

Databricks Certified Data Engineer Associate Exam Updated Torrent - Databricks-Certified-Data-Engineer-Associate exam pdf & Databricks Certified Data Engineer Associate Exam Practice questions



P.S. Free 2026 Databricks Databricks-Certified-Data-Engineer-Associate dumps are available on Google Drive shared by VCEPrep: <https://drive.google.com/open?id=1CeOLXe6dz7Qs7qtBU4X1wmq7vzbwcrXy>

Test your knowledge of the Databricks-Certified-Data-Engineer-Associate exam dumps with Databricks Databricks-Certified-Data-Engineer-Associate practice questions. The software is designed to help with Databricks Certified Data Engineer Associate Exam (Databricks-Certified-Data-Engineer-Associate) exam dumps preparation. Databricks Databricks-Certified-Data-Engineer-Associate Practice Test software can be used on devices that range from mobile devices to desktop computers.

Your purchase with VCEPrep is safe and fast. We use Paypal for payment and committed to keep your personal information secret and never share your information to the third part without your permission. In addition, our Databricks Databricks-Certified-Data-Engineer-Associate practice exam torrent can be available for immediate download after your payment. Besides, we guarantee you 100% pass for Databricks-Certified-Data-Engineer-Associate Actual Test, in case of failure, you can ask for full refund. The refund procedure is very easy. You just need to show us your Databricks-Certified-Data-Engineer-Associate failure certification, then after confirmation, we will deal with your case.

>> Valid Databricks-Certified-Data-Engineer-Associate Exam Bootcamp <<

Databricks-Certified-Data-Engineer-Associate Learning Material: Databricks Certified Data Engineer Associate Exam & Databricks-Certified-Data-Engineer-Associate Practice Test

Our worldwide after sale staff on the Databricks-Certified-Data-Engineer-Associate exam questions will be online and reassure your rows of doubts as well as exclude the difficulties and anxiety with all the customers. Just let us know your puzzles on Databricks-Certified-Data-Engineer-Associate study materials and we will figure out together. We can give you suggestion on Databricks-Certified-Data-Engineer-Associate training engine 24/7, as long as you contact us, no matter by email or online, you will be answered quickly and professionally!

The GAQM Databricks-Certified-Data-Engineer-Associate Certification is an essential credential for individuals and organizations working with big data using Databricks. It validates the candidate's skills and knowledge of data engineering concepts and their ability to design and implement data pipelines using Databricks. Databricks Certified Data Engineer Associate Exam certification is globally recognized and highly respected in the industry, making it a valuable asset for anyone looking to advance their career in the field of data engineering.

Databricks Certified Data Engineer Associate Exam Sample Questions

(Q121-Q126):

NEW QUESTION # 121

A data engineer has been using a Databricks SQL dashboard to monitor the cleanliness of the input data to a data analytics dashboard for a retail use case. The job has a Databricks SQL query that returns the number of store-level records where sales is equal to zero. The data engineer wants their entire team to be notified via a messaging webhook whenever this value is greater than 0.

Which of the following approaches can the data engineer use to notify their entire team via a messaging webhook whenever the number of stores with \$0 in sales is greater than zero?

- A. They can set up an Alert with a new email alert destination.
- B. They can set up an Alert with a custom template.
- C. They can set up an Alert with one-time notifications.
- D. They can set up an Alert without notifications.
- E. They can set up an Alert with a new webhook alert destination.

Answer: E

Explanation:

A webhook alert destination is a notification destination that allows Databricks to send HTTP POST requests to a third-party endpoint when an alert is triggered. This enables the data engineer to integrate Databricks alerts with their preferred messaging or collaboration platform, such as Slack, Microsoft Teams, or PagerDuty.

To set up a webhook alert destination, the data engineer needs to create and configure a webhook connector in their messaging platform, and then add the webhook URL to the Databricks notification destination. After that, the data engineer can create an alert for their Databricks SQL query, and select the webhook alert destination as the notification destination. The alert can be configured with a custom condition, such as when the number of stores with \$0 in sales is greater than zero, and a custom message template, such as "Alert:{number_of_stores} stores have \$0 in sales". The alert can also be configured with a recurrence interval, such as every hour, to check the query result periodically. When the alert condition is met, the data engineer and their team will receive a notification via the messaging webhook, with the custom message and a link to the Databricks SQL query. The other options are either not suitable for sending notifications via a messaging webhook (A, B, E), or not suitable for sending recurring notifications.

References: Databricks Documentation - Manage notification destinations, Databricks Documentation - Create alerts for Databricks SQL queries, Databricks Documentation - Configure alert conditions and messages.

NEW QUESTION # 122

A data engineer has realized that the data files associated with a Delta table are incredibly small. They want to compact the small files to form larger files to improve performance.

Which of the following keywords can be used to compact the small files?

- A. VACUUM
- B. REDUCE
- C. OPTIMIZE
- D. REPARTITION
- E. COMPACTION

Answer: C

Explanation:

The keyword that can be used to compact the small files associated with a Delta table is OPTIMIZE. The OPTIMIZE command performs file compaction on a Delta table by rewriting a set of small files into a set of larger files¹. This can improve the performance of queries that scan the table by reducing the number of files that need to be read and the amount of metadata that needs to be processed¹. The OPTIMIZE command can also optionally sort the data within each file by a given set of columns, which can further improve the query performance by enabling data skipping and predicate pushdown¹. The OPTIMIZE command can be applied to the whole table or to a specific partition of the table¹.

The other keywords are not suitable for compacting the small files associated with a Delta table. REDUCE is a keyword used in the SQL syntax for aggregating data using a user-defined function². COMPACTION is not a valid keyword in SQL or Python. REPARTITION is a keyword used in the Python syntax for changing the number of partitions of a DataFrame or an RDD³. VACUUM is a keyword used to remove files that are no longer referenced by a Delta table and are older than a retention threshold⁴.

References:

* 1: OPTIMIZE | Databricks on AWS

- * 2: REDUCE | Databricks on AWS
- * 3: repartition | Databricks on AWS
- * 4: VACUUM | Databricks on AWS

NEW QUESTION # 123

A data engineering team has noticed that their Databricks SQL queries are running too slowly when they are submitted to a non-running SQL endpoint. The data engineering team wants this issue to be resolved.

Which of the following approaches can the team use to reduce the time it takes to return results in this scenario?

- A. They can increase the maximum bound of the SQL endpoint's scaling range
- B. They can turn on the Serverless feature for the SQL endpoint and change the Spot Instance Policy to "Reliability Optimized."
- C. They can increase the cluster size of the SQL endpoint.
- D. They can turn on the Auto Stop feature for the SQL endpoint.
- E. They can turn on the Serverless feature for the SQL endpoint.

Answer: E

Explanation:

Option D is the correct answer because it enables the Serverless feature for the SQL endpoint, which allows the endpoint to automatically scale up and down based on the query load. This way, the endpoint can handle more concurrent queries and reduce the time it takes to return results. The Serverless feature also reduces the cold start time of the endpoint, which is the time it takes to start the cluster when a query is submitted to a non-running endpoint. The Serverless feature is available for both AWS and Azure Databricks platforms.

References: Databricks SQL Serverless, Serverless SQL endpoints, New Performance Improvements in Databricks SQL

NEW QUESTION # 124

A data engineer is designing an ETL pipeline to process both streaming and batch data from multiple sources. The pipeline must ensure data quality, handle schema evolution, and provide easy maintenance. The team is considering using Delta Live Tables (DLT) in Databricks to achieve these goals. They want to understand the key features and benefits of DLT that make it suitable for this use case.

Why is Delta Live Tables (DLT) an appropriate choice?

- A. Requires custom code for data quality checks, no support for streaming data, and complex pipeline maintenance
- B. Automatic data quality checks, built-in support for schema evolution, and declarative pipeline development
- C. Manual schema enforcement, high operational overhead, and limited scalability
- D. Supports only batch processing, no data versioning, and high infrastructure costs

Answer: B

NEW QUESTION # 125

Which of the following describes a benefit of creating an external table from Parquet rather than CSV when using a CREATE TABLE AS SELECT statement?

- A. Parquet files will become Delta tables
- B. Parquet files have a well-defined schema
- C. Parquet files can be partitioned
- D. Parquet files have the ability to be optimized
- E. CREATE TABLE AS SELECT statements cannot be used on files

Answer: B

Explanation:

Option C is the correct answer because Parquet files have a well-defined schema that is embedded within the data itself. This means that the data types and column names of the Parquet files are automatically detected and preserved when creating an external table from them. This also enables the use of SQL and other structured query languages to access and analyze the data. CSV files, on the other hand, do not have a schema embedded in them, and require specifying the schema explicitly or inferring it from the data when creating an external table from them. This can lead to errors or inconsistencies in the data types and column names, and also increase

BONUS!!! Download part of VCEPrep Databricks-Certified-Data-Engineer-Associate dumps for free:
<https://drive.google.com/open?id=1CeOLXe6dz7Qs7qtBU4X1wmq7vzbwcrXy>