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.pass4itsure.com/dp-600.html

100% Passing Guarantee
100% Money Back Assurance

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

QUESTION 1

Which type of data store should you recommend in the AnalyticsPOC workspace?

A. a data lake

B. a warehouse

C. a lakehouse

D. an external Hive metaStore

Correct Answer: C

Explanation: A lakehouse (C) should be recommended for the AnalyticsPOC workspace. It combines the capabilities of a data warehouse with the flexibility of a data lake. A lakehouse supports semi-structured and unstructured data and allows for T-SQL and Python read access, fulfilling the technical requirements outlined for Litware. References = For further understanding, Microsoft\\'s documentation on the lakehouse architecture provides insights into how it supports various data types and analytical operations.

QUESTION 2

You have a Fabric tenant that contains a lakehouse named Lakehouse1

Readings from 100 IoT devices are appended to a Delta table in Lakehouse1. Each set of readings is approximately 25 KB. Approximately 10 GB of data is received daily.

All the table and SparkSession settings are set to the default.

You discover that queries are slow to execute. In addition, the lakehouse storage contains data and log files that are no longer used.

You need to remove the files that are no longer used and combine small files into larger files with a target size of 1 GB per file.

What should you do? To answer, drag the appropriate actions to the correct requirements.

Each action may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.

NOTE: Each correct selection is worth one point.

Select and Place:

Correct Answer:

Remove the files: Run the VACUUM command on a schedule.

Combine the files: Set the optimizeWrite table setting. or Run the OPTIMIZE command on a schedule.

To remove files that are no longer used, the VACUUM command is used in Delta Lake to clean up invalid files from a table. To combine smaller files into larger ones, you can either set the optimizeWrite setting to combine files during write

operations or use the OPTIMIZE command, which is a Delta Lake operation used to compact small files into larger ones.

QUESTION 3

You have a Fabric tenant that contains a warehouse.

Several times a day. the performance of all warehouse queries degrades. You suspect that Fabric is throttling the compute used by the warehouse.

What should you use to identify whether throttling is occurring?

A. the Capacity settings

B. the Monitoring hub

C. dynamic management views (DMVs)

D. the Microsoft Fabric Capacity Metrics app

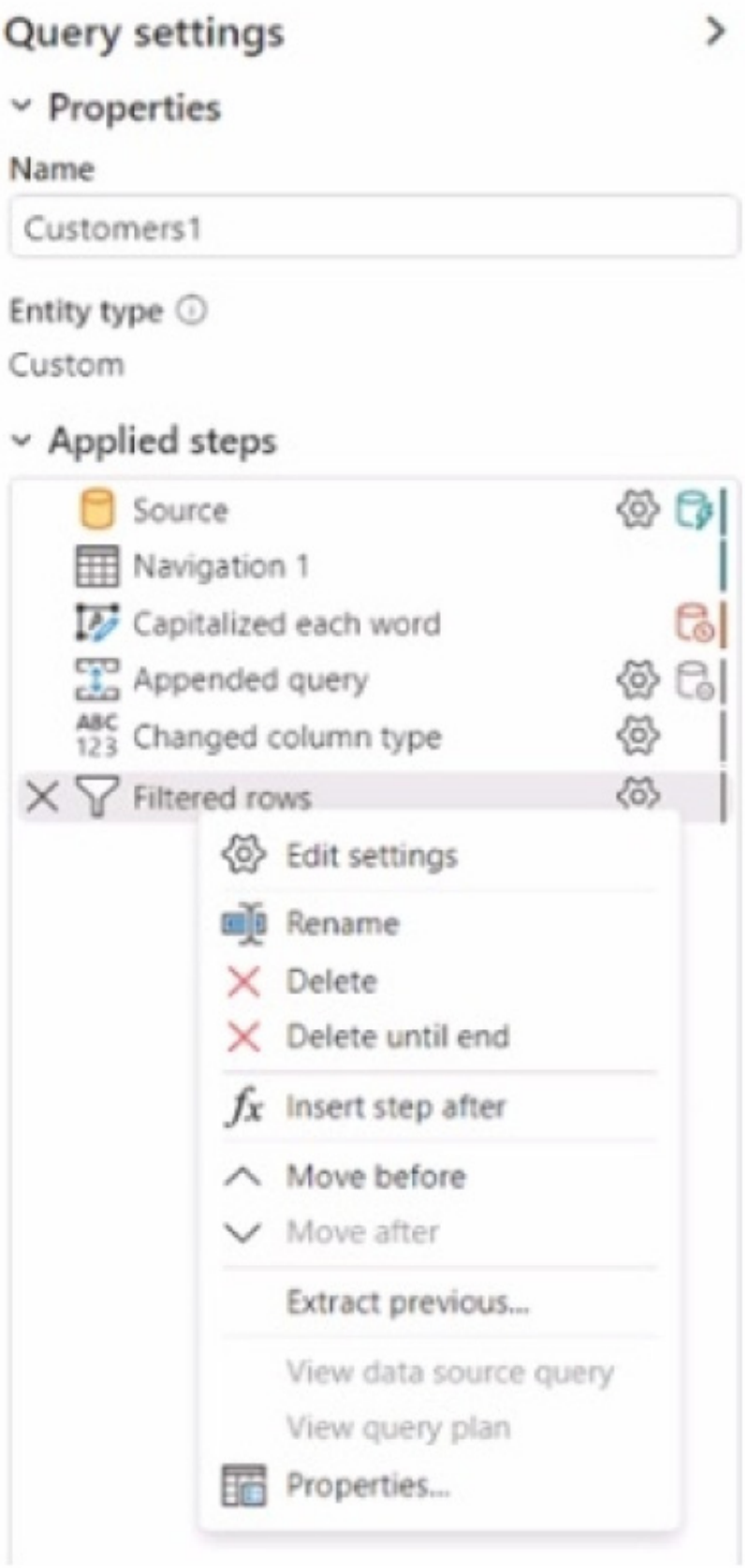
Correct Answer: B

Explanation: To identify whether throttling is occurring, you should use the Monitoring hub (B). This provides a centralized place where you can monitor and manage the health, performance, and reliability of your data estate, and see if the compute resources are being throttled. References = The use of the Monitoring hub for performance management and troubleshooting is detailed in the Azure Synapse Analytics documentation.

QUESTION 4

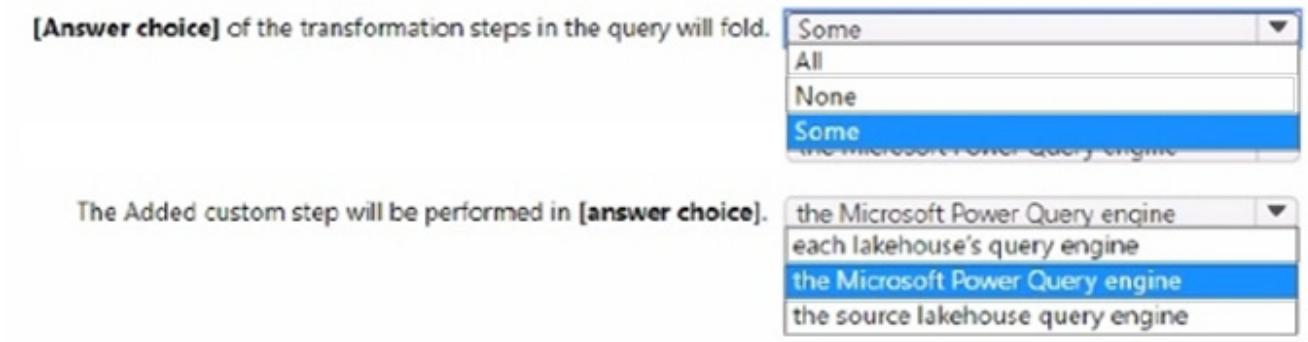
You have a Fabric tenant that contains two lakehouses.

You are building a dataflow that will combine data from the lakehouses. The applied steps from one of the queries in the dataflow is shown in the following exhibit.

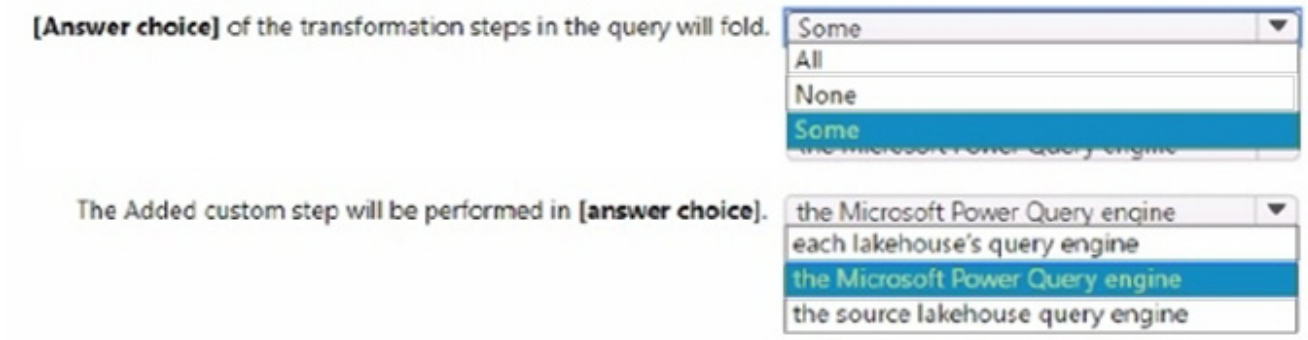


Use the drop-down menus to select the answer choice that completes each statement based on the information presented in the graphic. NOTE: Each correct selection is worth one point.

Hot Area:



Correct Answer:



Folding in Power Query refers to operations that can be translated into source queries. In this case, "some" of the steps can be folded, which means that some transformations will be executed at the data source level. The steps that cannot be folded will be executed within the Power Query engine. Custom steps, especially those that are not standard query operations, are usually executed within Power Query engine rather than being pushed down to the source system. References = Query folding in Power Query Power Query M formula language

QUESTION 5

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 Spark engine will read only the 'SalesOrderNumber', 'OrderDate','CustomerName', 'UnitPrice' columns from Sales_raw.csv.	○	○
Removing the partition will reduce the execution time of the query.	○	○
Adding `inferSchema='true'` to the options will increase the execution time of the query.	○	○

Correct Answer:

Statements	Yes	No
The Spark engine will read only the 'SalesOrderNumber', 'OrderDate','CustomerName', 'UnitPrice' columns from Sales_raw.csv.	○	○
Removing the partition will reduce the execution time of the query.	○	○
Adding `inferSchema='true'` to the options will increase the execution time of the query.	○	○

The Spark engine will read only the \\'SalesOrderNumber\\', \\'OrderDate\\', \\'CustomerName\\', \\'UnitPrice\\' columns from Sales_raw.csv. - Yes Removing the partition will reduce the execution time of the query. - No Adding inferSchema=\\'true\\' to the options will increase the execution time of the query. - Yes

The code specifies the selection of certain columns, which means only those columns will be read into the DataFrame. Partitions in Spark are a way to optimize the execution of queries by organizing the data into parts that can be processed in parallel. Removing the partition could potentially increase the execution time because Spark would no longer be able to process the data in parallel efficiently. The inferSchema option allows Spark to automatically detect the column data types, which can increase the execution time of the initial read operation because it requires Spark to read through the data to infer the schema.