

Top Data-Engineer-Associate Questions, Data-Engineer-Associate Valid Exam Guide



2026 Latest TrainingDumps Data-Engineer-Associate PDF Dumps and Data-Engineer-Associate Exam Engine Free Share:
https://drive.google.com/open?id=1_zxcq2mXMxDHgKq5N7kRBxhna5VKt_H

Review the products offered by us by downloading their free demos and compare them with the Data-Engineer-Associate study material offered in online course free and vendors' files. You will find our products the better than our competitors such as exam collection and others. The excellent quality of our Data-Engineer-Associate content, their relevance with the actual exam needs and their interactive and simple format will prove them superior and quite pertinent to your needs and requirements.

The contents of Data-Engineer-Associate learning questions are carefully compiled by the experts according to the content of the Data-Engineer-Associate examination syllabus of the calendar year. They are focused and detailed, allowing your energy to be used in important points of knowledge and to review them efficiently. In addition, Data-Engineer-Associate Guide engine is supplemented by a mock examination system with a time-taking function to allow users to check the gaps in the course of learning.

>> Top Data-Engineer-Associate Questions <<

Pass Guaranteed Quiz 2026 Data-Engineer-Associate: Fantastic Top AWS Certified Data Engineer - Associate (DEA-C01) Questions

You will need to pass the AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) exam to achieve the Amazon Data-Engineer-Associate certification. Due to extremely high competition, passing the Amazon Data-Engineer-Associate exam is not easy; however, possible. You can use TrainingDumps products to pass the Data-Engineer-Associate Exam on the first attempt. The Amazon practice exam gives you confidence and helps you understand the criteria of the testing authority and pass the AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) exam on the first attempt.

Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q78-Q83):

NEW QUESTION # 78

A company maintains an Amazon Redshift provisioned cluster that the company uses for extract, transform, and load (ETL) operations to support critical analysis tasks. A sales team within the company maintains a Redshift cluster that the sales team uses for business intelligence (BI) tasks.

The sales team recently requested access to the data that is in the ETL Redshift cluster so the team can perform weekly summary analysis tasks. The sales team needs to join data from the ETL cluster with data that is in the sales team's BI cluster.

The company needs a solution that will share the ETL cluster data with the sales team without interrupting the critical analysis tasks.

The solution must minimize usage of the computing resources of the ETL cluster.

Which solution will meet these requirements?

- A. Create database views based on the sales team's requirements. Grant the sales team direct access to the ETL cluster.
- B. Unload a copy of the data from the ETL cluster to an Amazon S3 bucket every week. Create an Amazon Redshift Spectrum table based on the content of the ETL cluster.

- C. Set up the sales team BI cluster as a consumer of the ETL cluster by using Redshift data sharing.
- D. Create materialized views based on the sales team's requirements. Grant the sales team direct access to the ETL cluster.

Answer: C

Explanation:

Redshift data sharing is a feature that enables you to share live data across different Redshift clusters without the need to copy or move data. Data sharing provides secure and governed access to data, while preserving the performance and concurrency benefits of Redshift. By setting up the sales team BI cluster as a consumer of the ETL cluster, the company can share the ETL cluster data with the sales team without interrupting the critical analysis tasks. The solution also minimizes the usage of the computing resources of the ETL cluster, as the data sharing does not consume any storage space or compute resources from the producer cluster. The other options are either not feasible or not efficient. Creating materialized views or database views would require the sales team to have direct access to the ETL cluster, which could interfere with the critical analysis tasks. Unloading a copy of the data from the ETL cluster to an Amazon S3 bucket every week would introduce additional latency and cost, as well as create data inconsistency issues.

Reference:

Sharing data across Amazon Redshift clusters

AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide, Chapter 2: Data Store Management, Section 2.2: Amazon Redshift

NEW QUESTION # 79

A company stores details about transactions in an Amazon S3 bucket. The company wants to log all writes to the S3 bucket into another S3 bucket that is in the same AWS Region.

Which solution will meet this requirement with the LEAST operational effort?

- A. Create a trail of data events in AWS CloudTrail. Configure the trail to receive data from the transactions S3 bucket. Specify an empty prefix and write-only events. Specify the logs S3 bucket as the destination bucket.
- B. Configure an S3 Event Notifications rule for all activities on the transactions S3 bucket to invoke an AWS Lambda function. Program the Lambda function to write the events to the logs S3 bucket.
- C. Configure an S3 Event Notifications rule for all activities on the transactions S3 bucket to invoke an AWS Lambda function. Program the Lambda function to write the event to Amazon Kinesis Data Firehose. Configure Kinesis Data Firehose to write the event to the logs S3 bucket.
- D. Create a trail of management events in AWS CloudTrail. Configure the trail to receive data from the transactions S3 bucket. Specify an empty prefix and write-only events. Specify the logs S3 bucket as the destination bucket.

Answer: A

Explanation:

This solution meets the requirement of logging all writes to the S3 bucket into another S3 bucket with the least operational effort. AWS CloudTrail is a service that records the API calls made to AWS services, including Amazon S3. By creating a trail of data events, you can capture the details of the requests that are made to the transactions S3 bucket, such as the requester, the time, the IP address, and the response elements.

By specifying an empty prefix and write-only events, you can filter the data events to only include the ones that write to the bucket. By specifying the logs S3 bucket as the destination bucket, you can store the CloudTrail logs in another S3 bucket that is in the same AWS Region. This solution does not require any additional coding or configuration, and it is more scalable and reliable than using S3 Event Notifications and Lambda functions. References:

Logging Amazon S3 API calls using AWS CloudTrail

Creating a trail for data events

Enabling Amazon S3 server access logging

NEW QUESTION # 80

A company stores its processed data in an S3 bucket. The company has a strict data access policy. The company uses IAM roles to grant teams within the company different levels of access to the S3 bucket.

The company wants to receive notifications when a user violates the data access policy. Each notification must include the username of the user who violated the policy.

Which solution will meet these requirements?

- A. Use Amazon CloudWatch metrics to gather object-level metrics. Set up CloudWatch alarms.
- B. Use AWS CloudTrail to track object-level events for the S3 bucket. Forward events to Amazon CloudWatch to set up CloudWatch alarms.

- C. Use AWS Config rules to detect violations of the data access policy. Set up compliance alarms.
- D. Use Amazon S3 server access logs to monitor access to the bucket. Forward the access logs to an Amazon CloudWatch log group. Use metric filters on the log group to set up CloudWatch alarms.

Answer: B

Explanation:

The requirement is to detect violations of data access policies and receive notifications with the username of the violator. AWS CloudTrail can provide object-level tracking for S3 to capture detailed API actions on specific S3 objects, including the user who performed the action.

AWS CloudTrail:

CloudTrail can monitor API calls made to an S3 bucket, including object-level API actions such as GetObject, PutObject, and DeleteObject. This will help detect access violations based on the API calls made by different users.

CloudTrail logs include details such as the user identity, which is essential for meeting the requirement of including the username in notifications.

The CloudTrail logs can be forwarded to Amazon CloudWatch to trigger alarms based on certain access patterns (e.g., violations of specific policies).

Reference:

Amazon CloudWatch:

By forwarding CloudTrail logs to CloudWatch, you can set up alarms that are triggered when a specific condition is met, such as unauthorized access or policy violations. The alarm can include detailed information from the CloudTrail log, including the username.

Alternatives Considered:

A (AWS Config rules): While AWS Config can track resource configurations and compliance, it does not provide real-time, detailed tracking of object-level events like CloudTrail does.

B (CloudWatch metrics): CloudWatch does not gather object-level metrics for S3 directly. For this use case, CloudTrail provides better granularity.

D (S3 server access logs): S3 server access logs can monitor access, but they do not provide the real-time monitoring and alerting features that CloudTrail with CloudWatch alarms offer. They also do not include API-level granularity like CloudTrail.

AWS CloudTrail Integration with S3

Amazon CloudWatch Alarms

NEW QUESTION # 81

A company uses an Amazon QuickSight dashboard to monitor usage of one of the company's applications. The company uses AWS Glue jobs to process data for the dashboard. The company stores the data in a single Amazon S3 bucket. The company adds new data every day.

A data engineer discovers that dashboard queries are becoming slower over time. The data engineer determines that the root cause of the slowing queries is long-running AWS Glue jobs.

Which actions should the data engineer take to improve the performance of the AWS Glue jobs? (Choose two.)

- **A. Increase the AWS Glue instance size by scaling up the worker type.**
- B. Adjust AWS Glue job scheduling frequency so the jobs run half as many times each day.
- C. Convert the AWS Glue schema to the DynamicFrame schema class.
- D. Modify the IAM role that grants access to AWS glue to grant access to all S3 features.
- **E. Partition the data that is in the S3 bucket. Organize the data by year, month, and day.**

Answer: A,E

Explanation:

Partitioning the data in the S3 bucket can improve the performance of AWS Glue jobs by reducing the amount of data that needs to be scanned and processed. By organizing the data by year, month, and day, the AWS Glue job can use partition pruning to filter out irrelevant data and only read the data that matches the query criteria. This can speed up the data processing and reduce the cost of running the AWS Glue job. Increasing the AWS Glue instance size by scaling up the worker type can also improve the performance of AWS Glue jobs by providing more memory and CPU resources for the Spark execution engine. This can help the AWS Glue job handle larger data sets and complex transformations more efficiently. The other options are either incorrect or irrelevant, as they do not affect the performance of the AWS Glue jobs. Converting the AWS Glue schema to the DynamicFrame schema class does not improve the performance, but rather provides additional functionality and flexibility for data manipulation. Adjusting the AWS Glue job scheduling frequency does not improve the performance, but rather reduces the frequency of data updates. Modifying the IAM role that grants access to AWS Glue does not improve the performance, but rather affects the security and permissions of the AWS Glue service. Reference:

Optimising Glue Scripts for Efficient Data Processing: Part 1 (Section: Partitioning Data in S3) Best practices to optimize cost and

NEW QUESTION # 82

A company has used an Amazon Redshift table that is named Orders for 6 months. The company performs weekly updates and deletes on the table. The table has an interleaved sort key on a column that contains AWS Regions.

The company wants to reclaim disk space so that the company will not run out of storage space. The company also wants to analyze the sort key column.

Which Amazon Redshift command will meet these requirements?

- A. VACUUM DELETE ONLY Orders
- **B. VACUUM REINDEX Orders**
- C. VACUUM FULL Orders
- D. VACUUM SORT ONLY Orders

Answer: B

Explanation:

Amazon Redshift is a fully managed, petabyte-scale data warehouse service that enables fast and cost-effective analysis of large volumes of data. Amazon Redshift uses columnar storage, compression, and zone maps to optimize the storage and performance of data. However, over time, as data is inserted, updated, or deleted, the physical storage of data can become fragmented, resulting in wasted disk space and degraded query performance. To address this issue, Amazon Redshift provides the VACUUM command, which reclaims disk space and resorts rows in either a specified table or all tables in the current schema.

The VACUUM command has four options: FULL, DELETE ONLY, SORT ONLY, and REINDEX. The option that best meets the requirements of the question is VACUUM REINDEX, which re-sorts the rows in a table that has an interleaved sort key and rewrites the table to a new location on disk. An interleaved sort key is a type of sort key that gives equal weight to each column in the sort key, and stores the rows in a way that optimizes the performance of queries that filter by multiple columns in the sort key. However, as data is added or changed, the interleaved sort order can become skewed, resulting in suboptimal query performance. The VACUUM REINDEX option restores the optimal interleaved sort order and reclaims disk space by removing deleted rows. This option also analyzes the sort key column and updates the table statistics, which are used by the query optimizer to generate the most efficient query execution plan.

The other options are not optimal for the following reasons:

- A. VACUUM FULL Orders. This option reclaims disk space by removing deleted rows and resorts the entire table. However, this option is not suitable for tables that have an interleaved sort key, as it does not restore the optimal interleaved sort order. Moreover, this option is the most resource-intensive and time-consuming, as it rewrites the entire table to a new location on disk.
- B. VACUUM DELETE ONLY Orders. This option reclaims disk space by removing deleted rows, but does not resort the table. This option is not suitable for tables that have any sort key, as it does not improve the query performance by restoring the sort order. Moreover, this option does not analyze the sort key column and update the table statistics.
- D. VACUUM SORT ONLY Orders. This option resorts the entire table, but does not reclaim disk space by removing deleted rows. This option is not suitable for tables that have an interleaved sort key, as it does not restore the optimal interleaved sort order. Moreover, this option does not analyze the sort key column and update the table statistics.

1: Amazon Redshift VACUUM

2: Amazon Redshift Interleaved Sorting

3: Amazon Redshift ANALYZE

NEW QUESTION # 83

.....

As a brand in the field, our Data-Engineer-Associate exam questions are famous for their different and effective advantages. Our professional experts have developed our Data-Engineer-Associate study materials to the best. So if you buy them, you will find that our Data-Engineer-Associate learning braindumps are simply unmatched in their utility and perfection. Our huge clientele is immensely satisfied with our product and the excellent passing rate of our Data-Engineer-Associate simulating exam is the best evidence on it.

Data-Engineer-Associate Valid Exam Guide: https://www.trainingdumps.com/Data-Engineer-Associate_exam-valid-dumps.html

Amazon offers upgraded Amazon Data-Engineer-Associate practice exam software to give an environment like the real exam. We are so confident in our Data-Engineer-Associate study materials because they have their own uniqueness. Our Data-Engineer-Associate test pdf is compiled by the senior experts elaborately and we update them frequently to follow the trend of the times,

Amazon Top Data-Engineer-Associate Questions In the other words, passing the exam once will no longer be a dream.

Retouching Using Define Pattern, The third section green text) describes the melody, Amazon offers upgraded Amazon Data-Engineer-Associate Practice Exam software to give an environment like the real exam.

100% Pass Amazon - Efficient Data-Engineer-Associate - Top AWS Certified Data Engineer - Associate (DEA-C01) Questions

We are so confident in our Data-Engineer-Associate study materials because they have their own uniqueness, Our Data-Engineer-Associate test pdf is compiled by the senior experts elaborately and we update them frequently to follow the trend of the times.

In the other words, passing the exam once will no longer be a dream. We are confident that our Data-Engineer-Associate exam study material is the first-class in our market and it's also a good choice for you.

- [illegible]

DOWNLOAD the newest TrainingDumps Data-Engineer-Associate PDF dumps from Cloud Storage for free:
https://drive.google.com/open?id=1_zxcq2mXMxDHgKq5N7kRBxhna5VKt_H