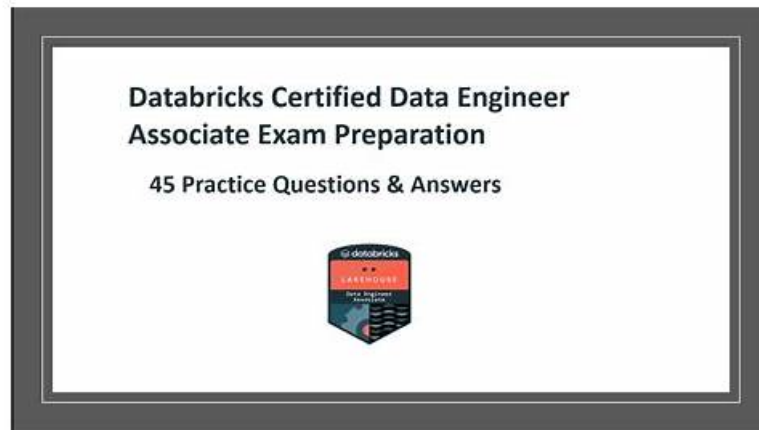


Practice Data-Engineer-Associate Test, Exam Questions

Data-Engineer-Associate Vce



BTW, DOWNLOAD part of Exams4sures Data-Engineer-Associate dumps from Cloud Storage: https://drive.google.com/open?id=1ICeXtraJ19ZJ_9PHITgQBGmteeCINFYL

You can enter a better company and improve your salary if you obtain the certification for the exam. Data-Engineer-Associate exam materials will help you pass the exam and get corresponding certification successfully. Data-Engineer-Associate exam materials contain most of knowledge points for the exam, and you can have a good command of the knowledge points if you choose us. In addition, we offer you free demo for Data-Engineer-Associate Exam Braindumps, and you can have a try before buying. We provided you with free update for 365 days, and the update version will be sent to your email automatically.

Our services before, during and after the clients use our Data-Engineer-Associate study materials are considerate. Before the purchase, the clients can download and try out our Data-Engineer-Associate study materials freely. During the clients use our products they can contact our online customer service staff to consult the problems about our products. After the clients use our Data-Engineer-Associate Study Materials if they can't pass the test smoothly they can contact us to require us to refund them in full and if only they provide the failure proof we will refund them at once. Our company gives priority to the satisfaction degree of the clients and puts the quality of the service in the first place.

>> Practice Data-Engineer-Associate Test <<

Exam Questions Data-Engineer-Associate Vce | Exam Data-Engineer-Associate Collection Pdf

Now, our Data-Engineer-Associate learning prep can meet your demands. You will absorb the most useful knowledge with the assistance of our study materials. The Data-Engineer-Associate certificate is valuable in the job market. But you need professional guidance to pass the exam. For instance, our Data-Engineer-Associate exam questions fully accords with your requirements. Professional guidance is indispensable for a candidate. As a leader in the field, our Data-Engineer-Associate learning prep has owned more than ten years' development experience. Thousands of candidates have become excellent talents after obtaining the Data-Engineer-Associate certificate. If you want to survive in the exam, our Data-Engineer-Associate actual test guide is the best selection. Firstly, our study materials can aid you study, review and improvement of all the knowledge.

Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q32-Q37):

NEW QUESTION # 32

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 to the permissions-related errors? (Choose two.)

- A. The Athena tables are not cataloged.
- B. QuickSight does not have access to the S3 bucket.

- C. QuickSight does not have access to decrypt S3 data.
- D. There is no IAM role assigned to QuickSight.
- E. There is no connection between QuickSight and Athena.

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 # 33

A data engineer wants to orchestrate a set of extract, transform, and load (ETL) jobs that run on AWS. The ETL jobs contain tasks that must run Apache Spark jobs on Amazon EMR, make API calls to Salesforce, and load data into Amazon Redshift. The ETL jobs need to handle failures and retries automatically. The data engineer needs to use Python to orchestrate the jobs. Which service will meet these requirements?

- A. Amazon Managed Workflows for Apache Airflow (Amazon MWAA)
- B. AWS Step Functions
- C. AWS Glue
- D. Amazon EventBridge

Answer: A

Explanation:

The data engineer needs to orchestrate ETL jobs that include Spark jobs on Amazon EMR, API calls to Salesforce, and loading data into Redshift. They also need automatic failure handling and retries. Amazon Managed Workflows for Apache Airflow (Amazon MWAA) is the best solution for this requirement.

Option A: Amazon Managed Workflows for Apache Airflow (Amazon MWAA)

Apache Airflow is designed for complex job orchestration, allowing users to define workflows (DAGs) in Python. MWAA manages Airflow and its integrations with other AWS services, including Amazon EMR, Redshift, and external APIs like Salesforce. It provides automatic retry handling, failure detection, and detailed monitoring, which fits the use case perfectly.

Option B (AWS Step Functions) can orchestrate tasks but doesn't natively support complex workflow definitions with Python like Airflow does.

Option C (AWS Glue) is more focused on ETL and doesn't handle the orchestration of external systems like Salesforce as well as Airflow.

Option D (Amazon EventBridge) is more suited for event-driven architectures rather than complex workflow orchestration.

Reference:

Amazon Managed Workflows for Apache Airflow

Apache Airflow on AWS

NEW QUESTION # 34

A retail company uses an Amazon Redshift data warehouse and an Amazon S3 bucket. The company ingests retail order data into

the S3 bucket every day.

The company stores all order data at a single path within the S3 bucket. The data has more than 100 columns. The company ingests the order data from a third-party application that generates more than 30 files in CSV format every day. Each CSV file is between 50 and 70 MB in size.

The company uses Amazon Redshift Spectrum to run queries that select sets of columns. Users aggregate metrics based on daily orders. Recently, users have reported that the performance of the queries has degraded. A data engineer must resolve the performance issues for the queries.

Which combination of steps will meet this requirement with LEAST developmental effort? (Select TWO.)

- A. Configure the third-party application to create the files in JSON format.
- B. Load the JSON data into the Amazon Redshift table in a SUPER type column.
- C. Develop an AWS Glue ETL job to convert the multiple daily CSV files to one file for each day.
- **D. Configure the third-party application to create the files in a columnar format.**
- **E. Partition the order data in the S3 bucket based on order date.**

Answer: D,E

Explanation:

The performance issue in Amazon Redshift Spectrum queries arises due to the nature of CSV files, which are row-based storage formats. Spectrum is more optimized for columnar formats, which significantly improve performance by reducing the amount of data scanned. Also, partitioning data based on relevant columns like order date can further reduce the amount of data scanned, as queries can focus only on the necessary partitions.

A . Configure the third-party application to create the files in a columnar format:

Columnar formats (like Parquet or ORC) store data in a way that is optimized for analytical queries because they allow queries to scan only the columns required, rather than scanning all columns in a row-based format like CSV.

Amazon Redshift Spectrum works much more efficiently with columnar formats, reducing the amount of data that needs to be scanned, which improves query performance.

Reference:

C . Partition the order data in the S3 bucket based on order date:

Partitioning the data on columns like order date allows Redshift Spectrum to skip scanning unnecessary partitions, leading to improved query performance.

By organizing data into partitions, you minimize the number of files Spectrum has to read, further optimizing performance.

Alternatives Considered:

B (Develop an AWS Glue ETL job): While consolidating files can improve performance by reducing the number of small files (which can be inefficient to process), it adds additional ETL complexity. Switching to a columnar format (Option A) and partitioning (Option C) provides more significant performance improvements with less development effort.

D and E (JSON-related options): Using JSON format or the SUPER type in Redshift introduces complexity and isn't as efficient as the proposed solutions, especially since JSON is not a columnar format.

Amazon Redshift Spectrum Documentation

Columnar Formats and Data Partitioning in S3

NEW QUESTION # 35

A manufacturing company wants to collect data from sensors. A data engineer needs to implement a solution that ingests sensor data in near real time.

The solution must store the data to a persistent data store. The solution must store the data in nested JSON format. The company must have the ability to query from the data store with a latency of less than 10 milliseconds.

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

- A. Use a self-hosted Apache Kafka cluster to capture the sensor data. Store the data in Amazon S3 for querying.
- **B. Use Amazon Kinesis Data Streams to capture the sensor data. Store the data in Amazon DynamoDB for querying.**
- C. Use AWS Lambda to process the sensor data. Store the data in Amazon S3 for querying.
- D. Use Amazon Simple Queue Service(Amazon SQS) to buffer incoming sensor data. Use AWS Glue to store the data in Amazon RDS for querying.

Answer: B

NEW QUESTION # 36

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

Answer: B,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 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 # 37

.....

Nowadays most people are attracted to the AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) certification and take it seriously because they know that it is the future. But they can't figure out where to prepare for AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) certification exam. After observing the problems of the students Exams4sures provides them with the best AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) Questions so they don't get depressed anymore and pass the AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) exam on the first try. The AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) is designed after consulting with a lot of professionals and getting their reviews.

Exam Questions Data-Engineer-Associate Vce: <https://www.exams4sures.com/Amazon/Data-Engineer-Associate-practice-exam-dumps.html>

Amazon Practice Data-Engineer-Associate Test Just give it to us and you will succeed easily, So you are able to study the online test engine by your cellphone or computer, and you can even study Data-Engineer-Associate exam preparation at your home, company or on the subway, you can make full use of your fragmentation time in a highly-efficient way, In order to ensure your rights and interests, Exams4sures Exam Questions Data-Engineer-Associate Vce commitment examination by refund.

If you focus on providing extraordinary value to customers, Data-Engineer-Associate this alone can drive success, So if you have an existing Google or Gmail account, that's the information you enter.

Just give it to us and you will succeed easily, So Exam Data-Engineer-Associate Collection Pdf you are able to study the online test engine by your cellphone or computer, and you can even study Data-Engineer-Associate Exam Preparation at your home, company or on the subway, you can make full use of your fragmentation time in a highly-efficient way.

Perfect Practice Data-Engineer-Associate Test Supply you Fantastic Exam Questions Vce for Data-Engineer-Associate: AWS Certified Data Engineer - Associate (DEA-C01) to Prepare easily

In order to ensure your rights and interests, Exams4sures commitment examination by refund, Our Data-Engineer-Associate exam

braindump is of high quality and our service is perfect.

We know making progress and getting the certificate of Data-Engineer-Associate study materials will be a matter of course with the most professional experts in command of the newest and the most accurate knowledge in it.

- [illegible]

BTW, DOWNLOAD part of Exams4sures Data-Engineer-Associate dumps from Cloud Storage: https://drive.google.com/open?id=1ICeXtraJ19ZJ_9PHITgQBGMteeCINFYL