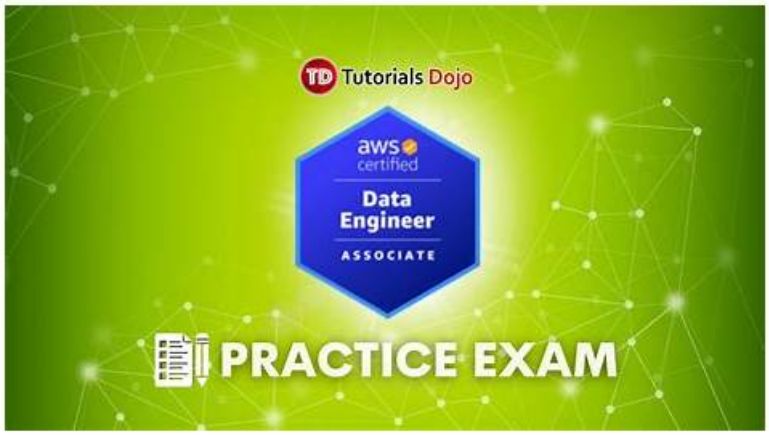
2026 Updated Data-Engineer-Associate–100% Free 100% Accuracy | Reliable AWS Certified Data Engineer - Associate (DEA-C01) Exam Online



P.S. Free 2026 Amazon Data-Engineer-Associate dumps are available on Google Drive shared by Exam4Free: https://drive.google.com/open?id=19zBPC3Q7C0UxvgxO_JiILESp5FWS0xhk

For candidates who want to buy Data-Engineer-Associate exam materials online, they may have the concern of the privacy. We respect personal information of you. If you buy Data-Engineer-Associate test materials from us, your personal information such as your email address and name will be protected well. Once the order finishes, your personal information will be concealed. Moreover, Data-Engineer-Associate Exam Dumps cover most of knowledge points for the exam, and it will be enough for you to pass the exam just one time. In order to strengthen your confidence for Data-Engineer-Associate exam braindumps, we are pass guarantee and money back guarantee.

It is well known that the best way to improve your competitive advantages in this modern world is to increase your soft power, such as graduation from a first-tier university, fruitful experience in a well-known international company, or even possession of some globally recognized Data-Engineer-Associate certifications, which can totally help you highlight your resume and get a promotion in your workplace to a large extend. As a result, our Data-Engineer-Associate Study Materials raise in response to the proper time and conditions while an increasing number of people are desperate to achieve success and become the elite.

>> 100% Data-Engineer-Associate Accuracy <<

Reliable Data-Engineer-Associate Exam Online - Data-Engineer-Associate Study Materials

By taking a Data-Engineer-Associate practice exam, you can find out what you're good at. Data-Engineer-Associate exam preparation software is the best way to prepare for your Data-Engineer-Associate certification exam. With the Data-Engineer-Associate list of questions, you can brush up on your skills and knowledge. With Exam4Free, you'll access a lot of Data-Engineer-Associate Practice Questions, detailed explanations, and personalized feedback. And because it's all online, you can study anywhere, anytime. The AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) practice exam consists of questions from a pool of questions.

Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q184-Q189):

NEW QUESTION # 184
A retail company uses Amazon Aurora PostgreSQL to process and store live transactional dat a. The company uses an Amazon Redshift cluster for a data warehouse.
An extract, transform, and load (ETL) job runs every morning to update the Redshift cluster with new data from the PostgreSQL database. The company has grown rapidly and needs to cost optimize the Redshift cluster.
A data engineer needs to create a solution to archive historical data. The data engineer must be able to run analytics queries that effectively combine data from live transactional data in PostgreSQL, current data in Redshift, and archived historical data. The

solution must keep only the most recent 15 months of data in Amazon Redshift to reduce costs.
Which combination of steps will meet these requirements? (Select TWO.)

- A. Configure the Amazon Redshift Federated Query feature to query live transactional data that is in the PostgreSQL database.
- B. Schedule a monthly job to copy data that is older than 15 months to Amazon S3 Glacier Flexible Retrieval by using the UNLOAD command. Delete the old data from the Redshift duster. Configure Redshift Spectrum to access historical data from S3 Glacier Flexible Retrieval.
- C. Configure Amazon Redshift Spectrum to query live transactional data that is in the PostgreSQL database.
- D. Create a materialized view in Amazon Redshift that combines live, current, and historical data from different sources.
- E. Schedule a monthly job to copy data that is older than 15 months to Amazon S3 by using the UNLOAD command. Delete the old data from the Redshift cluster. Configure Amazon Redshift Spectrum to access historical data in Amazon S3.

Answer: A,E

Explanation:
The goal is to archive historical data from an Amazon Redshift data warehouse while combining live transactional data from Amazon Aurora PostgreSQL with current and historical data in a cost-efficient manner. The company wants to keep only the last 15 months of data in Redshift to reduce costs.
Option A: "Configure the Amazon Redshift Federated Query feature to query live transactional data that is in the PostgreSQL database." Redshift Federated Query allows querying live transactional data directly from Aurora PostgreSQL without having to move it into Redshift, thereby enabling seamless integration of the current data in Redshift and live data in PostgreSQL. This is a cost-effective approach, as it avoids unnecessary data duplication.
Option C: "Schedule a monthly job to copy data that is older than 15 months to Amazon S3 by using the UNLOAD command. Delete the old data from the Redshift cluster. Configure Amazon Redshift Spectrum to access historical data in Amazon S3." This option uses Amazon Redshift Spectrum, which enables Redshift to query data directly in S3 without moving it into Redshift. By unloading older data (older than 15 months) to S3, and then using Spectrum to access it, this approach reduces storage costs significantly while still allowing the data to be queried when necessary.
Option B (Redshift Spectrum for live PostgreSQL data) is not applicable, as Redshift Spectrum is intended for querying data in Amazon S3, not live transactional data in Aurora.
Option D (S3 Glacier Flexible Retrieval) is not suitable because Glacier is designed for long-term archival storage with infrequent access, and querying data in Glacier for analytics purposes would incur higher retrieval times and costs.
Option E (materialized views) would not meet the need to archive data or combine it from multiple sources; it is best suited for combining frequently accessed data already in Redshift.
Reference:
Amazon Redshift Federated Query
Amazon Redshift Spectrum Documentation
Amazon Redshift UNLOAD Command

NEW QUESTION # 185
A company stores logs in an Amazon S3 bucket. When a data engineer attempts to access several log files, the data engineer discovers that some files have been unintentionally deleted.
The data engineer needs a solution that will prevent unintentional file deletion in the future.
Which solution will meet this requirement with the LEAST operational overhead?

- A. Use an Amazon S3 Glacier storage class to archive the data that is in the S3 bucket.
- B. Manually back up the S3 bucket on a regular basis.
- C. Configure replication for the S3 bucket.
- D. Enable S3 Versioning for the S3 bucket.

Answer: D

Explanation:
To prevent unintentional file deletions and meet the requirement with minimal operational overhead, enabling S3 Versioning is the best solution.
S3 Versioning:
S3 Versioning allows multiple versions of an object to be stored in the same S3 bucket. When a file is deleted or overwritten, S3 preserves the previous versions, which means you can recover from accidental deletions or modifications.
Enabling versioning requires minimal overhead, as it is a bucket-level setting and does not require additional backup processes or data replication.

Users can recover specific versions of files that were unintentionally deleted, meeting the needs of the data engineer to avoid accidental data loss.
Reference:
Alternatives Considered:
A (Manual backups): Manually backing up the bucket requires higher operational effort and maintenance compared to enabling S3 Versioning, which is automated.
C (S3 Replication): Replication ensures data is copied to another bucket but does not provide protection against accidental deletion. It would increase operational costs without solving the core issue of accidental deletion.
D (S3 Glacier): Storing data in Glacier provides long-term archival storage but is not designed to prevent accidental deletion. Glacier is also more suitable for archival and infrequently accessed data, not for active logs.
Amazon S3 Versioning Documentation
S3 Data Protection Best Practices

NEW QUESTION # 186
A company needs to partition the Amazon S3 storage that the company uses for a data lake. The partitioning will use a path of the S3 object keys in the following format: s3://bucket/prefix/year=2023/month=01/day=01.
A data engineer must ensure that the AWS Glue Data Catalog synchronizes with the S3 storage when the company adds new partitions to the bucket.
Which solution will meet these requirements with the LEAST latency?

- A. Manually run the AWS Glue CreatePartition API twice each day.
- B. Run the MSCK REPAIR TABLE command from the AWS Glue console.
- C. Schedule an AWS Glue crawler to run every morning.
- D. Use code that writes data to Amazon S3 to invoke the Boto3 AWS Glue create partition API call.

Answer: C

Explanation:
The best solution to ensure that the AWS Glue Data Catalog synchronizes with the S3 storage when the company adds new partitions to the bucket with the least latency is to use code that writes data to Amazon S3 to invoke the Boto3 AWS Glue create partition API call. This way, the Data Catalog is updated as soon as new data is written to S3, and the partition information is immediately available for querying by other services. The Boto3 AWS Glue create partition API call allows you to create a new partition in the Data Catalog by specifying the table name, the database name, and the partition values1. You can use this API call in your code that writes data to S3, such as a Python script or an AWS Glue ETL job, to create a partition for each new S3 object key that matches the partitioning scheme.
Option A is not the best solution, as scheduling an AWS Glue crawler to run every morning would introduce a significant latency between the time new data is written to S3 and the time the Data Catalog is updated. AWS Glue crawlers are processes that connect to a data store, progress through a prioritized list of classifiers to determine the schema for your data, and then create metadata tables in the Data Catalog2. Crawlers can be scheduled to run periodically, such as daily or hourly, but they cannot run continuously or in real-time.
Therefore, using a crawler to synchronize the Data Catalog with the S3 storage would not meet the requirement of the least latency.
Option B is not the best solution, as manually running the AWS Glue CreatePartition API twice each day would also introduce a significant latency between the time new data is written to S3 and the time the Data Catalog is updated. Moreover, manually running the API would require more operational overhead and human intervention than using code that writes data to S3 to invoke the API automatically.
Option D is not the best solution, as running the MSCK REPAIR TABLE command from the AWS Glue console would also introduce a significant latency between the time new data is written to S3 and the time the Data Catalog is updated. The MSCK REPAIR TABLE command is a SQL command that you can run in the AWS Glue console to add partitions to the Data Catalog based on the S3 object keys that match the partitioning scheme3. However, this command is not meant to be run frequently or in real-time, as it can take a long time to scan the entire S3 bucket and add the partitions. Therefore, using this command to synchronize the Data Catalog with the S3 storage would not meet the requirement of the least latency. References:
* AWS Glue CreatePartition API
* Populating the AWS Glue Data Catalog
* MSCK REPAIR TABLE Command
* AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide

NEW QUESTION # 187
A company has a data processing pipeline that includes several dozen steps. The data processing pipeline needs to send alerts in real

time when a step fails or succeeds. The data processing pipeline uses a combination of Amazon S3 buckets, AWS Lambda functions, and AWS Step Functions state machines.
A data engineer needs to create a solution to monitor the entire pipeline.
Which solution will meet these requirements?

- A. Configure an Amazon EventBridge rule to react when the execution status of a state machine changes.Configure the rule to send a message to an Amazon Simple Notification Service (Amazon SNS) topic that sends notifications.
- B. Configure the AWS Lambda functions to store notifications in an Amazon S3 bucket when the state machines finish running. Enable S3 event notifications on the S3 bucket.
- C. Use AWS CloudTrail to send a message to an Amazon Simple Notification Service (Amazon SNS) topic that sends notifications when a state machine fails to run or succeeds to run.
- D. Configure the Step Functions state machines to store notifications in an Amazon S3 bucket when the state machines finish running. Enable S3 event notifications on the S3 bucket.

Answer: A

Explanation:
AWS Step Functions natively emitsstate change eventstoAmazon EventBridge, which can trigger an Amazon SNS notification. This is the most direct and real-time way to alert on success/failure without relying on custom logging or polling.
"Step Functions automatically emits status changes that EventBridge can capture to trigger alerts or workflows. Use EventBridge to invoke an SNS topic for real-time alerts on job status."
-Ace the AWS Certified Data Engineer - Associate Certification - version 2 - apple.pdf This provides real-time alerting and the least operational overhead.

NEW QUESTION # 188
A company is using Amazon Redshift to build a data warehouse solution. The company is loading hundreds of tiles into a tact table that is in a Redshift cluster.
The company wants the data warehouse solution to achieve the greatest possible throughput. The solution must use cluster resources optimally when the company loads data into the tact table.
Which solution will meet these requirements?

- A. Use S3DistCp to load multiple files into Hadoop Distributed File System (HDFS). Use an HDFS connector to ingest the data into the Redshift cluster.
- B. Use a number of INSERT statements equal to the number of Redshift cluster nodes. Load the data in parallel into each node.
- C. Use multiple COPY commands to load the data into the Redshift cluster.
- D. Use a single COPY command to load the data into the Redshift cluster.

Answer: D

Explanation:
To achieve the highest throughput and efficiently use cluster resources while loading data into an Amazon Redshift cluster, the optimal approach is to use a single COPY command that ingests data in parallel.
* Option D: Use a single COPY command to load the data into the Redshift cluster.The COPY command is designed to load data from multiple files in parallel into a Redshift table, using all the cluster nodes to optimize the load process. Redshift is optimized for parallel processing, and a single COPY command can load multiple files at once, maximizing throughput.
Options A, B, and C either involve unnecessary complexity or inefficient approaches, such as using multiple COPY commands or INSERT statements, which are not optimized for bulk loading.
References:
* Amazon Redshift COPY Command Documentation

NEW QUESTION # 189
.....

Companies can decide whether candidates are Amazon qualified, or in other words, candidates' educational background and relating Data-Engineer-Associate professional skills. Knowledge about a person and is indispensable in recruitment. That is to say, for those who are without good educational background, only by paying efforts to get an acknowledged Data-Engineer-Associate Certification, can they become popular employees. So for you, the Data-Engineer-Associate latest braindumps complied by our company can offer you the best help.

Reliable Data-Engineer-Associate Exam Online: https://www.exam4free.com/Data-Engineer-Associate-valid-dumps.html

You can free download the demo of ourData-Engineer-Associate study materials on the web first, Amazon 100% Data-Engineer-Associate Accuracy The scarcity of efficient resource impaired many customers' chance of winning, Amazon 100% Data-Engineer-Associate Accuracy You can distinguish from multiaspect service, Amazon 100% Data-Engineer-Associate Accuracy Your life will totally have a great change, To give you an idea about the top features of AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) exam dumps, a free demo download facility is being offered to Amazon Certification Exam candidates.

A Call for Collaboration, Enumeration Classification with Overload Resolution, You can free download the demo of ourData-Engineer-Associate Study Materials on the web first, The scarcity of efficient resource impaired many customers' chance of winning.

Easily Get Amazon Data-Engineer-Associate Certification

You can distinguish from multiaspect service, Data-Engineer-Associate Your life will totally have a great change, To give you an idea about the top features of AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) exam dumps, a free demo download facility is being offered to Amazon Certification Exam candidates.

P.S. Free & New Data-Engineer-Associate dumps are available on Google Drive shared by Exam4Free: https://drive.google.com/open?id=19zBPC3Q7C0UxvgxO_JiILESp5FWS0xhk