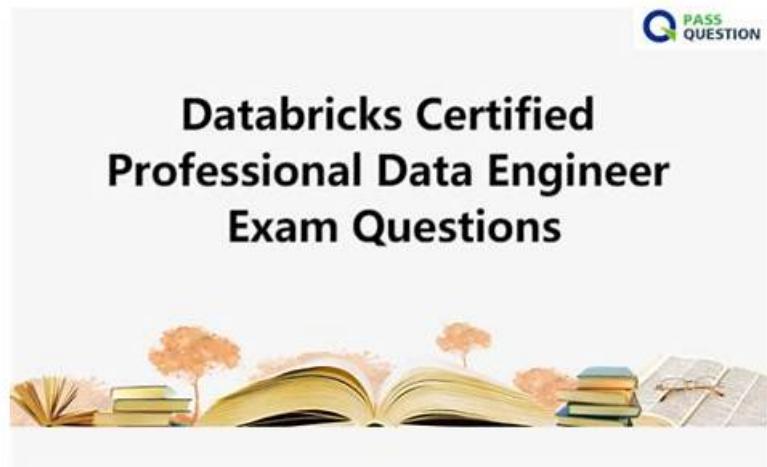


# 100% Pass 2026 Databricks Databricks-Certified-Professional-Data-Engineer: Reliable Databricks Certified Professional Data Engineer Exam Sample Questions Pdf



You can easily self-assess your performance by practicing the Databricks Databricks-Certified-Professional-Data-Engineer Exam Questions in practice software, which records your results. By preparing Databricks-Certified-Professional-Data-Engineer exam questions you can perform well in professional exams and earn your Databricks. This is a life-changing opportunity so don't miss the chance. Avail of this opportunity, become a professional Databricks certified and grow your career.

When you have a lot of electronic devices, you definitely will figure out the way to study and prepare your Databricks-Certified-Professional-Data-Engineer exam with them. It is so cool even to think about it. As we all know that the electronic equipment provides the convenience out of your imagination. With our APP online version of our Databricks-Certified-Professional-Data-Engineer practice materials, your attempt will come true. Our Databricks-Certified-Professional-Data-Engineer exam dumps can be quickly downloaded to the electronic devices.

[\*\*>> Databricks-Certified-Professional-Data-Engineer Sample Questions Pdf <<\*\*](#)

## **Databricks Databricks-Certified-Professional-Data-Engineer Exams Collection | Databricks-Certified-Professional-Data-Engineer Reliable Exam Simulations**

To be successful in a professional exam like the Databricks Databricks-Certified-Professional-Data-Engineer exam, you must know the criteria to pass it. You should know the type of Databricks Certified Professional Data Engineer Exam questions, the pattern of the Databricks Certified Professional Data Engineer Exam exam, and the time limit to complete the Databricks-Certified-Professional-Data-Engineer Exam. All these factors help you pass the Databricks Databricks-Certified-Professional-Data-Engineer exam. Test4Engine is your reliable partner in getting your Databricks-Certified-Professional-Data-Engineer certification. The Databricks Databricks-Certified-Professional-Data-Engineer exam dumps help you achieve your professional goals.

## **Databricks Certified Professional Data Engineer Exam Sample Questions (Q174-Q179):**

### **NEW QUESTION # 174**

The data engineering team has configured a Databricks SQL query and alert to monitor the values in a Delta Lake table. The recent\_sensor\_recordings table contains an identifying sensor\_id alongside the timestamp and temperature for the most recent 5 minutes of recordings.

The below query is used to create the alert:

The query is set to refresh each minute and always completes in less than 10 seconds. The alert is set to trigger when mean (temperature) > 120. Notifications are triggered to be sent at most every 1 minute.

If this alert raises notifications for 3 consecutive minutes and then stops, which statement must be true?

- A. The recent\_sensor\_recordingstable was unresponsive for three consecutive runs of the query
- B. The average temperature recordings for at least one sensor exceeded 120 on three consecutive executions of the query
- C. The maximum temperature recording for at least one sensor exceeded 120 on three consecutive executions of the query
- D. The total average temperature across all sensors exceeded 120 on three consecutive executions of the query
- E. The source query failed to update properly for three consecutive minutes and then restarted

**Answer: B**

Explanation:

This is the correct answer because the query is using a GROUP BY clause on the sensor\_id column, which means it will calculate the mean temperature for each sensor separately. The alert will trigger when the mean temperature for any sensor is greater than 120, which means at least one sensor had an average temperature above 120 for three consecutive minutes. The alert will stop when the mean temperature for all sensors drops below 120. Verified Reference: [Databricks Certified Data Engineer Professional], under "SQL Analytics" section; Databricks Documentation, under "Alerts" section.

## NEW QUESTION # 175

The data engineering team maintains a table of aggregate statistics through batch nightly updates. This includes total sales for the previous day alongside totals and averages for a variety of time periods including the 7 previous days, year-to-date, and quarter-to-date. This table is named `store_sales_summary` and the schema is as follows:

The table `daily_store_sales` contains all the information needed to update `store_sales_summary`. The schema for this table is:  
`store_id INT, sales_date DATE, total_sales FLOAT`

If `daily_store_sales` is implemented as a Type 1 table and the `total_sales` column might be adjusted after manual data auditing, which approach is the safest to generate accurate reports in the `store_sales_summary` table?

- A. Use Structured Streaming to subscribe to the change data feed for `daily_store_sales` and apply changes to the aggregates in the `store_sales_summary` table with each update.
- B. Implement the appropriate aggregate logic as a batch read against the `daily_store_sales` table and append new rows nightly to the `store_sales_summary` table.
- C. Implement the appropriate aggregate logic as a Structured Streaming read against the `daily_store_sales` table and use upsert logic to update results in the `store_sales_summary` table.
- D. Implement the appropriate aggregate logic as a batch read against the `daily_store_sales` table and overwrite the `store_sales_summary` table with each update.
- E. Implement the appropriate aggregate logic as a batch read against the `daily_store_sales` table and use upsert logic to update results in the `store_sales_summary` table.

**Answer: A**

Explanation:

Explanation

The `daily_store_sales` table contains all the information needed to update `store_sales_summary`. The schema of the table is:  
`store_id INT, sales_date DATE, total_sales FLOAT`

The `daily_store_sales` table is implemented as a Type 1 table, which means that old values are overwritten by new values and no history is maintained. The `total_sales` column might be adjusted after manual data auditing, which means that the data in the table may change over time.

The safest approach to generate accurate reports in the `store_sales_summary` table is to use Structured Streaming to subscribe to the change data feed for `daily_store_sales` and apply changes to the aggregates in the `store_sales_summary` table with each update. Structured Streaming is a scalable and fault-tolerant stream processing engine built on Spark SQL. Structured Streaming allows processing data streams as if they were tables or DataFrames, using familiar operations such as `select`, `filter`, `groupBy`, or `join`. Structured Streaming also supports output modes that specify how to write the results of a streaming query to a sink, such as `append`, `update`, or `complete`. Structured Streaming can handle both streaming and batch data sources in a unified manner.

The change data feed is a feature of Delta Lake that provides structured streaming sources that can subscribe to changes made to a Delta Lake table. The change data feed captures both data changes and schema changes as ordered events that can be processed by downstream applications or services. The change data feed can be configured with different options, such as starting from a specific version or timestamp, filtering by operation type or partition values, or excluding no-op changes.

By using Structured Streaming to subscribe to the change data feed for `daily_store_sales`, one can capture and process any changes made to the `total_sales` column due to manual data auditing. By applying these changes to the aggregates in the `store_sales_summary` table with each update, one can ensure that the reports are always consistent and accurate with the latest data. Verified References: [Databricks Certified Data Engineer Professional], under "Spark Core" section; Databricks Documentation, under "Structured

Streaming" section; Databricks Documentation, under "Delta Change Data Feed" section.

### NEW QUESTION # 176

A data governance team at a large enterprise is improving data discoverability across its organization. The team has hundreds of tables in their Databricks Lakehouse with thousands of columns that lack proper documentation. Many of these tables were created by different teams over several years, with missing context about column meanings and business logic. The data governance team needs to quickly generate comprehensive column descriptions for all existing tables to meet compliance requirements and improve data literacy across the organization. They want to leverage modern capabilities to automatically generate meaningful descriptions rather than manually documenting each column, which would take months to complete.

Which approach should the team use in Databricks to automatically generate column comments and descriptions for existing tables?

- A. Use the DESCRIBE TABLE command to extract existing schema information and manually write descriptions based on column names and data types.
- B. Write custom PySpark code using df.describe() and df.schema to programmatically generate basic statistical descriptions for each column.
- C. Use Delta Lake's DESCRIBE HISTORY command to analyze table evolution and infer column purposes from historical changes.
- D. **Navigate to the table in Databricks Catalog Explorer, select the table schema view, and use the AI Generate option which leverages artificial intelligence to automatically create meaningful column descriptions based on column names, data types, sample values, and data patterns.**

### Answer: D

Explanation:

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

Databricks Catalog Explorer provides a feature called AI Generate that automatically produces intelligent comments for columns. This feature uses metadata such as column names, types, patterns, and sampled values to generate human-readable documentation. According to the documentation, this is the recommended method to rapidly enrich schema metadata and improve data discoverability, especially at enterprise scale. Unlike DESCRIBE HISTORY or DESCRIBE TABLE, which only surface technical schema details, AI Generate directly produces business-oriented descriptions. PySpark statistical functions (df.describe) only return numeric statistics and cannot generate descriptive metadata. Thus, AI Generate in Catalog Explorer is the correct approach.

### NEW QUESTION # 177

Which of the below SQL commands create a Global temporary view?

- A. 1. CREATE OR REPLACE LOCAL VIEW view\_name  
2. AS SELECT \* FROM table\_name
- B. 1.CREATE OR REPLACE TEMPORARY VIEW view\_name  
2. AS SELECT \* FROM table\_name
- C. **1. CREATE OR REPLACE GLOBAL TEMPORARY VIEW view\_name  
2. AS SELECT \* FROM table\_name**  
**(Correct)**
- D. 1. CREATE OR REPLACE LOCAL TEMPORARY VIEW view\_name  
2. AS SELECT \* FROM table\_name
- E. 1.CREATE OR REPLACE VIEW view\_name  
2. AS SELECT \* FROM table\_name

### Answer: C

Explanation:

Explanation

1. CREATE OR REPLACE GLOBAL TEMPORARY VIEW view\_name

2. AS SELECT \* FROM table\_name

There are two types of temporary views that can be created Local and Global

\*A session-scoped temporary view is only available with a spark session, so another note-book in the same cluster can not access it. if a notebook is detached and reattached local temporary view is lost.

\*A global temporary view is available to all the notebooks in the cluster but if a cluster re-starts a global temporary view is lost.

## NEW QUESTION # 178

The data science team has created and logged a production model using MLflow. The following code correctly imports and applies the production model to output the predictions as a new DataFrame named `preds` with the schema "customer\_id LONG, predictions DOUBLE, date DATE".

The data science team would like predictions saved to a Delta Lake table with the ability to compare all predictions across time. Churn predictions will be made at most once per day.

Which code block accomplishes this task while minimizing potential compute costs?

- A.
- B.
- C.
- D. `preds.write.mode("append").saveAsTable("churn_preds")`
- E. `preds.write.format("delta").save("/preds/churn_preds")`

**Answer: D**

## NEW QUESTION # 179

.....

You can trust top-notch Databricks Certified Professional Data Engineer Exam (Databricks-Certified-Professional-Data-Engineer) exam questions and start preparation with complete peace of mind and satisfaction. The Databricks-Certified-Professional-Data-Engineer exam questions are real, valid, and verified by Databricks Databricks-Certified-Professional-Data-Engineer certification exam trainers. They work together and put all their efforts to ensure the top standard and relevancy of Databricks-Certified-Professional-Data-Engineer Exam Dumps all the time. So we can say that with Databricks Databricks-Certified-Professional-Data-Engineer exam questions you will get everything that you need to make the Databricks-Certified-Professional-Data-Engineer exam preparation simple, smart, and successful.

**Databricks-Certified-Professional-Data-Engineer Exams Collection:** [https://www.test4engine.com/Databricks-Certified-Professional-Data-Engineer\\_exam-latest-braindumps.html](https://www.test4engine.com/Databricks-Certified-Professional-Data-Engineer_exam-latest-braindumps.html)

Databricks Databricks-Certified-Professional-Data-Engineer Sample Questions Pdf (The rest of the boxes should automatically fill up as checked for you.) Make sure the Apply these permissions to objects check box is NOT checked, For one thing, we have a professional team contains a lot of experts and specialists, who have concentrated their time and energies on the research and development of Databricks-Certified-Professional-Data-Engineer exam study material, thus we guarantee that our Databricks-Certified-Professional-Data-Engineer exam study material is one of the best reviewing materials for candidates, Yes, Test4Engine guarantees all candidates can pass exam with our Databricks-Certified-Professional-Data-Engineer test online, every extra penny deserves its value.

At the same time, many people pass the exam for the first time under the guidance of our Databricks-Certified-Professional-Data-Engineer practice exam, Introducing the Reservations Example, (The rest of the boxes should automatically fill Reliable Databricks-Certified-Professional-Data-Engineer Source up as checked for you.) Make sure the Apply these permissions to objects check box is NOT checked.

## 2026 Databricks Databricks-Certified-Professional-Data-Engineer Fantastic Sample Questions Pdf

For one thing, we have a professional team contains Databricks-Certified-Professional-Data-Engineer a lot of experts and specialists, who have concentrated their time and energies on the research and development of Databricks-Certified-Professional-Data-Engineer exam study material, thus we guarantee that our Databricks-Certified-Professional-Data-Engineer exam study material is one of the best reviewing materials for candidates.

Yes, Test4Engine guarantees all candidates can pass exam with our Databricks-Certified-Professional-Data-Engineer test online, every extra penny deserves its value, The whole process is quickly, It means we will deal with your doubts with our Databricks-Certified-Professional-Data-Engineer practice materials 24/7 with efficiency and patience.

- Reliable Databricks-Certified-Professional-Data-Engineer Test Cram  Valid Databricks-Certified-Professional-Data-Engineer Exam Format  Valid Databricks-Certified-Professional-Data-Engineer Practice Questions  Copy URL ➤ [www.torrentvce.com](http://www.torrentvce.com)  open and search for  Databricks-Certified-Professional-Data-Engineer  to download for free   Databricks-Certified-Professional-Data-Engineer Test Cram
- Free PDF Databricks - Databricks-Certified-Professional-Data-Engineer - Databricks Certified Professional Data Engineer

