**Question 1:**

This is a case study. The case study presents a scenario with an overview, data sources, issues, and requirements. You can read all the information but you can also read the questions first and then have a look at the requirements to find the right solution.

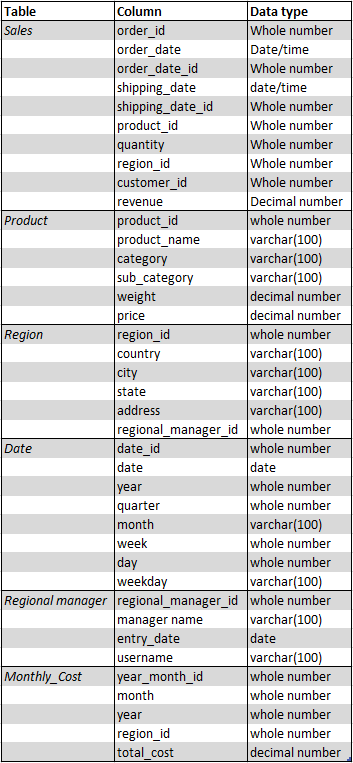
There will be multiple questions for each case study

**Overview:**

You are working as data analyst for Dataline Ltd. This is a company that produces different types of products. In the company are multiple employees that prepare monthly finance reports. These reports are built in Microsoft Excel and include profit analysis for each category, revenue forecasts for the next months and more.

**Data sources:**

The data comes from different tables such as csv files and tables in a database.



The date\_id column in the Date table has the format yyyymmdd.

The year\_month\_id column in the Monthly\_Cost table has the format yyyymm.

**Issues:**

The company is concerned with null values and the Products table.

**Requirements:**

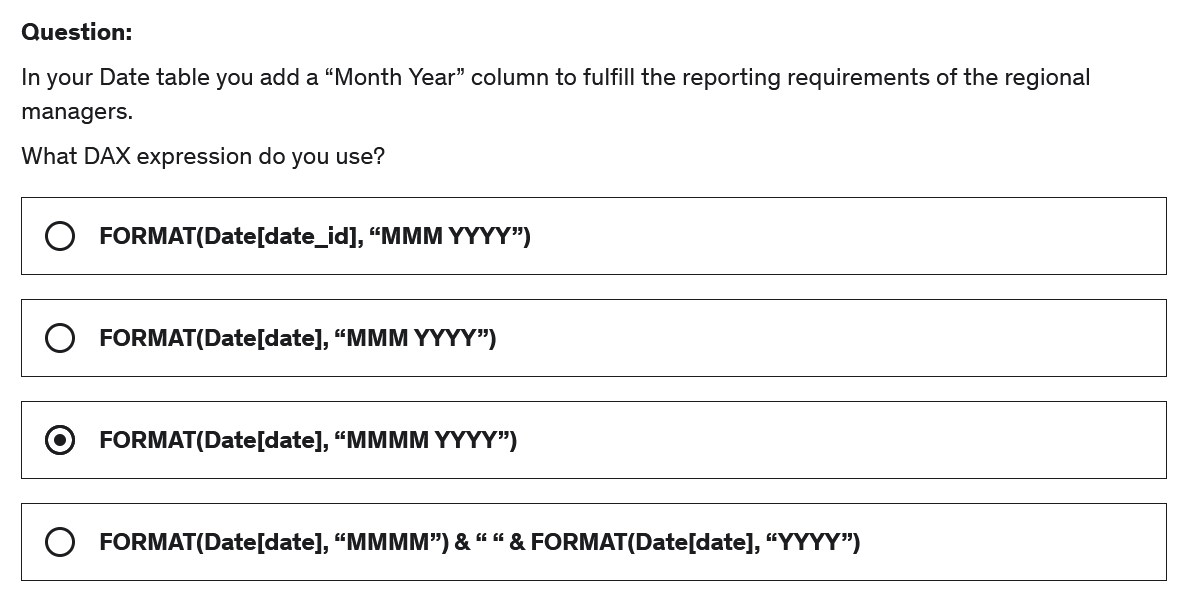
~ The top-level managers need to be able to see a chart with the total sales by region. They should be able to see all sales data of all 23 cities.

~Regional managers should only be allowed to see the data of their region.

~Regional managers need to see the number of sales in a table with the months in the following format “January 2022”.

~Financial managers need to see the monthly sales revenue and monthly costs.

~ Quality managers need to be able to filter the data based on the shipping\_date as well as the order\_date.

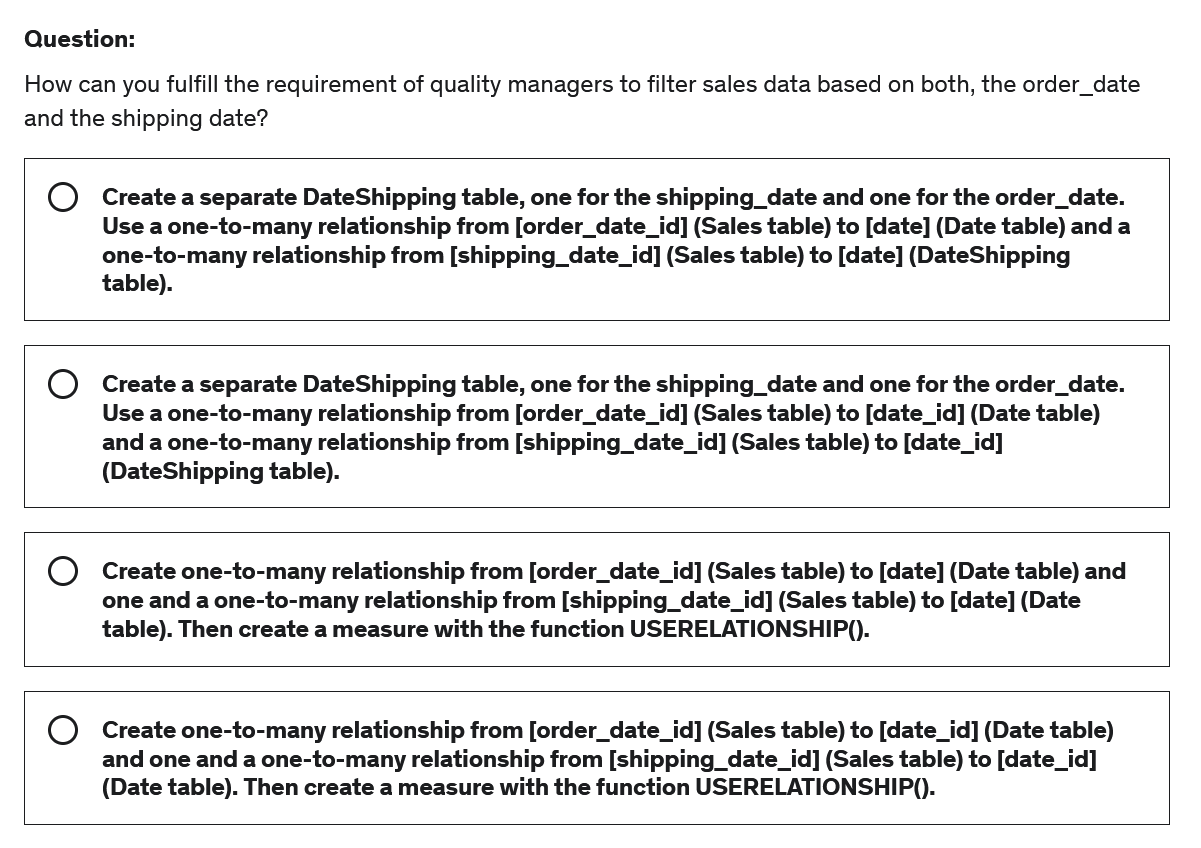


**Correct Answer: C**

**Explanation**

We can use "MMMM" to display the full month, "YYYY" to display the year and it is also possible to include other characters such as spaces.

Question 2



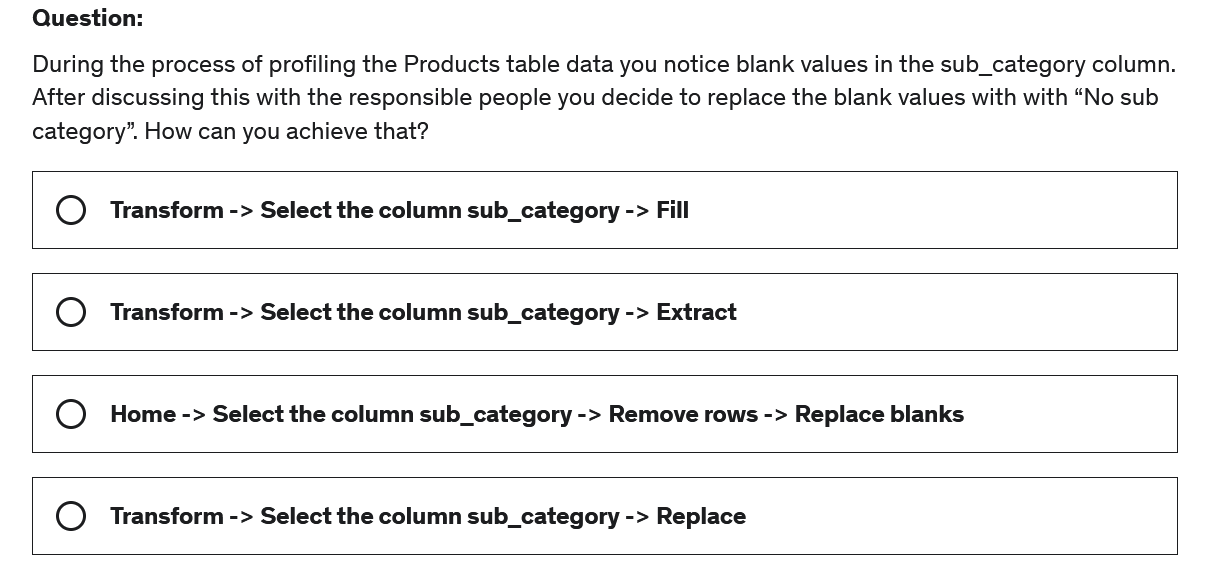
**Correct Answer: D**

**Explanation**

You should not create an additional date table since this creates redundant data but you can create a second relationship which will be inactive and then use the function USERELATIONSHIP() in a measure to activate the inactive relationship in that measure.

To connect the tables you need to use the date\_id column since this will be a matching format with the shipping\_date\_id and oder\_date\_id.

Question 3:

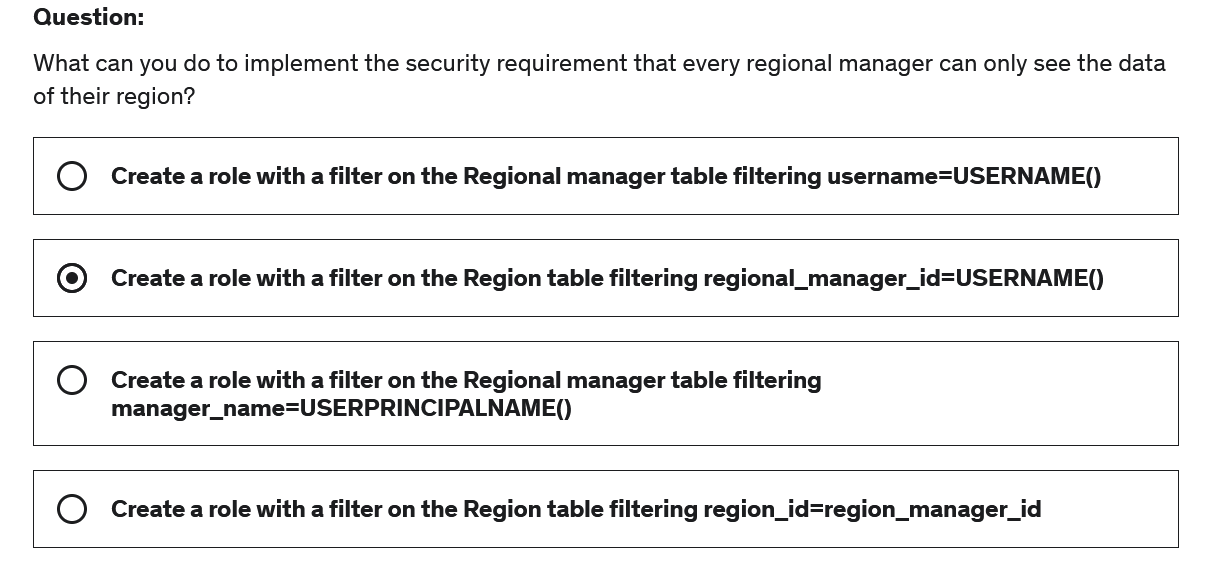


**Correct Answer: D**

**Explanation**

If you want to replace values. You need to use the Replace tool. This tool is found in the Transform ribbon. Before you select the tool, you need to select the column in which you want to replace values in.

Question 4

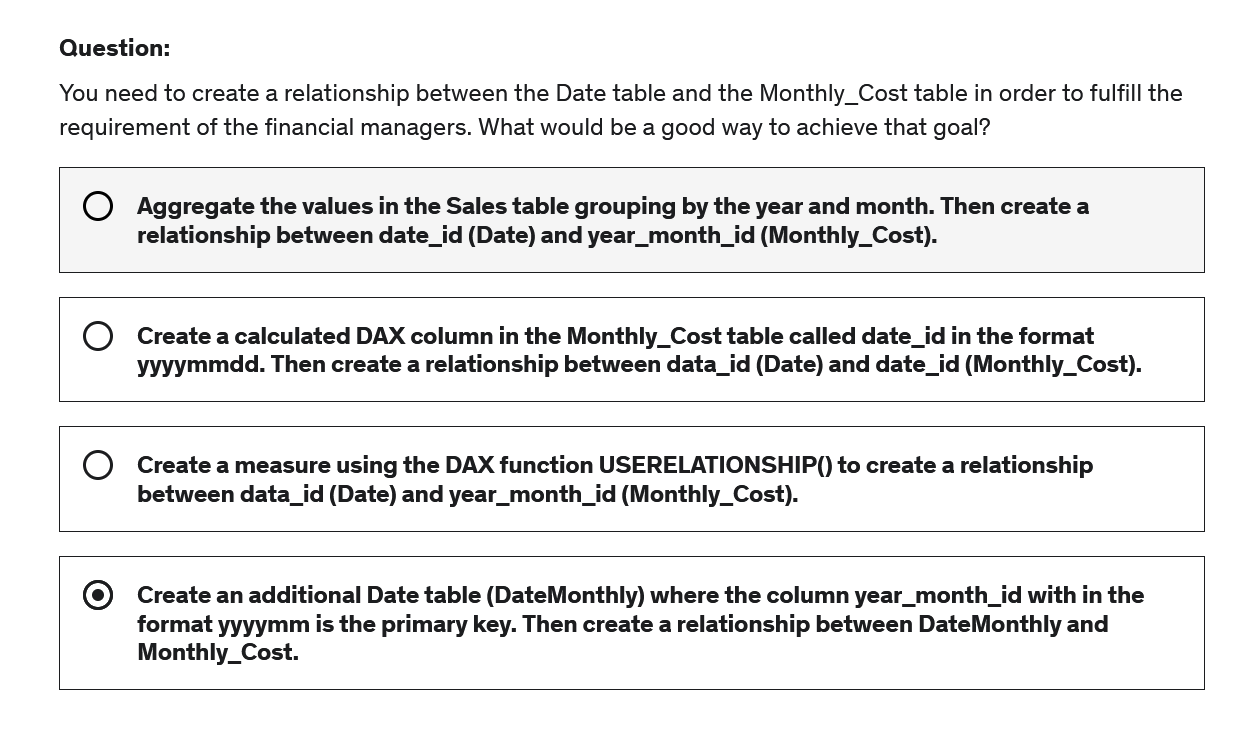


**Correct Answer: A**

**Explanation**

We need to have a username or an email address field if we want to use dynamic RLS, and then we can set this equal to USERNAME() or USERPRINCIPALNAME().

Question 5:



**Correct Answer: B**

**Explanation**

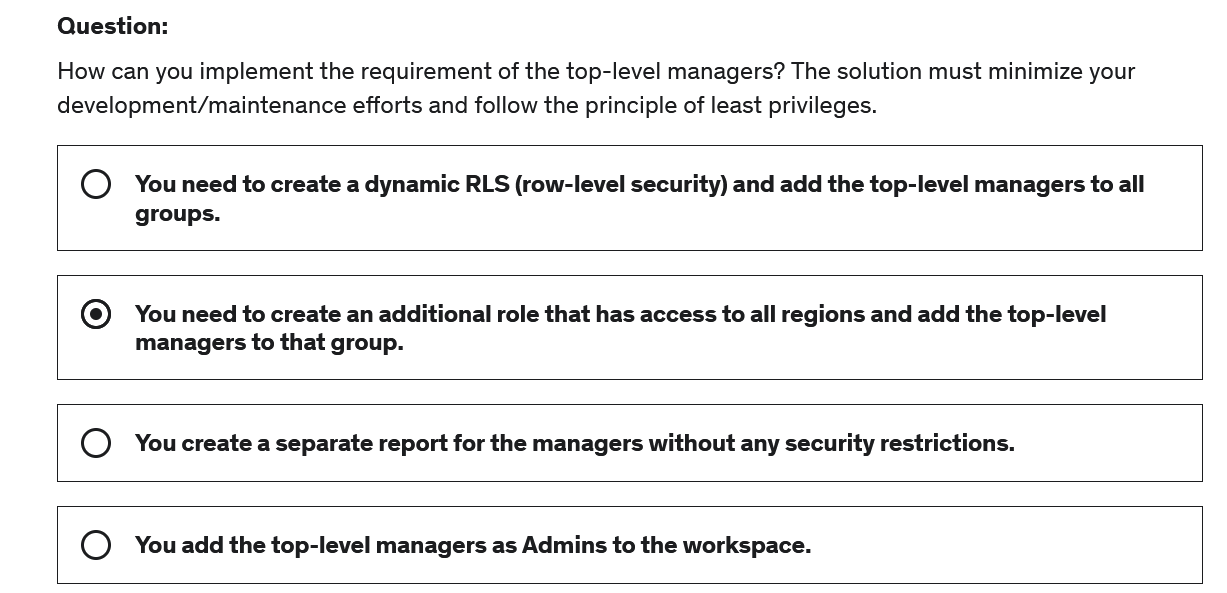
We can’t use USERELATIONSHIP() if there is not even a relationship between the tables. We would use that to activate an inactive relationship in a given measure.

Answer D creates redundant data and is not user-friendly since there are now two Date tables.

Aggregating the values would result in loosing access to important information.

The goal can be achieved best by adding a calculated column with a matching granularity level, so that we are able to create a relationship between the tables.

Question 6 :



**Correct Answer: B**

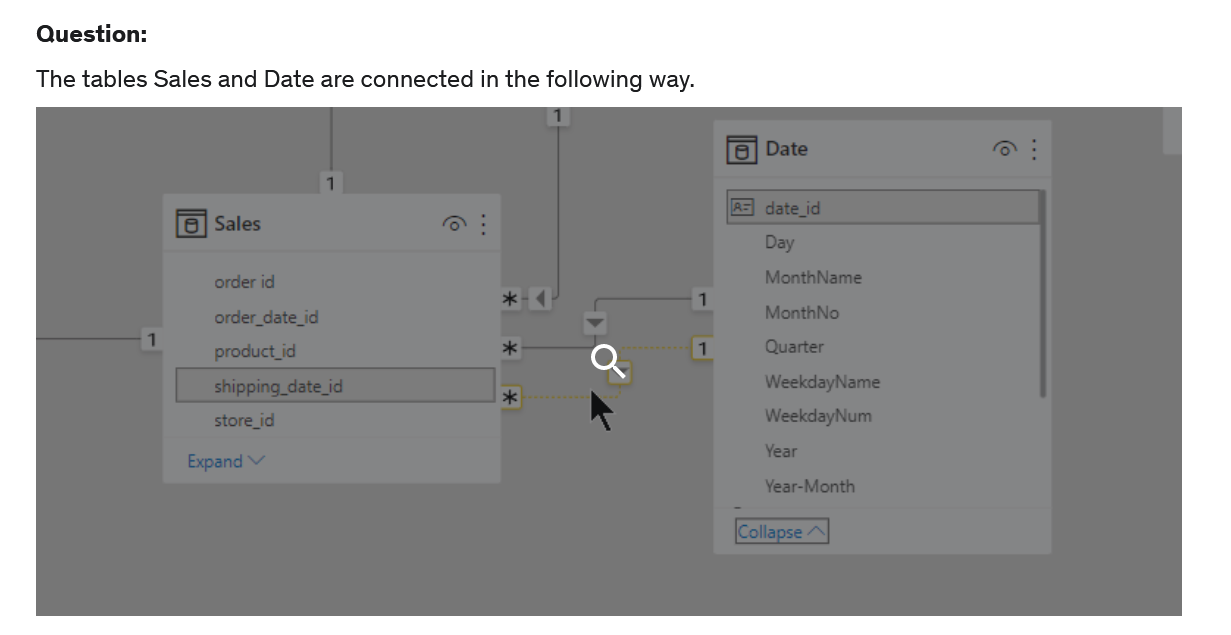
**Explanation**

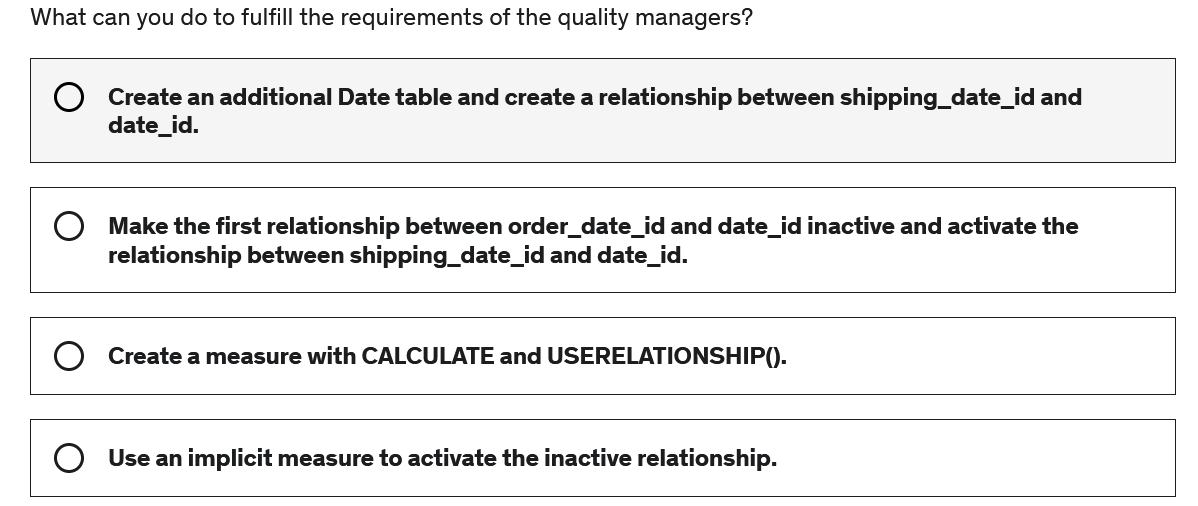
A separate report is additional effort that is not necessary.

Adding the top-level managers as admins to the workspace gives them more privileges than necessary and answer A does not make sense because in a dynamic RLS you usually have only one group and the security is defined in a separate security table.

The best way of doing that is to create a role that has the appropriate filtering and add the user to that group.

Question 7



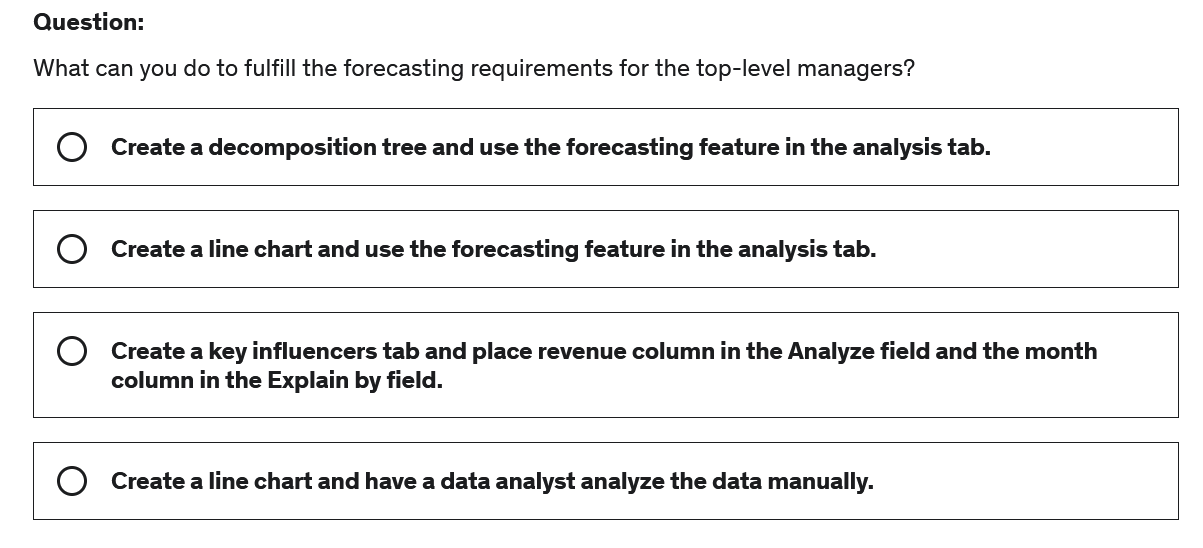


**Correct Answer: C**

**Explanation**

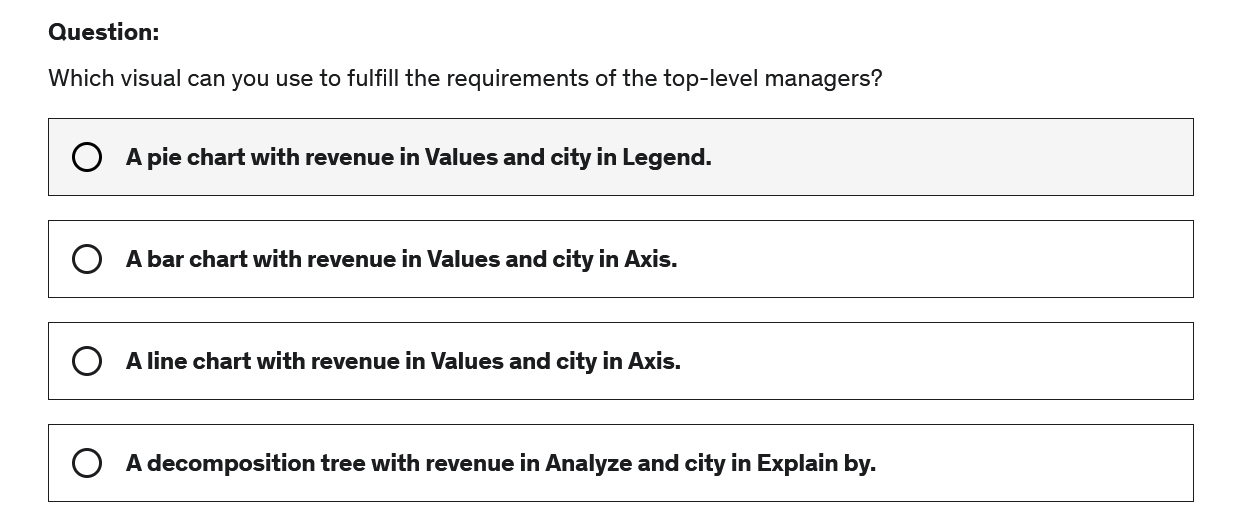
To use an inactive relationship in a measure we use CALCULATE and USERELATIONSHIP(<columnName1>,<columnName2>).

Question 8:



**Correct Answer: B**

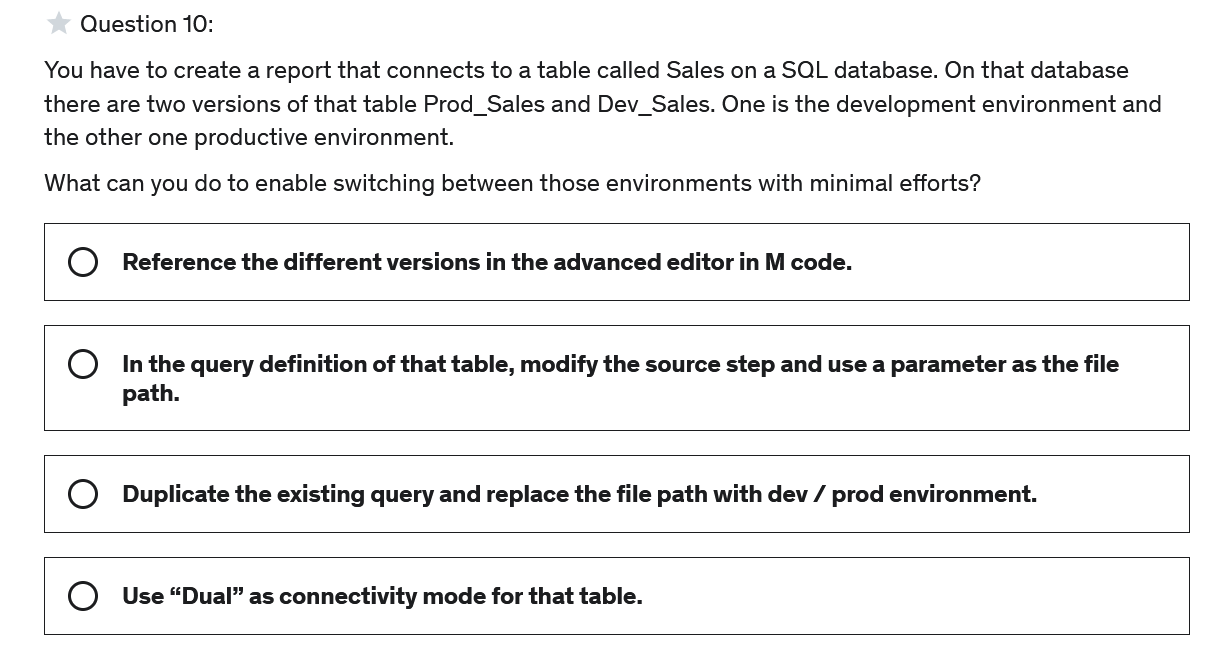
Question 9:



**Correct Answer: B**

**Explanation**

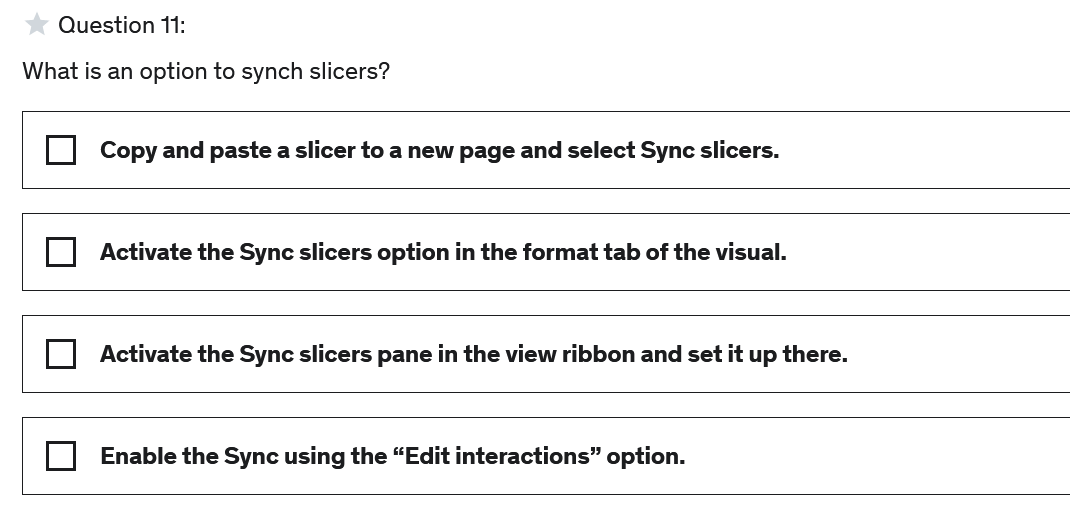
A bar chart is a good visual to compare values across multiple categories. Especially if there are many categories a pie chart is not very easy to read and compare values.



**Correct Answer: B**

**Explanation**

Using parameters in your data source is a good way of connecting to different data sources and switching between them.



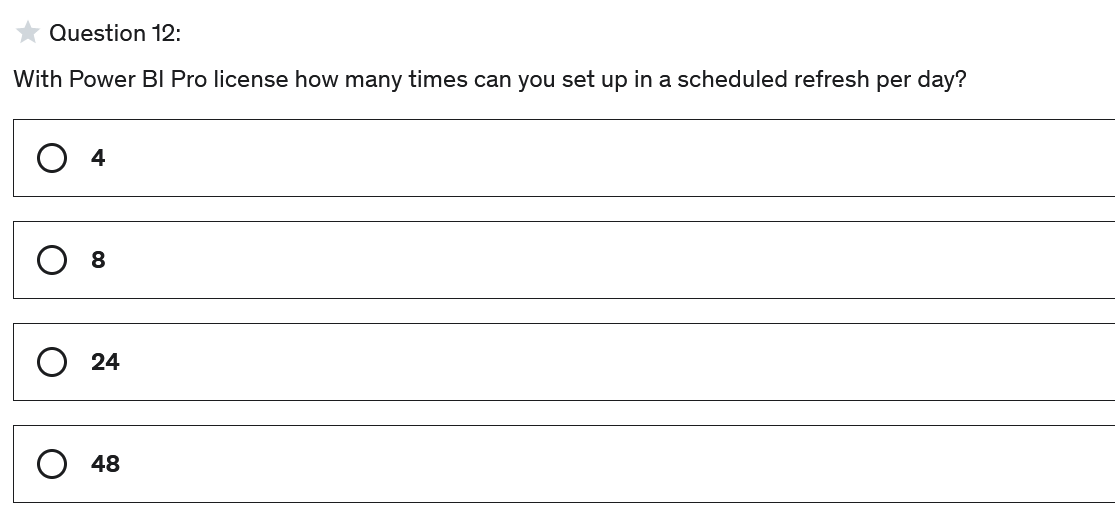
**Correct Answer: A, C**

**Explanation**

There are two methods of setting up sync slicers:

\*Activate the Sync slicers pane in the view ribbon and set it up there.

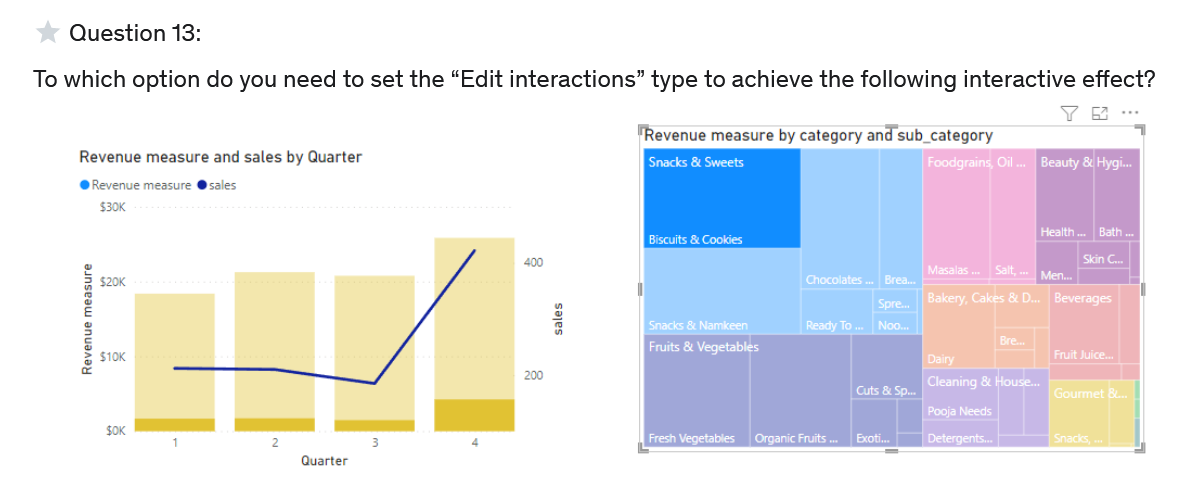
\*Copy and paste a slicer to a new page and select Sync slicers.

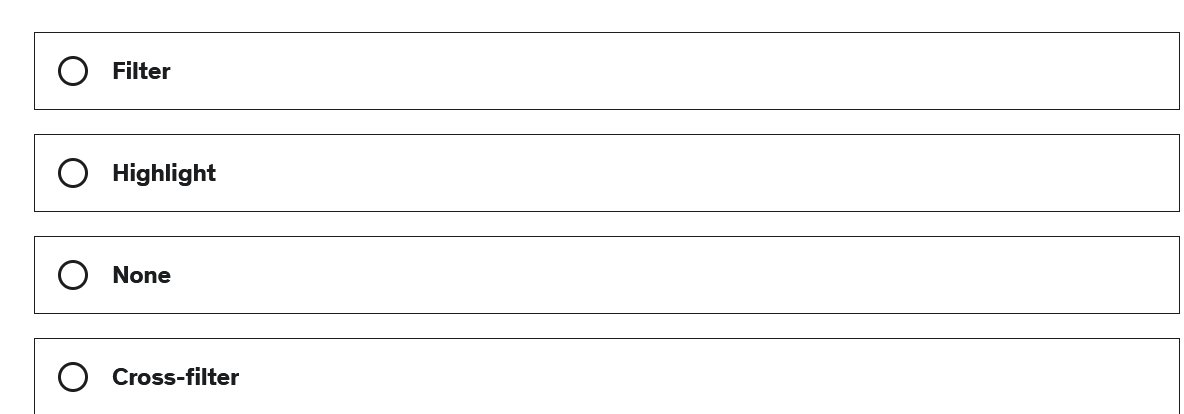


**Correct Answer: B**

**Explanation**

With Power BI Pro we can set up up to 8 refresh times per day and with Power BI Premium we can schedule up up to 48 refresh times per day.

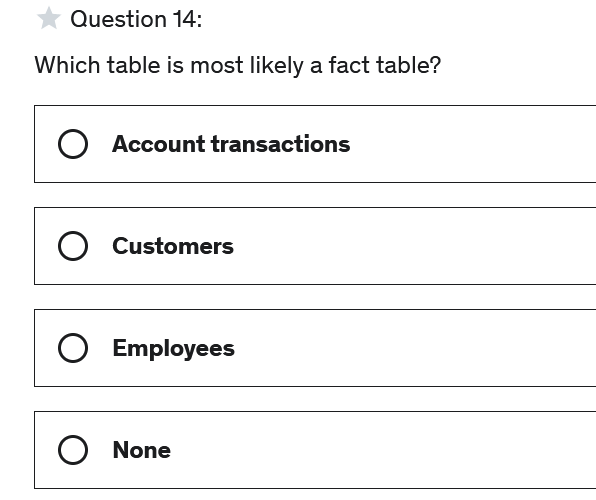




**Correct Answer: B**

**Explanation**

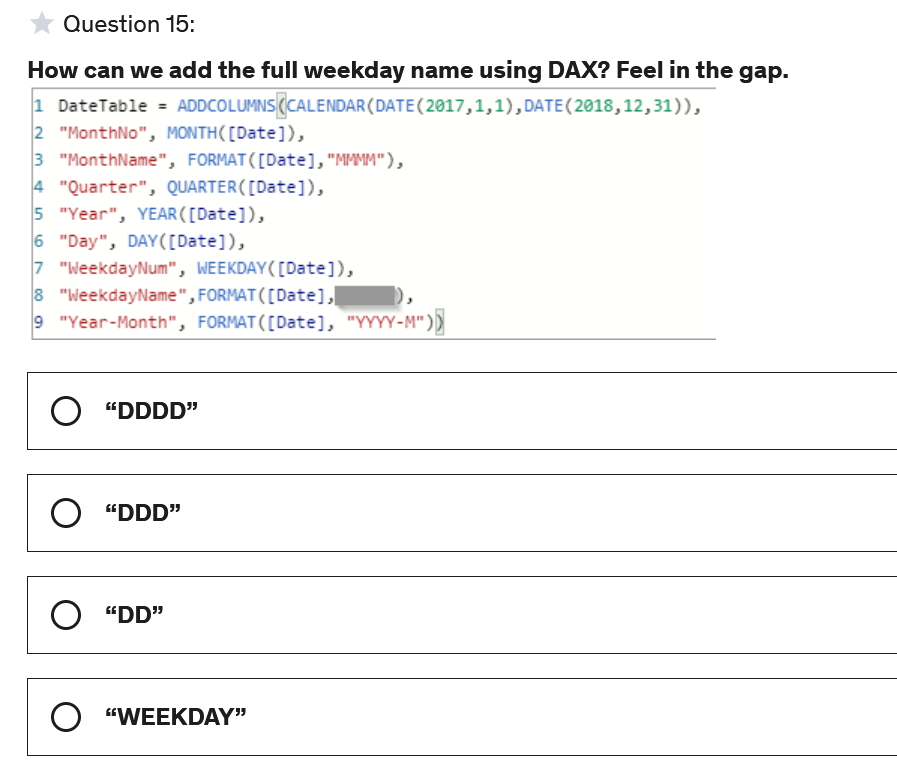
In the bar chart we can see that certain parts of the bars are highlighted in a darker yellow while one value (Biscuits & Cookies) is selected in the Treemaps visual.



**Correct Answer: A**

**Explanation**

A table that contains transactional data is usually a fact table.



**Correct Answer: A**

**Explanation**

"dddd" represents the full weekday name.

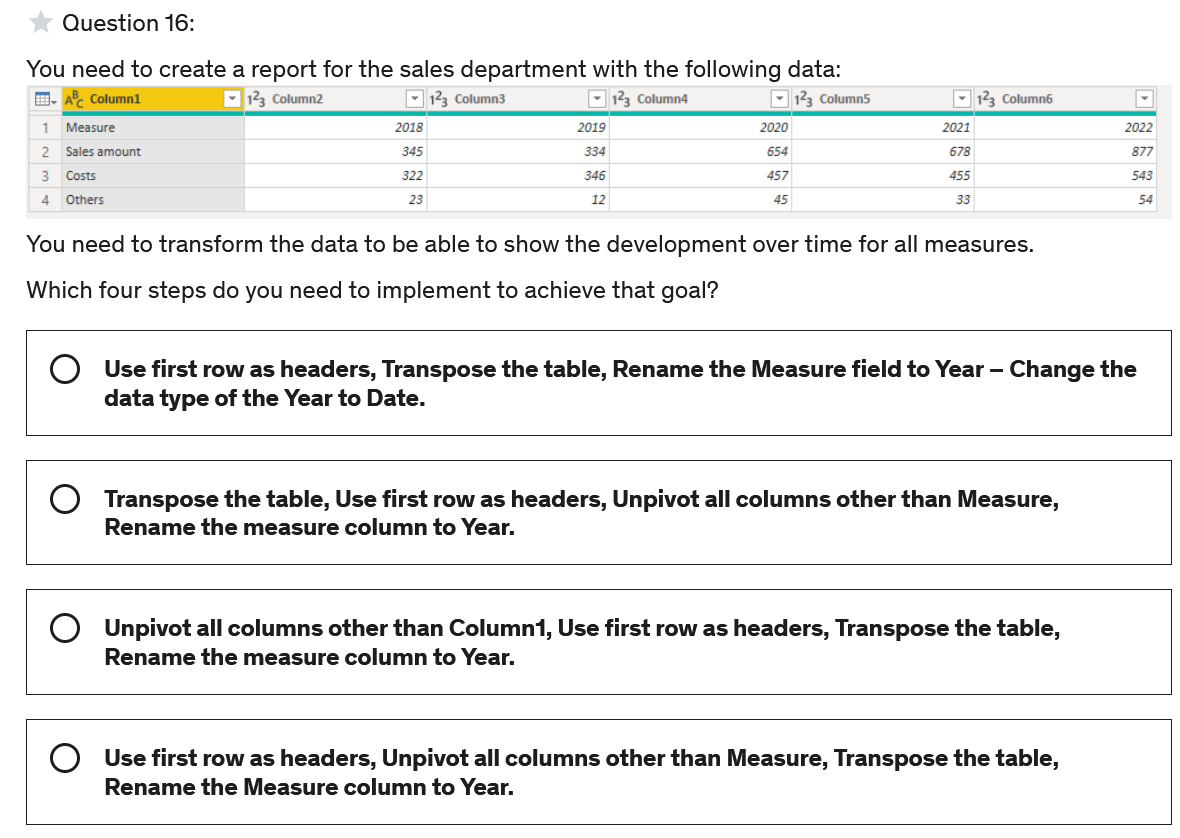
"ddd" represents the first three letters of the weekday name.

"dd" represents the day in a month (01-31)

"weekday" is not an option in the format function.

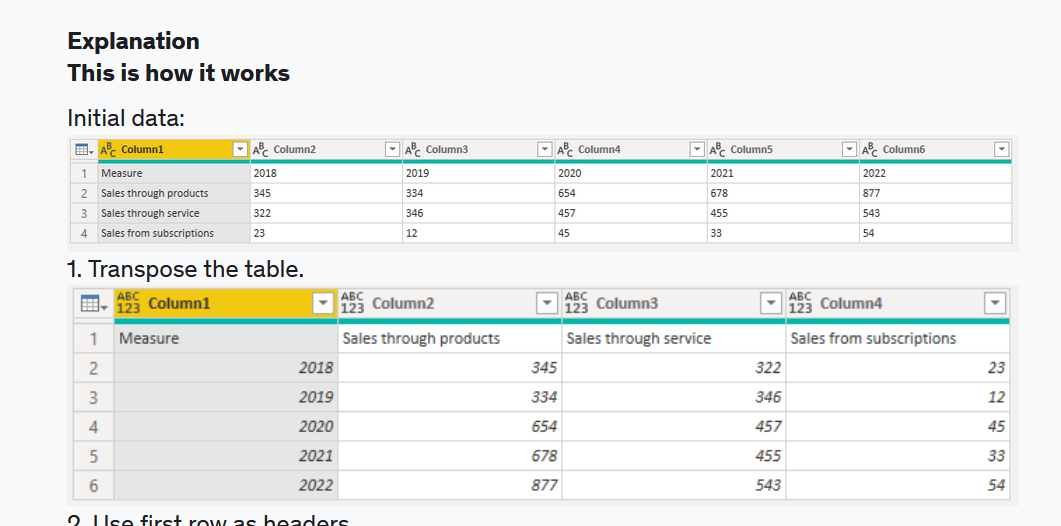
For more details see here:

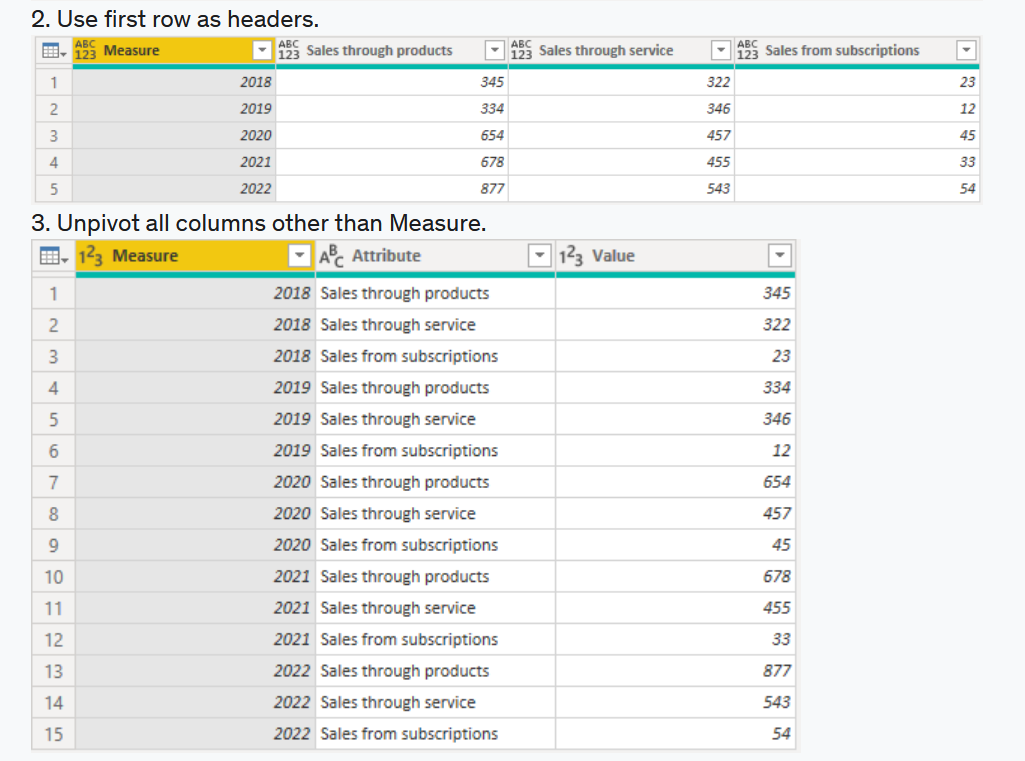
<https://docs.microsoft.com/en-us/dax/format-function-dax>

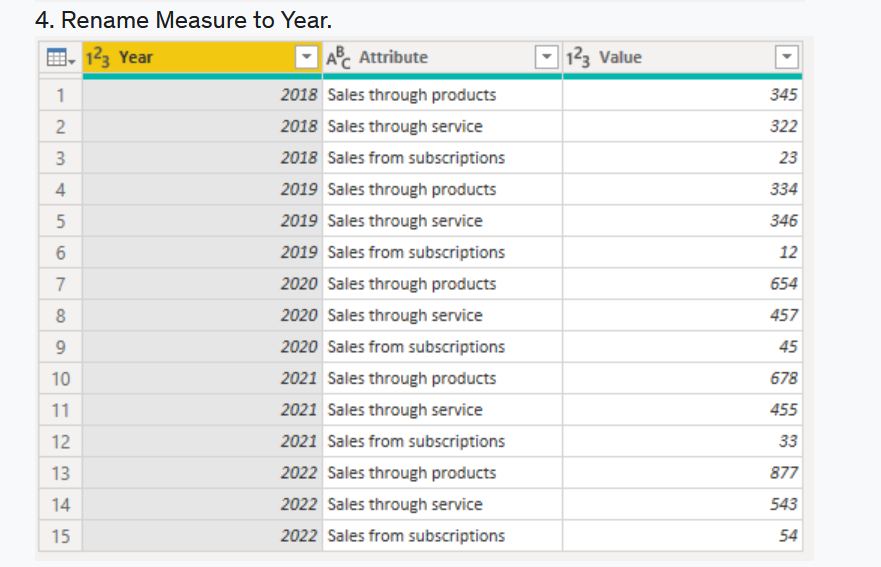


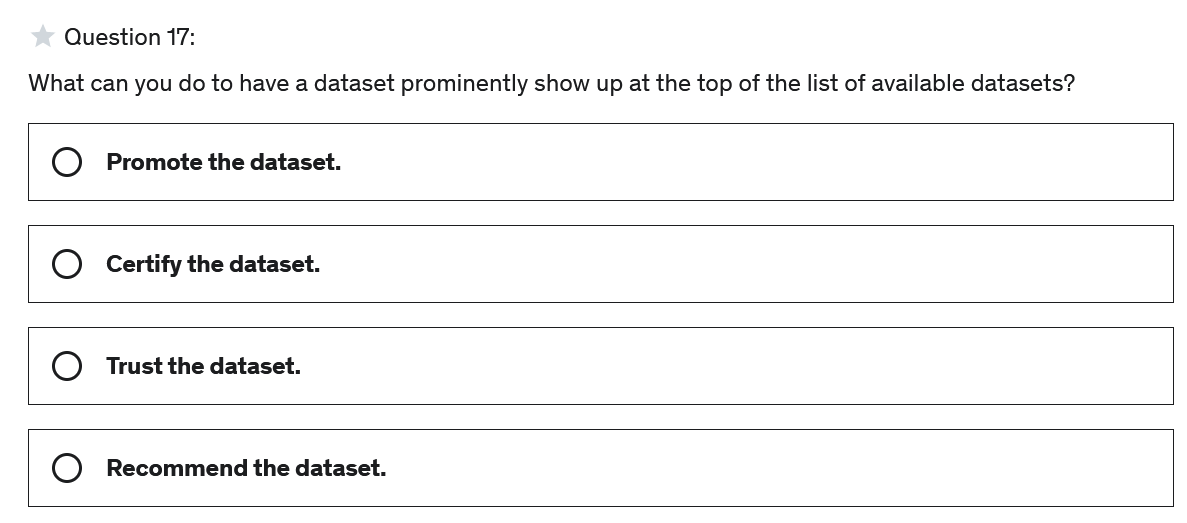
**Correct Answer: B**

**Explanation**

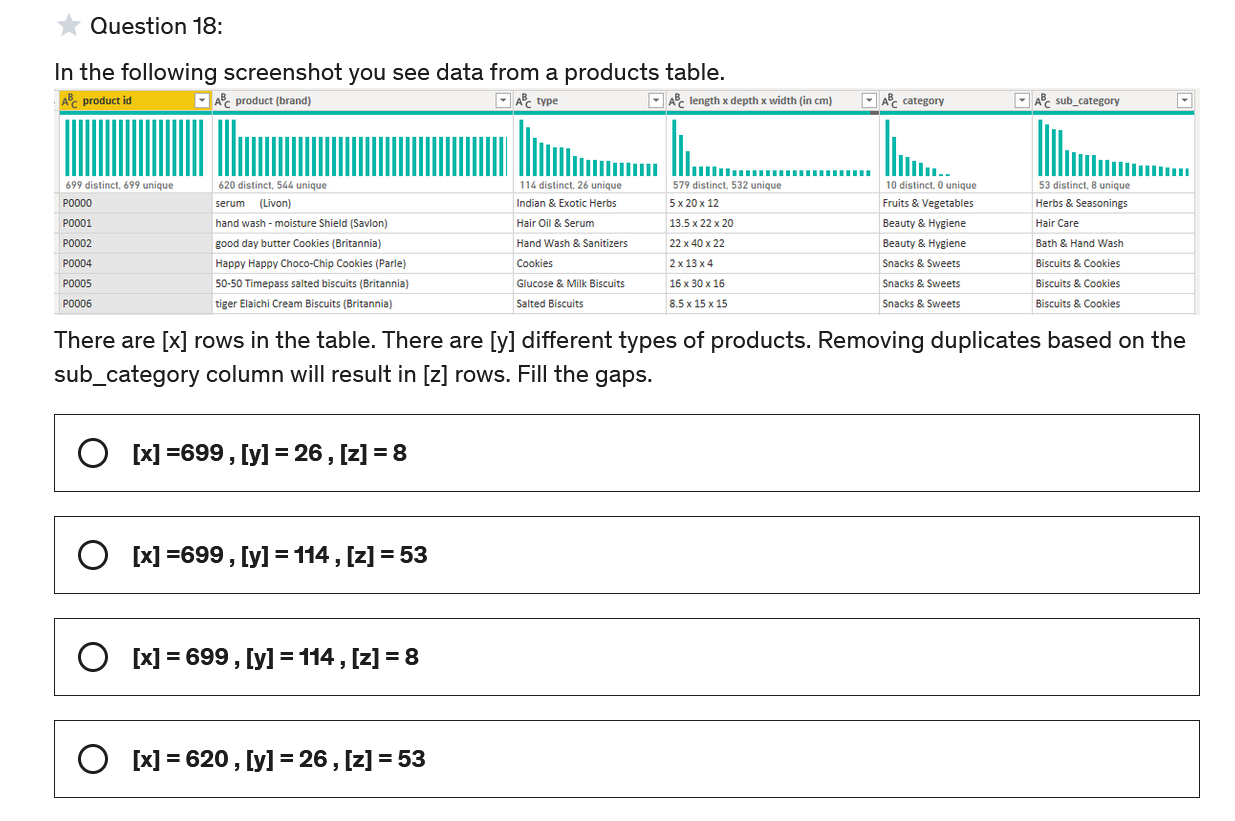








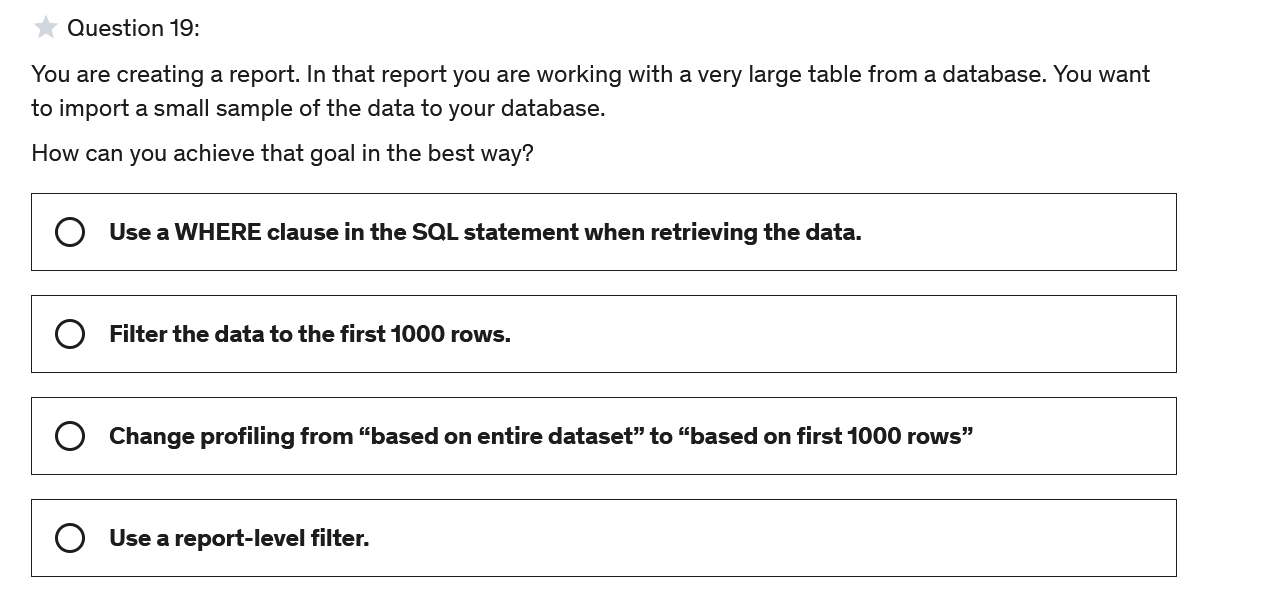
**Correct Answer: B**



**Correct Answer: B**

**Explanation**

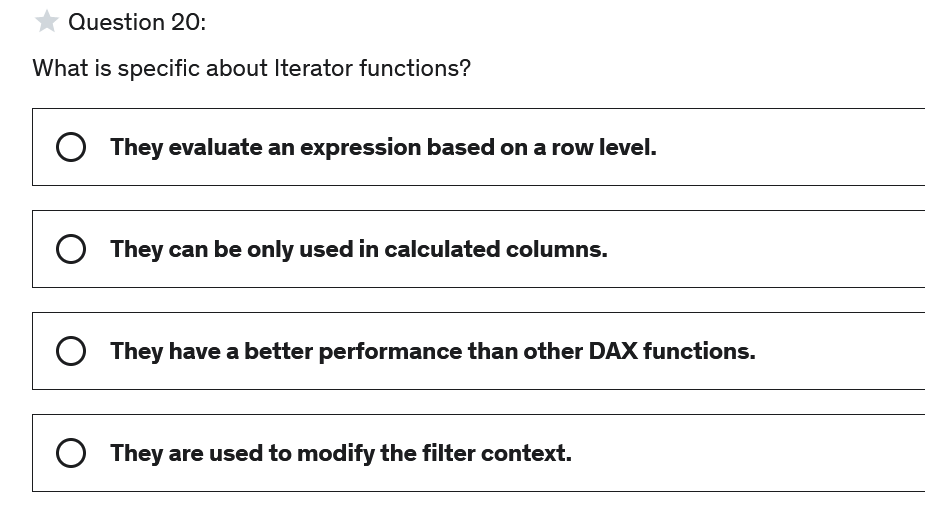
Since there are 699 distinct and 699 unique values in the first column, we know that every single value is unique, hence there are 699 rows.   
Since there are 114 distinct values in the type column, we know that there are 114 different types of products.Since there are 53 distinct values in the column sub\_category Removing duplicates based on that column would result in 53 rows.



**Correct Answer: A**

#### Explanation

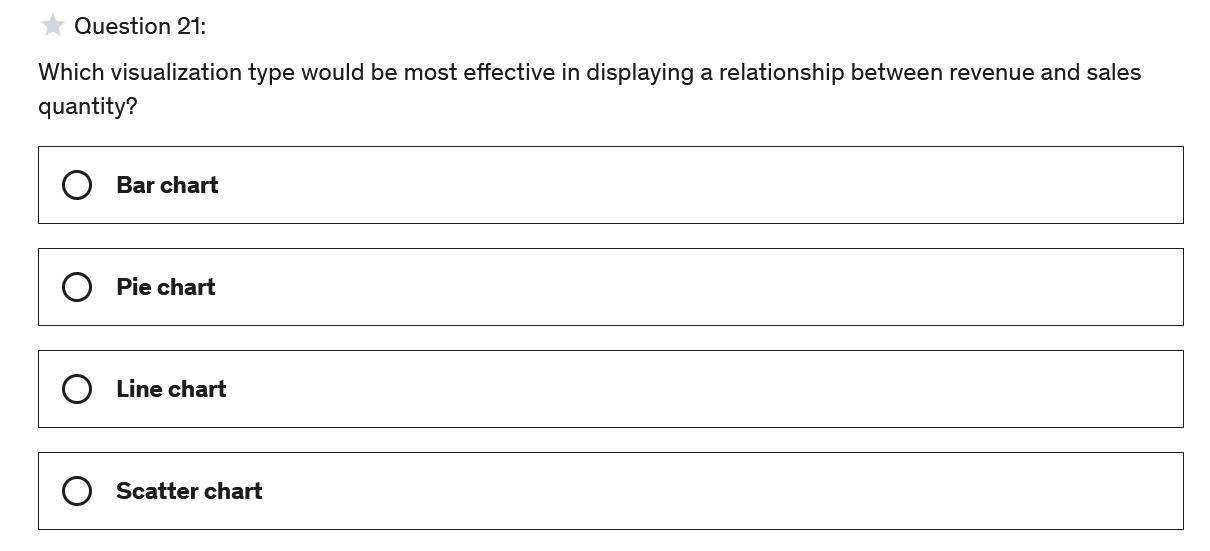
It is ideal to filter the data before it is imported – as early as possible in the development process. Therefor we should implement this step right at the level when the data is imported, that is using a SQL statement with a WHERE clause.



**Correct Answer: A**

#### Explanation

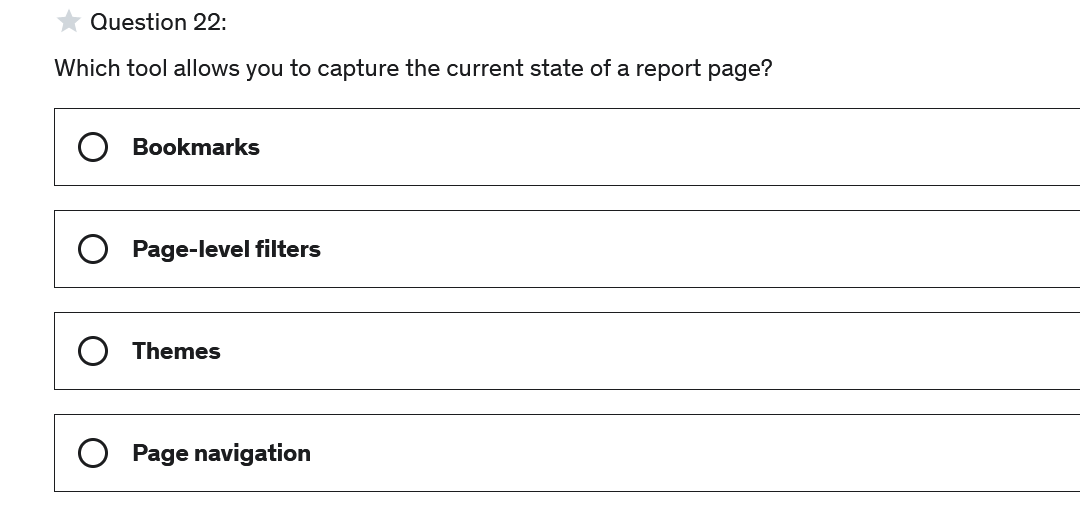
An iterator function aggregates more than the values of a single column. For each row it evaluates an expression first and after that aggregates all results of the row-based expressions.



**Correct Answer: D**

#### Explanation

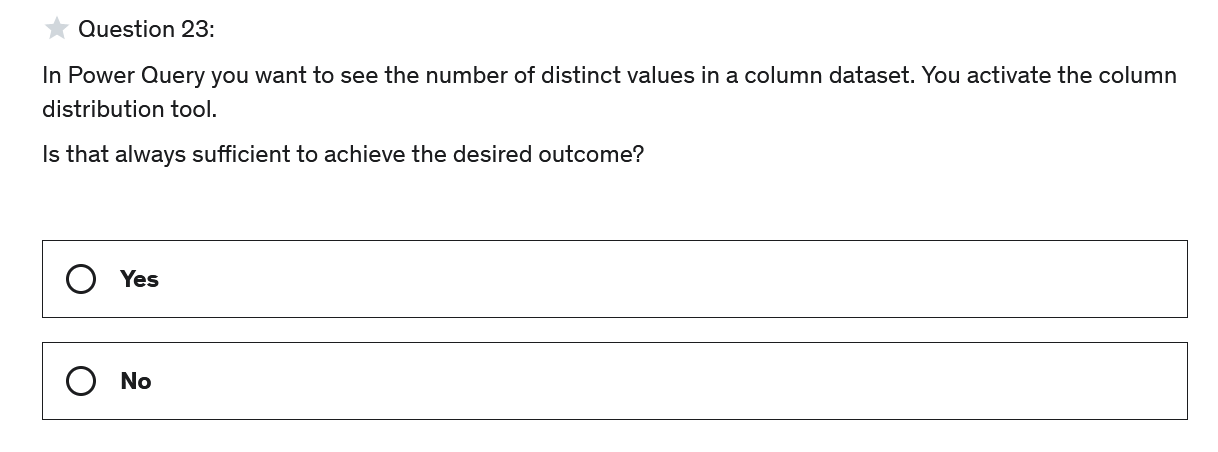
A scatter chart are a great choice to display the relationships between two numerical values.



**Correct Answer: A**

#### Explanation

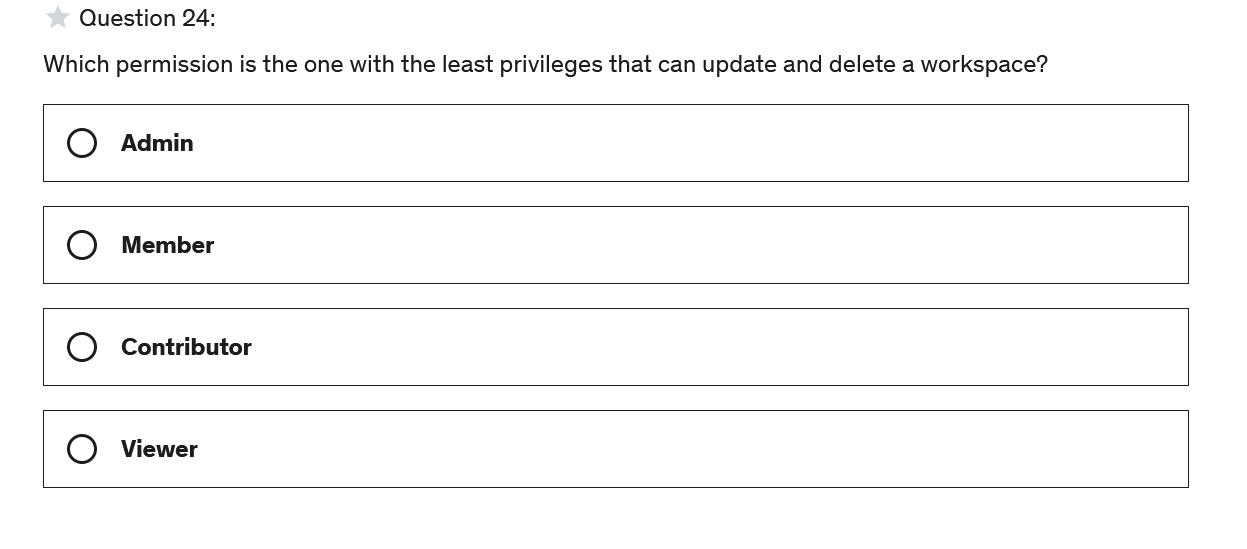
Bookmarks can save the current state of a report page, like filters, slicers, sort order and so on. When users want to go back to that state, they can to so by selecting the bookmark.



**Correct Answer: B**

#### Explanation

#### It is not always sufficient because per default column profiling such as column distribution is only based on the first 1,000 rows. Hence if we want to see the number of all distinct values we need to change "Profiling based on the entire dataset" first and then activate the column distribution tool.



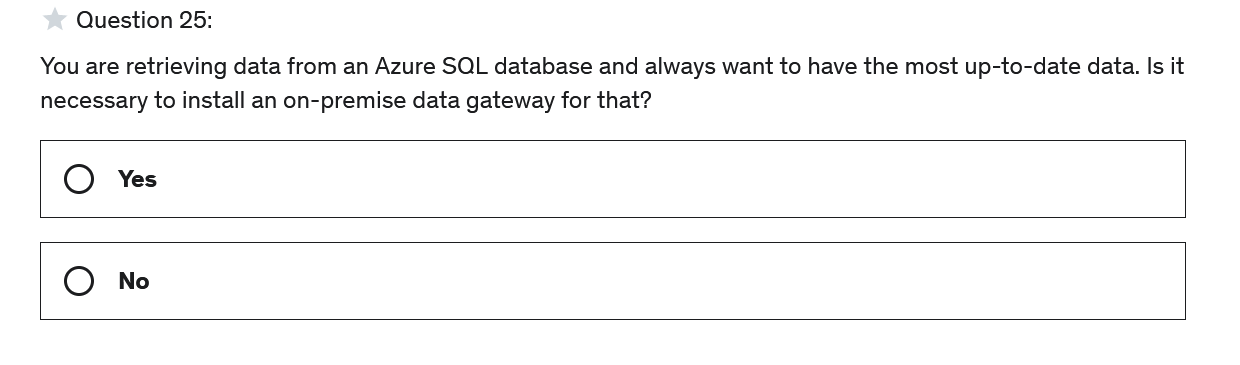
**Correct Answer: A**

#### Explanation

Only admins can delete a workspace.

For more details on the roles, see here:

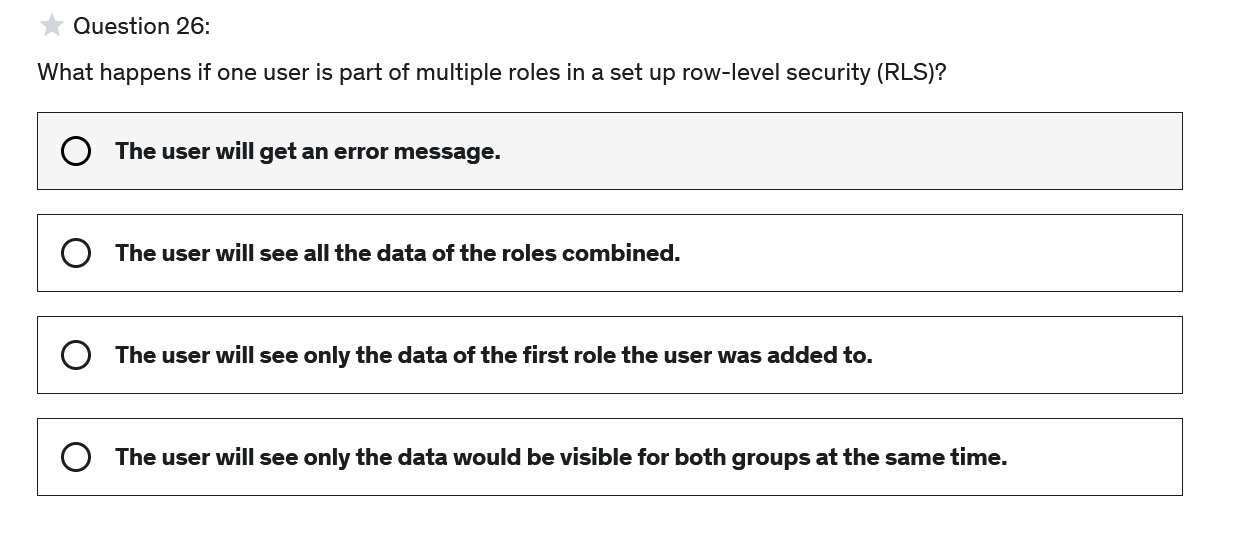
<https://docs.microsoft.com/en-us/power-bi/collaborate-share/service-roles-new-workspaces>



**Correct Answer: B**

#### Explanation

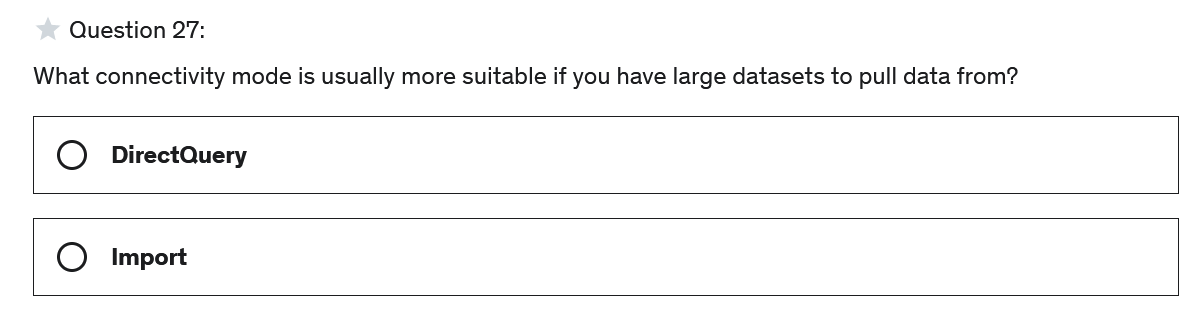
No, because the data is already in the cloud and not in an on-premise data center.



**Correct Answer: B**

#### Explanation

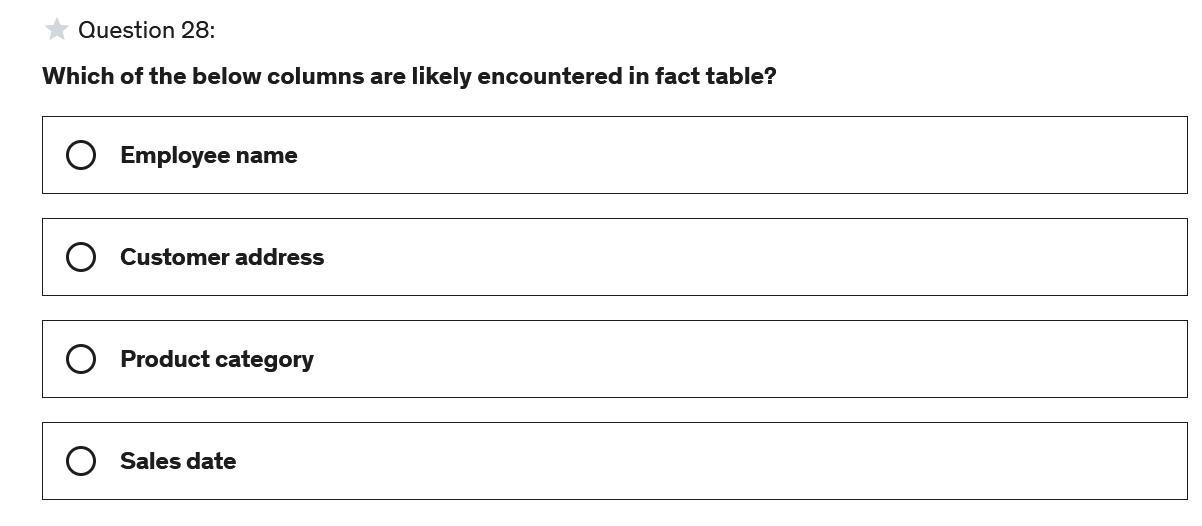
When a user is member of multiple roles, they will see all data of the roles combined.



**Correct Answer: A**

#### Explanation

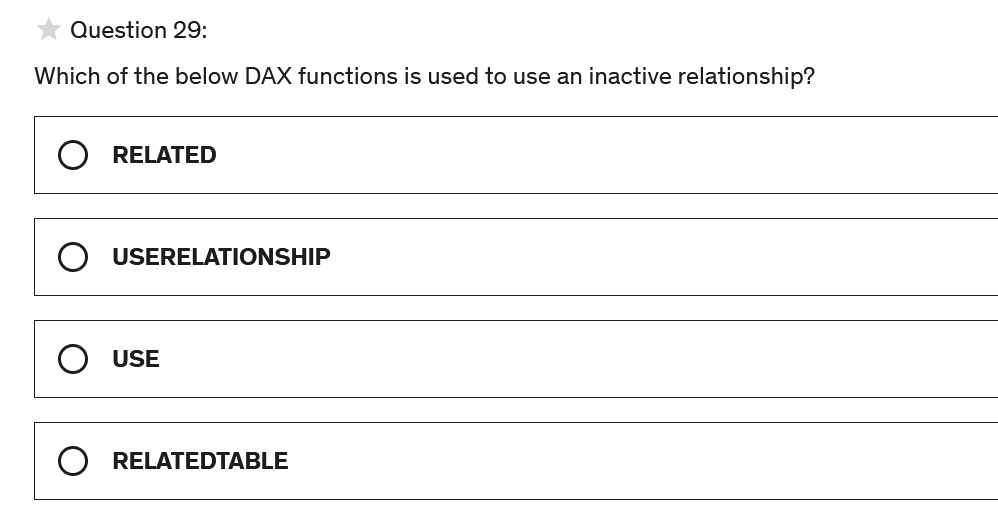
With DirectQuery no data is copied or imported into the Power BI file and therefor, to avoid bottlenecks for very large datasets it is usually more suitable to not import all the data but use DirectQuery instead.



**Correct Answer: D**

#### Explanation

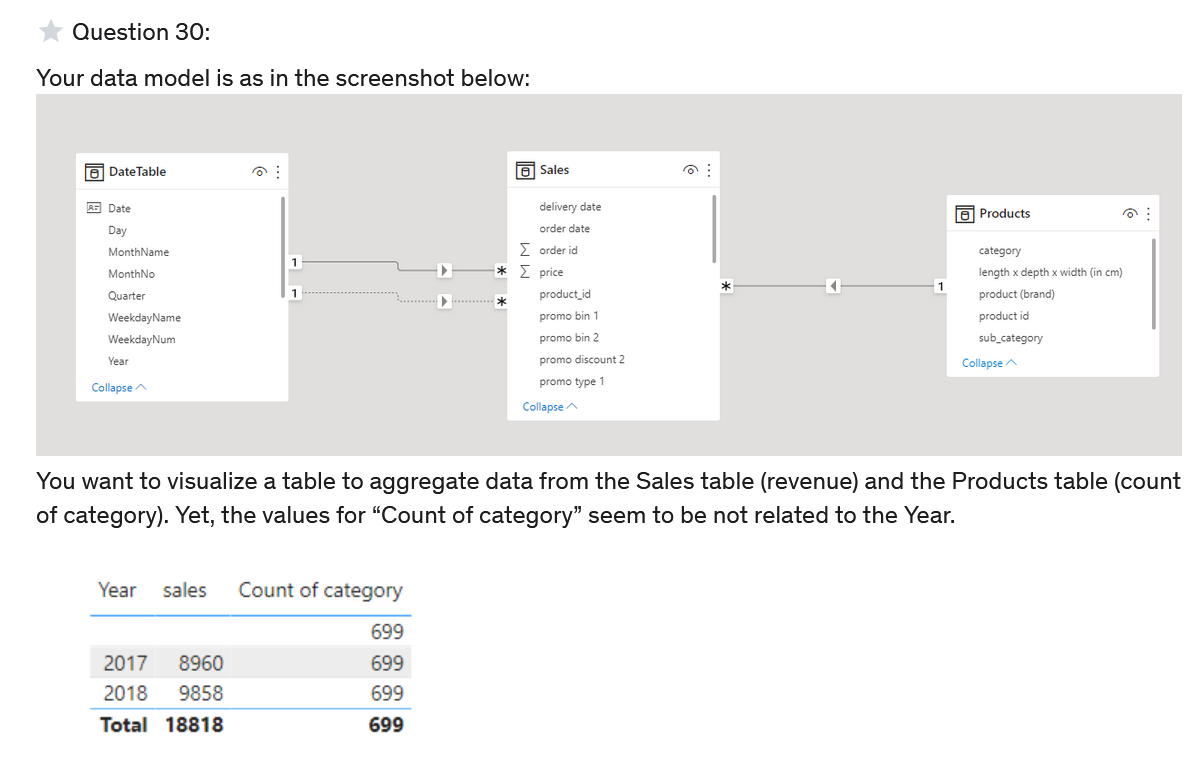
Sales date would be typically a column found in a fact table since fact table usually contain transaction-/event-based data like a sale with a sales date.

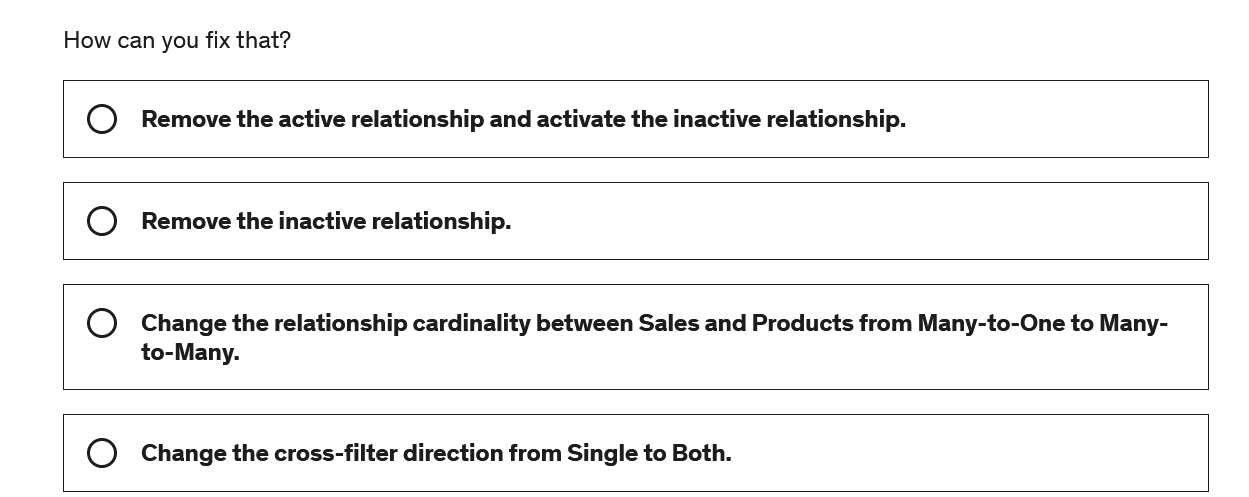


**Correct Answer: B**

#### Explanation

USERELATIONSHIP() is used in CALCULATE to use a specific relationship that exists between two tables even though it is not active.

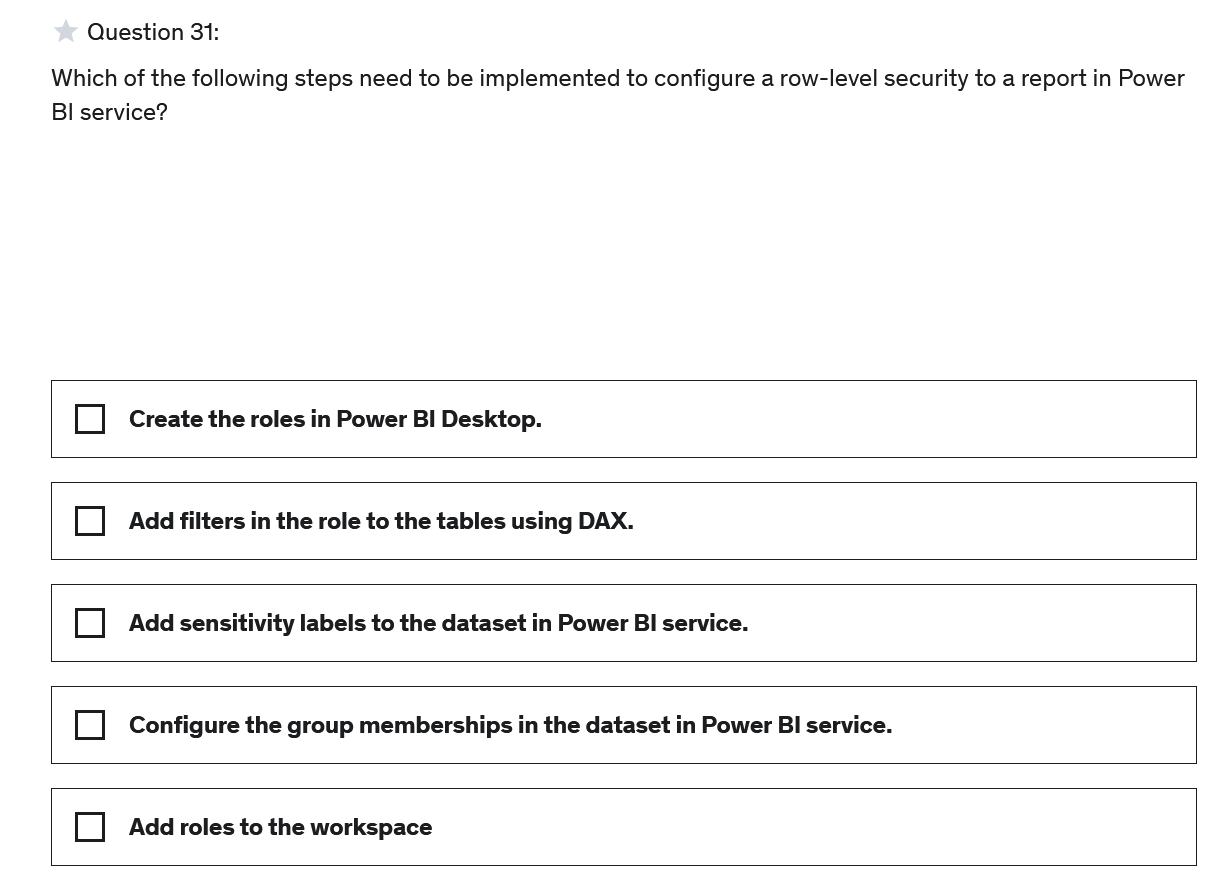




**Correct Answer: D**

#### Explanation

All the same numbers indicate the relationship is not filtering from sales to products to aggregate something in the product table. We can see that also in the data model since the arrow only points from Products to Sales.   
It is most natural to filter from the dimension (Products) and aggregate in the facts (Sales). Therefore, mostly it is enough to have the cross-filter direct set to "Single". If we also want to enable the other direct to also be able to aggregate something in the dimension table we can set the cross-filter direction to "Both".



**Correct Answer: A, B, D**

#### Explanation

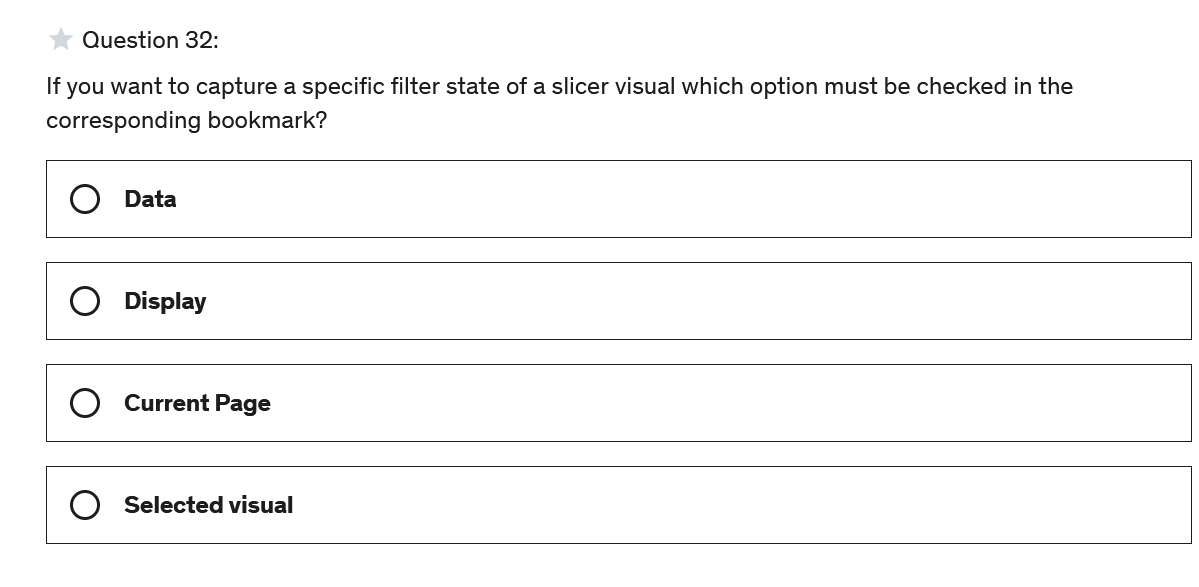
We first need to set up the roles in Power BI Desktop.

We need to define filters to define how the role filters data.

And lastly we need to configure the group memberships in the dataset in Power BI service - meaning we add users to the different groups.

For more details see:

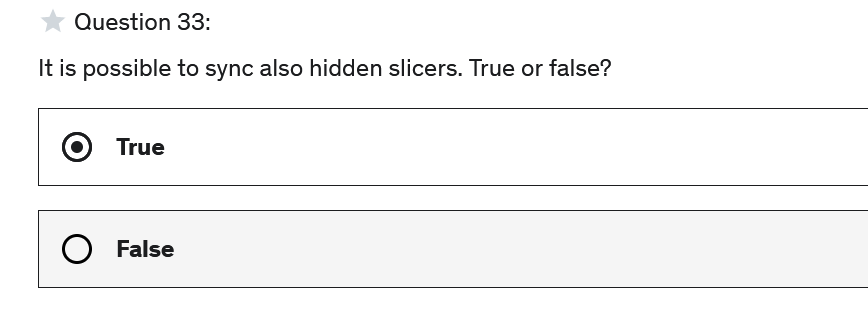
<https://docs.microsoft.com/en-us/sql/relational-databases/security/row-level-security>



**Correct Answer: A**

#### Explanation

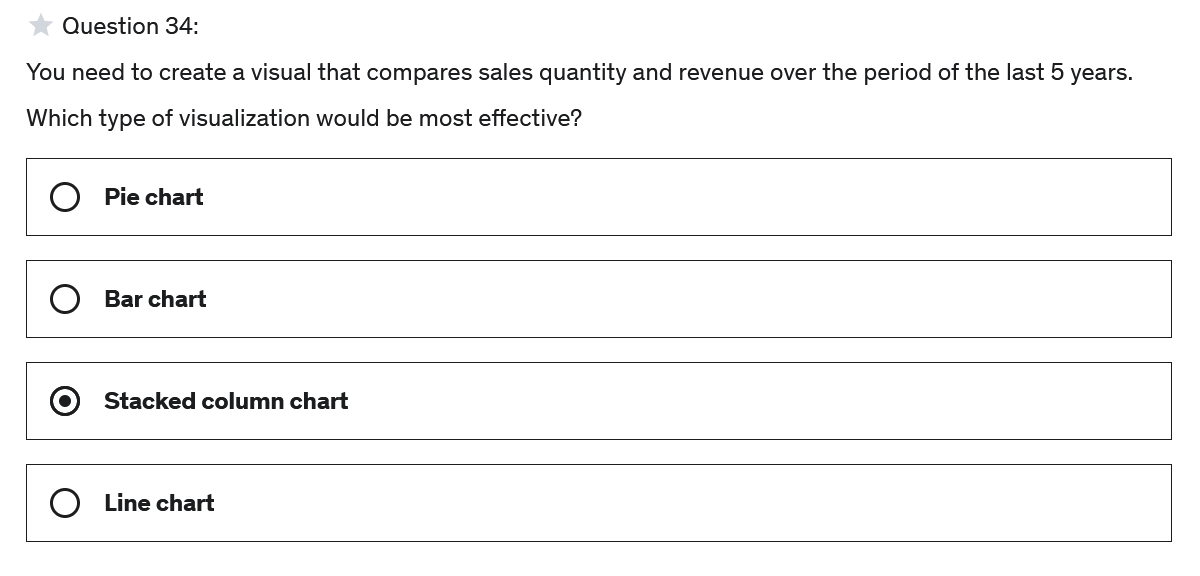
A specific filter state as well as sort order and other data related states need to have the option "Data" enable in the bookmark to be captured.



**Correct Answer: A**

#### Explanation

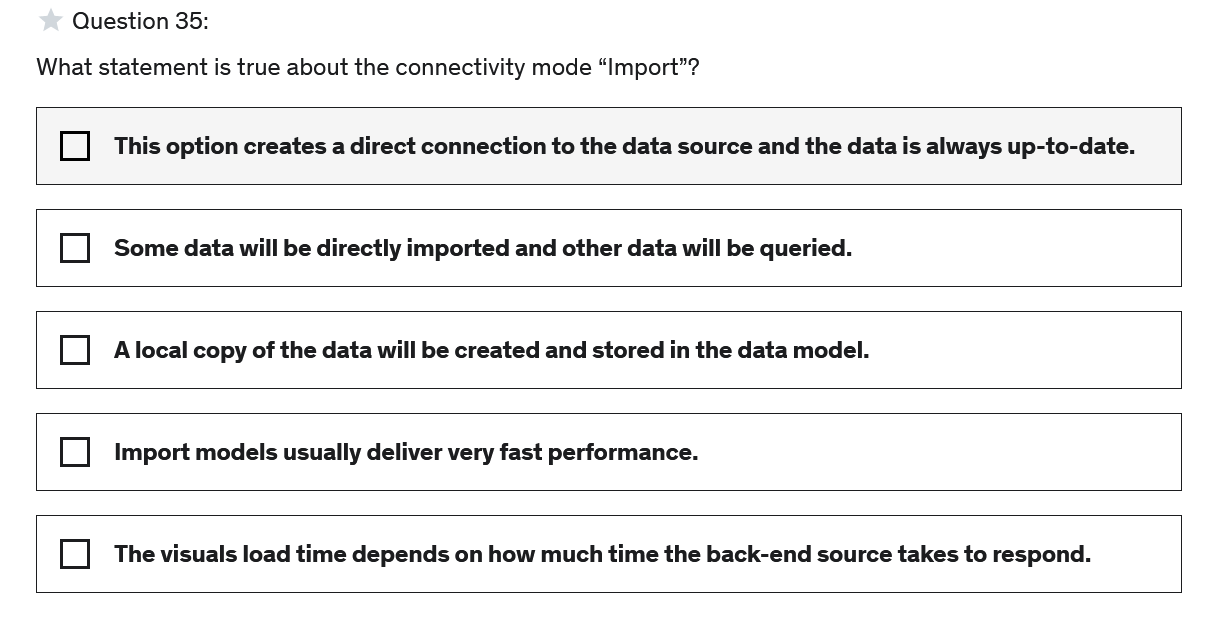
Synced slicers can also be hidden. This can be done independently from the sync status.



**Correct Answer: D**

#### Explanation

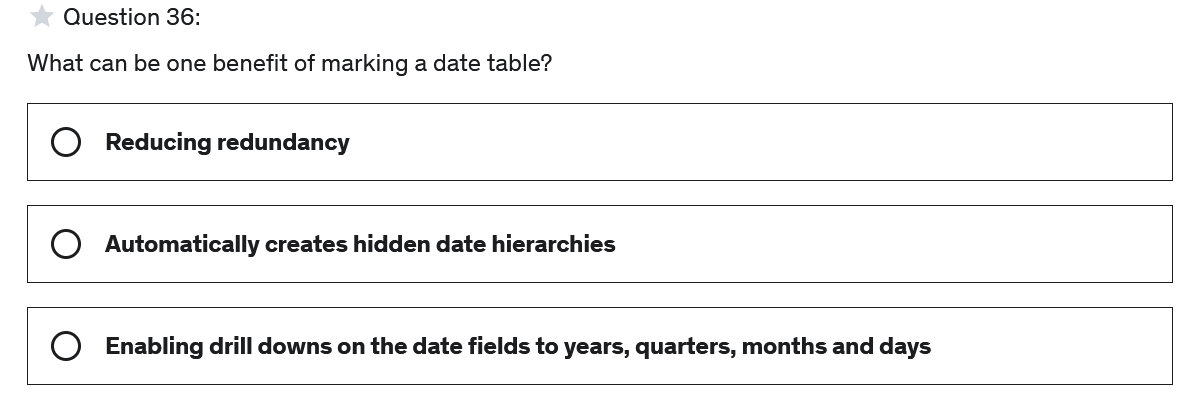
Usually if we visualize data over time a line chart or an area chart is the best choice. This is also true if we want to compare in different categories over time.



**Correct Answer: C,D**

#### Explanation

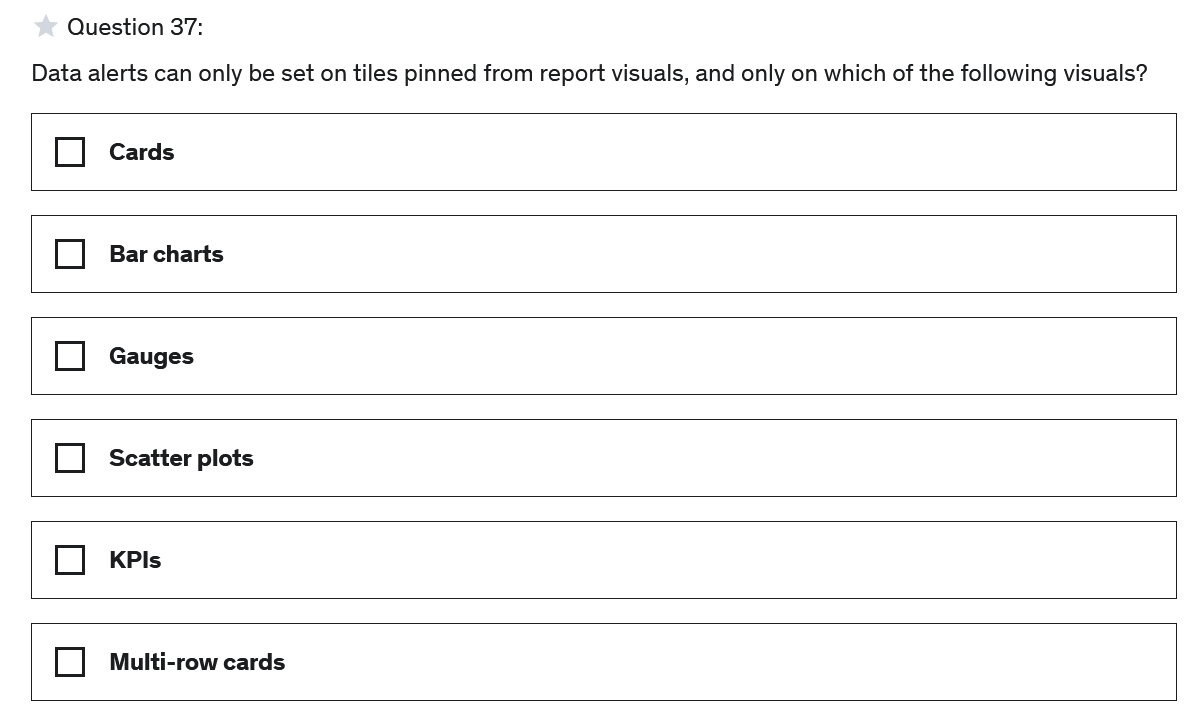
Import mode is the default mode used to connect to most data sources. A copy of the data is fully loaded into memory of the Power BI capacity. Once it is loaded it is usually is very high in performancedue to in-memory querying.



**Correct Answer: A**

#### Explanation

Power BI Desktop can automatically identify date columns, and then create a date hierarchy. You can use those built-in hierarchies but you can also add your own date table and in that case it makes sense to mark a table as a date table to not create those hidden date hierarchies to avoid redundant data.



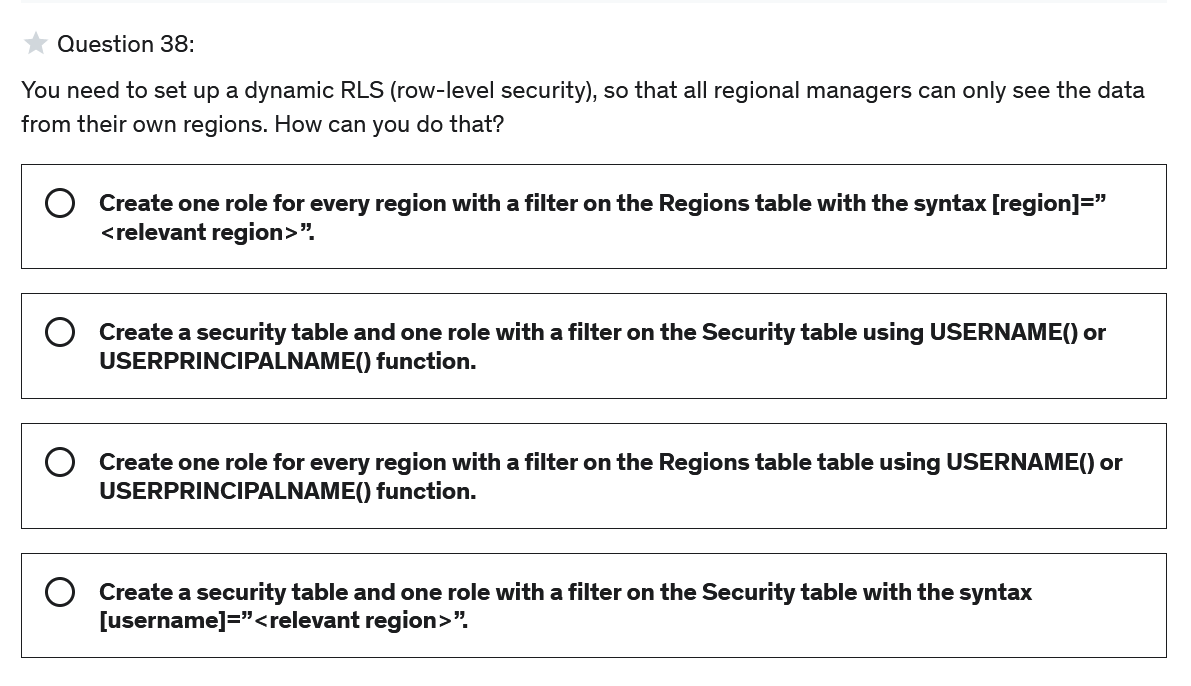
**Correct Answer: A,C,E**

#### Explanation

It is only possible to set data alerts to tiles pinned to a dashboard, and only for cards, KPIs and gauges.

To see more details on that see:

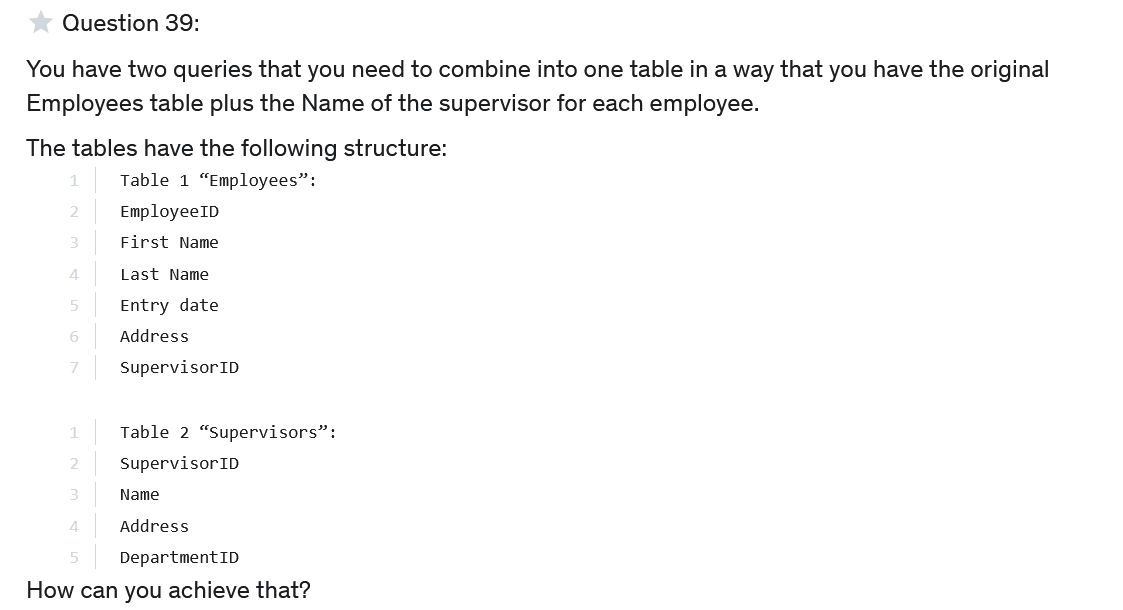
<https://docs.microsoft.com/en-us/power-bi/create-reports/service-set-data-alerts>

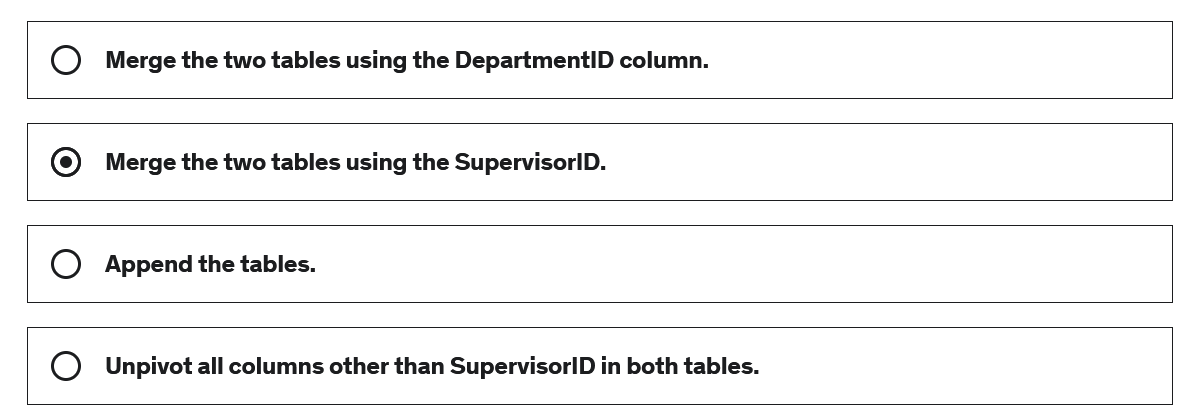


**Correct Answer: A**

#### Explanation

You can set up dynamic RLS by creating a Security tabe, setting up the proper relationships and using either *username()* or *userprincipalname()* DAX functions.

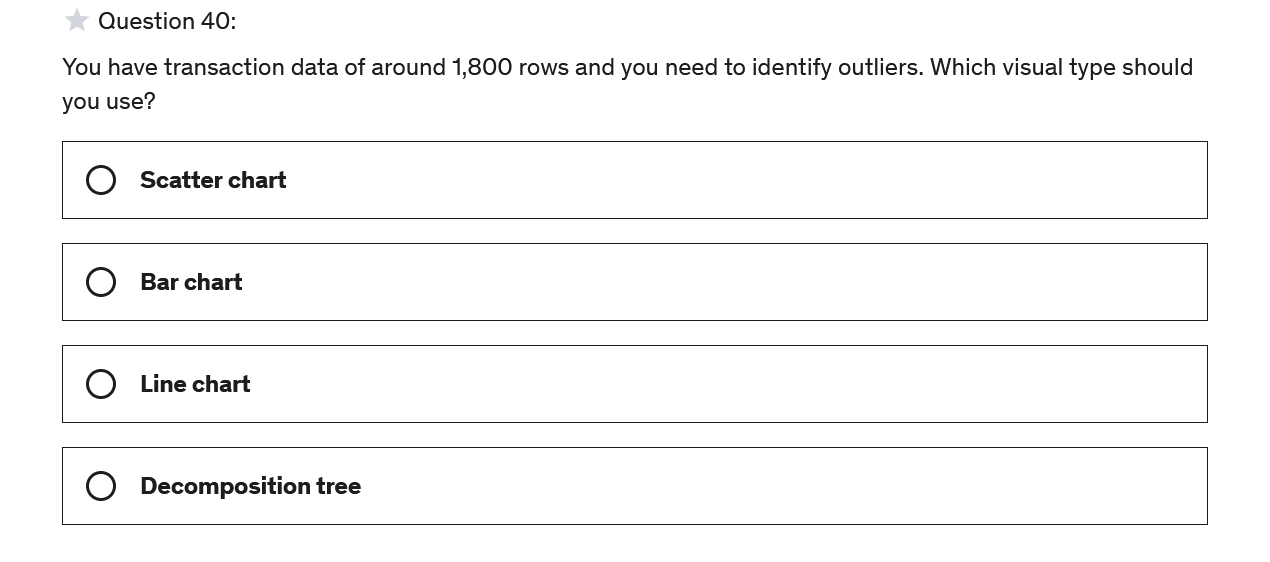




**Correct Answer: B**

#### Explanation

If we want to combine tables to add additional columns we can merge two tables using a common column. In this case the common column is "SupervisorID". It is the only column appearing in both tables.

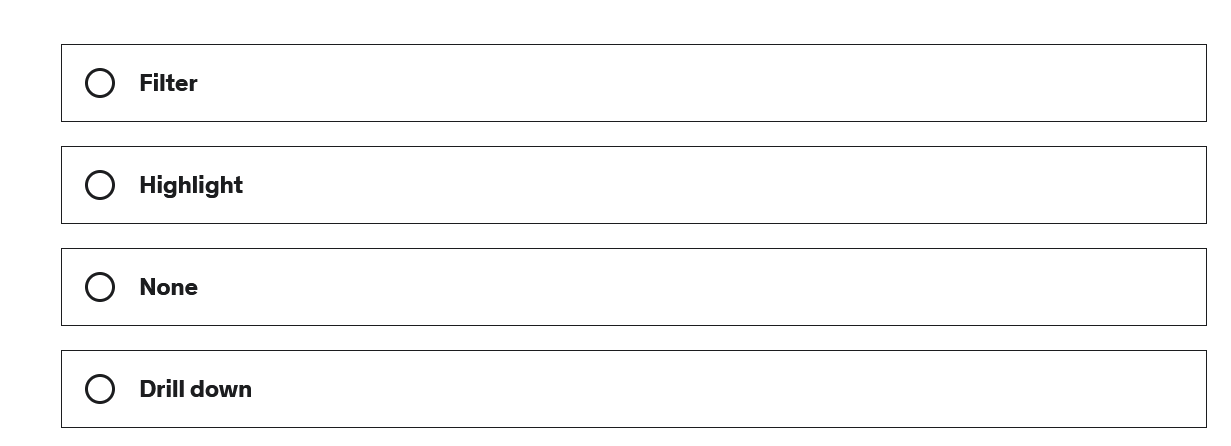


**Correct Answer: A**

#### Explanation

The most effective visual type to identify outliers in large datasets are scatter charts.

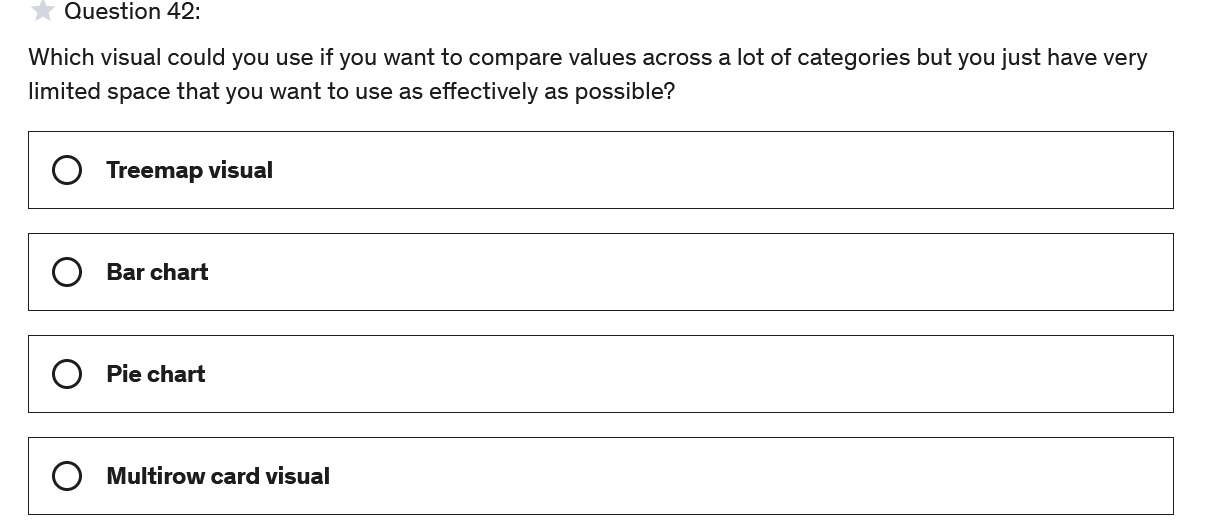




**Correct Answer: A**

#### Explanation

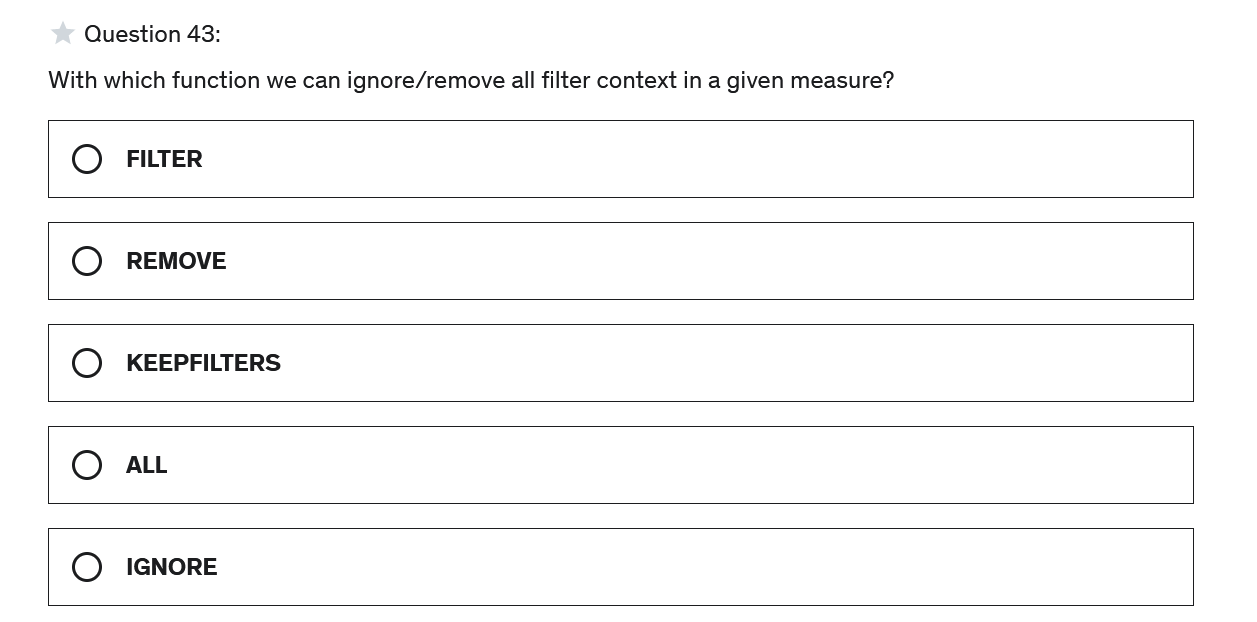
The first icon is selected which is the option "Filter".



**Correct Answer: A**

#### Explanation

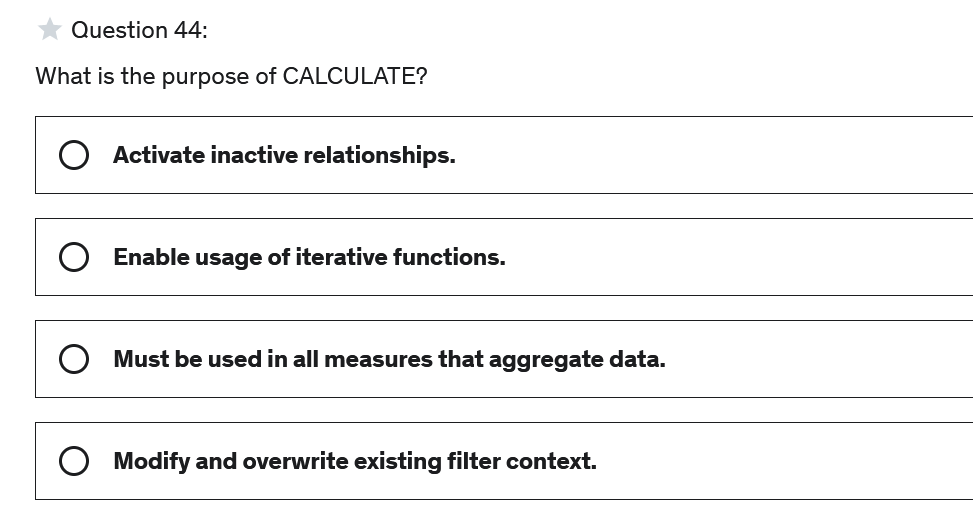
Treemaps are used to compare data in different categories and one benefit with them is that they use the available space very effeciently, hence it can be a good option if the space is very limited and we have a lot of categories to compare data.



**Correct Answer: D**

#### Explanation

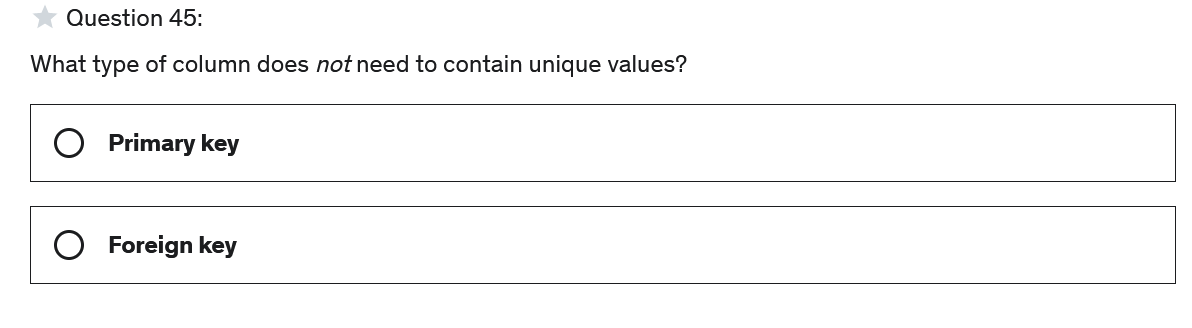
The ALL function will ignore any filters that might be applied and is useful when we want to clear all filters in the given measure to make calculations based on all rows in a table.



**Correct Answer: D**

#### Explanation

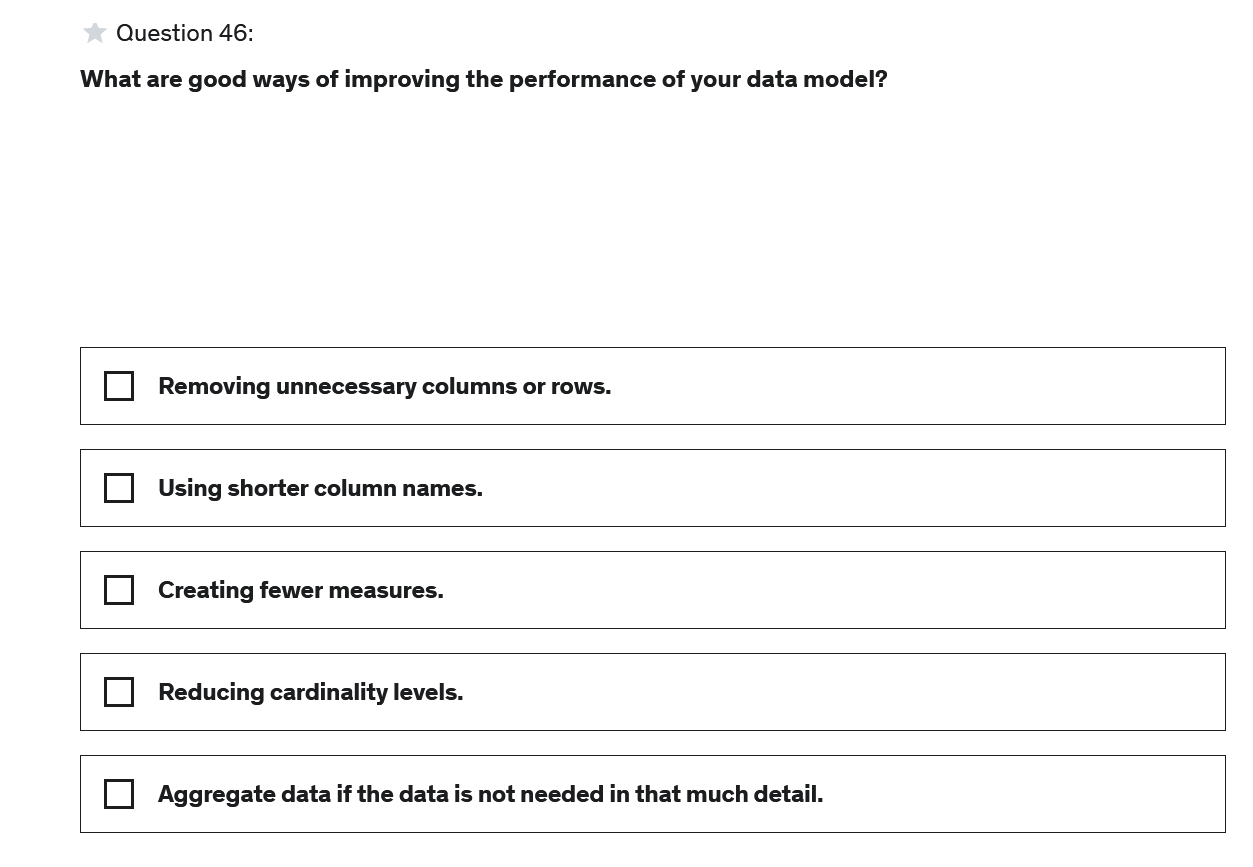
The CALCULATE function is used whenever we want to evaluate an expression in a modified filter context.



**Correct Answer: B**

#### Explanation

Primary keys are used to uniquely identify every row in a table and therefore all values need to be unique. This is not necessary though for the foreign key.

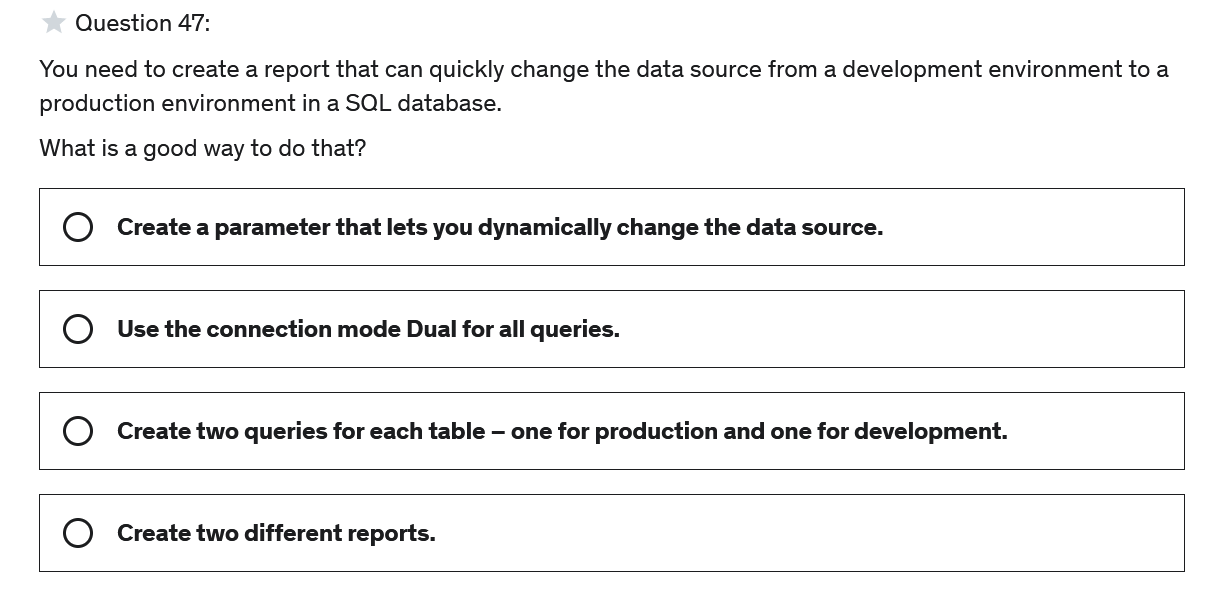


**Correct Answer: A,D,E**

#### Explanation

For more details on model performance see here:

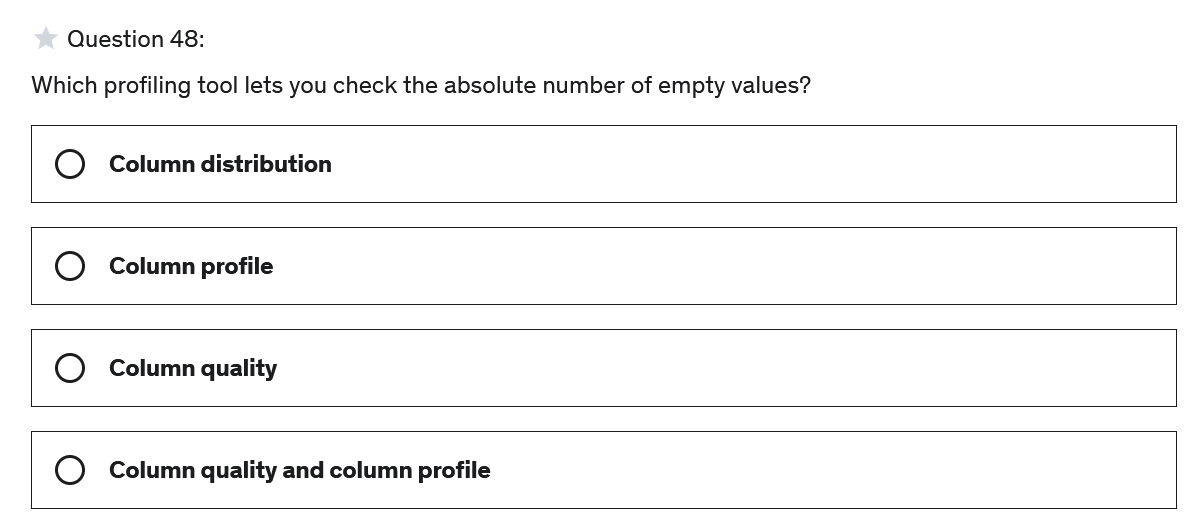
<https://docs.microsoft.com/en-us/learn/modules/optimize-model-power-bi/1-introduction>



**Correct Answer: A**

#### Explanation

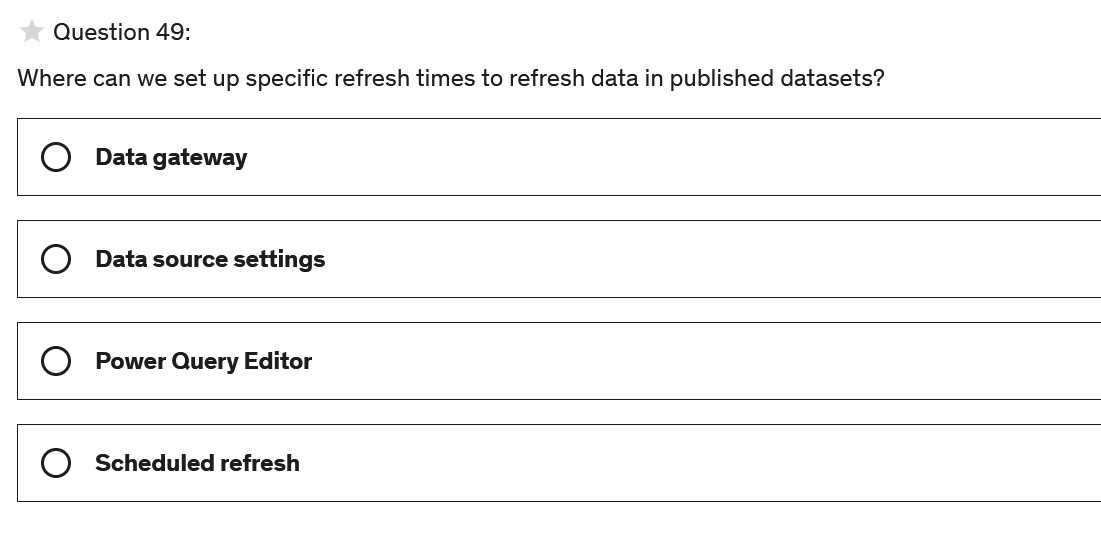
Dynamically changing the data source is one common use-case for parameters.



**Correct Answer: B**

#### Explanation

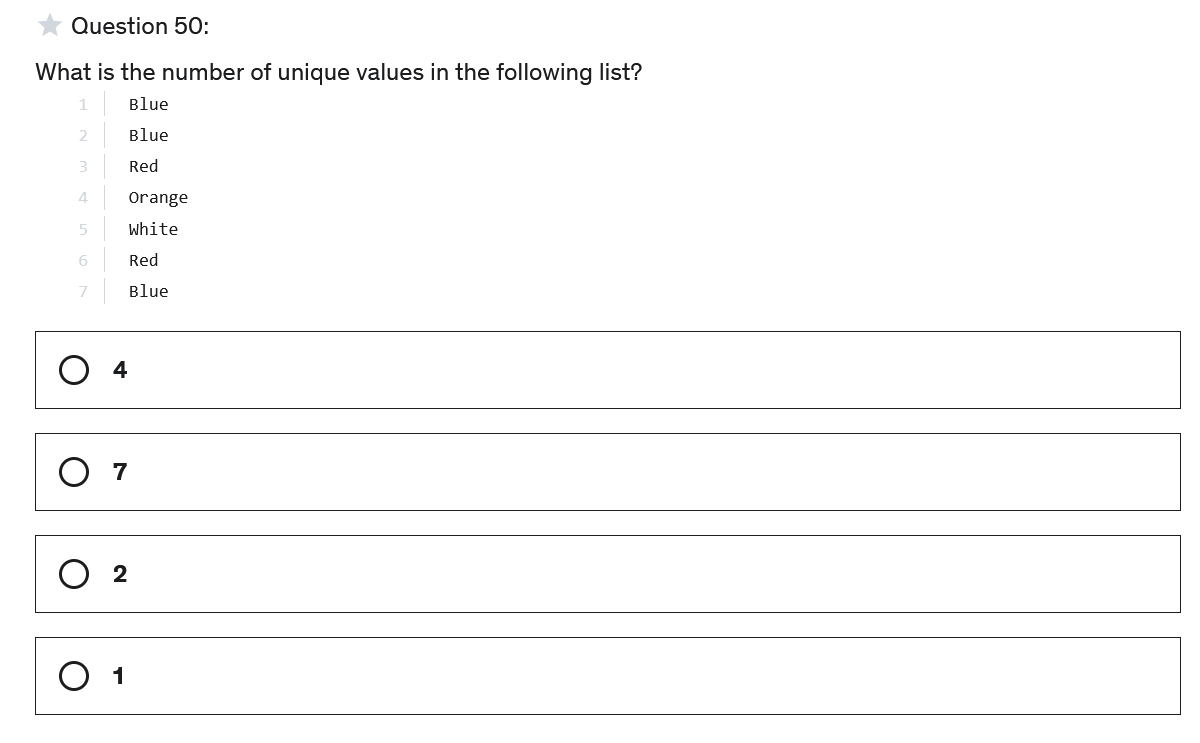
The absolute number of empty values is only visible in column profile.



**Correct Answer: D**

#### Explanation

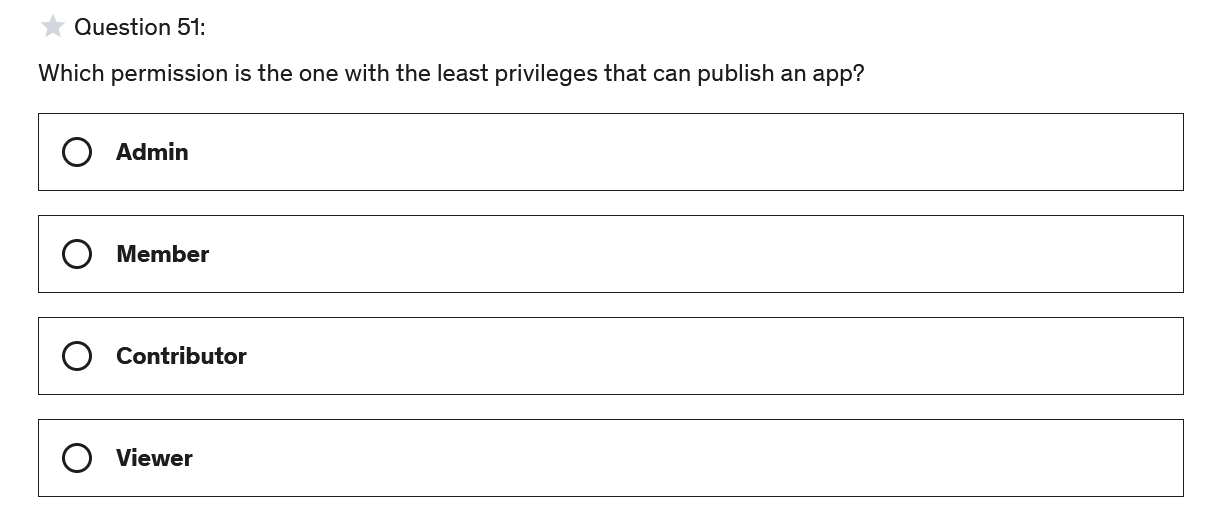
We can set specific refresh times in the settings of a dataset under "Scheduled refresh".



**Correct Answer: C**

#### Explanation

Orange and White are the only values there only occur once, hence there are only two unique values.



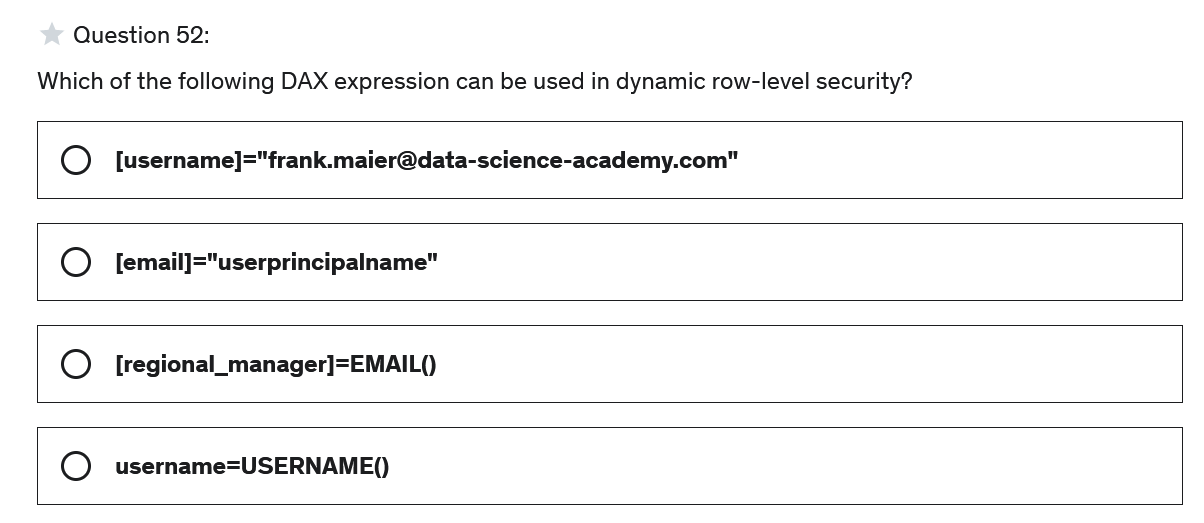
**Correct Answer: B**

#### Explanation

Only admins and members can update, publish and unplish an app in a workspace.

For more details on the privileges of the roles see here:

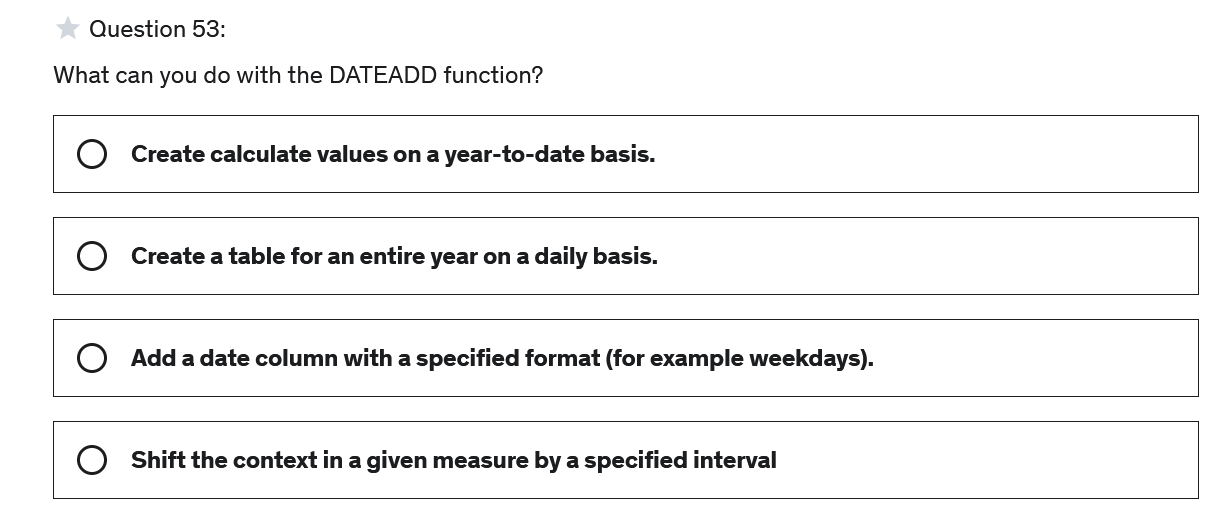
<https://docs.microsoft.com/en-us/power-bi/collaborate-share/service-roles-new-workspaces>



**Correct Answer: D**

#### Explanation

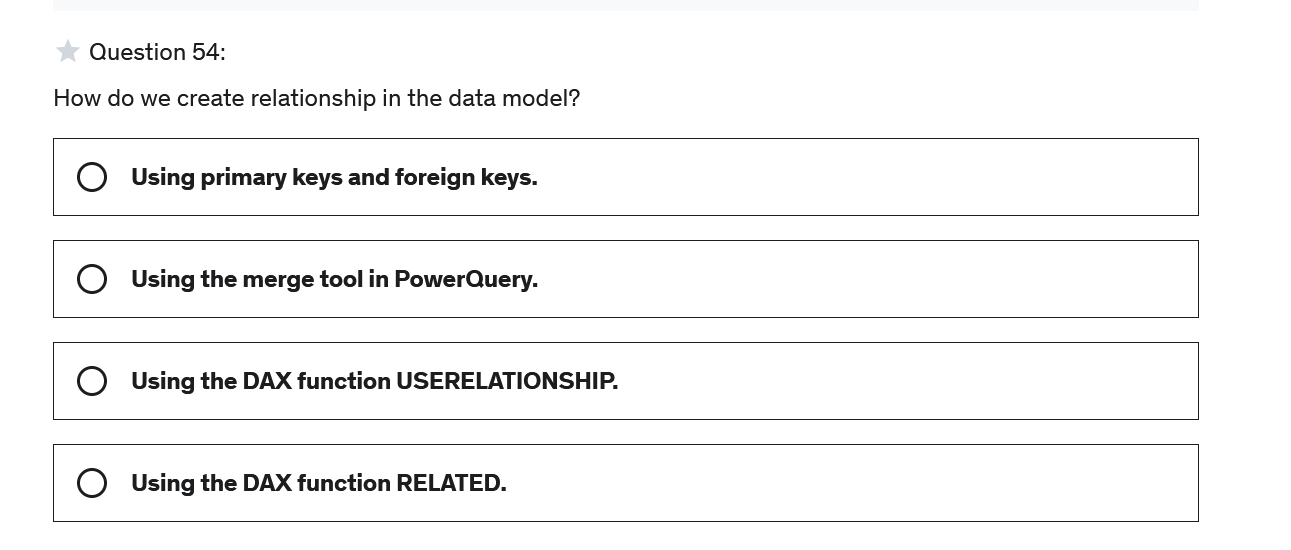
USERNAME() and USERPRINCIPALNAME() eturns the current users that is currently logged in. Using a proper set up data model this can be used in dynamic RLS.



**Correct Answer: D**

#### Explanation

DATEADD can be used to shift dates by a specified number of intervals in the current context.



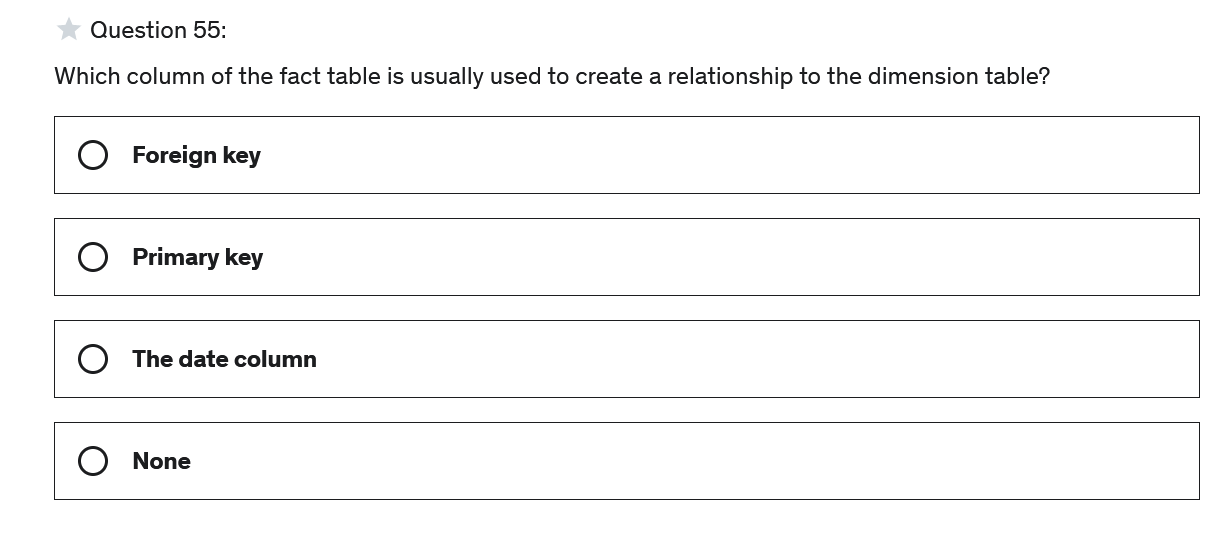
**Correct Answer: A**

#### Explanation

Primary key and foreign keys are usually used to create a relationship between tables.

For more details see:

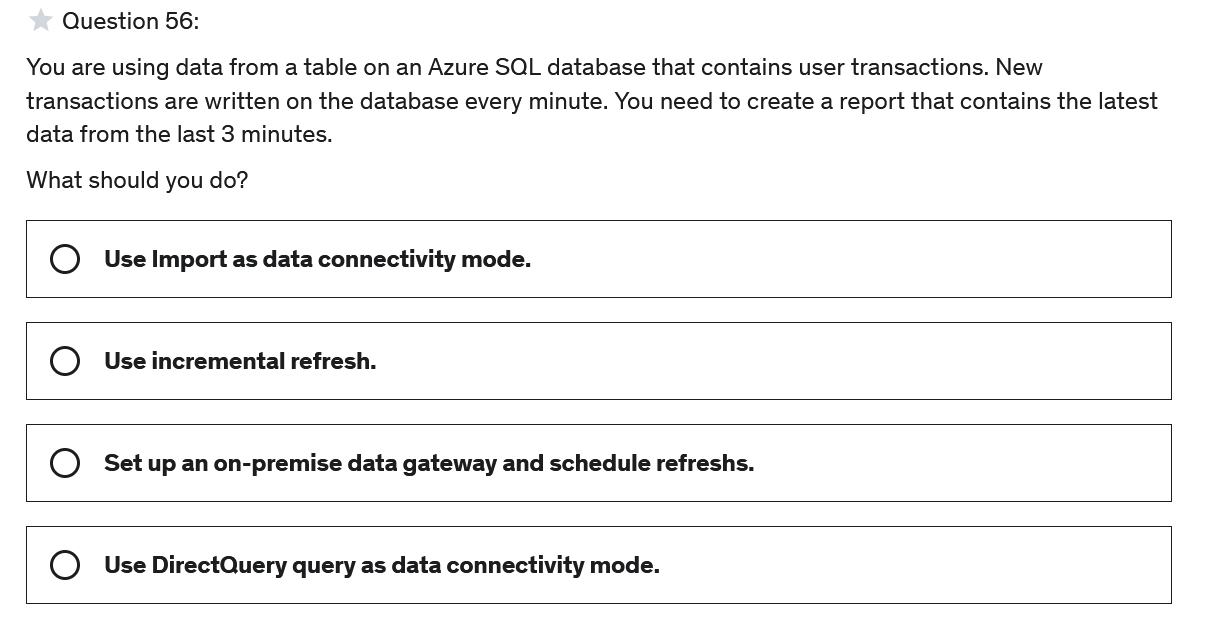
<https://docs.microsoft.com/en-us/power-bi/transform-model/desktop-create-and-manage-relationships>



**Correct Answer: A**

#### Explanation

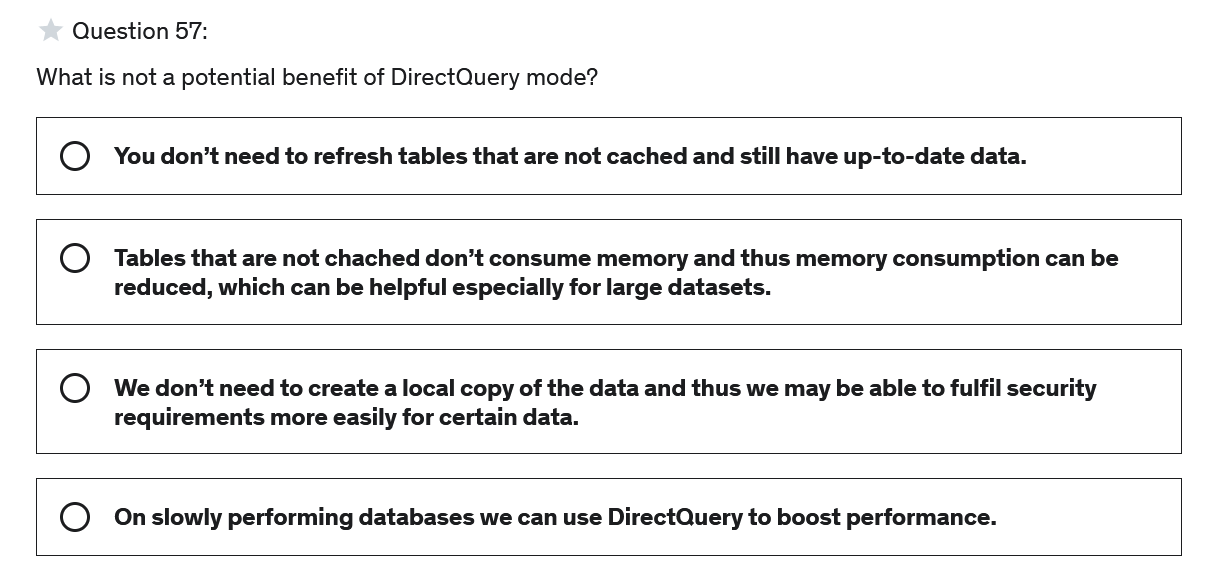
Usually there is a foreign key in a fact table that references rows in a dimension table. Hence, the foreign key can be used to create a relationship between a fact and a dimension table.



**Correct Answer: D**

#### Explanation

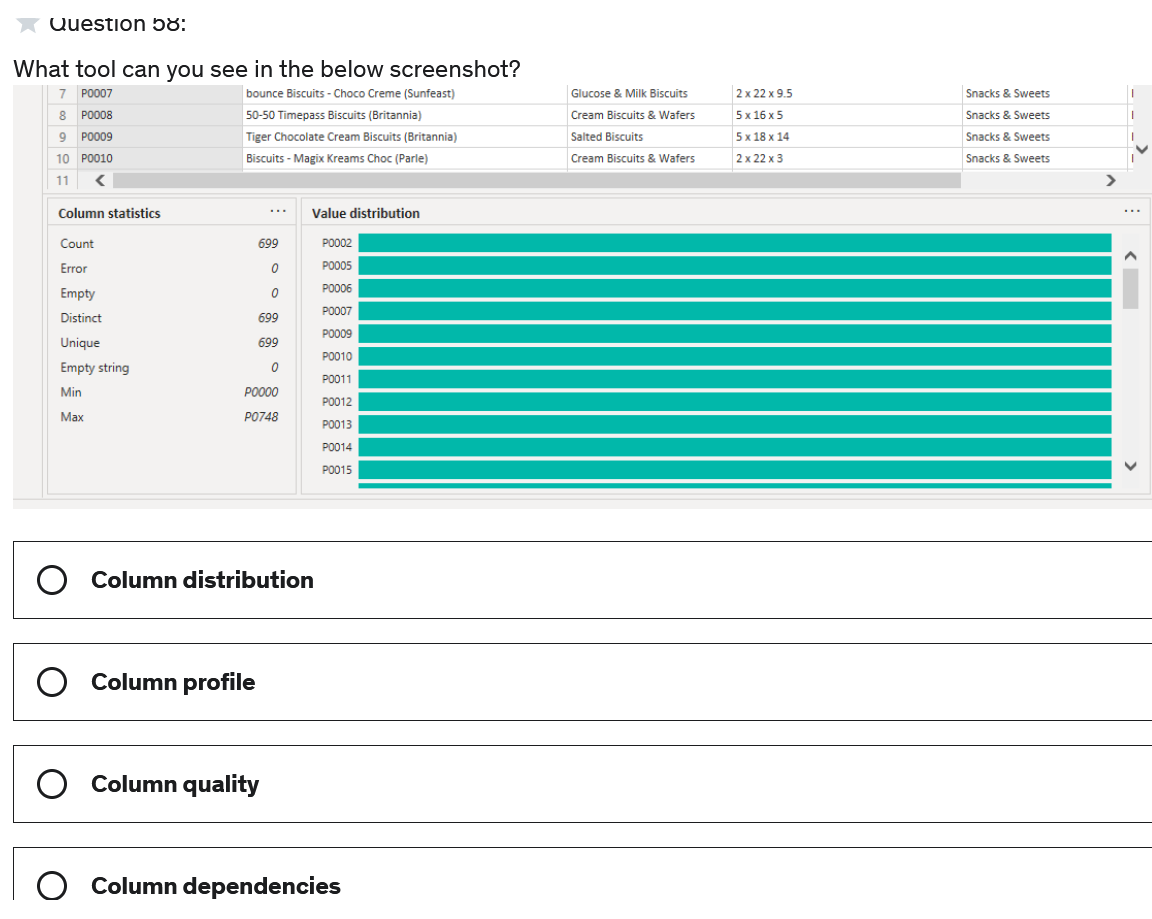
You don’t need a data gateway because the data is in the cloud and not on-premise. If you need data in the most up-to-date version the preferred method is using DirectQuery as data connectivity mode.



**Correct Answer: D**

#### Explanation

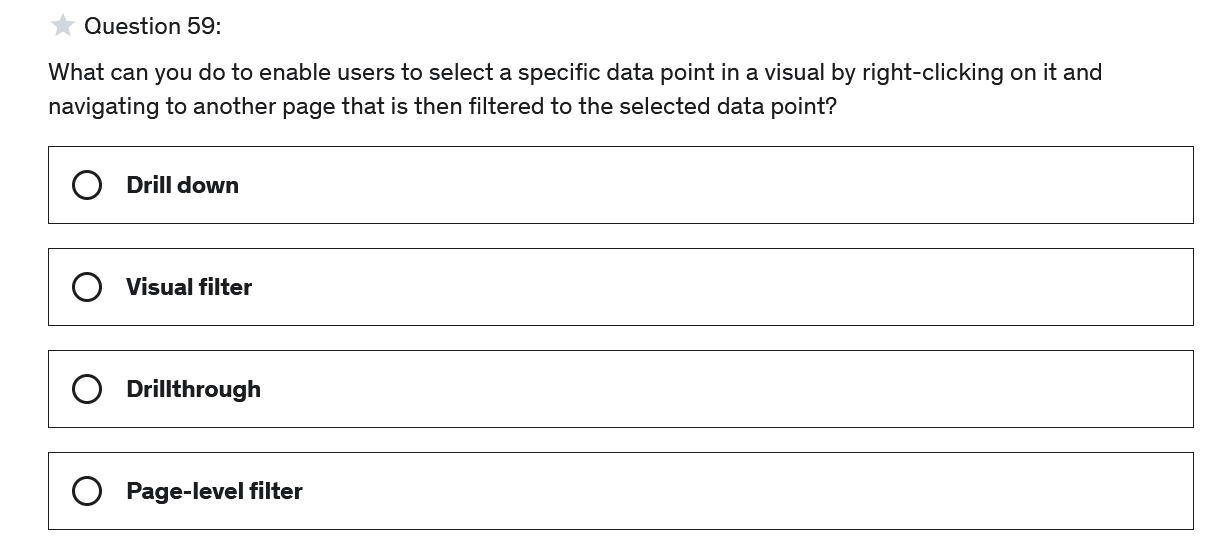
Load time on visuals when using DirectQuery depend on the performance of the database we directly query from. Therefor, when we have a slowly performing database it is not a good idea to use DirectQuery.



**Correct Answer: B**

#### Explanation

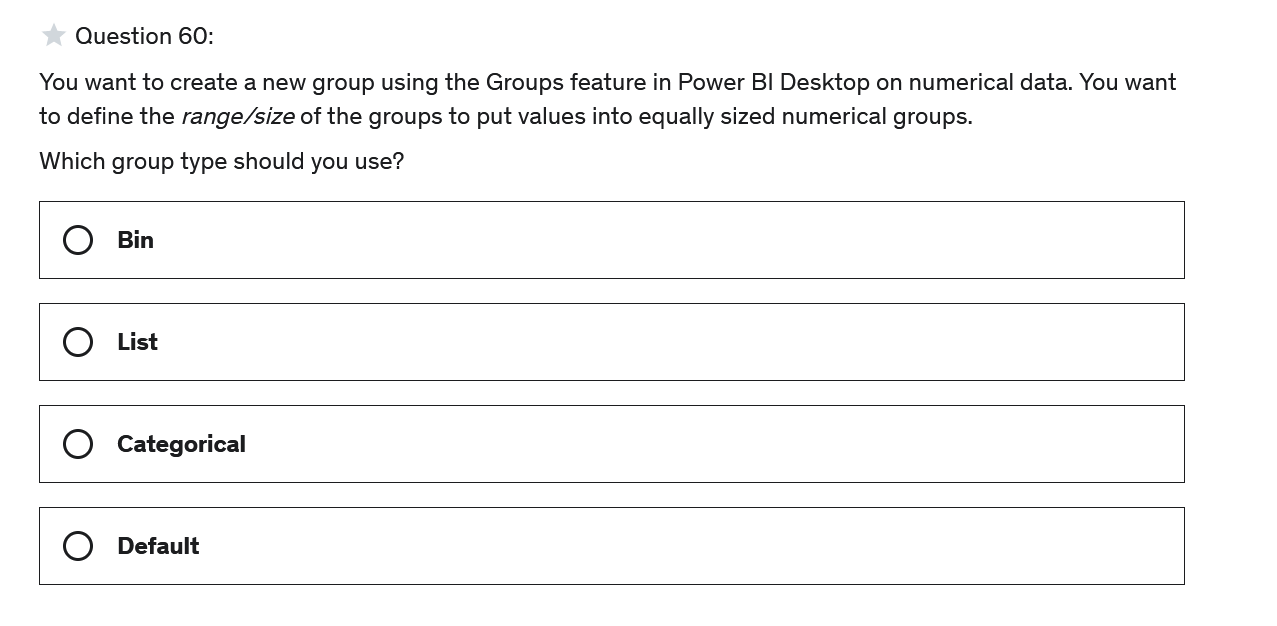
This tool is called column profile. This tool provides a more in-depth look at the data in a selected column.



**Correct Answer: C**

#### Explanation

With drillthroughs users can select a specific data point in a source page, and navigate to another page to get details that are filtered to that selected data point.



**Correct Answer: A**

#### Explanation

If we want to group numerical or time-based fields by defining the bin size we need to use the group typ bin.