

Data-Engineer-Associate Exam Voucher | Data-Engineer-Associate Latest Exam Notes



What's more, part of that Dumbleader Data-Engineer-Associate dumps now are free: https://drive.google.com/open?id=1X_cVIt1b-v3JA35LFhGxsyN7SA6QSaKj

Many students often start to study as the exam is approaching. Time is very valuable to these students, and for them, one extra hour of study may mean 3 points more on the test score. If you are one of these students, then AWS Certified Data Engineer - Associate (DEA-C01) exam tests are your best choice. Because students often purchase materials from the Internet, there is a problem that they need transport time, especially for those students who live in remote areas. When the materials arrive, they may just have a little time to read them before the exam. However, with Data-Engineer-Associate Exam Questions, you will never encounter such problems, because our materials are distributed to customers through emails. After you have successfully paid, you can immediately receive Data-Engineer-Associate test guide from our customer service staff, and then you can start learning immediately.

You can choose the number of AWS Certified Data Engineer - Associate (DEA-C01) (Data-Engineer-Associate) questions and time frame of the Data-Engineer-Associate Desktop practice exam software as per your learning needs. Performance reports of Amazon Data-Engineer-Associate Practice Test will be useful for tracking your progress and identifying areas for further study.

>> Data-Engineer-Associate Exam Voucher <<

Data-Engineer-Associate Latest Exam Notes - Data-Engineer-Associate Latest Study Notes

Our online test engine and the windows software of the Data-Engineer-Associate study materials can evaluate your exercises of the virtual exam and practice exam intelligently. Our calculation system of the Data-Engineer-Associate study materials is designed subtly. Our evaluation process is absolutely correct. We are strictly in accordance with the detailed grading rules of the real exam. The point of every question is set separately. Once you submit your exercises of the Data-Engineer-Associate Study Materials, the calculation system will soon start to work.

Amazon AWS Certified Data Engineer - Associate (DEA-C01) Sample Questions (Q51-Q56):

NEW QUESTION # 51

Two developers are working on separate application releases. The developers have created feature branches named Branch A and Branch B by using a GitHub repository's master branch as the source.

The developer for Branch A deployed code to the production system. The code for Branch B will merge into a master branch in the following week's scheduled application release.

Which command should the developer for Branch B run before the developer raises a pull request to the master branch?

- A. git diff branchB master
git commit -m <message>
- B. git rebase master
- C. git pull master

- D. git fetch -b master

Answer: B

Explanation:

To ensure that Branch B is up to date with the latest changes in the master branch before submitting a pull request, the correct approach is to perform a git rebase. This command rewrites the commit history so that Branch B will be based on the latest changes in the master branch.

git rebase master:

This command moves the commits of Branch B to be based on top of the latest state of the master branch. It allows the developer to resolve any conflicts and create a clean history.

Reference:

Alternatives Considered:

A (git diff): This will only show differences between Branch B and master but won't resolve conflicts or bring Branch B up to date.

B (git pull master): Pulling the master branch directly does not offer the same clean history management as rebase.

D (git fetch -b): This is an incorrect command.

Git Rebase Best Practices

NEW QUESTION # 52

A retail company stores data from a product lifecycle management (PLM) application in an on-premises MySQL database. The PLM application frequently updates the database when transactions occur.

The company wants to gather insights from the PLM application in near real time. The company wants to integrate the insights with other business datasets and to analyze the combined dataset by using an Amazon Redshift data warehouse.

The company has already established an AWS Direct Connect connection between the on-premises infrastructure and AWS.

Which solution will meet these requirements with the LEAST development effort?

- A. Run a scheduled AWS Glue extract, transform, and load (ETL) job to get the MySQL database updates by using a Java Database Connectivity (JDBC) connection. Set Amazon Redshift as the destination for the ETL job.
- **B. Run a full load plus CDC task in AWS Database Migration Service (AWS DMS) to continuously replicate the MySQL database changes. Set Amazon Redshift as the destination for the task.**
- C. Use the Amazon AppFlow SDK to build a custom connector for the MySQL database to continuously replicate the database changes. Set Amazon Redshift as the destination for the connector.
- D. Run scheduled AWS DataSync tasks to synchronize data from the MySQL database. Set Amazon Redshift as the destination for the tasks.

Answer: B

Explanation:

* Problem Analysis:

* The company needs near real-time replication of MySQL updates to Amazon Redshift.

* Minimal development effort is required for this solution.

* Key Considerations:

* AWS DMS provides a full load + CDC (Change Data Capture) mode for continuous replication of database changes.

* DMS integrates natively with both MySQL and Redshift, simplifying setup.

* Solution Analysis:

* Option A: AWS Glue Job

* Glue is batch-oriented and does not support near real-time replication.

* Option B: DMS with Full Load + CDC

* Efficiently handles initial database load and continuous updates.

* Requires minimal setup and operational overhead.

* Option C: AppFlow SDK

* AppFlow is not designed for database replication. Custom connectors increase development effort.

* Option D: DataSync

* DataSync is for file synchronization and not suitable for database updates.

* Final Recommendation:

* Use AWS DMS in full load + CDC mode for continuous replication.

:

AWS Database Migration Service Documentation

Setting Up DMS with Redshift

NEW QUESTION # 53

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

Answer: A

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 # 54

A company needs to build a data lake in AWS. The company must provide row-level data access and column-level data access to specific teams. The teams will access the data by using Amazon Athena, Amazon Redshift Spectrum, and Apache Hive from Amazon EMR.

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

- A. Use Amazon S3 for data lake storage. Use Apache Ranger through Amazon EMR to restrict data access by rows and columns. Provide data access by using Apache Pig.
- B. Use Amazon Redshift for data lake storage. Use Redshift security policies to restrict data access by rows and columns. Provide data access by using Apache Spark and Amazon Athena federated queries.
- C. Use Amazon S3 for data lake storage. Use AWS Lake Formation to restrict data access by rows and columns. Provide data access through AWS Lake Formation.
- D. Use Amazon S3 for data lake storage. Use S3 access policies to restrict data access by rows and columns. Provide data access through Amazon S3.

Answer: C

Explanation:

Option D is the best solution to meet the requirements with the least operational overhead because AWS Lake Formation is a fully managed service that simplifies the process of building, securing, and managing data lakes. AWS Lake Formation allows you to define granular data access policies at the row and column level for different users and groups. AWS Lake Formation also integrates with Amazon Athena, Amazon Redshift Spectrum, and Apache Hive on Amazon EMR, enabling these services to access the data in the data lake through AWS Lake Formation.

Option A is not a good solution because S3 access policies cannot restrict data access by rows and columns.

S3 access policies are based on the identity and permissions of the requester, the bucket and object ownership, and the object prefix and tags. S3 access policies cannot enforce fine-grained data access control at the row and column level.

Option B is not a good solution because it involves using Apache Ranger and Apache Pig, which are not fully managed services and require additional configuration and maintenance. Apache Ranger is a framework that provides centralized security administration for data stored in Hadoop clusters, such as Amazon EMR.

Apache Ranger can enforce row-level and column-level access policies for Apache Hive tables. However, Apache Ranger is not a

native AWS service and requires manual installation and configuration on Amazon EMR clusters. Apache Pig is a platform that allows you to analyze large data sets using a high-level scripting language called Pig Latin. Apache Pig can access data stored in Amazon S3 and process it using Apache Hive. However, Apache Pig is not a native AWS service and requires manual installation and configuration on Amazon EMR clusters.

Option C is not a good solution because Amazon Redshift is not a suitable service for data lake storage.

Amazon Redshift is a fully managed data warehouse service that allows you to run complex analytical queries using standard SQL. Amazon Redshift can enforce row-level and column-level access policies for different users and groups. However, Amazon Redshift is not designed to store and process large volumes of unstructured or semi-structured data, which are typical characteristics of data lakes. Amazon Redshift is also more expensive and less scalable than Amazon S3 for data lake storage.

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

What Is AWS Lake Formation? - AWS Lake Formation

Using AWS Lake Formation with Amazon Athena - AWS Lake Formation

Using AWS Lake Formation with Amazon Redshift Spectrum - AWS Lake Formation Using AWS Lake Formation with Apache Hive on Amazon EMR - AWS Lake Formation Using Bucket Policies and User Policies - Amazon Simple Storage Service Apache Ranger Apache Pig What Is Amazon Redshift? - Amazon Redshift

NEW QUESTION # 55

A company is building an inventory management system and an inventory reordering system to automatically reorder products. Both systems use Amazon Kinesis Data Streams. The inventory management system uses the Amazon Kinesis Producer Library (KPL) to publish data to a stream. The inventory reordering system uses the Amazon Kinesis Client Library (KCL) to consume data from the stream. The company configures the stream to scale up and down as needed.

Before the company deploys the systems to production, the company discovers that the inventory reordering system received duplicated data.

Which factors could have caused the reordering system to receive duplicated data? (Select TWO.)

- A. The producer experienced network-related timeouts.
- B. The AggregationEnabled configuration property was set to true.
- C. There was a change in the number of shards, record processors, or both.
- D. The stream's value for the IteratorAgeMilliseconds metric was too high.
- E. The max_records configuration property was set to a number that was too high.

Answer: A,C

Explanation:

* Problem Analysis:

* The company uses Kinesis Data Streams for both inventory management and reordering.
* The Kinesis Producer Library (KPL) publishes data, and the Kinesis Client Library (KCL) consumes data.

* Duplicate records were observed in the inventory reordering system.

* Key Considerations:

* Kinesis streams are designed for durability but may produce duplicates under certain conditions.
* Factors such as network timeouts, shard splits, or changes in record processors can cause duplication.

* Solution Analysis:

* Option A: Network-Related Timeouts

* If the producer (KPL) experiences network timeouts, it retries data submission, potentially causing duplicates.

* Option B: High IteratorAgeMilliseconds

* High iterator age suggests delays in processing but does not directly cause duplication.

* Option C: Changes in Shards or Processors

* Changes in the number of shards or record processors can lead to re-processing of records, causing duplication.

* Option D: AggregationEnabled Set to True

* AggregationEnabled controls the aggregation of multiple records into one, but it does not cause duplication.

* Option E: High max_records Value

* A high max_records value increases batch size but does not lead to duplication.

* Final Recommendation:

* Network-related timeouts and changes in shards or processors are the most likely causes of duplicate data in this scenario.

:

Amazon Kinesis Data Streams Best Practices

Kinesis Producer Library (KPL) Overview

Kinesis Client Library (KCL) Overview

NEW QUESTION # 56

.....

Our Data-Engineer-Associate exam guide is suitable for everyone whether you are a business man or a student, because you just need 20-30 hours to practice it that you can attend to your exam. There is no doubt that you can get a great grade. If you follow our learning pace, you will get unexpected surprises. Only when you choose our Data-Engineer-Associate Guide Torrent will you find it easier to pass this significant Data-Engineer-Associate examination and have a sense of brand new experience of preparing the Data-Engineer-Associate exam.

Data-Engineer-Associate Latest Exam Notes: https://www.dumpleader.com/Data-Engineer-Associate_exam.html

Our researchers and experts are working hard to develop the newest version Data-Engineer-Associate study materials, The Data-Engineer-Associate exam questions are categorized into specific areas, letting you focus on the Amazon Data-Engineer-Associate subject areas you need to work on, Amazon Data-Engineer-Associate Exam Voucher They personally attest that time is money, Amazon Data-Engineer-Associate Exam Voucher They will instruct you on efficient points of knowledge to get familiar and remember high-effective.

Logging Videotape Prior to Capture, Best, Worst, and Average Case, Our researchers and experts are working hard to develop the newest version Data-Engineer-Associate Study Materials.

The Data-Engineer-Associate exam questions are categorized into specific areas, letting you focus on the Amazon Data-Engineer-Associate subject areas you need to work on, They personally attest that time is money.

Data-Engineer-Associate Exam Voucher - 100% Fantastic Questions Pool

They will instruct you on efficient points of knowledge Data-Engineer-Associate to get familiar and remember high-effective, You can have a practice through different versions.

- Pass Guaranteed Quiz Data-Engineer-Associate - High-quality AWS Certified Data Engineer - Associate (DEA-C01) Exam Voucher Open ➡ www.examdiscuss.com enter 「 Data-Engineer-Associate 」 and obtain a free download Data-Engineer-Associate Reliable Exam Sims
- Data-Engineer-Associate Test Vce Data-Engineer-Associate Latest Dumps Questions Valid Data-Engineer-Associate Exam Answers Open ➤ www.pdfvce.com and search for [Data-Engineer-Associate] to download exam materials for free Data-Engineer-Associate Reliable Test Voucher
- Free PDF Amazon Data-Engineer-Associate - AWS Certified Data Engineer - Associate (DEA-C01) Exam Voucher Copy URL ➤ www.prepawaypdf.com open and search for ➤ Data-Engineer-Associate to download for free Valid Exam Data-Engineer-Associate Registration
- Data-Engineer-Associate Exam Forum Data-Engineer-Associate Reliable Test Braindumps Data-Engineer-Associate Certification Sample Questions The page for free download of ✓ Data-Engineer-Associate ✓ on ➡ www.pdfvce.com will open immediately Data-Engineer-Associate Exam Forum
- Latest Data-Engineer-Associate Study Plan Data-Engineer-Associate Certification Exam Dumps Latest Data-Engineer-Associate Study Plan Search for ✓ Data-Engineer-Associate ✓ and download it for free on ➡ www.practicevce.com website Valid Exam Data-Engineer-Associate Registration
- Valid Data-Engineer-Associate Exam Answers Data-Engineer-Associate Certification Exam Dumps Data-Engineer-Associate Reliable Test Voucher Enter ⚡ www.pdfvce.com and search for « Data-Engineer-Associate » to download for free Data-Engineer-Associate Reliable Test Braindumps
- Data-Engineer-Associate Certification Exam Dumps Data-Engineer-Associate Latest Braindumps Sheet Data-Engineer-Associate Latest Braindumps Sheet Open ➡ www.vceengine.com enter « Data-Engineer-Associate » and obtain a free download Data-Engineer-Associate Certification Exam Dumps
- Data-Engineer-Associate Latest Exam Cost Data-Engineer-Associate Certification Sample Questions Exam Data-Engineer-Associate Actual Tests Enter ➡ www.pdfvce.com and search for ➡ Data-Engineer-Associate ⇄ to download for free Data-Engineer-Associate Exam Dumps Demo
- Data-Engineer-Associate Exam Dumps Demo Reliable Data-Engineer-Associate Braindumps Data-Engineer-Associate Reliable Exam Sims Open ⚡ www.troytecdumps.com ⚡ enter ➡ Data-Engineer-Associate ⇄ and obtain a free download Valid Exam Data-Engineer-Associate Registration
- Free PDF Amazon Data-Engineer-Associate - AWS Certified Data Engineer - Associate (DEA-C01) Exam Voucher Go to website 「 www.pdfvce.com 」 open and search for ➡ Data-Engineer-Associate to download for free ↴Data-Engineer-Associate Valid Learning Materials
- Valid Data-Engineer-Associate Exam Answers Reliable Data-Engineer-Associate Braindumps Data-Engineer-Associate Exam Score Search for Data-Engineer-Associate and obtain a free download on ➤ www.validtorrent.com Valid Data-Engineer-Associate Exam Cram

P.S. Free 2026 Amazon Data-Engineer-Associate dumps are available on Google Drive shared by Dumpleader:
https://drive.google.com/open?id=1X_cVIt1b-v3JA35LFhGxsyN7SA6QSaKj