

# Microsoft

## Exam Questions DP-600

Implementing Analytics Solutions Using Microsoft Fabric



**NEW QUESTION 1**

- (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 2**

HOTSPOT - (Topic 2)

You have a Fabric tenant that contains a warehouse named Warehouse1. Warehouse1 contains a fact table named FactSales that has one billion rows. You run the following T-SQL statement.

CREATE TABLE test.FactSales AS CLONE OF Dbo.FactSales;

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
A replica of dbo.Sales is created in the test schema by copying the metadata only.	<input type="radio"/>	<input type="radio"/>
Additional schema changes to dbo.FactSales will also apply to test.FactSales.	<input type="radio"/>	<input type="radio"/>
Additional data changes to dbo.FactSales will also apply to test.FactSales.	<input type="radio"/>	<input type="radio"/>

- A. Mastered
- B. Not Mastered

**Answer: A**

**Explanation:**

? A replica of dbo.Sales is created in the test schema by copying the metadata only.

- No

? Additional schema changes to dbo.FactSales will also apply to test.FactSales. - No

? Additional data changes to dbo.FactSales will also apply to test.FactSales. - Yes

The CREATE TABLE AS CLONE statement creates a copy of an existing table, including its data and any associated data structures, like indexes. Therefore, the statement does not merely copy metadata; it also copies the data. However, subsequent schema changes to the original table do not automatically propagate to the cloned table. Any data changes in the original table after the clone operation will not be reflected in the clone unless explicitly updated.

References =

? CREATE TABLE AS SELECT (CTAS) in SQL Data Warehouse

**NEW QUESTION 3**

- (Topic 2)

You have a Fabric tenant that contains a lakehouse. You plan to use a visual query to merge two tables.

You need to ensure that the query returns all the rows that are present in both tables. Which type of join should you use?

- A. left outer
- B. right anti
- C. full outer
- D. left anti
- E. right outer
- F. inner

**Answer: C**

**Explanation:**

When you need to return all rows that are present in both tables, you use a full outer join. This type of join combines the results of both left and right outer joins and returns all rows from both tables, with matching rows from both sides where available. If there is no match, the result is NULL on the side of the join where there is no match. References: Information about joins and their use in querying data in a lakehouse can be typically found in the SQL and data processing documentation of the Fabric tenant or lakehouse solutions.

**NEW QUESTION 4**

- (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 5**

- (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.

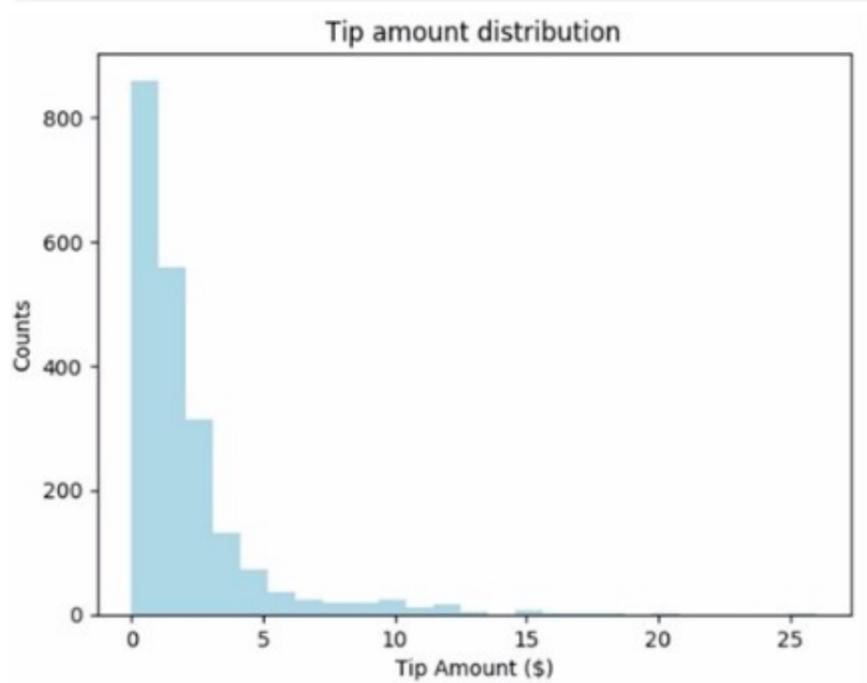
**NEW QUESTION 6**

- (Topic 2)

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

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

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



Which type of analytics are you performing?

- A. predictive
- B. descriptive
- C. prescriptive
- D. diagnostic

**Answer: B**

**Explanation:**

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

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

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

### NEW QUESTION 7

- (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 8

- (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 9

HOTSPOT - (Topic 2)

You have a Fabric workspace that uses the default Spark starter pool and runtime version 1,2.

You plan to read a CSV file named Sales.raw.csv in a lakehouse, select columns, and save the data as a Delta table to the managed area of the lakehouse.

Sales\_raw.csv contains 12 columns.

You have the following code.

```
from pyspark.sql.functions import year

(spark
 .read
 .format("csv")
 .option("header", 'true')
 .load("Files/sales_raw.csv")
 .select('SalesOrderNumber', 'OrderDate', 'CustomerName', 'UnitPrice')
 .withColumn("Year", year("OrderDate"))
 .write
 .partitionBy('Year')
 .saveAsTable("sales")
)
```

For each of the following statements, select Yes if the statement is true. Otherwise, select No. NOTE: Each correct selection is worth one point.

Answer Area

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

- A. Mastered
- B. Not Mastered

**Answer: A**

#### Explanation:

? The Spark engine will read only the 'SalesOrderNumber', 'OrderDate', 'CustomerName', 'UnitPrice' columns from Sales\_raw.csv. - Yes

? Removing the partition will reduce the execution time of the query. - No

? Adding inferSchema='true' to the options will increase the execution time of the query. - Yes

The code specifies the selection of certain columns, which means only those columns will be read into the DataFrame. Partitions in Spark are a way to optimize the execution of queries by organizing the data into parts that can be processed in parallel. Removing the partition could potentially increase the execution time because Spark would no longer be

able to process the data in parallel efficiently. The inferSchema option allows Spark to automatically detect the column data types, which can increase the execution time of the initial read operation because it requires Spark to read through the data to infer the schema.

**NEW QUESTION 10**

- (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 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 12**

HOTSPOT - (Topic 2)

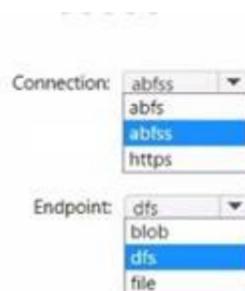
You have a Fabric workspace named Workspace1 and an Azure Data Lake Storage Gen2 account named storage!. Workspace1 contains a lakehouse named Lakehouse1.

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



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 13**

- (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

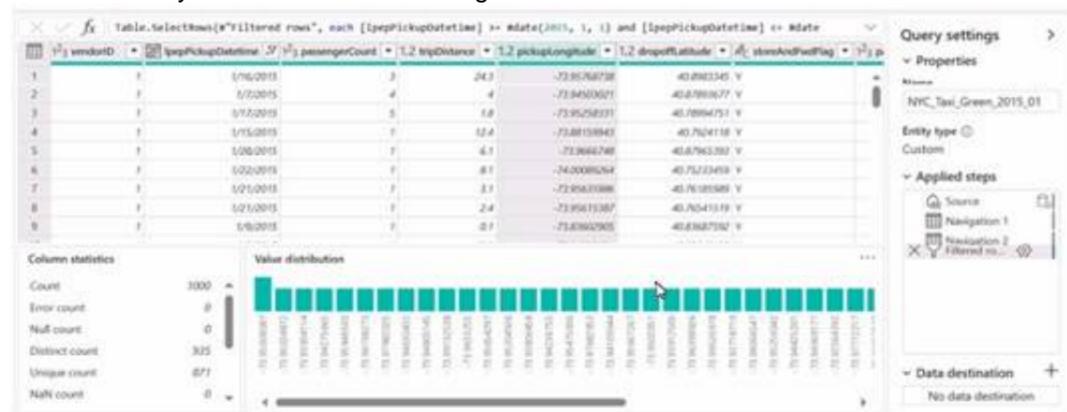
**Explanation:**

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 17**

- (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 21**

- (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 25

- (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 26

- (Topic 2)

You are creating a semantic model in Microsoft Power BI Desktop.

You plan to make bulk changes to the model by using the Tabular Model Definition Language (TMDL) extension for Microsoft Visual Studio Code.

You need to save the semantic model to a file. Which file format should you use?

- A. PBIP
- B. PBIX
- C. PBIT
- D. PBIDS

**Answer:** B

#### Explanation:

When saving a semantic model to a file that can be edited using the Tabular Model Scripting Language (TMSL) extension for Visual Studio Code, the PBIX (Power BI Desktop) file format is the correct choice. The PBIX format contains the report, data model, and queries, and is the primary file format for editing in Power BI Desktop. References = Microsoft's documentation on Power BI file formats and Visual Studio Code provides further clarification on the usage of PBIX files.

#### NEW QUESTION 27

- (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 31

- (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 34

- (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 36**

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 two dropdown menus for completing the SQL query. 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 37**

.....

## **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](#)**