

Databricks-Certified-Professional-Data-Engineer Learning Materials: Databricks Certified Professional Data Engineer Exam & Databricks-Certified- Professional-Data-Engineer Test Braindumps



In order to have better life, attending certification exams and obtaining certifications will be essential on the path to success. Databricks-Certified-Professional-Data-Engineer latest test cram sheet will help you achieve your goal. Only if you receive the certificate the companies require you can have the opportunities for raising-salary and promotion. Thousands of companies think highly of this certification. You will be popular if you pass exam with Databricks-Certified-Professional-Data-Engineer Latest Test Cram sheet.

Passing the Databricks Certified Professional Data Engineer exam can be a significant achievement for individuals pursuing a career in big data and cloud computing. Databricks Certified Professional Data Engineer Exam certification demonstrates that the candidate has a deep understanding of Apache Spark and can apply that knowledge to design and implement scalable big data solutions. Additionally, this certification is recognized by many companies in the industry and can improve the candidate's chances of obtaining a high-paying job.

Databricks Certified Professional Data Engineer Exam is a comprehensive exam that covers a wide range of topics related to data engineering. These topics include data modeling, data ingestion, data integration, data transformation, data storage, and data analysis. Candidates are required to demonstrate their knowledge and skills in these areas by completing a series of tasks and exercises.

>> Databricks-Certified-Professional-Data-Engineer Simulation Questions <<

Free PDF Quiz 2025 Databricks-Certified-Professional-Data-Engineer: Databricks Certified Professional Data Engineer Exam – Efficient Simulation Questions

It is a truism that an internationally recognized Databricks-Certified-Professional-Data-Engineer certification can totally mean you have a good command of the knowledge in certain areas and showcase your capacity to a considerable extend. If you are overwhelmed by workload heavily and cannot take a breath from it, why not choose our Databricks-Certified-Professional-Data-Engineer Preparation torrent? We are specialized in providing our customers with the most reliable and accurate exam materials and help them pass their exams by achieve their satisfied scores. With our Databricks-Certified-Professional-Data-Engineer practice materials, your exam will be a piece of cake.

Databricks Certified Professional Data Engineer Exam Sample Questions (Q61-Q66):

NEW QUESTION # 61

The view updates represents an incremental batch of all newly ingested data to be inserted or updated in the customers table. The following logic is used to process these records.

```

MERGE INTO customers
USING (
SELECT updates.customer_id as merge_key, updates.*
FROM updates
UNION ALL
SELECT NULL as merge_key, updates.*
FROM updates JOIN customers
ON updates.customer_id = customers.customer_id
WHERE customers.current = true AND updates.address <> customers.address ) staged_updates ON customers.customer_id =
mergekey WHEN MATCHED AND customers.current = true AND customers.address <> staged_updates.address THEN
UPDATE SET current = false, end_date = staged_updates.effective_date WHEN NOT MATCHED THEN INSERT
(customer_id, address, current, effective_date, end_date) VALUES (staged_updates.customer_id, staged_updates.address, true,
staged_updates.effective_date, null) Which statement describes this implementation?

```

- A. The customers table is implemented as a Type 0 table; all writes are append only with no changes to existing values.
- B. The customers table is implemented as a Type 1 table; old values are overwritten by new values and no history is maintained.
- **C. The customers table is implemented as a Type 2 table; old values are maintained but marked as no longer current and new values are inserted.**
- D. The customers table is implemented as a Type 2 table; old values are overwritten and new customers are appended.

Answer: C

Explanation:

The provided MERGE statement is a classic implementation of a Type 2 SCD in a data warehousing context.

In this approach, historical data is preserved by keeping old records (marking them as not current) and adding new records for changes. Specifically, when a match is found and there's a change in the address, the existing record in the customers table is updated to mark it as no longer current (`current = false`), and an end date is assigned (`end_date = staged_updates.effective_date`). A new record for the customer is then inserted with the updated information, marked as current. This method ensures that the full history of changes to customer information is maintained in the table, allowing for time-based analysis of customer data. References: Databricks documentation on implementing SCDs using Delta Lake and the MERGE statement (<https://docs.databricks.com/delta/delta-update.html#upsert-into-a-table-using-merge>).

NEW QUESTION # 62

A data architect has determined that a table of the following format is necessary:

Which of the following code blocks uses SQL DDL commands to create an empty Delta table in the above format regardless of whether a table already exists with this name?

- A. 1. CREATE TABLE table_name AS
2. SELECT id STRING, birthDate DATE, avgRating FLOAT
- **B. 1. CREATE OR REPLACE TABLE table_name (id STRING, birthDate DATE, avgRating FLOAT)**
- C. 1. CREATE OR REPLACE TABLE table_name AS
2. SELECT id STRING, birthDate DATE, avgRating FLOAT USING DELTA
- D. 1. CREATE OR REPLACE TABLE table_name
2. WITH COLUMNS (id STRING, birthDate DATE, avgRating FLOAT) USING DELTA
- E. 1. CREATE TABLE IF NOT EXISTS table_name (id STRING, birthDate DATE, avgRating FLOAT)

Answer: B

NEW QUESTION # 63

Which method can be used to determine the total wall-clock time it took to execute a query?

- A. In the Spark UI, take the job duration of the longest-running job associated with that query.
- B. In the Spark UI, take the sum of all task durations that ran across all stages for all jobs associated with that query.
- **C. Open the Query Profiler associated with that query and use the Total wall-clock duration metric.**
- D. Open the Query Profiler associated with that query and use the Aggregated task time metric.

Answer: C

Explanation:

Comprehensive and Detailed Explanation From Exact Extract of Databricks Data Engineer Documents:

The Query Profiler in Databricks SQL and notebooks provides a detailed breakdown of query performance metrics. The "Total wall-clock duration" metric directly represents the total elapsed time from query start to completion, including all execution, planning, and waiting stages. In contrast, "Aggregated task time" reflects the cumulative duration across all parallel tasks, which does not equal the total elapsed wall time since tasks often run concurrently. Using job duration from Spark UI can underestimate or overestimate runtime when queries span multiple jobs. Therefore, the Query Profiler's total wall-clock duration is the officially documented method to determine actual query execution time.

NEW QUESTION # 64

Which statement describes Delta Lake Auto Compaction?

- A. An asynchronous job runs after the write completes to detect if files could be further compacted; if yes, an optimize job is executed toward a default of 1 GB.
- B. Before a Jobs cluster terminates, optimize is executed on all tables modified during the most recent job.
- C. An asynchronous job runs after the write completes to detect if files could be further compacted; if yes, an optimize job is executed toward a default of 128 MB.
- D. Data is queued in a messaging bus instead of committing data directly to memory; all data is committed from the messaging bus in one batch once the job is complete.
- E. Optimized writes use logical partitions instead of directory partitions; because partition boundaries are only represented in metadata, fewer small files are written.

Answer: C

Explanation:

Explanation

This is the correct answer because it describes the behavior of Delta Lake Auto Compaction, which is a feature that automatically optimizes the layout of Delta Lake tables by coalescing small files into larger ones. Auto Compaction runs as an asynchronous job after a write to a table has succeeded and checks if files within a partition can be further compacted. If yes, it runs an optimize job with a default target file size of 128 MB.

Auto Compaction only compacts files that have not been compacted previously. Verified References:

[Databricks Certified Data Engineer Professional], under "Delta Lake" section; Databricks Documentation, under "Auto Compaction for Delta Lake on Databricks" section.

"Auto compaction occurs after a write to a table has succeeded and runs synchronously on the cluster that has performed the write.

Auto compaction only compacts files that haven't been compacted previously."

<https://learn.microsoft.com/en-us/azure/databricks/delta/tune-file-size>

NEW QUESTION # 65

A junior data engineer seeks to leverage Delta Lake's Change Data Feed functionality to create a Type 1 table representing all of the values that have ever been valid for all rows in a bronze table created with the property `delta.enableChangeDataFeed = true`. They plan to execute the following code as a daily job:

```
from pyspark.sql.functions import col

(spark.read.format("delta")
 .option("readChangeFeed", "true")
 .option("startingVersion", 0)
 .table("bronze")
 .filter(col("_change_type").isin(["update_postimage", "insert"]))
 .write
 .mode("append")
 .table("bronze_history_type1")
 )
```

Which statement describes the execution and results of running the above query multiple times?

- A. Each time the job is executed, newly updated records will be merged into the target table, overwriting previous values with the same primary keys.
- B. Each time the job is executed, the entire available history of inserted or updated records will be appended to the target table, resulting in many duplicate entries.
- C. Each time the job is executed, the differences between the original and current versions are calculated; this may result in duplicate entries for some records.
- **D. Each time the job is executed, the target table will be overwritten using the entire history of inserted or updated records, giving the desired result.**
- E. Each time the job is executed, only those records that have been inserted or updated since the last execution will be appended to the target table giving the desired result.

Answer: D

Explanation:

Explanation

This is the correct answer because it describes the execution and results of running the above query multiple times. The query uses the readChanges function to read all change events from a bronze table that has enabled change data feed. The readChanges function takes two arguments: version and options. The version argument specifies which version of the table to read changes from, and can be either a specific version number or -1 to indicate all versions. The options argument specifies additional options for reading changes, such as whether to include deletes or not. In this case, the query reads all changes from all versions of the bronze table and filters out delete events by setting includeDeletes to false. Then, it uses write.format("delta").mode("overwrite") to overwrite a target table using the entire history of inserted or updated records, giving the desired result of a Type 1 table representing all values that have ever been valid for all rows in the bronze table. Verified References: [Databricks Certified Data Engineer Professional], under

"Delta Lake" section; Databricks Documentation, under "Read changes in batch queries" section.

NEW QUESTION # 66

.....

Are you still overwhelmed by the low-production and low-efficiency in your daily life? If your answer is yes, please pay attention to our Databricks-Certified-Professional-Data-Engineer guide torrent, because we will provide well-rounded and first-tier services for you, thus supporting you obtain your dreamed Databricks-Certified-Professional-Data-Engineer certificate and have a desired occupation. There are some main features of our products and we believe you will be satisfied with our Databricks-Certified-Professional-Data-Engineer test questions. And once you have a try on our Databricks-Certified-Professional-Data-Engineer exam questions, you will love it.

Databricks-Certified-Professional-Data-Engineer Positive Feedback: <https://www.getvalidtest.com/Databricks-Certified-Professional-Data-Engineer-exam.html>

- Databricks-Certified-Professional-Data-Engineer Reliable Exam Guide □ Best Databricks-Certified-Professional-Data-Engineer Preparation Materials □ Exam Databricks-Certified-Professional-Data-Engineer Simulator □ Copy URL 【 www.examdiscuss.com 】 open and search for □ Databricks-Certified-Professional-Data-Engineer □ to download for free □ Best Databricks-Certified-Professional-Data-Engineer Preparation Materials
- Wonderful Databricks-Certified-Professional-Data-Engineer Learning Questions: Databricks Certified Professional Data Engineer Exam are form the latest Exam Brain Dumps - Pdfvce □ Download ► Databricks-Certified-Professional-Data-Engineer ◀ for free by simply entering [www.pdfvce.com] website □ Best Databricks-Certified-Professional-Data-Engineer Preparation Materials
- Test Databricks-Certified-Professional-Data-Engineer Tutorials □ Databricks-Certified-Professional-Data-Engineer Exam Answers □ New Databricks-Certified-Professional-Data-Engineer Test Tips □ Simply search for 【 Databricks-Certified-Professional-Data-Engineer 】 for free download on ☼ www.passtestking.com □ ☼ □ New Databricks-Certified-Professional-Data-Engineer Test Tips
- Databricks-Certified-Professional-Data-Engineer Test Dumps Pdf □ Reliable Databricks-Certified-Professional-Data-Engineer Exam Prep □ Best Databricks-Certified-Professional-Data-Engineer Preparation Materials □ Download 「 Databricks-Certified-Professional-Data-Engineer 」 for free by simply entering 《 www.pdfvce.com 》 website □ □ Updated Databricks-Certified-Professional-Data-Engineer CBT
- 100% Pass Quiz Trustable Databricks - Databricks-Certified-Professional-Data-Engineer - Databricks Certified Professional Data Engineer Exam Simulation Questions □ Search for ► Databricks-Certified-Professional-Data-Engineer ◄ and download it for free immediately on □ www.examsreviews.com □ □ Databricks-Certified-Professional-Data-Engineer Study Material
- Databricks Databricks-Certified-Professional-Data-Engineer Exam is Easy with Our Verified Databricks-Certified-

Professional-Data-Engineer Simulation Questions: Databricks Certified Professional Data Engineer Exam ☐ Open 
www.pdfvce.com ☐ enter (Databricks-Certified-Professional-Data-Engineer) and obtain a free download ☐
☐ Databricks-Certified-Professional-Data-Engineer Reliable Exam Guide