

Microsoft

Exam Questions DP-600

Implementing Analytics Solutions Using Microsoft Fabric



NEW QUESTION 1

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.

Answer Area

Authentication method:

Service principal authentication

Basic authentication

Service principal authentication

Single sign-on (SSO) authentication

Mode:

DirectQuery

Direct Lake

DirectQuery

Import

- 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 2

- (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 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 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

Rolling 12 Overall Satisfaction =

VAR NumberOfMonths = 12

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

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

VAR Result =

CALCULATE (

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

)

RETURN

Result

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 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

GRANT

ALTER

ALTER

CONNECT

EXECUTE

ON

SCHEMA :: schemaA

DATABASE :: schemaA

OBJECT :: schemaA

SCHEMA :: schemaA

TO User1;

- 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 BI 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

DRAG DROP - (Topic 2)

You have a Fabric tenant that contains a semantic model. The model contains data about retail stores.

You need to write a DAX query that will be executed by using the XMLA endpoint The query must return a table of stores that have opened since December 1,2023.

How should you complete the DAX expression? 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.

Values

DEFINE

EVALUATE

FILTER

SUMMARIZE

TABLE

Answer Area

VAR _SalesSince =

DATE (2023, 12, 01)

FILTER (

{ Store, Store[Name], Store[OpenDate] },

Store[OpenDate] >= _SalesSince

)

A. Mastered
 B. Not Mastered

Answer: A

Explanation:

The correct order for the DAX expression would be:

? DEFINE VAR _SalesSince = DATE (2023, 12, 01)

? EVALUATE

? FILTER (

? SUMMARIZE (Store, Store[Name], Store[OpenDate]),

? Store[OpenDate] >= _SalesSince)

In this DAX query, you're defining a variable _SalesSince to hold the date from which you want to filter the stores. EVALUATE starts the definition of the query. The FILTER function is used to return a table that filters another table or expression. SUMMARIZE creates a summary table for the stores, including the Store[Name] and Store[OpenDate] columns, and the filter expression Store[OpenDate] >= _SalesSince ensures only stores opened on or after December 1, 2023, are included in the results.

Passing Certification Exams Made Easy

visit - <https://www.surepassexam.com>

References =
? DAX FILTER Function
? DAX SUMMARIZE Function

NEW QUESTION 8

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.



The screenshot shows a drag-and-drop interface. On the left, under the heading "Actions", there is a list of seven actions in a scrollable container: "Set the autoCompact table setting.", "Set the optimizeWrite table setting.", "Run the VACUUM command on a schedule.", "Set the autoCompact SparkSession setting.", "Run the OPTIMIZE command on a schedule.", and "Set the parallelDelete SparkSession setting.". On the right, under the heading "Answer Area", there are two requirements: "Remove the files:" followed by a text input field, and "Combine the files:" followed by a text input field.

- 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

- (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

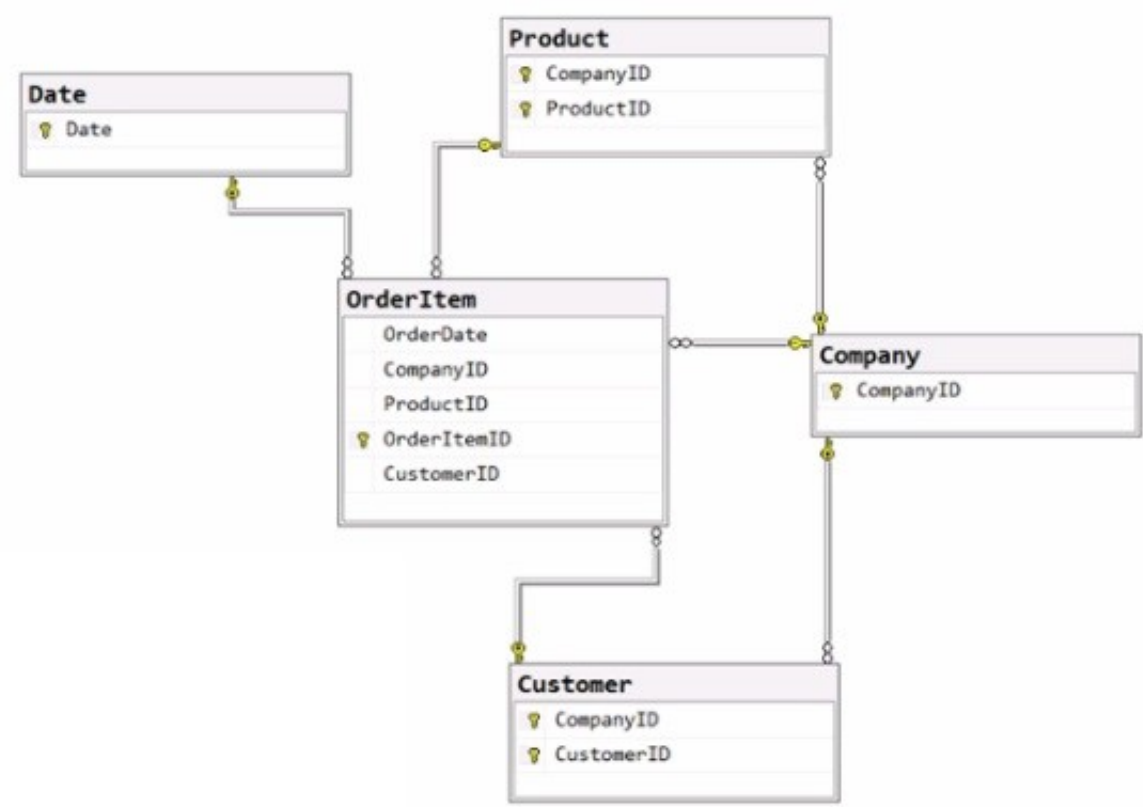
Explanation:

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 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.

The relationship between OrderItem and Product must be based on:

The Company entity must be:

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

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 11

- (Topic 2)
 You have a Fabric tenant that contains a Microsoft Power BI 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 14

- (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 17

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

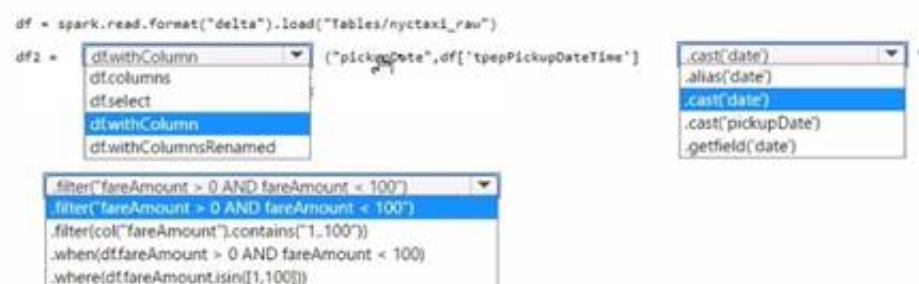
You create a Fabric notebook and attach it to lakehouse1.

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 18

- (Topic 2)

You have a Fabric tenant that contains a semantic model. The model contains 15 tables.

You need to programmatically change each column that ends in the word Key to meet the following requirements:

- Hide the column.
- Set Nullable to False.
- Set Summarize By to None
- Set Available in MDX to False.
- Mark the column as a key column. What should you use?

- A. Microsoft Power BI Desktop
- B. Tabular Editor
- C. ALM Toolkit
- D. DAX Studio

Answer: B

Explanation:

Tabular Editor is an advanced tool for editing Tabular models outside of Power BI Desktop that allows you to script out changes and apply them across multiple columns or tables. To accomplish the task programmatically, you would:

? Open the model in Tabular Editor.

? Create an Advanced Script using C# to iterate over all tables and their respective columns.

? Within the script, check if the column name ends with 'Key'.

? For columns that meet the condition, set the properties accordingly: IsHidden = true, IsNullable = false, SummarizeBy = None, IsAvailableInMDX = false.

? Additionally, mark the column as a key column.

? Save the changes and deploy them back to the Fabric tenant.

References: The ability to batch-edit properties using scripts in Tabular Editor is well- documented in the tool's official documentation and user community resources.

NEW QUESTION 23

- (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.summary()`

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 28

DRAG DROP - (Topic 2)

You create a semantic model by using Microsoft Power BI Desktop. The model contains one security role named SalesRegionManager and the following tables:

- Sales
- SalesRegion
- Sales Address

You need to modify the model to ensure that users assigned the SalesRegionManager role cannot see a column named Address in Sales Address.

Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

Actions

⋮ Open the model in Power BI Desktop.

⋮ Set Object Level Security to **Default** for SalesRegionManager.

⋮ Set the Hidden property to **True**.

⋮ Open the model in Tabular Editor.

⋮ Select the **Address** column in SalesAddress.

⋮ Set Object Level Security to **None** for SalesRegionManager.

Answer Area

A. Mastered

B. Not Mastered

Answer: A

Explanation:

To ensure that users assigned the SalesRegionManager role cannot see the Address column in the SalesAddress table, follow these steps in sequence:

? Open the model in Tabular Editor.

? Select the Address column in SalesAddress.

? Set Object Level Security to None for SalesRegionManager.

NEW QUESTION 30

- (Topic 2)

You have a Fabric workspace named Workspace1 that contains a data flow named Dataflow1. Dataflow1 contains a query that returns the data shown in the following exhibit.



You need to transform the date columns into attribute-value pairs, where columns become rows.
 You select the VendorID column.
 Which transformation should you select from the context menu of the VendorID column?

- A. Group by
- B. Unpivot columns
- C. Unpivot other columns
- D. Split column
- E. Remove other columns

Answer: B

Explanation:

The transformation you should select from the context menu of the VendorID column to transform the date columns into attribute-value pairs, where columns become rows, is Unpivot columns (B). This transformation will turn the selected columns into rows with two new columns, one for the attribute (the original column names) and one for the value (the data from the cells). References = Techniques for unpivoting columns are covered in the Power Query documentation, which explains how to use the transformation in data modeling.

NEW QUESTION 32

- (Topic 2)

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

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

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

- A. Split OrderDateTime into separate date and time columns.
- B. Replace TotalQuantity with a calculated column.
- C. Convert Quantity into the Text data type.
- D. Replace TotalSalesAmount with a measure.

Answer: AD

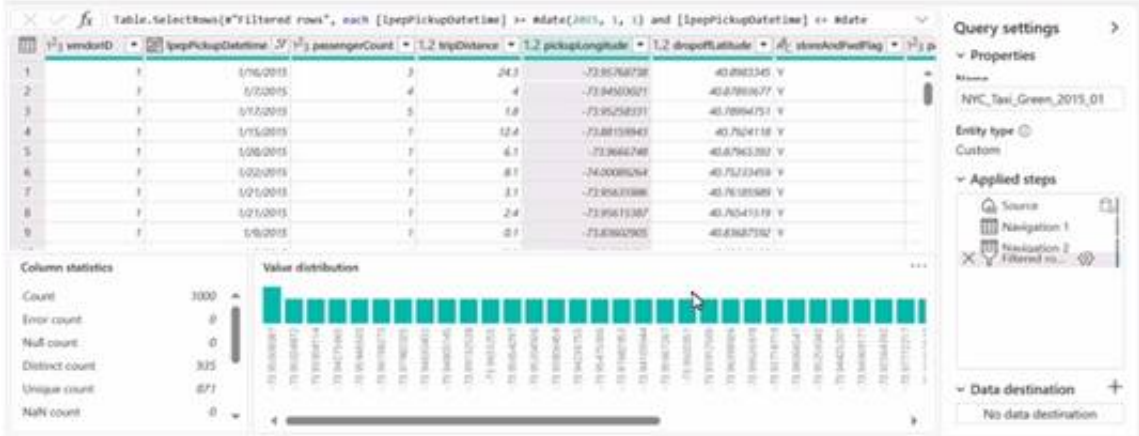
Explanation:

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

NEW QUESTION 35

- (Topic 2)

You have a Fabric workspace named Workspace 1 that contains a dataflow named Dataflow1. Dataflow1 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 39

- (Topic 2)

You have a Microsoft Power BI semantic model that contains measures. The measures use multiple calculate functions and a filter function.

You are evaluating the performance of the measures.

In which use case will replacing the filter function with the keepfilters function reduce execution time?

- A. when the filter function uses a nested calculate function
- B. when the filter function references a column from a single table that uses Import mode
- C. when the filter function references columns from multiple tables
- D. when the filter function references a measure

Answer: A

Explanation:

The KEEPFILTERS function modifies the way filters are applied in calculations done through the CALCULATE function. It can be particularly beneficial to replace the FILTER function with KEEPFILTERS when the filter context is being overridden by nested CALCULATE functions, which may remove filters that are being applied on a column. This can potentially reduce execution time because KEEPFILTERS maintains the existing filter context and allows the nested CALCULATE functions to be evaluated more efficiently. References: This information is based on the DAX reference and performance optimization guidelines in the Microsoft Power BI documentation.

NEW QUESTION 42

- (Topic 2)

You have a Fabric tenant tha1 contains a lakehouse 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:

REFRESH TABLE customer Does this meet the goal?

- A. Yes
- B. No

Answer: B

Explanation:

No, the REFRESH TABLE statement does not provide information on whether maintenance tasks were performed. It only updates the metadata of a table to reflect any changes on the data files. References = The use and effects of the REFRESH TABLE command are explained in the Spark SQL documentation.

NEW QUESTION 46

- (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 52

- (Topic 2)

You have a Microsoft Power BI semantic model.

You need to identify any surrogate key columns in the model that have the Summarize By property set to a value other than to None. The solution must minimize effort.

What should you use?

- A. DAX Formatter in DAX Studio
- B. Model view in Microsoft Power BI Desktop
- C. Model explorer in Microsoft Power BI Desktop
- D. Best Practice Analyzer in Tabular Editor

Answer: D

Explanation:

To identify surrogate key columns with the "Summarize By" property set to a value other than "None," the Best Practice Analyzer in Tabular Editor is the most efficient tool. The Best Practice Analyzer can analyze the entire model and provide a report on all columns that do not meet a specified best practice, such as having the "Summarize By" property set correctly for surrogate key columns. Here's how you would proceed:

? Open your Power BI model in Tabular Editor.

? Go to the Advanced Scripting window.

? Write or use an existing script that checks the "Summarize By" property of each column.

? Execute the script to get a report on the surrogate key columns that do not have their "Summarize By" property set to "None".

? You can then review and adjust the properties of the columns directly within the Tabular Editor.

References: The functionality of the Best Practice Analyzer in Tabular Editor is documented in the community and learning resources for Power BI.

NEW QUESTION 55

HOTSPOT - (Topic 2)

You have a Fabric tenant that contains a lakehouse.

You are using a Fabric notebook to save a large DataFrame by using the following code.

```
df.write.partitionBy("year", "month", "day").mode("overwrite").parquet("Files/SalesOrder")
```

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 results will form a hierarchy of folders for each partition key.	<input type="radio"/>	<input type="radio"/>
The resulting file partitions can be read in parallel across multiple nodes.	<input type="radio"/>	<input type="radio"/>
The resulting file partitions will use file compression.	<input type="radio"/>	<input type="radio"/>

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

? The results will form a hierarchy of folders for each partition key. - Yes

? The resulting file partitions can be read in parallel across multiple nodes. - Yes

? The resulting file partitions will use file compression. - No

Partitioning data by columns such as year, month, and day, as shown in the DataFrame write operation, organizes the output into a directory hierarchy that reflects the partitioning structure. This organization can improve the performance of read operations, as queries that filter by the partitioned columns can scan only the

relevant directories. Moreover, partitioning facilitates parallelism because each partition can be processed independently across different nodes in a distributed system like Spark. However, the code snippet provided does not explicitly specify that file compression should be used, so we cannot assume that the output will be compressed without additional context.

References =

? DataFrame write partitionBy

? Apache Spark optimization with partitioning

NEW QUESTION 59

- (Topic 2)

You have a Fabric tenant that contains a warehouse.

A user discovers that a report that usually takes two minutes to render has been running for 45 minutes and has still not rendered.

You need to identify what is preventing the report query from completing. Which dynamic management view (DMV) should you use?

- A. sys.dm-exec_requests
- B. sys.dm_exec_sessions
- C. sys.dm_exec_connections
- D. sys.dm_pdw_exec_requests

Answer: D

Explanation:

The correct DMV to identify what is preventing the report query from completing is sys.dm_pdw_exec_requests (D). This DMV is specific to Microsoft Analytics Platform System (previously known as SQL Data Warehouse), which is the environment assumed to be used here. It provides information about all queries and load commands currently running or that have recently run. References = You can find more about DMVs in the Microsoft documentation for Analytics Platform System.

NEW QUESTION 60

- (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 62

- (Topic 2)

You have a Fabric tenant that contains a Microsoft Power BI 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 66

- (Topic 2)

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

- Ensure that the developers can deploy items to the workspaces for Development and Test.
- Prevent the developers from deploying items to the workspace for Production.
- Follow the principle of least privilege.

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

- A. Build permission to the production semantic models
- B. Admin access to the deployment pipeline
- C. Viewer access to the Development and Test workspaces
- D. Viewer access to the Production workspace

- E. Contributor access to the Development and Test workspaces
- F. Contributor access to the Production workspace

Answer: BDE

Explanation:

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

NEW QUESTION 71

- (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 75

DRAG DROP - (Topic 2)

You have a Fabric tenant that contains a Microsoft Power BI report named Report 1. Report1 is slow to render. You suspect that an inefficient DAX query is being executed.

You need to identify the slowest DAX query, and then review how long the query spends in the formula engine as compared to the storage engine.

Which five actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.

Actions

View the Server Timings tab.

From Performance analyzer, capture a recording.

Enable Query Timings and Server Timings. Run the query.

View the Query Timings tab.

Sort the Duration (ms) column in descending order by DAX query time.

Copy the first query to DAX Studio.

Answer Area

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

To identify the slowest DAX query and analyze the time it spends in the formula engine compared to the storage engine, you should perform the following actions in sequence:

- ? From Performance analyzer, capture a recording.
- ? View the Server Timings tab.
- ? Enable Query Timings and Server Timings. Run the query.
- ? View the Query Timings tab.
- ? Sort the Duration (ms) column in descending order by DAX query time.

NEW QUESTION 78

.....

Thank You for Trying Our Product

We offer two products:

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

2nd - Questions 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](#)