



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 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

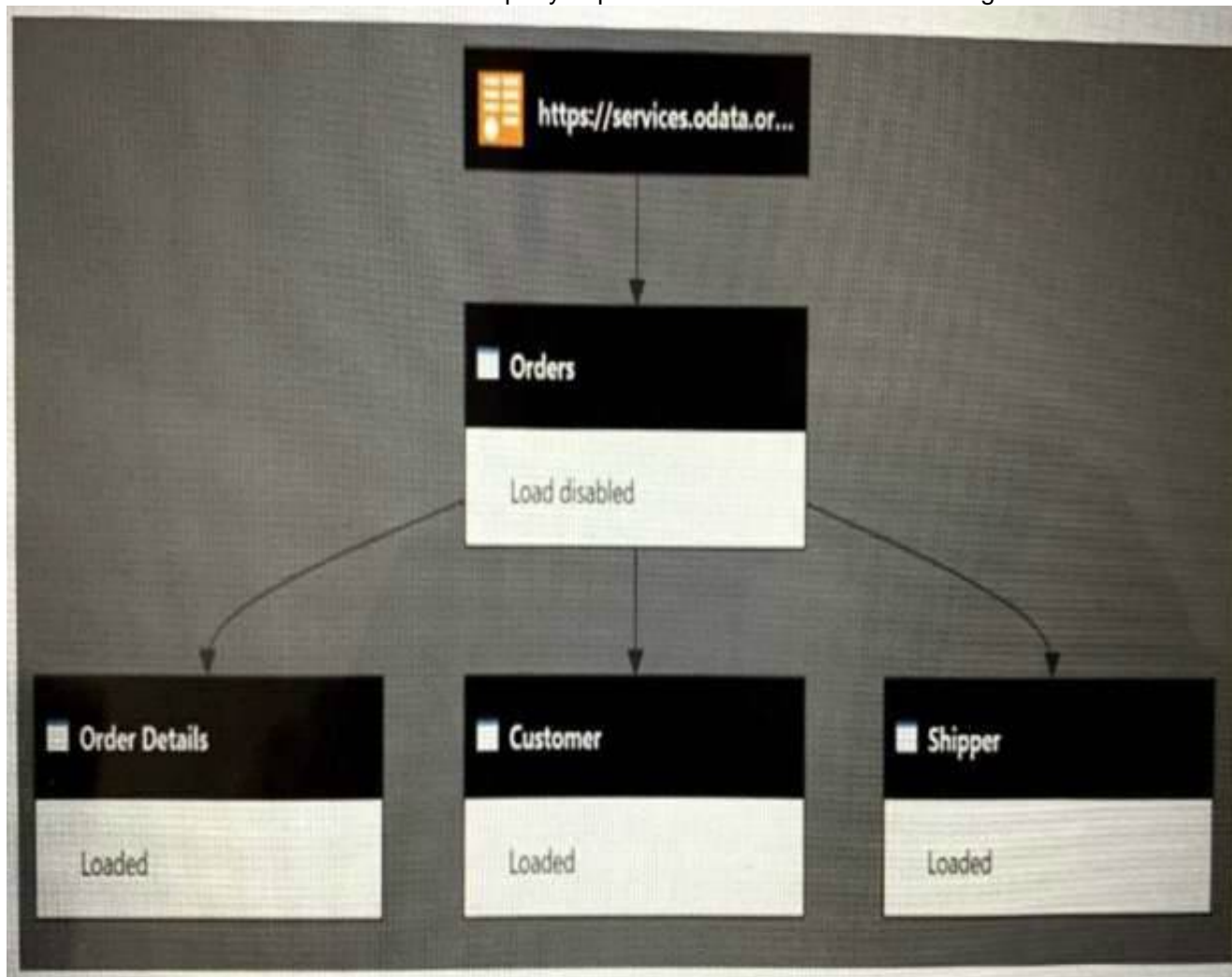
Avoid using these colors together in a chart, or on the same report page.

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

NEW QUESTION 2

- (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 3

- (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>

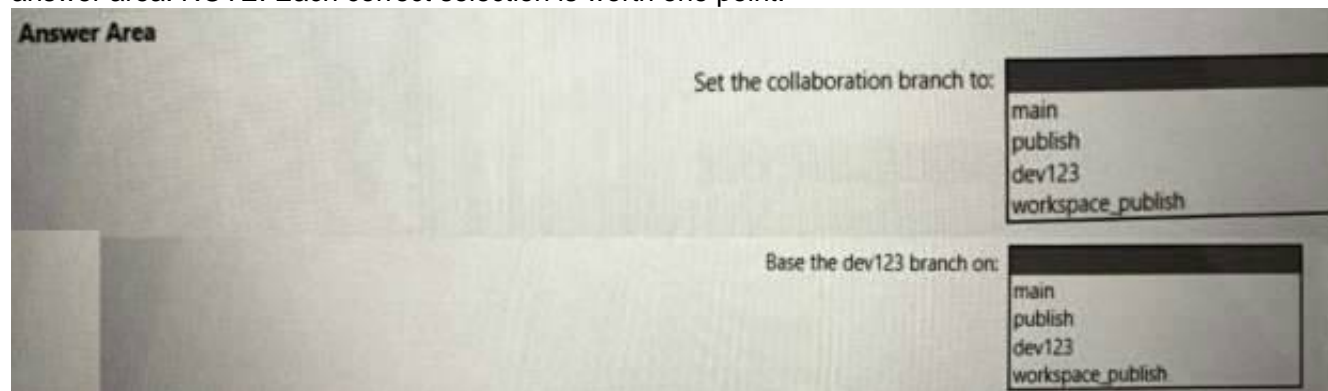
<https://powerbi.microsoft.com/en-us/blog/five-new-power-bi-premium-capacity-settings-is-available-on-the-por>

NEW QUESTION 4

- (Exam Topic 3)

You need to configure a source control solution for Azure Synapse Analytics. The solution must meet the following requirements:

- Code must always be merged to the main branch before being published, and the main branch must be used for publishing resource
- The workspace templates must be stored in the publish branch.
- A branch named dev123 will be created to support the development of a new feature. What should you do? 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: main

Code must always be merged to the main branch before being published, and the main branch must be used for publishing resources.

Collaboration branch - Your Azure Repos collaboration branch that is used for publishing. By default, its master. Change this setting in case you want to publish resources from another branch. You can select existing branches or create new.

Each Git repository that's associated with a Synapse Studio has a collaboration branch. (main or master is the default collaboration branch).

Box 2: workspace_publish

A branch named dev123 will be created to support the development of a new feature. The workspace templates must be stored in the publish branch.

Creating feature branches

Users can also create feature branches by clicking + New Branch in the branch dropdown.

By default, Synapse Studio generates the workspace templates and saves them into a branch called workspace_publish. To configure a custom publish branch, add a publish_config.json file to the root folder in the collaboration branch.

Reference: <https://docs.microsoft.com/en-us/azure/synapse-analytics/cicd/source-control>

NEW QUESTION 5

- (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

NEW QUESTION 6

- (Exam Topic 3)

You use Vertipaq Analyzer to analyze a model.

The Relationships tab contains the results shown in the following exhibit.

| IsRowNumber | Cardinality (Filter) | Relationship Type | Max From Cardinality | Max to Cardinality | 1:M Ratio | % Missing Keys | Invalid Rows | Relationships Size | Bid. Filters | MMR |
|-------------|----------------------|-------------------|----------------------|--------------------|-----------|----------------|--------------|--------------------|--------------|-----|
| | | | 84 | 2,557 | 3044.05% | 0 | 0 | 4,056 | | |
| | | | 90 | 327 | 0.69% | 22 | | 184 | | |
| | | | 26 | 164 | 0.34% | 0 | 0 | 32 | | |
| | | | 90 | 327 | 0.69% | 21 | 1,804 | 112 | | |
| | | | 7 | 6 | 0.01% | 1 | 6,577 | 8 | | |
| | | | 2 | 2 | 0.00% | 0 | 0 | 8 | | |
| | | | 16 | 84 | 0.18% | 0 | 0 | 24 | | |
| | | | 90 | 2,557 | 3044.05% | 27 | | 4,320 | | |

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

NOTE: Each correct selection is worth one point.

Answer Area

The [answer choice] table is missing records needed by the Fact table.

▼

BU Key

Customer

Date

Scenario

There are [answer choice] blank values created by missing dimensional relationships.

▼

22

1,804

6,577

8,381

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: Customer

There are 1804 invalid rows (records) in the Customer table. Box 2: 22

There are 22 missing keys.











Note: VertiPaq Analyzer in DAX Studio is useful in identifying referential integrity violations which slow

down your DAX codes. It helps you determine which table or column needs to be optimized and improved. Reference: <https://blog.enterprisedna.co/vertipaq-analyzer-tutorial-relationships-referential-integrity/>

NEW QUESTION 7

- (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 8

- (Exam Topic 3)

You have a file named File1.txt that has the following characteristics:

- A header row
- Tab delimited values
- UNIX-style line endings

You need to read File1.txt by using an Azure Synapse Analytics serverless SQL pool. Which query should you execute?

- A. `SELECT*`
`FROM OPENROWSET (`
`BULK 'file1.txt',`
`DATA_SOURCE = 'Sql1',`
`FORMAT = 'CSV', PARSER_VERSION = '2.0',`
`FIELDTERMINATOR = '\t',`
`ROWTERMINATOR = '0x0a',`
`FIRSTROW= 2`
`)`
- B. `SELECT*`
`FROM OPENROWSET (`
`BULK 'file1.txt',`
`DATA_SOURCE = 'Sql1',`
`FORMAT = 'CSV', PARSER_VERSION = '2.0',`
`FIELDTERMINATOR = ',' ,`
`ROWTERMINATOR = '\n',`
`FIRSTROW= 2`
`)`
- C. `SELECT*`
`FROM OPENROWSET (`
`BULK 'file1.txt',`
`DATA_SOURCE = 'Sql1',`
`FORMAT = 'CSV', PARSER_VERSION = '2.0',`
`FIELDTERMINATOR = ',' ,`
`ROWTERMINATOR = '0x0a',`
`FIRSTROW= 2`
`)`
- D. `SELECT*`
`FROM OPENROWSET (`
`BULK 'file1.txt',`
`DATA_SOURCE = 'Sql1',`
`FORMAT = 'CSV', PARSER_VERSION = '2.0',`
`FIELDTERMINATOR = '\t',`
`ROWTERMINATOR = '0x0a',`
`FIRSTROW= 1`
`)`

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

Answer: A

Explanation:

Use FIELDTERMINATOR = '\t' for tab.

Use ROWTERMINATOR = '0x0A ' for UNIX-style line endings Use FIRSTROW= 2 for a header row

Note: Using Row Terminators

The row terminator can be the same character as the terminator for the last field. Generally, however, a distinct row terminator is useful. For example, to produce tabular output, terminate the last field in each row with the newline character (\n) and all other fields with the tab character (\t).

If you want to output a line feed character only (LF) as the row terminator - as is typical on Unix and Linux computers - use hexadecimal notation to specify the LF row terminator. For example:

`bcp -r '0x0A' FIRSTROW`

`FIRSTROW=first_row` Specifies the number of the first row to load. The default is 1. This indicates the first row in the specified data file. The row numbers are determined by counting the row terminators. `FIRSTROW` is 1-based.

Reference:

<https://docs.microsoft.com/en-us/sql/relational-databases/import-export/specify-field-and-row-terminators-sql-se>

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

NEW QUESTION 9

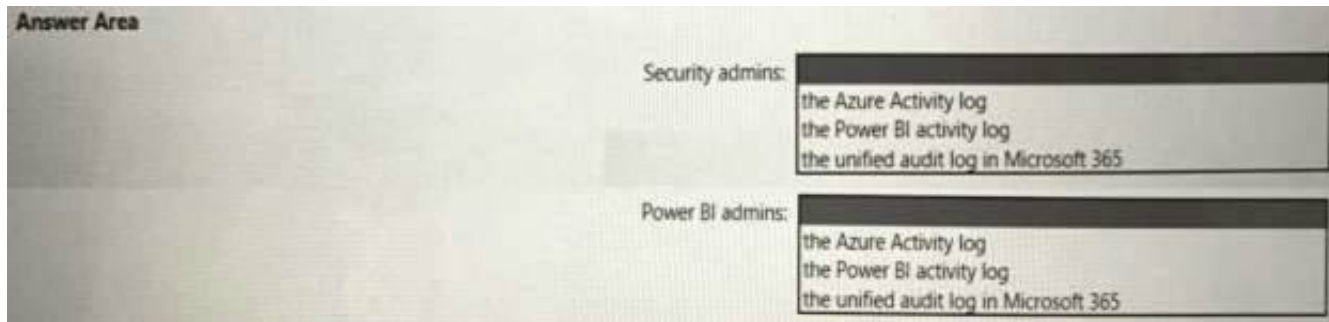
- (Exam Topic 3)

You need to recommend an automated solution to monitor Power BI user activity. The solution must meet the following requirements:

- Security admins must identify when users export reports from Power BI within five days of a new sensitivity label being applied to the artifacts in Power BI.
- Power BI admins must identify updates or changes to the Power BI capacity.
- The principle of least privilege must be used.

Which log should you include in the recommendation for each group? 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: the unified audit log in Microsoft 365

Security admins must identify when users export reports from Power BI within five days of a new sensitivity label being applied to the artifacts in Power BI.

Use the audit log

If your task is to track user activities across Power BI and Microsoft 365, you work with auditing in Microsoft 365 compliance or use PowerShell. Auditing relies on functionality in Exchange Online, which automatically supports Power BI.

You can filter the audit data by date range, user, dashboard, report, dataset, and activity type. You can also download the activities in a csv (comma-separated value) file to analyze offline.

Box 2: Power BI activity log

Power BI admins must identify updates or changes to the Power BI capacity. Use the activity log

Power BI administrators can analyze usage for all Power BI resources at the tenant level by using custom reports that are based on the Power BI activity log.

Reference: <https://docs.microsoft.com/en-us/power-bi/admin/service-admin-auditing>

NEW QUESTION 10

- (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 10

- (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 15

- (Exam Topic 3)

You have a Power BI workspace named Workspace1 that contains five dataflows.

You need to configure Workspace1 to store the dataflows in an Azure Data Lake Storage Gen2 account. What should you do first?

- A. Delete the dataflow queries.
- B. From the Power BI Admin portal, enable tenant-level storage.
- C. Disable load for all dataflow queries.
- D. Change the Data source settings in the dataflow queries.

Answer: B

Explanation:

Configuring Azure connections is an optional setting with additional properties that can optionally be set:

* Tenant Level storage, which lets you set a default, and/or

* Workspace-level storage, which lets you specify the connection per workspace

You can optionally configure tenant-level storage if you want to use a centralized data lake only, or want this to be the default option.

Reference:

<https://docs.microsoft.com/en-us/power-bi/transform-model/dataflows/dataflows-azure-data-lake-storage-integra>

NEW QUESTION 20

- (Exam Topic 3)

You plan to modify a Power BI dataset.

You open the Impact analysis panel for the dataset and select Notify contacts. Which contacts will be notified when you use the Notify contacts feature?

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

Answer: D

Explanation:

Notify contacts

If you've made a change to a dataset or are thinking about making a change, you might want to contact the relevant users to tell them about it. When you notify contacts, an email is sent to the contact lists of all the impacted workspaces. Your name appears on the email so the contacts can find you and reply back in a new email thread.

Reference: <https://docs.microsoft.com/en-us/power-bi/collaborate-share/service-dataset-impact-analysis>

NEW QUESTION 22

- (Exam Topic 3)

You are creating a Power BI single-page report.

Some users will navigate the report by using a keyboard, and some users will navigate the report by using a screen reader.

You need to ensure that the users can consume content on a report page in a logical order. What should you configure on the report page?

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

Answer: D

Explanation:

Tab order is the order in which users interact with the items on a page using the keyboard. Generally, we want tab order to be predictable and to closely match the visual order on the page (unless there is a good reason to deviate).

Note: If you are using the keyboard to navigate in a Power BI report, the order in which you arrive at visuals will not follow your vision unless you set the new tab order property. If you have low or no vision, this becomes an even bigger issue because you may not be able to see that you are navigating visuals out of visual order because the screen reader just reads whatever comes next.

Reference: <https://datasavvy.me/2018/12/26/tab-order-enhances-power-bi-report-accessibility/>

NEW QUESTION 23

- (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 Power BI Desktop, you group the measures in a display folder.

Does this meet the goal?

- A. Yes
- B. No

Answer: B

Explanation:

Solution: From DAX Studio, you write a query that uses grouping sets.

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

NEW QUESTION 24

- (Exam Topic 3)

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

You need to publish the reports to the Power BI service and configure a daily refresh of datasets. What should you include in the solution?

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

Answer: D

Explanation:

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

Reference: <https://docs.microsoft.com/en-us/power-bi/connect-data/desktop-r-in-query-editor>

NEW QUESTION 25

- (Exam Topic 3)

You manage a Power BI dataset that queries a fact table named SalesDetails. SalesDetails contains three date columns named OrderDate, CreatedOnDate, and ModifiedDate.

You need to implement an incremental refresh of SalesDetails. The solution must ensure that OrderDate starts on or after the beginning of the prior year.

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

NOTE: More than one order of answer choices is correct. You will receive credit for any of the correct orders you select.

Actions

Answer Area

- Create RangeStart and RangeEndTime parameters.
- Configure an incremental refresh to archive data that starts one year before the refresh date.
- Add an applied step that filters OrderDate to the start of the prior year.
- Configure an incremental refresh to archive data that starts two years before the refresh date.
- Add an applied step that adds a custom date filter where OrderDate is between RangeStart and RangeEnd.



- 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 RangeEndTime parameters.

When configuring incremental refresh in Power BI Desktop, you first create two Power Query date/time parameters with the reserved, case-sensitive names RangeStart and RangeEnd. These parameters, defined in the Manage Parameters dialog in Power Query Editor are initially used to filter the data loaded into the Power BI Desktop model table to include only those rows with a date/time within that period.

Step 2: Add an applied step that adds a custom date filter OrderDate is Between RangeStart and RangeEnd. With RangeStart and RangeEnd parameters defined, you then apply custom Date filters on your table's date column. The filters you apply select a subset of data that will be loaded into the model when you click Apply.

Step 3: Configure an incremental refresh to archive data that starts two years before the refresh date.

After filters have been applied and a subset of data has been loaded into the model, you then define an incremental refresh policy for the table. After the model is published to the service, the policy is used by the service to create and manage table partitions and perform refresh operations. To define the policy, you will use the Incremental refresh and real-time data dialog box to specify both required settings and optional settings.

Step 4: Add an applied step that filters OrderDate to the start of the prior year.

Reference: <https://docs.microsoft.com/en-us/power-bi/connect-data/incremental-refresh-overview>

NEW QUESTION 28

- (Exam Topic 3)

You are configuring Azure Synapse Analytics pools to support the Azure Active Directory groups shown in the following table.

| Name | Requirement |
|--------|--|
| Group1 | Analyze data to create and train machine learning models in Synapse Analytics. |
| Group2 | Execute complex queries with multiple joins against relational data. Results will be exported by using PolyBase. |
| Group3 | Query and load data from Apache Parquet files stored in Azure Data Lake Storage Gen2. Costs must be based on the amount of data processed. |

Which type of pool should each group use? To answer, drag the appropriate pool types to the groups. Each pool type may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.
NOTE: Each correct selection is worth one point.

Pool Types

Apache Spark pool

Dedicated SQL pool

Serverless SQL pool

Answer Area

Group1:

Pool Type

Group2:

Pool Type

Group3:

Pool Type

- A. Mastered
- B. Not Mastered

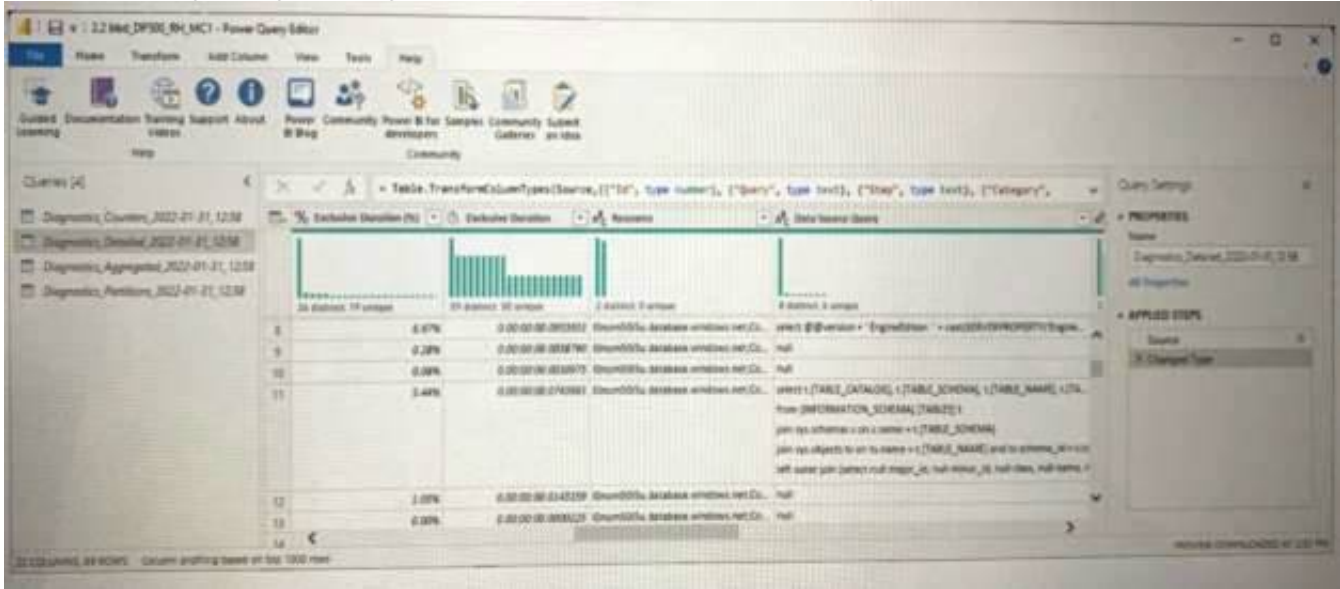
Answer: A

Explanation:

Box 1: Apache Spark pool
An Apache Spark pool provides open-source big data compute capabilities. After you've created an Apache Spark pool in your Synapse workspace, data can be loaded, modeled, processed, and distributed for faster analytic insight.
Box 2: Dedicated SQL Pool
Dedicated SQL Pool - Data is stored in relational tables Box 3: Serverless SQL pool
Serverless SQL pool - Cost is incurred for the data processed per query
Reference:
<https://docs.microsoft.com/en-us/azure/synapse-analytics/quickstart-create-apache-spark-pool-portal>
<https://www.royalcyber.com/blog/data-services/dedicated-sql-pool-vs-serverless-sql/>

NEW QUESTION 33

- (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 34

- (Exam Topic 3)

You have a deployment pipeline for a Power BI workspace. The workspace contains two datasets that use import storage mode.

A database administrator reports a drastic increase in the number of queries sent from the Power BI service to an Azure SQL database since the creation of the deployment pipeline.

An investigation into the issue identifies the following:

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

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

What should you recommend?

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

Answer: C

Explanation:

Previously in Power BI Desktop, when you used a DirectQuery in a report, no other data connections, whether DirectQuery or import, were allowed for that report. With composite models, that restriction is removed. A report can seamlessly include data connections from more than one DirectQuery or import data connection, in any combination you choose.

The composite models capability in Power BI Desktop consists of three related features:

* Composite models: Allows a report to have two or more data connections from different source groups, such as one or more DirectQuery connections and an import connection, two or more DirectQuery connections, or any combination thereof.

* Etc.

Reference: <https://docs.microsoft.com/en-us/power-bi/transform-model/desktop-composite-models>

NEW QUESTION 35

- (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 40

- (Exam Topic 3)

You are implementing a reporting solution that has the following requirements:

- Reports for external customers must support 500 concurrent requests. The data for these reports is approximately 7 GB and is stored in Azure Synapse Analytics.
- Reports for the security team use data that must have local security rules applied at the database level to restrict access. The data being reviewed is 2 GB.

Which storage mode provides the best response time for each group of users?

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

Answer: A

Explanation:

With DirectQuery, queries are sent back to your Azure Synapse Analytics in real time as you explore the data. Real-time queries, combined with the scale of Synapse Analytics enables users to create dynamic reports in minutes against terabytes of data.

Need import for the security team for local security rules. Reference:

<https://docs.microsoft.com/en-us/power-bi/connect-data/service-azure-sql-data-warehouse-with-direct-connect>

NEW QUESTION 44

- (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 45

- (Exam Topic 3)

You are building a Power BI dataset that contains a table named Calendar. Calendar contains the following calculated column.

pfFlag = IF('Calendar'[Date] < TODAY(), "Past", "Future")

You need to create a measure that will perform a fiscal prior year-to-date calculation that meets the following requirements:

- Returns the fiscal prior year-to-date value for [Sales Amount]
- Uses a fiscal year end of June 30
- Produces no result for dates in the future

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

Answer Area

PYTD =

CALCULATE (

[Sales Amount],

CALCULATETABLE

CROSSJOIN

SUMMARIZECOLUMNS

UNION

)

DATEADD

FIRSTDATE

PARALLELPERIOD

SAMEPERIODLASTYEAR

(DATESYTD ('Calendar'[Date], "6/30/2022")),

'Calendar'[pfFlag] =

"Future"

NOW()

"Past"

TODAY()

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: CALCULATETABLE

CALCULATETABLE evaluates a table expression in a modified filter context. Syntax: CALCULATETABLE(<expression>[, <filter1> [, <filter2> [, ...]]]) Incorrect:

* SUMMARIZECOLUMNS

SUMMARIZECOLUMNS returns a summary table over a set of groups.

Syntax: SUMMARIZECOLUMNS(<groupBy_columnName> [, <groupBy_columnName >]..., [<filterTable>]...[, <name>, <expression>]...)

* CROSSJOIN returns a table that contains the Cartesian product of all rows from all tables in the arguments. The columns in the new table are all the columns in all the argument tables.

Syntax: CROSSJOIN(<table>, <table>[, <table>]...)

* UNION creates a union (join) table from a pair of tables.

Syntax: UNION(<table_expression1>, <table_expression2> [,<table_expression>]...)

Box 2: SAMEPERIODLASTYEAR
SAMEPERIODLASTYEAR returns a table that contains a column of dates shifted one year back in time from the dates in the specified dates column, in the current context.

Syntax: SAMEPERIODLASTYEAR(<dates>)

The dates returned are the same as the dates returned by this equivalent formula: DATEADD(dates, -1, year) Example:

The following sample formula creates a measure that calculates the previous year sales of Reseller sales.

= CALCULATE(SUM(ResellerSales_USD[SalesAmount_USD]), SAMEPERIODLASTYEAR(DateTime[DateKey]))

Box 3: TODAY()

TODAY() returns the current date.

The TODAY function is useful when you need to have the current date displayed on a worksheet, regardless of when you open the workbook. It is also useful for calculating intervals.

Example:

The following sample formula creates a measure that calculates the 'Running Total' for Internet sales.

= CALCULATE(SUM(InternetSales_USD[SalesAmount_USD]), DATESYTD(DateTime[DateKey])) Reference: <https://docs.microsoft.com/en-us/dax/calculatetable-function-dax> <https://docs.microsoft.com/en-us/dax/sameperiodlastyear-function-dax>

<https://docs.microsoft.com/en-us/dax/datesytd-function-dax>

NEW QUESTION 46

- (Exam Topic 3)

You have a group of data scientists who must create machine learning models and run periodic experiments on a large dataset.

You need to recommend an Azure Synapse Analytics pool for the data scientists. The solution must minimize costs.

Which type of pool should you recommend?

- A. a Data Explorer pool
- B. an Apache Spark pool
- C. a dedicated SQL pool
- D. a serverless SQL pool

Answer: B

Explanation:

In Azure Synapse, training machine learning models can be performed on the Apache Spark Pools with tools like PySpark/Python, Scala, or .NET.

Reference:

<https://docs.microsoft.com/en-us/azure/synapse-analytics/machine-learning/what-is-machine-learning>

NEW QUESTION 51

- (Exam Topic 3)

You need to provide users with a reproducible method to connect to a data source and transform the data by using an AI function. The solution must meet the following requirement

- Minimize development effort.
- Avoid including data in the file. Which type of file should you create?

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

Answer: C

Explanation:

A PBIT file is a template created by Power BI Desktop, a Microsoft application used to create reports and visualizations. It contains queries, visualization settings, data models, reports, and other data added by the user.

A PBIT file acts as a Power BI template. It doesn't include any data from your source systems. Reference: <https://docs.microsoft.com/en-us/power-bi/connect-data/desktop-data-sources>

NEW QUESTION 52

- (Exam Topic 3)

You develop a solution that uses a Power BI Premium capacity. The capacity contains a dataset that is expected to consume 50 GB of memory.

Which two actions should you perform to ensure that you can publish the model successfully to the Power BI service? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

- A. Increase the Max Offline Dataset Size setting.
- B. Invoke a refresh to load historical data based on the incremental refresh policy.
- C. Restart the capacity.
- D. Publish an initial dataset that is less than 10 GB.
- E. Publish the complete dataset.

Answer: BE

Explanation:

Enable large datasets

Steps here describe enabling large datasets for a new model published to the service. For existing datasets, only step 3 is necessary.
 Create a model in Power BI Desktop. If your dataset will become larger and progressively consume more memory, be sure to configure Incremental refresh.
 Publish the model as a dataset to the service.
 In the service > dataset > Settings, expand Large dataset storage format, set the slider to On, and then select Apply.
 Enable large dataset slider
 Invoke a refresh to load historical data based on the incremental refresh policy. The first refresh could take a while to load the history. Subsequent refreshes should be faster, depending on your incremental refresh policy.
 Reference: <https://docs.microsoft.com/en-us/power-bi/enterprise/service-premium-large-models>

NEW QUESTION 55

- (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 59

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

Manage aggregations

Aggregations accelerate query performance to unlock big-data sets. [Learn more](#)

Aggregation table

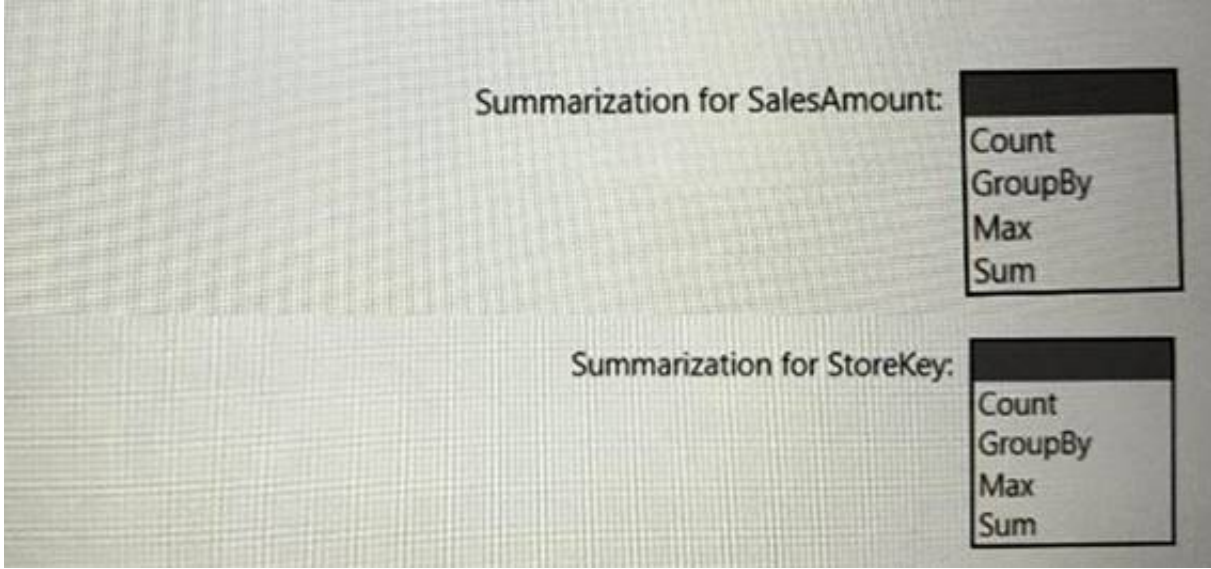
FactSales(Agg)

Precedence ⓘ

0

| | | | | |
|---------------|------------------------|--|--|--|
| ProductKey | Select Summarizatio... | | | |
| PromotionKey | Select Summarizatio... | | | |
| SalesAmount | Select Summarizatio... | | | |
| SalesQuantity | Select Summarizatio... | | | |
| StoreKey | Select Summarizatio... | | | |
| TotalCost | Select Summarizatio... | | | |

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



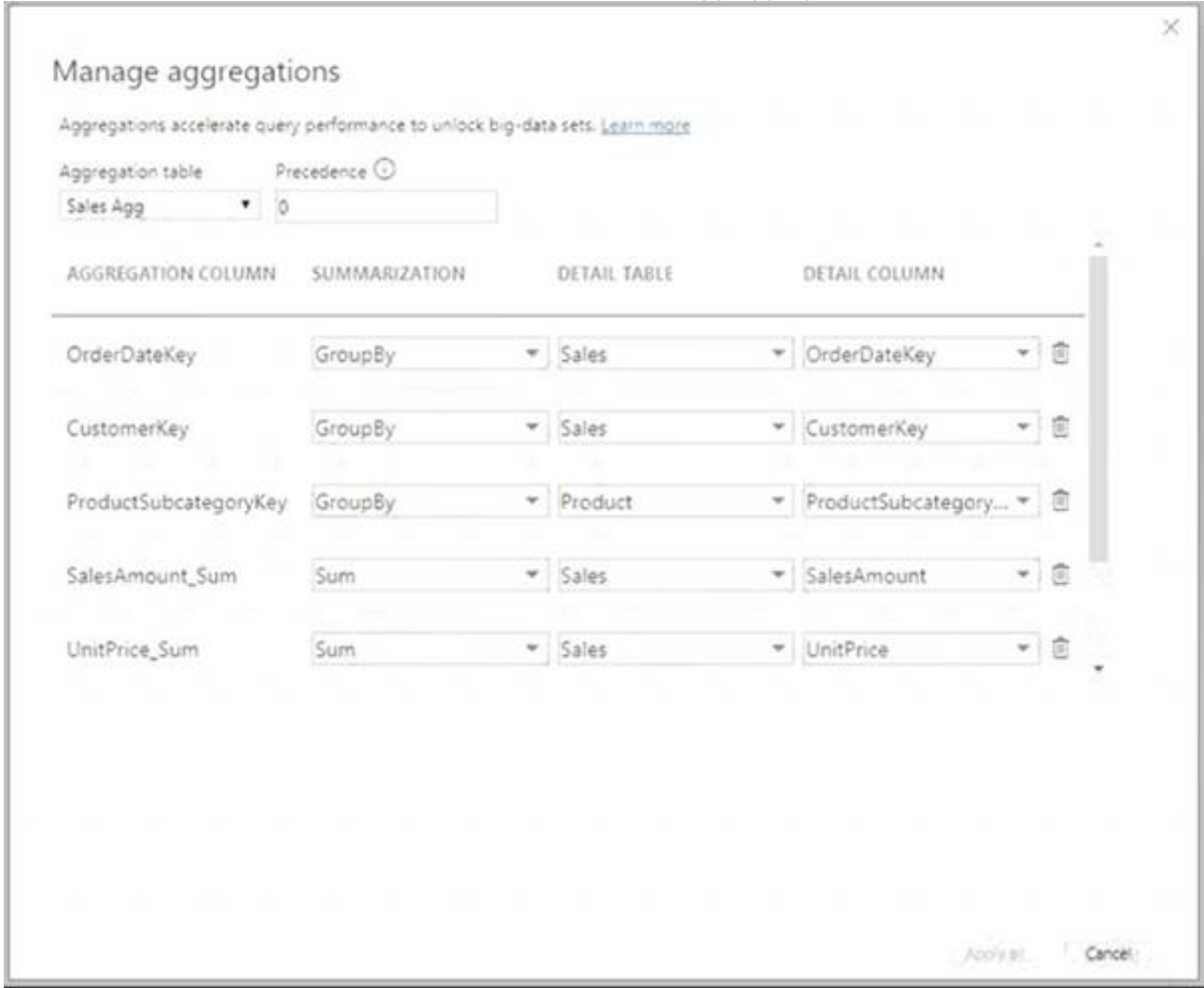
- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: Sum

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



Box 2: GroupBy

Reference: <https://docs.microsoft.com/en-us/power-bi/transform-model/aggregations-advanced>

NEW QUESTION 62

- (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 64

- (Exam Topic 3)
 You are optimizing a Power BI data model by using DAX Studio.
 You need to capture the query events generated by a Power BI Desktop report. What should you use?

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

Answer: C

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

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

NEW QUESTION 68

- (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

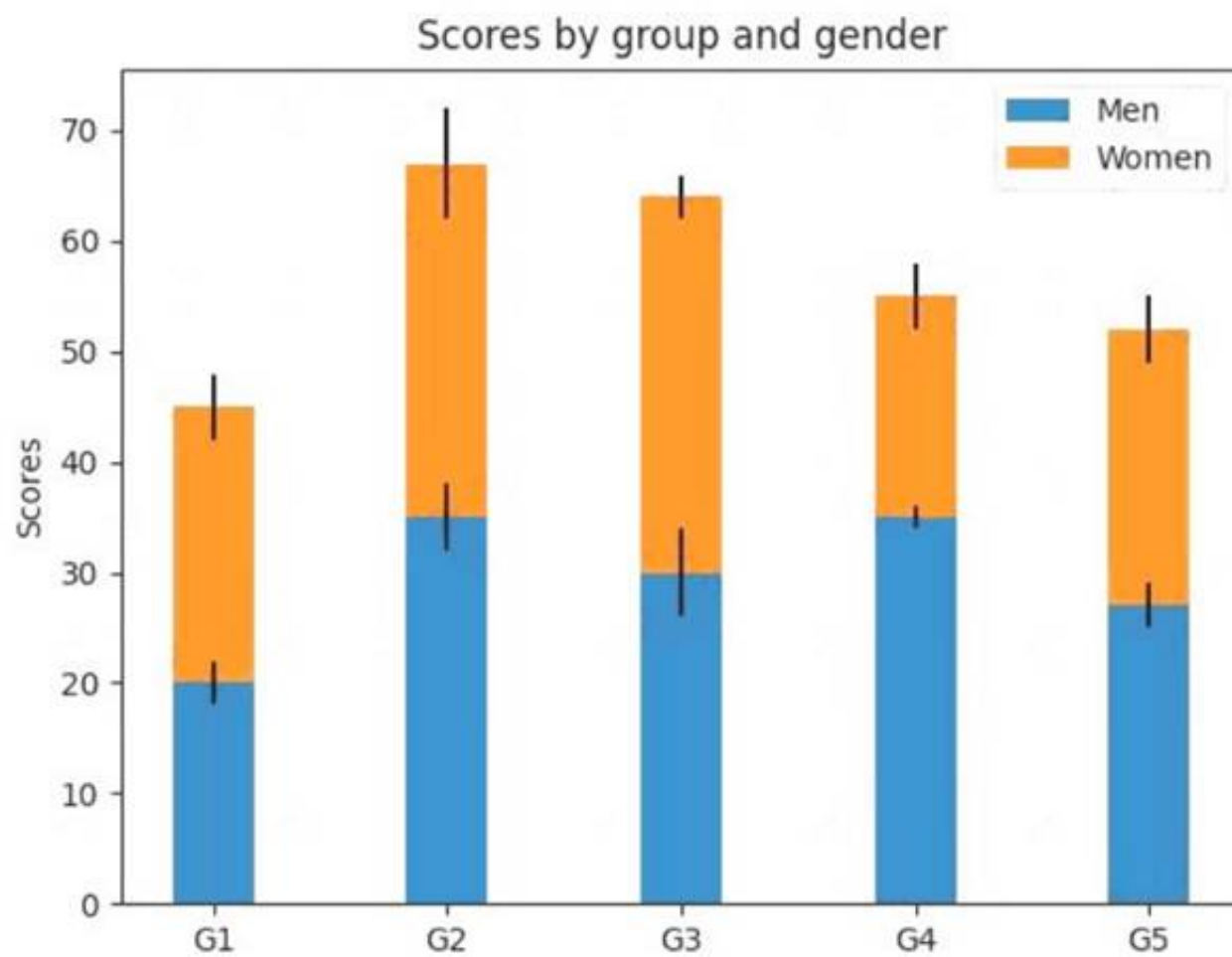
seven items

- 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 year 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/gallery/lines_bars_and_markers/bar_stacked.html
https://matplotlib.org/stable/api/legend_api.html

NEW QUESTION 73

- (Exam Topic 3)

You have a Power BI data model.

You need to refresh the data from the source every 15 minutes. What should you do first?

- A. Enable the XMLA endpoint.
- B. Define an incremental refresh policy.
- C. Change the storage mode of the dataset.
- D. Configure a scheduled refresh.

Answer: D

Explanation:

To get to the Scheduled refresh screen:

- * 1. In the navigation pane, under Datasets, select More options (...) next to a dataset listed.
- * 2. Select Schedule refresh.

Reference: <https://docs.microsoft.com/en-us/power-bi/connect-data/refresh-scheduled-refresh>

NEW QUESTION 77

- (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 80

- (Exam Topic 3)

You are using an Azure Synapse Analytics serverless SQL pool to query network traffic logs in the Apache Parquet format. A sample of the data is shown in the following table.

| source | | destination | |
|-----------|-------------|-------------|---------|
| name | ip | name | ip |
| Network01 | 192.168.0.1 | Internet | 0.0.0.0 |

You need to create a Transact-SQL query that will return the source IP address.

Which function should you use in the select statement to retrieve the source IP address?

- A. JSON_VALUE
- B. FOR.JSON
- C. CONVERT
- D. FIRST VALUE

Answer: A

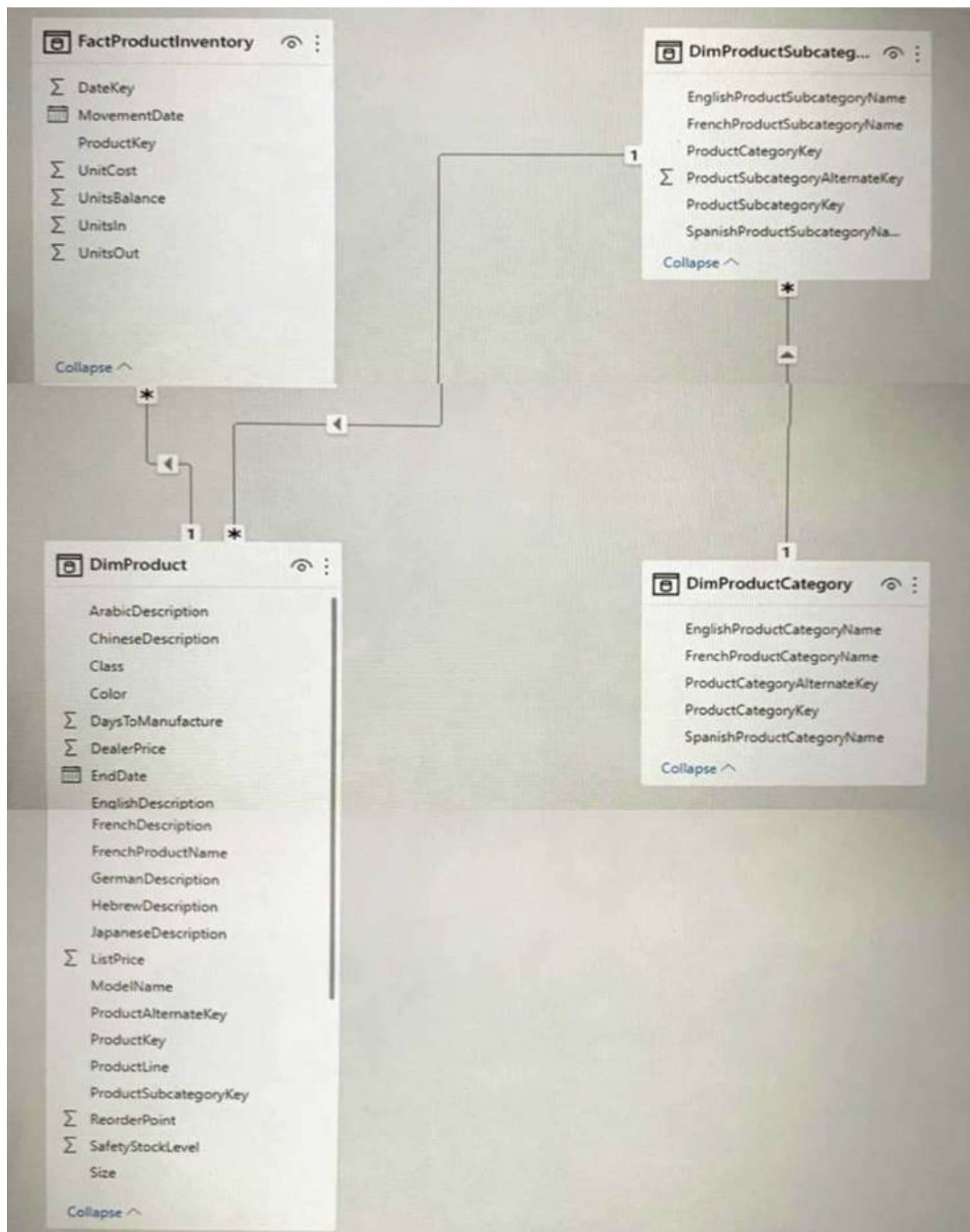
NEW QUESTION 81

- (Exam Topic 3)

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

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

You have the Power BI data model shown in the exhibit (Click the Exhibit tab.)



Users indicate that when they build reports from the data model, the reports take a long time to load. You need to recommend a solution to reduce the load times of the reports.

Solution: You recommend denormalizing the data model. Does this meet the goal?

- A. Yes
- B. No

Answer: A

Explanation:

Denormalize For Performance.

Even though it might mean storing a bit of redundant data, schema denormalization can sometimes provide better query performance. The only question then becomes is the extra space used worth the performance benefit.

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

NEW QUESTION 85

- (Exam Topic 3)

You have a Power BI dataset that uses DirectQuery against an Azure SQL database.

Multiple reports use the dataset.

A database administrator reports that too many queries are being sent from Power BI to the database. You need to reduce the number of queries sent to the database. The solution must meet the following requirements:

- DirectQuery must continue to be used.
- Visual interactions in all the reports must remain as they are configured currently.
- Consumers of the reports must only be allowed to apply filters from the Filter pane. Which two settings should you select? Each correct answer presents part of

the solution. NOTE: Each correct selection is worth one point.

- A. Disabling cross highlighting/filtering by default
- B. Add a single Apply button to the filter pane to apply changes at once
- C. Add an Apply button to each slicer to apply changes when you're ready
- D. Add Apply buttons to all basic filters to apply changes when you're ready
- E. Ignore the Privacy Levels and potentially improve performance

Answer: BC

Explanation:

Reduce queries

Reduce the number of queries sent by Power BI using the Query reduction settings. For slicers, select the "Add an Apply button to each slicer to apply changes when you're ready" option. For filters, select "Add a single Apply button to the filter pane to apply changes at once (preview)."

Reference: <https://maqsoftware.com/insights/power-bi-best-practices>

NEW QUESTION 86

- (Exam Topic 3)

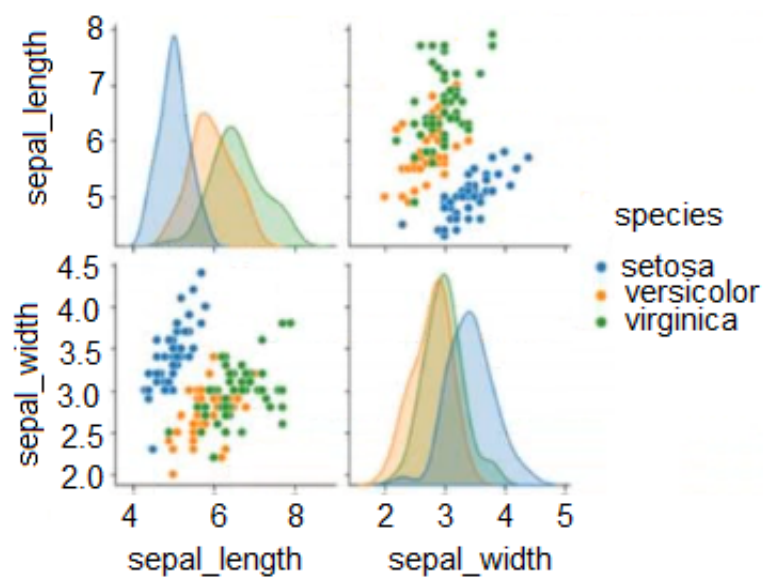
You are using an Azure Synapse notebook to create a Python visual. You run the following code cell to import a dataset named Iris.

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

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

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

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



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

Answer Area

```
sns. boxplot (iris, hue= 'sepal_length', height=2.5)

plt.show()
```

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

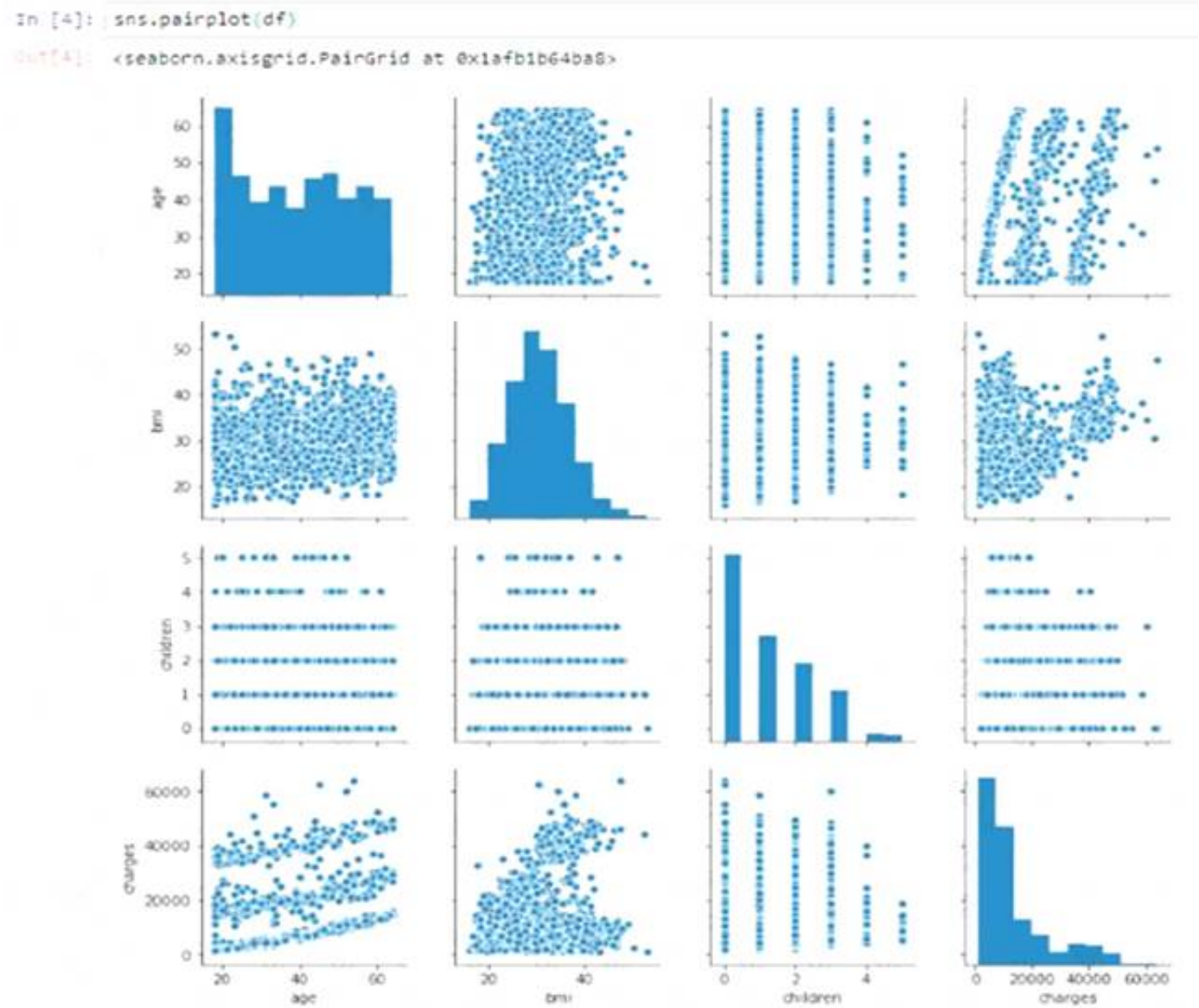
Box 1: pairplot

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 87

- (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 92

- (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 94

- (Exam Topic 2)

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

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

NOTE: Each correct selection is worth one point.

```
DECLARE @model varbinary(max) = (
    SELECT native_model_object
    FROM ml_models
    WHERE model_name = 'rxLinMod'
    AND model_version = 'v1');
SELECT d.*, p.*
FROM [ ] (MODEL = @model, DATA = dbo.rx_linMod as lm)
    EVALUATE
    PIVOT
    PREDICT
    SCORE

go [ ] (model_outcome float, trade_volume float, price_Pred float) as p;
    AS
    CONTAINS
    FROM
    GROUP BY
    WITH
```

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: PREDICT

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

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

-- Query for ML predictions SELECT d.*, p.Score

FROM PREDICT(MODEL = (SELECT Model FROM Models WHERE Id = 1),

DATA = dbo.mytable AS d, RUNTIME = ONNX) WITH (Score float) AS p; Box 2: WITH

Reference:

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

NEW QUESTION 98

- (Exam Topic 2)

You need to identify the root cause of the data refresh issue. What should you use?

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

Answer: B

Explanation:

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

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

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

NEW QUESTION 99

- (Exam Topic 2)

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

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

| Actions | Answer Area |
|--|-------------|
| Create a perspective. | |
| Enable bidirectional filtering. | |
| Create a DAX expression. | |
| Create row-level security (RLS) roles. | |
| Add members to row-level security (RLS) roles. | |

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

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

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

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

Deploy prebuilt datasets to Power BI to simplify the query experience of the customers. Step 2: Create a DAX expression

Consider a model with two roles: The first role, named Workers, restricts access to all Payroll table rows by using the following rule expression: FALSE()

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

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

TRUE()

Take care: Should a report user map to both roles, they'll see all Payroll table rows. Step 3: Add members to row-level security (RLS) roles

Configure role mappings

Once [the model is] published to Power BI, you must map members to dataset roles. Reference: <https://docs.microsoft.com/en-us/power-bi/guidance/rls-guidance>

NEW QUESTION 102

- (Exam Topic 2)

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

What should you include in the recommendation?

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

Answer: D

Explanation:

Refresh data from an on-premises SQL Server database

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

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

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

NEW QUESTION 105

- (Exam Topic 1)

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

- A. an incremental refresh of Product where the ModifiedDate value is during the last three days.
- B. an incremental refresh of dbo.SalesTransactions where the SalesDate value is during the last three days.
- C. a full refresh of all the tables
- D. an incremental refresh of dbo.SalesTransactions where the SalesDate value is during the last hour.

Answer: B

Explanation:

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

NEW QUESTION 110

- (Exam Topic 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 113

- (Exam Topic 1)
You need to populate the CustomersWithProductScore table.
How should you complete the stored procedure? To answer, select the appropriate options in the answer area. NOTE: Each correct selection is worth one point.

Answer Area

```
DECLARE @model
SELECT model
FROM MLModel
WHERE model_name = 'PredictPurchase'
);
INSERT INTO CustomersWithProductScore (
    CustomerID
    ,CustomerEmail
    ,ProductID
    ,ProductName
    ,Score
)
SELECT d.CustomerID
    ,d.CustomerEmail
    ,d.ProductID
    ,d.ProductName
    ,p.score
FROM PREDICT(MODEL = @model, DATA =
    WITH (score FLOAT) AS p;
```

Box 1:
BIT
FLOAT
NVARCHAR(1000)
VARBINARY(max)

Box 2:
dbo.Customer
dbo.CustomerPurchases
dbo.CustomersWithProductScore
dbo.Product

AS d)

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: FLOAT
Identify which customers should receive promotional emails based on their likelihood of purchasing promoted products.
FLOT is used in the last statement of the code: WITH (score FLOAT) as p; From syntax: MODEL
The MODEL parameter is used to specify the model used for scoring or prediction. The model is specified as a variable or a literal or a scalar expression.
Box 2: dbo.CustomerWithProductScore
Identify which customers should receive promotional emails based on their likelihood of purchasing promoted products.
Only table CustomerWithProductScore has the required filed score.
From the syntax: DATA
The DATA parameter is used to specify the data used for scoring or prediction. Data is specified in the form of a table source in the query. Table source can be a table, table alias, CTE alias, view, or table-valued function.
Reference: <https://docs.microsoft.com/en-us/sql/t-sql/queries/predict-transact-sql>

NEW QUESTION 118

.....

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](#)