

SPS-C01 Exam Duration - Latest SPS-C01 Test Pass4sure



There are three different Snowflake SPS-C01 questions format that is being provided to applicants from Itexamguide. Anyone can download a free SPS-C01 exam dumps demo to evaluate this product before shopping. These Snowflake Certified SnowPro Specialty - Snowpark (SPS-C01) latest questions formats are Snowflake SPS-C01 PDF dumps format, web-based Snowflake Certified SnowPro Specialty - Snowpark (SPS-C01) practice tests, and desktop-based Snowflake SPS-C01 practice test software is provided to customers.

To do this you just need to pass SPS-C01 exam, which is quite challenging and demands thorough Snowflake Certified SnowPro Specialty - Snowpark (SPS-C01) exam preparation. For the complete, comprehensive and quick SPS-C01 Exam Preparation, the Itexamguide SPS-C01 Dumps questions are ideal. You should not ignore it and must try Itexamguide SPS-C01 exam questions for preparation today.

>> SPS-C01 Exam Duration <<

Latest SPS-C01 Test Pass4sure - SPS-C01 Valid Real Exam

The data that come up with our customers who have bought our SPS-C01 actual exam and provided their scores show that our high pass rate is 98% to 100%. This is hard to find and compare with in the market. And numerous enthusiastic feedbacks from our worthy clients give high praises not only on our SPS-C01 study torrent, but also on our sincere and helpful 24 hours customer services on SPS-C01 exam questions online. All of these prove that we are the first-class vendor in this career and have authority to ensure your success in your first try on SPS-C01 exam.

Snowflake Certified SnowPro Specialty - Snowpark Sample Questions (Q177-Q182):

NEW QUESTION # 177

Consider two Snowpark DataFrames, 'employees' and 'departments', with the following schemas: 'employees': (employee_id: Integer Type, employee_name: StringType, department_id: Integer Type, salary: IntegerType) 'departments': (department_id: Integer Type, department_name: StringType, location: StringType) You want to find the highest salary within each department, along with the department name and location, and display the results in a Snowpark DataFrame. Which of the following Snowpark Python code snippets correctly achieves this?

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

Answer: B

Explanation:

Option A correctly groups the 'employees' DataFrame by 'department_id', calculates the maximum salary for each department using and then joins the resulting DataFrame with the 'departments' DataFrame on Finally, it selects the required columns: department name, location, and the calculated maximum salary. Option B is not correct as it calculates the rank of all salaries within each department but it does not aggregate on 'department_id'. Option C incorrectly attempts to group by columns from both tables before joining. Option D incorrectly uses the un-joined employees table where rank_salaries are from department_id. Option E will not work due to table names during select. It tries to use 'joined_df' table for columns which doesn't have it.

NEW QUESTION # 178

You're working with Snowpark and want to load data from a Pandas DataFrame into a Snowpark DataFrame. The Pandas DataFrame, 'customer_data', contains columns with mixed data types (integers, strings, dates). Some columns also contain NULL values. You need to ensure that the data types are correctly inferred by Snowpark, NULL values are handled appropriately, and the resulting Snowpark DataFrame 'snowpark_customers' can be used for further transformations. What is the best approach to achieve this with minimal code and maximum performance?

- A. Use 'session.write_pandas' because its optimized for large pandas dataframe.
- B. Use 'session.createDataFrame(customer_data)' and rely on Snowpark to automatically infer the schema and handle NULL values implicitly. Convert any problematic columns after the Snowpark DataFrame is created.
- C. Infer the schema explicitly before creating the Snowpark DataFrame using Pandas DataFrame column types. For the string columns, define them to be StringType().
- D. Explicitly define the schema with StructType and StructField, specifying the column names and data types based on the Pandas DataFrame, converting null values to Snowflake's null representation during DataFrame creation.
- E. First, replace all NA/NaN values in Pandas DataFrame with None, then create Snowpark DataFrame using 'session.createDataFrame(customer_data)'.

Answer: A

Explanation:

Relying on schema inference (option C) might not always guarantee the correct data types, especially with dates or mixed-type columns. Explicitly defining the schema (option B) can be verbose and error-prone. Replacing NA/NaN with None and using 'createDataFrame' (option D) is a functional approach, but might not be as performant as the optimized method of 'write_pandas'. Inferring the schema (Option A) might not be fully accurate. Using session.write_pandas leverages internal Snowflake optimizations for data transfer and type handling from Pandas to Snowpark, making it the most efficient.

NEW QUESTION # 179

You are developing a Snowpark application that processes high-volume event data stored in a Snowflake table named 'raw events'. The application aggregates data by session ID. You observe significant performance degradation during peak hours. Analyzing Snowflake query history reveals that the 'session_id' column has high cardinality and data skew. Which of the following strategies, or combination of strategies, would be MOST effective in optimizing the aggregation performance?

- A. Increase the warehouse size to a larger tier (e.g., from X-Small to Small).
- B. Use a 'GROUP' clause in the Snowpark DataFrame to perform the aggregation.
- C. Implement a custom UDF (User-Defined Function) in Python to perform the aggregation and then apply the 'GROUP' clause in the Snowpark DataFrame.
- D. Use a 'GROUP BY' clause in the Snowpark DataFrame combined with a 'hint' to specify the 'for optimized parallel processing'.
- E. Pre-aggregate the raw event data into smaller batches using a scheduled task before the main Snowpark application runs, and then aggregate the pre-aggregated data in the Snowpark application.

Answer: D,E

Explanation:

Using BUCKET_ID hint improves parallel processing and mitigates data skew in Snowpark. Pre-aggregation reduces the amount of data processed by the Snowpark application, thus improving performance. Increasing the warehouse size (A) might help but doesn't address data skew. UDFs (D) can introduce overhead if not optimized. GROUP BY alone (B) will not address the data skew problem

NEW QUESTION # 180

You're developing a Snowpark application that reads data from a Snowflake table, performs several transformations, and then writes the results back to a different table. You want to ensure that the entire process is executed as a single atomic transaction, even if it involves multiple Snowpark DataFrames and operations. Which of the following actions are required to achieve this transactional behavior?

- A. Ensure that the target table for writing the results has the 'TRANSIENT' property set to 'TRUE'.
- **B. All Snowpark operations within a single session are automatically executed as a single atomic transaction by default; no additional configuration is required.**
- C. Leverage the 'CREATE OR REPLACE TABLE AS SELECT' statement within a Stored Procedure called from your Snowpark code. All DML operations done as part of stored proc is transactional
- D. Explicitly start a transaction using 'session.beginTransaction()' at the beginning of the Snowpark application and commit it using 'session.commitTransaction()' at the end.
- E. Configure the Snowpark session with the parameter set to ' FALSE

Answer: B

Explanation:

Snowflake inherently provides transactional consistency. All operations within a single Snowpark session are automatically executed as a single atomic transaction by default. This is a core feature of Snowflake and doesn't require explicit transaction management in most common scenarios. Options A, B and C are incorrect as Snowflake handles transaction automatically. E describes a possible solution, however, it isn't required.

NEW QUESTION # 181

You have a Snowpark DataFrame named 'products' with columns 'product_id' (INT), 'product_name' (STRING), and 'price' (DOUBLE). You want to apply a transformation to calculate a 'discounted_price' column, which is the 'price' reduced by 10% if the price is greater than \$100.00. Which of the following code snippets is the most efficient way to achieve this using Snowpark Python?

- **A.**
- B.
- **C.**
- D.
- E.

Answer: A,C

Explanation:

The most efficient ways are B and C. Option B directly uses the 'when' and 'otherwise' functions in Snowpark, which are optimized for execution within Snowflake. Option C is similar to B but explicitly uses 'lit' to represent the numeric literal, ensuring proper type handling in Snowpark. IJDFs (Option A) are generally less efficient than built-in functions. Option D attempts to use RDDs, which are not part of the Snowpark API. Option E is not valid Snowpark python syntax. Therefore, B and C are the correct answers.

NEW QUESTION # 182

.....

It is very important for us to keep pace with the changeable world and update our knowledge if we want to get a good job, a higher standard of life and so on. First, we need to get a good SPS-C01 quiz prep. Because we only pass SPS-C01 exam and get a certificate, we can have the chance to get a decent job and make more money. But there are question is that how you can pass the SPS-C01 Exam and get a certificate. The best answer is to download and learn our SPS-C01 quiz torrent. Our products will help you get what you want in a short time.

Latest SPS-C01 Test Pass4sure: https://www.itexamguide.com/SPS-C01_braindumps.html

