



**Free Questions for Databricks-Certified-Data-Engineer-Associate by certsinside**

**Shared by Booth on 24-05-2024**

**For More Free Questions and Preparation Resources**

**Check the Links on Last Page**

# Question 1

---

## Question Type: MultipleChoice

---

A new data engineering team team has been assigned to an ELT project. The new data engineering team will need full privileges on the table sales to fully manage the project.

Which command can be used to grant full permissions on the database to the new data engineering team?

### Options:

---

- A- grant all privileges on table sales TO team;
- B- GRANT SELECT ON TABLE sales TO team;
- C- GRANT SELECT CREATE MODIFY ON TABLE sales TO team;
- D- GRANT ALL PRIVILEGES ON TABLE team TO sales;

### Answer:

---

A

### Explanation:

---

To grant full privileges on a table such as 'sales' to a group like 'team', the correct SQL command in Databricks is:

```
GRANT ALL PRIVILEGES ON TABLE sales TO team;
```

This command assigns all available privileges, including SELECT, INSERT, UPDATE, DELETE, and any other data manipulation or definition actions, to the specified team. This is typically necessary when a team needs full control over a table to manage and manipulate it as part of a project or ongoing maintenance.

Reference: Databricks documentation on SQL permissions: [SQL Permissions in Databricks](#)

## Question 2

---

**Question Type:** MultipleChoice

---

A data engineer needs access to a table new\_table, but they do not have the correct permissions. They can ask the table owner for permission, but they do not know who the table owner is.

Which approach can be used to identify the owner of new\_table?

**Options:**

---

- A- There is no way to identify the owner of the table
- B- Review the Owner field in the table's page in the cloud storage solution
- C- Review the Permissions tab in the table's page in Data Explorer
- D- Review the Owner field in the table's page in Data Explorer

**Answer:**

---

D

**Explanation:**

---

To find the owner of a table in Databricks, one can utilize the Data Explorer feature. The Data Explorer provides detailed information about various data objects, including tables. By navigating to the specific table's page in Data Explorer, a data engineer can review the Owner field, which identifies the individual or role that owns the table. This information is crucial for obtaining the necessary permissions or for any administrative actions related to the table.

Reference: Databricks documentation on Data Explorer: Using Data Explorer in Databricks

## Question 3

---

**Question Type:** MultipleChoice

---

A data engineer wants to schedule their Databricks SQL dashboard to refresh every hour, but they only want the associated SQL endpoint to be running when it is necessary. The dashboard has multiple queries on multiple datasets associated with it. The data that feeds the dashboard is automatically processed using a Databricks Job.

Which approach can the data engineer use to minimize the total running time of the SQL endpoint used in the refresh schedule of their dashboard?

### Options:

---

- A-  They can reduce the cluster size of the SQL endpoint.
- B-  They can turn on the Auto Stop feature for the SQL endpoint.
- C-  They can set up the dashboard's SQL endpoint to be serverless.
- D-  They can ensure the dashboard's SQL endpoint matches each of the queries' SQL endpoints.

### Answer:

---

B

### Explanation:

---

To minimize the total running time of the SQL endpoint used in the refresh schedule of a dashboard in Databricks, the most effective approach is to utilize the Auto Stop feature. This feature allows the SQL endpoint to automatically stop after a period of inactivity,

ensuring that it only runs when necessary, such as during the dashboard refresh or when actively queried. This minimizes resource usage and associated costs by ensuring the SQL endpoint is not running idle outside of these operations.

Reference: Databricks documentation on SQL endpoints: SQL Endpoints in Databricks

## Question 4

---

**Question Type:** MultipleChoice

---

A data engineer and data analyst are working together on a data pipeline. The data engineer is working on the raw, bronze, and silver layers of the pipeline using Python, and the data analyst is working on the gold layer of the pipeline using SQL. The raw source of the pipeline is a streaming input. They now want to migrate their pipeline to use Delta Live Tables.

Which change will need to be made to the pipeline when migrating to Delta Live Tables?

### Options:

---

- A-** The pipeline can have different notebook sources in SQL & Python.
- B-** The pipeline will need to be written entirely in SQL.
- C-** The pipeline will need to be written entirely in Python.

**D-** The pipeline will need to use a batch source in place of a streaming source.

### **Answer:**

---

A

### **Explanation:**

---

When migrating to Delta Live Tables (DLT) with a data pipeline that involves different programming languages across various data layers, the migration does not require unifying the pipeline into a single language. Delta Live Tables support multi-language pipelines, allowing data engineers and data analysts to work in their preferred languages, such as Python for data engineering tasks (raw, bronze, and silver layers) and SQL for data analytics tasks (gold layer). This capability is particularly beneficial in collaborative settings and leverages the strengths of each language for different stages of data processing.

Reference: Databricks documentation on Delta Live Tables: [Delta Live Tables Guide](#)

## **Question 5**

---

**Question Type:** MultipleChoice

---

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.

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

### Options:

---

- A-** All datasets will be updated once and the pipeline will shut down. The compute resources will be terminated.
- B-** All datasets will be updated at set intervals until the pipeline is shut down. The compute resources will persist to allow for additional testing.
- C-** All datasets will be updated once and the pipeline will shut down. The compute resources will persist to allow for additional testing.
- D-** 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.

### Answer:

---

D

### Explanation:

---

In Delta Live Tables (DLT), when configured to run in Continuous Pipeline Mode, particularly in a production environment, the system is designed to continuously process and update data as it becomes available. This mode keeps the compute resources active to handle ongoing data processing and automatically updates all datasets defined in the pipeline at predefined intervals. Once the pipeline is



manually stopped, the compute resources are terminated to conserve resources and reduce costs. This mode is suitable for production environments where datasets need to be kept up-to-date with the latest data.

Reference: Databricks documentation on Delta Live Tables: Delta Live Tables Guide

## Question 6

---

**Question Type:** MultipleChoice

---

Which tool is used by Auto Loader to process data incrementally?

### Options:

---

- A- Spark Structured Streaming
- B- Unity Catalog
- C- Checkpointing
- D- Databricks SQL

### Answer:

---

A

**Explanation:**

---

Auto Loader in Databricks utilizes Spark Structured Streaming for processing data incrementally. This allows Auto Loader to efficiently ingest streaming or batch data at scale and to recognize new data as it arrives in cloud storage. Spark Structured Streaming provides the underlying engine that supports various incremental data loading capabilities like schema inference and file notification mode, which are crucial for the dynamic nature of data lakes.

Reference: [Databricks documentation on Auto Loader: Auto Loader Overview](#)

**To Get Premium Files for Databricks-Certified-Data-Engineer-Associate Visit**

**<https://www.p2pexams.com/products/databricks-certified-data-engineer-associate>**

**For More Free Questions Visit**

**<https://www.p2pexams.com/databricks/pdf/databricks-certified-data-engineer-associate>**

