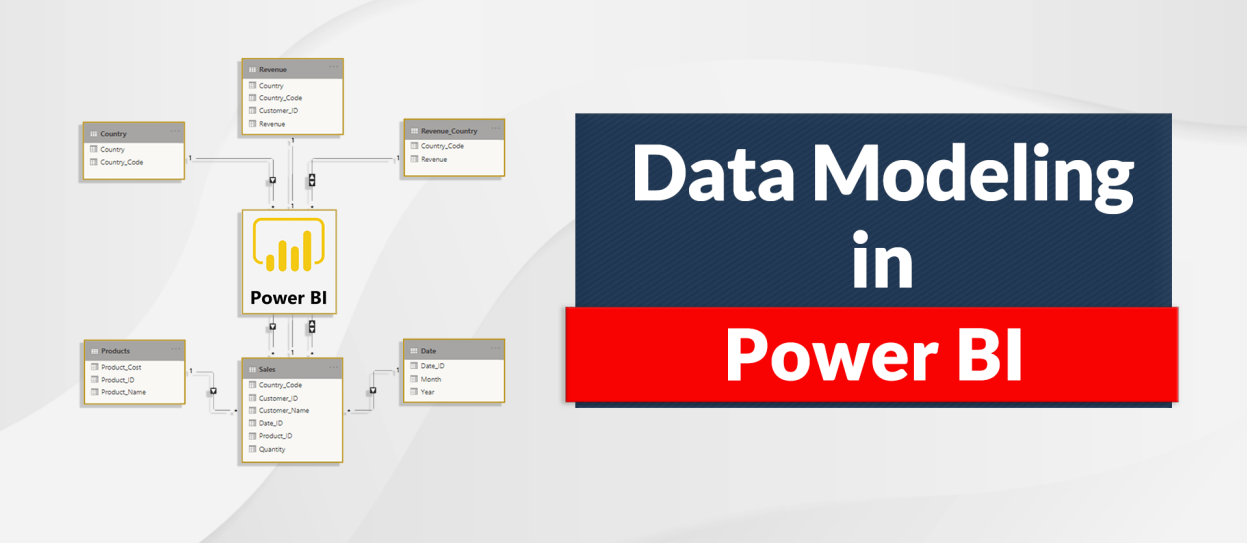
Data Modeling in Power Bi
(Relationships,Calculated Columns and Measures)

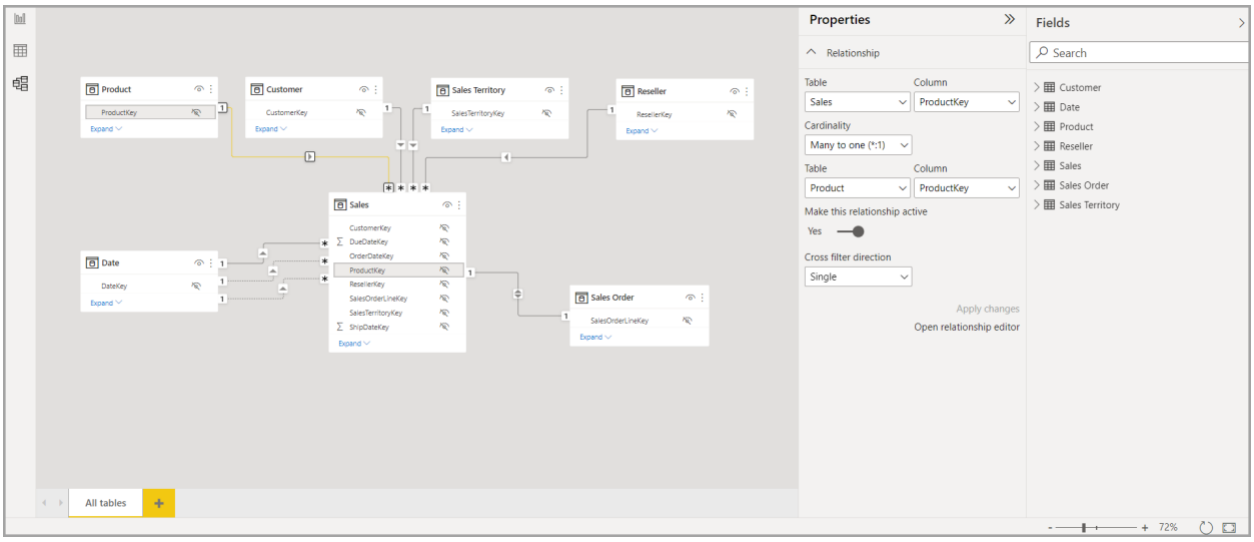
By Mayar ElAshmawy



1.Relationships:

- Relationships allow you to link tables in Power BI, propagate filters across tables which means that filters applied to one table can automatically apply to related tables and allow for cross table calculations.
- Power BI automatically guesses relationships based on column names, but customizing relationships can be manually done as well.

Mentorness Article
Task1



Relationships in Power BI are established based on **keys**, which are columns or sets of columns that uniquely identify rows in a table.
There are two primary types of keys:

1- Natural keys, which exist within the dataset and uniquely identify entities (e.g., email addresses)
2- Surrogate keys, which are artificially generated and incrementally increase for each row (e.g., an identity column).

Power BI requires single-column relationships, which can be challenging when dealing with surrogate keys. To overcome this limitation, we can utilize composite keys. Composite keys are created by concatenating multiple text fields, typically composed of natural keys. For instance, a composite key might combine the name and birth year to uniquely identify individuals.

Cardinality refers to the nature of relationships between tables in a data model. There are four cardinality types, each representing the characteristics of the columns involved in the relationship.

One-to-many (1:*): In this type, the column on the "one" side contains unique values, while the column on the "many" side can contain duplicate values.
Many-to-one (*:1): Here, the column on the "many" side contains unique values, and the column on the "one" side may have duplicate values.
One-to-one (1:1): Both columns involved in the relationship contain unique values.
Many-to-many (:): Both columns can contain duplicate values.
It's crucial to ensure that a "one" side column indeed contains unique values. If a data refresh operation attempts to load duplicate values into such a column, the entire refresh process will fail.

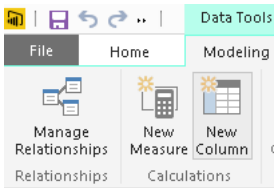
When creating relationships in Power BI Desktop, the software automatically detects and sets the cardinality type based on the data characteristics. However, there might be instances where Power BI Desktop misinterprets the cardinality type, especially when tables are yet to be loaded with data or when columns unexpectedly contain unique values. In such cases, you can manually update the cardinality type, provided that the "one" side columns contain unique values or the table is yet to be populated with rows of data.

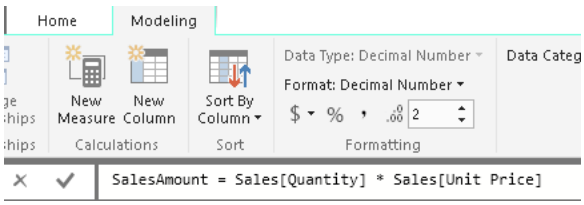
Calculated Columns:

- Calculated columns are a feature in Power Pivot for Excel, Analysis Services Tabular, and Power BI Desktop that allow users to extend tables by creating new columns based on DAX expressions. Here's a breakdown of key points regarding calculated columns:

- Creation Process: In Power BI Desktop, a new column is created by clicking the "New Column" button, and the column name is included in the formula textbox.

- User Interface: The user interface differs between tools, but the purpose remains the same: to create new columns based on DAX expressions.





- Functionality: Calculated columns are distinguished from native columns (those from the data source) and are used to extend tables in the data model.They behave like regular columns and can be utilized in reports, including defining relationships if necessary.

- Calculation Process: Calculated columns are computed during database processing and stored in the model, unlike SQL-computed columns that are computed at query time.This computation during database processing ensures a better user experience by reducing query time, but it also consumes memory.

Measures:

Measures, another method for defining calculations in a DAX model, are particularly useful when you want to aggregate values from multiple rows in a table rather than compute values for each row individually.
Implicit vs Explicit Measures:

Implicit Measures:
Automatically created by Power BI from the database.
Example: Dragging "Sales" to the values of a table, Power BI automatically sums it.
Explicit Measures:
Explicitly written by the user.
Example: Calculating "Total Sales" as the sum of all sales.

Why Explicit Measures are Preferred:

Reduced Confusion:
Implicit measures may lead to confusion as their exact meaning is unclear.

Explicit measures provide clarity as they are explicitly defined.

Reusability:
Explicit measures can be reused in other measures, enhancing efficiency and reducing redundancy.
Custom names can be given to explicit measures to explain their functionality.
Best Practices

Organized Structure:
Group DAX measures together in a separate table for better organization.
Unlike calculated columns, measures can be moved to any table, so keeping them together improves manageability**.**

Conclusion:

Calculated Columns	Measures
1- Evaluate at a row level 2- Add a new column to an existing table. 3- Calculated during data load or refresh.	1- Enable creation of complex calculations for data analysis. 2- Aggregate multiple rows and add a new field to visualizations. 3- Calculated at query time, dynamically updating as you interact and filter data. 4- More efficient as calculations are only performed when the measure is used in visualizations.