Exam Questions DP-500

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

https://www.2passeasy.com/dumps/DP-500/

NEW QUESTION 1
- (Exam Topic 3)
You have a deployment pipeline for a Power BI workspace. The workspace contains two datasets that use import storage mode.
A database administrator reports a drastic increase in the number of queries sent from the Power BI service to an Azure SQL database since the creation of the deployment pipeline.
An investigation into the issue identifies the following:

- One of the datasets is larger than 1 GB and has a fact table that contains more than 500 million rows.
- When publishing dataset changes to development, test, or production pipelines, a refresh is triggered against the entire dataset.

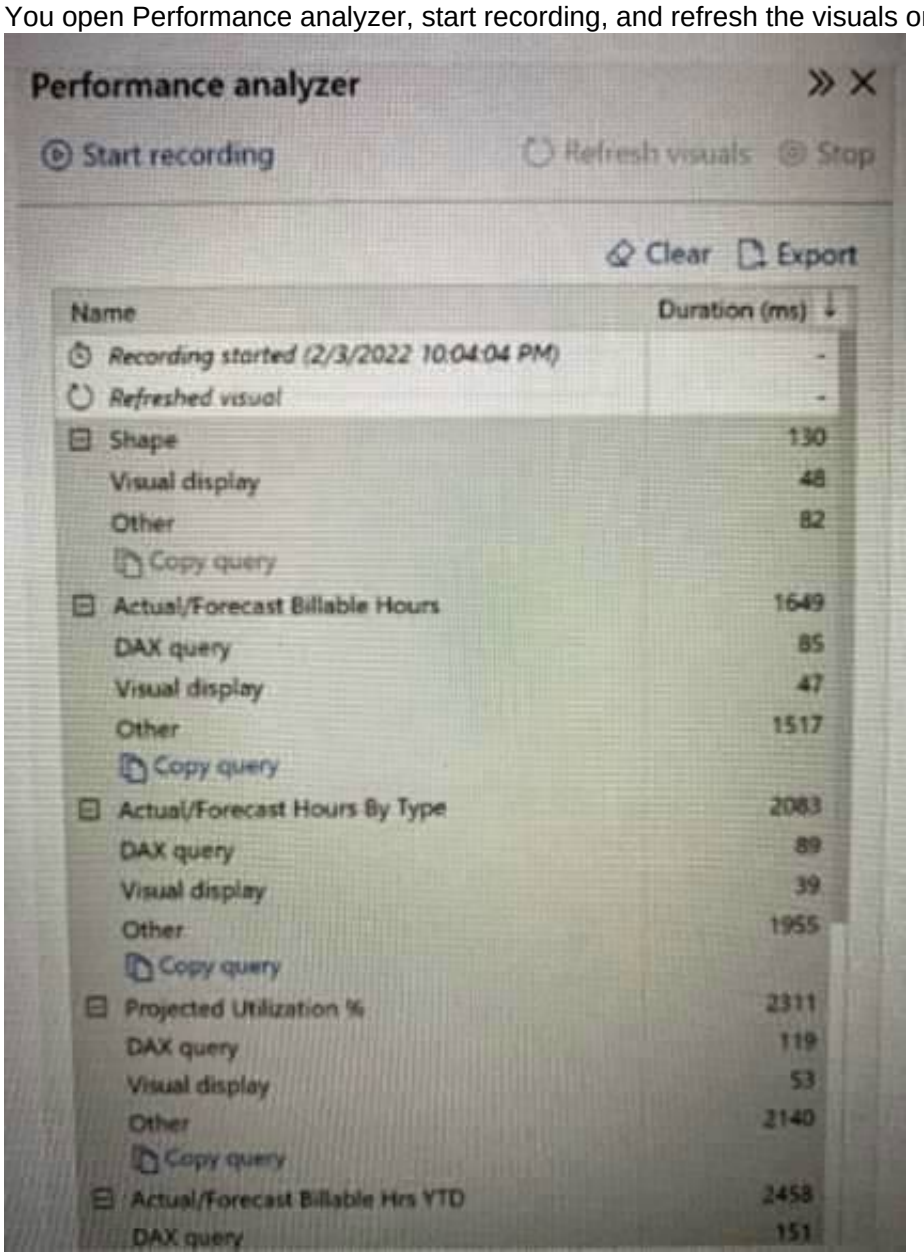
You need to recommend a solution to reduce the size of the queries sent to the database when the dataset changes are published to development, test, or production.
What should you recommend?

A. From Capacity settings in the Power BI Admin portal, reduce the Max Intermediate Row Set Count setting.
B. Configure the dataset to use a composite model that has a DirectQuery connection to the fact table.
C. Enable the large dataset storage format for workspace.
D. From Capacity settings in the Power BI Admin portal, increase the Max Intermediate Row Set Count setting.

Answer: B

Explanation:
A composite model in Power BI means part of your model can be a DirectQuery connection to a data source (for example, SQL Server database), and another part as Import Data (for example, an Excel file). Previously, when you used DirectQuery, you couldn't even add another data source into the model.
DirectQuery and Import Data have different advantages.
Now the Composite Model combines the good things of both Import and DirectQuery into one model. Using the Composite Model, you can work with big data tables using DirectQuery, and still import smaller tables using Import Data.
Reference:
https://radacad.com/composite-model-directquery-and-import-data-combined-evolution-begins-in-power-bi
https://powerbi.microsoft.com/en-us/blog/five-new-power-bi-premium-capacity-settings-is-available-on-the-por

NEW QUESTION 2
- (Exam Topic 3)
You open a Power Bl Desktop report that contains an imported data model and a single report page.
You open Performance analyzer, start recording, and refresh the visuals on the page. The recording produces the results shown in the following exhibit



What can you identify from the results?

A. The Actual/Forecast Hours by Type visual takes a long time to render on the report page when the data is cross-filtered.
B. The Actual/Forecast Billable Hrs YTD visual displays the most data.
C. Unoptimized DAX queries cause the page to load slowly.
D. When all the visuals refresh simultaneously, the visuals spend most of the time waiting on other processes to finish.

Answer: D

Explanation:
Most time is spent in the category Other - time required by the visual for preparing queries, waiting for other visuals to complete, or performing other background

processing.
Note: Each visual's log information includes the time spent (duration) to complete the following categories of tasks:
DAX query - if a DAX query was required, this is the time between the visual sending the query, and for Analysis Services to return the results.
Visual display - time required for the visual to draw on the screen, including time required to retrieve any web images or geocoding.
Other - time required by the visual for preparing queries, waiting for other visuals to complete, or performing other background processing.
Reference: https://docs.microsoft.com/en-us/power-bi/create-reports/desktop-performance-analyzer

NEW QUESTION 3
- (Exam Topic 3)
You use the Vertipaq Analyzer to analyze tables in a dataset as shown in the Tables exhibit. (Click the Tables tab.)

VertiPaq Analyzer Metrics

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.)

VertiPaq Analyzer Metrics

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/

NEW QUESTION 4
- (Exam Topic 3)
You have a Power BI tenant.
You plan to register the tenant in an Azure Purview account.
You need to ensure that you can scan the tenant by using Azure Purview.
Which two actions should you perform? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point.

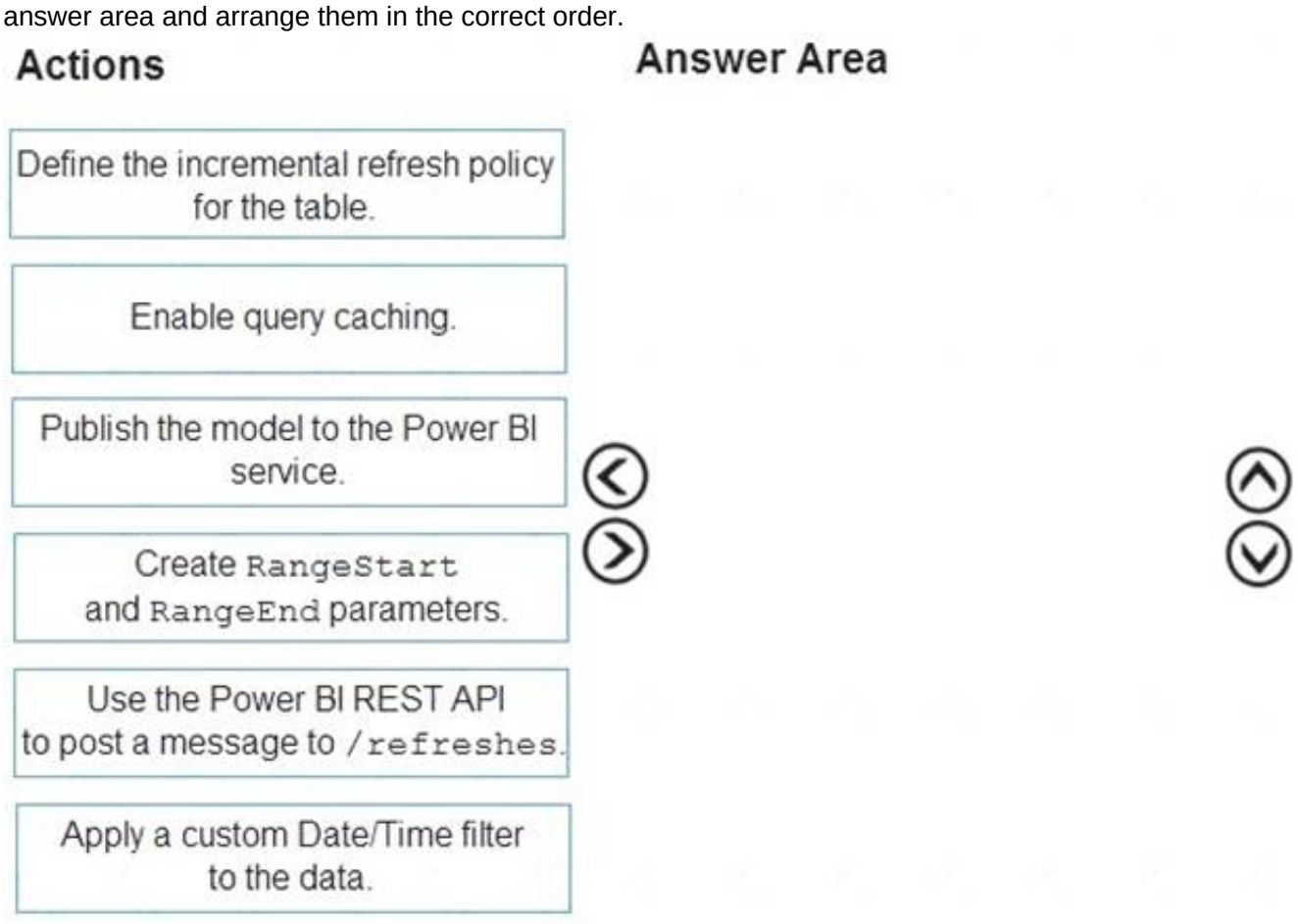
A. From the Microsoft 365 admin center, create a Microsoft 365 group.
B. From the Power BI Admin center, set Allow live connections to Enabled.
C. From the Power BI Admin center, set Allow service principals to use read-only Power BI admin APIs toEnabled.
D. From the Azure Active Directory admin center, create a security group.
E. From the Power BI Admin center, set Share content with external users to Enabled.

Answer: CD

Explanation:
Scan same-tenant Power BI using Azure IR and Managed Identity in public network. Make sure Power BI and Microsoft Purview accounts are in the same tenant.
Make sure Power BI tenant Id is entered correctly during the registration.
From Azure portal, validate if Microsoft Purview account Network is set to public access.
From Power BI tenant Admin Portal, make sure Power BI tenant is configured to allow public network.
(D) In Azure Active Directory tenant, create a security group.
From Azure Active Directory tenant, make sure Microsoft Purview account MSI is member of the new security group.
On the Power BI Tenant Admin portal, validate if Allow service principals to use read-only Power BI admin APIs is enabled for the new security group.
Associate the security group with Power BI tenant Log into the Power BI admin portal.

Select the Tenant settings page.
(C) Select Admin API settings > Allow service principals to use read-only Power BI admin APIs (Preview). Select Specific security groups.
Select Admin API settings > Enhance admin APIs responses with detailed metadata > Enable the toggle to allow Microsoft Purview Data Map automatically discover the detailed metadata of Power BI datasets as part of its scans.
Reference: https://docs.microsoft.com/en-us/azure/purview/register-scan-power-bi-tenant

NEW QUESTION 5
- (Exam Topic 3)
You have a Power Bl dataset. The dataset contains data that is updated frequently. You need to improve the performance of the dataset by using incremental refreshes.
Which four actions should you perform in sequence to enable the incremental refreshes? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.



A. Mastered
B. Not Mastered

Answer: A

Explanation:
Graphical user interface, text, application, chat or text message Description automatically generated
Step 1: Create RangeStart and RangeEnd parameters. Create parameters
In this task, use Power Query Editor to create RangeStart and RangeEnd parameters with default values. The default values apply only when filtering the data to be loaded into the model in Power BI Desktop. The values you enter should include only a small amount of the most recent data from your data source. When published to the service, these values are overridden by the incremental refresh policy.
Step 2: Apply a custom Date/Time filter to the data. Filter data
With RangeStart and RangeEnd parameters defined, apply a filter based on conditions in the RangeStart and RangeEnd parameters.
Before continuing with this task, verify your source table has a date column of Date/Time data type. Step 3: Define the incremental refresh policy for the table.
Define policy
After you've defined RangeStart and RangeEnd parameters, and filtered data based on those parameters, you define an incremental refresh policy. The policy is applied only after the model is published to the service and a manual or scheduled refresh operation is performed.
Step 4: Publish the model to the Power BI service. Save and publish to the service
When your RangeStart and RangeEnd parameters, filtering, and refresh policy settings are complete, be sure to save your model, and then publish to the service.
Reference: https://docs.microsoft.com/en-us/power-bi/connect-data/incremental-refresh-configure

NEW QUESTION 6
- (Exam Topic 3)
You have a Power Bl report that contains one visual.
You need to provide users with the ability to change the visual type without affecting the view for other users. What should you do?

A. From Report setting, select Personalize visuals.
B. From Tabular Editor, create a new perspective.
C. From the Bookmarks pane, select Focus mode, and then select Add.
D. From Visual options in Report settings, select Use the modern visual header with updated styling options.

Answer: A

Explanation:
Enable personalization in a report
You can enable the feature either in Power BI Desktop or the Power BI service. You can also enable it in embedded reports.
To enable the feature in Power BI Desktop, go to File > Options and settings > Options > Current file > Report settings. Make sure Personalize visuals is turned on.

NEW QUESTION 7

- (Exam Topic 3)
You have a file named File1.txt that has the following characteristics:
• A header row
• Tab delimited values
• UNIX-style line endings
You need to read File1.txt by using an Azure Synapse Analytics serverless SQL pool. Which query should you execute?

A.
```
SELECT*
FROM OPENROWSET(
  BULK 'file1.txt',
  DATA_SOURCE = 'Sql1',
  FORMAT = 'CSV', PARSER_VERSION = '2.0',
  FIELDTERMINATOR = '\t',
  ROWTERMINATOR = '0x0a',
  FIRSTROW= 2
)
```

B.
```
SELECT*
FROM OPENROWSET(
  BULK 'file1.txt',
  DATA_SOURCE = 'Sql1',
  FORMAT = 'CSV', PARSER_VERSION = '2.0',
  FIELDTERMINATOR = ',',
  ROWTERMINATOR = '\n',
  FIRSTROW= 2
)
```

C.
```
SELECT*
FROM OPENROWSET(
  BULK 'file1.txt',
  DATA_SOURCE = 'Sql1',
  FORMAT = 'CSV', PARSER_VERSION = '2.0',
  FIELDTERMINATOR = ',',
  ROWTERMINATOR = '0x0a',
  FIRSTROW= 2
)
```

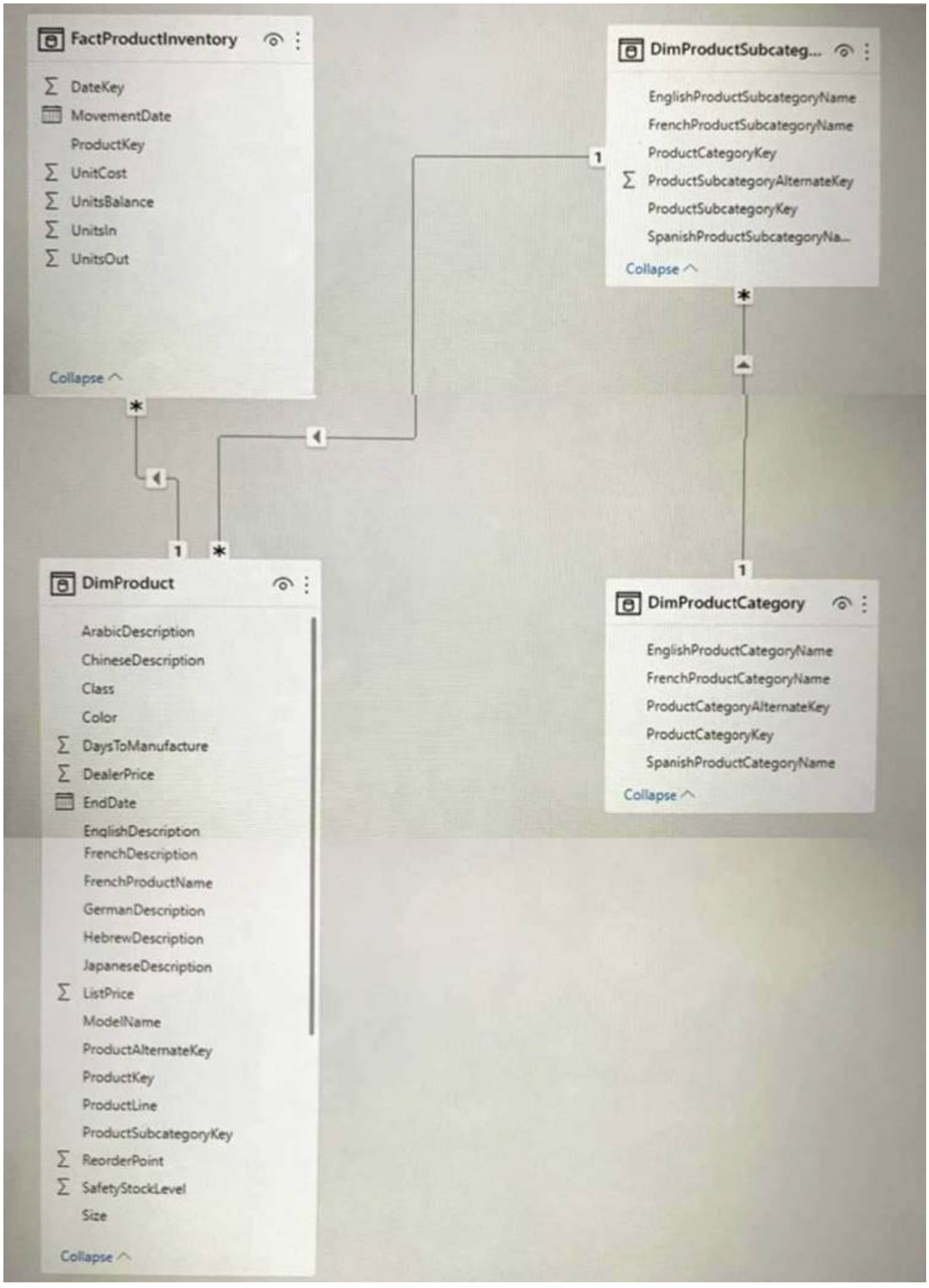
D.
```
SELECT*
FROM OPENROWSET(
  BULK 'file1.txt',
  DATA_SOURCE = 'Sql1',
  FORMAT = 'CSV', PARSER_VERSION = '2.0',
  FIELDTERMINATOR = '\t',
  ROWTERMINATOR = '0x0a',
  FIRSTROW= 1
)
```

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

Answer: A

Explanation:
Use FIELDTERMINATOR ='\t' for tab.
Use ROWTERMINATOR ='\0x0A ' for UNIX-style line endings Use FIRSTROW= 2 for a header row
Note: Using Row Terminators
The row terminator can be the same character as the terminator for the last field. Generally, however, a distinct row terminator is useful. For example, to produce tabular output, terminate the last field in each row with the newline character (\n) and all other fields with the tab character (\t).
If you want to output a line feed character only (LF) as the row terminator - as is typical on Unix and Linux computers - use hexadecimal notation to specify the LF row terminator. For example:
bcp -r '0x0A' FIRSTROW
FIRSTROW =first_row Specifies the number of the first row to load. The default is 1. This indicates the first row in the specified data file. The row numbers are determined by counting the row terminators. FIRSTROW
is 1-based.
Reference:
https://docs.microsoft.com/en-us/sql/relational-databases/import-export/specify-field-and-row-terminators-sql-se
https://docs.microsoft.com/en-us/sql/t-sql/functions/openrowset-transact-sql

NEW QUESTION 8
- (Exam Topic 3)
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 the Power Bl data model shown in the exhibit. (Click the Exhibit tab.)



Users indicate that when they build reports from the data model, the reports take a long time to load. You need to recommend a solution to reduce the load times of the reports.
Solution: You recommend normalizing the data model. Does this meet the goal?

A. Yes
B. No

Answer: B

Explanation:
Instead denormalize For Performance.
Even though it might mean storing a bit of redundant data, schema denormalization can sometimes provide better query performance. The only question then becomes is the extra space used worth the performance benefit.
Reference: https://www.mssqltips.com/sqlservertutorial/3211/denormalize-for-performance/

NEW QUESTION 9
- (Exam Topic 3)

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 are using an Azure Synapse Analytics serverless SQL pool to query a collection of Apache Parquet files by using automatic schema inference. The files contain more than 40 million rows of UTF-8-encoded business names, survey names, and participant counts. The database is configured to use the default collation.
The queries use open row set and infer the schema shown in the following table.

name	system_type_name	max_length
businessName	varchar(8000)	8000
surveyName	varchar(8000)	8000
participants	int	4

You need to recommend changes to the queries to reduce I/O reads and tempdb usage.
Solution: You recommend using openrowset with to explicitly define the collation for businessName and surveyName as Latim_Generai_100_BiN2_UTF8.
Does this meet the goal?

A. Yes
B. No

Answer: A

Explanation:
Query Parquet files using serverless SQL pool in Azure Synapse Analytics. Important
Ensure you are using a UTF-8 database collation (for example Latin1_General_100_BIN2_UTF8) because string values in PARQUET files are encoded using UTF-8 encoding. A mismatch between the text encoding in the PARQUET file and the collation may cause unexpected conversion errors. You can easily change the default collation of the current database using the following T-SQL statement: alter database current collate Latin1_General_100_BIN2_UTF8'.
Note: If you use the Latin1_General_100_BIN2_UTF8 collation you will get an additional performance boost compared to the other collations. The Latin1_General_100_BIN2_UTF8 collation is compatible with parquet string sorting rules. The SQL pool is able to eliminate some parts of the parquet files that will not contain data needed in the queries (file/column-segment pruning). If you use other collations, all data from the parquet files will be loaded into Synapse SQL and the filtering is happening within the SQL process. The Latin1_General_100_BIN2_UTF8 collation has additional performance optimization that works only for parquet and CosmosDB. The downside is that you lose fine-grained comparison rules like case insensitivity.
Reference: https://docs.microsoft.com/en-us/azure/synapse-analytics/sql/query-parquet-files

NEW QUESTION 10
- (Exam Topic 3)
You have a Power Bl workspace that contains one dataset and four reports that connect to the dataset. The dataset uses Import storage mode and contains the following data sources:
• A CSV file in an Azure Storage account
• An Azure Database for PostgreSQL database
You plan to use deployment pipelines to promote the content from development to test to production. There will be different data source locations for each stage.
What should you include in the deployment pipeline to ensure that the appropriate data source locations are used during each stage?

A. parameter rules
B. selective deployment
C. auto-binding across pipelines
D. data source rules

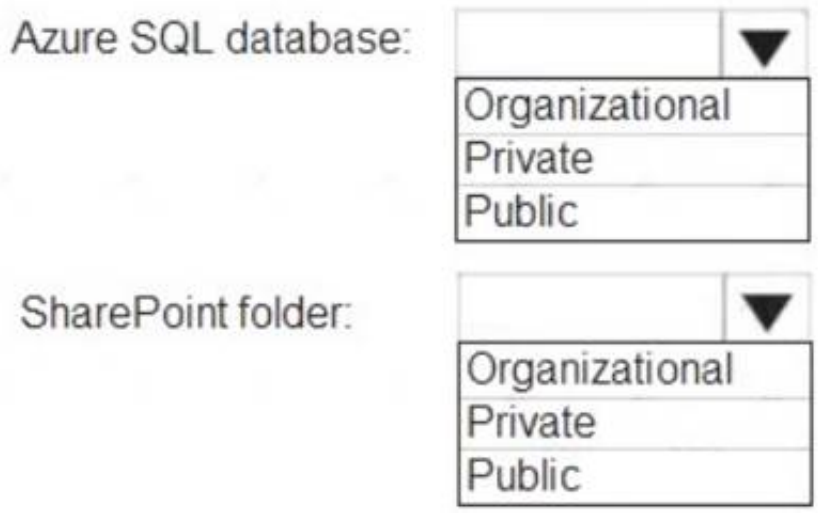
Answer: A

Explanation:
Note: 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.

NEW QUESTION 10
- (Exam Topic 3)
You manage a dataset that contains the two data sources as shown in the following table.

Data source	Type of data	Privacy level
Azure SQL database	Sensitive company data	Private
Microsoft SharePoint folder	Non-sensitive company data	Private

When you attempt to refresh the dataset in powerbi.com, you receive the following error message: “[Unable to combine data] Add Columns is accessing data sources that have privacy levels which cannot be used together. Please rebuild this data combination.”
You discover that the dataset contains queries that fold data from the SharePoint folder to the Azure SQL database.
You need to resolve the error. The solution must provide the highest privacy possible.
Which privacy level should you select for each data source? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.



A. Mastered
B. Not Mastered

Answer: A

Explanation:
Box 1: Private
This Formula.Firewall error is the result of Power Query's Data Privacy Firewall (aka the Firewall)
Note: Folding is a term that refers to converting expressions in M (such as filters, renames, joins, and so on) into operations against a raw data source (such as SQL, OData, and so on).
Box 2: Organizational
Organizational Limits the visibility of a data source to a trusted group of people. It is isolated from all Public data sources, but is visible to other Organizational data sources. A common example is a Microsoft Word document on an intranet SharePoint site with permissions enabled for a trusted group.
Reference:
https://support.microsoft.com/en-us/office/set-privacy-levels-power-query-cc3ede4d-359e-4b28-bc72-9bee7900

NEW QUESTION 11
- (Exam Topic 3)
You are creating an external table by using an Apache Spark pool in Azure Synapse Analytics. The table will contain more than 20 million rows partitioned by date. The table will be shared with the SQL engines.
You need to minimize how long it takes for a serverless SQL pool to execute a query data against the table. In which file format should you recommend storing the table data?

A. JSON
B. Apache Parquet
C. CSV
D. Delta

Answer: B

Explanation:
Prepare files for querying
If possible, you can prepare files for better performance:
* Convert large CSV and JSON files to Parquet. Parquet is a columnar format. Because it's compressed, its file sizes are smaller than CSV or JSON files that contain the same data. Serverless SQL pool skips the columns and rows that aren't needed in a query if you're reading Parquet files. Serverless SQL pool needs less time and fewer storage requests to read it.
Reference: https://docs.microsoft.com/en-us/azure/synapse-analytics/sql/best-practices-serverless-sql-pool https://stackoverflow.com/questions/65320949/parquet-vs-delta-format-in-azure-data-lake-gen-2-store

NEW QUESTION 13
- (Exam Topic 3)
You are using DAX Studio to analyze a slow-running report query. You need to identify inefficient join operations in the query. What should you review?

A. the query statistics
B. the query plan
C. the query history
D. the server timings

Answer: B

Explanation:

Open DAX Studio.
Paste the query there, enable Query Plan display and Server Timings, run your query (with clear cache), and then study the query plan for large row counts. Once the culprit is identified you can decide how to rewrite your DAX to make that part faster.
Reference: https://community.powerbi.com/t5/Power-Query/DAX-Query-taking-longer-time/td-p/1171961 https://www.sqlbi.com/wp-content/uploads/DAX-Query-Plans.pdf

NEW QUESTION 15
- (Exam Topic 3)
You plan to modify a Power BI dataset.
You open the Impact analysis panel for the dataset and select Notify contacts. Which contacts will be notified when you use the Notify contacts feature?

A. any users that accessed a report that uses the dataset within the last 30 days
B. the workspace admins of any workspace that uses the dataset
C. the Power Bl admins
D. all the workspace members of any workspace that uses the dataset

Answer: D

Explanation:
Notify contacts
If you've made a change to a dataset or are thinking about making a change, you might want to contact the relevant users to tell them about it. When you notify contacts, an email is sent to the contact lists of all the impacted workspaces. Your name appears on the email so the contacts can find you and reply back in a new email thread.
Reference: https://docs.microsoft.com/en-us/power-bi/collaborate-share/service-dataset-impact-analysis

NEW QUESTION 17
- (Exam Topic 3)
You are creating a Power 81 single-page report.
Some users will navigate the report by using a keyboard, and some users will navigate the report by using a screen reader.
You need to ensure that the users can consume content on a report page in a logical order. What should you configure on the report page?

A. the bookmark order
B. the X position
C. the layer order
D. the tab order

Answer: D

Explanation:
Tab order is the order in which users interact with the items on a page using the keyboard. Generally, we want tab order to be predictable and to closely match the visual order on the page (unless there is a good reason to deviate).
Note: If you are using the keyboard to navigate in a Power BI report, the order in which you arrive at visuals will not follow your vision unless you set the new tab order property. If you have low or no vision, this becomes an even bigger issue because you may not be able to see that you are navigating visuals out of visual order because the screen reader just reads whatever comes next.
Reference: https://datasavvy.me/2018/12/26/tab-order-enhances-power-bi-report-accessibility/

NEW QUESTION 22
- (Exam Topic 3)
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 Power Bl dataset named Dataset1.
In Dataset1, you currently have 50 measures that use the same time intelligence logic. You need to reduce the number of measures, while maintaining the current functionality. Solution: From Power Bl Desktop, you create a hierarchy.
Does this meet the goal?

A. Yes
B. No

Answer: B

Explanation:
Instead use the solution: From DAX Studio, you write a query that uses grouping sets. A grouping is a set of discrete values that are used to group measure fields.
Note: A hierarchy is an ordered set of values that are linked to the level above. An example of a hierarchy could be Country, State, and City. Cities are in a State, and States make up a Country. In Power BI visuals can handle hierarchy data and provide controls for the user to navigate up and down the hierarchy.
Reference: https://docs.microsoft.com/en-us/power-bi/developer/visuals/capabilities https://powerbi.tips/2018/09/how-to-navigate-hierarchies/

NEW QUESTION 23
- (Exam Topic 3)
You have an Azure Synapse Analytics serverless SQL pool and an Azure Data Lake Storage Gen2 account. You need to query all the files in the ‘csv/taxi/’ folder and all its subfolders. All the files are in CSV format
and have a header row.
How should you complete the query? To answer, select the appropriate options in the answer area.
NOTE: Each correct selection is worth one point.

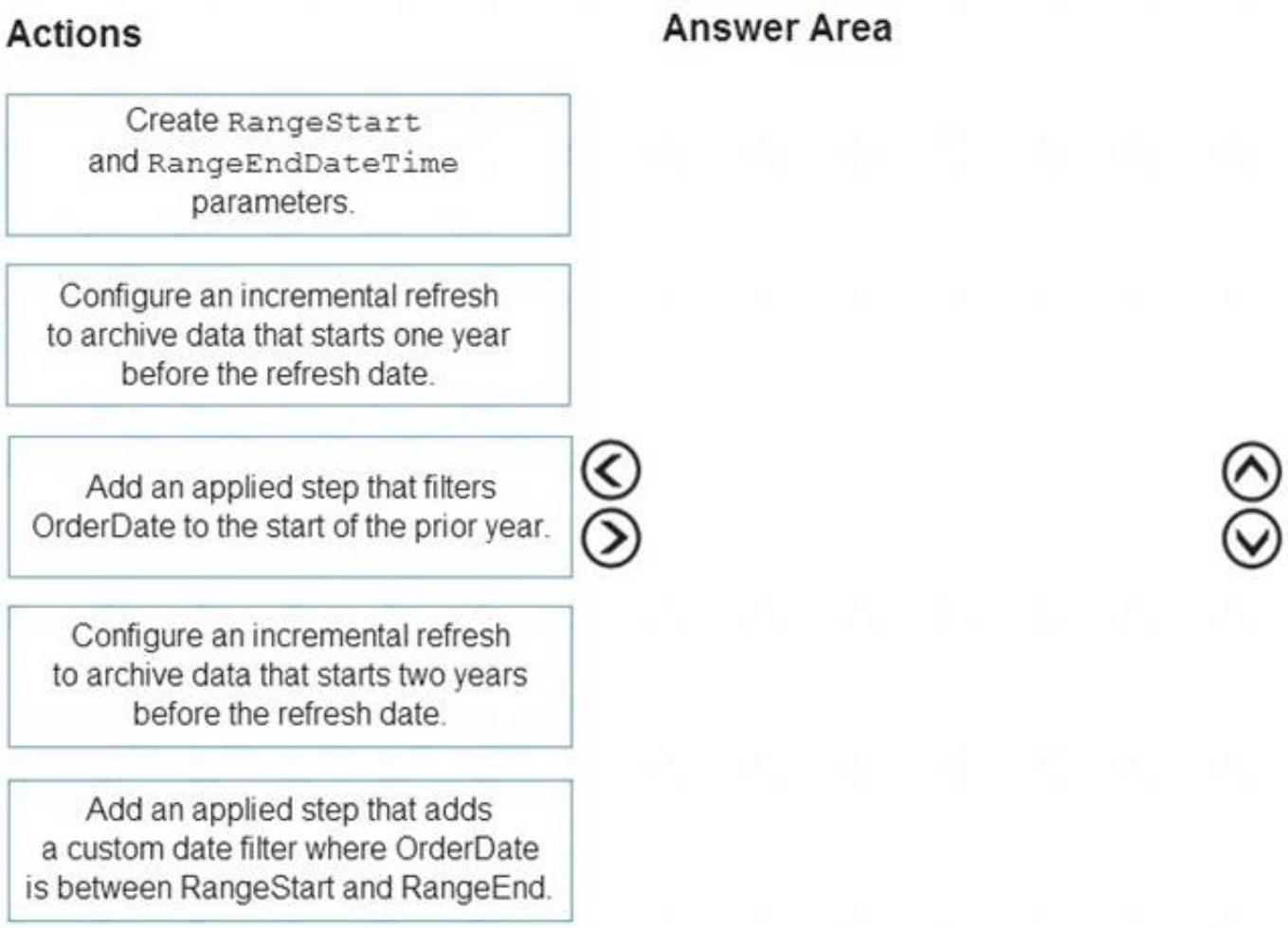
```
SELECT*
FROM OPENROWSET(
        [BULK 'csv/taxi', | BULK 'csv/taxi/**', | BULK 'csv/taxi/*.csv', | BULK 'csv/taxi/',]
        DATA_SOURCE = 'datalake',
        FORMAT = 'CSV', PARSER_VERSION = '2.0',
        [FIRSTROW = 0 | FIRSTROW = 1 | FIRSTROW = –1 | FIRSTROW = 2]
)
WITH (
      pickup_datetime DATETIME2,
      passenger_count INT,
      trip_distance FLOAT,
      total_amount FLOAT
) AS nyc;
```

A. Mastered
B. Not Mastered

Answer: A

Explanation:
Box 1: BULK 'csv/taxi*.CSV',
*.CSV to get all the CSV files. Box 2: FIRSTROW=2
As there is a header we should read from the second line. Note: FIRSTROW = 'first_row'
Specifies the number of the first row to load. The default is 1 and indicates the first row in the specified data file. The row numbers are determined by counting the row terminators. FIRSTROW is 1-based.
Incorrect:
Not FIRSTROW=1. FIRSTROW=1 is used when there is no header.
Reference: https://docs.microsoft.com/en-us/azure/synapse-analytics/sql/develop-openrowset

NEW QUESTION 28
- (Exam Topic 3)
You manage a Power BI dataset that queries a fact table named SalesDetails. SalesDetails contains three date columns named OrderDate, CreatedOnDate, and ModifiedDate.
You need to implement an incremental refresh of SalesDetails. The solution must ensure that OrderDate starts on or after the beginning of the prior year.
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.
NOTE: More than one order of answer choices is correct. You will receive credit for any of the correct orders you select.



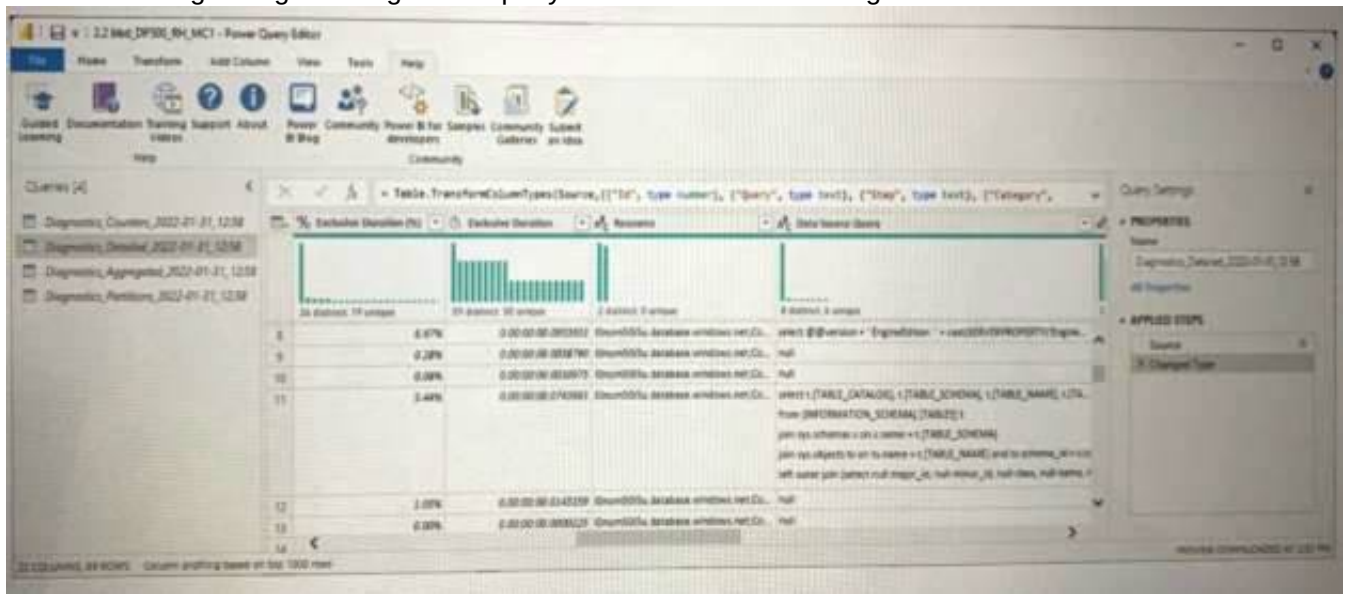
A. Mastered

B. Not Mastered

Answer: A

Explanation:
Graphical user interface, text, application, chat or text message Description automatically generated
Step 1: Create RangeStart and RangeEndDateTime parameters.
When configuring incremental refresh in Power BI Desktop, you first create two Power Query date/time parameters with the reserved, case-sensitive names RangeStart and RangeEnd. These parameters, defined in the Manage Parameters dialog in Power Query Editor are initially used to filter the data loaded into the Power BI Desktop model table to include only those rows with a date/time within that period.
Step 2: Add an applied step that adds a custom date filter OrderDate is Between RangeStart and RangeEnd. With RangeStart and RangeEnd parameters defined, you then apply custom Date filters on your table's date
column. The filters you apply select a subset of data that will be loaded into the model when you click Apply.
Step 3: Configure an incremental refresh to archive data that starts two years before the refresh date.
After filters have been applied and a subset of data has been loaded into the model, you then define an incremental refresh policy for the table. After the model is published to the service, the policy is used by the service to create and manage table partitions and perform refresh operations. To define the policy, you will use the Incremental refresh and real-time data dialog box to specify both required settings and optional settings.
Step 4: Add an applied step that filters OrderDate to the start of the prior year.
Reference: https://docs.microsoft.com/en-us/power-bi/connect-data/incremental-refresh-overview

NEW QUESTION 29
- (Exam Topic 3)
You are running a diagnostic against a query as shown in the following exhibit.



What can you identify from the diagnostics query?

A. All the query steps are folding.
B. Elevated permissions are being used to query records.
C. The query is timing out.
D. Some query steps are folding.

Answer: A

Explanation:
Understanding folding with Query Diagnostics
One of the most common reasons to use Query Diagnostics is to have a better understanding of what operations were 'pushed down' by Power Query to be performed by the back-end data source, which is also known as 'folding'. If we want to see what folded, we can look at what is the 'most specific' query, or queries, that get sent to the back-end data source. We can look at this for both ODATA and SQL.
Reference: https://docs.microsoft.com/en-us/power-query/querydiagnosticsfolding

NEW QUESTION 31
- (Exam Topic 3)
You have a Power Bl tenant that contains 10 workspaces.
You need to create dataflows in three of the workspaces. The solution must ensure that data engineers can access the resulting data by using Azure Data Factory.
Which two actions should you perform? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point

A. Associate the Power Bl tenant to an Azure Data Lake Storage account.
B. Add the managed identity for Data Factory as a member of the workspaces.
C. Create and save the dataflows to an Azure Data Lake Storage account.
D. Create and save the dataflows to the internal storage of Power BI

Answer: AC

Explanation:
Data used with Power BI is stored in internal storage provided by Power BI by default. With the integration of dataflows and Azure Data Lake Storage Gen 2 (ADLS Gen2), you can store your dataflows in your organization's Azure Data Lake Storage Gen2 account. This essentially allows you to "bring your own storage" to Power BI dataflows, and establish a connection at the tenant or workspace level.
Reference:
https://docs.microsoft.com/en-us/power-bi/transform-model/dataflows/dataflows-azure-data-lake-storage-integra

NEW QUESTION 34
- (Exam Topic 3)
You have a deployment pipeline for a Power BI workspace. The workspace contains two datasets that use import storage mode.
A database administrator reports a drastic increase in the number of queries sent from the Power Bi service to an Azure SQL database since the creation of the

deployment pipeline.
An investigation into the issue identifies the following:

- One of the datasets is larger than 1 GB and has a fact table that contains more than 500 million rows.
- When publishing dataset changes to development, test, or production pipelines, a refresh is triggered against the entire dataset.

You need to recommend a solution to reduce the size of the queries sent to the database when the dataset changes are published to development, test, or production.
What should you recommend?

A. Request the authors of the deployment pipeline datasets to reduce the number of datasets republished during development.
B. In the dataset, delete the fact table.
C. Configure the dataset to use a composite model that has a DirectQuery connection to the fact table.
D. From Capacity settings in the Power Bi Admin portal, reduce the Max Intermediate Row Set Count setting.

Answer: C

Explanation:
Previously in Power BI Desktop, when you used a DirectQuery in a report, no other data connections, whether DirectQuery or import, were allowed for that report. With composite models, that restriction is removed. A report can seamlessly include data connections from more than one DirectQuery or import data connection, in any combination you choose.
The composite models capability in Power BI Desktop consists of three related features:
* Composite models: Allows a report to have two or more data connections from different source groups, such as one or more DirectQuery connections and an import connection, two or more DirectQuery connections, or any combination thereof.
* Etc.
Reference: https://docs.microsoft.com/en-us/power-bi/transform-model/desktop-composite-models

NEW QUESTION 37
- (Exam Topic 3)
You have a 2-GB Power Bl dataset.
You need to ensure that you can redeploy the dataset by using Tabular Editor. The solution must minimize how long it will take to apply changes to the dataset from powerbi.com.
Which two actions should you perform in powerbi.com? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point

A. Enable service principal authentication for read-only admin APIs.
B. Turn on Large dataset storage format.
C. Connect the target workspace to an Azure Data Lake Storage Gen2 account.
D. Enable XMLA read-write.

Answer: BD

Explanation:
Optimize datasets for write operations by enabling large models
When using the XMLA endpoint for dataset management with write operations, it's recommended you enable the dataset for large models. This reduces the overhead of write operations, which can make them considerably faster. For datasets over 1 GB in size (after compression), the difference can be significant.
Tabular Editor supports Azure Analysis Services and Power BI Premium Datasets through XMLA read/write. Note: Tabular Editor - An open-source tool for creating, maintaining, and managing tabular models using an
intuitive, lightweight editor. A hierarchical view shows all objects in your tabular model. Objects are
organized by display folders with support for multi-select property editing and DAX syntax highlighting. XMLA read-only is required for query operations. Read-write is required for metadata operations.
Reference: https://docs.microsoft.com/en-us/power-bi/enterprise/service-premium-connect-tools https://tabulareditor.github.io/

NEW QUESTION 42
- (Exam Topic 3)
You are using GitHub as a source control solution for an Azure Synapse Studio workspace. You need to modify the source control solution to use an Azure DevOps Git repository. What should you do first?

A. Disconnect from the GitHub repository.
B. Create a new pull request.
C. Change the workspace to live mode.
D. Change the active branch.

Answer: A

Explanation:
By default, Synapse Studio authors directly against the Synapse service. If you have a need for collaboration using Git for source control, Synapse Studio allows you to associate your workspace with a Git repository, Azure DevOps, or GitHub.
Prerequisites
Users must have the Azure Contributor (Azure RBAC) or higher role on the Synapse workspace to configure, edit settings and disconnect a Git repository with Synapse.
Reference: https://docs.microsoft.com/en-us/azure/synapse-analytics/cicd/source-control

NEW QUESTION 46
- (Exam Topic 3)
You have an Azure Synapse Analytics serverless SQL pool.
You need to return a list of files and the number of rows in each file.
How should you complete the Transact-SQL statement? To answer, drag the appropriate values to the 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

- APPROX_COUNT_DISTINCT
- COUNT_BIG
- OPENDATASOURCE
- OPENJSON
- OPENQUERY
- OPENROWSET

Answer Area

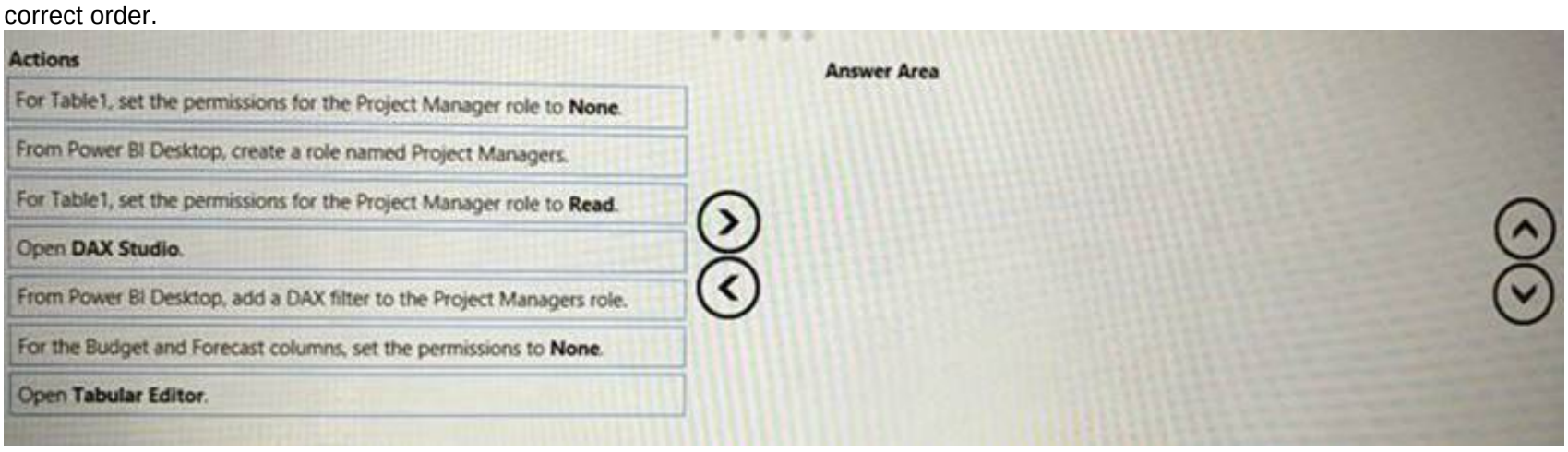
```
SELECT
    asa.filename() AS [filename]
    , [          ](*) AS [rows]
FROM
    [          ]
    BULK 'parquet/production/year=2017/month=9/*.parquet',
    DATA_SOURCE = 'DataLake1',
    FORMAT= 'PARQUET'
) asa
GROUP BY [filename]
```

A. Mastered
B. Not Mastered

Answer: A

Explanation:
Box 1: APPROX_COUNT_DISTINCT
The APPROX_COUNT_DISTINCT function returns the approximate number of unique non-null values in a group.
Box 2: OPENROWSET
OPENROWSET function in Synapse SQL reads the content of the file(s) from a data source. The data source is an Azure storage account and it can be explicitly referenced in the OPENROWSET function or can be dynamically inferred from URL of the files that you want to read. The OPENROWSET function can optionally contain a DATA_SOURCE parameter to specify the data source that contains files.
The OPENROWSET function can be referenced in the FROM clause of a query as if it were a table name OPENROWSET. It supports bulk operations through a built-in BULK provider that enables data from a file to be read and returned as a rowset.
Reference: https://docs.microsoft.com/en-us/sql/t-sql/functions/approx-count-distinct-transact-sql https://docs.microsoft.com/en-us/azure/synapse-analytics/sql/develop-openrowset

NEW QUESTION 50
- (Exam Topic 3)
You have a Power Bl dataset that contains two tables named Table1 and Table2. The dataset is used by one report.
You need to prevent project managers from accessing the data in two columns in Table1 named Budget and Forecast.
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.



A. Mastered
B. Not Mastered

Answer: A

Explanation:
Step 1: From Power BI Desktop, create a role named Project Managers. Create roles
You can define roles within Power BI Desktop. Step 2: Open Tabular Editor
Under Tables, select the table to which you want to apply a DAX rule.
In the Table filter DAX expression box, enter the DAX expressions. This expression returns a value of true or false. For example: [Entity ID] = “Value”.
Step 3: From Power BI Desktop, add a DAX filter to the Project Managers role. Step 4: For Table1, the Budget and Forecast columns, set the permissions to None.
Reference: https://docs.microsoft.com/en-us/power-bi/guidance/rls-guidance

NEW QUESTION 53
- (Exam Topic 3)
You are implementing a reporting solution that has the following requirements:
• Reports for external customers must support 500 concurrent requests. The data for these reports is approximately 7 GB and is stored in Azure Synapse Analytics.
• Reports for the security team use data that must have local security rules applied at the database level to restrict access. The data being reviewed is 2 GB.
Which storage mode provides the best response time for each group of users?

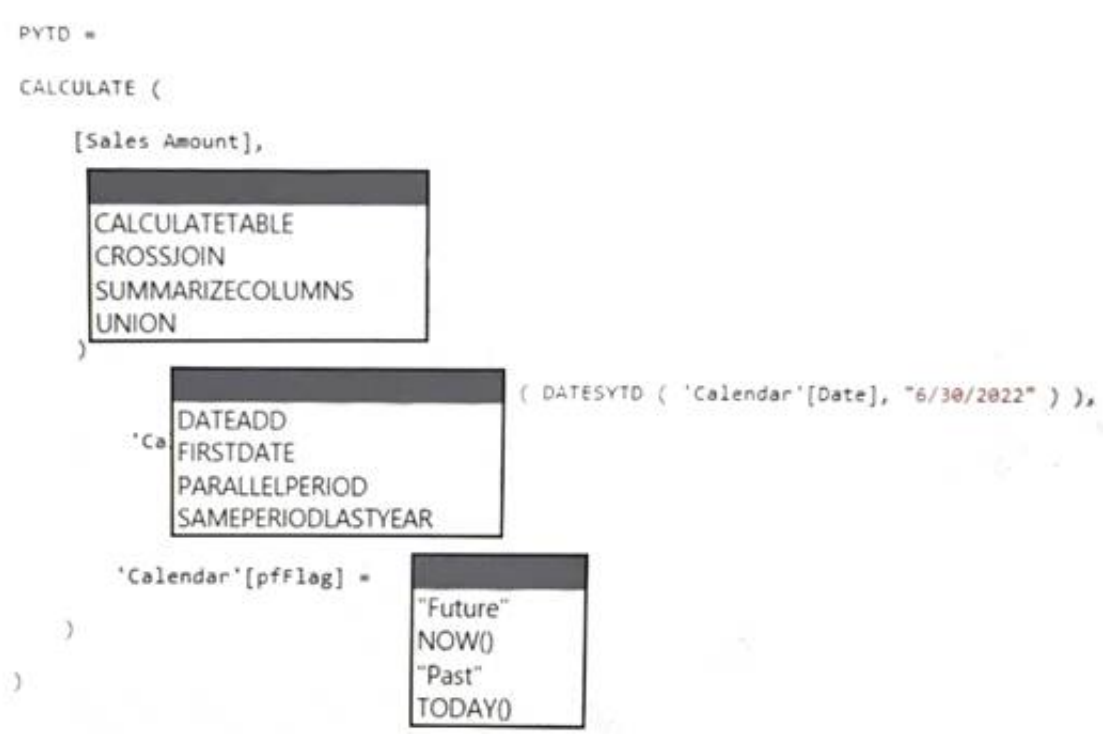
A. DirectQuery for the external customers and import for the security team.

B. DirectQuery for the external customers and DirectQuery for the security team.
C. Import for the external customers and DirectQuery for the security team.
D. Import for the external customers and import for the security team.

Answer: A

Explanation:
With DirectQuery, queries are sent back to your Azure Synapse Analytics in real time as you explore the data. Real-time queries, combined with the scale of Synapse Analytics enables users to create dynamic reports in minutes against terabytes of data.
Need import for the security team for local security rules. Reference:
https://docs.microsoft.com/en-us/power-bi/connect-data/service-azure-sql-data-warehouse-with-direct-connect

NEW QUESTION 56
- (Exam Topic 3)
You are building a Power Bl dataset that contains a table named Calendar. Calendar contains the following calculated column.
pfflag = IF('Calendar'[Date] < TOOAYQ, "Past", "Future")
You need to create a measure that will perform a fiscal prior year-to-date calculation that meets the following requirements:
• Returns the fiscal prior year-to-date value for [sales Amount]
• Uses a fiscal year end of June 30
• Produces no result for dates in the future
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



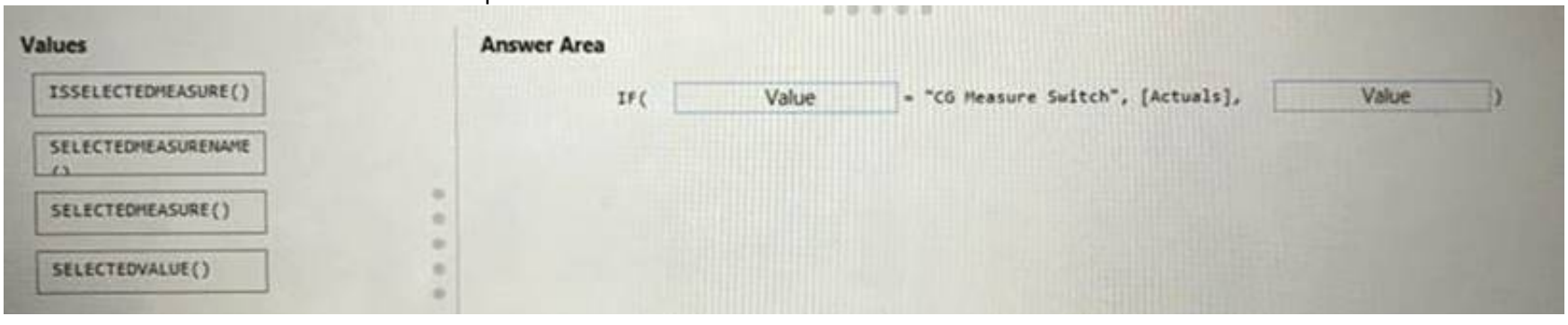
A. Mastered
B. Not Mastered

Answer: A

Explanation:
Box 1: CALCULATETABLE
CALCULATETABLE evaluates a table expression in a modified filter context. Syntax: CALCULATETABLE(<expression>[, <filter1> [, <filter2> [, …]]]) Incorrect:
* SUMMARIZECOLUMNS
SUMMARIZECOLUMNS returns a summary table over a set of groups.
Syntax: SUMMARIZECOLUMNS(<groupBy_columnName> [, < groupBy_columnName >]…, [<filterTable>]…[, <name>, <expression>]…)
* CROSSJOIN returns a table that contains the Cartesian product of all rows from all tables in the arguments. The columns in the new table are all the columns in all the argument tables.
Syntax: CROSSJOIN(<table>, <table>[, <table>]…)
* UNION creates a union (join) table from a pair of tables.
Syntax: UNION(<table_expression1>, <table_expression2> [,<table_expression>]…) Box 2: SAMEPERIODLASTYEAR
SAMEPERIODLASTYEAR returns a table that contains a column of dates shifted one year back in time from the dates in the specified dates column, in the current context.
Syntax: SAMEPERIODLASTYEAR(<dates>)
The dates returned are the same as the dates returned by this equivalent formula: DATEADD(dates, -1, year) Example:
The following sample formula creates a measure that calculates the previous year sales of Reseller sales.
= CALCULATE(SUM(ResellerSales_USD[SalesAmount_USD]), SAMEPERIODLASTYEAR(DateTime[DateKey]))
Box 3: TODAY()
TODAY() returns the current date.
The TODAY function is useful when you need to have the current date displayed on a worksheet, regardless of when you open the workbook. It is also useful for calculating intervals.
Example:
The following sample formula creates a measure that calculates the 'Running Total' for Internet sales.
= CALCULATE(SUM(InternetSales_USD[SalesAmount_USD]), DATESYTD(DateTime[DateKey])) Reference: https://docs.microsoft.com/en-us/dax/calculatetable-function-dax https://docs.microsoft.com/en-us/dax/sameperiodlastyear-function-dax
https://docs.microsoft.com/en-us/dax/datesytd-function-dax

NEW QUESTION 59
- (Exam Topic 3)
You have a Power Bl dataset that contains the following measures:

• Budget
• Actuals
• Forecast
You create a report that contains 10 visuals.
You need provide users with the ability to use a slicer to switch between the measures in two visuals only. You create a dedicated measure named cg Measure switch.
How should you complete the DAX expression for the Actuals measure? To answer, drag the appropriate values to the targets. Each value may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.
NOTE: Each correct selection is worth one point.



A. Mastered
B. Not Mastered

Answer: A

Explanation:
Box 1: SELECTEDMEASURENAME()
SELECTEDMEASURENAME is used by expressions for calculation items to determine the measure that is in context by name.
Syntax: SELECTEDMEASURENAME()
No parameters. Example:
The following calculation item expression checks if the current measure is Expense Ratio and conditionally applies calculation logic. Since the check is based on a string comparison, it is not subject to formula fixup and will not benefit from object renaming being automatically reflected. For a similar comparison that would benefit from formula fixup, please see the ISSLECTEDMEASURE function instead.
IF (
SELECTEDMEASURENAME = "Expense Ratio", SELECTEDMEASURE (),
DIVIDE (SELECTEDMEASURE (), COUNTROWS (DimDate))
)
Box 2: SELECTEDVALUE()
SELECTEDVALUE returns the value when the context for columnName has been filtered down to one distinct value only. Otherwise returns alternateResult.
Syntax:
SELECTEDVALUE(<columnName>[, <alternateResult>]) M1, M2, ... - A list of measures.
Reference: https://docs.microsoft.com/en-us/dax/selectedmeasurename-function-dax https://docs.microsoft.com/en-us/dax/selectedvalue-function

NEW QUESTION 62
- (Exam Topic 3)
You need to provide users with a reproducible method to connect to a data source and transform the data by using an Al function. The solution must meet the following requirement
• Minimize development effort.
• Avoid including data in the file. Which type of file should you create?

A. PBIDS
B. PBIX
C. PBIT

Answer: C

Explanation:
A PBIT file is a template created by Power BI Desktop, a Microsoft application used to create reports and visualizations. It contains queries, visualization settings, data models, reports, and other data added by the user.
A PBIT file acts as a Power BI template. It doesn't include any data from your source systems. Reference: https://docs.microsoft.com/en-us/power-bi/connect-data/desktop-data-sources

NEW QUESTION 63
- (Exam Topic 3)
You have a Power Bl dataset that contains the following measure.

```
YTD Year-over-Year Var =
DIVIDE (
  (
    [Sales Amount]
      - CALCULATE (
        [Sales],
        SAMEPERIODLASTYEAR ( 'Calendar'[Date] ),
        'Calendar'[Flag] = "YTD"
      )
  ),
  CALCULATE (
    [Sales],
    SAMEPERIODLASTYEAR ( 'Calendar'[Date] ),
    'Calendar'[Flag] = "YTD"
  ),
  BLANK()
)
```

You need to improve the performance of the measure without affecting the logic or the results. What should you do?

A. Replace both calculate functions by using a variable that contains the calculate function.
B. Remove the alternative result of blank( ) from the divide function.
C. Create a variable and replace the values for [sales Amount].
D. Remove "calendar'[Flag] = "YTD" from the code.

Answer: A

NEW QUESTION 66
- (Exam 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. rollup
B. cov
C. explain
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

NEW QUESTION 71
- (Exam Topic 3)
You have the following Python code in an Apache Spark notebook.

```
import matplotlib.pyplot as plt
import numpy as np
ys = 300 + np.random.randn(100)
x = [x for x in range(len(ys))]
plt.plot(x, ys, '-')
plt.fill_between(x, ys, 395, where=(ys > 395), facecolor='g', alpha=0.5)
plt.title("Chart Sample")
plt.show()
```

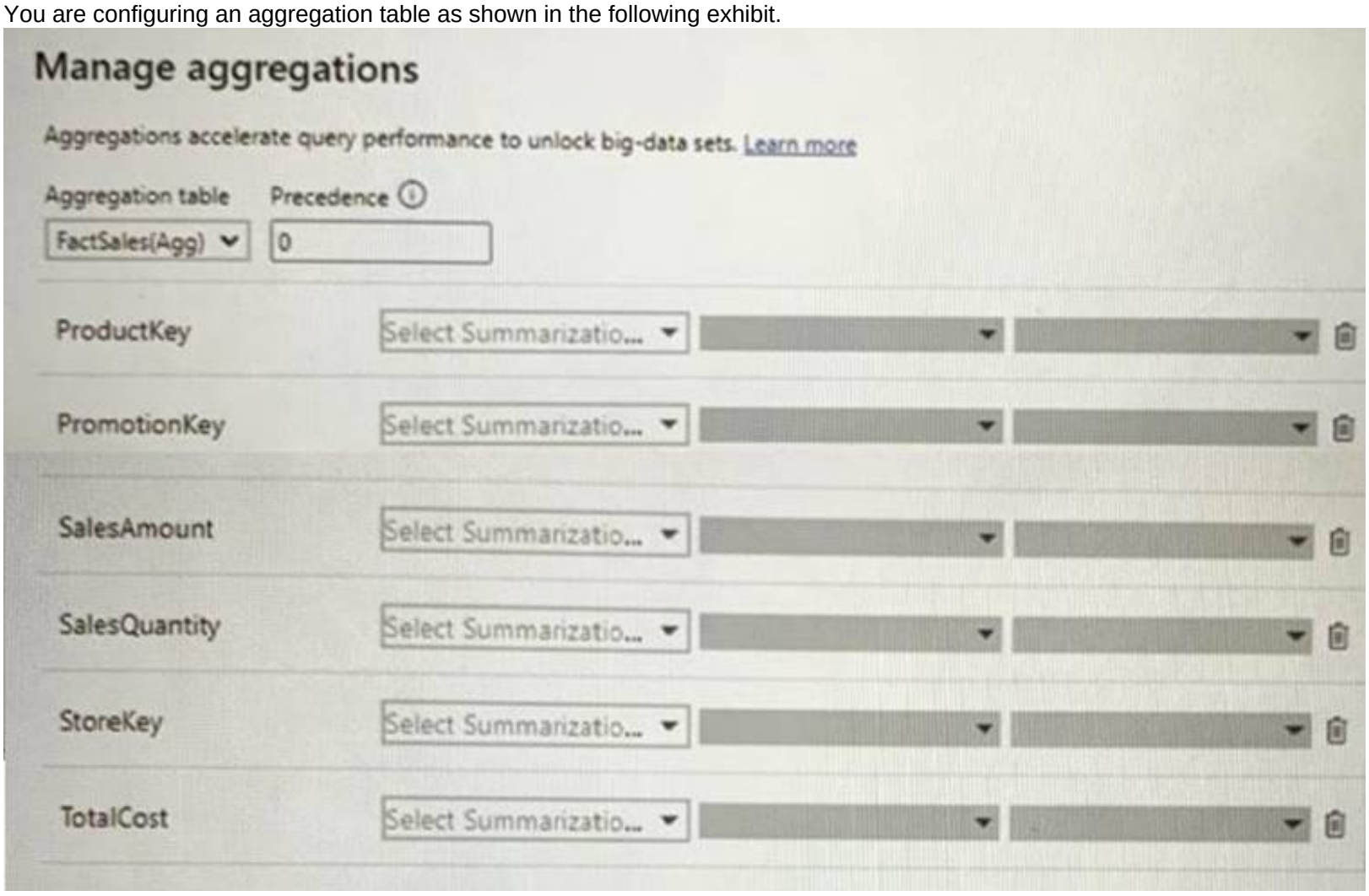
Which type of chart will the code produce?

A. a stacked bar chart
B. a pie chart
C. a bar chart
D. an area chart

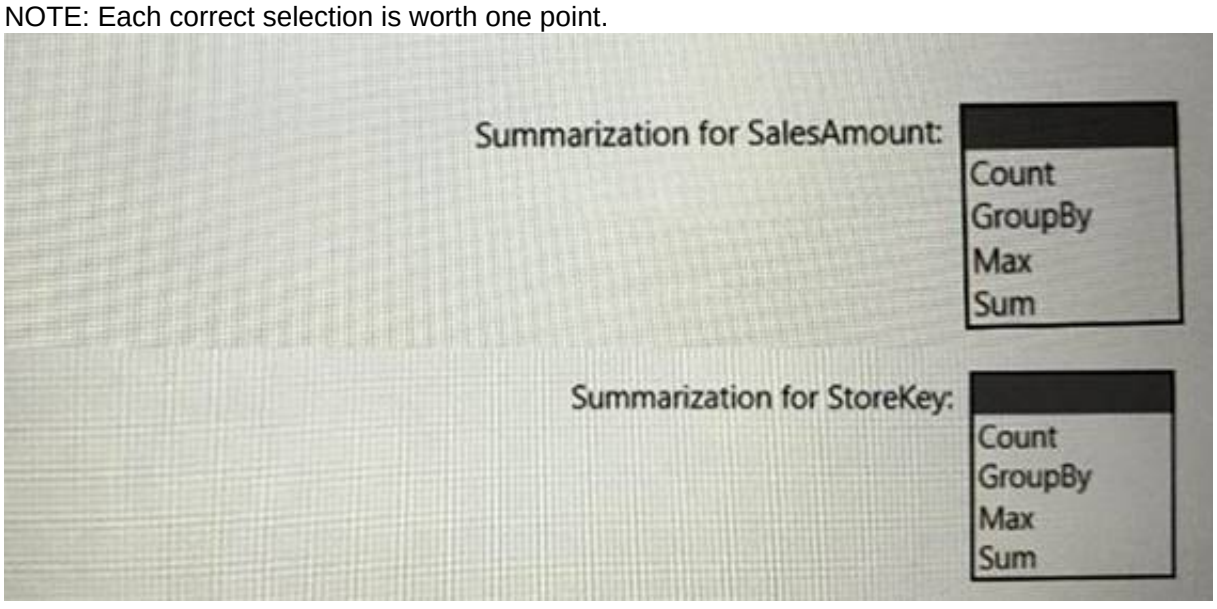
Answer: D

Explanation:
The matplotlib.pyplot.fill_between function fills the area between two horizontal curves.
The curves are defined by the points (x, y1) and (x, y2). This creates one or multiple polygons describing the filled area.
Reference: https://matplotlib.org/3.5.0/api/_as_gen/matplotlib.pyplot.fill_between.html

NEW QUESTION 75
- (Exam Topic 3)
You are configuring an aggregation table as shown in the following exhibit.



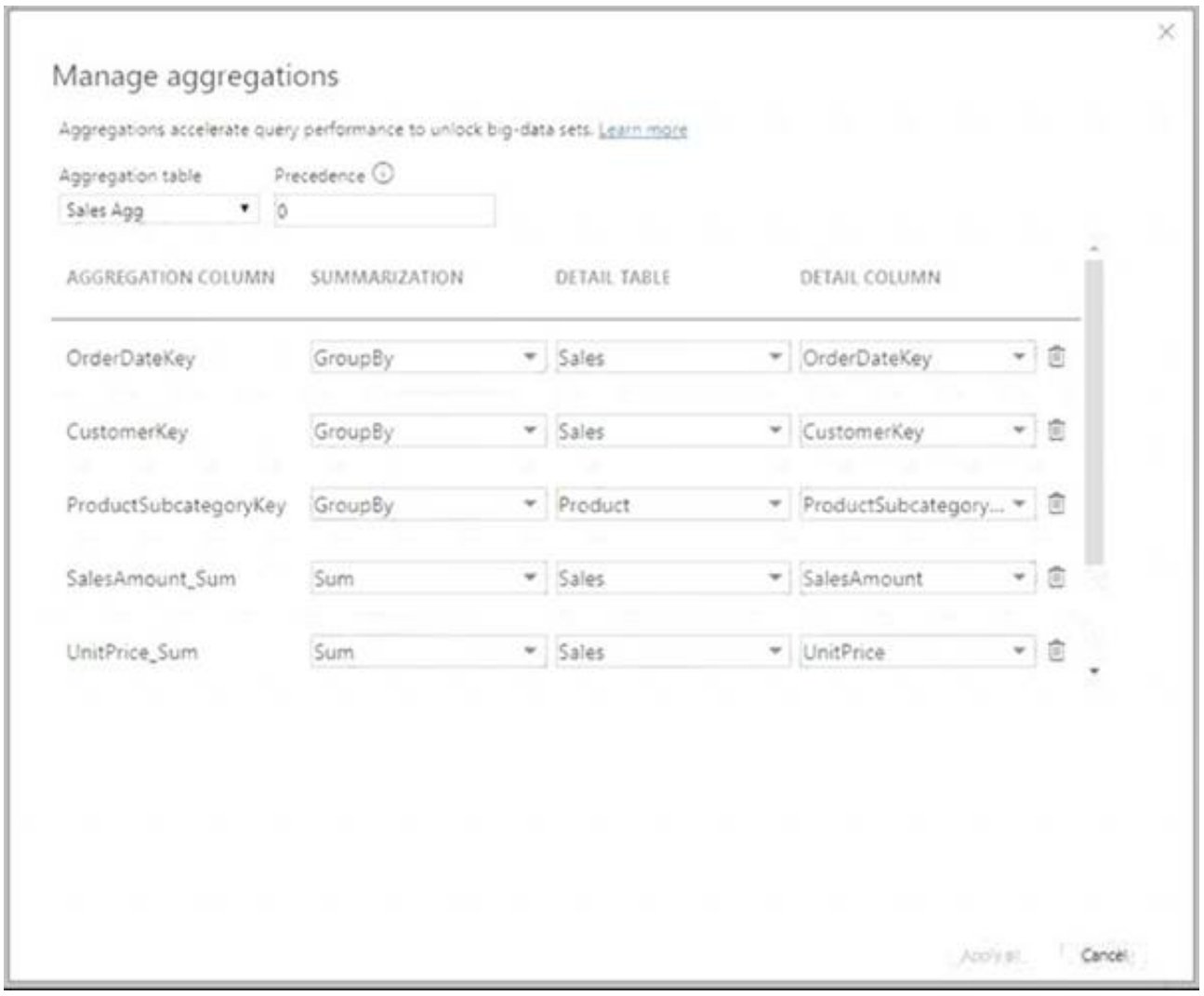
The detail table is named FactSales and the aggregation table is named FactSales(Agg). You need to aggregate SalesAmount for each store.
Which type of summarization should you use for SalesAmount and StoreKey? To answer, select the appropriate options in the answer area,
NOTE: Each correct selection is worth one point.



A. Mastered
B. Not Mastered

Answer: A

Explanation:
Box 1: Sum
The Manage aggregations dialog shows a row for each column in the table, where you can specify the aggregation behavior. In the following example, queries to the Sales detail table are internally redirected to the Sales Agg aggregation table.



Box 2: GroupBy
Reference: https://docs.microsoft.com/en-us/power-bi/transform-model/aggregations-advanced

NEW QUESTION 79
- (Exam Topic 3)
You are optimizing a dataflow in a Power Bl Premium capacity. The dataflow performs multiple joins. You need to reduce the load time of the dataflow.
Which two actions should you perform? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point.

A. Reduce the memory assigned to the dataflows.
B. Execute non-foldable operations before foldable operations.
C. Execute foldable operations before non-foldable operations.
D. Place the ingestion operations and transformation operations in a single dataflow.
E. Place the ingestion operations and transformation operations in separate dataflows.

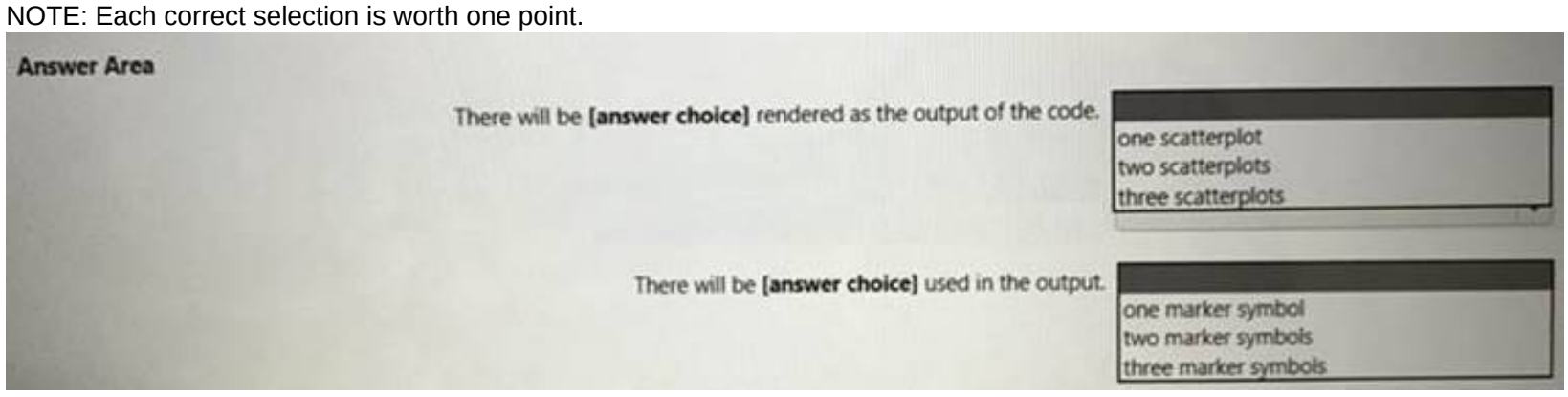
Answer: CE

Explanation:
Using the compute engine to improve performance
Take the following steps to enable workloads trigger the compute engine, and always improve performance: For computed and linked entities in the same workspace:
Ensure you perform the operations that fold, such as merges, joins, conversion, and others.
For ingestion focus on getting the data into the storage as fast as possible, using filters only if they reduce the overall dataset size. It's best practice to keep your transformation logic separate from this step, and allow the engine to focus on the initial gathering of ingredients. Next, separate your transformation and business logic into a separate dataflow in the same workspace, using linked or computed entities; doing so allows for the engine to activate and accelerate your computations. In our analogy, it's like food preparation in the kitchen: food preparation is typically a separate and distinct step from gathering your raw ingredients, and a pre-requisite for putting the food in the oven. Similarly, your logic needs to be prepared separately before it can take advantage of the compute engine.
Reference:
https://docs.microsoft.com/en-us/power-bi/transform-model/dataflows/dataflows-premium-workload-configurati

NEW QUESTION 82
- (Exam Topic 3)
You have the following code in an Azure Synapse notebook.

```
import matplotlib.pyplot as plt
x1 = [2, 3, 4]
y1 = [5, 5, 5]
x2 = [1, 2, 3, 4, 5]
y2 = [2, 3, 2, 3, 4]
y3 = [6, 8, 7, 8, 7]
plt.scatter(x1, y1)
plt.scatter(x2, y2, marker='v', color='r')
plt.scatter(x2, y3, marker='^', color='m')
plt.title('Scatter Plot')
plt.show()
```

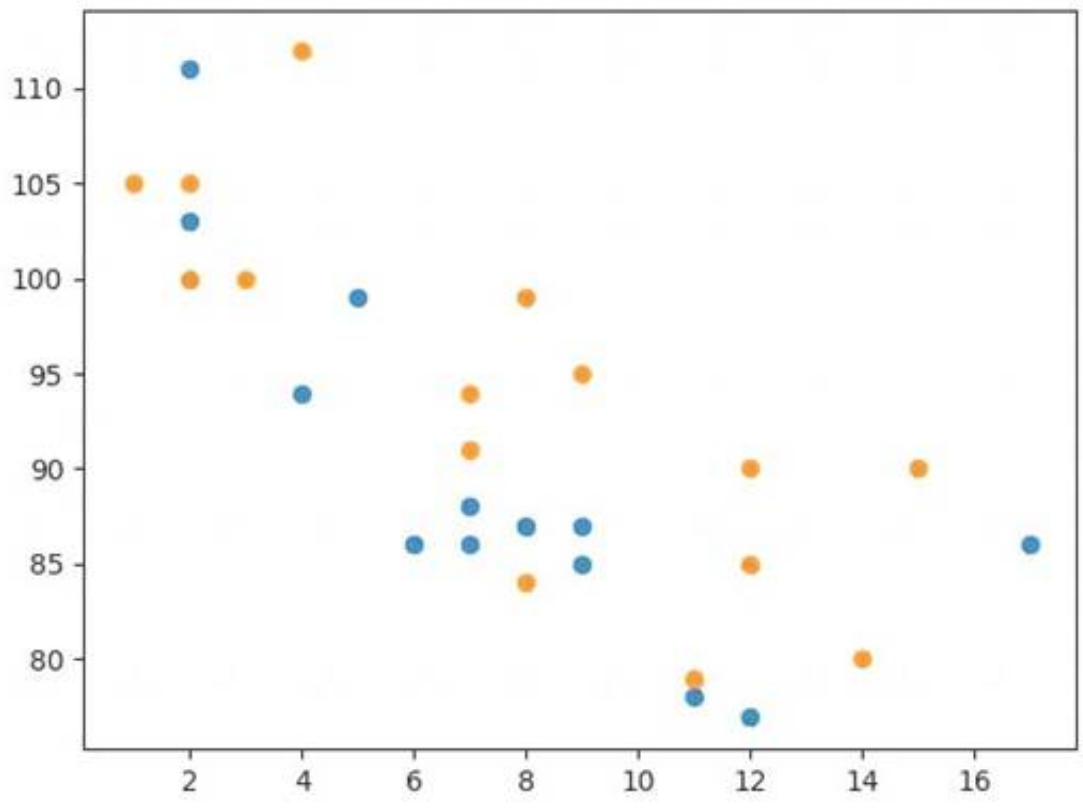
Use the drop-down menus to select the answer choice that completes each statement based on the information presented in the code.
NOTE: Each correct selection is worth one point.



A. Mastered
B. Not Mastered

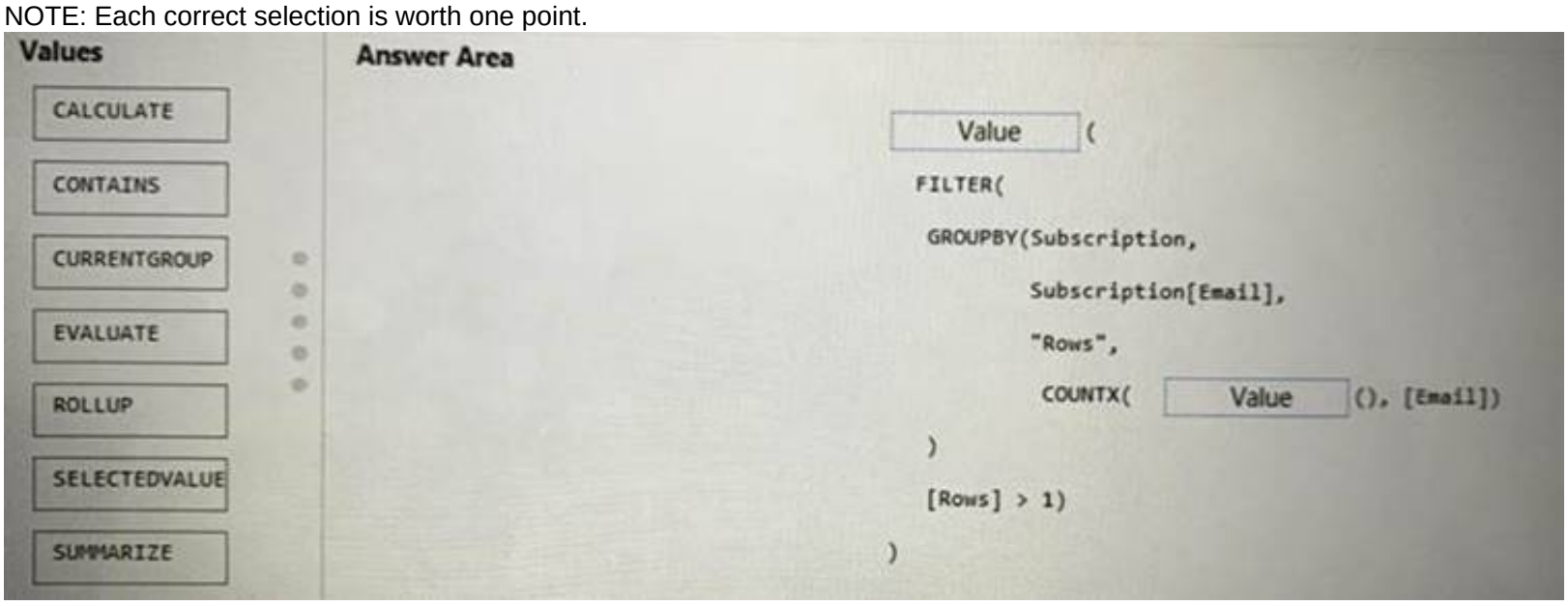
Answer: A

Explanation:
Box 1: three scatterplots Compare Plots
Example, Draw two plots on the same figure: import matplotlib.pyplot as plt
import numpy as np
#day one, the age and speed of 13 cars:
x = np.array([5,7,8,7,2,17,2,9,4,11,12,9,6])
y = np.array([99,86,87,88,111,86,103,87,94,78,77,85,86])
plt.scatter(x, y)
#day two, the age and speed of 15 cars:
x = np.array([2,2,8,1,15,8,12,9,7,3,11,4,7,14,12])
y = np.array([100,105,84,105,90,99,90,95,94,100,79,112,91,80,85])
plt.scatter(x, y) plt.show() Result:
Chart, scatter chart Description automatically generated



Box 2: three marker symbols
One for each scatterplot. One default, and two defined.
Default is point.
v is triangle down.
^ is triangle up.
Reference: https://www.w3schools.com/python/matplotlib_scatter.asp https://matplotlib.org/stable/api/markers_api.html

NEW QUESTION 84
- (Exam Topic 3)
You are using DAX Studio to query an XMLA endpoint.
You need to identify the duplicate values in a column named Email in a table named Subscription.
How should you complete the DAX expression? To answer, drag the appropriate values to the targets. Each value may be used once, more than once. may need to drag the split bar between panes or scroll to view content.
NOTE: Each correct selection is worth one point.



A. Mastered
B. Not Mastered

Answer: A

Explanation:
Box 1: CALCULATE
Box 2: CURRENTGROUP
CURRENTGROUP returns a set of rows from the table argument of a GROUPBY expression that belong to the current row of the GROUPBY result.
Remarks
This function can only be used within a GROUPBY expression.
This function takes no arguments and is only supported as the first argument to one of the following aggregation functions: AVERAGEX, COUNTAX, COUNTX, GEOMEANX, MAXX, MINX, PRODUCTX, STDEVX.S, STDEVX.P, SUMX, VARX.S, VARX.P.
Note: COUNTX counts the number of rows that contain a non-blank value or an expression that evaluates to a non-blank value, when evaluating an expression over a table.
Reference: https://docs.microsoft.com/en-us/dax/currentgroup-function-dax

NEW QUESTION 86
- (Exam Topic 3)
You are using an Azure Synapse Analytics serverless SQL pool to query network traffic logs in the Apache Parquet format. A sample of the data is shown in the following table.

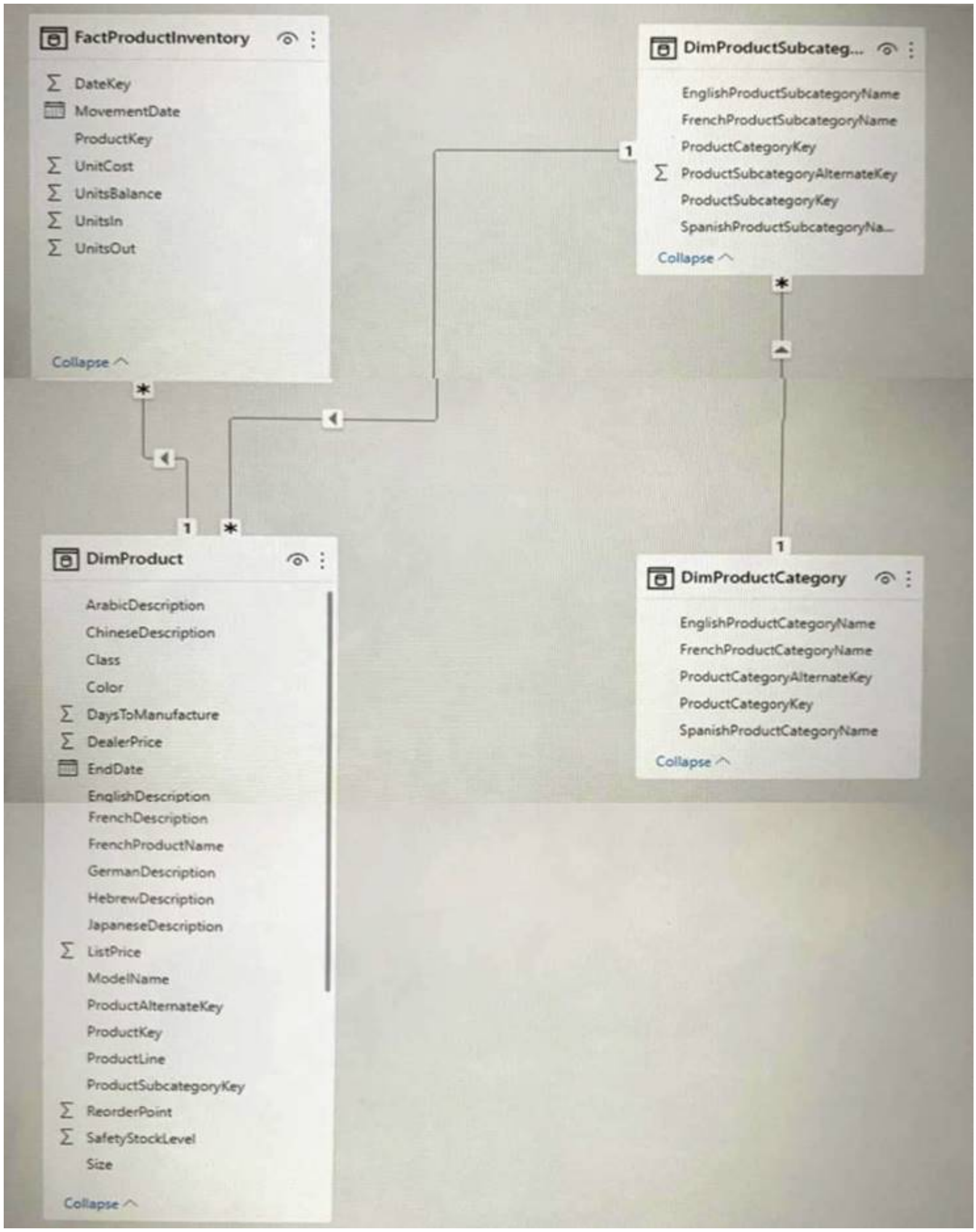
source		destination	
name	**ip**	**name**	**ip**
Network01	192.168.0.1	Internet	0.0.0.0

You need to create a Transact-SQL query that will return the source IP address.
Which function should you use in the select statement to retrieve the source IP address?

A. JS0N_VALUE
B. FOR.JSON
C. CONVERT
D. FIRST VALUE

Answer: A

NEW QUESTION 88
- (Exam Topic 3)
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 the Power BI data model shown in the exhibit (Click the Exhibit tab.)



Users indicate that when they build reports from the data model, the reports take a long time to load. You need to recommend a solution to reduce the load times of the reports.
Solution: You recommend denormalizing the data model. Does this meet the goal?

A. Yes
B. No

Answer: A

Explanation:
Denormalize For Performance.
Even though it might mean storing a bit of redundant data, schema denormalization can sometimes provide better query performance. The only question then becomes is the extra space used worth the performance benefit.
Reference: https://www.mssqltips.com/sqlservertutorial/3211/denormalize-for-performance/

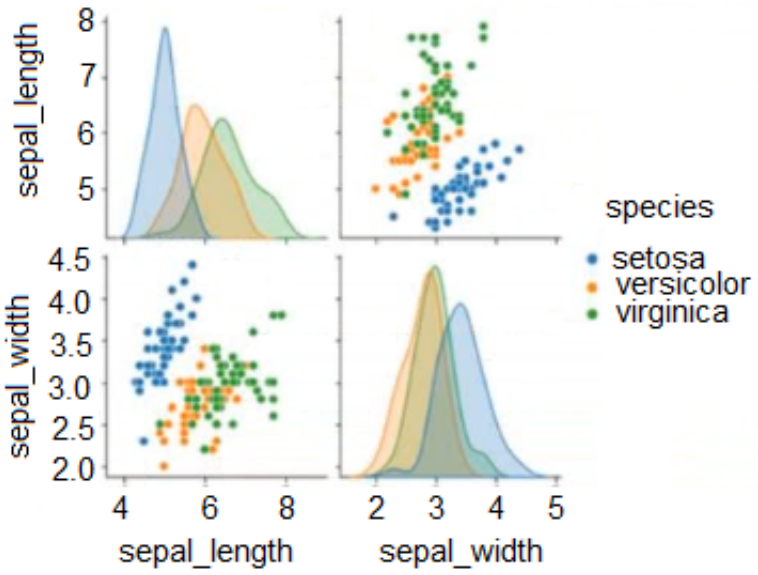
NEW QUESTION 92
- (Exam Topic 3)
You are using an Azure Synapse notebook to create a Python visual. You run the following code cell to import a dataset named Iris.

```
iris = sns.load_dataset("iris")
iris.head()
```

A sample of the data is shown in the following table.

index	sepal_length	sepal_width	species
0	5.1	3.5	setosa
2	4.9	3	setosa
145	6.7	3	virginica
156	6.3	2.5	virginica

You need to create the visual shown in the exhibit. (Click the Exhibit tab.)



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

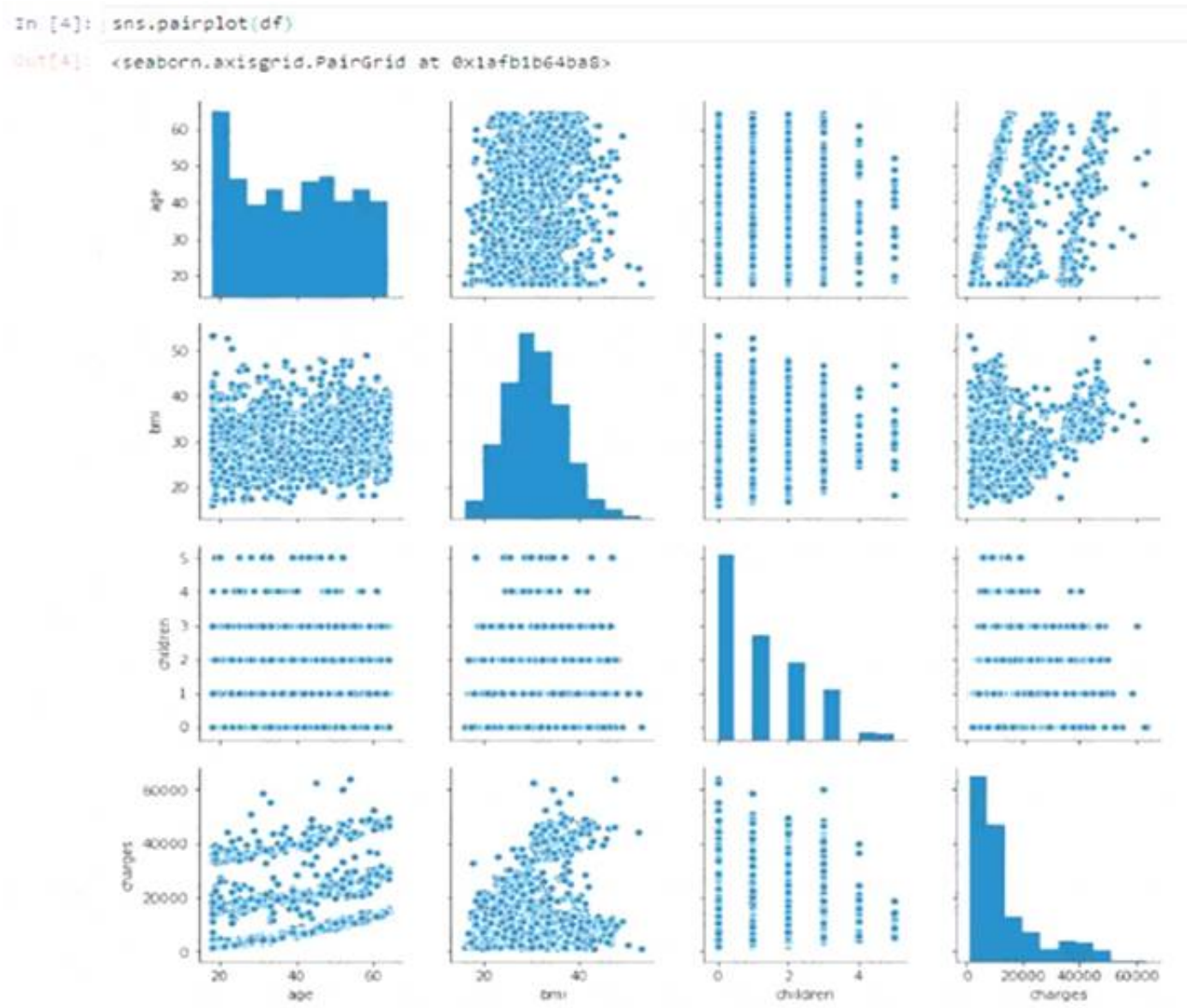
Answer Area

```
sns. [boxplot / implot / pairplot / replot] (iris, hue= ' [sepal_length / sepal_width / species] ', height=2.5)
plt.show()
```

A. Mastered
B. Not Mastered

Answer: A

Explanation:
Box 1: pairplot
A pairs plot allows us to see both distribution of single variables and relationships between two variables. Pair plots are a great method to identify trends for follow-up analysis and, fortunately, are easily implemented in Python!
Example, let's plot data using pairplot:
From the picture below, we can observe the variations in each plot. The plots are in matrix format where the row name represents x axis and column name represents the y axis. The main-diagonal subplots are the univariate histograms (distributions) for each attribute.
A picture containing diagram Description automatically generated



Box 2: sepal_width
sepal_width is displayed with a height of 2.5 (between 2.0 and 4.5).
Reference: https://medium.com/analytics-vidhya/pairplot-visualization-16325cd725e6

NEW QUESTION 93
- (Exam 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 Area

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.	○	○

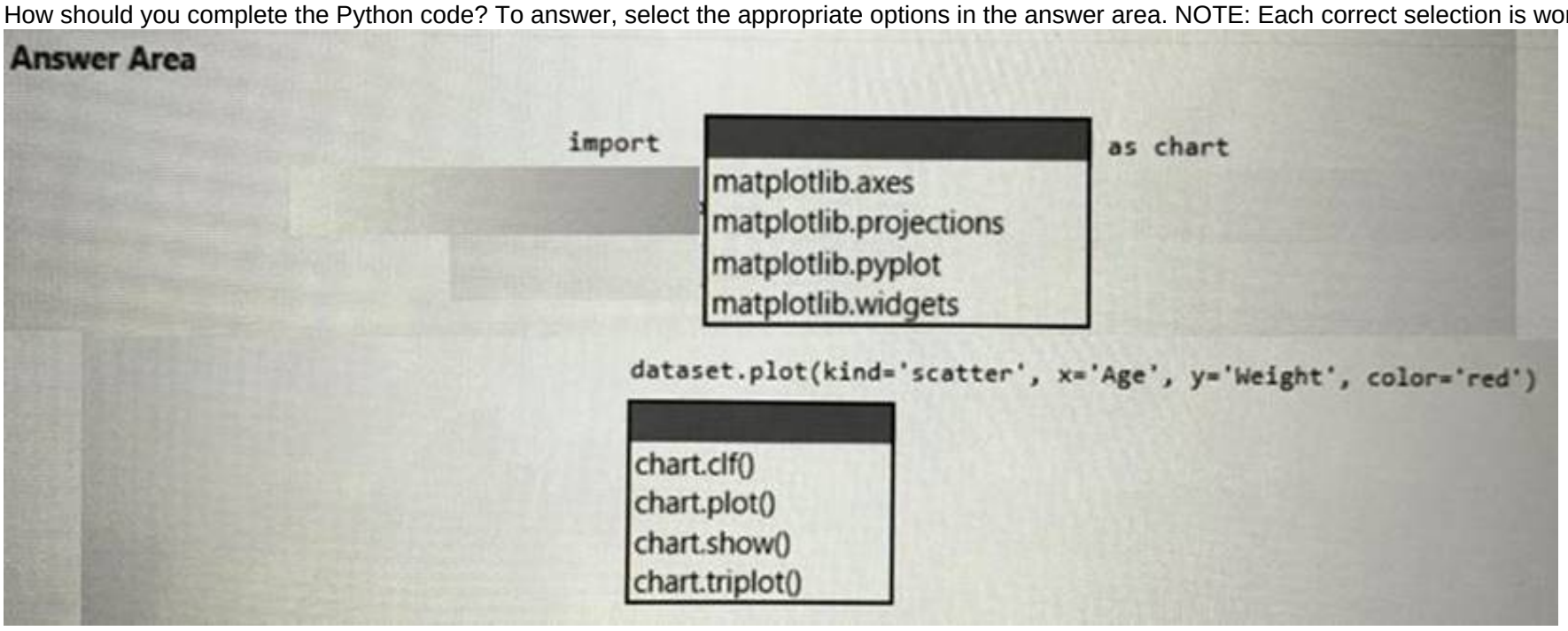
A. Mastered
B. Not Mastered

Answer: A

Explanation:
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
Reference: https://docs.microsoft.com/en-us/azure/storage/blobs/data-lake-storage-namespace https://docs.microsoft.com/en-us/powerquery-m/table-selectrows https://docs.microsoft.com/en-us/azure/data-lake-store/data-lake-store-comparison-with-blob-storage

NEW QUESTION 98
- (Exam Topic 3)
You are creating a Power BI Desktop report. You add a Python visual to the report page.
You plan to create a scatter chart to visualize the data. You add Python code to the Python script editor.
You need to create the scatter chart.
How should you complete the Python code? To answer, select the appropriate options in the answer area. NOTE: Each correct selection is worth one point.



A. Mastered
B. Not Mastered

Answer: A

Explanation:
Box 1: matplotlib.pyplot
Create a scatter plot
Let's create a scatter plot to see if there's a correlation between age and weight. Under Paste or type your script code here, enter this code:
import matplotlib.pyplot as plt
dataset.plot(kind='scatter', x='Age', y='Weight', color='red') plt.show()
Box 2: chart.show()
Reference:
https://docs.microsoft.com/en-us/power-bi/connect-data/desktop-python-visuals#create-a-scatter-plot

NEW QUESTION 102
- (Exam Topic 3)
You have a dataset that contains a table named UserPermissions. UserPermissions contains the following data.

User	Region
CONTOSO\User1	1
CONTOSO\User2	2
CONTOSO\User3	1
CONTOSO\User4	3
CONTOSO\User4	5

You plan to create a security role named User Security for the dataset. You need to filter the dataset based on the current users. What should you include in the DAX expression?

A. [UserPermissions] - USERNAME()
B. [UserPermissions] - USERPRINCIPALNAME()
C. [User] = USERPRINCIPALNAME()
D. [User] = USERNAME()
E. [User] = USEROBJECTID()

Answer: D

Explanation:
USERNAME() returns the domain name and username from the credentials given to the system at connection time.
It should be compared to column name of User, which in DAX is expressed through [User]. Reference: https://docs.microsoft.com/en-us/dax/username-function-dax

NEW QUESTION 103
- (Exam Topic 3)
You have a deployment pipeline for a Power BI workspace. The workspace contains two datasets that use import storage mode.
A database administrator reports a drastic increase in the number of queries sent from the Power BI service to an Azure SQL database since the creation of the deployment pipeline.
An investigation into the issue identifies the following:
One of the datasets is larger than 1 GB and has a fact table that contains more than 500 million rows.
When publishing dataset changes to development, test, or production pipelines, a refresh is triggered against the entire dataset.
You need to recommend a solution to reduce the size of the queries sent to the database when the dataset changes are published to development, test, or production.
What should you recommend?

A. Turn off auto refresh when publishing the dataset changes to the Power BI service.
B. In the datase
C. change the fact table from an import table to a hybrid table.
D. Enable the large dataset storage format for workspace.
E. Create a dataset parameter to reduce the fact table row count in the development and test pipelines.

Answer: B

Explanation:
Hybrid tables
Hybrid tables are tables with incremental refresh that can have both import and direct query partitions. During a clean deployment, both the refresh policy and the hybrid table partitions are copied. When deploying to a pipeline stage that already has hybrid table partitions, only the refresh policy is copied. To update the partitions, refresh the table.
Refreshes are faster - Only the most recent data that has changed needs to be refreshed.
Reference: https://docs.microsoft.com/en-us/power-bi/create-reports/deployment-pipelines-best-practices

NEW QUESTION 106
- (Exam 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 [ ▼ ] (MODEL = @model, DATA = dbo.rx_linMod as lm)
go [ ▼ ] (model_outcome float, trade_volume float, price_Pred float) as p;
```

FROM: EVALUATE | PIVOT | PREDICT | SCORE
go: AS | CONTAINS | FROM | GROUP BY | WITH

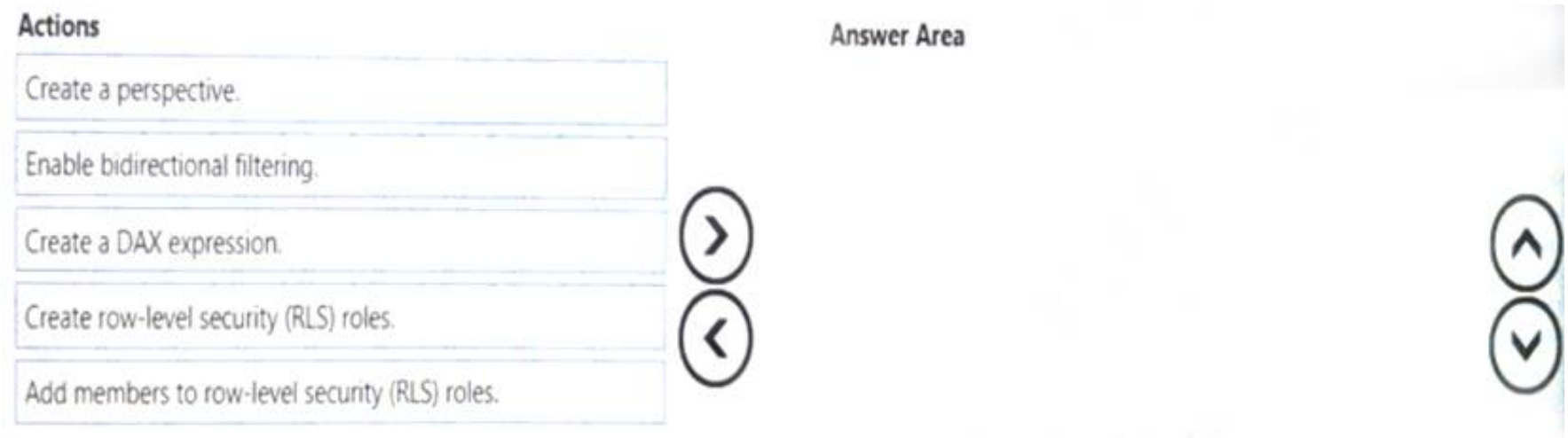
A. Mastered
B. Not Mastered

Answer: A

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
Reference:
https://docs.microsoft.com/en-us/azure/synapse-analytics/sql-data-warehouse/sql-data-warehouse-predict

NEW QUESTION 108

- (Exam 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.



A. Mastered
B. Not Mastered

Answer: A

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. Reference: https://docs.microsoft.com/en-us/power-bi/guidance/rls-guidance

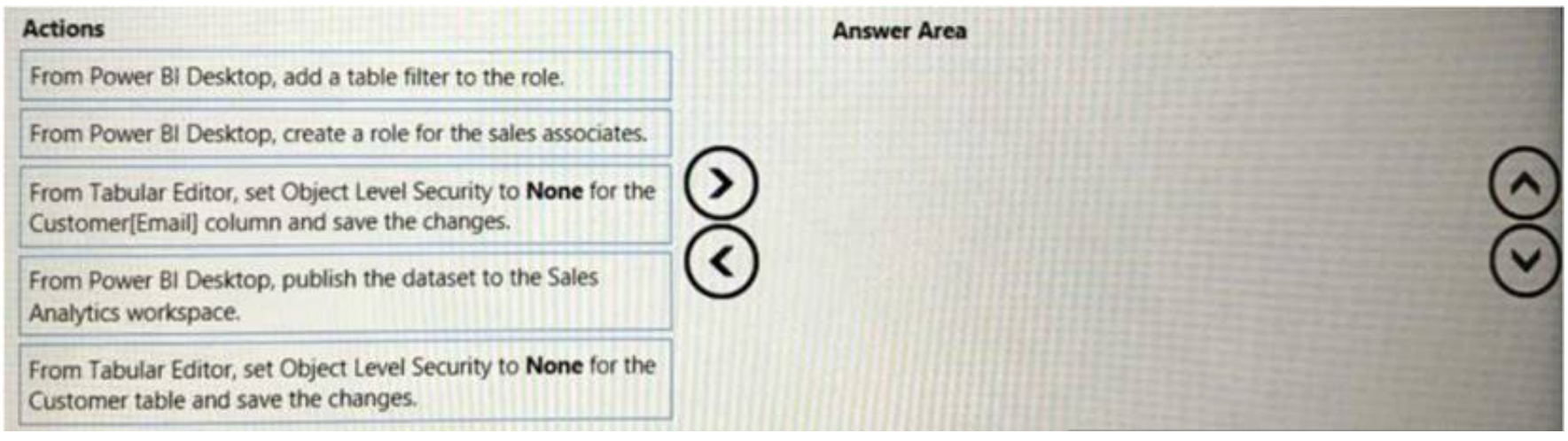
NEW QUESTION 109
- (Exam Topic 1)
How should you configure the Power BI dataset refresh for the dbo.SalesTransactions table?

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.

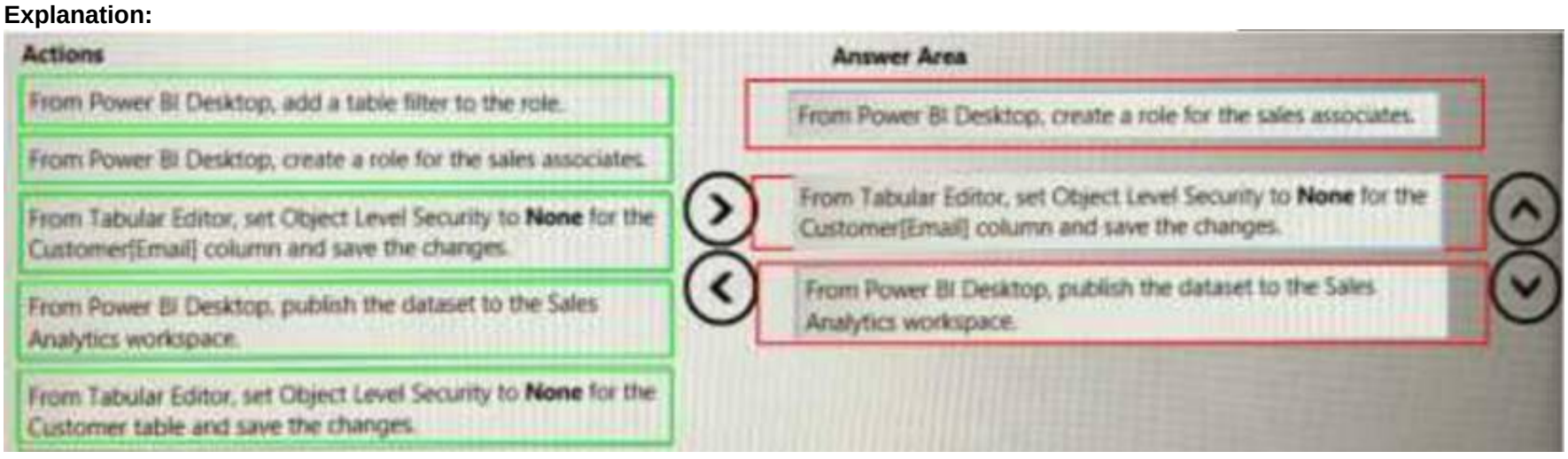
NEW QUESTION 110
- (Exam Topic 1)
You need to implement object-level security (OLS) in the Power Bl dataset for the sales associates.
Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.



A. Mastered
B. Not Mastered

Answer: A

Explanation:



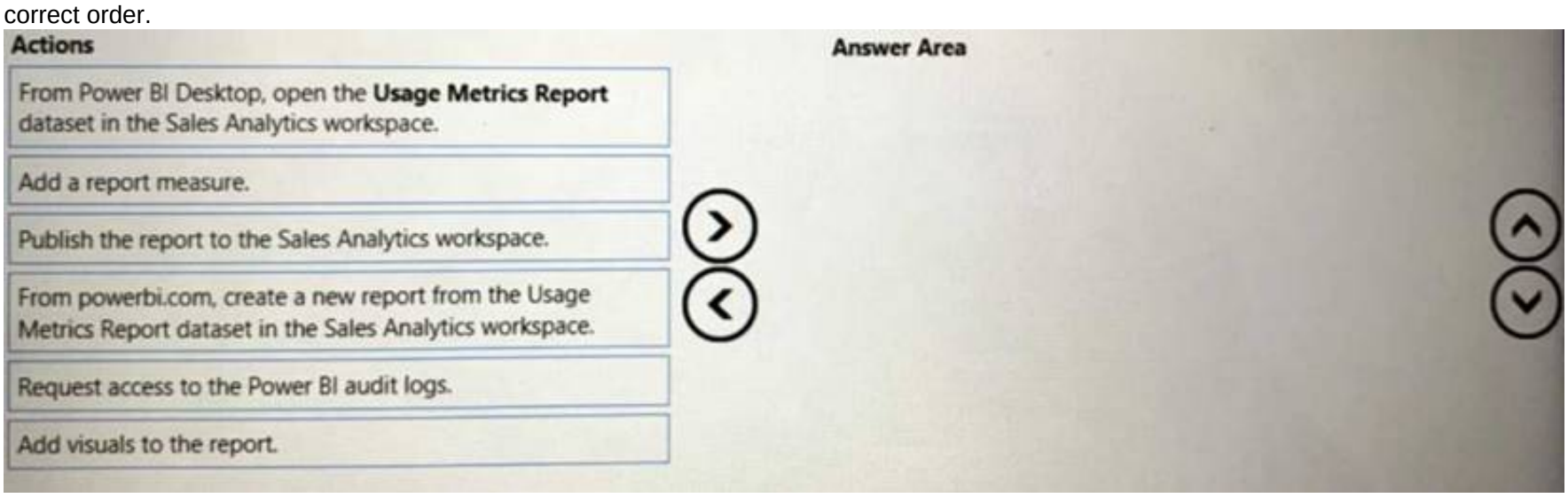
NEW QUESTION 111
- (Exam Topic 1)
You need to configure the Sales Analytics workspace to meet the ad hoc reporting requirements. What should you do?

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

NEW QUESTION 116
- (Exam Topic 1)
You need to create the customized Power BI 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.



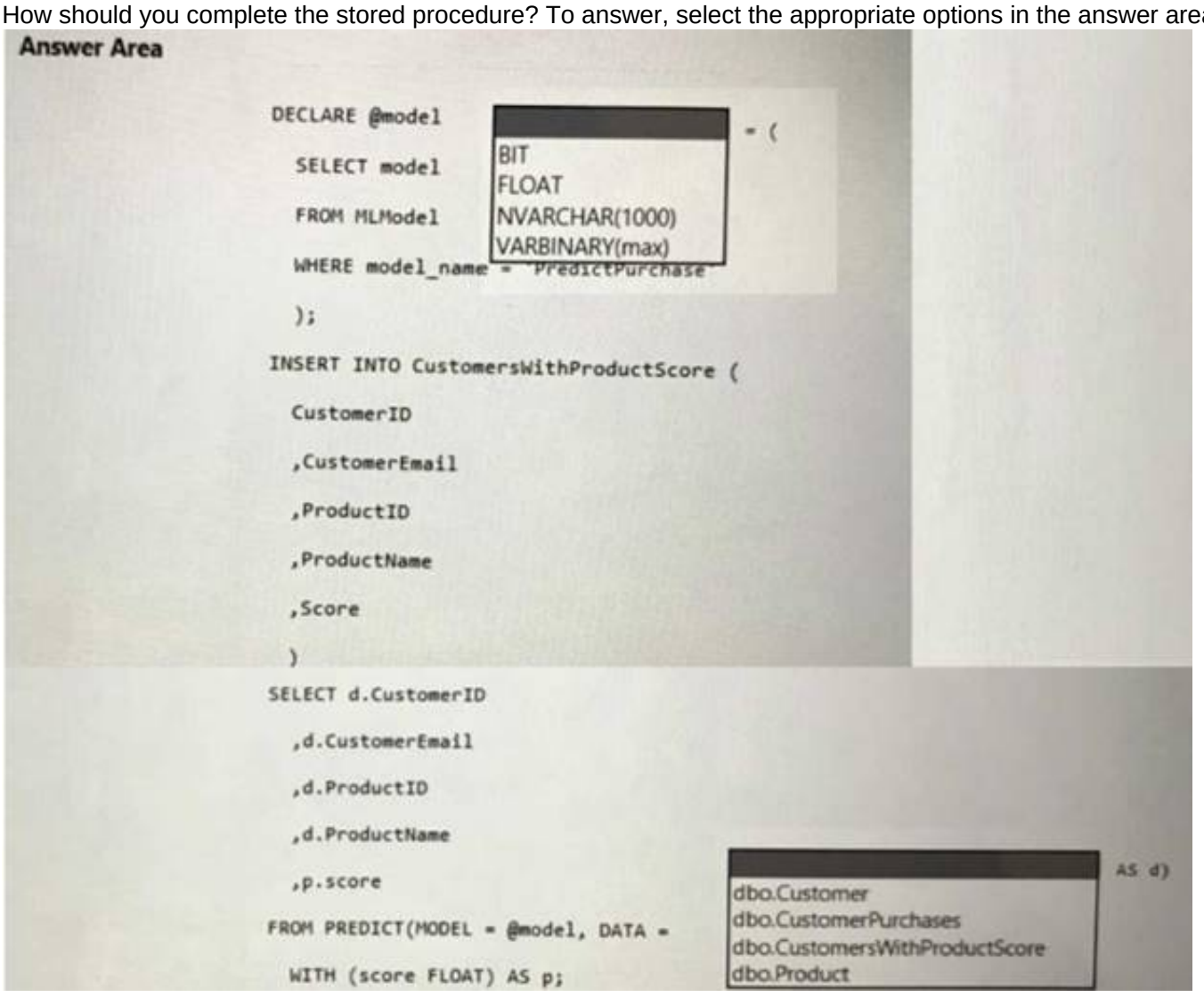
A. Mastered
B. Not Mastered

Answer: A

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
Reference: https://docs.microsoft.com/en-us/power-bi/transform-model/desktop-measures

NEW QUESTION 119

- (Exam 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.



A. Mastered
B. Not Mastered

Answer: A

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.
Reference: https://docs.microsoft.com/en-us/sql/t-sql/queries/predict-transact-sql

NEW QUESTION 123
- (Exam 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-deplo

NEW QUESTION 126

.....

THANKS FOR TRYING THE DEMO OF OUR PRODUCT

Visit Our Site to Purchase the Full Set of Actual DP-500 Exam Questions With Answers.

We Also Provide Practice Exam Software That Simulates Real Exam Environment And Has Many Self-Assessment Features. Order the DP-500 Product From:

https://www.2passeasy.com/dumps/DP-500/

Money Back Guarantee

DP-500 Practice Exam Features:

* DP-500 Questions and Answers Updated Frequently

* DP-500 Practice Questions Verified by Expert Senior Certified Staff

* DP-500 Most Realistic Questions that Guarantee you a Pass on Your FirstTry

* DP-500 Practice Test Questions in Multiple Choice Formats and Updatesfor 1 Year