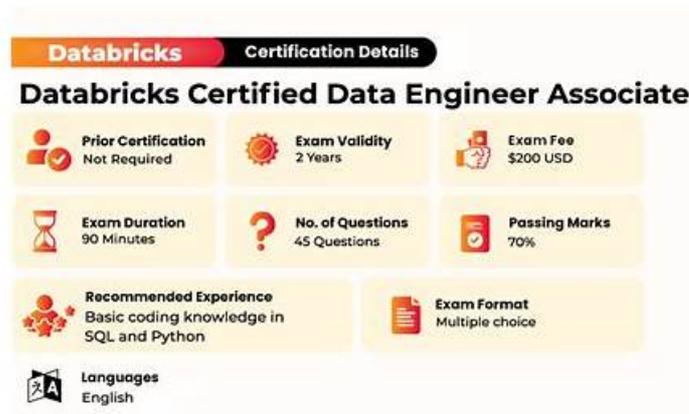


Latest Databricks Databricks-Certified-Data-Engineer-Associate Exam Questions in Three Formats



2026 Latest VCETorrent Databricks-Certified-Data-Engineer-Associate PDF Dumps and Databricks-Certified-Data-Engineer-Associate Exam Engine Free Share: https://drive.google.com/open?id=1UYml_rMF6T4cbGYZfPaqlu-rRHilHYkc

They make an effort to find reliable and current Databricks Databricks-Certified-Data-Engineer-Associate practice questions for the difficult Databricks Databricks-Certified-Data-Engineer-Associate exam. More challenging than just passing the Databricks Databricks-Certified-Data-Engineer-Associate Certification are the intense anxiety and heavy workload that the candidate must endure to be eligible for the Databricks Databricks-Certified-Data-Engineer-Associate certification.

The GAQM Databricks-Certified-Data-Engineer-Associate Certification is an excellent way for professionals to demonstrate their competence and expertise in working with Databricks. Databricks Certified Data Engineer Associate Exam certification validates your knowledge and skills in designing, building, and maintaining data pipelines with Databricks, which is an essential skill for any organization that is looking to leverage big data for business growth.

>> **Databricks-Certified-Data-Engineer-Associate Download Free Dumps** <<

Quiz Databricks - The Best Databricks-Certified-Data-Engineer-Associate Download Free Dumps

The Databricks Databricks-Certified-Data-Engineer-Associate exam questions of VCETorrent mainly come in three formats: Databricks Certified Data Engineer Associate Exam (Databricks-Certified-Data-Engineer-Associate) dumps PDF, Web-Based Databricks-Certified-Data-Engineer-Associate Practice Exam and Desktop Databricks Certified Data Engineer Associate Exam (Databricks-Certified-Data-Engineer-Associate) Practice Test Software. With thousands of satisfied customers, you can start your preparation for Databricks Certified Data Engineer Associate Exam (Databricks-Certified-Data-Engineer-Associate) certification with VCETorrent.

Databricks Certified Data Engineer Associate Exam Sample Questions (Q32-Q37):

NEW QUESTION # 32

Which of the following data lakehouse features results in improved data quality over a traditional data lake?

- A. A data lakehouse supports ACID-compliant transactions.
- B. A data lakehouse provides storage solutions for structured and unstructured data.
- C. A data lakehouse enables machine learning and artificial intelligence workloads.
- D. A data lakehouse stores data in open formats.
- E. A data lakehouse allows the use of SQL queries to examine data.

Answer: A

Explanation:

Explanation

One of the key features of a data lakehouse that results in improved data quality over a traditional data lake is its support for ACID (Atomicity, Consistency, Isolation, Durability) transactions. ACID transactions provide data integrity and consistency guarantees, ensuring that operations on the data are reliable and that data is not left in an inconsistent state due to failures or concurrent access. In a traditional data lake, such transactional guarantees are often lacking, making it challenging to maintain data quality, especially in scenarios involving multiple data writes, updates, or complex transformations. A data lakehouse, by offering ACID compliance, helps maintain data quality by providing strong consistency and reliability, which is crucial for data pipelines and analytics.

NEW QUESTION # 33

A data engineer needs to create a table in Databricks using data from a CSV file at location /path/to/csv. They run the following command:

```
CREATE TABLE new_table
____
OPTIONS (
  header = "true",
  delimiter = "|"
)
LOCATION "path/to/csv"
```

Which of the following lines of code fills in the above blank to successfully complete the task?

- A. FROM "path/to/csv"
- B. USING DELTA
- C. None of these lines of code are needed to successfully complete the task
- D. FROM CSV
- E. USING CSV

Answer: A

Explanation:

A data lakehouse is a new paradigm that can be used to simplify and unify siloed data architectures that are specialized for specific use cases. A data lakehouse combines the best of both data lakes and data warehouses, providing a single platform that supports diverse data types, open standards, low-cost storage, high-performance queries, ACID transactions, schema enforcement, and governance. A data lakehouse enables data engineers to build reliable and scalable data pipelines that can serve various downstream applications and users, such as data science, machine learning, analytics, and reporting. A data lakehouse leverages the power of Delta Lake, a storage layer that brings reliability and performance to data lakes. References: What is a data lakehouse?, Delta Lake, Lakehouse: A New Generation of Open Platforms that Unify Data Warehousing and Advanced Analytics

NEW QUESTION # 34

A data engineer has configured a Structured Streaming job to read from a table, manipulate the data, and then perform a streaming write into a new table.

The code block used by the data engineer is below:

```
(spark.readStream
  .table("sales")
  .withColumn("avg_price", col("sales") / col("units"))
  .writeStream
  .option("checkpointLocation", checkpointPath)
  .outputMode("complete")
  ._____
  .table("new_sales")
)
```

If the data engineer only wants the query to process all of the available data in as many batches as required, which of the following lines of code should the data engineer use to fill in the blank?

- A. `trigger(continuous="once")`
- B. `trigger(processingTime="once")`
- C. `trigger(parallelBatch=True)`
- **D. `trigger(availableNow=True)`**
- E. `processingTime(1)`

Answer: D

Explanation:

<https://spark.apache.org/docs/latest/api/python/reference/pyspark.ss/api/pyspark.sql.streaming.DataStreamWriter.trigger.html>

NEW QUESTION # 35

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

employeeId	startDate	avgRating
a1	2009-01-06	5.5
a2	2018-11-21	7.1
...

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?

```
CREATE TABLE IF NOT EXISTS table_name (  
    employeeId STRING,  
A.   startDate DATE,  
    avgRating FLOAT  
)  
  
CREATE OR REPLACE TABLE table_name AS  
SELECT  
B.   employeeId STRING,  
    startDate DATE,  
    avgRating FLOAT  
USING DELTA  
  
CREATE OR REPLACE TABLE table_name WITH COLUMNS (  
    employeeId STRING,  
C.   startDate DATE,  
    avgRating FLOAT  
) USING DELTA  
  
CREATE TABLE table_name AS  
SELECT  
D.   employeeId STRING,  
    startDate DATE,  
    avgRating FLOAT  
  
CREATE OR REPLACE TABLE table_name (  
    employeeId STRING,  
E.   startDate DATE,  
    avgRating FLOAT  
)  
  
databricks
```

- A. Option A
- **B. Option E**
- C. Option C
- D. Option D
- E. Option B

Answer: B

NEW QUESTION # 36

A data engineer is working with two tables. Each of these tables is displayed below in its entirety.

sales		
customer_id	spend	units
a1	28.94	7
a3	874.12	23
a4	8.99	1

favorite_stores	
customer_id	store_id
a1	s1
a2	s1
a4	s2

The data engineer runs the following query to join these tables together:

```
SELECT
    sales.customer_id,
    sales.spend,
    favorite_stores.store_id
FROM sales
LEFT JOIN favorite_stores
ON sales.customer_id = favorite_stores.customer_id;
```

Which of the following will be returned by the above query?

A.

customer_id	spend	store_id
a1	28.94	s1
a4	8.99	s2

B.

customer_id	spend	units	store_id
a1	28.94	7	s1
a4	8.99	1	s2

C.

customer_id	spend	store_id
a1	28.94	s1
a3	874.12	NULL
a4	8.99	s2

D.

customer_id	spend	store_id
a1	28.94	s1
a2	NULL	s1
a3	874.12	NULL
a4	8.99	s2

E.

customer_id	spend	store_id
a1	28.94	s1
a2	NULL	s1
a4	8.99	s2

- A. Option A
- B. Option C
- C. Option E
- D. Option D
- E. Option B

Answer: A

Explanation:

Option A is the correct answer because it shows the result of an INNER JOIN between the two tables. An INNER JOIN returns only the rows that have matching values in both tables based on the join condition. In this case, the join condition is ON a.customer_id = c.customer_id, which means that only the rows that have the same customer ID in both tables will be included in the output. The output will have four columns:

customer_id, name, account_id, and overdraft_amt. The output will have four rows, corresponding to the four customers who have accounts in the account table.

The use of INNER JOIN can be referenced from Databricks documentation on SQL JOIN or from other sources like W3Schools or GeeksforGeeks.

NEW QUESTION # 37

.....

You only need 20-30 hours to learn our Databricks-Certified-Data-Engineer-Associate test torrents and prepare for the exam.

