

Databricks-Certified-Data-Analyst-Associate認證 - Databricks-Certified-Data-Analyst-Associate考試內容



Fast2test為Databricks Databricks-Certified-Data-Analyst-Associate 認證考試準備的培訓包括Databricks Databricks-Certified-Data-Analyst-Associate認證考試的模擬測試題和當前的考試真題。在互聯網上你也可以看到幾個也提供相關的培訓的網站，但是你比較之後，你就會發現Fast2test的關於Databricks Databricks-Certified-Data-Analyst-Associate 認證考試的培訓比較有針對性，不僅品質是最高的，而且內容是最全面的。

Databricks Databricks-Certified-Data-Analyst-Associate 考試大綱：

主題	簡介
主題 1	Analytics applications: It describes key moments of statistical distributions, data enhancement, and the blending of data between two source applications. Moroeover, the topic also explains last-mile ETL, a scenario in which data blending would be beneficial, key statistical measures, descriptive statistics, and discrete and continuous statistics.
主題 2	SQL in the Lakehouse: It identifies a query that retrieves data from the database, the output of a SELECT query, a benefit of having ANSI SQL, access, and clean silver-level data. It also compares and contrasts MERGE INTO, INSERT TABLE, and COPY INTO. Lastly, this topic focuses on creating and applying UDFs in common scaling scenarios.

主題 3	• Databricks SQL: This topic discusses key and side audiences, users, Databricks SQL benefits, complementing a basic Databricks SQL query, schema browser, Databricks SQL dashboards, and the purpose of Databricks SQL endpoints • warehouses. Furthermore, the delves into Serverless Databricks SQL endpoint • warehouses, trade-off between cluster size and cost for Databricks SQL endpoints • warehouses, and Partner Connect. Lastly it discusses small-file upload, connecting Databricks SQL to visualization tools, the medallion architecture, the gold layer, and the benefits of working with streaming data.
主題 4	• Data Visualization and Dashboarding: Sub-topics of this topic are about of describing how notifications are sent, how to configure and troubleshoot a basic alert, how to configure a refresh schedule, the pros and cons of sharing dashboards, how query parameters change the output, and how to change the colors of all of the visualizations. It also discusses customized data visualizations, visualization formatting, Query Based Dropdown List, and the method for sharing a dashboard.
主題 5	• Data Management: The topic describes Delta Lake as a tool for managing data files, Delta Lake manages table metadata, benefits of Delta Lake within the Lakehouse, tables on Databricks, a table owner's responsibilities, and the persistence of data. It also identifies management of a table, usage of Data Explorer by a table owner, and organization-specific considerations of PII data. Lastly, the topic it explains how the LOCATION keyword changes, usage of Data Explorer to secure data.

>> Databricks-Certified-Data-Analyst-Associate認證 <<

Databricks-Certified-Data-Analyst-Associate考試內容，Databricks-Certified-Data-Analyst-Associate題庫更新資訊

對於為了進入大中型IT公司的畢業生來說，不想花費太多的錢去補習，只能借助於最新的 Fast2test Databricks-Certified-Data-Analyst-Associate 考題准備考試，因為這類考題在網路中的價錢不是很高。我們會持續不斷從世界各地使用者在他們的地區參加 Databricks 考試而獲得回饋，輕鬆的了解最新考試資訊，及時的更新 Databricks-Certified-Data-Analyst-Associate 題庫，這使我們的其他用戶可以共用這些考試資訊，參加考試更有信心！這是通過 Databricks-Certified-Data-Analyst-Associate 考試最有效的方法之一。

最新的 Data Analyst Databricks-Certified-Data-Analyst-Associate 免費考試真題 (Q61-Q66):

問題 #61

A data analyst is attempting to drop a table my_table. The analyst wants to delete all table metadata and data.

They run the following command:

```
DROP TABLE IF EXISTS my_table;
```

While the object no longer appears when they run SHOW TABLES, the data files still exist.

Which of the following describes why the data files still exist and the metadata files were deleted?

- A. The table was external
- B. The table's data was larger than 10 GB
- C. The table was managed
- D. The table did not have a location
- E. The table's data was smaller than 10 GB

答案：A

解題說明：

An external table is a table that is defined in the metastore, but its data is stored outside of the Databricks environment, such as in S3, ADLS, or GCS. When an external table is dropped, only the metadata is deleted from the metastore, but the data files are not affected. This is different from a managed table, which is a table whose data is stored in the Databricks environment, and whose data files are deleted when the table is dropped. To delete the data files of an external table, the analyst needs to specify the PURGE option in the DROP TABLE command, or manually delete the files from the storage system. Reference: DROP TABLE, Drop Delta table features, Best practices for dropping a managed Delta Lake table

問題 #62

A data team has been given a series of projects by a consultant that need to be implemented in the Databricks Lakehouse Platform. Which of the following projects should be completed in Databricks SQL?

- A. Automating complex notebook-based workflows with multiple tasks
- **B. Combining two data sources into a single, comprehensive dataset**
- C. Segmenting customers into like groups using a clustering algorithm
- D. Testing the quality of data as it is imported from a source
- E. Tracking usage of feature variables for machine learning projects

答案: B

解題說明:

Databricks SQL is a service that allows users to query data in the lakehouse using SQL and create visualizations and dashboards¹. One of the common use cases for Databricks SQL is to combine data from different sources and formats into a single, comprehensive dataset that can be used for further analysis or reporting². For example, a data analyst can use Databricks SQL to join data from a CSV file and a Parquet file, or from a Delta table and a JDBC table, and create a new table or view that contains the combined data³. This can help simplify the data management and governance, as well as improve the data quality and consistency. Reference:

Databricks SQL overview

Databricks SQL use cases

Joining data sources

問題 #63

A data analyst runs the following command:

```
INSERT INTO stakeholders.suppliers TABLE stakeholders.new_suppliers;
```

What is the result of running this command?

- A. The suppliers table now contains only the data from the new suppliers table.
- **B. The command fails because it is written incorrectly.**
- C. The suppliers table now contains both the data it had before the command was run and the data from the new suppliers table, including any duplicate data.
- D. The suppliers table now contains the data from the new suppliers table, and the new suppliers table now contains the data from the suppliers table.
- E. The suppliers table now contains both the data it had before the command was run and the data from the new suppliers table, and any duplicate data is deleted.

答案: B

解題說明:

The command `INSERT INTO stakeholders.suppliers TABLE stakeholders.new_suppliers` is not a valid syntax for inserting data into a table in Databricks SQL. According to the documentation^{1,2}, the correct syntax for inserting data into a table is either: `INSERT { OVERWRITE | INTO } [ TABLE ] table_name [ PARTITION clause ] [ ( column_name [, ...] ) | BY NAME ] query` or `INSERT INTO [ TABLE ] table_name REPLACE WHERE predicate query`. The command in the question is missing the `OVERWRITE` or `INTO` keyword, and the query part that specifies the source of the data to be inserted. The `TABLE` keyword is optional and can be omitted. The `PARTITION` clause and the column list are also optional and depend on the table schema and the data source. Therefore, the command in the question will fail with a syntax error.

Reference:

INSERT | Databricks on AWS

INSERT - Azure Databricks - Databricks SQL | Microsoft Learn

問題 #64

A stakeholder has provided a data analyst with a lookup dataset in the form of a 50-row CSV file. The data analyst needs to upload this dataset for use as a table in Databricks SQL.

Which approach should the data analyst use to quickly upload the file into a table for use in Databricks SQL?

- A. Create a table by manually copying and pasting the data values into cloud storage and then importing the data to

Databricks.

- B. Create a table by uploading the file using the Create page within Databricks SQL
- C. Create a table via a connection between Databricks and the desktop facilitated by Partner Connect.
- D. Create a table by uploading the file to cloud storage and then importing the data to Databricks.

答案： B

解題說明：

Databricks provides a user-friendly interface that allows data analysts to quickly upload small datasets, such as a 50-row CSV file, and create tables within Databricks SQL. The steps are as follows:

Access the Data Upload Interface:

In the Databricks workspace, navigate to the sidebar and click on New > Add or upload data.

Select Create or modify a table.

Upload the CSV File:

Click on the browse button or drag and drop the CSV file directly onto the designated area.

The interface supports uploading up to 10 files simultaneously, with a total size limit of 2 GB.

Configure Table Settings:

After uploading, a preview of the data is displayed.

Specify the table name, select the appropriate schema, and configure any additional settings as needed.

Create the Table:

Once all configurations are set, click on the Create Table button to finalize the process.

This method is efficient for quickly importing small datasets without the need for additional tools or complex configurations. Options B, C, and D involve more complex or manual processes that are unnecessary for this task.

問題 #65

What describes the variance of a set of values?

- A. Variance is a measure of how far a set of values is spread out from the sets central value.
- B. Variance is a measure of how far an observed value is from the variable's maximum or minimum value.
- C. Variance is a measure of how far a single observed value is from a set of values.
- D. Variance is a measure of central tendency of a set of values.

答案： A

解題說明：

Variance is a statistical measure that quantifies the dispersion or spread of a set of values around their mean (central value). It is calculated by taking the average of the squared differences between each value and the mean of the dataset. A higher variance indicates that the data points are more spread out from the mean, while a lower variance suggests that they are closer to the mean. This measure is fundamental in statistics to understand the degree of variability within a dataset. Wikipedia Wikipedia+1 Investopedia+1

問題 #66

.....

Fast2test為你提供了一個明確而優秀的選擇，為你減少煩惱。想早點成功嗎？早點拿到Databricks Databricks-Certified-Data-Analyst-Associate認證考試的證書嗎？快點將Fast2test加入購物車吧。Fast2test會給你很好的指導，能確保你通過考試。使用Fast2test你可以很快獲得你想要的證書。

Databricks-Certified-Data-Analyst-Associate考試內容： <https://tw.fast2test.com/Databricks-Certified-Data-Analyst-Associate-premium-file.html>

- Databricks-Certified-Data-Analyst-Associate考題套裝 □ Databricks-Certified-Data-Analyst-Associate權威考題 □ Databricks-Certified-Data-Analyst-Associate更新 □ 透過 www.newdumps.pdf.com □ 搜索 ☼ Databricks-Certified-Data-Analyst-Associate □ ☼ □ 免費下載考試資料 Databricks-Certified-Data-Analyst-Associate 證照
- Databricks-Certified-Data-Analyst-Associate考試資訊 □ Databricks-Certified-Data-Analyst-Associate題庫資訊 □ Databricks-Certified-Data-Analyst-Associate考試心得 □ 在 **【 www.newdumps.pdf.com 】** 網站上查找 ☼ Databricks-Certified-Data-Analyst-Associate □ ☼ □ 的最新題庫 Databricks-Certified-Data-Analyst-Associate 認證題庫
- 下載 Databricks-Certified-Data-Analyst-Associate 認證表示您已在通過 Databricks Certified Data Analyst Associate Exam 的路上 □ 在 { www.pdfexamdumps.com } 搜索最新的 《 Databricks-Certified-Data-Analyst-Associate 》 題庫最

Databricks Databricks-Certified-Data-Analyst-Associate認證和Newdumpsdf- 資格考試的領導者和Databricks-Certified-Data-Analyst-Associate考試內容 在 (www.newdumpsdf.com) 網站上查找【 Databricks-Certified-Data-Analyst-Associate 】的最新題庫Databricks-Certified-Data-Analyst-Associate題庫資訊

- [illegible]