

# 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.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



https://www.passapply.com/databricks-certified-data-engineer-associate.htm 2024 Latest passapply 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



### 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 engineer has three tables in a Delta Live Tables (DLT) pipeline. They have configured the pipeline to drop invalid records at each table. They notice that some data is being dropped due to quality concerns at some point in the DLT pipeline. They would like to determine at which table in their pipeline the data is being dropped.

Which of the following approaches can the data engineer take to identify the table that is dropping the records?

- A. They can set up separate expectations for each table when developing their DLT pipeline.
- B. They cannot determine which table is dropping the records.
- C. They can set up DLT to notify them via email when records are dropped.
- D. They can navigate to the DLT pipeline page, click on each table, and view the data quality statistics.
- E. They can navigate to the DLT pipeline page, click on the "Error" button, and review the present errors.

Correct Answer: E

#### **QUESTION 2**

A data engineer wants to create a data entity from a couple of tables. The data entity must be used by other data engineers in other sessions. It also must be saved to a physical location.

Which of the following data entities should the data engineer create?

- A. Database
- B. Function
- C. View
- D. Temporary view
- E. Table

Correct Answer: C

#### **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.



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

- 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**

In order for Structured Streaming to reliably track the exact progress of the processing so that it can handle any kind of failure by restarting and/or reprocessing, which of the following two approaches is used by Spark to record the offset range of the data being processed in each trigger?

- A. Checkpointing and Write-ahead Logs
- B. Structured Streaming cannot record the offset range of the data being processed in each trigger.
- C. Replayable Sources and Idempotent Sinks
- D. Write-ahead Logs and Idempotent Sinks
- E. Checkpointing and Idempotent Sinks

Correct Answer: E

#### **QUESTION 5**

A data engineer only wants to execute the final block of a Python program if the Python variable day\_of\_week is equal to 1 and the Python variable review\_period is True.

Which of the following control flow statements should the data engineer use to begin this conditionally executed code block?

A. if day\_of\_week = 1 and review\_period:

B. if day\_of\_week = 1 and review\_period = "True":

C. if day\_of\_week == 1 and review\_period == "True":

D. if day\_of\_week == 1 and review\_period:

E. if day of week = 1 and review period: = "True":

Correct Answer: C

DATA-ENGINEER-ASSOCIATE VCE Dumps

**DATA-ENGINEER-**ASSOCIATE Practice Test

DATABRICKS-CERTIFIED- DATABRICKS-CERTIFIED- DATABRICKS-CERTIFIED-**DATA-ENGINEER-ASSOCIATE Study Guide**