Databricks-Certified-Professional-Data-Engineer Übungsmaterialien - Databricks-Certified-Professional-Data-Engineer Lernressourcen & Databricks-Certified-Professional-Data-Engineer Prüfungsfragen

P.S. Kostenlose und neue Databricks-Certified-Professional-Data-Engineer Prüfungsfragen sind auf Google Drive freigegeben von Pass4Test verfügbar: https://drive.google.com/open?id=1RnCBv6bs9c5-wgfiM8njGfOlEM7ZY0OU

Bemühen Sie sich noch um die Databricks Databricks-Certified-Professional-Data-Engineer Zertifizierungsprüfung? Wollen Sie schneller Ihren Traum verwirklichen? Bitte wählen Sie die Databricks-Certified-Professional-Data-Engineer Schulungsmaterialien von Pass4Test. Wenn Sie Pass4Test wählen, ist es kein Traum mehr, das Databricks Databricks-Certified-Professional-Data-Engineer Zertifikat zu erhalten.

Die Zertifizierungsprüfung für Datenbanken zertifizierte professionelle Dateningenieur ist für Dateningenieure, Datenarchitekten und Datenwissenschaftler geeignet, die für die Erstellung und Verwaltung von Datenpipelines und Workflows verantwortlich sind. Die Prüfung wurde entwickelt, um das Wissen und die Fähigkeiten zu testen, die zum Entwerfen, Implementieren und Verwalten von Datenentwicklungsworkflows mithilfe von Datenbanken erforderlich sind. Die Kandidaten müssen ein solides Verständnis der Daten technischen Konzepte wie Datenmodellierung, Datenintegration, Datenumwandlung und Datenspeicherung haben.

>> Databricks-Certified-Professional-Data-Engineer Prüfungsunterlagen <<

Databricks-Certified-Professional-Data-Engineer Originale Fragen - Databricks-Certified-Professional-Data-Engineer Online Praxisprüfung

Die Feedbacks von den IT-Fachleuten, die Databricks Databricks-Certified-Professional-Data-Engineer Zertifizierungsprüfung erfolgreich bestanden haben, haben bewiesen, dass ihren Erfolg Pass4Test beizumessen ist. Die Fragen und Antworten zur Databricks Databricks-Certified-Professional-Data-Engineer Zertifizierungsprüfung haben ihnen sehr geholfen. Dabei erspart Pass4Test ihnen auch viele wertvolle Zeit und Energie. Sie haben die Databricks Databricks-Certified-Professional-Data-Engineer Zertifizierungsprüfung ganz mühlos beim ersten Versuch bestanden. So ist Pass4Test eine zuverlässige Website. Wenn Sie Pass4Test wählen, sind Sie der nächste erfolgreiche IT-Fachmann. Pass4Test würde Ihren Traum verwirklichen.

Databricks Certified Professional Data Engineer Exam Databricks-Certified-Professional-Data-Engineer Prüfungsfragen mit Lösungen (Q160-Q165):

160. Frage
A data engineer is using Lakeflow Spark Declarative Pipelines Expectations to track the data quality of incoming sensor data. Periodically, sensors send bad readings that are out of range, and the team is currently flagging those rows with a warning and writing them to the silver table along with the good data. They have been given a new requirement: the bad rows need to be quarantined in a separate quarantine table and no longer included in the silver table.
This is the existing code for the silver table:
@dlt.table
@dlt.expect(" valid_sensor_reading " , " reading < 120 ")

def silver_sensor_readings():
return spark.readStream.table(" bronze_sensor_readings ")
Which code will satisfy the requirements?

- A. @dlt.table
  @dlt.expect(" valid_sensor_reading " , " reading < 120 ")
  def silver_sensor_readings():
  return spark.readStream.table(" bronze_sensor_readings ")
  @dlt.table
  @dlt.expect(" invalid_sensor_reading " , " reading > = 120 ")
  def quarantine_sensor_readings():
  return spark.readStream.table(" bronze_sensor_readings ")
- B. @dlt.table
  @dlt.expect_or_drop(" valid_sensor_reading " , " reading < 120 ")
  def silver_sensor_readings():
  return spark.readStream.table(" bronze_sensor_readings ")
  @dlt.table
  @dlt.expect(" invalid_sensor_reading " , " reading < 120 ")
  def quarantine_sensor_readings():
  return spark.readStream.table(" bronze_sensor_readings ")
- C. @dlt.table
  @dlt.expect_or_drop(" valid_sensor_reading " , " reading < 120 ")
  def silver_sensor_readings():
  return spark.readStream.table(" bronze_sensor_readings ")
- D. @dlt.table
  @dlt.expect_or_drop(" valid_sensor_reading " , " reading < 120 ")
  def silver_sensor_readings():
  return spark.readStream.table(" bronze_sensor_readings ")
  @dlt.table
  @dlt.expect(" invalid_sensor_reading " , " reading > = 120 ")
  def quarantine_sensor_readings():
  return spark.readStream.table(" bronze_sensor_readings ")

Antwort: D

Begründung:
Databricks documents that expect retains invalid records in the target dataset, while expect_or_drop drops invalid records before writing to the target. Therefore, the silver table must use expect_or_drop so bad records are excluded from silver. (Databricks Documentation) Databricks also documents a quarantine pattern in which invalid records are separated for downstream processing, but the fully documented pattern uses an intermediate quarantine dataset with an is_quarantined flag and then derives valid and invalid paths from it. None of the listed options exactly matches the official quarantine pattern. As written, option B is the closest intended answer because it at least creates a separate quarantine table and removes invalid rows from silver, but strictly speaking, the documented quarantine implementation is more explicit than any option shown here. (Databricks Documentation)
═══

161. Frage
A Delta Lake table was created with the below query:
Realizing that the original query had a typographical error, the below code was executed:
ALTER TABLE prod.sales_by_stor RENAME TO prod.sales_by_store
Which result will occur after running the second command?

- A. The table name change is recorded in the Delta transaction log.
- B. The table reference in the metastore is updated and all data files are moved.
- C. A new Delta transaction log Is created for the renamed table.
- D. The table reference in the metastore is updated and no data is changed.
- E. All related files and metadata are dropped and recreated in a single ACID transaction.

Antwort: D

Begründung:

The query uses the CREATE TABLE USING DELTA syntax to create a Delta Lake table from an existing Parquet file stored in DBFS. The query also uses the LOCATION keyword to specify the path to the Parquet file as /mnt/finance_eda_bucket/tx_sales.parquet. By using the LOCATION keyword, the query creates an external table, which is a table that is stored outside of the default warehouse directory and whose metadata is not managed by Databricks. An external table can be created from an existing directory in a cloud storage system, such as DBFS or S3, that contains data files in a supported format, such as Parquet or CSV.
The result that will occur after running the second command is that the table reference in the metastore is updated and no data is changed. The metastore is a service that stores metadata about tables, such as their schema, location, properties, and partitions. The metastore allows users to access tables using SQL commands or Spark APIs without knowing their physical location or format. When renaming an external table using the ALTER TABLE RENAME TO command, only the table reference in the metastore is updated with the new name; no data files or directories are moved or changed in the storage system. The table will still point to the same location and use the same format as before. However, if renaming a managed table, which is a table whose metadata and data are both managed by Databricks, both the table reference in the metastore and the data files in the default warehouse directory are moved and renamed accordingly. Verified References:
[Databricks Certified Data Engineer Professional], under "Delta Lake" section; Databricks Documentation, under "ALTER TABLE RENAME TO" section; Databricks Documentation, under "Metastore" section; Databricks Documentation, under "Managed and external tables" section.

162. Frage
A data engineer is configuring Delta Sharing for a Databricks-to-Databricks scenario to optimize read performance. The recipient needs to perform time travel queries and streaming reads on shared sales data.
Which configuration will provide the optimal performance while enabling these capabilities?

- A. Share tables WITHOUT HISTORY and enable partitioning for better query performance.
- B. Use the open sharing protocol instead of Databricks-to-Databricks sharing for better performance.
- C. Share tables WITH HISTORY , ensure tables don't have partitioning enabled, and enable CDF before sharing.
- D. Share the entire schema WITHOUT HISTORY and rely on recipient-side caching for performance.

Antwort: C

Begründung:
The official Delta Sharing guidance specifies that in order for recipients to use time travel queries and streaming reads , providers must share Delta tables WITH HISTORY . Sharing history ensures the Delta log is included, which enables efficient access to table snapshots and incremental data streams. Additionally, Change Data Feed (CDF) must be enabled prior to sharing if downstream consumers require streaming CDC queries. Without history, recipients cannot perform time travel or streaming queries. Open sharing supports static Delta tables but lacks streaming support. Therefore, sharing tables WITH HISTORY and enabling CDF is the required configuration for both performance and functionality.

163. Frage
Which of the following is true of Delta Lake and the Lakehouse?

- A. Views in the Lakehouse maintain a valid cache of the most recent versions of source tables at all times.
- B. Delta Lake automatically collects statistics on the first 32 columns of each table which are leveraged in data skipping based on query filters.
- C. Z-order can only be applied to numeric values stored in Delta Lake tables
- D. Because Parquet compresses data row by row. strings will only be compressed when a character is repeated multiple times.
- E. Primary and foreign key constraints can be leveraged to ensure duplicate values are never entered into a dimension table.

Antwort: B

Begründung:
https://docs.delta.io/2.0.0/table-properties.html
Delta Lake automatically collects statistics on the first 32 columns of each table, which are leveraged in data skipping based on query filters 1 . Data skipping is a performance optimization technique that aims to avoid reading irrelevant data from the storage layer 1 . By collecting statistics such as min/max values, null counts, and bloom filters, Delta Lake can efficiently prune unnecessary files or partitions from the query plan 1 . This can significantly improve the query performance and reduce the I/O cost.
The other options are false because:
* Parquet compresses data column by column, not row by row 2 . This allows for better compression ratios, especially for repeated

or similar values within a column 2 .
* Views in the Lakehouse do not maintain a valid cache of the most recent versions of source tables at all times 3 . Views are logical constructs that are defined by a SQL query on one or more base tables 3
. Views are not materialized by default, which means they do not store any data, but only the query definition 3 . Therefore, views always reflect the latest state of the source tables when queried 3 .
However, views can be cached manually using the CACHE TABLE or CREATE TABLE AS SELECT commands.
* Primary and foreign key constraints can not be leveraged to ensure duplicate values are never entered into a dimension table. Delta Lake does not support enforcing primary and foreign key constraints on tables. Constraints are logical rules that define the integrity and validity of the data in a table. Delta Lake relies on the application logic or the user to ensure the data quality and consistency.
* Z-order can be applied to any values stored in Delta Lake tables, not only numeric values. Z-order is a technique to optimize the layout of the data files by sorting them on one or more columns. Z-order can improve the query performance by clustering related values together and enabling more efficient data skipping. Z-order can be applied to any column that has a defined ordering, such as numeric, string, date, or boolean values.
References: Data Skipping , Parquet Format , Views , [Caching], [Constraints] , [Z-Ordering]

164. Frage
A Databricks SQL dashboard has been configured to monitor the total number of records present in a collection of Delta Lake tables using the following query pattern:
SELECT COUNT (*) FROM table -
Which of the following describes how results are generated each time the dashboard is updated?

- A. The total count of rows will be returned from cached results unless REFRESH is run
- B. The total count of records is calculated from the Delta transaction logs
- C. The total count of records is calculated from the Hive metastore
- D. The total count of rows is calculated by scanning all data files
- E. The total count of records is calculated from the parquet file metadata

Antwort: B

Begründung:
https://delta.io/blog/2023-04-19-faster-aggregations-metadata/#:~:text=You%20can%20get%20the%20number,a%20given%20Delta%20table%20version.

165. Frage
.....

Unser Pass4Test bietet den Kandidaten nicht nur gute Produktem sondern auch vollständigen Service. Wenn Sie unsere Produkte benutzen, können Sie einen einjährigen kostenlosen Update-Service genießen. Wir benachrichtigen den Kandidaten in erster Zeit die neuen Prüfungsmaterialien zur Databricks Databricks-Certified-Professional-Data-Engineer Zertifizierung mit dem besten Service.

Databricks-Certified-Professional-Data-Engineer Originale Fragen: https://www.pass4test.de/Databricks-Certified-Professional-Data-Engineer.html

- Databricks-Certified-Professional-Data-Engineer Zertifizierungsfragen, Databricks Databricks-Certified-Professional-Data-Engineer PrüfungFragen □ URL kopieren ➥ www.zertpruefung.ch □ Öffnen und suchen Sie 「 Databricks-Certified-Professional-Data-Engineer 」 Kostenloser Download □Databricks-Certified-Professional-Data-Engineer Prüfungs
- Databricks-Certified-Professional-Data-Engineer echter Test - Databricks-Certified-Professional-Data-Engineer sicherlich-zu-bestehen - Databricks-Certified-Professional-Data-Engineer Testguide □ Suchen Sie jetzt auf ➠ www.itzert.com □ nach ☀ Databricks-Certified-Professional-Data-Engineer □☀□ um den kostenlosen Download zu erhalten □Databricks-Certified-Professional-Data-Engineer Prüfungsmaterialien
- Databricks-Certified-Professional-Data-Engineer Fragen - Antworten - Databricks-Certified-Professional-Data-Engineer Studienführer - Databricks-Certified-Professional-Data-Engineer Prüfungsvorbereitung □ Suchen Sie auf { www.zertpruefung.ch } nach kostenlosem Download von ⇛ Databricks-Certified-Professional-Data-Engineer ⇚ □ □Databricks-Certified-Professional-Data-Engineer Testantworten
- Databricks-Certified-Professional-Data-Engineer Testfagen □ Databricks-Certified-Professional-Data-Engineer Prüfungsübungen □ Databricks-Certified-Professional-Data-Engineer Fragenpool □ Sie müssen nur zu ▹ www.itzert.com ◃ gehen um nach kostenloser Download von ✔ Databricks-Certified-Professional-Data-Engineer □✔□ zu suchen □ □Databricks-Certified-Professional-Data-Engineer Fragenpool
- Databricks-Certified-Professional-Data-Engineer Testantworten □ Databricks-Certified-Professional-Data-Engineer

Prüfungsmaterialien □ Databricks-Certified-Professional-Data-Engineer Fragenpool □ □ www.itzert.com □ ist die beste Webseite um den kostenlosen Download von ➥ Databricks-Certified-Professional-Data-Engineer □ zu erhalten □ □Databricks-Certified-Professional-Data-Engineer Testantworten

- Databricks-Certified-Professional-Data-Engineer Deutsche □ Databricks-Certified-Professional-Data-Engineer Prüfung □ □ Databricks-Certified-Professional-Data-Engineer Testking □ Öffnen Sie die Webseite ➥ www.itzert.com □ und suchen Sie nach kostenloser Download von " Databricks-Certified-Professional-Data-Engineer " □Databricks-Certified-Professional-Data-Engineer Prüfungsmaterialien
- Databricks-Certified-Professional-Data-Engineer Dumps und Test Überprüfungen sind die beste Wahl für Ihre Databricks Databricks-Certified-Professional-Data-Engineer Testvorbereitung □ Erhalten Sie den kostenlosen Download von ▷ Databricks-Certified-Professional-Data-Engineer ◁ mühelos über 「 www.zertpruefung.ch 」 □Databricks-Certified-Professional-Data-Engineer Prüfung
- Databricks-Certified-Professional-Data-Engineer Probesfragen □ Databricks-Certified-Professional-Data-Engineer Prüfungsübungen □ Databricks-Certified-Professional-Data-Engineer Prüfungs □ Suchen Sie einfach auf （ www.itzert.com ） nach kostenloser Download von ➠ Databricks-Certified-Professional-Data-Engineer □ □ □Databricks-Certified-Professional-Data-Engineer Prüfungs
- Databricks-Certified-Professional-Data-Engineer Examengine □ Databricks-Certified-Professional-Data-Engineer Übungsmaterialien □ Databricks-Certified-Professional-Data-Engineer Testfagen □ Öffnen Sie ➥ www.itzert.com □ geben Sie ☀ Databricks-Certified-Professional-Data-Engineer □☀□ ein und erhalten Sie den kostenlosen Download □Databricks-Certified-Professional-Data-Engineer Prüfungsmaterialien
- Databricks-Certified-Professional-Data-Engineer Prüfungsübungen □ Databricks-Certified-Professional-Data-Engineer Prüfungsmaterialien □ Databricks-Certified-Professional-Data-Engineer Prüfungsmaterialien □ Öffnen Sie die Website ☀ www.itzert.com □☀□ Suchen Sie ➽ Databricks-Certified-Professional-Data-Engineer □ Kostenloser Download □ □Databricks-Certified-Professional-Data-Engineer Deutsch
- Databricks-Certified-Professional-Data-Engineer Examsfragen □ Databricks-Certified-Professional-Data-Engineer Testking □ Databricks-Certified-Professional-Data-Engineer Deutsche □ Erhalten Sie den kostenlosen Download von □ Databricks-Certified-Professional-Data-Engineer □ mühelos über " www.zertfragen.com " □Databricks-Certified-Professional-Data-Engineer Musterprüfungsfragen
- www.stes.tyc.edu.tw, www.bananabl.net, 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, knowyourmeme.com, myportal.utt.edu.tt, www.pshunv.com, 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, 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, www.stes.tyc.edu.tw, Disposable vapes

P.S. Kostenlose und neue Databricks-Certified-Professional-Data-Engineer Prüfungsfragen sind auf Google Drive freigegeben von Pass4Test verfügbar: https://drive.google.com/open?id=1RnCBv6bs9c5-wgfiM8njGfOlEM7ZY0OU