100% Pass Quiz Exam DEA-C02 Outline - First-grade SnowPro Advanced: Data Engineer (DEA-C02) Test Dumps Pdf

With limited time for your preparation, many exam candidates can speed up your pace of making progress. Our DEA-C02 practice materials will remedy your faults of knowledge understanding for our DEA-C02 exam questions contain everything you need in the real DEA-C02 exam. You won't regret your decision of choosing our DEA-C02 training guide. In contrast, they will inspire your potential without obscure content to feel. After getting our DEA-C02 exam prep, you will not live under great stress during the exam period.

Snowflake DEA-C02 latest exam lab questions are collected and arranged based on latest exam questions and new information materials. It covers a range wide and includes latest exam knowledge points. If you are urgent to pass exam DEA-C02 Latest Exam lab questions will be the best preparation materials for you. Complete and valid exam study learning materials will help you save time cost and economic cost, then clear exam easily.

>> Exam DEA-C02 Outline <<

DEA-C02 Test Dumps Pdf, Reliable DEA-C02 Exam Pdf

When you're in pain, it is best to learn things. Learning will make you invincible. DumpsTorrent Snowflake DEA-C02 Exam Training materials can also help you to be invincible. With this training materials, you will receive the Snowflake DEA-C02 certification which recognized and accepted internationally. Then all of your life, including money and position, will improve a lot. Until then, will you still feel painful? No, you will be very happy. You should thanks DumpsTorrent which provide you with a good training materials. It can help you when you lost, and let you not only improve your own quality, but also demonstratethe value of your perfect life.

Snowflake SnowPro Advanced: Data Engineer (DEA-C02) Sample Questions (Q142-Q147):

NEW QUESTION # 142

You're loading data into a Snowflake table using 'COPY INTO'. You notice that some rows are being rejected due to data validation errors (e.g., data type mismatch, uniqueness constraint violations). You want to implement a strategy to capture these rejected rows for further analysis and correction. Which of the following approaches offers the MOST efficient and reliable method for capturing and storing the rejected rows, minimizing performance impact during the data loading process? Assume no staging table exists before loading data to production table.

- Enable the `VALIDATE` function with a large number of rows (e.g., `VALIDATE(1000)`) before running the `COPY INTO` command. Analyze the validation results and then use `COPY INTO` with `ON_ERROR = 'SKIP_FILE'` to avoid rejected rows.
- Set `ON_ERROR = 'CONTINUE'` in the `COPY INTO` command. After the load, query the `VALIDATE` function on the target table with `SAMPLE` to identify potentially problematic rows.
- Utilize the `COPY INTO ... REJECTED_RECORD_COUNT` functionality after each load, then implement a query to retrieve the rejected records from `SYSTEM$LAST_COPY_LOAD_SUMMARY()`.
- Implement `COPY INTO` with the `ON_ERROR = 'ABORT_STATEMENT'` parameter to rollback the entire load if any errors are encountered, ensuring data consistency at the cost of potential downtime.
- Use the `COPY INTO ... ERROR_INTEGRATION` parameter to automatically capture rejected rows in a separate error table or stage. Define an error notification integration to alert when errors occur.

- A. Option E
- B. Option C
- C. Option B
- D. Option D
- E. Option A

Answer: A

Explanation:

Option E, utilizing 'ERROR INTEGRATION', is the most efficient and reliable. It automatically captures rejected rows during the 'COPY INTO' process and stores them in a designated error table or stage, minimizing performance impact and providing a structured way to analyze and correct errors. Options A, B, C, and D have drawbacks. A requires pre-validation, adding overhead. B uses sampling, which might not identify all errors. C only provides a record count, not the actual rejected rows. D aborts the entire statement, impacting availability.

NEW QUESTION # 143

You have an external table named in Snowflake that points to a set of CSV files in an AWS S3 bucket. The CSV files have a header row, and the data is comma-separated. However, some of the files in the S3 bucket are gzipped. You need to define the external table to correctly read both compressed and uncompressed files. Which of the following SQL statements BEST achieves this?

- CREATE OR REPLACE EXTERNAL TABLE ext_sales_data (id NUMBER, product STRING, sales NUMBER) LOCATION = @my_s3_stage FILE_FORMAT = (TYPE = CSV, SKIP_HEADER = 1);
- CREATE OR REPLACE EXTERNAL TABLE ext_sales_data (id NUMBER, product STRING, sales NUMBER) LOCATION = @my_s3_stage FILE_FORMAT = (TYPE = CSV, SKIP_HEADER = 1, COMPRESSION = AUTO);
- CREATE OR REPLACE EXTERNAL TABLE ext_sales_data (id NUMBER, product STRING, sales NUMBER) LOCATION = @my_s3_stage FILE_FORMAT = (TYPE = CSV, SKIP_HEADER = 1, COMPRESSION = GZIP);
- CREATE OR REPLACE EXTERNAL TABLE ext_sales_data (id NUMBER, product STRING, sales NUMBER) LOCATION = @my_s3_stage FILE_FORMAT = (TYPE = CSV, SKIP_HEADER = 1) REFRESH;
- CREATE OR REPLACE EXTERNAL TABLE ext_sales_data (id NUMBER, product STRING, sales NUMBER) LOCATION = @my_s3_stage AUTO_REFRESH = TRUE FILE_FORMAT = (TYPE = CSV, SKIP_HEADER = 1, COMPRESSION = AUTO);

- A. Option C
- B. Option D
- C. Option E
- D. Option B
- E. Option A

Answer: D

Explanation:

The 'COMPRESSION = AUTO' parameter in the file format definition allows Snowflake to automatically detect and decompress gzipped files while also reading uncompressed files. The 'SKIP HEADER = 1 ' parameter ensures that the header row in the CSV files is skipped.

NEW QUESTION # 144
A Snowflake data engineer is troubleshooting a slow-running query that joins two large tables, 'ORDERS' (1 billion rows) and 'CUSTOMER' (10 million rows), using the 'CUSTOMER ID' column. The query execution plan shows a significant amount of data spilling to local disk. The query is as follows:

```
SELECT
    O.ORDER_ID,
    C.CUSTOMER_NAME,
    SUM(O.ORDER_AMOUNT)
FROM
    ORDERS O
JOIN
    CUSTOMER C ON O.CUSTOMER_ID = C.CUSTOMER_II
WHERE C.CUSTOMER_REGION = 'West'
GROUP BY
    O.ORDER_ID, C.CUSTOMER_NAME;
```

Which of the following are the MOST likely root causes of the disk spilling and the best corresponding solutions? Select two options that directly address the disk spilling issue.

- A. The statistics on the tables are outdated. Run 'ANALYZE TABLE ORDERS' and 'ANALYZE TABLE CUSTOMER to update the statistics.
- B. The virtual warehouse is undersized for the amount of data being processed. Increase the virtual warehouse size to provide more memory.
- C. The query is performing a full table scan on the 'ORDERS' table. Add an index on the 'CUSTOMER ID column in the 'ORDERS table.
- D. The join operation is resulting in a large intermediate result set that exceeds the available memory. Apply a filter on the 'ORDERS' table to reduce the data volume before the join.
- E. The 'CUSTOMER_ID column is not properly clustered in either the 'ORDERS' or 'CUSTOMER table. Define a clustering key on 'CUSTOMER_ID for both tables.

Answer: B,D

Explanation:
Options A and D are the most direct solutions for disk spilling. A undersized warehouse directly impacts available memory, leading to disk spilling. Increasing the warehouse size (option A) provides more memory for the operation. When data spill happens increasing the warehouse size is the primary action to take. Option D correctly addresses the root cause of the spill an overly large intermediate result set. Reducing the data volume before the join minimizes the memory required. Option B could improve query performance overall, but doesn't directly address disk spilling. Option C is incorrect, as Snowflake does not support manual indexes. Option E would improve the accuracy of the query optimizer's decisions, which could indirectly improve performance, but is less direct than options A and D.

NEW QUESTION # 145
You are designing a Snowflake data pipeline that continuously ingests clickstream dat a. You need to monitor the pipeline for latency and throughput, and trigger notifications if these metrics fall outside acceptable ranges. Which of the following combinations of Snowflake features and techniques would be MOST effective for achieving this goal?

- A. Implement a combination of Snowflake Streams, Tasks, and external functions. Streams capture changes, Tasks process the changes, and external functions send notifications to a monitoring service when latency or throughput issues are detected.
- B. Create a custom dashboard using a BI tool that connects to Snowflake via JDBC/ODBC and visualizes data ingestion and processing metrics. Manually monitor the dashboard for anomalies.
- C. Use Snowflake's Event Tables and Event Notifications to capture events related to data ingestion and processing. Configure alerts based on event patterns that indicate latency or throughput issues.
- D. Rely on Snowflake's default resource monitors to track warehouse usage. If warehouse usage exceeds a certain threshold, assume there are performance issues and send a notification.
- E. Use Snowflake's 'QUERY_HISTORY view to track query execution times and implement a scheduled task that queries this view, calculates latency and throughput, and sends email notifications using Snowflake's built-in email integration if thresholds are exceeded.

Answer: A,C

Explanation:
Options B and D offer the most effective solutions. Option B provides a granular approach using Streams, Tasks, and external functions for real-time monitoring and notification. Option D leverages Event Tables and Event Notifications, enabling a reactive

approach based on specific event patterns. Option A is less precise as it relies on query history, which may not accurately reflect real-time latency. Option C is too general. Option E requires manual monitoring, which is not ideal for continuous pipelines.

NEW QUESTION # 146

You are tasked with implementing a Row Access Policy (RAP) on a table 'customer_data' that contains Personally Identifiable Information (PII). The policy must meet the following requirements: 1. Data analysts with the 'ANALYST role should only see anonymized customer data (e.g., masked email addresses, hashed names). 2. Data engineers with the 'ENGINEER role should see the full, unmasked customer data for data processing purposes. 3. No other roles should have access to the data'. You create the following UDFs: 'MASK EMAIL(email address VARCHAR)': Returns an anonymized version of the email address. 'HASH NAME(name VARCHAR): Returns a hash of the customer name. Which of the following is the most efficient and secure way to implement this RAP, assuming minimal performance impact is desired?

) ```sql CREATE OR REPLACE ROW ACCESS POLICY customer_rap AS (email_address VARCHAR, name VARCHAR) RETURNS BOOLEAN -> CASE
VHEN IS_ROLE_IN_SESSION('ANALYST') THEN MASK_EMAIL(email_address) IS NOT NULL AND HASH_NAME(name) IS NOT NULL WHEN
S_ROLE_IN_SESSION('ENGINEER') THEN TRUE ELSE FALSE END; ALTER TABLE customer_data ADD ROW ACCESS POLICY customer_rap ON (email,
iame); ```

) ```sql CREATE OR REPLACE ROW ACCESS POLICY customer_rap AS (email_address VARCHAR, name VARCHAR) RETURNS BOOLEAN -> CASE
VHEN CURRENT_ROLE() = 'ANALYST' THEN MASK_EMAIL(email_address) IS NOT NULL AND HASH_NAME(name) IS NOT NULL WHEN
:URRENT_ROLE() = 'ENGINEER' THEN TRUE ELSE FALSE END; ALTER TABLE customer_data ADD ROW ACCESS POLICY customer_rap ON (email,
iame); ```

) ```sql CREATE OR REPLACE ROW ACCESS POLICY customer_rap AS (email_address VARCHAR, name VARCHAR) RETURNS BOOLEAN -> CASE
VHEN IS_ROLE_IN_SESSION('ANALYST') THEN CURRENT_ROLE() = 'ANALYST' WHEN IS_ROLE_IN_SESSION('ENGINEER') THEN CURRENT_ROLE()
ENGINEER' ELSE FALSE END; ALTER TABLE customer_data ADD ROW ACCESS POLICY customer_rap ON (email, name); ```

) ```sql CREATE OR REPLACE ROW ACCESS POLICY customer_rap AS (email_address VARCHAR, name VARCHAR) RETURNS BOOLEAN -> CASE
VHEN IS_ROLE_IN_SESSION('ANALYST') THEN TRUE WHEN IS_ROLE_IN_SESSION('ENGINEER') THEN TRUE ELSE FALSE END; ALTER TABLE
:ustomer_data ADD ROW ACCESS POLICY customer_rap ON (email, name); CREATE OR REPLACE VIEW analyst_view AS SELECT MASK_EMAIL(email),
HASH_NAME(name), other_columns FROM customer_data; ```

) ```sql CREATE OR REPLACE ROW ACCESS POLICY customer_rap AS (email_address VARCHAR, name VARCHAR) RETURNS BOOLEAN ->
S_ROLE_IN_SESSION('ANALYST') OR IS_ROLE_IN_SESSION('ENGINEER'); ALTER TABLE customer_data ADD ROW ACCESS POLICY customer_rap ON
email, name); ```

- A. Option C
- B. Option B
- C. Option D
- D. Option E
- E. Option A

Answer: C

Explanation:

Option D is the most efficient because it filters access based on roles in the RAP without applying expensive UDFs within the policy itself. This minimizes the performance impact of the RAP. The view 'analyst_view' then applies the masking/hashing for analysts. Options A and B apply the UDFs within the RAP, which will significantly degrade performance. The 'MASK EMAIL(email_address) IS NOT NULL' conditions are also incorrect as they are not validating the email. Option C doesn't implement the required masking/hashing for analysts at all, and also is not as effecient. Option E allows both roles to see all data which does not meet requirement 1.

NEW QUESTION # 147

......

Because industry of information technology is fast-moving. To excel in this advanced industry, pass the DEA-C02 exam of the Snowflake DEA-C02 certification. Hundreds of applicants have faced issues in updated dumps material to crack the Snowflake DEA-C02 examination in one go.

This option unlocks one of the best Family Sharing DEA-C02 features: Ask to Buy, Optimize performance with caching, Become part of the huge group of people who used DumpsTorrent DEA-C02 Latest Study Notes for the DEA-C02 Snowflake certification and passing with flying colors.

Easiest and Quick Way to Pass Snowflake DEA-C02 Exam

"How" you may ask: simple, our easy to download exams DEA-C02 Test Dumps Pdf are examples from the actual Snowflake certification exam, After finishing your task, you can reviewthem plenty of times and find out the wrong items, some Reliable DEA-C02 Exam Pdf questions may have explanations for your understanding, and you can practice many times day to day.

IT-Tests has rich experience in DEA-C02 certification exam, If it is ok, don't hesitate to sign up for the exam.