



VCE & PDF

Pass4itSure.com

[https://www.pass4itsure.com/databricks-certified-data-engineer-associate.h](https://www.pass4itsure.com/databricks-certified-data-engineer-associate.html)  
2024 Latest pass4itsure DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE PDF and VCE dumps Download

---

# DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE<sup>Q&As</sup>

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.pass4itsure.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



VCE & PDF

Pass4itSure.com

<https://www.pass4itsure.com/databricks-certified-data-engineer-associate.h>  
2024 Latest pass4itsure DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE PDF and VCE dumps Download

- ⚙️ **Instant Download** After Purchase
- ⚙️ **100% Money Back** Guarantee
- ⚙️ **365 Days** Free Update
- ⚙️ **800,000+** Satisfied Customers





### QUESTION 1

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 2

Which of the following Git operations must be performed outside of Databricks Repos?

- A. Commit
- B. Pull
- C. Push
- D. Clone
- E. Merge

Correct Answer: D

Reference: <https://docs.databricks.com/repos/repos-setup.html>

---

### QUESTION 3

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

#### QUESTION 4

An engineering manager wants to monitor the performance of a recent project using a Databricks SQL query. For the first week following the project's release, the manager wants the query results to be updated every minute. However, the manager is concerned that the compute resources used for the query will be left running and cost the organization a lot of money beyond the first week of the project's release.

Which of the following approaches can the engineering team use to ensure the query does not cost the organization any money beyond the first week of the project's release?

- A. They can set a limit to the number of DBUs that are consumed by the SQL Endpoint.
- B. They can set the query's refresh schedule to end after a certain number of refreshes.
- C. They cannot ensure the query does not cost the organization money beyond the first week of the project's release.
- D. They can set a limit to the number of individuals that are able to manage the query's refresh schedule.
- E. They can set the query's refresh schedule to end on a certain date in the query scheduler.

Correct Answer: E

#### QUESTION 5

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.table("sales")
  .withColumn("avg_price", col("sales") / col("units"))
  .writeStream
  .option("checkpointLocation", checkpointPath)
  .outputMode("complete")
  .trigger(Trigger.Once())
  .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?

- A. `trigger("5 seconds")`
- B. `trigger()`
- C. `trigger(once="5 seconds")`
- D. `trigger(processingTime="5 seconds")`
- E. `trigger(continuous="5 seconds")`

Correct Answer: D

[DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE Practice Test](#)

[DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE Exam Questions](#)

[DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE Braindumps](#)