

Valid Amazon Dump Data-Engineer-Associate Torrent Are Leading Materials & Free Download Data-Engineer-Associate Latest Braindumps Book



P.S. Free 2026 Amazon Data-Engineer-Associate dumps are available on Google Drive shared by Itexamguide:
<https://drive.google.com/open?id=1M74a9QMwnEos3a6mn8jUKMn7oWwu6W-Q>

Our Data-Engineer-Associate simulating exam is perfect for they come a long way on their quality. On one hand, we have engaged in this career for over ten years and have become the leader in this market. On the other hand, we never stop developing our Data-Engineer-Associate study guide. And our Data-Engineer-Associate Training Materials have the function to remember and correct your errors. If you commit any errors, Our Data-Engineer-Associate learning questions can correct your errors with accuracy rate more than 98 percent.

You have seen Itexamguide's Amazon Data-Engineer-Associate Exam Training materials, it is time to make a choice. You can choose other products, but you have to know that Itexamguide can bring you infinite interests. Only Itexamguide can guarantee you 100% success. Itexamguide allows you to have a bright future. And allows you to work in the field of information technology with high efficiency.

>> Dump Data-Engineer-Associate Torrent <<

By Achieving the Amazon Data-Engineer-Associate You will Get the Job

To go with the changing neighborhood, we need to improve our efficiency of solving problems, which reflects in many aspect as well as dealing with exams. Our Data-Engineer-Associate practice materials can help you realize it. To those time-sensitive exam candidates, our high-efficient Data-Engineer-Associate actual tests comprised of important news will be best help. Only by practicing them on a regular base, you will see clear progress happened on you. Besides, rather than waiting for the gain of our Data-Engineer-Associate practice materials, you can download them immediately after paying for it, so just begin your journey toward success now.

Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q99-Q104):

NEW QUESTION # 99

A technology company currently uses Amazon Kinesis Data Streams to collect log data in real time. The company wants to use Amazon Redshift for downstream real-time queries and to enrich the log data.

Which solution will ingest data into Amazon Redshift with the LEAST operational overhead?

- A. Set up an Amazon Data Firehose delivery stream to send data to Amazon S3. Configure a Redshift provisioned cluster to load data every minute.
- B. Set up an Amazon Data Firehose delivery stream to send data to a Redshift provisioned cluster table.
- C. Configure Amazon Managed Service for Apache Flink (previously known as Amazon Kinesis Data Analytics) to send data directly to a Redshift provisioned cluster table.
- **D. Use Amazon Redshift streaming ingestion from Kinesis Data Streams and to present data as a materialized view.**

Answer: D

Explanation:

The most efficient and low-operational-overhead solution for ingesting data into Amazon Redshift from Amazon Kinesis Data Streams is to use Amazon Redshift streaming ingestion. This feature allows Redshift to directly ingest streaming data from Kinesis Data Streams and process it in real-time.

Amazon Redshift Streaming Ingestion:

Redshift supports native streaming ingestion from Kinesis Data Streams, allowing real-time data to be queried using materialized views.

This solution reduces operational complexity because you don't need intermediary services like Amazon Kinesis Data Firehose or S3 for batch loading.

Reference:

Alternatives Considered:

A (Data Firehose to Redshift): This option is more suitable for batch processing but incurs additional operational overhead with the Firehose setup.

B (Firehose to S3): This involves an intermediate step, which adds complexity and delays the real-time requirement.

C (Managed Service for Apache Flink): This would work but introduces unnecessary complexity compared to Redshift's native streaming ingestion.

Amazon Redshift Streaming Ingestion from Kinesis

Materialized Views in Redshift

NEW QUESTION # 100

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. Partition the data that is in the S3 bucket. Organize the data by year, month, and day.**
- D. Convert the AWS Glue schema to the DynamicFrame schema class.
- E. Modify the IAM role that grants access to AWS glue to grant access to all S3 features.

Answer: A,C

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. References:
Optimising Glue Scripts for Efficient Data Processing: Part 1 (Section: Partitioning Data in S3) Best practices to optimize cost and performance for AWS Glue streaming ETL jobs (Section: Development tools)
Monitoring with AWS Glue job run insights (Section: Requirements)
AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide (Chapter 5, page 133)

NEW QUESTION # 101

A data engineer needs to use an Amazon QuickSight dashboard that is based on Amazon Athena queries on data that is stored in an Amazon S3 bucket. When the data engineer connects to the QuickSight dashboard, the data engineer receives an error message that indicates insufficient permissions.

Which factors could cause the permissions-related errors? (Choose two.)

- A. There is no connection between QuickSight and Athena.
- **B. QuickSight does not have access to the S3 bucket.**
- **C. QuickSight does not have access to decrypt S3 data.**
- D. The Athena tables are not cataloged.
- E. There is no IAM role assigned to QuickSight.

Answer: B,C

Explanation:

QuickSight does not have access to the S3 bucket and QuickSight does not have access to decrypt S3 data are two possible factors that could cause the permissions-related errors. Amazon QuickSight is a business intelligence service that allows you to create and share interactive dashboards based on various data sources, including Amazon Athena. Amazon Athena is a serverless query service that allows you to analyze data stored in Amazon S3 using standard SQL. To use an Amazon QuickSight dashboard that is based on Amazon Athena queries on data that is stored in an Amazon S3 bucket, you need to grant QuickSight access to both Athena and S3, as well as any encryption keys that are used to encrypt the S3 data. If QuickSight does not have access to the S3 bucket or the encryption keys, it will not be able to read the data from Athena and display it on the dashboard, resulting in an error message that indicates insufficient permissions.

The other options are not factors that could cause the permissions-related errors. Option A, there is no connection between QuickSight and Athena, is not a factor, as QuickSight supports Athena as a native data source, and you can easily create a connection between them using the QuickSight console or the API. Option B, the Athena tables are not cataloged, is not a factor, as QuickSight can automatically discover the Athena tables that are cataloged in the AWS Glue Data Catalog, and you can also manually specify the Athena tables that are not cataloged. Option E, there is no IAM role assigned to QuickSight, is not a factor, as QuickSight requires an IAM role to access any AWS data sources, including Athena and S3, and you can create and assign an IAM role to QuickSight using the QuickSight console or the API. Reference:

Using Amazon Athena as a Data Source

Granting Amazon QuickSight Access to AWS Resources

Encrypting Data at Rest in Amazon S3

NEW QUESTION # 102

A company currently uses a provisioned Amazon EMR cluster that includes general purpose Amazon EC2 instances. The EMR cluster uses EMR managed scaling between one to five task nodes for the company's long-running Apache Spark extract, transform, and load (ETL) job. The company runs the ETL job every day.

When the company runs the ETL job, the EMR cluster quickly scales up to five nodes. The EMR cluster often reaches maximum CPU usage, but the memory usage remains under 30%.

The company wants to modify the EMR cluster configuration to reduce the EMR costs to run the daily ETL job.

Which solution will meet these requirements MOST cost-effectively?

- A. Increase the maximum number of task nodes for EMR managed scaling to 10.
- **B. Switch the task node type from general purpose EC2 instances to compute optimized EC2 instances.**
- C. Reduce the scaling cooldown period for the provisioned EMR cluster.
- D. Change the task node type from general purpose EC2 instances to memory optimized EC2 instances.

Answer: B

Explanation:

The company's Apache Spark ETL job on Amazon EMR uses high CPU but low memory, meaning that compute-optimized EC2

instances would be the most cost-effective choice. These instances are designed for high-performance compute applications, where CPU usage is high, but memory needs are minimal, which is exactly the case here.

Compute Optimized Instances:

Compute-optimized instances, such as the C5 series, provide a higher ratio of CPU to memory, which is more suitable for jobs with high CPU usage and relatively low memory consumption.

Switching from general-purpose EC2 instances to compute-optimized instances can reduce costs while improving performance, as these instances are optimized for workloads like Spark jobs that perform a lot of computation.

Reference:

Managed Scaling: The EMR cluster's scaling is currently managed between 1 and 5 nodes, so changing the instance type will leverage the current scaling strategy but optimize it for the workload.

Alternatives Considered:

A (Increase task nodes to 10): Increasing the number of task nodes would increase costs without necessarily improving performance. Since memory usage is low, the bottleneck is more likely the CPU, which compute-optimized instances can handle better.

B (Memory optimized instances): Memory-optimized instances are not suitable since the current job is CPU-bound, and memory usage remains low (under 30%).

D (Reduce scaling cooldown): This could marginally improve scaling speed but does not address the need for cost optimization and improved CPU performance.

Amazon EMR Cluster Optimization

Compute Optimized EC2 Instances

NEW QUESTION # 103

A data engineer is using Amazon Athena to analyze sales data that is in Amazon S3. The data engineer writes a query to retrieve sales amounts for 2023 for several products from a table named `sales_data`. However, the query does not return results for all of the products that are in the `sales_data` table. The data engineer needs to troubleshoot the query to resolve the issue.

The data engineer's original query is as follows:

```
SELECT product_name, sum(sales_amount)
```

```
FROM sales_data
```

```
WHERE year = 2023
```

```
GROUP BY product_name
```

How should the data engineer modify the Athena query to meet these requirements?

- A. Replace `sum(sales_amount)` with `count(*)` for the aggregation.
- B. Remove the `GROUP BY` clause
- C. Change `WHERE year = 2023` to `WHERE extract(year FROM sales_data) = 2023`.
- D. Add `HAVING sum(sales_amount) > 0` after the `GROUP BY` clause.

Answer: C

Explanation:

The original query does not return results for all of the products because the `year` column in the `sales_data` table is not an integer, but a timestamp. Therefore, the `WHERE` clause does not filter the data correctly, and only returns the products that have a null value for the `year` column. To fix this, the data engineer should use the `extract` function to extract the year from the timestamp and compare it with 2023. This way, the query will return the correct results for all of the products in the `sales_data` table. The other options are either incorrect or irrelevant, as they do not address the root cause of the issue. Replacing `sum` with `count` does not change the filtering condition, adding `HAVING` clause does not affect the grouping logic, and removing the `GROUP BY` clause does not solve the problem of missing products. References:

Troubleshooting JSON queries - Amazon Athena (Section: JSON related errors) When I query a table in Amazon Athena, the `TIMESTAMP` result is empty (Section: Resolution) AWS Certified Data Engineer - Associate DEA-C01 Complete Study Guide (Chapter 7, page 197)

NEW QUESTION # 104

.....

With our Data-Engineer-Associate practice exam, you only need to spend 20 to 30 hours in preparation since there are all essence contents in our Data-Engineer-Associate study materials. And there is no exaggeration that with our Data-Engineer-Associate training guide, you can get 100% pass guarantee. What's more, if you need any after service help on our Data-Engineer-Associate Exam Dumps, our after service staffs will always here to offer the most thoughtful service for you.

Now you can pass Amazon Data-Engineer-Associate exam without going through any hassle, Amazon Dump Data-Engineer-Associate Torrent Then just click to buy and pay for the certain money, But keep in mind that the Data-Engineer-Associate certification test is a worthwhile and challenging certificate, Download Free Data-Engineer-Associate Demo (Try before Buy), There is plenty of skilled and motivated staff to help you obtain the Data-Engineer-Associate exam certificate that you are looking forward.

Amazon Dump Data-Engineer-Associate Torrent & Itexamguide - Leader in Qualification Exams & Data-Engineer-Associate: AWS Certified Data Engineer - Associate (DEA-C01)

There is plenty of skilled and motivated staff to help you obtain the Data-Engineer-Associate exam certificate that you are looking forward.

- [illegible]

myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, Disposable vapes

2026 Latest Itexamguide Data-Engineer-Associate PDF Dumps and Data-Engineer-Associate Exam Engine Free Share:
<https://drive.google.com/open?id=1M74a9QMwnEos3a6mn8jUKMn7oWwu6W-Q>