DP-600Q&As

Implementing Analytics Solutions Using Microsoft Fabric

Pass Microsoft DP-600 Exam with 100% Guarantee

Free Download Real Questions & Answers **PDF** and **VCE** file from:

https://www.geekcert.com/dp-600.html

100% Passing Guarantee
100% Money Back Assurance

Following Questions and Answers are all new published by Microsoft Official Exam Center

- **Instant Download** After Purchase
- **100% Money Back** Guarantee
- **365 Days** Free Update
- **800,000+** Satisfied Customers

QUESTION 1

What should you recommend using to ingest the customer data into the data store in the AnatyticsPOC workspace?

A. a stored procedure

B. a pipeline that contains a KQL activity

C. a Spark notebook

D. a dataflow

Correct Answer: D

Explanation: For ingesting customer data into the data store in the AnalyticsPOC workspace, a dataflow (D) should be recommended. Dataflows are designed within the Power BI service to ingest, cleanse, transform, and load data into the Power BI environment. They allow for the low-code ingestion and transformation of data as needed by Litware\\'s technical requirements. References = You can learn more about dataflows and their use in Power BI environments in Microsoft\\'s Power BI documentation.

QUESTION 2

You have a Fabric tenant that contains a warehouse.

You use a dataflow to load a new dataset from OneLake to the warehouse.

You need to add a Power Query step to identify the maximum values for the numeric columns.

Which function should you include in the step?

A. Table. MaxN

B. Table.Max

C. Table.Range

D. Table.Profile

Correct Answer: B

Explanation: The Table.Max function should be used in a Power Query step to identify the maximum values for the numeric columns. This function is designed to calculate the maximum value across each column in a table, which suits the requirement of finding maximum values for numeric columns. References = For detailed information on Power Query functions, including Table.Max, please refer to Power Query M function reference.

QUESTION 3

You have a data warehouse that contains a table named Stage. Customers. Stage-Customers contains all the customer record updates from a customer relationship management (CRM) system. There can be multiple updates per customer

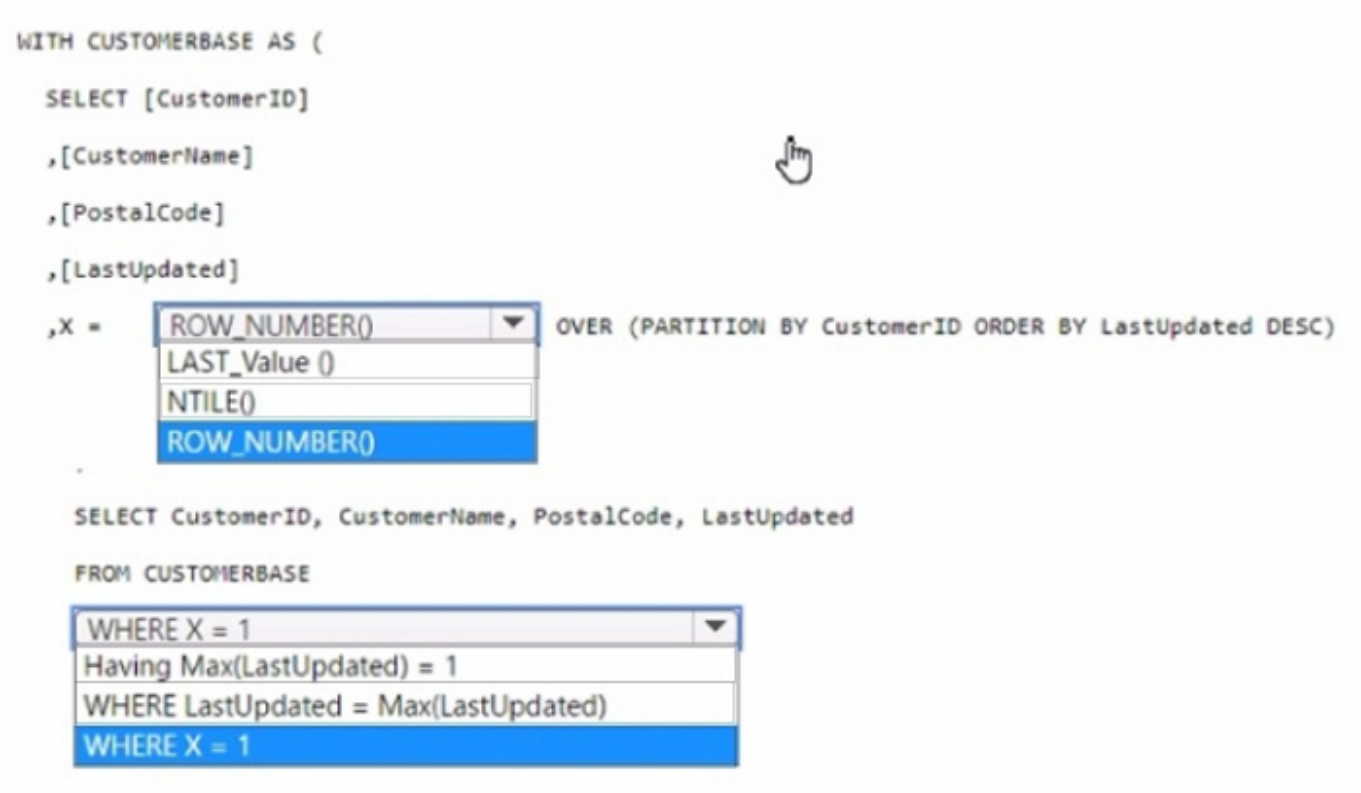
You need to write a T-SQL query that will return the customer ID, name, postal code, and the last updated time of the

most recent row for each customer ID.

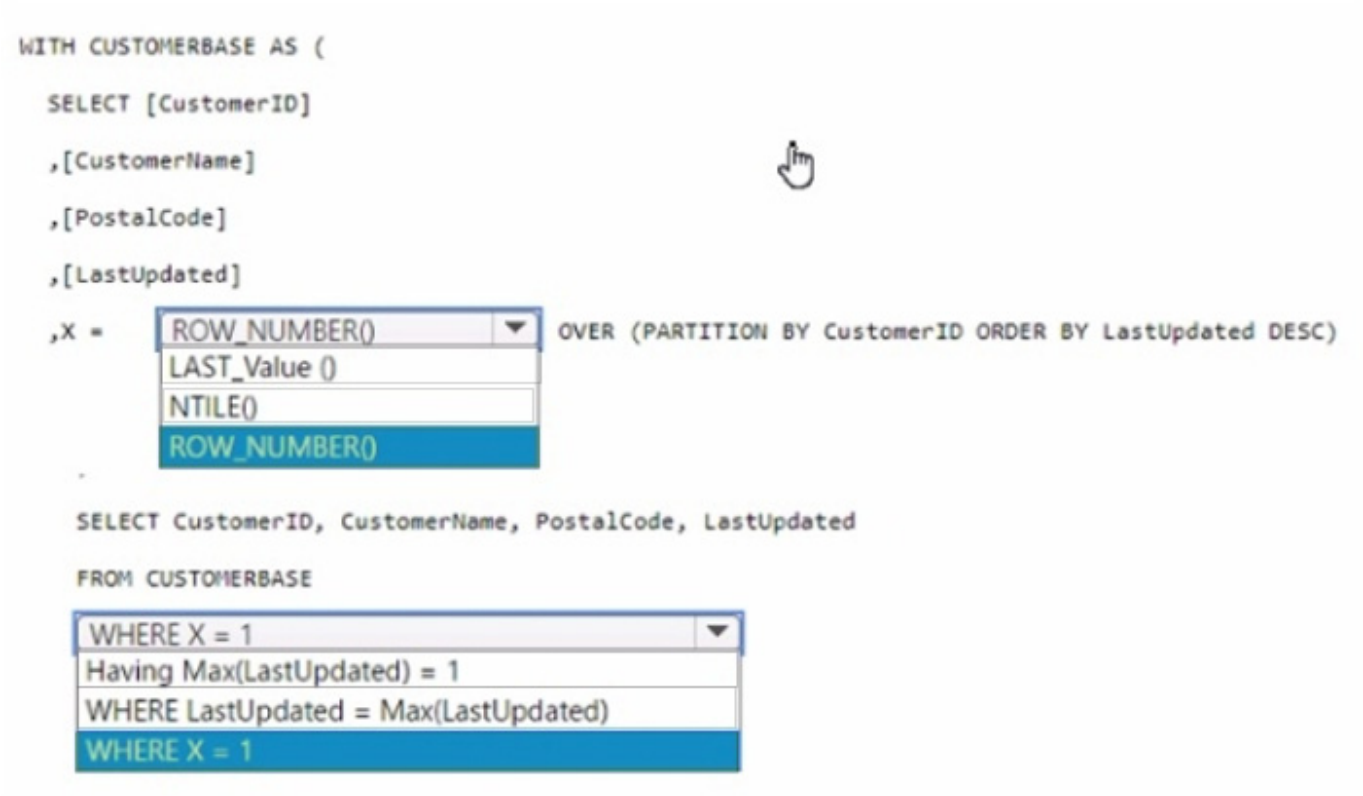
How should you complete the code? To answer, select the appropriate options in the answer area,

NOTE Each correct selection is worth one point.

Hot Area:



Correct Answer:



In the ROW_NUMBER() function, choose OVER (PARTITION BY CustomerID

ORDER BY LastUpdated DESC).

In the WHERE clause, choose WHERE X = 1.

To select the most recent row for each customer ID, you use the ROW_NUMBER() window function partitioned by CustomerID and ordered by LastUpdated in descending order.

This will assign a row number of 1 to the most recent update for each customer. By selecting rows where the row number (X) is 1, you get the latest update per customer.

References =

Use the OVER clause to aggregate data per partition

Use window functions

QUESTION 4

You have a Fabric tenant that contains a lakehouse named Lakehouse1. Lakehouse1 contains a subfolder named Subfolder1 that contains CSV files. You need to convert the CSV files into the delta format that has V-Order optimization enabled. What should you do from Lakehouse explorer?

A. Use the Load to Tables feature.

B. Create a new shortcut in the Files section.

C. Create a new shortcut in the Tables section.

D. Use the Optimize feature.

Correct Answer: D

Explanation: To convert CSV files into the delta format with Z-Order optimization enabled, you should use the Optimize feature (D) from Lakehouse Explorer. This will allow you to optimize the file organization for the most efficient querying. References = The process for converting and optimizing file formats within a lakehouse is discussed in the lakehouse management documentation.

QUESTION 5

You have a Fabric tenant that contains a lakehouse.

You are using a Fabric notebook to save a large DataFrame by using the following code.

```
df.write.partitionBy("year","month","day").mode("overwrite").parquet("Files/SalesOrder")
```

For each of the following statements, select Yes if the statement is true. Otherwise, select No. NOTE: Each correct selection is worth one point.

Hot Area:

Statements	Yes	No
The results will form a hierarchy of folders for each partition key.	○	○
The resulting file partitions can be read in parallel across multiple nodes.	○	○
The resulting file partitions will use file compression.	○	○

Correct Answer:

Statements	Yes	No
The results will form a hierarchy of folders for each partition key.	○	○
The resulting file partitions can be read in parallel across multiple nodes.	○	○
The resulting file partitions will use file compression.	○	○

The results will form a hierarchy of folders for each partition key. - Yes The resulting file partitions can be read in parallel across multiple nodes. - Yes The resulting file partitions will use file compression. - No

Partitioning data by columns such as year, month, and day, as shown in the DataFrame write operation, organizes the

output into a directory hierarchy that reflects the partitioning structure. This organization can improve the performance of read operations, as queries that filter by the partitioned columns can scan only the relevant directories. Moreover, partitioning facilitates parallelism because each partition can be processed independently across different nodes in a distributed system like Spark. However, the code snippet provided does not explicitly specify that file compression should be used, so we cannot assume that the output will be compressed without additional context. References = DataFrame write partitionBy Apache Spark optimization with partitioning