

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

- (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 5

- (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 6

- (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 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

```
df = spark.read.format("delta").load("Tables/nyctaxi_raw")
df2 = df.withColumn("pickupDate", df["pickupDateTime"].cast("date"))
df2 = 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 8

- (Topic 2)

You have a Fabric tenant that contains a semantic model. The model uses Direct Lake mode.

You suspect that some DAX queries load unnecessary columns into memory. You need to identify the frequently used columns that are loaded into memory.

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

- A. Use the Analyze in Excel feature.
- B. Use the Vertipaq Analyzer tool.
- C. Query the `$system.discovered_STORAGE_TABLE_COLUMN-IN_SEGMENTS` dynamic management view (DMV).
- D. Query the `discover_hehory6Rant` dynamic management view (DMV).

Answer: BC

Explanation:

The Vertipaq Analyzer tool (B) and querying the `$system.discovered_STORAGE_TABLE_COLUMNS_IN_SEGMENTS` dynamic management view (DMV) (C) can help identify which columns are frequently loaded into memory. Both methods provide insights into the storage and retrieval aspects of the semantic model. References = The Power BI documentation on Vertipaq Analyzer and DMV queries offers detailed guidance on how to use these tools for performance analysis.

NEW QUESTION 9

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 10

- (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 10

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: abfss ▼

- abfs
- abfss
- https

Endpoint: dfs ▼

- blob
- dfs
- file

- 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 14

- (Topic 2)

You have a Fabric tenant that 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: 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 17

- (Topic 2)

You have a Fabric tenant that contains a workspace named Workspace^ 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 18

- (Topic 2)

You are the administrator of a Fabric workspace that contains a lakehouse named Lakehouse1. Lakehouse1 contains the following tables:

- Table1: A Delta table created by using a shortcut
- Table2: An external table created by using Spark
- Table3: A managed table

You plan to connect to Lakehouse1 by using its SQL endpoint. What will you be able to do after connecting to Lakehouse1?

- A. ReadTable3.
- B. Update the data Table3.
- C. ReadTable2.
- D. Update the data in Table1.

Answer: D

NEW QUESTION 20

- (Topic 2)

You have a Fabric tenant that 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 25

- (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 30

- (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 32

- (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. displayHTML
- 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 35

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 39

- (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 40

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 42

- (Topic 2)

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

Name	Data type	Nullable
ProductID	Integer	No
ProductName	Varchar(30)	No
SalesDate	Datetime2(6)	No
WholesalePrice	Decimal(18, 2)	Yes
Amount	Decimal(18, 2)	Yes

You need to write a T-SQL query that will return data for the year 2023 that displays ProductID and ProductName and has a summarized Amount that is higher than 10,000. Which query should you use?

A)

```
SELECT ProductID, ProductName, SUM(Amount) AS TotalAmount
FROM Staging.Sales
WHERE DATEPART(YEAR,SaleDate) = '2023'
GROUP BY ProductID, ProductName
HAVING SUM(Amount) > 10000
```

B)

```
SELECT ProductID, ProductName, SUM(Amount) AS TotalAmount
FROM Staging.Sales
GROUP BY ProductID, ProductName
HAVING DATEPART(YEAR,SaleDate) = '2023' AND SUM(Amount) > 10000
```

C)

```
SELECT ProductID, ProductName, SUM(Amount) AS TotalAmount
FROM Staging.Sales
WHERE DATEPART(YEAR,SaleDate) = '2023' AND SUM(Amount) > 10000
```

D)

```
SELECT ProductID, ProductName, SUM(Amount) AS TotalAmount
FROM Staging.Sales
WHERE DATEPART(YEAR,SaleDate) = '2023'
GROUP BY ProductID, ProductName
HAVING TotalAmount > 10000
```

- A. Option A
- B. Option B
- C. Option C
- D. Option D

Answer: B

Explanation:

The correct query to use in order to return data for the year 2023 that displays ProductID, ProductName, and has a summarized Amount greater than 10,000 is Option B. The reason is that it uses the GROUP BY clause to organize the data by ProductID and ProductName and then filters the result using the HAVING clause to only include groups where the sum of Amount is greater than 10,000. Additionally, the DATEPART(YEAR, SaleDate) = '2023' part of the HAVING clause ensures that only records from the year 2023 are included. References = For more information, please visit the official documentation on T-SQL queries and the GROUP BY clause at T-SQL GROUP BY.

NEW QUESTION 47

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 51

- (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 52

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	Answer Area
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.	

- 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 56

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,
  (ListPrice, WholesalePrice, AgentPrice) AS HighestSellingPrice,
  (AgentPrice, WholesalePrice, ListPrice) AS TradePrice
FROM Sales.Products
```

For the first dropdown menu, the options are: GREATEST, COALESCE, GREATEST, IIF, MAX. The selected option is GREATEST.

For the second dropdown menu, the options are: COALESCE, CHOOSE, COALESCE, IIF, MAX. The selected option is COALESCE.

- 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 58

.....

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](#)