Vendor: Databricks

Exam Code: Certified-Data-Analyst-Associate

Exam Name: Databricks Certified Data Analyst Associate

Version: DEMO

QUESTION 1
A data analyst created and is the owner of the managed table my_ table. They now want to change ownership of the table to a single other user using Data Explorer.

Which of the following approaches can the analyst use to complete the task?

A. Edit the Owner field in the table page by removing their own account
B. Edit the Owner field in the table page by selecting All Users
C. Edit the Owner field in the table page by selecting the new owner's account
D. Edit the Owner field in the table page by selecting the Admins group
E. Edit the Owner field in the table page by removing all access

Answer: C
Explanation:
The Owner field in the table page shows the current owner of the table and allows the owner to change it to another user or group. To change the ownership of the table, the owner can click on the Owner field and select the new owner from the drop-down list. This will transfer the ownership of the table to the selected user or group and remove the previous owner from the list of table access control entries.

QUESTION 2
A data analyst has a managed table table_name in database database_name. They would now like to remove the table from the database and all of the data files associated with the table. The rest of the tables in the database must continue to exist.

Which of the following commands can the analyst use to complete the task without producing an error?

A. DROP DATABASE database_name;
B. DROP TABLE database_name.table_name;
C. DELETE TABLE database_name.table_name;
D. DELETE TABLE table_name FROM database_name;
E. DROP TABLE table_name FROM database_name;

Answer: B
Explanation:
The DROP TABLE command removes a table from the metastore and deletes the associated data files. The syntax for this command is DROP TABLE [IF EXISTS] [database_name.]table_name;. The optional IF EXISTS clause prevents an error if the table does not exist. The optional database_name. prefix specifies the database where the table resides. If not specified, the current database is used. Therefore, the correct command to remove the table table_name from the database database_name and all of the data files associated with it is DROP TABLE database_name.table_name;.

QUESTION 3
A data analyst runs the following command:

```
SELECT age, country
FROM my_table
WHERE age >= 75 AND country = 'canada';
```

Which of the following tables represents the output of the above command?

A.

age	country
80	canada
NULL	canada
90	*NULL*

B.

age	country
80	*NULL*
75	*NULL*
90	*NULL*

C.

id	age	country
900	80	canada
901	75	canada
902	90	canada

D.

age	country
80	canada
14	canada
90	canada

E.

age	country
80	canada
75	canada
90	canada

Answer: E
Explanation:
The SQL query provided is designed to filter out records from "my_table" where the age is 75 or above and the country is Canada. Since I can't view the content of the links provided directly, I need to rely on the image attached to this question for context. Based on that, Option E (the image attached) represents a table with columns "age" and "country", showing records where age is 75 or above and country is Canada.

QUESTION 4
A data analyst runs the following command:

```
INSERT INTO stakeholders.suppliers TABLE stakeholders.new_suppliers;
```

What is the result of running this command?

A. The suppliers table now contains both the data it had before the command was run and the data from the new suppliers table, and any duplicate data is deleted.
B. The command fails because it is written incorrectly.
C. The suppliers table now contains both the data it had before the command was run and the data from the new suppliers table, including any duplicate data.
D. The suppliers table now contains the data from the new suppliers table, and the new suppliers table now contains the data from the suppliers table.
E. The suppliers table now contains only the data from the new suppliers table.

Answer: B
Explanation:
The command INSERT INTO stakeholders.suppliers TABLE stakeholders.new_suppliers is not a valid syntax for inserting data into a table in Databricks SQL. According to the documentation, the correct syntax for inserting data into a table is either:

```
INSERT { OVERWRITE | INTO } [ TABLE ] table_name [ PARTITION clause ]
[ ( column_name [, ...] ) | BY NAME ] query
INSERT INTO [ TABLE ] table_name REPLACE WHERE predicate query
```

The command in the question is missing the OVERWRITE or INTO keyword, and the query part that specifies the source of the data to be inserted. The TABLE keyword is optional and can be omitted. The PARTITION clause and the column list are also optional and depend on the table schema and the data source. Therefore, the command in the question will fail with a syntax error.

QUESTION 5
A data engineer is working with a nested array column products in table transactions. They want to expand the table so each unique item in products for each row has its own row where the transaction_id column is duplicated as necessary.

They are using the following incomplete command:

```
SELECT
    transaction_id,
    _______ AS product
FROM transactions;
```

Which of the following lines of code can they use to fill in the blank in the above code block so that it successfully completes the task?

A. array distinct(produces)
B. explode(produces)
C. reduce(produces)
D. array(produces)
E. flatten(produces)

Answer: B
Explanation:

The explode function is used to transform a DataFrame column of arrays or maps into multiple rows, duplicating the other column's values. In this context, it will be used to expand the nested array column products in the transactions table so that each unique item in products for each row has its own row and the transaction_id column is duplicated as necessary.

QUESTION 6
A data analysis team is working with the table_bronze SQL table as a source for one of its most complex projects. A stakeholder of the project notices that some of the downstream data is duplicative. The analysis team identifies table_bronze as the source of the duplication.

Which of the following queries can be used to deduplicate the data from table_bronze and write it to a new table table_silver?

A.
```
CREATE TABLE table_silver AS
SELECT DISTINCT *
FROM table_bronze;
```
B.
```
CREATE TABLE table_silver AS
INSERT *
FROM table_bronze;
```
C.
```
CREATE TABLE table_silver AS
MERGE DEDUPLICATE *
FROM table_bronze;
```
D.
```
INSERT INTO TABLE table_silver
SELECT * FROM table_bronze;
```
E.
```
INSERT OVERWRITE TABLE table_silver
SELECT * FROM table_bronze;
```

Answer: A
Explanation:
Option A uses the SELECT DISTINCT statement to remove duplicate rows from the table_bronze and create a new table table_silver with the deduplicated data. This is the correct way to deduplicate data using Spark SQL.

Thank You for Trying Our Product

Passleader Certification Exam Features:

★ More than **99,900** Satisfied Customers Worldwide.

★ Average **99.9%** Success Rate.

★ **Free Update** to match latest and real exam scenarios.

★ **Instant Download** Access! No Setup required.

★ Questions & Answers are downloadable in **PDF** format and **VCE** test engine format.

★ Multi-Platform capabilities - **Windows, Laptop, Mac, Android, iPhone, iPod, iPad**.

★ **100%** Guaranteed Success or **100%** Money Back Guarantee.

★ **Fast**, helpful support **24x7**.

View list of all certification exams: http://www.passleader.com/all-products.html

10% Discount Coupon Code: **ASTR14**