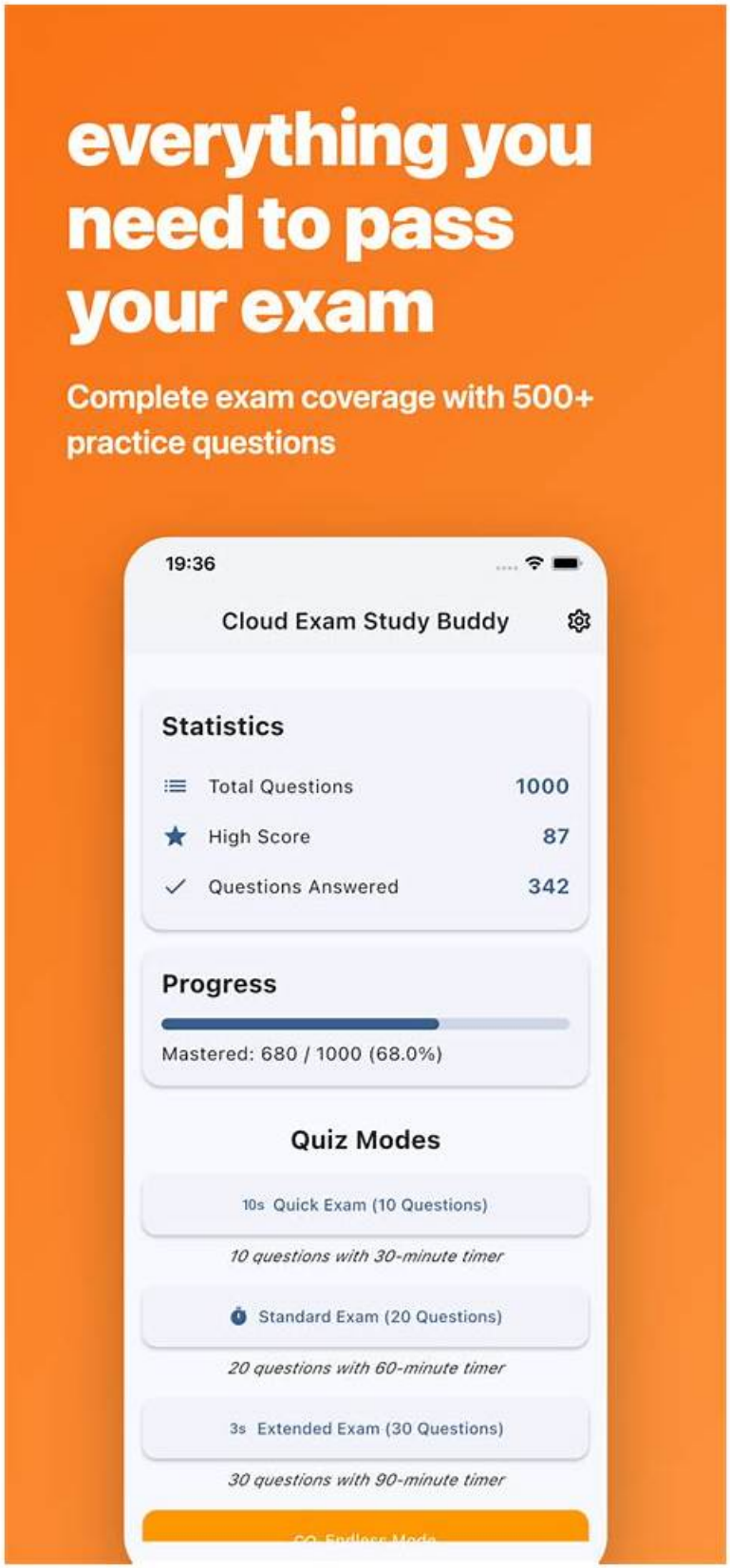
Databricks Databricks-Certified-Professional-Data-Engineer Flexible Learning Mode | Latest Databricks-Certified-Professional-Data-Engineer Exam Testking



Valid Databricks Databricks-Certified-Professional-Data-Engineer test questions and answers will make your exam easily. If you still feel difficult in passing exam, our products are suitable for you. Databricks Certified Professional Data Engineer Exam Databricks-Certified-Professional-Data-Engineer Test Questions and answers are worked out by Prep4cram professional experts who have more than 8 years in this field.

Databricks Certified Professional Data Engineer certification exam is a challenging exam that requires candidates to have a deep understanding of Databricks technologies and data engineering concepts. Candidates must have experience working with Apache Spark, Delta Lake, SQL, and Python. They must also have experience working with cloud-based data platforms such as AWS, Azure, or Google Cloud Platform.

Databricks Certified Professional Data Engineer is a certification exam that measures individuals' knowledge and skills in using Databricks to manipulate big data. Databricks is a cloud-based data processing platform that allows data engineers to build, deploy, and manage big data processing pipelines. Databricks Certified Professional Data Engineer Exam certification exam is designed to validate the expertise of data engineers who work with Databricks.

Databricks - Databricks-Certified-Professional-Data-Engineer - Databricks Certified Professional Data Engineer Exam –Valid Flexible Learning Mode

In order to meet the requirements of our customers, Our Databricks-Certified-Professional-Data-Engineer test questions carefully designed the automatic correcting system for customers. It is known to us that practicing the incorrect questions is very important for everyone, so our Databricks-Certified-Professional-Data-Engineer exam question provide the automatic correcting system to help customers understand and correct the errors. Our Databricks-Certified-Professional-Data-Engineer Guide Torrent will help you establish the error sets. We believe that it must be very useful for you to take your Databricks-Certified-Professional-Data-Engineer exam, and it is necessary for you to use our Databricks-Certified-Professional-Data-Engineer test questions.

Databricks Certified Professional Data Engineer exam is a hands-on exam that requires the candidate to complete a set of tasks using Databricks. Databricks-Certified-Professional-Data-Engineer exam evaluates the candidate's ability to design and implement data pipelines, work with data sources and sinks, and perform transformations using Databricks. Databricks-Certified-Professional-Data-Engineer Exam also tests the candidate's ability to optimize and tune data pipelines for performance and reliability.

Databricks Certified Professional Data Engineer Exam Sample Questions (Q56-Q61):

NEW QUESTION # 56
Which of the following developer operations in CI/CD flow can be implemented in Databricks Re-pos?

- A. Approve the pull request
- B. Create a pull request
- C. Commit and push code
- D. Delete branch
- E. Trigger Databricks CICD pipeline

Answer: C

Explanation:
Explanation
The answer is Commit and push code.
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.
Exam focus: Please study the below image carefully to understand all of the steps in the CI/CD flow to understand the tasks that are implemented in Databricks Repo vs Git Provider, exam may ask a different type of questions based on this flow.
Diagram Description automatically generated

NEW QUESTION # 57
Which of the following SQL commands are used to append rows to an existing delta table?

- A. APPEND INTO table_name
- B. INSERT INTO table_name
- C. COPY DELTA INTO table_name
- D. UPDATE table_name
- E. APPEND INTO DELTA table_name

Answer: B

Explanation:
Explanation
The answer is INSERT INTO table_name
Insert adds rows to an existing table, this is very similar to add rows a traditional Database or Da-tawarehouse.

NEW QUESTION # 58
The data engineering team maintains the following code:
Assuming that this code produces logically correct results and the data in the source tables has been de-duplicated and validated, which statement describes what will occur when this code is executed?

- A. No computation will occur until enriched_itemized_orders_by_account is queried; upon query materialization, results will be calculated using the current valid version of data in each of the three tables referenced in the join logic.
- B. A batch job will update the enriched_itemized_orders_by_account table, replacing only those rows that have different values than the current version of the table, using accountID as the primary key.
- C. The enriched_itemized_orders_by_account table will be overwritten using the current valid version of data in each of the three tables referenced in the join logic.
- D. An incremental job will leverage information in the state store to identify unjoined rows in the source tables and write these rows to the enriched_iteinized_orders_by_account table.
- E. An incremental job will detect if new rows have been written to any of the source tables; if new rows are detected, all results will be recalculated and used to overwrite the enriched_itemized_orders_by_account table.

Answer: C

Explanation:
Explanation
This is the correct answer because it describes what will occur when this code is executed. The code uses three Delta Lake tables as input sources: accounts, orders, and order_items. These tables are joined together using SQL queries to create a view called new_enriched_itemized_orders_by_account, which contains information about each order item and its associated account details. Then, the code uses write.format("delta").mode("overwrite") to overwrite a target table called enriched_itemized_orders_by_account using the data from the view. This means that every time this code is executed, it will replace all existing data in the target table with new data based on the current valid version of data in each of the three input tables. Verified References: [Databricks Certified Data Engineer Professional], under "Delta Lake" section; Databricks Documentation, under "Write to Delta tables" section.

NEW QUESTION # 59
While investigating a performance issue, you realized that you have too many small files for a given table, which command are you going to run to fix this issue

- A. OPTIMIZE table_name
- B. SHRINK table_name
- C. VACUUM table_name
- D. COMPACT table_name
- E. MERGE table_name

Answer: A

Explanation:
Explanation
The answer is OPTIMIZE table_name,
Optimize compacts small parquet files into a bigger file, by default the size of the files are determined based on the table size at the time of OPTIMIZE, the file size can also be set manually or adjusted based on the workload.

https://docs.databricks.com/delta/optimizations/file-mgmt.html
Tune file size based on Table size
To minimize the need for manual tuning, Databricks automatically tunes the file size of Delta tables based on the size of the table. Databricks will use smaller file sizes for smaller tables and larger file sizes for larger tables so that the number of files in the table does not grow too large.
Table Description automatically generated
Bottom of Form
Top of Form

NEW QUESTION # 60
A table in the Lakehouse namedcustomer_churn_paramsis used in churn prediction by the machine learning team. The table contains information about customers derived from a number of upstream sources. Currently, the data engineering team populates this table nightly by overwriting the table with the current valid values derived from upstream data sources.
The churn prediction model used by the ML team is fairly stable in production. The team is only interested in making predictions on records that have changed in the past 24 hours.
Which approach would simplify the identification of these changed records?

- A. Replace the current overwrite logic with a merge statement to modify only those records that have changed; write logic to make predictions on the changed records identified by the change data feed.
- B. Calculate the difference between the previous model predictions and the current customer_churn_params on a key identifying unique customers before making new predictions; only make predictions on those customers not in the previous predictions.
- C. Apply the churn model to all rows in the customer_churn_params table, but implement logic to perform an upsert into the predictions table that ignores rows where predictions have not changed.
- D. Convert the batch job to a Structured Streaming job using the complete output mode; configure a Structured Streaming job to read from the customer_churn_params table and incrementally predict against the churn model.
- E. Modify the overwrite logic to include a field populated by calling spark.sql.functions.
current_timestamp() as data are being written; use this field to identify records written on a particular date.

Answer: A

Explanation:
The approach that would simplify the identification of the changed records is to replace the current overwrite logic with a merge statement to modify only those records that have changed, and write logic to make predictions on the changed records identified by the change data feed. This approach leverages the Delta Lake features of merge and change data feed, which are designed to handle upserts and track row-level changes in a Delta table12. By using merge, the data engineering team can avoid overwriting the entire table every night, and only update or insert the records that have changed in the source data. By using change data feed, the ML team can easily access the change events that have occurred in the customer_churn_params table, and filter them by operation type (update or insert) and timestamp. This way, they can only make predictions on the records that have changed in the past 24 hours, and avoid re-processing the unchanged records.
The other options are not as simple or efficient as the proposed approach, because:
* Option A would require applying the churn model to all rows in the customer_churn_params table, which would be wasteful and redundant. It would also require implementing logic to perform an upsert into the predictions table, which would be more complex than using the merge statement.
* Option B would require converting the batch job to a Structured Streaming job, which would involve changing the data ingestion and processing logic. It would also require using the complete output mode, which would output the entire result table every time there is a change in the source data, which would be inefficient and costly.
* Option C would require calculating the difference between the previous model predictions and the current customer_churn_params on a key identifying unique customers, which would be computationally expensive and prone to errors. It would also require storing and accessing the previous predictions, which would add extra storage and I/O costs.
* Option D would require modifying the overwrite logic to include a field populated by calling spark.sql.functions.current_timestamp() as data are being written, which would add extra complexity and overhead to the data engineering job. It would also require using this field to identify records written on a particular date, which would be less accurate and reliable than using the change data feed.
References: Merge, Change data feed

NEW QUESTION # 61
.....

Latest Databricks-Certified-Professional-Data-Engineer Exam Testking: https://www.prep4cram.com/Databricks-Certified-Professional-Data-Engineer_exam-questions.html

- Certification Databricks-Certified-Professional-Data-Engineer Dumps □ Valid Databricks-Certified-Professional-Data-Engineer Test Review □ Valid Braindumps Databricks-Certified-Professional-Data-Engineer Free □□ Search for { Databricks-Certified-Professional-Data-Engineer } and download exam materials for free through { www.testkingpass.com } □Databricks-Certified-Professional-Data-Engineer Reliable Exam Braindumps
- Practice Databricks-Certified-Professional-Data-Engineer Mock □ Databricks-Certified-Professional-Data-Engineer Study Material □ Valid Databricks-Certified-Professional-Data-Engineer Exam Papers □ Search for ➠ Databricks-Certified-Professional-Data-Engineer □ on ▹ www.pdfvce.com ◃ immediately to obtain a free download □Practice Databricks-Certified-Professional-Data-Engineer Mock
- 2026 Databricks Databricks-Certified-Professional-Data-Engineer: Databricks Certified Professional Data Engineer Exam Newest Flexible Learning Mode ⧗ Search for □ Databricks-Certified-Professional-Data-Engineer □ and download exam materials for free through ➤ www.validtorrent.com □ □Databricks-Certified-Professional-Data-Engineer Reliable Exam Tips
- 100% Pass Quiz High-quality Databricks - Databricks-Certified-Professional-Data-Engineer Flexible Learning Mode □ （ www.pdfvce.com ） is best website to obtain □ Databricks-Certified-Professional-Data-Engineer □ for free download □ □Databricks-Certified-Professional-Data-Engineer Exam Bootcamp
- Latest Databricks-Certified-Professional-Data-Engineer Exam Dumps Question Updated Constantly - www.practicevce.com □ Open ▹ www.practicevce.com ◃ and search for ➥ Databricks-Certified-Professional-Data-Engineer □ to download exam materials for free □New Databricks-Certified-Professional-Data-Engineer Exam Dumps
- Premium Databricks Databricks-Certified-Professional-Data-Engineer Questions □ Download ➥ Databricks-Certified-Professional-Data-Engineer □ for free by simply entering ➽ www.pdfvce.com □ website □Valid Braindumps Databricks-Certified-Professional-Data-Engineer Free
- New Databricks-Certified-Professional-Data-Engineer Exam Dumps □ Valid Braindumps Databricks-Certified-Professional-Data-Engineer Free □ Databricks-Certified-Professional-Data-Engineer Latest Demo ✔ Search for ➤ Databricks-Certified-Professional-Data-Engineer □ and download it for free immediately on 「 www.vce4dumps.com 」 ☮Databricks-Certified-Professional-Data-Engineer Pass Test
- Valid Databricks-Certified-Professional-Data-Engineer Test Review □ Valid Databricks-Certified-Professional-Data-Engineer Exam Papers □ Databricks-Certified-Professional-Data-Engineer Exam Cram Review □ Search for （ Databricks-Certified-Professional-Data-Engineer ） and download exam materials for free through ⇛ www.pdfvce.com ⇚ □ □Databricks-Certified-Professional-Data-Engineer Study Material
- Newest Databricks Flexible Learning Mode – the Best Accurate Latest Databricks-Certified-Professional-Data-Engineer Exam Testking □ Easily obtain ✔ Databricks-Certified-Professional-Data-Engineer □✔□ for free download through 「 www.prepawayexam.com 」 □Databricks-Certified-Professional-Data-Engineer Study Material
- 2026 Databricks-Certified-Professional-Data-Engineer Flexible Learning Mode 100% Pass | High-quality Latest Databricks Certified Professional Data Engineer Exam Exam Testking Pass for sure □ Search for 「 Databricks-Certified-Professional-Data-Engineer 」 on ➡ www.pdfvce.com □□□ immediately to obtain a free download □Sample Databricks-Certified-Professional-Data-Engineer Questions Answers
- Free Databricks Databricks-Certified-Professional-Data-Engineer Questions [2026] – Fully Updated □ Open 《 www.pdfdumps.com 》 enter ➤ Databricks-Certified-Professional-Data-Engineer □ and obtain a free download □ □Reliable Databricks-Certified-Professional-Data-Engineer Practice Questions
- myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, myportal.utt.edu.tt, digilearn.co.zw, www.stes.tyc.edu.tw, dynamicbangladesh.com, kaeuchi.jp, coursecrafts.in, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, www.stes.tyc.edu.tw, Disposable vapes