

SPS-C01合格資料 & SPS-C01資格講座



Snowflake SPS-C01

SnowPro Specialty: Snowpark Certification Exam

Questions & Answers PDF
(Demo Version Limited Content)

For More Information Visit link below:

<https://p2pexam.com/>

Visit us at: <https://p2pexam.com/spc-c01>

ちなみに、Fast2test SPS-C01の一部をクラウドストレージからダウンロードできます：<https://drive.google.com/open?id=1ueR-iRP82Us-Ra16dDYYjY8IU29EbEFn>

Snowflakeお客様との持続可能な関係に高い価値を置いているため、SPS-C01準備ガイドのヘルプの下で最高の証明書学習体験をお楽しみいただけます。まず、5~10分でお支払いが完了すると、短納期で、オンラインでSPS-C01ガイドトレントをお送りします。加えて、当社のSPS-C01試験トレントの使用中に技術的および運用上の問題に対処するのに問題がある場合は、すぐにご連絡ください。24時間のオンラインサービスは、Snowflake Certified SnowPro Specialty - Snowpark問題をすぐに解決するための努力です。

長年のマーケティングを通じて、当社のSPS-C01最新の認定ガイドは多くのお客様のサポートを獲得しています。最も明白なデータは、当社の製品が毎年徐々に増加していることであり、当社の製品開発のおかげでこのような大きな成功を達成するための大きな努力です。まず、資料の更新を研究する上で非常に良い仕事をしました。さらに、SPS-C01の実際のSPS-C01学習ガイド教材の品質は、教師によって厳密に管理されています。だから、私たちは正しい選択だと信じています。SPS-C01学習教材について質問がある場合は、ご相談ください。

>> SPS-C01合格資料 <<

効果的SPS-C01 | 高品質なSPS-C01合格資料試験 | 試験の準備方法 Snowflake Certified SnowPro Specialty - Snowpark資格講座

Fast2testのSnowflakeのSPS-C01試験トレーニング資料を手に入れるなら、君が他の人の一半の努力で、同じSnowflakeのSPS-C01認定試験を簡単に合格できます。あなたはFast2testのSnowflakeのSPS-C01問題集を購入した

後、私たちは一年間で無料更新サービスを提供することができます。もしうちのSnowflakeのSPS-C01問題集は問題があれば、或いは試験に不合格になる場合は、全額返金することを保証いたします。

Snowflake Certified SnowPro Specialty - Snowpark 認定 SPS-C01 試験問題 (Q77-Q82):

質問 # 77

You are developing a Snowpark application in Python to perform sentiment analysis on customer reviews stored in a Snowflake table named 'CUSTOMER_REVIEWS'. The table has columns 'REVIEW ONT), 'REVIEW TEXT (VARCHAR), and 'SENTIMENT SCORE (FLOAT). You want to define a UDF using Snowpark that leverages a pre-trained sentiment analysis model from the 'nltk' library (already uploaded to a stage). The UDF should take 'REVIEW TEXT' as input and return the sentiment score. Which of the following code snippets will correctly define and register the UDF, ensuring it's accessible for use in Snowpark DataFrames, taking into account potential serialization issues with 'nltk' models?

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

正解: B

解説:

Option E is correct because it utilizes the '@udf decorator combined with to ensure the 'nltk' library is available within the UDF's execution environment. Importantly, the analyzer is initialized within the function to avoid serialization issues, and all necessary imports are present, including specifying the data types. The nltk import is included inside the function due to the nature of the UDF and the package import. Option A is incorrect because it does not address the dependency on 'nltk' within the Snowflake environment. Option B is incorrect since the @udf decorator is not used correctly and doesn't load the dependencies correctly, and does not explicitly state the Snowflake data types. Option C is incorrect as it uses 'session.add_import' which is deprecated and not the recommended way to add packages to the session, packages option is the recommended method. Option D is incorrect since it does not explicitly state the Snowflake data types, and has the udf.register which is not a decorator, and also not a good approach.

質問 # 78

You need to perform a set difference operation between two DataFrames in Snowpark Python. 'df1' contains customer IDs from a marketing campaign, and 'df2' contains customer IDs from a recent purchase event. You want to identify customers who were targeted in the campaign but did not make a recent purchase. Both DataFrames have a column named 'customer id'. Which of the following approaches provides the most efficient way to accomplish this task in Snowpark?

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

正解: A

解説:

Option C, using a 'left_anti' join, is the most efficient way to perform a set difference operation between two DataFrames in Snowpark. A join returns only the rows from the left DataFrame (df1) where the join condition is not met in the right DataFrame (df2). This leverages Snowflake's query optimizer for optimal performance. Option A, 'subtract(df2)', is equivalent to 'exceptAll(df2)' (Option B) and removes duplicate rows. While functionally correct, join is often more performant, especially for larger datasets. Option D is highly inefficient as it collects the 'customer_id' from 'df2' to the driver, it should be avoided. Option E calculates intersection, not difference.

質問 # 79

You have a Snowpark application that utilizes a vectorized Python UDF to perform complex calculations on a large dataset. You notice that the performance is still not optimal. You suspect that the bottleneck might be related to how the data is being partitioned and processed by Snowflake. Which of the following actions, when performed in conjunction with vectorization, would MOST likely

improve performance?

- A. Repartition the Snowpark DataFrame using to align the data distribution with the computational needs of the UDF.
- B. Ensure that the data is pre-sorted according to the primary key of the table before applying the UDF.
- C. Convert the DataFrame to a Pandas DataFrame before applying the UDF.
- D. Broadcast the DataFrame to all compute nodes before applying the UDF.
- E. Increase the number of UDF worker threads within the UDF definition.

正解: A

解説:

Repartitioning the DataFrame using allows you to control how the data is distributed across compute nodes. This can improve performance by ensuring that related data is processed together, reducing data shuffling and improving data locality. Pre-sorting data (A) might help in some cases, but it doesn't guarantee optimal data distribution for parallel processing. Broadcasting the DataFrame (C) is suitable for smaller datasets, not large ones where it can lead to memory issues. Converting the DataFrame to a Pandas DataFrame (D) defeats the purpose of using Snowpark for distributed processing and introduces a single-node bottleneck. There's no direct control over the number of UDF worker threads in Snowflake.

質問 # 80

You have a Snowpark DataFrame named containing daily sales transactions. The DataFrame includes columns like 'transaction_id', 'product_id', 'sale_date', and 'sale_amount'. You need to perform the following transformations and persist the results: (1) Calculate the total sales amount for each product on a daily basis. (2) Store the aggregated results into a new table named , partitioned by 'sale_date'. (3) Ensure that if the table already exists, the new data is appended to the existing table. Which of the following code blocks achieve these requirements in the most efficient and correct manner?

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

正解: B

解説:

Option A is the most efficient and correct. It first aggregates the data as required, then uses 'mode('append')' to add new data without overwriting existing data, and for partitioning. Option B will overwrite existing data. Option C is equivalent to option A, but is a less common coding style. Option D renames the default aggregate column name which is less readable but also works correctly. Option E, 'mergeSchema' isn't a valid option to be passed.

質問 # 81

A data scientist has developed a Snowpark Python stored procedure named 'model_training'. This procedure utilizes a large machine learning model and requires significant compute resources. The data scientist wants to optimize the cost and performance of running this stored procedure. Which of the following strategies would be the MOST effective for achieving this goal?

- A. Convert the Python stored procedure to a SQL stored procedure to leverage Snowflake's SQL optimization engine.
- B. Split the stored procedure into multiple smaller procedures and execute them sequentially on a smaller warehouse.
- C. Register the stored procedure with the '@sproc' decorator without specifying any warehouse size, letting Snowflake automatically manage the warehouse.
- D. Run the stored procedure on a larger Snowflake warehouse to reduce execution time, regardless of potential idle time.
- E. Specify a warehouse size using the 'warehouse' parameter within the '@sproc' decorator and leverage auto-suspend and auto-resume features to minimize costs when the procedure is idle.

正解: E

解説:

Specifying a warehouse size and using auto-suspend and auto-resume provides a balance between performance and cost. Option A might increase costs due to idle time. Option B relies on Snowflake's default warehouse, which might not be optimal. Option D could increase overall execution time due to overhead. Option E may not be feasible or efficient if the logic is heavily dependent on Python libraries. Therefore, specifically assigning an appropriate warehouse size with auto-suspend/resume is the most effective approach.

myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt,
myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt,
ronaldtxkh223891.bloggactivo.com, shaunaygjk026341.blogoxo.com, Disposable vapes

P.S. Fast2testがGoogle Driveで共有している無料かつ新しいSPS-C01ダンプ: <https://drive.google.com/open?id=1ueR-iRP82Us-Ral6dDYYjY8IU29EbEFn>