

First-hand Databricks Databricks-Certified-Professional-Data-Engineer Valid Test Voucher - Databricks-Certified-Professional-Data-Engineer Test Databricks Certified Professional Data Engineer Exam Dumps



The Databricks Certified Professional Data Engineer Exam (Databricks-Certified-Professional-Data-Engineer) practice exam consists of a Databricks Certified Professional Data Engineer Exam (Databricks-Certified-Professional-Data-Engineer) PDF dumps format, Desktop-based Databricks-Certified-Professional-Data-Engineer practice test software and a Web-based Databricks Certified Professional Data Engineer Exam (Databricks-Certified-Professional-Data-Engineer) practice exam. Each of the ActualtestPDF Databricks Databricks-Certified-Professional-Data-Engineer Exam Dumps formats excels in its way and carries actual Databricks Certified Professional Data Engineer Exam (Databricks-Certified-Professional-Data-Engineer) exam questions for optimal preparation.

Databricks Certified Professional Data Engineer certification exam is intended for data engineers, data architects, and other IT professionals who work with big data technologies. Databricks-Certified-Professional-Data-Engineer Exam covers a wide range of topics, including data ingestion, data transformation, data storage, and data analysis. It also covers the use of Databricks tools and technologies such as Databricks Delta, Databricks Runtime, and Apache Spark.

>> **Databricks-Certified-Professional-Data-Engineer Valid Test Voucher <<**

Test Databricks-Certified-Professional-Data-Engineer Dumps - Certification Databricks-Certified-Professional-Data-Engineer Dumps

By adhering to the principle of “quality first, customer foremost”, and “mutual development and benefit”, our company will provide first class service for our customers. The exam prepare materials of ActualtestPDF is high quality and high pass rate, it is completed by our experts who have a good understanding of Real Databricks-Certified-Professional-Data-Engineer Exam and have many years of experience writing study materials. They know very well what candidates really need most when they prepare for the Databricks-Certified-Professional-Data-Engineer exam. They also understand the real exam situation very well. We will let you know what a real exam is like.

Databricks Certified Professional Data Engineer Exam Sample Questions (Q203-Q208):

NEW QUESTION # 203

A data engineer wants to create a relational object by pulling data from two tables. The relational object must be used by other data engineers in other sessions. In order to save on storage costs, the data engineer wants to avoid copying and storing physical data.

Which of the following relational objects should the data engineer create?

- A. Database
- B. Temporary view
- C. Delta Table

- D. Spark SQL Table
- E. View

Answer: E

NEW QUESTION # 204

A data architect has designed a system in which two Structured Streaming jobs will concurrently write to a single bronze Delta table. Each job is subscribing to a different topic from an Apache Kafka source, but they will write data with the same schema. To keep the directory structure simple, a data engineer has decided to nest a checkpoint directory to be shared by both streams.

The proposed directory structure is displayed below:

Which statement describes whether this checkpoint directory structure is valid for the given scenario and why?

- A. No; each of the streams needs to have its own checkpoint directory.
- B. Yes; Delta Lake supports infinite concurrent writers.
- C. Yes; both of the streams can share a single checkpoint directory.
- D. No; only one stream can write to a Delta Lake table.
- E. No; Delta Lake manages streaming checkpoints in the transaction log.

Answer: A

Explanation:

This is the correct answer because checkpointing is a critical feature of Structured Streaming that provides fault tolerance and recovery in case of failures. Checkpointing stores the current state and progress of a streaming query in a reliable storage system, such as DBFS or S3. Each streaming query must have its own checkpoint directory that is unique and exclusive to that query. If two streaming queries share the same checkpoint directory, they will interfere with each other and cause unexpected errors or data loss. Verified Reference: [Databricks Certified Data Engineer Professional], under "Structured Streaming" section; Databricks Documentation, under "Checkpointing" section.

NEW QUESTION # 205

A task orchestrator has been configured to run two hourly tasks. First, an outside system writes Parquet data to a directory mounted at /mnt/raw_orders/. After this data is written, a Databricks job containing the following code is executed:

```
(spark.readStream
  .format("parquet")
  .load("/mnt/raw_orders/")
  .withWatermark("time", "2 hours")
  .dropDuplicates(["customer_id", "order_id"])
  .writeStream
  .trigger(once=True)
  .table("orders")
)
```

Assume that the fields customer_id and order_id serve as a composite key to uniquely identify each order, and that the time field indicates when the record was queued in the source system. If the upstream system is known to occasionally enqueue duplicate entries for a single order hours apart, which statement is correct?

- A. All records will be held in the state store for 2 hours before being deduplicated and committed to the orders table.
- B. Duplicate records enqueued more than 2 hours apart may be retained and the orders table may contain duplicate records with the same customer_id and order_id.
- C. The orders table will not contain duplicates, but records arriving more than 2 hours late will be ignored and missing from the table.
- D. The orders table will contain only the most recent 2 hours of records and no duplicates will be present.

Answer: C

Explanation:

Comprehensive and Detailed Explanation From Exact Extract:

Exact extract: "dropDuplicates with watermark performs stateful deduplication on the keys within the watermark delay." Exact extract: "Records older than the event-time watermark are considered late and may be dropped." Exact extract: "trigger(once) processes all available data once and then stops."

NEW QUESTION # 206

A data engineer is designing an append-only pipeline that needs to handle both batch and streaming data in Delta Lake. The team wants to ensure that the streaming component can efficiently track which data has already been processed.

Which configuration should be set to enable this?

- A. mergeSchema
- B. overwriteSchema
- C. **checkpointLocation**
- D. partitionBy

Answer: C

Explanation:

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

When working with Delta Lake streaming ingestion, checkpointing is critical for maintaining fault tolerance and ensuring exactly-once data processing semantics.

The checkpointLocation parameter defines the directory where Spark Structured Streaming stores progress information, offsets, and metadata. This allows the engine to resume processing from the last committed offset without reprocessing previously ingested data. Without checkpointing, each stream restart would reprocess all data, leading to duplicates. Parameters like partitionBy or schema options (mergeSchema / overwriteSchema) affect table structure, not data lineage tracking. Therefore, the correct and required configuration for efficient streaming state management is checkpointLocation.

NEW QUESTION # 207

Which of the below commands can be used to drop a DELTA table?

- A. DROP table_name
- B. DROP DELTA table_name
- C. **DROP TABLE table_name**
- D. DROP TABLE table_name FORMAT DELTA

Answer: C

NEW QUESTION # 208

.....

You can install and use ActualtestPDF Databricks exam dumps formats easily and start Databricks Databricks-Certified-Professional-Data-Engineer exam preparation right now. The ActualtestPDF Databricks-Certified-Professional-Data-Engineer desktop practice test software and web-based practice test software both are the mock Databricks Certified Professional Data Engineer Exam (Databricks-Certified-Professional-Data-Engineer) exam that stimulates the actual exam format and content.

Test Databricks-Certified-Professional-Data-Engineer Dumps: <https://www.actualtestpdf.com/Databricks/Databricks-Certified-Professional-Data-Engineer-practice-exam-dumps.html>

- Databricks Databricks-Certified-Professional-Data-Engineer Valid Test Voucher - First-Grade Test Databricks-Certified-Professional-Data-Engineer Dumps and Pass-Sure Certification Databricks Certified Professional Data Engineer Exam Dumps □ Enter ➔ www.prepawaypdf.com □□□ and search for (Databricks-Certified-Professional-Data-Engineer) to download for free □ Databricks-Certified-Professional-Data-Engineer Practice Questions
- Reliable Databricks Databricks-Certified-Professional-Data-Engineer Exam Study Material from Pdfvce □ Search for ➔ Databricks-Certified-Professional-Data-Engineer □□□ and easily obtain a free download on □ www.pdfvce.com □ □ □ Databricks-Certified-Professional-Data-Engineer Valid Torrent
- Databricks-Certified-Professional-Data-Engineer Latest Dumps Book □ Databricks-Certified-Professional-Data-Engineer Test Collection □ Databricks-Certified-Professional-Data-Engineer Exam Questions Vce □ Go to website ➔ www.examcollectionpass.com □ open and search for { Databricks-Certified-Professional-Data-Engineer } to download for free □ Databricks-Certified-Professional-Data-Engineer Latest Dumps Book
- Study Databricks-Certified-Professional-Data-Engineer Materials □ Databricks-Certified-Professional-Data-Engineer Exam Questions Vce □ Databricks-Certified-Professional-Data-Engineer Training Questions □ Download 【 Databricks-Certified-Professional-Data-Engineer 】 for free by simply entering ✎ www.pdfvce.com ✎ ✎ website □

