

Exam Questions DP-600

Implementing Analytics Solutions Using Microsoft Fabric

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



NEW QUESTION 1

- (Topic 1)

Which type of data store should you recommend in the AnalyticsPOC workspace?

- A. a data lake
- B. a warehouse
- C. a lakehouse
- D. an external Hive metaStore

Answer: C

Explanation:

A lakehouse (C) should be recommended for the AnalyticsPOC workspace. It combines the capabilities of a data warehouse with the flexibility of a data lake. A lakehouse supports semi-structured and unstructured data and allows for T-SQL and Python read access, fulfilling the technical requirements outlined for Litware. References = For further understanding, Microsoft's documentation on the lakehouse architecture provides insights into how it supports various data types and analytical operations.

NEW QUESTION 2

HOTSPOT - (Topic 1)

You need to resolve the issue with the pricing group classification.

How should you complete the T-SQL statement? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

```
CREATE [ ] [dbo].[ProductsWithPricingGroup]
AS
SELECT ProductId,
       ProductName,
       ProductCategory,
       ListPrice,
       [ ]
       WHEN ListPrice <= 50 THEN 'low'
       [ ]
END AS PricingGroup
FROM dbo.Products
```

Answer Area

```
CREATE [VIEW] [ ] [dbo].[ProductsWithPricingGroup]
AS
SELECT ProductId,
       ProductCategory,
       ListPrice,
       [CASE] [ ] THEN 'low'
       [ ]
       WHEN (ListPrice >= 50 AND ListPrice < 1000) THEN 'medium'
       [ ]
       WHEN ListPrice BETWEEN 50 AND 1000 ) THEN 'medium'
END AS PricingGroup
FROM dbo.Products
```

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

C:\Users\Waqas Shahid\Desktop\Mudassir\Untitled.jpg

? You should use CREATE VIEW to make the pricing group logic available for T- SQL queries.

? The CASE statement should be used to determine the pricing group based on the list price.

The T-SQL statement should create a view that classifies products into pricing groups based on the list price. The CASE statement is the correct conditional logic to assign each product to the appropriate pricing group. This view will standardize the pricing group logic across different databases and semantic models.

NEW QUESTION 3

- (Topic 1)

You need to implement the date dimension in the data store. The solution must meet the technical requirements.

What are two ways to achieve the goal? Each correct answer presents a complete solution. NOTE: Each correct selection is worth one point.

- A. Populate the date dimension table by using a dataflow.
- B. Populate the date dimension table by using a Stored procedure activity in a pipeline.
- C. Populate the date dimension view by using T-SQL.
- D. Populate the date dimension table by using a Copy activity in a pipeline.

Answer: AB

Explanation:

Both a dataflow (A) and a Stored procedure activity in a pipeline (B) are capable of creating and populating a date dimension table. A dataflow can perform the transformation needed to create the date dimension, and it aligns with the preference for using low-code tools for data ingestion when possible. A Stored procedure could be written to generate the necessary date dimension data and executed within a pipeline, which also adheres to the technical requirements for the PoC.

NEW QUESTION 4

- (Topic 1)

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

What should you do?

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

Answer: A

Explanation:

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

NEW QUESTION 5

- (Topic 2)

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

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

Answer: A

Explanation:

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

NEW QUESTION 6

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.

- 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 7

- (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 8

- (Topic 2)

You have a semantic model named Model 1. Model 1 contains five tables that all use Import mode. Model1 contains a dynamic row-level security (RLS) role named HR. The HR role filters employee data so that HR managers only see the data of the department to which they are assigned.

You publish Model1 to a Fabric tenant and configure RLS role membership. You share the model and related reports to users.

An HR manager reports that the data they see in a report is incomplete. What should you do to validate the data seen by the HR Manager?

- A. Ask the HR manager to open the report in Microsoft Power BI Desktop.
- B. Select Test as role to view the data as the HR role.
- C. Select Test as role to view the report as the HR manager,
- D. Filter the data in the report to match the intended logic of the filter for the HR department.

Answer: B

Explanation:

To validate the data seen by the HR manager, you should use the 'Test as role' feature in Power BI service. This allows you to see the data exactly as it would appear for the HR role, considering the dynamic RLS setup. Here is how you would proceed:

- ? Navigate to the Power BI service and locate Model1.
- ? Access the dataset settings for Model1.
- ? Find the security/RLS settings where you configured the roles.
- ? Use the 'Test as role' feature to simulate the report viewing experience as the HR role.
- ? Review the data and the filters applied to ensure that the RLS is functioning correctly.
- ? If discrepancies are found, adjust the RLS expressions or the role membership as needed.

References: The 'Test as role' feature and its use for validating RLS in Power BI is covered in the Power BI documentation available on Microsoft's official documentation.

NEW QUESTION 9

DRAG DROP - (Topic 2)

You have a Fabric tenant that contains a lakehouse named Lakehouse1

Readings from 100 IoT devices are appended to a Delta table in Lakehouse1. Each set of readings is approximately 25 KB. Approximately 10 GB of data is received daily.

All the table and SparkSession settings are set to the default.

You discover that queries are slow to execute. In addition, the lakehouse storage contains data and log files that are no longer used.

You need to remove the files that are no longer used and combine small files into larger files with a target size of 1 GB per file.

What should you do? To answer, drag the appropriate actions to the correct requirements. Each action may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.

NOTE: Each correct selection is worth one point.

Actions	Answer Area
Set the autoCompact table setting.	Remove the files: <input type="text"/>
Set the optimizeWrite table setting.	Combine the files: <input type="text"/>
Run the VACUUM command on a schedule.	
Set the autoCompact SparkSession setting.	
Run the OPTIMIZE command on a schedule.	
Set the parallelDelete SparkSession setting.	

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

? Remove the files: Run the VACUUM command on a schedule.

? Combine the files: Set the optimizeWrite table setting. or Run the OPTIMIZE command on a schedule.

To remove files that are no longer used, the VACUUM command is used in Delta Lake to clean up invalid files from a table. To combine smaller files into larger ones, you can either set the optimizeWrite setting to combine files during write operations or use the OPTIMIZE command, which is a Delta Lake operation used to compact small files into larger ones.

NEW QUESTION 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 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(
FROM Sales.Orders
weekday
day
dayofyear
weekday
, OrderDate) AS DayName
```

- 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 14

- (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 15

- (Topic 2)

You have a Fabric tenant named Tenant1 that contains a workspace named WS1. WS1 uses a capacity named C1 and contains a dataset named DS1. You need to ensure read- write access to DS1 is available by using the XMLA endpoint. What should be modified first?

- A. the DS1 settings
- B. the WS1 settings
- C. the C1 settings
- D. the Tenant1 settings

Answer: C

Explanation:

To ensure read-write access to DS1 is available by using the XMLA endpoint, the C1 settings (which refer to the capacity settings) should be modified first. XMLA endpoint configuration is a capacity feature, not specific to individual datasets or workspaces. References = The configuration of XMLA endpoints in Power BI capacities is detailed in the Power BI documentation on dataset management.

NEW QUESTION 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

```
df = spark.read.format("delta").load("Tables/nyctaxi_raw")
df2 = df.withColumn("pickupDate", df["pickupDateTime"].cast("date"))
df2.filter("fareAmount > 0 AND fareAmount < 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 22

- (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 24

- (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 27

- (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 31

- (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 35

- (Topic 2)

You have a Fabric tenant that contains a lakehouse.

You plan to query sales data files by using the SQL endpoint. The files will be in an Amazon Simple Storage Service (Amazon S3) storage bucket.

You need to recommend which file format to use and where to create a shortcut. Which two actions should you include in the recommendation? Each correct answer presents part of the solution.

NOTE: Each correct answer is worth one point.

- A. Create a shortcut in the Files section.
- B. Use the Parquet format
- C. Use the CSV format.
- D. Create a shortcut in the Tables section.
- E. Use the delta format.

Answer: BD

Explanation:

You should use the Parquet format (B) for the sales data files because it is optimized for performance with large datasets in analytical processing and create a shortcut in the Tables section (D) to facilitate SQL queries through the lakehouse's SQL endpoint. References = The best practices for working with file formats and shortcuts in a lakehouse environment are covered in the lakehouse and SQL endpoint documentation provided by the cloud data platform services.

NEW QUESTION 40

- (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 45

- (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

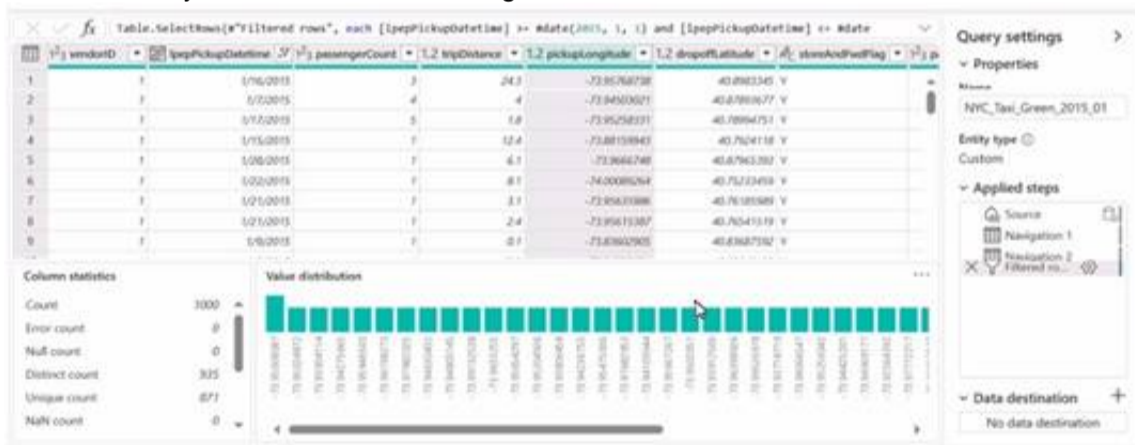
Explanation:

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 50

- (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 51

DRAG DROP - (Topic 2)

You are implementing two dimension tables named Customers and Products in a Fabric warehouse. You need to use slowly changing dimension (SCD) to manage the versioning of data. The solution must meet the requirements shown in the following table.

Table	Change action
Customers	Create a new version of the row.
Products	Overwrite the existing value in the latest row.

Which type of SCD should you use for each table? To answer, drag the appropriate SCD types to the correct tables. Each SCD type may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.

NOTE: Each correct selection is worth one point.

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

For the Customers table, where the requirement is to create a new version of the row, you would use:

? Type 2 SCD: This type allows for the creation of a new record each time a change occurs, preserving the history of changes over time.

For the Products table, where the requirement is to overwrite the existing value in the latest row, you would use:

? Type 1 SCD: This type updates the record directly, without preserving historical data.

NEW QUESTION 53

- (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:

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 55

- (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 56

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 59

- (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 64

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 (ListPrice, WholesalePrice, AgentPrice) AS HighestSellingPrice,

COALESCE

GREATEST

IIF

MAX

COALESCE (AgentPrice, WholesalePrice, ListPrice) AS TradePrice

FROM

COALESCE

CHOOSE

COALESCE

IIF

MAX

- 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 66

.....

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