DP-600 Q&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.leads4pass.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

You have a Fabric tenant that contains a new semantic model in OneLake.

You use a Fabric notebook to read the data into a Spark DataFrame.

You need to evaluate the data to calculate the min, max, mean, and standard deviation values for all the string and numeric columns.

Solution: You use the following PySpark expression:

df.explain()

Does this meet the goal?

A. Yes

B. No

Correct Answer: B

Explanation: The df.explain() method does not meet the goal of evaluating data to calculate statistical functions. It is used to display the physical plan that Spark will execute. References = The correct usage of the explain() function can be found in the PySpark documentation.

QUESTION 2

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 3

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 4

You have a Fabric tenant that contains a data pipeline.

You need to ensure that the pipeline runs every four hours on Mondays and Fridays.

To what should you set Repeat for the schedule?

A. Daily

B. By the minute

C. Weekly

D. Hourly

Correct Answer: C

Explanation: You should set Repeat for the schedule to Weekly (C). This allows you to specify the pipeline to run on specific days of the week, in this case, every four hours on Mondays and Fridays. References = Scheduling options for data pipelines are available in the Azure Data Factory documentation, which includes details on configuring recurring triggers.

QUESTION 5

You are analyzing the data in a Fabric notebook.

You have a Spark DataFrame assigned to a variable named df.

You need to use the Chart view in the notebook to explore the data manually.

Which function should you run to make the data available in the Chart view?

A. displayMTML

B. show

C. write

D. display

Correct Answer: D

Explanation: The display function is the correct choice to make the data available in the Chart view within a Fabric notebook. This function is used to visualize Spark DataFrames in various formats including charts and graphs directly within the notebook environment. References = Further explanation of the display function can be found in the official documentation on Azure Synapse Analytics notebooks.

QUESTION 6

You have a Fabric tenant.

You plan to create a Fabric notebook that will use Spark DataFrames to generate Microsoft Power BI visuals.

You run the following code.

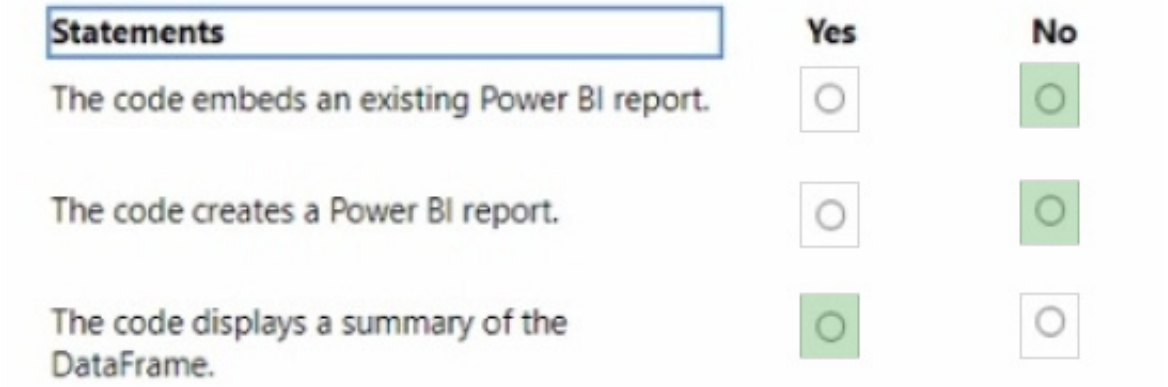
```
from powerbiclient import QuickVisualize, get_dataset_config, Report

PBI_visualize = QuickVisualize(get_dataset_config(df))
PBI_visualize
```

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 code embeds an existing Power BI report.	○	○
The code creates a Power BI report.	○	○
The code displays a summary of the DataFrame.	○	○

Correct Answer:



Statements	Yes	No
The code embeds an existing Power BI report.	○	○
The code creates a Power BI report.	○	○
The code displays a summary of the DataFrame.	○	○

The code embeds an existing Power BI report. - No The code creates a Power BI report. - No The code displays a summary of the DataFrame. - Yes

The code provided seems to be a snippet from a SQL query or script which is neither creating nor embedding a Power BI report directly. It appears to be setting up a DataFrame for use within a larger context, potentially for visualization in Power BI, but the code itself does not perform the creation or embedding of a report. Instead, it\\'s likely part of a data processing step that summarizes data.

References = Introduction to DataFrames - Spark SQL Power BI and Azure Databricks

You have a Fabric workspace that uses the default Spark starter pool and runtime version 1,2.

You plan to read a CSV file named Sales.raw.csv in a lakehouse, select columns, and save the data as a Delta table to the managed area of the lakehouse. Sales_raw.csv contains 12 columns.

You have the following code.

```
from pyspark.sql.functions import year

(spark
  .read
  .format("csv")
  .option("heade" ,'true')
  .load("Files/sales_raw.csv")
  .select('SalesOrderNumber', 'OrderDate','CustomerName', 'UnitPrice')
  .withColumn("Year",year("OrderDate"))
  .write
  .partitionBy('Year')
  .saveAsTable("sales")
)
```

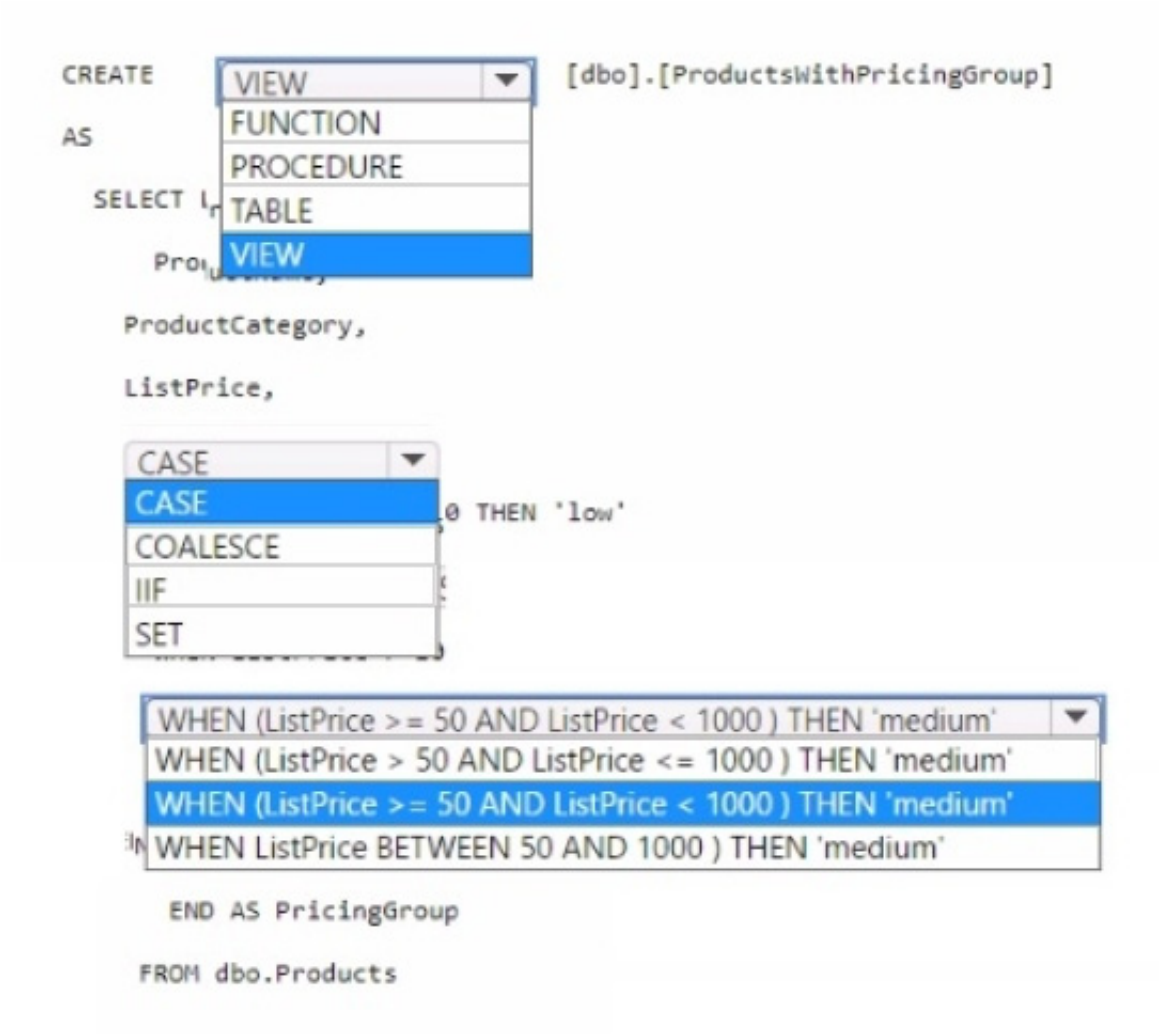
QUESTION 7

You need to resolve the issue with the pricing group classification.

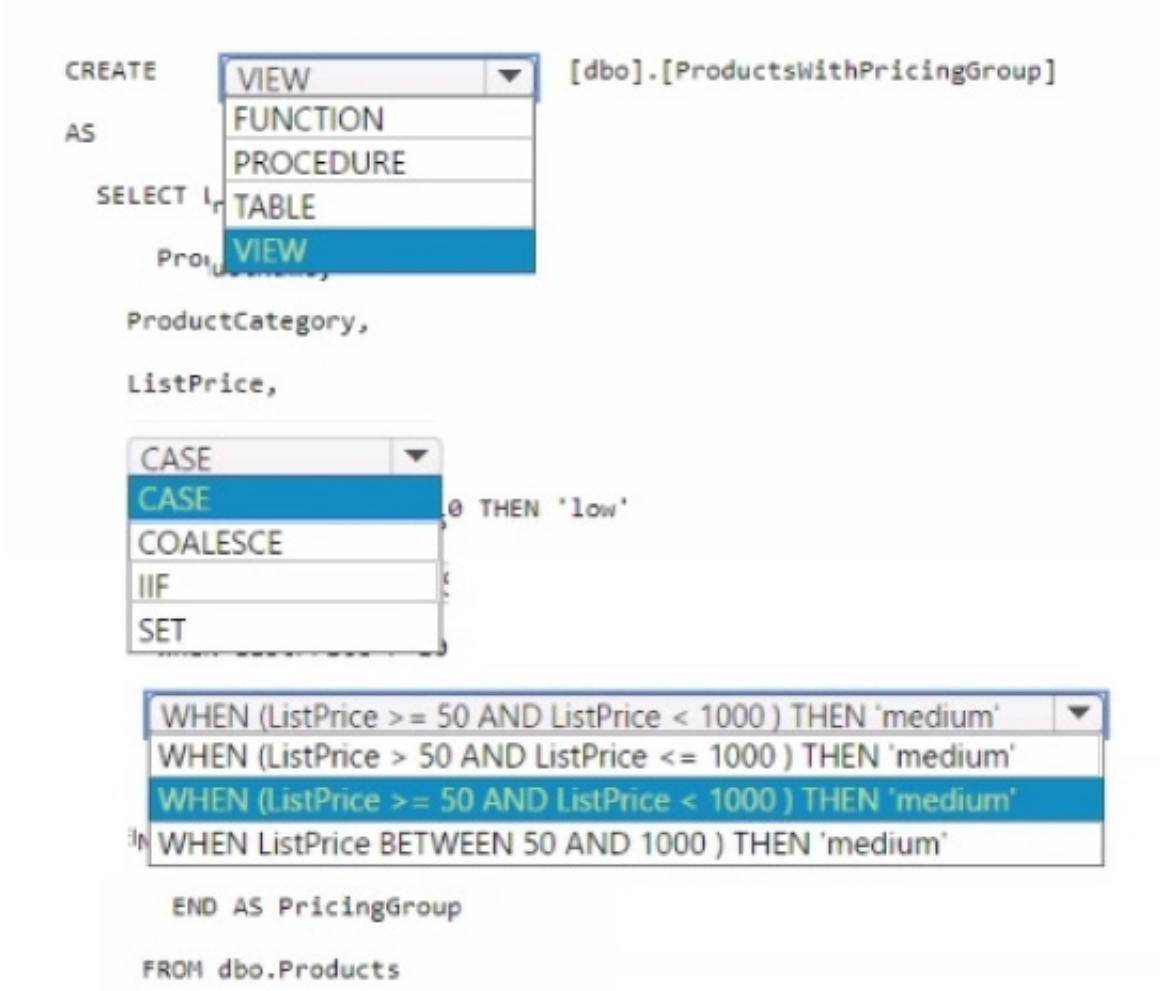
How should you complete the T-SQL statement? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:



Correct Answer:

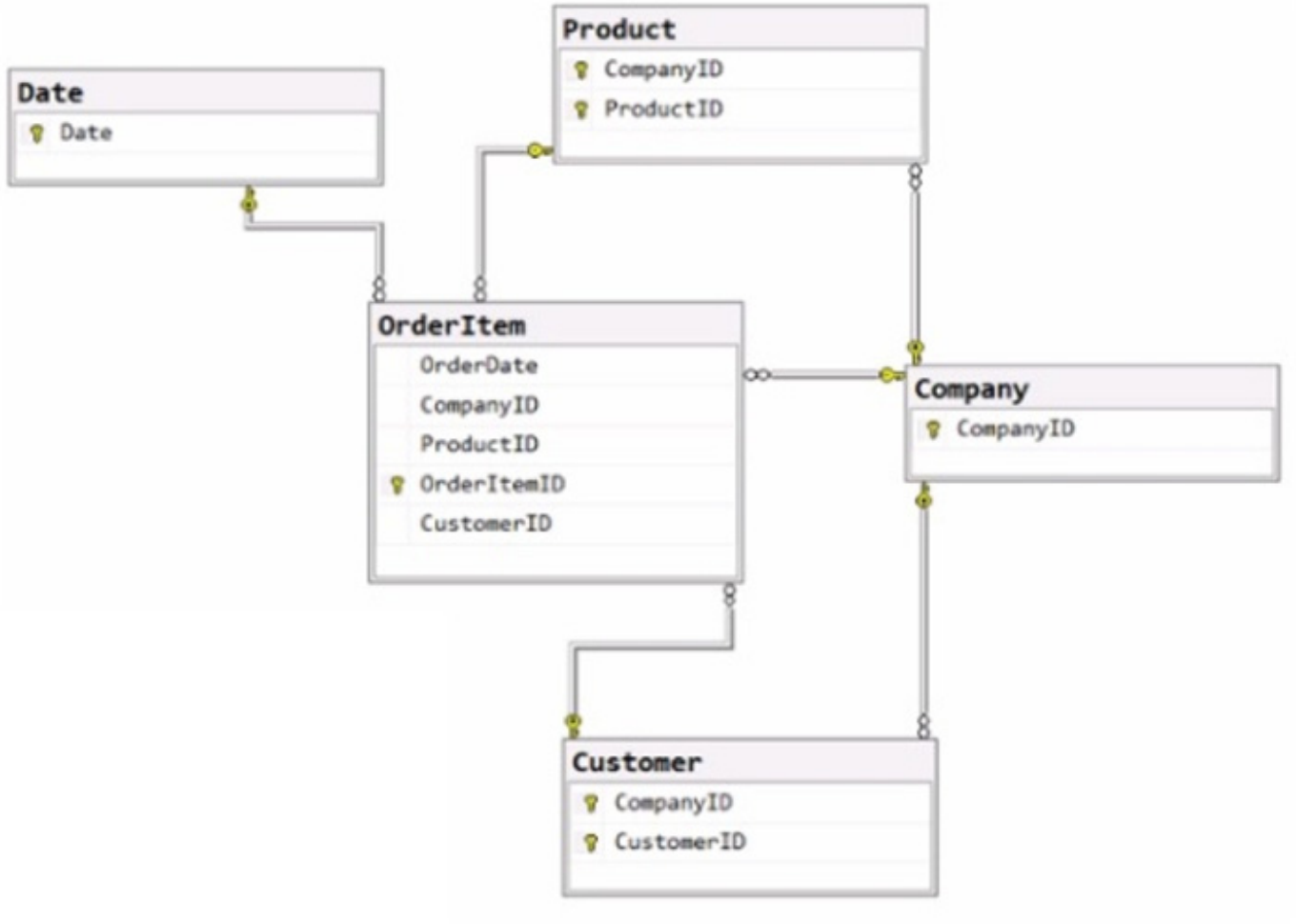


You should use CREATE VIEW to make the pricing group logic available for TSQL queries. The CASE statement should be used to determine the pricing group based on the list price. The T-SQL statement should create a view that classifies products into pricing groups based on the list price. The CASE statement is the correct conditional logic to assign each product to the appropriate pricing group. This view will

standardize the pricing group logic across different databases and semantic models.

QUESTION 8

You have the source data model shown in the following exhibit.



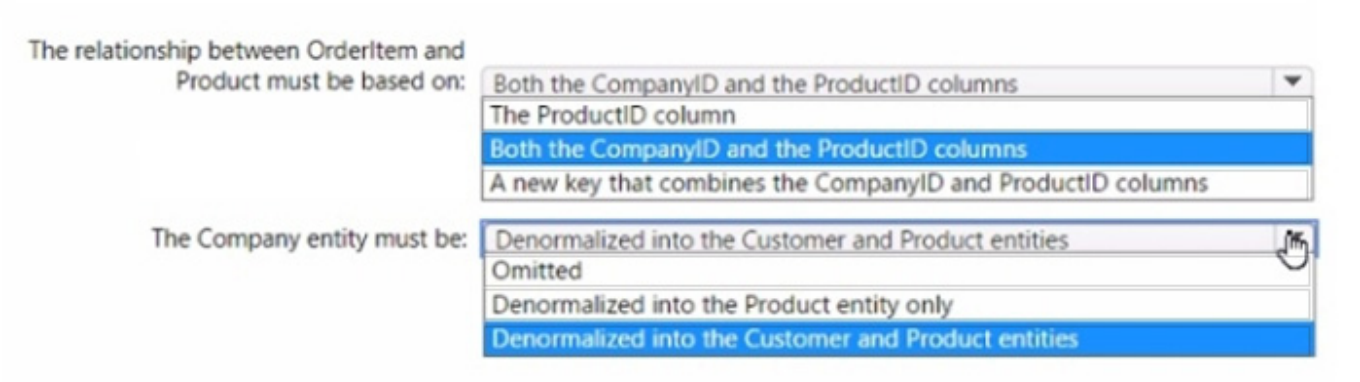
The primary keys of the tables are indicated by a key symbol beside the columns involved in each key.

You need to create a dimensional data model that will enable the analysis of order items by date, product, and customer.

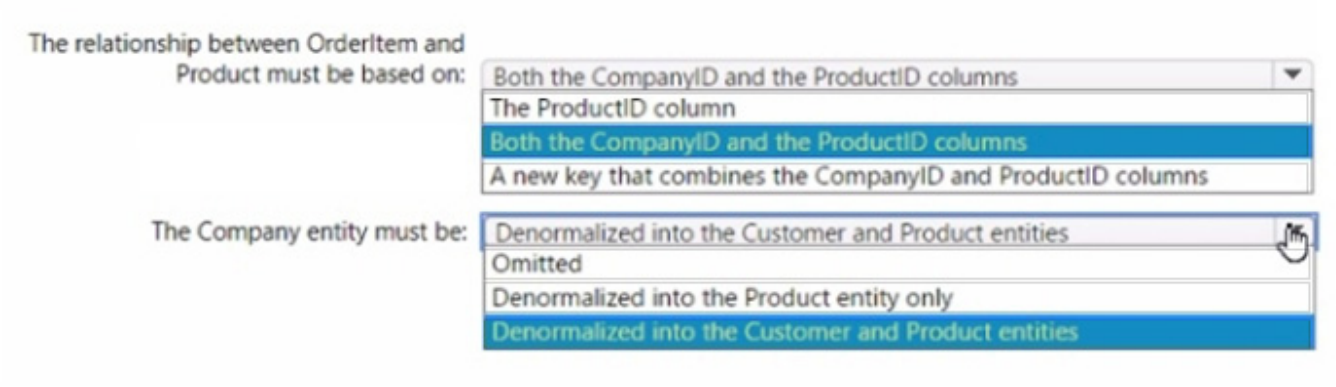
What should you include in the solution? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:



Correct Answer:



The relationship between OrderItem and Product must be based on: Both the CompanyID and the ProductID columns
The Company entity must be: Denormalized into the Customer and Product entities

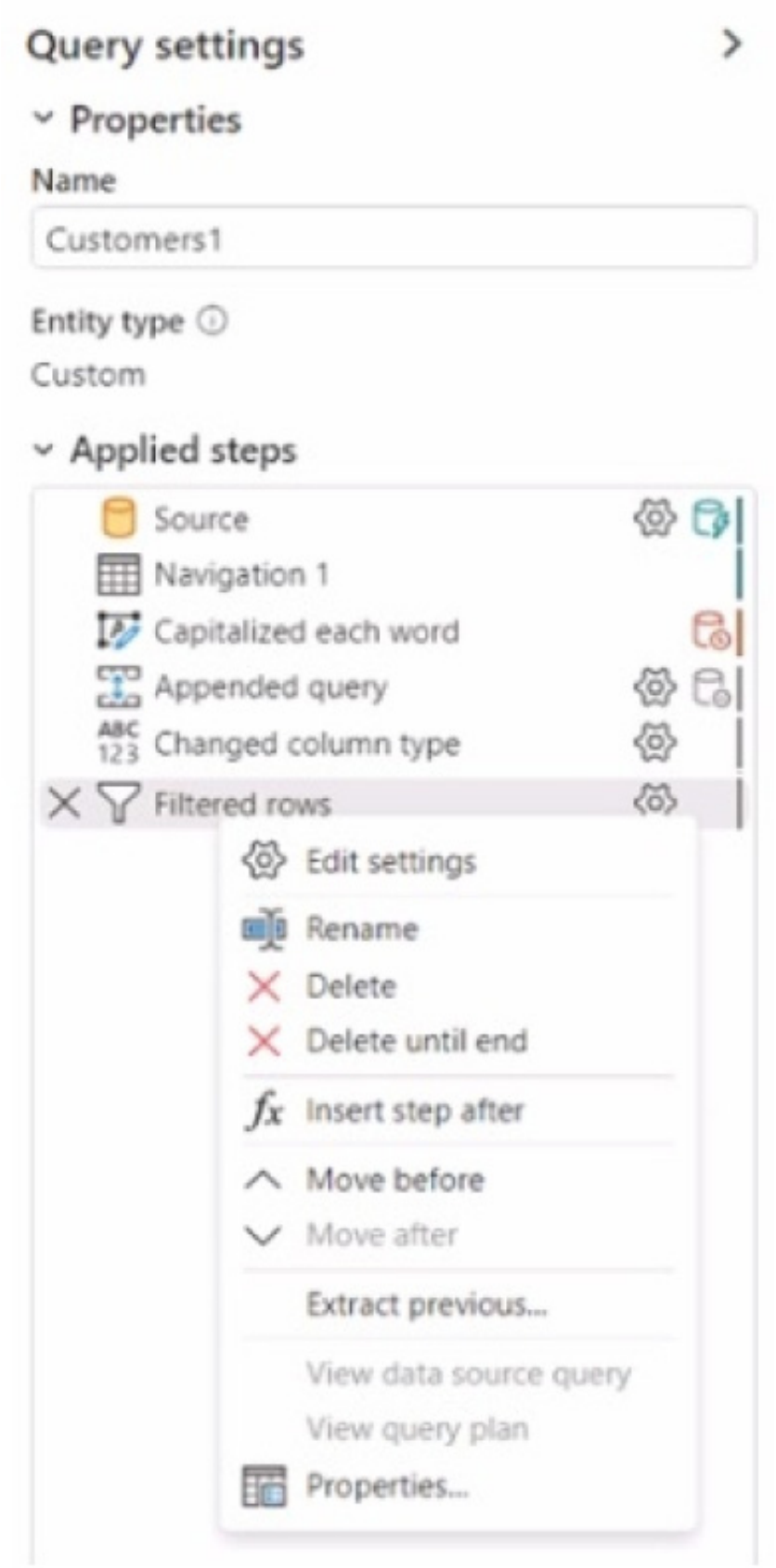
In a dimensional model, the relationships are typically based on foreign key constraints between the fact table (OrderItem) and dimension tables (Product, Customer, Date). Since CompanyID is present in both the OrderItem and Product tables, it acts as a foreign key in the relationship. Similarly, ProductID is a foreign key that relates these two tables. To enable analysis by date, product, and customer, the Company entity would need to be denormalized into the Customer and Product entities to ensure that the relevant company information is available within those dimensions for querying and reporting purposes.

References = Dimensional modeling Star schema design

QUESTION 9

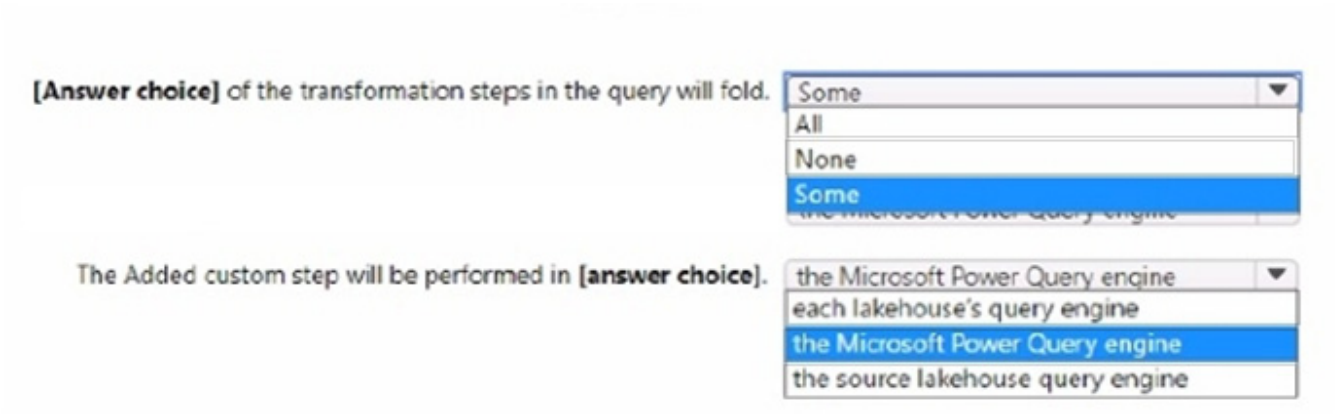
You have a Fabric tenant that contains two lakehouses.

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.

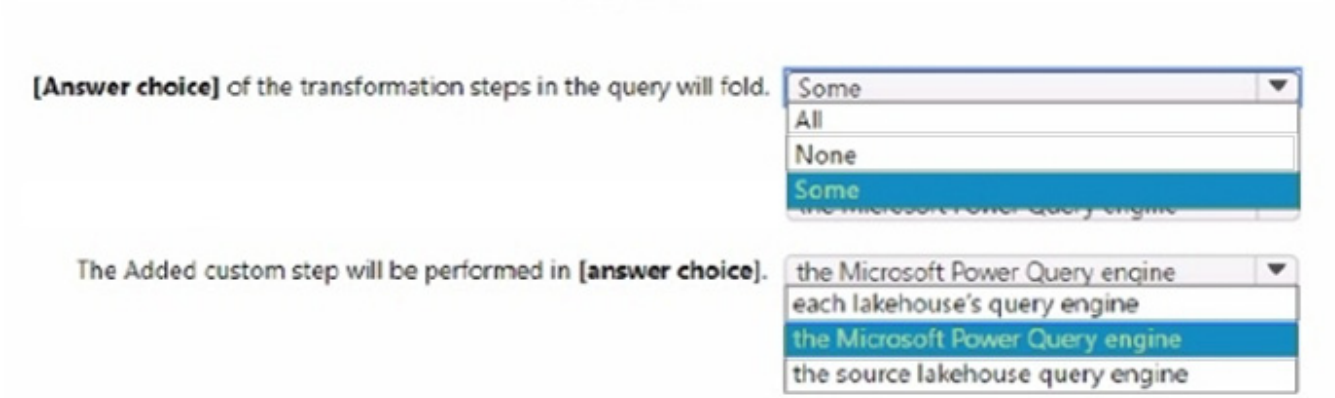


Use the drop-down menus to select the answer choice that completes each statement based on the information presented in the graphic. NOTE: Each correct selection is worth one point.

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 10

You have a Fabric tenant that contains a lakehouse named lakehouse1. Lakehouse1 contains an unpartitioned table named Table1.

You plan to copy data to Table1 and partition the table based on a date column in the source data.

You create a Copy activity to copy the data to Table1.

You need to specify the partition column in the Destination settings of the Copy activity.

What should you do first?

A. From the Destination tab, set Mode to Append.

B. From the Destination tab, select the partition column,

C. From the Source tab, select Enable partition discovery

D. From the Destination tab, set Mode to Overwrite.

Correct Answer: A

Explanation: Before specifying the partition column in the Destination settings of the Copy activity, you should set Mode to Append (A). This will allow the Copy activity to add data to the table while taking the partition column into account. References = The configuration options for Copy activities and partitioning in Azure Data Factory, which are applicable to Fabric dataflows, are outlined in the official Azure Data Factory documentation.