

Realistic Databricks Databricks-Certified-Professional-Data-Engineer Updated Testkings - Databricks Certified Professional Data Engineer Exam Original Questions

100% Pass Quiz



Our Databricks Certified Professional Data Engineer Exam (Databricks-Certified-Professional-Data-Engineer) exam dumps are top-notch and designed to help students pass the Databricks Certified Professional Data Engineer Exam (Databricks-Certified-Professional-Data-Engineer) test on the first try. ActualTestsIT offers three formats of preparation material for the Databricks-Certified-Professional-Data-Engineer exam: Databricks Databricks-Certified-Professional-Data-Engineer Pdf Dumps format, desktop-based Databricks-Certified-Professional-Data-Engineer practice exam software, and web-based Databricks Certified Professional Data Engineer Exam (Databricks-Certified-Professional-Data-Engineer) practice test. These Databricks-Certified-Professional-Data-Engineer exam dumps formats are designed to suit the needs of different types of students.

We can promise that our Databricks-Certified-Professional-Data-Engineer exam questions are always the latest and valid for we are always trying to do better for our worthy customers. The first and the most important thing is to make sure the high-quality of our Databricks-Certified-Professional-Data-Engineer learning guide and keep it updated on time. Once any new question is found, we will send you a link to download a new version of the Databricks-Certified-Professional-Data-Engineer Training Materials. So don't worry if you are left behind the trend. Experts in our company won't let this happen.

>> **Databricks-Certified-Professional-Data-Engineer Updated Testkings** <<

Pass Guaranteed 2026 Databricks Authoritative Databricks-Certified-Professional-Data-Engineer: Databricks Certified Professional Data Engineer Exam Updated Testkings

Do you want to find a job that really fulfills your ambitions? That's because you haven't found an opportunity to improve your ability to lay a solid foundation for a good career. Our Databricks-Certified-Professional-Data-Engineer quiz torrent can help you get out of trouble regain confidence and embrace a better life. Our Databricks-Certified-Professional-Data-Engineer exam question can help you learn effectively and ultimately obtain the authority certification of Databricks, which will fully prove your ability and let you stand out in the labor market. We have the confidence and ability to make you finally have rich rewards. Our Databricks-Certified-Professional-Data-Engineer Learning Materials provide you with a platform of knowledge to help you achieve your wishes.

Databricks Certified Professional Data Engineer certification exam is designed for data engineers who work with Databricks. Databricks-Certified-Professional-Data-Engineer exam tests the candidate's ability to design, build, and maintain data pipelines, as well as their knowledge of various data engineering tools and techniques. Databricks-Certified-Professional-Data-Engineer Exam is intended to validate the candidate's proficiency in using Databricks for data engineering tasks.

Databricks Certified Professional Data Engineer Exam Sample Questions (Q116-Q121):

NEW QUESTION # 116

A data team's Structured Streaming job is configured to calculate running aggregates for item sales to update a downstream

marketing dashboard. The marketing team has introduced a new field to track the number of times this promotion code is used for each item. A junior data engineer suggests updating the existing query as follows: Note that proposed changes are in bold.
Which step must also be completed to put the proposed query into production?

- A. Remove `.option (mergeSchema', true')` from the streaming write
- B. Increase the shuffle partitions to account for additional aggregates
- C. Run `REFRESH TABLE delta, /item_agg'`
- **D. Specify a new checkpointlocation**

Answer: D

Explanation:

When introducing a new aggregation or a change in the logic of a Structured Streaming query, it is generally necessary to specify a new checkpoint location. This is because the checkpoint directory contains metadata about the offsets and the state of the aggregations of a streaming query. If the logic of the query changes, such as including a new aggregation field, the state information saved in the current checkpoint would not be compatible with the new logic, potentially leading to incorrect results or failures. Therefore, to accommodate the new field and ensure the streaming job has the correct starting point and state information for aggregations, a new checkpoint location should be specified.

Reference:

Databricks documentation on Structured Streaming: <https://docs.databricks.com/spark/latest/structured-streaming/index.html>

Databricks documentation on streaming checkpoints: <https://docs.databricks.com/spark/latest/structured-streaming/production.html#checkpointing>

NEW QUESTION # 117

A team of data engineer are adding tables to a DLT pipeline that contain repetitive expectations for many of the same data quality checks.

One member of the team suggests reusing these data quality rules across all tables defined for this pipeline.

What approach would allow them to do this?

- A. Use global Python variables to make expectations visible across DLT notebooks included in the same pipeline.
- B. Maintain data quality rules in a separate Databricks notebook that each DLT notebook of file.
- **C. Maintain data quality rules in a Delta table outside of this pipeline's target schema, providing the schema name as a pipeline parameter.**
- D. Add data quality constraints to tables in this pipeline using an external job with access to pipeline configuration files.

Answer: C

Explanation:

Maintaining data quality rules in a centralized Delta table allows for the reuse of these rules across multiple DLT (Delta Live Tables) pipelines. By storing these rules outside the pipeline's target schema and referencing the schema name as a pipeline parameter, the team can apply the same set of data quality checks to different tables within the pipeline. This approach ensures consistency in data quality validations and reduces redundancy in code by not having to replicate the same rules in each DLT notebook or file.

:

Databricks Documentation on Delta Live Tables: Delta Live Tables Guide

NEW QUESTION # 118

Which REST API call can be used to review the notebooks configured to run as tasks in a multi-task job?

- A. `/jobs/runs/list`
- B. `/jobs/runs/get-output`
- **C. `/jobs/get`**
- D. `/jobs/runs/get`
- E. `/jobs/list`

Answer: C

Explanation:

This is the correct answer because it is the REST API call that can be used to review the notebooks configured to run as tasks in a multi-task job. The REST API is an interface that allows programmatically interacting with Databricks resources, such as clusters,

jobs, notebooks, or tables. The REST API uses HTTP methods, such as GET, POST, PUT, or DELETE, to perform operations on these resources. The /jobs/get endpoint is a GET method that returns information about a job given its job ID. The information includes the job settings, such as the name, schedule, timeout, retries, email notifications, and tasks. The tasks are the units of work that a job executes. A task can be a notebook task, which runs a notebook with specified parameters; a jar task, which runs a JAR uploaded to DBFS with specified main class and arguments; or a python task, which runs a Python file uploaded to DBFS with specified parameters. A multi-task job is a job that has more than one task configured to run in a specific order or in parallel. By using the /jobs/get endpoint, one can review the notebooks configured to run as tasks in a multi-task job. Verified Reference: [Databricks Certified Data Engineer Professional], under "Databricks Jobs" section; Databricks Documentation, under "Get" section; Databricks Documentation, under "JobSettings" section.

NEW QUESTION # 119

The DevOps team has configured a production workload as a collection of notebooks scheduled to run daily using the Jobs UI. A new data engineering hire is onboarding to the team and has requested access to one of these notebooks to review the production logic.

What are the maximum notebook permissions that can be granted to the user without allowing accidental changes to production code or data?

- A. No permissions
- B. Can Run
- C. Can Manage
- D. Can Edit
- E. **Can Read**

Answer: E

Explanation:

This is the correct answer because it is the maximum notebook permissions that can be granted to the user without allowing accidental changes to production code or data. Notebook permissions are used to control access to notebooks in Databricks workspaces. There are four types of notebook permissions: Can Manage, Can Edit, Can Run, and Can Read. Can Manage allows full control over the notebook, including editing, running, deleting, exporting, and changing permissions. Can Edit allows modifying and running the notebook, but not changing permissions or deleting it. Can Run allows executing commands in an existing cluster attached to the notebook, but not modifying or exporting it. Can Read allows viewing the notebook content, but not running or modifying it. In this case, granting Can Read permission to the user will allow them to review the production logic in the notebook without allowing them to make any changes to it or run any commands that may affect production data. Verified Reference: [Databricks Certified Data Engineer Professional], under "Databricks Workspace" section; Databricks Documentation, under "Notebook permissions" section.

NEW QUESTION # 120

You had AUTO LOADER to process millions of files a day and noticed slowness in load process, so you scaled up the Databricks cluster but realized the performance of the Auto loader is still not improving, what is the best way to resolve this.

- A. Copy the data from cloud storage to local disk on the cluster for faster access
- B. Setup a second AUTO LOADER process to process the data
- C. Merge files to one large file
- D. AUTO LOADER is not suitable to process millions of files a day
- E. **Increase the maxFilesPerTrigger option to a sufficiently high number**

Answer: E

Explanation:

Explanation

The default value of maxFilesPerTrigger is 1000 it can be increased to a much higher number but will require a much larger compute to process.

Graphical user interface, text, application, email Description automatically generated

<https://docs.databricks.com/ingestion/auto-loader/options.html>

NEW QUESTION # 121

Our Databricks Databricks-Certified-Professional-Data-Engineer practice exam simulator mirrors the Databricks-Certified-Professional-Data-Engineer exam experience, so you know what to anticipate on Databricks-Certified-Professional-Data-Engineer certification exam day. Our Databricks Certified Professional Data Engineer Exam (Databricks-Certified-Professional-Data-Engineer) practice test software features various question styles and levels, so you can customize your Databricks Databricks-Certified-Professional-Data-Engineer exam questions preparation to meet your needs.

Databricks-Certified-Professional-Data-Engineer Original Questions: <https://www.actualtestsit.com/Databricks/Databricks-Certified-Professional-Data-Engineer-exam-prep-dumps.html>