

Microsoft

Exam Questions DP-500

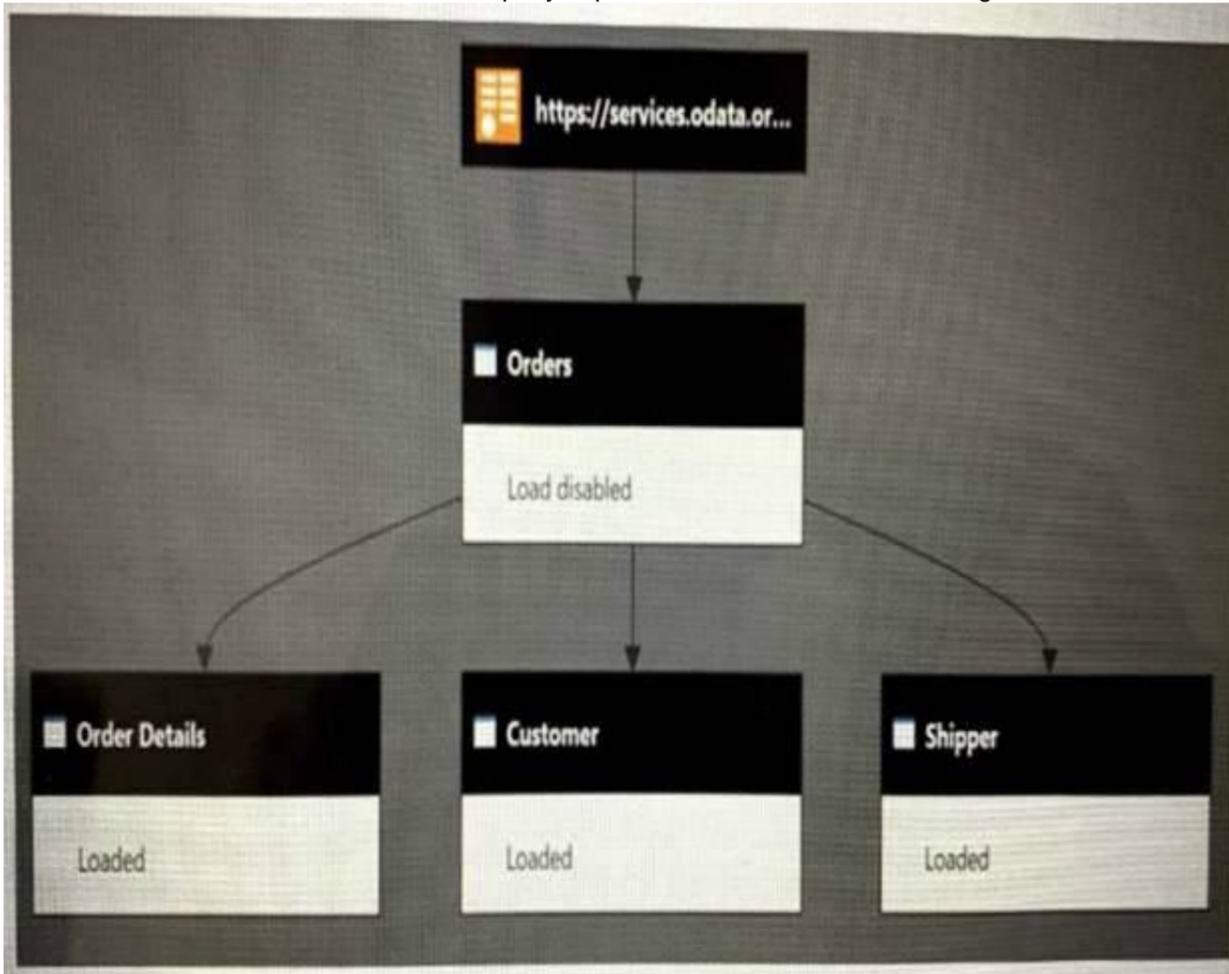
Designing and Implementing Enterprise-Scale Analytics Solutions Using Microsoft Azure and Microsoft Power BI



NEW QUESTION 1

- (Exam Topic 3)

You have a Power BI dataset that has the query dependencies shown in the following exhibit.



Use the drop-down menus to select the answer choice that completes each statement based on the information presented in the graphic.

NOTE: Each correct selection is worth one point.

Answer Area

When the dataset refreshes, the orders query will be executed [answer choice] times.

[answer choice] will reduce data refresh times for this model.

0
1
3

Duplicating the Orders query instead of referencing the query
 Replacing the Orders query with a dataflow
 Using Table.Buffer in the Orders query

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: 3
 Power Query doesn't start at the first query and work down, it starts at the bottom (last) query and works backwards, so 3 tables from 1 will cause it to process that first source table 3 times.

Box 2: Using Table.Buffer in the Orders query

Table.Buffer buffers a table in memory, isolating it from external changes during evaluation. Buffering is shallow. It forces the evaluation of any scalar cell values, but leaves non-scalar values (records, lists, tables, and so on) as-is.

Note that using this function might or might not make your queries run faster. In some cases, it can make your queries run more slowly due to the added cost of reading all the data and storing it in memory, as well as the fact that buffering prevents downstream folding.

Example 1

Load all the rows of a SQL table into memory, so that any downstream operations will no longer be able to query the SQL server.

Usage let

Source = Sql.Database("SomeSQLServer", "MyDb"), MyTable = Source[Item="MyTable"][Data], BufferMyTable = Table.Buffer(dbo_MyTable)

in BufferMyTable Output

Reference: <https://radacad.com/performance-tip-for-power-bi-enable-load-sucks-memory-up> <https://docs.microsoft.com/en-us/powerquery-m/table-buffer>

NEW QUESTION 2

- (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 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 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.

Create RangeStart and RangeEnd parameters.

Use the Power BI REST API to post a message to /refreshes.

Apply a custom Date/Time filter to the data.



- 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 5

- (Exam Topic 3)

You have a Power BI report that contains the table shown in the following exhibit.

Store ID	Store	Returns	
6	Leo	\$6,108	●
5	Fama	\$6,097	●
13	Contoso	\$5,214	●
11	Pomum	\$4,968	●
7	VanArsdel	\$4,964	●
10	Pirum	\$4,644	●
2	Aliqui	\$4,479	●
1	Abbas	\$4,070	●
8	Natura	\$3,376	●
14	Victoria	\$2,317	●
4	Salvus	\$2,296	●
12	Quibus	\$2,208	●
3	Barba	\$1,601	●
Total		\$52,342	

The table contains conditional formatting that shows which stores are above, near, or below the monthly quota for returns. You need to ensure that the table is accessible to consumers of reports who have color vision deficiency. What should you do?

- A. Add alt text to explain the information that each color conveys.
- B. Move the conditional formatting icons to a tooltip report.
- C. Change the icons to use a different shape for each color.
- D. Remove the icons and use red, yellow, and green background colors instead.

Answer: A

Explanation:

Report accessibility checklist, All Visuals.

* Ensure alt text is added to all non-decorative visuals on the page.

* Avoid using color as the only means of conveying information. Use text or icons to supplement or replace the color.

* Check that your report page works for users with color vision deficiency.

* Etc.

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

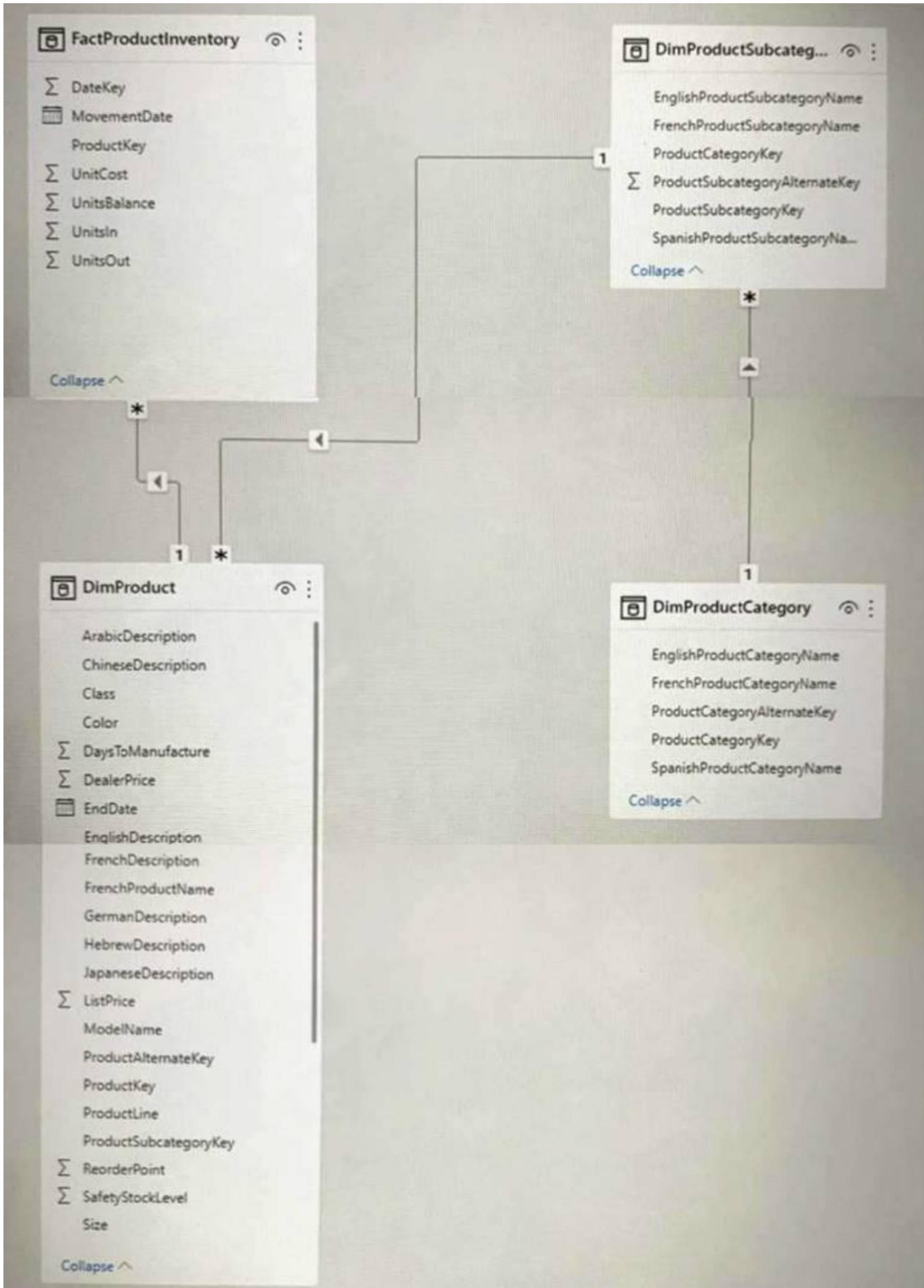
NEW QUESTION 6

- (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 Analytics serverless SQL pool.
 You need to catalog the serverless SQL pool by using Azure Purview.
 Which three actions should you perform? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point.

- A. Create a managed identity in Azure Active Directory (Azure AD).
- B. Assign the Storage Blob Data Reader role to the Azure Purview managed service identity (MSI) for the storage account associated to the Synapse Analytics workspace.
- C. Assign the Owner role to the Azure Purview managed service identity (MSI) for the Azure Purview resource group.
- D. Register a data source.
- E. Assign the Reader role to the Azure Purview managed service identity (MSI) for the Synapse Analytics workspace.

Answer: ABE

Explanation:

Authentication for enumerating serverless SQL database resources
 There are three places you'll need to set authentication to allow Microsoft Purview to enumerate your serverless SQL database resources:
 The Azure Synapse workspace
 The associated storage
 The Azure Synapse serverless databases
 The steps below will set permissions for all three. Azure Synapse workspace
 In the Azure portal, go to the Azure Synapse workspace resource. On the left pane, select Access Control (IAM).
 Select the Add button.
 Set the Reader role and enter your Microsoft Purview account name, which represents its managed service identity (MSI).
 Select Save to finish assigning the role
 Azure Synapse Analytics serverless SQL pool catalog Purview Azure Purview managed service identity Storage account
 In the Azure portal, go to the Resource group or Subscription that the storage account associated with the Azure Synapse workspace is in.
 On the left pane, select Access Control (IAM). Select the Add button.
 Set the Storage blob data reader role and enter your Microsoft Purview account name (which represents its MSI) in the Select box.
 Select Save to finish assigning the role. Azure Synapse serverless database
 Go to your Azure Synapse workspace and open the Synapse Studio. Select the Data tab on the left menu.
 Select the ellipsis (...) next to one of your databases, and then start a new SQL script.
 Add the Microsoft Purview account MSI (represented by the account name) on the serverless SQL databases. You do so by running the following command in your SQL script:
 SQL
 CREATE LOGIN [PurviewAccountName] FROM EXTERNAL PROVIDER;
 Apply permissions to scan the contents of the workspace
 You can set up authentication for an Azure Synapse source in either of two ways. Select your scenario below for steps to apply permissions.
 Use a managed identity Use a service principal
 Reference: <https://docs.microsoft.com/en-us/azure/purview/register-scan-synapse-workspace?tabs=MI>

NEW QUESTION 8

- (Exam Topic 3)

You are attempting to configure certification for a Power BI dataset and discover that the certification setting for the dataset is unavailable.
 What are two possible causes of the issue? Each correct answer presents a complete solution. NOTE: Each correct selection is worth one point.

- A. The workspace is in shared capacity.
- B. You have insufficient permissions.
- C. Dataset certification is disabled for the Power BI tenant.
- D. The sensitivity level for the dataset is set to Highly Confidential.
- E. Row-level security (RLS) is missing from the dataset.

Answer: BC

Explanation:

Reference: <https://docs.microsoft.com/en-us/power-bi/admin/service-admin-setup-certification> <https://docs.microsoft.com/en-us/power-bi/collaborate-share/service-endorse-content>

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 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 (
    [Box 1]
    DATA_SOURCE = 'datalake',
    FORMAT = 'CSV', PARSER_VERSION = '2.0',
    [Box 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:

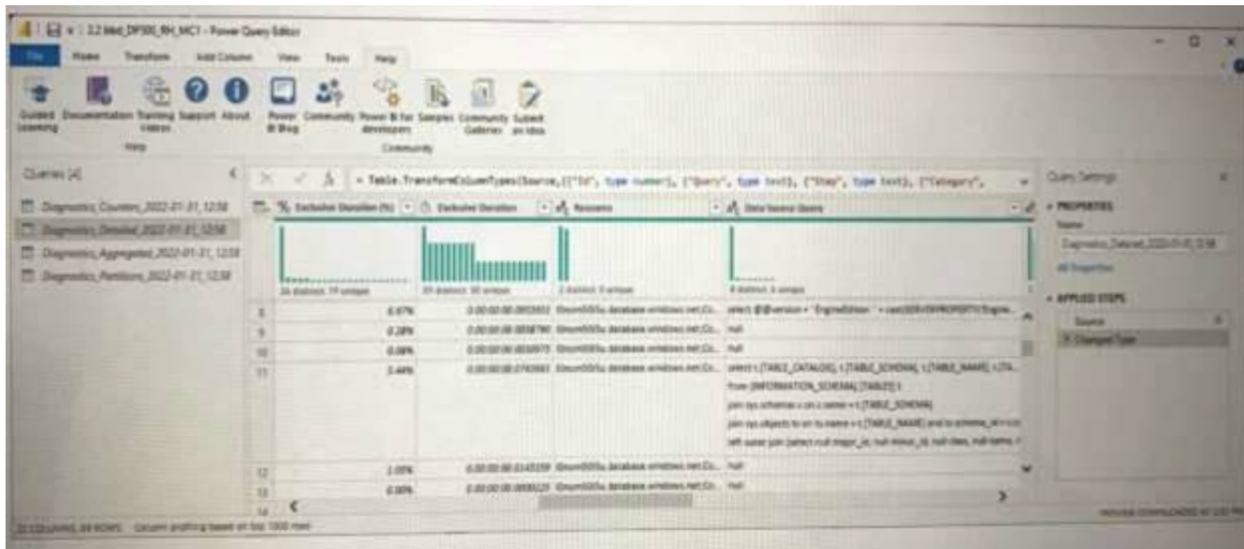
Not FIRSTROW=1. FIRSTROW=1 is used when there is no header.

Reference: <https://docs.microsoft.com/en-us/azure/synapse-analytics/sql/develop-openrowset>

NEW QUESTION 10

- (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>

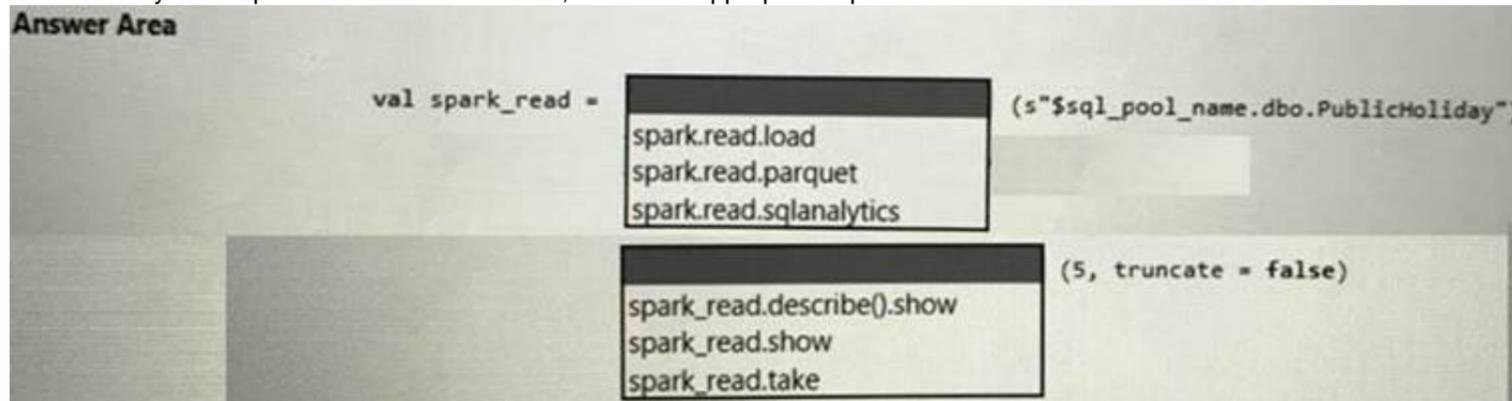
NEW QUESTION 12

- (Exam Topic 3)

You have an Azure Synapse workspace named Workspace1.

You need to use PySpark in a notebook to read data from a SQL pool as an Apache Spark DataFrame and display the top five

How should you complete the 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:

Graphical user interface, text, application Description automatically generated

Box 1: sqlanalytics

Read from a SQL Pool table with Spark

// Read the table we just created in the sql pool as a Spark dataframe val spark_read = spark.read. sqlanalytics(s"\$sql_pool_name.dbo.PublicHoliday")

spark_read.show(5, truncate = false)

Box 2: spark_read.show Sample output:

Graphical user interface, text, application Description automatically generated

spark_read: org.apache.spark.sql.DataFrame = [countryOrRegion: string, holidayName: string ... 3 more fields]

countryOrRegion	holidayName	normalizeHolidayName	isPaidTimeOff	countryRegionCode
Czech	Den české státnosti	Den české státnosti	null	CZ
Norway	Søndag	Søndag	null	NO
Sweden	Söndag	Söndag	null	SE
India	Gandhi Jayanti	Gandhi Jayanti	true	IN
Germany	Tag der Deutschen Einheit	Tag der Deutschen Einheit	null	DE

only showing top 5 rows

Reference:

<https://github.com/Azure-Samples/Synapse/blob/main/Notebooks/Scala/03%20Read%20and%20write%20from>

NEW QUESTION 16

- (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 defining a data source and view for the Parquet files. You recommend updating the query to use the view.

Does this meet the goal?

- A. Yes
- B. No

Answer: B

Explanation:

Solution: You recommend using OPENROWSET WITH to explicitly specify the maximum length for businessName and surveyName.

The size of the varchar(8000) columns are too big. Better reduce their size.

A SELECT...FROM OPENROWSET(BULK...) statement queries the data in a file directly, without importing the data into a table. SELECT...FROM OPENROWSET(BULK...) statements can also list bulk-column aliases by using a format file to specify column names, and also data types.

Reference: <https://docs.microsoft.com/en-us/sql/t-sql/functions/openrowset-transact-sql>

NEW QUESTION 18

- (Exam Topic 3)

You have a 2-GB Power BI 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 22

- (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 26

- (Exam Topic 3)

You have a Power BI 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 30

- (Exam Topic 3)

You have a Power BI dataset that has only the necessary fields visible for report development. You need to ensure that end users see only 25 specific fields that they can use to personalize visuals. What should you do?

- A. From Tabular Editor, create a new role.
- B. Hide all the fields in the dataset.
- C. Configure object-level security (OLS).
- D. From Tabular Editor, create a new perspective.

Answer: B

NEW QUESTION 35

- (Exam Topic 3)

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

Product	Sales
Amarilla	17,747,116.06
Carretera	13,815,307.89
Montana	15,390,801.88
Paseo	33,011,143.95
Velo	18,250,059.47
VTT	20,511,921.02
Total	118,726,350.26

You need to make the visual more accessible to users who have color vision deficiency. What should you do?

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

Answer: A

Explanation:

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

green and brown blue and purple green and blue
 light green and yellow blue and grey
 green and grey

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

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

NEW QUESTION 36

- (Exam Topic 3)

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

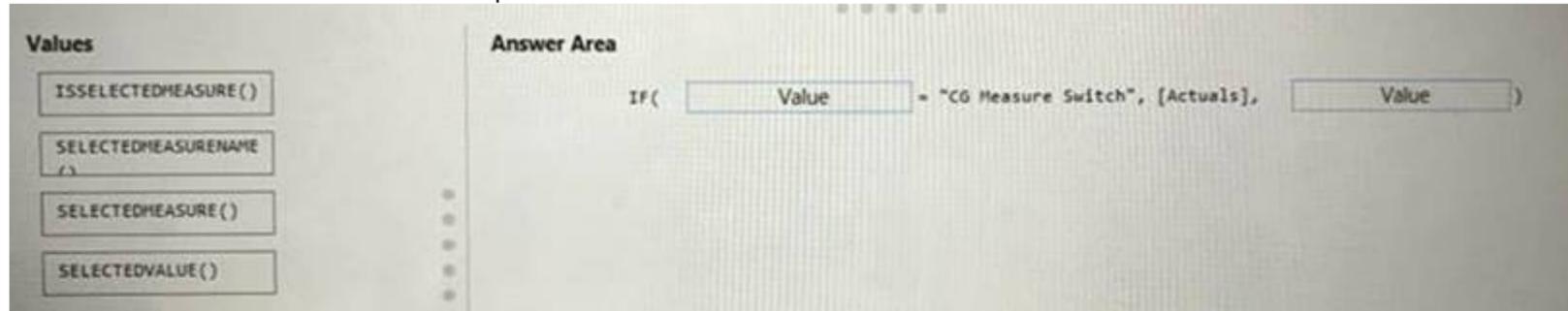
- Budget
- Actuals
- Forecast

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 38

- (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 42

- (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>

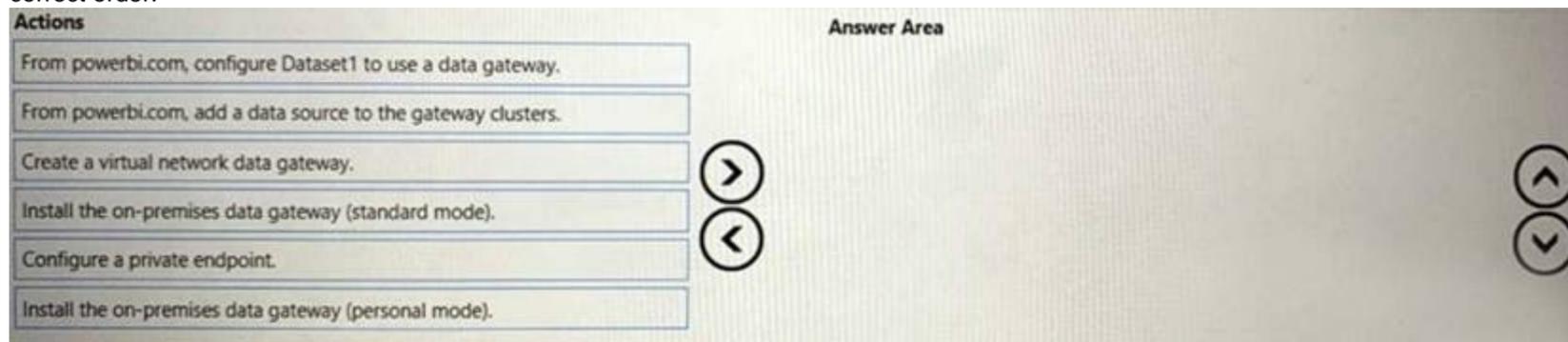
NEW QUESTION 45

- (Exam Topic 3)

You have a shared dataset in Power BI named Dataset1.

You have an on-premises Microsoft SQL Server database named DB1. You need to ensure that Dataset1 refreshes data from DB1.

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

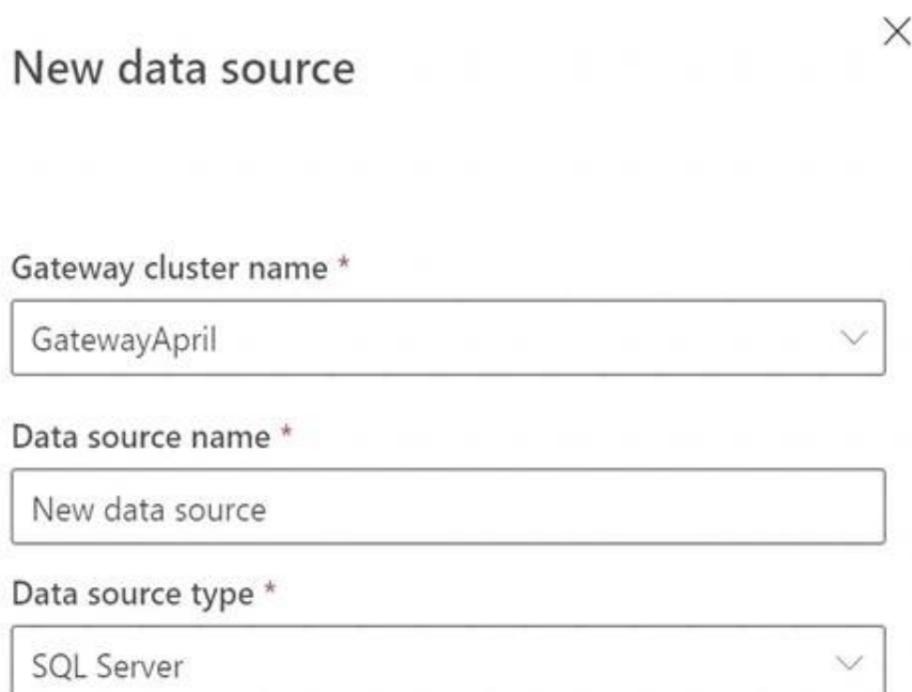
Explanation:

Step 1: Install the on-premises data gateway (standard mode)

The personal mode is only for a single user, not to be used for a shared dataset. Step 2: From powerbi.com, add a data source to the gateway clusters

After you install the on-premises data gateway, you can add data sources that can be used with the gateway. Add a data source

Under Data Source Type, select SQL Server.



After you fill in everything, select Create. You can now use this data source for scheduled refresh or DirectQuery against a SQL Server that's on-premises. You see Created New data source if it succeeded.

Step 3: From powerbi.com, configure Dataset1 to use a data gateway. Connect a dataset to a SQL Server database

In Power BI Desktop, you connected directly to your on-premises SQL Server database, but the Power BI service requires a data gateway to act as a bridge between the cloud and your on-premises network. Follow these steps to add your on-premises SQL Server database as a data source to a gateway and then

connect your dataset to this data source.

- > Sign in to Power BI. In the upper-right corner, select the settings gear icon and then select Settings.
- > On the Datasets tab, select the dataset AdventureWorksProducts, so you can connect to your on-premises SQL Server database through a data gateway.
- > Expand Gateway connection and verify that at least one gateway is listed.
- > Under Actions, expand the toggle button to view the data sources and select the Add to gateway link.
- > On the Gateways management page, on the Data Source Settings tab, enter and verify the following information, and select Add.
- > On the Datasets tab, expand the Gateway connection section again. Select the data gateway you configured, which shows a Status of running on the machine where you installed it, and select Apply.

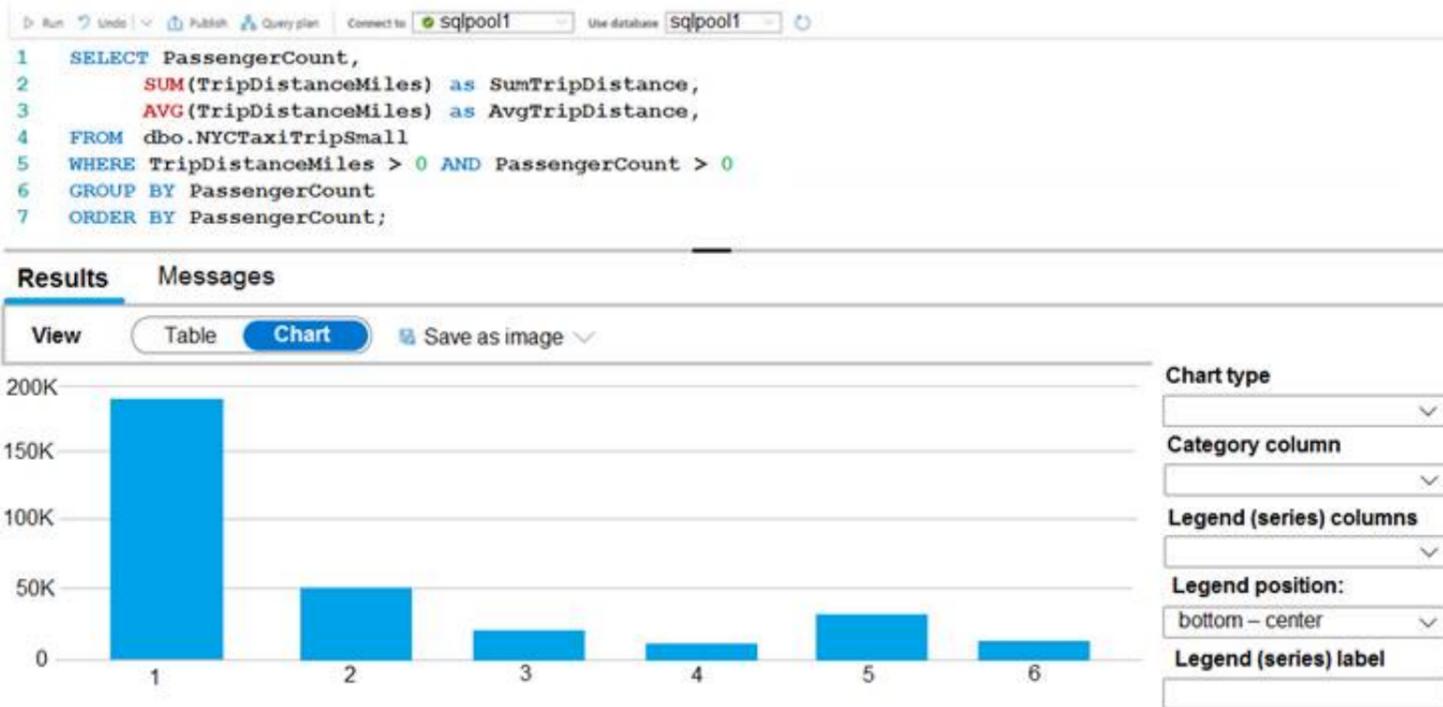
Reference: <https://docs.microsoft.com/en-us/power-bi/connect-data/service-gateway-personal-mode> <https://docs.microsoft.com/en-us/power-bi/connect-data/service-gateway-sql-tutorial> <https://docs.microsoft.com/en-us/power-bi/connect-data/service-gateway-enterprise-manage-sql>

NEW QUESTION 50

- (Exam Topic 3)

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

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



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

Answer Area

Category column:

[Dropdown menu with options: AvgTripDistance, PassengerCount, SumTripDistance, TripDistanceMiles]

Legend (series) column:

[Dropdown menu with options: AvgTripDistance, PassengerCount, SumTripDistance, TripDistanceMiles]

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Answer Area

Category column:

	▼
AvgTripDistance	
PassengerCount	
SumTripDistance	
TripDistanceMiles	

Legend (series) column:

	▼
AvgTripDistance	
PassengerCount	
SumTripDistance	
TripDistanceMiles	

NEW QUESTION 55

- (Exam Topic 3)

You plan to create a Power BI report that will use an OData feed as the data source. You will retrieve all the entities from two different collections by using the same service root

The OData feed is still in development. The location of the feed will change once development is complete. The report will be published before the OData feed development is complete.

You need to minimize development effort to change the data source once the location changes.

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

Actions

- Create a parameter that contains the service root URI.
- From Advanced Editor, duplicate the query and change the resource path in the URL.
- Get data from an OData feed source and use the parameter to populate the first part of the URL.
- From Advanced Editor, get data from an OData feed source and use the parameter to populate the last part of the URL.
- From Advanced Editor, reference the query and change the resource path in the URL.

Answer Area

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Graphical user interface, text, application Description automatically generated

Step 1: Create a parameter that contains the service root URI

Step 2: Get data from OData feed source and use the parameter to populate the first part of the URL. The URI is in the first part of the query.

Example: let

Source = OData.Feed

("https://analytics.dev.azure.com/{organization}/{project}/_odata/v3.0-preview/WorkItemSnapshot? "

&"\$apply=filter("

&"WorkItemType eq 'Bug' "

&"and StateCategory ne 'Completed' "

&"and startswith(Area/AreaPath,'{areapath}') "

&"and DateValue ge {startdate} "

&") "

&"/groupby("

&"(DateValue,State,WorkItemType,Priority,Severity,Area/AreaPath,Iteration/IterationPath,AreaSK), "

&"aggregate(\$count as Count) "

&") "

,null, [Implementation="2.0",OmitValues = ODataOmitValues.Nulls,ODataVersion = 4]) in

Source

Box 3: From Advanced Editor, duplicate the query and change the resource path in the URL. Choose Get Data, and then Blank Query.

From the Power BI Query editor, choose Advanced Editor. The Advanced Editor window opens.

Edit the query. Etc.

Reference: <https://docs.microsoft.com/en-us/azure/devops/report/powerbi/odataquery-connect>

NEW QUESTION 58

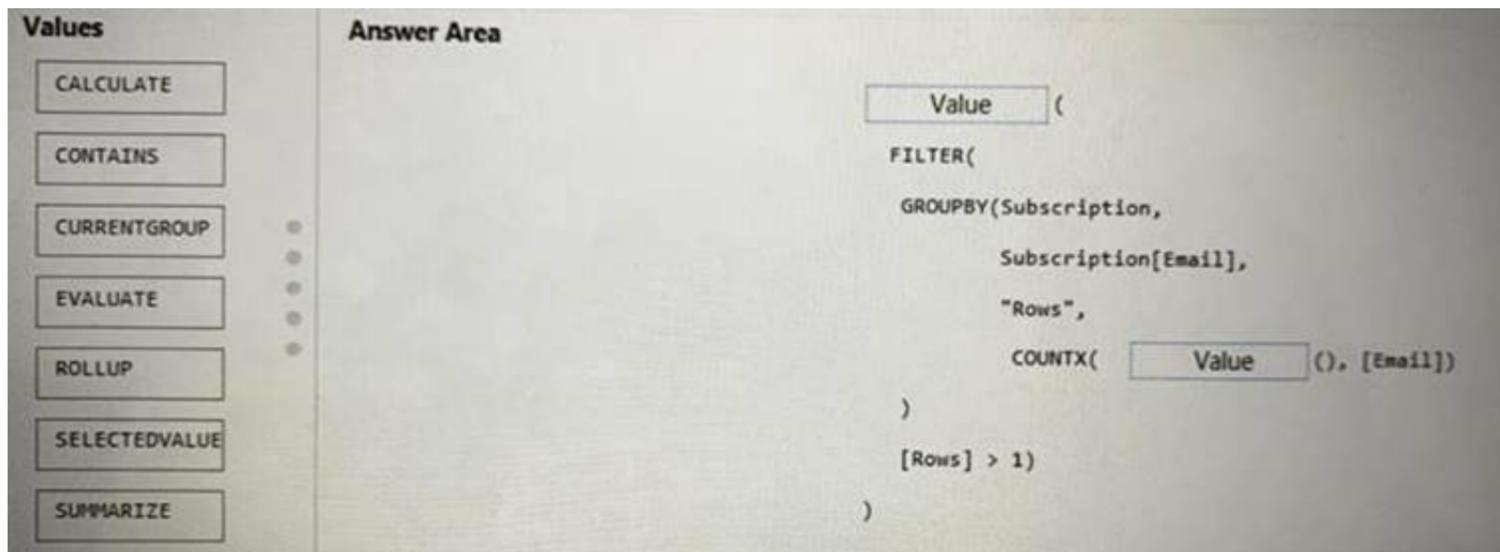
- (Exam Topic 3)

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 59

- (Exam Topic 3)

You are creating a Python visual in Power BI Desktop.

You need to retrieve the value of a column named Unit Price from a DataFrame. How should you reference the Unit Price column in the Python code?

- A. pandas.DataFrame('Unit Price')
- B. dataset['Unit Price']
- C. data = [Unit Price]
- D. ('Unit Price')

Answer: A

Explanation:

You can retrieve a column in a pandas DataFrame object by using the DataFrame object name, followed by the label of the column name in brackets.

So if the DataFrame object name is dataframe1 and the column we are trying to retrieve the 'X' column, then we retrieve the column using the statement, dataframe1['X'].

Here's a simple Python script that imports pandas and uses a data frame: import pandas as pd

data = [['Alex',10],['Bob',12],['Clarke',13]]

df = pd.DataFrame(data,columns=['Name','Age'],dtype=float) print (df)

When run, this script returns: Name Age

0 Alex 10.0

1 Bob 12.0

2 Clarke 13.0 Reference:

<http://www.learningaboutelectronics.com/Articles/How-to-retrieve-a-column-from-a-pandas-dataframe-object-i>

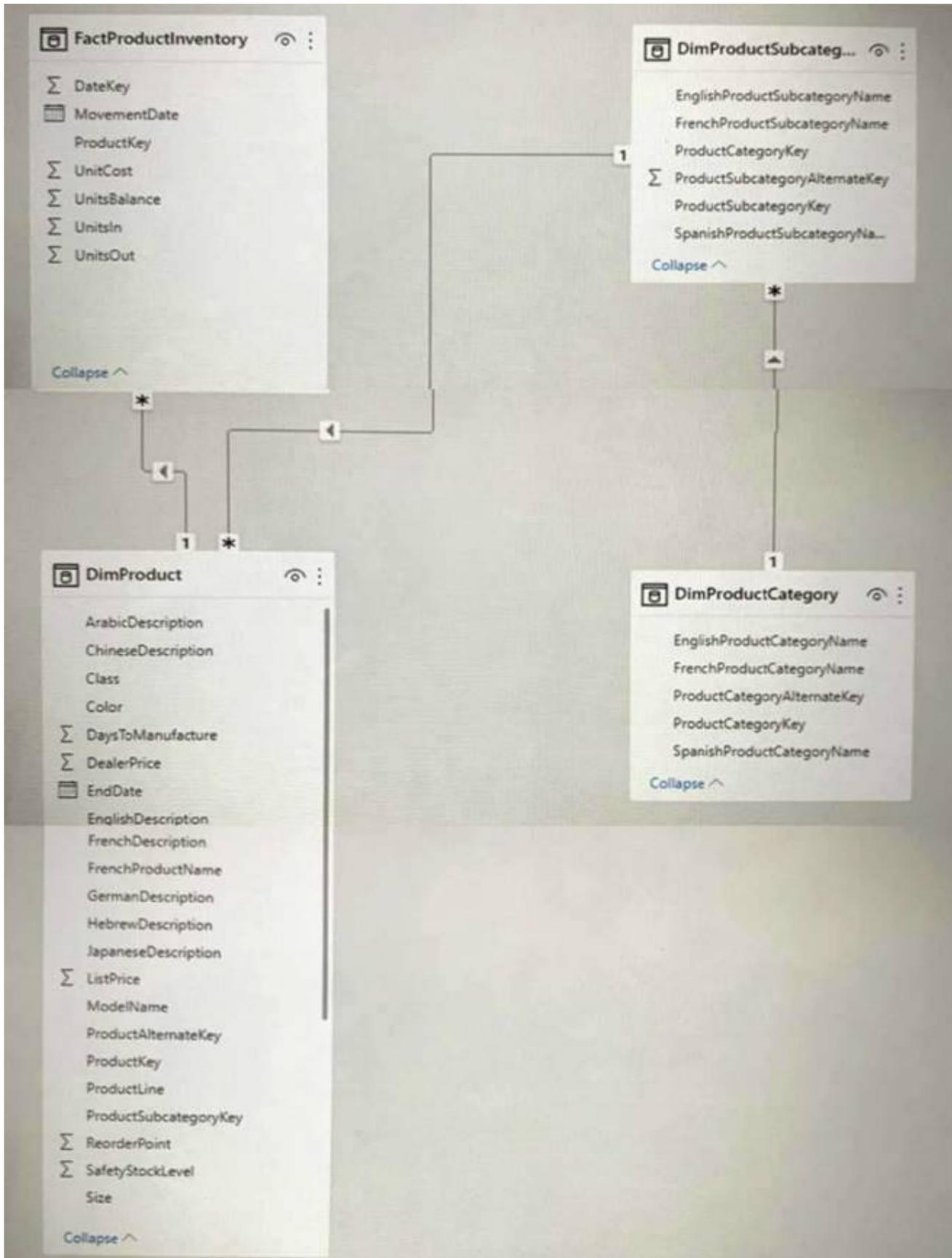
NEW QUESTION 63

- (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 creating a perspective that contains the commonly used fields. 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 67

- (Exam Topic 3)

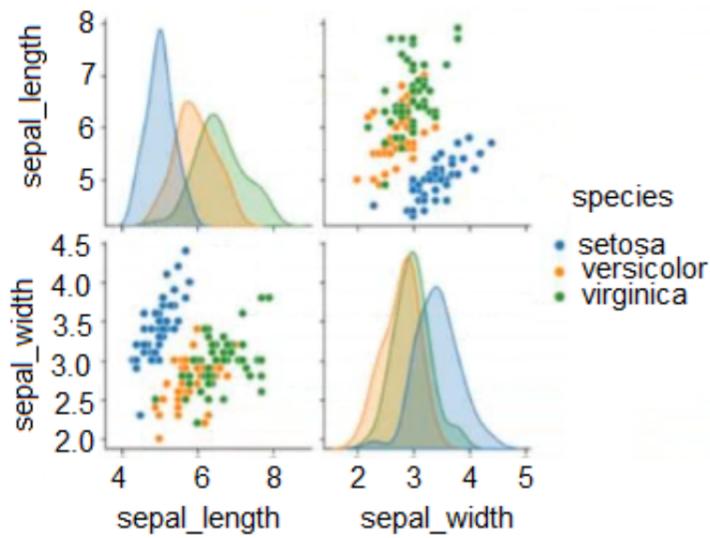
You are using an Azure Synapse notebook to create a Python visual. You run the following code cell to import a dataset named Iris.

```
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.  (iris, hue= '', height=2.5)

plt.show()
```

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

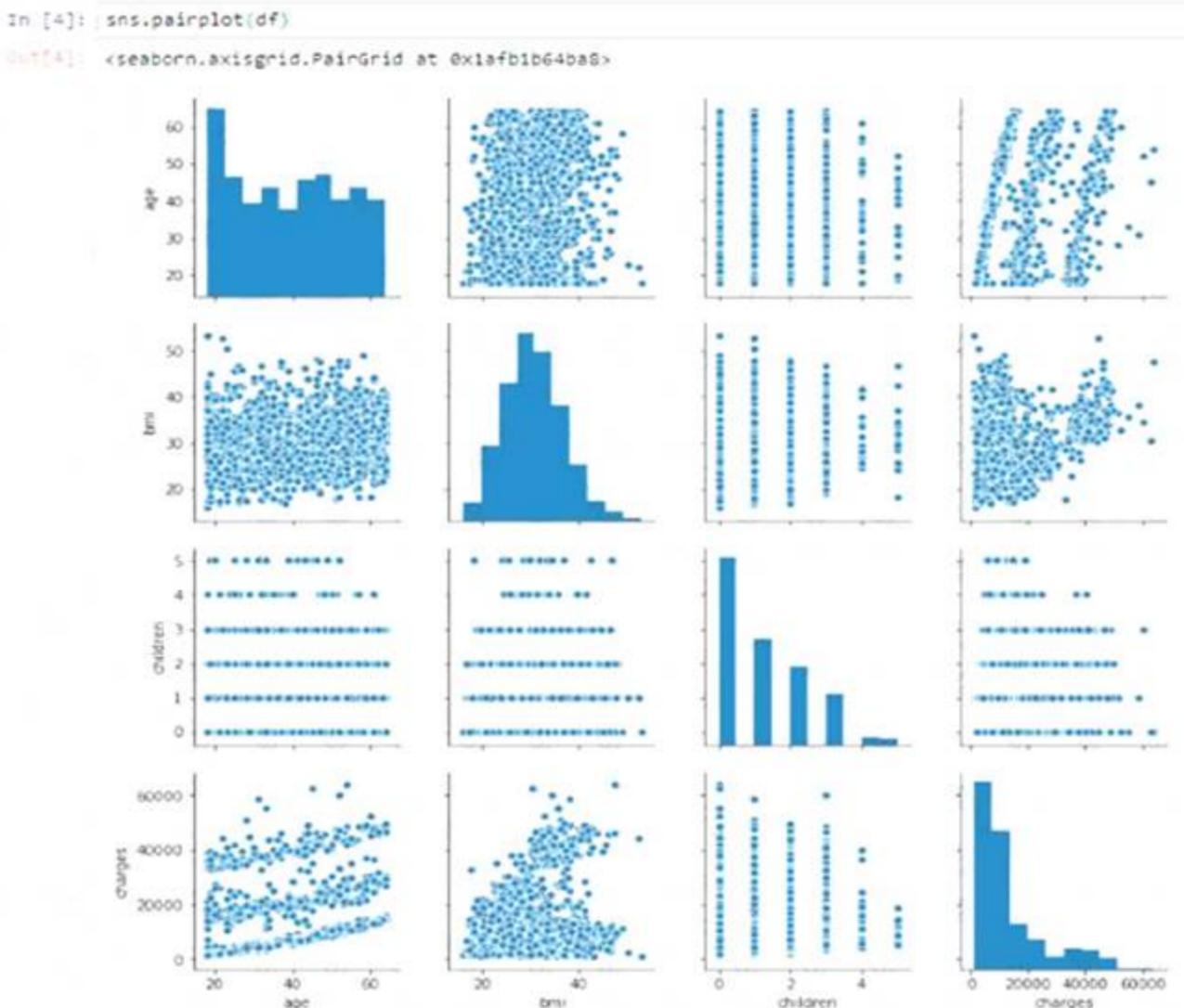
Box 1: pairplot

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 68

- (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.	<input type="radio"/>	<input type="radio"/>
The query uses a custom function to load file data.	<input type="radio"/>	<input type="radio"/>
Changing the source to use AzureStorage.DataLake will reduce the load time of the query.	<input type="radio"/>	<input type="radio"/>

- 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 71

- (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 75

- (Exam Topic 3)

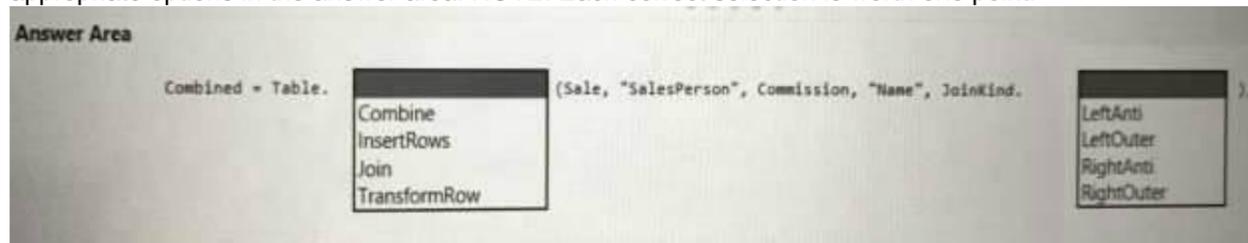
You use Advanced Editor in Power Query Editor to edit a query that references two tables named Sales and Commission. A sample of the data in the Sales table is shown in the following table.

OrderID	SalesPerson	Amount
101	Tom	199.99
103	Eileen	279.99
108	Enrique	333.42

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

Person	Commission
Tom	0.04
Eileen	0.05

You need to merge the tables by using Power Query Editor without losing any rows in the Sales table. How should you complete the query? 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: Join

Box 2: LeftOuter Left outer join

One of the join kinds available in the Merge dialog box in Power Query is a left outer join, which keeps all the rows from the left table and brings in any matching rows from the right table.

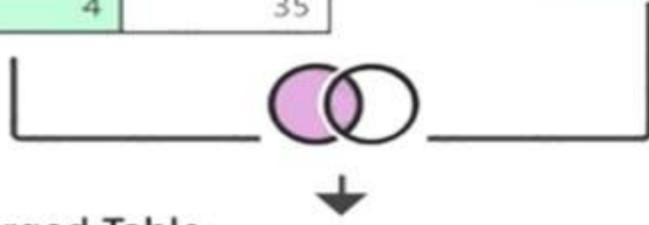
Diagram, table Description automatically generated

Left Table

Date	CountryID	Units
1/1/2020	1	40
1/2/2020	1	25
1/3/2020	3	30
1/4/2020	4	35

Right Table

ID	Country
1	USA
2	Canada
3	Panama



Merged Table

Date	CountryID	Units	Country
1/1/2020	1	40	USA
1/2/2020	1	25	USA
1/3/2020	3	30	Panama
1/4/2020	4	35	<i>null</i>

Reference: <https://docs.microsoft.com/en-us/power-query/merge-queries-left-outer>

NEW QUESTION 78

- (Exam Topic 2)

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

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

Actions

- Create an Apache Spark data source.
- Merge columns.
- Create a web data source.
- Expand the attributes.
- Publish the model.

Answer Area

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Actions

- Create an Apache Spark data source.
- Merge columns.
- Create a web data source.
- Expand the attributes.
- Publish the model.

Answer Area

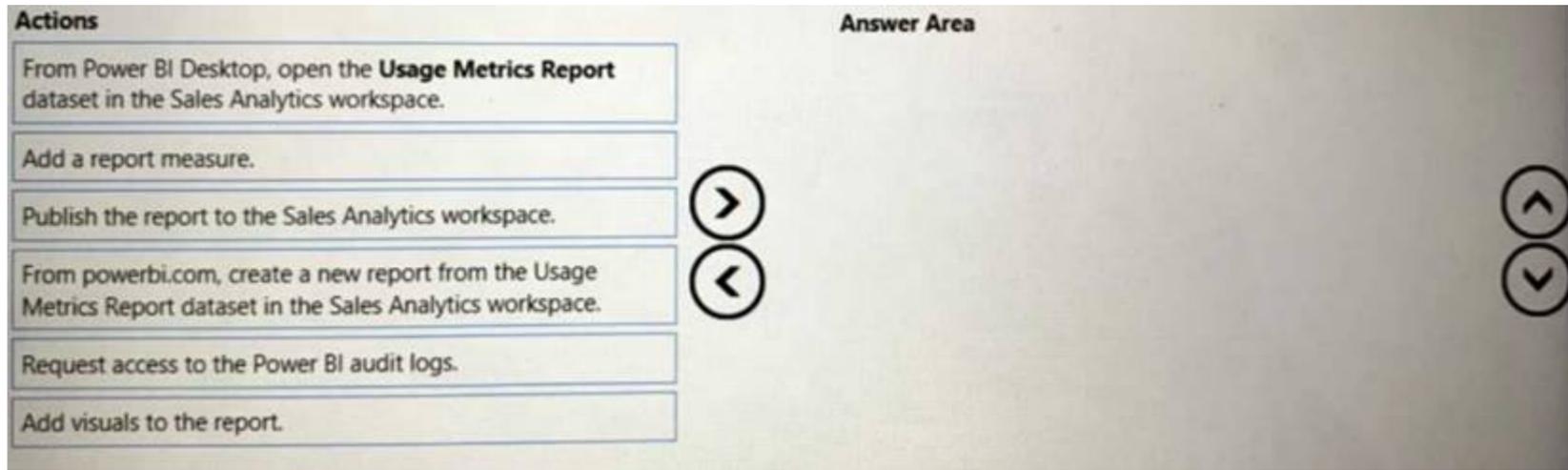
- Create a web data source.
- Expand the attributes.
- Publish the model.

NEW QUESTION 80

- (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 81

- (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 83

.....

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](#)