

Databricks-Certified-Professional-Data-Engineer Real Question | Databricks-Certified-Professional-Data-Engineer Free Sample Questions



Our company made these Databricks-Certified-Professional-Data-Engineer practice materials with accountability. We understand you can have more chances being accepted by other places and getting higher salary or acceptance. Our Databricks Certified Professional Data Engineer Exam training materials are made by our responsible company which means you can gain many other benefits as well. We offer Databricks-Certified-Professional-Data-Engineer free demos for your reference, and send you the new updates if our experts make them freely. If you fail the exam after using our Databricks-Certified-Professional-Data-Engineer exam prep unfortunately, we will switch other versions for you or return full refund.

Databricks Certified Professional Data Engineer (Databricks-Certified-Professional-Data-Engineer) certification exam is designed for data professionals who want to validate their skills and knowledge in building and deploying data engineering solutions using Databricks. Databricks is a unified data analytics platform that provides a collaborative environment for data engineers, data scientists, and business analysts to work together on big data projects. Databricks Certified Professional Data Engineer Exam certification exam covers a range of topics such as data ingestion, data processing, data transformation, and data storage using Databricks.

>> **Databricks-Certified-Professional-Data-Engineer Real Question** <<

Actual Databricks Databricks-Certified-Professional-Data-Engineer Exam Dumps - Achieve Success In Exam

Many candidates know our exam bootcamp materials are valid and enough to help them clear Databricks Databricks-Certified-Professional-Data-Engineer exams. But they are afraid that purchasing on internet is not safe, money unsafe and information unsafe. In fact you may worry too much. Online sale is very common. Every year there are thousands of candidates choose our Databricks-Certified-Professional-Data-Engineer Exam Bootcamp materials and pass exam surely. Money is certainly safe. PayPal will guarantee your money and your benefits safe. We have strict information secret system to guarantee that your information is safe too.

Databricks Certified Professional Data Engineer Exam Sample Questions (Q147-Q152):

NEW QUESTION # 147

A healthcare analytics team is implementing a dimensional model in Delta Lake for patient care analysis. They have a date dimension table and are evaluating design options to ensure it supports a wide range of time-based analyses.

Which design approach for the date dimension will support efficient time-based querying and aggregation?

- A. Create separate dimension tables for different calendar systems (fiscal, academic, etc.).
- B. Pre-calculate attributes like `fiscal_period`, `quarter`, `month_name`, `day_of_week`, and `holiday`.
- C. Store the date as a string in the format YYYY-MM-DD for readability.

- D. Store only the date value and calculate all time attributes dynamically in queries.

Answer: B

Explanation:

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

In dimensional modeling, Databricks recommends denormalized, attribute-rich dimension tables for performance and usability. A date dimension should include all commonly used derived time attributes such as fiscal period, quarter, month, weekday, and holiday flags. Precomputing these attributes ensures consistent business logic, eliminates repeated calculations during query time, and enables efficient filtering and aggregation. The documentation for Delta Lake and Lakehouse design explicitly advises precomputing these attributes for analytical workloads that depend heavily on time-based slicing. Options A and C degrade performance and consistency, while maintaining multiple calendar-specific dimension tables (B) complicates the model unnecessarily.

NEW QUESTION # 148

A data architect has heard about lake's built-in versioning and time travel capabilities. For auditing purposes they have a requirement to maintain a full of all valid street addresses as they appear in the customers table.

The architect is interested in implementing a Type 1 table, overwriting existing records with new values and relying on Delta Lake time travel to support long-term auditing. A data engineer on the project feels that a Type 2 table will provide better performance and scalability.

Which piece of information is critical to this decision?

- A. Data corruption can occur if a query fails in a partially completed state because Type 2 tables requires
- B. Delta Lake time travel cannot be used to query previous versions of these tables because Type 1 changes modify data files in place.
- C. Delta Lake time travel does not scale well in cost or latency to provide a long-term versioning solution.
- D. Shallow clones can be combined with Type 1 tables to accelerate historic queries for long-term versioning.

Answer: C

Explanation:

Setting multiple fields in a single update.

Explanation:

Delta Lake's time travel feature allows users to access previous versions of a table, providing a powerful tool for auditing and versioning. However, using time travel as a long-term versioning solution for auditing purposes can be less optimal in terms of cost and performance, especially as the volume of data and the number of versions grow. For maintaining a full history of valid street addresses as they appear in a customers table, using a Type 2 table (where each update creates a new record with versioning) might provide better scalability and performance by avoiding the overhead associated with accessing older versions of a large table. While Type 1 tables, where existing records are overwritten with new values, seem simpler and can leverage time travel for auditing, the critical piece of information is that time travel might not scale well in cost or latency for long-term versioning needs, making a Type 2 approach more viable for performance and scalability.

Reference:

Databricks Documentation on Delta Lake's Time Travel: Delta Lake Time Travel Databricks Blog on Managing Slowly Changing Dimensions in Delta Lake: Managing SCDs in Delta Lake

NEW QUESTION # 149

You are currently working with the second team and both teams are looking to modify the same notebook, you noticed that the second member is copying the notebooks to the personal folder to edit and replace the collaboration notebook, which notebook feature do you recommend to make the process easier to collaborate.

- A. Databricks notebook can be exported as HTML and imported at a later time
- B. Databricks notebooks should be copied to a local machine and setup source control locally to version the notebooks
- C. Databricks notebooks support automatic change tracking and versioning
- D. Databricks notebooks can be exported into dbc archive files and stored in data lake
- E. Databricks Notebooks support real-time coauthoring on a single notebook

Answer: E

Explanation:

Explanation

Answer is Databricks Notebooks support real-time coauthoring on a single notebook. Every change is saved, and a notebook can be changed by multiple users.

NEW QUESTION # 150

Where in the Spark UI can one diagnose a performance problem induced by not leveraging predicate push-down?

- A. In the Delta Lake transaction log, by noting the column statistics
- B. In the Storage Detail screen, by noting which RDDs are not stored on disk
- **C. In the Query Detail screen, by interpreting the Physical Plan**
- D. In the Executor's log file, by grepping for "predicate push-down"
- E. In the Stage's Detail screen, in the Completed Stages table, by noting the size of data read from the Input column

Answer: C

Explanation:

Explanation

This is the correct answer because it is where in the Spark UI one can diagnose a performance problem induced by not leveraging predicate push-down. Predicate push-down is an optimization technique that allows filtering data at the source before loading it into memory or processing it further. This can improve performance and reduce I/O costs by avoiding reading unnecessary data. To leverage predicate push-down, one should use supported data sources and formats, such as Delta Lake, Parquet, or JDBC, and use filter expressions that can be pushed down to the source. To diagnose a performance problem induced by not leveraging predicate push-down, one can use the Spark UI to access the Query Detail screen, which shows information about a SQL query executed on a Spark cluster. The Query Detail screen includes the Physical Plan, which is the actual plan executed by Spark to perform the query. The Physical Plan shows the physical operators used by Spark, such as Scan, Filter, Project, or Aggregate, and their input and output statistics, such as rows and bytes. By interpreting the Physical Plan, one can see if the filter expressions are pushed down to the source or not, and how much data is read or processed by each operator. Verified References: [Databricks Certified Data Engineer Professional], under "Spark Core" section; Databricks Documentation, under "Predicate pushdown" section; Databricks Documentation, under "Query detail page" section.

NEW QUESTION # 151

Which of the following is correct for the global temporary view?

- **A. global temporary views can be still accessed even if the notebook is detached and attached**
- B. global temporary views can be still accessed even if the cluster is restarted
- C. global temporary views are created in a database called temp database
- D. global temporary views cannot be accessed once the notebook is detached and attached
- E. global temporary views can be accessed across many clusters

Answer: A

Explanation:

Explanation

The answer is global temporary views can be still accessed even if the notebook is detached and attached. There are two types of temporary views that can be created: Local and Global.

* A local temporary view is only available with a Spark session, so another notebook in the same cluster cannot access it. If a notebook is detached and reattached, the local temporary view is lost.

* A global temporary view is available to all the notebooks in the cluster, even if the notebook is detached and reattached; it can still be accessible, but if a cluster is restarted, the global temporary view is lost.

NEW QUESTION # 152

.....

With precious time passing away, many exam candidates are making progress with high speed and efficiency with the help of our Databricks-Certified-Professional-Data-Engineer study guide. You cannot lag behind and with our Databricks-Certified-Professional-Data-Engineer preparation materials, and your goals will be easier to fix. So stop idling away your precious time and begin your review with the help of our Databricks-Certified-Professional-Data-Engineer learning quiz as soon as possible, and you will pass the exam in the least time.

Databricks-Certified-Professional-Data-Engineer Free Sample Questions: <https://www.testkingpass.com/Databricks-Certified-Professional-Data-Engineer-testking-dumps.html>

- [illegible]