



VCE & PDF

GeekCert.com

<https://www.geekcert.com/databricks-certified-data-engineer-associate.html>

2024 Latest geekcert DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE PDF and VCE dumps Download

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.geekcert.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

GeekCert.com

<https://www.geekcert.com/databricks-certified-data-engineer-associate.html>

2024 Latest geekcert 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 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(_____)
  .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

QUESTION 2

Which of the following tools is used by Auto Loader process data incrementally?

- A. Checkpointing
- B. Spark Structured Streaming
- C. Data Explorer
- D. Unity Catalog
- E. Databricks SQL

Correct Answer: B



QUESTION 3

Which of the following Structured Streaming queries is performing a hop from a Silver table to a Gold table?

A.

```
(spark.readStream.load(rawSalesLocation)
  .writeStream
  .option("checkpointLocation", checkpointPath)
  .outputMode("append")
  .table("newSales")
)
```

B.

```
(spark.read.load(rawSalesLocation)
  .writeStream
  .option("checkpointLocation", checkpointPath)
  .outputMode("append")
  .table("newSales")
)
```

C.

```
(spark.table("sales")
  .withColumn("avgPrice", col("sales") / col("units"))
  .writeStream
  .option("checkpointLocation", checkpointPath)
  .outputMode("append")
  .table("newSales")
)
```

D.

```
(spark.table("sales")
  .filter(col("units") > 0)
  .writeStream
  .option("checkpointLocation", checkpointPath)
  .outputMode("append")
  .table("newSales")
)
```

E.

```
(spark.table("sales")
  .groupBy("store")
  .agg(sum("sales"))
  .writeStream
  .option("checkpointLocation", checkpointPath)
  .outputMode("complete")
  .table("newSales")
)
```

A. Option A

B. Option B

C. Option C



D. Option D

E. Option E

Correct Answer: E

QUESTION 4

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?

- 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

QUESTION 5

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

[Latest DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE](#) | [DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE](#) | [DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE](#)



VCE & PDF

GeekCert.com

<https://www.geekcert.com/databricks-certified-data-engineer-associate.html>

2024 Latest geekcert DATABRICKS-CERTIFIED-DATA-ENGINEER-ASSOCIATE PDF and VCE dumps Download

[ASSOCIATE Dumps](#)

[ASSOCIATE Practice Test](#)

[ASSOCIATE Study Guide](#)