

비중인 여러분들한테 유용하고 필요한 시험가이드를 제공합니다. 만약 KoreaDumps의 제품을 구매하려면, 우리 KoreaDumps에서는 아주 디테일 한 설명과 최신버전 최고품질의 자료를 즉석중율이 높은 문제와 답을 제공합니다. Amazon Data-Engineer-Associate 자료는 충분한 시험대비자료가 될 것입니다. 안심하시고 KoreaDumps가 제공하는 상품을 사용하시고, 100% 통과율을 확신합니다.

최신 AWS Certified Data Engineer Data-Engineer-Associate 무료 샘플문제 (Q47-Q52):

질문 # 47

A company needs to generate a one-time performance report by joining data that is stored in Amazon DynamoDB, Amazon RDS, Amazon Redshift, and Amazon S3. The company wants to avoid unnecessary data movement and to minimize query execution time. Which solution will meet these requirements?

- A. Set up an AWS Glue ETL pipeline to extract, transform, and centralize data in Amazon S3. Use Amazon Athena to run analytical queries.
- B. Capture data from DynamoDB by using DynamoDB Streams. Migrate data from Amazon RDS by using AWS DMS. Export Amazon Redshift data. Store all data in Amazon S3. Use Redshift Spectrum to run queries.
- C. Deploy an Amazon EMR cluster powered by Apache Spark to ingest, process, and merge datasets from multiple sources. Run analytical workloads on the merged data.
- **D. Use Amazon Athena Federated Query to perform one-time joins and analysis across DynamoDB, Amazon RDS, Amazon Redshift, and Amazon S3.**

정답: D

설명:

Comprehensive and Detailed Explanation (150-250 words)

Amazon Athena Federated Query enables SQL queries across multiple data sources without moving data.

Athena can directly query Amazon DynamoDB, Amazon RDS, Amazon Redshift, and Amazon S3 using data source connectors. This capability is ideal for one-time or ad hoc analytical queries, where setting up ingestion pipelines would introduce unnecessary cost and delay.

Because the requirement is for a one-time performance report, the solution should minimize operational overhead and data duplication. Athena Federated Query executes queries in place, eliminating ETL processes, data exports, and staging layers. This significantly reduces query preparation time and avoids storage costs associated with copying data into a centralized location. The other options require extensive data movement and setup. DynamoDB Streams, AWS DMS, Glue ETL pipelines, and Amazon EMR clusters are appropriate for ongoing analytics pipelines, not for a single analytical report. These approaches increase operational complexity and execution time.

Athena Federated Query is purpose-built for this exact scenario and provides the fastest path to insight with minimal overhead.

질문 # 48

A data engineer is designing a new data lake architecture for a company. The data engineer plans to use Apache Iceberg tables and AWS Glue Data Catalog to achieve fast query performance and enhanced metadata handling. The data engineer needs to query historical data for trend analysis and optimize storage costs for a large volume of event data.

Which solution will meet these requirements with the LEAST development effort?

- A. Define partitioning schemes based on event type and event date.
- B. Use AWS Glue Data Catalog to automatically optimize Iceberg storage.
- C. Run a custom AWS Glue job to compact Iceberg table data files.
- **D. Store Iceberg table data files in Amazon S3 Intelligent-Tiering.**

정답: D

설명:

Amazon S3 Intelligent-Tiering is designed to optimize storage costs by automatically moving objects between access tiers based on access patterns. Since Apache Iceberg works with S3 storage, using Intelligent-Tiering provides cost-efficiency without the need for custom development or jobs.

* Option B improves performance but doesn't optimize cost automatically.

* Option C is not a real AWS Glue feature - Glue does not automatically optimize Iceberg storage.

* Option D requires custom development effort, which is contrary to the requirement.

"S3 Intelligent-Tiering is ideal for data lakes and analytics use cases that access data irregularly." Reference: AWS Documentation - S3 Intelligent-Tiering

질문 # 49

A company is migrating its database servers from Amazon EC2 instances that run Microsoft SQL Server to Amazon RDS for Microsoft SQL Server DB instances. The company's analytics team must export large data elements every day until the migration is complete. The data elements are the result of SQL joins across multiple tables. The data must be in Apache Parquet format. The analytics team must store the data in Amazon S3.

Which solution will meet these requirements in the MOST operationally efficient way?

- A. Schedule SQL Server Agent to run a daily SQL query that selects the desired data elements from the EC2 instance-based SQL Server databases. Configure the query to direct the output .csv objects to an S3 bucket. Create an S3 event that invokes an AWS Lambda function to transform the output format from .csv to Parquet.
- B. Use a SQL query to create a view in the EC2 instance-based SQL Server databases that contains the required data elements. Create and run an AWS Glue crawler to read the view. Create an AWS Glue job that retrieves the data and transfers the data in Parquet format to an S3 bucket. Schedule the AWS Glue job to run every day.
- C. Create an AWS Lambda function that queries the EC2 instance-based databases by using Java Database Connectivity (JDBC). Configure the Lambda function to retrieve the required data, transform the data into Parquet format, and transfer the data into an S3 bucket. Use Amazon EventBridge to schedule the Lambda function to run every day.
- **D. Create a view in the EC2 instance-based SQL Server databases that contains the required data elements. Create an AWS Glue job that selects the data directly from the view and transfers the data in Parquet format to an S3 bucket. Schedule the AWS Glue job to run every day.**

정답: D

설명:

Option A is the most operationally efficient way to meet the requirements because it minimizes the number of steps and services involved in the data export process. AWS Glue is a fully managed service that can extract, transform, and load (ETL) data from various sources to various destinations, including Amazon S3. AWS Glue can also convert data to different formats, such as Parquet, which is a columnar storage format that is optimized for analytics. By creating a view in the SQL Server databases that contains the required data elements, the AWS Glue job can select the data directly from the view without having to perform any joins or transformations on the source data. The AWS Glue job can then transfer the data in Parquet format to an S3 bucket and run on a daily schedule.

Option B is not operationally efficient because it involves multiple steps and services to export the data. SQL Server Agent is a tool that can run scheduled tasks on SQL Server databases, such as executing SQL queries. However, SQL Server Agent cannot directly export data to S3, so the query output must be saved as .csv objects on the EC2 instance. Then, an S3 event must be configured to trigger an AWS Lambda function that can transform the .csv objects to Parquet format and upload them to S3. This option adds complexity and latency to the data export process and requires additional resources and configuration.

Option C is not operationally efficient because it introduces an unnecessary step of running an AWS Glue crawler to read the view. An AWS Glue crawler is a service that can scan data sources and create metadata tables in the AWS Glue Data Catalog. The Data Catalog is a central repository that stores information about the data sources, such as schema, format, and location. However, in this scenario, the schema and format of the data elements are already known and fixed, so there is no need to run a crawler to discover them. The AWS Glue job can directly select the data from the view without using the Data Catalog. Running a crawler adds extra time and cost to the data export process.

Option D is not operationally efficient because it requires custom code and configuration to query the databases and transform the data. An AWS Lambda function is a service that can run code in response to events or triggers, such as Amazon EventBridge. Amazon EventBridge is a service that can connect applications and services with event sources, such as schedules, and route them to targets, such as Lambda functions. However, in this scenario, using a Lambda function to query the databases and transform the data is not the best option because it requires writing and maintaining code that uses JDBC to connect to the SQL Server databases, retrieve the required data, convert the data to Parquet format, and transfer the data to S3. This option also has limitations on the execution time, memory, and concurrency of the Lambda function, which may affect the performance and reliability of the data export process.

Reference:

AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide

AWS Glue Documentation

Working with Views in AWS Glue

Converting to Columnar Formats

질문 # 50

A data engineer uses Amazon Kinesis Data Streams to ingest and process records that contain user behavior data from an application every day.

The data engineer notices that the data stream is experiencing throttling because hot shards receive much more data than other shards in the data stream.

How should the data engineer resolve the throttling issue?

- A. Use a random partition key to distribute the ingested records.
- **B. Limit the number of records that are sent each second by the producer to match the capacity of the stream.**
- C. Increase the number of shards in the data stream. Distribute the records across the shards.
- D. Decrease the size of the records that the producer sends to match the capacity of the stream.

정답: B

질문 # 51

A manufacturing company uses AWS Glue jobs to process IoT sensor data to generate predictive maintenance models. A data engineer needs to implement automated data quality checks to identify temperature readings that are outside the expected range of 50°C to 150°C. The data quality checks must also identify records that are missing timestamp values.

The data engineer needs a solution that requires minimal coding and can automatically flag the specified issues.

Which solution will meet these requirements?

- **A. Create an AWS Glue DataBrew project to profile the sensor data. Define completeness rules for timestamps. Set up numeric range validation for temperature values.**
- B. Use AWS Glue's Data Quality rules and machine learning (ML)-based anomaly detection to identify missing timestamps and to detect temperature anomalies.
- C. Create an AWS Glue DynamicFrame that uses a custom data quality operator to profile the sensor data. Use Amazon SageMaker Data Wrangler transforms to validate timestamps and temperature ranges.
- D. Create an AWS Lambda function to scan the sensor data files to validate temperature ranges. Use AWS Glue Data Catalog tables to check timestamp completeness.

정답: A

설명:

AWS Glue DataBrew provides a no-code data preparation and validation interface. It allows you to set data profiling, completeness checks, and numeric range validations directly through its UI-ideal for IoT validation use cases.

"AWS Glue DataBrew enables users to define validation rules such as completeness and value range checks without writing code."

- Ace the AWS Certified Data Engineer - Associate Certification - version 2 - apple.pdf This fulfills the requirement for minimal coding and automatic data quality flagging.

질문 # 52

.....

Amazon인증 Data-Engineer-Associate시험을 등록했는데 마땅한 공부자료가 없어 고민중 이시라면 KoreaDumps의 Amazon인증 Data-Engineer-Associate덤프를 추천해드립니다. KoreaDumps의 Amazon인증 Data-Engineer-Associate덤프는 거의 모든 시험문제를 커버하고 있어 시험패스율이 100%입니다. KoreaDumps제품을 선택하시면 어려운 시험공부도 한결 가벼워집니다.

Data-Engineer-Associate인기시험덤프: https://www.koreadumps.com/Data-Engineer-Associate_exam-braindumps.html

가장 최신 시험에 대비하여 제작된 Amazon Data-Engineer-Associate인기시험덤프 Data-Engineer-Associate인기시험덤프 덤프는 응시자분들의 시험준비에 대한 모든 로망을 만족해드립니다, IT국제공인자격증 Amazon Data-Engineer-Associate시험대비덤프를 제공하는 전문적인 사이트로서 회원님의 개인정보를 철저하게 보호해드리고페이팔을 통한 결제라 안전한 결제를 진행할수 있습니다, 이는 응시자가 확실하고도 빠르게 Data-Engineer-Associate 시험출제 경향을 마스터하고 AWS Certified Data Engineer - Associate (DEA-C01)시험을 패스할수 있도록 하는 또 하나의 보장입니다, 지금 같은 상황에서 몇년간 Amazon Data-Engineer-Associate시험자격증만 소지한다면 일상생활에서많은 도움이 될것입니다.

조르쥬가 자신의 성별을 여자라고 밝혔음에도 불구하고, 여자들이 조르쥬를Data-Engineer-Associate보는 눈빛은 그리 변하지 않았다, 함부로 건드려서는 안 되지만, 은근한 자극으로 수사에 집중을 못하게끔 혼란을 주는 건 얼마든지 가능했다.

최신버전 Data-Engineer-Associate최신 덤프데모 덤프자료

