

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

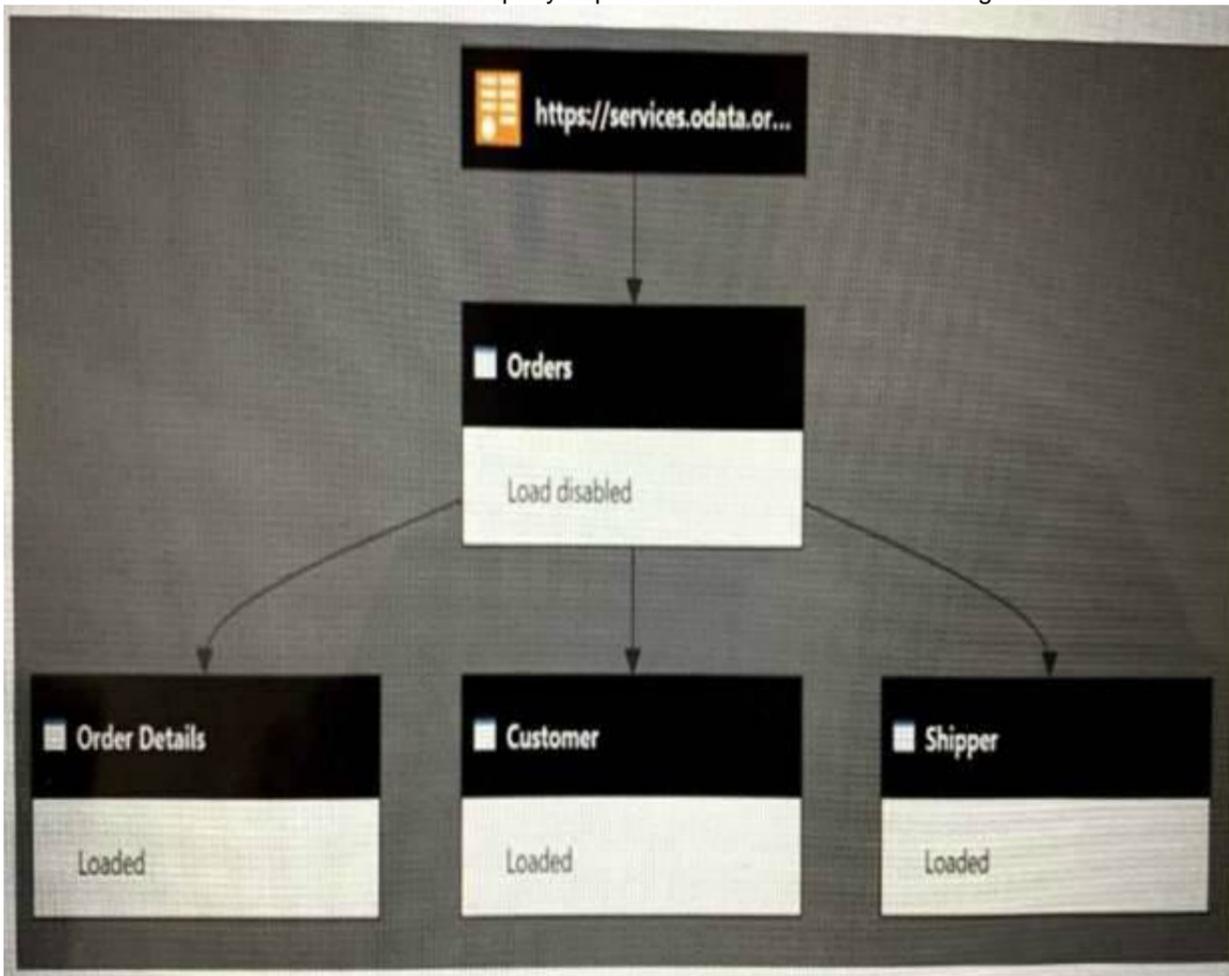
Avoid using these colors together in a chart, or on the same report page.

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

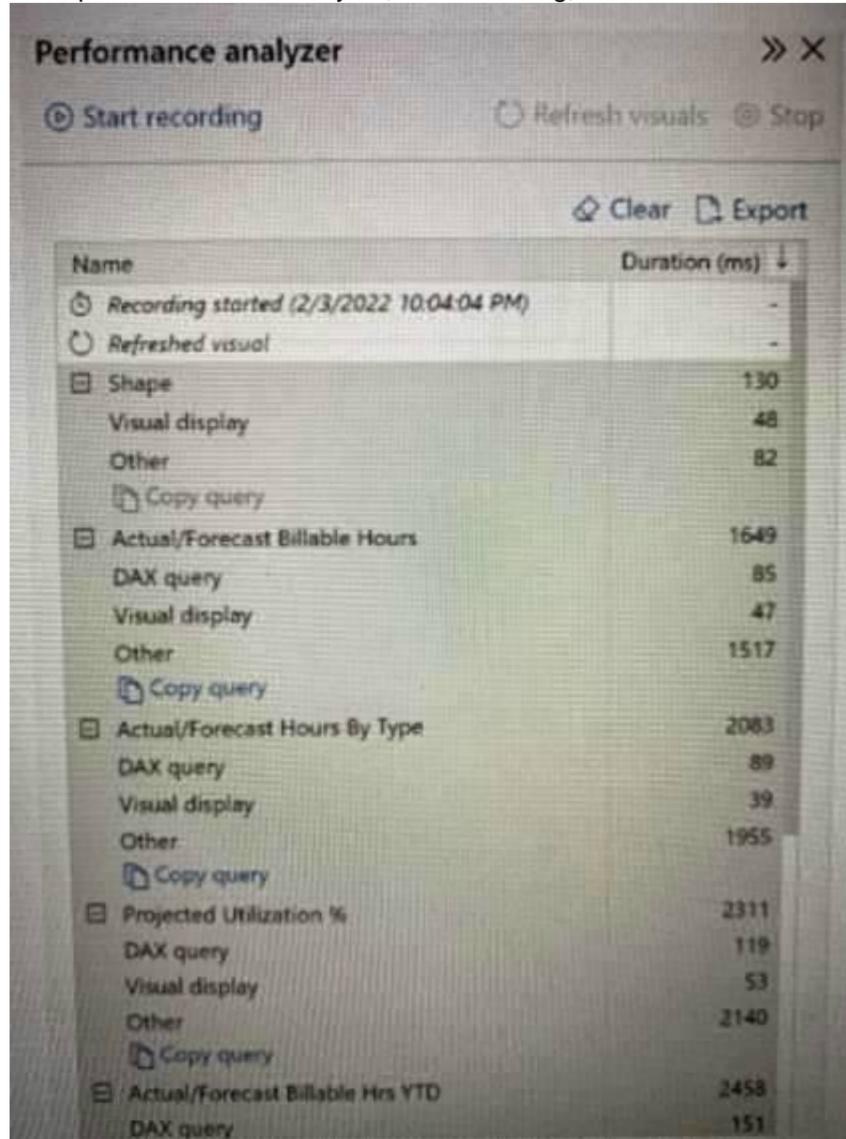
Reference: <https://radacad.com/performance-tip-for-power-bi-enable-load-sucks-memory-up> <https://docs.microsoft.com/en-us/powerquery-m/table-buffer>

NEW QUESTION 3

- (Exam Topic 3)

You open a Power BI Desktop report that contains an imported data model and a single report page.

You open Performance analyzer, start recording, and refresh the visuals on the page. The recording produces the results shown in the following exhibit



What can you identify from the results?

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

Answer: D

Explanation:

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

Note: Each visual's log information includes the time spent (duration) to complete the following categories of tasks:

DAX query - if a DAX query was required, this is the time between the visual sending the query, and for Analysis Services to return the results.

Visual display - time required for the visual to draw on the screen, including time required to retrieve any web images or geocoding.

Other - time required by the visual for preparing queries, waiting for other visuals to complete, or performing other background processing.

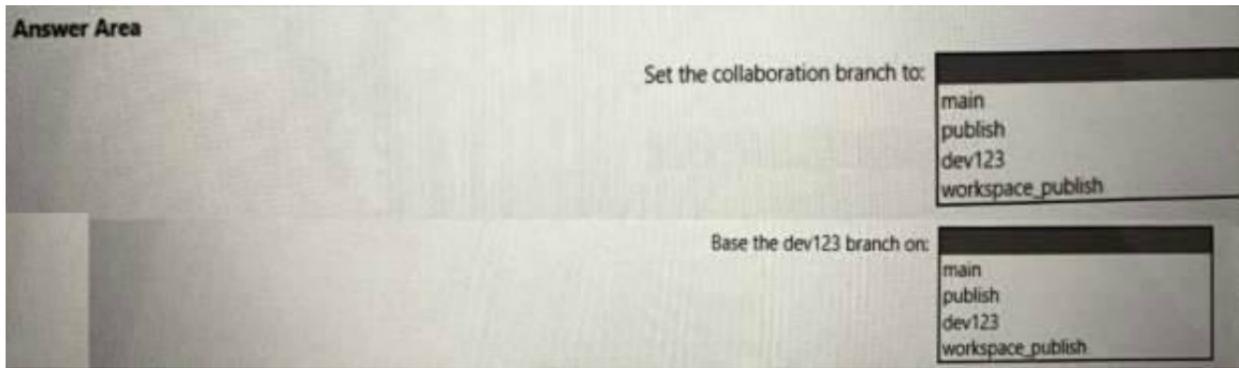
Reference: <https://docs.microsoft.com/en-us/power-bi/create-reports/desktop-performance-analyzer>

NEW QUESTION 4

- (Exam Topic 3)

You 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 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 6

- (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	Red
5	Fama	\$6,097	Red
13	Contoso	\$5,214	Red
11	Pomum	\$4,968	Yellow
7	VanArsdel	\$4,964	Yellow
10	Pirum	\$4,644	Yellow
2	Aliqui	\$4,479	Yellow
1	Abbas	\$4,070	Yellow
8	Natura	\$3,376	Yellow
14	Victoria	\$2,317	Green
4	Salvus	\$2,296	Green
12	Quibus	\$2,208	Green
3	Barba	\$1,601	Green
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 7

- (Exam Topic 3)

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

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

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

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

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

Answer: A

Explanation:

Use FIELDTERMINATOR = '\t' for tab.

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

Note: Using Row Terminators

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

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

bcp -r '0x0A' FIRSTROW

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

Reference:

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

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

NEW QUESTION 8

- (Exam Topic 3)

You have an Azure Synapse Analytics dedicated SQL pool.

You need to ensure that the SQL pool is scanned by Azure Purview. What should you do first?

- A. Register a data source.
 B. Search the data catalog.
 C. Create a data share connection.
 D. Create a data policy.

Answer: B

NEW QUESTION 9

- (Exam Topic 3)

After you answer a question in this section, you will NOT be able to return to it. As a result, these questions will not appear in the review screen.
 You are using an Azure Synapse Analytics serverless SQL pool to query a collection of Apache Parquet files by using automatic schema inference. The files contain more than 40 million rows of UTF-8-encoded business names, survey names, and participant counts. The database is configured to use the default collation.

The queries use open row set and infer the schema shown in the following table.

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

You need to recommend changes to the queries to reduce I/O reads and tempdb usage.

Solution: You recommend using openrowset with to explicitly define the collation for businessName and surveyName as 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 have a Power BI workspace that contains one dataset and four reports that connect to the dataset. The dataset uses Import storage mode and contains the following data sources:

- A CSV file in an Azure Storage account
- An Azure Database for PostgreSQL database

You plan to use deployment pipelines to promote the content from development to test to production. There will be different data source locations for each stage. What should you include in the deployment pipeline to ensure that the appropriate data source locations are used during each stage?

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

Answer: A

Explanation:

Note: Create deployment rules

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

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

NEW QUESTION 10

- (Exam Topic 3)

You 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 13

- (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 14

- (Exam Topic 3)

You are creating an external table by using an Apache Spark pool in Azure Synapse Analytics. The table will contain more than 20 million rows partitioned by date. The table will be shared with the SQL engines.

You need to minimize how long it takes for a serverless SQL pool to execute a query data against the table. In which file format should you recommend storing the table data?

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

Answer: B

Explanation:

Prepare files for querying

If possible, you can prepare files for better performance:

* Convert large CSV and JSON files to Parquet. Parquet is a columnar format. Because it's compressed, its file sizes are smaller than CSV or JSON files that contain the same data. Serverless SQL pool skips the columns and rows that aren't needed in a query if you're reading Parquet files. Serverless SQL pool needs less time and fewer storage requests to read it.

Reference: <https://docs.microsoft.com/en-us/azure/synapse-analytics/sql/best-practices-serverless-sql-pool> <https://stackoverflow.com/questions/65320949/parquet-vs-delta-format-in-azure-data-lake-gen-2-store>

NEW QUESTION 19

- (Exam Topic 3)

You are using DAX Studio to analyze a slow-running report query. You need to identify inefficient join operations in the query. What should you review?

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

Answer: B

Explanation:

Open DAX Studio.

Paste the query there, enable Query Plan display and Server Timings, run your query (with clear cache), and then study the query plan for large row counts. Once the culprit is identified you can decide how to rewrite your DAX to make that part faster.

Reference: <https://community.powerbi.com/t5/Power-Query/DAX-Query-taking-longer-time/td-p/1171961> <https://www.sqlbi.com/wp-content/uploads/DAX-Query-Plans.pdf>

NEW QUESTION 21

- (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. corr
- C. summary
- D. rollup

Answer: B

Explanation:

pandas.DataFrame.corr computes pairwise correlation of columns, excluding NA/null values. Incorrect:

* freqItems pyspark.sql.DataFrame.freqItems

Finding frequent items for columns, possibly with false positives. Using the frequent element count algorithm described in <https://doi.org/10.1145/762471.762473>, proposed by Karp, Schenker, and Papadimitriou.'

* summary is used for index.

* There is no panda method for rollup. Rollup would not be correct anyway. Reference: <https://pandas.pydata.org/docs/reference/api/pandas.DataFrame.corr.html>

NEW QUESTION 22

- (Exam Topic 3)

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

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

You have a Power 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 create a hierarchy.

Does this meet the goal?

- A. Yes
- B. No

Answer: B

Explanation:

Instead use the solution: From DAX Studio, you write a query that uses grouping sets. A grouping is a set of discrete values that are used to group measure fields. Note: A hierarchy is an ordered set of values that are linked to the level above. An example of a hierarchy could be Country, State, and City. Cities are in a State, and States make up a Country. In Power BI visuals can handle hierarchy data and provide controls for the user to navigate up and down the hierarchy.

Reference: <https://docs.microsoft.com/en-us/power-bi/developer/visuals/capabilities> <https://powerbi.tips/2018/09/how-to-navigate-hierarchies/>

NEW QUESTION 23

- (Exam Topic 3)

You have an Azure Synapse Analytics serverless SQL pool and an Azure Data Lake Storage Gen2 account. You need to query all the files in the 'csv/taxi/' folder and all its subfolders. All the files are in CSV format and have a header row.

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

NOTE: Each correct selection is worth one point.

SELECT*

FROM OPENROWSET (

	▼
BULK 'csv/taxi',	
BULK 'csv/taxi/**',	
BULK 'csv/taxi/*.csv',	
BULK 'csv/taxi/',	

DATA_SOURCE = 'datalake',
 FORMAT = 'CSV', PARSER_VERSION = '2.0',

	▼
FIRSTROW = 0	
FIRSTROW = 1	
FIRSTROW = -1	
FIRSTROW = 2	

)
 WITH (
 pickup_datetime DATETIME2,
 passenger_count INT,
 trip_distance FLOAT,
 total_amount FLOAT
) AS nyc;

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: BULK 'csv/taxi*.CSV',

*.CSV to get all the CSV files. Box 2: FIRSTROW=2

As there is a header we should read from the second line. Note: FIRSTROW = 'first_row'

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

Incorrect:

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

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

NEW QUESTION 26

- (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.



- 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 30

- (Exam Topic 3)

Your company is migrating its current, custom-built reporting solution to Power BI. The Power BI tenant must support the following scenarios:

- > 40 reports that will be embedded in external websites. The websites control their own security. The reports will be consumed by 50 users monthly.
- > Forty-five users that require access to the workspaces and apps in the Power BI Admin portal. Ten of the users must publish and consume datasets that are larger than 1 GB.
- > Ten developers that require Text Analytics transformations and paginated reports for datasets. An additional 15 users will consume the reports.

You need to recommend a licensing solution for the company. The solution must minimize costs.

Which two Power BI license options should you include in the recommendation? Each correct answer presents part of the solution.

NOTE: Each correct selection is worth one point.

- A. 70 Premium per user
- B. one Premium
- C. 70 Pro
- D. one Embedded
- E. 35 Pro
- F. 35 Premium per user

Answer: BF

Explanation:

B:

Free - 40 reports that will be embedded in external websites. The websites control their own security. Free - The reports will be consumed by 50 users monthly.

Free + 1 Premium for the Workspace -Forty-five users that require access to the workspaces and apps in the Power BI Admin portal.

F: Ten of the users must publish and consume datasets that are larger than 1 GB.

Ten developers that require Text Analytics transformations and paginated reports for datasets. An additional 15 users will consume the reports.

Power BI Premium per user features and capabilities

* Pixel perfect paginated reports are available for operational reporting capabilities based on SSRS technology. Users can create highly formatted reports in various formats such as PDF and PPT, which are embeddable in applications and are designed to be printed or shared.

Note: There are three kinds of Power BI per-user licenses: Free, Pro, and Premium Per User. Power BI (free): Access to content in My Workspace

Power BI (free) + Workspace is Premium: Consume content shared with them

Power BI Pro: Publish content to other workspaces, share dashboards, subscribe to dashboards and reports, share with users who have a Pro license

Power BI Pro + Workspace is Premium: Distribute content to users who have free licenses

Power BI Premium Per User: Publish content to other workspaces, share dashboards, subscribe to dashboards and reports, share with users who have a Premium Per User license

Power BI Premium Per User + Workspace is Premium: Distribute content to users who have free and Pro licenses

Reference: <https://docs.microsoft.com/en-us/power-bi/fundamentals/service-features-license-type>

NEW QUESTION 32

- (Exam Topic 3)

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

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

An investigation into the issue identifies the following:

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

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

What should you recommend?

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

Answer: C

Explanation:

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

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

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

* Etc.

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

NEW QUESTION 37

- (Exam Topic 3)

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 40

- (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 44

- (Exam Topic 3)

You use an Apache Spark notebook in Azure Synapse Analytics to filter and transform data. You need to review statistics for a DataFrame that includes:

The column name
 The column type

The number of distinct values

Whether the column has missing values
 Which function should you use?

- A. displayHTML()
- B. display(df, summary=true)
- C. %%configure
- D. display(df)
- E. %%lsmagic

Answer: B

Explanation:

display(df) statistic details

You can use display(df, summary = true) to check the statistics summary of a given Apache Spark DataFrame that include the column name, column type, unique values, and missing values for each column. You can also select on specific column to see its minimum value, maximum value, mean value and standard deviation.

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

NEW QUESTION 45

- (Exam Topic 3)

You have an Azure Synapse Analytics serverless SQL pool.

You need to return a list of files and the number of rows in each file.

How should you complete the Transact-SQL statement? To answer, drag the appropriate values to the targets. Each value may be used once, more than once, or not at all. You may need to drag the split bar between panes or scroll to view content.

NOTE: Each correct selection is worth one point.

Values	Answer Area
APPROX_COUNT_DISTINCT	<pre> SELECT asa.filename() AS [filename] , [] (*) AS [rows] FROM [] BULK 'parquet/production/year=2017/month=9/*.parquet', DATA_SOURCE = 'DataLake1', FORMAT= 'PARQUET') asa GROUP BY [filename]</pre>
COUNT_BIG	
OPENDATASOURCE	
OPENJSON	
OPENQUERY	
OPENROWSET	

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: APPROX_COUNT_DISTINCT

The APPROX_COUNT_DISTINCT function returns the approximate number of unique non-null values in a group.

Box 2: OPENROWSET

OPENROWSET function in Synapse SQL reads the content of the file(s) from a data source. The data source is an Azure storage account and it can be explicitly referenced in the OPENROWSET function or can be dynamically inferred from URL of the files that you want to read. The OPENROWSET function can optionally contain a DATA_SOURCE parameter to specify the data source that contains files.

The OPENROWSET function can be referenced in the FROM clause of a query as if it were a table name OPENROWSET. It supports bulk operations through a built-in BULK provider that enables data from a file to be read and returned as a rowset.

Reference: <https://docs.microsoft.com/en-us/sql/t-sql/functions/approx-count-distinct-transact-sql> <https://docs.microsoft.com/en-us/azure/synapse-analytics/sql/develop-openrowset>

NEW QUESTION 50

- (Exam Topic 3)

You have a Power 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 53

- (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] < TOOAYQ, "Past", "Future")

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

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

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

Answer Area

PYTD =

CALCULATE (

[Sales Amount],

CALCULATETABLE
CROSSJOIN
SUMMARIZECOLUMNS
UNION

)

DATEADD
FIRSTDATE
PARALLELPERIOD
SAMEPERIODLASTYEAR

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

'Ca

"Future"
NOW()
"Past"
TODAY()

'Calendar'[pfFlag] =

)

)

- 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 56

- (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>

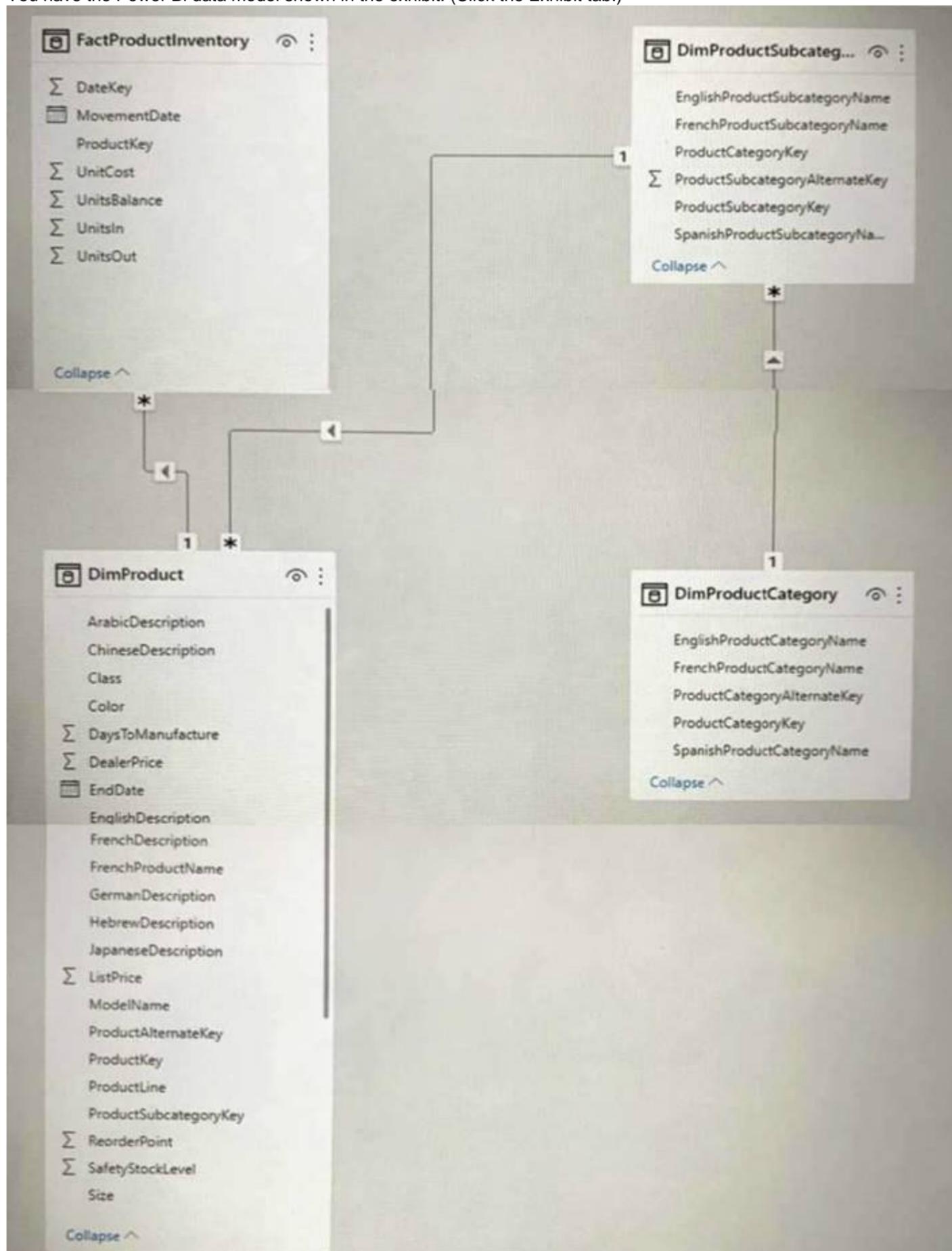
NEW QUESTION 61

- (Exam Topic 3)

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

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

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



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

of the reports.

Solution: You recommend moving all the measures to a calculation group. Does this meet the goal?

- A. Yes
- B. No

Answer: B

Explanation:

Instead denormalize For Performance.

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

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

NEW QUESTION 66

- (Exam Topic 3)

You have a kiosk that displays a Power BI report page. The report uses a dataset that uses Import storage mode. You need to ensure that the report page updates all the visuals every 30 minutes. Which two actions should you perform? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point.

- A. Enable Power BI embedded.
- B. Configure the data sources to use DirectQuery.
- C. Configure the data sources to use a streaming dataset
- D. Select Auto page refresh.
- E. Enable the XMIA endpoint.
- F. Add a Microsoft Power Automate visual to the report page.

Answer: BD

Explanation:

Automatic page refresh in Power BI enables your active report page to query for new data, at a predefined cadence, for DirectQuery sources.

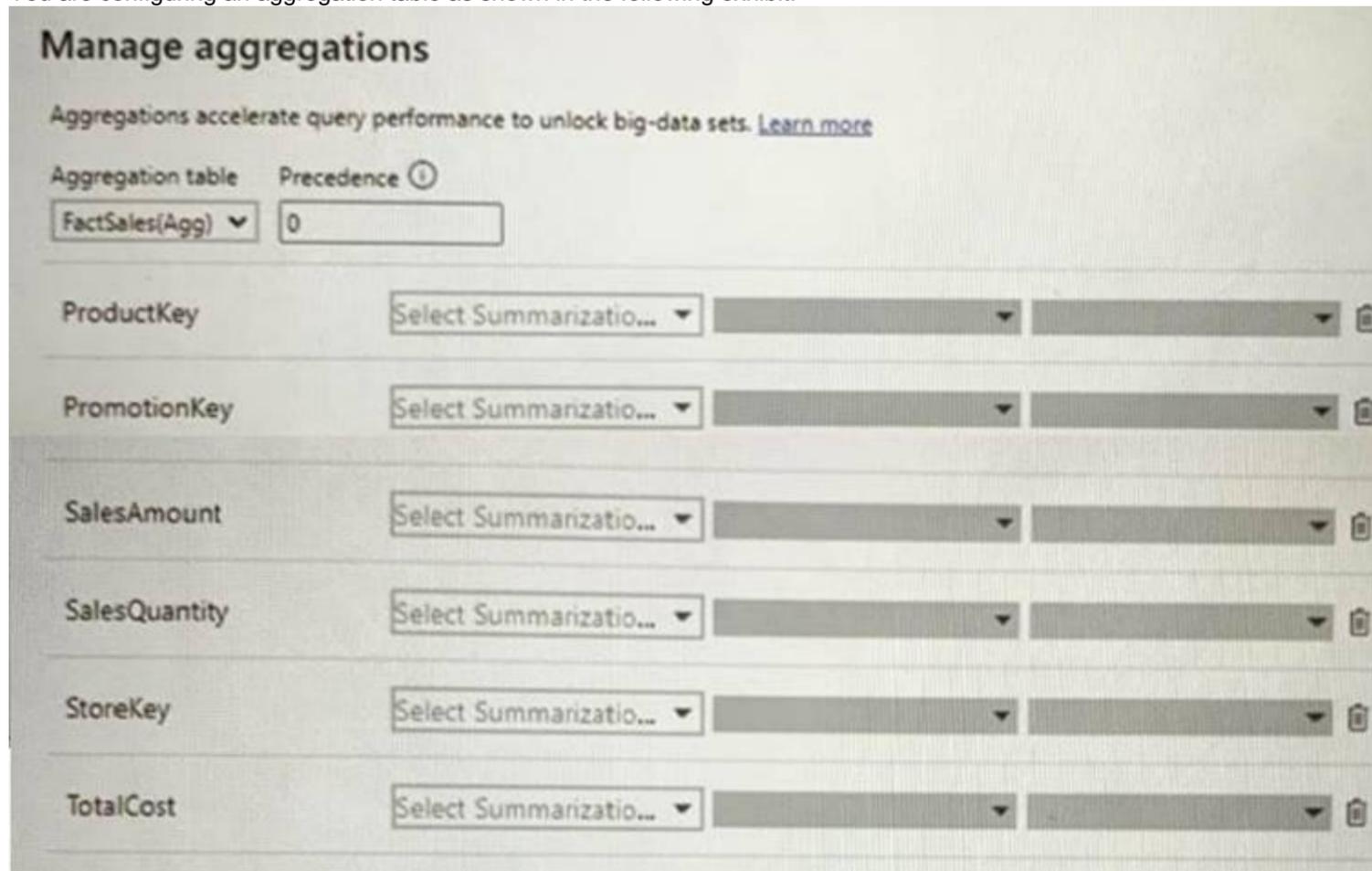
Automatic page refresh is available for DirectQuery sources and some LiveConnect scenarios, so it will only be available when you are connected to a supported data source. This restriction applies to both automatic page refresh types.

Reference: <https://docs.microsoft.com/en-us/power-bi/create-reports/desktop-automatic-page-refresh>

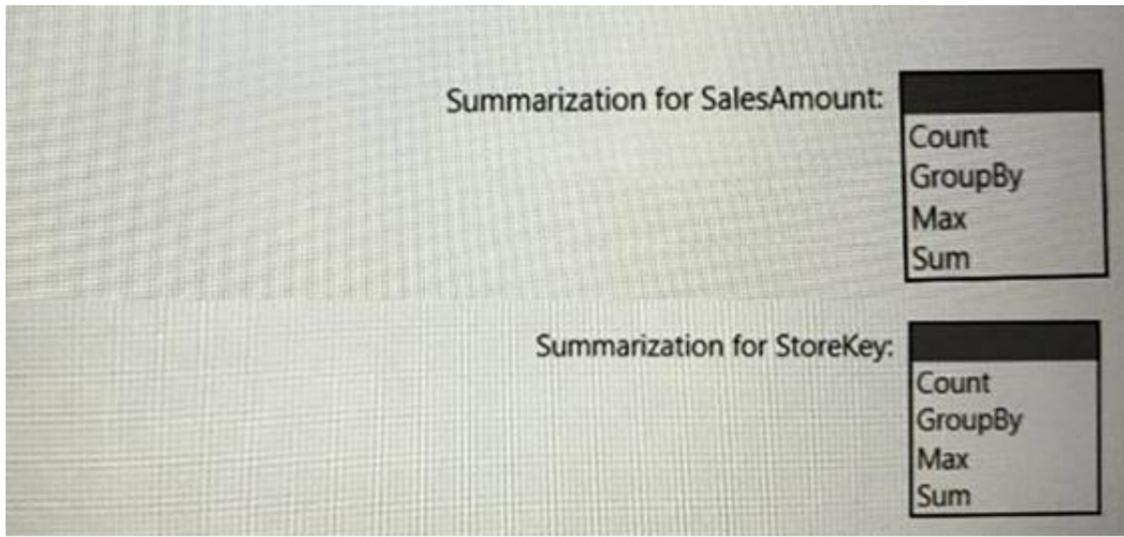
NEW QUESTION 69

- (Exam Topic 3)

You are configuring an aggregation table as shown in the following exhibit.



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



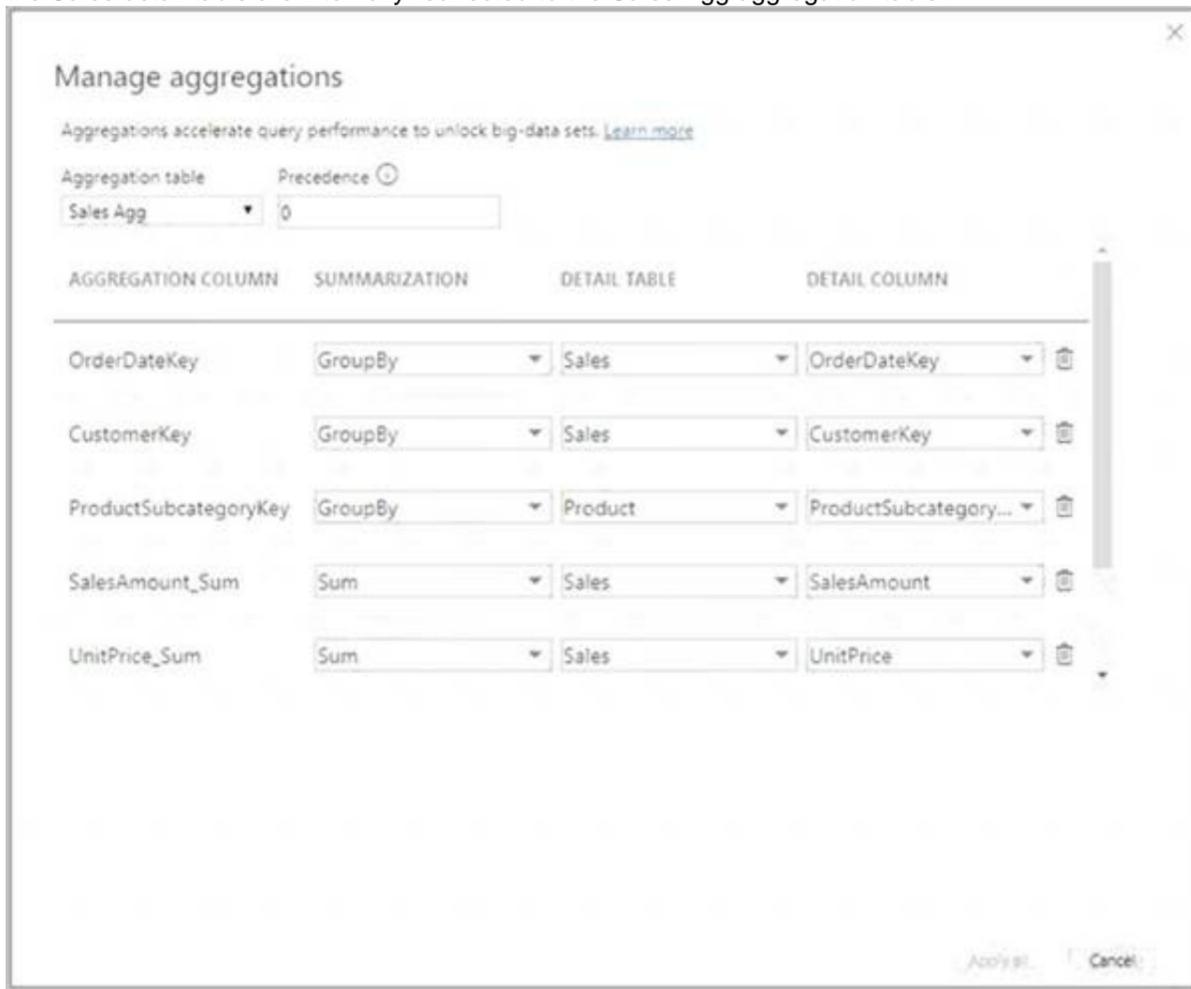
- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: Sum

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



Box 2: GroupBy

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

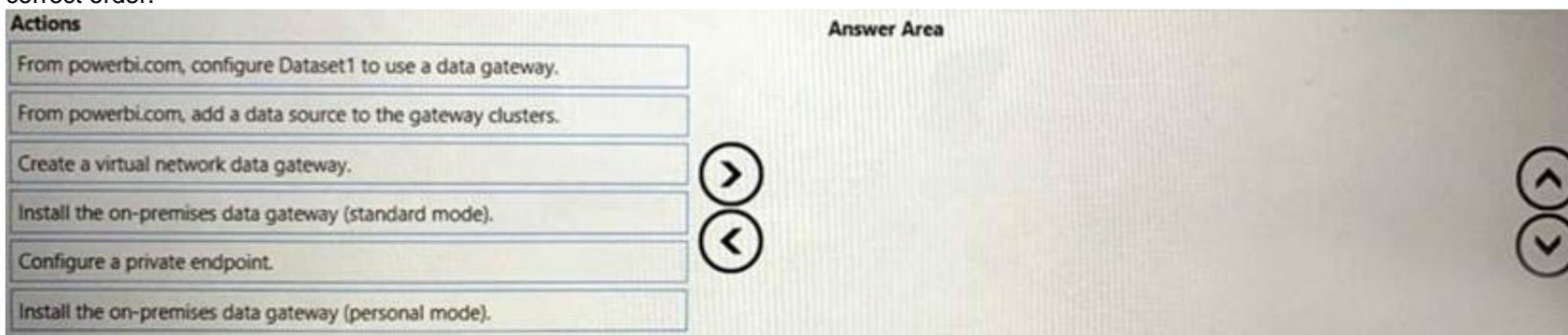
NEW QUESTION 74

- (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.

New data source

Gateway cluster name *

Data source name *

Data source type *

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.

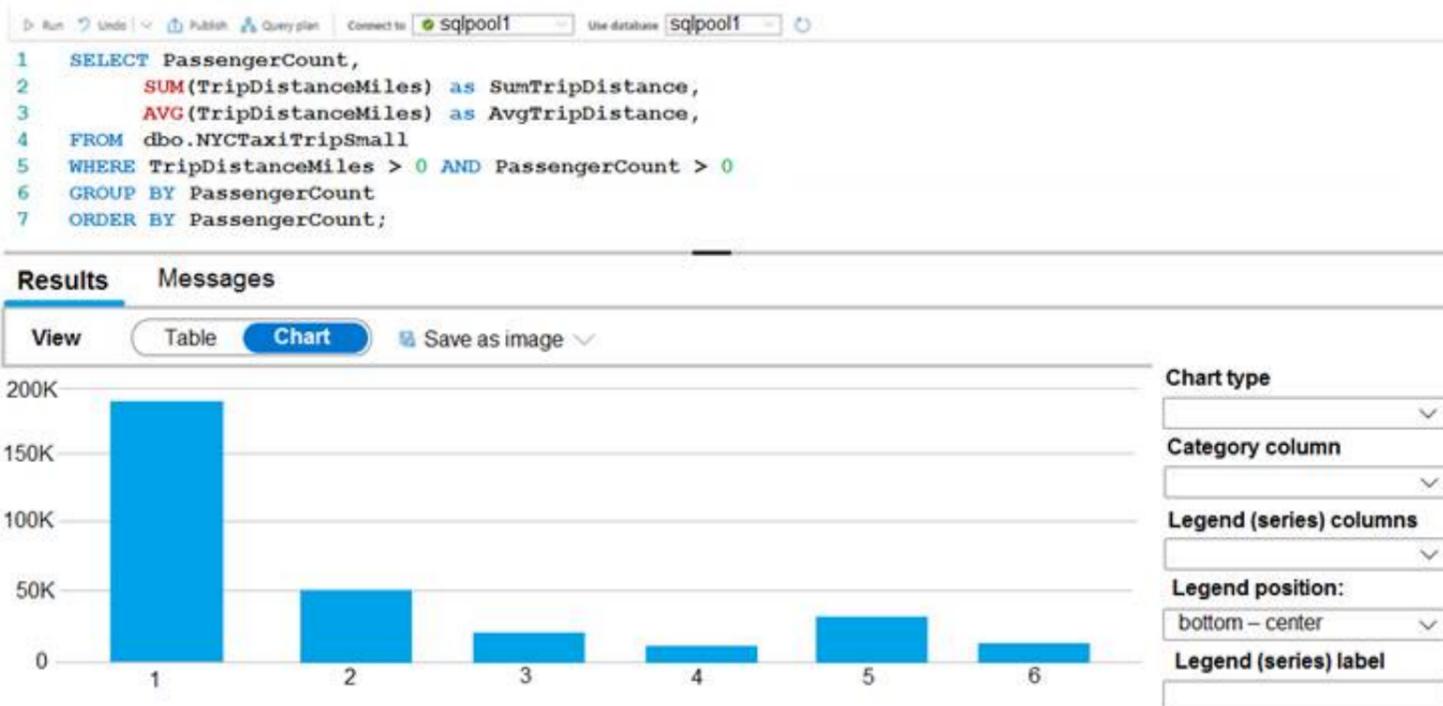
Reference: <https://docs.microsoft.com/en-us/power-bi/connect-data/service-gateway-personal-mode> <https://docs.microsoft.com/en-us/power-bi/connect-data/service-gateway-sql-tutorial> <https://docs.microsoft.com/en-us/power-bi/connect-data/service-gateway-enterprise-manage-sql>

NEW QUESTION 76

- (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:

	▼
AvgTripDistance	
PassengerCount	
SumTripDistance	
TripDistanceMiles	

Legend (series) column:

	▼
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 80

- (Exam Topic 3)

You have an Azure Synapse Analytics dataset that contains data about jet engine performance. You need to score the dataset to identify the likelihood of an engine failure. Which function should you use in the query?

- A. PIVOT
- B. GROUPING
- C. PREDICT
- D. CAST

Answer: A

NEW QUESTION 84

- (Exam Topic 3)

You have new security and governance protocols for Power BI reports and datasets. The new protocols must meet the following requirements.

- New reports can be embedded only in locations that require authentication.
- Live connections are permitted only for workspaces that use Premium capacity datasets.

Which three actions should you recommend performing in the Power BI Admin portal? Each correct answer presents part of the solution. NOTE: Each correct selection is worth one point.

- A. From Tenant settings, disable Allow XMLA endpoints and Analyze in Excel with on-premises datasets.
- B. From the Premium per user settings, set XMLA Endpoint to Off.
- C. From Embed Codes, delete all the codes.
- D. From Capacity settings, set XMLA Endpoint to Read Write.
- E. From Tenant settings, set Publish to web to Disable.

Answer: ADE

Explanation:

Reference: <https://docs.microsoft.com/en-us/power-bi/enterprise/service-premium-connect-tools> <https://powerbi.microsoft.com/en-us/blog/power-bi-february-service-update>

NEW QUESTION 89

- (Exam Topic 3)

You are planning a Power BI solution for a customer.

The customer will have 200 Power BI users. The customer identifies the following requirements:

- Ensure that all the users can create paginated reports.
- Ensure that the users can create reports containing AI visuals.
- Provide autoscaling of the CPU resources during heavy usage spikes.

You need to recommend a Power BI solution for the customer. The solution must minimize costs. What should you recommend?

- A. Power BI Premium per user
- B. a Power BI Premium per capacity
- C. Power BI Pro per user
- D. Power BI Report Server

Answer: A

Explanation:

Announcing Power BI Premium Per User general availability and autoscale preview for Gen2. Power BI Premium per user features and capabilities

* Pixel perfect paginated reports are available for operational reporting capabilities based on SSRS technology. Users can create highly formatted reports in various formats such as PDF and PPT, which are embeddable in applications and are designed to be printed or shared.

* Automated machine learning (AutoML) in Power BI enables business users to build ML models to predict outcomes without having to write any code.

* Etc. Note:

Power BI empowers every business user and business analyst to get amazing insights with AI infused experiences. With Power BI Premium, we enable business analysts to not only analyze and visualize their data, but to also build an end-to-end data platform through drag and drop experiences. Everything from ingesting and transforming data at scale, to building automated machine learning models, and analyzing massive volumes of data is now possible for our millions of business analysts.

Reference:

<https://powerbi.microsoft.com/nl-be/blog/announcing-power-bi-premium-per-user-general-availability-and-auto>

NEW QUESTION 90

- (Exam Topic 3)

You have the following code in an Azure Synapse notebook.

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

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

NOTE: Each correct selection is worth one point.

Answer Area

There will be [answer choice] rendered as the output of the code.

- one scatterplot
- two scatterplots
- three scatterplots

There will be [answer choice] used in the output.

- one marker symbol
- two marker symbols
- three marker symbols

- A. Mastered
- B. Not Mastered

Answer: A

Explanation:

Box 1: three scatterplots Compare Plots

Example, Draw two plots on the same figure: import matplotlib.pyplot as plt

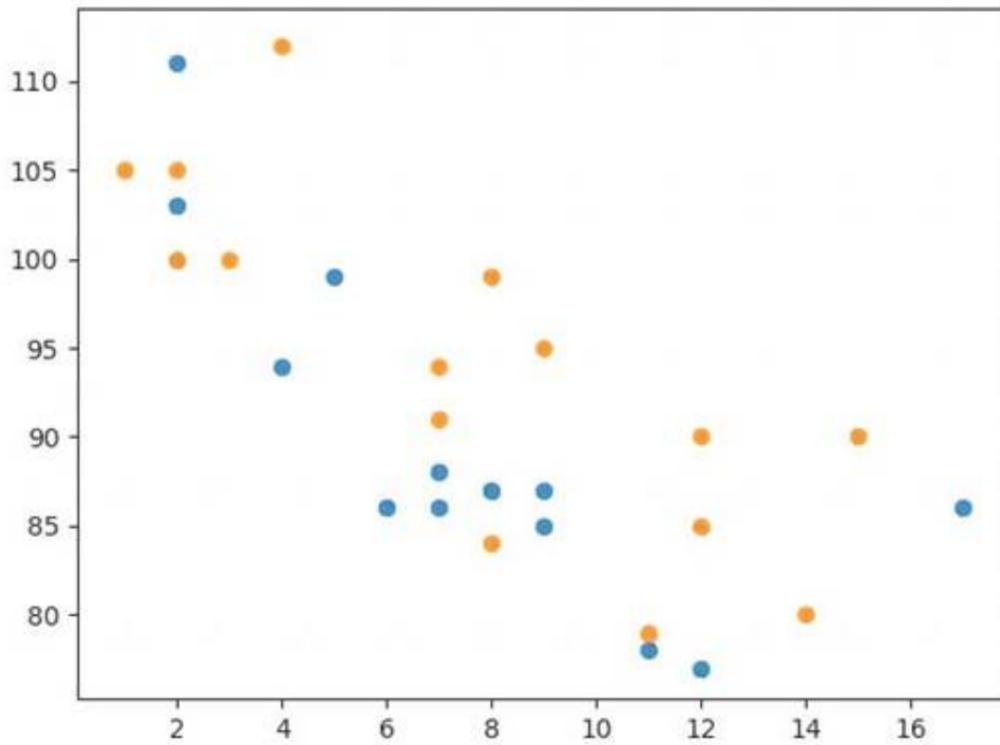
import numpy as np

#day one, the age and speed of 13 cars:

x = np.array([5,7,8,7,2,17,2,9,4,11,12,9,6])

y = np.array([99,86,87,88,111,86,103,87,94,78,77,85,86])

```
plt.scatter(x, y)
#day two, the age and speed of 15 cars:
x = np.array([2,2,8,1,15,8,12,9,7,3,11,4,7,14,12])
y = np.array([100,105,84,105,90,99,90,95,94,100,79,112,91,80,85])
plt.scatter(x, y) plt.show() Result:
Chart, scatter chart Description automatically generated
```



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

NEW QUESTION 91

- (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 96

- (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>

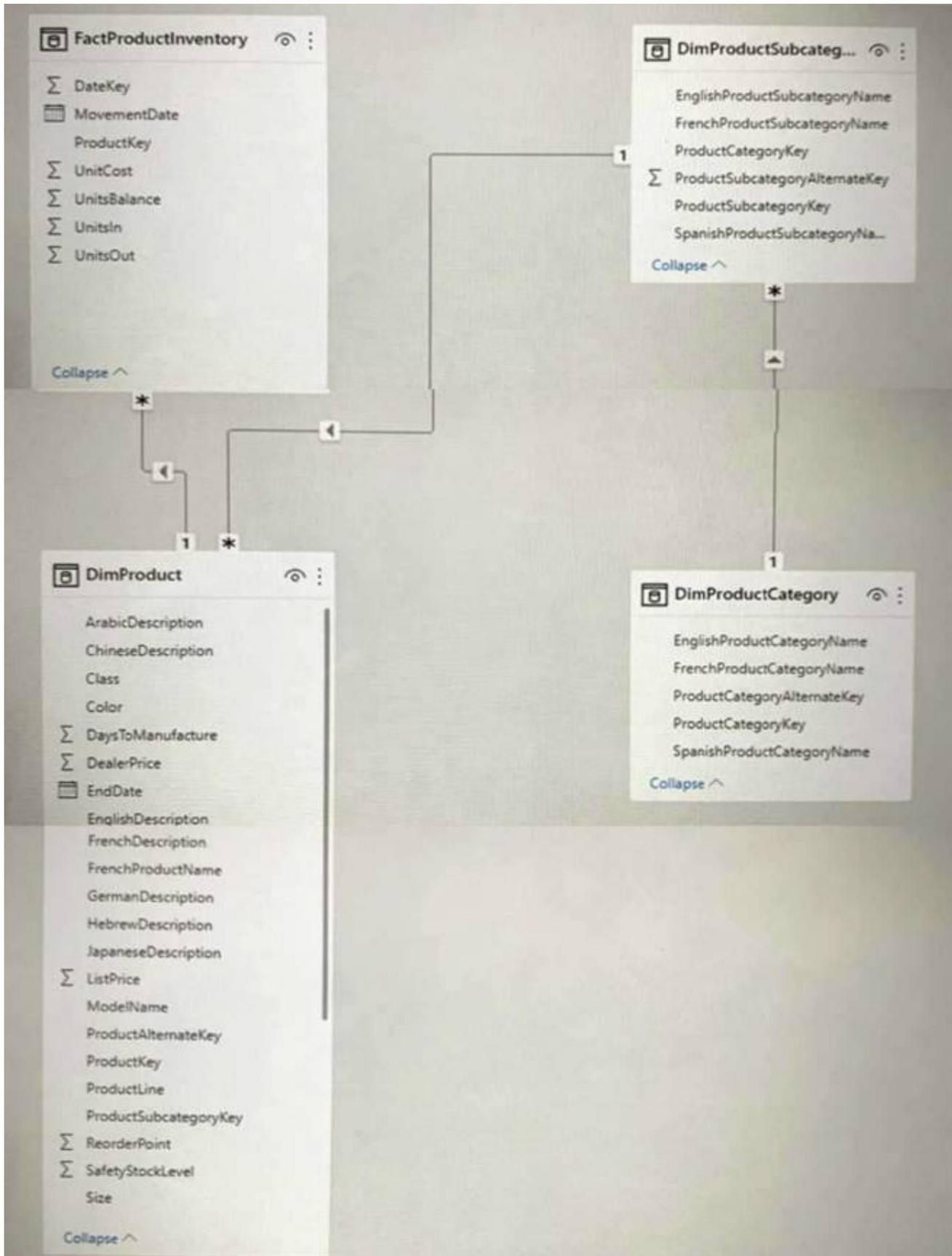
NEW QUESTION 97

- (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/>

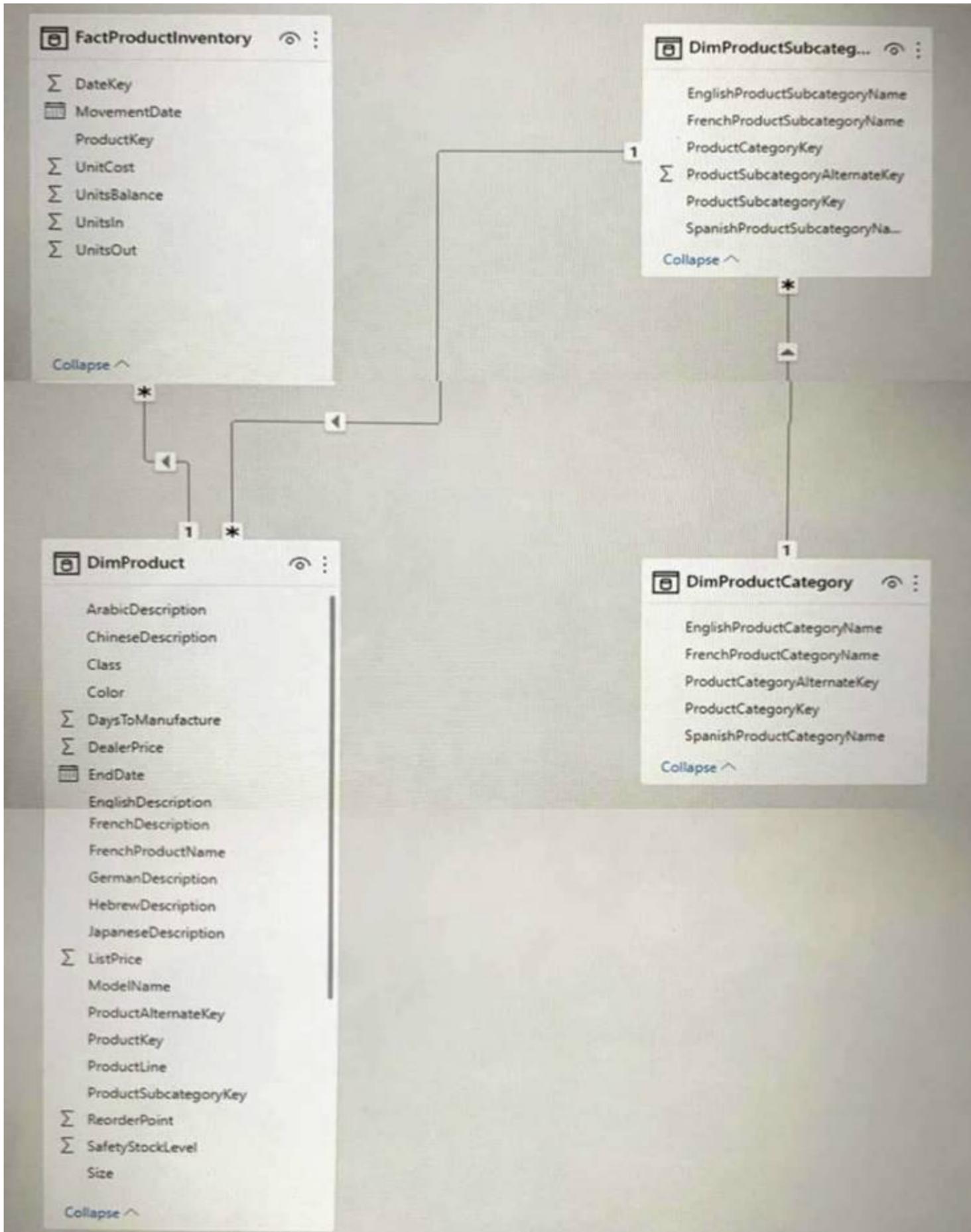
NEW QUESTION 99

- (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 102

- (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://maqssoftware.com/insights/power-bi-best-practices>

NEW QUESTION 105

- (Exam Topic 3)

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

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

An investigation into the issue identifies the following:

One of the datasets is larger than 1 GB and has a fact table that contains more than 500 million rows.

When publishing dataset changes to development, test, or production pipelines, a refresh is triggered against the entire dataset.

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

What should you recommend?

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

Answer: B

Explanation:

Hybrid tables

Hybrid tables are tables with incremental refresh that can have both import and direct query partitions. During a clean deployment, both the refresh policy and the hybrid table partitions are copied. When deploying to a pipeline stage that already has hybrid table partitions, only the refresh policy is copied. To update the partitions, refresh the table.

Refreshes are faster - Only the most recent data that has changed needs to be refreshed.

Reference: <https://docs.microsoft.com/en-us/power-bi/create-reports/deployment-pipelines-best-practices>

NEW QUESTION 106

- (Exam Topic 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 111

- (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.


```

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;

```

- 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 120

.....

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](#)