Data-Engineer-Associate test vce practice & Data-Engineer-Associate exam training files & Data-Engineer-Associate updated prep exam

Today, in an era of fierce competition, how can we occupy a place in a market where talent is saturated? The answer is a certificate. What the certificate main? All kinds of the test Data-Engineer-Associate certification, prove you through all kinds of qualification certificate, it is not hard to find, more and more people are willing to invest time and effort on the Data-Engineer-Associate Exam Guide, because get the test Data-Engineer-Associate certification is not an easy thing, so, a lot of people are looking for an efficient learning method. Our Data-Engineer-Associate exam questions are the right tool for you to pass the Data-Engineer-Associate exam.

After your purchase of Data-Engineer-Associate learning engine, our system will send a link to your email in 5 to 10 minutes. You can contact our staff anytime and anywhere during the learning process. The staff of Data-Engineer-Associate study materials is online 24 hours a day, seven days a week. Our staff is really serious and responsible. We just want to provide you with the best service. I hope you enjoy using Data-Engineer-Associate Exam Materials.

>> **Reliable Data-Engineer-Associate Real Exam** <<

Valid Data-Engineer-Associate Test Pass4sure - Data-Engineer-Associate Relevant Answers

We constantly improve and update our Data-Engineer-Associate study materials and infuse new blood into them according to the development needs of the times and the change of the trend in the industry. We try our best to teach the learners all of the related knowledge about the test Data-Engineer-Associate Certification in the most simple, efficient and intuitive way. We pay our experts high remuneration to let them play their biggest roles in producing our Data-Engineer-Associate study materials.

Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q152-Q157):

NEW QUESTION # 152

A company's data engineer needs to optimize the performance of table SQL queries. The company stores data in an Amazon Redshift cluster. The data engineer cannot increase the size of the cluster because of budget constraints.

The company stores the data in multiple tables and loads the data by using the EVEN distribution style. Some tables are hundreds of gigabytes in size. Other tables are less than 10 MB in size.

Which solution will meet these requirements?

- A. Use the ALL distribution style for rarely updated small tables. Specify primary and foreign keys for all tables.
- B. Keep using the EVEN distribution style for all tables. Specify primary and foreign keys for all tables.
- C. Specify a combination of distribution, sort, and partition keys for all tables.
- D. Use the ALL distribution style for large tables. Specify primary and foreign keys for all tables.

Answer: A

Explanation:
This solution meets the requirements of optimizing the performance of table SQL queries without increasing the size of the cluster. By using the ALL distribution style for rarely updated small tables, you can ensure that the entire table is copied to every node in the cluster, which eliminates the need for data redistribution during joins. This can improve query performance significantly, especially for frequently joined dimension tables.

However, using the ALL distribution style also increases the storage space and the load time, so it is only suitable for small tables that are not updated frequently orextensively. By specifying primary and foreign keys for all tables, you can help the query optimizer to generate better query plans and avoid unnecessary scans or joins. You can also use the AUTO distribution style to let Amazon Redshift choose the optimal distribution style based on the table size and the query patterns. References:
Choose the best distribution style
Distribution styles
Working with data distribution styles

NEW QUESTION # 153

A company needs to load customer data that comes from a third party into an Amazon Redshift data warehouse. The company stores order data and product data in the same data warehouse. The company wants to use the combined dataset to identify potential new customers.

A data engineer notices that one of the fields in the source data includes values that are in JSON format.

How should the data engineer load the JSON data into the data warehouse with the LEAST effort?

- A. Use an AWS Lambda function to flatten the JSON data. Store the data in Amazon S3.
- B. Use AWS Glue to flatten the JSON data and ingest it into the Amazon Redshift table.
- C. Use the SUPER data type to store the data in the Amazon Redshift table.
- D. Use Amazon S3 to store the JSON data. Use Amazon Athena to query the data.

Answer: C

Explanation:
In Amazon Redshift, the SUPER data type is designed specifically to handle semi-structured data like JSON, Parquet, ORC, and others. By using the SUPER data type, Redshift can ingest and query JSON data without requiring complex data flattening processes, thus reducing the amount of preprocessing required before loading the data. The SUPER data type also works seamlessly with Redshift Spectrum, enabling complex queries that can combine both structured and semi-structured datasets, which aligns with the company's need to use combined datasets to identify potential new customers.

Using the SUPER data type also allows for automatic parsing and query processing of nested data structures through Amazon Redshift's PARTITION BY and JSONPATH expressions, which makes this option the most efficient approach with the least effort involved. This reduces the overhead associated with using tools like AWS Glue or Lambda for data transformation.
Amazon Redshift Documentation - SUPER Data Type
AWS Certified Data Engineer - Associate Training: Building Batch Data Analytics Solutions on AWS AWS Certified Data Engineer
- Associate Study Guide By directly leveraging the capabilities of Redshift with the SUPER data type, the data engineer ensures streamlined JSON ingestion with minimal effort while maintaining query efficiency.

NEW QUESTION # 154

A data engineer needs to build an enterprise data catalog based on the company's Amazon S3 buckets and Amazon RDS databases. The data catalog must include storage format metadata for the data in the catalog.

Which solution will meet these requirements with the LEAST effort?

- A. Use Amazon Macie to build a data catalog and to identify sensitive data elements. Collect the data format information from Macie.
- B. Use an AWS Glue crawler to build a data catalog. Use AWS Glue crawler classifiers to recognize the format of data and store the format in the catalog.
- C. Use an AWS Glue crawler to scan the S3 buckets and RDS databases and build a data catalog. Use data stewards to inspect the data and update the data catalog with the data format.
- D. Use scripts to scan data elements and to assign data classifications based on the format of the data.

Answer: B

Explanation:
To build an enterprise data catalog with metadata for storage formats, the easiest and most efficient solution is using an AWS Glue crawler. The Glue crawler can scan Amazon S3 buckets and Amazon RDS databases to automatically create a data catalog that includes metadata such as the schema and storage format (e.g., CSV, Parquet, etc.). By using AWS Glue crawler classifiers, you can configure the crawler to recognize the format of the data and store this information directly in the catalog.
Option B: Use an AWS Glue crawler to build a data catalog. Use AWS Glue crawler classifiers to recognize the format of data and store the format in the catalog.
This option meets the requirements with the least effort because Glue crawlers automate the discovery and cataloging of data from multiple sources, including S3 and RDS, while recognizing various file formats via classifiers.
Other options (A, C, D) involve additional manual steps, like having data stewards inspect the data, or using services like Amazon Macie that focus more on sensitive data detection rather than format cataloging.
Reference:
AWS Glue Crawler Documentation
AWS Glue Classifiers

NEW QUESTION # 155
A company uploads .csv files to an Amazon S3 bucket. The company's data platform team has set up an AWS Glue crawler to perform data discovery and to create the tables and schemas.
An AWS Glue job writes processed data from the tables to an Amazon Redshift database. The AWS Glue job handles column mapping and creates the Amazon Redshift tables in the Redshift database appropriately.
If the company reruns the AWS Glue job for any reason, duplicate records are introduced into the Amazon Redshift tables. The company needs a solution that will update the Redshift tables without duplicates.
Which solution will meet these requirements?

- A. Modify the AWS Glue job to copy the rows into a staging Redshift table. Add SQL commands to update the existing rows with new values from the staging Redshift table.
- B. Use the AWS Glue ResolveChoice built-in transform to select the value of the column from the most recent record.
- C. Modify the AWS Glue job to load the previously inserted data into a MySQL database. Perform an upsert operation in the MySQL database. Copy the results to the Amazon Redshift tables.
- D. Use Apache Spark's DataFrame dropDuplicates() API to eliminate duplicates. Write the data to the Redshift tables.

Answer: A

Explanation:
To avoid duplicate records in Amazon Redshift, the most effective solution is to perform the ETL in a way that first loads the data into a staging table and then uses SQL commands like MERGE or UPDATE to insert new records and update existing records without introducing duplicates.
* Using Staging Tables in Redshift:
* The AWS Glue job can write data to a staging table in Redshift. Once the data is loaded, SQL commands can be executed to compare the staging data with the target table and update or insert records appropriately. This ensures no duplicates are introduced during re-runs of the Glue job.
Reference:Amazon Redshift Best Practices
Alternatives Considered:
B (MySQL upsert): This introduces unnecessary complexity by involving another database (MySQL).
C (Spark dropDuplicates): While Spark can eliminate duplicates, handling duplicates at the Redshift level with a staging table is a more reliable and Redshift-native solution.
D (AWS Glue ResolveChoice): The ResolveChoice transform in Glue helps with column conflicts but does not handle record-level duplicates effectively.
References:
Amazon Redshift MERGE Statements

Staging Tables in Amazon Redshift

NEW QUESTION # 156
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. Partition the data that is in the S3 bucket. Organize the data by year, month, and day.
- 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. Increase the AWS Glue instance size by scaling up the worker type.
- E. Modify the 1AM role that grants access to AWS glue to grant access to all S3 features.

Answer: A,D

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 # 157
......

Free 1 year Amazon Data-Engineer-Associate Dumps Updates

ALL candidates can receive our pass guide AWS Certified Data Engineer - Associate (DEA-C01) dumps Data-Engineer-Associate Valid Test Blueprint materials soon after payment, You just take 20-30 hours to learn it, Now our pass rate has reached 99 percent.

In order to provide the latest and the most Data-Engineer-Associate Accurate Study Material, our IT experts are doing their best to update the Data-Engineer-Associate exam training pdf to make sure that our customers can have a full knowledge about Data-Engineer-Associate actual examination.

- Pass Guaranteed 2026 Amazon Data-Engineer-Associate –Valid Reliable Real Exam 🌳 Open [www.practicevce.com] and search for [Data-Engineer-Associate] to download exam materials for free 📁Exam Dumps Data-Engineer-Associate Free
- Data-Engineer-Associate Free Exam 🌝 Valid Data-Engineer-Associate Exam Camp Pdf 🦋 Authorized Data-Engineer-Associate Certification 🚚 Open website ☀ www.pdfvce.com 🌞️ and search for { Data-Engineer-Associate } for free download 🧛Reliable Data-Engineer-Associate Dumps Sheet
- Valid Data-Engineer-Associate Exam Dumps 🔫 Data-Engineer-Associate Reliable Dumps Free 😅 Data-Engineer-Associate Formal Test 🍀 Open [www.validtorrent.com] enter 🔆 Data-Engineer-Associate 🍝 and obtain a free download 🚛Testking Data-Engineer-Associate Exam Questions
- Exam Dumps Data-Engineer-Associate Free 🔱 Reliable Data-Engineer-Associate Exam Pattern 🏆 Verified Data-Engineer-Associate Answers 😕 Search on （ www.pdfvce.com ） for 《 Data-Engineer-Associate 》 to obtain exam materials for free download 🚘Data-Engineer-Associate Study Demo
- Achieve Your Goals with Real Amazon Data-Engineer-Associate Exam Questions � Enter ✔ www.prep4sures.top 🍡✔️ and search for ⇒ Data-Engineer-Associate ⇐ to download for free 🥁Verified Data-Engineer-Associate Answers
- 100% Pass Quiz Amazon - High Hit-Rate Data-Engineer-Associate - Reliable AWS Certified Data Engineer - Associate (DEA-C01) Real Exam 🏯 Copy URL ▶ www.pdfvce.com ◀ open and search for ➡ Data-Engineer-Associate 🍑 to download for free 🛸Testking Data-Engineer-Associate Exam Questions
- 2026 Data-Engineer-Associate – 100% Free Reliable Real Exam | High-quality Valid Data-Engineer-Associate Test Pass4sure 🦐 Search for 《 Data-Engineer-Associate 》 on ▶ www.torrentvce.com ◀ immediately to obtain a free download 🎍Data-Engineer-Associate Test Questions Vce
- 100% Pass Quiz Amazon - Data-Engineer-Associate - Reliable AWS Certified Data Engineer - Associate (DEA-C01) Real Exam 🔈 Search for 🍏 Data-Engineer-Associate 🦨 on 🐲 www.pdfvce.com 🌼 immediately to obtain a free download 🕴🧑Authorized Data-Engineer-Associate Certification
- Valid Data-Engineer-Associate Real Test 🔥 Data-Engineer-Associate Formal Test 🏤 Data-Engineer-Associate Reliable Test Sims 🐾 Immediately open ☀ www.pdfdumps.com 🌞️ and search for ➡ Data-Engineer-Associate 🍚 to obtain a free download 🍭Dumps Data-Engineer-Associate Collection
- Practical Reliable Data-Engineer-Associate Real Exam - Guaranteed Amazon Data-Engineer-Associate Exam Success with Useful Valid Data-Engineer-Associate Test Pass4sure 🚶 Open ▷ www.pdfvce.com ◁ and search for 《 Data-Engineer-Associate 》 to download exam materials for free 🥔Data-Engineer-Associate Free Exam
- Achieve Your Goals with Real Amazon Data-Engineer-Associate Exam Questions 🔝 ➤ www.exam4labs.com 🡸 is best website to obtain （ Data-Engineer-Associate ） for free download 🌮Reliable Data-Engineer-Associate Exam Pattern
- myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, bbs.t-firefly.com, the-businesslounge.com, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, Disposable vapes

P.S. Free 2026 Amazon Data-Engineer-Associate dumps are available on Google Drive shared by PassCollection: https://drive.google.com/open?id=1_KIV9w89u1GLDT99bbZxOHW0HCCzKYvE