

AWS Certified Data Analytics – Specialty (DAS-C01) 시험 안내서

소개

AWS Certified Data Analytics – Specialty (DAS-C01) 시험은 데이터 분석 역할을 수행하는 개인을 대상으로 합니다. 이 시험은 AWS 서비스를 사용하여 데이터로부터 인사이트를 제공하는 분석 솔루션을 설계, 구축, 보안 및 유지 관리하는 방법에 대한 응시자의 종합적인 이해를 검증합니다.

또한 이 시험에서는 응시자가 다음 태스크를 완료할 수 있는지도 확인합니다.

- AWS 데이터 분석 서비스를 정의하고 이러한 서비스가 서로 어떻게 통합되는지 이해
- AWS 데이터 분석 서비스가 수집, 스토리지, 처리 및 시각화의 데이터 수명 주기에 어떻게 적합한지 설명

대상 응시자 설명

대상 응시자는 공통 데이터 분석 기술에 대해 최소 5년 이상의 경험을 보유하고 있어야 합니다. 또한 대상 응시자는 AWS 서비스를 사용하여 분석 솔루션을 설계, 구축, 보안 및 유지 관리한 최소 2년 이상의 실무 경험과 전문 지식을 보유하고 있어야 합니다.

대상 응시자가 갖추지 않아도 되는 것은 무엇입니까?

다음은 대상 응시자가 수행하지 않아도 될 것으로 예상되는 관련 작업 태스크의 목록입니다(전체 목록은 아님). 다음 항목은 시험 범위에 포함되지 않는 것으로 간주됩니다.

- 기계 학습 알고리즘 설계 및 구현
- 컨테이너 기반 솔루션 구현
- 고성능 컴퓨팅(HPC) 활용
- OLTP(온라인 트랜잭션 처리) 데이터베이스 솔루션 설계

시험에서 다룰 수 있는 특정 도구 및 기술에 대한 자세한 목록과 범위 포함 여부가 표시된 AWS 서비스 목록은 부록을 참조하십시오.

시험 콘텐츠

응답 유형

이 시험의 문항은 두 가지 유형으로 제공됩니다.

- **선다형:** 정답 1 개와 오답 3 개(정답 이외의 답)가 있습니다.
- **다답형:** 5 개 이상의 응답 항목 중에 2 개 이상의 정답이 있습니다.

문장을 가장 잘 완성하거나 질문에 대한 답으로 가장 적합한 응답을 하나 이상 선택합니다. 정답 이외의 답 또는 오답은 지식이나 기술이 부족한 응시자가 선택할 가능성이 큰 응답 항목입니다. 정답 이외의 답은 일반적으로 콘텐츠 영역에 부합하여 맞아 보이는 응답입니다.

답을 하지 않은 문항은 오답으로 처리됩니다. 추측에 따른 불이익은 없습니다. 시험에는 점수에 반영되는 50 개의 문항이 있습니다.

채점 대상이 아닌 콘텐츠

시험에는 점수에 반영되지 않아 채점 대상이 아닌 15 개의 문항이 포함되어 있습니다. AWS 는 채점 대상이 아닌 질문에 대한 응시자 성과 정보를 수집하여 추후 채점 대상 질문으로 사용할 수 있도록 이러한 질문을 평가합니다. 이러한 채점 대상이 아닌 질문은 시험에서 식별되지 않습니다.

시험 결과

AWS Certified Data Analytics – Specialty(DAS-C01) 시험은 합격 또는 불합격이 결정되는 시험입니다. AWS 전문가가 자격증 분야 모범 사례 및 지침에 따라 설정한 최소 표준을 기준으로 시험 점수를 매깁니다.

시험 결과는 100~1,000 기준의 스케일링된 점수로 채점됩니다. 합격 최소 점수는 750 점입니다. 응시자의 점수는 전반적인 시험 성적과 합격 여부를 보여줍니다. 스케일링된 점수 모델은 난이도가 조금씩 다를 수 있는 여러 시험 형식에 걸쳐 점수를 균등하게 조정하는 데 도움이 됩니다.

점수 보고서에는 섹션 레벨별로 성적 분류표가 포함될 수 있습니다. 이 정보는 시험 성적에 대한 일반적인 피드백을 제공하기 위한 것입니다. 시험은 보상 점수 모델을 사용하므로 각 섹션에서 합격 점수를 얻을 필요는 없습니다. 전체 시험에만 합격하면 됩니다.

시험의 섹션마다 특정 가중치가 적용되므로 일부 섹션은 다른 섹션보다 문항 수가 많습니다. 표에는 응시자의 장단점을 보여주는 일반 정보가 포함되어 있습니다. 섹션 레벨 피드백을 검토할 때 주의하시기 바랍니다.

콘텐츠 개요

이 시험 가이드는 시험의 가중치, 테스트 영역 및 목표를 제공하며, 이 시험에 대한 종합적인 콘텐츠 목록은 아닙니다. 하지만 시험을 준비하는 데 도움이 되는 각 목표에 대한 추가 배경 정보가 있습니다. 다음 표에는 주요 콘텐츠 영역과 가중치가 나열되어 있습니다. 이 표는 추가 배경 정보가 포함되어 있는 전체 시험 콘텐츠 개요 앞에 나옵니다. 각 영역의 백분율은 채점되는 콘텐츠만 나타냅니다.

영역	시험 비율(%)
영역 1: 수집	18%
영역 2: 스토리지 및 데이터 관리	22%
영역 3: 처리	24%
영역 4: 분석 및 시각화	18%
영역 5: 보안	18%
합계	100%

영역 1: 수집

1.1 수집 시스템의 작동 특성 파악

- 장애가 발생할 경우 데이터 손실이 허용 한도 내에 있는지 평가
- 다양한 소스에서 수집 시스템으로의 데이터 수집, 전송 및 프로비저닝과 관련된 비용 평가(예: 네트워킹, 대역폭, ETL/데이터 마이그레이션 비용)
- 수집 시스템에 발생할 수 있는 장애 시나리오 평가 및 영향에 따라 수정 조치 수행
- 다양한 데이터 캡처 지점에서 데이터 지속성 파악
- 수집 시스템의 대기 시간 특성 식별

1.2 빈도, 볼륨 및 데이터 원본을 처리하는 수집 시스템 선택

- 수신 데이터(스트리밍, 트랜잭션, 배치)의 볼륨 및 흐름 특성을 설명하고 규명
- 데이터의 흐름 특성을 잠재적 솔루션에 일치
- 확장성, 비용, 내결함성, 대기 시간 등을 고려하여 다양한 수집 서비스 간의 장단점 평가
- 다양한 데이터 수집 유형의 처리량 설명 및 병목 현상 식별
- 소스 데이터 시스템의 연결 제약 조건을 충족하는 수집 솔루션 선택

1.3 순서, 형식 및 압축과 같은 데이터의 주요 속성을 다루는 수집 시스템 선택

- 소스에서 데이터 변경 사항을 캡처하는 방법 설명

- 데이터 구조 및 형식, 적용된 압축 방식 및 암호화 요구 사항 논의
- 비순차적 데이터 전송 및 중복 데이터 전송이 미치는 영향과 최대 한 번, 정확히 한 번, 최소 한 번 처리 간의 장단점 구별
- 수집 프로세스 중에 데이터를 변환하고 필터링하는 방법 설명

영역 2: 스토리지 및 데이터 관리

2.1 분석을 위한 스토리지 솔루션의 운영 특성 파악

- 비용 대비 성능을 기준으로 적합한 스토리지 서비스 결정
- 요구 사항에 따라 스토리지 솔루션의 내구성, 안정성 및 대기 시간 특성 이해
- 스토리지 시스템의 강력한 일관성과 최종 일관성을 위한 시스템 요구 사항 파악
- 데이터 새로 고침 요구 사항 해결에 적합한 스토리지 솔루션 결정

2.2 데이터 액세스 및 검색 패턴 파악

- 업데이트 패턴(예: 벌크, 트랜잭션, 마이크로 배치)에 따라 적합한 스토리지 솔루션 결정
- 액세스 패턴(예: 순차적 액세스, 랜덤 액세스, 연속 사용, 임시)에 따라 적합한 스토리지 솔루션 결정
- 데이터의 변경 특성(추가 전용 변경 대 업데이트) 해결에 적합한 스토리지 솔루션 결정
- 장기 저장과 일시 저장에 적합한 스토리지 솔루션 결정
- 구조화된 데이터와 반구조화된 데이터에 적합한 스토리지 솔루션 결정
- 쿼리 대기 시간 요구 사항 해결에 적합한 스토리지 솔루션 결정

2.3 적합한 데이터 레이아웃, 스키마, 구조 및 형식 선택

- 스키마 혁신 요구 사항 해결에 적합한 메커니즘 결정
- 태스크에 사용할 스토리지 형식 선택
- 선택한 스토리지 형식에 대한 압축/인코딩 전략 선택
- 효율적인 데이터 액세스를 위한 데이터 정렬 및 배포 전략과 스토리지 레이아웃 선택
- 비용 및 성능이 다양한 데이터 배포, 레이아웃 및 형식(예: 파일 크기 및 수)에 미치는 영향 설명
- 데이터 최적화 분석을 위한 데이터 형식 및 파티셔닝 스키마 구현

2.4 사용 패턴 및 비즈니스 요구 사항에 따라 데이터 수명 주기 정의

- 데이터 수명 주기 요구 사항 해결을 위한 전략 결정
- 수명 주기 및 데이터 보존 정책을 다양한 스토리지 솔루션에 적용

2.5 데이터 카탈로그 작성 및 메타데이터 관리에 적합한 시스템 결정

- 새 데이터 원본 및 업데이트된 데이터 원본의 검색을 위한 메커니즘 평가
- 데이터 카탈로그 및 메타데이터의 생성 및 업데이트를 위한 메커니즘 평가

- 데이터 카탈로그 및 메타데이터의 검색을 위한 메커니즘 설명
- 데이터 태깅 및 분류를 위한 메커니즘 설명

영역 3: 처리

3.1 적합한 데이터 처리 솔루션 요구 사항 파악

- 데이터 준비 및 사용 요구 사항 이해
- 다양한 유형의 데이터 원본 및 대상 이해
- 성능 및 오케스트레이션 요구 사항 평가
- 비용, 확장성 및 가용성을 위해 적합한 서비스 평가

3.2 분석을 위한 데이터 변환 및 준비 솔루션 설계

- 배치 및 실시간 워크로드에 적합한 ETL/ELT 기술 적용
- 장애 조치, 확장 및 복제 메커니즘 구현
- 동시성 요구 사항 해결을 위한 기술 구현
- 비용 최적화 효율성 향상을 위한 기술 구현
- 오케스트레이션 워크플로 적용
- 다운스트림 소비를 위한 데이터 집계 및 보강

3.3 데이터 처리 솔루션 자동화 및 운용

- 반복 가능한 워크플로를 위한 자동화된 기술 구현
- 처리 장애를 식별하고 복구하는 방법 적용
- 감사 및 추적 가능성을 지원할 로깅 및 모니터링 솔루션 배포

영역 4: 분석 및 시각화

4.1 분석 및 시각화 솔루션의 운영 특성 파악

- 분석 및 시각화와 관련된 비용 파악
- 분석과 관련된 확장성 파악
- RPO/RTO 내의 장애 조치 복구 및 내결함성 파악
- 분석 도구의 가용성 특성 파악
- 동적, 대화형 및 정적 데이터 프레젠테이션 평가
- 성능 요구 사항을 적합한 시각화 접근 방식으로 변환(정적 데이터 사전 컴퓨팅 및 소비 대 동적 데이터 소비)

4.2 특정 시나리오에 적합한 데이터 분석 솔루션 선택

- 분석 솔루션 평가 및 비교
- 고객 사용 사례(스트리밍, 대화형, 협업, 운영)에 따라 적합한 분석 유형 선택

4.3 특정 시나리오에 적합한 데이터 시각화 솔루션 선택

- 특정 분석 솔루션의 출력 기능 평가(지표, KPI, 표 형식, API)
- 적합한 데이터 전송 방법 선택(예: 웹, 모바일, 이메일, 협업 노트북)
- 적합한 데이터 새로 고침 일정 선택 및 정의
- 다양한 데이터 새로 고침 요구 사항에 적합한 도구 선택(예: Amazon Elasticsearch Service, Amazon QuickSight, Amazon EMR Notebooks)
- 대화형 사용 사례를 위한 시각화 도구의 기능 이해(예: 드릴다운, 드릴스루, 피벗)
- 적합한 데이터 액세스 메커니즘 구현(예: 인메모리, 직접 액세스)
- 여러 이기종 데이터 원본에서 통합 솔루션 구현

영역 5: 보안

5.1 적합한 인증 및 권한 부여 메커니즘 선택

- 적합한 인증 방법 구현(예: 페더레이션 액세스, SSO, IAM)
- 적합한 권한 부여 방법 구현(예: 정책, ACL, 테이블/열 수준 권한)
- 적합한 액세스 제어 메커니즘 구현(예: 보안 그룹, 역할 기반 제어)

5.2 데이터 보호 및 암호화 기술 적용

- 데이터 암호화 및 마스킹 요구 사항 파악
- 다양한 암호화 접근 방식 적용(서버 측 암호화, 클라이언트 측 암호화, AWS KMS, AWS CloudHSM)
- 저장 데이터 및 전송 데이터 암호화 메커니즘 구현
- 데이터 난독화 및 마스킹 기법 구현
- 키 교체 및 비밀 관리의 기본 원칙 적용

5.3 데이터 거버넌스 및 규정 준수 제어 적용

- 데이터 거버넌스 및 규정 준수 요구 사항 파악
- 데이터 분석 서비스 전반의 액세스 및 감사 로깅 이해 및 구성
- 규정 준수 요구 사항 충족에 적합한 제어 구현

부록

시험에서 다를 수 있는 주요 도구, 기술 및 개념은 무엇입니까?

다음은 시험에서 다를 수 있는 도구 및 기술 목록입니다(전체 목록은 아님). 이 목록은 변경될 수 있으며 시험에서 다루는 서비스, 기능 또는 기술의 일반적인 범위를 이해하는 데 도움이 됩니다. AWS 서비스는

기본 기능에 따라 그룹화됩니다. 이러한 기술 중 일부는 시험에서 다른 기술보다 더 많이 다룰 수 있지만, 이 목록에서 순서 및 배치는 상대적인 비중이나 중요성을 나타내지 않습니다.

AWS 서비스 및 기능

분석:

- Amazon Athena
- Amazon CloudSearch
- Amazon Elasticsearch Service(Amazon ES)
- Amazon EMR
- AWS Glue
- Amazon Kinesis(Kinesis Video Streams 제외)
- AWS Lake Formation
- Amazon Managed Streaming for Apache Kafka
- Amazon QuickSight
- Amazon Redshift

애플리케이션 통합:

- Amazon MQ
- Amazon Simple Notification Service(Amazon SNS)
- Amazon Simple Queue Service(Amazon SQS)
- AWS Step Functions

컴퓨팅:

- Amazon EC2
- Elastic Load Balancing
- AWS Lambda

고객 지원:

- Amazon Simple Email Service(Amazon SES)

데이터베이스:

- Amazon DocumentDB(MongoDB 호환)
- Amazon DynamoDB
- Amazon ElastiCache
- Amazon Neptune
- Amazon RDS
- Amazon Redshift

- Amazon Timestream

관리 및 거버넌스:

- AWS Auto Scaling
- AWS CloudFormation
- AWS CloudTrail
- Amazon CloudWatch
- AWS Trusted Advisor

기계 학습:

- Amazon SageMaker

마이그레이션 및 전송:

- AWS Database Migration Service(AWS DMS)
- AWS DataSync
- AWS Snowball
- AWS Transfer for SFTP

네트워킹 및 콘텐츠 전송:

- Amazon API Gateway
- AWS Direct Connect
- Amazon VPC(및 관련 기능)

보안, 자격 증명 및 규정 준수:

- AWS AppSync
- AWS Artifact
- AWS Certificate Manager(ACM)
- AWS CloudHSM
- Amazon Cognito
- AWS Identity and Access Management(IAM)
- AWS Key Management Service(AWS KMS)
- Amazon Macie
- AWS Secrets Manager
- AWS Single Sign-On

스토리지:

- Amazon Elastic Block Store(Amazon EBS)
- Amazon S3

- Amazon S3 Glacier

대상 범위가 아닌 AWS 서비스 및 기능

다음은 시험에서 다루지 않는 AWS 서비스 및 기능의 목록입니다(전체 목록은 아님). 여기에 나열된 서비스와 기능이 시험 콘텐츠에서 제외되는 모든 AWS 서비스 및 기능을 나타내지는 않습니다. 시험의 대상 직무 역할과 전혀 관련이 없는 서비스 또는 기능은 관련성이 없는 것으로 간주되므로 이 목록에서 제외됩니다.

대상 범위가 아닌 AWS 서비스 및 기능에는 다음이 포함됩니다.

- AWS IoT Core