

Well-Prepared Valid Databricks-Certified-Data-Engineer-Professional Test Cost - Effective Databricks-Certified-Data-Engineer-Professional Exam Tool Guarantee Purchasing Safety



The passing rate of our Databricks-Certified-Data-Engineer-Professional exam torrent is up to 98 to 100 percent, and this is a striking outcome staged anywhere in the world. They are appreciated with passing rate up to 98 percent among the former customers. So they are in ascendant position in the market. If you choose our Databricks-Certified-Data-Engineer-Professional question materials, you can get success smoothly. Besides, they are effective Databricks-Certified-Data-Engineer-Professional guide tests to fight against difficulties emerged on your way to success.

According to the needs of all people, the experts and professors in our company designed three different versions of the Databricks-Certified-Data-Engineer-Professional study materials for all customers. The three versions are very flexible for all customers to operate. According to your actual need, you can choose the version for yourself which is most suitable for you to preparing for the coming exam. All the Databricks-Certified-Data-Engineer-Professional Study Materials of our company can be found in the three versions. It is very flexible for you to use the three versions of the Databricks-Certified-Data-Engineer-Professional study materials to preparing for your coming exam.

>> Valid Databricks-Certified-Data-Engineer-Professional Test Cost <<

Latest updated Valid Databricks-Certified-Data-Engineer-Professional Test Cost & Leader in Qualification Exams & Professional Databricks-Certified-Data-Engineer-Professional: Databricks Certified Data Engineer Professional Exam

Prep4pass Databricks Databricks-Certified-Data-Engineer-Professional Exam Training materials can help you to come true your dreams. Because it contains all the questions of Databricks Databricks-Certified-Data-Engineer-Professional examination. With Prep4pass, you could throw yourself into the exam preparation completely. With high quality training materials by Prep4pass provided, you will certainly pass the exam. Prep4pass can give you a brighter future.

Databricks Certified Data Engineer Professional Exam Sample Questions (Q21-Q26):

NEW QUESTION # 21

The view updates represents an incremental batch of all newly ingested data to be inserted or Get Latest & Actual Certified-Data-Engineer-Professional Exam's Question and Answers from updated in the customers table.

The following logic is used to process these records.

```
MERGE INTO customers
```

```
USING (
```

```
SELECT updates.customer_id as merge_ey, updates.*
```

```

FROM updates
UNION ALL
SELECT NULL as merge_key, updates.*
FROM updates JOIN customers
ON updates.customer_id = customers.customer_id
WHERE customers.current = true AND updates.address <> customers.address ) staged_updates ON customers.customer_id =
mergekey WHEN MATCHED AND customers.current = true AND customers.address <> staged_updates.address THEN
UPDATE SET current = false, end_date = staged_updates.effective_date WHEN NOT MATCHED THEN INSERT
(customer_id, address, current, effective_date, end_date) VALUES (staged_updates.customer_id, staged_updates.address, true,
staged_updates.effective_date, null) Which statement describes this implementation?

```

- A. The customers table is implemented as a Type 2 table; old values are overwritten and new customers are appended.
- **B. The customers table is implemented as a Type 2 table; old values are maintained but marked as no longer current and new values are inserted.**
- C. The customers table is implemented as a Type 1 table; old values are overwritten by new values and no history is maintained.
- D. The customers table is implemented as a Type 0 table; all writes are append only with no changes to existing values.

Answer: B

Explanation:

The provided MERGE statement is a classic implementation of a Type 2 SCD in a data warehousing context. In this approach, historical data is preserved by keeping old records (marking them as not current) and adding new records for changes. Specifically, when a match is found and there's a change in the address, the existing record in the customers table is updated to mark it as no longer current (`current = false`), and an end date is assigned (`end_date = staged_updates.effective_date`). A new record for the customer is then inserted with the updated information, marked as current. This method ensures that the full history of changes to customer information is maintained in the table, allowing for time-based analysis of customer data.

NEW QUESTION # 22

Which approach demonstrates a modular and testable way to use DataFrame transform for ETL code in PySpark?

- A.

```

def upper_value (df):
    return df.withColumn("value_upper", upper(col("value")))

def filter_positive(df):
    return df.filter (df["id"] > 0)

pipeline_df = df.transform(upper_value).transform
(filter_positive)

```

```

def upper_transform(df):
    return df.withColumn("value_upper", upper(col("value")))

actual = test_input.transform(upper_transform)
assertDataFrameEqual(actual, expected)

```

- B.
- C.

```

def transform_data(input_df):
    # transformation logic here
    return output_df

test_input = spark.createDataFrame([(1, "a")], ["id", "value"])
assertDataFrameEqual(transform_data(test_input), expected)

```

- D.

```

class Pipeline:
    def transform(self, df):
        return df.withColumn("value_upper", upper(col("value")))

pipeline = Pipeline()
assertDataFrameEqual (pipeline.transform(test_input), expected)

```

Answer: B

Explanation:

Using DataFrame.transform with a pure transformation function promotes modular, reusable, and easily testable ETL logic. Each transformation is encapsulated as a standalone function, can be independently unit tested, and composed cleanly in a pipeline without coupling to orchestration or class state.

NEW QUESTION # 23

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.

Original query:

Get Latest & Actual Certified-Data-Engineer-Professional Exam's Question and Answers from

```

df.groupBy("item")
  .agg(count("item").alias("total_count"),
        mean("sale_price").alias("avg_price"))
  .writeStream
  .outputMode("complete")
  .option("checkpointLocation", "/item_agg/__checkpoint")
  .start("/item_agg")

```

Proposed query:

```

df.groupBy("item")
  .agg(count("item").alias("total_count"),
        mean("sale_price").alias("avg_price"))
  .writeStream
  .outputMode("complete")
  .option("checkpointLocation", "/item_agg/__checkpoint")
  .start("/item_agg")

```

Proposed query:

```
.start("/item_agg")
```

Which step must also be completed to put the proposed query into production?

- A. Increase the shuffle partitions to account for additional aggregates
- B. Register the data in the '/item_agg' directory to the Hive metastore
- C. Remove .option('mergeSchema', true) from the streaming write
- **D. Specify a new checkpointLocation**
- E. Run REFRESH TABLE delta, '/item_agg'

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.

NEW QUESTION # 24

A data company uses Databricks Unity Catalog and has multiple enterprise data sources, including PostgreSQL, Snowflake, and SQL Server. The central data platform team wants to configure Lakehouse Federation so analysts can query external tables directly in Databricks using Databricks SQL, without duplicating data. Which steps are necessary to configure Lakehouse Federation in a secure and governed manner?

- A. Create external locations and storage credentials to connect to each database, then register foreign tables in Unity Catalog.
- B. Mirror the external datasets into Delta Lake using Auto Loader, and govern them using Data Lineage and System Tables.
- C. Configure connections and foreign catalog in Unity Catalog, then grant access to foreign catalogs, schemas, and tables using Unity Catalog permissions.
- D. Use Partner Connect to create linked datasets, and apply table ACLs at the source system to govern access through Databricks.

Answer: C

Explanation:

Lakehouse Federation is configured by defining secure connections to external data sources and registering them as foreign catalogs in Unity Catalog. Access is then governed using Unity Catalog permissions at the catalog, schema, and table levels, enabling analysts to query external tables securely without data duplication.

NEW QUESTION # 25

A member of the data engineering team has submitted a short notebook that they wish to schedule as part of a larger data pipeline. Assume that the commands provided below produce the logically correct results when run as presented.

Get Latest & Actual Certified-Data-Engineer-Professional Exam's Question and Answers from

Cmd 1

```
rawDF = spark.table("raw_data")
```

Cmd 2

```
rawDF.printSchema()
```

Cmd 3

```
flattenedDF = rawDF.select("*", "values.*")
```

Cmd 4

```
finalDF = flattenedDF.drop("values")
```

Cmd 5

```
finalDF.explain()
```

Cmd 6

```
display(finalDF)
```

Cmd 7

```
finalDF.write.mode("append").saveAsTable("flat_data")
```

Which command should be removed from the notebook before scheduling it as a job?

- A. Cmd 6
- B. Cmd 2
- C. Cmd 5
- D. Cmd 3
- E. Cmd 4

Answer: A

Explanation:

When scheduling a Databricks notebook as a job, it's generally recommended to remove or modify commands that involve displaying output, such as using the `display()` function. Displaying data using `display()` is an interactive feature designed for exploration and visualization within the notebook interface and may not work well in a production job context.

The `finalDF.explain()` command, which provides the execution plan of the DataFrame transformations and actions, is often useful for debugging and optimizing queries. While it doesn't Get Latest & Actual Certified-Data-Engineer-Professional Exam's Question and Answers from display interactive visualizations like `display()`, it can still be informative for understanding how Spark is executing the operations on your DataFrame.

NEW QUESTION # 26

.....

The Prep4pass Databricks Databricks-Certified-Data-Engineer-Professional exam questions is 100% verified and tested. Prep4pass Databricks Databricks-Certified-Data-Engineer-Professional exam practice questions and answers is the practice test software. In Prep4pass, you will find the best exam preparation material. The material including practice questions and answers. The information we have could give you the opportunity to practice issues, and ultimately achieve your goal that through Databricks Databricks-Certified-Data-Engineer-Professional Exam Certification.

Databricks-Certified-Data-Engineer-Professional Test Registration: https://www.prep4pass.com/Databricks-Certified-Data-Engineer-Professional_exam-braindumps.html

The Learning Path includes a mix of courses and hands-on labs that offer practical experience on the Databricks-Certified-Data-Engineer-Professional Test Registration platform, So you don't need to worry about the quality of our Databricks-Certified-Data-Engineer-Professional training torrent, You can prepare well with the help of our Databricks-Certified-Data-Engineer-Professional training material, But passing Databricks certification Databricks-Certified-Data-Engineer-Professional exam is not so simple, For there are three versions of Databricks-Certified-Data-Engineer-Professional learning materials and are not limited by the device.

André LaMothe has been involved in the computing industry and technology for more Databricks-Certified-Data-Engineer-Professional than a quarter century, Policies for a Web application are enforced no matter what is configured on individual sites that live within the web application.

100% Pass Quiz Databricks - Databricks-Certified-Data-Engineer-Professional The Best Valid Test Cost

The Learning Path includes a mix of courses and hands-on labs that offer practical experience on the Databricks Certification platform, So you don't need to worry about the quality of our Databricks-Certified-Data-Engineer-Professional training torrent.

You can prepare well with the help of our Databricks-Certified-Data-Engineer-Professional training material, But passing Databricks certification Databricks-Certified-Data-Engineer-Professional exam is not so simple, For there are three versions of Databricks-Certified-Data-Engineer-Professional learning materials and are not limited by the device.

- Valid Databricks-Certified-Data-Engineer-Professional Test Cost | Pass-Sure Databricks-Certified-Data-Engineer-Professional: Databricks Certified Data Engineer Professional Exam 100% Pass Download Databricks-Certified-Data-Engineer-Professional for free by simply searching on (www.pdf.dumps.com) Databricks-Certified-Data-Engineer-Professional Reliable Test Labs
- Databricks-Certified-Data-Engineer-Professional Related Certifications Detail Databricks-Certified-Data-Engineer-Professional Explanation New Databricks-Certified-Data-Engineer-Professional Test Discount Download { Databricks-Certified-Data-Engineer-Professional } for free by simply searching on \Rightarrow www.pdfvce.com \Leftarrow Hot Databricks-Certified-Data-Engineer-Professional Questions
- Valid Databricks-Certified-Data-Engineer-Professional Test Cost | Ready to Pass The Databricks Certified Data Engineer Professional Exam The page for free download of Databricks-Certified-Data-Engineer-Professional on www.troytecdumps.com will open immediately Online Databricks-Certified-Data-Engineer-Professional Test

