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.certbus.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

You need to create a DAX measure to calculate the average overall satisfaction score.

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

NOTE: Each correct selection is worth one point.

Hot Area:

```
Rolling 12 Overall Satisfaction =                   .

VAR NumberOfMonths = 12

VAR LastCurrentDate = MAX ( 'Date'[Date] )

VAR Period = DATESINPERIOD ( 'Date'[Date], LastCurrentDate, - NumberOfMonths, MONTH )

VAR Result =

    CALCULATE (

        AVERAGE('Survey'[Response Value]),
        AVERAGE('Survey'[Response Value]),
        AVERAGEA('Question'[Question Text]),
        AVERAGEX(VALUES('Survey'[Customer Key]),

        NumberOfMonths,
        LastCurrentDate,
        NumberOfMonths,
        Period,

        'Survey Question'[Question Title] = "Overall Satisfaction"

    )

RETURN

    Result
```

Correct Answer:

```
Rolling 12 Overall Satisfaction =

VAR NumberOfMonths = 12

VAR LastCurrentDate = MAX ( 'Date'[Date] )

VAR Period = DATESINPERIOD ( 'Date'[Date], LastCurrentDate, - NumberOfMonths, MONTH )

VAR Result =

    CALCULATE (
```

AVERAGE('Survey'[Response Value]),	▼
AVERAGE('Survey'[Response Value]),	
AVERAGEA('Question'[Question Text]),	
AVERAGEX(VALUES('Survey'[Customer Key]),	

NumberOfMonths,	▼
LastCurrentDate,	
NumberOfMonths,	
Period,	
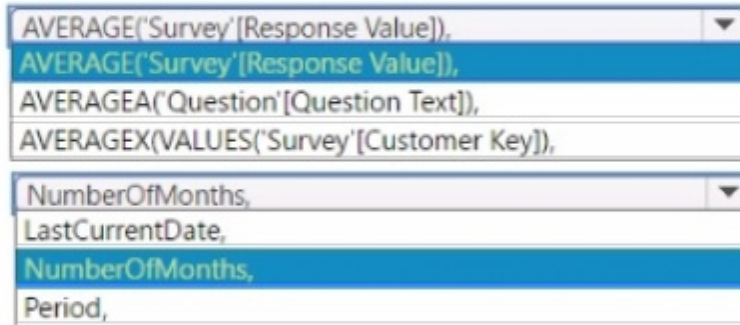
```
    'Survey Question'[Question Title] = "Overall Satisfaction"

  )

RETURN

  Result
```

The measure should use the AVERAGE function to calculate the average value.

It should reference the Response Value column from the \\'Survey\\' table.

The \\'Number of months\\' should be used to define the period for the average calculation.

To calculate the average overall satisfaction score using DAX, you would need to use the AVERAGE function on the response values related to satisfaction questions. The DATESINPERIOD function will help in calculating the rolling average

over the last 12 months.

QUESTION 2

You have a Fabric tenant that contains a semantic model named Model1. Model1 uses Import mode. Model1 contains a table named Orders. Orders has 100 million rows and the following fields.

Name	Data type	Description
OrderId	Integer	Column imported from the source
OrderDateTime	Date/time	Column imported from the source
Quantity	Integer	Column imported from the source
Price	Decimal	Column imported from the source
TotalSalesAmount	Decimal	Calculated column that multiplies Quantity and Price
TotalQuantity	Integer	Measure

You need to reduce the memory used by Model! and the time it takes to refresh the model. Which two actions should you perform? Each correct answer presents part of the solution. NOTE: Each correct answer is worth one point.

A. Split OrderDateTime into separate date and time columns.

B. Replace TotalQuantity with a calculated column.

C. Convert Quantity into the Text data type.

D. Replace TotalSalesAmount with a measure.

Correct Answer: AD

Explanation: To reduce memory usage and refresh time, splitting the OrderDateTime into separate date and time columns (A) can help optimize the model because date/time data types can be more memory-intensive than separate date and time columns. Moreover, replacing TotalSalesAmount with a measure (D) instead of a calculated column ensures that the calculation is performed at query time, which can reduce the size of the model as the value is not stored but calculated on the fly. References = The best practices for optimizing Power BI models are detailed in the Power BI documentation, which recommends using measures for calculations that don\\'t need to be stored and adjusting data types to improve performance.

QUESTION 3

You have the source data model shown in the following exhibit.
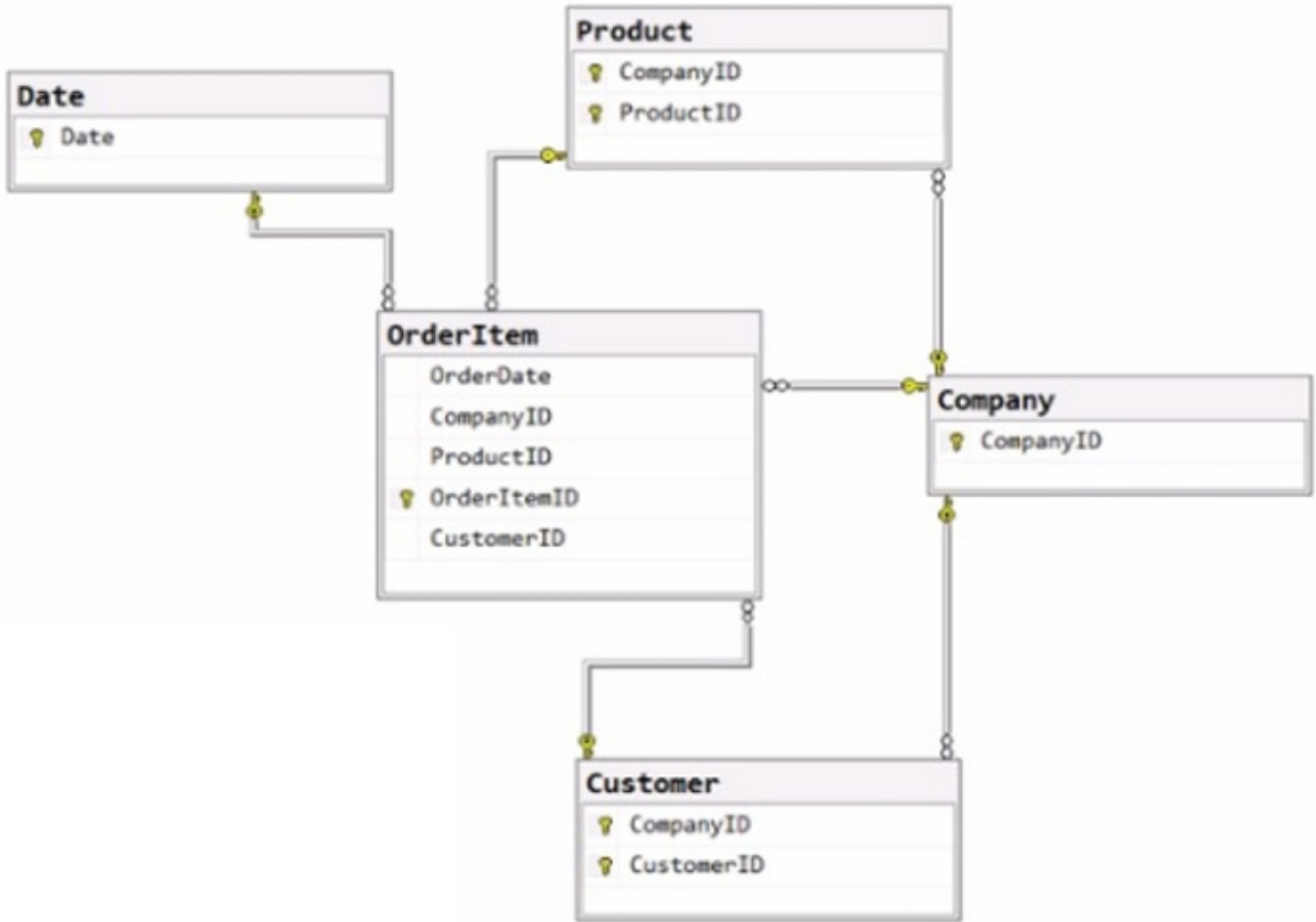
The primary keys of the tables are indicated by a key symbol beside the columns involved in each key.

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 4

You are analyzing customer purchases in a Fabric notebook by using PySpanc You have the following DataFrames:

- transactions: Contains five columns named transaction_id, customer_id, product_id, amount, and date and has 10 million rows, with each row representing a transaction
- customers: Contains customer details in 1,000 rows and three columns named customer_id, name, and country

You need to join the DataFrames on the customer_id column. The solution must minimize data shuffling. You write the following code.

```
from pyspark.sql import functions as F

results =
```

Which code should you run to populate the results DataFrame? A)

```
transactions.join(F.broadcast(customers), transactions.customer_id == customers.customer_id)
```

B)

```
transactions.join(customers, transactions.customer_id == customers.customer_id).distinct()
```

C)

```
transactions.join(customers, transactions.customer_id == customers.customer_id)
```

D)

```
transactions.crossJoin(customers).where(transactions.customer_id == customers.customer_id)
```

A. Option A

B. Option B

C. Option C

D. Option D

Correct Answer: A

Explanation: The correct code to populate the results DataFrame with minimal data shuffling is Option A. Using the broadcast function in PySpark is a way to minimize data movement by broadcasting the smaller DataFrame (customers) to each node in the cluster. This is ideal when one DataFrame is much smaller than the other, as in this case with customers. References = You can refer to the official Apache Spark documentation for more details on joins and the broadcast hint.

QUESTION 5

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.

Query settings >

˅ Properties

Name

Customers1

Entity type ⓘ
Custom

˅ Applied steps

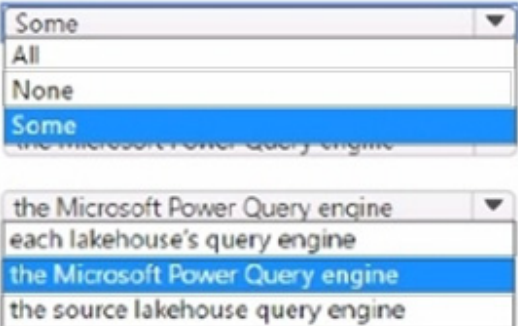|		
🛢 Source	⚙ 🗗	
▦ Navigation 1		
📝 Capitalized each word	🔒	
⬇ Appended query	⚙ 🔒	
ABC 123 Changed column type	⚙	
✕ ▽ Filtered rows	⚙	

⚙ Edit settings

🔲 Rename

✕ Delete

✕ Delete until end

ƒx Insert step after

⌃ Move before

∨ Move after

Extract previous...

View data source query

View query plan

🔳 Properties...

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:
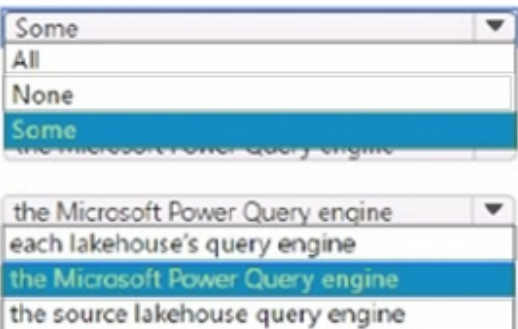
Correct Answer:

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 6

You need to provide Power BI developers with access to the pipeline. The solution must meet the following requirements:

Ensure that the developers can deploy items to the workspaces for Development and Test.

Prevent the developers from deploying items to the workspace for Production.

Follow the principle of least privilege.

Which three levels of access should you assign to the developers? Each correct answer presents part of the solution. NOTE: Each correct answer is worth one point.

A. Build permission to the production semantic models

B. Admin access to the deployment pipeline

C. Viewer access to the Development and Test workspaces

D. Viewer access to the Production workspace

E. Contributor access to the Development and Test workspaces

F. Contributor access to the Production workspace

Correct Answer: BDE

Explanation: To meet the requirements, developers should have Admin access to the deployment pipeline (B), Contributor access to the Development and Test workspaces (E), and Viewer access to the Production workspace (D). This setup ensures they can perform necessary actions in development and test environments without having the ability to affect production. References = The Power BI documentation on workspace access levels and deployment pipelines provides guidelines on assigning appropriate permissions.

QUESTION 7

You are creating a dataflow in Fabric to ingest data from an Azure SQL database by using a T-SQL statement.

You need to ensure that any foldable Power Query transformation steps are processed by the Microsoft SQL Server engine.

How should you complete the code? To answer, drag the appropriate values to the correct targets. Each value 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:

Values		Answer Area
| | | ```
let

    Source = Sql.Databases(

        "server.database.windows.net"

    ),

    Database = Source{[Name = "db"]}[Data],

    Query =  [Table]  -  [NativeQuery]  (

        Database,

        "  SELECT * FROM customer WHERE country IN ('USA', 'UK')",

        null,

        [  [EnableFolding]  = true]

    )

in

    Query
``` |

Values box contains: Optimize, Record, StopFolding, Value

You should complete the code as follows: Table NativeQuery EnableFolding

In Power Query, using Table before the SQL statement ensures that the result of the SQL query is treated as a table. NativeQuery allows a native database query to be passed through from Power Query to the source database. The EnableFolding option ensures that any subsequent transformations that can be folded will be sent back and executed at the source database (Microsoft SQL Server engine in this case).

QUESTION 8

You have a Fabric tenant named Tenant1 that contains a workspace named WS1. WS1 uses a capacity named C1 and contains a dawset named DS1. You need to ensure read- write access to DS1 is available by using the XMLA endpoint. What should be modified first?

A. the DS1 settings

B. the WS1 settings

C. the C1 settings

D. the Tenant1 settings

Correct Answer: C

Explanation: To ensure read-write access to DS1 is available by using the XMLA endpoint, the C1 settings (which refer to the capacity settings) should be modified first. XMLA endpoint configuration is a capacity feature, not specific to individual datasets or workspaces. References = The configuration of XMLA endpoints in Power BI capacities is detailed in the Power BI documentation on dataset management.

QUESTION 9

You have a Fabric workspace named Workspace1 and an Azure Data Lake Storage Gen2 account named storage"!. Workspace1 contains a lakehouse named Lakehouse1.

You need to create a shortcut to storage! in Lakehouse1.

Which connection and endpoint should you specify? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Hot Area:

Connection:	abfss ▼
	abfs
	abfss
	https

Endpoint:	dfs ▼
	blob
	dfs
	file

Correct Answer:

Connection:	abfss ▼
	abfs
	abfss
	https

Endpoint:	dfs ▼
	blob
	dfs
	file

When creating a shortcut to an Azure Data Lake Storage Gen2 account in a lakehouse, you should use the abfss (Azure Blob File System Secure) connection string and the dfs (Data Lake File System) endpoint. The abfss is used for secure

access to Azure Data Lake Storage, and the dfs endpoint indicates that the Data Lake Storage Gen2 capabilities are to be used.

QUESTION 10

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 .sumary ()

Does this meet the goal?

A. Yes

B. No

Correct Answer: A

Explanation: Yes, the df.summary() method does meet the goal. This method is used to compute specified statistics for numeric and string columns. By default, it provides statistics such as count, mean, stddev, min, and max. References = The PySpark API documentation details the summary() function and the statistics it provides.

DP-600 PDF Dumps   DP-600 Study Guide   DP-600 Braindumps