

# Microsoft

## Exam Questions DP-500

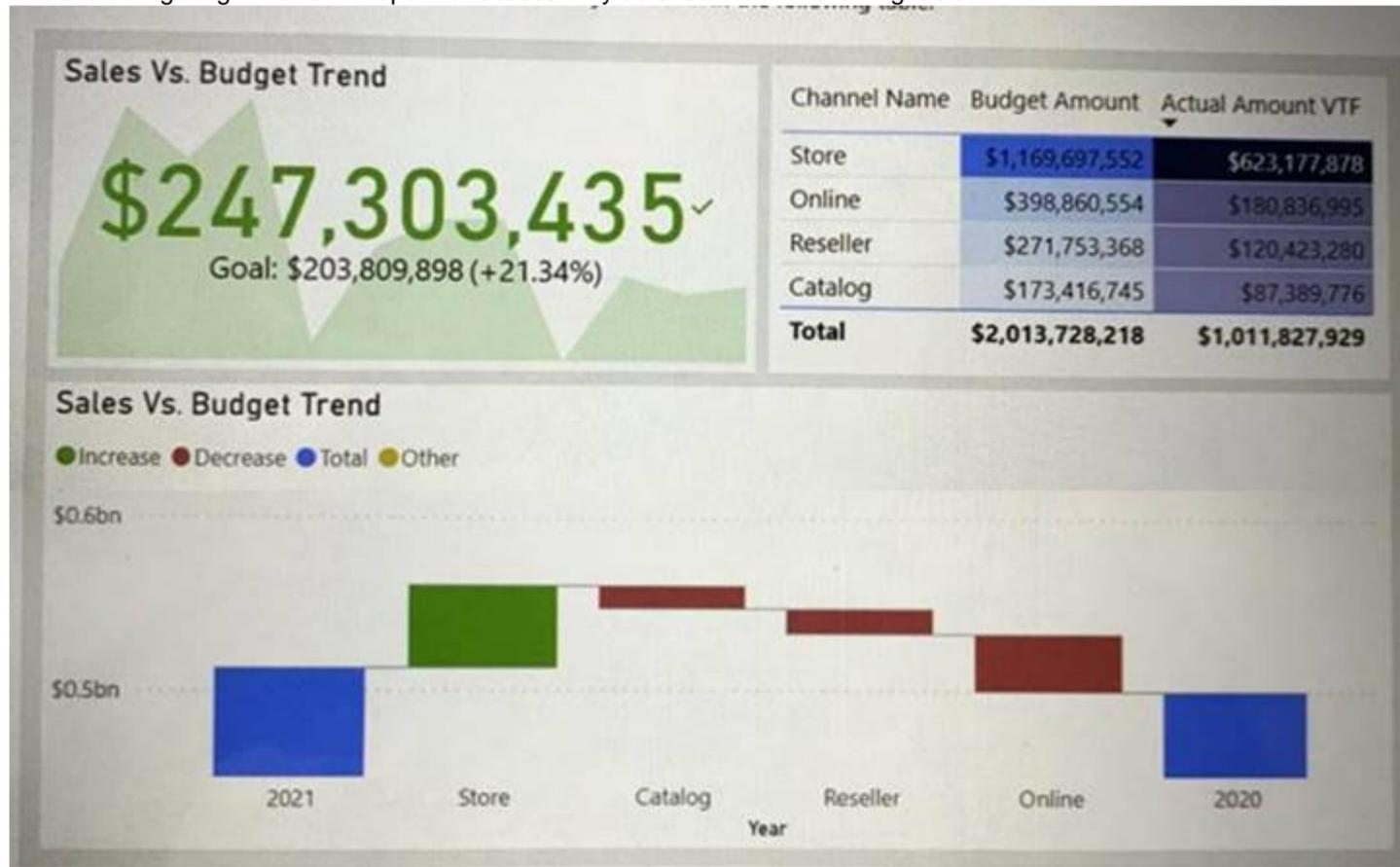
Designing and Implementing Enterprise-Scale Analytics Solutions Using Microsoft Azure and Microsoft Power BI



**NEW QUESTION 1**

- (Exam Topic 3)

You are configuring a Power BI report for accessibility as shown in the following table.



You need to change the default colors of all three visuals to make the report more accessible to users who have color vision deficiency. Which two settings should you configure in the Customize theme window? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point.

- A. Theme colors
- B. Sentiment colors
- C. Divergent colors
- D. First-level elements colors

**Answer:** AB

**Explanation:**

Reference: <https://docs.microsoft.com/en-us/power-bi/create-reports/desktop-report-themes>

**NEW QUESTION 2**

- (Exam Topic 3)

You plan to generate a line chart to visualize and compare the last six months of sales data for two departments. You need to increase the accessibility of the visual. What should you do?

- A. Replace long text with abbreviations and acronyms.
- B. Configure a unique marker for each series.
- C. Configure a distinct color for each series.
- D. Move important information to a tooltip.

**Answer:** C

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

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>

**NEW QUESTION 3**

- (Exam Topic 3)

You open a Power BI 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

Name	Duration (ms)
Recording started (2/3/2022 10:04:04 PM)	-
Refreshed visual	-
Shape	130
Visual display	48
Other	82
Copy query	
Actual/Forecast Billable Hours	1649
DAX query	85
Visual display	47
Other	1517
Copy query	
Actual/Forecast Hours By Type	2083
DAX query	89
Visual display	39
Other	1955
Copy query	
Projected Utilization %	2311
DAX query	119
Visual display	53
Other	2140
Copy query	
Actual/Forecast Billable Hrs YTD	2458
DAX query	151

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 4**

- (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. freqItems
- 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](https://pandas.pydata.org/docs/getting_started/intro_tutorials/06_calculate_statistics.html)

**NEW QUESTION 5**

- (Exam Topic 3)

You have a Power BI 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.

Actions	Answer Area
Define the incremental refresh policy for the table.	
Enable query caching.	
Publish the model to the Power BI service.	<div style="text-align: center;">⬅</div>
Create RangeStart and RangeEnd parameters.	<div style="text-align: center;">➡</div>
Use the Power BI REST API to post a message to /refreshes.	<div style="text-align: center;">⬆</div>
Apply a custom Date/Time filter to the data.	<div style="text-align: center;">⬇</div>

- 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.

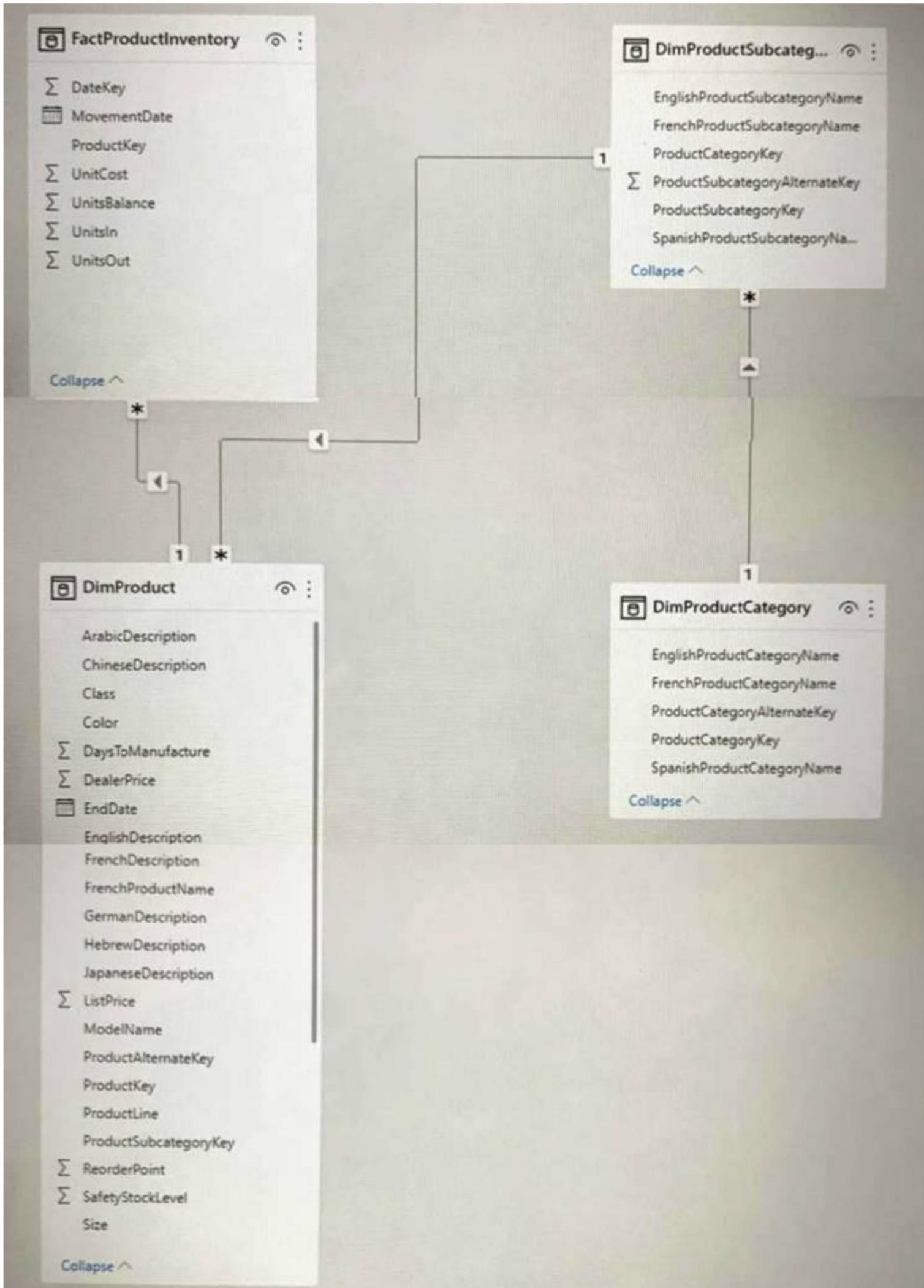
**Step 2: Apply a custom Date/Time filter to the data. Filter data**  
 With RangeStart and RangeEnd parameters defined, apply a filter based on conditions in the RangeStart and RangeEnd parameters.  
 Before continuing with this task, verify your source table has a date column of Date/Time data type. **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)

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 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.

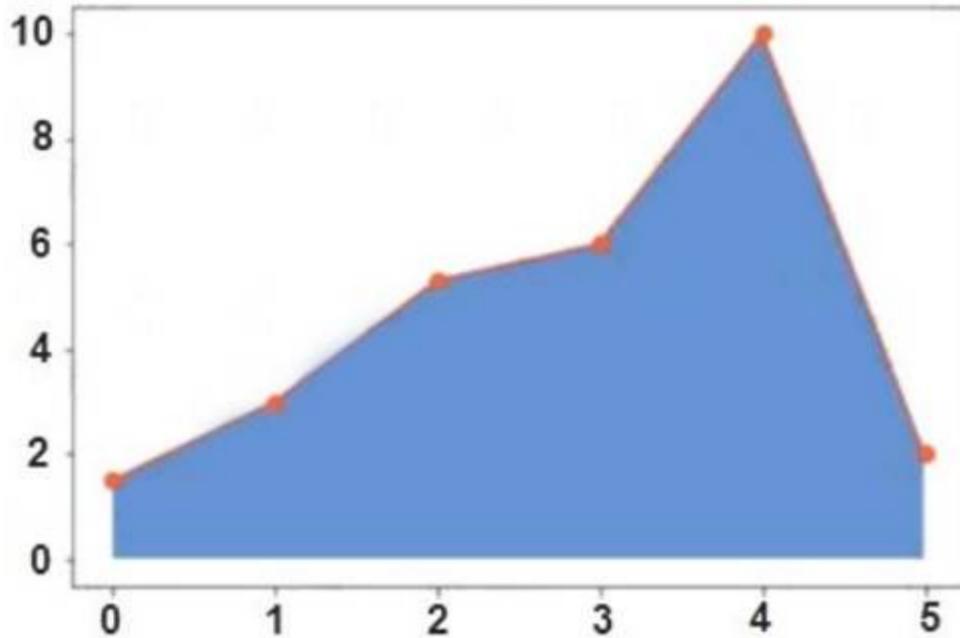
Reference: <https://www.mssqltips.com/sqlservertutorial/3211/denormalize-for-performance/>

**NEW QUESTION 7**

- (Exam Topic 3)

You have an Azure Synapse notebook.  
 You need to create the visual shown in the following exhibit.

chart 1



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

```
import matplotlib.pyplot as plt
x = [0, 1, 2, 3, 4, 5]
y = [1.5, 3, 5.3, 6, 10, 2]
plt.plot(x, y, '-o', color='red')
plt.  (x, y)
plt.  ('chart 1', fontweight='bold')
plt.show()
```

- A. Mastered
- B. Not Mastered

Answer: A

**Explanation:**

Box 1: fill\_between  
 matplotlib.pyplot.fill\_between 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.  
 Box 2: subtitle  
 Set the title of the visual.  
 subtitle adds a centred title to the figure. Reference:  
[https://matplotlib.org/3.1.1/api/\\_as\\_gen/matplotlib.pyplot.fill\\_between.html#matplotlib.pyplot.fill\\_between](https://matplotlib.org/3.1.1/api/_as_gen/matplotlib.pyplot.fill_between.html#matplotlib.pyplot.fill_between)  
[https://matplotlib.org/3.1.1/api/\\_as\\_gen/matplotlib.pyplot.subtitle.html#matplotlib.pyplot.subtitle](https://matplotlib.org/3.1.1/api/_as_gen/matplotlib.pyplot.subtitle.html#matplotlib.pyplot.subtitle)

**NEW QUESTION 8**

- (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 Latin1\_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 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 specify the maximum length for businessName and surveyName. Does this meet the goal?

- A. Yes
- B. No

**Answer:** B

**Explanation:**

Instead use Solution: You recommend using OPENROWSET WITH to explicitly define the collation for businessName and surveyName as Latin1\_General\_100\_BIN2\_UTF8. 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 BI 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 BI 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 10**

- (Exam Topic 3)

You use Azure Synapse Analytics and Apache Spark notebooks to You need to use PySpark to gain access to the visual libraries. Which Python libraries should you use?

- A. Seaborn only
- B. Matplotlib and Seaborn
- C. Matplotlib only
- D. Matplotlib and TensorFlow
- E. TensorFlow only
- F. Seaborn and TensorFlow

**Answer: B**

**Explanation:**

Matplotlib

You can render standard plotting libraries, like Matplotlib, using the built-in rendering functions for each library.

Matplotlib is a plotting library for the Python programming language and its numerical mathematics extension NumPy.

Additional libraries

Beyond these libraries, the Azure Synapse Analytics Runtime also includes the following set of libraries that are often used for data visualization:

Seaborn

Seaborn is a Python data visualization library based on matplotlib. It provides a high-level interface for drawing attractive and informative statistical graphics.

Reference: <https://docs.microsoft.com/en-us/azure/synapse-analytics/spark/apache-spark-data-visualization> <https://seaborn.pydata.org/>

**NEW QUESTION 13**

- (Exam Topic 3)

You have a Power BI dataset that contains the following measures:

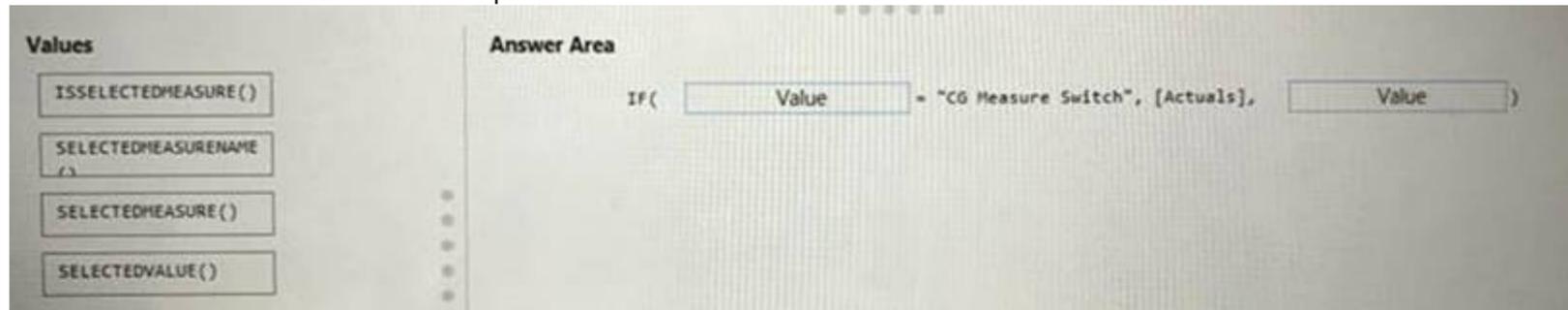
- Budget
- Actuals
- Forecast

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 17**

- (Exam Topic 3)

You have a Power BI 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 20**

- (Exam Topic 3)

You are building a Power BI dataset that will use two data sources.

The dataset has a query that uses a web data source. The web data source uses anonymous authentication. You need to ensure that the query can be used by all the other queries in the dataset.

Which privacy level should you select for the data source?

- A. Public
- B. Organizational
- C. Private
- D. None

**Answer:** A

**Explanation:**

A Public data source gives everyone visibility to the data contained in the data source. Only files, internet data sources, or workbook data can be marked Public. Data from a Public data source may be freely folded to other sources.

Reference: <https://docs.microsoft.com/en-us/power-bi/enterprise/desktop-privacy-levels>

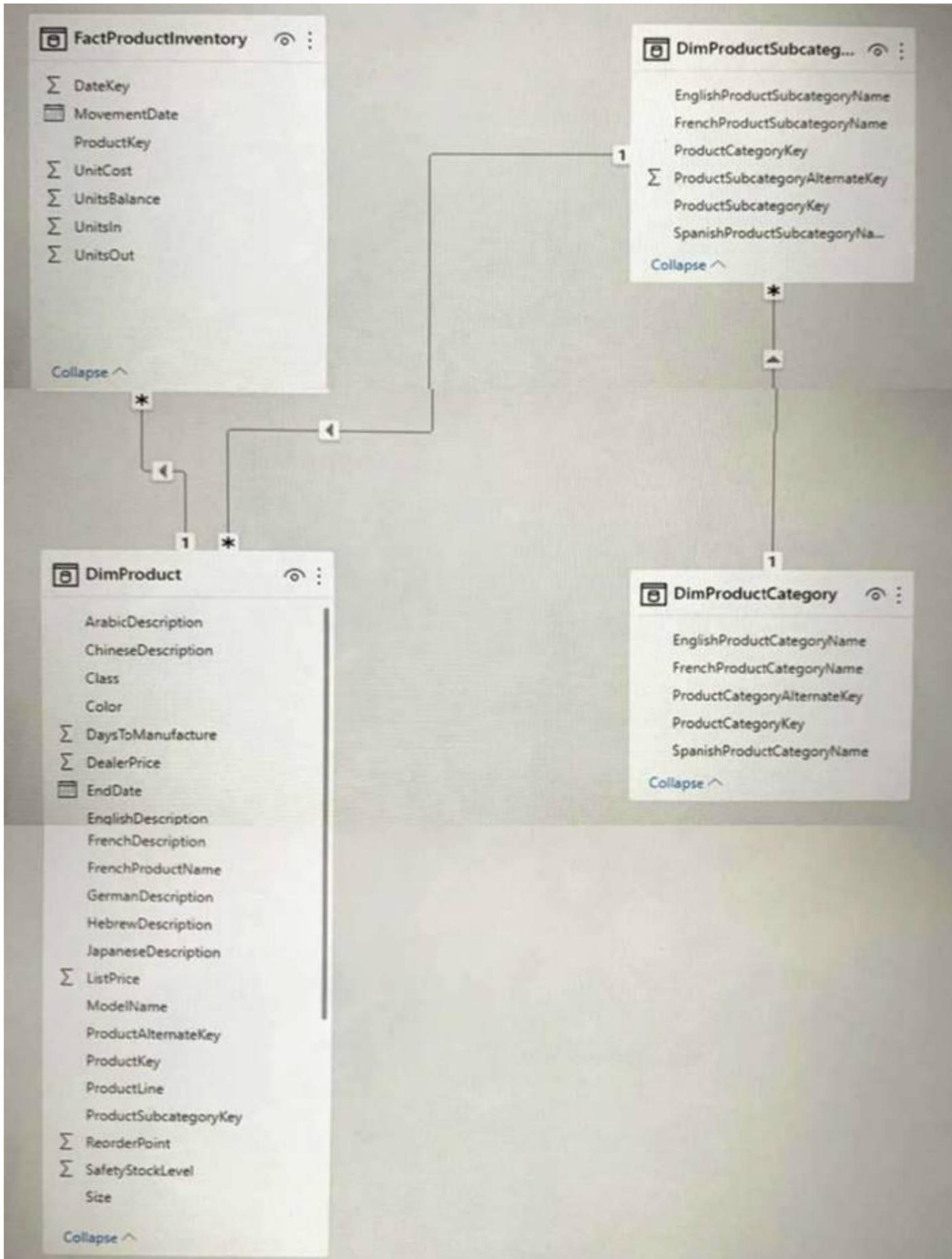
**NEW QUESTION 21**

- (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 moving all the measures to a calculation group. 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 26**

- (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](https://pandas.pydata.org/docs/getting_started/intro_tutorials/06_calculate_statistics.html)

**NEW QUESTION 30**

- (Exam Topic 3)

You are optimizing a dataflow in a Power BI 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 34**

- (Exam Topic 3)

You have a sales report as shown in the following exhibit.



The sales report has the following characteristics: The measures are optimized.

The dataset uses import storage mode.

Data points, hierarchies, and fields cannot be removed or filtered from the report page. From powerbi.com, users experience slow load times when viewing the report.

You need to reduce how long it takes for the report to load without affecting the data displayed in the report. Which two actions should you perform? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

- A. Change the report theme to monochromatic.
- B. Replace the single-value cards with a multi-row card.
- C. Replace the product category charts with a bar chart for sales and a hierarchy of Category and Sub Category on the axis.
- D. Replace all the filters on the Filters pane with visual slicers on the report page.

**Answer:** BC

**NEW QUESTION 37**

- (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 BI 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 DAX Studio, you write a query that uses grouping sets.

Does this meet the goal?

- A. Yes
- B. No

**Answer:** A

**Explanation:**

A grouping is a set of discrete values that are used to group measure fields. Reference: <https://docs.microsoft.com/en-us/power-bi/developer/visuals/capabilities>

**NEW QUESTION 40**

- (Exam Topic 3)

You have the following code in an Azure Synapse notebook.

```
import matplotlib.pyplot as plt
x1 = [1, 3, 4, 5, 6, 7, 9]
y1 = [4, 7, 2, 4, 7, 8, 3]
x2 = [2, 4, 6, 8, 10]
y2 = [5, 6, 2, 6, 2]
plt.bar(x1, y1, label="Blue Item", color='b')
plt.bar(x2, y2, label="Green Item", color='g')
plt.plot()
plt.xlabel("Number")
plt.ylabel("Height")
plt.title("My Chart")
plt.legend()
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.

**Answer Area**

Running the code will create a [answer choice] in the output cell.

- clustered bar chart
- histogram
- line chart
- stacked bar chart

The legend for the resulting chart will list [answer choice] in the legend.

- one item
- two items
- five items
- seven items

- A. Mastered
- B. Not Mastered

**Answer:** A

**Explanation:**

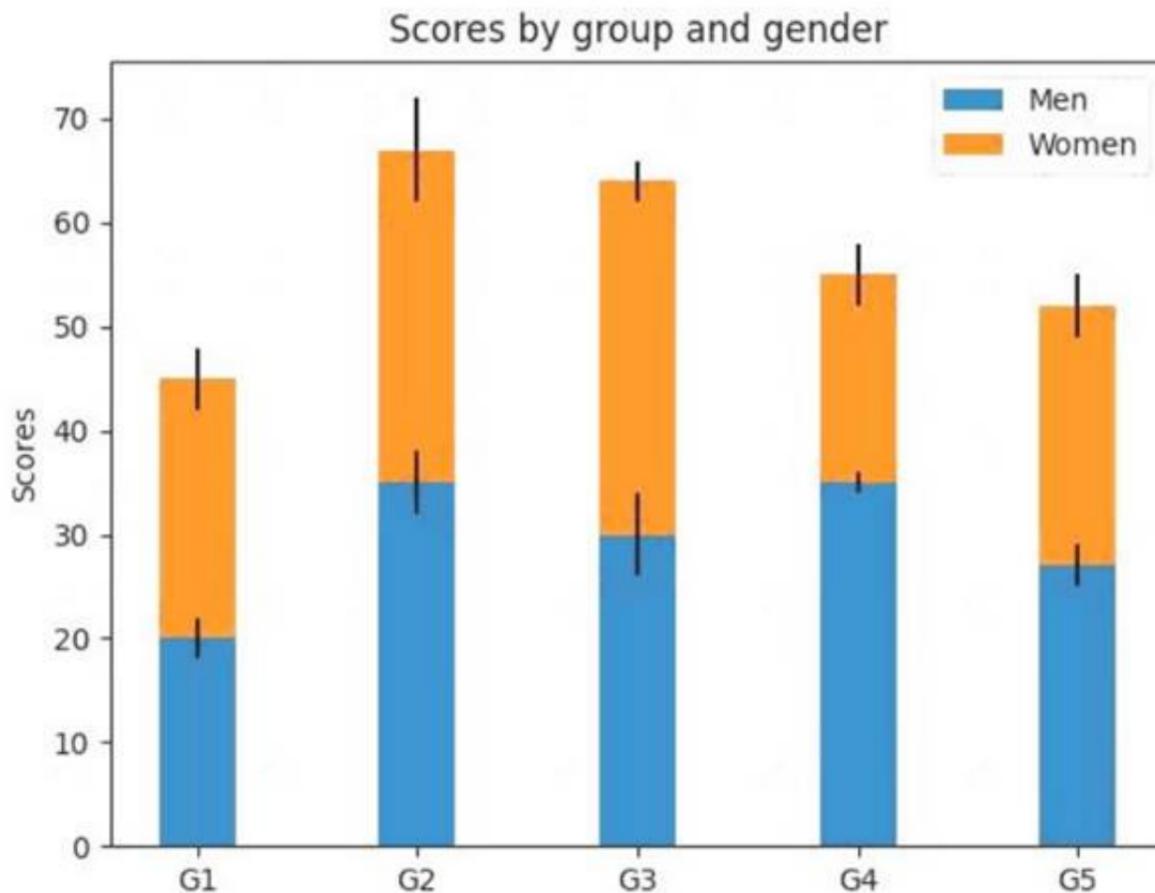
Box 1: stacked bar chart matplotlib.pyplot.bar makes a bar plot.

The bars are positioned at x with the given alignment. Their dimensions are given by height and width. The vertical baseline is bottom (default 0).

Many parameters can take either a single value applying to all bars or a sequence of values, one for each bar.

Stacked bars can be achieved by passing individual bottom values per bar. Stacked bar chart

This is an example of creating a stacked bar plot with error bars using bar. Note the parameters yeer used for error bars, and bottom to stack the women's bars on top of the men's bars.



```
import matplotlib.pyplot as plt
labels = ['G1', 'G2', 'G3', 'G4', 'G5']
men_means = [20, 35, 30, 35, 27]
women_means = [25, 32, 34, 20, 25]
men_std = [2, 3, 4, 1, 2]
women_std = [3, 5, 2, 3, 3]
width = 0.35 # the width of the bars: can also be len(x) sequence
fig, ax = plt.subplots()
ax.bar(labels, men_means, width, yerr=men_std, label='Men')
ax.bar(labels, women_means, width, yerr=women_std, bottom=men_means, label='Women')
ax.set_ylabel('Scores')
ax.set_title('Scores by group and gender')
ax.legend()
plt.show()
```

Box 2: two items

Blue item and Green Item. matplotlib.legend

The legend module defines the Legend class, which is responsible for drawing legends associated with axes and/or figures.

Note: A Diagram Legend is an element that you can add to your diagram to provide information about the colors and/or line thicknesses and styles that have been used in the current diagram, where those colors and other styles have some particular meaning.

Reference: [https://matplotlib.org/stable/api/\\_as\\_gen/matplotlib.pyplot.bar.html](https://matplotlib.org/stable/api/_as_gen/matplotlib.pyplot.bar.html) [https://matplotlib.org/stable/gallery/lines\\_bars\\_and\\_markers/bar\\_stacked.html](https://matplotlib.org/stable/gallery/lines_bars_and_markers/bar_stacked.html) [https://matplotlib.org/stable/api/legend\\_api.html](https://matplotlib.org/stable/api/legend_api.html)

**NEW QUESTION 45**

- (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.

**Answer Area**

```
import matplotlib.axes  
matplotlib.projections  
matplotlib.pyplot  
matplotlib.widgets as chart
```

```
dataset.plot(kind='scatter', x='Age', y='Weight', color='red')
```

```
chart.clf()  
chart.plot()  
chart.show()  
chart.triplot()
```

- 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 48**

- (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. sample
- B. describe
- C. freqItems
- D. explain

**Answer:** B

**Explanation:**

pandas.DataFrame.describe

Descriptive statistics include those that summarize the central tendency, dispersion and shape of a dataset's distribution, excluding NaN values.

Analyzes both numeric and object series, as well as DataFrame column sets of mixed data types. The output will vary depending on what is provided.

Reference: <https://pandas.pydata.org/pandas-docs/stable/reference/api/pandas.DataFrame.describe.html>

**NEW QUESTION 51**

- (Exam Topic 3)

You have a Power BI dataset named Dataset1 that uses DirectQuery against an Azure SQL database named DB1. DB1 is a transactional database in the third normal form.

You need to recommend a solution to minimize how long it takes to execute the query. The solution must maintain the current functionality. What should you include in the recommendation?

- A. Create calculated columns in Dataset1.
- B. Remove the relationships from Dataset1.
- C. Normalize the tables in DB1.
- D. Denormalize the tables in DB1.

**Answer:** D

**Explanation:**

Denormalize to improve query performance.

Note: Normalization prevents data duplications, preserves disk space, and improves the performance of the disk I/O operations. The downside of the normalization is that the queries based on these normalized tables require more table joins.

Schema denormalization (i.e. consolidation of some dimension tables) for such databases can significantly reduce costs of the analytical queries and improve the performance.

Reference:

<https://www.mssqltips.com/sqlservertip/7114/denormalization-dimensions-synapse-mapping-data-flow/>

**NEW QUESTION 53**

- (Exam 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:** AD

**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>

**NEW QUESTION 55**

- (Exam 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 BI Desktop

**Answer:** BD

**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.

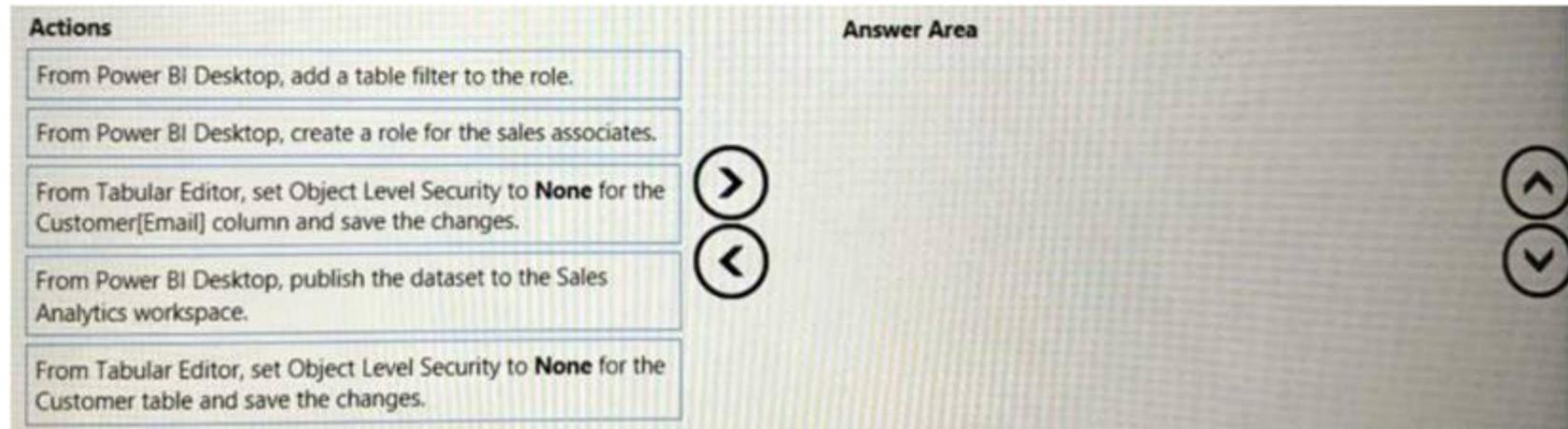
D: Use Power BI Desktop Performance Analyzer to optimize the report or model. Reference: <https://www.sqlbi.com/tv/analyzing-a-slow-report-query-in-dax-studio/>  
<https://docs.microsoft.com/en-us/power-bi/guidance/report-performance-troubleshoot>

**NEW QUESTION 60**

- (Exam Topic 1)

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

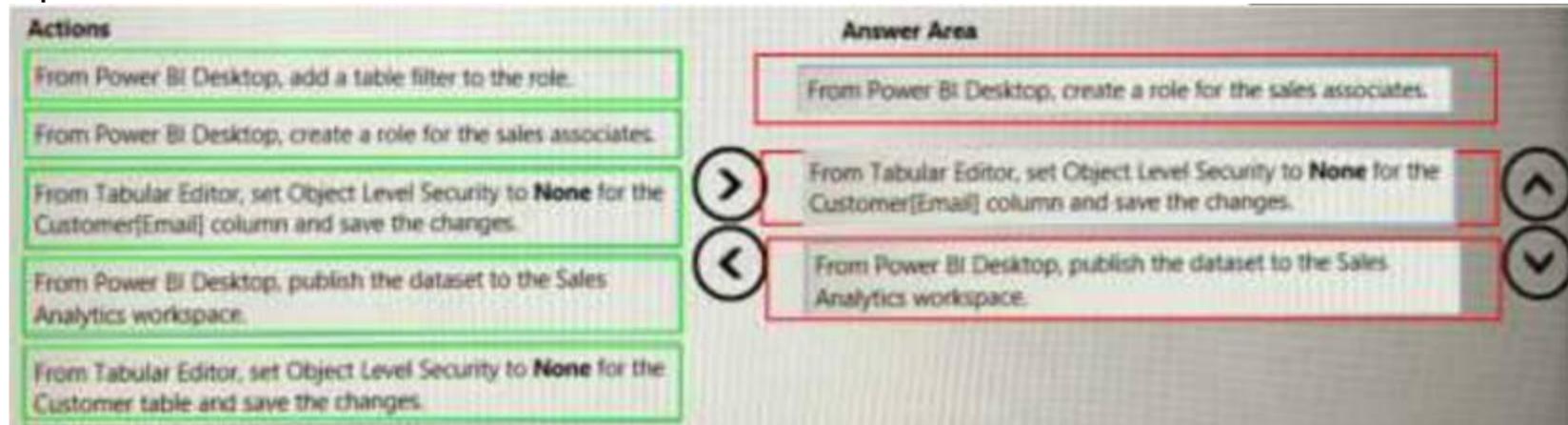
Which three actions should you perform in sequence? To answer, move the appropriate actions from the list of actions to the answer area and arrange them in the correct order.



- A. Mastered
- B. Not Mastered

**Answer:** A

**Explanation:**



**NEW QUESTION 65**

- (Exam 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 BI Admin portal, set Allow users to apply sensitivity labels for Power BI content to Enabled.
- B. From the Power BI Admin portal, set Apply sensitivity labels from data sources to their data in Power BI to Enabled.
- C. In SQLD
- D. apply sensitivity labels to the columns in the Customer and CustomersWithProductScore tables.
- E. In the Power BI datasets, apply sensitivity labels to the columns in the Customer and CustomersWithProductScore tables.
- F. From the Power BI Admin portal, set Make certified content discoverable to Enabled.

**Answer:** ADE

**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/>

**NEW QUESTION 70**

.....

## **Thank You for Trying Our Product**

### **We offer two products:**

1st - We have Practice Tests Software with Actual Exam Questions

2nd - Questions and Answers in PDF Format

### **DP-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

**100% Actual & Verified — Instant Download, Please Click**  
[Order The DP-500 Practice Test Here](#)