DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE^{Q&As}

Databricks Certified Data Engineer Associate Exam

Pass Databricks DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

https://www.passapply.com/databricks-certified-data-engineer-associate.html

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Databricks Official Exam Center

DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE VCE Dumps | DATABRICKS-CERTIFIED-DATA-
ENGINEER-ASSOCIATE Practice Test | DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE Exam
Questions

1 / 6

DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE VCE Dumps | DATABRICKS-CERTIFIED-DATA-
ENGINEER-ASSOCIATE Practice Test | DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE Exam
Questions

2 / 6

https://www.passapply.com/databricks-certified-data-engineer-associate.htm
2024 Latest passapply DATABRICKS-CERTIFIED-DATA-ENGINEER-
ASSOCIATE PDF and VCE dumps Download

QUESTION 1

A data engineering team has two tables. The first table march_transactions is a collection of all retail transactions in the month of March. The second table april_transactions is a collection of all retail transactions in the month of April. There are no duplicate records between the tables.

Which of the following commands should be run to create a new table all_transactions that contains all records from march_transactions and april_transactions without duplicate records?

A. CREATE TABLE all_transactions AS SELECT * FROM march_transactions INNER JOIN SELECT * FROM april_transactions;

B. CREATE TABLE all_transactions AS SELECT * FROM march_transactions UNION SELECT * FROM april_transactions;

C. CREATE TABLE all_transactions AS SELECT * FROM march_transactions OUTER JOIN SELECT * FROM april_transactions;

D. CREATE TABLE all_transactions AS SELECT * FROM march_transactions INTERSECT SELECT * from april_transactions;

E. CREATE TABLE all_transactions AS SELECT * FROM march_transactions MERGE SELECT * FROM april_transactions;

Correct Answer: B

Explanation: To create a new table all_transactions that contains all records from march_transactions and april_transactions without duplicate records, you should use the UNION operator, as shown in option B. This operator combines the result sets of the two tables while automatically removing duplicate records.

QUESTION 2

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 cade block used by the data engineer is below:

```
(spark.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 execute a micro-batch to process data every 5 seconds, which of the following lines of code should the data engineer use to fill in the blank?

DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE VCE Dumps | DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE Practice Test | DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE Exam Questions
3 / 6

A. trigger("5 seconds")

B. trigger()

C. trigger(once="5 seconds")

D. trigger(processingTime="5 seconds")

E. trigger(continuous="5 seconds")

Correct Answer: D

Explanation: # ProcessingTime trigger with two-seconds micro-batch interval df.writeStream \

format("console") \

trigger(processingTime=\\'2 seconds\\') \

start()

https://spark.apache.org/docs/latest/structured-streaming-programming-guide.html#triggers

QUESTION 3

A data engineer needs to create a table in Databricks using data from their organization\\'s existing SQLite database.
They run the following command:

```
CREATE TABLE jdbc_customer360
USING _____
OPTIONS (
  url "jdbc:sqlite:/customers.db",
  dbtable "customer360"
)
```

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

A. org.apache.spark.sql.jdbc

B. autoloader

C. DELTA

D. sqlite

E. org.apache.spark.sql.sqlite

Correct Answer: A

CREATE TABLE new_employees_table

USING JDBC

OPTIONS (

url "",

dbtable "",

user \\'\\',

password \\'\\'

) AS

SELECT * FROM employees_table_vw

https://docs.databricks.com/external-data/jdbc.html#language-sql

---

QUESTION 4

A data engineer is running code in a Databricks Repo that is cloned from a central Git repository. A colleague of the data engineer informs them that changes have been made and synced to the central Git repository. The data engineer now needs to sync their Databricks Repo to get the changes from the central Git repository.

Which of the following Git operations does the data engineer need to run to accomplish this task?

A. Merge

B. Push

C. Pull

D. Commit

E. Clone

Correct Answer: C

Explanation: From the docs:

In Databricks Repos, you can use Git functionality to:

Clone, push to, and pull from a remote Git repository. Create and manage branches for development work, including merging, rebasing, and resolving conflicts.

Create notebooksandmdash;including IPYNB notebooksandmdash;and edit them and other files.

Visually compare differences upon commit and resolve merge conflicts.

Source: https://docs.databricks.com/en/repos/index.html

QUESTION 5

A data engineer is maintaining a data pipeline. Upon data ingestion, the data engineer notices that the source data is starting to have a lower level of quality. The data engineer would like to automate the process of monitoring the quality

---

[DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE VCE Dumps](#) | [DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE Practice Test](#) | [DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE Exam Questions](#)

5 / 6

level.

Which of the following tools can the data engineer use to solve this problem?

A. Unity Catalog

B. Data Explorer

C. Delta Lake

D. Delta Live Tables

E. Auto Loader

Correct Answer: D

Explanation: https://docs.databricks.com/delta-live-tables/expectations.html Delta Live Tables is a tool provided by Databricks that can help data engineers automate the monitoring of data quality. It is designed for managing data pipelines, monitoring data quality, and automating workflows. With Delta Live Tables, you can set up data quality checks and alerts to detect issues and anomalies in your data as it is ingested and processed in real-time. It provides a way to ensure that the data quality meets your desired standards and can trigger actions or notifications when issues are detected. While the other tools mentioned may have their own purposes in a data engineeringenvironment, Delta Live Tables is specifically designed for data quality monitoring and automation within the Databricks ecosystem.

DATABRICKS-CERTIFIED-
DATA-ENGINEER-
ASSOCIATE VCE Dumps

DATABRICKS-CERTIFIED-
DATA-ENGINEER-
ASSOCIATE Practice Test

DATABRICKS-CERTIFIED-
DATA-ENGINEER-
ASSOCIATE Exam
Questions