Microsoft

Exam Questions DP-600

Implementing Analytics Solutions Using Microsoft Fabric

NEW QUESTION 1
- (Topic 1)
You need to recommend a solution to prepare the tenant for the PoC.
Which two actions should you recommend performing from the Fabric Admin portal? Each correct answer presents part of the solution.
NOTE: Each correct answer is worth one point.

A. Enable the Users can try Microsoft Fabric paid features option for specific security groups.
B. Enable the Allow Azure Active Directory guest users to access Microsoft Fabric option for specific security groups.
C. Enable the Users can create Fabric items option and exclude specific security groups.
D. Enable the Users can try Microsoft Fabric paid features option for the entire organization.
E. Enable the Users can create Fabric items option for specific security groups.

Answer: AE

Explanation:
The PoC is planned to be completed using a Fabric trial capacity, which implies that users involved in the PoC should be able to try paid features. However, this should be limited to specific security groups involved in the PoC to prevent the entire organization from accessing these features before the trial is proven successful (A). The ability for users to create Fabric items should also be enabled for specific security groups to ensure that only the relevant team members participating in the PoC can create items in the Fabric environment (E).

NEW QUESTION 2
HOTSPOT - (Topic 1)
You need to design a semantic model for the customer satisfaction report.
Which data source authentication method and mode should you use? 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 semantic model design required for the customer satisfaction report, the choices for data source authentication method and mode should be made based on security and performance considerations as per the case study provided.
Authentication method: The data should be accessed securely, and given that row-level security (RLS) is required for users executing T-SQL queries, you should use an authentication method that supports RLS. Service principal authentication is suitable for automated and secure access to the data, especially when the access needs to be controlled programmatically and is not tied to a specific user's credentials.
Mode: The report needs to show data as soon as it is updated in the data store, and it should only contain data from the current and previous year. DirectQuery mode allows for real-time reporting without importing data into the model, thus meeting the need for up-to- date data. It also allows for RLS to be implemented and enforced at the data source level, providing the necessary security measures.
Based on these considerations, the selections should be:
? Authentication method: Service principal authentication
? Mode: DirectQuery

NEW QUESTION 3
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

DataEngineers:	Build Reports on the default dataset ▼
	Build Reports on the default dataset
	Read All Apache Spark
	Read All SQL analytics endpoint data
DataAnalysts:	Read All Apache Spark ▼
	Build Reports on the default dataset
	Read All Apache Spark
	Read All SQL analytics endpoint data
DataScientists:	Read All SQL analytics endpoint data ▼
	Build Reports on the default dataset
	Read All Apache Spark
	Read All SQL analytics endpoint data

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 4
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

```
CREATE        [_____▼]   [dbo].[ProductsWithPricingGroup]

AS

    SELECT ProductId,

        ProductName,

        ProductCategory,

    ListPrice,

    [_____▼]

        WHEN ListPrice <= 50 THEN 'low'

    [_____▼]


    END AS PricingGroup

    FROM dbo.Products
```

Answer Area

```
CREATE   [VIEW          ▼]   [dbo].[ProductsWithPricingGroup]
          FUNCTION
AS        PROCEDURE
   SELECT L TABLE
          VIEW
      Pro
      ProductCategory,

   ListPrice,

   [CASE          ▼]
    CASE              0 THEN 'low'
    COALESCE
    IIF
    SET

   [WHEN (ListPrice >= 50 AND ListPrice < 1000 ) THEN 'medium'  ▼]
    WHEN (ListPrice > 50 AND ListPrice <= 1000 ) THEN 'medium'
    WHEN (ListPrice >= 50 AND ListPrice < 1000 ) THEN 'medium'
    WHEN ListPrice BETWEEN 50 AND 1000 ) THEN 'medium'

      END AS PricingGroup

   FROM dbo.Products
```

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 5
- (Topic 1)
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

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.

NEW QUESTION 6
- (Topic 2)
You have a Fabric tenant that uses a Microsoft tower BI 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 7
- (Topic 2)
You have a Fabric tenant that contains a warehouse. The warehouse uses row-level security (RLS). You create a Direct Lake semantic model that uses the Delta tables and RLS of the warehouse. When users interact with a report built from the model, which mode will be used by the DAX queries?

A. DirectQuery
B. Dual
C. Direct Lake
D. Import

Answer: A

Explanation:
When users interact with a report built from a Direct Lake semantic model that uses row-level security (RLS), the DAX queries will operate in DirectQuery mode (A). This is because the model directly queries the underlying data source without importing data into Power BI. References = The Power BI documentation on DirectQuery provides detailed explanations of how RLS and DAX queries function in this mode.

NEW QUESTION 8
DRAG DROP - (Topic 2)
You have a Fabric tenant that contains a lakehouse named Lakehouse1
Readings from 100 loT 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.

Actions	Answer Area
Set the autoCompact table setting.	Remove the files:
Set the optimizeWrite table setting.	Combine the files:
Run the VACUUM command on a schedule.	
Set the autoCompact SparkSession setting.	
Run the OPTIMIZE command on a schedule.	
Set the parallelDelete SparkSession setting.	

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 9
- (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
HOTSPOT - (Topic 2)
You have the source data model shown in the following exhibit.
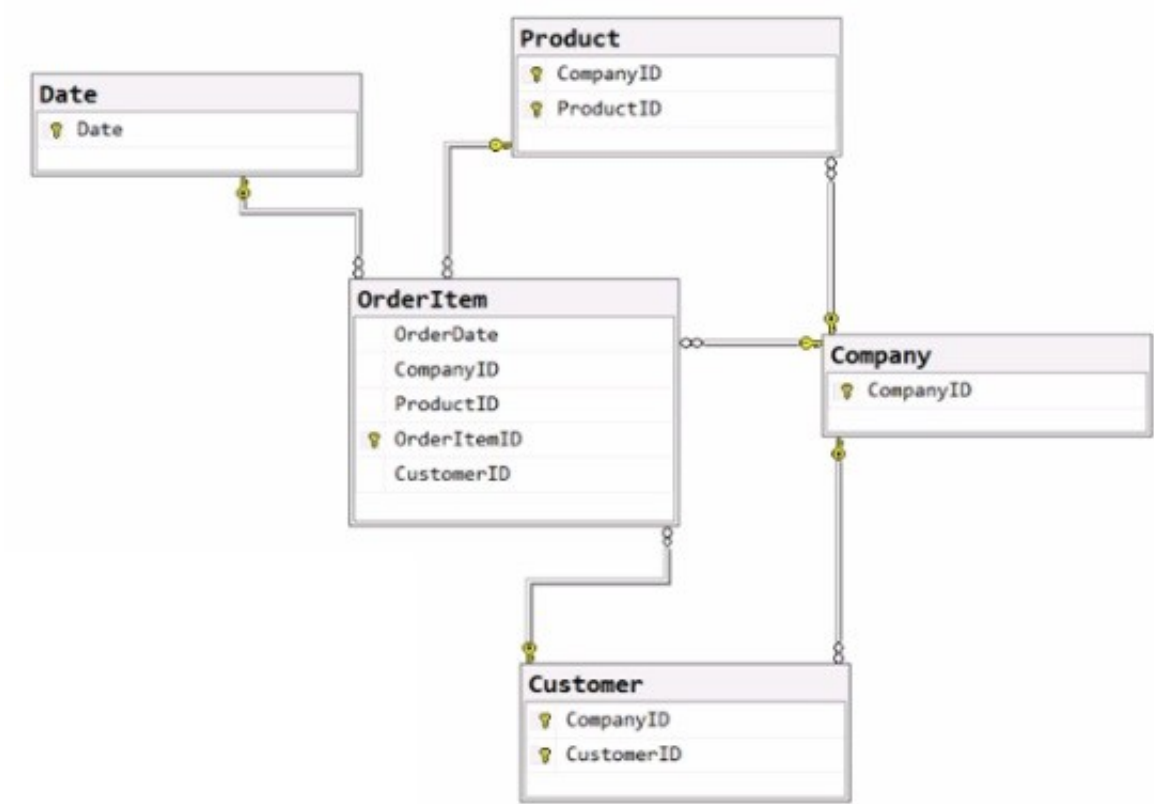
The primary keys of the tables are indicated by a key symbol beside the columns involved in each key.
You need to create a dimensional data model that will enable the analysis of order items by date, product, and customer.
What should you include in the solution? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.

Answer Area

The relationship between OrderItem and Product must be based on: Both the CompanyID and the ProductID columns ▼
The ProductID column
Both the CompanyID and the ProductID columns
A new key that combines the CompanyID and ProductID columns

The Company entity must be: Denormalized into the Customer and Product entities ▼
Omitted
Denormalized into the Product entity only
Denormalized into the Customer and Product entities

A. Mastered
B. Not Mastered

Answer: A

Explanation:
? 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

NEW QUESTION 10
- (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 12
- (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 15
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

```
df = spark.read.format("delta").load("Tables/nyctaxi_raw")
df2 = [ df.withColumn        ▼]  ("pickupDate",df['tpepPickupDateTime']   [.cast('date')        ▼]  \
         df.columns                                                        .alias('date')
         df.select                                                         .cast('date')
         df.withColumn                                                     .cast('pickupDate')
         df.withColumnsRenamed                                             .getfield('date')

         [.filter("fareAmount > 0 AND fareAmount < 100")        ▼]
          .filter("fareAmount > 0 AND fareAmount < 100")
          .filter(col("fareAmount").contains("1..100"))
          .when(df.fareAmount > 0 AND fareAmount < 100)
          .where(df.fareAmount.isin([1,100]))
```

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 19
DRAG DROP - (Topic 2)
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.

A. Mastered
B. Not Mastered

Answer: A

Explanation:
 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).

NEW QUESTION 21
- (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 24
- (Topic 2)
You have a Microsoft Power BI report named Report1 that uses a Fabric semantic model. Users discover that Report1 renders slowly.
You open Performance analyzer and identify that a visual named Orders By Date is the slowest to render. The duration breakdown for Orders By Date is shown in the following table.

Name	Duration (ms)
DAX query	27
Visual display	39
Other	1047

What will provide the greatest reduction in the rendering duration of Report1?

A. Change the visual type of Orders By Dale.
B. Enable automatic page refresh.
C. Optimize the DAX query of Orders By Date by using DAX Studio.
D. Reduce the number of visuals in Report1.

Answer: C

Explanation:
Based on the duration breakdown provided, the major contributor to the rendering duration is categorized as "Other," which is significantly higher than DAX Query and Visual display times. This suggests that the issue is less likely with the DAX calculation or visual rendering times and more likely related to model performance or the complexity of the visual. However, of the options provided, optimizing the DAX query can be a crucial step, even if "Other" factors are dominant. Using DAX Studio, you can analyze and optimize the DAX queries that power your visuals for performance improvements. Here's how you might proceed:
? Open DAX Studio and connect it to your Power BI report.
? Capture the DAX query generated by the Orders By Date visual.
? Use the Performance Analyzer feature within DAX Studio to analyze the query.
? Look for inefficiencies or long-running operations.
? Optimize the DAX query by simplifying measures, removing unnecessary calculations, or improving iterator functions.
? Test the optimized query to ensure it reduces the overall duration.
References: The use of DAX Studio for query optimization is a common best practice for improving Power BI report performance as outlined in the Power BI documentation.

NEW QUESTION 27
- (Topic 2)
You have a Fabric tenant that contains a machine learning model registered in a Fabric workspace. You need to use the model to generate predictions by using the predict function in a fabric notebook. Which two languages can you use to perform model scoring? Each correct answer presents a complete solution. NOTE: Each correct answer is worth one point.

A. T-SQL
B. DAX EC.
C. Spark SQL
D. PySpark

Answer: CD

Explanation:
 The two languages you can use to perform model scoring in a Fabric notebook using the predict function are Spark SQL (option C) and PySpark (option D). These are both part of the Apache Spark ecosystem and are supported for machine learning tasks in a Fabric environment. References = You can find more information about model scoring and supported languages in the context of Fabric notebooks in the official documentation on Azure Synapse Analytics.

NEW QUESTION 32
HOTSPOT - (Topic 2)
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")
)
```

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

A. Mastered
B. Not Mastered

Answer: A

Explanation:
? The Spark engine will read only the 'SalesOrderNumber', 'OrderDate', 'CustomerName', 'UnitPrice' columns from Sales_raw.csv. - Yes

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

NEW QUESTION 33
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 36
- (Topic 2)
You have a Fabric workspace named Workspace 1 that contains a dataflow named Dataflow1. Dataflow! has a query that returns 2.000 rows. You view the query in Power Query as shown in the following exhibit.
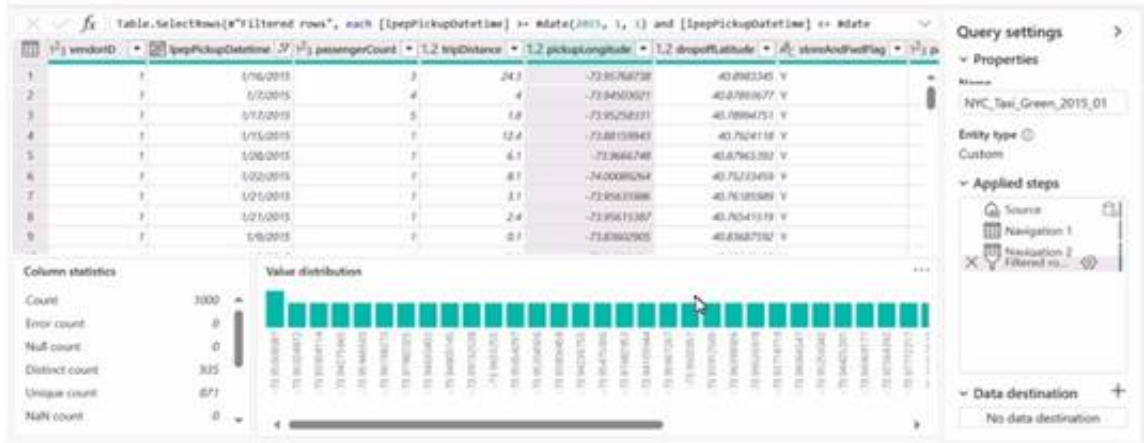
What can you identify about the pickupLongitude column?

A. The column has duplicate values.
B. All the table rows are profiled.
C. The column has missing values.
D. There are 935 values that occur only once.

Answer: B

Explanation:
The pickupLongitude column has duplicate values. This can be inferred because the 'Distinct count' is 935 while the 'Count' is 1000, indicating that there are repeated values within the column. References = Microsoft Power BI documentation on data profiling could provide further insights into understanding and interpreting column statistics like these.

NEW QUESTION 40
- (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 BI 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 43
- (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 44
- (Topic 2)
You have a Fabric tenant that contains a lakehouse named Lakehouse1.
You need to prevent new tables added to Lakehouse1 from being added automatically to the default semantic model of the lakehouse.
What should you configure? (5)

A. the semantic model settings
B. the Lakehouse1 settings
C. the workspace settings
D. the SQL analytics endpoint settings

Answer: A

Explanation:
To prevent new tables added to Lakehouse1 from being automatically added to the default semantic model, you should configure the semantic model settings. There should be an option within the settings of the semantic model to include or exclude new tables by default. By adjusting these settings, you can control the automatic inclusion of new tables.
References: The management of semantic models and their settings would be covered under the documentation for the semantic layer or modeling features of the Fabric tenant's lakehouse solution.

NEW QUESTION 47
- (Topic 2)
You are creating a semantic model in Microsoft Power BI 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 50
- (Topic 2)
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

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.
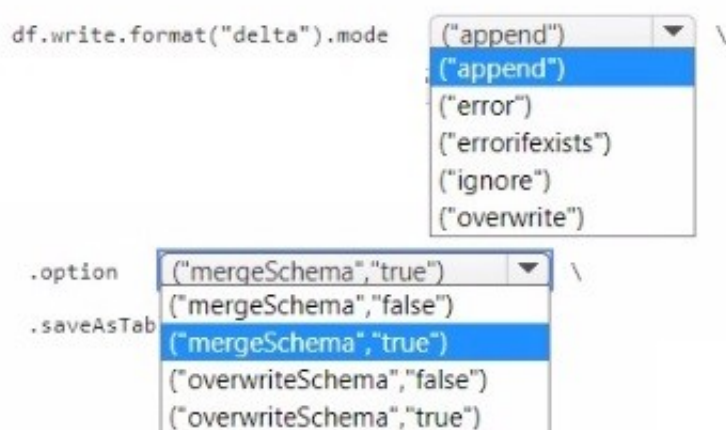
NEW QUESTION 53
HOTSPOT - (Topic 2)
You have a Fabric tenant that contains lakehouse named Lakehousel. Lakehousel contains a Delta table with eight columns. You receive new data that contains the same eight columns and two additional columns.
You create a Spark DataFrame and assign the DataFrame to a 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

```
df.write.format("delta").mode    ("append")            \
                                  ("append")
                                  ("error")
                                  ("errorifexists")
                                  ("ignore")
                                  ("overwrite")

              .option    ("mergeSchema","true")         \
              .saveAsTab  ("mergeSchema","false")
                          ("mergeSchema","true")
                          ("overwriteSchema","false")
                          ("overwriteSchema","true")
```

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 55
- (Topic 2)
You have a Fabric tenant that contains a lakehouse named lakehouse1. Lakehouse1 contains a table named Table1.
You are creating a new data pipeline.
You plan to copy external data to Table1. The schema of the external data changes regularly.
You need the copy operation to meet the following requirements:
• Replace Table1 with the schema of the external data.
• Replace all the data in Table1 with the rows in the external data.
You add a Copy data activity to the pipeline. What should you do for the Copy data activity?

A. From the Source tab, add additional columns.
B. From the Destination tab, set Table action to Overwrite.
C. From the Settings tab, select Enable staging.
D. From the Source tab, select Enable partition discovery
E. From the Source tab, select Recursively

Answer: B

Explanation:
For the Copy data activity, from the Destination tab, setting Table action to Overwrite (B) will ensure that Table1 is replaced with the schema and rows of the external data, meeting the requirements of replacing both the schema and data of the destination table. References = Information about Copy data activity and table actions in Azure Data Factory, which can be applied to data pipelines in Fabric, is available in the Azure Data Factory documentation.

NEW QUESTION 59
- (Topic 2)
You have a Fabric tenant tha1 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:
EXPLAIN TABLE customer Does this meet the goal?

A. Yes
B. No

Answer: B

Explanation:

No, the EXPLAIN TABLE statement does not identify whether maintenance tasks were performed on a table. It shows the execution plan for a query. References = The usage and output of the EXPLAIN command can be found in the Spark SQL documentation.

NEW QUESTION 60
HOTSPOT - (Topic 2)
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.

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 62
- (Topic 2)
You have a Fabric workspace that contains a DirectQuery semantic model. The model queries a data source that has 500 million rows.
You have a Microsoft Power BI report named Report1 that uses the model. Report! contains visuals on multiple pages.
You need to reduce the query execution time for the visuals on all the pages.
What are two features that you can use? Each correct answer presents a complete solution.
NOTE: Each correct answer is worth one point.

A. user-defined aggregations
B. automatic aggregation
C. query caching
D. OneLake integration

Answer: AB

Explanation:
User-defined aggregations (A) and query caching (C) are two features that can help reduce query execution time. User-defined aggregations allow precalculation of large datasets, and query caching stores the results of queries temporarily to speed up future queries. References = Microsoft Power BI documentation on performance optimization offers in-depth knowledge on these features.

NEW QUESTION 64
......

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

100% Actual & Verified — Instant Download, Please Click
[Order The DP-600 Practice Test Here](https://www.certshared.com/exam/DP-600/)