

Certification Provider: Microsoft

Exam Code: DP-600

Exam Name: Implementing Analytics Solutions Using Microsoft Fabric

Collated By: Million Zhang

Collated Date: 24th May 2024

Number of Question:

Topic	Number of Questions
Case Study 1 - Contoso, Ltd.	8
Case Study 2 - Litware, Inc.	9
Solution Challenge	11
Common Question	81
Total	109

Case Study 1 – Contoso, Ltd.

Overview -

Contoso, Ltd. is a US-based health supplements company. Contoso has two divisions named Sales and Research. The Sales division contains two departments named Online Sales and Retail Sales. The Research division assigns internally developed product lines to individual teams of researchers and analysts.

Existing Environment -

Identity Environment -

Contoso has a Microsoft Entra tenant named contoso.com. The tenant contains two groups named ResearchReviewersGroup1 and ResearchReviewersGroup2.

Data Environment -

Contoso has the following data environment:

The Sales division uses a Microsoft Power BI Premium capacity.

The semantic model of the Online Sales department includes a fact table named Orders that uses Import mode. In the system of origin, the OrderID value represents the sequence in which orders are created.

The Research department uses an on-premises, third-party data warehousing product. Fabric is enabled for contoso.com.

An Azure Data Lake Storage Gen2 storage account named storage1 contains Research division data for a product line named Productline1. The data is in the delta format.

A Data Lake Storage Gen2 storage account named storage2 contains Research division data for a product line named Productline2. The data is in the CSV format.

Requirements -

Planned Changes -

Contoso plans to make the following changes:

Enable support for Fabric in the Power BI Premium capacity used by the Sales division.

Make all the data for the Sales division and the Research division available in Fabric.

For the Research division, create two Fabric workspaces named Productline1ws and Productline2ws.

In Productline1ws, create a lakehouse named Lakehouse1.

In Lakehouse1, create a shortcut to storage1 named ResearchProduct.

Data Analytics Requirements -

Contoso identifies the following data analytics requirements:

All the workspaces for the Sales division and the Research division must support all Fabric experiences.

The Research division workspaces must use a dedicated, on-demand capacity that has per-minute billing.

The Research division workspaces must be grouped together logically to support OneLake data hub filtering based on the department name.

For the Research division workspaces, the members of ResearchReviewersGroup1 must be able to read lakehouse and warehouse data and shortcuts by using SQL endpoints.

For the Research division workspaces, the members of ResearchReviewersGroup2 must be able to read lakehouse data by using Lakehouse explorer.

All the semantic models and reports for the Research division must use version control that supports branching.

Data Preparation Requirements -

Contoso identifies the following data preparation requirements:

The Research division data for Productline1 must be retrieved from Lakehouse1 by using Fabric notebooks.

All the Research division data in the lakehouses must be presented as managed tables in Lakehouse explorer.

Semantic Model Requirements -

Contoso identifies the following requirements for implementing and managing semantic models:

The number of rows added to the Orders table during refreshes must be minimized.

The semantic models in the Research division workspaces must use Direct Lake mode.

General Requirements -

Contoso identifies the following high-level requirements that must be considered for all solutions:

Follow the principle of least privilege when applicable.

Minimize implementation and maintenance effort when possible.

Question 1

You need to ensure that Contoso can use version control to meet the data analytics requirements and the general requirements.

What should you do?

- A. Store at the semantic models and reports in Data Lake Gen2 storage.
- B. Modify the settings of the Research workspaces to use a GitHub repository.
- C. Modify the settings of the Research division workspaces to use an Azure Repos repository.
- D. Store all the semantic models and reports in Microsoft OneDrive.

Answer: C

Question 2

You need to recommend a solution to group the Research division workspaces.

What should you include in the recommendation? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area



Grouping method:

	▼
Capacity	
Domain	
Tenant	




Tool:

	▼
OneLake data hub	
The Fabric Admin portal	
The Microsoft Entra admin center	

Answer:

 **Momoanwar** Highly Voted  3 months ago

Domain\Fabric Admin Portal

   upvoted 21 times

Question 3

You need to refresh the Orders table of the Online Sales department. The solution must meet the semantic model requirements.

What should you include in the solution?

- A. an Azure Data Factory pipeline that executes a Stored procedure activity to retrieve the maximum value of the OrderID column in the destination lakehouse
- B. an Azure Data Factory pipeline that executes a Stored procedure activity to retrieve the minimum value of the OrderID column in the destination lakehouse
- C. an Azure Data Factory pipeline that executes a dataflow to retrieve the minimum value of the OrderID column in the destination lakehouse
- D. an Azure Data Factory pipeline that executes a dataflow to retrieve the maximum value of the OrderID column in the destination lakehouse

Answer: D

Question 4

Which syntax should you use in a notebook to access the Research division data for Productline1?

- A. `spark.read.format("delta").load("Tables/productline1/ResearchProduct")`
- B. `spark.sql("SELECT * FROM Lakehouse1.ResearchProduct ")`
- C. `external_table('Tables/ResearchProduct)`
- D. `external_table(ResearchProduct)`

Answer: B

Question 5

You need to recommend which type of Fabric capacity SKU meets the data analytics requirements for the Research division.

What should you recommend?

- A. A
- B. EM
- C. P
- D. F

Answer: D

Question 6

You need to migrate the Research division data for Productline1. The solution must meet the data preparation requirements.

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('csv').options(header="true",inferSchema="true").load("fs1@storage1.dfs.core.windows.net/files/productline1")
```

```
df.write.mode("overwrite").format(" ").save(" ")
```

▼

csv
delta
parquet

▼

productline1
Tables/productline1
Tables/research/productline1

Answer:

Answer Area

```
df=spark.read.format('csv').options(header="true",inferSchema="true").load("fs1@storage1.dfs.core.windows.net/files/productline1")
```

```
df.write.mode("overwrite").format(" ").save(" ")
```

▼

csv
delta
parquet

▼

productline1
Tables/productline1
Tables/research/productline1

Question 7

What should you use to implement calculation groups for the Research division semantic models?

- A. Microsoft Power BI Desktop
- B. the Power BI service
- C. DAX Studio
- D. Tabular Editor


Answer: D

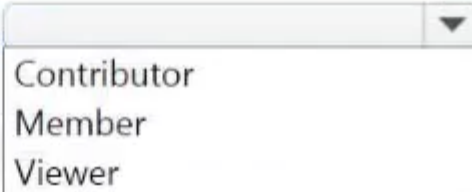
Question 8

Which workspace role assignments should you recommend for ResearchReviewersGroup1 and ResearchReviewersGroup2? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

ResearchReviewersGroup1: 

ResearchReviewersGroup2: 

Answer:

Answer Area

ResearchReviewersGroup1: 

ResearchReviewersGroup2: 

Case Study 2 - Litware, Inc.

Overview -

Litware, Inc. is a manufacturing company that has offices throughout North America. The analytics team at Litware contains data engineers, analytics engineers, data analysts, and data scientists.

Existing Environment -

Fabric Environment -

Litware has been using a Microsoft Power BI tenant for three years. Litware has NOT enabled any Fabric capacities and features.

Available Data -

Litware has data that must be analyzed as shown in the following table.

Description	Original source	Total size
Customer data	Customer relationship management (CRM) system	50 MB
Product data	Customer relationship management (CRM) system	200 MB
Customer satisfaction surveys	SurveyMonkey	500 GB

The Product data contains a single table and the following columns.

Name	Data type
ProductID	Integer
ProductName	String
ProductCategory	String
ListPrice	Decimal

The customer satisfaction data contains the following tables:

- Survey -
- Question -
- Response -

For each survey submitted, the following occurs:

One row is added to the Survey table.

One row is added to the Response table for each question in the survey.

The Question table contains the text of each survey question. The third question in each survey response is an overall satisfaction score. Customers can submit a survey after each purchase.

User Problems -

The analytics team has large volumes of data, some of which is semi-structured. The team wants to use Fabric to create a new data store.

Product data is often classified into three pricing groups: high, medium, and low. This logic is implemented in several databases and semantic models, but the logic does NOT always match across implementations.

Requirements -

Planned Changes -

Litware plans to enable Fabric features in the existing tenant. The analytics team will create a new data store as a proof of concept (PoC). The remaining Liware users will only get access to the Fabric features once the PoC is complete. The PoC will be completed by using a Fabric trial capacity

The following three workspaces will be created:

AnalyticsPOC: Will contain the data store, semantic models, reports pipelines, dataflow, and notebooks used to populate the data store

DataEngPOC: Will contain all the pipelines, dataflows, and notebooks used to populate OneLake

DataSciPOC: Will contain all the notebooks and reports created by the data scientists

The following will be created in the AnalyticsPOC workspace:

- A data store (type to be decided)
- A custom semantic model -
- A default semantic model -

Interactive reports -

The data engineers will create data pipelines to load data to OneLake either hourly or daily depending on the data source. The analytics engineers will create processes to ingest, transform, and load the data to the data store in the AnalyticsPOC workspace daily.

Whenever possible, the data engineers will use low-code tools for data ingestion. The choice of which data cleansing and transformation tools to use will be at the data engineers' discretion.

All the semantic models and reports in the Analytics POC workspace will use the data store as the sole data source.

Technical Requirements -

The data store must support the following:

Read access by using T-SQL or Python

Semi-structured and unstructured data

Row-level security (RLS) for users executing T-SQL queries

Files loaded by the data engineers to OneLake will be stored in the Parquet format and will meet Delta Lake specifications.

Data will be loaded without transformation in one area of the AnalyticsPOC data store. The data will then be cleansed, merged, and transformed into a dimensional model

The data load process must ensure that the raw and cleansed data is updated completely before populating the dimensional model

The dimensional model must contain a date dimension. There is no existing data source for the date dimension. The Litware fiscal year matches the calendar year. The date dimension must always contain dates from 2010 through the end of the current year.

The product pricing group logic must be maintained by the analytics engineers in a single location. The pricing group data must be made available in the data store for T-SQL

queries and in the default semantic model. The following logic must be used:

List prices that are less than or equal to 50 are in the low pricing group.

List prices that are greater than 50 and less than or equal to 1,000 are in the medium pricing group.

List prices that are greater than 1,000 are in the high pricing group.

Security Requirements -

Only Fabric administrators and the analytics team must be able to see the Fabric items created as part of the PoC.

Litware identifies the following security requirements for the Fabric items in the AnalyticsPOC workspace:

Fabric administrators will be the workspace administrators.

The data engineers must be able to read from and write to the data store. No access must be granted to datasets or reports.

The analytics engineers must be able to read from, write to, and create schemas in the data store. They also must be able to create and share semantic models with the data analysts and view and modify all reports in the workspace.

The data scientists must be able to read from the data store, but not write to it. They will access the data by using a Spark notebook

The data analysts must have read access to only the dimensional model objects in the data store. They also must have access to create Power BI reports by using the semantic models created by the analytics engineers.

The date dimension must be available to all users of the data store.

The principle of least privilege must be followed.

Both the default and custom semantic models must include only tables or views from the dimensional model in the data store. Litware already has the following Microsoft Entra security groups:

FabricAdmins: Fabric administrators

AnalyticsTeam: All the members of the analytics team

DataAnalysts: The data analysts on the analytics team

DataScientists: The data scientists on the analytics team

DataEngineers: The data engineers on the analytics team

AnalyticsEngineers: The analytics engineers on the analytics team

Report Requirements -

The data analysts must create a customer satisfaction report that meets the following requirements:

Enables a user to select a product to filter customer survey responses to only those who have purchased that product.

Displays the average overall satisfaction score of all the surveys submitted during the last 12 months up to a selected date.

Shows data as soon as the data is updated in the data store.

Ensures that the report and the semantic model only contain data from the current and previous year.

Ensures that the report respects any table-level security specified in the source data store.

Minimizes the execution time of report queries.

Question 1

You need to 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:	<div><div></div><div>▼</div><div>Build Reports on the default dataset Read All Apache Spark Read All SQL analytics endpoint data</div></div>
DataAnalysts:	<div><div></div><div>▼</div><div>Build Reports on the default dataset Read All Apache Spark Read All SQL analytics endpoint data</div></div>
DataScientists:	<div><div></div><div>▼</div><div>Build Reports on the default dataset Read All Apache Spark Read All SQL analytics endpoint data</div></div>

Answer:

Data Engineers: **Read all Apache Spark** - because they need to be able to work with Spark for Data curation.

Data Analysts: **Build Reports on the default dataset** - because they are report builders

Data Scientists: **Read SQL Endpoints** - They leverage curated data (by engineers) to do predictive analytics.

[Exam DP-600 topic 1 question 5 discussion - ExamTopics](#)

Question 2

You need to create a DAX measure to calculate the average overall satisfaction score.
How should you complete the DAX code? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

Answer Area

```
Rolling 12 Overall Satisfaction =  
VAR NumberOfMonths = 12  
VAR LastCurrentDate = MAX ( 'Date' [Date] )  
VAR Period = DATESINPERIOD ( 'Date' [Date], LastCurrentDate,  
- NumberOfMonths, MONTH )  
VAR Result =  
    CALCULATE (
```

▼

AVERAGE('Survey'[Response Value]),
AVERAGEA('Question'[Question Text]),
AVERAGEX(VALUE('Survey'[Customer Key]),

▼

LastCurrentDate,
NumberOfMonths,
Period,

```
        'Survey Question' [Question Title] = "Overall Satisfaction"  
    )  
RETURN  
    Result
```

Answer:

AVERAGE('Survey'[Response Value])

Period

Question 3

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]

FUNCTION
PROCEDURE
TABLE
VIEW

AS

SELECT ProductId,
ProductName,
ProductCategory,
ListPrice,

CASE
COALESCE
IIF
SET

WHEN ListPrice <= 50 THEN 'low'

WHEN(ListPrice > 50 AND ListPrice < 1000) THEN 'medium'
WHEN(ListPrice >= 50 AND ListPrice < 1000) THEN 'medium'
WHEN ListPrice BETWEEN 50 AND 1000) THEN 'medium'

WHEN ListPrice > 1000 THEN 'high'

END AS PricingGroup
FROM dbo.Products

Answer:

VIEW

CASE

WHEN ListPrice BETWEEN 50 AND 1000 THEN 'medium'

Question 4

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

- A. a stored procedure
- B. a pipeline that contains a KQL activity
- C. a Spark notebook
- D. a dataflow

Answer: D

Question 5

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

Question 6

You need to recommend a solution to prepare the tenant for the PoC.

Which two actions should you recommend performing from the Fabric Admin portal? Each correct answer presents part of the solution.

NOTE: Each correct answer is worth one point.

- A. Enable the Users can try Microsoft Fabric paid features option for the entire organization.
- B. Enable the Users can try Microsoft Fabric paid features option for specific security groups.
- C. Enable the Allow Azure Active Directory guest users to access Microsoft Fabric option for specific security groups.
- D. Enable the Users can create Fabric items option and exclude specific security groups.
- E. Enable the Users can create Fabric items option for specific security groups.

Answer: BE

Question 7

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:

▼

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

Mode:

▼

Direct Lake
DirectQuery
Import

Answer:

SSO/ Service Principal

Direct Lake

[Exam DP-600 topic 1 question 61 discussion - ExamTopics](#)

Question 8

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

Answer: AD

Question 9

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

Answer: D

Solution Challenge

Question 1 2 3 4

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

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.

- A. Solution: You use the following PySpark expression: `df.explain()`
- B. Solution: You use the following PySpark expression: `df.show()`
- C. Solution: You use the following PySpark expression: `df.summary()`
- D. Solution: You use the following PySpark expression: `df.explain().show()`

Answer: C

Question 5 6 7 8

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

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.

- A. Solution: You run the following Spark SQL statement: DESCRIBE HISTORY customer –
- B. Solution: You run the following Spark SQL statement: REFRESH TABLE customer –
- C. Solution: You run the following Spark SQL statement: EXPLAIN TABLE customer –
- D. Solution: You run the following Spark SQL statement: DESCRIBE DETAIL customer –

Answer: A

Question 9 10 11

Note: This question is part of a series of questions that present the same scenario. Each question in the series contains a unique solution that might meet the stated goals. Some question sets might have more than one correct solution, while others might not have a correct solution.

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.

You have a Fabric tenant that contains a semantic model named Model1.

You discover that the following query performs slowly against Model1.

```
1 EVALUATE
2   FILTER (
3     VALUES ( Customer[Customer Name] ),
4     CALCULATE ( COUNTROWS ( 'Order Item' ) ) > 0
5   )
6   ORDER BY Customer[Customer Name]
```

You need to reduce the execution time of the query.

- A. Solution: You replace line 4 by using the following code: ISEMPTY (RELATEDTABLE ('Order Item'))
- B. Solution: You replace line 4 by using the following code: NOT ISEMPTY (CALCULATETABLE ('Order Item'))
- C. Solution: You replace line 4 by using the following code: CALCULATE (COUNTROWS ('Order Item')) >= 0

Answer: B

Common Question

Question 1

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?

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

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

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

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

Answer: A

## Question 2

HOTSPOT -

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 = OVER (PARTITION BY CustomerID
 ORDER BY LastUpdated DESC)
 FROM [LakehousePOC].[dbo].[CustomerChanges]
)
 SELECT CustomerID, CustomerName, PostalCode, LastUpdated
 FROM CUSTOMERBASE

```

Answer:

Box 1: ROW\_NUMBER()

Box 2: WHERE X = 1

## Question 3

HOTSPOT -

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"/> |

Answer:

 **Momoanwar** Highly Voted 3 months ago

I think : No yes no

To display summary its describe() or sumarry()

   upvoted 21 times

 **TashaP** Highly Voted 2 months, 3 weeks ago

Answer is No, Yes, Yes

The code is displaying a summary of the data frame using the visualizations. Point 3 is a technicality and I think to some extent depends on your interpretation of the question. I am going with Yes.

   upvoted 8 times

 **dp600** 3 weeks, 5 days ago

I understand your reasoning but i would go with No in 3, when I think in a summary of a dataframe I'm imagining something different.

   upvoted 2 times

[Exam DP-600 topic 1 question 12 discussion - ExamTopics](#)

## Question 4

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. Read Table3.
- B. Update the data Table3.
- C. Read Table2.
- D. Update the data in Table1.

Answer: A

## Question 5

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 PowerQuery step to identify the maximum values for the numeric columns.

Which function should you include in the step?

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

ET: B

Gemini: B

Copilot: B

ET User:



[Exam DP-600 topic 1 question 14 discussion - ExamTopics](#)

## Question 6

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
- C. Spark SQL
- D. PySpark

Answer: CD



## Question 7

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

## Question 8

You have a Fabric tenant that contains a Microsoft Power BI report named Report1. Report1 includes a Python visual.

Data displayed by the visual is grouped automatically and duplicate rows are NOT displayed.

You need all rows to appear in the visual.

What should you do?

- A. Reference the columns in the Python code by index.
- B. Modify the Sort Column By property for all columns.
- C. Add a unique field to each row.
- D. Modify the Summarize By property for all columns.

ET: A

Gemini: C

Copilot: C

ET User: 

[Exam DP-600 topic 1 question 17 discussion - ExamTopics](#)

## Question 9

DRAG DROP -

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.

| Values    | Answer Area                                                                                                    |
|-----------|----------------------------------------------------------------------------------------------------------------|
| DEFINE    | <input type="text"/>                                                                                           |
| EVALUATE  | VAR_SalesSince =<br>DATE (2023, 12, 01)                                                                        |
| FILTER    | <input type="text"/>                                                                                           |
| SUMMARIZE | FILTER (<br><input type="text"/> (Store, Store[Name], Store[OpenDate]),<br>Store[OpenDate] >= _SalesSince<br>) |
| TABLE     |                                                                                                                |

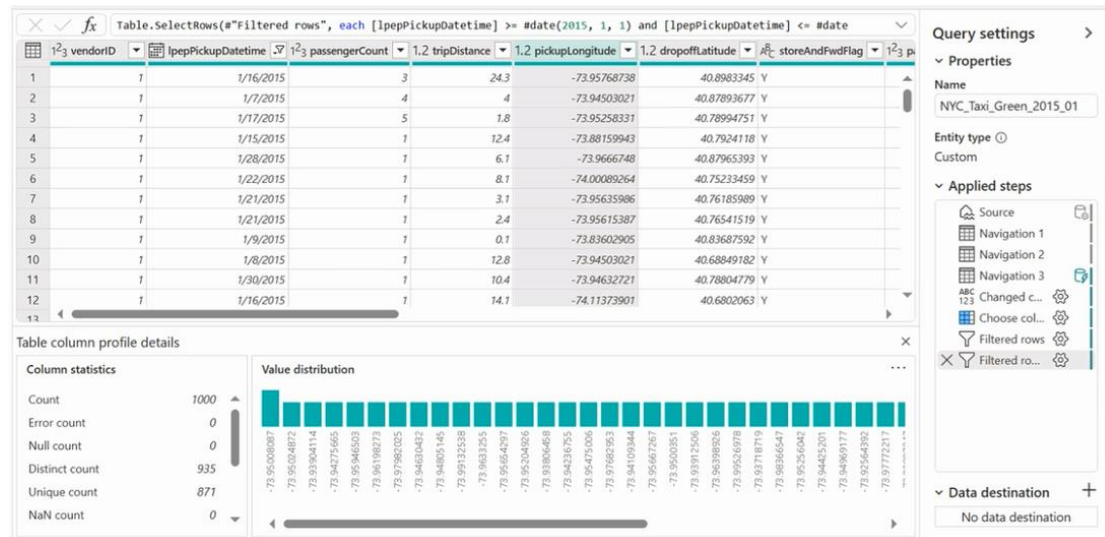
Answer:

- 1) DEFINE
- 2) EVALUATE
- 3) SUMMARIZE

## Question 10

You have a Fabric workspace named Workspace1 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: A

## Question 11

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 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

## Question 12

You have a Fabric tenant that contains a workspace named Workspace1. Workspace1 is assigned to a Fabric capacity.

You need to recommend a solution to provide users with the ability to create and publish custom Direct Lake semantic models by using external tools. The solution must follow the principle of least privilege.

Which three actions in the Fabric Admin portal should you include in the recommendation?

Each correct answer presents part of the solution.

NOTE: Each correct answer is worth one point.

- A. From the Tenant settings, set Allow XMLA Endpoints and Analyze in Excel with on-premises datasets to Enabled.
- B. From the Tenant settings, set Allow Azure Active Directory guest users to access Microsoft Fabric to Enabled.
- C. From the Tenant settings, select Users can edit data model 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.

ET ACD

Gemini: CDE

Copilot: ADE

ET User:



[Exam DP-600 topic 1 question 21 discussion - ExamTopics](#)

## Question 13

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

ET: B

Gemini: C

Copilot: B

ET User:

A (78%)

C (22%)

[Exam DP-600 topic 1 question 22 discussion - ExamTopics](#)

## Question 14

HOTSPOT -

You have a Fabric tenant that contains a warehouse named Warehouse1. Warehouse1 contains three schemas named schemaA, schemaB, and schemaC.

You need to ensure that a user named User1 can truncate tables in schemaA only.

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

|       |                                          |    |                                                                      |           |
|-------|------------------------------------------|----|----------------------------------------------------------------------|-----------|
| GRANT | <div><div></div><div>▼</div></div>       | ON | <div><div></div><div>▼</div></div>                                   | TO User1; |
|       | <div>ALTER<br/>CONNECT<br/>EXECUTE</div> |    | <div>DATABASE::schemaA<br/>OBJECT::schemaA<br/>SCHEMA::schemaA</div> |           |

Answer:

### Answer Area

|       |                                          |    |                                                                      |           |
|-------|------------------------------------------|----|----------------------------------------------------------------------|-----------|
| GRANT | <div><div></div><div>▼</div></div>       | ON | <div><div></div><div>▼</div></div>                                   | TO User1; |
|       | <div>ALTER<br/>CONNECT<br/>EXECUTE</div> |    | <div>DATABASE::schemaA<br/>OBJECT::schemaA<br/>SCHEMA::schemaA</div> |           |



## Question 15

You plan to deploy Microsoft Power BI items by using Fabric deployment pipelines. You have a deployment pipeline that contains three stages named Development, Test, and Production. A workspace is assigned to each stage.

You need to provide Power BI developers with access to the pipeline. The solution must meet the following requirements:

Ensure that the developers can deploy items to the workspaces for Development and Test.

Prevent the developers from deploying items to the workspace for Production.

Follow the principle of least privilege.

Which three levels of access should you assign to the developers? Each correct answer presents part of the solution.

NOTE: Each correct answer is worth one point.

- A. Build permission to the production semantic models
- B. Admin access to the deployment pipeline
- C. Viewer access to the Development and Test workspaces
- D. Viewer access to the Production workspace
- E. Contributor access to the Development and Test workspaces
- F. Contributor access to the Production workspace

ET: ADE

Gemini: CE

Copilot: BDE

ET User:



[Exam DP-600 topic 1 question 24 discussion - ExamTopics](#)

## Question 16

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

ET: BD

Gemini: BC

Copilot: AC

ET User:

AB (79%)

AC (17%) 4%

[Exam DP-600 topic 1 question 25 discussion - ExamTopics](#)

## Question 17

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 Model1 that uses the CSV files as a data source. You configure incremental refresh for Model1 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 delta type of the column used to partition the data has changed.

ET: D

Gemini: E

Copilot: D

ET User:



[Exam DP-600 topic 1 question 26 discussion - ExamTopics](#)

## Question 18

You have a Fabric tenant that uses a Microsoft Power BI Premium capacity.  
You need to enable scale-out for a semantic model.

What should you do first?

- A. At the semantic model level, set Large dataset storage format to Off.
- B. At the tenant level, set Create and use Metrics to Enabled.
- C. At the semantic model level, set Large dataset storage format to On.
- D. At the tenant level, set Data Activator to Enabled.

Answer: C

## Question 19

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

ET: D

Gemini: A

Copilot: A

ET User:



[Exam DP-600 topic 1 question 28 discussion - ExamTopics](#)

## Question 20

You have a Fabric tenant that contains a complex semantic model. The model is based on a star schema and contains many tables, including a fact table named Sales.

You need to create a diagram of the model. The diagram must contain only the Sales table and related tables.

What should you use from Microsoft Power BI Desktop?

- A. data categories
- B. Data view
- C. Model view
- D. DAX query view

Answer: C

## Question 21

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.

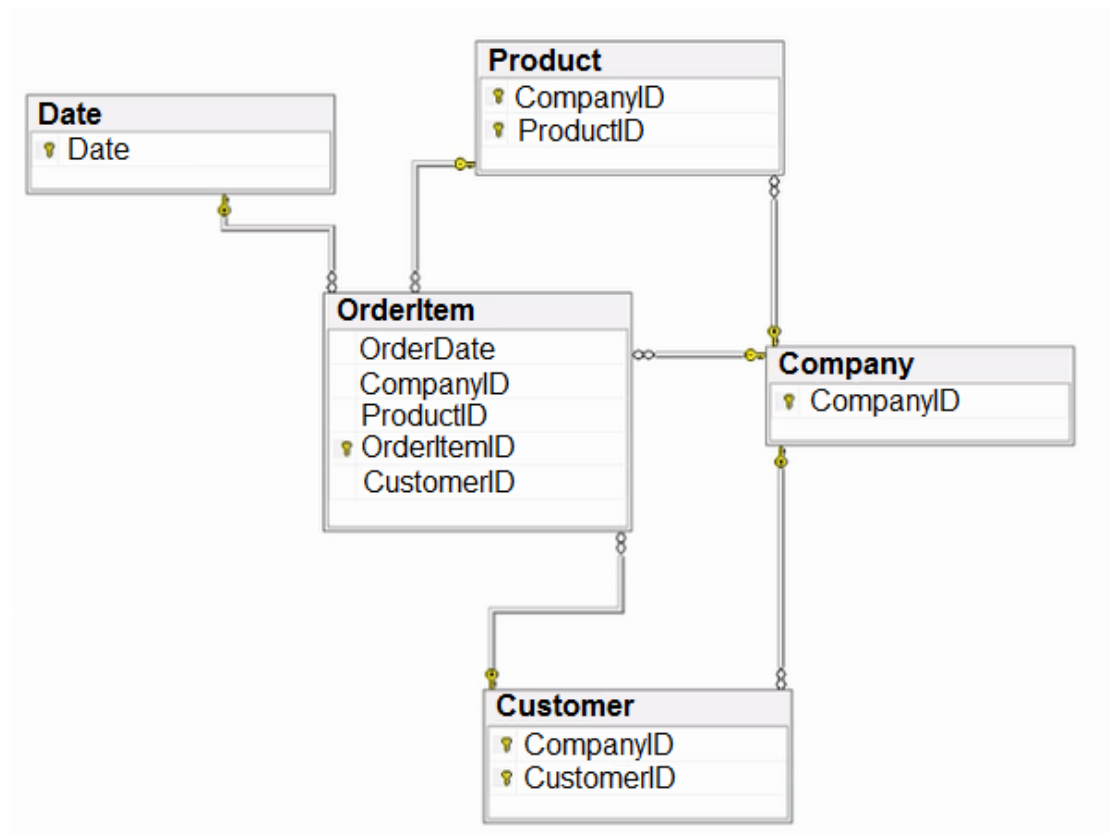
- A. Use the Analyze in Excel feature.
- B. Use the Vertipaq Analyzer tool.
- C. Query the \$System.DISCOVER\_STORAGE\_TABLE\_COLUMN\_SEGMENTS dynamic management view (DMV).
- D. Query the DISCOVER\_MEMORYGRANT dynamic management view (DMV).

Answer: BC

## Question 22

HOTSPOT -

You have the source data model shown in the following exhibit.



The primary keys of the tables are indicated by a key symbol beside the columns involved in each key.

You need to create a dimensional data model that will enable the analysis of order items by date, product, and customer.

What should you include in the solution? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

### Answer Area

The relationship between OrderItem and Product must be based on:

|                                                             |   |
|-------------------------------------------------------------|---|
|                                                             | ▼ |
| The ProductID column                                        |   |
| Both the CompanyID and the ProductID columns                |   |
| A new key that combines the CompanyID and ProductID columns |   |

The Company entity must be:

|                                                     |   |
|-----------------------------------------------------|---|
|                                                     | ▼ |
| Omitted                                             |   |
| Denormalized into the Product entity only           |   |
| Denormalized into the Customer and Product entities |   |

ET Answer: 争议



### Answer Area

The relationship between OrderItem and Product must be based on:

Both the CompanyID and the ProductID columns

The Company entity must be:



- Omitted
- Denormalized into the Product entity only
- Denormalized into the Customer and Product entities

  **thuss** Highly Voted  2 months, 3 weeks ago

I think what they are trying to get at is that you denormalize the company information into both tables so that you do not need the CompanyID anymore, which would then make the relationship between OrderItem and Product only based on ProductID. Hence I think A and C.

But honestly that whole model is just weird.

upvoted 15 times

  **d47320d** 4 weeks, 1 day ago

Agree up to the denormalization part.

But that does not imply that the relationship between OrderItem and Product should only be based on ProductID. The Product table has two columns as a primary key, ProductID and CompanyID.

After denormalization, a two columns join should be performed in order to establish the proper relationships and proper analysis.

So correct answers are B, C i.e. the given one.

   upvoted 2 times

  **metiii** Highly Voted  2 months, 2 weeks ago

The given answer is correct.

The company entity should be denormalized into both product and customer tables. Company ID is part of the customer table and you cannot just denormalize it into the product table.



and for the relationship between OrderItem and Product, you cannot just reference a part of the primary key. This is not how referencing works, your relationship would be wrong, if you only use product id and a product id is repeated for two different company IDs your relationship would consider both of them as the same product.

   upvoted 7 times

  2dc6125 **Most Recent**  1 week ago

The requirement for Analysis "date, product, and customer" so we only need productId and nothing about the company so Omitted. I'm interested on why many answers go for including company in the model? would love to hear about that.

   upvoted 1 times

  **2dc6125** 2 weeks ago

The request about creating dimensional data model don't include Company so in my opinion the model can be achieved by A (ProductID Column) and A(Omitted) so no need to include it.

upvoted 1 times

  **stilferx** 2 weeks, 1 day ago

IMHO, B & C

[Exam DP-600 topic 1 question 31 discussion - ExamTopics](#)

## Question 23

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 Model1 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

## Question 24

You have a Fabric tenant that contains a semantic model.

You need to prevent report creators from populating visuals by using implicit measures.

What are two tools that you can use to achieve the goal? Each correct answer presents a complete solution.

NOTE: Each correct answer is worth one point.

- A. Microsoft Power BI Desktop
- B. Tabular Editor
- C. Microsoft SQL Server Management Studio (SSMS)
- D. DAX Studio

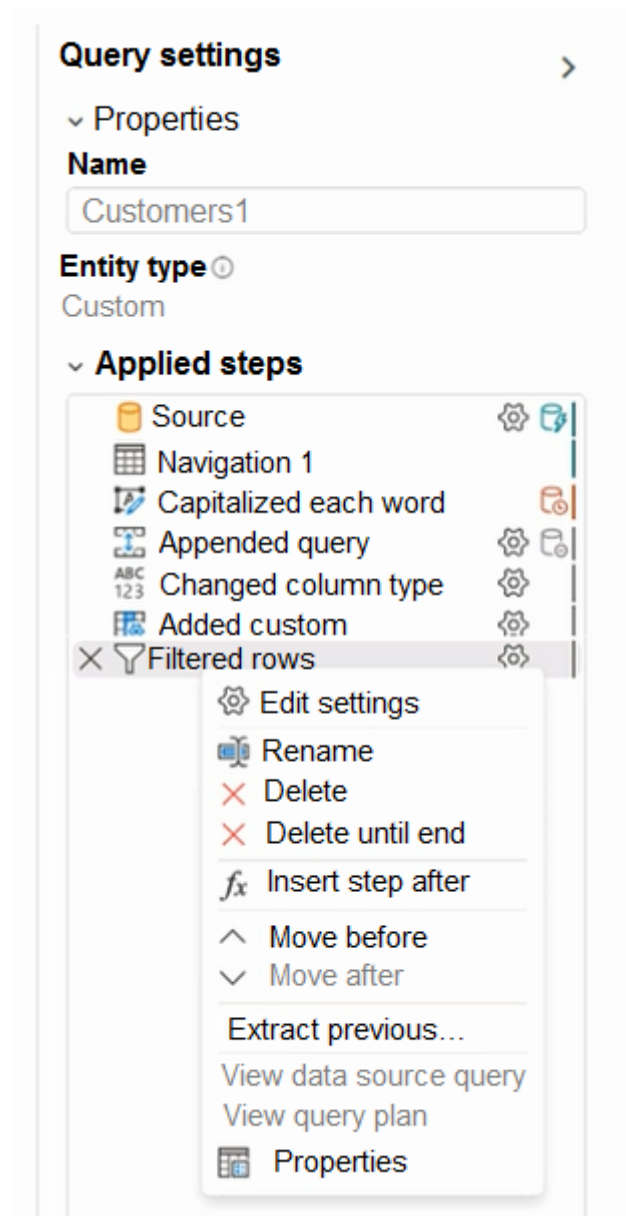
Answer: AB

## Question 25

HOTSPOT -

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.

▼

All  
None  
Some

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

▼

each lakehouse's query engine  
the Microsoft Power Query engine  
the source lakehouse query engine

Answer:

## Answer Area

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

▼

All  
None  
Some

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

▼

each lakehouse's query engine  
the Microsoft Power Query engine  
the source lakehouse query engine

## Question 26

You have a Fabric tenant that contains a lakehouse named Lakehouse1. Lakehouse1 contains a table named Table1.

You are creating a new data pipeline.

You plan to copy external data to Table1. The schema of the external data changes regularly.

You need the copy operation to meet the following requirements:

Replace Table1 with the schema of the external data.

Replace all the data in Table1 with the rows in the external data.

You add a Copy data activity to the pipeline.

What should you do for the Copy data activity?

- A. From the Source tab, add additional columns.
- B. From the Destination tab, set Table action to Overwrite.
- C. From the Settings tab, select Enable staging.
- D. From the Source tab, select Enable partition discovery.
- E. From the Source tab, select Recursively.

Answer: B

## Question 27

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.

ET: CD

Gemini: BD

Copilot: BD

ET User:

DE (57%)

BD (43%)

[Exam DP-600 topic 1 question 36 discussion - ExamTopics](#)

## Question 28

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: A



## Question 29

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 tabs, set Mode to Overwrite.

Answer: D

## Question 30

HOTSPOT -

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

Answer:

### Answer Area

| Statements                                                                                               | Yes                              | No                               |
|----------------------------------------------------------------------------------------------------------|----------------------------------|----------------------------------|
| A replica of <code>dbo.Sales</code> is created in the test schema by copying the metadata only.          | <input checked="" type="radio"/> | <input type="radio"/>            |
| Additional schema changes to <code>dbo.FactSales</code> will also apply to <code>test.FactSales</code> . | <input type="radio"/>            | <input checked="" type="radio"/> |
| Additional schema changes to <code>dbo.FactSales</code> will also apply to <code>test.FactSales</code> . | <input type="radio"/>            | <input checked="" type="radio"/> |

## Question 31

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

## Question 32

HOTSPOT -

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"/> |

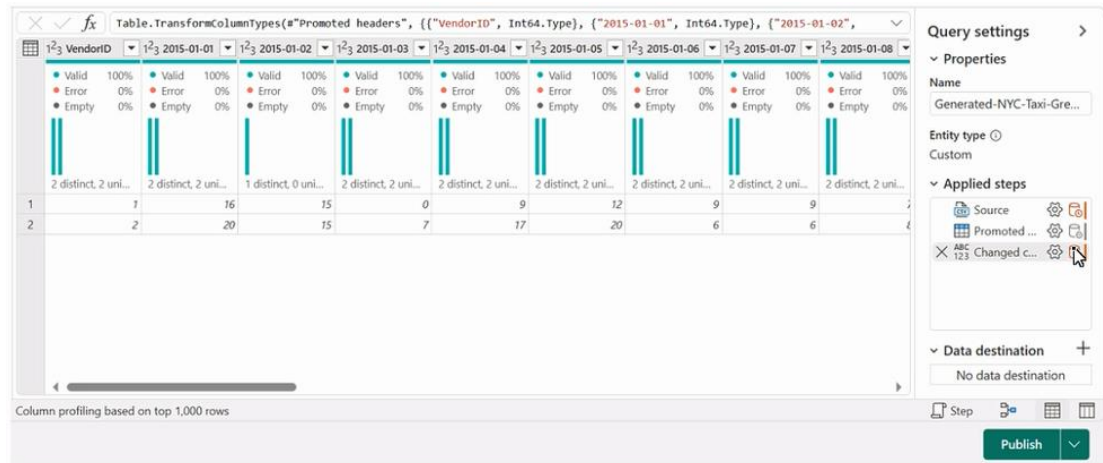
ET User: YYY

ET Answer: YYN

[Exam DP-600 topic 1 question 41 discussion - ExamTopics](#)

## Question 33

You have a Fabric workspace named Workspace1 that contains a data flow named Dataflow1 contains a query that returns the data shown in the following exhibit.



You need to transform the data 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: C

## Question 34

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

## Question 35

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: D

## Question 36

HOTSPOT -

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"/> |

ET Answer: YNY 争议

ET User: NYY/NNY

[Exam DP-600 topic 1 question 45 discussion - ExamTopics](#)



## Question 37

You have a Fabric tenant that contains a warehouse.

A user discovers that a report that usually takes two minutes to render has been running for 45 minutes and has still not rendered.

You need to identify what is preventing the report query from completing.

Which dynamic management view (DMV) should you use?

- A. sys.dm\_exec\_requests
- B. sys.dm\_exec\_sessions
- C. sys.dm\_exec\_connections
- D. sys.dm\_pdw\_exec\_requests

Answer: A

## Question 38

DRAG DROP -

You are creating a data flow in Fabric to ingest data from an Azure SQL database by using a T-SQL statement.

You need to ensure that any foldable Power Query transformation steps are processed by the Microsoft SQL Server engine.

How should you complete the code? To answer, drag the appropriate values to the correct targets. Each value may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.

NOTE: Each correct selection is worth one point.

| Values        | Answer Area                                                                                                                                                                                                                                              |
|---------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| EnableFolding | let                                                                                                                                                                                                                                                      |
| NativeQuery   | Source = Sql.Databases(<br>"server.database.windows.net"<br>),<br>Database = Source{[Name = "db"]}[Data],<br>Query = [ ] . [ ] (<br>Database,<br>" SELECT * FROM customer WHERE country IN ('USA', 'UK')",<br>null,<br>[ [ ] = true]<br>)<br>in<br>Query |
| Optimize      |                                                                                                                                                                                                                                                          |
| Record        |                                                                                                                                                                                                                                                          |
| StopFolding   |                                                                                                                                                                                                                                                          |
| Table         |                                                                                                                                                                                                                                                          |
| Value         |                                                                                                                                                                                                                                                          |

Answer:

**Answer Area**

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

## Question 39

DRAG DROP -

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.

Set the optimizeWrite  
table setting.

Run the VACUUM  
command on a schedule.

Set the autoCompact  
SparkSession setting.

Run the OPTIMIZE  
command on a schedule.

Set the parallelDelete  
SparkSession setting.

Remove the files:

Combine the files:

Answer:

## Answer Area

Remove the files:

Run the VACUUM  
command on a schedule.

Combine the files:

Run the OPTIMIZE  
command on a schedule.

## Question 40

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

## Question 41

HOTSPOT -

You have a Fabric workspace named Workspace1 and an Azure Data Lake Storage Gen2 account named storage1. Workspace1 contains a lakehouse named Lakehouse1.

You need to create a shortcut to storage1 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: 

|       |   |
|-------|---|
|       | ▼ |
| abfs  |   |
| abfss |   |
| https |   |

Endpoint: 

|      |   |
|------|---|
|      | ▼ |
| blob |   |
| dfs  |   |
| file |   |

Answer:

Box 1: abfss/**https**

Box 2: dfs

[Exam DP-600 topic 1 question 50 discussion - ExamTopics](#)

[Create an Azure Data Lake Storage Gen2 shortcut - Microsoft Fabric | Microsoft Learn](#)

## Question 42

You are analyzing customer purchases in a Fabric notebook by using PySpark.

You have the following DataFrames:

transactions: Contains five columns named transaction\_id, customer\_id, product\_id, amount, and date and has 10 million rows, with each row representing a transaction. customers: Contains customer details in 1,000 rows and three columns named customer\_id, name, and country.

You need to join the DataFrames on the customer\_id column. The solution must minimize data shuffling.

You write the following code.

```
from pyspark.sql import functions as F
results =
```

Which code should you run to populate the results DataFrame?

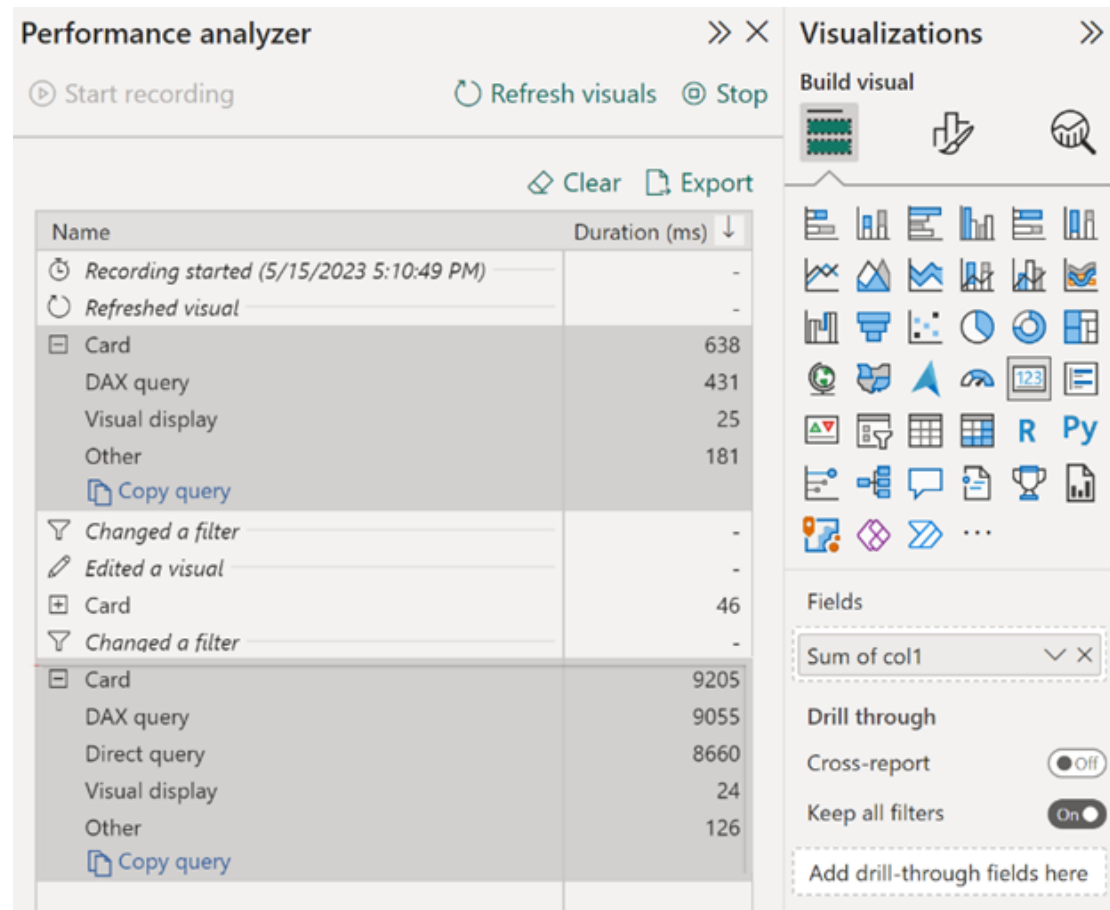
- A. `transactions.join(F.broadcast(customers), transactions.customer_id == customers.customer_id)`
- B. `transactions.join(customers, transactions.customer_id == customers.customer_id).distinct()`
- C. `transactions.join(customers, transactions.customer_id == customers.customer_id)`
- D. `transactions.crossJoin(customers).where(transactions.customer_id == customers.customer_id)`

Answer: A

## Question 43

HOTSPOT -

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

▼

Automatic  
DirectLakeOnly  
DirectQueryOnly

The query for the table visual is executed by using

▼

the composite model  
Direct Lake  
Direct Query

Answer: Automatic and Direct Lake



## Question 44

HOTSPOT -

You have a Fabric tenant that contains a lakehouse named Lakehouse1. Lakehouse1 contains a table named Nyctaxi\_raw. Nyctaxi\_raw contains the following table:

| 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= ("pickupDate", df['tpepPickupDateTime'])
```

df.columns

df.select

df.withColumn

df.withColumnsRenamed

.alias('date')

.cast('date')

.cast('pickupDate')

.getfield('date')

.filter("fareAmount > 0 AND fareAmount < 100")

.filter(col("fareAmount").contains("1..100"))

.when(df.fareAmount > 0 AND fareAmount < 100)

.where(df.fareAmount.isin([1,100]))

Answer: .withcolumn

.cast(date)

.filter(fareamount >0 and <100)

## Question 45

HOTSPOT -

You have a Fabric tenant.

You need to configure OneLake security for users shown in the following table.

| Name  | Required access                 |
|-------|---------------------------------|
| User1 | Read all the Spark data.        |
| User2 | Read all the SQL endpoint data. |

The solution must follow the principle of least privilege.

Which permission should you assign to each user? To answer, select the appropriate options in the answer area.

NOTE: Each correct selection is worth one point.

### Answer Area

User1:   
Read  
ReadAll  
ReadData

User2:   
Read  
ReadAll  
ReadData

Answer:

### Answer Area

User1:   
Read  
**ReadAll**  
ReadData

User2:   
Read  
ReadAll  
**ReadData**

## Question 46

You have an Azure Repos repository named Repo1 and a Fabric-enabled Microsoft Power BI Premium capacity. The capacity contains two workspaces named Workspace1 and Workspace2. Git integration is enabled at the workspace level.

You plan to use Microsoft Power BI Desktop and Workspace1 to make version-controlled changes to a semantic model stored in Repo1. The changes will be built and deployed to Workspace2 by using Azure Pipelines.

You need to ensure that report and semantic model definitions are saved as individual text files in a folder hierarchy. The solution must minimize development and maintenance effort.

In which file format should you save the changes?

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

Answer: A

## Question 47

DRAG DROP -

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.




| Triggers                             | Answer Area                                  |
|--------------------------------------|----------------------------------------------|
| A pipeline Copy activity             | Orchestration pipeline: <input type="text"/> |
| A pipeline Dataflow activity         | Bronze layer: <input type="text"/>           |
| A pipeline Stored procedure activity | Silver layer: <input type="text"/>           |
| A schedule                           | Gold layer: <input type="text"/>             |
| A Spark job definition               |                                              |
| An Invoke pipeline activity          |                                              |


Answer:

| Answer Area             |                                      |
|-------------------------|--------------------------------------|
| Orchestration pipeline: | A schedule                           |
| Bronze layer:           | A pipeline Copy activity             |
| Silver layer:           | A pipeline Dataflow activity         |
| Gold layer:             | A pipeline Stored procedure activity |

 **c8f5bdr** Highly Voted 3 weeks, 3 days ago




Bronze Layer answer is wrong, the Bronze Layer already has pipeline defined with it so you need an invoke pipeline activity to call them

   upvoted 7 times

 **Nefirs** Highly Voted 1 month ago

I'd say answer is correct.

However, IMO there is no suitable item for "orchestration pipeline". "A schedule" seems the most appropriate. But I'd imagine "data pipeline" would be a better option since one would use an data pipeline to orchestrate the different activities.

   upvoted 5 times

[Exam DP-600 topic 1 question 70 discussion - ExamTopics](#)

## Question 48

DRAG DROP -

You are building a solution by using a Fabric notebook.

You have a Spark DataFrame assigned to a variable named df. The DataFrame returns four columns.

You need to change the data type of a string column named Age to integer. The solution must return a DataFrame that includes all the columns.

How should you complete the code? To answer, drag the appropriate values to the correct targets. Each value may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.

NOTE: Each correct selection is worth one point.

| Values                | Answer Area                                                              |
|-----------------------|--------------------------------------------------------------------------|
| <div>cast</div>       | df. <div></div> ("age", <div></div> ("age"). <div></div> ("int")).show() |
| <div>col</div>        |                                                                          |
| <div>get</div>        |                                                                          |
| <div>select</div>     |                                                                          |
| <div>selectExpr</div> |                                                                          |
| <div>transform</div>  |                                                                          |
| <div>withColumn</div> |                                                                          |

Answer:

| Answer Area                                                                               |
|-------------------------------------------------------------------------------------------|
| df. <div>withColumn</div> ("age", <div>col</div> ("age"). <div>cast</div> ("int")).show() |

## Question 49

HOTSPOT -

You have an Azure Data Lake Storage Gen2 account named storage1 that contains a Parquet file named sales.parquet.

You have a Fabric tenant that contains a workspace named Workspace1.

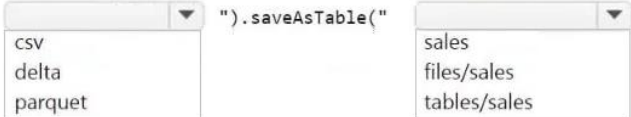
Using a notebook in Workspace1, you need to load the content of the file to the default lakehouse. The solution must ensure that the content will display automatically as a table named Sales in Lakehouse explorer.

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.parquet("abfss://fs1@storage1.dfs.core.windows.net/files/sales.parquet")
df.write.mode("overwrite").format(" ").saveAsTable(" ")
```






Answer: 争议


### Answer Area




```
df = spark.read.parquet("abfss://fs1@storage1.dfs.core.windows.net/files/sales.parquet")
df.write.mode("overwrite").format(" ").saveAsTable(" ")
```







 **taphyoe**  1 month, 2 weeks ago

delta  
sales  
   upvoted 16 times




 **neoverma** 1 month, 2 weeks ago

why? please provide some explanation  
   upvoted 1 times

 **harshalt10** 1 month, 2 weeks ago

I think it should only be sales because if saveastable is used, the argument should only be table name. Link: <https://learn.microsoft.com/en-us/fabric/data-engineering/lakehouse-notebook-load-data>.  
   upvoted 9 times

 **dp600dataroots**  1 month, 2 weeks ago

delta  
tables/sales  
   upvoted 7 times

[Exam DP-600 topic 1 question 72 discussion - ExamTopics](#)

## Question 50

You have a Fabric workspace named Workspace1 that contains a lakehouse named Lakehouse1.

In Workspace1, you create a data pipeline named Pipeline1.

You have CSV files stored in an Azure Storage account.

You need to add an activity to Pipeline1 that will copy data from the CSV files to Lakehouse1. The activity must support Power Query M formula language expressions.

Which type of activity should you add?

- A. Dataflow
- B. Notebook
- C. Script
- D. Copy data

Answer: A

## Question 51

HOTSPOT -

You have a Fabric tenant that contains lakehouse named Lakehouse1. Lakehouse1 contains a Delta table with eight columns.

You receive new data that contains the same eight columns and two additional columns.

You create a Spark DataFrame and assign the DataFrame to a variable named df. The DataFrame contains the new data.

You need to add the new data to the Delta table to meet the following requirements:

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

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

NOTE: Each correct selection is worth one point.

### Answer Area

```
df.write.format("delta").mode
```





(  
"append")  
("error")  
("errorifexists")  
("ignore")  
("overwrite")


```
.option
```

(  
"mergeSchema","false")  
("mergeSchema","true")  
("overwriteSchema","false")  
("overwriteSchema","true")

```
.saveAsTable("table")
```

Answer:

 **SJSull** Highly Voted 1 month, 2 weeks ago  
append, ("mergeschema", "true") - schema evolution, same columns and 2 added new columns  
   upvoted 20 times




 **9878eb9** Highly Voted 1 month, 1 week ago  
- append  
- ("mergeschema", "true")

<https://learn.microsoft.com/en-us/azure/databricks/delta/update-schema>

Add columns with automatic schema update

Columns that are present in the DataFrame but missing from the table are automatically added as part of a write transaction when:

write or writeStream have .option("mergeSchema", "true")  
spark.databricks.delta.schema.autoMerge.enabled is true

   upvoted 8 times



## Question 52

HOTSPOT -

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       |
| ProductID  | Integer        | 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 the OrderID value                                                                                            |
| CustomerID      | Returns the CustomerID value                                                                                         |
| ProductID       | Returns the ProductID value                                                                                          |
| ChargedQuantity | Returns the Weight value if present, otherwise returns the Quantity value if present, otherwise returns a value of 1 |
| OrderPrice      | Returns the lowest value from ListPrice and SalePrice<br>Ignores NULL values in SalePrice                            |

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, ProductID,
```

CHOOSE  
COALESCE  
IIF

```
(Weight, Quantity, 1) AS ChargedQuantity,
```

COALESCE  
LEAST  
MIN


```
(ListPrice, SalePrice) AS OrderPrice
```

```
FROM Sales.Orders;
```

Answer:

### Answer Area

```
SELECT OrderID, CustomerID, ProductID,
 (Weight, Quantity, 1) AS ChargedQuantity,
 CHOOSE
 COALESCE
 IIF
 (ListPrice, SalePrice) AS OrderPrice
 COALESCE
 LEAST
 MIN
FROM Sales.Orders;
```


 **Nefirs** Highly Voted 1 month ago

IMHO answer is correct:

<https://learn.microsoft.com/de-de/sql/t-sql/language-elements/coalesce-transact-sql?view=sql-server-ver16>

<https://learn.microsoft.com/de-de/sql/t-sql/functions/logical-functions-least-transact-sql?view=sql-server-ver16>

   upvoted 9 times


 **72bd3bc** Most Recent 3 days, 8 hours ago

strange question. In the second box coalesce is not correct because it doesn't choose the smallest. Least also has problems - it doesn't ignore null: If one or more arguments aren't NULL, then NULL arguments are ignored during comparison. If all arguments are NULL, then LEAST returns NULL. <https://learn.microsoft.com/en-us/sql/t-sql/functions/logical-functions-least-transact-sql?view=sql-server-ver16>




MIN compares all the rows in one column and choose the smallest, so it is also the wrong choice.

I feel confused.

   upvoted 1 times

 **serch\_engine** 3 days, 22 hours ago

IMHO answer is correct: COALESCE, LEAST

   upvoted 1 times

[Exam DP-600 topic 1 question 75 discussion - ExamTopics](#)

## Question 53

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 in both tables.

Which type of join should you use?

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

Answer: B

## Question 54

You have a Fabric tenant that contains a lakehouse named Lakehouse1. Lakehouse1 contains a Delta table that has one million Parquet files.

You need to remove files that were NOT referenced by the table during the past 30 days.

The solution must ensure that the transaction log remains consistent, and the ACID properties of the table are maintained.

What should you do?

- A. From OneLake file explorer, delete the files.
- B. Run the OPTIMIZE command and specify the Z-order parameter.
- C. Run the OPTIMIZE command and specify the V-order parameter.
- D. Run the VACUUM command.

Answer: D

## Question 55

DRAG DROP -

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 |   | Answer Area                     |
|-----------|---|---------------------------------|
| Type 0    |   | Customers: <input type="text"/> |
| Type 1    | ● | Products: <input type="text"/>  |
| Type 2    | ● |                                 |
| Type 3    | ● |                                 |

Answer:

| Answer Area |                                     |
|-------------|-------------------------------------|
| Customers:  | <input type="text" value="Type 2"/> |
| Products:   | <input type="text" value="Type 1"/> |

## Question 56

You have a Fabric workspace named Workspace1 and an Azure SQL database.

You plan to create a dataflow that will read data from the database, and then transform the data by performing an inner join.

You need to ignore spaces in the values when performing the inner join. The solution must minimize development effort.

What should you do?

- A. Append the queries by using fuzzy matching.
- B. Merge the queries by using fuzzy matching.
- C. Append the queries by using a lookup table.
- D. Merge the queries by using a lookup table.

Answer: B

## Question 57

You have a Fabric tenant that contains a warehouse named Warehouse1. Warehouse1 contains two schemas name schema1 and schema2 and a table named schema1.city. You need to make a copy of schema1.city in schema2. The solution must minimize the copying of data.

Which T-SQL statement should you run?

- A. `INSERT INTO schema2.city SELECT * FROM schema1.city;`
- B. `SELECT * INTO schema2.city FROM schema1.city;`
- C. `CREATE TABLE schema2.city AS CLONE OF schema1.city;`
- D. `CREATE TABLE schema2.city AS SELECT * FROM schema1.city;`

Answer: C

## Question 58

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?

- A. the SQL analytics endpoint settings
- B. the semantic model settings
- C. the workspace settings
- D. the Lakehouse1 settings

Answer: A



## Question 59

HOTSPOT -

You have a Fabric tenant that contains a workspace named Workspace1. Workspace1 contains a lakehouse named Lakehouse1 and a warehouse named Warehouse1.

You need to create a new table in Warehouse1 named POSCustomers by querying the customer table in Lakehouse1.

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 TABLE dbo.POSCustomers  
CREATE TABLE dbo.POSCustomers AS CLONE OF  
CREATE TABLE dbo.POSCustomers AS SELECT

DISTINCT customerid,  
customer,  
postalcode,  
category

FROM dbo.Customer  
FROM dbo.POSCustomers  
FROM lakehouse1.dbo.customer

Answer:

### Answer Area

CREATE TABLE dbo.POSCustomers  
CREATE TABLE dbo.POSCustomers AS CLONE OF  
CREATE TABLE dbo.POSCustomers AS SELECT

DISTINCT customerid,  
customer,  
postalcode,  
category

FROM dbo.Customer  
FROM dbo.POSCustomers  
FROM lakehouse1.dbo.customer

## Question 60

You have a Fabric tenant.

You are creating an Azure Data Factory pipeline.

You have a stored procedure that returns the number of active customers and their average sales for the current month.

You need to add an activity that will execute the stored procedure in a warehouse. The returned values must be available to the downstream activities of the pipeline.

Which type of activity should you add?

- A. Switch
- B. Copy data
- C. Append variable
- D. Lookup

Answer: D

## Question 61

You have a Fabric tenant that contains two workspaces named Workspace1 and Workspace2. Workspace1 contains a lakehouse named Lakehouse1. Workspace2 contains a lakehouse named Lakehouse2. Lakehouse1 contains a table named `dbo.Sales`. Lakehouse2 contains a table named `dbo.Customers`.

You need to ensure that you can write queries that reference both `dbo.Sales` and `dbo.Customers` in the same SQL query without making additional copies of the tables.

What should you use?

- A. a shortcut
- B. a dataflow
- C. a view
- D. a managed table

Answer: A

## Question 62

You have a Fabric tenant that contains a warehouse.

You are designing a star schema model that will contain a customer dimension. The customer dimension table will be a Type 2 slowly changing dimension (SCD).

You need to recommend which columns to add to the table. The columns must NOT already exist in the source.

Which three types of columns should you recommend? Each correct answer presents part of the solution.

NOTE: Each correct answer is worth one point.

- A. a foreign key
- B. a natural key
- C. an effective end date and time
- D. a surrogate key
- E. an effective start date and time

Answer: CDE

## Question 63

You have a Fabric tenant.

You plan to create a data pipeline named Pipeline1. Pipeline1 will include two activities that will execute in sequence.

You need to ensure that a failure of the first activity will NOT block the second activity.

Which conditional path should you configure between the first activity and the second activity?

- A. Upon Failure
- B. Upon Completion
- C. Upon Skip
- D. Upon Skip

Answer: B

## Question 64

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 explorer in Microsoft Power BI Desktop
- C. Model view in Microsoft Power BI Desktop
- D. Best Practice Analyzer in Tabular Editor

Answer: D

## Question 65

DRAG DROP -

You have a Fabric tenant that contains a Microsoft Power BI report named Report1. 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 Query Timings tab.                                          |             |
| From Performance analyzer, capture a recording.                      |             |
| Sort the Duration (ms) column in descending order.                   |             |
| Enable Query Timings and Server Timings. Run the query.              |             |
| Sort the Duration (ms) column in descending order by DAX query time. |             |
| Copy the first query to DAX Studio.                                  |             |
| View the Server Timings tab.                                         |             |

Answer:

| Answer Area                                             |
|---------------------------------------------------------|
| From Performance analyzer, capture a recording.         |
| Sort the Duration (ms) column in descending order.      |
| Copy the first query to DAX Studio.                     |
| Enable Query Timings and Server Timings. Run the query. |
| View the Server Timings tab.                            |

554b579 **Highly Voted** 1 month, 2 weeks ago  
Good Video: <https://www.youtube.com/watch?v=C5HBhlUFsE&t=176s>  
From performance analyzer capture code.  
Sort duration descending by DAX query time.  
Copy the first query over to Dax Studio  
Enable query and server timings and run the query  
View Server timings  
upvoted 16 times

hello2tomoki **Most Recent** 1 month ago  
Here are the steps you should follow to identify the slowest DAX query and review how long the query spends in the formula engine compared to the storage engine:  
  
1.Open the report in Power BI Desktop. Go to the 'View' tab and click on 'Performance Analyzer' and Click on 'Start Recording', then 'Refresh Visuals'  
2.Review the DAX query and duration in the performance analyzer pane. Sort the Duration (ms) column in descending order  
3.Copy the first query to DAX Studio  
4.In DAX Studio, enable Query Timings and Server Timings. Run the query  
5.View the Server Timings tab to see data for the formula engine (FE) and the storage engine (SE)  
upvoted 1 times

[Exam DP-600 topic 1 question 88 discussion - ExamTopics](#)

## Question 66

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. ALM Toolkit
- C. Tabular Editor
- D. DAX Studio

Answer: C



## Question 67

HOTSPOT -

You have a Microsoft Power BI semantic model.

You plan to implement calculation groups.

You need to create a calculation item that will change the context from the selected date to month-to-date (MTD).

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

NOTE: Each correct selection is worth one point.

### Answer Area

(

CALCULATE  
GENERATE  
MEASURE

( ),

COMBINEVALUES  
SELECTEDMEASURE  
SELECTEDVALUE

DATESMTD ( 'Date'[Date] )

)

Answer:

### Answer Area

(

CALCULATE  
GENERATE  
MEASURE

( ),

COMBINEVALUES  
SELECTEDMEASURE  
SELECTEDVALUE

DATESMTD ( 'Date'[Date] )

)

## Question 68

You have a Microsoft Power BI report named Report1 that uses a Fabric semantic model. Users discover that Report1 renders slowly.

You open Performance analyzer and identify that a visual named Orders By Date is the slowest to render. The duration breakdown for Orders By Date is shown in the following table.

| Name           | Duration (ms) |
|----------------|---------------|
| DAX query      | 27            |
| Visual display | 39            |
| Other          | 1047          |

What will provide the greatest reduction in the rendering duration of Report1?

- A. Enable automatic page refresh.
- B. Optimize the DAX query of Orders By Date by using DAX Studio.
- C. Change the visual type of Orders By Date.
- D. Reduce the number of visuals in Report1.

Answer: D

## Question 69

You have a custom Direct Lake semantic model named Model1 that has one billion rows of data.

You use Tabular Editor to connect to Model1 by using the XMLA endpoint.

You need to ensure that when users interact with reports based on Model1, their queries always use Direct Lake mode.

What should you do?

- A. From Model, configure the Default Mode option.
- B. From Partitions, configure the Mode option.
- C. From Model, configure the Storage Location option.
- D. From Model, configure the Direct Lake Behavior option.

Answer: D

## Question 70

DRAG DROP -

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
- SalesAddress

You need to modify the model to ensure that users assigned the SalesRegionManager role cannot see a column named Address in SalesAddress.

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                                                             | Answer Area |
|---------------------------------------------------------------------|-------------|
| Open the model in Power BI Desktop.                                 |             |
| Open the model in Tabular Editor.                                   |             |
| Select the <b>Address</b> column in SalesAddress.                   |             |
| Set the Hidden property to <b>True</b> .                            |             |
| Set Object Level Security to <b>Default</b> for SalesRegionManager. |             |
| Set Object Level Security to <b>None</b> for SalesRegionManager.    |             |

Answer:

| Answer Area                                                      |
|------------------------------------------------------------------|
| Open the model in Tabular Editor.                                |
| Select the <b>Address</b> column in SalesAddress.                |
| Set Object Level Security to <b>None</b> for SalesRegionManager. |

554b579 Highly Voted 1 month, 2 weeks ago

Setting OLS to None:

When you set OLS to None for a table or column:

- For viewers without the appropriate access levels, it's as if the secured tables or columns don't exist.
- The table or column will be hidden from that role.
- Users won't be able to see or interact with the secured data.

How To:

<https://www.insightsarena.com/post/mastering-object-level-security-in-power-bi-fixing-broken-visuals#:~:text=Configure%20Object-Level%20Security%3A%20In%20Tabular%20Editor%2C%20navigate,need%20to%20be%20restricted%20for%20a%20particular%20role.>

- In Tabular Editor, navigate to Tables and select the specific column names.
- Under Object-Level Security, set the security level to "None" for the role that needs restrictions.

- 1 Open the model in Tabular Editor
- 2 Select the address column in SalesAddress
- 3 Set the Object Level Security to None for the sales manager role.

upvoted 15 times

e4e2ec8 Most Recent 3 weeks, 2 days ago

Answer is correct; ref: <https://learn.microsoft.com/en-us/fabric/security/service-admin-object-level-security?tabs=table>

upvoted 3 times

## Question 71

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 measure
- C. when the FILTER function references columns from multiple tables
- D. when the FILTER function references a column from a single table that uses Import mode

ET: A

Gemini: C

Copilot: B

ET User:



### [Exam DP-600 topic 1 question 94 discussion - ExamTopics](#)

**trietnv** 6 days, 20 hours ago

D. ref: <https://learn.microsoft.com/en-us/dax/best-practices/dax-avoid-avoid-filter-as-filter-argument>  
 upvoted 1 times

**lpuech** 6 days, 21 hours ago

**Selected Answer: D**

Explanation here: <https://learn.microsoft.com/en-us/dax/best-practices/dax-avoid-avoid-filter-as-filter-argument>  
FILTER returns a table whereas KEEPFILTERS returns a Boolean. So, A, B and C are limitations of uses of Boolean expressions.  
 upvoted 2 times

**dp600** 3 weeks, 4 days ago

**Selected Answer: A**

I'm a bit confused, in any case I will go with A because context transition  
 upvoted 1 times

**Unbounded** 4 weeks ago

This is a very confusing question.  
If a FILTER references the entire table rather than a single table, then using KEEPFILTER is a better choice, but none of the options actually specify that.  
 upvoted 1 times

**hello2tomoki** 1 month ago



Replacing the FILTER function with the KEEPFILTERS function will reduce execution time in the use case where the FILTER function references columns from multiple tables (Option C). This is because KEEPFILTERS is more efficient when it intersects with an existing filter context, which can occur when conditions involve multiple columns (<https://www.sqlbi.com/articles/using-keepfilters-in-dax-updated/>)  
 upvoted 1 times

  **emmanuelkech** 1 month, 1 week ago

**Selected Answer: B**

When you replace FILTER with KEEPFILTERS, you avoid re-evaluating the entire table for each row. Instead, it leverages the existing context, resulting in better performance.

   upvoted 1 times

  **thisiston** 1 month, 1 week ago

**Selected Answer: A**

A is correct because

Option A (when the FILTER function uses a nested calculate function): The FILTER function in DAX filters a table to the rows that meet the condition provided in its arguments. When FILTER is used within a CALCULATE function, it typically modifies the context in which the calculation is performed. If FILTER uses a nested CALCULATE function, it creates a context transition. Using KEEPFILTERS in this scenario can improve performance by modifying the behavior of context transition to ensure that existing filters on a table are not unintentionally removed or overridden, which can otherwise increase the complexity and computational load of the query.



   upvoted 3 times

  **pppppppppie** 1 month, 1 week ago




he KEEPFILTERS function in Power BI is primarily designed to work with single-table or single-column filter contexts. When you use the FILTER function to reference columns from multiple tables, KEEPFILTERS may not work directly because it operates within the context of a single table or column.



A is the ans

   upvoted 3 times

  **dp600** 1 month, 2 weeks ago




view this --> <https://community.fabric.microsoft.com/t5/DAX-Commands-and-Tips/1-Diff-between-Filter-and-KeepFilters/m-p/2694220>



   upvoted 1 times

  **neoverma** 1 month, 2 weeks ago




**Selected Answer: C**

In general, every time you have to preserve the combination of columns coming from different tables, KEEPFILTERS is probably needed. Even though this could be required for columns of the same table, KEEPFILTERS is not required whenever there is a column that represents a cardinality higher than the intersection of the other columns – such as Date[Calendar Year Month] instead of the intersection of Date[Month] and Date[Calendar Year].  
<https://www.sqlbi.com/articles/when-to-use-keepfilters-over-iterators/>

   upvoted 3 times

  **Unbounded** 4 weeks ago

You cannot reference columns from multiple tables within the same FILTER

   upvoted 1 times

## Question 72

You have a semantic model named Model1. Model1 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. Select Test as role to view the data as the HR role.
- B. Filter the data in the report to match the intended logic of the filter for the HR department.
- C. Select Test as role to view the report as the HR manager.
- D. Ask the HR manager to open the report in Microsoft Power BI Desktop.

Answer: C, 争议

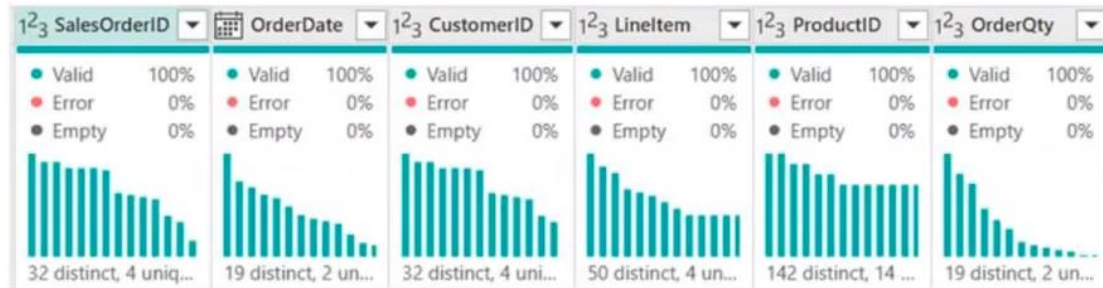
[Exam DP-600 topic 1 question 95 discussion - ExamTopics](#)

## Question 73

You have a Microsoft Fabric tenant that contains a dataflow.

You are exploring a new semantic model.

From Power Query, you need to view column information as shown in the following exhibit.



Which three Data view options should you select? Each correct answer presents part of the solution.

- A. Show column value distribution
- B. Enable details pane
- C. Enable column profile
- D. Show column quality details
- E. Show column profile in details pane

Answer: ACD



## Question 74

HOTSPOT -

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.

NOTE: Each correct answer is worth one point.

### Answer Area

```
SELECT ProductID,
 (ListPrice, WholesalePrice, AgentPrice) AS HighestSellingPrice,
 COALESCE
 GREATEST
 IIF
 MAX

 (ListPrice, WholesalePrice, AgentPrice) AS TradePrice
 CHOOSE
 COALESCE
 IIF
 MAX
FROM Sales.Products;
```

Answer:

## Answer Area

SELECT ProductID,

COALESCE  
**GREATEST**  
IIF  
MAX

(ListPrice, WholesalePrice, AgentPrice) AS HighestSellingPrice,

CHOOSE  
**COALESCE**  
IIF  
MAX

(ListPrice, WholesalePrice, AgentPrice) AS TradePrice

FROM Sales.Products;

**clux** Highly Voted 1 month, 2 weeks ago

GREATEST, COALESCE

upvoted 10 times

**woliveiras** Most Recent 3 weeks, 5 days ago

Greatest is ok; however, if we have all of them present, list price will return first if you use coalesce, which is not what the question requests. The question requests AgentPrice first. I am not sure about Coalesce.

upvoted 3 times

**72bd3bc** 2 days, 18 hours ago

Yes, it seems like there is no correct answer for the second box

upvoted 1 times

**Nefirs** 1 month ago

answer is correct

upvoted 2 times

**taphyoe** 1 month, 2 weeks ago

MAX  
COALESCE

upvoted 1 times

**neoverma** 1 month, 2 weeks ago

MAX() accepts one argument; GREATEST() accepts multiple arguments

upvoted 3 times

**taphyoe** 1 month, 2 weeks ago

correct - GREATEST and COALESCE

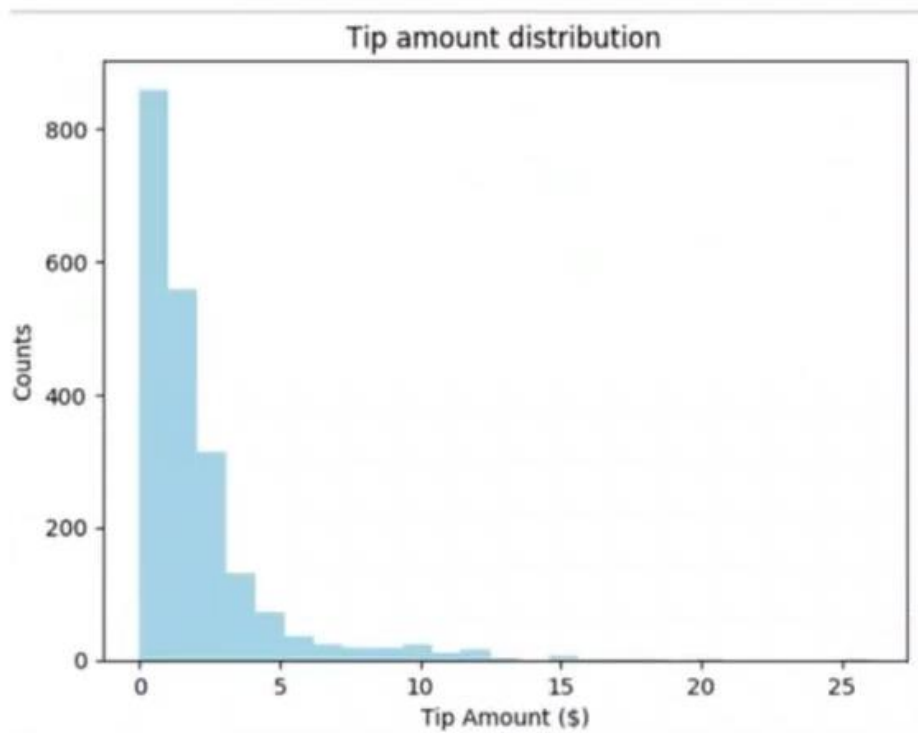
upvoted 5 times

## Question 75

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

Answer: A

## Question 76

HOTSPOT -

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 answer is worth one point.

### Answer Area

```
SELECT OrderID, CustomerID,
```

DATE\_BUCKET  
DATEFROMPARTS  
DATEPART  
DATETRUNC

```
(month, OrderDate) AS PeriodDate,
```

```
DATENAME(
```

day  
dayofyear  
weekday

```
, OrderDate) AS DayName
```

```
FROM Sales.Orders;
```

Answer:

### Answer Area

SELECT OrderID, CustomerID,

DATE\_BUCKET  
DATEFROMPARTS  
DATEPART  
**DATETRUNC**

(month, OrderDate) AS PeriodDate,

DATENAME(

day  
dayofyear  
**weekday**

, OrderDate) AS DayName

FROM Sales.Orders;

## Question 77

You have a Fabric tenant that contains JSON files in OneLake. The files have one billion items.

You plan to perform time series analysis of the items.

You need to transform the data, visualize the data to find insights, perform anomaly detection, and share the insights with other business users. The solution must meet the following requirements:

- Use parallel processing.
- Minimize the duplication of data.
- Minimize how long it takes to load the data.

What should you use to transform and visualize the data?

- A. the PySpark library in a Fabric notebook
- B. the pandas library in a Fabric notebook
- C. a Microsoft Power BI report that uses core visuals

Answer: A

## Question 78

You have a Fabric tenant that contains customer churn data stored as Parquet files in OneLake. The data contains details about customer demographics and product usage. You create a Fabric notebook to read the data into a Spark DataFrame. You then create column charts in the notebook that show the distribution of retained customers as compared to lost customers based on geography, the number of products purchased, age, and customer tenure.

Which type of analytics are you performing?

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

Answer: B

## Question 79

HOTSPOT -

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 the total amount of sales from the same period last year.

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

NOTE: Each correct selection is worth one point.

### Answer Area

EVALUATE

VAR \_LYSales =

CALCULATE  
CALCULATETABLE  
FILTER  
SUMMARIZE  
SUMMARIZECOLUMNS

( [Total Sales], SAMEPERIODLASTYEAR ( 'Orders'[Order Date] ) )

RETURN

[Total Sales]  
\_LYSales  
{ \_LYSales }  
VAR

Answer:

CALCULATE

\_LYSales



## Question 80

You have a Fabric workspace named Workspace1 that contains a dataflow named Dataflow1. Dataflow1 returns 500 rows of data.

You need to identify the min and max values for each column in the query results.

Which three Data view options should you select? Each correct answer presents part of the solution.

NOTE: Each correct answer is worth one point.

- A. Show column value distribution
- B. Enable column profile
- C. Show column profile in details pane
- D. Show column quality details
- E. Enable details pane

Answer: BCE

## Question 81

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. Table.schema
- B. Table.view
- C. Table.FuzzyGroup
- D. Table.Profile

Answer: D