

Free SPS-C01 Practice Exams | SPS-C01 Hot Spot Questions



DOWNLOAD the newest TestSimulate SPS-C01 PDF dumps from Cloud Storage for free: https://drive.google.com/open?id=1o9YmC1q_T1hb0UXMrcKFZ02IIGSz1233

When you are hesitating whether to purchase our SPS-C01 exam software, why not try our free demo of SPS-C01. Once you have tried our free demo, you will ensure that our product can guarantee that you successfully Pass SPS-C01 Exam. Our professional IT team of TestSimulate continues updating and improving SPS-C01 exam dumps in order to guarantee you win the exam while you are preparing for the exam.

SPS-C01 practice materials are typically seen as the tools of reviving, practicing and remembering necessary exam questions for the exam, spending much time on them you may improve the chance of winning. However, our SPS-C01 training materials can offer better condition than traditional practice materials and can be used effectively. We treat it as our major responsibility to offer help so our SPS-C01 Practice Guide can provide so much help, the most typical one is their efficiency.

>> **Free SPS-C01 Practice Exams** <<

SPS-C01 Hot Spot Questions, Exam SPS-C01 Reviews

When dealing with any kind of exams, the most important thing is to find a scientific way to review effectively. Our SPS-C01 practice materials compiled by the most professional experts. Till now, we have over tens of thousands of customers around the world supporting our SPS-C01 exam torrent. If you are unfamiliar with our SPS-C01 Study Materials, please download the free demos for your reference. To some unlearned exam candidates, you can master necessities by our SPS-C01 practice materials quickly So our materials are elemental materials you cannot miss.

Snowflake Certified SnowPro Specialty - Snowpark Sample Questions (Q211-Q216):

NEW QUESTION # 211

You are tasked with optimizing the performance of a Snowpark Python application that performs complex data transformations on a large dataset of IoT sensor readings. The application uses a Snowpark-optimized warehouse. You notice that the application is consistently slow, with CPU utilization on the warehouse fluctuating significantly. Which of the following actions would be MOST effective in addressing this performance issue? Assume the dataset is partitioned on the 'sensor_id' column within Snowflake.

- A. Enable auto-scaling on the warehouse with a minimum of 2 and maximum of 5 clusters. This will allow the warehouse to dynamically adjust capacity based on workload.
- B. Increase the warehouse size to a larger instance (e.g., from X-Small to Small). This will provide more CPU and memory resources.
- C. Repartition the Snowpark DataFrame using `partition_expression='sensor_id'` before applying transformations. Then, explicitly colocate similar operations.
- D. Ensure the Snowpark DataFrame transformations are pushed down to Snowflake as much as possible by avoiding actions

like 'collect()' until absolutely necessary and leverage stored procedures.

- E. Rewrite the Snowpark DataFrame transformations using only built-in Snowpark functions and avoid using User-Defined Functions (UDFs) written in Python.

Answer: C,D,E

Explanation:

Repartitioning allows for improved parallelism and reduces data skew, especially when the initial data distribution is uneven. Avoiding Python UDFs improves performance because they execute outside of Snowflake's optimized engine. Pushing down transformations and leveraging stored procedures minimizes data transfer between Snowpark and Snowflake, and leverages Snowflake's processing capabilities. Increasing warehouse size or enabling auto-scaling might help, but addressing data skew and UDF overhead will likely provide more significant performance gains.

NEW QUESTION # 212

You have a Snowpark Python application that uses a UDF to perform custom data transformations. The UDF relies on a large, read-only lookup table that is stored as a CSV file on a Snowflake stage. Which of the following strategies would be the MOST efficient way to access the lookup table within the UDF?

- A. Load the CSV file into a Snowflake stage, and in the python UDF code, use the `get_stage_file` API from session object to read the file once. Then the data cached in-memory within the UDF module, and reuse the cached data for subsequent calls.
- B. Load the CSV file into a Snowflake table and then query the table from within the UDF using `'session.sql(V`.
- C. Read the CSV file from the stage every time the UDF is called using `'snowflake.connector.connect()' and then load the data into a Pandas DataFrame within the UDF function.`
- D. Use the `'cachetools` library with a Least Recently Used (LRU) cache to store the lookup table in memory. The UDF will check the cache before reading the CSV file, and update the cache if necessary. The CSV file is read with `get_stage_file` API from session.
- E. Read the CSV file from the stage once when the UDF is first called, cache the data in a global variable within the UDF module, and then reuse the cached data for subsequent calls.

Answer: B,D

Explanation:

Both options C and E are efficient. Option C leverages Snowflake's internal storage and query capabilities, avoiding repeated file reads. Although reading a file from stage only at once is good, it would impact the first call. Option E avoids hitting the stage every time, and only when the key is not present in cache. Options A and B suffer from performance bottlenecks due to repeated file access. In general, reading data within the Snowflake environment (C) is more performant than reading data from external sources within the UDF, especially when Snowflake's query optimizer can be used. Caching with LRU using the `'cachetools` library is effective, as cache would contain some values from large data. `get_stage_file` API from session object is efficient way.

NEW QUESTION # 213

You have a Snowflake table named 'CUSTOMER DATA' with columns 'CUSTOMER ID', 'NAME', 'CITY', and 'ORDER COUNT'. You want to create a Snowpark DataFrame named 'customer_df' containing only customers from 'New York' with an 'ORDER COUNT' greater than 10. Which of the following code snippets is the MOST efficient and correct way to achieve this, minimizing data transfer and maximizing pushdown optimization to Snowflake?

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

Answer: C

Explanation:

Option A is the most efficient. It directly uses the `'filter'` method with the combined condition, allowing Snowpark to push down the entire filter to Snowflake for execution. Option B uses `'session.sqr`, which might not be as optimized as using the table API directly. Option C pulls the entire table into Pandas, which is highly inefficient for large tables, defeating the purpose of Snowpark. Option D, while functional, is less readable than A, and the multiple `'where'` calls are logically equivalent to a single `'filter'` with combined conditions. Option E, while functional, contains redundant `'select'` `'V` operation.

NEW QUESTION # 214

You are developing a Snowpark Python application that needs to process large datasets. You want to optimize performance by leveraging user-defined functions (UDFs) to perform complex calculations in parallel across the Snowflake data warehouse. Which of the following statements regarding Snowpark UDFs are TRUE?

- A. To ensure optimal performance, it is recommended to always use the default Snowflake Anaconda channel for UDF dependencies, as custom channels may introduce latency.
- **B. Snowpark UDFs can be defined as either scalar UDFs (processing one row at a time) or vectorized UDFs (processing batches of rows), offering different performance characteristics.**
- C. Snowpark Python UDFs are always executed in a single process on the Snowflake warehouse, limiting their parallel processing capabilities.
- D. Snowpark UDFs can be defined using either Python or Java, providing flexibility in choosing the programming language best suited for the task. The Java UDF creation method will allow faster execution speeds.
- **E. Snowpark UDFs automatically distribute the data and computation across multiple nodes in the Snowflake warehouse, but the distribution strategy cannot be controlled by the developer.**

Answer: B,E

Explanation:

Snowpark UDFs can be either scalar or vectorized, offering different performance tradeoffs. Vectorized UDFs are generally more efficient for large datasets as they process batches of rows. Snowpark UDFs do distribute the data and computation across multiple nodes automatically; however, the distribution strategy, while not directly controlled, is influenced by how the UDF is applied to the data and the inherent distribution of the underlying data itself. Python is the primary UDF language. Option A is false because UDFs are designed for parallel processing. Option C is not always true; custom channels might be necessary for specific dependencies. Option E is partially correct in the older releases but Python is used primarily now.

NEW QUESTION # 215

You are tasked with setting up secure authentication for your Snowpark application. You want to use key pair authentication for a service user. Which of the following steps are necessary and in the correct order?

- A. 1. Generate an RSA key pair (private and public key). 2. Store the public key securely on the client machine. 3. Provide the path to the public key file in the Snowpark session configuration. 4. Associate the private key with the Snowflake user using the `SALTER USER` command.
- B. 1. Generate an RSA key pair (private and public key). 2. Store the private key securely on the client machine. 3. Provide the path to the private key file in the
- C. 1. Generate an RSA key pair (private and public key). 2. Store the private key in a database table. 3. Use database credentials in the Snowpark session configuration. 4. Associate the public key with the Snowflake user using the `'ALTER USER` command.
- D. 1. Generate an RSA key pair (private and public key). 2. Store the private key securely on the client machine. 3. Associate the private key with the Snowflake user using the `SALTER USER` command. 4. Provide the public key in the Snowpark session configuration.
- **E. 1. Generate an RSA key pair (private and public key). 2. Store the private key securely on the client machine. 3. Provide the path to the private key file and passphrase (if any) in the Snowpark session configuration. 4. Associate the public key with the Snowflake user using the `'ALTER USER` command.**

Answer: E

Explanation:

The correct steps for key pair authentication involve generating an RSA key pair, securely storing the private key on the client, providing the path to the private key (and passphrase, if used) in the Snowpark session configuration, and associating the public key with the Snowflake user using the `'ALTER USER` command. Storing the private key in the database (option E) is a security risk. Options B and C have incorrect key associations.

NEW QUESTION # 216

.....

You choosing TestSimulate to help you pass Snowflake certification SPS-C01 exam is a wise choice. You can first online free

myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, Disposable vapes

What's more, part of that TestSimulate SPS-C01 dumps now are free: https://drive.google.com/open?id=1o9YmC1q_T1hb0UXMrckFZ02IGSz1233