use an existing service role

1. 드롭다운 목록에서 기존 서비스 역할 이름을 선택합니다.

   역할을 나열할 권한이 있는 경우 역할 목록이 표시됩니다.

   역할을 나열할 수 있는 권한이 없는 경우 사용하려는 역할의 Amazon 리소스 이름(ARN)을 입력할 수 있습니다.

2. IAM에서 보기 외부 링크를 선택하여 서비스 역할을 확인합니다.

   기존 서비스 역할이 없는 경우 기존 서비스 역할 사용 옵션을 사용할 수 없습니다.

   기본적으로 AWS Clean Rooms 는 필요한 권한을 추가하기 위해 기존 역할 정책을 업데이트하려고 시도하지 않습니다.

> **Note**
>
> - AWS Clean Rooms 에는 분석 규칙에 따라 쿼리할 수 있는 권한이 필요합니다. 권한에 대한 자세한 내용은 섹션을 AWS Clean Rooms참조하세요AWS 에 대한 관리형 정책 AWS Clean Rooms.
> - 역할에 충분한 권한이 없는 경우 역할에 충분한 권한이 없다는 오류 메시지가 AWS Clean Rooms표시됩니다 AWS Clean Rooms. 계속하기 전에 역할 정책을 추가해야 합니다.
> - 역할 정책을 수정할 수 없는 경우에서 서비스 역할에 대한 정책을 찾을 AWS Clean Rooms 수 없다는 오류 메시지가 표시됩니다.

6. 변경 사항 저장을 선택합니다.
7. 업데이트된 쿼리 결과 설정은 공동 작업 세부 정보 페이지에 표시됩니다.

작업 결과 설정의 기본값 편집

> **Note**
>
> Spark 분석 엔진을 사용하는 경우 Amazon S3의 결과 대상은 데이터 소스와 동일한 S3 버킷 내에 있을 수 없습니다.

결과를 받을 수 있는 구성원은 AWS Clean Rooms 콘솔에서 작업 결과 설정의 기본값을 편집할 수 있습니다.

작업 결과 설정의 기본값을 편집하려면

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 구성원 능력 상태가 결과 받기인 공동 작업을 선택합니다.
4. 분석 탭의 결과 설정 기본값에서 편집을 선택합니다.
5. 결과 설정 편집 기본값 페이지에서 필요에 따라 다음 중 하나를 수정합니다.
    a. 작업 결과에서 Amazon S3의 결과 대상을 수정합니다.
    b. 서비스 액세스에서 기존 서비스 역할 이름을 수정합니다.
6. 변경 사항 저장을 선택합니다.
7. 업데이트된 작업 결과 설정은 공동 작업 세부 정보 페이지에 표시됩니다.

다른에서 쿼리 출력 사용 AWS 서비스

SQL 쿼리 출력은 Clean Rooms ML 모델의 시드 데이터에 사용할 수 있습니다. 자세한 내용은 AWS Clean Rooms ML을 참조하세요.

의 쿼리 출력 AWS Clean Rooms 은 콘솔(콘솔을 사용하여 쿼리를 실행하는 경우)에서 사용할 수 있으며 지정된 Amazon S3 버킷에 다운로드됩니다. 여기에서 해당 서비스가 Amazon S3의 데이터를 사용하는 방식에 따라 Amazon QuickSight 및 Amazon SageMaker AI AWS 서비스와 같은 다른에서 쿼리 출력을 사용할 수 있습니다. Amazon S3

Amazon QuickSight에 대한 자세한 내용은 Amazon QuickSight 설명서를 참조하세요.

Amazon SageMaker AI에 대한 자세한 내용은 Amazon SageMaker AI 설명서를 참조하세요.

훈련 데이터 공급자로 AWS Clean Rooms ML 모델 생성

유사 모델은 훈련 데이터 공급자의 데이터 모델로, 시드 데이터 공급자는 이를 통해 시드 데이터와 가장 유사한 훈련 데이터 공급자 데이터 세그먼트를 만들 수 있습니다. 공동 작업에 사용할 수 있는 유사 모델을 만들려면 훈련 데이터를 가져와서 유사 모델을 만들고 유사 모델을 구성한 다음 이를 공동 작업에 연결해야 합니다.

유사 모델을 사용하려면 훈련 데이터 공급자와 시드 데이터 공급자라는 두 당사자가 순차적으로에서 작업 AWS Clean Rooms 하여 데이터를 공동 작업으로 가져와야 합니다. 훈련 데이터 공급자가 먼저 완료해야 하는 워크플로는 다음과 같습니다.

1. 훈련 데이터 공급자의 데이터는 사용자-항목 상호 작용의 AWS Glue 데이터 카탈로그 테이블에 저장되어야 합니다. 훈련 데이터에는 최소한 사용자 ID 열, 상호 작용 ID 열 및 타임스탬프 열이 포함되어야 합니다.
2. 훈련 데이터 공급자는 훈련 데이터를에 등록합니다 AWS Clean Rooms.
3. 훈련 데이터 공급자는 여러 시드 데이터 공급자와 공유할 수 있는 유사 모델을 생성합니다. 유사 모델은 신경망이며, 훈련하는 데 최대 24시간이 걸릴 수 있습니다. 자동으로 재훈련되지 않으므로 매주 모델을 재훈련하는 것이 좋습니다.
4. 훈련 데이터 공급자는 관련성 지표 공유 여부 및 출력 세그먼트의 Amazon S3 위치를 포함하여 유사 모델을 구성합니다. 훈련 데이터 공급자는 단일 유사 모델에서 구성된 유사 모델을 여러 개 생성할 수 있습니다.
5. 훈련 데이터 공급자는 구성된 대상 모델을 시드 데이터 공급자와 공유하는 공동 작업에 연결합니다.

훈련 데이터 공급자가 ML 모델 생성을 완료한 후 시드 데이터 공급자는 유사 세그먼트를 생성하고 내보낼 수 있습니다.

주제

- 학습 데이터 가져오기
- 유사 모델 생성
- 유사 모델 구성
- 구성된 유사 모델 연결
- 구성된 유사 모델 업데이트

학습 데이터 가져오기

> **Note**
>
> Amazon S3에 데이터가 저장된 Clean Rooms ML 유사 모델에서를 사용하기 위한 훈련 데이터 세트만 제공할 수 있습니다. 그러나 지원되는 모든 데이터 소스에 저장된 데이터에서 실행되는 SQL을 사용하여 유사 모델의 시드 데이터를 제공할 수 있습니다.

유사 모델을 생성하기 전에 훈련 데이터가 포함된 AWS Glue 테이블을 지정해야 합니다. Clean Rooms ML은 이 데이터의 복사본을 저장하지 않고 데이터에 액세스할 수 있는 메타데이터만 저장합니다.

에서 훈련 데이터를 가져오려면 AWS Clean Rooms

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 AWS ML 모델을 선택합니다.
3. 훈련 데이터 세트 탭에서 훈련 데이터 세트 생성을 선택합니다.
4. 훈련 데이터 세트 생성 페이지의 훈련 데이터 세트 세부 정보 에 이름과 설명(선택 사항)을 입력합니다.
5. 드롭다운 목록에서 구성하려는 데이터베이스 및 테이블을 선택하여 훈련 데이터 소스를 선택합니다.

   > **Note**
   >
   > 테이블이 올바른지 확인하려면 다음 중 하나를 수행합니다.
   >
   > - 보기를 AWS Glue 선택합니다.
   > - 스키마를 보려면 스키마 보기를 켜세요.

6. 훈련 세부 정보의 드롭다운 목록에서 사용자 식별자 열, 항목 식별자 열, 타임스탬프 열을 선택합니다. 훈련 데이터에는 이 세 개의 열이 포함되어야 합니다. 훈련 데이터에 포함하려는 다른 열을 선택할 수도 있습니다.

   타임스탬프 열의 데이터는 Unix epoch 시간(초) 형식이어야 합니다.
7. (선택 사항) 훈련할 추가 열이 있는 경우 드롭다운 목록에서 열 이름 및 유형을 선택합니다.

8. 서비스 액세스에서 데이터에 액세스할 수 있는 서비스 역할을 지정하고 데이터가 암호화된 경우 KMS 키를 제공해야 합니다. 새 서비스 역할 생성 및 사용을 선택하면 Clean Rooms ML에서 자동으로 서비스 역할을 생성하고 필요한 권한 정책을 추가합니다. 사용하려는 특정 서비스 역할이 있는 경우 기존 서비스 역할 사용을 선택하고 서비스 역할 이름 필드에 입력합니다.

   데이터가 암호화된 경우 AWS KMS key 필드에 KMS 키를 입력하거나 AWS KMS key생성을 클릭하여 새 KMS 키를 생성합니다.
9. 훈련 데이터 세트의 태그를 사용하려면 새 태그 추가를 선택한 다음 키와 값 쌍을 입력합니다.
10. 훈련 데이터 세트 생성을 선택합니다.

해당 API 작업에서 CreateTrainingDataset를 확인하세요.

유사 모델 생성

훈련 데이터 세트를 만들었으면 유사 모델을 만들 준비가 된 것입니다. 단일 훈련 데이터 세트에서 유사 모델을 여러 개 만들 수 있습니다.

에서 기본 데이터베이스를 생성 AWS Glue Data Catalog 하거나 제공된 역할에 `glue:createDatabase` 권한을 포함해야 합니다.

에서 유사 모델을 생성하려면 AWS Clean Rooms

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 AWS ML 모델을 선택합니다.
3. 유사 모델 탭에서 유사 모델 생성을 선택합니다.
4. 유사 모델 생성 페이지의 유사 모델 내보내기 세부 정보에 이름과 설명(선택 사항)을 입력합니다.

   a. 드롭다운 목록에서 모델링하고 싶은 훈련 데이터 세트를 선택합니다.

   > **Note**
   >
   > 올바른 훈련 데이터 세트인지 확인하려면 훈련 데이터세트 세부 정보 보기를 켜서 세부 정보를 확인합니다.
   > 훈련 데이터 세트를 새로 생성하려면 훈련 데이터 세트 생성을 선택합니다.

   b. (선택 사항) 훈련 기간을 입력합니다.

5. 유사 모델에 대한 사용자 지정 암호화 설정을 사용하려면 암호화 설정 사용자 지정을 선택한 다음 KMS 키를 입력합니다.
6. 유사 모델의 태그를 사용하려면 새 태그 추가를 선택한 다음 키와 값 쌍을 입력합니다.
7. 유사 모델 생성을 선택합니다.

> **Note**
>
> 모델 훈련에는 몇 시간에서 길게는 이틀이 걸릴 수 있습니다.

해당 API 작업에서 CreateAudienceModel을 확인하세요.

유사 모델 구성

유사 모델이 생성되면 공동 작업에 유사 모델을 사용하기 위해 구성할 준비가 된 것입니다. 단일 유사 모델에서 구성된 유사 모델을 여러 개 생성할 수 있습니다.

에서 유사 모델을 구성하려면 AWS Clean Rooms

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 AWS ML 모델을 선택합니다.
3. 유사 모델 구성 탭에서 유사 모델 생성을 선택합니다.
4. 유사 모델 구성 페이지의 구성된 유사 모델 세부 정보에 이름과 설명(선택 사항)을 입력합니다.

   a. 드롭다운 목록에서 구성하고 싶은 유사 모델을 선택합니다.

   > **Note**
   >
   > 올바른 유사 모델인지 확인하려면 유사 모델 세부 정보 보기를 켜서 세부 정보를 확인합니다.
   > 유사 모델을 새로 생성하려면 유사 모델 생성을 선택합니다.

   b. 원하는 최소 매칭 시드 크기를 선택합니다. 이는 시드 데이터 공급자 데이터에 있는 사용자 중 훈련 데이터의 사용자와 겹치는 최소 사용자 수입니다. 이 값은 0보다 커야 합니다.
5. 다른 구성원과 공유하기 위한 지표에서 공동 작업의 시드 데이터 공급자가 관련성 점수를 포함한 모델 지표를 수신하도록 할지 여부를 선택합니다.

6. 유사 세그먼트 대상 위치에 유사 세그먼트를 내보낼 Amazon S3 버킷을 입력합니다. 이 버킷은 다른 리소스와 동일한 리전에 있어야 합니다.
7. 서비스 액세스에서 이 테이블에 액세스하는 데 사용할 기존 서비스 역할 이름을 선택합니다.
8. 고급 빈 크기 구성에서 대상 크기 유형을 절대 수 또는 백분율로 지정합니다.
9. 구성된 테이블 리소스에 대해 태그를 활성화하려면 새 태그 추가를 선택한 다음 키와 값 쌍을 입력합니다.
10. 유사 모델 구성을 선택합니다.

해당 API 작업에서 CreateConfiguredAudienceModel을 확인하세요.

구성된 유사 모델 연결

유사 모델이 구성되면 해당 모델을 공동 작업에 연결할 수 있습니다.

에서 구성된 유사 모델을 연결하려면 AWS Clean Rooms

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 활성 멤버십 포함 탭에서 공동 작업을 선택합니다.
4. ML 모델 탭의 Ready-to-use 가능한 유사 모델에서 유사 모델 연결을 선택합니다.
5. 구성된 유사 모델 연결 페이지의 구성된 유사 모델 연결 세부 정보에서 다음을 수행합니다.
   a. 연결된 구성 대상 모델의 이름을 입력합니다.
   b. 테이블의 설명을 입력합니다.

      설명을 입력하면 비슷한 이름을 가진 다른 연결된 구성 대상 모델을 구분하는 데 도움이 됩니다.
6. 구성된 유사 모델의 드롭다운 목록에서 구성된 유사 모델을 선택합니다.
7. 연결을 선택합니다.

해당 API 작업에서 CreateConfiguredAudienceModelAssociation을 확인하세요.

구성된 유사 모델 업데이트

유사 모델을 구성한 후 이를 업데이트하여 이름, 공유할 지표 또는 출력 Amazon S3 위치와 같은 정보를 변경할 수 있습니다.

에서 연결된 구성된 유사 모델을 업데이트하려면 AWS Clean Rooms

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 AWS ML 모델을 선택합니다.
3. 구성된 유사 모델 탭의 바로 사용할 수 있는 유사 모델에서 구성된 유사 모델을 선택하고 편집을 선택합니다.
4. 편집 페이지의 구성된 유사 모델 연결 세부 정보에서 다음을 수행합니다.
    a. 이름과 설명(선택 사항)을 업데이트합니다.
    b. 드롭다운 목록에서 구성하려는 유사 모델을 선택합니다.
    c. 원하는 최소 매칭 시드 크기를 선택합니다. 이는 시드 데이터 공급자 데이터에 있는 사용자 중 훈련 데이터의 사용자와 겹치는 최소 사용자 수입니다. 이 값은 0보다 커야 합니다.
5. 다른 구성원과 공유하기 위한 지표에서 공동 작업의 시드 데이터 공급자가 관련성 점수를 포함한 모델 지표를 수신하도록 할지 여부를 선택합니다.
6. 유사 세그먼트 대상 위치에는 유사 세그먼트를 내보내는 Amazon S3 버킷을 입력합니다. 이 버킷은 다른 리소스와 동일한 리전에 있어야 합니다.
7. 서비스 액세스에서 이 테이블에 액세스하는 데 사용할 기존 서비스 역할 이름을 선택합니다.
8. 고급 빈 크기 구성에서 대상 빈 크기를 구성할 방법을 선택합니다.
9. 변경 사항 저장을 선택합니다.

해당 API 작업은 UpdateConfiguredAudienceModel을 참조하세요.

AWS Clean Rooms ML 모델을 시드 데이터 공급자로 생성

훈련 데이터 공급자가 ML 모델 생성을 완료한 후 시드 데이터 공급자는 유사 세그먼트를 생성하고 내보낼 수 있습니다. 유사 세그먼트는 시드 데이터와 가장 유사한 훈련 데이터의 하위 집합입니다.

다음은 시드 데이터 공급자가 완료해야 하는 워크플로입니다.

1. 시드 데이터 공급자의 데이터는 Amazon S3 버킷에 저장하거나 쿼리 결과에서 가져올 수 있습니다.
2. 시드 데이터 공급자는 훈련 데이터 공급자와 공유하는 공동 작업을 엽니다.
3. 시드 데이터 공급자는 공동 작업 페이지의 Clean Rooms ML 탭에서 유사 세그먼트를 만듭니다.
4. 시드 데이터 공급자는 관련성 지표가 공유된 경우 이를 평가하고 유사 세그먼트를 내보내 AWS Clean Rooms외부에서 사용할 수 있습니다.

주제

- 유사 세그먼트 생성
- 유사 세그먼트 내보내기

유사 세그먼트 생성

> **Note**
>
> Amazon S3에 데이터가 저장된 Clean Rooms ML 유사 모델에서를 사용하기 위한 훈련 데이터 세트만 제공할 수 있습니다. 그러나 지원되는 모든 데이터 소스에 저장된 데이터에서 실행되는 SQL을 사용하여 유사 모델의 시드 데이터를 제공할 수 있습니다.

유사 세그먼트는 시드 데이터와 가장 유사한 훈련 데이터의 하위 집합입니다.

에서 유사 세그먼트를 생성하려면 AWS Clean Rooms

1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 활성 멤버십 포함 탭에서 공동 작업을 선택합니다.
4. ML 모델 탭에서 유사 세그먼트 생성을 선택합니다.

5. 유사 세그먼트 생성 페이지의 연결된 구성된 유사 모델에서이 유사 세그먼트에 사용할 연결된 구성된 유사 모델을 선택합니다.

6. 유사 세그먼트 세부 정보에 이름과 설명(선택 사항)을 입력합니다.

7. 시드 프로파일에서 옵션을 선택한 다음 권장 작업을 수행하여 시드 방식을 선택합니다.

옵션	권장 조치
Amazon S3 경로	1. Amazon S3 위치를 선택합니다. 2. (선택 사항) 출력에 시드 프로파일 포함을 선택합니다.
SQL 쿼리	SQL 쿼리를 작성하고 결과를 시드 데이터로 사용합니다.
분석 템플릿	드롭다운 목록에서 분석 템플릿을 선택하고 분석 템플릿에서 생성된 결과를 사용합니다.

8. 이 데이터 채널을 생성할 때 사용할 작업자 유형 및 작업자 수를 선택합니다.
9. 서비스 액세스에서 이 테이블에 액세스하는 데 사용할 기존 서비스 역할 이름을 선택합니다.
10. 훈련 데이터 세트의 태그를 사용하려면 새 태그 추가를 선택한 다음 키와 값 쌍을 입력합니다.
11. 유사 세그먼트 생성을 선택합니다.

해당 API 작업에서 StartAudienceGenerationJob을 확인하세요.

유사 세그먼트 내보내기

유사 세그먼트를 생성한 후 데이터를 Amazon S3 버킷으로 내보낼 수 있습니다.

에서 유사 세그먼트를 내보내려면 AWS Clean Rooms

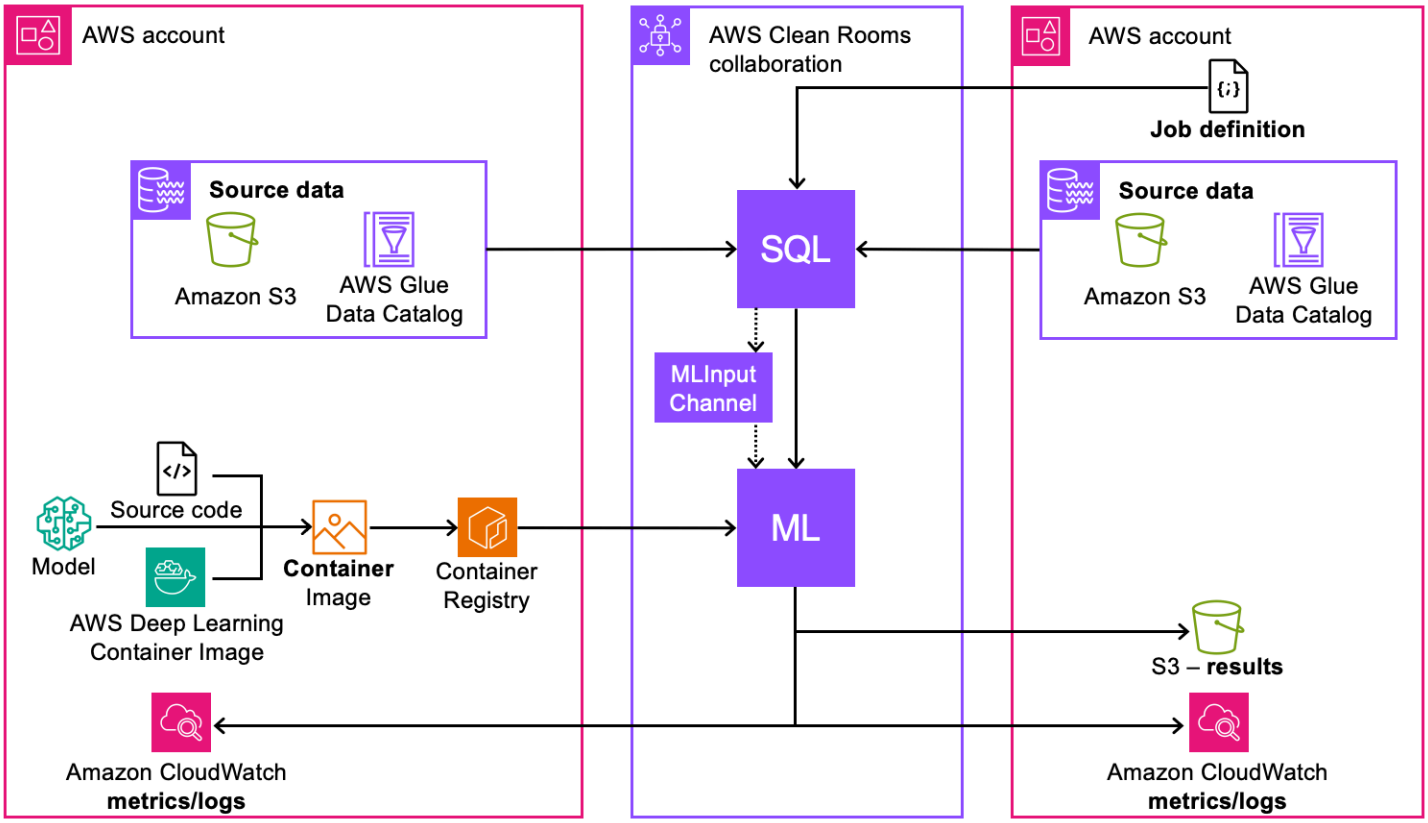
1. 에 로그인 AWS Management Console 하고를 사용하여 AWS Clean Rooms 콘솔을 엽니다 AWS 계정 (아직 열지 않은 경우).
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 활성 멤버십 포함 탭에서 공동 작업을 선택합니다.
4. ML 모델 탭에서 유사 세그먼트를 선택하고 내보내기를 선택합니다.

5. 유사 모델 내보내기의 유사 모델 내보내기 세부 정보에서 이름 및 필요에 따라 설명을 입력합니다.
6. 세그먼트 크기에서 내보낸 세그먼트에 사용할 크기를 선택합니다.
7. 내보내기를 선택합니다.

해당 API 작업에서 StartAudienceExportJob을 확인하세요.

AWS Clean Rooms ML의 사용자 지정 모델링

기술적 관점에서 다음 다이어그램은 AWS Clean Rooms ML에서 사용자 지정 ML 모델링이 작동하는 방식을 설명합니다.



다음은 Clean Rooms ML에서 사용자 지정 ML 모델링이 작동하는 방식입니다.

1. 데이터 소스 구성
   - 소스 데이터는 Amazon S3 카탈로그, AWS Glue Data Catalog또는 Snowflake에 저장할 수 있습니다.
   - AWS Glue Data Catalog 는 구성 및 카탈로그 작성에 사용됩니다.
   - 여러의 데이터를 동일한 공동 작업 내에서 사용할 AWS 계정 수 있습니다.
2. SQL 쿼리 및 데이터 처리
   - SQL 쿼리는 소스 데이터에 액세스하고 처리하는 데 사용됩니다.
   - 쿼리는 AWS Clean Rooms 공동 작업 경계 내에서 실행됩니다.
   - 모델 훈련을 위해 ML 입력 채널로 데이터 피드 처리
3. ML 모델 개발

- AWS 딥 러닝 컨테이너 이미지를 사용하여 모델의 소스 코드를 개발할 수 있습니다.
- 사용자 지정 컨테이너 이미지를 생성하여 Amazon Elastic Container Registry에 저장해야 합니다.

4. 인프라 구성 요소
   - Amazon Elastic Container Registry는 ML 모델 컨테이너를 저장하고 관리합니다.
   - ML 처리는 보안 AWS Clean Rooms 협업 환경 내에서 이루어집니다.
5. 모니터링 및 로깅
   - Amazon CloudWatch는 두 공동 작업에 지표와 로그를 제공합니다.
   - 공동 작업에 AWS 계정 관련된에서 모니터링을 사용할 수 있습니다.
   - 관련 당사자가 성능 지표 및 운영 로그에 액세스할 수 있음
6. 결과 관리
   - 결과에 대한 액세스는 공동 작업 권한에 따라 제어됩니다.

시작하기 전에 사용자 지정 ML 모델링 사전 조건 및 훈련 컨테이너에 대한 모델 작성 지침에서 자세한 내용을 참조하세요.

주제

- AWS Clean Rooms ML에서 공동 작업 생성 및 참여
- AWS Clean Rooms ML에서 훈련 데이터 기여
- AWS Clean Rooms ML에서 모델 알고리즘 구성
- AWS Clean Rooms ML에서 구성된 모델 알고리즘 연결
- AWS Clean Rooms ML에서 ML 입력 채널 생성
- AWS Clean Rooms ML에서 훈련된 모델 생성
- AWS Clean Rooms ML에서 모델 아티팩트 내보내기
- AWS Clean Rooms ML에서 훈련된 모델에 대한 추론 실행

AWS Clean Rooms ML에서 공동 작업 생성 및 참여

공동 작업 생성자는 공동 작업을 생성하고, 구성원을 초대하고, 역할을 할당할 책임이 있습니다. 초대된 구성원은 공동 작업에 참여하여 결과 설정, 훈련된 모델 아티팩트 대상 설정을 지정하고 공동 작업 설정 방식에 따라 결제 책임을 수락합니다.

기계 학습을 위한 공동 작업 생성

다음 절차에서는 기계 학습을 위한 공동 작업을 생성하고, 하나 이상의 구성원을 초대하고, 모델 훈련을 시작하고, 결과를 수신하고, 모델 아티팩트 및 지표를 포함한 훈련된 모델 결과를 수신하고, 모델 추론 결과를 수신할 수 있는 구성원을 할당하는 방법을 보여줍니다. 또한 공동 작업 생성자는 쿼리 컴퓨팅, 모델 훈련 및 모델 추론 비용을 지불할 구성원을 할당합니다.

Console

기계 학습을 위한 공동 작업을 생성하려면(콘솔)

1. 공동 작업을 생성하고 하나 이상의 구성원을 공동 작업에 참여하도록 초대합니다.
2. 쿼리 및 작업을 사용하여 분석에 대해 다음 멤버 기능을 할당합니다.
   - 모델 훈련을 시작할 구성원에게 실행 쿼리를 할당합니다.
   - 쿼리 결과를 받을 구성원에게 분석 결과 수신을 할당합니다.
3. 특별히 구축된 워크플로를 사용하여 ML 모델링에 대해 다음 멤버 기능을 할당합니다.
   - 모델 아티팩트 및 지표를 포함하여 훈련된 모델 결과를 받을 구성원에게 훈련된 모델의 수신 출력을 할당합니다.
   - 모델 추론 결과를 수신할 구성원에게 모델 추론의 수신 출력을 할당합니다.
4. 결제 구성에서 쿼리 컴퓨팅, 모델 훈련 및 모델 추론 비용을 지불할 멤버를 지정합니다. 이러한 각 비용은 동일하거나 다른 멤버에게 할당할 수 있습니다. 초대된 구성원이 결제 비용을 지불할 책임이 있는 구성원인 경우 공동 작업에 참여하기 전에 결제 책임을 수락해야 합니다.
5. 멤버십 구성에서 공동 작업 생성자는 지금 멤버십에 가입하거나 나중에 멤버십을 생성하기로 결정할 수 있습니다. 그러면 공동 작업 생성자가 ML 구성을 설정해야 합니다.
   a. 공동 작업 생성자가 결과 수신자이기도 한 경우 결과 설정 기본값에서 쿼리 결과 대상 및 형식도 지정해야 합니다.
   b. ML 구성은 Clean Rooms ML이 지표를에 게시하는 역할을 제공합니다 AWS 계정. 공동 작업 생성자가 훈련된 모델 아티팩트도 수신하는 경우 결과를 수신하는 데 사용되는 Amazon S3 버킷을 지정할 수 있습니다.
   c. ML 구성 섹션에서 ML 구성 생성을 선택한 다음 Amazon S3의 모델 출력 대상과이 위치에 액세스하는 데 필요한 서비스 액세스 역할을 지정합니다.
   d. 공동 작업 생성자가 결제 비용을 지불할 책임이 있는 구성원인 경우 공동 작업을 생성하기 전에 결제 책임을 수락해야 합니다.

API

기계 학습을 위한 공동 작업을 생성하려면(API)

1. 공동 작업을 생성하고 하나 이상의 구성원을 공동 작업에 초대합니다.
2. 공동 작업 구성원에게 다음 역할을 할당합니다.
   - CAN_QUERY - 모델 훈련 및 추론을 시작할 구성원에게 할당됩니다.
   - CAN_RECEIVE_MODEL_OUTPUT - 훈련된 모델 결과를 받을 구성원에게 할당됩니다.
   - CAN_RECEIVE_INFERENCE_OUTPUT - 모델 추론 결과를 받을 멤버에게 할당됩니다.

   공동 작업 생성자가 결과 수신자이기도 한 경우 공동 작업 생성 중에 쿼리 결과 대상 및 형식도 지정해야 합니다. 또한 서비스 역할 Amazon 리소스 이름(ARN)에 쿼리 결과 대상에 결과를 쓸 수 있는 권한을 부여합니다.
3. 쿼리 컴퓨팅, 모델 훈련 및 모델 추론 비용을 지불할 멤버를 지정합니다. 이러한 각 비용은 동일하거나 다른 멤버에게 할당할 수 있습니다. 초대된 구성원이 결제 비용을 지불할 책임이 있는 구성원인 경우 공동 작업에 참여하기 전에 결제 책임을 수락해야 합니다.
4. 다음 코드는 공동 작업을 생성하고, 쿼리를 실행하고 결과를 수신할 수 있는 멤버를 초대하고, 공동 작업 생성자를 모델 아티팩트 수신기로 지정합니다.

```
import boto3
acr_client= boto3.client('cleanrooms')

collaboration = a_acr_client.create_collaboration(
    members=[
        {
         'accountId': 'invited_member_accountId',
         'memberAbilities':["CAN_QUERY","CAN_RECEIVE_RESULTS"],
         'displayName': 'member_display_name'
        }
    ],
    name='collaboration_name',
    description=collaboration_description,
    creatorMLMemberAbilities= {
        'customMLMemberAbilities':["CAN_RECEIVE_MODEL_OUTPUT",
 "CAN_RECEIVE_INFERENCE_OUTPUT"],
    },
    creatorDisplayName='creator_display_name',
    queryLogStatus="ENABLED",
    analyticsEngine="SPARK",
    creatorPaymentConfiguration={
```

```
        "queryCompute": {
            "isResponsible": True
        },
        "machineLearning": {
            "modelTraining": {
                "isResponsible": True
            },
            "modelInference": {
                "isResponsible": True
            }
        }
    }
)

collaboration_id = collaboration['collaboration']['id']
print("collaborationId: {collaboration_id}")

member_membership = a_acr_client.create_membership(
    collaborationIdentifier = collaboration_id,
    queryLogStatus = 'ENABLED',
    paymentConfiguration={
        "queryCompute": {
            "isResponsible": True
        },
        "machineLearning": {
            "modelTraining": {
                "isResponsible": True
            },
            "modelInference": {
                "isResponsible": True
            }
        }
    }
)
```

5. 그러면 공동 작업 생성자가 ML 구성을 설정해야 합니다. ML 구성은 Clean Rooms ML이 지표와 로그를에 게시하는 역할을 제공합니다 AWS 계정. 공동 작업 생성자가 결과(모델 아티팩트 또는 추론 결과)도 수신하는 경우 결과를 수신하는 데 사용되는 Amazon S3 버킷을 지정할 수 있습니다.

```
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"
                }
            }
        }
    )
```

공동 작업 참여

공동 작업 생성자가 작업을 완료한 후 초대된 구성원은 작업을 완료해야 합니다.

Console

멤버십을 생성하고 공동 작업에 참여하려면(콘솔)

1. 초대된 구성원은 멤버십을 생성하고 공동 작업에 참여합니다.
2. 초대된 구성원이 쿼리 컴퓨팅, 모델 훈련 및 모델 추론 비용을 포함하여 비용을 지불할 책임이 있는 구성원인 경우 공동 작업에 참여하기 전에 지불 책임을 수락해야 합니다.
3. 초대된 구성원은 Clean Rooms ML이 모델 지표를에 게시하는 역할을 제공하는 ML 구성을 설정합니다 AWS 계정. 또한 훈련된 모델 아티팩트를 수신하는 구성원인 경우 훈련된 모델 아티팩트가 저장되는 Amazon S3 버킷을 제공해야 합니다.

API

멤버십을 생성하고 공동 작업에 참여하려면(API)

1. 초대된 구성원이 결과를 받을 수 있는 구성원인 경우, 초대된 구성원은 쿼리 결과 대상 및 형식을 지정합니다. 또한 서비스가 쿼리 결과 대상에 쓸 수 있도록 서비스 역할 ARN을 제공합니다.

   초대된 구성원이 쿼리 컴퓨팅, 모델 훈련 및 모델 추론 비용을 포함하여 비용을 지불할 책임이 있는 구성원인 경우 공동 작업에 참여하기 전에 지불 책임을 수락해야 합니다.

   초대된 구성원이 사용자 지정 모델링을 위한 모델 훈련 및 모델 추론 비용을 지불할 책임이 있는 구성원인 경우 공동 작업에 참여하기 전에 지불 책임을 수락해야 합니다.

다음 코드는 쿼리 로깅이 활성화된 멤버십을 생성합니다.

```
import boto3
acr_client= boto3.client('cleanrooms')

acr_client.create_membership(
    membershipIdentifier='membership_id',
    queryLogStatus='ENABLED'
)
```

2. 초대된 구성원은 Clean Rooms ML이 모델 지표를에 게시하는 역할을 제공하는 ML 구성을 설정합니다 AWS 계정. 또한 훈련된 모델 아티팩트를 수신하는 구성원인 경우 훈련된 모델 아티팩트가 저장되는 Amazon S3 버킷을 제공해야 합니다.

```
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"
                }
            }
        }
    )
```

AWS Clean Rooms ML에서 훈련 데이터 기여

공동 작업 생성자가 공동 작업을 생성하고 초대된 구성원이 참여하면 공동 작업에 훈련 데이터를 제공할 준비가 된 것입니다. 모든 구성원이 훈련 데이터를 제공할 수 있습니다.

Console

훈련 데이터를 제공하려면(콘솔)

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/cleanrooms AWS Clean Rooms 콘솔을 엽니다.

2. 왼쪽 탐색 창에서 테이블을 선택합니다.
3. 테이블 페이지에서 새 테이블 구성을 선택합니다.
4. 새 테이블 구성의 데이터 소스에서 Amazon S3, Amazon Athena 또는 Snowflake를 선택하고 데이터 소스에 따라 다음 단계를 완료합니다.

를 사용하는 경우	Then
Amazon S3	1. 드롭다운 목록에서 데이터베이스를 선택한 다음 데이터베이스에서 테이블을 선택합니다. 2. 공동 작업에서 허용되는 열의 경우 모든 열 또는 사용자 지정 목록을 선택합니다. 3. 구성된 테이블 세부 정보에이 테이블에 대한 이름 및 선택적 설명을 입력합니다. 4. 모델 지표를 보고하려면 지표의 이름과 출력 로그를 검색하여 지표를 찾을 정규식 문을 입력합니다. 5. 구성된 테이블 리소스에 대해 태그를 활성화하려면 새 태그 추가를 선택한 다음 키와 값 쌍을 입력합니다.
Amazon Athena	1. 드롭다운 목록에서 데이터베이스를 선택한 다음 데이터베이스에서 테이블을 선택합니다. 2. 공동 작업에서 허용되는 열의 경우 모든 열 또는 사용자 지정 목록을 선택합니다. 3. 구성된 테이블 세부 정보에이 테이블에 대한 이름 및 선택적 설명을 입력합니다. 4. 모델 지표를 보고하려면 지표의 이름과 출력 로그를 검색하여 지표를 찾을 정규식 문을 입력합니다. 5. 구성된 테이블 리소스에 대해 태그를 활성화하려면 새 태그 추가를 선택한 다음 키와 값 쌍을 입력합니다.

를 사용하는 경우	Then
Snowflake	1. 기존 보안 암호 ARN을 사용하거나이 테이블에 대한 새 보안 암호를 저장하여 Snowflake 보안 인증 정보를 지정합니다. 2. Snowflake 테이블 및 스키마 세부 정보에 세부 정보를 수동으로 입력하거나 세부 정보를 자동으로 가져옵니다. 3. 스키마에 열 이름을 입력하고 드롭다운 목록에서 데이터 유형을 선택합니다. 4. 구성된 테이블 리소스에 대해 태그를 활성화하려면 새 태그 추가를 선택한 다음 키와 값 쌍을 입력합니다.

5. 새 클라이언트 구성를 선택합니다.
6. 테이블 세부 정보 페이지에서 분석 규칙 구성을 선택하여이 테이블에 대한 사용자 지정 분석 규칙을 구성합니다. 사용자 지정 분석 규칙은 데이터에 대한 액세스를 제한합니다. 데이터에 대해 특정 사전 승인된 쿼리 집합을 허용하거나 특정 계정 집합이 데이터를 쿼리하도록 허용할 수 있습니다.

   a. 분석 규칙 유형에서 사용자 지정을 선택하고 생성 방법에서 안내 흐름을 선택합니다.

   b. 다음을 선택합니다.

   c. 분석 컨트롤 지정에서 각 새 분석 검토와 특정 공동 작업자의 분석 허용 중에서 선택합니다.

   d. 다음을 선택합니다.

   e. (선택 사항) 분석 결과 컨트롤 지정에서 출력에 허용되지 않는 열에 대해 출력에서 열을 제외할지 여부를 지정합니다. 없음을 선택하면 출력에서 열이 제외되지 않습니다. 사용자 지정 목록을 선택하면 출력에서 제거할 특정 열을 지정할 수 있습니다.

   f. 출력에 적용된 추가 분석의 경우 결과가 생성되기 전에 추가 분석을 허용, 거부 또는 요구할지 여부를 지정합니다.

   g. 다음을 선택합니다.

   h. (선택 사항) 차등 프라이버시 설정에서 끄기를 선택합니다.

   i. 다음을 선택합니다.

   j. 검토 및 구성 페이지의 정보를 검토한 다음 분석 규칙 구성을 선택합니다.

7. 테이블 세부 정보 페이지에서 공동 작업에 연결을 선택합니다.
8. 테이블 연결 대화 상자에서이 테이블을 연결할 공동 작업을 선택하고 공동 작업 선택을 선택합니다.
9. 테이블 연결 페이지에서 테이블 연결 세부 정보, 서비스 액세스 및 태그의 정보를 검토하고 확인합니다. 연결 테이블을 선택합니다.
10. 연결된 테이블 테이블에서 방금 연결한 테이블 옆의 라디오 버튼을 선택합니다. 작업 메뉴에서 협업 분석 규칙 그룹에서 구성을 선택합니다.
11. 공동 작업 분석 규칙 구성 페이지의 허용된 추가 분석에서 공동 작업 구성원 또는 특정 공동 작업 구성원이 추가 분석을 수행할 수 있는지 여부를 선택합니다.

    결과 전송에서 쿼리 출력에서 결과를 수신할 수 있는 멤버를 선택합니다.
12. 분석 규칙 구성을 선택합니다.

API

훈련 데이터를 제공하려면(API)

1. 사용할 수 있는 AWS Glue 테이블과 열을 AWS Clean Rooms 제공하여에서 사용할 기존 테이블을 구성합니다.

   특정 파라미터로 다음 코드를 실행합니다.

```
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. 데이터에 대한 액세스를 제한하는 사용자 지정 분석 규칙을 구성합니다. 데이터에 대해 특정 사전 승인된 쿼리 집합을 허용하거나 특정 계정 집합이 데이터를 쿼리하도록 허용할 수 있습니다.

특정 파라미터로 다음 코드를 실행합니다.

```
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"
            }
        }
    }
)
```

이 예제에서는 특정 계정이 데이터에 대한 쿼리를 실행할 수 있으며 추가 분석이 필요합니다.

3. 구성된 테이블을 공동 작업에 연결하고 AWS Glue 테이블에 서비스 액세스 역할을 제공합니다.

   특정 파라미터로 다음 코드를 실행합니다.

```
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**
>
> 이 서비스 역할에는 테이블에 대한 권한이 있습니다. 서비스 역할은 쿼리할 수 있는 구성원을 대신하여 허용된 쿼리를 실행 AWS Clean Rooms 하기 위해 에서만 수임할 수 있습니다. 공동 작업 구성원(데이터 소유자 제외)은 공동 작업의 기본 테이블에 액세스할 수

없습니다. 데이터 소유자는 차등 프라이버시를 해제하여 다른 멤버가 테이블을 쿼리할 수 있도록 할 수 있습니다.

4. 마지막으로 구성된 테이블 연결에 분석 규칙을 추가합니다.

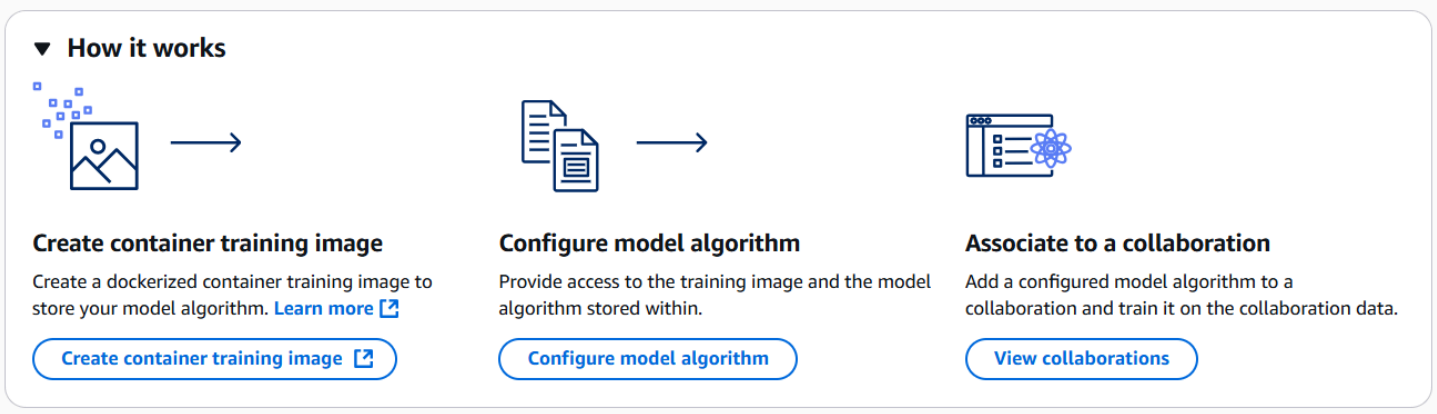
   특정 파라미터로 다음 코드를 실행합니다.

```
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']
            }
        }
    }
)
```

AWS Clean Rooms ML에서 모델 알고리즘 구성

컨테이너 훈련 이미지를 생성한 후에는 모델 알고리즘을 구성해야 합니다. 모델 알고리즘을 구성하면 공동 작업에 연결할 수 있습니다.

다음 이미지는 컨테이너 훈련 이미지를 생성한 후 공동 작업과 연결하기 전에 수행되는 단계로 모델 알고리즘을 구성하는 것을 보여줍니다.



Console

사용자 지정 ML 모델 알고리즘을 구성하려면(콘솔)

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/cleanrooms AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 사용자 지정 ML 모델을 선택합니다.
3. 사용자 지정 ML 모델 페이지에서 모델 알고리즘 구성을 선택합니다.
4. 모델 알고리즘 구성 페이지의 모델 알고리즘 세부 정보에 이름 및 선택적 설명을 입력합니다.
5. 모델 훈련을 수행하려면 훈련 이미지 ECR 컨테이너 세부 정보에서
   a. 훈련 이미지 URI 지정 확인란을 선택합니다.
   b. 드롭다운 목록에서 훈련 모델, 추론 컨테이너 또는 둘 다 있는 리포지토리를 선택합니다.
   c. 이미지를 선택합니다.
   d. (선택 사항) 진입점의 값을 입력하여 훈련 이미지에 액세스합니다.
   e. (선택 사항) 인수의 값을 입력합니다.
6. (선택 사항) 모델 지표를 보고하려면 훈련 지표에 지표 이름과 출력 로그를 검색하여 지표를 찾을 정규식 문을 입력합니다.
7. 모델 추론을 수행하려면 추론 이미지 ECR 컨테이너 세부 정보에서
   a. 추론 이미지 URI 지정 확인란을 선택합니다.
   b. 드롭다운 목록에서 리포지토리를 선택합니다.
   c. 이미지를 선택합니다.
8. 서비스 액세스에서 이 테이블에 액세스하는 데 사용할 기존 서비스 역할 이름을 선택합니다.

9. 암호화에서 암호화 설정 사용자 지정을 선택하여 자체 KMS 키 및 관련 정보를 지정합니다. 그렇지 않으면 Clean Rooms ML이 암호화를 관리합니다.
10. 태그를 활성화하려면 새 태그 추가를 선택한 다음 키 및 값 페어를 입력합니다.
11. 모델 알고리즘 구성을 선택합니다.

API

사용자 지정 ML 모델 알고리즘(API)을 구성하려면

1. SageMaker AI 호환 도커 이미지를 생성합니다. Clean Rooms ML은 SageMaker AI 호환 도커 이미지만 지원합니다.
2. SageMaker AI 호환 도커 이미지를 생성한 후 Amazon ECR을 사용하여 훈련 이미지를 생성합니다. Amazon Elastic Container Registry 사용 설명서의 지침에 따라 컨테이너 훈련 이미지를 생성합니다.
3. Clean Rooms ML에서 사용할 모델 알고리즘을 구성합니다. 다음 정보를 제공해야 합니다.
   - 모델을 훈련하고 추론을 실행하기 위한 Amazon ECR 리포지토리 링크 및 추가 인수입니다. Clean Rooms ML은 추론 컨테이너에서 배치 변환 작업 실행을 지원합니다.
   - Clean Rooms ML이 리포지토리에 액세스할 수 있도록 허용하는 서비스 액세스 역할입니다.
   - (선택 사항) 추론 컨테이너입니다. 별도의 구성된 모델 알고리즘에서 제공할 수 있지만 훈련 컨테이너와 추론 컨테이너가 모두 동일한 리소스의 일부로 관리되도록이 단계에서 제공하는 것이 좋습니다.

   특정 파라미터로 다음 코드를 실행합니다.

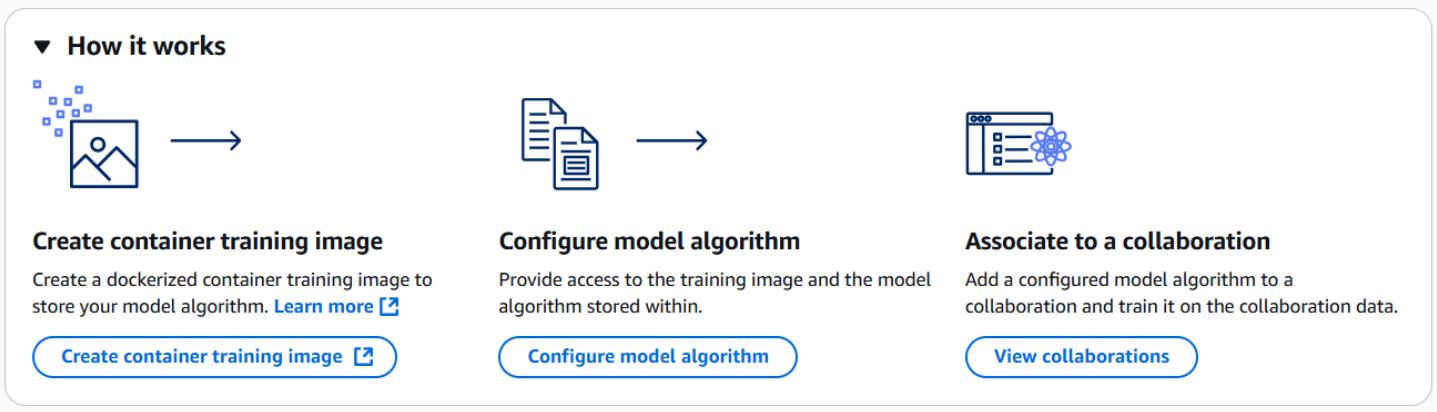
```
import boto3
acr_ml_client= boto3.client('cleanroomsml')

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'
)
```

AWS Clean Rooms ML에서 구성된 모델 알고리즘 연결

모델 알고리즘을 구성한 후에는 모델 알고리즘을 공동 작업에 연결할 준비가 된 것입니다. 모델 알고리즘을 연결하면 공동 작업의 모든 구성원이 모델 알고리즘을 사용할 수 있습니다.

다음 이미지는 컨테이너 훈련 이미지를 생성하고 모델 알고리즘을 구성한 후 구성된 모델 알고리즘을 마지막 단계로 연결하는 것을 보여줍니다.



Console

사용자 지정 ML 모델 알고리즘을 연결하려면(콘솔)

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/cleanrooms AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 사용자 지정 ML 모델을 선택합니다.
3. 사용자 지정 ML 모델 페이지에서 공동 작업에 연결할 구성된 모델 알고리즘을 선택하고 공동 작업에 연결을 클릭합니다.
4. 구성된 모델 알고리즘 연결 창에서 연결할 공동 작업을 선택합니다.
5. 공동 작업 선택을 선택합니다.

API

사용자 지정 ML 모델 알고리즘(API)을 연결하려면

특정 파라미터로 다음 코드를 실행합니다.

또한 여러 로그에 액세스할 수 있는 사용자를 정의하고, 고객이 정규식을 정의하도록 허용하며, 훈련 모델 출력 또는 추론 결과에서 내보낼 수 있는 데이터의 양을 정의하는 개인 정보 보호 정책을 제공합니다.

> **Note**
>
> 구성된 모델 알고리즘 연결은 변경할 수 없습니다.

```
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
                ]
            }
        }
    }
)
```

구성된 모델 알고리즘이 공동 작업에 연결된 후 훈련 데이터 공급자는 테이블에 공동 작업 분석 규칙을 추가해야 합니다. 이 규칙을 사용하면 구성된 모델 알고리즘 연결이 구성된 테이블에 액세스할 수 있습니다. 기여하는 모든 훈련 데이터 공급자는 다음 코드를 실행해야 합니다.

```
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**
>
> 구성된 모델 알고리즘 연결은 변경할 수 없으므로 사용자 지정 모델 구성의 처음 몇 번의 반복 `allowedAdditionalAnalyses` 중에에서 와일드카드를 사용하기 위해 모델을 허용 목록에 추가하려는 데이터 공급자를 훈련하는 것이 좋습니다. 이를 통해 모델 공급자는 다른 훈련 공급자가 업데이트된 모델 코드를 데이터로 훈련하기 전에 다시 연결할 필요 없이 코드를 반복할 수 있습니다.

AWS Clean Rooms ML에서 ML 입력 채널 생성

사전 조건:

- 에 액세스할 수 AWS 계정 있는 AWS Clean Rooms
- ML 입력 채널을 생성하려는 AWS Clean Rooms 에 설정된 공동 작업
- 공동 작업에서 데이터를 쿼리하고 ML 입력 채널을 생성할 수 있는 권한.
- (선택 사항) ML 입력 채널과 연결할 기존 모델 알고리즘 또는 새 입력 채널을 생성할 수 있는 권한
- (선택 사항) 지정된 모델에 대해 실행할 수 있는 분석 규칙이 있는 테이블입니다.
- (선택 사항) 데이터 세트를 생성하는 데 사용할 기존 SQL 쿼리 또는 분석 템플릿
- (선택 사항) 적절한 권한이 있는 기존 서비스 역할 또는 새 서비스 역할을 생성할 수 있는 권한
- (선택 사항) 자체 암호화 AWS KMS 키를 사용하려는 경우의 사용자 지정 키
- 공동 작업에서 ML 모델을 생성하고 관리할 수 있는 적절한 권한

ML 입력 채널은 특정 데이터 쿼리에서 생성되는 데이터 세트입니다. 데이터를 쿼리할 수 있는 구성원은 ML 입력 채널을 생성하여 훈련 및 추론을 위해 데이터를 준비할 수 있습니다. ML 입력 채널을 생성하면 동일한 공동 작업 내의 다양한 훈련 모델에서 데이터를 사용할 수 있습니다. 훈련 및 추론을 위해 별도의 ML 입력 채널을 생성해야 합니다.

ML 입력 채널을 생성하려면 입력 데이터를 쿼리하고 ML 입력 채널을 생성하는 데 사용되는 SQL 쿼리를 지정해야 합니다. 이 쿼리의 결과는 어떤 멤버와도 공유되지 않으며 Clean Rooms ML의 경계 내에 유지됩니다. 참조 Amazon 리소스 이름(ARN)은 다음 단계에서 모델을 훈련하거나 추론을 실행하는 데 사용됩니다.

Console

ML 입력 채널을 생성하려면(콘솔)

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/cleanrooms AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업 페이지에서 ML 입력 채널을 생성할 공동 작업을 선택합니다.
4. 공동 작업이 열리면 ML 모델 탭을 선택합니다.
5. 사용자 지정 ML 모델의 ML 입력 채널 섹션에서 ML 입력 채널 생성을 선택합니다.
6. ML 입력 채널 생성 페이지의 ML 입력 채널 세부 정보에서 다음을 수행합니다.

   a. 이름에 채널의 고유한 이름을 입력합니다.

   b. (선택 사항) 설명에 채널에 대한 설명을 입력합니다.

   c. 연결된 모델 알고리즘에서 사용할 알고리즘을 선택합니다.

      모델 알고리즘 연결을 선택하여 새 알고리즘을 추가합니다.
7. 데이터 세트에서 훈련 데이터 세트를 생성할 메서드를 선택합니다.

   - SQL 쿼리의 결과를 훈련 데이터 세트로 사용하려면 SQL 쿼리를 선택합니다.

     SQL 쿼리를 선택한 경우 SQL 쿼리 필드에 쿼리를 입력합니다.

     (선택 사항) 최근에 사용한 쿼리를 가져오려면 최근 쿼리에서 가져오기를 선택합니다.
   - 분석 템플릿을 선택하여 분석 템플릿의 결과를 훈련 데이터 세트로 사용합니다.

     분석 템플릿을 선택한 경우 원하는 분석 템플릿을 지정합니다.

   a. 연결된 테이블이 없는 경우 테이블 연결을 선택하여 지정된 모델에 대해 실행할 수 있는 분석 규칙이 있는 테이블을 추가합니다.

   b. 작업자 유형에서 사용할 작업자 유형을 선택합니다. 기본값은 CR.1X입니다.

   c. 작업자 수에서이 데이터 채널을 생성할 때 사용할 작업자 수를 선택합니다. 기본값은 16입니다.

   d. 일 단위의 데이터 보존에 데이터를 유지할 일 수를 입력합니다.

   e. 결과 형식에서 ML 입력 채널이 사용해야 하는 데이터 형식으로 CSV 또는 Parquet을 선택합니다.

8. 서비스 액세스에서이 테이블에 액세스하는 데 사용할 기존 서비스 역할 이름을 선택하거나 새 서비스 역할 생성 및 사용을 선택합니다.
9. 암호화에서 사용자 지정 KMS 키로 보안 암호 암호화를 선택하여 자체 KMS 키 및 관련 정보를 지정합니다. 그렇지 않으면 Clean Rooms ML이 암호화를 관리합니다.
10. ML 입력 채널 생성을 선택합니다.

    ML 입력 채널을 생성하는 데 몇 분 정도 걸립니다. ML 입력 탭에서 ML 입력 채널 목록을 볼 수 있습니다.

> **Note**
>
> ML 입력 채널이 생성된 후에는 편집할 수 없습니다.

API

ML 입력 채널(API)을 생성하려면

특정 파라미터로 다음 코드를 실행합니다.

```
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"
                "computeConfiguration": {
                    "worker": {
                          "type": "CR.1X",
                          "number": 16
                        }
                    },
                "resultFormat": "PARQUET"
```

```
                }
            }
        },
        "roleArn": "arn:aws:iam::111122223333:role/ezcrc-ctm-role"
    }
)
channel_arn = resp['ML Input Channel ARN']
```

AWS Clean Rooms ML에서 훈련된 모델 생성

사전 조건:

- 에 액세스할 수 AWS 계정 있는 AWS Clean Rooms
- 에 설정된 공동 작업 AWS Clean Rooms
- 공동 작업과 연결된 구성된 모델 알고리즘
- 구성된 ML 입력 채널 하나 이상
- 공동 작업에서 ML 모델을 생성하고 관리할 수 있는 적절한 권한

구성된 모델 알고리즘을 공동 작업에 연결한 다음 ML 입력 채널을 생성하고 구성한 후에는 훈련된 모델을 생성할 준비가 된 것입니다. 훈련된 모델은 공동 작업 구성원이 데이터를 공동으로 분석하는 데 사용됩니다.

다음 절차를 사용하여 훈련된 모델을 생성할 수 있습니다.

또는 증분 훈련을 사용하여 새 데이터로 기존 모델을 개선하거나 분산 훈련을 사용하여 여러 컴퓨팅 인스턴스에서 모델을 훈련할 수 있습니다.

주제

- AWS Clean Rooms ML에서 증분 훈련 사용
- AWS Clean Rooms ML에서 분산 훈련 사용

Console

훈련된 모델을 생성하려면(콘솔)

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/cleanrooms AWS Clean Rooms 콘솔을 엽니다.

2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업 페이지에서 훈련된 모델을 생성할 공동 작업을 선택합니다.
4. 공동 작업이 열리면 ML 모델 탭을 선택합니다.
5. 사용자 지정 ML 모델의 훈련된 모델 섹션에서 훈련된 모델 생성을 선택합니다.
6. 훈련된 모델 생성 페이지의 연결된 모델 알고리즘에 알고리즘을 지정합니다.
7. 훈련된 모델 세부 정보에 다음을 입력합니다.

   a. 이름에 공동 작업에서 모델의 고유한 이름을 입력합니다.

   b. (선택 사항) 설명에 훈련된 모델에 대한 설명을 입력합니다.

   c. 훈련 데이터 입력 모드에서 다음 중 하나를 선택합니다.

      - ML 스토리지 볼륨에 맞는 더 작은 데이터 세트가 있고 훈련 스크립트에 대한 기존 파일 시스템 액세스를 선호하는 경우 파일을 선택합니다.
      - 대규모 데이터 세트의 경우 파이프를 선택하여 S3에서 직접 데이터를 스트리밍하므로 모든 것을 디스크에 다운로드할 필요가 없으므로 훈련 속도를 개선하고 스토리지 요구 사항을 줄일 수 있습니다.
      - 특히 순차적으로 데이터를 읽거나 더 적은 파일을 처리하여 시작 시간을 단축할 때 S3 스트리밍의 이점을 파일 시스템 액세스와 결합하려면 FastFile을 선택합니다.

8. ML 입력 채널 세부 정보에서 다음을 수행합니다.

   a. ML 입력 채널의 경우 모델 알고리즘에 데이터를 제공하는 ML 입력 채널을 지정합니다.

      다른 채널을 추가하려면 다른 ML 입력 채널 추가를 선택합니다. 최대 19개의 ML 입력 채널을 추가할 수 있습니다.

   b. 채널 이름에 ML 입력 채널의 이름을 입력합니다.

   c. Amazon S3 데이터 배포 유형에서 다음 중 하나를 선택합니다.

      - 완전히 복제됨을 선택하여 각 훈련 인스턴스에 데이터 세트의 전체 사본을 제공합니다. 이는 데이터 세트가 메모리에 들어갈 만큼 작거나 각 인스턴스가 모든 데이터에 액세스해야 하는 경우에 가장 적합합니다.
      - S3 키로 샤딩을 선택하여 S3 키를 기반으로 훈련 인스턴스 간에 데이터 세트를 분할합니다. 각 인스턴스는 총 S3 객체의 약 1/n을 수신합니다. 여기서 'n'은 인스턴스 수입니다. 이는 병렬로 처리하려는 대규모 데이터 세트에 가장 적합합니다.

> **Note**
>
> 배포 유형을 선택할 때 데이터 세트 크기와 훈련 요구 사항을 고려하세요. 완전 복제는 완전한 데이터 액세스를 제공하지만 더 많은 스토리지가 필요한 반면, S3 키로 샤딩하면 대규모 데이터 세트를 분산 처리할 수 있습니다.

9. 최대 훈련 기간에서 모델을 훈련할 최대 시간을 선택합니다.
10. 하이퍼파라미터에서 알고리즘별 파라미터와 해당 값을 지정합니다. 하이퍼파라미터는 훈련 중인 모델에 고유하며 모델 훈련을 미세 조정하는 데 사용됩니다.
11. 환경 변수에서 알고리즘별 변수와 해당 값을 지정합니다. 환경 변수는 Docker 컨테이너에서 설정됩니다.
12. 암호화에서 사용자 지정를 사용하려면 사용자 지정 KMS 키로 보안 암호 암호화 확인란을 AWS KMS key선택합니다.
13. EC2 리소스 구성에서 모델 훈련에 사용되는 컴퓨팅 리소스에 대한 정보를 지정합니다.
    a. 인스턴스 유형에서 실행할 인스턴스 유형을 선택합니다.
    b. 인스턴스 수에 인스턴스 수를 입력합니다.
    c. 볼륨 크기 GB에 ML 스토리지 볼륨 크기를 입력합니다.
14. 훈련된 모델 생성을 선택합니다.

API

훈련된 모델을 생성하려면(API)

모델을 훈련할 수 있는 구성원은 ML 입력 채널과 모델 알고리즘을 선택하여 훈련을 시작합니다.

특정 파라미터로 다음 코드를 실행합니다.

```
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"
        }
    ]
)
```

> **Note**
>
> 훈련된 모델을 생성한 후에는 편집할 수 없습니다. 변경하려면 훈련된 모델을 삭제하고 새 모델을 생성합니다.

AWS Clean Rooms ML에서 증분 훈련 사용

사전 조건:

- 에 액세스할 수 AWS 계정 있는 AWS Clean Rooms
- 공동 작업의 기존 훈련된 모델
- 증분 훈련을 위한 새 데이터 세트 또는 업데이트된 데이터 세트
- 공동 작업에서 ML 모델을 생성하고 관리할 수 있는 적절한 권한
- 기존 모델의 하이퍼파라미터 및 구성에 대한 지식

증분 훈련을 사용하면 기존 모델의 아티팩트와 업데이트된 데이터 세트를 사용하여 새 모델을 훈련할 수 있습니다. 점진적 훈련은 시간 및 리소스를 둘 다 절약할 수 있습니다.

점진적 훈련 사용:

- 이전 훈련에서 고려하지 않은 기본 패턴이 있는 확장된 데이터 세트를 사용하여 새 모델을 훈련합니다.
- 다른 하이퍼파라미터를 사용하거나 다른 데이터 세트를 사용하여 모델의 여러 변형을 훈련합니다.

Console

증분 훈련 작업을 실행하려면(콘솔)

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/cleanrooms AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업 페이지에서 증분 훈련에 사용할 모델 아티팩트가 있는 공동 작업을 선택합니다.
4. 공동 작업이 열리면 ML 모델 탭을 선택합니다.
5. 사용자 지정 ML 모델의 훈련된 모델 섹션에서 점진적으로 훈련하려는 훈련된 모델 옆에 있는 라디오 버튼을 선택합니다.
6. 개요 페이지의 버전에서

   a. 점진적으로 훈련하려는 훈련된 모델 옆에 있는 라디오 버튼을 선택합니다.

   b. 버전에서 훈련을 선택합니다.
7. 버전에서 훈련된 모델 생성 페이지의 훈련된 모델 버전에서 버전을 선택합니다.

   기본 모델 버전이 자동으로 선택됩니다. 다른 버전이 있는 경우이 버전을 변경할 수 있습니다.
8. 훈련된 모델 세부 정보에 다음을 입력합니다.

   a. 이름에 공동 작업에서 모델의 고유한 이름을 입력합니다.

   b. (선택 사항) 설명에 훈련된 모델에 대한 설명을 입력합니다.

   c. 훈련 데이터 입력 모드에서 다음 중 하나를 선택합니다.

      - ML 스토리지 볼륨에 맞는 더 작은 데이터 세트가 있고 훈련 스크립트에 대한 기존 파일 시스템 액세스를 선호하는 경우 파일을 선택합니다.
      - 대규모 데이터 세트의 경우 파이프를 선택하여 S3에서 직접 데이터를 스트리밍하므로 모든 데이터를 디스크에 다운로드할 필요가 없으므로 훈련 속도를 개선하고 스토리지 요구 사항을 줄일 수 있습니다.
      - 특히 순차적으로 데이터를 읽거나 시작 시간을 단축하기 위해 더 적은 파일을 처리할 때 S3에서 스트리밍하는 이점을 파일 시스템 액세스와 결합하려면 FastFile을 선택합니다.

d. 증분 훈련 채널 이름에 증분 훈련 채널의 이름을 입력합니다.

> **Note**
>
> 버전 ID 없이 증분 훈련 채널 이름을 지정하면 시스템은 증분 훈련에 기본 모델을 사용합니다.

9. ML 입력 채널 세부 정보에서 다음을 수행합니다.

   a. ML 입력 채널의 경우 모델 알고리즘에 데이터를 제공하는 ML 입력 채널을 지정합니다.

   다른 채널을 추가하려면 다른 ML 입력 채널 추가를 선택합니다. 최대 19개의 ML 입력 채널을 추가할 수 있습니다.

   b. 채널 이름에 ML 입력 채널의 이름을 입력합니다.

   c. Amazon S3 데이터 배포 유형에서 다음 중 하나를 선택합니다.

   - 완전히 복제됨을 선택하여 각 훈련 인스턴스에 데이터 세트의 전체 사본을 제공합니다. 이는 데이터 세트가 메모리에 들어갈 만큼 작거나 각 인스턴스가 모든 데이터에 액세스해야 하는 경우에 가장 적합합니다.
   - S3 키로 샤딩을 선택하여 S3 키를 기반으로 훈련 인스턴스 간에 데이터 세트를 분할합니다. 각 인스턴스는 총 S3 객체의 약 1/n을 수신합니다. 여기서 'n'은 인스턴스 수입니다. 이는 병렬로 처리하려는 대규모 데이터 세트에 가장 적합합니다.

> **Note**
>
> 배포 유형을 선택할 때 데이터 세트 크기와 훈련 요구 사항을 고려하세요. 완전 복제는 완전한 데이터 액세스를 제공하지만 더 많은 스토리지가 필요한 반면, S3 키로 샤딩하면 대규모 데이터 세트를 분산 처리할 수 있습니다.

10. 최대 훈련 기간에서 모델을 훈련할 최대 시간을 선택합니다.
11. 하이퍼파라미터에서 알고리즘별 파라미터와 해당 값을 지정합니다. 하이퍼파라미터는 훈련 중인 모델에 고유하며 모델 훈련을 미세 조정하는 데 사용됩니다.
12. 환경 변수에서 알고리즘별 변수와 해당 값을 지정합니다. 환경 변수는 Docker 컨테이너에서 설정됩니다.
13. 암호화에서 사용자 지정를 사용하려면 사용자 지정 KMS 키로 보안 암호 암호화 확인란을 AWS KMS key선택합니다.

14. EC2 리소스 구성에서 모델 훈련에 사용되는 컴퓨팅 리소스에 대한 정보를 지정합니다.
    a. 인스턴스 유형에서 실행할 인스턴스 유형을 선택합니다.
    b. 인스턴스 수에 인스턴스 수를 입력합니다.
    c. 볼륨 크기 GB에 ML 스토리지 볼륨 크기를 입력합니다.
15. 버전에서 훈련된 모델 생성을 선택합니다.

API

증분 훈련 작업(API)을 실행하려면

특정 파라미터로 다음 코드를 실행합니다.

```
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
    },
    incrementalTrainingDataChannels=[
        {
            'trainedModelArn': trained_model_arn,
            'channelName': 'channel_name'
        },
    ]
    dataChannels=[
        {
            'mlInputChannelArn': channel_arn_1,
            'channelName': 'channel_name'
        },
        {
            'mlInputChannelArn': channel_arn_2,
            'channelName': 'channel_name'
        }
    ]
```

```
)
```

> **Note**
>
> 제한: 최대 총 20개의 채널(`dataChannels` 및 모두 포함`incrementalTrainingDataChannels`).

> **Note**
>
> 훈련된 모델을 생성한 후에는 편집할 수 없습니다. 변경하려면 훈련된 모델을 삭제하고 새 모델을 생성합니다.

AWS Clean Rooms ML에서 분산 훈련 사용

사전 조건:

- 에 액세스할 수 AWS 계정 있는 AWS Clean Rooms
- 에 설정된 공동 작업 AWS Clean Rooms
- 분산 훈련을 지원하는 구성된 모델 알고리즘
- 분산 처리에 적합한 대규모 데이터 세트
- 공동 작업에서 ML 모델을 생성하고 관리할 수 있는 적절한 권한
- 분산 훈련을 위해 여러 인스턴스를 실행하기에 충분한 Amazon EC2 할당량

분산 훈련은 병렬로 작동하는 많은 컴퓨팅 노드의 성능을 활용하여 대량의 데이터를 처리하고 모델 파라미터를 효율적으로 업데이트합니다.

분산 훈련에 대한 자세한 내용은 Amazon SageMaker AI 개발자 안내서의 분산 훈련 개념을 참조하세요.

Console

분산 훈련 작업을 실행하려면(콘솔)

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/cleanrooms AWS Clean Rooms 콘솔을 엽니다.

2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업 페이지에서 훈련된 모델을 생성할 공동 작업을 선택합니다.
4. 공동 작업이 열리면 ML 모델 탭을 선택합니다.
5. 사용자 지정 ML 모델의 훈련된 모델 섹션에서 훈련된 모델 생성을 선택합니다.
6. 훈련된 모델 생성 페이지의 연결된 모델 알고리즘에 알고리즘을 지정합니다.
7. 훈련된 모델 세부 정보에 다음을 입력합니다.

   a. 이름에 공동 작업에서 모델의 고유한 이름을 입력합니다.

   b. (선택 사항) 설명에 훈련된 모델에 대한 설명을 입력합니다.

   c. 훈련 데이터 입력 모드에서 다음 중 하나를 선택합니다.

      - ML 스토리지 볼륨에 맞는 더 작은 데이터 세트가 있고 훈련 스크립트에 대한 기존 파일 시스템 액세스를 선호하는 경우 파일을 선택합니다.
      - 대규모 데이터 세트의 경우 파이프를 선택하여 S3에서 직접 데이터를 스트리밍하므로 모든 것을 디스크에 다운로드할 필요가 없으므로 훈련 속도를 개선하고 스토리지 요구 사항을 줄일 수 있습니다.
      - 특히 순차적으로 데이터를 읽거나 더 적은 파일을 처리하여 시작 시간을 단축할 때 S3 스트리밍의 이점을 파일 시스템 액세스와 결합하려면 FastFile을 선택합니다.
8. ML 입력 채널 세부 정보에서 다음을 수행합니다.

   a. ML 입력 채널의 경우 모델 알고리즘에 데이터를 제공하는 ML 입력 채널을 지정합니다.

      다른 채널을 추가하려면 다른 ML 입력 채널 추가를 선택합니다. 최대 19개의 ML 입력 채널을 추가할 수 있습니다.

   b. 채널 이름에 ML 입력 채널의 이름을 입력합니다.

   c. Amazon S3 데이터 배포 유형에서 다음 중 하나를 선택합니다.

      - 완전히 복제됨을 선택하여 각 훈련 인스턴스에 데이터 세트의 전체 사본을 제공합니다. 이는 데이터 세트가 메모리에 들어갈 만큼 작거나 각 인스턴스가 모든 데이터에 액세스해야 하는 경우에 가장 적합합니다.
      - S3 키로 샤딩을 선택하여 S3 키를 기반으로 훈련 인스턴스 간에 데이터 세트를 분할합니다. 각 인스턴스는 총 S3 객체의 약 1/n을 수신합니다. 여기서 'n'은 인스턴스 수입니다. 이는 병렬로 처리하려는 대규모 데이터 세트에 가장 적합합니다.

> **Note**
>
> 배포 유형을 선택할 때 데이터 세트 크기와 훈련 요구 사항을 고려하세요. 완전 복제는 완전한 데이터 액세스를 제공하지만 더 많은 스토리지가 필요한 반면, S3 키로 샤딩하면 대규모 데이터 세트를 분산 처리할 수 있습니다.

9. 최대 훈련 기간에서 모델을 훈련할 최대 시간을 선택합니다.
10. 하이퍼파라미터에서 알고리즘별 파라미터와 해당 값을 지정합니다. 하이퍼파라미터는 훈련 중인 모델에 고유하며 모델 훈련을 미세 조정하는 데 사용됩니다.
11. 환경 변수에서 알고리즘별 변수와 해당 값을 지정합니다. 환경 변수는 Docker 컨테이너에서 설정됩니다.
12. 암호화에서 사용자 지정를 사용하려면 사용자 지정 KMS 키로 보안 암호 암호화 확인란을 AWS KMS key선택합니다.
13. EC2 리소스 구성에서 모델 훈련에 사용되는 컴퓨팅 리소스에 대한 정보를 지정합니다.

    a. 인스턴스 유형에서 실행할 인스턴스 유형을 선택합니다.

       분산 훈련에 지원되는 인스턴스 유형은 다음과 같습니다.

       - ml.m5.4xlarge
       - ml.m5.12xlarge
       - ml.m5.2xlarge
       - ml.g5.12xlarge
       - ml.g5.24xlarge

    b. 인스턴스 수에 인스턴스 수를 입력합니다.

    c. 볼륨 크기 GB에 ML 스토리지 볼륨 크기를 입력합니다.
14. 훈련된 모델 생성을 선택합니다.

API

분산 훈련 작업(API)을 실행하려면

특정 파라미터로 다음 코드를 실행합니다.

```
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',
    trainingInputMode: "File",
    resourceConfig={
        'instanceCount': "3"
        'instanceType': "ml.m5.xlarge",
        'volumeSizeInGB': 3
    },
    dataChannels=[
        {
            "mlInputChannelArn": channel_arn_1,
            "channelName": "channel_name",
            "S3DataDistributionType:" "FullyReplicated"
        }

    ]
)
```

> **Note**
>
> 훈련된 모델을 생성한 후에는 편집할 수 없습니다. 변경하려면 훈련된 모델을 삭제하고 새 모델을 생성합니다.

AWS Clean Rooms ML에서 모델 아티팩트 내보내기

이 작업은 선택 사항이며 공동 작업의 CAN_RECEIVE_MODEL_OUTPUT 구성원에게 구성원 기능을 할당한 경우 완료해야 합니다.

모델 훈련이 완료되면 모델을 훈련한 구성원이 모델 아티팩트 내보내기를 시작할 수 있습니다. 모델을 훈련한 구성원은 모델 아티팩트를 받을 사람을 선택합니다. 단, 해당 구성원이 결과와 유효한 ML 구성을 받을 수 있어야 합니다.

Console

사용자 지정 ML 모델 알고리즘을 구성하려면(콘솔)

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/cleanrooms AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업 페이지에서 내보낼 사용자 지정 모델이 포함된 공동 작업을 선택합니다.
4. 공동 작업이 열리면 ML 모델 탭을 선택한 다음 사용자 지정 훈련된 모델 테이블에서 모델을 선택합니다.
5. 사용자 지정 훈련된 모델 세부 정보 페이지에서 모델 출력 내보내기를 클릭합니다.
6. 모델 출력 내보내기의 모델 출력 세부 정보 내보내기에 이름 및 선택적 설명을 입력합니다.

   공동 작업 구성원에게 내보낸 모델 출력 드롭다운 목록에서 모델 아티팩트를 수신할 구성원을 선택합니다.
7. 내보내기를 선택합니다.

   결과는 ML 구성에 지정된 Amazon S3 위치의 다음 경로로 내보내집니다`yourSpecifiedS3Path/collaborationIdentifier/trainedModelName/callerAccountId/jobName`. 구성된 모델 알고리즘을 연결할 때 선택한 최대 파일 크기까지만 내보냅니다.

API

사용자 지정 ML 모델 알고리즘(API)을 구성하려면

다음 코드를 실행하여 모델 내보내기를 시작합니다.

```
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'
)
```

결과는 ML 구성에 지정된 Amazon S3 위치의 다음 경로로 내보내집니다`yourSpecifiedS3Path/collaborationIdentifier/trainedModelName/callerAccountId/jobName`. 구성된 모델 알고리즘을 연결할 때 선택한 만 `maxSize` 지정된 `filesToExport`까지 내보냅니다.

AWS Clean Rooms ML에서 훈련된 모델에 대한 추론 실행

쿼리를 실행할 수 있는 구성원은 훈련 작업이 완료되면 추론 작업을 시작할 수도 있습니다. 추론을 실행하려는 추론 데이터 세트를 선택하고 추론 컨테이너를 실행하려는 훈련된 모델 출력을 참조합니다.

추론 출력을 수신할 멤버에게 멤버 기능을 부여해야 합니다`CAN_RECEIVE_INFERENCE_OUTPUT`.

Console

모델 추론 작업을 생성하려면(콘솔)

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/cleanrooms AWS Clean Rooms 콘솔을 엽니다.
2. 왼쪽 탐색 창에서 공동 작업을 선택합니다.
3. 공동 작업 페이지에서 추론 작업을 생성할 사용자 지정 모델이 포함된 공동 작업을 선택합니다.
4. 공동 작업이 열리면 ML 모델 탭을 선택한 다음 사용자 지정 훈련된 모델 테이블에서 모델을 선택합니다.
5. 사용자 지정 훈련된 모델 세부 정보 페이지에서 추론 작업 시작을 클릭합니다.
6. 추론 작업 시작의 추론 작업 세부 정보에 이름 및 선택적 설명을 입력합니다.

   다음 정보를 입력합니다.

   - 연결된 모델 알고리즘 - 추론 작업 중에 사용되는 연결된 모델 알고리즘입니다.
   - ML 입력 채널 세부 정보 -이 추론 작업에 대한 데이터를 제공하는 ML 입력 채널입니다.
   - 리소스 변환 - 추론 작업의 변환 함수를 수행하는 데 사용되는 컴퓨팅 인스턴스입니다.
   - 출력 구성 - 추론 작업 출력과 출력의 MIME 유형을 수신할 사용자입니다.

- 암호화 - 암호화 설정 사용자 지정을 선택하여 자체 KMS 키 및 관련 정보를 지정합니다. 그렇지 않으면 Clean Rooms ML이 암호화를 관리합니다.
- 변환 작업 세부 정보 - 추론 작업의 최대 페이로드입니다. 단위는 MB입니다.
- 환경 변수 - 추론 작업 컨테이너 이미지에 액세스하는 데 필요한 모든 환경 변수입니다.

7. 추론 작업 시작을 선택합니다.

   결과는 ML 구성에 지정된 Amazon S3 위치의 다음 경로로 내보내집니다`yourSpecifiedS3Path/collaborationIdentifier/trainedModelName/callerAccountId/jobName`.

API

모델 추론 작업을 생성하려면(API)

다음 코드를 실행하여 추론 작업을 시작합니다.

```
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'
            }
        ]
    }
)
```

결과는 ML 구성에 지정된 Amazon S3 위치의 다음 경로로 내보내집니다`yourSpecifiedS3Path/collaborationIdentifier/trainedModelName/callerAccountId/jobName`.

문제 해결 AWS Clean Rooms

이 섹션에서는를 사용할 때 발생할 수 있는 몇 가지 일반적인 문제 AWS Clean Rooms 와 이를 해결하는 방법을 설명합니다.

문제

- 쿼리에서 참조하는 하나 이상의 테이블은 관련 서비스 역할로 액세스할 수 없습니다. 테이블/역할 소유자는 서비스 역할에 테이블에 대한 액세스 권한을 부여해야 합니다.
- 기본 데이터 세트 중 하나에 지원되지 않는 파일 형식이 있습니다.
- Clean Rooms에 대한 암호화 컴퓨팅을 사용하는 경우 쿼리 결과가 예상과 다릅니다.

쿼리에서 참조하는 하나 이상의 테이블은 관련 서비스 역할로 액세스할 수 없습니다. 테이블/역할 소유자는 서비스 역할에 테이블에 대한 액세스 권한을 부여해야 합니다.

- 서비스 역할에 대한 권한이 필요에 따라 설정되었는지 확인합니다. 자세한 정보는 설 AWS Clean Rooms정을 참조하세요.

기본 데이터 세트 중 하나에 지원되지 않는 파일 형식이 있습니다.

- 데이터 세트가 지원되는 파일 형식 중 하나인지 확인하세요.
  - Parquet
  - RCFile
  - TextFile
  - SequenceFile
  - RegexSerde
  - OpenCSV
  - AVRO
  - JSON

  자세한 내용은 의 데이터 형식 AWS Clean Rooms 단원을 참조하십시오.

Clean Rooms에 대한 암호화 컴퓨팅을 사용하는 경우 쿼리 결과가 예상과 다릅니다.

Clean Rooms용 암호화 컴퓨팅(C3R)을 사용하는 경우 쿼리에서 암호화된 열을 올바르게 사용하는지 확인하세요.

- sealed 열은 SELECT 조항에만 사용됩니다.
- fingerprint열은 JOIN 조항(및 특정 조건의 GROUP BY 조항)에만 사용됩니다.
- 공동 작업 설정에 필요한 경우 동일한 이름을 가진 JOINing fingerprint 열만 사용할 수 있다는 것입니다.

자세한 내용은 the section called “암호화 컴퓨팅” 및 the section called “열 유형” 단원을 참조하세요.

의 보안 AWS Clean Rooms

의 클라우드 보안 AWS 이 최우선 순위입니다. AWS 고객은 보안에 가장 민감한 조직의 요구 사항을 충족하도록 구축된 데이터 센터 및 네트워크 아키텍처의 이점을 누릴 수 있습니다.

보안은 AWS 와 사용자 간의 공동 책임입니다. 공동 책임 모델은 이 사항을 클라우드의 보안 및 클라우드 내 보안으로 설명합니다.

- 클라우드 보안 - AWS 는 AWS 클라우드에서 AWS 서비스를 실행하는 인프라를 보호할 책임이 있습니다. AWS 또한는 안전하게 사용할 수 있는 서비스를 제공합니다. 타사 감사자는 AWS 규정 준수 프로그램 일환으로 보안의 효과를 정기적으로 테스트하고 확인합니다. 에 적용되는 규정 준수 프로그램에 대한 자세한 내용은 규정 준수 프로그램 제공 범위 내 AWS 서비스규정 준수 프로그램 AWS Clean Rooms참조하세요.
- 클라우드의 보안 - 사용자의 책임은 사용하는 AWS 서비스에 따라 결정됩니다. 또한 여러분은 데이터의 민감도, 회사 요구 사항, 관련 법률 및 규정을 비롯한 기타 요소에 대해서도 책임이 있습니다

이 설명서는를 사용할 때 공동 책임 모델을 적용하는 방법을 이해하는 데 도움이 됩니다 AWS Clean Rooms. 보안 및 규정 준수 목표를 충족하도록 AWS Clean Rooms 를 구성하는 방법을 보여줍니다. 또한 리소스를 모니터링하고 보호하는 AWS Clean Rooms 데 도움이 되는 다른 AWS 서비스를 사용하는 방법을 알아봅니다.

내용

- 의 데이터 보호 AWS Clean Rooms
- 의 데이터 보존 AWS Clean Rooms
- 의 데이터 협업 모범 사례 AWS Clean Rooms
- 에 대한 자격 증명 및 액세스 관리 AWS Clean Rooms
- 에 대한 규정 준수 검증 AWS Clean Rooms
- 의 복원력 AWS Clean Rooms
- 의 인프라 보안 AWS Clean Rooms
- 인터페이스 엔드포인트를 사용하여 AWS Clean Rooms 또는 AWS Clean Rooms ML에 액세스 (AWS PrivateLink)

의 데이터 보호 AWS Clean Rooms

AWS 공동 책임 모델의 데이터 보호에 적용됩니다 AWS Clean Rooms. 이 모델에 설명된 대로 AWS 는 모든를 실행하는 글로벌 인프라를 보호할 책임이 있습니다 AWS 클라우드. 사용자는 인프라에서 호스팅되는 콘텐츠를 관리해야 합니다. 사용하는 AWS 서비스 의 보안 구성과 관리 태스크에 대한 책임도 사용자에게 있습니다. 데이터 프라이버시에 대한 자세한 내용은 데이터 프라이버시 FAQ를 참조하세요. 유럽의 데이터 보호에 대한 자세한 내용은 AWS 보안 블로그의 AWS 공동 책임 모델 및 GDPR 블로그 게시물을 참조하세요.

데이터 보호를 위해 자격 증명을 보호하고 AWS 계정 AWS IAM Identity Center 또는 AWS Identity and Access Management (IAM)를 사용하여 개별 사용자를 설정하는 것이 좋습니다. 이렇게 하면 개별 사용자에게 자신의 직무를 충실히 이행하는 데 필요한 권한만 부여됩니다. 또한 다음과 같은 방법으로 데이터를 보호하는 것이 좋습니다.

- 각 계정에 다중 인증(MFA)을 사용하세요.
- SSL/TLS를 사용하여 AWS 리소스와 통신합니다. TLS 1.2는 필수이며 TLS 1.3을 권장합니다.
- 를 사용하여 API 및 사용자 활동 로깅을 설정합니다 AWS CloudTrail. CloudTrail 추적을 사용하여 AWS 활동을 캡처하는 방법에 대한 자세한 내용은 AWS CloudTrail 사용 설명서의 CloudTrail 추적 작업을 참조하세요.
- 내의 모든 기본 보안 제어와 함께 AWS 암호화 솔루션을 사용합니다 AWS 서비스.
- Amazon S3에 저장된 민감한 데이터를 검색하고 보호하는 데 도움이 되는 Amazon Macie와 같은 고급 관리형 보안 서비스를 사용하세요.
- 명령줄 인터페이스 또는 API를 AWS 통해에 액세스할 때 FIPS 140-3 검증 암호화 모듈이 필요한 경우 FIPS 엔드포인트를 사용합니다. 사용 가능한 FIPS 엔드포인트에 대한 자세한 내용은 Federal Information Processing Standard(FIPS) 140-3을 참조하세요.

고객의 이메일 주소와 같은 기밀 정보나 중요한 정보는 태그나 이름 필드와 같은 자유 형식 텍스트 필드에 입력하지 않는 것이 좋습니다. 여기에는 AWS Clean Rooms 또는 기타 AWS 서비스 에서 콘솔, API AWS CLI또는 AWS SDKs를 사용하여 작업하는 경우가 포함됩니다. 이름에 사용되는 태그 또는 자유 형식 텍스트 필드에 입력하는 모든 데이터는 청구 또는 진단 로그에 사용될 수 있습니다. 외부 서버에 URL을 제공할 때 해당 서버에 대한 요청을 검증하기 위해 보안 인증 정보를 URL에 포함시켜서는 안 됩니다.

저장된 데이터 암호화

AWS Clean Rooms 는 추가 구성 없이 항상 모든 서비스 메타데이터를 암호화합니다. 이 암호화는를 사용할 때 자동으로 수행됩니다 AWS Clean Rooms.

Clean Rooms ML은 저장된 서비스 내에 저장된 모든 데이터를 암호화합니다 AWS KMS. 자체 KMS 키를 제공하기로 선택하면 유사 모델 및 유사 세그먼트 생성 작업의 콘텐츠가 KMS 키를 사용하여 유휴 상태에서 암호화됩니다.

AWS Clean Rooms 사용자 지정 ML 모델을 사용하는 경우 서비스는 저장 시 저장된 모든 데이터를 암호화합니다 AWS KMS.는 사용자가 저장 시 데이터를 암호화하기 위해 생성, 소유 및 관리하는 대칭 고객 관리형 키의 사용을 AWS Clean Rooms 지원합니다. 고객 관리형 키를 지정하지 않으면 AWS 소유 키 가 기본적으로 사용됩니다.

AWS Clean Rooms 는 권한 부여 및 키 정책을 사용하여 고객 관리형 키에 액세스합니다. 언제든지 권한 부여에 대한 액세스 권한을 취소하거나 고객 관리형 키에 대한 서비스 액세스를 제거할 수 있습니다. 이렇게 하면 고객 관리형 키로 암호화된 데이터에 액세스할 수 AWS Clean Rooms 없으며, 이는 해당 데이터에 의존하는 작업에 영향을 미칩니다. 예를 들어가 액세스할 AWS Clean Rooms 수 없는 암호화된 ML 입력 채널에서 훈련된 모델을 생성하려고 하면 작업이 `ValidationException` 오류를 반환합니다.

> **Note**
>
> Amazon S3의 암호화 옵션을 사용하여 저장 데이터를 보호할 수 있습니다.
> 자세한 내용은 Amazon S3 사용 설명서의 Amazon S3 암호화 지정을 참조하세요.

내에서 ID 매핑 테이블을 사용하는 경우 AWS Clean Rooms서비스는 저장 시 저장된 모든 데이터를 암호화합니다 AWS KMS. 자체 KMS 키를 제공하기로 선택하면 ID 매핑 테이블의 콘텐츠가 KMS 키를 통해 저장 시 암호화됩니다 AWS Entity Resolution. ID 매핑 워크플로로 암호화 작업을 하는 데 필요한 권한에 대한 자세한 내용은 AWS Entity Resolution 사용 설명서의 AWS Entity Resolution용 워크플로 작업 역할 생성을 참조하세요.

전송 중 암호화

AWS Clean Rooms 는 전송 중 암호화에 TLS(전송 계층 보안)를 사용합니다. 와의 통신 AWS Clean Rooms 은 항상 HTTPS를 통해 수행되므로 데이터가 Amazon S3, Amazon Athena 또는 Snowflake에 저장되었는지 여부에 관계없이 전송 중에 항상 암호화됩니다. Clean Rooms ML을 사용할 경우 전송 중인 모든 데이터가 여기에 포함됩니다.

기본 데이터 암호화

기본 데이터를 암호화하는 방법에 대한 자세한 내용은 Clean Rooms에 대한 암호화 컴퓨팅 섹션을 참조하세요.

키 정책

키 정책에서는 고객 관리형 키에 대한 액세스를 제어합니다. 모든 고객 관리형 키에는 키를 사용할 수 있는 사람과 키를 사용하는 방법을 결정하는 문장이 포함된 정확히 하나의 키 정책이 있어야 합니다. 고객 관리형 키를 만들 때 키 정책을 지정할 수 있습니다. 자세한 내용은 AWS Key Management Service 개발자 안내서의 고객 관리형 키에 대한 액세스 관리를 참조하세요.

사용자 AWS Clean Rooms 지정 ML 모델에서 고객 관리형 키를 사용하려면 키 정책에서 다음 API 작업을 허용해야 합니다.

- `kms:DescribeKey` -가 키를 AWS Clean Rooms 검증할 수 있도록 고객 관리형 키 세부 정보를 제공합니다.
- `kms:Decrypt` - 암호화된 데이터를 복호화하고 관련 작업에 사용할 AWS Clean Rooms 수 있는에 대한 액세스를 제공합니다.
- `kms:CreateGrant` - Clean Rooms ML은 Amazon ECR에 대한 권한 부여를 생성하여 Amazon ECR에서 저장된 훈련 및 추론 이미지를 암호화합니다. 자세한 내용은 Amazon ECR의 저장 데이터 암호화를 참조하세요. 또한 Clean Rooms ML은 Amazon SageMaker AI를 사용하여 훈련 및 추론 작업을 실행하고 SageMaker AI에 대한 권한 부여를 생성하여 인스턴스에 연결된 Amazon EBS 볼륨과 Amazon S3의 출력 데이터를 암호화합니다. 자세한 내용은 Amazon SageMaker AI의 암호화를 사용하여 저장 데이터 보호를 참조하세요.
- `kms:GenerateDataKey` - Clean Rooms ML은를 사용한 서버 측 암호화를 사용하여 Amazon S3에 저장된 저장 데이터를 암호화합니다 AWS KMS keys. 자세한 내용은 Amazon S3에서 AWS KMS keys (SSE-KMS)를 사용한 서버 측 암호화 사용을 참조하세요.

다음은 다음 리소스 AWS Clean Rooms 에 대해 추가할 수 있는 정책 설명 예제입니다.

ML 입력 채널

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
        "Sid": "Allow access to principals authorized to use ",
        "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  service principal",
        "Effect": "Allow",
        "Principal": {
            "Service": "cleanrooms-ml.amazonaws.com"
        },
        "Action": [
            "kms:DescribeKey",
            "kms:GenerateDataKey",
            "kms:Decrypt"
        ],
        "Resource": "*"
    }
  ]
}
```

훈련된 모델 작업 또는 훈련된 모델 추론 작업

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
        "Sid": "Allow access to principals authorized to use ",
        "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  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 ML은 고객 관리형 키 정책에서 서비스 암호화 컨텍스트 또는 소스 컨텍스트 지정을 지원하지 않습니다. 서비스에서 내부적으로 사용하는 암호화 컨텍스트는 CloudTrail의 고객에게 표시됩니다.

의 데이터 보존 AWS Clean Rooms

AWS Clean Rooms 공동 작업에 일시적으로 읽는 모든 데이터는 쿼리가 완료된 후 삭제됩니다.

유사 모델을 만들면 Clean Rooms ML은 훈련 데이터를 읽고 이를 ML 모델에 적합한 형식으로 변환한 다음 훈련된 모델 파라미터를 Clean Rooms ML 내에 저장합니다. Clean Rooms ML은 훈련 데이터의 사본을 보관하지 않습니다. AWS Clean Rooms SQL 쿼리는 쿼리가 실행된 후 데이터를 보관하지 않습니다. 그런 다음 Clean Rooms ML은 훈련된 모델을 사용하여 모든 사용자의 동작을 요약합니다. Clean Rooms ML은 유사 모델이 활성화되어 있는 한 각 사용자에 대한 사용자 수준 데이터 세트를 데이터에 저장합니다.

유사 세그먼트 생성 작업을 시작하면 Clean Rooms ML은 시드 데이터를 읽고, 연결된 유사 모델에서 동작 요약을 읽고, AWS Clean Rooms 서비스 내에 저장된 유사 세그먼트를 생성합니다. Clean Rooms ML은 시드 데이터의 사본을 보관하지 않습니다. Clean Rooms ML은 작업이 활성 상태인 한 작업의 사용자 수준 출력을 저장합니다.

시드 데이터가 SQL 쿼리에서 나온 경우, 해당 쿼리의 출력은 작업 기간 동안만 서비스에 저장됩니다. 쿼리 결과는 암호화된 상태로 저장 및 전송됩니다.

유사 모델 또는 유사 세그먼트 생성 작업 데이터를 제거하려면 API를 사용하여 삭제합니다. Clean Rooms ML은 모델 또는 작업과 관련된 모든 데이터를 비동기적으로 삭제합니다. 이 프로세스가 완료

되면 Clean Rooms ML은 모델 또는 작업에 대한 메타데이터를 삭제하며 API에 더 이상 표시되지 않습니다. Clean Rooms ML은 재해 복구 방지를 위해 삭제된 데이터를 3일 동안 보관합니다. 작업 또는 모델이 API에 더 이상 표시되지 않고 3일이 지나면 모델 또는 작업과 관련된 모든 데이터가 영구적으로 삭제됩니다.

의 데이터 협업 모범 사례 AWS Clean Rooms

이 항목에서는 AWS Clean Rooms에서 데이터 공동 작업을 수행하는 모범 사례를 설명합니다.

AWS Clean Rooms 는 AWS 공동 책임 Model. AWS Clean Rooms offers 분석 규칙을 따르며, 이를 통해 공동 작업에서 민감한 데이터를 보호하는 능력을 강화할 수 있습니다. 에서 구성하는 분석 규칙 AWS Clean Rooms 은 구성한 제한(쿼리 제어 및 쿼리 출력 제어)을 적용합니다. 제한을 결정하고 그에 따라 분석 규칙을 구성하는 것은 사용자의 책임입니다.

데이터 공동 작업에는 단순한 사용 이상의 작업이 포함될 수 있습니다 AWS Clean Rooms. 데이터 공동 작업의 이점을 극대화하려면 AWS Clean Rooms 를 사용하고 특히 분석 규칙을 사용하여 다음 모범 사례를 수행하는 것이 좋습니다.

주제

- 의 모범 사례 AWS Clean Rooms
- AWS Clean Rooms에서 분석 규칙을 사용하는 모범 사례

의 모범 사례 AWS Clean Rooms

각 데이터 공동 작업의 위험을 평가하고 이를 외부 및 내부 규정 준수 프로그램 및 정책과 같은 개인 정보 보호 요구 사항과 비교하는 것은 귀하의 책임입니다. 사용 시 추가 조치를 취하는 것이 좋습니다 AWS Clean Rooms. 이러한 조치는 위험을 추가로 관리하고 제3자가 데이터를 재식별하려는 시도(예: 차별화 공격 또는 부채널 공격)를 방지하는 데 도움이 될 수 있습니다.

예를 들어, 공동 작업에 참여하기 전에 다른 협업자를 대상으로 실사를 실시하고 이들과 법적 계약을 체결하는 것을 고려해 보세요. 데이터 사용을 모니터링하려면 AWS Clean Rooms을 사용하여 다른 감사 메커니즘을 채택하는 것도 고려해 볼 수 있습니다.

AWS Clean Rooms에서 분석 규칙을 사용하는 모범 사례

의 분석 규칙을 AWS Clean Rooms 사용하면 구성된 테이블에 쿼리 제어를 설정하여 실행할 수 있는 쿼리를 제한할 수 있습니다. 예를 들어 구성된 테이블을 조인하는 방법과 선택할 수 있는 열에 대한 쿼

리 컨트롤을 설정할 수 있습니다. 또한 출력 행의 집계 임계값과 같은 쿼리 결과 제어를 설정하여 쿼리 출력을 제한할 수 있습니다. 서비스는 구성원이 쿼리에서 구성한 테이블에 대한 분석 규칙을 준수하지 않는 쿼리를 거부하고 해당 행을 제거합니다.

구성된 테이블에 분석 규칙을 사용할 때는 다음 10가지 모범 사례를 따르는 것이 좋습니다.

- 별도의 쿼리 사용 사례(예: 대상 계획 또는 어트리뷰션)에 맞게 구성된 테이블을 별도로 생성하세요. 동일한 기본 AWS Glue 테이블을 사용하여 구성된 테이블을 여러 개 만들 수 있습니다.
- 공동 작업에서 쿼리에 필요한 분석 규칙의 열(예: 차원 열, 목록 열, 조인 열)을 지정합니다. 이렇게 하면 차별화 공격의 위험을 완화하거나 다른 구성원이 데이터를 리버스 엔지니어링할 수 있습니다. 열 허용 목록 기능을 사용하여 나중에 쿼리할 수 있는 다른 열을 기록해 둡니다. 특정 공동 작업에 사용할 수 있는 열을 사용자 지정하려면 동일한 기본 테이블로 구성된 AWS Glue 테이블을 추가로 생성합니다.
- 공동 작업에서 분석에 필요한 기능을 분석 규칙에 지정합니다. 이렇게 하면 드문 기능 오류로부터의 리스크를 완화하는 데 도움이 될 수 있습니다. 이러한 오류는 개별 데이터 포인트에 관한 정보를 노출시킬 수 있습니다. 특정 공동 작업에 사용할 수 있는 함수를 사용자 지정하려면 동일한 기본 AWS Glue 테이블을 사용하여 구성된 테이블을 추가로 생성합니다.
- 해당 행의 값이 민감한 열에는 집계 제약 조건을 추가합니다. 이는 구성된 테이블에 있는 열 뿐만 아니라 다른 공동 작업 멤버의 테이블과 분석 규칙에서도 집계 제약 조건으로 사용되는 열을 포함합니다. 여기에는 쿼리할 수 없는 구성 테이블의 열, 즉 구성된 테이블에는 있지만 분석 규칙에는 없는 열도 포함됩니다. 집계 제약 조건은 쿼리 결과를 공동 작업 외부 데이터와 상호 연관시키는 데 따른 위험을 완화하는 데 도움이 될 수 있습니다.
- 테스트 공동 작업 및 분석 규칙을 생성하여 지정된 분석 규칙으로 생성된 테스트 제한을 테스트할 수 있습니다.
- 구성된 테이블에 대한 공동 작업자 구성 테이블 및 구성원의 분석 규칙을 검토하여 공동 작업을 위해 합의한 내용과 일치하는지 확인하세요. 이렇게 하면 다른 구성원이 합의되지 않은 쿼리를 실행하기 위해 자신의 데이터를 엔지니어링하는 위험을 완화할 수 있습니다.
- 분석 규칙을 설정한 후 구성된 테이블에서 활성화된 제공된 예제 쿼리(콘솔만 해당)를 검토합니다.

> **Note**
>
> 제공된 예제 쿼리 외에도 분석 규칙, 기타 공동 작업 멤버 테이블 및 분석 규칙을 기반으로 다른 쿼리를 사용할 수 있습니다.

- 공동 작업에서 구성된 테이블에 대한 분석 규칙을 추가 또는 업데이트할 수 있습니다. 그런 다음 구성된 테이블이 연결된 모든 공동 작업과 그에 따른 영향을 검토합니다. 이렇게 하면 더 이상 사용되지 않는 분석 규칙을 사용하는 공동 작업이 없도록 할 수 있습니다.

- 공동 작업에서 실행되는 쿼리를 검토하여 쿼리가 공동 작업을 위해 합의된 사용 사례 또는 쿼리와 일치하는지 확인하세요. (쿼리 로깅 기능이 켜져 있으면 쿼리 로그에서 쿼리를 사용할 수 있습니다.) 이렇게 하면 합의되지 않은 분석을 실행하는 구성원의 위험과 사이드 채널 공격과 같은 잠재적 공격을 완화하는 데 도움이 될 수 있습니다.
- 공동 작업 구성원의 분석 규칙 및 쿼리에 사용되는 구성된 테이블 열을 검토하여 공동 작업에서 합의한 내용과 일치하는지 확인하세요. (해당 기능이 켜져 있으면 쿼리 로그에서 쿼리를 사용할 수 있습니다.) 이렇게 하면 다른 구성원이 합의되지 않은 쿼리를 수행하기 위해 자신의 데이터를 조작하는 위험을 완화할 수 있습니다.

에 대한 자격 증명 및 액세스 관리 AWS Clean Rooms

AWS Identity and Access Management (IAM)는 관리자가 AWS 리소스에 대한 액세스를 안전하게 제어하는 데 도움이 AWS 서비스 되는 입니다. IAM 관리자는 누가 AWS Clean Rooms 리소스를 사용할 수 있는 인증(로그인) 및 권한(권한 있음)을 받을 수 있는지 제어합니다. IAM은 추가 비용 없이 사용할 수 AWS 서비스 있는 입니다.

주제

- 대상
- ID를 통한 인증
- 정책을 사용하여 액세스 관리
- AWS Clean Rooms 에서 IAM을 사용하는 방법
- 에 대한 자격 증명 기반 정책 예제 AWS Clean Rooms
- AWS 에 대한 관리형 정책 AWS Clean Rooms
- AWS Clean Rooms 자격 증명 및 액세스 문제 해결
- 교차 서비스 혼동된 대리자 예방
- AWS Clean Rooms ML에 대한 IAM 동작
- Clean Rooms ML Custom 모델의 IAM 동작

대상

AWS Identity and Access Management (IAM)를 사용하는 방법은 수행하는 작업에 따라 다릅니다 AWS Clean Rooms.

서비스 사용자 - AWS Clean Rooms 서비스를 사용하여 작업을 수행하는 경우 필요한 자격 증명과 권한을 관리자가 제공합니다. 더 많은 AWS Clean Rooms 기능을 사용하여 작업을 수행하게 되면 추가 권한이 필요할 수 있습니다. 액세스 권한 관리 방법을 이해하면 관리자에게 올바른 권한을 요청하는 데 도움이 됩니다. AWS Clean Rooms의 기능에 액세스할 수 없는 경우 AWS Clean Rooms 자격 증명 및 액세스 문제 해결을 참조하세요.

서비스 관리자 - 회사에서 AWS Clean Rooms 리소스를 책임지고 있는 경우에 대한 전체 액세스 권한을 가지고 있을 것입니다 AWS Clean Rooms. 서비스 관리자는 서비스 사용자가 액세스해야 하는 AWS Clean Rooms 기능과 리소스를 결정합니다. 그런 다음 IAM 관리자에게 요청을 제출하여 서비스 사용자의 권한을 변경해야 합니다. 이 페이지의 정보를 검토하여 IAM의 기본 개념을 이해하세요. 회사가에서 IAM을 사용하는 방법에 대한 자세한 내용은 섹션을 AWS Clean Rooms참조하세요AWS Clean Rooms 에서 IAM을 사용하는 방법.

IAM 관리자 - IAM 관리자라면 AWS Clean Rooms에 대한 액세스 권한 관리 정책 작성 방법을 자세히 알고 싶을 것입니다. IAM에서 사용할 수 있는 자격 AWS Clean Rooms 증명 기반 정책 예제를 보려면 섹션을 참조하세요에 대한 자격 증명 기반 정책 예제 AWS Clean Rooms.

ID를 통한 인증

인증은 자격 증명 자격 증명을 AWS 사용하여에 로그인하는 방법입니다. , AWS 계정 루트 사용자 IAM 사용자 또는 IAM 역할을 수임하여 인증(로그인 AWS)되어야 합니다.

자격 증명 소스를 통해 제공된 자격 증명을 사용하여 페더레이션 자격 증명 AWS 으로에 로그인할 수 있습니다. AWS IAM Identity Center (IAM Identity Center) 사용자 또는 회사의 Single Sign-On 인증은 페더레이션 자격 증명의 예입니다. 페더레이션형 ID로 로그인할 때 관리자가 이전에 IAM 역할을 사용하여 ID 페더레이션을 설정했습니다. 페더레이션을 사용하여 AWS 에 액세스하면 간접적으로 역할을 수임하게 됩니다.

사용자 유형에 따라 AWS Management Console 또는 AWS 액세스 포털에 로그인할 수 있습니다. 로그인에 대한 자세한 내용은 AWS 로그인 사용 설명서의에 로그인하는 방법을 AWS참조하세요. AWS 계정

AWS 프로그래밍 방식으로에 액세스하는 경우는 자격 증명을 사용하여 요청에 암호화 방식으로 서명할 수 있는 소프트웨어 개발 키트(SDK)와 명령줄 인터페이스(CLI)를 AWS 제공합니다. AWS 도구를 사용하지 않는 경우 요청에 직접 서명해야 합니다. 권장 방법을 사용하여 요청에 직접 서명하는 방법에 대한 자세한 내용은 AWS 일반 참조의 Signature Version 4 서명 프로세스를 참조하세요.

사용하는 인증 방법에 상관없이 추가 보안 정보를 제공해야 할 수도 있습니다. 예를 들어, 다중 인증(MFA)을 사용하여 계정의 보안을 강화하는 것이 AWS 좋습니다. 자세한 내용은 AWS IAM Identity Center 사용 설명서의 다중 인증 및 IAM 사용 설명서의 AWS에서 다중 인증(MFA) 사용을 참조하세요.

AWS 계정 루트 사용자

를 생성할 때 계정의 모든 AWS 서비스 및 리소스에 대한 완전한 액세스 권한이 있는 하나의 로그인 자격 증명으로 AWS 계정시작합니다. 이 ID는 AWS 계정 루트 사용자라고 하며, 계정을 생성할 때 사용한 이메일 주소와 암호로 로그인하여 액세스합니다. 일상적인 작업에 루트 사용자를 사용하지 않을 것을 강력히 권장합니다. 루트 사용자 보안 인증 정보를 보호하고 루트 사용자만 수행할 수 있는 작업을 수행하는 데 사용합니다. 루트 사용자로 로그인해야 하는 전체 작업 목록을 보려면 AWS 일반 참조의 AWS 계정 루트 사용자 보안안 인증 및 IAM 자격 증명을 참조하세요.

페더더레이션 자격 증명

가장 좋은 방법은 관리자 액세스가 필요한 사용자를 포함한 인간 사용자에게 자격 증명 공급자와의 페더레이션을 사용하여 임시 자격 증명을 사용하여 AWS 서비스 에 액세스하도록 요구하는 것입니다.

페더레이션 자격 증명은 엔터프라이즈 사용자 디렉터리, 웹 자격 증명 공급자, AWS Directory Service, Identity Center 디렉터리 또는 자격 증명 소스를 통해 제공된 자격 증명을 사용하여 AWS 서비스 에 액세스하는 모든 사용자의 사용자입니다. 페더레이션 자격 증명에 액세스할 때 역할을 AWS 계정수임하고 역할은 임시 자격 증명을 제공합니다.

중앙 집중식 액세스 관리를 위해 AWS IAM Identity Center을(를) 사용하는 것이 좋습니다. IAM Identity Center에서 사용자 및 그룹을 생성하거나 모든 및 애플리케이션에서 사용할 수 있도록 자체 ID 소스의 사용자 AWS 계정 및 그룹 집합에 연결하고 동기화할 수 있습니다. IAM Identity Center에 대한 자세한 내용은 AWS IAM Identity Center 사용 설명서에서 IAM Identity Center란 무엇인가요?를 참조하세요.

IAM 사용자 및 그룹

IAM 사용자는 단일 사용자 또는 애플리케이션에 대한 특정 권한이 AWS 계정 있는 내의 자격 증명입니다. 가능하면 암호 및 액세스 키와 같은 장기 자격 증명이 있는 IAM 사용자를 생성하는 대신 임시 자격 증명을 사용하는 것이 좋습니다. 하지만 IAM 사용자의 장기 자격 증명이 필요한 특정 사용 사례가 있는 경우, 액세스 키를 교체하는 것이 좋습니다. 자세한 내용은 IAM 사용 설명서의 장기 보안 인증이 필요한 사용 사례의 경우, 정기적으로 액세스 키 교체를 참조하세요.

IAM 그룹은 IAM 사용자 컬렉션을 지정하는 자격 증명입니다. 사용자는 그룹으로 로그인할 수 없습니다. 그룹을 사용하여 여러 사용자의 권한을 한 번에 지정할 수 있습니다. 그룹을 사용하면 대규모 사용자 집합의 권한을 더 쉽게 관리할 수 있습니다. 예를 들어, IAMAdmins라는 그룹이 있고 이 그룹에 IAM 리소스를 관리할 권한을 부여할 수 있습니다.

사용자는 역할과 다릅니다. 사용자는 한 사람 또는 애플리케이션과 고유하게 연결되지만, 역할은 해당 역할이 필요한 사람이라면 누구나 수임할 수 있습니다. 사용자는 영구적인 장기 자격 증명을 가지고 있

지만, 역할은 임시 보안 인증만 제공합니다. 자세한 내용은 IAM 사용 설명서에서 IAM 사용자 사용 사례를 참조하세요.

IAM 역할

IAM 역할은 특정 권한이 AWS 계정 있는 내의 자격 증명입니다. IAM 사용자와 유사하지만, 특정 개인과 연결되지 않습니다. 에서 IAM 역할을 일시적으로 수임하려면 사용자에서 IAM 역할(콘솔)로 전환할 AWS Management Console수 있습니다. https://docs.aws.amazon.com/IAM/latest/UserGuide/id_roles_use_switch-role-console.html 또는 AWS API 작업을 호출하거나 사용자 지정 URL을 AWS CLI 사용하여 역할을 수임할 수 있습니다. 역할 사용 방법에 대한 자세한 내용은 IAM 사용 설명서의 역할 수임 방법을 참조하세요.

임시 보안 인증이 있는 IAM 역할은 다음과 같은 상황에서 유용합니다.

- 페더레이션 사용자 액세스 - 페더레이션 ID에 권한을 부여하려면 역할을 생성하고 해당 역할의 권한을 정의합니다. 페더레이션 ID가 인증되면 역할이 연결되고 역할에 정의된 권한이 부여됩니다. 페더레이션 관련 역할에 대한 자세한 내용은 IAM 사용 설명서의 Create a role for a third-party identity provider (federation)를 참조하세요. IAM Identity Center를 사용하는 경우, 권한 집합을 구성합니다. 인증 후 ID가 액세스할 수 있는 항목을 제어하기 위해 IAM Identity Center는 권한 집합을 IAM의 역할과 연관짓습니다. 권한 집합에 대한 자세한 내용은 AWS IAM Identity Center 사용 설명서의 권한 집합을 참조하세요.
- 임시 IAM 사용자 권한 - IAM 사용자 또는 역할은 IAM 역할을 수임하여 특정 작업에 대한 다양한 권한을 임시로 받을 수 있습니다.
- 교차 계정 액세스 - IAM 역할을 사용하여 다른 계정의 사용자(신뢰할 수 있는 보안 주체)가 내 계정의 리소스에 액세스하도록 허용할 수 있습니다. 역할은 계정 간 액세스를 부여하는 기본적인 방법입니다. 그러나 일부 에서는 정책을 리소스에 직접 연결할 AWS 서비스수 있습니다(역할을 프록시로 사용하는 대신). 교차 계정 액세스에 대한 역할과 리소스 기반 정책의 차이점을 알아보려면 IAM 사용 설명서의 IAM의 교차 계정 리소스 액세스를 참조하세요.
- 교차 서비스 액세스 - 일부는 다른에서 기능을 AWS 서비스 사용합니다 AWS 서비스. 예를 들어, 서비스에서 호출하면 일반적으로 해당 서비스는 Amazon EC2에서 애플리케이션을 실행하거나 Amazon S3에 객체를 저장합니다. 서비스는 직접적으로 호출하는 위탁자의 권한을 사용하거나, 서비스 역할을 사용하거나, 또는 서비스 연결 역할을 사용하여 이 작업을 수행할 수 있습니다.
  - 전달 액세스 세션(FAS) - IAM 사용자 또는 역할을 사용하여에서 작업을 수행하는 경우 AWS보안 주체로 간주됩니다. 일부 서비스를 사용하는 경우, 다른 서비스에서 다른 작업을 시작하는 작업을 수행할 수 있습니다. FAS는를 호출하는 보안 주체의 권한을 다운스트림 서비스에 AWS 서비스 대한 요청과 AWS 서비스함께 사용합니다. FAS 요청은 서비스가 완료하기 위해 다른 AWS 서비스 또는 리소스와의 상호 작용이 필요한 요청을 수신할 때만 이루어집니다. 이 경우, 두 작업을 모두

수행할 수 있는 권한이 있어야 합니다. FAS 요청 시 정책 세부 정보는 전달 액세스 세션을 참조하세요.

- 서비스 역할 - 서비스 역할은 서비스가 사용자를 대신하여 작업을 수행하기 위해 맡는 IAM 역할입니다. IAM 관리자는 IAM 내에서 서비스 역할을 생성, 수정 및 삭제할 수 있습니다. 자세한 정보는 IAM 사용 설명서의 Create a role to delegate permissions to an AWS 서비스를 참조하세요.
- 서비스 연결 역할 - 서비스 연결 역할은에 연결된 서비스 역할의 한 유형입니다 AWS 서비스. 서비스는 사용자를 대신하여 작업을 수행하기 위해 역할을 수임할 수 있습니다. 서비스 연결 역할은에 나타나 AWS 계정 며 서비스가 소유합니다. IAM 관리자는 서비스 링크 역할의 권한을 볼 수 있지만 편집은 할 수 없습니다.

- Amazon EC2에서 실행되는 애플리케이션 - IAM 역할을 사용하여 EC2 인스턴스에서 실행되고 AWS CLI 또는 AWS API 요청을 수행하는 애플리케이션의 임시 자격 증명을 관리할 수 있습니다. 이는 EC2 인스턴스 내에 액세스 키를 저장할 때 권장되는 방법입니다. EC2 인스턴스에 AWS 역할을 할당하고 모든 애플리케이션에서 사용할 수 있도록 하려면 인스턴스에 연결된 인스턴스 프로파일을 생성합니다. 인스턴스 프로필에는 역할이 포함되어 있으며 EC2 인스턴스에서 실행되는 프로그램이 임시 보안 인증을 얻을 수 있습니다. 자세한 정보는 IAM 사용 설명서의 IAM 역할을 사용하여 Amazon EC2 인스턴스에서 실행되는 애플리케이션에 권한 부여를 참조하세요.

정책을 사용하여 액세스 관리

정책을 AWS 생성하고 자격 증명 또는 리소스에 연결하여 AWS 에서 액세스를 제어합니다. 정책은 자격 증명 또는 리소스와 연결 AWS 될 때 권한을 정의하는의 객체입니다.는 보안 주체(사용자, 루트 사용자 또는 역할 세션)가 요청할 때 이러한 정책을 AWS 평가합니다. 정책에서 권한은 요청이 허용되거나 거부되는 지를 결정합니다. 대부분의 정책은에 JSON 문서 AWS 로 저장됩니다. JSON 정책 문서의 구조와 콘텐츠에 대한 자세한 정보는 IAM 사용 설명서의 JSON 정책 개요를 참조하세요.

관리자는 AWS JSON 정책을 사용하여 누가 무엇에 액세스할 수 있는지 지정할 수 있습니다. 즉, 어떤 보안 주체가 어떤 리소스와 어떤 조건에서 작업을 수행할 수 있는지를 지정할 수 있습니다.

모든 IAM 개체(사용자 또는 역할)는 처음에는 권한이 없습니다. 기본적으로 사용자는 아무 작업도 수행할 수 없으며, 자신의 암호를 변경할 수도 없습니다. 사용자에게 태스크를 수행할 권한을 부여하기 위해 관리자는 사용자에게 권한 정책을 연결해야 합니다. 또한 관리자는 의도한 권한을 가지고 있는 그룹에 사용자를 추가할 수 있습니다. 관리자가 그룹에 권한을 부여하면 그룹의 모든 사용자가 해당 권한을 받습니다.

IAM 정책은 태스크를 수행하기 위해 사용하는 방법과 상관없이 작업에 대한 권한을 정의합니다. 예를 들어, `iam:GetRole` 작업을 허용하는 정책이 있다고 가정합니다. 해당 정책이 있는 사용자는 AWS Management Console AWS CLI, 또는 AWS API에서 역할 정보를 가져올 수 있습니다.

ID 기반 정책

ID 기반 정책은 IAM 사용자, 사용자 그룹 또는 역할과 같은 ID에 연결할 수 있는 JSON 권한 정책 문서입니다. 이러한 정책은 사용자 및 역할이 어떤 리소스와 어떤 조건에서 어떤 작업을 수행할 수 있는지를 제어합니다. 자격 증명 기반 정책을 생성하는 방법을 알아보려면 IAM 사용 설명서에서 고객 관리형 정책으로 사용자 지정 IAM 권한 정의를 참조하세요.

ID 기반 정책은 인라인 정책 또는 관리형 정책으로 한층 더 분류할 수 있습니다. 인라인 정책은 단일 사용자, 그룹 또는 역할에 직접 포함됩니다. 관리형 정책은 AWS 계정에 속한 다수의 사용자, 그룹 및 역할에 독립적으로 추가할 수 있는 정책입니다. 관리형 정책에는 AWS 관리형 정책 및 고객 관리형 정책이 포함됩니다. 관리형 정책 또는 인라인 정책을 선택하는 방법을 알아보려면 IAM 사용 설명서의 관리형 정책과 인라인 정책의 선택을 참조하세요.

리소스 기반 정책

리소스 기반 정책은 리소스에 연결하는 JSON 정책 설명서입니다. 리소스 기반 정책의 예제는 IAM 역할 신뢰 정책과 Amazon S3 버킷 정책입니다. 리소스 기반 정책을 지원하는 서비스에서 서비스 관리자는 이러한 정책을 사용하여 특정 리소스에 대한 액세스를 통제할 수 있습니다. 정책이 연결된 리소스의 경우 정책은 지정된 위탁자가 해당 리소스와 어떤 조건에서 어떤 작업을 수행할 수 있는지를 정의합니다. 리소스 기반 정책에서 위탁자를 지정해야 합니다. 보안 주체에는 계정, 사용자, 역할, 페더레이션 사용자 또는가 포함될 수 있습니다 AWS 서비스.

리소스 기반 정책은 해당 서비스에 있는 인라인 정책입니다. 리소스 기반 정책에서는 IAM의 AWS 관리형 정책을 사용할 수 없습니다.

기타 정책 타입

AWS 는 덜 일반적인 추가 정책 유형을 지원합니다. 이러한 정책 타입은 더 일반적인 정책 유형에 따라 사용자에게 부여되는 최대 권한을 설정할 수 있습니다.

- 권한 경계 – 권한 경계는 ID 기반 정책에 따라 IAM 엔터티(IAM 사용자 또는 역할)에 부여할 수 있는 최대 권한을 설정하는 고급 기능입니다. 엔터티에 대한 권한 경계를 설정할 수 있습니다. 그 결과로 얻는 권한은 개체의 자격 증명 기반 정책과 그 권한 경계의 교집합입니다. `Principal` 필드에서 사용자나 역할을 보안 주체로 지정하는 리소스 기반 정책은 권한 경계를 통해 제한되지 않습니다. 이러한 정책 중 하나에 포함된 명시적 거부는 허용을 재정의합니다. 권한 경계에 대한 자세한 정보는 IAM 사용 설명서의 IAM 엔티티에 대한 권한 경계를 참조하세요.
- 서비스 제어 정책(SCPs) - SCPs는의 조직 또는 조직 단위(OU)에 대한 최대 권한을 지정하는 JSON 정책입니다 AWS Organizations. AWS Organizations 는 비즈니스가 소유 AWS 계정 한 여러을 그

룹화하고 중앙에서 관리하기 위한 서비스입니다. 조직에서 모든 기능을 활성화할 경우, 서비스 제어 정책(SCP)을 임의의 또는 모든 계정에 적용할 수 있습니다. SCP는 각각을 포함하여 멤버 계정의 엔터티에 대한 권한을 제한합니다 AWS 계정 루트 사용자. 조직 및 SCP에 대한 자세한 정보는 AWS Organizations 사용 설명서의 SCP 작동 방식을 참조하십시오.

- 세션 정책 – 세션 정책은 역할 또는 페더레이션 사용자에 대해 임시 세션을 프로그래밍 방식으로 생성할 때 파라미터로 전달하는 고급 정책입니다. 결과적으로 얻는 세션의 권한은 사용자 또는 역할의 ID 기반 정책의 교차와 세션 정책입니다. 또한 권한을 리소스 기반 정책에서 가져올 수도 있습니다. 이러한 정책 중 하나에 포함된 명시적 거부는 허용을 재정의합니다. 자세한 정보는 IAM 사용 설명서의 세션 정책을 참조하세요.

여러 정책 유형

여러 정책 유형이 요청에 적용되는 경우, 결과 권한은 이해하기가 더 복잡합니다. 에서 여러 정책 유형이 관련될 때 요청을 허용할지 여부를 AWS 결정하는 방법을 알아보려면 IAM 사용 설명서의 정책 평가 로직을 참조하세요.

AWS Clean Rooms 에서 IAM을 사용하는 방법

IAM을 사용하여에 대한 액세스를 관리하기 전에 사용할 수 있는 IAM 기능에 대해 AWS Clean Rooms 알아봅니다 AWS Clean Rooms.

에서 사용할 수 있는 IAM 기능 AWS Clean Rooms

IAM 기능	AWS Clean Rooms 지원
ID 기반 정책	예
리소스 기반 정책	부분
정책 작업	예
정책 리소스	예
정책 조건 키(서비스별)	부분
ACLs	아니요
ABAC(정책의 태그)	예

IAM 기능	AWS Clean Rooms 지원
임시 보안 인증	예
전달 액세스 세션(FAS)	예
서비스 역할	예
서비스 연결 역할	아니요

AWS Clean Rooms 및 기타에서 대부분의 IAM 기능을 AWS 서비스 사용하는 방법을 개괄적으로 알아보려면 IAM 사용 설명서의 AWS 서비스 IAM으로 작업하는 섹션을 참조하세요.

에 대한 자격 증명 기반 정책 AWS Clean Rooms

ID 기반 정책 지원: 예

ID 기반 정책은 IAM 사용자, 사용자 그룹 또는 역할과 같은 ID에 연결할 수 있는 JSON 권한 정책 문서입니다. 이러한 정책은 사용자 및 역할이 어떤 리소스와 어떤 조건에서 어떤 작업을 수행할 수 있는지를 제어합니다. 자격 증명 기반 정책을 생성하는 방법을 알아보려면 IAM 사용 설명서에서 고객 관리형 정책으로 사용자 지정 IAM 권한 정의를 참조하세요.

IAM ID 기반 정책을 사용하면 허용되거나 거부되는 작업과 리소스뿐 아니라 작업이 허용되거나 거부되는 조건을 지정할 수 있습니다. ID 기반 정책에서는 위탁자가 연결된 사용자 또는 역할에 적용되므로 위탁자를 지정할 수 없습니다. JSON 정책에서 사용하는 모든 요소에 대해 알아보려면 IAM 사용 설명서의 IAM JSON 정책 요소 참조를 참조하세요.

에 대한 자격 증명 기반 정책 예제 AWS Clean Rooms

자격 AWS Clean Rooms 증명 기반 정책의 예를 보려면 섹션을 참조하세요에 대한 자격 증명 기반 정책 예제 AWS Clean Rooms.

내의 리소스 기반 정책 AWS Clean Rooms

리소스 기반 정책 지원: 부분적

리소스 기반 정책은 리소스에 연결하는 JSON 정책 설명서입니다. 리소스 기반 정책의 예제는 IAM 역할 신뢰 정책과 Amazon S3 버킷 정책입니다. 리소스 기반 정책을 지원하는 서비스에서 서비스 관리자는 이러한 정책을 사용하여 특정 리소스에 대한 액세스를 통제할 수 있습니다. 정책이 연결된 리소스의

경우 정책은 지정된 위탁자가 해당 리소스와 어떤 조건에서 어떤 작업을 수행할 수 있는지를 정의합니다. 리소스 기반 정책에서 위탁자를 지정해야 합니다. 보안 주체에는 계정, 사용자, 역할, 페더레이션 사용자 또는가 포함될 수 있습니다 AWS 서비스.

교차 계정 액세스를 활성화하려는 경우, 전체 계정이나 다른 계정의 IAM 개체를 리소스 기반 정책의 위탁자로 지정할 수 있습니다. 리소스 기반 정책에 크로스 계정 보안 주체를 추가하는 것은 트러스트 관계 설정의 절반밖에 되지 않는다는 것을 유념하세요. 보안 주체와 리소스가 다른 경우 신뢰할 수 AWS 계정있는 계정의 IAM 관리자는 보안 주체 엔터티(사용자 또는 역할)에게 리소스에 액세스할 수 있는 권한도 부여해야 합니다. 엔터티에 ID 기반 정책을 연결하여 권한을 부여합니다. 하지만 리소스 기반 정책이 동일 계정의 위탁자에 액세스를 부여하는 경우, 추가 자격 증명 기반 정책이 필요하지 않습니다. 자세한 내용은 IAM 사용 설명서의 교차 계정 리소스 액세스를 참조하세요.

이 AWS Clean Rooms 서비스는 구성된 유사 모델에 연결된 구성된 유사 모델 관리형 리소스 정책이라는 한 가지 유형의 리소스 기반 정책만 지원합니다. 이 정책은 구성된 유사 모델에 대해 작업을 수행할 수 있는 보안 주체를 정의합니다.

구성된 유사 모델에 리소스 기반 정책을 연결하는 방법은 AWS Clean Rooms ML에 대한 IAM 동작 단원을 참조하세요.

에 대한 정책 작업 AWS Clean Rooms

정책 작업 지원: 예

관리자는 AWS JSON 정책을 사용하여 누가 무엇에 액세스할 수 있는지 지정할 수 있습니다. 즉, 어떤 위탁자가 어떤 리소스와 어떤 조건에서 작업을 수행할 수 있는지를 지정할 수 있습니다.

JSON 정책의 `Action` 요소는 정책에서 액세스를 허용하거나 거부하는 데 사용할 수 있는 작업을 설명합니다. 정책 작업은 일반적으로 연결된 AWS API 작업과 이름이 동일합니다. 일치하는 API 작업이 없는 권한 전용 작업 같은 몇 가지 예외도 있습니다. 정책에서 여러 작업이 필요한 몇 가지 작업도 있습니다. 이러한 추가 작업을 일컬어 종속 작업이라고 합니다.

연결된 작업을 수행할 수 있는 권한을 부여하기 위한 정책에 작업을 포함하세요.

AWS Clean Rooms 작업 목록을 보려면 서비스 승인 참조의에서 정의한 작업을 AWS Clean Rooms 참조하세요.

의 정책 작업은 작업 앞에 다음 접두사를 AWS Clean Rooms 사용합니다.

```
cleanrooms
```

단일 문에서 여러 작업을 지정하려면 다음과 같이 쉼표로 구분합니다.

```
"Action": [
      "cleanrooms:action1",
      "cleanrooms:action2"
         ]
```

자격 AWS Clean Rooms 증명 기반 정책의 예를 보려면 섹션을 참조하세요에 대한 자격 증명 기반 정책 예제 AWS Clean Rooms.

에 대한 정책 리소스 AWS Clean Rooms

정책 리소스 지원: 예

관리자는 AWS JSON 정책을 사용하여 누가 무엇에 액세스할 수 있는지 지정할 수 있습니다. 즉, 어떤 보안 주체가 어떤 리소스와 어떤 조건에서 작업을 수행할 수 있는지를 지정할 수 있습니다.

`Resource` JSON 정책 요소는 작업이 적용되는 하나 이상의 객체를 지정합니다. 문에는 `Resource`또는 `NotResource`요소가 반드시 추가되어야 합니다. 모범 사례에 따라 Amazon 리소스 이름(ARN)을 사용하여 리소스를 지정합니다. 리소스 수준 권한이라고 하는 특정 리소스 유형을 지원하는 작업에 대해 이를 수행할 수 있습니다.

작업 나열과 같이 리소스 수준 권한을 지원하지 않는 작업의 경우, 와일드카드(*)를 사용하여 해당 문이 모든 리소스에 적용됨을 나타냅니다.

```
"Resource": "*"
```

AWS Clean Rooms 리소스 유형 및 해당 ARNs 목록을 보려면 서비스 승인 참조의에서 정의한 리소스를 AWS Clean Rooms 참조하세요. 각 리소스의 ARN을 지정할 수 있는 작업을 알아보려면 AWS Clean Rooms가 정의한 작업을 참조하십시오.

자격 AWS Clean Rooms 증명 기반 정책의 예를 보려면 섹션을 참조하세요에 대한 자격 증명 기반 정책 예제 AWS Clean Rooms.

에 대한 정책 조건 키 AWS Clean Rooms

서비스별 정책 조건 키 지원: 부분

관리자는 AWS JSON 정책을 사용하여 누가 무엇에 액세스할 수 있는지 지정할 수 있습니다. 즉, 어떤 보안 주체가 어떤 리소스와 어떤 조건에서 작업을 수행할 수 있는지를 지정할 수 있습니다.

`Condition` 요소(또는 `Condition` 블록)를 사용하면 정책이 발효되는 조건을 지정할 수 있습니다. `Condition` 요소는 옵션입니다. 같거나 작음과 같은 조건 연산자를 사용하여 정책의 조건을 요청의 값과 일치시키는 조건식을 생성할 수 있습니다.

한 문에서 여러 `Condition` 요소를 지정하거나 단일 `Condition` 요소에서 여러 키를 지정하는 경우, AWS 는 논리적 `AND` 작업을 사용하여 평가합니다. 단일 조건 키에 여러 값을 지정하는 경우는 논리적 `OR` 작업을 사용하여 조건을 AWS 평가합니다. 문의 권한을 부여하기 전에 모든 조건을 충족해야 합니다.

조건을 지정할 때 자리 표시자 변수를 사용할 수도 있습니다. 예를 들어, IAM 사용자에게 IAM 사용자 이름으로 태그가 지정된 경우에만 리소스에 액세스할 수 있는 권한을 부여할 수 있습니다. 자세한 내용은 IAM 사용 설명서의 IAM 정책 요소: 변수 및 태그를 참조하세요.

AWS 는 전역 조건 키와 서비스별 조건 키를 지원합니다. 모든 AWS 전역 조건 키를 보려면 IAM 사용 설명서의 AWS 전역 조건 컨텍스트 키를 참조하세요.

AWS Clean Rooms ML이 정책 조건 키를 사용하는 방법을 알아보려면 섹션을 참조하세요AWS Clean Rooms ML에 대한 IAM 동작.

ACLs AWS Clean Rooms

ACL 지원: 아니요

액세스 제어 목록(ACL)은 어떤 위탁자(계정 멤버, 사용자 또는 역할)가 리소스에 액세스할 수 있는 권한을 가지고 있는지를 제어합니다. ACL은 JSON 정책 문서 형식을 사용하지 않지만 리소스 기반 정책과 유사합니다.

를 사용한 ABAC AWS Clean Rooms

ABAC 지원(정책의 태그): 예

속성 기반 액세스 제어(ABAC)는 속성에 근거하여 권한을 정의하는 권한 부여 전략입니다. 에서는 AWS이러한 속성을 태그라고 합니다. IAM 엔터티(사용자 또는 역할) 및 많은 AWS 리소스에 태그를 연결할 수 있습니다. ABAC의 첫 번째 단계로 개체 및 리소스에 태그를 지정합니다. 그런 다음 위탁자의 태그가 액세스하려는 리소스의 태그와 일치할 때 작업을 허용하도록 ABAC 정책을 설계합니다.

ABAC는 빠르게 성장하는 환경에서 유용하며 정책 관리가 번거로운 상황에 도움이 됩니다.

태그에 근거하여 액세스를 제어하려면 `aws:ResourceTag/key-name`, `aws:RequestTag/key-name` 또는 `aws:TagKeys` 조건 키를 사용하여 정책의 조건 요소에 태그 정보를 제공합니다.

서비스가 모든 리소스 유형에 대해 세 가지 조건 키를 모두 지원하는 경우, 값은 서비스에 대해 예입니다. 서비스가 일부 리소스 유형에 대해서만 세 가지 조건 키를 모두 지원하는 경우, 값은 부분적입니다.

ABAC에 대한 자세한 내용은 IAM 사용 설명서의 ABAC 권한 부여를 통한 권한 정의를 참조하세요. ABAC 설정 단계가 포함된 자습서를 보려면 IAM 사용 설명서의 속성 기반 액세스 제어(ABAC) 사용을 참조하세요.

에서 임시 자격 증명 사용 AWS Clean Rooms

임시 자격 증명 지원: 예

일부 AWS 서비스 는 임시 자격 증명을 사용하여 로그인할 때 작동하지 않습니다. 임시 자격 증명으로 AWS 서비스 작업하는를 비롯한 추가 정보는 AWS 서비스 IAM 사용 설명서의 IAM으로 작업하는 섹션을 참조하세요.

사용자 이름과 암호를 제외한 방법을 AWS Management Console 사용하여에 로그인하는 경우 임시 자격 증명을 사용합니다. 예를 들어 회사의 SSO(Single Sign-On) 링크를 AWS 사용하여에 액세스하면 해당 프로세스가 임시 자격 증명을 자동으로 생성합니다. 또한 콘솔에 사용자로 로그인한 다음 역할을 전환할 때 임시 자격 증명을 자동으로 생성합니다. 역할 전환에 대한 자세한 내용은 IAM 사용 설명서의 사용자에서 IAM 역할로 전환(콘솔)을 참조하세요.

AWS CLI 또는 AWS API를 사용하여 임시 자격 증명을 수동으로 생성할 수 있습니다. 그런 다음 이러한 임시 자격 증명을 사용하여 장기 액세스 키를 사용하는 대신 임시 자격 증명을 동적으로 생성하는 access AWS. AWS recommends에 액세스할 수 있습니다. 자세한 정보는 IAM의 임시 보안 자격 증명 섹션을 참조하세요.

에 대한 전달 액세스 세션 AWS Clean Rooms

전달 액세스 세션(FAS) 지원: 예

IAM 사용자 또는 역할을 사용하여에서 작업을 수행하는 경우 AWS보안 주체로 간주됩니다. 일부 서비스를 사용하는 경우, 다른 서비스에서 다른 작업을 시작하는 작업을 수행할 수 있습니다. FAS는를 호출하는 보안 주체의 권한을 다운스트림 서비스에 AWS 서비스 대한 요청과 AWS 서비스함께 사용합니다. FAS 요청은 서비스가 완료하기 위해 다른 AWS 서비스 또는 리소스와의 상호 작용이 필요한 요청을 수신할 때만 이루어집니다. 이 경우, 두 작업을 모두 수행할 수 있는 권한이 있어야 합니다. FAS 요청 시 정책 세부 정보는 전달 액세스 세션을 참조하세요.

AWS Clean Rooms의 서비스 역할

서비스 역할 지원: 예

서비스 역할은 서비스가 사용자를 대신하여 작업을 수행하는 것으로 가정하는 IAM 역할입니다. IAM 관리자는 IAM 내에서 서비스 역할을 생성, 수정 및 삭제할 수 있습니다. 자세한 정보는 IAM 사용 설명서의 Create a role to delegate permissions to an AWS 서비스를 참조하세요.

> **Warning**
>
> 서비스 역할에 대한 권한을 변경하면 AWS Clean Rooms 기능이 중단될 수 있습니다. 에서 관련 지침을 AWS Clean Rooms 제공하는 경우에만 서비스 역할을 편집합니다.

에 대한 서비스 연결 역할 AWS Clean Rooms

서비스 링크 역할 지원: 아니요

서비스 연결 역할은에 연결된 서비스 역할의 한 유형입니다 AWS 서비스. 서비스는 사용자를 대신하여 작업을 수행하기 위해 역할을 수임할 수 있습니다. 서비스 연결 역할은에 표시 AWS 계정 되며 서비스가 소유합니다. IAM 관리자는 서비스 링크 역할의 권한을 볼 수 있지만 편집은 할 수 없습니다.

서비스 연결 역할 생성 또는 관리에 대한 자세한 내용은 IAM으로 작업하는AWS 서비스를 참조하세요. 서비스 연결 역할 열에서 `Yes`이(가) 포함된 서비스를 테이블에서 찾습니다. 해당 서비스에 대한 서비스 연결 역할 설명서를 보려면 예 링크를 선택합니다.

에 대한 자격 증명 기반 정책 예제 AWS Clean Rooms

기본적으로 사용자 및 역할에는 AWS Clean Rooms 리소스를 생성하거나 수정할 수 있는 권한이 없습니다. 또한 AWS Management Console, AWS Command Line Interface (AWS CLI) 또는 AWS API를 사용하여 작업을 수행할 수 없습니다. 사용자에게 사용자가 필요한 리소스에서 작업을 수행할 권한을 부여하려면 IAM 관리자가 IAM 정책을 생성하면 됩니다. 그런 다음 관리자가 IAM 정책을 역할에 추가하고, 사용자가 역할을 맡을 수 있습니다.

이러한 예제 JSON 정책 문서를 사용하여 IAM 자격 증명 기반 정책을 생성하는 방법을 알아보려면 IAM 사용 설명서의 IAM 정책 생성(콘솔)을 참조하세요.

각 리소스 유형에 대한 ARNs 형식을 포함하여 AWS Clean Rooms에서 정의한 작업 및 리소스 유형에 대한 자세한 내용은 서비스 승인 참조의 에 대한 작업, 리소스 및 조건 키를 AWS Clean Rooms 참조하세요.

주제

- 정책 모범 사례
- AWS Clean Rooms 콘솔 사용
- 사용자가 자신의 고유한 권한을 볼 수 있도록 허용

정책 모범 사례

자격 증명 기반 정책에 따라 계정에서 사용자가 AWS Clean Rooms 리소스를 생성, 액세스 또는 삭제할 수 있는지 여부가 결정됩니다. 이 작업으로 인해 AWS 계정에 비용이 발생할 수 있습니다. ID 기반 정책을 생성하거나 편집할 때는 다음 지침과 권장 사항을 따릅니다.

- AWS 관리형 정책을 시작하고 최소 권한으로 전환 - 사용자 및 워크로드에 권한 부여를 시작하려면 많은 일반적인 사용 사례에 대한 권한을 부여하는 AWS 관리형 정책을 사용합니다. 에서 사용할 수 있습니다 AWS 계정. 사용 사례에 맞는 AWS 고객 관리형 정책을 정의하여 권한을 추가로 줄이는 것이 좋습니다. 자세한 정보는 IAM 사용 설명서의 AWS 관리형 정책 또는 AWS 직무에 대한 관리형 정책을 참조하세요.
- 최소 권한 적용 – IAM 정책을 사용하여 권한을 설정하는 경우, 작업을 수행하는 데 필요한 권한만 부여합니다. 이렇게 하려면 최소 권한으로 알려진 특정 조건에서 특정 리소스에 대해 수행할 수 있는 작업을 정의합니다. IAM을 사용하여 권한을 적용하는 방법에 대한 자세한 정보는 IAM 사용 설명서에 있는 IAM의 정책 및 권한을 참조하세요.
- IAM 정책의 조건을 사용하여 액세스 추가 제한 – 정책에 조건을 추가하여 작업 및 리소스에 대한 액세스를 제한할 수 있습니다. 예를 들어, SSL을 사용하여 모든 요청을 전송해야 한다고 지정하는 정책 조건을 작성할 수 있습니다. AWS 서비스와 같은 특정를 통해 사용되는 경우 조건을 사용하여 서비스 작업에 대한 액세스 권한을 부여할 수도 있습니다 AWS CloudFormation. 자세한 정보는 IAM 사용 설명서의 IAM JSON 정책 요소: 조건을 참조하세요.
- IAM Access Analyzer를 통해 IAM 정책을 확인하여 안전하고 기능적인 권한 보장 - IAM Access Analyzer에서는 IAM 정책 언어(JSON)와 모범 사례가 정책에서 준수되도록 새로운 및 기존 정책을 확인합니다. IAM Access Analyzer는 100개 이상의 정책 확인 항목과 실행 가능한 추천을 제공하여 안전하고 기능적인 정책을 작성하도록 돕습니다. 자세한 내용은 IAM 사용 설명서의 IAM Access Analyzer에서 정책 검증을 참조하세요.
- 다중 인증(MFA) 필요 -에서 IAM 사용자 또는 루트 사용자가 필요한 시나리오가 있는 경우 추가 보안을 위해 MFA를 AWS 계정켭니다. API 작업을 직접 호출할 때 MFA가 필요하면 정책에 MFA 조건을 추가합니다. 자세한 내용은 IAM 사용 설명서의 MFA를 통한 보안 API 액세스를 참조하세요.

IAM의 모범 사례에 대한 자세한 내용은 IAM 사용 설명서의 IAM의 보안 모범 사례를 참조하세요.

AWS Clean Rooms 콘솔 사용

AWS Clean Rooms 콘솔에 액세스하려면 최소 권한 집합이 있어야 합니다. 이러한 권한은의 AWS Clean Rooms 리소스에 대한 세부 정보를 나열하고 볼 수 있도록 허용해야 합니다 AWS 계정. 최소 필수 권한보다 더 제한적인 ID 기반 정책을 생성하는 경우, 콘솔이 해당 정책에 연결된 엔티티(사용자 또는 역할)에 대해 의도대로 작동하지 않습니다.

AWS CLI 또는 AWS API만 호출하는 사용자에게 최소 콘솔 권한을 허용할 필요는 없습니다. 대신 수행하려는 API 작업과 일치하는 작업에만 액세스할 수 있도록 합니다.

사용자와 역할이 AWS Clean Rooms 콘솔을 계속 사용할 수 있도록 하려면 AWS Clean Rooms *FullAccess* 또는 *ReadOnly* AWS 관리형 정책도 엔터티에 연결합니다. 자세한 내용은 IAM 사용 설명서의 사용자에게 권한 추가를 참조하세요.

사용자가 자신의 고유한 권한을 볼 수 있도록 허용

이 예제는 IAM 사용자가 자신의 사용자 ID에 연결된 인라인 및 관리형 정책을 볼 수 있도록 허용하는 정책을 생성하는 방법을 보여줍니다. 이 정책에는 콘솔에서 또는 AWS CLI 또는 AWS API를 사용하여 프로그래밍 방식으로이 작업을 완료할 수 있는 권한이 포함됩니다.

```
{
    "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 에 대한 관리형 정책 AWS Clean Rooms

AWS 관리형 정책은에서 생성하고 관리하는 독립 실행형 정책입니다 AWS. AWS 관리형 정책은 사용자, 그룹 및 역할에 권한 할당을 시작할 수 있도록 많은 일반적인 사용 사례에 대한 권한을 제공하도록 설계되었습니다.

AWS 관리형 정책은 모든 AWS 고객이 사용할 수 있으므로 특정 사용 사례에 대해 최소 권한을 부여하지 않을 수 있습니다. 사용 사례에 고유한 고객 관리형 정책을 정의하여 권한을 줄이는 것이 좋습니다.

AWS 관리형 정책에 정의된 권한은 변경할 수 없습니다. 가 관리형 정책에 정의된 권한을 AWS 업데이트하는 AWS 경우 업데이트는 정책이 연결된 모든 보안 주체 자격 증명(사용자, 그룹 및 역할)에 영향을 줍니다. AWS 서비스 는 새가 시작되거나 기존 서비스에 새 API 작업을 사용할 수 있게 될 때 AWS 관리형 정책을 업데이트할 가능성이 AWS 높습니다.

자세한 내용은 IAM 사용 설명서의 AWS 관리형 정책을 참조하세요.

AWS 관리형 정책: **AWSCleanRoomsReadOnlyAccess**

`AWSCleanRoomsReadOnlyAccess`를 IAM 보안 주체에 연결할 수 있습니다.

이 정책은 `AWSCleanRoomsReadOnlyAccess` 공동 작업의 리소스 및 메타데이터에 대해 읽기 전용 액세스 권한을 부여합니다.

권한 세부 정보

이 정책에는 다음 권한이 포함되어 있습니다.

- `CleanRoomsRead`— 보안 주체가 서비스에 대한 읽기 전용 액세스를 허용합니다.

- `ConsoleDisplayTables` - 보안 주체가 콘솔의 기본 AWS Glue 테이블에 대한 데이터를 표시하는 데 필요한 AWS Glue 메타데이터에 대한 읽기 전용 액세스를 허용합니다.
- `ConsoleLogSummaryQueryLogs`— 보안 주체가 쿼리 로그를 볼 수 있도록 허용합니다.
- `ConsoleLogSummaryObtainLogs`— 보안 주체가 로그 결과를 검색할 수 있습니다.

정책 세부 정보의 JSON 목록은 AWS 관리형 정책 참조 안내서의 AWSCleanRoomsReadOnlyAccess를 참조하세요.

AWS 관리형 정책: **`AWSCleanRoomsFullAccess`**

`AWSCleanRoomsFullAccess`를 IAM 보안 주체에 연결할 수 있습니다.

이 정책은 AWS Clean Rooms 공동 작업의 리소스 및 메타데이터에 대한 전체 액세스(읽기, 쓰기 및 업데이트)를 허용하는 관리 권한을 부여합니다. 이 정책에는 쿼리를 수행할 수 있는 액세스 권한이 포함됩니다.

권한 세부 정보

이 정책에는 다음 권한이 포함되어 있습니다.

- `CleanRoomsAccess` - 모든 리소스의 모든 작업에 대한 전체 액세스 권한을 부여합니다 AWS Clean Rooms.
- `PassServiceRole`— 이름에 "cleanrooms“이 있는 서비스(`PassedToService`조건)에만 서비스 역할을 전달할 수 있는 액세스 권한을 부여합니다.
- `ListRolesToPickServiceRole` - 보안 주체가 사용 시 서비스 역할을 선택하기 위해 모든 역할을 나열할 수 있도록 허용합니다 AWS Clean Rooms.
- `GetRoleAndListRolePoliciesToInspectServiceRole`— 보안 주체가 IAM의 서비스 역할과 해당 정책을 볼 수 있도록 허용합니다.
- `ListPoliciesToInspectServiceRolePolicy`— 보안 주체가 IAM의 서비스 역할과 해당 정책을 볼 수 있도록 허용합니다.
- `GetPolicyToInspectServiceRolePolicy`— 주체가 IAM의 서비스 역할과 해당 정책을 볼 수 있도록 허용합니다.
- `ConsoleDisplayTables` - 보안 주체가 콘솔의 기본 AWS Glue 테이블에 대한 데이터를 표시하는 데 필요한 AWS Glue 메타데이터에 대한 읽기 전용 액세스를 허용합니다.
- `ConsolePickQueryResultsBucketListAll`— 보안 주체가 사용 가능한 모든 S3 버킷 목록에서 Amazon S3 버킷(쿼리 결과가 작성되는)을 선택할 수 있도록 허용합니다.
- `SetQueryResultsBucket`— 보안 주체가 쿼리 결과를 기록되는 S3 버킷을 선택할 수 있습니다.

- `ConsoleDisplayQueryResults`— 보안 주체가 S3 버킷에서 읽은 쿼리 결과를 고객에게 표시할 수 있습니다.
- `WriteQueryResults`— 보안 주체가 쿼리 결과를 고객 소유의 S3 버킷에 쓸 수 있습니다.
- `EstablishLogDeliveries`— 보안 주체가 고객의 Amazon CloudWatch Logs 로그 그룹에 쿼리 로그를 전송할 수 있습니다.
- `SetupLogGroupsDescribe`— 보안 주체가 Amazon CloudWatch Logs 로그 그룹 생성 프로세스를 사용할 수 있습니다.
- `SetupLogGroupsCreate` – 보안 주체가 Amazon CloudWatch Logs 로그 그룹을 생성, 관리 및 설명할 수 있습니다.
- `SetupLogGroupsResourcePolicy`— 보안 주체가 Amazon CloudWatch Logs 로그 그룹에 대한 리소스 정책을 설정할 수 있습니다.
- `ConsoleLogSummaryQueryLogs`— 보안 주체가 쿼리 로그를 볼 수 있도록 허용합니다.
- `ConsoleLogSummaryObtainLogs`— 보안 주체가 로그 결과를 검색할 수 있습니다.

정책 세부 정보의 JSON 목록은 AWS 관리형 정책 참조 안내서의 AWSCleanRoomsFullAccess를 참조하세요.

AWS 관리형 정책: **`AWSCleanRoomsFullAccessNoQuerying`**

`AWSCleanRoomsFullAccessNoQuerying`을 IAM principals에 첨부할 수 있습니다.

이 정책은 AWS Clean Rooms 공동 작업의 리소스 및 메타데이터에 대한 전체 액세스(읽기, 쓰기 및 업데이트)를 허용하는 관리 권한을 부여합니다. 이 정책은 쿼리를 수행할 수 있는 액세스를 제외합니다.

권한 세부 정보

이 정책에는 다음 권한이 포함되어 있습니다.

- `CleanRoomsAccess` - 공동 작업에서 쿼리를 AWS Clean Rooms제외하고에 대한 모든 리소스의 모든 작업에 대한 전체 액세스 권한을 부여합니다.
- `CleanRoomsNoQuerying`— 쿼리를 방지하기 위해 `StartProtectedQuery`와 `UpdateProtectedQuery`를 명시적으로 거부합니다.
- `PassServiceRole`— 이름에 "cleanrooms“이 있는 서비스(`PassedToService`조건)에만 서비스 역할을 전달할 수 있는 액세스 권한을 부여합니다.
- `ListRolesToPickServiceRole` - 보안 주체가 사용 시 서비스 역할을 선택하기 위해 모든 역할을 나열할 수 있도록 허용합니다 AWS Clean Rooms.

- `GetRoleAndListRolePoliciesToInspectServiceRole`— 보안 주체가 IAM의 서비스 역할과 해당 정책을 볼 수 있도록 허용합니다.
- `ListPoliciesToInspectServiceRolePolicy`— 보안 주체가 IAM의 서비스 역할과 해당 정책을 볼 수 있도록 허용합니다.
- `GetPolicyToInspectServiceRolePolicy`— 주체가 IAM의 서비스 역할과 해당 정책을 볼 수 있도록 허용합니다.
- `ConsoleDisplayTables` - 보안 주체가 콘솔의 기본 AWS Glue 테이블에 대한 데이터를 표시하는 데 필요한 AWS Glue 메타데이터에 대한 읽기 전용 액세스를 허용합니다.
- `EstablishLogDeliveries`— 보안 주체가 고객의 Amazon CloudWatch Logs 로그 그룹에 쿼리 로그를 전송할 수 있습니다.
- `SetupLogGroupsDescribe`— 보안 주체가 Amazon CloudWatch Logs 로그 그룹 생성 프로세스를 사용할 수 있습니다.
- `SetupLogGroupsCreate` – 보안 주체가 Amazon CloudWatch Logs 로그 그룹을 생성, 관리 및 설명할 수 있습니다.
- `SetupLogGroupsResourcePolicy`— 보안 주체가 Amazon CloudWatch Logs 로그 그룹에 대한 리소스 정책을 설정할 수 있습니다.
- `ConsoleLogSummaryQueryLogs`— 보안 주체가 쿼리 로그를 볼 수 있도록 허용합니다.
- `ConsoleLogSummaryObtainLogs`— 보안 주체가 로그 결과를 검색할 수 있습니다.
- `cleanrooms` — AWS Clean Rooms 서비스 내에서 공동 작업, 분석 템플릿, 구성된 테이블, 멤버십 및 관련 리소스를 관리합니다. 이러한 리소스에 대한 정보를 생성, 업데이트, 삭제, 나열 및 검색하는 등의 다양한 작업을 수행합니다.
- `iam` - 이름이 "`cleanrooms`"인 서비스 역할을 AWS Clean Rooms 서비스에 전달합니다. 역할, 정책을 나열하고 서비스와 관련된 AWS Clean Rooms 서비스 역할 및 정책을 검사합니다.
- `glue` - 데이터베이스, 테이블, 파티션 및 스키마에 대한 정보를 검색합니다 AWS Glue. 이는 AWS Clean Rooms 서비스가 기본 데이터 소스를 표시하고 상호 작용하는 데 필요합니다.
- `logs` — CloudWatch Logs에 대한 로그 전송, 로그 그룹 및 리소스 정책을 관리합니다. AWS Clean Rooms 서비스와 관련된 로그를 쿼리하고 검색합니다. 이러한 권한은 서비스 내에서 모니터링, 감사 및 문제 해결을 위해 필요합니다.

이 정책은 또한 사용자가 보호된 쿼리를 직접 실행하거나 업데이트하지 못하도록 `cleanrooms:StartProtectedQuery` 및 `cleanrooms:UpdateProtectedQuery` 동작을 명시적으로 거부합니다. 이 동작은 AWS Clean Rooms 제어 메커니즘을 통해 수행해야 합니다.

정책 세부 정보의 JSON 목록은 AWS 관리형 정책 참조 안내서의 AWSCleanRoomsFullAccessNoQuerying을 참조하세요.

AWS 관리형 정책: **AWSCleanRoomsMLReadOnlyAccess**

AWSCleanRoomsMLReadOnlyAccess를 IAM 보안 주체에 연결할 수 있습니다.

이 정책은 AWSCleanRoomsMLReadOnlyAccess 공동 작업의 리소스 및 메타데이터에 대해 읽기 전용 액세스 권한을 부여합니다.

이 정책에는 다음 권한이 포함되어 있습니다.

- CleanRoomsConsoleNavigation - AWS Clean Rooms 콘솔의 화면을 볼 수 있는 액세스 권한을 부여합니다.
- CleanRoomsMLRead — 보안 주체가 Clean Rooms ML 서비스에 대한 읽기 전용 액세스를 허용합니다.
- PassCleanRoomsResources - 지정된 AWS Clean Rooms 리소스를 전달할 수 있는 액세스 권한을 부여합니다.

정책 세부 정보의 JSON 목록은 AWS 관리형 정책 참조 안내서의 AWSCleanRoomsMLReadOnlyAccess를 참조하세요.

AWS 관리형 정책: **AWSCleanRoomsMLFullAccess**

AWSCleanRoomsMLFullAcces를 IAM 보안 주체에 연결할 수 있습니다. 이 정책은 Clean Rooms ML에서 필요한 리소스와 메타데이터에 대한 전체 액세스(읽기, 쓰기, 업데이트)를 허용하는 관리 권한을 부여합니다.

권한 세부 정보

이 정책에는 다음 권한이 포함되어 있습니다.

- CleanRoomsMLFullAccess — 모든 Clean Rooms ML 작업에 대한 액세스 권한을 부여합니다.
- PassServiceRole— 이름에 "cleanrooms-ml“이 있는 서비스(PassedToService조건)에만 서비스 역할을 전달할 수 있는 액세스 권한을 부여합니다.
- CleanRoomsConsoleNavigation - AWS Clean Rooms 콘솔의 화면을 볼 수 있는 액세스 권한을 부여합니다.
- CollaborationMembershipCheck - 공동 작업 내에서 대상 생성(유사 세그먼트) 작업을 시작하면 Clean Rooms ML 서비스가를 호출ListMembers하여 공동 작업이 유효한지, 호출자가 활성 멤

버인지, 구성된 대상 모델 소유자가 활성 멤버인지 확인합니다. 이 권한은 항상 필요하며 콘솔 탐색 SID는 콘솔 사용자에게만 필요합니다.

- `PassCleanRoomsResources` - 지정된 AWS Clean Rooms 리소스를 전달할 수 있는 액세스 권한을 부여합니다.
- `AssociateModels` — 보안 주체가 Clean Rooms ML 모델을 공동 작업과 연결할 수 있습니다.
- `TagAssociations`— 보안 주체가 유사 모델과 공동 작업 간의 연관성에 태그를 추가할 수 있습니다.
- `ListRolesToPickServiceRole` - 보안 주체가 사용 시 서비스 역할을 선택하기 위해 모든 역할을 나열할 수 있도록 허용합니다 AWS Clean Rooms.
- `GetRoleAndListRolePoliciesToInspectServiceRole`— 보안 주체가 IAM의 서비스 역할과 해당 정책을 볼 수 있도록 허용합니다.
- `ListPoliciesToInspectServiceRolePolicy`— 보안 주체가 IAM의 서비스 역할과 해당 정책을 볼 수 있도록 허용합니다.
- `GetPolicyToInspectServiceRolePolicy`— 주체가 IAM의 서비스 역할과 해당 정책을 볼 수 있도록 허용합니다.
- `ConsoleDisplayTables` - 보안 주체가 콘솔의 기본 AWS Glue 테이블에 대한 데이터를 표시하는 데 필요한 AWS Glue 메타데이터에 대한 읽기 전용 액세스를 허용합니다.
- `ConsolePickOutputBucket`— 보안 주체가 구성된 대상 모델 출력에 대한 Amazon S3 버킷을 선택할 수 있습니다.
- `ConsolePickS3Location`— 보안 주체가 구성된 대상 모델 출력에 대한 위치를 버킷 내에서 선택할 수 있습니다.
- `ConsoleDescribeECRRepositories` - 보안 주체가 Amazon ECR 리포지토리 및 이미지를 설명할 수 있도록 허용합니다.

정책 세부 정보의 JSON 목록은 AWS 관리형 정책 참조 안내서의 AWSCleanRoomsMLFullAccess를 참조하세요.

AWS Clean RoomsAWS 관리형 정책에 대한 업데이트

이 서비스가 이러한 변경 사항을 추적하기 시작한 AWS Clean Rooms 이후부터의 AWS 관리형 정책 업데이트에 대한 세부 정보를 봅니다. 이 페이지의 변경 사항에 대한 자동 알림을 받으려면 AWS Clean Rooms 문서 기록 페이지에서 RSS 피드를 구독하세요.

변경 사항	설명	날짜
AWSCleanRoomsMLReadOnlyAccess - 기존 정책에 대한 업데이트 AWSCleanRoomsMLFullAccess - 기존 정책에 대한 업데이트	AWSCleanRoomsMLReadOnlyAccess에 PassCleanRoomsResources이(가) 추가되었습니다. AWSCleanRoomsMLFullAccess에 PassCleanRoomsResources 및 ConsoleDescribeECRRepositories 추가.	2025년 1월 10일
AWSCleanRoomsFullAccessNoQuerying - 기존 정책에 대한 업데이트	CleanRoomsAccess에 cleanrooms:BatchGetSchemaAnalysisRule이(가) 추가되었습니다.	2024년 5월 13일
AWSCleanRoomsFullAccess - 기존 정책에 대한 업데이트	콘솔을 사용하거나 사용하지 않고 쿼리 결과 버킷을 설정하는 데 권한이 필요하므로 권한을 더 잘 나타내기 위해 이 정책에서 AWSCleanRoomsFullAccess의 Statement ID를 ConsolePickQueryResultsBucket에서 SetQueryResultsBucket으로 업데이트했습니다.	2024년 3월 21일
AWSCleanRoomsMLReadOnlyAccess - 새 정책 AWSCleanRoomsMLFullAccess - 새 정책	AWS Clean Rooms ML을 지원하기 AWSCleanRoomsMLFullAccess 위해 AWSCleanRoomsMLReadOnlyAccess 및를 추가했습니다.	2023년 11월 29일
AWSCleanRoomsFullAccessNoQuerying - 기존 정책에 대한 업데이트	새로운 분석 템플릿 기능을 활성화하기 위해 cleanrooms:CreateAnalysisTemplate, cleanrooms:GetAnalysisTemplate, cleanrooms:UpdateAnalysisTemplate, cleanrooms:DeleteAnalysisTemplate, cleanrooms:ListAnalysisTemplates, cleanrooms:GetCollaborationAnalysisTemplate, cleanrooms:BatchGetCollaborationAnalysisTemplate, cleanrooms:ListCollaboratio	2023년 7월 31일

변경 사항	설명	날짜
	nAnalysisTemplates를 CleanRoom sAccess에 추가했습니다.	
AWSCleanRoomsFullAccessNoQuerying - 기존 정책에 대한 업데이트	리소스 태깅을 활성화하기 위해 cleanrooms:ListTagsForResource. cleanrooms:UntagResource 및 cleanrooms:TagResource를 CleanRoomsAccess에 추가했습니다.	2023년 3월 21일
AWS Clean Rooms 에서 변경 내용 추적 시작	AWS Clean Rooms 가 AWS 관리형 정책에 대한 변경 사항 추적을 시작했습니다.	2023년 1월 12일

AWS Clean Rooms 자격 증명 및 액세스 문제 해결

다음 정보를 사용하여 및 IAM으로 작업할 때 발생할 수 있는 일반적인 문제를 진단 AWS Clean Rooms 하고 수정할 수 있습니다.

주제

- 에서 작업을 수행할 권한이 없음 AWS Clean Rooms
- iam:PassRole을 수행하도록 인증되지 않음
- 내 외부의 사람이 내 AWS Clean Rooms 리소스에 액세스 AWS 계정 하도록 허용하고 싶습니다.

에서 작업을 수행할 권한이 없음 AWS Clean Rooms

작업을 수행할 수 있는 권한이 없다는 오류가 수신되면 작업을 수행할 수 있도록 정책을 업데이트해야 합니다.

다음 예제 오류는 `mateojackson` IAM 사용자가 콘솔을 사용하여 가상 *`my-example-widget`* 리소스에 대한 세부 정보를 보려고 하지만 가상 `cleanrooms:`*`GetWidget`* 권한이 없을 때 발생합니다.

```
User: arn:aws:iam::123456789012:user/mateojackson is not authorized to perform:
 cleanrooms:GetWidget on resource: my-example-widget
```

이 경우 Mateo의 정책은 `cleanrooms:GetWidget` 작업을 사용하여 `my-example-widget` 리소스에 액세스하도록 허용하도록 업데이트해야 합니다.

도움이 필요한 경우 AWS 관리자에게 문의하세요. 관리자는 로그인 자격 증명을 제공한 사람입니다.

iam:PassRole을 수행하도록 인증되지 않음

`iam:PassRole` 작업을 수행할 수 있는 권한이 없다는 오류가 수신되면 AWS Clean Rooms에 역할을 전달할 수 있도록 정책을 업데이트해야 합니다.

일부 AWS 서비스 에서는 새 서비스 역할 또는 서비스 연결 역할을 생성하는 대신 기존 역할을 해당 서비스에 전달할 수 있습니다. 이렇게 하려면 사용자가 서비스에 역할을 전달할 수 있는 권한을 가지고 있어야 합니다.

다음 예 오류는 `marymajor`라는 IAM 사용자가 콘솔을 사용하여 AWS Clean Rooms에서 작업을 수행하려고 하는 경우에 발생합니다. 하지만 작업을 수행하려면 서비스 역할이 부여한 권한이 서비스에 있어야 합니다. Mary는 서비스에 역할을 전달할 수 있는 권한을 가지고 있지 않습니다.

```
User: arn:aws:iam::123456789012:user/marymajor is not authorized to perform:
 iam:PassRole
```

이 경우, Mary가 `iam:PassRole` 작업을 수행할 수 있도록 Mary의 정책을 업데이트해야 합니다.

도움이 필요한 경우 AWS 관리자에게 문의하세요. 관리자는 로그인 자격 증명을 제공한 사람입니다.

내 외부의 사람이 내 AWS Clean Rooms 리소스에 액세스 AWS 계정 하도록 허용하고 싶습니다.

다른 계정의 사용자 또는 조직 외부의 사람이 리소스에 액세스할 때 사용할 수 있는 역할을 생성할 수 있습니다. 역할을 수임할 신뢰할 수 있는 사람을 지정할 수 있습니다.

자세히 알아보려면 다음을 참조하세요.

- 에서 이러한 기능을 AWS Clean Rooms 지원하는지 여부를 알아보려면 섹션을 참조하세요AWS Clean Rooms 에서 IAM을 사용하는 방법.
- 소유 AWS 계정 한의 리소스에 대한 액세스 권한을 제공하는 방법을 알아보려면 IAM 사용 설명서의 소유한 다른의 IAM 사용자에게 액세스 권한 제공을 참조 AWS 계정 하세요.
- 타사에 리소스에 대한 액세스 권한을 제공하는 방법을 알아보려면 IAM 사용 설명서의 타사가 AWS 계정 소유한에 대한 액세스 권한 제공을 AWS 계정참조하세요.

- ID 페더레이션을 통해 액세스 권한을 제공하는 방법을 알아보려면 IAM 사용 설명서의 외부에서 인증된 사용자에게 액세스 권한 제공(ID 페더레이션)을 참조하십시오.
- 크로스 계정 액세스를 위한 역할과 리소스 기반 정책 사용의 차이점을 알아보려면 IAM 사용 설명서의 IAM 역할과 리소스 기반 정책의 차이를 참조하세요.

교차 서비스 혼동된 대리자 예방

혼동된 대리자 문제는 작업을 수행할 권한이 없는 엔터티가 권한이 더 많은 엔터티에게 작업을 수행하도록 강요할 수 있는 보안 문제입니다. 에서 AWS교차 서비스 가장은 혼동된 대리자 문제를 초래할 수 있습니다. 교차 서비스 가장은 한 서비스(호출하는 서비스)가 다른 서비스(호출되는 서비스)를 직접적으로 호출할 때 발생할 수 있습니다. 직접적으로 호출하는 서비스는 다른 고객의 리소스에 대해 액세스 권한이 없는 방식으로 작동하게 권한을 사용하도록 조작될 수 있습니다. 이를 방지하기 위해 AWS 에서는 계정의 리소스에 대한 액세스 권한이 부여된 서비스 보안 주체를 사용하여 모든 서비스에 대한 데이터를 보호하는 데 도움이 되는 도구를 제공합니다.

리소스 정책에서 `aws:SourceArn` 글로벌 조건 컨텍스트 키를 사용하여 AWS Clean Rooms이 다른 서비스가 리소스에 부여하는 권한을 제한하는 것이 좋습니다. 하나의 리소스만 교차 서비스 액세스와 연결되도록 허용하려는 경우 `aws:SourceArn`를 사용하세요.

혼동된 대리인 문제로부터 보호하는 가장 효과적인 방법은 리소스의 전체 ARN이 포함된 `aws:SourceArn` 글로벌 조건 컨텍스트 키를 사용하는 것입니다. 에서는 `sts:ExternalId` 조건 키와 AWSClean Rooms도 비교해야 합니다.

`aws:SourceArn`의 값은 수임된 역할 멤버십의 ARN으로 설정해야 합니다.

다음 예는 Clean Rooms에서 `aws:SourceArn` 및 AWS글로벌 조건 컨텍스트 키를 사용하여 혼동된 대리자 문제를 방지하는 방법을 보여줍니다.

> **Note**
>
> 예제 정책은가 구성된 테이블의 데이터 및 메타데이터에 액세스하는 데 AWS Clean Rooms 사용하는 서비스 역할의 신뢰 정책에 적용됩니다.
> *<query-runner-membership-id>*의 값을 쿼리 실행기의 멤버십 ID로 설정해야 합니다.
> 공동 작업의 모든 구성원은 구성된 테이블 matadata를 볼 수 있으므로 각 멤버십 ARN이 멤버십 ARNs 목록에 포함되어야 합니다.

JSON

```
{
    "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:*/<query-
runner-membership-id>*"
                }
            }
        },
        {
            "Sid": "AllowIfSourceArnMatches",
            "Effect": "Allow",
            "Principal": {
                "Service": "cleanrooms.amazonaws.com"
            },
            "Action": "sts:AssumeRole",
            "Condition": {
                "ForAnyValue:ArnEquals": {
                    "aws:SourceArn": [
                        "arn:aws:cleanrooms:<aws-region>:<member-1-account-
id>:membership/<member-1-membership-id>",
                        "arn:aws:cleanrooms:<aws-region>:<member-2-account-
id>:membership/<member-2-membership-id>"
                    ]
                }
            }
        }
    ]
}
```

AWS Clean Rooms ML에 대한 IAM 동작

교차 계정 작업

Clean Rooms ML을 사용하면 한에서 생성한 특정 리소스 AWS 계정 에 다른 계정에서 안전하게 액세스할 수 있습니다 AWS 계정. AWS 계정 A의 클라이언트가 AWS 계정 B가 소유한 `ConfiguredAudienceModel` 리소스`StartAudienceGenerationJob`에서를 호출하면 Clean Rooms ML은 작업에 대해 두 개의 ARNs을 생성합니다. AWS 계정 A의 ARN 하나와 AWS 계정 B의 ARN 하나. ARNs은 ARN을 제외하고 동일합니다 AWS 계정.

Clean Rooms ML은 두 계정의 작업에 자체 IAM 정책을 적용할 수 있도록 작업에 대해 ARN 2개를 생성합니다. 예를 들어 두 계정 모두 태그 기반 액세스 제어를 사용하고 AWS 조직의 정책을 적용할 수 있습니다. 작업은 두 계정의 데이터를 모두 처리하므로 두 계정에서 작업 및 관련 데이터를 삭제할 수 있습니다. 두 계정 모두 다른 계정이 작업을 삭제하지 못하도록 차단할 수는 없습니다.

작업은 한 번만 실행되며 두 계정 모두 `ListAudienceGenerationJobs`를 호출할 때 작업을 볼 수 있습니다. 두 계정 모두 자체 AWS 계정 ID가 있는 ARN을 사용하여 작업에서 `GetDelete`, 및 `Export` APIs를 호출할 수 있습니다.

AWS 계정 는 다른 AWS 계정 ID와 함께 ARN을 사용할 때 작업에 액세스할 수 없습니다.

작업 이름은 AWS 계정내에서 고유해야 합니다. AWS 계정 B의 이름은 *`$accountA-$name`*입니다. AWS 계정 B에서 작업을 볼 때 AWS 계정 A가 선택한 이름 앞에 A가 AWS 계정 붙습니다.

교차 계정이 `StartAudienceGenerationJob` 성공하려면 AWS 계정 다음 예제와 유사한 리소스 정책을 사용하여 AWS 계정 B의 새 작업과 AWS 계정 B의 `ConfiguredAudienceModel`에 대해 해당 작업을 허용해야 합니다.

JSON

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

AWS Clean Rooms ML API를 사용하여가 true로 manageResourcePolicies 설정된 구성된 유사 모델을 생성하는 경우가이 정책을 자동으로 AWS Clean Rooms 생성합니다.

또한 AWS 계정 A에 있는 호출자의 자격 증명 정책에는에 대한 StartAudienceGenerationJob 권한이 필요합니다arn:aws:cleanrooms-ml:us-west-1:AccountA:audience-generation-job/*. 따라서 작업을 위한 세 가지 IAM 리소스StartAudienceGenerationJob는 AWS 계정 A 작업, AWS 계정 B 작업 및 AWS 계정 B입니다ConfiguredAudienceModel.

> **Warning**
>
> 작업을 시작한 AWS 계정 는 작업에 대한 AWS CloudTrail 감사 로그 이벤트를 수신합니다. ConfiguredAudienceModel을 소유한 AWS 계정 는 AWS CloudTrail 감사 로그 이벤트를 수신할 수 없습니다.

작업에 태그 지정

CreateConfiguredAudienceModel의 childResourceTagOnCreatePolicy=FROM_PARENT_RESOURCE 파라미터를 설정하면 구성된 유사 모델에서 생성된 계정 내 모든 유사 세그먼트 생성 작업은 구성된 유사 모델과 동일한 태그를 기본으로 사용합니다. 구성된 유사 모델은 상위 모델이고 유사 세그먼트 생성 작업은 하위 모델입니다.

자신의 계정 내에서 작업을 생성하는 경우 작업의 요청 태그가 상위 태그를 재정의합니다. 다른 계정에서 생성한 작업은 계정에 태그를 생성할 수 없습니다.

childResourceTagOnCreatePolicy=FROM_PARENT_RESOURCE를 설정하고 다른 계정에서 작업을 생성하면 해당 작업의 복사본이 2개가 됩니다. 계정의 사본에는 상위 리소스 태그가 있고 작업 제출자 계정의 사본에는 요청에서 얻은 태그가 있습니다.

공동 작업자 검증

AWS Clean Rooms 공동 작업의 다른 구성원에게 권한을 부여할 때는 리소스 정책에 조건 키가 포함되어야 합니다cleanrooms-ml:CollaborationId. 이렇게 하면 해당 collaborationId 파라미터가 StartAudienceGenerationJob 요청에 포함됩니다. collaborationId 파라미터가 요청에 포함되면 Clean Rooms ML은 공동 작업이 있는지, 작업 제출자가 공동 작업의 활성 구성원인지, 구성된 유사 모델 소유자가 공동 작업의 활성 구성원인지 검증합니다.

가 구성된 유사 모델 리소스 정책을 AWS Clean Rooms 관리할 때(manageResourcePolicies파라미터가 TRUE CreateConfiguredAudienceModelAssociation 요청에 있음)이 조건 키는 리소스 정책에 설정됩니다. 따라서 StartAudienceGenerationJob에서 collaborationId를 지정해야 합니다.

크로스 계정 액세스

계정 전반에서 StartAudienceGenerationJob만 호출할 수 있습니다. 다른 모든 Clean Rooms ML API는 자체 계정의 리소스에서만 사용할 수 있습니다. 이렇게 하면 훈련 데이터, 유사 모델 구성 및 기타 정보를 비공개로 유지할 수 있습니다.

Clean Rooms ML은 계정 전체에서 Amazon S3 또는 AWS Glue 위치를 공개하지 않습니다. 훈련 데이터 위치, 구성된 유사 모델 출력 위치, 유사 세그먼트 생성 작업 시드 위치는 계정 전반에서 절대 볼 수 없습니다. 공동 작업에서 쿼리 로깅을 활성화하지 않는 한 시드 데이터가 SQL 쿼리에서 왔는지 여부와 쿼리 자체는 계정 전체에 표시되지 않습니다. 다른 계정에서 제출한 대상 생성 작업을 Get으로 처리한 경우 서비스는 시드 위치를 표시하지 않습니다.

Clean Rooms ML Custom 모델의 IAM 동작

교차 계정 작업

Clean Rooms ML을 사용하면 한에서 생성한 공동 작업과 연결된 특정 리소스 AWS 계정 에 다른 계정에서 안전하게 액세스할 수 있습니다 AWS 계정. 멤버가 쿼리를 실행할 수 있는 AWS 계정 A의 클라이언트는 공동 작업에서 다른 멤버가 소유한 ConfiguredModelAlgorithmAssociation 리소스StartTrainedModelInferenceJob에 대해 CreateMLInputChannel, 또는 CreateTrainedModel를 호출할 수 있습니다. 단, 로 생성된 사용자 지정 분석 규칙에 의해 ConfiguredModelAlgorithmAssociation가 허용됩니다CreateConfiguredTableAnalysisRule.

또한 공동 작업의 활성 구성원은 DeleteTrainedModelOutput 및 DeleteMLInputChannelData APIs.

크로스 계정 액세스

Clean Rooms ML을 사용하면 사용자가 GetCollaboration 및 ListCollaboration APIs. Clean Rooms ML은 KMS 키 ARNs, 태그, 환경 변수 또는 하이퍼파라미터(TrainedModel작업용)를 다른 계정에 공개하지 않습니다.

멤버십 및 공동 작업 액세스

Clean Rooms ML 사용자 지정 모델의 맥락에서 멤버십 및 공동 작업 리소스에 액세스할 때 사용자의 자격 증명 정책에는 작업 cleanrooms:PassMembership, cleanrooms:PassCollaboration또는 둘 다에 대한 권한이 필요합니다. 를 수락하는 모든 APIs에는 cleanrooms:PassMembership 권한이 membershipId 필요하고를 수락하는 모든 APIs에는 cleanrooms:PassCollaboration 권한이 collaborationId 필요합니다. 공동 작업 ID의 컨텍스트createTrainedModel에서 호출할 수 있는 멤버십 ID의 컨텍스트GetCollaborationTrainedModel에서 호출할 수 있는 역할에 대한 샘플 자격 증명 정책이 제공됩니다.

JSON

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

에 대한 규정 준수 검증 AWS Clean Rooms

AWS 서비스 가 특정 규정 준수 프로그램의 범위 내에 있는지 알아보려면 AWS 서비스 규정 준수 프로그램 제공 범위 섹션을 참조하고 관심 있는 규정 준수 프로그램을 선택합니다. 일반 정보는 AWS 규정 준수 프로그램.

를 사용하여 타사 감사 보고서를 다운로드할 수 있습니다 AWS Artifact. 자세한 내용은 에서 보고서 다운로드 AWS Artifact에서 .

사용 시 규정 준수 책임은 데이터의 민감도, 회사의 규정 준수 목표 및 관련 법률과 규정에 따라 AWS 서비스 결정됩니다.는 규정 준수를 지원하기 위해 다음 리소스를 AWS 제공합니다.

- 보안 규정 준수 및 거버넌스 - 이러한 솔루션 구현 가이드에서는 아키텍처 고려 사항을 설명하고 보안 및 규정 준수 기능을 배포하는 단계를 제공합니다.
- HIPAA 적격 서비스 참조 - HIPAA 적격 서비스가 나열되어 있습니다. 모두 HIPAA 자격이 AWS 서비스 있는 것은 아닙니다.
- AWS 규정 준수 리소스 -이 워크북 및 가이드 모음은 산업 및 위치에 적용될 수 있습니다.
- AWS 고객 규정 준수 가이드 - 규정 준수의 관점에서 공동 책임 모델을 이해합니다. 이 가이드에는 여러 프레임워크(미국 국립표준기술연구소(NIST), 결제 카드 산업 보안 표준 위원회(PCI), 국제표준화기구(ISO) 포함)의 보안 제어에 대한 지침을 보호하고 AWS 서비스 매핑하는 모범 사례가 요약되어 있습니다.
- AWS Config 개발자 안내서의 규칙을 사용하여 리소스 평가 -이 AWS Config 서비스는 리소스 구성이 내부 관행, 업계 지침 및 규정을 얼마나 잘 준수하는지 평가합니다.
- AWS Security Hub - 내 보안 상태에 대한 포괄적인 보기를 AWS 서비스 제공합니다 AWS. Security Hub는 보안 컨트롤을 사용하여 AWS 리소스를 평가하고 보안 업계 표준 및 모범 사례에 대한 규정 준수를 확인합니다. 지원되는 서비스 및 제어 목록은 Security Hub 제어 참조를 참조하세요.

- Amazon GuardDuty - 의심스러운 활동과 악의적인 활동이 있는지 환경을 모니터링하여 사용자, AWS 계정워크로드, 컨테이너 및 데이터에 대한 잠재적 위협을 AWS 서비스 탐지합니다. GuardDuty는 특정 규정 준수 프레임워크에서 요구하는 침입 탐지 요구 사항을 충족하여 PCI DSS와 같은 다양한 규정 준수 요구 사항을 따르는 데 도움을 줄 수 있습니다.
- AWS Audit Manager - 이를 AWS 서비스 통해 AWS 사용량을 지속적으로 감사하여 위험과 규정 및 업계 표준 준수를 관리하는 방법을 간소화할 수 있습니다.

의 복원력 AWS Clean Rooms

AWS 글로벌 인프라는 AWS 리전 및 가용 영역을 중심으로 구축됩니다. 리전은 물리적으로 분리되고 격리된 다수의 가용 영역을 제공하며, 이러한 영역은 짧은 지연 시간, 높은 처리량 및 높은 중복성을 갖춘 네트워크를 통해 연결되어 있습니다. 가용 영역을 사용하면 중단 없이 영역 간에 자동으로 장애 극복 조치가 이루어지는 애플리케이션 및 데이터베이스를 설계하고 운영할 수 있습니다. 가용 영역은 기존의 단일 또는 다중 데이터 센터 인프라보다 가용성, 내결함성, 확장성이 뛰어납니다.

AWS 리전 및 가용 영역에 대한 자세한 내용은 AWS 글로벌 인프라를 참조하세요.

의 인프라 보안 AWS Clean Rooms

관리형 서비스인는 AWS 글로벌 네트워크 보안으로 보호 AWS Clean Rooms 됩니다. AWS 보안 서비스 및가 인프라를 AWS 보호하는 방법에 대한 자세한 내용은 AWS 클라우드 보안을 참조하세요. 인프라 보안 모범 사례를 사용하여 AWS 환경을 설계하려면 보안 원칙 AWS Well-Architected Framework의 인프라 보호를 참조하세요.

AWS 에서 게시한 API 호출을 사용하여 네트워크를 AWS Clean Rooms 통해에 액세스합니다. 고객은 다음을 지원해야 합니다.

- Transport Layer Security(TLS) TLS 1.2는 필수이며 TLS 1.3을 권장합니다.
- DHE(Ephemeral Diffie-Hellman) 또는 ECDHE(Elliptic Curve Ephemeral Diffie-Hellman)와 같은 완전 전송 보안(PFS)이 포함된 암호 제품군 Java 7 이상의 최신 시스템은 대부분 이러한 모드를 지원합니다.

또한 요청은 액세스 키 ID 및 IAM 위탁자와 관련된 보안 암호 액세스 키를 사용하여 서명해야 합니다. 또는 AWS Security Token Service(AWS STS)를 사용하여 임시 보안 인증 정보를 생성하여 요청에 서명할 수 있습니다.

네트워크 보안

쿼리 실행 중에가 S3 버킷에서 AWS Clean Rooms 읽으면와 Amazon S3 간의 AWS Clean Rooms 트래픽이 AWS 프라이빗 네트워크를 통해 안전하게 라우팅됩니다. 이동 중인 트래픽은 Amazon Signature Version 4 프로토콜(SIGv4)로 서명되고 HTTPS를 사용하여 암호화됩니다. 이 트래픽에는 구성된 테이블에 대해 설정한 IAM 서비스 역할을 기반으로 권한이 부여됩니다.

엔드포인트를 AWS Clean Rooms 통해 프로그래밍 방식으로에 연결할 수 있습니다. 서비스 엔드포인트 목록은 AWS 일반 참조에서 AWS Clean Rooms 엔드포인트 및 할당량을 참조하세요.

모든 서비스 엔드포인트는 HTTPS 전용입니다. VPC AWS Clean Rooms 에서에 연결하고 인터넷에 연결하지 않으려는 경우 Amazon Virtual Private Cloud(VPC) 엔드포인트를 사용할 수 있습니다. 자세한 내용은 AWS PrivateLink 가이드의를 통한 AWS 서비스 액세스를 AWS PrivateLink 참조하세요.

aws:SourceVpce 컨텍스트 키를 사용하여 IAM 보안 주체가 인터넷이 아닌 VPC 엔드포인트를 AWS Clean Rooms 통해서만를 호출할 수 있도록 제한하는 IAM 정책을 IAM 보안 주체에 할당할 수 있습니다.

인터페이스 엔드포인트를 사용하여 AWS Clean Rooms 또는 AWS Clean Rooms ML에 액세스(AWS PrivateLink)

AWS PrivateLink 를 사용하여 Virtual Private Cloud(VPC)와 AWS Clean Rooms 또는 AWS Clean Rooms ML 간에 프라이빗 연결을 생성할 수 있습니다. 인터넷 게이트웨이, NAT 디바이스, VPN 연결 또는 AWS Direct Connect 연결을 사용하지 않고 VPC에 있는 것처럼 AWS Clean Rooms 또는 AWS Clean Rooms ML에 액세스할 수 있습니다. VPC의 인스턴스에서 AWS Clean Rooms API에 액세스하는 데는 퍼블릭 IP 주소가 필요하지 않습니다.

AWS PrivateLink에서 제공되는 인터페이스 엔드포인트를 생성하여 이 프라이빗 연결을 설정합니다. 인터페이스 엔드포인트에 대해 사용 설정하는 각 서브넷에서 엔드포인트 네트워크 인터페이스를 생성합니다. 이는 AWS Clean Rooms로 향하는 트래픽의 진입점 역할을 하는 요청자 관리형 네트워크 인터페이스입니다.

자세한 내용은AWS PrivateLink 가이드의 AWS PrivateLink를 통해 AWS 서비스 에 액세스를 참조하세요.

에 대한 고려 사항 AWS Clean Rooms

에 대한 인터페이스 엔드포인트를 설정하기 전에 AWS PrivateLink 가이드의 고려 사항을 AWS Clean Rooms검토하세요.

AWS Clean Rooms 및 AWS Clean Rooms ML은 인터페이스 엔드포인트를 통해 모든 API 작업을 호출할 수 있도록 지원합니다.

AWS Clean Rooms 또는 AWS Clean Rooms ML에는 VPC 엔드포인트 정책이 지원되지 않습니다. 기본적으로 인터페이스 엔드포인트를 통해 AWS Clean Rooms 및 AWS Clean Rooms ML에 대한 전체 액세스가 허용됩니다. 또는 보안 그룹을 엔드포인트 네트워크 인터페이스와 연결하여 인터페이스 엔드포인트를 통해 AWS Clean Rooms 또는 AWS Clean Rooms ML에 대한 트래픽을 제어할 수 있습니다.

에 대한 인터페이스 엔드포인트 생성 AWS Clean Rooms

Amazon VPC 콘솔 또는 AWS Command Line Interface ()를 사용하여 AWS Clean Rooms 또는 AWS Clean Rooms ML에 대한 인터페이스 엔드포인트를 생성할 수 있습니다AWS CLI. 자세한 내용은 AWS PrivateLink 설명서의 인터페이스 엔드포인트 생성을 참조하세요.

다음 서비스 이름을 AWS Clean Rooms 사용하여 용 인터페이스 엔드포인트를 생성합니다.

```
com.amazonaws.region.cleanrooms
```

다음 서비스 이름을 사용하여 AWS Clean Rooms ML용 인터페이스 엔드포인트를 생성합니다.

```
com.amazonaws.region.cleanrooms-ml
```

인터페이스 엔드포인트에 프라이빗 DNS를 사용하도록 설정하는 경우, 리전에 대한 기본 DNS 이름(예: AWS Clean Rooms)을 사용하여 에 API 요청을 할 수 있습니다. 예를 들어 `cleanrooms-ml.us-east-1.amazonaws.com`입니다.

모니터링 AWS Clean Rooms

모니터링은 AWS Clean Rooms 및 다른 AWS 솔루션의 안정성, 가용성 및 성능을 유지하는 데 중요한 부분입니다.는 다음과 같은 모니터링 도구를 AWS 제공하여 문제 발생 시 모니터링하고 AWS Clean Rooms보고하며 적절한 경우 자동 조치를 취합니다.

- Amazon CloudWatch Logs를 사용하면 Amazon EC2 인스턴스 및 기타 소스에서 로그 파일을 모니터링 AWS CloudTrail, 저장 및 액세스할 수 있습니다. Amazon CloudWatch Logs는 로그 파일의 정보를 모니터링하고 특정 임계값에 도달하면 사용자에게 알릴 수 있습니다. 또한 매우 내구력 있는 스토리지에 로그 데이터를 저장할 수 있습니다. 자세한 내용은 Amazon CloudWatch Logs 사용 설명서를 참조하세요.

  Clean Rooms ML은 특정 API 작업에 대해 교차 계정 작업을 허용합니다. 작업을 시작한 AWS 계정는 작업에 대한 AWS CloudTrail 감사 로그 이벤트를 수신합니다. 자세한 내용은 AWS Clean Rooms ML에 대한 IAM 동작 섹션을 참조하세요.
- AWS CloudTrail는에 의해 또는를 대신하여 수행된 API 호출 및 관련 이벤트를 캡처 AWS 계정 하고 사용자가 지정한 Amazon S3 버킷에 로그 파일을 전송합니다. 호출한 사용자 및 계정 AWS, 호출이 수행된 소스 IP 주소, 호출이 발생한 시기를 식별할 수 있습니다. 자세한 내용은 AWS CloudTrail 사용 설명서를 참조하십시오.

의 분석 로깅 AWS Clean Rooms

분석 로깅은의 기능입니다 AWS Clean Rooms. 공동 작업을 생성하고 분석 로깅을 켜면 멤버는 Amazon CloudWatch Logs의 작업에서 쿼리 또는 로그의 관련 로그를 저장할 수 있습니다.

쿼리 로그 및 작업 로그를 사용하여 구성원은 쿼리가 분석 규칙을 준수하고 협업 계약에 부합하는지 확인할 수 있습니다. 또한 쿼리 로그는 감사를 지원하는 데 도움이 됩니다.

AWS Clean Rooms 콘솔에서 분석 로깅 옵션이 켜져 있는 경우 쿼리 로그에는 다음이 포함됩니다.

- `analysisRule`— 구성된 테이블의 분석 규칙.
- `analysisTemplateArn`— 실행된 분석 템플릿(분석 규칙에 따라 표시됨).
- `collaborationId`— 쿼리가 실행된 공동 작업의 고유 식별자입니다.
- `configuredTableID`— 쿼리에서 참조되는 구성 테이블의 고유 식별자입니다.
- `directQueryAnalysisRulePolicy.custom.allowedAnalysis`— 구성된 테이블에서 실행할 수 있는 분석 템플릿(분석 규칙에 따라 표시됨).

- `directQueryAnalysisRulePolicy.v1.custom.allowedAnalysisProviders`— 쿼리 생성을 허용한 쿼리 제공자(분석 규칙에 따라 표시됨).
- `errorCode` — 쿼리가 제대로 실행되지 않은 경우의 오류 코드입니다.
- `errorMessage` — 쿼리가 제대로 실행되지 않은 경우의 오류 메시지입니다.
- `eventID` – 쿼리 실행의 고유 식별자입니다. 2023년 8월 31일 이후에는 고유 식별자가 `protectedQueryID`와 동일합니다.
- `eventTimestamp`— 쿼리 실행 시간.
- `parameters.parametervalue`— 매개 변수 값(쿼리 텍스트에 따라 표시됨).
- `queryText`— 쿼리 실행의 SQL 정의. 매개변수가 있는 경우 `:parametervalue`와 같이 레이블이 지정됩니다.
- `queryValidationErrors`— 쿼리 검증 시 쿼리 오류가 발생했습니다.
- `schemaName`— 쿼리에서 참조되는 구성된 테이블 연결의 이름입니다.
- `status` — 쿼리의 실행 상태입니다.

쿼리 및 작업 로그 수신

쿼리 로그 및 작업 로그를 설정 AWS Clean Rooms 하기 위해 외부에서 작업을 수행할 필요가 없습니다.는 각 공동 작업 구성원이 멤버십을 AWS Clean Rooms 생성한 후 공동 작업을 위한 로그 그룹을 생성합니다. ???

쿼리할 수 있는 멤버, 쿼리 및 작업을 실행할 수 있는 멤버, 결과를 받을 수 있는 멤버, 쿼리에서 구성 테이블이 참조되는 멤버는 쿼리 로그 또는 작업 로그를 받게 됩니다.

쿼리할 수 있는 구성원과 결과를 받을 수 있는 구성원은 쿼리에서 참조되는 각 구성 테이블에 대한 쿼리 로그를 받게 됩니다. 구성된 테이블을 소유하지 않으면 구성된 테이블 ID(`configuredTableID`)를 볼 수 없습니다.

쿼리 및 작업을 실행할 수 있는 구성원과 결과를 받을 수 있는 구성원은 작업에서 참조되는 구성된 각 테이블에 대한 작업 로그를 수신합니다. 구성된 테이블을 소유하지 않으면 구성된 테이블 ID(`configuredTableID`)를 볼 수 없습니다.

쿼리에서 참조되는 구성된 테이블 연결이 여러 개 있는 구성원이 있는 경우 구성원은 구성된 각 테이블에 대한 쿼리 로그를 받게 됩니다.

멤버가 작업에서 참조되는 여러 개의 구성된 테이블 연결을 가지고 있는 경우 구성된 각 테이블에 대한 작업 로그를 받게 됩니다.

AWS Clean Rooms에서 지원되지 않는 SQL과 지원되는 SQL이 포함된 쿼리에 대한 로그가 생성됩니다. 자세한 내용은 AWS Clean Rooms SQL 참조를 참조하세요.

쿼리 또는 작업이 공동 작업과 연결되지 않은 구성된 테이블을 참조할 때도 로그가 생성됩니다.

에서 잘못된 SQL에 대한 로그가 생성되지 않습니다 AWS Clean Rooms.

쿼리 및 작업 로그는 쿼리의 상태를 나타내지만 쿼리 출력이 전달되었는지 여부는 보고하지 않습니다. 쿼리할 수 있는 구성원이 쿼리 또는 작업을 제출했는지 확인합니다. 또한 쿼리 로그는 쿼리에에서 지원되는 SQL이 포함되어 있고 공동 작업 AWS Clean Rooms 과 연결된 구성된 테이블을 참조하는지 확인합니다.

Example

예를 들어, 분석 규칙 준수 여부를 AWS Clean Rooms 확인한 후 쿼리 처리 중에 쿼리가 취소된 경우 로그가 생성되지 않습니다.

로그 그룹을 삭제하는 경우 동일한 로그 그룹 이름(공동 작업의 공동 작업 ID)을 사용하여 로그 그룹을 수동으로 다시 만들어야 합니다. 또는 멤버십에서 로그오프 및 로그온을 해제할 수 있습니다.

분석 로깅을 켜는 방법에 대한 자세한 내용은 섹션을 참조하세요공동 작업 생성.

Amazon CloudWatch Logs에 대한 자세한 내용은 Amazon CloudWatch Logs 사용자 설명서를 참조하세요.

쿼리 및 작업 로그에 대한 권장 작업

회원은 정기적으로 다음 조치를 취하는 것이 좋습니다.

- 쿼리와 작업이 공동 작업에 대해 합의된 사용 사례 또는 쿼리와 일치하는지 확인하려면 공동 작업에서 실행되는 쿼리와 작업을 검토합니다.

  최근 쿼리를 보는 방법에 대한 자세한 내용은 최근 쿼리 보기를 참조하세요.

  최근 작업을 보는 방법에 대한 자세한 내용은 섹션을 참조하세요최근 작업 보기.
- 구성된 테이블 열이 공동 작업에 대해 합의된 내용과 일치하는지 확인하려면 공동 작업 구성원의 분석 규칙 및 쿼리에 사용되는 구성된 테이블 열을 검토하세요.

  구성된 열을 보는 방법에 대한 자세한 내용은 테이블 및 분석 규칙 보기를 참조하세요.

를 사용하여 AWS Clean Rooms API 호출 로깅 AWS CloudTrail

AWS Clean Rooms 는 사용자 AWS CloudTrail, 역할 또는 AWS 서비스 의가 수행한 작업에 대한 레코드를 제공하는 서비스와 통합됩니다 AWS Clean Rooms. CloudTrail은에 대한 모든 API 호출을 이벤트 AWS Clean Rooms 로 캡처합니다. 캡처된 호출에는 AWS Clean Rooms 콘솔의 호출과 AWS Clean Rooms API 작업에 대한 코드 호출이 포함됩니다. 추적을 생성하면 이벤트를 포함하여 CloudTrail 이벤트를 Amazon S3 버킷으로 지속적으로 전송할 수 있습니다 AWS Clean Rooms. 추적을 구성하지 않은 경우에도 CloudTrail 콘솔의 이벤트 기록에서 최신 이벤트를 볼 수 있습니다. CloudTrail에서 수집한 정보를 사용하여 AWS Clean Rooms에 수행된 요청, 요청이 수행된 IP 주소, 요청을 수행한 사람, 요청이 수행된 시간 및 추가 세부 정보를 확인할 수 있습니다.

CloudTrail에 대한 자세한 설명은 AWS CloudTrail 사용자 가이드를 참조하십시오.

AWS Clean Rooms CloudTrail의 정보

CloudTrail은 계정을 생성할 AWS 계정 때에서 활성화됩니다. 에서 활동이 발생하면 AWS Clean Rooms해당 활동이 이벤트 기록의 다른 이벤트와 함께 CloudTrail AWS 서비스 이벤트에 기록됩니다. AWS 계정에서 최신 이벤트를 확인, 검색 및 다운로드할 수 있습니다. 자세한 내용은 CloudTrail 이벤트 기록을 사용하여 이벤트 보기를 참조하세요.

에 대한 이벤트를 AWS 계정포함하여에 이벤트를 지속적으로 기록하려면 추적을 AWS Clean Rooms 생성합니다. CloudTrail은 추적을 사용하여 Amazon S3 버킷으로 로그 파일을 전송할 수 있습니다. 콘솔에서 추적을 생성하면 기본적으로 모든 AWS 리전에 추적이 적용됩니다. 추적은 AWS 파티션의 모든 리전에서 이벤트를 로깅하고 지정한 Amazon S3 버킷으로 로그 파일을 전송합니다. 또는 CloudTrail 로그에서 수집된 이벤트 데이터를 추가 분석 및 처리하도록 다른 AWS 서비스 (을)를 구성할 수 있습니다. 자세한 설명은 다음을 참조하세요.

- 추적 생성 개요
- CloudTrail 지원 서비스 및 통합
- CloudTrail에 대한 Amazon SNS 알림 구성
- 여러 리전에서 CloudTrail 로그 파일 수신
- 여러 계정에서 CloudTrail 로그 파일 수신

모든 AWS Clean Rooms 작업은 CloudTrail에서 로깅되며 AWS Clean Rooms API 참조에 문서화됩니다.

모든 이벤트 또는 로그 항목에는 요청을 생성했던 사용자에 대한 정보가 포함됩니다. 자격 증명을 이용하면 다음을 쉽게 판단할 수 있습니다.

- 요청을 루트 사용자로 했는지 아니면 IAM 사용자 자격 증명으로 했는지 여부.
- 역할 또는 페더레이션 사용자에 대한 임시 보안 인증을 사용하여 요청이 생성되었는지 여부.
- 다른 AWS 서비스에서 요청했는지 여부.

자세한 내용은 CloudTrail userIdentity 요소를 참조하세요.

AWS Clean Rooms 로그 파일 항목 이해

추적이란 지정한 Amazon S3 버킷에 이벤트를 로그 파일로 입력할 수 있게 하는 구성입니다. CloudTrail 로그 파일에는 하나 이상의 로그 항목이 포함될 수 있습니다. 이벤트는 모든 소스로부터의 단일 요청을 나타내며 요청 작업, 작업 날짜와 시간, 요청 파라미터 등에 대한 정보가 들어 있습니다. CloudTrail 로그 파일은 퍼블릭 API 직접 호출의 주문 스택 트레이스가 아니므로 특정 순서로 표시되지 않습니다.

Example AWS Clean Rooms CloudTrail 이벤트

다음 예는 다음에 대한 CloudTrail 이벤트를 보여줍니다.

주제

- StartProtectedQuery(성공)
- StartProtectedQuery (failed)

StartProtectedQuery(성공)


```
{
    "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 (failed)

```
{
    "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"
}
```

를 사용하여 AWS Clean Rooms 리소스 생성 AWS CloudFormation

AWS Clean Rooms 는 AWS 리소스를 모델링하고 설정하는 데 도움이 되는 AWS CloudFormation서비스와 통합됩니다. 이러한 통합을 통해 리소스 및 인프라를 생성하고 관리하는 데 소요되는 시간을 줄일 수 있습니다. 원하는 모든 AWS 리소스를 설명하는 템플릿을 생성하면가 해당 리소스를 AWS CloudFormation 프로비저닝하고 구성합니다. 리소스의 예로는 공동 작업, 구성된 테이블, 구성된 테이블 연결, 멤버십 등이 있습니다.

를 사용하면 템플릿을 재사용하여 AWS Clean Rooms 리소스를 일관되고 반복적으로 설정할 AWS CloudFormation수 있습니다. 리소스를 한 번 설명한 다음 여러 AWS 계정 및에서 동일한 리소스를 반복적으로 프로비저닝합니다 AWS 리전.

AWS Clean Rooms 및 AWS CloudFormation 템플릿

AWS Clean Rooms 및 관련 서비스에 대한 리소스를 프로비저닝하고 구성하려면 AWS CloudFormation 템플릿을 이해해야 합니다. 템플릿은 JSON 또는 YAML로 서식 지정된 텍스트 파일입니다. 이러한 템플릿은 AWS CloudFormation 스택에서 프로비저닝하려는 리소스를 설명합니다. JSON 또는 YAML에 익숙하지 않은 경우 AWS CloudFormation Designer를 사용하여 AWS CloudFormation 템플릿을 시작할 수 있습니다. 자세한 내용은 AWS CloudFormation 사용 설명서에서 AWS CloudFormation Designer이란 무엇입니까?를 참조하세요.

AWS Clean Rooms 는 공동 작업, 구성된 테이블, 구성된 테이블 연결 및 멤버십 생성을 지원합니다 AWS CloudFormation. 공동 작업, 구성된 테이블, 구성된 테이블 연결 및 멤버십을 위한 JSON 및 YAML 템플릿의 예를 비롯한 자세한 내용은 AWS CloudFormation 사용 설명서의 AWS Clean Rooms 및 AWS Clean Rooms ML 리소스 유형 참조를 참조하세요.

에 대해 자세히 알아보기 AWS CloudFormation

에 대해 자세히 알아보려면 다음 리소스를 AWS CloudFormation참조하세요.

- AWS CloudFormation
- AWS CloudFormation 사용 설명서
- AWS CloudFormation API Reference
- AWS CloudFormation 명령줄 인터페이스 사용 설명서

에 대한 할당량 AWS Clean Rooms

AWS 계정 에는 각에 대해 이전에 제한이라고 하는 기본 할당량이 있습니다 AWS 서비스. 달리 명시되지 않는 한 각 할당량은에 고유합니다 AWS 리전. 일부 할당량에 대한 증가를 요청할 수 있으며 다른 할당량은 늘릴 수 없습니다.

할당량을 보려면 Service Quotas 콘솔을 AWS Clean Rooms엽니다. 탐색 창에서 AWS 서비스(AWS services)를 선택하고 AWS Clean Rooms을 선택합니다.

할당량 증가를 요청하려면 Service Quotas 사용 설명서의 할당량 증가 요청을 참조하세요. Service Quotas에서 아직 할당량을 사용할 수 없는 경우 서비스 한도 증가 양식을 사용합니다.

주제

- AWS Clean Rooms 할당량
- AWS Clean Rooms ML 할당량

AWS Clean Rooms 할당량

AWS 계정 에는 다음과 관련된 할당량이 있습니다 AWS Clean Rooms.

명칭	기본값	조정 가능	설명
분석 규칙 크기	지원되는 각 리전: 100KB	아니요	분석 규칙을 위한 최대 JSON 크기
멤버십당 분석 템플릿 수	지원되는 각 지역: 25	예	멤버십당 최대 분석 템플릿 수
계정당 생성된 협업 수	지원되는 각 리전: 10	예	계정당 생성되는 최대 협업 수

명칭	기본값	조정 가능	설명
구성된 테이블 허용목록당 열 수	지원되는 각 리전: 100	예	구성된 테이블당 허용 목록에 추가할 수 있는 최대 열 수
계정당 동시 PySpark 작업 vCPU 사용량	us-east-1: 1,024 지원되는 다른 각 리전: 512	예	계정당 동시에 실행되는 모든 PySpark 작업의 최대 총 vCPU 사용량
계정당 동시 PySpark 작업	us-east-1: 5 각각의 지원되는 다른 리전: 2	예	계정당 동시에 실행되는 PySpark 작업의 최대 수
계정당 동시 SQL 쿼리	us-east-1: 5 각각의 지원되는 다른 리전: 2	예	계정당 동시에 실행되는 SQL 쿼리의 최대 수
계정당 동시 SQL 쿼리 vCPU 사용량	지원되는 각 리전: 512	예	계정당 동시에 실행되는 모든 SQL 쿼리의 최대 총 vCPU 사용량
멤버십당 동시 진행 중인 작업	지원되는 각 리전: 1	아니요	멤버십당 진행 중인 동시 작업의 최대 수
멤버십당 진행 중인 동시 쿼리 수	지원되는 각 지역: 5개	아니요	멤버십당 최대 동시 진행 중인 쿼리 수
멤버십당 구성된 대상 모델 연결	지원되는 각 리전: 5	예	멤버십당 구성된 대상 모델 연결의 최대 수

명칭	기본값	조정 가능	설명
계정당 구성된 테이블	지원되는 각 리전: 250개	예	계정당 생성된 최대 구성된 테이블 수
보호된 쿼리당 구성된 테이블 수	지원되는 각 리전: 15개	예	보호된 쿼리의 최대 구성 테이블 수
멤버십당 ID 매핑 테이블	지원되는 각 리전: 5	예	멤버십당 최대 ID 매핑 테이블 수
멤버십당 ID 네임스페이스 연결	지원되는 각 리전: 10	예	멤버십당 최대 ID 네임스페이스 연결 수
협업당 초대된 구성원 수	지원되는 각 리전: 5	예	협업당 초대되는 최대 구성원 수
계정당 멤버십	지원되는 각 리전: 100	예	계정당 최대 멤버십 수
쿼리 텍스트 길이(차등 프라이버시 사용)	지원되는 각 리전: 8KB	아니요	차등 프라이버시를 사용하는 SQL 쿼리 문의 최대 텍스트 길이
Clean Rooms SQL 분석 엔진의 쿼리 텍스트 길이	지원되는 각 리전: 90KB	아니요	Clean Rooms SQL 분석 엔진에서 SQL 쿼리 문의 최대 텍스트 길이
Spark 분석 엔진의 쿼리 텍스트 길이	지원되는 각 리전: 500킬로바이트	아니요	Spark 분석 엔진의 SQL 쿼리 문의 최대 텍스트 길이
BatchGetSchema 요청률	지원되는 각 리전: 5	예	초당 최대 BatchGetSchema API 직접 호출 수

명칭	기본값	조정 가능	설명
공동 작업 요청 생성률	지원되는 각 리전: 5	예	초당 최대 CreateCollaboration API 직접 호출 수
CreateConfiguredTable 요청률	지원되는 각 리전: 5	예	초당 최대 CreateConfiguredTable API 직접 호출 수
CreateConfiguredTableAnalysisRule 요청률	지원되는 각 리전: 5	예	초당 최대 CreateConfiguredTableAnalysisRule API 직접 호출 수
CreateConfiguredTableAssociation 요청률	지원되는 각 리전: 5	예	초당 최대 CreateConfiguredTableAssociation API 직접 호출 수
CreateMembership 요청률	지원되는 각 리전: 5	예	초당 최대 CreateMembership API 직접 호출 수
DeleteCollaboration 요청률	지원되는 각 리전: 5	예	초당 최대 DeleteCollaboration API 직접 호출 수
DeleteConfiguredTable 요청률	지원되는 각 리전: 5	예	초당 최대 DeleteConfiguredTable API 직접 호출 수
DeleteConfiguredTableAnalysisRule 요청률	지원되는 각 리전: 5	예	초당 최대 DeleteConfiguredTableAnalysisRule API 직접 호출 수
DeleteConfiguredTableAssociation 요청률	지원되는 각 리전: 5	예	초당 최대 DeleteConfiguredTableAssociation API 직접 호출 수

명칭	기본값	조정 가능	설명
DeleteMember 요청률	지원되는 각 리전: 5	예	초당 최대 DeleteMember API 직접 호출 수
DeleteMembership 요청률	지원되는 각 리전: 5	예	초당 최대 DeleteMembership API 직접 호출 수
GetCollaboration 요청률	지원되는 각 리전: 5	예	초당 최대 GetCollaboration API 직접 호출 수
GetConfiguredTable 요청률	지원되는 각 리전: 20개	예	초당 최대 GetConfiguredTable API 직접 호출 수
GetConfiguredTableAnalysisRule 요청률	지원되는 각 리전: 5	예	초당 최대 GetConfiguredTableAnalysisRule API 직접 호출 수
GetConfiguredTableAssociation 요청률	지원되는 각 리전: 5	예	초당 최대 GetConfiguredTableAssociation API 직접 호출 수
GetMembership 요청률	지원되는 각 리전: 5	예	초당 최대 GetMembership API 직접 호출 수
GetProtectedJob 요청 속도	지원되는 각 리전: 5	예	초당 최대 GetProtectedJob API 호출 수
GetProtectedQuery 요청률	지원되는 각 리전: 20개	예	초당 최대 GetProtectedQuery API 직접 호출 수
GetSchema 요청률	지원되는 각 리전: 5	예	초당 최대 GetSchema API 직접 호출 수

명칭	기본값	조정 가능	설명
GetSchemaAnalysisRule 요청률	지원되는 각 리전: 5	예	초당 최대 GetSchemaAnalysisRule API 직접 호출 수
ListCollaborations 요청률	지원되는 각 리전: 5	예	초당 최대 ListCollaborations API 직접 호출 수
ListConfiguredTableAssociations 요청률	지원되는 각 리전: 5	예	초당 최대 ListConfiguredTableAssociations API 직접 호출 수
ListConfiguredTables 요청률	지원되는 각 리전: 5	예	초당 최대 ListConfiguredTables API 직접 호출 수
ListMembers 요청률	지원되는 각 리전: 5	예	초당 최대 ListMembers API 직접 호출 수
ListMemberships 요청률	지원되는 각 리전: 5	예	초당 최대 ListMemberships API 직접 호출 수
ListProtectedJobs 요청 속도	지원되는 각 리전: 5	예	초당 최대 ListProtectedJobs API 호출 수
ListProtectedQueries 요청률	지원되는 각 리전: 5	예	초당 최대 ListProtectedQueries API 직접 호출 수
ListSchemas 요청 속도	지원되는 각 리전: 5	예	초당 최대 ListSchemas API 직접 호출 수
StartProtectedJob 요청 속도	지원되는 각 리전: 5	예	초당 최대 StartProtectedJob API 호출 수

명칭	기본값	조정 가능	설명
StartProtectedQuery 요청률	지원되는 각 리전: 5	예	초당 최대 StartProtectedQuery API 직접 호출 수
UpdateCollaboration 요청률	지원되는 각 리전: 5	예	초당 최대 UpdateCollaboration API 직접 호출 수
UpdateConfiguredTable 요청률	지원되는 각 리전: 5	예	초당 최대 UpdateConfiguredTable API 직접 호출 수
UpdateConfiguredTableAnalysisRule 요청률	지원되는 각 리전: 5	예	초당 최대 UpdateConfiguredTableAnalysisRule API 직접 호출 수
UpdateConfiguredTableAssociation 요청률	지원되는 각 리전: 5	예	초당 최대 UpdateConfiguredTableAssociation API 직접 호출 수
UpdateProtectedJob 요청 속도	지원되는 각 리전: 5	예	초당 최대 UpdateProtectedJob API 호출 수
UpdateProtectedQuery 요청 속도	지원되는 각 리전: 5	예	초당 최대 UpdateProtectedQuery API 호출 수
맵버십당 테이블 연관 수	지원되는 각 지역: 25	예	멤버십당 최대 테이블 연관 수

AWS Clean Rooms 리소스 파라미터 제한

리소스	Default	설명
쿼리 텍스트 길이	90KB	SQL 쿼리문의 최대 텍스트 길이
쿼리 텍스트 길이(차등 프라이버시 사용)	90KB	차등 프라이버시를 사용하는 SQL 쿼리 문의 최대 텍스트 길이
쿼리 실행 시간	12시간	제한 시간이 초과되기 전까지 쿼리가 실행되는 최대 시간

AWS Clean Rooms API 제한 할당량

AWS 계정 에는 다음 리소스에 대한 엔드포인트 할당량당 계정당 다음과 같은 초당 API 트랜잭션(TPS)이 있습니다.

- AnalysisTemplate
- ConfiguredAudienceModelAssociation
- PrivacyBudgetTempate
- CollaborationConfiguredAudienceModelAssociation

리소스	비율 제한	설명
`BatchGetCollaborationAnalysisTemplate` 요청 비율	5TPS	초당 최대 `BatchGetCollaborationAnalysisTemplate` API 호출 수
`CreateAnalysisTemplate` 요청 비율	5TPS	초당 최대 `CreateAnalysisTemplate` API 호출 수

리소스	비율 제한	설명
CreateConfiguredAudienceModelAssociation 요청 비율	5TPS	초당 최대 CreateConfiguredAudienceModelAssociation 호출 수
CreatePrivacyBudgetTempate 요청 비율	5TPS	초당 최대 CreatePrivacyBudgetTemplate 호출 수
DeleteAnalysisTemplate 요청 비율	5TPS	초당 최대 DeleteAnalysisTemplate 호출 수
DeleteConfiguredAudienceModelAssociation 요청 비율	5TPS	초당 최대 DeleteConfiguredAudienceModelAssociation 호출 수
DeletePrivacyBudgetTemplate 요청 비율	5TPS	초당 최대 DeletePrivacyBudgetTemplate 호출 수
GetAnalysisTemplate 요청 비율	5TPS	초당 최대 GetAnalysisTemplate 호출 수
GetCollaborationConfiguredAudienceModelAssociation 요청 비율	5TPS	초당 최대 GetCollaborationConfiguredAudienceModelAssociation 호출 수
GetCollaborationPrivacyBudgetTemplate 요청 비율	5TPS	초당 최대 GetCollaborationPrivacyBudgetTemplate 호출 수
GetConfiguredAudienceModelAssociation 요청 비율	5TPS	초당 최대 GetConfiguredAudienceModelAssociation 호출 수

리소스	비율 제한	설명
GetPrivacyBudgetTemplate 요청 비율	5TPS	초당 최대 GetPrivacyBudgetTemplate 호출 수
ListAnalysisTemplates 요청 비율	5TPS	초당 최대 ListAnalysisTemplates 호출 수
ListCollaborationConfiguredAudienceModelAssociations 요청 비율	5TPS	초당 최대 ListCollaborationConfiguredAudienceModelAssociations 호출 수
ListCollaborationPrivacyBudgets 요청 비율	5TPS	초당 최대 ListCollaborationPrivacyBudgets 호출 수
ListCollaborationPrivacyBudgetTemplates 요청 비율	5TPS	초당 최대 ListCollaborationPrivacyBudgetTemplates 호출 수
ListConfiguredAudienceModelAssociations 요청 비율	5TPS	초당 최대 ListConfiguredAudienceModelAssociations 호출 수
ListPrivacyBudgets 요청 비율	5TPS	초당 최대 ListPrivacyBudgets 호출 수
ListPrivacyBudgetTemplates 요청 비율	5TPS	초당 최대 ListPrivacyBudgetTemplates 호출 수
UpdateAnalysisTemplate 요청 비율	5TPS	초당 최대 UpdateAnalysisTemplate 호출 수
UpdateConfiguredAudienceModelAssociation 요청 비율	5TPS	초당 최대 UpdateConfiguredAudienceModelAssociation 호출 수

리소스	비율 제한	설명
UpdatePrivacyBudgetTemplate 요청 비율	5TPS	초당 최대 UpdatePrivacyBudgetTemplate 호출 수

AWS Clean Rooms ML 할당량

AWS 계정 에는 AWS Clean Rooms ML에 대한 이전 제한이라고 하는 기본 할당량이 있습니다.

AWS Clean Rooms ML에 대한 서비스 할당량을 보려면 다음 중 하나를 수행합니다.

- Service Quotas 사용 설명서의 서비스 할당량 보기 단계에 따라 AWS Clean Rooms ML을 서비스로 선택합니다.
- 의 AWS Clean Rooms ML 서비스 할당량을 참조하세요Amazon Web Services 일반 참조.

서비스 성능을 유지하고 AWS Clean Rooms ML의 적절한 사용을 보장하기 위해 계정에 할당된 기본 할당량은 리전 요인, 결제 기록, 사기성 사용 및/또는 할당량 증가 요청 승인에 따라 업데이트될 수 있습니다.

명칭	기본값	조정 가능	설명
대상 생성 작업당 활성 대상 내보내기 작업	지원되는 각 리전: 25개	아니요	대상 생성 작업에 대한 최대 활성 대상 내보내기 작업 수
멤버십당 구성된 활성 모델 알고리즘 연결	지원되는 각 리전: 1,000	예	멤버십당 구성된 활성 모델 알고리즘 연결의 최대 수
멤버십당 활성 구성 모델 알고리즘	지원되는 각 리전: 1,000	예	멤버십당 활성 구성 모델 알고리즘의 최대 수

명칭	기본값	조정 가능	설명
멤버십당 활성 사용자 지정 모델 입력 채널	지원되는 각 리전: 100	예	멤버십당 활성 사용자 지정 모델 입력 채널의 최대 수
계정당 최대 활성 훈련 인스턴스 수	지원되는 각 리전: 10	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 활성 훈련 인스턴스 수입니다.
훈련된 모델당 최대 active/pending/in 훈련된 모델 버전 수	지원되는 각 리전: 100	예	훈련된 모델당 생성할 수 있는 active/pending/in 훈련된 모델 버전 수입니다.
계정당 최대 ml.c4.2xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.c4.2xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.c4.4xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.c4.4xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.c4.8xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.c4.8xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.c4.xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.c4.xlarge 훈련 인스턴스 수입니다.

명칭	기본값	조정 가능	설명
계정당 최대 ml.c5.18xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.c5.18xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.c5.2xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.c5.2xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.c5.4xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.c5.4xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.c5.9xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.c5.9xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.c5.xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.c5.xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.c5n.18xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.c5n.18xlarge 훈련 인스턴스 수입니다.

명칭	기본값	조정 가능	설명
계정당 최대 ml.c5n.2xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.c5n.2xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.c5n.4xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.c5n.4xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.c5n.9xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.c5n.9xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.c5n.xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.c5n.xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.c6i.12xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.c6i.12xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.c6i.16xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.c6i.16xlarge 훈련 인스턴스 수입니다.

명칭	기본값	조정 가능	설명
계정당 최대 ml.c6i.24xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.c6i.24xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.c6i.2xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.c6i.2xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.c6i.32xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.c6i.32xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.c6i.4xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.c6i.4xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.c6i.8xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.c6i.8xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.c6i.xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.c6i.xlarge 훈련 인스턴스 수입니다.

명칭	기본값	조정 가능	설명
계정당 최대 ml.g4dn.12xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.g4dn.12xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.g4dn.16xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.g4dn.16xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.g4dn.2xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.g4dn.2xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.g4dn.4xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.g4dn.4xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.g4dn.8xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.g4dn.8xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.g4dn.xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.g4dn.xlarge 훈련 인스턴스 수입니다.

명칭	기본값	조정 가능	설명
계정당 최대 ml.g5.12xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.g5.12xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.g5.16xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.g5.16xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.g5.24xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.g5.24xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.g5.2xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.g5.2xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.g5.48xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.g5.48xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.g5.4xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.g5.4xlarge 훈련 인스턴스의 최대 수입니다.

명칭	기본값	조정 가능	설명
계정당 최대 ml.g5.8xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.g5.8xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.g5.xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.g5.xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.m4.10xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.m4.10xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.m4.16xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.m4.16xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.m4.2xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.m4.2xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.m4.4xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.m4.4xlarge 훈련 인스턴스의 최대 수입니다.

명칭	기본값	조정 가능	설명
계정당 최대 ml.m4.xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.m4.xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.m5.12xlarge 훈련 인스턴스 수	지원되는 각 리전: 3	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.m5.12xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.m5.24xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.m5.24xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.m5.2xlarge 훈련 인스턴스 수	지원되는 각 리전: 3	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.m5.2xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.m5.4xlarge 훈련 인스턴스 수	지원되는 각 리전: 3	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.m5.4xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.m5.large 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.m5.large 훈련 인스턴스 수입니다.

명칭	기본값	조정 가능	설명
계정당 최대 ml.m5.xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.m5.xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.m6i.12xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.m6i.12xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.m6i.16xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.m6i.16xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.m6i.24xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.m6i.24xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.m6i.2xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.m6i.2xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.m6i.32xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.m6i.32xlarge 훈련 인스턴스의 최대 수입니다.

명칭	기본값	조정 가능	설명
계정당 최대 ml.m6i.4xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.m6i.4xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.m6i.8xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.m6i.8xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.m6i.large 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.m6i.large 훈련 인스턴스 수입니다.
계정당 최대 ml.m6i.xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.m6i.xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.p2.16xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.p2.16xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.p2.8xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.p2.8xlarge 훈련 인스턴스 수입니다.

명칭	기본값	조정 가능	설명
계정당 최대 ml.p2.xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.p2.xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.p3.16xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.p3.16xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.p3.2xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.p3.2xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.p3.8xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.p3.8xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.p3dn.24xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.p3dn.24xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.p4d.24xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.p4d.24xlarge 훈련 인스턴스의 최대 수입니다.

명칭	기본값	조정 가능	설명
계정당 최대 ml.p4de.24xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.p4de.24xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.p5.48xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.p5.48xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.r5.12xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.r5.12xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.r5.16xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.r5.16xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.r5.24xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.r5.24xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.r5.2xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	훈련된 모델을 생성하는 데 계정당 사용할 수 있는 ml.r5.2xlarge 훈련 인스턴스의 최대 수입니다.

명칭	기본값	조정 가능	설명
계정당 최대 ml.r5.4xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.r5.4xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.r5.8xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.r5.8xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.r5.large 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.r5.large 훈련 인스턴스 수입니다.
계정당 최대 ml.r5.xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.r5.xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.r5d.12xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.r5d.12xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.r5d.16xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.r5d.16xlarge 훈련 인스턴스 수입니다.

명칭	기본값	조정 가능	설명
계정당 최대 ml.r5d.24xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.r5d.24xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.r5d.2xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.r5d.2xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.r5d.4xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.r5d.4xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.r5d.8xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.r5d.8xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.r5d.large 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.r5d.large 훈련 인스턴스 수입니다.
계정당 최대 ml.r5d.xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.r5d.xlarge 훈련 인스턴스 수입니다.

명칭	기본값	조정 가능	설명
계정당 최대 ml.t3.2xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.t3.2xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.t3.large 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.t3.large 훈련 인스턴스 수입니다.
계정당 최대 ml.t3.medium 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.t3.medium 훈련 인스턴스 수입니다.
계정당 최대 ml.t3.xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.t3.xlarge 훈련 인스턴스 수입니다.
계정당 최대 ml.trn1.2xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.trn1.2xlarge 훈련 인스턴스의 최대 수입니다.
계정당 최대 ml.trn1.32xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 ml.trn1.32xlarge 훈련 인스턴스의 최대 수입니다.

명칭	기본값	조정 가능	설명
계정당 최대 ml.trn1n.32xlarge 훈련 인스턴스 수	지원되는 각 리전: 0	예	계정당 훈련된 모델을 생성하는 데 사용할 수 있는 최대 ml.trn1n.32xlarge 훈련 인스턴스 수입니다.
고객당 보류 중/진행 중인 대상 내보내기 작업	지원되는 각 지역: 20	아니요	고객당 보류 중/진행 중인 대상 내보내기 작업의 최대 개수
고객당 보류 중/진행 중인 대상 생성 작업	지원되는 각 리전: 10	예	고객당 보류 중/진행 중인 대상 생성 작업의 최대 개수
고객당 보류 중/진행 중인 대상 모델	지원되는 각 지역: 2	예	고객당 보류 중/진행 중인 대상 모델 훈련 작업의 최대 개수
계정당 보류 중/진행 중인 사용자 지정 모델 추론 작업	지원되는 각 리전: 10	예	계정당 보류 중/진행 중인 사용자 지정 모델 추론 작업의 최대 수
멤버십당 보류 중/진행 중인 사용자 지정 모델 추론 작업	지원되는 각 리전: 5	예	멤버십당 보류 중/진행 중인 사용자 지정 모델 추론 작업의 최대 수
계정당 보류 중/진행 중인 사용자 지정 모델 훈련 작업	지원되는 각 리전: 10	예	계정당 보류 중/진행 중인 사용자 지정 모델 훈련 작업의 최대 수
멤버십당 보류 중/진행 중인 사용자 지정 모델 훈련 작업	지원되는 각 리전: 5	예	멤버십당 보류 중/진행 중인 사용자 지정 모델 훈련 작업의 최대 수

AWS Clean Rooms ML에는 아래에 추가 할당량이 나열되어 있습니다.

Clean Rooms ML 할당량

리소스	Default	설명
사용자 지정 모델링 추론 작업의 최대 기간	25시간	
최대 상호작용 수	200억	훈련 데이터에 허용되는 최대 상호작용 수입니다. 입력값이 클수록 샘플링은 줄어듭니다.
최소 상호작용 수	100만	
유사 모델 훈련을 위한 최대 고유한 사용자 수	1억	더 많이 포함하면 상호작용 횟수를 기준으로 순위가 매겨진 상위 1억 개만 사용됩니다.
유사 모델 훈련을 위한 최소 고유한 사용자 수	100,000건	
유사 세그먼트(대상) 내보내기 작업의 최소 사용자 수	10,000	
모델 훈련에 사용되는 최대 고유한 사용자 수	100만	최대 5천만 개의 항목을 포함할 수 있지만 가장 인기 있는 1백만 개만 사용됩니다.
훈련 데이터 세트의 최대 특징 열 수입니다.	10	
사용자당 최소 고유 항목 수	2	AWS Clean Rooms ML에서는 각 행 또는 사용자에게 반복 항목을 포함하여 두 개 이상의 항목이 있어야 합니다.
시드 대상의 최대 크기	500,000	

리소스	Default	설명
시드 대상의 최소 크기	500	훈련 데이터 공급자는 이 값을 25까지 낮게 설정할 수 있습니다.
API	고객당	
총 활성 훈련 데이터 세트 수	500	
활성 유사 모델의 총 수(대상 모델)	500	
활성 구성 유사 모델의 총 수(대상 모델)	10,000	
완료된 유사 세그먼트(대상) 생성 작업의 총 수	제한 없음	
완료된 유사 세그먼트(대상) 내보내기 작업의 총 수	제한 없음	
유사 모델(대상 모델) 생성 작업의 최대 기간	1일(24시간)	
유사 세그먼트(대상) 생성 작업의 최대 기간	10시간	시드를 제공한 후 Clean Rooms ML은 유사 세그먼트를 생성하는 데 최대 10시간이 걸립니다. SQL 쿼리를 시드 데이터로 사용하는 경우 쿼리를 실행하는 데 최대 12시간이 걸리고 유사 세그먼트를 생성하는 데 10시간이 걸릴 수 있습니다.
세그먼트(대상) 크기 빈의 최소 비율	1%	
세그먼트(대상) 크기 빈의 최대 비율	20%	

리소스	Default	설명
세그먼트(대상) 크기 빈의 최소 절대 크기	개별 사용자 수의 1%	
세그먼트(대상) 크기 빈의 최대 절대 크기	개별 사용자 수의 20%	

Clean Rooms ML API 제한 할당량

AWS 계정 에는 엔드포인트 할당량당 계정당 다음과 같은 초당 API 트랜잭션(TPS)이 있습니다.

리소스	비율 제한	설명
`CreateAudienceModel` 요청 비율	1TPS 속도, 3TPS 버스트	초당 최대 `CreateAudienceModel` API 호출 수
`CreateConfiguredAudienceModel` 요청 비율	10TPS	초당 최대 `CreateConfiguredAudienceModel` API 호출 수
`CreateTrainingDataset` 요청 비율	10TPS	초당 최대 `CreateTrainingDataset` API 호출 수
`DeleteAudienceGenerationJob` 요청 비율	2TPS 속도, 10TPS 버스트	초당 최대 `DeleteAudienceGenerationJob` API 호출 수
`DeleteAudienceModel` 요청 비율	2TPS 속도, 10TPS 버스트	초당 최대 `DeleteAudienceModel` API 호출 수
`DeleteConfiguredAudienceModel` 요청 비율	10TPS	초당 최대 `DeleteConfiguredAudienceModel` API 호출 수

리소스	비율 제한	설명
DeleteConfiguredAudienceModelPolicy 요청 비율	25TPS	초당 최대 DeleteConfiguredAudienceModelPolicy API 호출 수
DeleteTrainingDataset 요청 비율	10TPS	초당 최대 DeleteTrainingDataset API 호출 수
GetAudienceGenerationJob 요청 비율	50TPS	초당 최대 GetAudienceGenerationJob API 호출 수
GetAudienceModel 요청 비율	50TPS	초당 최대 GetAudienceModel API 호출 수
GetConfiguredAudienceModel 요청 비율	50TPS	초당 최대 GetConfiguredAudienceModel API 호출 수
GetConfiguredAudienceModelPolicy 요청 비율	50TPS	초당 최대 GetConfiguredAudienceModelPolicy API 호출 수
GetTrainingDataset 요청 비율	50TPS	초당 최대 GetTrainingDataset API 호출 수
ListAudienceExportJobs 요청 비율	50TPS	초당 최대 ListAudienceExportJobs API 호출 수
ListAudienceGenerationJobs 요청 비율	50TPS	초당 최대 ListAudienceGenerationJobs API 호출 수
ListAudienceModels 요청 비율	50TPS	초당 최대 ListAudienceModels API 호출 수

리소스	비율 제한	설명
ListConfiguredAudienceModels 요청 비율	50TPS	초당 최대 ListConfiguredAudienceModels API 호출 수
ListTagsForResource 요청 비율	50TPS	초당 최대 ListTagsForResource API 호출 수
ListTrainingDatasets 요청 비율	50TPS	초당 최대 ListTrainingDatasets API 호출 수
PutConfiguredAudienceModelPolicy 요청 비율	25TPS	초당 최대 PutConfiguredAudienceModelPolicy API 호출 수
StartAudienceExportJob 요청 비율	1TPS 속도, 3TPS 버스트	초당 최대 StartAudienceExportJob API 호출 수
StartAudienceGenerationJob 요청 비율	1TPS 속도, 5TPS 버스트	초당 최대 StartAudienceGenerationJob API 호출 수
TagResource 요청 비율	10TPS	초당 최대 TagResource API 호출 수
UntagResource 요청 비율	50TPS	초당 최대 UntagResource API 호출 수
UpdateConfiguredAudienceModel 요청 비율	10TPS	초당 최대 UpdateConfiguredAudienceModel API 호출 수
CreateConfiguredModelAlgorithm 요청 비율	10TPS	초당 최대 CreateConfiguredModelAlgorithm API 호출 수입니다.

리소스	비율 제한	설명
CreateConfiguredModelAlgorithmAssociation 요청 비율	10TPS	초당 최대 CreateConfiguredModelAlgorithmAssociaton API 호출 수입니다.
PutMLConfiguration 요청 비율	10TPS	초당 최대 PutMLConfiguration API 호출 수입니다.
CreateTrainedModel 요청 비율	1TPS 속도, 3TPS 버스트	초당 최대 CreateTrainedModel API 호출 수입니다.
StartTrainedModelExportJob 요청 비율	10TPS	초당 최대 StartTrainedModelExportJob API 호출 수입니다.
StartTrainedModelInferenceJob 요청 비율	1TPS 속도, 3TPS 속도	초당 최대 StartTrainedModelInferenceJob API 호출 수입니다.
GetConfiguredModelAlgorithm 요청 속도	50TPS	초당 최대 GetConfiguredModelAlgorithm API 호출 수입니다.
GetConfiguredModelAlgorithmAssociation 요청 속도	50TPS	초당 최대 GetConfiguredModelAlgorithmAssociaton API 호출 수입니다.
GetTrainedModel 요청 비율	50TPS	초당 최대 GetTrainedModel API 호출 수입니다.
GetMLConfiguration 요청 비율	50TPS	초당 최대 GetMLConfiguration API 호출 수입니다.

리소스	비율 제한	설명
GetTrainedModelInferenceJob 요청 비율	50TPS	초당 최대 GetTrainedModelInferenceJob API 호출 수입니다.
ListConfiguredModelAlgorithm 요청 비율	50TPS	초당 최대 ListConfiguredModelAlgorithm API 호출 수입니다.
ListConfiguredModelAlgorithmAssociations 요청 비율	50TPS	초당 최대 ListConfiguredModelAlgorithmAssociatons API 호출 수입니다.
ListTrainedModels 요청 비율	50TPS	초당 최대 ListTrainedModels API 호출 수입니다.
ListCollaborationTrainedModelExportJobs 요청 비율	50TPS	초당 최대 ListCollaborationTrainedModelExportJobs API 호출 수입니다.
ListCollaborationTrainedModelInferenceJobs 요청 비율	50TPS	초당 최대 ListCollaborationTrainedModelInferenceJobs API 호출 수입니다.
DeleteConfiguredModelAlgorithm 요청 비율	2TPS 속도, 10TPS 버스트	초당 최대 DeleteConfiguredModelAlgorithm API 호출 수입니다.
DeleteConfiguredModelAlgorithmAssociation 요청 비율	2TPS 속도, 10TPS 버스트	초당 최대 DeleteConfiguredModelAlgorithmAssociaton API 요청 수입니다.

리소스	비율 제한	설명
DeleteMLConfiguration 요청 비율	2TPS 속도, 10TPS 버스트	초당 최대 DeleteMLConfiguration API 요청 수입니다.
DeleteTrainedModelOutput 요청 비율	2TPS 속도, 10TPS 버스트	초당 최대 DeleteTrainedModelOutput API 요청 수입니다.

AWS Clean Rooms 사용 설명서의 문서 기록

다음 표에서는에 대한 설명서 릴리스를 설명합니다 AWS Clean Rooms.

이 설명서에 대한 업데이트 알림을 받으려면 RSS 피드에 가입하면 됩니다. RSS 업데이트를 구독하려면 사용 중인 브라우저에서 RSS 플러그인을 활성화해야 합니다.

변경 사항	설명	날짜
ML 입력 채널에서 Parquet 파일 유형 지원	이제 고객은 ML 입력 채널에서 사용할 수 있는 데이터 형식으로 Parquet 파일 유형을 선택할 수 있습니다.	2025년 7월 17일
Spark SQL 분석 엔진의 차등 프라이버시 지원	이제 고객은 Spark SQL 분석 엔진에서 차등 프라이버시를 사용할 수 있습니다. 또한 차등 프라이버시를 사용하는 SQL 쿼리 문의 최대 텍스트 길이는 이제 90KB입니다.	2025년 7월 16일
AWS Clean Rooms SQL 분석 엔진에 대한 지원 종료 타임라인	Clean Rooms SQL 분석 엔진을 사용한 새로운 공동 작업 지원은 2025년 7월 16일에 종료됩니다. 2025년 7월 17일부터 레거시 Clean Rooms SQL 분석 엔진과 새 공동 작업을 생성할 수 없습니다.는 2025년 12월 17일에 Clean Rooms SQL 분석 엔진에 대한 지원을 AWS Clean Rooms 종료합니다.	2025년 7월 16일
증분 및 분산 훈련 지원	이제 고객은 사용자 지정 모델에 증분 훈련 및 분산 훈련을 사용할 수 있습니다.	2025년 7월 1일
Clean Rooms SQL 분석 엔진의 지원 종료 타임라인	지금부터 2025년 7월 16일까지 레거시 Clean Rooms SQL	2025년 5월 22일

	분석 엔진을 사용한 새로운 공동 작업에는 한도 증가가 필요합니다. 2025년 7월 17일부터 이 엔진으로 새 공동 작업을 생성할 수 없습니다.는 2025년 12월 17일에 Clean Rooms SQL 분석 엔진에 대한 지원을 AWS Clean Rooms 종료합니다.	
PySpark 분석 템플릿 업데이트	이제 고객은 유효한 Python 파일 이름으로 사용자 스크립트 파일의 이름을 지정할 수 있습니다.	2025년 5월 15일
여러 결과 수신기 지원	이제 고객은 여러 멤버를 선택하여 단일 쿼리에서 쿼리 결과를 받을 수 있습니다.	2025년 4월 30일
Spark SQL로 공동 작업 마이그레이션 지원	AWS Clean Rooms SQL은 이제 사용자 지정 분석 규칙 외에도 집계 및 목록 분석 규칙을 지원합니다. 또한 고객은 기존 공동 작업을 업데이트하여 Spark SQL을 지원하는 Spark 분석 엔진을 사용할 수 있습니다.	2025년 4월 2일
PySpark 작업 지원	이제 고객은 승인된 PySpark 분석 템플릿을 사용하여 작업을 실행하여 데이터를 분석할 수 있습니다.	2025년 3월 18일

기존 정책 업데이트	AWSCleanRoomsMLReadOnlyAccess 관리형 정책에 다음과 같은 새 권한이 있습니다 - PassCleanRoomsResources AWSCleanRoomsMLFullAccess 관리형 정책에 PassCleanRoomsResources 및 라는 새 권한이 추가되었습니다ConsoleDescribeECRRepositories .	2025년 1월 10일
여러 컴퓨팅 작업자 지원	이제 고객은 유사 세그먼트를 생성할 때 프로비저닝할 컴퓨팅 작업자 유형과 수를 지정할 수 있습니다.	2024년 12월 17일
여러 데이터 소스 및 클라우드 지원	이제 고객은 Amazon Athena 및 Snowflake와 같은 여러 데이터 소스와 클라우드를 사용하여 파트너의 데이터 세트와 협업할 수 있습니다.	2024년 12월 1일
Clean Rooms ML Custom Modeling 사용 가능	이제 고객은 공동 작업에서 자체 사용자 지정 ML 모델을 사용할 수 있습니다.	2024년 11월 7일
새 분석 엔진	이제 데이터 세트가 큰 고객은 Spark SQL 분석 엔진에서 지원하는 SQL 함수를 사용하여 복잡한 쿼리를 실행할 수 있습니다.	2024년 10월 29일

향상된 프라이버시 보호, 유사 대상 구축, 여러 결과 수신자 선택	추가 분석 및 공동 작업 분석 규칙을 사용하여 복잡한 활성화 쿼리를 허용하면서 데이터를 보호할 수 있습니다. SQL 쿼리 또는 분석 템플릿에서 유사 대상 모델을 구축할 수 있습니다. 결과를 수신할 구성원을 여러 명 선택할 수 있습니다.	2024년 7월 24일
의 개체 확인 AWS Clean Rooms	AWS Entity Resolution 를 입력 AWS Clean Rooms하면 두 ID 네임스페이스 간에 ID 매핑 테이블을 생성하여 서로 다른 ID 공간에서 이벤트 데이터를 쿼리할 수 있습니다.	2024년 7월 23일
기존 정책 업데이트	`AWSCleanRoomsFullAccessNoQuerying` 관리형 정책에 다음과 같은 새 권한이 있습니다 - `cleanrooms:BatchGetSchemaAnalysisRule`	2024년 5월 13일
이제 AWS Clean Rooms ML을 완전히 사용할 수 있습니다.	AWS Clean Rooms ML은 두 당사자가 데이터를 서로 공유할 필요 없이 데이터에서 유사한 사용자를 식별할 수 있는 개인 정보 보호 개선 방법을 제공합니다.	2024년 4월 3일
기존 정책 업데이트	권한을 더 잘 나타내기 위해 `AWSCleanRoomsFullAccess` 관리형 정책의 Statement ID가 ConsolePickQueryResultsBucket에서 SetQueryResultsBucket으로 업데이트되었습니다.	2024년 3월 21일

AWS Clean Rooms ML에 대한 새로운 관리형 정책	`AWSCleanRoomsMLReadOnlyAccess` , `AWSCleanRoomsMLFullAccess` 라는 2개의 새로운 관리형 정책이 추가되었습니다.	2023년 11월 29일
AWS Clean Rooms ML(미리 보기)	AWS Clean Rooms ML은 두 당사자가 데이터를 서로 공유할 필요 없이 데이터에서 유사한 사용자를 식별할 수 있는 개인 정보 보호 개선 방법을 제공합니다.	2023년 11월 29일
AWS Clean Rooms 차등 프라이버시(미리 보기)	이제 고객은 AWS Clean Rooms 차등 프라이버시를 사용하여 사용자의 개인 정보를 보호할 수 있습니다.	2023년 11월 29일
결제 구성	이제 공동 작업 생성자는 쿼리를 실행할 수 있는 구성원 또는 공동 작업의 다른 구성원에게 쿼리 컴퓨팅 비용을 청구하도록 구성할 수 있습니다.	2023년 11월 14일
쿼리 실행 시간 - 업데이트	제한 시간이 초과되기 전까지 쿼리가 실행되는 최대 시간이 4시간에서 12시간으로 업데이트되었습니다.	2023년 10월 6일

AWS CloudFormation 리소스 - 업데이트	AWS Clean Rooms 에 AWS::CleanRooms::Membership ProtectedQueryOutputConfiguration , AWS::CleanRooms::MembershipProtectedQueryResultConfiguration 및 리소스가 추가되었습니다AWS::CleanRooms::Membership ProtectedQueryS3OutputConfiguration .	2023년 9월 7일
AWS CloudFormation 리소스 - 업데이트	AWS Clean Rooms 에 AWS::CleanRooms::AnalysisTemplate 및 리소스가 추가되었습니다AWS::CleanRooms::ConfiguredTable AnalysisRuleCustom .	2023년 8월 31일
구성원 권한을 분리하세요.	공동 작업 생성자는 이제 한 구성원을 쿼리할 수 있는 구성원으로 지정하고 다른 구성원을 결과를 받을 수 있는 구성원으로 지정할 수 있습니다. 이렇게 하면 공동 작업 생성자는 쿼리를 할 수 있는 구성원이 쿼리 결과에 접근할 수 없도록 할 수 있습니다.	2023년 8월 30일
AWS Clean Rooms 용어집	AWS Clean Rooms 용어집을 추가하기 위한 설명서 전용 업데이트.	2023년 8월 30일

Apache Iceberg 테이블 지원(미리 보기)	AWS Clean Rooms 는 이제 Apache Iceberg 테이블(미리 보기)을 지원합니다.	2023년 8월 25일
할당량 업데이트	새 계정별 멤버십 기본 할당량을 반영하도록 할당량 섹션을 업데이트했습니다.	2023년 8월 9일
기존 정책에 대한 업데이트	`AWSCleanRoomsFullAccessNoQuerying` 관리형 정책에 다음과 같은 새 권한이 추가되었습니다: `cleanrooms:CreateAnalysisTemplate`, `cleanrooms:GetAnalysisTemplate`, `cleanrooms:UpdateAnalysisTemplate`, `cleanrooms:DeleteAnalysisTemplate`, `cleanrooms:ListAnalysisTemplates`, `cleanrooms:GetCollaborationAnalysisTemplate`, `cleanrooms:BatchGetCollaborationAnalysisTemplate`, 및 `cleanrooms:ListCollaborationAnalysisTemplates`.	2023년 7월 31일

분석 템플릿 및 사용자 지정 분석 규칙	AWS Clean Rooms 는 이제 분석 템플릿과 사용자 지정 분석 규칙을 지원합니다. 분석 템플릿을 사용하면 공동 작업자가 공동 작업에 사용할 사용자 지정 SQL 쿼리를 직접 작성하거나 가져올 수 있습니다. 사용자 지정 분석 규칙을 사용하면 테이블 소유자가 구성된 테이블에 대한 사용자 지정 SQL 쿼리를 승인할 수 있습니다.	2023년 7월 31일
분석 규칙은 OR 논리적 조건을 지원합니다	AWS Clean Rooms 이제 분석 규칙OR이 JOIN 절의 논리적 조건을 지원합니다.	2023년 6월 29일
CloudFormation 통합	AWS Clean Rooms 이제가와 통합됩니다 AWS CloudFormation.	2023년 6월 15일
분석 빌더	쿼리하고 결과를 받을 수 있는 구성원은 이제 SQL 코드를 작성하지 않고도 분석 빌더 UI를 사용하여 일부 테이블에 대해 쿼리를 실행할 수 있습니다.	2023년 6월 15일
SQL 함수	지원되는 SQL 함수를 명확히 하기 위한 설명서 전용 업데이트.	2023년 5월 5일
문제 해결	일반적인 문제에 대한 문제 해결 섹션을 추가하기 위한 설명서 전용 업데이트입니다.	2023년 4월 27일
에 지원되는 데이터 형식 AWS Clean Rooms	지원되는 AWS Glue Data Catalog 데이터 유형을 나열하는 새 섹션을 추가하기 위한 설명서 전용 업데이트입니다.	2023년 4월 26일

AWS CloudTrail 이벤트의 예	문서 전용 업데이트를 통해 StartProtectedQuery(성공) 및 StartProtectedQuery(실패)에 대한 CloudTrail 이벤트의 예제를 추가했습니다.	2023년 4월 20일
기존 정책에 대한 업데이트	`AWSCleanRoomsFullAccessNoQuerying` 관리형 정책에 다음과 같은 새 권한이 추가되었습니다: `cleanrooms:ListTagsForResource`, `cleanrooms:UntagResource`, 및 `cleanrooms:TagResource`. 자세한 내용은 AWS 관리형 정책 섹션을 참조하세요.	2023년 3월 21일
정식 출시	AWS Clean Rooms 이제를 정식 버전으로 사용할 수 있습니다.	2023년 3월 21일
미리 보기 릴리스	AWS Clean Rooms 사용 설명서의 미리 보기 릴리스	2023년 1월 12일

AWS Clean Rooms 용어집

AWS Clean Rooms에 사용되는 용어에 익숙해지려면 이 용어집을 참조하세요.

집계 분석 규칙

선택적 차원에 따라 COUNT, SUM, 또는 AVG 함수를 사용하여 분석을 집계하는 쿼리를 허용하는 쿼리 제한입니다. 이러한 쿼리는 행 수준 정보를 나타내지 않습니다.

캠페인 계획, 미디어 도달 범위, 빈도, 전환 측정과 같은 사용 사례를 지원합니다.

다른 유형의 분석 규칙으로는 사용자 지정 및 목록 분석 규칙이 있습니다.

분석 규칙

특정 유형의 쿼리를 승인하는 쿼리 제한.

분석 규칙 유형에 따라 구성된 테이블에서 실행할 수 있는 분석의 종류가 결정됩니다. 각 유형에는 사전 정의된 쿼리 구조가 있습니다. 쿼리 컨트롤을 통해 구조에서 테이블 열을 사용하는 방법을 제어할 수 있습니다.

분석 규칙 유형은 집계, 목록 및 사용자 지정입니다.

분석 템플릿

재사용할 수 있는 사전 승인된 공동 작업별 쿼리입니다.

지원되는 형식: Spark용 SQL 코드 또는 Python 코드.

SQL을 사용하는 경우 분석 템플릿에는 일반적으로 SQL 쿼리에 리터럴 값이 나타날 수 있는 모든 파라미터가 포함될 수 있습니다. 지원되는 파라미터 유형에 대한 자세한 내용은AWS Clean Rooms SQL 참조의 데이터 유형을 참조하세요.

분석 템플릿은 사용자 지정 분석 규칙에서만 작동합니다.

AWS Clean Rooms SQL 분석 엔진

사용자가에서 지원하는 SQL 함수를 사용하여 Amazon S3에 저장된 데이터를 쿼리할 수 AWS Clean Rooms 있는에 내장된 쿼리 처리 시스템입니다 AWS Clean Rooms. 다양한 데이터 형식을 지원하고 차등 프라이버시와 같은 기능을 포함하여 데이터 프라이버시 및 제어를 유지하면서 협업 데이터 세트

에서 SQL 쿼리를 실행하는 기능을 제공합니다. 이 엔진은 사용 사례에 맞게 AWS Clean Rooms 조정되어 SQL 기능, 데이터 개인 정보 보호 기능 및 다른 AWS Clean Rooms 기능과의 통합의 균형을 제공하므로 Spark SQL 분석 엔진의 고급 기능이나 규모가 필요하지 않은 사용자에게 적합합니다.

CreateCollaboration API를 사용하여 공동 작업을 생성할 때 AWS Clean Rooms SQL 분석 엔진의 값은 입니다`CLEAN_ROOMS_SQL`.

C3R 암호화 클라이언트

Clean Rooms (C3R) 암호화용 암호화 컴퓨팅 클라이언트.

데이터를 암호화 및 해독할 때 사용되는 C3R은 명령줄 인터페이스를 통한 클라이언트측 암호화 SDK입니다.

일반 텍스트 열

JOIN 또는 SELECT SQL 구문에 대해 암호로 보호되지 않는 열입니다.

일반 텍스트 열은 SQL 쿼리의 모든 부분에서 사용할 수 있습니다.

공동 작업

멤버 AWS Clean Rooms 가 구성된 테이블에서 SQL 쿼리를 수행할 수 있는 보안 논리적 경계입니다.

공동 작업은 공동 작업 생성자가 생성합니다.

공동 작업에 초대된 구성원만 공동 작업에 참여할 수 있습니다.

공동 작업에는 데이터를 쿼리할 수 있는 구성원 한 명 또는 쿼리 및 작업을 실행할 수 있는 구성원 한 명만 있을 수 있습니다.

공동 작업에는 결과를 받을 수 있는 멤버가 한 명만 있을 수 있습니다.

공동 작업에는 쿼리 컴퓨팅 비용을 지불하는 멤버 한 명 또는 쿼리 및 작업 컴퓨팅 비용을 지불하는 멤버 한 명만 있을 수 있습니다.

모든 구성원은 공동 작업에 참여하기 전에 공동 작업에 초대된 참여자 목록을 볼 수 있습니다.

공동 작업 생성자

공동 작업을 생성하는 구성원.

공동 작업당 공동 작업 생성자는 한 명뿐입니다.

공동 작업 생성자만 공동 작업에서 구성원을 제거하거나 공동 작업을 삭제할 수 있습니다.

구성된 테이블

구성된 각 테이블은에서 사용하도록 AWS Glue Data Catalog 구성된의 기존 테이블에 대한 참조를 나타냅니다 AWS Clean Rooms. 구성된 테이블에는 데이터 사용 방법을 결정하는 분석 규칙이 포함되어 있습니다.

현재는 카탈로그화된 Amazon Simple Storage Service(Amazon S3)에 저장된 데이터 연결을 AWS Clean Rooms 지원합니다 AWS Glue.

에 대한 자세한 내용은 AWS Glue 개발자 안내서를 AWS Glue참조하세요.

구성된 테이블을 하나 이상의 공동 작업에 연결할 수 있습니다.

> **Note**
>
> AWS Clean Rooms 는 현재에 등록된 Amazon S3 버킷 위치를 지원하지 않습니다 AWS Lake Formation.

사용자 지정 분석 규칙

특정 사전 승인된 쿼리 세트(분석 템플릿)를 허용하거나 데이터를 사용하는 쿼리 또는 작업을 제공할 수 있는 특정 계정 세트를 허용하는 쿼리 제한입니다.

퍼스트 터치 어트리뷰션, 증분 분석, 고객 발굴 분석과 같은 사용 사례를 지원합니다.

차등 프라이버시를 지원합니다.

다른 유형의 분석 규칙은 집계 및 목록입니다.

해독

암호화된 데이터를 원래 형태로 다시 변환하는 프로세스입니다. 암호 해독은 비밀 키에 대한 액세스 권한이 있는 경우에만 을 수행할 수 있습니다.

차등 프라이버시

수학적으로 엄격한 기술이며, 특정 개인에 대해 훈련한 결과를 수신할 수 있는 구성원에서 얻은 공동 작업 데이터를 보호합니다.

암호화

키라는 비밀 값을 사용하여 데이터를 무작위로 나타나는 형태로 인코딩하는 프로세스입니다. 키에 액세스하지 않고는 원본 평문을 확인할 수 없습니다.

핑거프린트 컬럼

JOIN SQL 구성에 대해 암호로 보호되는 열.

ID 매핑 워크플로 방법

ID 매핑을 수행할 방법입니다.

다음과 같은 두 가지 ID 매핑 워크플로 방법이 있습니다.

- 규칙 기반 ID 매핑 - 일치하는 규칙을 사용하여 소스의 퍼스트 파티 데이터를 ID 매핑 워크플로의 대상으로 변환하는 방법입니다.
- 공급자 서비스 ID 매핑 - 공급자 서비스를 사용하여 소스의 서드 파티 인코딩 데이터를 ID 매핑 워크플로의 대상으로 변환하는 방법입니다.

  AWS Clean Rooms 는 현재 공급자 서비스 기반 ID 매핑 워크플로 방법으로 LiveRamp를 지원합니다. 이 방법을 사용하려면를 통해 LiveRamp AWS Data Exchange 를 구독해야 합니다. 자세한 내용은 AWS Entity Resolution 사용 설명서의 AWS Data Exchange에서 공급자 서비스 구독을 참조하세요.

ID 매핑 테이블

공동 작업에서 자사 매칭 규칙 또는 다자간 자격 증명 트랜스코딩을 활성화 AWS Clean Rooms 하는의 리소스입니다.

ID 매핑 테이블은 AWS Glue Data Catalog의 기존 테이블에 대한 참조입니다. 여기에는 AWS Clean Rooms에서 데이터를 쿼리할 수 있는 방법을 결정하는 ID 매핑 테이블 분석 규칙이 포함됩니다. ID 매핑 테이블은 하나 이상의 공동 작업에 연결할 수 있습니다.

ID 매핑 테이블 분석 규칙

AWS Clean Rooms에서 관리하며 서로 다른 ID 데이터를 조인하여 쿼리를 쉽게 수행할 수 있도록 지원하는 분석 규칙 유형입니다. ID 매핑 테이블에 자동으로 추가되며 편집할 수 없습니다. 또한 공동 작업 내 다른 분석 규칙의 동작을 상속합니다. 단, 해당 분석 규칙들이 동질적이어야 합니다.

ID 매핑 워크플로

지정된 ID 매핑 워크플로 방법에 따라 데이터를 소스에서 대상으로 매핑하는 데이터 처리 작업입니다. ID 매핑 테이블을 생성합니다.

ID 네임스페이스

여러의 데이터 세트를 설명하는 메타데이터 AWS 계정 와 ID 매핑 워크플로에서 이러한 데이터 세트를 사용하는 방법을 AWS Clean Rooms 포함하는의 리소스입니다.

ID 네임스페이스 연결

ID 매핑 워크플로에 대한 입력을 검색하는 데 도움이 되는 ID 네임스페이스 리소스 연결입니다.

작업

지원되는 함수, 클래스 및 변수 세트를 사용하여 공동 작업에서 구성된 테이블에 액세스하고 분석하는 방법입니다.

AWS Clean Rooms 는 현재 PySpark 작업 유형을 지원합니다.

AWS Clean Rooms 는 현재 PySpark 분석 템플릿을 사용하여 실행 중인 작업을 지원합니다.

목록 분석 규칙

이 테이블과 쿼리할 수 있는 멤버의 테이블 간의 중복에 대한 행 수준 속성 분석을 출력하는 쿼리를 허용하는 쿼리 제한입니다.

강화, 대상 구축 또는 억제와 같은 사용 사례를 지원합니다.

다른 유형의 분석 규칙은 집계 및 사용자 지정입니다.

유사 모델

훈련 데이터 공급자의 데이터 모델로, 시드 데이터 공급자가 시드 데이터와 가장 유사한 훈련 데이터 공급자의 데이터 세그먼트를 생성할 수 있습니다시드 데이터.

유사 세그먼트

시드 데이터와 가장 유사한 훈련 데이터의 하위 집합입니다.

Member

공동 작업의 참가자인 AWS 고객입니다.

구성원은 AWS 계정를 사용하여 식별됩니다.

모든 구성원이 데이터를 제공할 수 있습니다.

쿼리할 수 있는 회원

공동 작업에서 데이터를 쿼리할 수 있는 구성원.

공동 작업당 쿼리할 수 있는 구성원은 한 명이며 해당 구성원은 변경할 수 없습니다.

관리 사용자는 AWS Identity and Access Management (IAM) 권한을 사용하여 공동 작업에서 데이터를 쿼리할 수 있는 IAM 보안 주체(예: 사용자 또는 역할)를 제어할 수 있습니다. 자세한 내용은 Amazon S3에서 데이터를 읽을 서비스 역할 생성 단원을 참조하십시오.

쿼리 및 작업을 실행할 수 있는 구성원

공동 작업의 데이터에 대해 쿼리 및 작업을 실행할 수 있는 구성원입니다.

공동 작업당 쿼리 및 작업을 실행할 수 있는 구성원은 하나뿐이며 해당 구성원은 변경할 수 없습니다.

관리 사용자는 AWS Identity and Access Management (IAM) 권한을 사용하여 공동 작업에서 쿼리 및 작업을 실행할 수 있는 IAM 보안 주체(예: 사용자 또는 역할)를 제어할 수 있습니다. 자세한 내용은 Amazon S3에서 데이터를 읽을 서비스 역할 생성 단원을 참조하십시오.

결과를 받을 수 있는 구성원

쿼리 결과를 받을 수 있는 멤버입니다. 결과를 수신할 수 있는 멤버는 Amazon S3 대상 및 쿼리 결과 형식(CSV 또는 Parquet)에 대한 쿼리 결과 설정을 지정합니다.

Spark 분석 엔진을 사용한 분석의 경우 결과를 받을 수 있는 구성원은 파일을 단일 파일로 출력할지 아니면 여러 파일로 출력할지도 지정합니다.

공동 작업에서 결과를 받을 수 있는 멤버가 두 명 이상 있을 수 있습니다.

구성원은 쿼리 컴퓨팅 비용을 지불합니다.

쿼리 컴퓨팅 비용을 지불할 책임이 있는 구성원.

공동 작업당 쿼리 컴퓨팅 비용을 지불해야 하는 구성원은 한 명뿐이며 해당 구성원은 변경할 수 없습니다.

공동 작업 생성자가 쿼리 컴퓨팅 비용을 지불하는 구성원으로 누구를 지정하지 않은 경우에는 쿼리를 할 수 있는 구성원이 기본 지불자가 됩니다.

쿼리 컴퓨팅 비용을 지불하는 구성원은 공동 작업에서 실행된 쿼리에 대한 청구서를 받습니다.

쿼리 및 작업 컴퓨팅 비용을 지불하는 구성원

쿼리 및 작업 컴퓨팅 비용을 지불할 책임이 있는 구성원입니다.

공동 작업당 쿼리 및 작업 컴퓨팅 비용을 지불할 책임이 있는 구성원은 한 명뿐이며 해당 구성원은 변경할 수 없습니다.

공동 작업 생성자가 쿼리 및 작업 컴퓨팅 비용을 지불하는 구성원으로 아무도 지정하지 않은 경우 쿼리할 수 있는 구성원이 기본 지급인입니다.

쿼리 및 작업 컴퓨팅 비용을 지불하는 구성원은 공동 작업에서 실행된 쿼리에 대한 청구서를 받습니다.

멤버십

구성원이 공동 작업에 참여할 때 생성되는 리소스입니다.

구성원이 공동 작업에 연결하는 모든 리소스는 멤버십의 일부이거나 멤버십과 연결되어 있습니다.

멤버십을 소유한 구성원만 해당 멤버십에서 리소스를 추가, 제거 또는 편집할 수 있습니다.

봉인 열

SELECT SQL 구성에 대해 암호로 보호되는 열.

시드 데이터

유사 세그먼트를 생성하는 데 사용되는 시드 데이터 공급자의 데이터입니다. 시드 데이터는 직접 제공하거나 AWS Clean Rooms 쿼리 결과에서 가져올 수 있습니다. 유사 세그먼트는 시드 사용자와 가장 유사한 훈련 데이터의 사용자 집합입니다.

Spark 분석 엔진

고객이 Apache Spark SQL 함수를 사용하여 Amazon S3, Amazon Athena 또는 Snowflake에 저장된 대규모 데이터 세트에서 복잡한 쿼리를 실행할 수 AWS Clean Rooms 있는의 분석 옵션입니다. SQL AWS Clean Rooms 분석 엔진의 대안 역할을 하며에서 PySpark 분석도 지원합니다 AWS Clean Rooms.

CreateCollaboration API를 사용하여 공동 작업을 생성할 때 Spark 분석 엔진의 값은 입니다`SPARK`.

Query

지원되는 함수, 클래스 및 변수 세트를 사용하여 공동 작업에서 구성된 테이블에 액세스하고 분석하는 방법입니다.

AWS Clean Rooms 는 현재 SQL 쿼리 언어를 지원합니다.

AWS Clean Rooms 는 현재 직접 SQL 쿼리 실행 또는 SQL 분석 템플릿을 사용한 쿼리 실행을 지원합니다.

기계 번역으로 제공되는 번역입니다. 제공된 번역과 원본 영어의 내용이 상충하는 경우에는 영어 버전이 우선합니다.