

Exam Questions DP-600

Implementing Analytics Solutions Using Microsoft Fabric

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



NEW QUESTION 1

HOTSPOT - (Topic 1)

You need to design a semantic model for the customer satisfaction report.

Which data source authentication method and mode should you use? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

Authentication method:

- Service principal authentication
- Basic authentication
- Service principal authentication
- Single sign-on (SSO) authentication

Mode:

- DirectQuery
- Direct Lake
- DirectQuery
- Import

- A. Mastered
- B. Not Mastered

Answer: A**Explanation:**

For the semantic model design required for the customer satisfaction report, the choices for data source authentication method and mode should be made based on security and performance considerations as per the case study provided.

Authentication method: The data should be accessed securely, and given that row-level security (RLS) is required for users executing T-SQL queries, you should use an authentication method that supports RLS. Service principal authentication is suitable for automated and secure access to the data, especially when the access needs to be controlled programmatically and is not tied to a specific user's credentials.

Mode: The report needs to show data as soon as it is updated in the data store, and it should only contain data from the current and previous year. DirectQuery mode allows for real-time reporting without importing data into the model, thus meeting the need for up-to-date data. It also allows for RLS to be implemented and enforced at the data source level, providing the necessary security measures.

Based on these considerations, the selections should be:

? Authentication method: Service principal authentication

? Mode: DirectQuery

NEW QUESTION 2

- (Topic 1)

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 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 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

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 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

- (Topic 2)

You have a Fabric tenant that contains a warehouse. Several times a day, the performance of all warehouse queries degrades. You suspect that Fabric is throttling the compute used by the warehouse. What should you use to identify whether throttling is occurring?

- A. the Capacity settings
- B. the Monitoring hub
- C. dynamic management views (DMVs)
- D. the Microsoft Fabric Capacity Metrics app

Answer: B

Explanation:

To identify whether throttling is occurring, you should use the Monitoring hub (B). This provides a centralized place where you can monitor and manage the health, performance, and reliability of your data estate, and see if the compute resources are being throttled. References = The use of the Monitoring hub for performance management and troubleshooting is detailed in the Azure Synapse Analytics documentation.

NEW QUESTION 8

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.

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

You should complete the code as follows:

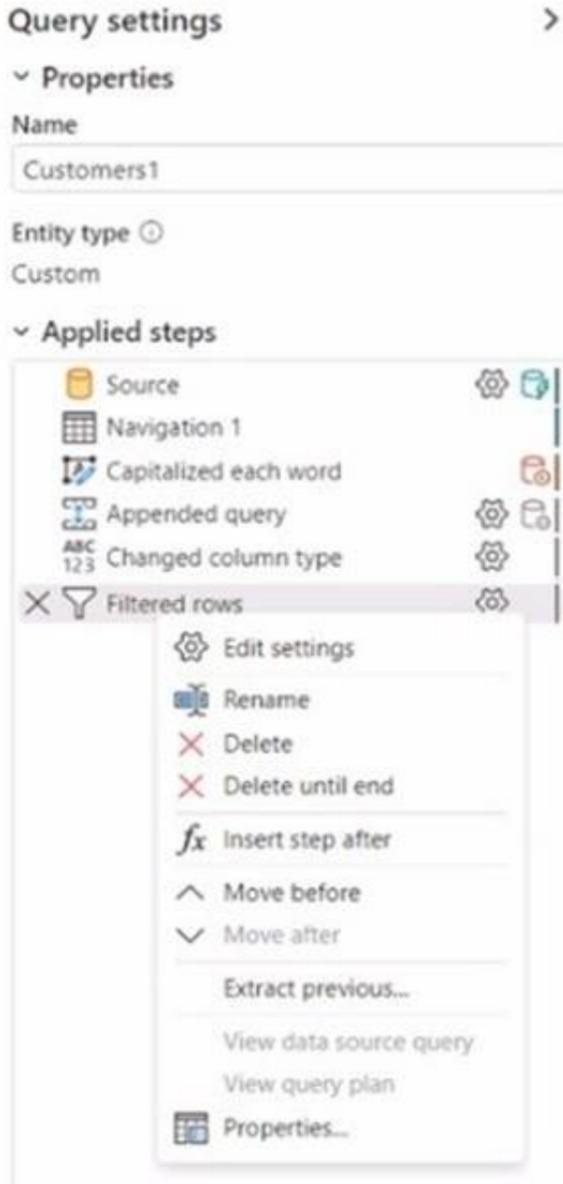
- ? Table
- ? NativeQuery
- ? EnableFolding

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 9

HOTSPOT - (Topic 2)

You have a Fabric tenant that contains two lakehouses. You are building a dataflow that will combine data from the lakehouses. The applied steps from one of the queries in the dataflow is shown in the following exhibit.



Use the drop-down menus to select the answer choice that completes each statement based on the information presented in the graphic. NOTE: Each correct selection is worth one point.

Answer Area

[Answer choice] of the transformation steps in the query will fold.

The Added custom step will be performed in [answer choice].

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Folding in Power Query refers to operations that can be translated into source queries. In this case, "some" of the steps can be folded, which means that some transformations will be executed at the data source level. The steps that cannot be folded will be executed within the Power Query engine. Custom steps, especially those that are not standard query operations, are usually executed within Power Query engine rather than being pushed down to the source system.

- References =
- ? Query folding in Power Query
- ? Power Query M formula language

NEW QUESTION 10

- (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 10

- (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 13

- (Topic 2)

You have a Fabric tenant that contains a workspace named Workspace^ Workspacel 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 16

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.

SCD Types

Type 0

Type 1

Type 2

Type 3

Answer Area

Customers:

Products:

- 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 20

- (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 22

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 23

- (Topic 2)

You have a Fabric tenant that contains a warehouse.

You use a dataflow to load a new dataset from OneLake to the warehouse.

You need to add a Power Query step to identify the maximum values for the numeric columns.

Which function should you include in the step?

- A. Tabl
- B. MaxN
- C. Table.Max
- D. Table.Range
- E. Table.Profile

Answer: B

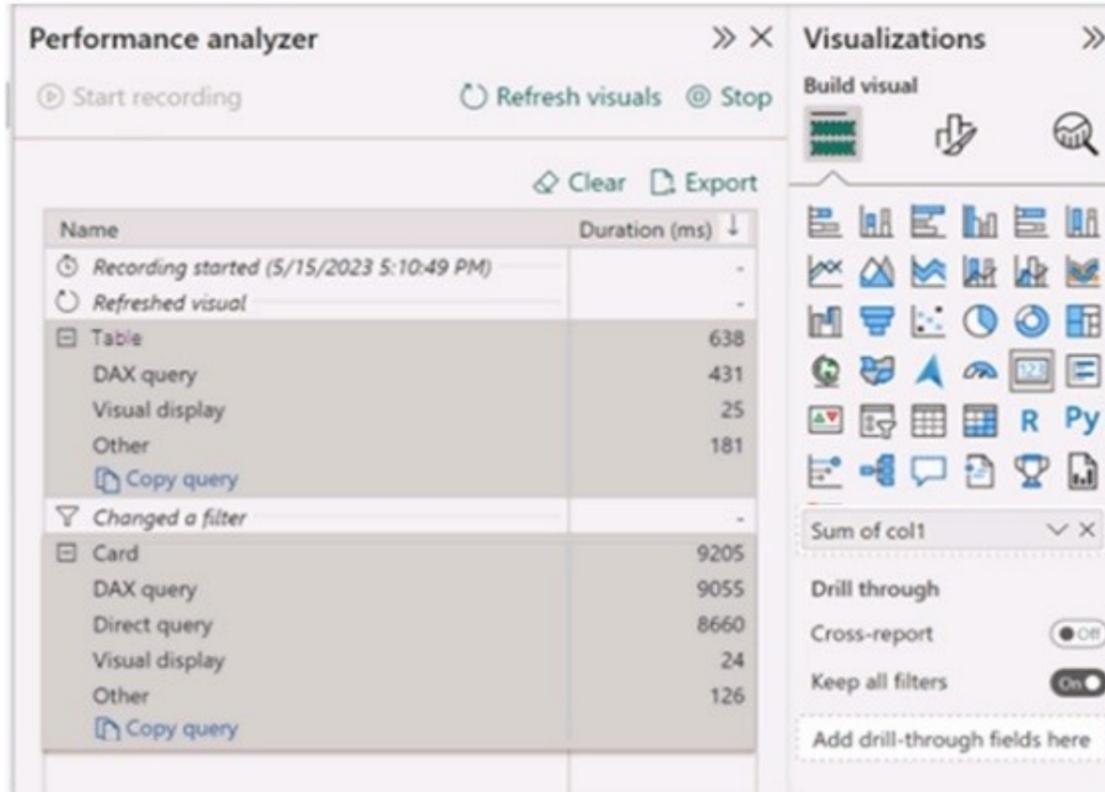
Explanation:

The Table.Max function should be used in a Power Query step to identify the maximum values for the numeric columns. This function is designed to calculate the maximum value across each column in a table, which suits the requirement of finding maximum values for numeric columns. References = For detailed information on Power Query functions, including Table.Max, please refer to Power Query M function reference.

NEW QUESTION 25

HOTSPOT - (Topic 2)

You have a Microsoft Power BI report and a semantic model that uses Direct Lake mode. From Power BI Desktop, you open Performance analyzer as shown in the following exhibit.



Use the drop-down menus to select the answer choice that completes each statement based on the information presented in the graphic. NOTE: Each correct selection is worth one point.

Answer Area

The Direct Lake fallback behavior is set to [answer choice].

- DirectQueryOnly
- Automatic
- DirectLakeOnly
- DirectQueryOnly

The query for the table visual is executed by using [answer choice].

- the composite model
- Direct Lake
- DirectQuery

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

? The Direct Lake fallback behavior is set to: DirectQueryOnly
 ? The query for the table visual is executed by using: DirectQuery
 In the context of Microsoft Power BI, when using DirectQuery in Direct Lake mode, there is no caching of data and all queries are sent directly to the underlying data source. The Performance Analyzer tool shows the time taken for different operations, and from the options provided, it indicates that DirectQuery mode is being used for the visuals, which is consistent with the Direct Lake setting. DirectQueryOnly as the fallback behavior ensures that only DirectQuery will be used without reverting to import mode.

NEW QUESTION 30

- (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 33

- (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. Report1 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 38

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
```

The image shows a SQL editor interface with two dropdown menus. The first dropdown is for the HighestSellingPrice column and has options: GREATEST, COALESCE, GREATEST (highlighted), IIF, and MAX. The second dropdown is for the TradePrice column and has options: COALESCE (highlighted), CHOOSE, COALESCE, IIF, and 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 43

.....

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