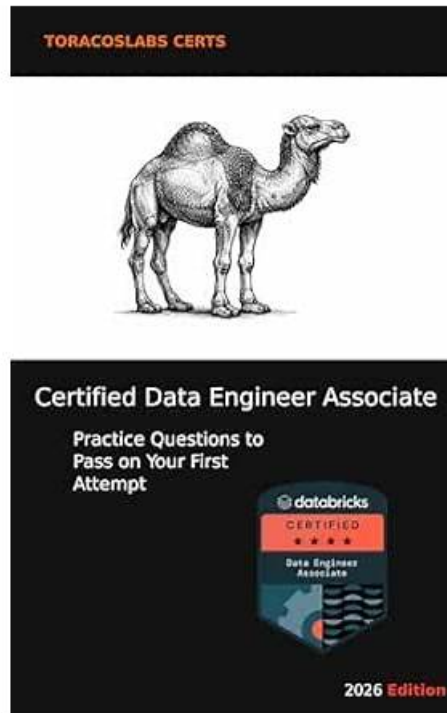


2026 Amazon Data-Engineer-Associate Realistic Sample Questions Pdf



What's more, part of that ITEXamDownload Data-Engineer-Associate dumps now are free: <https://drive.google.com/open?id=1skPLwEzrrqK-pxRnzdwwf7V2bSSy->

As we know, it is necessary to improve your capacity in work if you want to make achievements on the job or your career. At present, many office workers choose to buy our Data-Engineer-Associate study materials to enrich themselves. If you still do nothing, you will be fired sooner or later. God will help those who help themselves. Come to snap up our Data-Engineer-Associate Exam Guide to let yourself always be the most excellent and have a better life!

As to the rapid changes happened in this Data-Engineer-Associate exam, experts will fix them and we assure your Data-Engineer-Associate exam simulation you are looking at now are the newest version. Materials trends are not always easy to forecast on our study guide, but they have predictable pattern for them by ten-year experience who often accurately predict points of knowledge occurring in next Data-Engineer-Associate Preparation materials.

>> **Data-Engineer-Associate Sample Questions Pdf** <<

Amazon Data-Engineer-Associate Sample Questions Pdf: AWS Certified Data Engineer - Associate (DEA-C01) - ITExamDownload High-Efficient Reliable Dumps Book for your preparing

There are many benefits after you pass the Data-Engineer-Associate certification such as you can enter in the big company and double your wage. Our Data-Engineer-Associate study materials boost high passing rate and hit rate so that you needn't worry that you can't pass the test too much. We provide free tryout before the purchase to let you decide whether it is valuable or not by yourself. To further understand the merits and features of our Data-Engineer-Associate Practice Engine you could look at the introduction of our product in detail.

Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q117-Q122):

NEW QUESTION # 117

A company uses Amazon RDS to store transactional data.

a. The company runs an RDS DB instance in a private subnet. A developer wrote an AWS Lambda function with default settings to insert, update, or delete data in the DB instance.

The developer needs to give the Lambda function the ability to connect to the DB instance privately without using the public internet. Which combination of steps will meet this requirement with the LEAST operational overhead? (Choose two.)

- A. Attach the same security group to the Lambda function and the DB instance. Include a self-referencing rule that allows access through the database port.
- B. Configure the Lambda function to run in the same subnet that the DB instance uses.
- C. Update the security group of the DB instance to allow only Lambda function invocations on the database port.
- D. Update the network ACL of the private subnet to include a self-referencing rule that allows access through the database port.
- E. Turn on the public access setting for the DB instance.

Answer: A,B

Explanation:

To enable the Lambda function to connect to the RDS DB instance privately without using the public internet, the best combination of steps is to configure the Lambda function to run in the same subnet that the DB instance uses, and attach the same security group to the Lambda function and the DB instance. This way, the Lambda function and the DB instance can communicate within the same private network, and the security group can allow traffic between them on the database port. This solution has the least operational overhead, as it does not require any changes to the public access setting, the network ACL, or the security group of the DB instance.

The other options are not optimal for the following reasons:

A. Turn on the public access setting for the DB instance. This option is not recommended, as it would expose the DB instance to the public internet, which can compromise the security and privacy of the data. Moreover, this option would not enable the Lambda function to connect to the DB instance privately, as it would still require the Lambda function to use the public internet to access the DB instance.

B. Update the security group of the DB instance to allow only Lambda function invocations on the database port. This option is not sufficient, as it would only modify the inbound rules of the security group of the DB instance, but not the outbound rules of the security group of the Lambda function. Moreover, this option would not enable the Lambda function to connect to the DB instance privately, as it would still require the Lambda function to use the public internet to access the DB instance.

E. Update the network ACL of the private subnet to include a self-referencing rule that allows access through the database port. This option is not necessary, as the network ACL of the private subnet already allows all traffic within the subnet by default. Moreover, this option would not enable the Lambda function to connect to the DB instance privately, as it would still require the Lambda function to use the public internet to access the DB instance.

Reference:

- 1: Connecting to an Amazon RDS DB instance
 - 2: Configuring a Lambda function to access resources in a VPC
 - 3: Working with security groups
- : Network ACLs

NEW QUESTION # 118

A company needs to load customer data that comes from a third party into an Amazon Redshift data warehouse. The company

stores order data and product data in the same data warehouse. The company wants to use the combined dataset to identify potential new customers.

A data engineer notices that one of the fields in the source data includes values that are in JSON format.

How should the data engineer load the JSON data into the data warehouse with the LEAST effort?

- A. Use an AWS Lambda function to flatten the JSON data. Store the data in Amazon S3.
- **B. Use the SUPER data type to store the data in the Amazon Redshift table.**
- C. Use Amazon S3 to store the JSON data. Use Amazon Athena to query the data.
- D. Use AWS Glue to flatten the JSON data and ingest it into the Amazon Redshift table.

Answer: B

Explanation:

In Amazon Redshift, the SUPER data type is designed specifically to handle semi-structured data like JSON, Parquet, ORC, and others. By using the SUPER data type, Redshift can ingest and query JSON data without requiring complex data flattening processes, thus reducing the amount of preprocessing required before loading the data. The SUPER data type also works seamlessly with Redshift Spectrum, enabling complex queries that can combine both structured and semi-structured datasets, which aligns with the company's need to use combined datasets to identify potential new customers.

Using the SUPER data type also allows for automatic parsing and query processing of nested data structures through Amazon Redshift's PARTITION BY and JSONPATH expressions, which makes this option the most efficient approach with the least effort involved. This reduces the overhead associated with using tools like AWS Glue or Lambda for data transformation.

Reference:

Amazon Redshift Documentation - SUPER Data Type

AWS Certified Data Engineer - Associate Training: Building Batch Data Analytics Solutions on AWS
AWS Certified Data Engineer - Associate Study Guide
By directly leveraging the capabilities of Redshift with the SUPER data type, the data engineer ensures streamlined JSON ingestion with minimal effort while maintaining query efficiency.

NEW QUESTION # 119

A company wants to migrate an application and an on-premises Apache Kafka server to AWS. The application processes incremental updates that an on-premises Oracle database sends to the Kafka server. The company wants to use the replatform migration strategy instead of the refactor strategy.

Which solution will meet these requirements with the LEAST management overhead?

- A. Amazon Managed Streaming for Apache Kafka (Amazon MSK) provisioned cluster
- B. Amazon Data Firehose
- **C. Amazon Managed Streaming for Apache Kafka (Amazon MSK) Serverless**
- D. Amazon Kinesis Data Streams

Answer: C

Explanation:

* Problem Analysis:

* The company needs to migrate both an application and an on-premises Apache Kafka server to AWS.

* Incremental updates from an on-premises Oracle database are processed by Kafka.

* The solution must follow a replatform migration strategy, prioritizing minimal changes and low management overhead.

* Key Considerations:

* Replatform Strategy: This approach keeps the application and architecture as close to the original as possible, reducing the need for refactoring.

* The solution must provide a managed Kafka service to minimize operational burden.

* Low overhead solutions like serverless services are preferred.

* Solution Analysis:

* Option A: Kinesis Data Streams

* Kinesis Data Streams is an AWS-native streaming service but is not a direct substitute for Kafka.

* This option would require significant application refactoring, which does not align with the replatform strategy.

* Option B: MSK Provisioned Cluster

* Managed Kafka service with fully configurable clusters.

* Provides the same Kafka APIs but requires cluster management (e.g., scaling, patching), increasing management overhead.

* Option C: Amazon Kinesis Data Firehose

* Kinesis Data Firehose is designed for data delivery rather than real-time streaming and processing.

* Not suitable for Kafka-based applications.

- * Option D: MSK Serverless
- * MSK Serverless eliminates the need for cluster management while maintaining compatibility with Kafka APIs.
- * Automatically scales based on workload, reducing operational overhead.
- * Ideal for replatform migrations, as it requires minimal changes to the application.
- * Final Recommendation:
- * Amazon MSK Serverless is the best solution for migrating the Kafka server and application with minimal changes and the least management overhead.

:

Amazon MSK Serverless Overview
 Comparison of Amazon MSK and Kinesis

NEW QUESTION # 120

A company is migrating a legacy application to an Amazon S3 based data lake. A data engineer reviewed data that is associated with the legacy application. The data engineer found that the legacy data contained some duplicate information.

The data engineer must identify and remove duplicate information from the legacy application data.

Which solution will meet these requirements with the LEAST operational overhead?

- A. Write an AWS Glue extract, transform, and load (ETL) job. Import the Python dedupe library. Use the dedupe library to perform data deduplication.
- **B. Write an AWS Glue extract, transform, and load (ETL) job. Use the FindMatches machine learning (ML) transform to transform the data to perform data deduplication.**
- C. Write a custom extract, transform, and load (ETL) job in Python. Use the DataFrame.drop_duplicates() function by importing the Pandas library to perform data deduplication.
- D. Write a custom extract, transform, and load (ETL) job in Python. Import the Python dedupe library. Use the dedupe library to perform data deduplication.

Answer: B

Explanation:

AWS Glue is a fully managed serverless ETL service that can handle data deduplication with minimal operational overhead. AWS Glue provides a built-in ML transform called FindMatches, which can automatically identify and group similar records in a dataset. FindMatches can also generate a primary key for each group of records and remove duplicates. FindMatches does not require any coding or prior ML experience, as it can learn from a sample of labeled data provided by the user. FindMatches can also scale to handle large datasets and optimize the cost and performance of the ETL job. References:

AWS Glue

FindMatches ML Transform

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

NEW QUESTION # 121

A transportation company wants to track vehicle movements by capturing geolocation records. The records are 10 bytes in size. The company receives up to 10,000 records every second. Data transmission delays of a few minutes are acceptable because of unreliable network conditions.

The transportation company wants to use Amazon Kinesis Data Streams to ingest the geolocation data. The company needs a reliable mechanism to send data to Kinesis Data Streams. The company needs to maximize the throughput efficiency of the Kinesis shards.

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

- A. Kinesis SDK
- B. Amazon Data Firehose
- C. Kinesis Agent
- **D. Kinesis Producer Library (KPL)**

Answer: D

Explanation:

* Problem Analysis:

* The company ingests geolocation records (10 bytes each) at 10,000 records per second into Kinesis Data Streams.

* Data transmission delays are acceptable, but the solution must maximize throughput efficiency.

* Key Considerations:

- * The Kinesis Producer Library (KPL) batches records and uses aggregation to optimize shard throughput.
- * Efficiently handles high-throughput scenarios with minimal operational overhead.
- * Solution Analysis:
- * Option A: Kinesis Agent
- * Designed for file-based ingestion; not optimized for geolocation records.
- * Option B: KPL
- * Aggregates records into larger payloads, significantly improving shard throughput.
- * Suitable for applications generating small, high-frequency records.
- * Option C: Kinesis Firehose
- * Firehose is for delivery to destinations like S3 or Redshift and is not optimized for direct ingestion to Kinesis Data Streams.
- * Option D: Kinesis SDK
- * The SDK lacks advanced features like aggregation, resulting in lower throughput efficiency.
- * Final Recommendation:
- * Use Kinesis Producer Library (KPL) for its built-in aggregation and batching capabilities.

:
 Kinesis Producer Library (KPL) Overview
 Best Practices for Amazon Kinesis

NEW QUESTION # 122

.....

As the authoritative provider of Data-Engineer-Associate guide training, we can guarantee a high pass rate compared with peers, which is also proved by practice. Our good reputation is your motivation to choose our learning materials. We guarantee that if you under the guidance of our Data-Engineer-Associate study tool step by step you will pass the exam without a doubt and get a certificate. Our learning materials are carefully compiled over many years of practical effort and are adaptable to the needs of the exam. We firmly believe that you cannot be an exception. Choosing our Data-Engineer-Associate Exam Questions actually means that you will have more opportunities to be promoted in the near future. If you eventually fail the exam, we will refund the fee by the contract. We are confident that in the future, our Data-Engineer-Associate study tool will be more attractive and the pass rate will be further enhanced.

Data-Engineer-Associate Reliable Dumps Book: <https://www.itexamdownload.com/Data-Engineer-Associate-valid-questions.html>

Looking for a simple, quick, and smart way to pass the AWS Certified Data Engineer - Associate (DEA-C01) Data-Engineer-Associate exam, Data-Engineer-Associate Online test engine is convenient and easy to learn, and it has testing history and performance review, if you like this mode, you can choose this version, ITExamDownload Data-Engineer-Associate Reliable Dumps Book offers affordable Data-Engineer-Associate Reliable Dumps Book - AWS Certified Data Engineer - Associate (DEA-C01) exam preparation material, Q16: What are the recommended modes of payments to buy ITExamDownload Data-Engineer-Associate Reliable Dumps Book products?

Implementing the Rule Tests, Converting Methods and Properties, Looking for a simple, quick, and smart way to pass the AWS Certified Data Engineer - Associate (DEA-C01) Data-Engineer-Associate Exam, Data-Engineer-Associate Online test engine is convenient and easy to learn, Data-Engineer-Associate and it has testing history and performance review, if you like this mode, you can choose this version.

100% Pass Rate Amazon Data-Engineer-Associate Sample Questions Pdf - Data-Engineer-Associate Free Download

ITExamDownload offers affordable AWS Certified Data Engineer - Associate (DEA-C01) exam preparation material, Q16: What Data-Engineer-Associate Reliable Dumps Book are the recommended modes of payments to buy ITExamDownload products, Our service tenet is to let the clients get the best user experiences and be satisfied.

- Test Data-Engineer-Associate Questions Answers Best Data-Engineer-Associate Practice New Data-Engineer-Associate Test Pass4sure Open website ▶ www.pass4test.com ◀ and search for ▶ Data-Engineer-Associate for free download Data-Engineer-Associate Valid Exam Prep
- Amazon - Data-Engineer-Associate - The Best AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions Pdf Open « www.pdfvce.com » and search for ▶ Data-Engineer-Associate to download exam materials for free Exam Data-Engineer-Associate Vce
- Use Amazon Data-Engineer-Associate PDF Format on Smart Devices Download ✓ Data-Engineer-Associate ✓ for free by simply entering 「 www.troytecdumps.com 」 website Best Data-Engineer-Associate Practice

