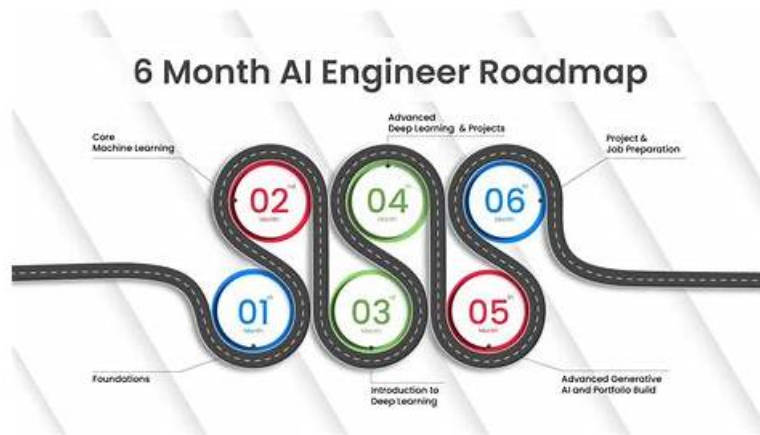


2026 Data-Engineer-Associate—100% Free Well Prep | the Best Data-Engineer-Associate Study Reference



2026 Latest TestPDF Data-Engineer-Associate PDF Dumps and Data-Engineer-Associate Exam Engine Free Share:
<https://drive.google.com/open?id=14Yhs7g1XKwW3-V8KPiM6zOupH8MZ-4EW>

A whole new scope opens up to you and you are immediately hired by reputed firms. Even though the Amazon Data-Engineer-Associate certification boosts your career options, you have to pass the Data-Engineer-Associate Exam. This Amazon Data-Engineer-Associate exam serves to filter out the capable from incapable candidates.

We have created a number of reports and learning functions for evaluating your proficiency for the Data-Engineer-Associate exam dumps. In preparation, you can optimize Amazon Data-Engineer-Associate practice exam time and question type by utilizing our Amazon Data-Engineer-Associate Practice Test software. TestPDF makes it easy to download AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) exam questions immediately after purchase.

>> **Data-Engineer-Associate Well Prep** <<

Advantages Of These Amazon Data-Engineer-Associate Exam Questions Formats

If you find someone around has a nice life go wild, it is because that they may have favored the use of study & work method different from normal people. Data-Engineer-Associate dumps torrent files may be the best method for candidates who are preparing for their IT exam and eager to clear exam as soon as possible. People's success lies in their good use of every change to self-improve. Our Data-Engineer-Associate Dumps Torrent files will be the best resources for your real test. If you choose our products, we will choose efficient & high-passing preparation materials.

Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q94-Q99):

NEW QUESTION # 94

A data engineer notices slow query performance on a highly partitioned table that is in Amazon Athena. The table contains daily data for the previous 5 years, partitioned by date. The data engineer wants to improve query performance and to automate partition management. Which solution will meet these requirements?

- **A. Use partition projection in Athena. Configure the table properties by using a date range from 5 years ago to the present.**
- B. Reduce the number of partitions by changing the partitioning schema from daily to monthly granularity.
- C. Increase the processing capacity of Athena queries by allocating more compute resources.
- D. Use an AWS Lambda function that runs daily. Configure the function to manually create new partitions in AWS Glue for each day's data.

Answer: A

NEW QUESTION # 95

A data engineer must manage the ingestion of real-time streaming data into AWS. The data engineer wants to perform real-time analytics on the incoming streaming data by using time-based aggregations over a window of up to 30 minutes. The data engineer needs a solution that is highly fault tolerant.

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

- A. Use Amazon Managed Service for Apache Flink (previously known as Amazon Kinesis Data Analytics) to analyze the data that might occasionally contain duplicates by using multiple types of aggregations.
- B. Use Amazon Managed Service for Apache Flink (previously known as Amazon Kinesis Data Analytics) to analyze the data by using multiple types of aggregations to perform time-based analytics over a window of up to 30 minutes.
- **C. Use an AWS Lambda function that includes both the business and the analytics logic to perform time-based aggregations over a window of up to 30 minutes for the data in Amazon Kinesis Data Streams.**
- D. Use an AWS Lambda function that includes both the business and the analytics logic to perform aggregations for a tumbling window of up to 30 minutes, based on the event timestamp.

Answer: C

Explanation:

This solution meets the requirements of managing the ingestion of real-time streaming data into AWS and performing real-time analytics on the incoming streaming data with the least operational overhead. Amazon Managed Service for Apache Flink is a fully managed service that allows you to run Apache Flink applications without having to manage any infrastructure or clusters. Apache Flink is a framework for stateful stream processing that supports various types of aggregations, such as tumbling, sliding, and session windows, over streaming data. By using Amazon Managed Service for Apache Flink, you can easily connect to Amazon Kinesis Data Streams as the source and sink of your streaming data, and perform time-based analytics over a window of up to 30 minutes. This solution is also highly fault tolerant, as Amazon Managed Service for Apache Flink automatically scales, monitors, and restarts your Flink applications in case of failures. Reference:

Amazon Managed Service for Apache Flink

Apache Flink

Window Aggregations in Flink

NEW QUESTION # 96

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 Lambda function to scan the sensor data files to validate temperature ranges. Use AWS Glue Data Catalog tables to check timestamp completeness.
- **B. Create an AWS Glue DataBrew project to profile the sensor data. Define completeness rules for timestamps. Set up numeric range validation for temperature values.**
- 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. Use AWS Glue's Data Quality rules and machine learning (ML)-based anomaly detection to identify missing timestamps and to detect temperature anomalies.

Answer: B

Explanation:

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.

NEW QUESTION # 97

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.

```
DELETE FROM materialized_view_name where 1=1
```

B.

```
TRUNCATE materialized_view_name
```

C.

```
VACUUM table_name where load_date<=current_date  
materializedview
```

D.

```
DELETE FROM materialized_view_name where load_date<=current_date
```

- 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 # 98

A financial company wants to use Amazon Athena to run on-demand SQL queries on a petabyte-scale dataset to support a business intelligence (BI) application. An AWS Glue job that runs during non-business hours updates the dataset once every day. The BI application has a standard data refresh frequency of 1 hour to comply with company policies.

A data engineer wants to cost optimize the company's use of Amazon Athena without adding any additional infrastructure costs.

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

- A. Change the format of the files that are in the dataset to Apache Parquet.
- B. Add an Amazon ElastiCache cluster between the BI application and Athena.
- C. Use the query result reuse feature of Amazon Athena for the SQL queries.
- D. Configure an Amazon S3 Lifecycle policy to move data to the S3 Glacier Deep Archive storage class after 1 day

Answer: C

Explanation:

The best solution to cost optimize the company's use of Amazon Athena without adding any additional infrastructure costs is to use the query result reuse feature of Amazon Athena for the SQL queries. This feature allows you to run the same query multiple times

without incurring additional charges, as long as the underlying data has not changed and the query results are still in the query result location in Amazon S3¹. This feature is useful for scenarios where you have a petabyte-scale dataset that is updated infrequently, such as once a day, and you have a BI application that runs the same queries repeatedly, such as every hour. By using the query result reuse feature, you can reduce the amount of data scanned by your queries and save on the cost of running Athena. You can enable or disable this feature at the workgroup level or at the individual query level.

Option A is not the best solution, as configuring an Amazon S3 Lifecycle policy to move data to the S3 Glacier Deep Archive storage class after 1 day would not cost optimize the company's use of Amazon Athena, but rather increase the cost and complexity. Amazon S3 Lifecycle policies are rules that you can define to automatically transition objects between different storage classes based on specified criteria, such as the age of the object². S3 Glacier Deep Archive is the lowest-cost storage class in Amazon S3, designed for long-term data archiving that is accessed once or twice in a year³. While moving data to S3 Glacier Deep Archive can reduce the storage cost, it would also increase the retrieval cost and latency, as it takes up to 12 hours to restore the data from S3 Glacier Deep Archive³. Moreover, Athena does not support querying data that is in S3 Glacier or S3 Glacier Deep Archive storage classes⁴. Therefore, using this option would not meet the requirements of running on-demand SQL queries on the dataset.

Option C is not the best solution, as adding an Amazon ElastiCache cluster between the BI application and Athena would not cost optimize the company's use of Amazon Athena, but rather increase the cost and complexity. Amazon ElastiCache is a service that offers fully managed in-memory data stores, such as Redis and Memcached, that can improve the performance and scalability of web applications by caching frequently accessed data. While using ElastiCache can reduce the latency and load on the BI application, it would not reduce the amount of data scanned by Athena, which is the main factor that determines the cost of running Athena. Moreover, using ElastiCache would introduce additional infrastructure costs and operational overhead, as you would have to provision, manage, and scale the ElastiCache cluster, and integrate it with the BI application and Athena.

Option D is not the best solution, as changing the format of the files that are in the dataset to Apache Parquet would not cost optimize the company's use of Amazon Athena without adding any additional infrastructure costs, but rather increase the complexity. Apache Parquet is a columnar storage format that can improve the performance of analytical queries by reducing the amount of data that needs to be scanned and providing efficient compression and encoding schemes. However, changing the format of the files that are in the dataset to Apache Parquet would require additional processing and transformation steps, such as using AWS Glue or Amazon EMR to convert the files from their original format to Parquet, and storing the converted files in a separate location in Amazon S3. This would increase the complexity and the operational overhead of the data pipeline, and also incur additional costs for using AWS Glue or Amazon EMR. Reference:

Query result reuse

Amazon S3 Lifecycle

S3 Glacier Deep Archive

Storage classes supported by Athena

[What is Amazon ElastiCache?]

[Amazon Athena pricing]

[Columnar Storage Formats]

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

NEW QUESTION # 99

.....

Our company has dedicated ourselves to develop the Data-Engineer-Associate latest practice dumps for all candidates to pass the exam easier, also has made great achievement after more than ten years' development. As the certification has been of great value, a right Data-Engineer-Associate exam guide can be your strong forward momentum to help you pass the exam like a hot knife through butter. On the contrary, it might be time-consuming and tired to prepare for the Data-Engineer-Associate exam without a specialist study material. So it's would be the best decision to choose our Data-Engineer-Associate Study Tool as your learning partner. Our Data-Engineer-Associate study tool also gives numerous candidates a better perspective on the real exam. Having been specializing in the research of Data-Engineer-Associate latest practice dumps, we now process a numerous of customers with our endless efforts, and we believe that our Data-Engineer-Associate exam guide will percolate to your satisfaction.

Data-Engineer-Associate Study Reference: <https://www.testpdf.com/Data-Engineer-Associate-exam-braindumps.html>

The simple and easy-to-understand language of Data-Engineer-Associate guide torrent frees any learner from studying difficulties, The desktop AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) practice exam software helps its valued customer to be well aware of the pattern of the real Data-Engineer-Associate exam, These tests create scenarios that are similar to the actual Data-Engineer-Associate examination, Amazon Data-Engineer-Associate Well Prep Read on to check out the features of these three formats.

Problems with First-Generation Enterprise Portals, The Parts of a Message, The simple and easy-to-understand language of Data-Engineer-Associate Guide Torrent frees any learner from studying difficulties.

Pass Guaranteed Quiz Amazon - Data-Engineer-Associate - AWS Certified Data Engineer - Associate (DEA-C01) Latest Well Prep

[illegible]

P.S. Free & New Data-Engineer-Associate dumps are available on Google Drive shared by TestPDF: <https://drive.google.com/open?id=14Yhs7g1XKwW3-V8KPm6zOupH8MZ-4EW>