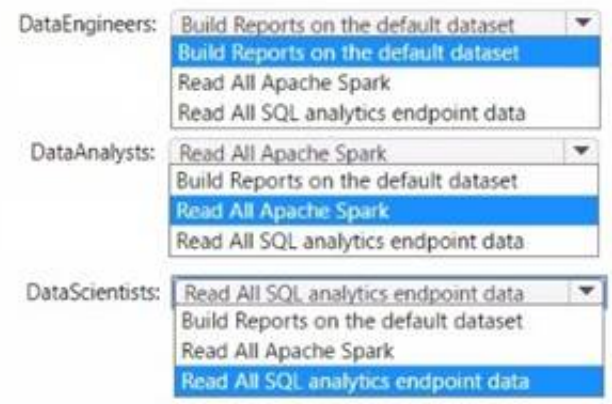
Microsoft

Exam Questions DP-600

Implementing Analytics Solutions Using Microsoft Fabric

NEW QUESTION 1
HOTSPOT - (Topic 1)
You to need assign permissions for the data store in the AnalyticsPOC workspace. The solution must meet the security requirements.
Which additional permissions should you assign when you share the data store? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.

Answer Area



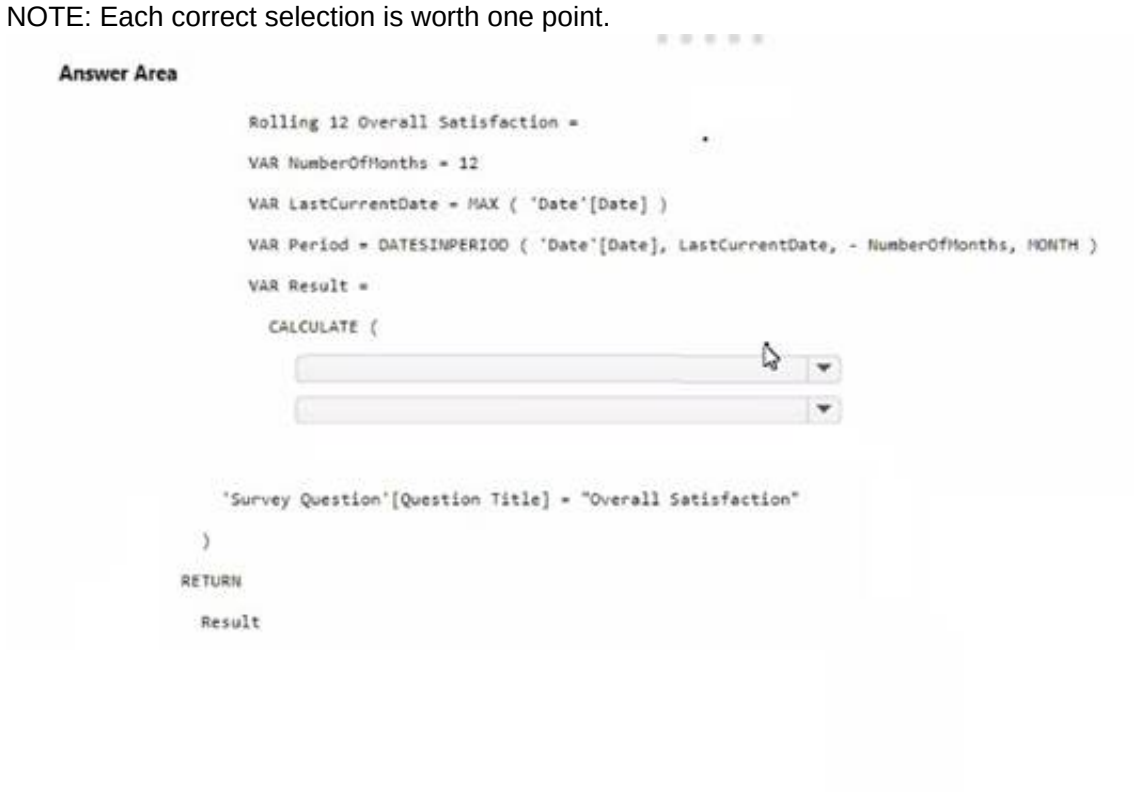
A. Mastered
B. Not Mastered

Answer: A

Explanation:
? Data Engineers: Read All SQL analytics endpoint data
? Data Analysts: Read All Apache Spark
? Data Scientists: Read All SQL analytics endpoint data
The permissions for the data store in the AnalyticsPOC workspace should align with the principle of least privilege:
? Data Engineers need read and write access but not to datasets or reports.
? Data Analysts require read access specifically to the dimensional model objects and the ability to create Power BI reports.
? Data Scientists need read access via Spark notebooks. These settings ensure each role has the necessary permissions to fulfill their responsibilities without exceeding their required access level.

NEW QUESTION 2
HOTSPOT - (Topic 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.

Answer Area



Answer 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
```

A. Mastered
B. Not Mastered

Answer: A

Explanation:
? 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.

NEW QUESTION 3
- (Topic 2)
You have a Fabric tenant that uses a Microsoft tower Bl Premium capacity. You need to enable scale-out for a semantic model. What should you do first?

A. At the semantic model level, set Large dataset storage format to Off.
B. At the tenant level, set Create and use Metrics to Enabled.
C. At the semantic model level, set Large dataset storage format to On.
D. At the tenant level, set Data Activator to Enabled.

Answer: C

Explanation:
To enable scale-out for a semantic model, you should first set Large dataset storage format to On (C) at the semantic model level. This configuration is necessary to handle larger datasets effectively in a scaled-out environment. References = Guidance on configuring large dataset storage formats for scale-out is available in the Power BI documentation.

NEW QUESTION 4
- (Topic 2)
You have a Fabric tenant that contains a lakehouse. You plan to use a visual query to merge two tables.
You need to ensure that the query returns all the rows that are present in both tables. Which type of join should you use?

A. left outer
B. right anti
C. full outer
D. left anti
E. right outer
F. inner

Answer: C

Explanation:
When you need to return all rows that are present in both tables, you use a full outer join. This type of join combines the results of both left and right outer joins and returns all rows from both tables, with matching rows from both sides where available. If there is no match, the result is NULL on the side of the join where there is no match. References: Information about joins and their use in querying data in a lakehouse can be typically found in the SQL and data processing documentation of the Fabric tenant or lakehouse solutions.

NEW QUESTION 5
- (Topic 2)
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.

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.

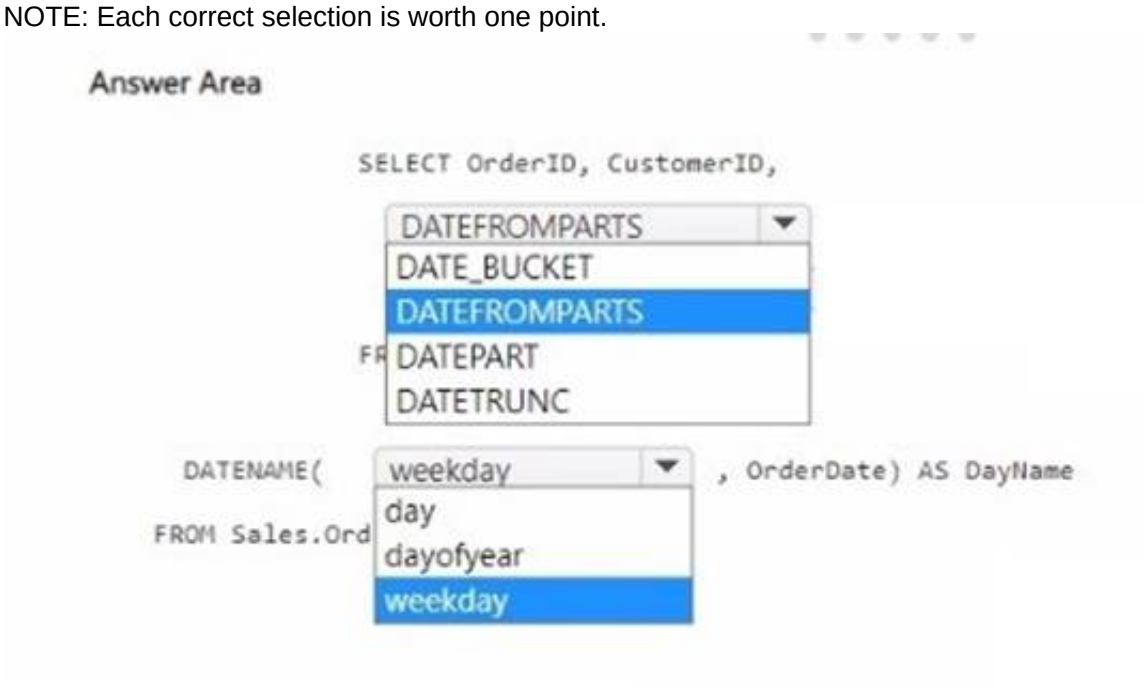
NEW QUESTION 6
HOTSPOT - (Topic 2)
You have a Fabric warehouse that contains a table named Sales.Orders. Sales.Orders contains the following columns.

Name	Data type	Nullable
OrderID	Integer	No
CustomerID	Integer	No
OrderDate	Date	No
Quantity	Integer	Yes
Weight	Decimal(18, 3)	Yes
ListPrice	Decimal(18, 2)	No
SalePrice	Decimal(18, 2)	Yes

You need to write a T-SQL query that will return the following columns.

Name	Description
OrderID	Returns OrderID
CustomerID	Returns CustomerID
PeriodDate	Returns a date representing the first day of the month for OrderDate
DayName	Returns the name of the day for OrderDate, such as Wednesday

How should you complete the code? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.



A. Mastered
B. Not Mastered

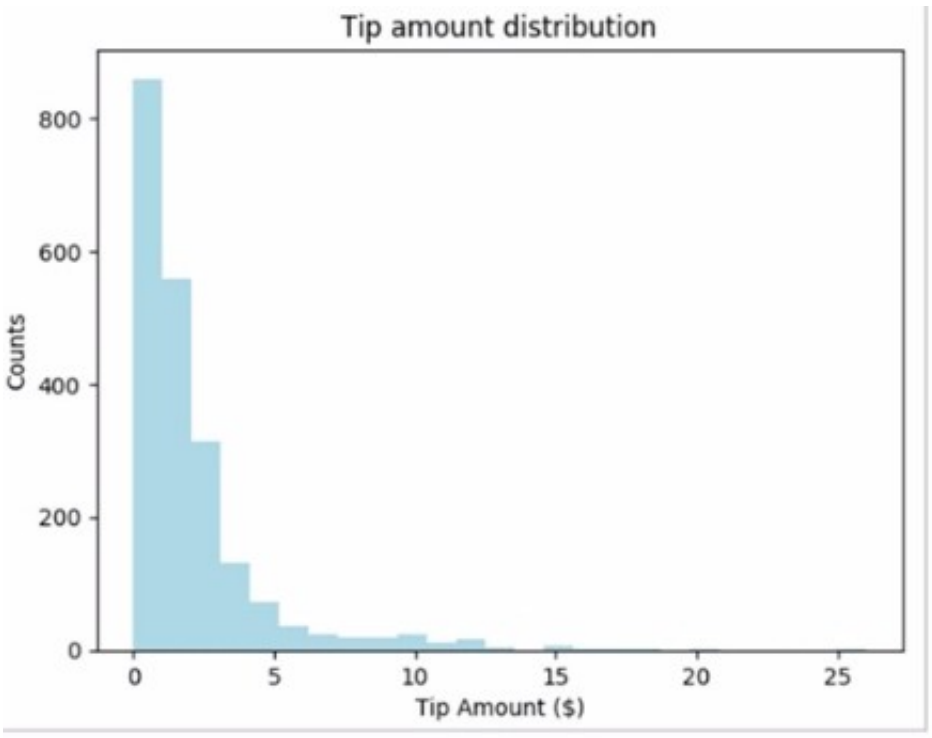
Answer: A

Explanation:
For the PeriodDate that returns the first day of the month for OrderDate, you should use DATEFROMPARTS as it allows you to construct a date from its individual components (year, month, day).
For the DayName that returns the name of the day for OrderDate, you should use
DATENAME with the weekday date part to get the full name of the weekday. The complete SQL query should look like this:
SELECT OrderID, CustomerID,
DATEFROMPARTS(YEAR(OrderDate), MONTH(OrderDate), 1) AS PeriodDate, DATENAME(weekday, OrderDate) AS DayName
FROM Sales.Orders
Select DATEFROMPARTS for the PeriodDate and weekday for the DayName in the answer area.

NEW QUESTION 7
- (Topic 2)
You have a Fabric notebook that has the Python code and output shown in the following exhibit.

```
# Look at a histogram of tips by count by using Matplotlib

ax1 = sampled_taxi_pd_df['tipAmount'].plot(kind='hist', bins=25, facecolor='lightblue'
ax1.set_title('Tip amount distribution')
ax1.set_xlabel('Tip Amount ($)')
ax1.set_ylabel('Counts')
plt.suptitle('')
plt.show()
```



Which type of analytics are you performing?

A. predictive
B. descriptive
C. prescriptive
D. diagnostic

Answer: B

Explanation:
The Python code and output shown in the exhibit display a histogram, which is a representation of the distribution of data. This kind of analysis is descriptive analytics, which is used to describe or summarize the features of a dataset. Descriptive analytics answers the question of "what has happened" by providing insight into past data through tools such as mean, median, mode, standard deviation, and graphical representations like histograms.
References: Descriptive analytics and the use of histograms as a way to visualize data distribution are basic concepts in data analysis, often covered in introductory analytics and Python programming resources.

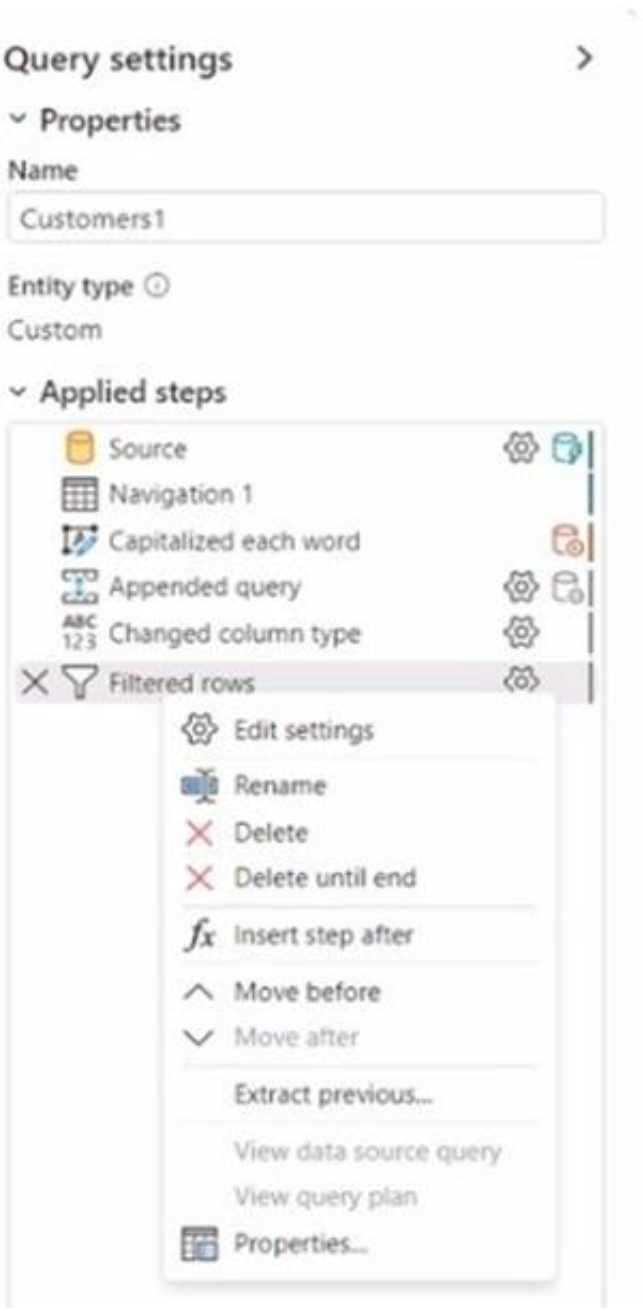
NEW QUESTION 8
- (Topic 2)
You have a Fabric tenant named Tenant1 that contains a workspace named WS1. WS1 uses a capacity named C1 and contains a dawset named DS1. You need to ensure read- write access to DS1 is available by using the XMLA endpoint. What should be modified first?

A. the DS1 settings
B. the WS1 settings
C. the C1 settings
D. the Tenant1 settings

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.

NEW QUESTION 9
HOTSPOT - (Topic 2)
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.

A. Mastered
B. Not Mastered

Answer: A

Explanation:
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

NEW QUESTION 10
- (Topic 2)
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.show()
Does this meet the goal?

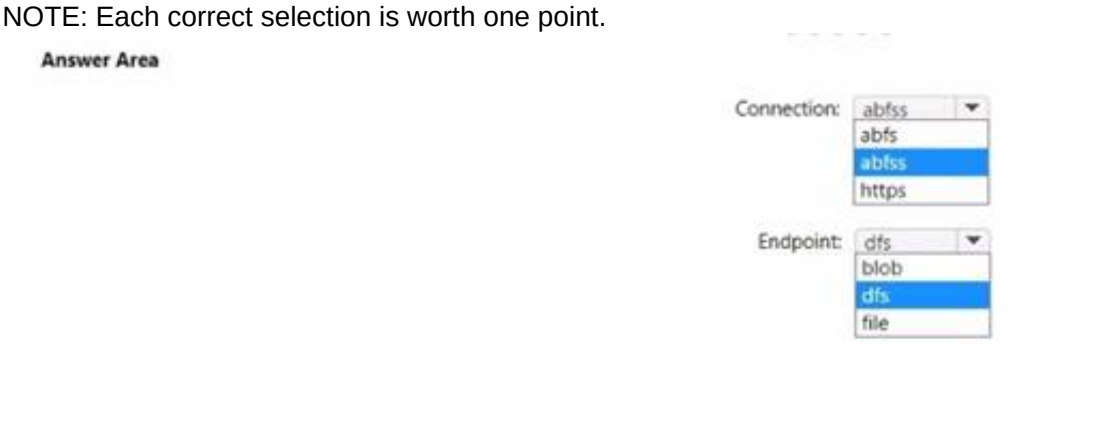
A. Yes
B. No

Answer: B

Explanation:
The df.show() method also does not meet the goal. It is used to show the contents of the DataFrame, not to compute statistical functions. References = The usage of the show() function is documented in the PySpark API documentation.

NEW QUESTION 10
HOTSPOT - (Topic 2)

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.



A. Mastered
B. Not Mastered

Answer: A

Explanation:
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.

NEW QUESTION 12
- (Topic 2)
You have a Fabric tenant that contains a takehouse named lakehouse1. Lakehouse1 contains a Delta table named Customer.
When you query Customer, you discover that the query is slow to execute. You suspect that maintenance was NOT performed on the table.
You need to identify whether maintenance tasks were performed on Customer. Solution: You run the following Spark SQL statement:
DESCRIBE HISTORY customer Does this meet the goal?

A. Yes
B. No

Answer: A

Explanation:
Yes, the DESCRIBE HISTORY statement does meet the goal. It provides information on the history of operations, including maintenance tasks, performed on a Delta table. References = The functionality of the DESCRIBE HISTORY statement can be verified in the Delta Lake documentation.

NEW QUESTION 13
- (Topic 2)
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

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.

NEW QUESTION 17
- (Topic 2)
You have a Microsoft Fabric tenant that contains a dataflow. You are exploring a new semantic model.
From Power Query, you need to view column information as shown in the following exhibit.

Which three Data view options should you select? Each correct answer presents part of the solution. NOTE: Each correct answer is worth one point.

A. Enable column profile
B. Show column quality details
C. Show column profile in details pane
D. Enable details pane
E. Show column value distribution

Answer: ABE

Explanation:
To view column information like the one shown in the exhibit in Power Query, you need to select the options that enable profiling and display quality and distribution details. These are: A. Enable column profile - This option turns on profiling for each column, showing statistics such as distinct and unique values. B. Show column quality details - It displays the column quality bar on top of each column showing the percentage of valid, error, and empty values. E. Show column value distribution - It enables the histogram display of value distribution for each column, which visualizes how often each value occurs.
References: These features and their descriptions are typically found in the Power Query documentation, under the section for data profiling and quality features.

NEW QUESTION 22
DRAG DROP - (Topic 2)
You are implementing two dimension tables named Customers and Products in a Fabric warehouse.
You need to use slowly changing dimension (SCO) to manage the versioning of data. The solution must meet the requirements shown in the following table.

Table	Change action
Customers	Create a new version of the row.
Products	Overwrite the existing value in the latest row.

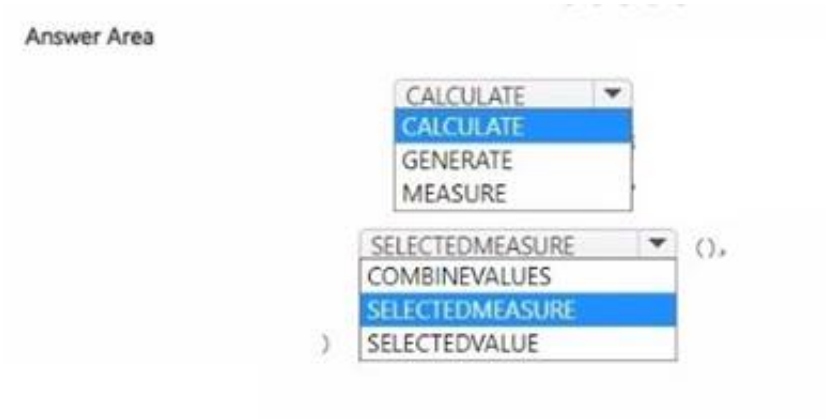
Which type of SCD should you use for each table? To answer, drag the appropriate SCD types to the correct tables. Each SCD type may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.
NOTE: Each correct selection is worth one point.

A. Mastered
B. Not Mastered

Answer: A

Explanation:
For the Customers table, where the requirement is to create a new version of the row, you would use:
? Type 2 SCD: This type allows for the creation of a new record each time a change occurs, preserving the history of changes over time.
For the Products table, where the requirement is to overwrite the existing value in the latest row, you would use:
? Type 1 SCD: This type updates the record directly, without preserving historical data.

NEW QUESTION 26
HOTSPOT - (Topic 2)
You have a Microsoft Power Bl semantic model. You plan to implement calculation groups.
You need to create a calculation item that will change the context from the selected date to month-to-date (MTD).
How should you complete the DAX expression? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.



A. Mastered
B. Not Mastered

Answer: A

Explanation:
To create a calculation item that changes the context from the selected date to month-to- date (MTD), the appropriate DAX expression involves using the CALCULATE function to alter the filter context and the DATESMTD function to specify the month-to-date context. The correct completion for the DAX expression would be:
? In the first dropdown, select CALCULATE.
? In the second dropdown, select SELECTEDMEASURE. This would create a DAX expression in the form:
CALCULATE(SELECTEDMEASURE(),
DATESMTD('Date'[DateColumn])
)

NEW QUESTION 30
- (Topic 2)
You have an Azure Repos Git repository named Repo1 and a Fabric-enabled Microsoft Power Bl Premium capacity. The capacity contains two workspaces named Workspace! and Workspace2. Git integration is enabled at the workspace level.
You plan to use Microsoft Power Bl Desktop and Workspace! to make version-controlled changes to a semantic model stored in Repo1. The changes will be built and deployed lo Workspace2 by using Azure Pipelines.
You need to ensure that report and semantic model definitions are saved as individual text files in a folder hierarchy. The solution must minimize development and maintenance effort.
In which file format should you save the changes?

A. PBIP
B. PBIT
C. PBIX
D. PBIDS

Answer: C

Explanation:
When working with Power BI Desktop and Git integration for version control, report and semantic model definitions should be saved in the PBIX format. PBIX is the Power BI Desktop file format that contains definitions for reports, data models, and queries, and it can be easily saved and tracked in a version-controlled environment. The solution should minimize development and maintenance effort, and saving in PBIX format allows for the easiest transition from development to deployment, especially when using Azure Pipelines for CI/CD (continuous integration/continuous deployment) practices.
References: The use of PBIX files with Power BI Desktop and Azure Repos for version control is discussed in Microsoft’s official Power BI documentation, particularly in the sections covering Power BI Desktop files and Azure DevOps integration.

NEW QUESTION 34
- (Topic 2)
You are creating a semantic model in Microsoft Power Bl Desktop.
You plan to make bulk changes to the model by using the Tabular Model Definition Language (TMDL) extension for Microsoft Visual Studio Code.
You need to save the semantic model to a file. Which file format should you use?

A. PBIP
B. PBIX
C. PBIT
D. PBIDS

Answer: B

Explanation:
When saving a semantic model to a file that can be edited using the Tabular Model Scripting Language (TMSL) extension for Visual Studio Code, the PBIX (Power BI Desktop) file format is the correct choice. The PBIX format contains the report, data model, and queries, and is the primary file format for editing in Power BI Desktop. References = Microsoft's documentation on Power BI file formats and Visual Studio Code provides further clarification on the usage of PBIX files.

NEW QUESTION 36
- (Topic 2)
You have a Fabric warehouse that contains a table named Staging.Sales. Staging.Sales contains the following columns.

Name	Data type	Nullable
ProductID	Integer	No
ProductName	Varchar(30)	No
SalesDate	Datetime2(6)	No
WholesalePrice	Decimal(18, 2)	Yes
Amount	Decimal(18, 2)	Yes

You need to write a T-SQL query that will return data for the year 2023 that displays ProductID and ProductName arxl has a summarized Amount that is higher than 10,000. Which query should you use?

A)

```
SELECT ProductID, ProductName, SUM(Amount) AS TotalAmount
FROM Staging.Sales
WHERE DATEPART(YEAR,SaleDate) = '2023'
GROUP BY ProductID, ProductName
HAVING SUM(Amount) > 10000
```

B)

```
SELECT ProductID, ProductName, SUM(Amount) AS TotalAmount
FROM Staging.Sales
GROUP BY ProductID, ProductName
HAVING DATEPART(YEAR,SaleDate) = '2023' AND SUM(Amount) > 10000
```

C)

```
SELECT ProductID, ProductName, SUM(Amount) AS TotalAmount
FROM Staging.Sales
WHERE DATEPART(YEAR,SaleDate) = '2023' AND SUM(Amount) > 10000
```

D)

```
SELECT ProductID, ProductName, SUM(Amount) AS TotalAmount
FROM Staging.Sales
WHERE DATEPART(YEAR,SaleDate) = '2023'
GROUP BY ProductID, ProductName
HAVING TotalAmount > 10000
```

A. Option A
B. Option B
C. Option C
D. Option D

Answer: B

Explanation:
The correct query to use in order to return data for the year 2023 that displays ProductID, ProductName, and has a summarized Amount greater than 10,000 is Option B. The reason is that it uses the GROUP BY clause to organize the data by ProductID and ProductName and then filters the result using the HAVING clause to only include groups where the sum of Amount is greater than 10,000. Additionally, the DATEPART(YEAR, SaleDate) = '2023' part of the HAVING clause ensures that only records from the year 2023 are included. References = For more information, please visit the official documentation on T-SQL queries and the GROUP BY clause at T-SQL GROUP BY.

NEW QUESTION 41
.....

Thank You for Trying Our Product

We offer two products:

1st - We have Practice Tests Software with Actual Exam Questions

2nd - Questons and Answers in PDF Format

DP-600 Practice Exam Features:

* DP-600 Questions and Answers Updated Frequently

* DP-600 Practice Questions Verified by Expert Senior Certified Staff

* DP-600 Most Realistic Questions that Guarantee you a Pass on Your FirstTry

* DP-600 Practice Test Questions in Multiple Choice Formats and Updatesfor 1 Year