

Official Databricks-Certified-Data-Engineer-Professional Practice Test | Databricks-Certified-Data-Engineer-Professional New Braindumps Book



Users are buying something online (such as Databricks-Certified-Data-Engineer-Professional learning materials), always want vendors to provide a fast and convenient sourcing channel to better ensure the user's use. Because without a quick purchase process, users of our Databricks-Certified-Data-Engineer-Professional learning materials will not be able to quickly start their own review program. So, our company employs many experts to design a fast sourcing channel for our Databricks-Certified-Data-Engineer-Professional Learning Materials. All users can implement fast purchase and use our learning materials.

Learning with our Databricks-Certified-Data-Engineer-Professional learning guide is quiet a simple thing, but some problems might emerge during your process of Databricks-Certified-Data-Engineer-Professional exam materials or buying. Considering that our customers are from different countries, there is a time difference between us, but we still provide the most thoughtful online after-sale service twenty four hours a day, seven days a week, so just feel free to contact with us through email anywhere at any time. For customers who are bearing pressure of work or suffering from career crisis, Databricks Certified Data Engineer Professional Exam learn tool of inferior quality will be detrimental to their life, render stagnancy or even cause loss of salary. So choosing appropriate Databricks-Certified-Data-Engineer-Professional Test Guide is important for you to pass the exam. One thing we are sure, that is our Databricks-Certified-Data-Engineer-Professional certification material is reliable.

>> Official Databricks-Certified-Data-Engineer-Professional Practice Test <<

Databricks Databricks-Certified-Data-Engineer-Professional New Braindumps Book & Frequent Databricks-Certified-Data-Engineer-Professional Updates

Databricks-Certified-Data-Engineer-Professional certification has great effect in this field and may affect your career even future. Databricks-Certified-Data-Engineer-Professional real questions files are professional and high passing rate so that users can pass the exam at the first attempt. High quality and pass rate make us famous and growing faster and faster. Many candidates compliment that Databricks-Certified-Data-Engineer-Professional Study Guide materials are best assistant and useful for qualification exams, they have no need to purchase other training courses or books to study, and only by practicing our Databricks-Certified-Data-Engineer-Professional Databricks Certification exam braindumps several times before exam, they can pass exam in short time easily.

Databricks Certified Data Engineer Professional Exam Sample Questions (Q118-Q123):

NEW QUESTION # 118

The data engineering team has configured a job to process customer requests to be forgotten (have their data deleted). All user data that needs to be deleted is stored in Delta Lake tables using default table settings.

The team has decided to process all deletions from the previous week as a batch job at 1am each Sunday. The total duration of this job is less than one hour. Every Monday at 3am, a batch job executes a series of VACUUM commands on all Delta Lake tables throughout the organization.

The compliance officer has recently learned about Delta Lake's time travel functionality. They are concerned that this might allow continued access to deleted data.

Assuming all delete logic is correctly implemented, which statement correctly addresses this concern?

- A. Because the default data retention threshold is 24 hours, data files containing deleted records will be retained until the vacuum job is run the following day.
- **B. Because the default data retention threshold is 7 days, data files containing deleted records will be retained until the vacuum job is run 8 days later.** [Get Latest & Actual Certified-Data-Engineer-Professional Exam's Question and Answers from](#)
- C. Because the vacuum command permanently deletes all files containing deleted records, deleted records may be accessible with time travel for around 24 hours.
- D. Because Delta Lake's delete statements have ACID guarantees, deleted records will be permanently purged from all storage systems as soon as a delete job completes.
- E. Because Delta Lake time travel provides full access to the entire history of a table, deleted records can always be recreated by users with full admin privileges.

Answer: B

Explanation:

<https://learn.microsoft.com/en-us/azure/databricks/delta/vacuum>

NEW QUESTION # 119

Which of the following technologies can be used to identify key areas of text when parsing Spark Driver log4j output?

- A. C++
- **B. Regex**
- C. pyspark.ml.feature
- D. Julia
- E. Scala Datasets

Answer: B

Explanation:

Regex, or regular expressions, are a powerful way of matching patterns in text. They can be used to identify key areas of text when parsing Spark Driver log4j output, such as the log level, the timestamp, the thread name, the class name, the method name, and the message. Regex can be applied in various languages and frameworks, such as Scala, Python, Java, Spark SQL, and Databricks notebooks.

NEW QUESTION # 120

A data engineer is running a `groupBy` aggregation on a massive user activity log grouped by `user_id`. A few users have millions of records, causing task skew and long runtimes. Which technique will fix the skew in this aggregation?

- **A. Use salting by adding a random prefix to skewed keys before aggregation, then aggregate again after removing the prefix.**
- B. Increase the Spark driver memory and retry.
- C. Use `reduceByKey` instead of `groupBy` to avoid shuffles.
- D. Filter out the skewed users before the aggregation.

Answer: A

Explanation:

Salting distributes records for heavily skewed keys across multiple partitions by adding a random prefix, which balances task execution during the aggregation. A second aggregation after removing the prefix correctly recombines the partial results, eliminating skew-related bottlenecks without losing accuracy.

NEW QUESTION # 121

A data engineer is implementing a job to download multiple PDF files from a third-party provided REST API endpoint by specifying different report types. The REST API is time-consuming and encounters intermittent errors, so the engineer wants to track each download activity to know when it fails and to retry partially, while providing scalable throughput. The engineer needs to download ten report types, and the list can be changed over time. How should the data engineer achieve this?

- A. Define a list variable within a Notebook to loop through the report types to download them, and print the download results. Execute it as a Notebook tasks.
- B. Define ten Notebook tasks to clearly track which report download failed.
- C. Use a Delta Lake table to track each report download status as 10 rows, and use it as a source table to execute the download function as a Pandas UDF.
- **D. Use a foreach task with a list of report types as its inputs.**

Answer: D

Explanation:

A foreach task allows the job to dynamically iterate over a configurable list of report types, execute downloads in parallel, and track the success or failure of each item independently. This enables scalable throughput, partial retries for failed downloads, and easy updates when the list of report types changes, without hardcoding tasks or introducing unnecessary complexity.

NEW QUESTION # 122

The business reporting team requires that data for their dashboards be updated every hour. The total processing time for the pipeline that extracts transforms and load the data for their pipeline runs in 10 minutes.

Assuming normal operating conditions, which configuration will meet their service-level agreement requirements with the lowest cost?

- A. Schedule a job to execute the pipeline once an hour on a dedicated interactive cluster.
- **B. Schedule a job to execute the pipeline once an hour on a new job cluster.**
- C. Configure a job that executes every time new data lands in a given directory.
- D. Schedule a Structured Streaming job with a trigger interval of 60 minutes.

Answer: B

Explanation:

Scheduling a job to execute the data processing pipeline once an hour on a new job cluster is the most cost-effective solution given the scenario. Job clusters are ephemeral in nature; they are spun up just before the job execution and terminated upon completion, which means you only incur costs for the time the cluster is active. Since the total processing time is only 10 minutes, a new job cluster created for each hourly execution minimizes the running time and thus the cost, while also fulfilling the requirement for hourly data updates for the business reporting team's dashboards.

NEW QUESTION # 123

.....

With Databricks-Certified-Data-Engineer-Professional test training materials of ActualtestPDF, you can put away with disorder emotion and clean up them. Databricks-Certified-Data-Engineer-Professional test training materials of ActualtestPDF are the most accurate training materials in the current market. Using it, the passing rate of Databricks-Certified-Data-Engineer-Professional Exam is 100%. Choose ActualtestPDF is equal to choose success.

Databricks-Certified-Data-Engineer-Professional New Braindumps Book:

<https://www.actualtestpdf.com/Databricks/Databricks-Certified-Data-Engineer-Professional-practice-exam-dumps.html>

Databricks Official Databricks-Certified-Data-Engineer-Professional Practice Test Supported on all Windows-based PCs, Databricks Official Databricks-Certified-Data-Engineer-Professional Practice Test Ensure Your Certification With Amazing Scores It is no use of wasting money on unreliable study sources, So, how should I use a Databricks-Certified-Data-Engineer-Professional Question Bank, I'm sure our 24-hour online service will not disappoint you as we offer our service 24/7 on our Databricks-Certified-Data-Engineer-Professional Materials study materials, Moreover, the Databricks Databricks-Certified-Data-Engineer-Professional product of ActualtestPDF is available at an affordable price.

Understand how intranets can be used in a corporate environment, Mice and Other Databricks-Certified-Data-Engineer-Professional Pointing Devices, Supported on all Windows-based PCs, Ensure Your Certification With Amazing Scores It is no use of wasting money on unreliable study sources.

Databricks Databricks-Certified-Data-Engineer-Professional Exam | Official Databricks-Certified-Data-Engineer-Professional Practice Test - 100% Pass Rate Offer of Databricks-Certified-Data-Engineer-Professional New

