Microsoft

Exam Questions DP-600

Implementing Analytics Solutions Using Microsoft Fabric

NEW QUESTION 1
- (Topic 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

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.

NEW QUESTION 2
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.

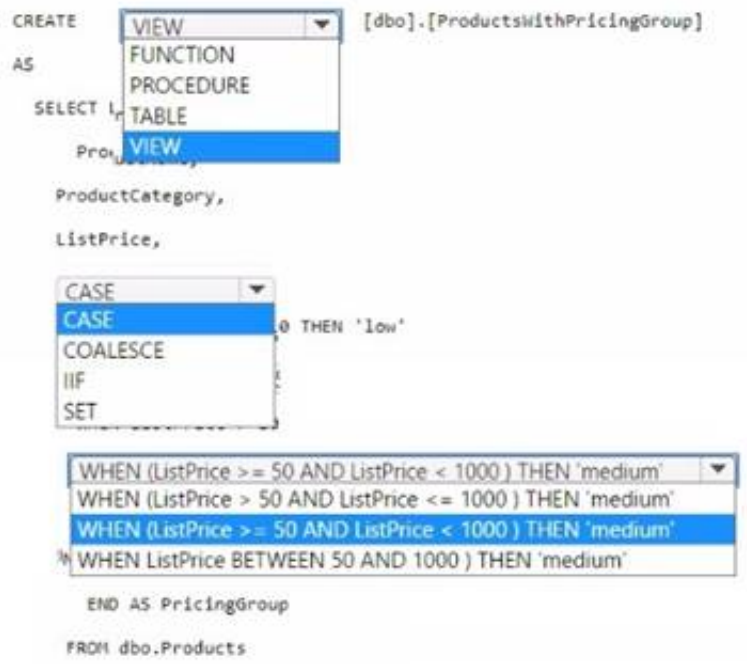
A. Mastered
B. Not Mastered

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 3
HOTSPOT - (Topic 1)
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.

Answer Area



A. Mastered
B. Not Mastered

Answer: A

Explanation:
C:\Users\Waqas Shahid\Desktop\Mudassir\Untitled.jpg
? You should use CREATE VIEW to make the pricing group logic available for T- SQL 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.

NEW QUESTION 4
- (Topic 1)
You need to ensure the data loading activities in the AnalyticsPOC workspace are executed in the appropriate sequence. The solution must meet the technical requirements.
What should you do?

A. Create a pipeline that has dependencies between activities and schedule the pipeline.
B. Create and schedule a Spark job definition.
C. Create a dataflow that has multiple steps and schedule the dataflow.
D. Create and schedule a Spark notebook.

Answer: A

Explanation:
To meet the technical requirement that data loading activities must ensure the raw and cleansed data is updated completely before populating the dimensional model, you would need a mechanism that allows for ordered execution. A pipeline in Microsoft Fabric with dependencies set between activities can ensure that activities are executed in a specific sequence. Once set up, the pipeline can be scheduled to run at the required intervals (hourly or daily depending on the data source).

NEW QUESTION 5
HOTSPOT - (Topic 2)
You have a Fabric tenant that contains a warehouse named Warehouse1. Warehouse1 contains three schemas named schemaA, schemaB. and schemaC
You need to ensure that a user named User1 can truncate tables in schemaA only.
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.

Answer Area

A. Mastered
B. Not Mastered

Answer: A

Explanation:
? GRANT ALTER ON SCHEMA::schemaA TO User1;
The ALTER permission allows a user to modify the schema of an object, and granting ALTER on a schema will allow the user to perform operations like TRUNCATE TABLE on any object within that schema. It is the correct permission to grant to User1 for truncating tables in schemaA.

References =
? GRANT Schema Permissions
? Permissions That Can Be Granted on a Schema

NEW QUESTION 6
- (Topic 2)
You have a Fabric tenant that contains 30 CSV files in OneLake. The files are updated daily.
You create a Microsoft Power Bl semantic model named Modell that uses the CSV files as a data source. You configure incremental refresh for Model 1 and publish the model to a Premium capacity in the Fabric tenant.
When you initiate a refresh of Model1, the refresh fails after running out of resources. What is a possible cause of the failure?

A. Query folding is occurring.
B. Only refresh complete days is selected.
C. XMLA Endpoint is set to Read Only.
D. Query folding is NOT occurring.
E. The data type of the column used to partition the data has changed.

Answer: E

Explanation:
A possible cause for the failure is that query folding is NOT occurring (D). Query folding helps optimize refresh by pushing down the query logic to the source system, reducing the amount of data processed and transferred, hence conserving resources. References = The Power BI documentation on incremental refresh and query folding provides detailed information on this topic.

NEW QUESTION 7
HOTSPOT - (Topic 2)
You have a Fabric tenant that contains a warehouse named Warehouse1. Warehouse1 contains a fact table named FactSales that has one billion rows. You run the following T- SQL statement.
CREATE TABLE test.FactSales AS CLONE OF Dbo.FactSales;
For each of the following statements, select Yes if the statement is true. Otherwise, select No. NOTE: Each correct selection is worth one point.
Answer Area

Statements	Yes	No
A replica of `dbo.Sales` is created in the test schema by copying the metadata only.	○	○
Additional schema changes to `dbo.FactSales` will also apply to `test.FactSales`.	○	○
Additional data changes to `dbo.FactSales` will also apply to `test.FactSales`.	○	○

A. Mastered
B. Not Mastered

Answer: A

Explanation:
? A replica of dbo.Sales is created in the test schema by copying the metadata only.
- No
? Additional schema changes to dbo.FactSales will also apply to test.FactSales. - No
? Additional data changes to dbo.FactSales will also apply to test.FactSales. - Yes
The CREATE TABLE AS CLONE statement creates a copy of an existing table, including its data and any associated data structures, like indexes. Therefore, the statement does not merely copy metadata; it also copies the data. However, subsequent schema changes to the original table do not automatically propagate to the cloned table. Any data changes in the original table after the clone operation will not be reflected in the clone unless explicitly updated.
References =
? CREATE TABLE AS SELECT (CTAS) in SQL Data Warehouse

NEW QUESTION 8
HOTSPOT - (Topic 2)
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.
Answer Area

```
WITH CUSTOMERBASE AS (
  SELECT [CustomerID]
  ,[CustomerName]
  ,[PostalCode]
  ,[LastUpdated]
  ,X = ROW_NUMBER()  OVER (PARTITION BY CustomerID ORDER BY LastUpdated DESC)
        LAST_Value ()
        NTILE()
        ROW_NUMBER()
  SELECT CustomerID, CustomerName, PostalCode, LastUpdated
  FROM CUSTOMERBASE
  WHERE X = 1
        Having Max(LastUpdated) = 1
        WHERE LastUpdated = Max(LastUpdated)
        WHERE X = 1
```

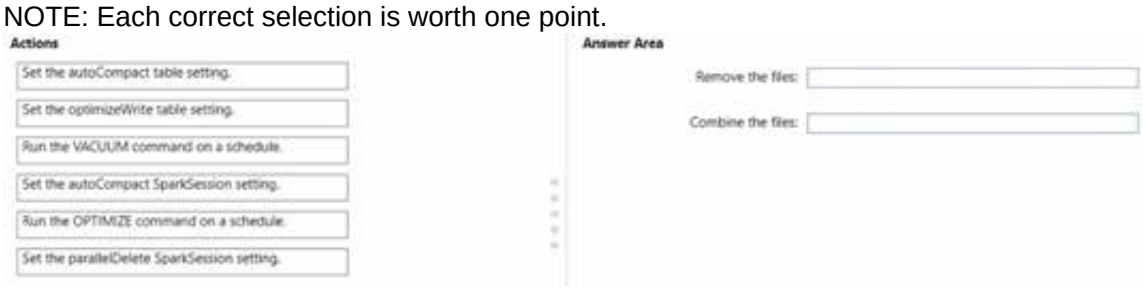
A. Mastered

B. Not Mastered

Answer: A

Explanation:
? 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

NEW QUESTION 9
DRAG DROP - (Topic 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.



A. Mastered
B. Not Mastered

Answer: A

Explanation:
? 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.

NEW QUESTION 10
- (Topic 2)
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.

Answer: B

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.

NEW QUESTION 10
- (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 12
- (Topic 2)
You have a Fabric tenant that contains a Microsoft Power Bl report. You are exploring a new semantic model.
You need to display the following column statistics:
• Count
• Average
• Null count
• Distinct count
• Standard deviation
Which Power Query function should you run?

A. Tabl
B. FuzzyGroup
C. Table.Profile
D. Table.View
E. Table.Schema

Answer: B

Explanation:
The Table.Profile function in Power Query is used to generate column statistics such as count, average, null count, distinct count, and standard deviation. You can use this function as follows:
? Invoke the Power Query Editor.
? Apply the Table.Profile function to your table.
? The result will be a table where each row represents a column from the original table, and each column in the result represents a different statistic such as those listed in the requirement.
References: The use of Table.Profile is part of Power Query M function documentation where it explains how to gather column statistics for a given table.

NEW QUESTION 13
- (Topic 2)
You have a Fabric tenant that contains a complex semantic model. The model is based on a star schema and contains many tables, including a fact table named Sales. You need to create a diagram of the model. The diagram must contain only the Sales table and related tables. What should you use from Microsoft Power Bl Desktop?

A. data categories
B. Data view
C. Model view
D. DAX query view

Answer: C

Explanation:
To create a diagram that contains only the Sales table and related tables, you should use the Model view (C) in Microsoft Power BI Desktop. This view allows you to visualize and manage the relationships between tables within your semantic model. References = Microsoft Power BI Desktop documentation outlines the functionalities available in Model view for managing semantic models.

NEW QUESTION 16
- (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 21
- (Topic 2)
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

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.

NEW QUESTION 23
- (Topic 2)
You have a Fabric tenant that contains a semantic model.
You need to prevent report creators from populating visuals by using implicit measures. What are two tools that you can use to achieve the goal? Each correct answer presents a complete solution.
NOTE: Each correct answer is worth one point.

A. Microsoft Power BI Desktop
B. Tabular Editor
C. Microsoft SQL Server Management Studio (SSMS)
D. DAX Studio

Answer: AB

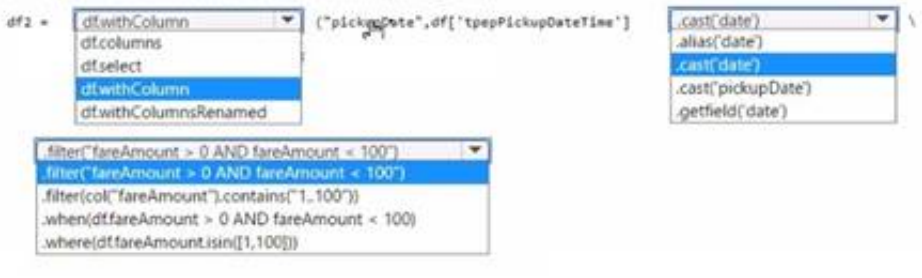
Explanation:
Microsoft Power BI Desktop (A) and Tabular Editor (B) are the tools you can use to prevent report creators from using implicit measures. In Power BI Desktop, you can define explicit measures which can be used in visuals. Tabular Editor allows for advanced model editing, where you can enforce the use of explicit measures. References = Guidance on using explicit measures and preventing implicit measures in reports can be found in the Power BI and Tabular Editor official documentation.

NEW QUESTION 28
HOTSPOT - (Topic 2)
You have a Fabric tenant that contains a lakehouse named Lakehouse1. Lakehouse1 contains a table named Nyctaxi_raw. Nyctaxi_raw contains the following columns.

Name	Data type
pickupDateTime	Timestamp
passengerCount	Integer
fareAmount	Double
paymentType	String
tipAmount	Double

You create a Fabric notebook and attach it to lakehouse1.
You need to use PySpark code to transform the data. The solution must meet the following requirements:
• Add a column named pickupDate that will contain only the date portion of pickupDateTime.
• Filter the DataFrame to include only rows where fareAmount is a positive number that is less than 100.
How should you complete the code? 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:
? Add the pickupDate column: .withColumn("pickupDate", df["pickupDateTime"].cast("date"))
? Filter the DataFrame: .filter("fareAmount > 0 AND fareAmount < 100")
In PySpark, you can add a new column to a DataFrame using the .withColumn method, where the first argument is the new column name and the second argument is the expression to generate the content of the new column. Here, we use the .cast("date") function to extract only the date part from a timestamp. To filter the DataFrame, you use the .filter method with a condition that selects rows where fareAmount is greater than 0 and less than 100, thus ensuring only positive values less than 100 are included.

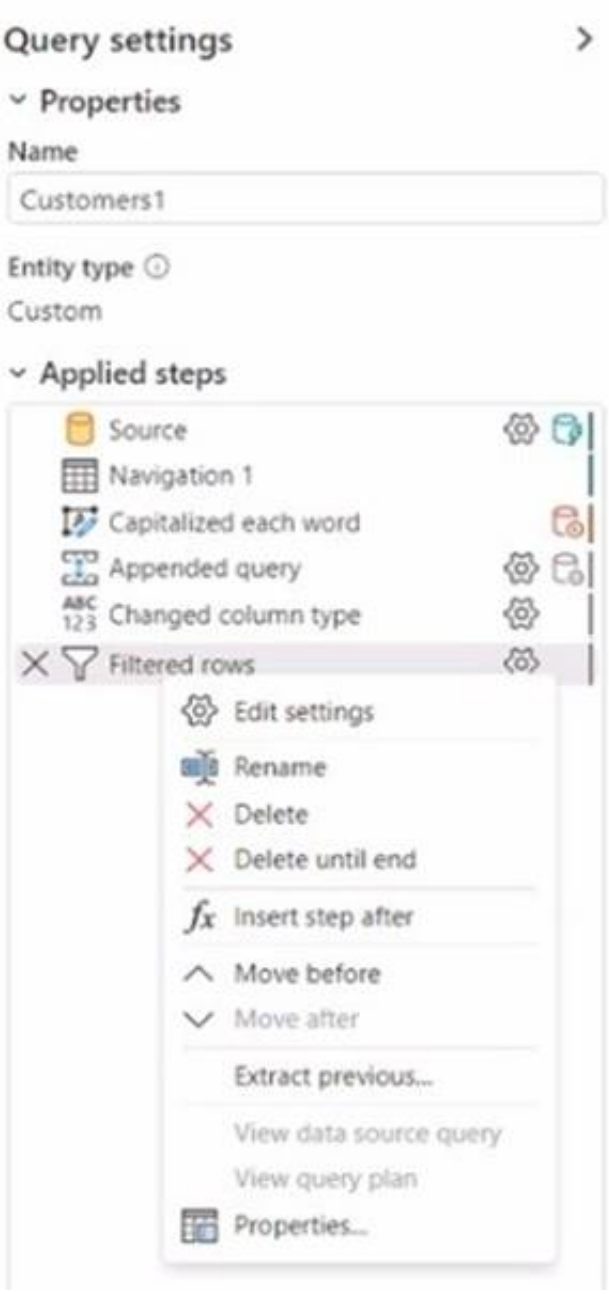
NEW QUESTION 30
- (Topic 2)
You have a Fabric tenant that contains a semantic model. The model uses Direct Lake mode.
You suspect that some DAX queries load unnecessary columns into memory. You need to identify the frequently used columns that are loaded into memory.
What are two ways to achieve the goal? Each correct answer presents a complete solution. NOTE: Each correct answer is worth one point.

A. Use the Analyze in Excel feature.
B. Use the Vertipaq Analyzer tool.
C. Query the $system.discovered_STORAGE_TABLE_COLUMN-iN_SEGMeNTS dynamic management view (DMV).
D. Query the discover_hehory6Rant dynamic management view (DMV).

Answer: BC

Explanation:
The Vertipaq Analyzer tool (B) and querying the $system.discovered_STORAGE_TABLE_COLUMNS_IN_SEGMENTS dynamic management view (DMV) (C) can help identify which columns are frequently loaded into memory. Both methods provide insights into the storage and retrieval aspects of the semantic model. References = The Power BI documentation on Vertipaq Analyzer and DMV queries offers detailed guidance on how to use these tools for performance analysis.

NEW QUESTION 33
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 35
- (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 .sumary ()
Does this meet the goal?

A. Yes

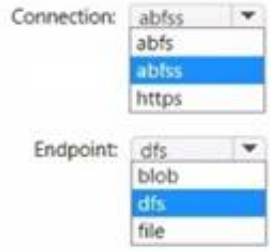
B. No

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.

NEW QUESTION 39
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.

Answer Area



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 43
- (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 44
- (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 48
- (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 50
- (Topic 2)
You have a Fabric tenant that contains a workspace named Workspace^ Workspacel is assigned to a Fabric capacity.
You need to recommend a solution to provide users with the ability to create and publish custom Direct Lake semantic models by using external tools. The solution must follow the principle of least privilege.
Which three actions in the Fabric Admin portal should you include in the recommendation? Each correct answer presents part of the solution.
NOTE: Each correct answer is worth one point.

A. From the Tenant settings, set Allow XMLA Endpoints and Analyze in Excel with on- premises datasets to Enabled
B. From the Tenant settings, set Allow Azure Active Directory guest users to access Microsoft Fabric to Enabled
C. From the Tenant settings, select Users can edit data models in the Power Bl service.
D. From the Capacity settings, set XMLA Endpoint to Read Write
E. From the Tenant settings, set Users can create Fabric items to Enabled
F. From the Tenant settings, enable Publish to Web

Answer: ACD

Explanation:
For users to create and publish custom Direct Lake semantic models using external tools, following the principle of least privilege, the actions to be included are enabling XMLA Endpoints (A), editing data models in Power BI service (C), and setting XMLA Endpoint to Read-Write in the capacity settings (D). References = More information can be found in the Admin portal of the Power BI service documentation, detailing tenant and capacity settings.

NEW QUESTION 54
- (Topic 2)
You are the administrator of a Fabric workspace that contains a lakehouse named Lakehouse1. Lakehouse1 contains the following tables:
• Table1: A Delta table created by using a shortcut
• Table2: An external table created by using Spark
• Table3: A managed table
You plan to connect to Lakehouse1 by using its SQL endpoint. What will you be able to do after connecting to Lakehouse1?

A. ReadTable3.
B. Update the data Table3.
C. ReadTable2.
D. Update the data in Table1.

Answer: D

NEW QUESTION 58
- (Topic 2)

You have a Fabric tenant that contains a warehouse.
You are designing a star schema model that will contain a customer dimension. The customer dimension table will be a Type 2 slowly changing dimension (SCD).
You need to recommend which columns to add to the table. The columns must NOT already exist in the source.
Which three types of columns should you recommend? Each correct answer presents part of the solution.
NOTE: Each correct answer is worth one point.

A. an effective end date and time
B. a foreign key
C. a surrogate key
D. a natural key
E. an effective start date and time

Answer: ACE

Explanation:
For a Type 2 slowly changing dimension (SCD), you typically need to add the following types of columns that do not exist in the source system:
? An effective start date and time (E): This column records the date and time from which the data in the row is effective.
? An effective end date and time (A): This column indicates until when the data in the row was effective. It allows you to keep historical records for changes over time.
? A surrogate key (C): A surrogate key is a unique identifier for each row in a table, which is necessary for Type 2 SCDs to differentiate between historical and current records.
References: Best practices for designing slowly changing dimensions in data warehousing solutions, which include Type 2 SCDs, are commonly discussed in data warehousing and business intelligence literature and would be part of the modeling guidance in a Fabric tenant's documentation.

NEW QUESTION 61
- (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?

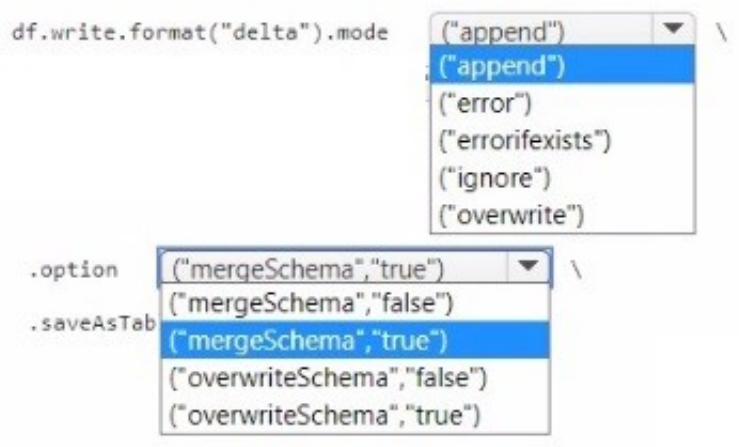
A. PBIP
B. PBIX
C. PBIT
D. PBIDS

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 66
HOTSPOT - (Topic 2)
You have a Fabric tenant that contains lakehouse named Lakehousel. Lakehousel contains a Delta table with eight columns. You receive new data that contains the same eight columns and two additional columns.
You create a Spark DataFrame and assign the DataFrame to a variable named df. The DataFrame contains the new data. You need to add the new data to the Delta table to meet the following requirements:
• Keep all the existing rows.
• Ensure that all the new data is added to the table.
How should you complete the code? To answer, select the appropriate options in the answer area.

Answer Area



A. Mastered
B. Not Mastered

Answer: A

Explanation:
o add new data to the Delta table while meeting the specified requirements:
? You should use the append mode to ensure that all new data is added to the table without affecting the existing rows.
? You should set the mergeSchema option to true to allow the schema of the Delta table to be updated with the new columns found in the DataFrame.
The completed code would look like this:
df.write.format("delta").mode("append") option("mergeSchema", "true") saveAsTable("Lakehouse1.TableName")

NEW QUESTION 67
- (Topic 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. Tabl
B. MaxN
C. Table.Max
D. Table.Range
E. Table.Profile

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.

NEW QUESTION 68
- (Topic 2)
You have a Fabric tenant.
You are creating a Fabric Data Factory pipeline.
You have a stored procedure that returns the number of active customers and their average sales for the current month.
You need to add an activity that will execute the stored procedure in a warehouse. The returned values must be available to the downstream activities of the pipeline.
Which type of activity should you add?

A. Stored procedure
B. Get metadata
C. Lookup
D. Copy data

Answer: C

Explanation:
In a Fabric Data Factory pipeline, to execute a stored procedure and make the returned values available for downstream activities, the Lookup activity is used. This activity can retrieve a dataset from a data store and pass it on for further processing. Here's how you would use the Lookup activity in this context:
? Add a Lookup activity to your pipeline.
? Configure the Lookup activity to use the stored procedure by providing the necessary SQL statement or stored procedure name.
? In the settings, specify that the activity should use the stored procedure mode.
? Once the stored procedure executes, the Lookup activity will capture the results and make them available in the pipeline's memory.
? Downstream activities can then reference the output of the Lookup activity. References: The functionality and use of Lookup activity within Azure Data Factory is documented in Microsoft's official documentation for Azure Data Factory, under the section
for pipeline activities.

NEW QUESTION 72
- (Topic 2)
You have a Fabric tenant that contains a Microsoft Power Bl report named Report 1. Report1 includes a Python visual. Data displayed by the visual is grouped automatically and duplicate rows are NOT displayed. You need all rows to appear in the visual. What should you do?

A. Reference the columns in the Python code by index.
B. Modify the Sort Column By property for all columns.
C. Add a unique field to each row.
D. Modify the Summarize By property for all columns.

Answer: C

Explanation:
To ensure all rows appear in the Python visual within a Power BI report, option C, adding a unique field to each row, is the correct solution. This will prevent automatic grouping by unique values and allow for all instances of data to be represented in the visual. References = For more on Power BI Python visuals and how they handle data, please refer to the Power BI documentation.

NEW QUESTION 73
HOTSPOT - (Topic 2)
You have a Fabric tenant.
You plan to create a Fabric notebook that will use Spark DataFrames to generate Microsoft Power Bl 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.

Answer 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.	○	○

A. Mastered
B. Not Mastered

Answer: A

Explanation:
? 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

NEW QUESTION 75
- (Topic 2)
You need to create a data loading pattern for a Type 1 slowly changing dimension (SCD).
Which two actions should you include in the process? Each correct answer presents part of the solution.
NOTE: Each correct answer is worth one point.

A. Update rows when the non-key attributes have changed.
B. Insert new rows when the natural key exists in the dimension table, and the non-key attribute values have changed.
C. Update the effective end date of rows when the non-key attribute values have changed.
D. Insert new records when the natural key is a new value in the table.

Answer: AD

Explanation:
For a Type 1 SCD, you should include actions that update rows when non- key attributes have changed (A), and insert new records when the natural key is a new value in the table (D). A Type 1 SCD does not track historical data, so you always overwrite the old data with the new data for a given key. References = Details on Type 1 slowly changing dimension patterns can be found in data warehousing literature and Microsoft's official documentation.

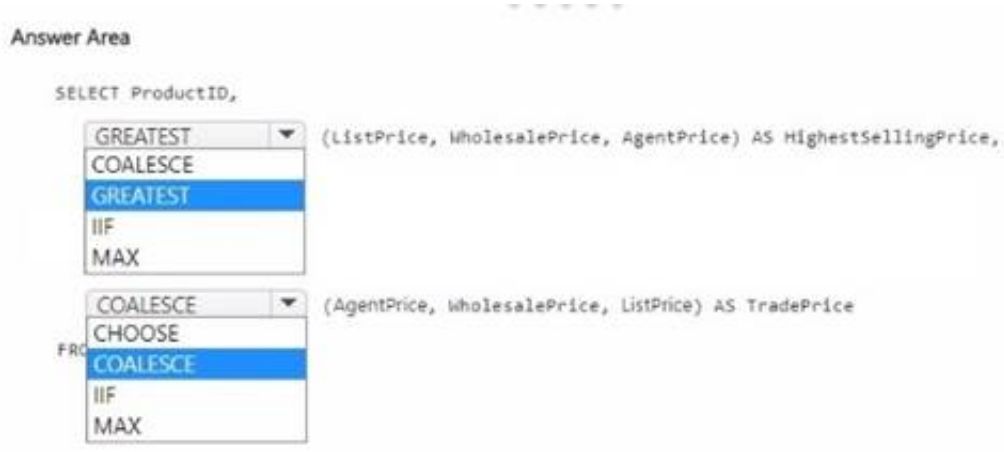
NEW QUESTION 76
HOTSPOT - (Topic 2)
You have a Fabric warehouse that contains a table named Sales.Products. Sales.Products contains the following columns.

Name	Data type	Nullable
ProductID	Integer	No
ProductName	Varchar(30)	No
ListPrice	Decimal(18, 2)	No
WholesalePrice	Decimal(18, 2)	Yes
AgentPrice	Decimal(18, 2)	Yes

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

Name	Description
ProductID	Return the ProductID value
HighestSellingPrice	Returns the highest value from ListPrice, WholesalePrice, and AgentPrice
TradePrice	Returns the AgentPrice value if present, otherwise returns the WholesalePrice value if present, otherwise returns the ListPrice value

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



A. Mastered
B. Not Mastered

Answer: A

Explanation:
? For the HighestSellingPrice, you should use the GREATEST function to find the highest value from the given price columns. However, T-SQL does not have a GREATEST function as found in some other SQL dialects, so you would typically use a CASE statement or an IIF statement with nested MAX functions. Since neither of those are provided in the options, you should select MAX as a placeholder to indicate the function that would be used to find the highest value if combining multiple MAX functions or a similar logic was available.
? For the TradePrice, you should use the COALESCE function, which returns the first non-null value in a list. The COALESCE function is the correct choice as it will return AgentPrice if it's not null; if AgentPrice is null, it will check WholesalePrice, and if that is also null, it will return ListPrice.
The complete code with the correct SQL functions would look like this:
SELECT ProductID,
MAX(ListPrice, WholesalePrice, AgentPrice) AS HighestSellingPrice, -- MAX is used as a placeholder
COALESCE(AgentPrice, WholesalePrice, ListPrice) AS TradePrice FROM Sales.Products
Select MAX for HighestSellingPrice and COALESCE for TradePrice in the answer area.

NEW QUESTION 77
.....

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