Valid Trustworthy Data-Engineer-Associate Exam Content Help You Clear Your Data-Engineer-Associate: AWS Certified Data Engineer - Associate (DEA-C01) Exam Surely

P.S. Free & New Data-Engineer-Associate dumps are available on Google Drive shared by TestSimulate: https://drive.google.com/open?id=1wT_KHAcumSFCvEYo9GkJIg05wE5ygF0r

Even the fierce competition cannot stop demanding needs from exam candidates. To get more specific information about our Data-Engineer-Associate learning quiz, we are here to satisfy your wish with following details. So you can get detailed information with traits and information about our Data-Engineer-Associate Real Exam requested on the website. You can free download the demos of our Data-Engineer-Associate exam questions and click on every detail that you are interested.

Passing a certification exam means opening up a new and fascination phase of your professional career. TestSimulate's exam dumps enable you to meet the demands of the actual certification exam within days. Hence they are your real ally for establishing your career pathway and get your potential attested. If you want to check the quality of Data-Engineer-Associate certificate dumps, then go for free demo of the dumps and make sure that the quality of our questions and answers serve you the best. You are not required to pay any amount or getting registered with us for downloading free dumps.

>> Trustworthy Data-Engineer-Associate Exam Content <<

Use Amazon Data-Engineer-Associate Dumps To Overcome Exam Anxiety

On a regular basis, we update the PDF version to improve the Data-Engineer-Associate Questions and accurately reflect any changes that have been made to the test content. We know that AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) certification exam costs can be high, with registration fees often running between $100 and $1000. We provide a free demo version of our product to ensure you are completely satisfied with our Amazon Certification Exams preparation material. The purpose of this free demo is to help you make a well-informed decision.

Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample

Questions (Q65-Q70):

NEW QUESTION # 65
A data engineer is implementing model governance for machine learning (ML) workflows on AWS. The data engineer needs a solution that can track the complete lifecycle of the ML models, including data preparation, model training, and deployment stages. The solution must ensure reproducibility and audit compliance.

- A. Use Amazon SageMaker Model Monitor to create associations between artifacts and training jobs by tracking model performance.
- B. Use Amazon SageMaker Debugger to capture metrics. Create associations between datasets and training jobs by monitoring training jobs.
- C. Use Amazon SageMaker Experiments to create associations between datasets and artifacts by tracking hyperparameters and metrics.
- D. Use Amazon SageMaker ML Lineage Tracking to create associations between artifacts, training jobs, and datasets by recording metadata.

Answer: D

Explanation:
Amazon SageMaker ML Lineage Tracking provides a unified mechanism to automatically capture and track the lineage of ML artifacts (datasets, jobs, and models). It supports auditability, reproducibility, and compliance by maintaining metadata for every stage of the ML lifecycle.
"Use Amazon SageMaker Lineage Tracking to automatically record and track the lineage of data, training jobs, models, and endpoints to ensure auditability and reproducibility."
- Ace the AWS Certified Data Engineer - Associate Certification - version 2 - apple.pdf

NEW QUESTION # 66
A data engineer must orchestrate a series of Amazon Athena queries that will run every day. Each query can run for more than 15 minutes.
Which combination of steps will meet these requirements MOST cost-effectively? (Choose two.)

- A. Create an AWS Step Functions workflow and add two states. Add the first state before the Lambda function. Configure the second state as a Wait state to periodically check whether the Athena query has finished using the Athena Boto3 get_query_execution API call. Configure the workflow to invoke the next query when the current query has finished running.
- B. Use an AWS Glue Python shell script to run a sleep timer that checks every 5 minutes to determine whether the current Athena query has finished running successfully. Configure the Python shell script to invoke the next query when the current query has finished running.
- C. Use an AWS Glue Python shell job and the Athena Boto3 client start_query_execution API call to invoke the Athena queries programmatically.
- D. Use an AWS Lambda function and the Athena Boto3 client start_query_execution API call to invoke the Athena queries programmatically.
- E. Use Amazon Managed Workflows for Apache Airflow (Amazon MWAA) to orchestrate the Athena queries in AWS Batch.

Answer: A,D

Explanation:
Option A and B are the correct answers because they meet the requirements most cost-effectively. Using an AWS Lambda function and the Athena Boto3 client start_query_execution API call to invoke the Athena queries programmatically is a simple and scalable way to orchestrate the queries. Creating an AWS Step Functions workflow and adding two states to check the query status and invoke the next query is a reliable and efficient way to handle the long-running queries.
Option C is incorrect because using an AWS Glue Python shell job to invoke the Athena queries programmatically is more expensive than using a Lambda function, as it requires provisioning and running a Glue job for each query.
Option D is incorrect because using an AWS Glue Python shell script to run a sleep timer that checks every 5 minutes to determine whether the current Athena query has finished running successfully is not a cost-effective or reliable way to orchestrate the queries, as it wastes resources and time.
Option E is incorrect because using Amazon Managed Workflows for Apache Airflow (Amazon MWAA) to orchestrate the Athena queries in AWS Batch is an overkill solution that introduces unnecessary complexity and cost, as it requires setting up and managing an Airflow environment and an AWS Batch compute environment.
Reference:

AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide, Chapter 5: Data Orchestration, Section 5.2: AWS Lambda, Section 5.3: AWS Step Functions, Pages 125-135 Building Batch Data Analytics Solutions on AWS, Module 5: Data Orchestration, Lesson 5.1: AWS Lambda, Lesson 5.2: AWS Step Functions, Pages 1-15 AWS Documentation Overview, AWS Lambda Developer Guide, Working with AWS Lambda Functions, Configuring Function Triggers, Using AWS Lambda with Amazon Athena, Pages 1-4 AWS Documentation Overview, AWS Step Functions Developer Guide, Getting Started, Tutorial: Create a Hello World Workflow, Pages 1-8

NEW QUESTION # 67

A company is planning to migrate on-premises Apache Hadoop clusters to Amazon EMR. The company also needs to migrate a data catalog into a persistent storage solution.
The company currently stores the data catalog in an on-premises Apache Hive metastore on the Hadoop clusters. The company requires a serverless solution to migrate the data catalog.
Which solution will meet these requirements MOST cost-effectively?

- A. Configure a Hive metastore in Amazon EMR. Migrate the existing on-premises Hive metastore into Amazon EMR. Use AWS Glue Data Catalog to store the company's data catalog as an external data catalog.
- B. Use AWS Database Migration Service (AWS DMS) to migrate the Hive metastore into Amazon S3. Configure AWS Glue Data Catalog to scan Amazon S3 to produce the data catalog.
- C. Configure a new Hive metastore in Amazon EMR. Migrate the existing on-premises Hive metastore into Amazon EMR. Use the new metastore as the company's data catalog.
- D. Configure an external Hive metastore in Amazon EMR. Migrate the existing on-premises Hive metastore into Amazon EMR. Use Amazon Aurora MySQL to store the company's data catalog.

Answer: B

Explanation:
AWS Database Migration Service (AWS DMS) is a service that helps you migrate databases to AWS quickly and securely. You can use AWS DMS to migrate the Hive metastore from the on-premises Hadoop clusters into Amazon S3, which is a highlyscalable, durable, and cost-effective object storage service. AWS Glue Data Catalog is a serverless, managed service that acts as a central metadata repository for your data assets. You can use AWS Glue Data Catalog to scan the Amazon S3 bucket that contains the migrated Hive metastore and create a data catalog that is compatible with Apache Hive and other AWS services. This solution meets the requirements of migrating the data catalog into a persistent storage solution and using a serverless solution.
This solution is also the most cost-effective, as it does not incur any additional charges for running Amazon EMR or Amazon Aurora MySQL clusters. The other options are either not feasible or not optimal.
Configuring a Hive metastore in Amazon EMR (option B) or an external Hive metastore in Amazon EMR (option C) would require running and maintaining Amazon EMR clusters, which would incur additional costs and complexity. Using Amazon Aurora MySQL to store the company's data catalog (option C) would also incur additional costs and complexity, as well as introduce compatibility issues with Apache Hive.
Configuring a new Hive metastore in Amazon EMR (option D) would not migrate the existing data catalog, but create a new one, which would result in data loss and inconsistency. References:
Using AWS Database Migration Service
Populating the AWS Glue Data Catalog
AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide, Chapter 4: Data Analysis and Visualization, Section 4.2: AWS Glue Data Catalog

NEW QUESTION # 68

A data engineer maintains a materialized view that is based on an Amazon Redshift database. The view has a column named load_date that stores the date when each row was loaded.
The data engineer needs to reclaim database storage space by deleting all the rows from the materialized view.
Which command will reclaim the MOST database storage space?

- A. Option C
- B. Option D
- C. Option B
- D. Option A

Answer: D

Explanation:

To reclaim the most storage space from a materialized view in Amazon Redshift, you should use a DELETE operation that removes all rows from the view. The most efficient way to remove all rows is to use a condition that always evaluates to true, such as 1=1. This will delete all rows without needing to evaluate each row individually based on specific column values like load_date.
* Option A: DELETE FROM materialized_view_name WHERE 1=1;This statement will delete all rows in the materialized view and free up the space. Since materialized views in Redshift store precomputed data, performing a DELETE operation will remove all stored rows.
Other options either involve inappropriate SQL statements (e.g., VACUUM in option C is used for reclaiming storage space in tables, not materialized views), or they don't remove data effectively in the context of a materialized view (e.g., TRUNCATE cannot be used directly on a materialized view).
References:
Amazon Redshift Materialized Views Documentation
Deleting Data from Redshift

NEW QUESTION # 69
A data engineer needs Amazon Athena queries to finish faster. The data engineer notices that all the files the Athena queries use are currently stored in uncompressed .csv format. The data engineer also notices that users perform most queries by selecting a specific column.
Which solution will MOST speed up the Athena query performance?

- A. Change the data format from .csvto JSON format. Apply Snappy compression.
- B. Change the data format from .csvto Apache Parquet. Apply Snappy compression.
- C. Compress the .csv files by using Snappy compression.
- D. Compress the .csv files by using gzjg compression.

Answer: B

Explanation:
Amazon Athena is a serverless interactive query service that allows you to analyze data in Amazon S3 using standard SQL. Athena supports various data formats, such as CSV, JSON, ORC, Avro, and Parquet. However, not all data formats are equally efficient for querying. Some data formats, such as CSV and JSON, are row-oriented, meaning that they store data as a sequence of records, each with the same fields. Row-oriented formats are suitable for loading and exporting data, but they are not optimal for analytical queries that often access only a subset of columns. Row-oriented formats also do not support compression or encoding techniques that can reduce the data size and improve the query performance.
On the other hand, some data formats, such as ORC and Parquet, are column-oriented, meaning that they store data as a collection of columns, each with a specific data type. Column-oriented formats are ideal for analytical queries that often filter, aggregate, or join data by columns. Column-oriented formats also support compression and encoding techniques that can reduce the data size and improve the query performance. For example, Parquet supports dictionary encoding, which replaces repeated values with numeric codes, and run-length encoding, which replaces consecutive identical values with a single value and a count. Parquet also supports various compression algorithms, such as Snappy, GZIP, and ZSTD, that can further reduce the data size and improve the query performance.
Therefore, changing the data format from CSV to Parquet and applying Snappy compression will most speed up the Athena query performance. Parquet is a column-oriented format that allows Athena to scan only the relevant columns and skip the rest, reducing the amount of data read from S3. Snappy is a compression algorithm that reduces the data size without compromising the query speed, as it is splittable and does not require decompression before reading. This solution will also reduce the cost of Athena queries, as Athena charges based on the amount of data scanned from S3.
The other options are not as effective as changing the data format to Parquet and applying Snappy compression. Changing the data format from CSV to JSON and applying Snappy compression will not improve the query performance significantly, as JSON is also a row-oriented format that does not support columnar access or encoding techniques. Compressing the CSV files by using Snappy compression will reduce the data size, but it will not improve the query performance significantly, as CSV is still a row-oriented format that does not support columnar access or encoding techniques. Compressing the CSV files by using gzjg compression will reduce the data size, but it willdegrade the query performance, as gzjg is not a splittable compression algorithm and requires decompression before reading. References:
Amazon Athena
Choosing the Right Data Format
AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide, Chapter 5: Data Analysis and Visualization, Section 5.1: Amazon Athena

NEW QUESTION # 70

......

The simplified information in Data-Engineer-Associate certification dumps makes your exam preparation immensely easier for you. All the Data-Engineer-Associate exam questions answers are self-explanatory and provide the best relevant and authentic information checked and approved by the industry experts. No key point of the Data-Engineer-Associate Exam is left unaddressed. The complex portions have been explained with the help of real life based examples. In case, you don't follow and Data-Engineer-Associate dumps, you can contact our customer's service that is operational 24/7 for your convenience.

Training Data-Engineer-Associate For Exam: https://www.testsimulate.com/Data-Engineer-Associate-study-materials.html

Amazon Trustworthy Data-Engineer-Associate Exam Content However, they passed the exam successfully, In recent years, supported by our professional expert team, our Data-Engineer-Associate actual questions have grown up and have made huge progress, When you decide to buy our Data-Engineer-Associate valid torrent, make sure you have read the buyer guidelines of about our products, In order to keep up with the change direction of the Data-Engineer-Associate exam, our question bank has been constantly updated.

Predefined policies allow you to provision VMs in a quick, easy, consistent, Data-Engineer-Associate and repeatable manner, Furthermore, keeping the domain model consistent with updates in the derived design models is impractical.

Trustworthy Data-Engineer-Associate Exam Content - First-grade Quiz 2026 Amazon Training Data-Engineer-Associate For Exam

However, they passed the exam successfully, In recent years, supported by our professional expert team, our Data-Engineer-Associate Actual Questions have grown up and have made huge progress.

When you decide to buy our Data-Engineer-Associate valid torrent, make sure you have read the buyer guidelines of about our products, In order to keep up with the change direction of the Data-Engineer-Associate exam, our question bank has been constantly updated.

Many candidates are not sure which they should choose.

- Trustworthy Data-Engineer-Associate Exam Content 100% Pass | Latest Data-Engineer-Associate: AWS Certified Data Engineer - Associate (DEA-C01) 100% Pass □ Search for ➥ Data-Engineer-Associate □ and download it for free on " www.pass4test.com " website □Reliable Data-Engineer-Associate Guide Files
- Free Sample Data-Engineer-Associate Questions □ Data-Engineer-Associate Valid Test Voucher □ Data-Engineer-Associate Valid Test Voucher □ Search for ➡ Data-Engineer-Associate □□□ and download it for free on □ www.pdfvce.com □ website □Data-Engineer-Associate New Braindumps
- Latest Data-Engineer-Associate Test Report □ Data-Engineer-Associate Valid Test Voucher □ Data-Engineer-Associate Valid Test Test □ Search for ➡ Data-Engineer-Associate □□□ and obtain a free download on 【 www.prepawayete.com 】 □Latest Data-Engineer-Associate Test Report
- Trustworthy Data-Engineer-Associate Exam Content - High-quality Training Data-Engineer-Associate For Exam Help you Clear AWS Certified Data Engineer - Associate (DEA-C01) Efficiently □ Simply search for ➥ Data-Engineer-Associate □ for free download on □ www.pdfvce.com □ □Data-Engineer-Associate Test Dumps Demo
- Reliable Data-Engineer-Associate Braindumps Free □ Best Data-Engineer-Associate Vce ♥ Data-Engineer-Associate Real Exam Answers □ The page for free download of ➠ Data-Engineer-Associate □ on □ www.troytecdumps.com □ will open immediately □Data-Engineer-Associate Real Exam Answers
- Free PDF Data-Engineer-Associate - High Pass-Rate Trustworthy AWS Certified Data Engineer - Associate (DEA-C01) Exam Content □ Download { Data-Engineer-Associate } for free by simply entering [www.pdfvce.com] website □Data-Engineer-Associate Valid Test Test
- Valid Data-Engineer-Associate Exam Dumps □ Data-Engineer-Associate New Braindumps □ New Data-Engineer-Associate Test Simulator □ Search for ☀ Data-Engineer-Associate □☀□ and download it for free on 「 www.testkingpass.com 」 website □Reliable Data-Engineer-Associate Guide Files
- Data-Engineer-Associate Reliable Braindumps Ebook □ Data-Engineer-Associate Reliable Dumps Files □ Valid Data-Engineer-Associate Exam Dumps □ Search for ☀ Data-Engineer-Associate □☀□ and easily obtain a free download on 「 www.pdfvce.com 」 □Reliable Data-Engineer-Associate Braindumps Free
- Data-Engineer-Associate Prep Guide □ Data-Engineer-Associate Reliable Test Braindumps □ Data-Engineer-Associate Real Exam Answers □ Go to website 【 www.prepawayete.com 】 open and search for ☀ Data-Engineer-Associate □☀□ to download for free □Data-Engineer-Associate Reliable Dumps Files
- Data-Engineer-Associate Reliable Braindumps Ebook □ New Data-Engineer-Associate Mock Exam □ Data-Engineer-Associate Test Dumps Demo □ Enter ➥ www.pdfvce.com □ and search for ➡ Data-Engineer-Associate □□□ to download for free □Pdf Data-Engineer-Associate Braindumps
- New Data-Engineer-Associate Test Simulator ✏ Data-Engineer-Associate Valid Test Voucher □ Reliable Data-Engineer-

Associate Guide Files Search for 【 Data-Engineer-Associate 】 and obtain a free download on “ www.examcollectionpass.com ” New Data-Engineer-Associate Mock Exam