Practice Exam Questions

DP-500

Designing and Implementing Enterprise-Scale Analytics Solutions Using Microsoft Azure and Microsoft Power BI

EXAMKILLER

Help Pass Your Exam At First Try

Microsoft

Exam DP-500

Designing and Implementing Enterprise-Scale Analytics Solutions Using Microsoft Azure and Microsoft Power BI

Version: 4.1

[Total Questions: 114]

Topic break down

Topic	No. of Questions
Topic 1: Litware, Inc. Overview	7
Topic 2: Contoso, Ltd	8
Topic 3: Misc. Questions	99

Topic 1, Litware, Inc. Overview

Litware, Inc. is a retail company that sells outdoor recreational goods and accessories. The company sells goods both online and at its stores located in six countries.

Azure Resources

Litware has the following Azure resources:

- An Azure Synapse Analytics workspace named synapseworkspace1
- An Azure Data Lake Storage Gen2 account named datalake1 that is associated with synapseworkspace1
- A Synapse Analytics dedicated SQL pool named SQLDW

Dedicated SQL Pool

SQLDW contains a dimensional model that contains the following table.

Name	Relevant column	Description
dbo.Customer	CustomerKey, CustomerID, CustomerEmail	The table currently contains 250,000 rows. Each row identifies a unique customer.
dbo.Product	ProductKey, ProductID, ProductName, ProductCategory, IsActive	The table currently contains 2,500 rows. Each row identifies a unique product.
dbo.Date	Date, Month, Year	The table currently contains 3,653 rows. Each row identifies a unique date.
dbo.SalesTransactions	CustomerKey, ProductKey, SalesDate, SalesChannelKey, SalesAmount, QuantitySold	The table currently contains 75 million rows. Each row identifies the purchase of a single product in a sales transaction.
dbo.SalesChannel	SalesChannelKey, SalesChannel	The table currently contains two rows to identify whether a sale occurred online or in a store.

SQLDW contains the following additional tables.

Name	Relevant column	Description
MLModel	Model, Model_Name	The table contains a machine learning model named PredictPurchase that predicts the likelihood of customers purchasing a specific product based on their past purchases.
CustomersWithProductScore	CustomerID, CustomerEmail, ProductID, ProductName, Score	The Score column contains the results from calling the predictive model.

SQLDW contains a view named dbo.CustomerPurchases that creates a distinct list of values from dbo.Customer [customeriD], dbo.Customer [CustomerEmail], dbo.ProductfProductID] and dbo.Product[ProductName].

The sales data in SQLDW is updated every 30 minutes. Records in dbo.SalesTransactions are updated in SQLDW up to three days after being created. The records do NOT change after three days.

Power BI

Litware has a new Power Bl tenant that contains an empty workspace named Sales Analytics.

All users have Power B1 Premium per user licenses.

IT data analysts are workspace administrators. The IT data analysts will create datasets and reports.

A single imported dataset will be created to support the company's sales analytics goals. The dataset will be refreshed every 30 minutes.

Analytics Goals

Litware identifies the following analytics goals:

- Provide historical reporting of sales by product and channel over time.
- Allow sales managers to perform ad hoc sales reporting with minimal effort.
- Perform market basket analysis to understand which products are commonly purchased in the same transaction.
- Identify which customers should receive promotional emails based on their likelihood of purchasing promoted products.

Litware plans to monitor the adoption of Power Bl reports over time. The company wants custom Power Bl usage reporting that includes the percent change of users that view reports in the Sales Analytics workspace each month.

Security Requirements

Litware identifies the following security requirements for the analytics environment:

• All the users in the sales department and the marketing department must be able to see Power B1 reports that contain market basket analysis and data about which customers are likely to purchase a product.

• Customer contact data in SQLDW and the Power B1 dataset must be labeled as Sensitive. Records must be kept of any users that use the sensitive data.

• Sales associates must be prevented from seeing the CustomerEmail column in Power B1 reports.

• Sales managers must be prevented from modifying reports created by other users.

Development Process Requirements

Litware identifies the following development process requirements:

• SQLDW and datalake1 will act as the development environment. Once feature development is complete, all entities in synapseworkspace1 will be promoted to a test workspace, and then to a production workspace.

• Power Bl content must be deployed to test and production by using deployment pipelines.

• All SQL scripts must be stored in Azure Repos.

The IT data analysts prefer to build Power Bl reports in Synapse Studio.

Question No : 1 - (Topic 1)

How should you configure the Power BI dataset refresh for the dbo.SalesTransactions table?

A. an incremental refresh of Product where the ModifiedDate value is during the last three days.

B. an incremental refresh of dbo.SalesTransactions where the SalesDate value is during the last three days.

C. a full refresh of all the tables

D. an incremental refresh of dbo.SalesTransactions where the SalesDate value is during the last hour.

Answer: B

Explanation:

The sales data in SQLDW is updated every 30 minutes. Records in dbo.SalesTransactions are updated in SQLDW up to three days after being created. The records do NOT change after three days.

Question No : 2 - (Topic 1)

You need to configure the Sales Analytics workspace to meet the ad hoc reporting requirements.

What should you do?

A. Grant the sales managers the Build permission for the existing Power BI datasets.
B. Grant the sales managers admin access to the existing Power BI workspace.
C. Create a deployment pipeline and grant the sales managers access to the pipeline.
D. Create a PBIT file and distribute the file to the sales managers.

Answer: D

Explanation:

Allow sales managers to perform ad hoc sales reporting with minimal effort

Power BI report templates contain the following information from the report from which they were generated:

Report pages, visuals, and other visual elements
The data model definition, including the schema, relationships, measures, and other model definition items
All query definitions, such as queries, Query Parameters, and other query elements
What is not included in templates is the report's data.

Report templates use the file extension .PBIT (compare to Power BI Desktop reports, which use the .PBIX extension).

Note: With Power BI Desktop, you can create compelling reports that share insights across your entire organization. With Power BI Desktop templates, you can streamline your work by creating a report template, based on an existing template, which you or other users in your organization can use as a starting point for a new report's layout, data model, and queries. Templates in Power BI Desktop help you jump-start and standardize report creation.

Reference: https://docs.microsoft.com/en-us/power-bi/create-reports/desktop-templates

Question No : 3 - (Topic 1)

What should you configure in the deployment pipeline?

A. a backward deployment
B. a selective deployment
C. auto-binding
D. a data source rule

Answer: D

Explanation:
Development Process Requirements
Litware identifies the following development process requirements:
SQLDW and datalake1 will act as the development environment. Once feature development is complete, all entities in synapseworkspace1 will be promoted to a test workspace, and then to a production workspace.
Power BI content must be deployed to test and production by using deployment pipelines.

Create deployment rules
When working in a deployment pipeline, different stages may have different configurations. For example, each stage can have different databases or different query parameters. The development stage might query sample data from the database, while the test and production stages query the entire database.

When you deploy content between pipeline stages, configuring deployment rules enables you to allow changes to content, while keeping some settings intact. For example, if you want a dataset in a production stage to point to a production database, you can define a rule for this. The rule is defined in the production stage, under the appropriate dataset. Once the rule is defined, content deployed from test to production, will inherit the value as defined in the deployment rule, and will always apply as long as the rule is unchanged and valid.

You can configure data source rules and parameter rules.

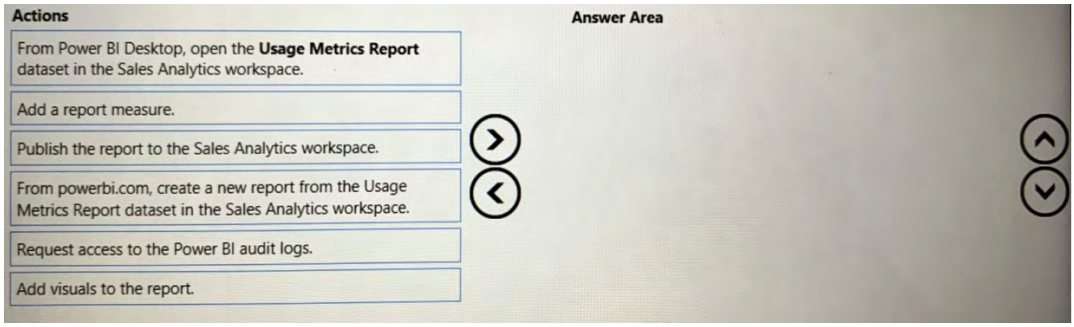
Incorrect:
Not B: if you already have a steady production environment, you can deploy it backward (to Test or Dev, based on your need) and set up the pipeline. The feature is not limited to any sequential orders.

Reference: https://docs.microsoft.com/en-us/power-bi/create-reports/deployment-pipelines-get-started#step-4---create-deployment-rules

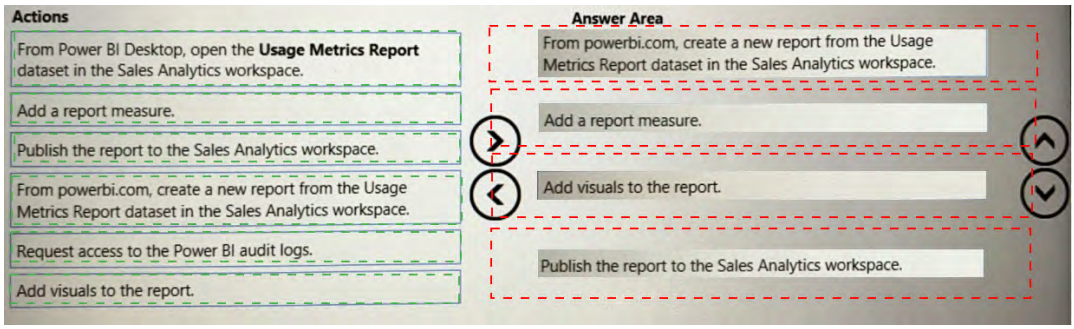
Question No : 4 DRAG DROP - (Topic 1)

You need to create the customized Power Bl usage reporting. The Usage Metrics Report dataset has already been created. The solution must minimize development and administrative effort.

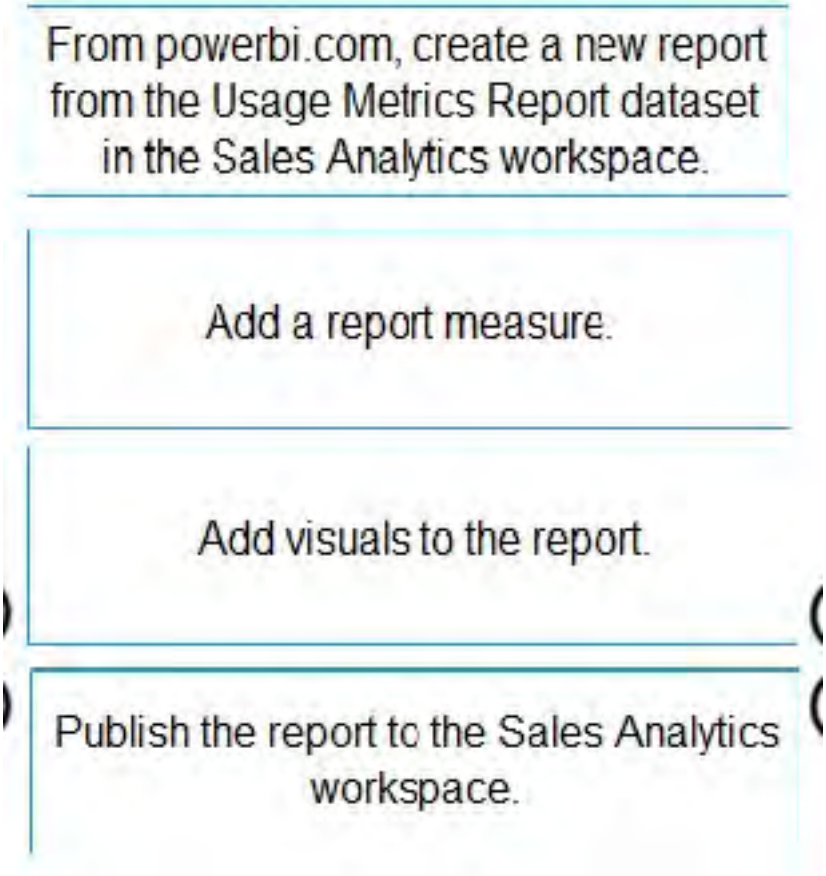
Which four actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.



Answer:



Explanation:



Step 1: From powerbi.com, create a new report..
The company wants custom Power BI usage reporting that includes the percent change of users that view reports in the Sales Analytics workspace each month.

Step 2: Add a report measure
Measures are used in some of the most common data analyses. Simple summarizations such as sums, averages, minimum, maximum and counts can be set through the Fields well. The calculated results of measures are always changing in response to your interaction with your reports, allowing for fast and dynamic ad-hoc data exploration.

Step 3: Add visuals to the report

Step 4: Publish the report to the Sales Analytics workspace

Question No : 5 - (Topic 1)

You need to recommend a solution to ensure that sensitivity labels are applied. The solution must minimize administrative effort.

Which three actions should you include in the recommendation? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

A. From the Power Bl Admin portal, set Allow users to apply sensitivity labels for Power Bl content to **Enabled.**
B. From the Power Bl Admin portal, set Apply sensitivity labels from data sources to their data in Power Bl to Enabled.
C. In SQLDW. apply sensitivity labels to the columns in the Customer and CustomersWithProductScore tables.
D. In the Power Bl datasets, apply sensitivity labels to the columns in the Customer and CustomersWithProductScore tables.
E. From the Power Bl Admin portal, set Make certified content discoverable to **Enabled.**

Answer: A,D,E

Explanation:

A Synapse Analytics dedicated SQL pool is named SQLDW.
Customer contact data in SQLDW and the Power BI dataset must be labeled as Sensitive.
Records must be kept of any users that use the sensitive data.

A (not B): Enable sensitivity labels
Sensitivity labels must be enabled on the tenant before they can be used in both the service and in Desktop.
To enable sensitivity labels on the tenant, go to the Power BI Admin portal, open the Tenant settings pane, and find the Information protection section.
In the Information Protection section, perform the following steps:

- Open Allow users to apply sensitivity labels for Power BI content.
- Enable the toggle.

D (not C): When data protection is enabled on your tenant, sensitivity labels appear in the sensitivity column in the list view of dashboards, reports, datasets, and dataflows.

E: Power BI Tenant Discovery Setting include Make certified content discoverable.

Reference: https://docs.microsoft.com/en-us/power-bi/enterprise/service-security-enable-data-sensitivity-labels

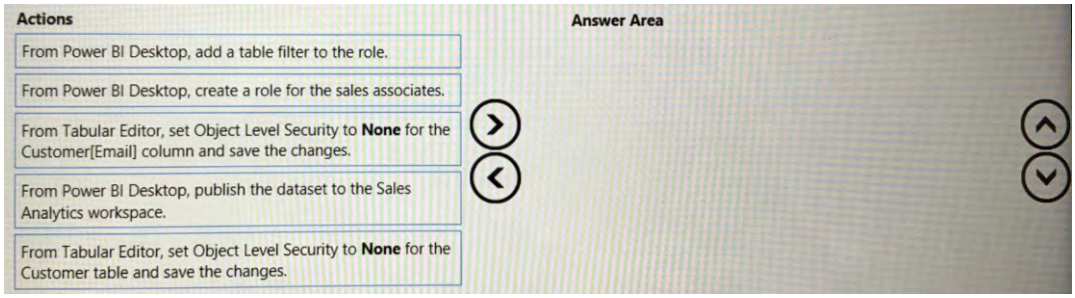
https://docs.microsoft.com/en-us/power-bi/enterprise/service-security-apply-data-sensitivity-labels

https://support.nhs.net/knowledge-base/power-bi-guidance/

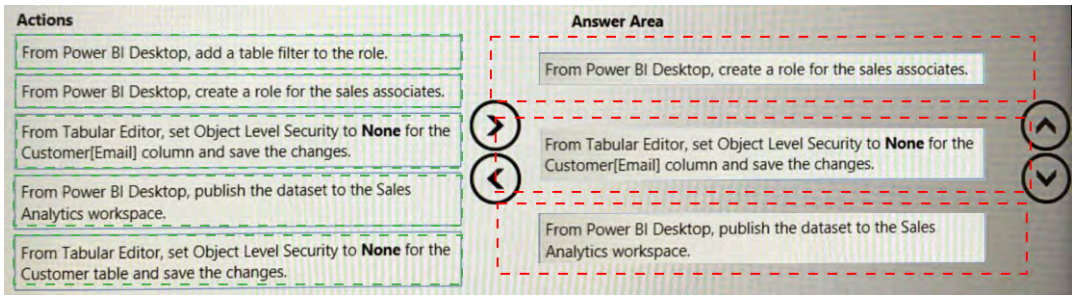
Question No : 6 DRAG DROP - (Topic 1)

You need to implement object-level security (OLS) in the Power BI dataset for the sales associates.

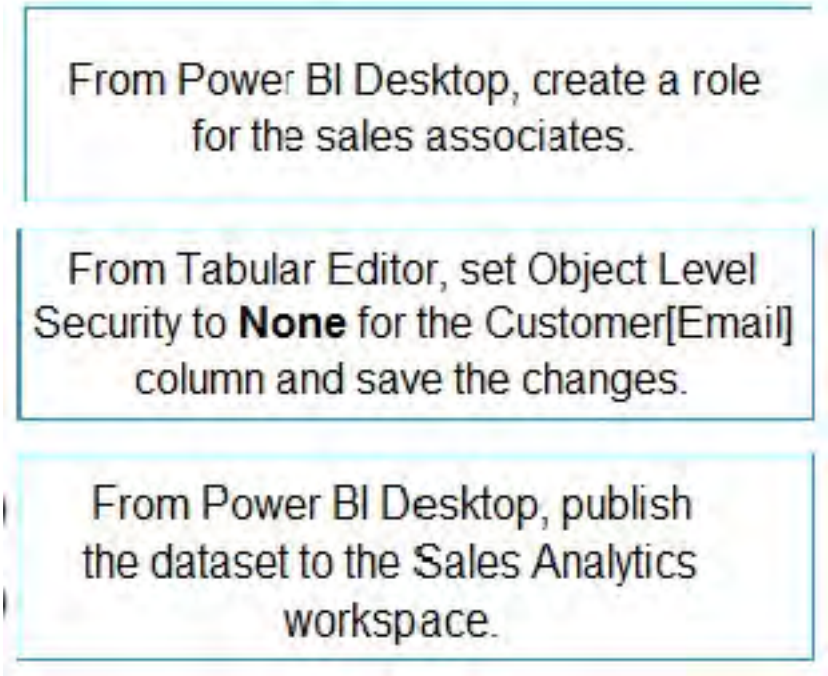
Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.



Answer:



Explanation:

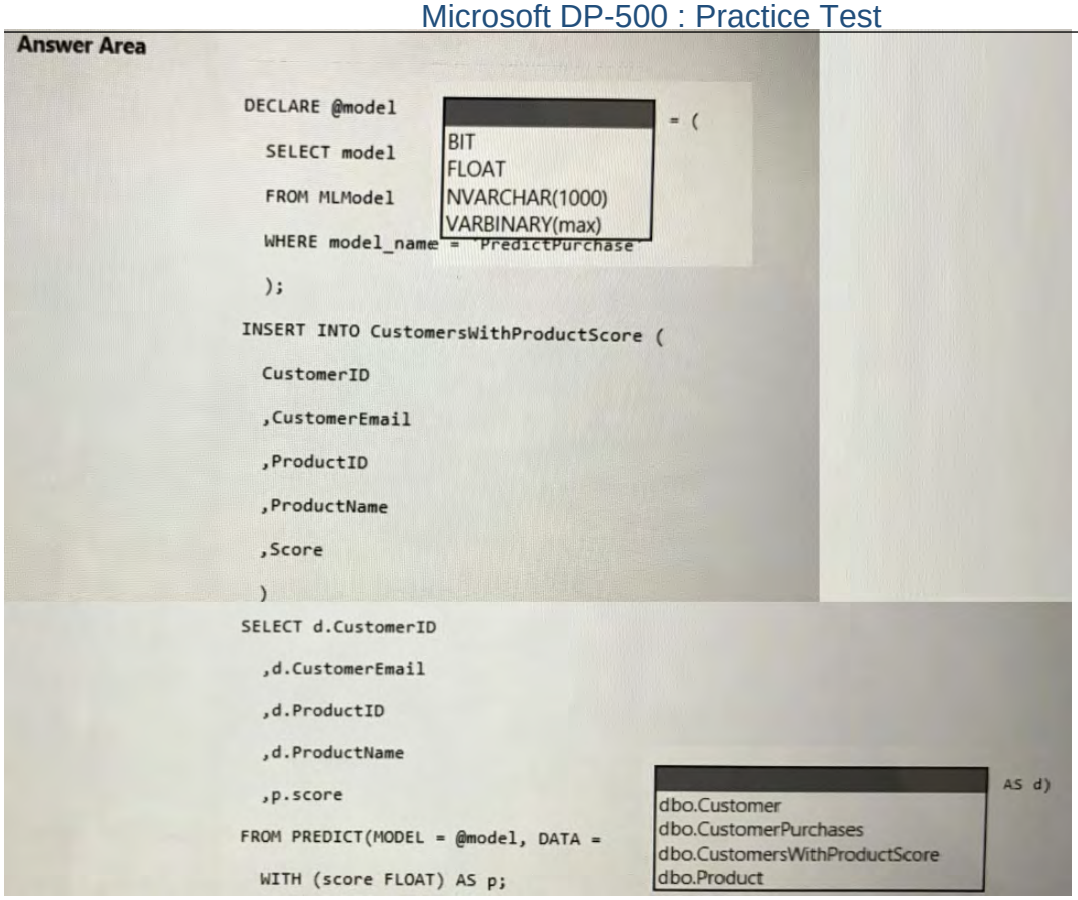


Question No : 7 HOTSPOT - (Topic 1)

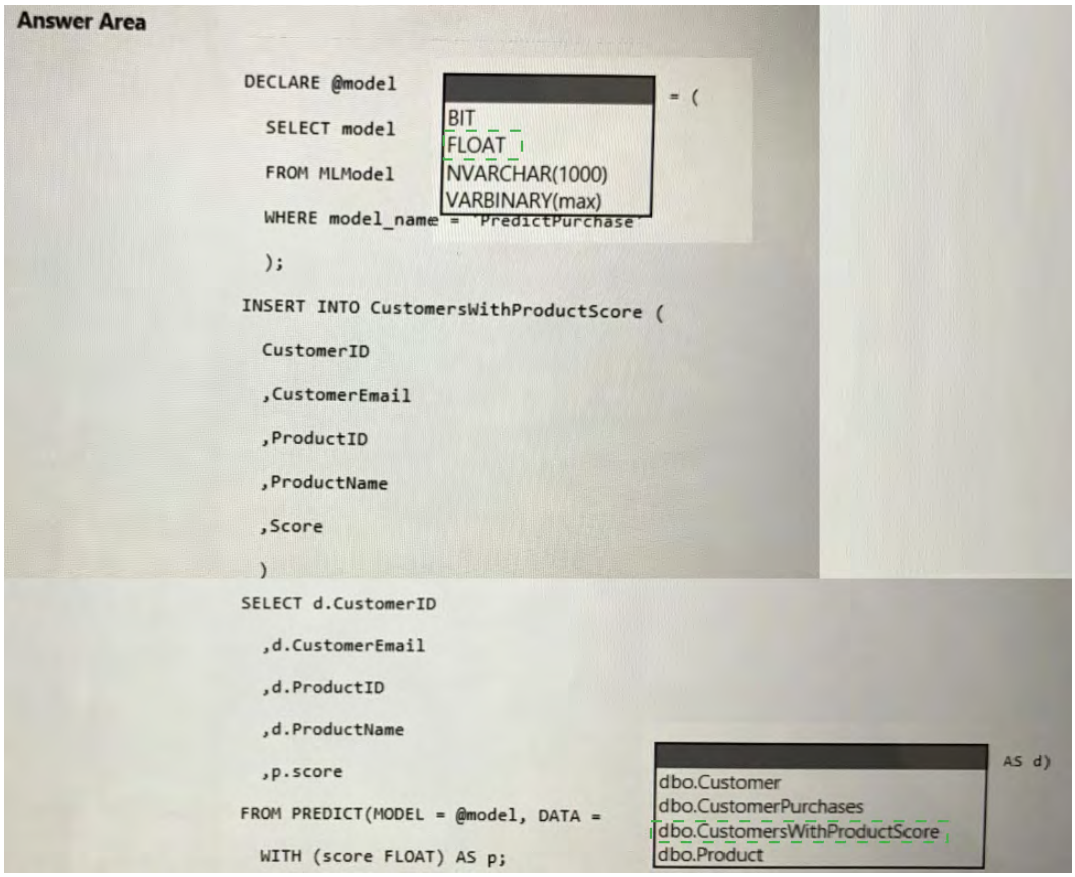
You need to populate the CustomersWithProductScore table.

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

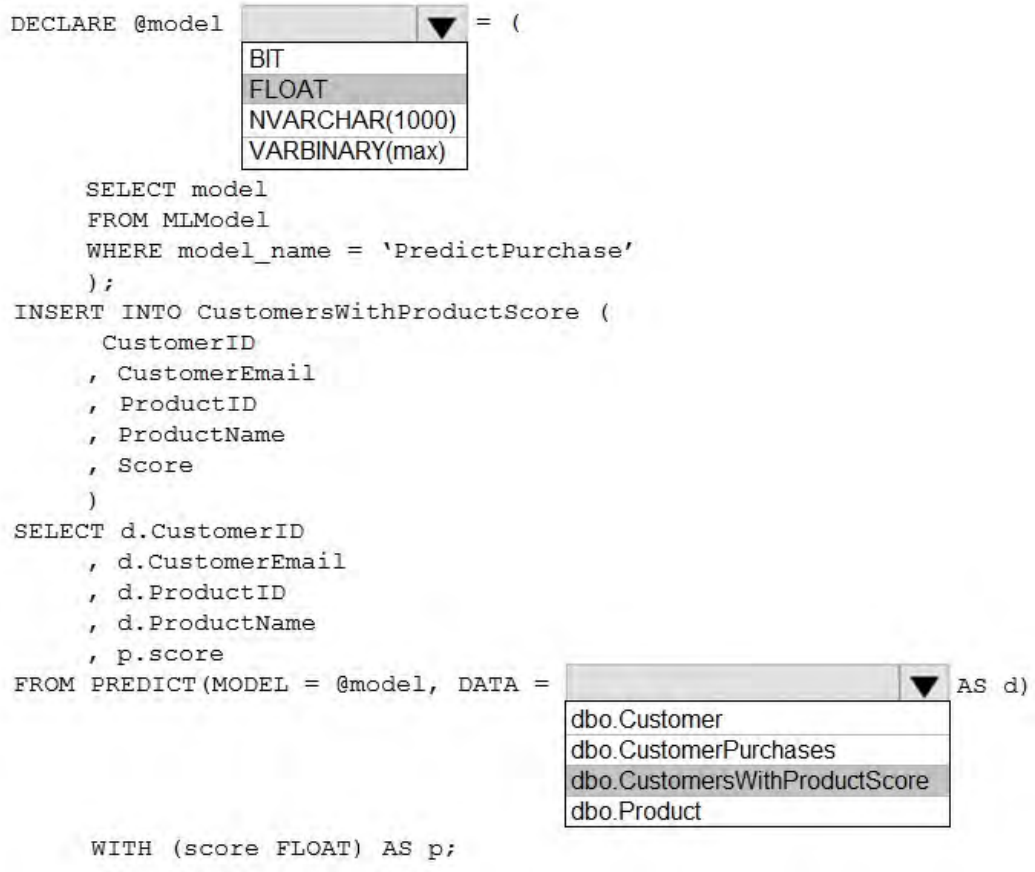
NOTE: Each correct selection is worth one point.



Answer:



Explanation:



Box 1: FLOAT
Identify which customers should receive promotional emails based on their likelihood of purchasing promoted products.

FLOT is used in the last statement of the code: WITH (score FLOAT) as p;

From syntax: MODEL

The MODEL parameter is used to specify the model used for scoring or prediction. The model is specified as a variable or a literal or a scalar expression.

Box 2: dbo.CustomerWithProductScore
Identify which customers should receive promotional emails based on their likelihood of purchasing promoted products.

Only table CustomerWithProductScore has the required filed score.

From the syntax:
DATA

The DATA parameter is used to specify the data used for scoring or prediction. Data is specified in the form of a table source in the query. Table source can be a table, table alias, CTE alias, view, or table-valued function.

Topic 2, Contoso, Ltd

Overview

Contoso, Ltd. is a company that sells enriched financial data to a variety of external customers.

Contoso has a main office in Los Angeles and two branch offices in New York and Seattle.

Data Infrastructure

Contoso has a 50-TB data warehouse that uses an instance of SQL Server on Azure Virtual Machines.

The data warehouse populates an Azure Synapse Analytics workspace that is accessed by the external customers. Currently, the customers can access alt the data.

Contoso has one Power Bl workspace named FinData that contains a single dataset. The dataset contains financial data from around the world. The workspace is used by 10 internal users and one external customer. The dataset has the following two data sources: the data warehouse and the Synapse Analytics serverless SQL pool.

Users frequently query the Synapse Analytics workspace by using Transact-SQL.

User Problems

Contoso identifies the following user issues:

• Some users indicate that the visuals in Power Bl reports are slow to render when making filter selections.

• Users indicate that queries against the serverless SQL pool fail occasionally because the size of tempdb has been exceeded.

• Users indicate that the data in Power Bl reports is stale. You discover that the refresh process of the Power Bl model occasionally times out

Planned Changes

Contoso plans to implement the following changes:

• Into the existing Power Bl dataset, integrate an external data source that is accessible by using the REST API.

• Build a new dataset in the FinData workspace by using data from the Synapse Analytics dedicated SQL pool.

• Provide all the customers with their own Power Bl workspace to create their own reports. Each workspace will use the new dataset in the FinData workspace.

• Implement subscription levels for the customers. Each subscription level will provide access to specific rows of financial data.

• Deploy prebuilt datasets to Power Bl to simplify the query experience of the customers.

• Provide internal users with the ability to incorporate machine learning models loaded to the dedicated SQL pool.

Question No : 8 - (Topic 2)

You need to recommend a solution for the customer workspaces to support the planned changes.

Which two configurations should you include in the recommendation? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

A. Set Use datasets across workspaces to Enabled
B. Publish the financial data to the web.
C. Grant the Build permission for the financial data to each customer.
D. Configure the FinData workspace to use a Power BI Premium capacity.

Answer: A,D

Explanation:

Build a new dataset in the FinData workspace by using data from the Synapse Analytics dedicated SQL pool.

Provide all the customers with their own Power BI workspace to create their own reports.

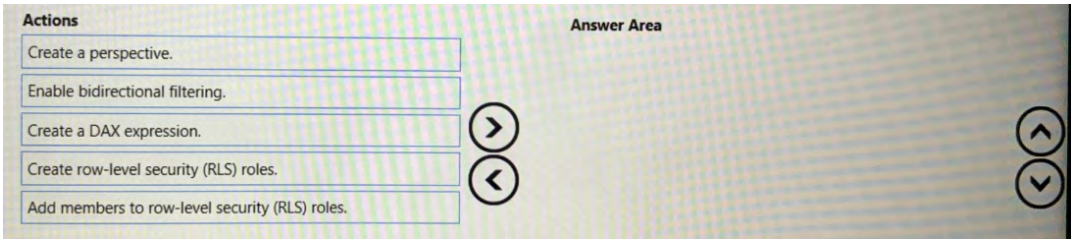
Each workspace will use the new dataset in the FinData workspace

Reference: https://docs.microsoft.com/en-us/power-bi/connect-data/service-datasets-admin-across-workspaces

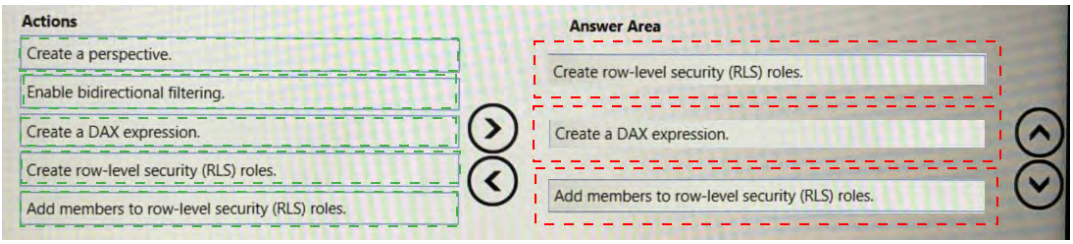
Question No : 9 DRAG DROP - (Topic 2)

You need to create Power BI reports that will display data based on the customers' subscription level.

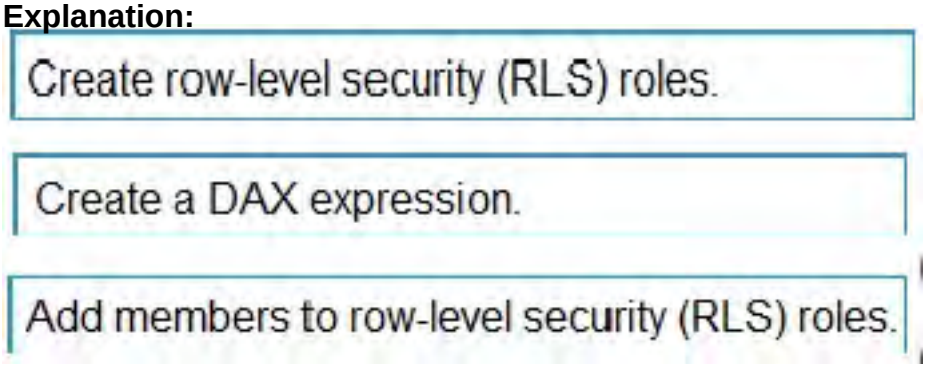
Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.



Answer:



Explanation:



Step 1: Create row-level security (RLS) roles
Create roles

Note: Provide all the customers with their own Power BI workspace to create their own reports. Each workspace will use the new dataset in the FinData workspace.
Implement subscription levels for the customers. Each subscription level will provide access to specific rows of financial data.
Deploy prebuilt datasets to Power BI to simplify the query experience of the customers.

Step 2: Create a DAX expression
Consider a model with two roles: The first role, named Workers, restricts access to all Payroll table rows by using the following rule expression:

FALSE()
Note: A rule will return no table rows when its expression evaluates to false.

Yet, a second role, named Managers, allows access to all Payroll table rows by using the following rule expression:

TRUE()
Take care: Should a report user map to both roles, they'll see all Payroll table rows.

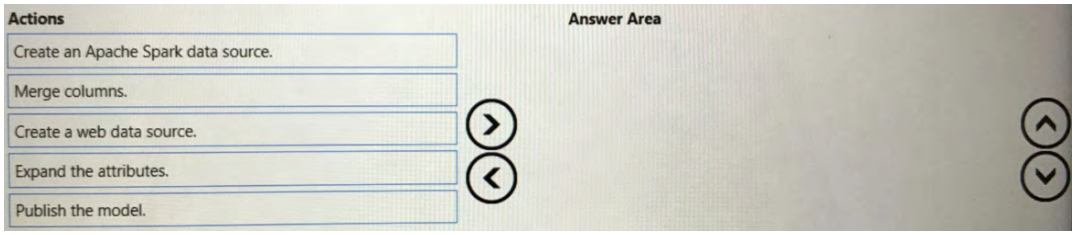
Step 3: Add members to row-level security (RLS) roles
Configure role mappings

Once [the model is] published to Power BI, you must map members to dataset roles.

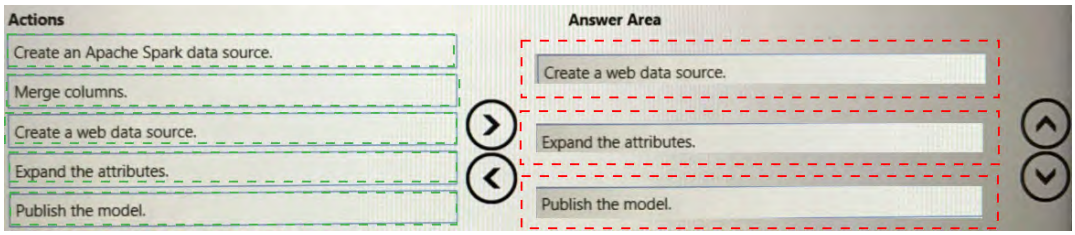
Question No : 10 DRAG DROP - (Topic 2)

You need to integrate the external data source to support the planned changes.

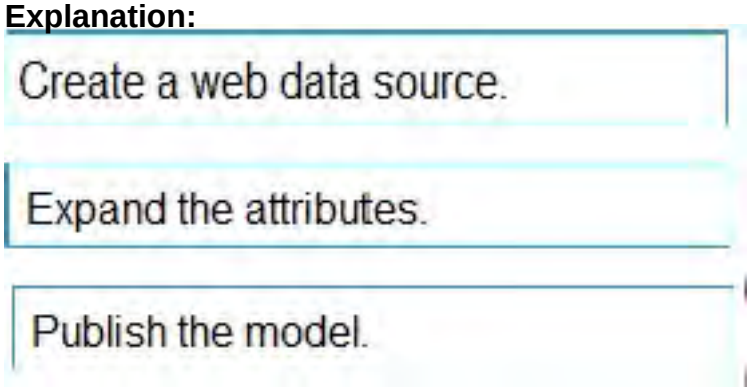
Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.



Answer:



Explanation:



Question No : 11 - (Topic 2)

You need to identify the root cause of the data refresh issue.

What should you use?

A. the Usage Metrics Report in powerbi.com
B. Query Diagnostics in Power Query Editor
C. Performance analyzer in Power Bl Desktop

Answer: B

Explanation:

Users indicate that the data in Power BI reports is stale. You discover that the refresh process of the Power BI model occasionally times out.

With Query Diagnostics, you can achieve a better understanding of what Power Query is doing at authoring and at refresh time in Power BI Desktop. While we'll be expanding on this feature in the future, including adding the ability to use it during full refreshes, at this time you can use it to understand what sort of queries you're emitting, what slowdowns you might run into during authoring refresh, and what kind of background events are happening.

Reference: https://docs.microsoft.com/en-us/power-query/querydiagnostics

Question No : 12 - (Topic 2)

Which two possible tools can you use to identify what causes the report to render slowly? Each correct answer presents a complete solution.

NOTE: Each correct selection is worth one point.

A. Synapse Studio
B. DAX Studio
C. Azure Data Studio
D. Performance analyzer in Power Bl Desktop

Answer: B,D

Explanation:

Some users indicate that the visuals in Power BI reports are slow to render when making filter selections.

B: You can investigate a slow query in a Power BI report using DAX Studio, looking at the query plan and the server timings.

D: Use Power BI Desktop Performance Analyzer to optimize the report or model.

Reference: https://www.sqlbi.com/tv/analyzing-a-slow-report-query-in-dax-studio/

https://docs.microsoft.com/en-us/power-bi/guidance/report-performance-troubleshoot

Question No : 13 - (Topic 2)

You need to recommend a solution to add new fields to the financial data Power Bl dataset with data from the Microsoft SQL Server data warehouse.

What should you include in the recommendation?

A. Azure Purview
B. Site-to-Site VPN
C. an XMLA endpoint
D. the on-premises data gateway

Answer: D

Explanation:

Refresh data from an on-premises SQL Server database

The SQL Server database must be accessed by Power BI through an on-premises data gateway.

You can install an on-premises data gateway on the same local computer as SQL Server (in production, it would typically be a different computer).

Reference: https://docs.microsoft.com/en-us/power-bi/connect-data/service-gateway-sql-tutorial

Question No : 14 - (Topic 2)

You need to recommend a solution to resolve the query issue of the serverless SQL pool.

The solution must minimize impact on the users.

What should you in the recommendation?

A. Update the statistics for the serverless SQL pool.
B. Move the data from the serverless SQL pool to a dedicated Apache Spark pool.
C. Execute the sp_sec_process_daca_limic stored procedure for the serverless SQL pool.
D. Move the data from the serverless SQL pool to a dedicated SQL pool.

Answer: D

Explanation:

Users indicate that queries against the serverless SQL pool fail occasionally because the size of tempdb has been exceeded.

In the dedicated SQL pool resource, temporary tables offer a performance benefit because their results are written to local rather than remote storage.

Temporary tables in serverless SQL pool.
Temporary tables in serverless SQL pool are supported but their usage is limited. They can't be used in queries which target files.

For example, you can't join a temporary table with data from files in storage. The number of temporary tables is limited to 100, and their total size is limited to 100 MB.

Reference: https://docs.microsoft.com/en-us/azure/synapse-analytics/sql/develop-tables-temporary

Question No : 15 HOTSPOT - (Topic 2)

You need to build a Transact-SQL query to implement the planned changes for the internal users.

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

NOTE: Each correct selection is worth one point.

```
DECLARE @model varbinary(max) = (
  SELECT native_model_object
  FROM ml_models
  WHERE model_name = 'rxLinMod'
  AND model_version = 'v1');
SELECT d.*, p.*
FROM [EVALUATE / PIVOT / PREDICT / SCORE] (MODEL = @model, DATA = dbo.rx_linMod as lm)
go [AS / CONTAINS / FROM / GROUP BY / WITH] (model_outcome float, trade_volume float, price_Pred float) as p;
```

Answer:

```
DECLARE @model varbinary(max) = (
  SELECT native_model_object
  FROM ml_models
  WHERE model_name = 'rxLinMod'
  AND model_version = 'v1');
SELECT d.*, p.*
FROM [EVALUATE / PIVOT / PREDICT / SCORE] (MODEL = @model, DATA = dbo.rx_linMod as lm)
go [AS / CONTAINS / FROM / GROUP BY / WITH] (model_outcome float, trade_volume float, price_Pred float) as p;
```

Explanation:

Box 1: PREDICT
Provide internal users with the ability to incorporate machine learning models loaded to the dedicated SQL pool.

The example below shows a sample query using prediction function. An additional column with name Score and data type float is created containing the prediction results. All the input data columns as well as output prediction columns are available to display with the select statement.

```
-- Query for ML predictions
SELECT d.*, p.Score
FROM PREDICT(MODEL = (SELECT Model FROM Models WHERE Id = 1),
DATA = dbo.mytable AS d, RUNTIME = ONNX) WITH (Score float) AS p;
```

Box 2: WITH

Topic 3, Misc. Questions

Question No : 16 - (Topic 3)

You are optimizing a Power Bl data model by using DAX Studio.

You need to capture the query events generated by a Power Bl Desktop report.

What should you use?

A. the DMV list
B. a Query Plan trace
C. an All Queries trace
D. a Server Timings trace

Answer: C

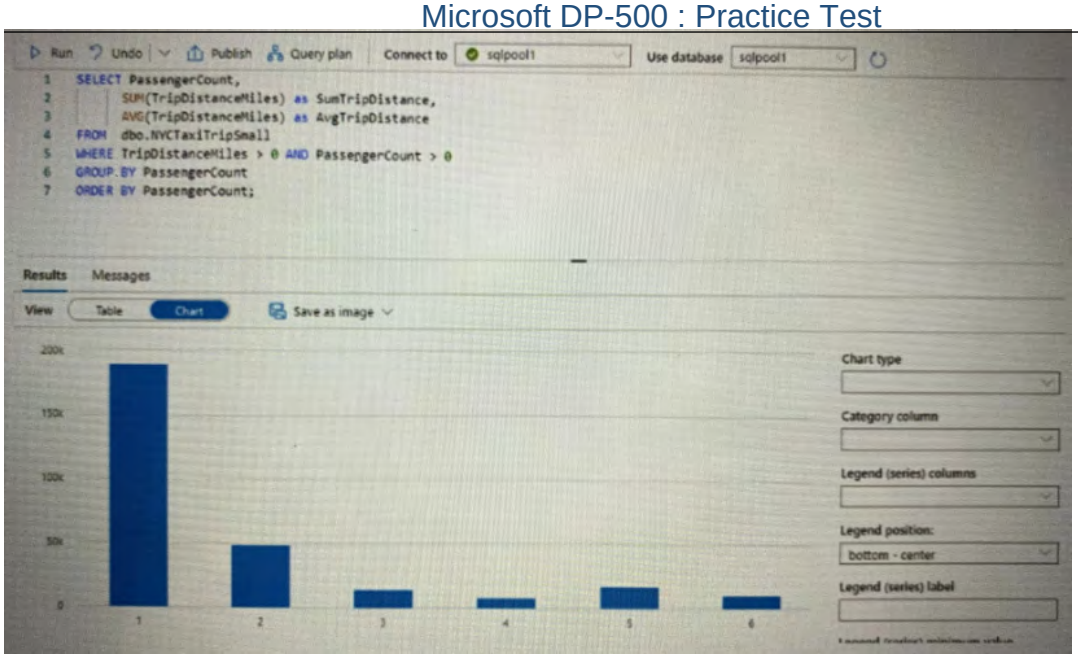
Explanation:
The All Queries trace in Dax Studio supports capturing the query events from all client tools (not just queries sent from DAX Studio like the Query Plan and Server Timings features do). The 'All Queries" trace is really useful when you wish to see the queries that are generated by a client tool like Power BI Desktop.

Reference: https://daxstudio.org/documentation/features/all-queries-trace/

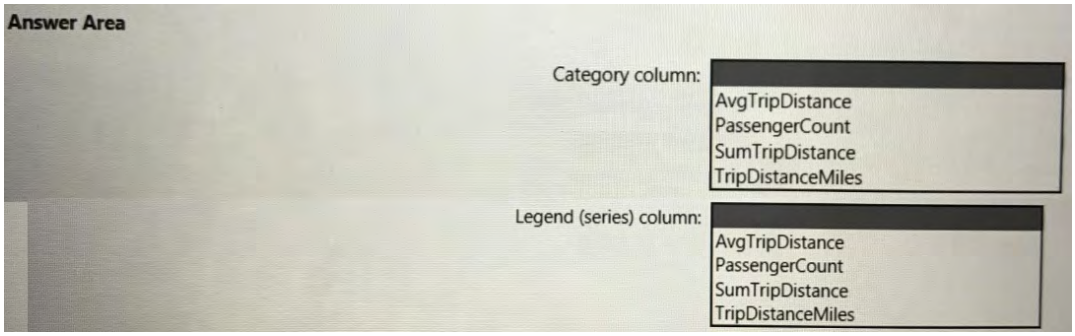
Question No : 17 HOTSPOT - (Topic 3)

You are using Azure Synapse Studio to explore a dataset that contains data about taxi trips.

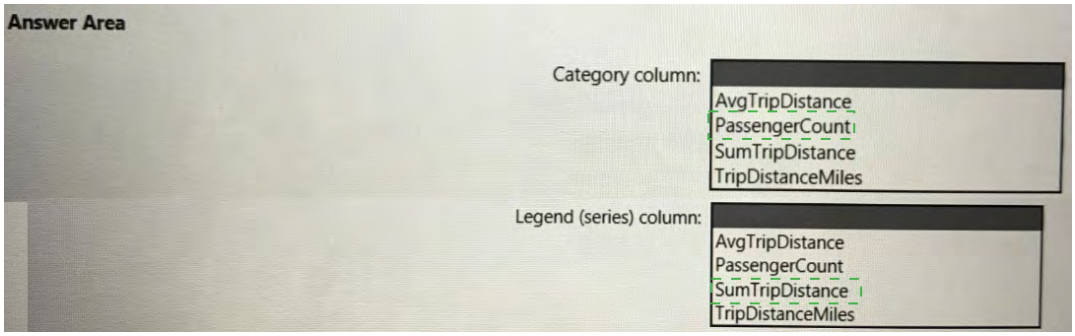
You need to create a chart that will show the total trip distance according to the number of passengers as shown in the following exhibit.

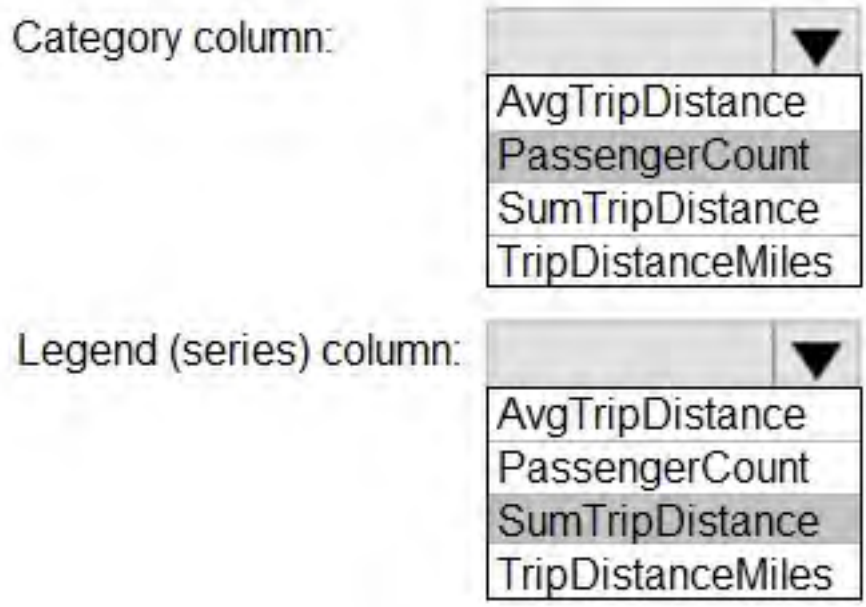


How should you configure the chart? To answer, select the appropriate options in the answer area. NOTE: Each correct selection is worth one point.



Answer:





Explanation:

Question No : 18 - (Topic 3)

You have a Power Bl report that contains the visual shown in the following exhibit.

You need to make the visual more accessible to users who have color vision deficiency.

What should you do?

A. Change the font color of values in the Sales column to white.
B. Change the red background color to orange.
C. Add icons to represent the sales status of each product.
D. Add additional measures to the table values.

Answer: A

Explanation:

Themes, contrast and colorblind-friendly colors
You should ensure that your reports have enough contrast between text and any background colors.
Certain color combinations are particularly difficult for users with color vision deficiencies to distinguish. These include the following combinations:

**---> green and black
green and red
green and brown
blue and purple
green and blue
light green and yellow
blue and grey
green and grey

Avoid using these colors together in a chart, or on the same report page.

Reference: https://docs.microsoft.com/en-us/power-bi/create-reports/desktop-accessibility-creating-reports

Question No : 19 - (Topic 3)

You are using a Python notebook in an Apache Spark pool in Azure Synapse Analytics.
You need to present the data distribution statistics from a DataFrame in a tabular view.
Which method should you invoke on the DataFrame?

A. freqltems
B. explain

C. rollup
D. describe

Answer: D

Explanation:

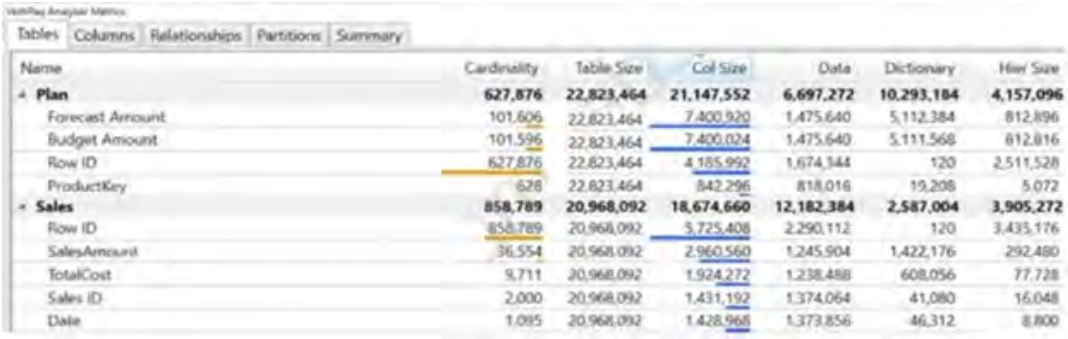
The aggregating statistic can be calculated for multiple columns at the same time with the describe function.

Example:
titanic[["Age", "Fare"]].describe()
Out[6]:
Age Fare
count 714.000000 891.000000
mean 29.699118 32.204208
std 14.526497 49.693429
min 0.420000 0.000000
25% 20.125000 7.910400
50% 28.000000 14.454200
75% 38.000000 31.000000
max 80.000000 512.329200

Reference:
https://pandas.pydata.org/docs/getting_started/intro_tutorials/06_calculate_statistics.html

Question No : 20 - (Topic 3)

You use the Vertipaq Analyzer to analyze tables in a dataset as shown in the Tables exhibit. (Click the Tables tab.)


Tables | Columns | Relationships | Partitions | Summary

Name	Cardinality	Table Size	Col Size	Data	Dictionary	Hier Size
Plan	**627,876**	**22,823,464**	**21,147,552**	**6,697,272**	**10,293,184**	**4,157,096**
Forecast Amount	101,606	22,823,464	7,400,920	1,475,640	5,112,384	812,896
Budget Amount	101,596	22,823,464	7,400,024	1,475,640	5,111,568	812,816
Row ID	627,876	22,823,464	4,185,992	1,674,344	120	2,511,528
ProductKey	628	22,823,464	842,296	818,016	19,208	5,072
Sales	**858,789**	**20,968,092**	**18,674,660**	**12,182,384**	**2,587,004**	**3,905,272**
Row ID	858,789	20,968,092	5,725,408	2,290,112	120	3,435,176
SalesAmount	36,554	20,968,092	2,960,560	1,245,904	1,422,176	292,480
TotalCost	9,711	20,968,092	1,924,272	1,238,488	608,056	77,728
Sales ID	2,000	20,968,092	1,431,192	1,374,064	41,080	16,048
Date	1,095	20,968,092	1,428,968	1,373,856	46,312	8,800

The table relationships for the dataset are shown in the Relationships exhibit. (Click the Relationships tab.)

Tables | Columns | Relationships | Partitions | Summary

Table / Relationship	Size	Max From Cardinality	Max To Cardinality	1:M Ratio %	Missing Keys
Plan	**1,675,912**	**627,876**	**858,789**	**136.78%**	**7**
Plan[ProductKey] ∞–1 Product[ProductKey]	848	628	629	0.10%	0
Plan[StoreKey] ∞–1 Store[Store Key]	360	306	299	0.05%	7
Plan[GeographyKey] ∞–1 Geography[GeographyKey]	312	263	263	0.04%	0
Plan[DateKey] ∞–1 Month & Year Distinct[Date]	32	36	36	0.01%	0
Sales	**2,293,432**	**858,789**	**1,095**	**0.13%**	**858,793**
Sales[Date] ∞–1 Calendar[Date]	1,760	1,095	1,095	0.13%	0
Sales[GeographyKey] ∞–1 Geography[GeographyKey]	312	263	263	0.03%	0
Sales[PromotionKey] ∞–1 Promotion[Promotion Key]	24	28	28	0.00%	0
Sales[channelKey] ∞–1 Channel[ChannelKey]	8	4	4	0.00%	0
Sales[Row ID] ∞–1 Plan Header Details[Row ID]	0	858,789	3	0.00%	858,786

You need to reduce the model size by eliminating invalid relationships.

Which column should you remove?

A. Sales[Sales Amount]
B. Sales[RowID]
C. Sales[Sales ID]
D. Plan[RowID]

Answer: B

Explanation:

Sales[Row ID] has 858,786 missing keys and 858,789 Max From Cardinality.
Note: The Max From Cardinality column defines the cost of the relationship which is the amount of time DAX needs to transfer the filters from the dimensions table to the fact table.

Reference: https://blog.enterprisedna.co/vertipaq-analyzer-tutorial-relationships-referential-integrity/

Question No : 21 HOTSPOT - (Topic 3)

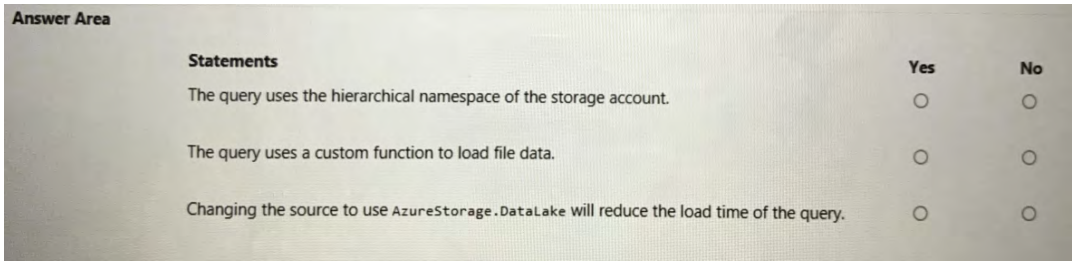
You have an Azure Data Lake Storage Gen 2 container that stores more than 300,000 files representing hourly telemetry data. The data is organized in folders by the year, month, and day according to when the telemetry was captured.

You have the following query in Power Query Editor.

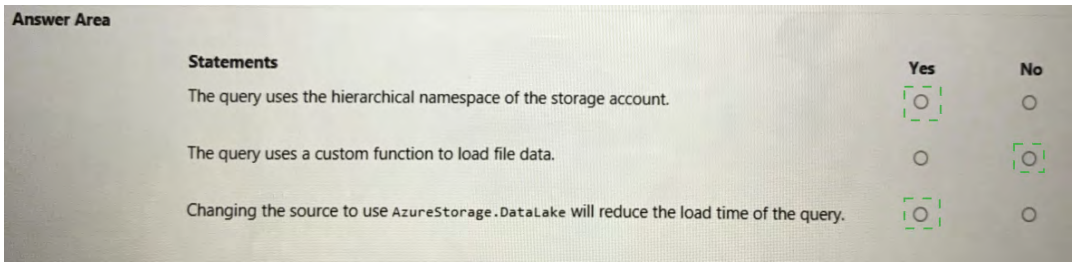
```
let
  Source = AzureStorage.Blobs("https://tmppbie01.blob.core.windows.net/logs/"),
  Filtered = Table.SelectRows(Source, each Text.StartsWith([Name], "2019/12/")
    and [Extension] = ".csv"),
  Transformed = Table.AddColumn(Filtered, "Transformed", each TransformFiles([Content])),
  Limited = Table.SelectColumns(Transformed, "Transformed"),
  Expanded = Table.ExpandTableColumn(Limited, "Transformed", {"Date", "Name", "Activity"}),
  Final = Table.TransformColumnTypes(Expanded,
    {{"Date", type date}, {"Name", type text}, {"Activity", type text}})
in
  Final
```

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:



Explanation:

Statements	Yes	No
The query uses the hierarchical namespace of the storage account.	○	○
The query uses a custom function to load file data.	○	○
Changing the source to use `AzureStorage.DataLake` will reduce the load time of the query.	○	○

Box 1: Yes

A key mechanism that allows Azure Data Lake Storage Gen2 to provide file system

performance at object storage scale and prices is the addition of a hierarchical namespace. This allows the collection of objects/files within an account to be organized into a hierarchy of directories and nested subdirectories in the same way that the file system on your computer is organized. With a hierarchical namespace enabled, a storage account becomes capable of providing the scalability and cost-effectiveness of object storage, with file system semantics that are familiar to analytics engines and frameworks.

Box 2: No
Table.SelectRows returns a table of rows from the table, that matches the selection condition.

Box 3: Yes
Azure Data Lake Storage has higher throughput and IOPS.

Note: Azure Blob Storage is a general purpose, scalable object store that is designed for a wide variety of storage scenarios. Azure Data Lake Storage is a hyper-scale repository that is optimized for big data analytics workloads.

Azure Data Lake Storage use Cases: Batch, interactive, streaming analytics and machine learning data such as log files, IoT data, click streams, large datasets

Question No : 22 - (Topic 3)

You have five Power BI reports that contain R script data sources and R visuals.

You need to publish the reports to the Power BI service and configure a daily refresh of datasets.

What should you include in the solution?

A. a Power BI Embedded capacity
B. an on-premises data gateway (standard mode)
C. a workspace that connects to an Azure Data Lake Storage Gen2 account
D. an on-premises data gateway (personal mode)

Answer: D

Explanation:
To schedule refresh of your R visuals or dataset, enable scheduled refresh and install an on-premises data gateway (personal mode) on the computer containing the workbook and R.