

Exam Questions DP-600

Implementing Analytics Solutions Using Microsoft Fabric

<https://www.2passeasy.com/dumps/DP-600/>



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

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.

References =

? DAX FILTER Function

? DAX SUMMARIZE Function

NEW QUESTION 4

- (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.

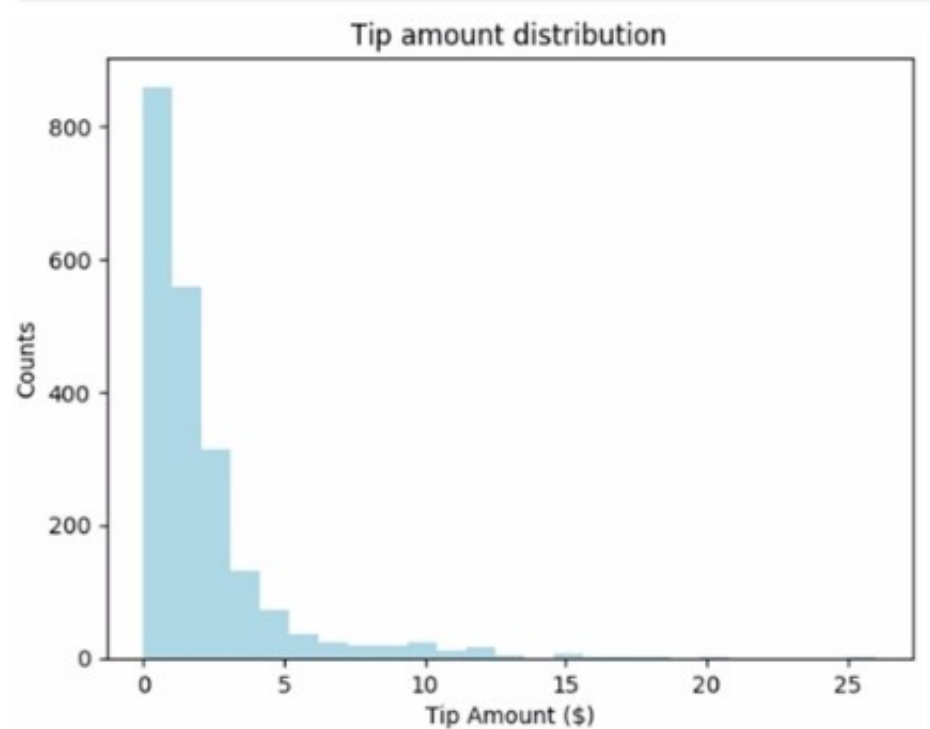
NEW QUESTION 5

- (Topic 2)

You have a Fabric notebook that has the Python code and output shown in the following exhibit.

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

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



Which type of analytics are you performing?

A. predictive

B. descriptive

C. prescriptive

D. diagnostic

Answer: B

Explanation:

The Python code and output shown in the exhibit display a histogram, which is a representation of the distribution of data. This kind of analysis is descriptive

analytics, which is used to describe or summarize the features of a dataset. Descriptive analytics answers the question of "what has happened" by providing insight into past data through tools such as mean, median, mode, standard deviation, and graphical representations like histograms.

References: Descriptive analytics and the use of histograms as a way to visualize data distribution are basic concepts in data analysis, often covered in introductory analytics and Python programming resources.

NEW QUESTION 6

- (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 BI 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 7

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

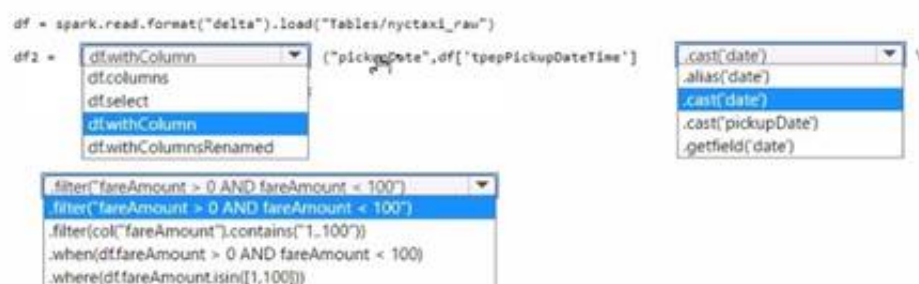
You create a Fabric notebook and attach it to lakehouse1.

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 8

- (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 9

- (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 10

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.

Statements	Yes	No
The Spark engine will read only the 'SalesOrderNumber', 'OrderDate','CustomerName', 'UnitPrice' columns from Sales_raw.csv.	<input type="radio"/>	<input type="radio"/>
Removing the partition will reduce the execution time of the query.	<input type="radio"/>	<input type="radio"/>
Adding inferSchema=true to the options will increase the execution time of the query.	<input type="radio"/>	<input type="radio"/>

- 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 10

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 Ad dress

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 14

- (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

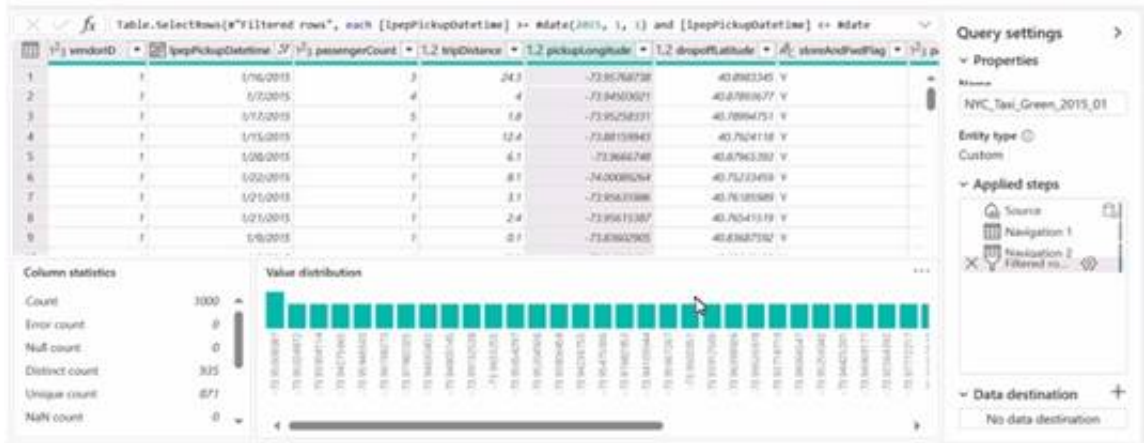
Explanation:

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 18

- (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 23

- (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 27

- (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 32

- (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 37

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.

Answer Area

SELECT ProductID,

GREATEST
COALESCE
GREATEST
IIF
MAX

(ListPrice, WholesalePrice, AgentPrice) AS HighestSellingPrice,

FROM

COALESCE
CHOOSE
COALESCE
IIF
MAX

(AgentPrice, WholesalePrice, ListPrice) AS TradePrice

- 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 41

.....

THANKS FOR TRYING THE DEMO OF OUR PRODUCT

Visit Our Site to Purchase the Full Set of Actual DP-600 Exam Questions With Answers.

We Also Provide Practice Exam Software That Simulates Real Exam Environment And Has Many Self-Assessment Features. Order the DP-600 Product From:

<https://www.2passeasy.com/dumps/DP-600/>

Money Back Guarantee

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