

Test SPS-C01 Preparation, SPS-C01 Exam Questions Answers



Unlike other kinds of exam files which take several days to wait for delivery from the date of making a purchase, our SPS-C01 study materials can offer you immediate delivery after you have paid for them. The moment you money has been transferred to our account, and our system will send our SPS-C01 training dumps to your mail boxes so that you can download SPS-C01 exam questions directly. It is fast and convenient out of your imagination.

With the ever-increasing competition, people take Snowflake SPS-C01 certification to exhibit their experience, skills, and abilities in a better way. Having Snowflake Certified SnowPro Specialty - Snowpark SPS-C01 certificate shows that you have better exposure than others. So, SPS-C01 Certification also gives you an advantage in the industry when employers seek candidates for job opportunities. However, preparing for the Snowflake SPS-C01 exam can be a difficult and time-consuming process.

>> Test SPS-C01 Preparation <<

SPS-C01 Exam Questions Answers & Study SPS-C01 Demo

Market is a dynamic place because a number of variables keep changing, so is the practice materials field of the SPS-C01 practice exam. Our SPS-C01 exam dumps are indispensable tool to pass it with high quality and low price. By focusing on how to help you effectively, we encourage exam candidates to buy our SPS-C01 practice test with high passing rate up to 98 to 100 percent all these years. Our Snowflake exam dumps almost cover everything you need to know about the exam. As long as you practice our SPS-C01 Test Question, you can pass exam quickly and successfully. By using them, you can not only save your time and money, but also pass SPS-C01 practice exam without any stress.

Snowflake Certified SnowPro Specialty - Snowpark Sample Questions (Q121-Q126):

NEW QUESTION # 121

You're using Snowpark in Python and need to execute a complex SQL query. The query involves several joins and aggregations, and you want to optimize its performance. You are using 'session.sql(query)' to execute the query. Which of the following strategies, applied before executing 'session.sql(query)', would likely lead to the most significant performance improvement for a very large

dataset?

- A. Create a view of the underlying data source instead of directly querying the table.
- B. Reduce the size of the data by filtering the DataFrame returned by 'session.sql(queryy using 'where()' before executing any further operations.
- C. Ensure that the SQL query includes appropriate comments to improve readability.
- D. Use the method on the DataFrame returned by 'session.sql(queryy.
- E. Convert the SQL query into a series of Snowpark DataFrame operations (e.g., 'groupBy()', 'agg()').

Answer: E

Explanation:

Option A provides the most significant improvement because Snowpark DataFrame operations allow Snowflake's query optimizer to leverage pushdown optimizations. When you express your logic as DataFrame operations, Snowpark translates these into SQL that is specifically tailored for Snowflake's engine. This gives Snowflake more control over the execution plan compared to simply passing in a pre-written SQL query via 'session.sql(queryy. DataFrame operations allow the query optimizer to push down operations such as filters and aggregations to the data source, significantly reducing the amount of data transferred and processed. Option B is incorrect because comments only improve readability, not performance. Option C, , can help if the DataFrame is used multiple times, but it doesn't address the initial optimization of the query itself. Option D could help, but converting to DataFrame operations provides more comprehensive optimization. Option E can assist, but often DataFrame creation and optimal query plan generation can be better using Option A.

NEW QUESTION # 122

You have two Snowpark DataFrames, 'customers' and 'orders'. The 'customers' DataFrame has columns and 'customer name'. The 'orders' DataFrame has columns 'order id', 'customer id', and 'order amount'. You need to find all customers who have NOT placed any orders. Which of the following Snowpark set operations correctly implements this?

- A. ☐
- B. ☐
- C. ☐
- D. ☒
- E. ☐

Answer: D

Explanation:

Option D is correct. It first selects the distinct 'customer_id' from both DataFrames. Then, it uses the 'minus' (or 'except_') set operation to find the difference between the customer IDs in the 'customers' DataFrame and the customer IDs in the 'orders' DataFrame. This effectively returns the customer IDs of customers who have not placed any orders. The 'join' operation is not a set operation, and options B and C are not valid syntax for Snowpark. Option E is syntactically correct in Snowpark, and equivalent to Option D. However, since the question has to have one answer, Option D is kept.

NEW QUESTION # 123

You have a Snowpark Python stored procedure that performs complex data transformations. This stored procedure needs to read data from a large table ('TRANSACTIONS') and write the transformed data to another table 'PROCESSED TRANSACTIONS'. You want to optimize the performance of this stored procedure by leveraging Snowpark's features for parallel processing. Which of the following approaches can significantly improve the performance of the stored procedure, assuming sufficient warehouse resources are available?

- A. Load the data from 'TRANSACTIONS' table into a temporary table within the stored procedure, then use standard SQL queries on the temporary table for transformations, finally using snowpark DataFrame API to write it back to the 'PROCESSED_TRANSACTIONS' table.
- B. Use Snowflake's standard SQL queries within the stored procedure to read and transform the data. Write the results to the 'PROCESSED_TRANSACTIONS' table using 'INSERT' statements.
- C. Read the entire 'TRANSACTIONS' table into a Pandas DataFrame within the stored procedure and perform the transformations using Pandas functions. Then, write the transformed data back to the table using Snowpark's 'createDataFrame' and 'write' methods.
- D. Use Snowpark's 'sprocs' decorator with appropriate 'packages' and leverage the Snowpark DataFrame API with vectorized UDFs to transform the data. Use 'session.write_pandas' to write the Pandas DataFrame to the

'PROCESSED_TRANSACTIONS' table after the transformation.

- E. Use the Snowpark DataFrame API to read the 'TRANSACTIONS' table and apply transformations using vectorized UDFs. Then, use the 'write' method to write the transformed data to the 'PROCESSED_TRANSACTIONS' table.

Answer: E

Explanation:

Using Snowpark DataFrame API along with vectorized UDFs leverages Snowflake's distributed processing capabilities for parallel execution, greatly enhancing performance. Reading the entire table into a Pandas DataFrame (Option B) limits parallelism and can lead to memory issues with large datasets. While SQL queries (Option C) work, they don't fully leverage Snowpark's optimized data transfer and processing. Option D refers to 'session.write_pandas' which isn't accurate in the context of writing transformed Snowpark data to Snowflake tables within the stored procedure. Using a temporary table and standard SQL queries, while functional, doesn't harness the full potential of Snowpark's distributed execution engine as effectively as using the DataFrame API directly (Option E).

NEW QUESTION # 124

A data engineering team is using Snowpark Python to build a complex ETL pipeline. They notice that certain transformations are not being executed despite being defined in the code. Which of the following are potential reasons why transformations in Snowpark might not be executed immediately, reflecting the principle of lazy evaluation? Select TWO correct answers.

- A. The 'eager_execution' session parameter is set to 'True'.
- B. Snowpark operations are only executed when an action (e.g., 'collect()', 'show()', is called on the DataFrame or when the DataFrame is materialized.
- C. Snowpark employs lazy evaluation to optimize query execution by delaying the execution of transformations until the results are actually required.
- D. The size of the data being processed exceeds Snowflake's memory limits, causing transformations to be skipped.
- E. Snowpark automatically executes all transformations as soon as they are defined, regardless of whether the results are needed.

Answer: B,C

Explanation:

Snowpark employs lazy evaluation, which means transformations are not executed until an action is performed on the DataFrame. This allows Snowflake to optimize the entire query plan before execution. Setting 'eager_execution' to True does NOT exist in Snowpark Python. Data size exceeding Snowflake's limits would result in an error, not skipped transformations.

NEW QUESTION # 125

A Snowpark Python application is experiencing significant performance degradation when processing a large dataset (100GB+) stored in Snowflake. The application performs a complex series of transformations, including window functions and joins with smaller lookup tables. You suspect data skew is contributing to the issue. Which of the following strategies would be MOST effective in mitigating the impact of data skew and improving performance?

- A. Disable query result caching to ensure fresh data is always used.
- B. Implement salting or pre-partitioning of the data based on a hash of the skewed column before performing the joins and window functions.
- C. Convert the Snowpark DataFrame to a Pandas DataFrame before performing transformations.
- D. Increase the warehouse size to a larger instance (e.g., from X-Small to Large).
- E. Cache the smaller lookup tables using 'session.createDataFrame' and broadcast them to all worker nodes.

Answer: B

Explanation:

Salting or pre-partitioning addresses data skew directly by distributing the skewed values more evenly across partitions. Increasing warehouse size (A) might help to some extent but doesn't solve the underlying skew issue. Broadcasting small tables (C) is a good optimization, but it's less effective if the larger dataset is skewed. Disabling query result caching (D) is irrelevant to data skew. Converting to Pandas (E) will likely make performance worse for large datasets due to data transfer overhead and limitations of single-node processing.

• • • • •

SPS-C01 Exam Questions Answers: <https://www.torrentvalid.com/SPS-C01-valid-braindumps-torrent.html>

Scott lives in the California Bay Area where SPS-C01 his penchant for gardening can be indulged year round, Internet crime is real, We can assure you that our SPS-C01 practice dumps will make a significant difference to you as long as you want to change your status quo.

Pass with ease, In the fast-developing this industry, Study SPS-C01 Demo more and more technology standard and the knowledge have emerged every month, With the help of SPS-C01 lead4pass review, you can not only solve difficulties of questions in real exam, but also increase your efficiency level for SPS-C01 actual test.

[illegible]