



Up-to-date Questions and Answers from authentic resources to improve knowledge and pass the exam at very first attempt. ----- Guaranteed.



DCDEP Dumps
DCDEP Braindumps
DCDEP Real Questions
DCDEP Practice Test
DCDEP Actual Questions



killexams.com

Databrick

DCDEP

Databricks Certified Data Engineer Professional

ORDER FULL VERSION

<https://killexams.com/pass4sure/exam-detail/DCDEP>



Question: 21

A DELTA LIVE TABLE pipelines can be scheduled to run in two different modes, what are these two different modes?

- A. Triggered, Incremental
- B. Once, Continuous
- C. Triggered, Continuous
- D. Once, Incremental
- E. Continuous, Incremental

Answer: C

Explanation:

The answer is Triggered, Continuous

<https://docs.microsoft.com/en-us/azure/databricks/data-engineering/delta-live-tables/delta-live-tables-concepts#--continuous-and-triggered-pipelines>

â€¢ Triggered pipelines update each table with whatever data is currently available and then stop the cluster running the pipeline. Delta Live Tables automatically analyzes the dependencies between your tables and starts by computing those that read from external sources. Tables within the pipeline are updated after their dependent data sources have been updated.

â€¢ Continuous pipelines update tables continuously as input data changes. Once an update is started, it continues to run until manually stopped. Continuous pipelines require an always-running cluster but ensure that downstream consumers have the most up-to-date data.

Question: 22

Which of the following developer operations in CI/CD flow can be implemented in Databricks Re-pos?

- A. Merge when code is committed
- B. Pull request and review process
- C. Trigger Databricks Repos API to pull the latest version of code into production folder
- D. Resolve merge conflicts
- E. Delete a branch

Answer: C

Explanation:

See the below diagram to understand the role Databricks Repos and Git provider plays when building a CI/CD workflow.

All the steps highlighted in yellow can be done Databricks Repo, all the steps highlighted in Gray are done in a git provider like Github or Azure DevOps

Question: 23

Identify one of the below statements that can query a delta table in PySpark Dataframe API

- A. Spark.read.mode("delta").table("table_name")
- B. Spark.read.table.delta("table_name")
- C. Spark.read.table("table_name")
- D. Spark.read.format("delta").LoadTableAs("table_name")
- E. Spark.read.format("delta").TableAs("table_name")

Answer: C

Question: 24

How VACCUM and OPTIMIZE commands can be used to manage the DELTA lake?

- A. VACCUM command can be used to compact small parquet files, and the OP-TIMZE command can be used to delete parquet files that are marked for deletion/unused.
- B. VACCUM command can be used to delete empty/blank parquet files in a delta table. OPTIMIZE command can be used to update stale statistics on a delta table.
- C. VACCUM command can be used to compress the parquet files to reduce the size of the table, OPTIMIZE command can be used to cache frequently delta tables for better performance.
- D. VACCUM command can be used to delete empty/blank parquet files in a delta table, OPTIMIZE command can be used to cache frequently delta tables for better performance.
- E. OPTIMIZE command can be used to compact small parquet files, and the VAC-CUM command can be used to delete parquet files that are marked for deletion/unused. (Correct)

Answer: E

Explanation:

VACCUM:

You can remove files no longer referenced by a Delta table and are older than the retention threshold by running the vacuum command on the table. vacuum is not triggered automatically. The default retention threshold for the files is 7 days. To change this behavior, see Configure data retention for time travel.

OPTIMIZE:

Using OPTIMIZE you can compact data files on Delta Lake, this can improve the speed of read queries on the table. Too many small files can significantly degrade the performance of the query.

Question: 25

Which of the following statements are correct on how Delta Lake implements a lake house?

- A. Delta lake uses a proprietary format to write data, optimized for cloud storage
- B. Using Apache Hadoop on cloud object storage
- C. Delta lake always stores meta data in memory vs storage
- D. Delta lake uses open source, open format, optimized cloud storage and scalable meta data
- E. Delta lake stores data and meta data in computes memory

Answer: D

Explanation:

Delta lake is

â€ Open source

â€ Builds up on standard data format

â€ Optimized for cloud object storage

â€ Built for scalable metadata handling Delta lake is not

â€ Proprietary technology

â€ Storage format

â€ Storage medium

â€ Database service or data warehouse

Question: 26

What are the different ways you can schedule a job in Databricks workspace?

- A. Continuous, Incremental
- B. On-Demand runs, File notification from Cloud object storage
- C. Cron, On Demand runs
- D. Cron, File notification from Cloud object storage
- E. Once, Continuous

Answer: C

Explanation:

The answer is, Cron, On-Demand runs

Supports running job immediately or using can be scheduled using CRON syntax

Question: 27

Which of the following type of tasks cannot setup through a job?

- A. Notebook
- B. DELTA LIVE PIPELINE
- C. Spark Submit
- D. Python
- E. Databricks SQL Dashboard refresh

Answer: E

Question: 28

Which of the following describes how Databricks Repos can help facilitate CI/CD workflows on the Databricks Lakehouse Platform?

- A. Databricks Repos can facilitate the pull request, review, and approval process before merging branches
- B. Databricks Repos can merge changes from a secondary Git branch into a main Git branch
- C. Databricks Repos can be used to design, develop, and trigger Git automation pipelines
- D. Databricks Repos can store the single-source-of-truth Git repository
- E. Databricks Repos can commit or push code changes to trigger a CI/CD process

Answer: E

Explanation:

Answer is Databricks Repos can commit or push code changes to trigger a CI/CD process See below diagram to understand the role Databricks Repos and Git provider plays when building a CI/CD workflow.

All the steps highlighted in yellow can be done Databricks Repo, all the steps highlighted in Gray are done in a git provider like Github or Azure Devops.

Diagram

Description automatically generated

SAMPLE QUESTIONS



*These questions are for demo purpose only. **Full version** is up to date and contains actual questions and answers.*

Killexams.com is an online platform that offers a wide range of services related to certification exam preparation. The platform provides actual questions, exam dumps, and practice tests to help individuals prepare for various certification exams with confidence. Here are some key features and services offered by Killexams.com:



Actual Exam Questions: *Killexams.com provides actual exam questions that are experienced in test centers. These questions are updated regularly to ensure they are up-to-date and relevant to the latest exam syllabus. By studying these actual questions, candidates can familiarize themselves with the content and format of the real exam.*

Exam Dumps: *Killexams.com offers exam dumps in PDF format. These dumps contain a comprehensive collection of questions and answers that cover the exam topics. By using these dumps, candidates can enhance their knowledge and improve their chances of success in the certification exam.*

Practice Tests: *Killexams.com provides practice tests through their desktop VCE exam simulator and online test engine. These practice tests simulate the real exam environment and help candidates assess their readiness for the actual exam. The practice tests cover a wide range of questions and enable candidates to identify their strengths and weaknesses.*

Guaranteed Success: *Killexams.com offers a success guarantee with their exam dumps. They claim that by using their materials, candidates will pass their exams on the first attempt or they will refund the purchase price. This guarantee provides assurance and confidence to individuals preparing for certification exams.*

Updated Content: *Killexams.com regularly updates its question bank and exam dumps to ensure that they are current and reflect the latest changes in the exam syllabus. This helps candidates stay up-to-date with the exam content and increases their chances of success.*

Technical Support: *Killexams.com provides free 24x7 technical support to assist candidates with any queries or issues they may encounter while using their services. Their certified experts are available to provide guidance and help candidates throughout their exam preparation journey.*

For More exams visit <https://killexams.com/vendors-exam-list>
Kill your exam at First Attempt....Guaranteed!