

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 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 [dbo].[ProductsWithPricingGroup]

AS

SELECT ProductId,

ProductName,

ProductCategory,

ListPrice,

WHEN ListPrice <= 50 THEN 'low'

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 2

- (Topic 1)

You need to ensure the data loading activities in the AnalyticsPOC workspace are executed in the appropriate sequence. The solution must meet the technical requirements.

What should you do?

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

Answer: A

Explanation:

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

source).

NEW QUESTION 3

- (Topic 1)

What should you recommend using to ingest the customer data into the data store in the AnalyticsPOC 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 4

- (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 5

- (Topic 2)

You have source data in a folder on a local computer.

You need to create a solution that will use Fabric to populate a data store. The solution must meet the following requirements:

- Support the use of dataflows to load and append data to the data store.
- Ensure that Delta tables are V-Order optimized and compacted automatically. Which type of data store should you use?

- A. a lakehouse
- B. an Azure SQL database
- C. a warehouse
- D. a KQL database

Answer: A

Explanation:

A lakehouse (A) is the type of data store you should use. It supports dataflows to load and append data and ensures that Delta tables are Z-Order optimized and compacted automatically. References = The capabilities of a lakehouse and its support for Delta tables are described in the lakehouse and Delta table documentation.

NEW QUESTION 6

- (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.explain()`

Does this meet the goal?

- A. Yes
- B. No

Answer: B

Explanation:

The `df.explain()` method does not meet the goal of evaluating data to calculate statistical functions. It is used to display the physical plan that Spark will execute. References = The correct usage of the `explain()` function can be found in the PySpark documentation.

NEW QUESTION 7

HOTSPOT - (Topic 2)

You have a data warehouse that contains a table named Stage. Customers. Stage- Customers contains all the customer record updates from a customer relationship management (CRM) system. There can be multiple updates per customer

You need to write a T-SQL query that will return the customer ID, name, postal code, and the last updated time of the most recent row for each customer ID. How should you complete the code? To answer, select the appropriate options in the answer area, NOTE Each correct selection is worth one point.

Answer Area

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

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

? In the ROW_NUMBER() function, choose OVER (PARTITION BY CustomerID ORDER BY LastUpdated DESC).

? In the WHERE clause, choose WHERE X = 1.

To select the most recent row for each customer ID, you use the ROW_NUMBER() window function partitioned by CustomerID and ordered by LastUpdated in descending order. This will assign a row number of 1 to the most recent update for each customer. By selecting rows where the row number (X) is 1, you get the latest update per customer. References =

? Use the OVER clause to aggregate data per partition

? Use window functions

NEW QUESTION 8

- (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 9

HOTSPOT - (Topic 2)

You have a Fabric warehouse that contains a table named Sales.Orders. Sales.Orders contains the following columns.

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

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

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

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

Answer Area

SELECT OrderID, CustomerID,

DATEFROMPARTS

DATE_BUCKET

DATEFROMPARTS

DATEPART

DATETRUNC

DATENAME(weekday, OrderDate) AS DayName

FROM Sales.Orders

day

dayofyear

weekday

- A. Mastered
B. Not Mastered

Answer: A

Explanation:

For the PeriodDate that returns the first day of the month for OrderDate, you should use DATEFROMPARTS as it allows you to construct a date from its individual components (year, month, day).

For the DayName that returns the name of the day for OrderDate, you should use

DATENAME with the weekday date part to get the full name of the weekday. The complete SQL query should look like this:

```
SELECT OrderID, CustomerID,
```

DATEFROMPARTS(YEAR(OrderDate), MONTH(OrderDate), 1) AS PeriodDate, DATENAME(weekday, OrderDate) AS DayName

FROM Sales.Orders

Select DATEFROMPARTS for the PeriodDate and weekday for the DayName in the answer area.

NEW QUESTION 10

- (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 10

NEW Q&A
- (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 13

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

df.columns

df.select

df.withColumn

df.withColumnsRenamed

```
df2 = df.withColumn("pickupDate", df["tpepPickupDateTime"]
```

.cast("date")

.alias("date")

.cast("date")

.cast("pickupDate")

.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 17

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.

Values

EnableFolding

NativeQuery

Optimize

Record

StopFolding

Table

Value

Answer Area

```
let
    Source = Sql.Databases(
        "server.database.windows.net"
    ),
    Database = Source[Name = "db"] [Data],
    Query = [ ] . [ ] (
        Database,
        " SELECT * FROM customer WHERE country IN ('USA', 'UK')",
        null,
        [ ] = true
    )
in
    Query
```

- 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 20

- (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 22

- (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 27

- (Topic 2)

You have a Fabric tenant that contains a new semantic model in OneLake. You use a Fabric notebook to read the data into a Spark DataFrame.

You need to evaluate the data to calculate the min, max, mean, and standard deviation values for all the string and numeric columns.

Solution: You use the following PySpark expression: df.show()

Does this meet the goal?

A. Yes

B. No

Answer: B

Explanation:

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

NEW QUESTION 32

HOTSPOT - (Topic 2)

You have a Fabric workspace named Workspace1 and an Azure Data Lake Storage Gen2 account named storage!. Workspace1 contains a lakehouse named Lakehouse1.

You need to create a shortcut to storage! in Lakehouse1.

Which connection and endpoint should you specify? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

Connection:

Endpoint:

- 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 37

- (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 41

- (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 that contains a workspace named Workspace1. Workspace1 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 44

- (Topic 2)

You have a Fabric tenant that contains a data pipeline.

You need to ensure that the pipeline runs every four hours on Mondays and Fridays. To what should you set Repeat for the schedule?

- A. Daily
- B. By the minute
- C. Weekly
- D. Hourly

Answer: C

Explanation:

You should set Repeat for the schedule to Weekly (C). This allows you to specify the pipeline to run on specific days of the week, in this case, every four hours on Mondays and Fridays. References = Scheduling options for data pipelines are available in the Azure Data Factory documentation, which includes details on configuring recurring triggers.

NEW QUESTION 49

HOTSPOT - (Topic 2)

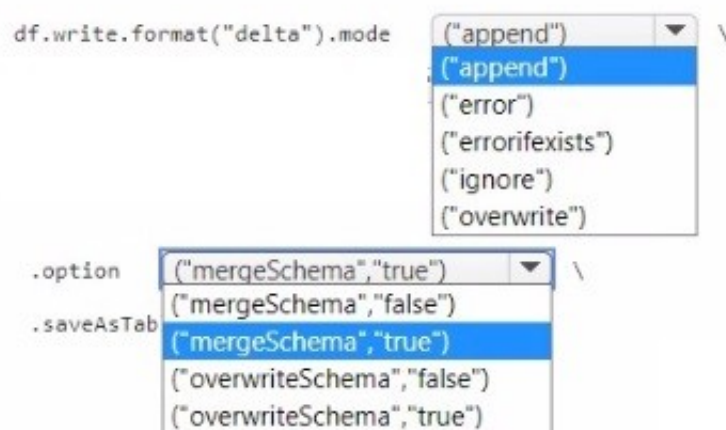
You have a Fabric tenant that contains lakehouse named Lakehouse1. Lakehouse1 contains a Delta table with eight columns. You receive new data that contains the same eight columns and two additional columns.

You create a Spark DataFrame and assign the DataFrame to a variable named df. The DataFrame contains the new data. You need to add the new data to the Delta table to meet the following requirements:

- Keep all the existing rows.
- Ensure that all the new data is added to the table.

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

Answer Area



- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

To 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 52

DRAG DROP - (Topic 2)

You are implementing a medallion architecture in a single Fabric workspace. You have a lakehouse that contains the Bronze and Silver layers and a warehouse that contains the Gold layer. You create the items required to populate the layers as shown in the following table.

Layer	Data integration tool
Bronze	Pipelines with Copy activities
Silver	Dataflows
Gold	Stored procedures

You need to ensure that the layers are populated daily in sequential order such that Silver is populated only after Bronze is complete, and Gold is populated only after Silver is complete. The solution must minimize development effort and complexity. What should you use to execute each set of items? To answer, drag the appropriate options to the correct items. Each option 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.

Execution Methods

A pipeline Copy activity

A pipeline Dataflow activity

A pipeline Stored procedure activity

A schedule

A Spark job definition

An Invoke pipeline activity

Answer Area

Orchestration pipeline:

Bronze layer:

Silver layer:

Gold layer:

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

To execute each set of items in sequential order with minimized development effort and complexity, you should use the following options:
? Orchestration pipeline: Use a pipeline with an Invoke pipeline activity. This allows for orchestrating and scheduling the execution of other pipelines, ensuring they run in the correct sequence.
? Bronze layer: Implement a pipeline Copy activity. This aligns with the table indicating that the Bronze layer uses pipelines with Copy activities for data integration.
? Silver layer: Implement a pipeline Dataflow activity. The table specifies that Dataflows are used for the Silver layer.
? Gold layer: Implement a pipeline Stored procedure activity. Stored procedures are specified for the Gold layer according to the table.

NEW QUESTION 56

- (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 57

- (Topic 2)

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

- A. Stored procedure
- B. Get metadata

- C. Lookup
- D. Copy data

Answer: C

Explanation:

In a Fabric Data Factory pipeline, to execute a stored procedure and make the returned values available for downstream activities, the Lookup activity is used. This activity can retrieve a dataset from a data store and pass it on for further processing. Here's how you would use the Lookup activity in this context:

- ? Add a Lookup activity to your pipeline.
- ? Configure the Lookup activity to use the stored procedure by providing the necessary SQL statement or stored procedure name.
- ? In the settings, specify that the activity should use the stored procedure mode.
- ? Once the stored procedure executes, the Lookup activity will capture the results and make them available in the pipeline's memory.
- ? Downstream activities can then reference the output of the Lookup activity. References: The functionality and use of Lookup activity within Azure Data Factory is documented in Microsoft's official documentation for Azure Data Factory, under the section for pipeline activities.

NEW QUESTION 61

- (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 65

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.	<input type="radio"/>	<input type="radio"/>
The code creates a Power BI report.	<input type="radio"/>	<input type="radio"/>
The code displays a summary of the DataFrame.	<input type="radio"/>	<input type="radio"/>

- 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 68

- (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 72

.....

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