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.pass2lead.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 PDF Dumps | DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE VCE Dumps | DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE Study Guide

1 / 5

![Pass2Lead Logo](https://Pass2Lead.com)
- Instant Download After Purchase
- 100% Money Back Guarantee
- 365 Days Free Update
- 800,000+ Satisfied Customers

QUESTION 1

A Delta Live Table pipeline includes two datasets defined using STREAMING LIVE TABLE. Three datasets are defined against Delta Lake table sources using LIVE TABLE.

The table is configured to run in Production mode using the Continuous Pipeline Mode.

Assuming previously unprocessed data exists and all definitions are valid, what is the expected outcome after clicking Start to update the pipeline?

A. All datasets will be updated at set intervals until the pipeline is shut down. The compute resources will persist to allow for additional testing.

B. All datasets will be updated once and the pipeline will persist without any processing. The compute resources will persist but go unused.

C. All datasets will be updated at set intervals until the pipeline is shut down. The compute resources will be deployed for the update and terminated when the pipeline is stopped.

D. All datasets will be updated once and the pipeline will shut down. The compute resources will be terminated.

E. All datasets will be updated once and the pipeline will shut down. The compute resources will persist to allow for additional testing.

Correct Answer: E

QUESTION 2

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

DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE PDF Dumps | DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE VCE Dumps | DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE Study Guide

3 / 5

E. org.apache.spark.sql.sqlite

Correct Answer: E

QUESTION 3

A data engineer has a single-task Job that runs each morning before they begin working. After identifying an upstream data issue, they need to set up another task to run a new notebook prior to the original task.

Which of the following approaches can the data engineer use to set up the new task?

A. They can clone the existing task in the existing Job and update it to run the new notebook.

B. They can create a new task in the existing Job and then add it as a dependency of the original task.

C. They can create a new task in the existing Job and then add the original task as a dependency of the new task.

D. They can create a new job from scratch and add both tasks to run concurrently.

E. They can clone the existing task to a new Job and then edit it to run the new notebook.

Correct Answer: E

QUESTION 4

A data analyst has created a Delta table sales that is used by the entire data analysis team. They want help from the data engineering team to implement a series of tests to ensure the data is clean. However, the data engineering team uses Python for its tests rather than SQL.

Which of the following commands could the data engineering team use to access sales in PySpark?

A. SELECT * FROM sales

B. There is no way to share data between PySpark and SQL.

C. spark.sql("sales")

D. spark.delta.table("sales")

E. spark.table("sales")

Correct Answer: D

QUESTION 5

A dataset has been defined using Delta Live Tables and includes an expectations clause:

CONSTRAINT valid_timestamp EXPECT (timestamp > \\'2020-01-01\\') ON VIOLATION DROP ROW

What is the expected behavior when a batch of data containing data that violates these constraints is processed?

DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE PDF Dumps | DATABRICKS-CERTIFIED-DATA-
ENGINEER-ASSOCIATE VCE Dumps | DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE Study Guide

4 / 5

![Pass2Lead](https://Pass2Lead.com)
A. Records that violate the expectation are dropped from the target dataset and loaded into a quarantine table.

B. Records that violate the expectation are added to the target dataset and flagged as invalid in a field added to the target dataset.

C. Records that violate the expectation are dropped from the target dataset and recorded as invalid in the event log.

D. Records that violate the expectation are added to the target dataset and recorded as invalid in the event log.

E. Records that violate the expectation cause the job to fail.

Correct Answer: D

DATABRICKS-CERTIFIED-
DATA-ENGINEER-
ASSOCIATE PDF Dumps

DATABRICKS-CERTIFIED-
DATA-ENGINEER-
ASSOCIATE VCE Dumps

DATABRICKS-CERTIFIED-
DATA-ENGINEER-
ASSOCIATE Study Guide

DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE PDF Dumps | DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE VCE Dumps | DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE Study Guide

5 / 5