Data Analytics using Power BI

Vaidehi Takalikar

Corporate Trainer | Tableau | Python | Power BI

+91 - 9664353290

vaidehi89@gmail.com

bit.ly/vaidehi_Takalikar

www.linkedin.com/in/vaidehitakalikar

Course Overview

- Power BI is a cloud-based business analytics service that helps create live operational dashboards from on-premises and cloud data in one central location that you can access across a range of devices.
- Power BI helps you stay up to date with the information that matters to you. You can connect to multiple datasets to bring all of the relevant data together in one place.
- With Power BI, dashboards help you keep a finger on the pulse of your business. Your dashboards display tiles that you can click to explore further with reports.
- Power BI Desktop puts visual analytics at your fingertips with intuitive report authoring. You can drag-and-drop to place content exactly where you want it on the flexible and fluid canvas.

Course Takeaway

- Master concepts like Data Visualization and Integration
- Learn about Power Pivot, Power View, Power Map etc.
- Understand DAX
- Use and implement Custom Visuals
- Explore Power BI Q&A
- Perform Data Binding and Formatting
- Master Power BI Embedded and Workspace Collection
- Understand and use Data Gateways, Content packs, Power BI Report Server

Course Contents

- **Introduction to Data Analytics**
  - Overview of Data Analysis
  - Roles in Data
  - Tasks of Data Analyst
  - Introduction to Power BI
  - Components of Power BI
- **Prepare data for analysis**
  - Get data into Power Bi from various data connections
  - Clean, transform, and load data in Power BI

Course Contents

- **Model data in Power BI**
  - Design a data model in Power BI
  - Creating measures using DAX in Power BI
  - Optimize a model for performance in Power BI
- **Visualize data in Power BI**
  - Creating visuals using Report view
  - Create a data-driven story with Power BI reports
  - Create dashboards in Power BI
  - Create paginated reports

Course Contents

- **Data analysis in Power BI**
  - Explore statistical summary - Identify outliers with Power BI visuals
  - Group and bin data for analysis, Apply clustering techniques
  - Conduct time series analysis
  - Use the Analyze feature
  - Work with AI visuals in Power BI
- **Manage workspaces and datasets in Power BI**
  - Distribute a report or dashboard, Monitor usage and performance
  - Recommend a development life cycle strategy, Troubleshoot data by viewing its lineage
  - Configure data protection
  - Manage datasets in Power BI
  - Implement row-level security

Introduction to Data Analytics

Learning Objective –

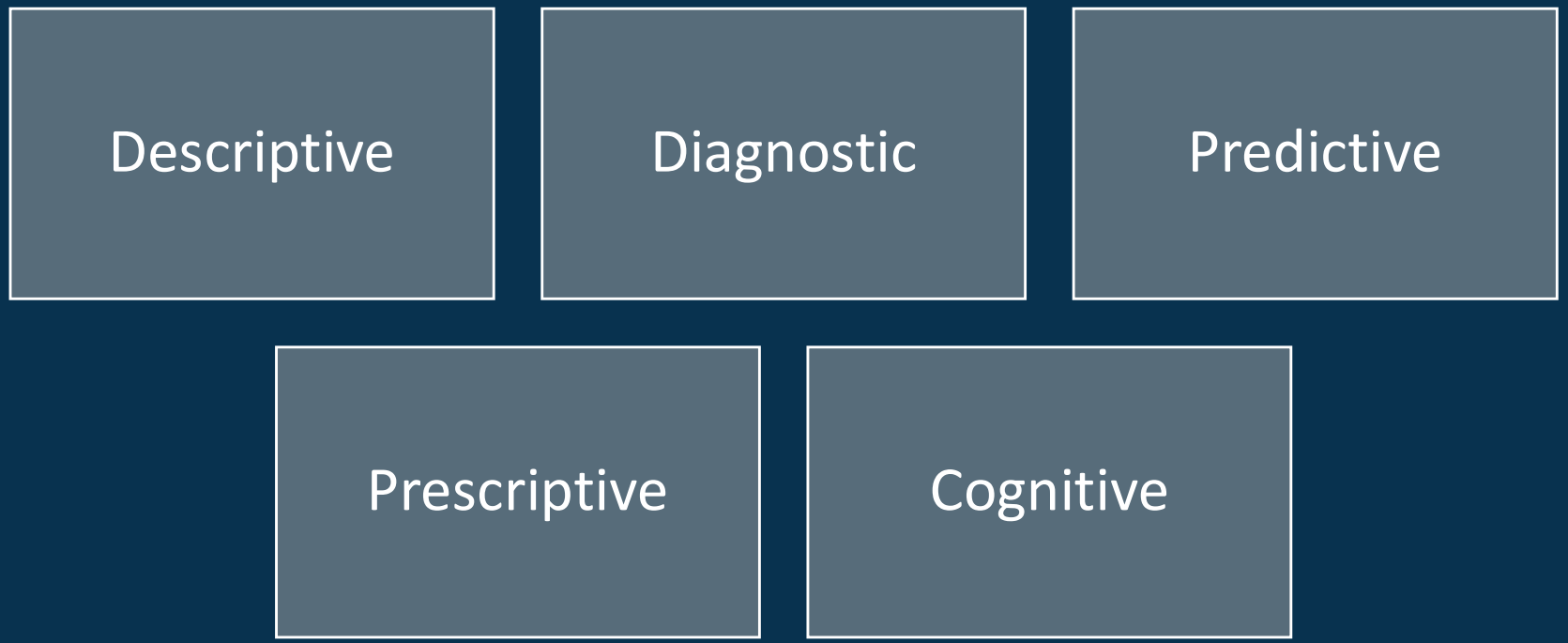
- Overview of Data Analysis
- Visual Data Analytics
- Getting familiarized with Power BI Desktop Software
- Power BI Components
- Steps to build a Dashboard

Overview of Data Analysis

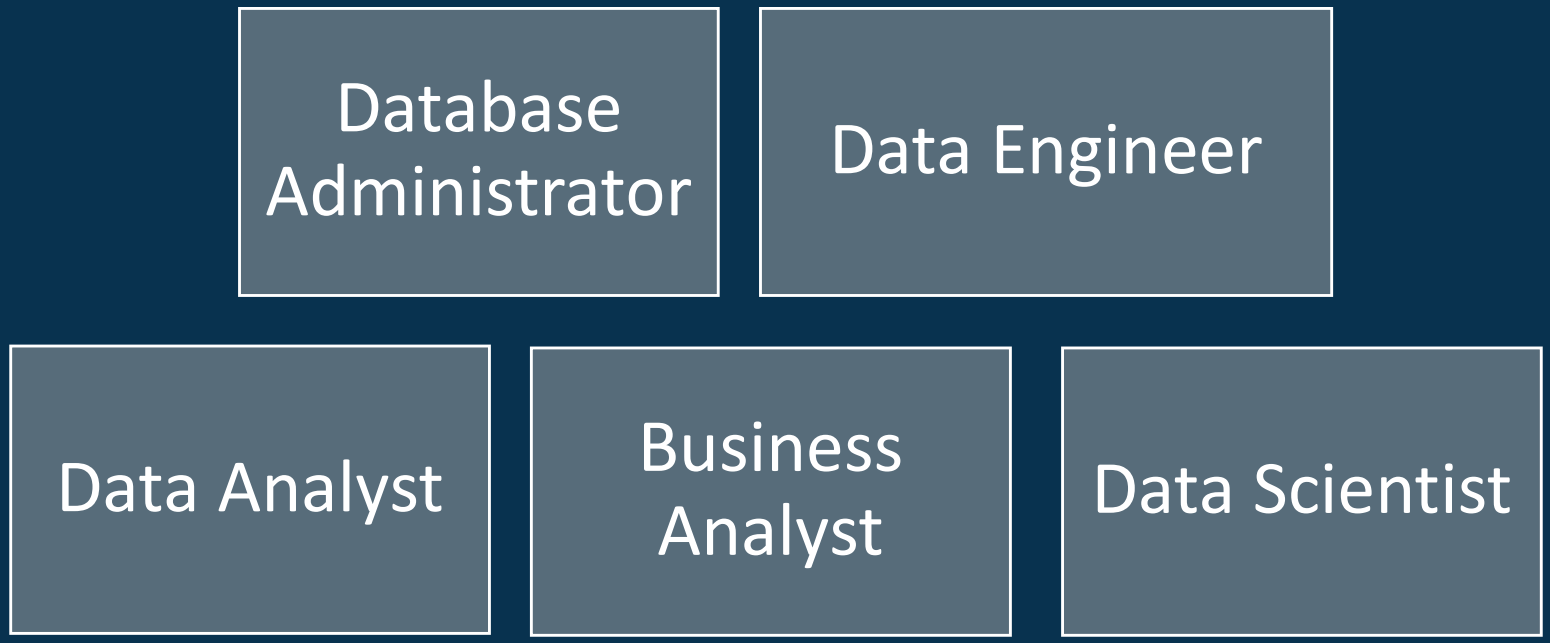
Overview of Data Analysis

- Before data can be used to tell a story, it must be run through a process that makes it usable in the story.
- Data analysis is the process of identifying, cleaning, transforming, and modeling data to discover meaningful and useful information.
- The data is then crafted into a story through reports for analysis to support the critical decision-making process.

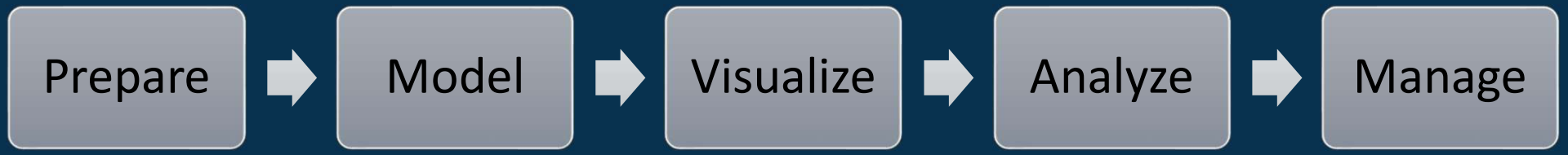
Data Analytics Categories



Roles in Data



Tasks of a Data Analyst



Visual Data Analytics

Data Visualization

› Visualization allows visual access to huge amounts of data in easily understandable visuals

› Data Visualization involve the graphical representation of the data in the form of charts, graph, and maps which make the analysis of the complex data easier to the user.

Visualization is used for –

Getting familiarized with Power BI Desktop Software

Power BI Installation

- Which version to install?
  - For a 32 bit machine, you need to install the 32bit Power BI Desktop
  - For a 32 bit Office installed (regardless of your machine), you need to install the 32 bit Power BI Desktop
  - Otherwise, you can install the 64 bit PBI Desktop
- Minimum requirements
  - Windows 7 / Windows Server 2008 R2, or later
  - .NET 4.5
  - Internet Explorer 10 or later

What is Power BI?

- It is a BI tool
- Reports and dashboards can be easily created using Power BI
- Data cleaning and data modeling can be achieved using inbuilt tool - Power Query
- It helps in visualizing data

Power BI Products

Power BI Desktop

Power BI Pro

Power BI Premium

Power BI Mobile

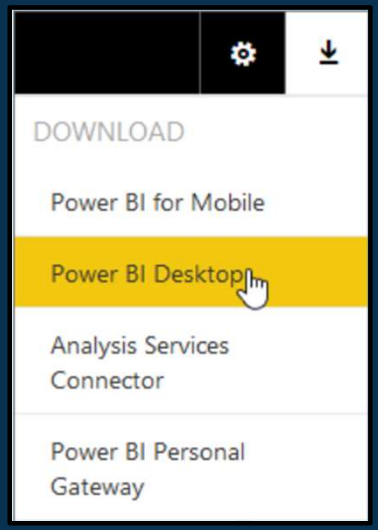
Power BI Embedded

Power BI Report Server

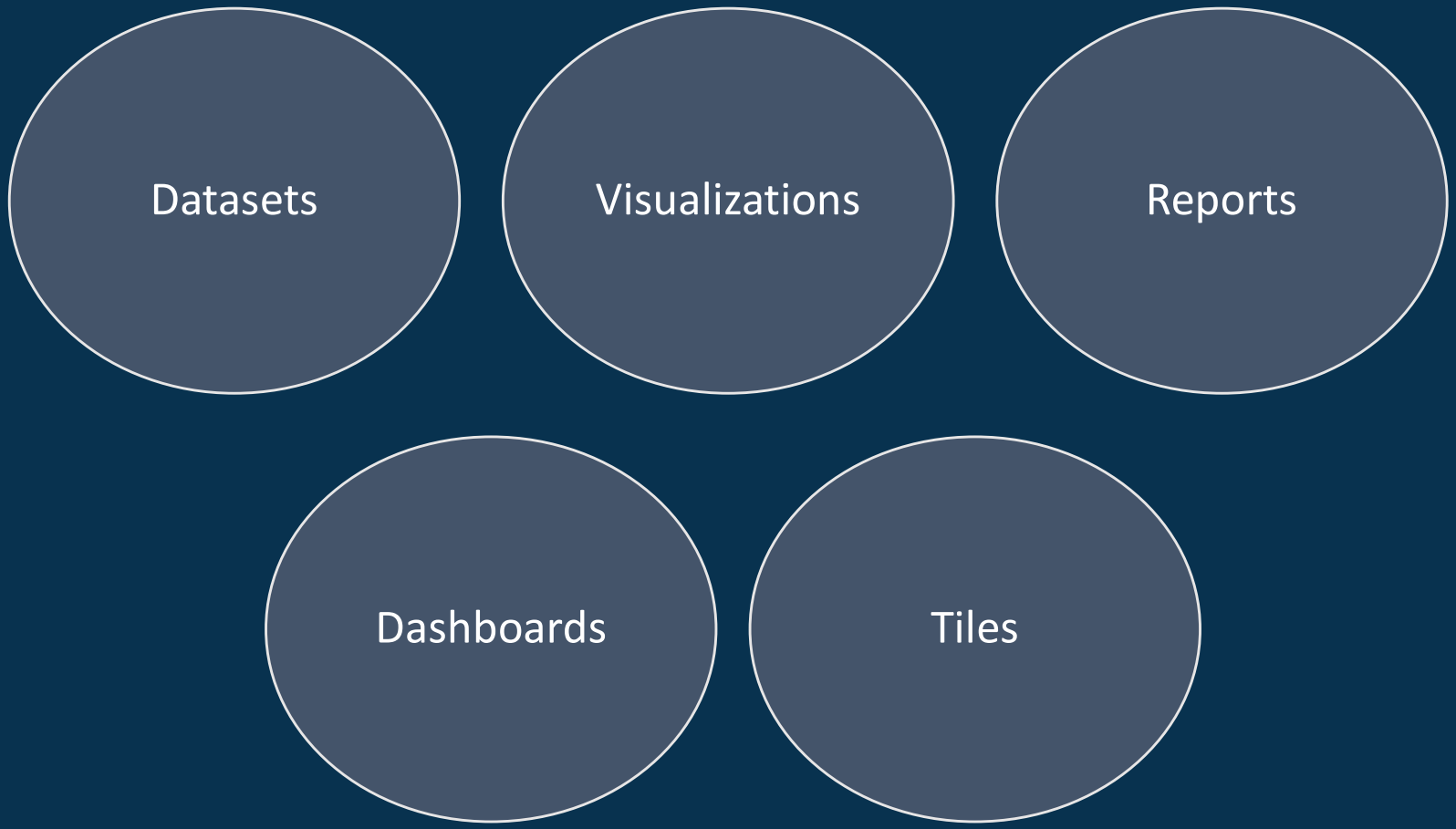
Power BI Installation

- Download and install Power BI Desktop
- Download and install the latest version of Power BI Desktop from

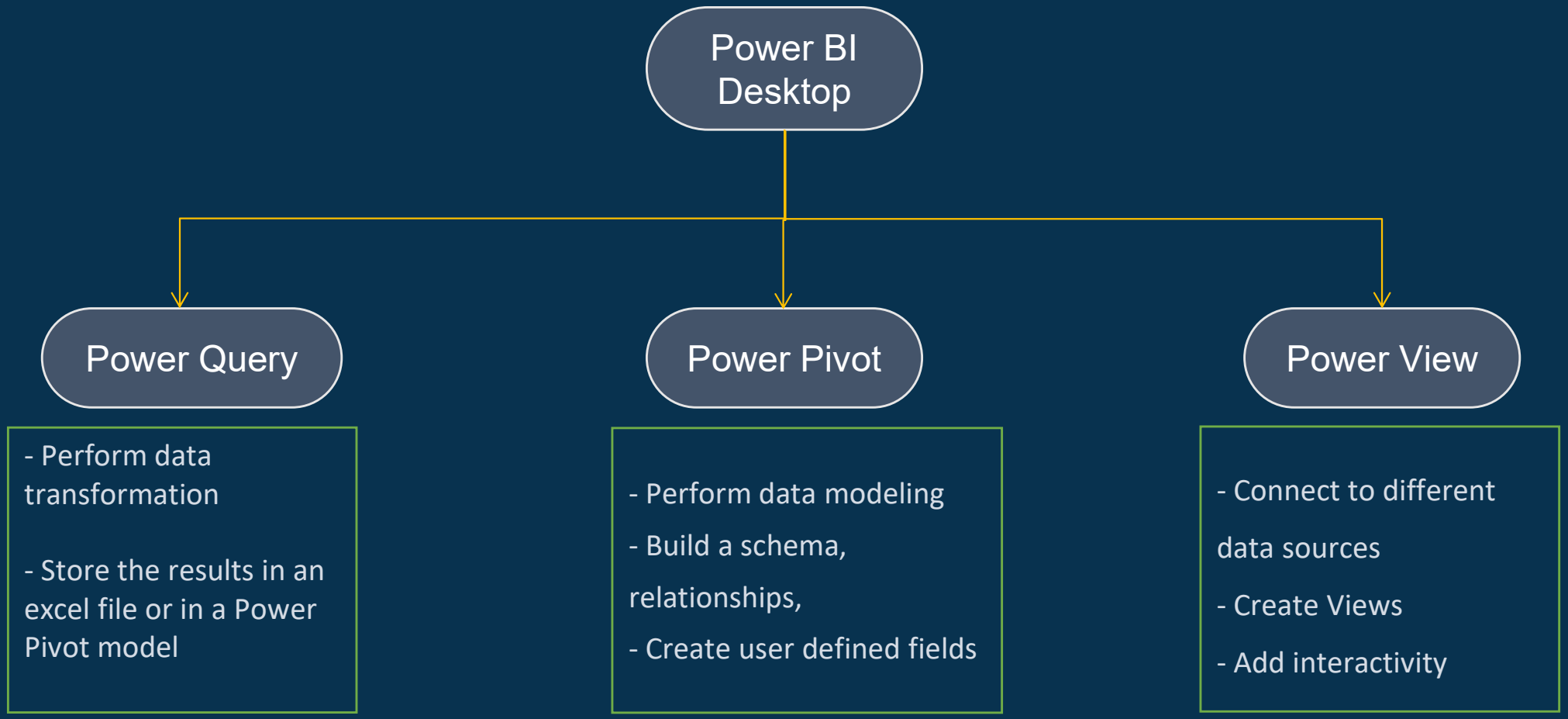
  https://powerbi.microsoft.com/en-us/desktop/?wt.mc_id=DXLEX_EDX_DAT207X
- To select which version to download, click Advanced download options
- Or, from Power BI, in Power BI, click the Downloads > Power BI Desktop.



Building Blocks of Power BI



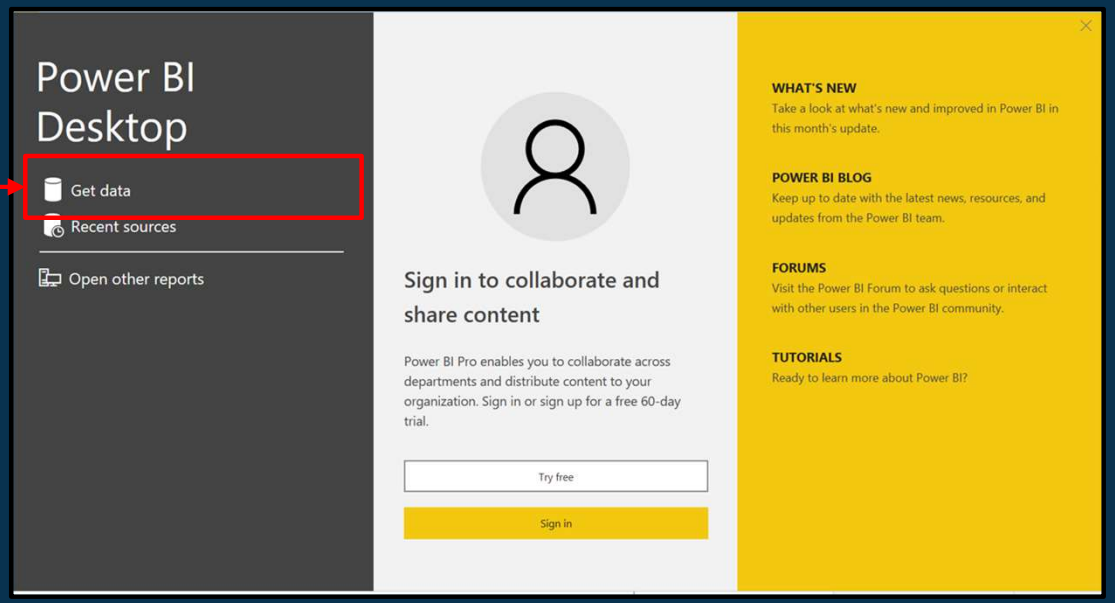
Power BI Components



Steps to Build a Dashboard

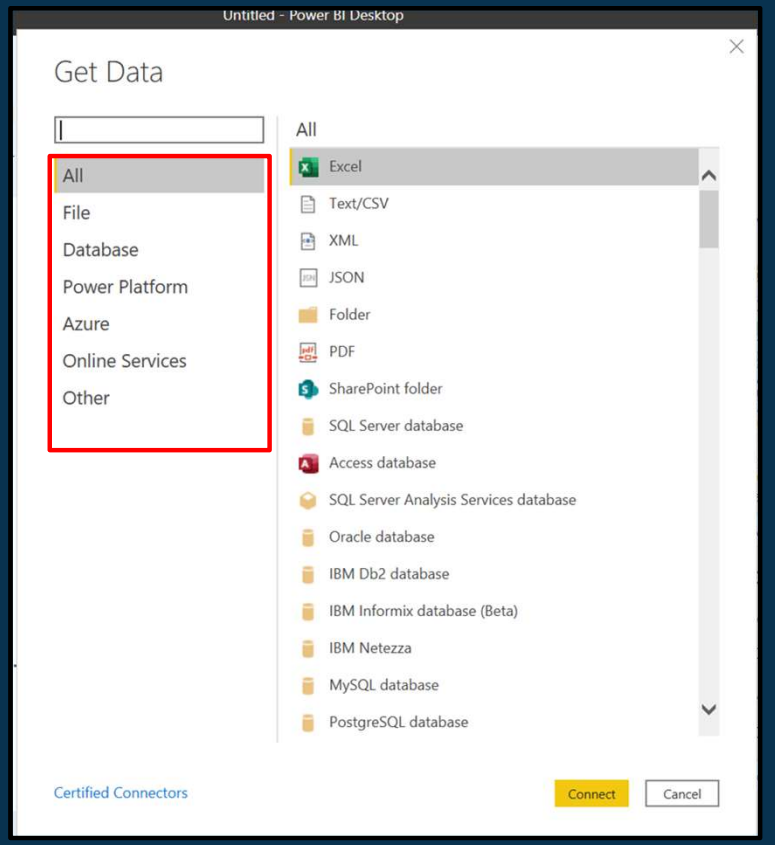
Data Connection

Connect to data sources



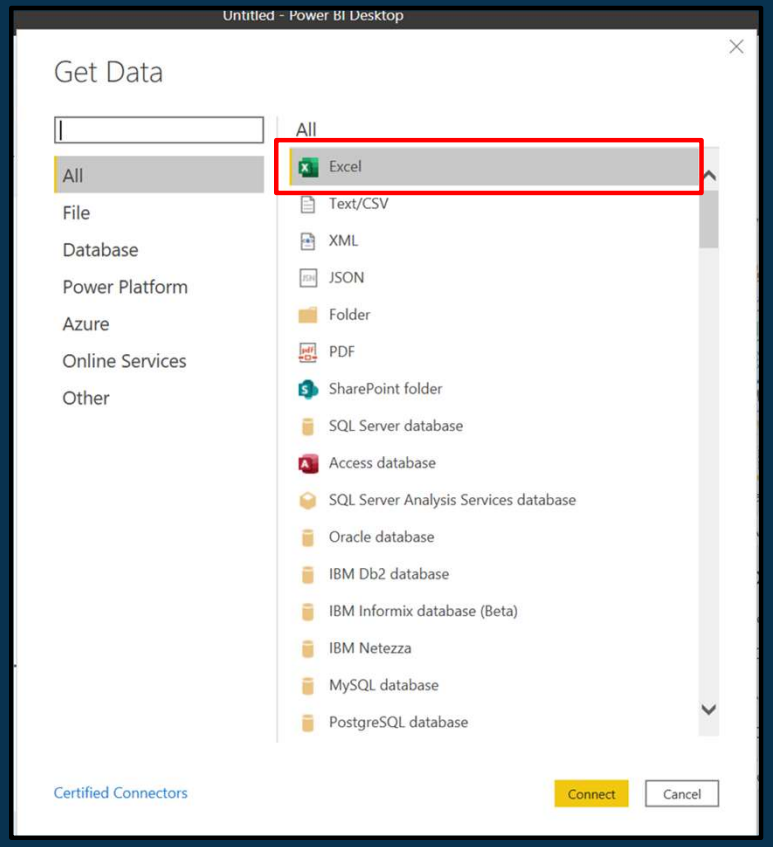
Data Connection

- Power BI supports 99+ data connection including file sources, database, cloud based
- It also supports extracting data from website



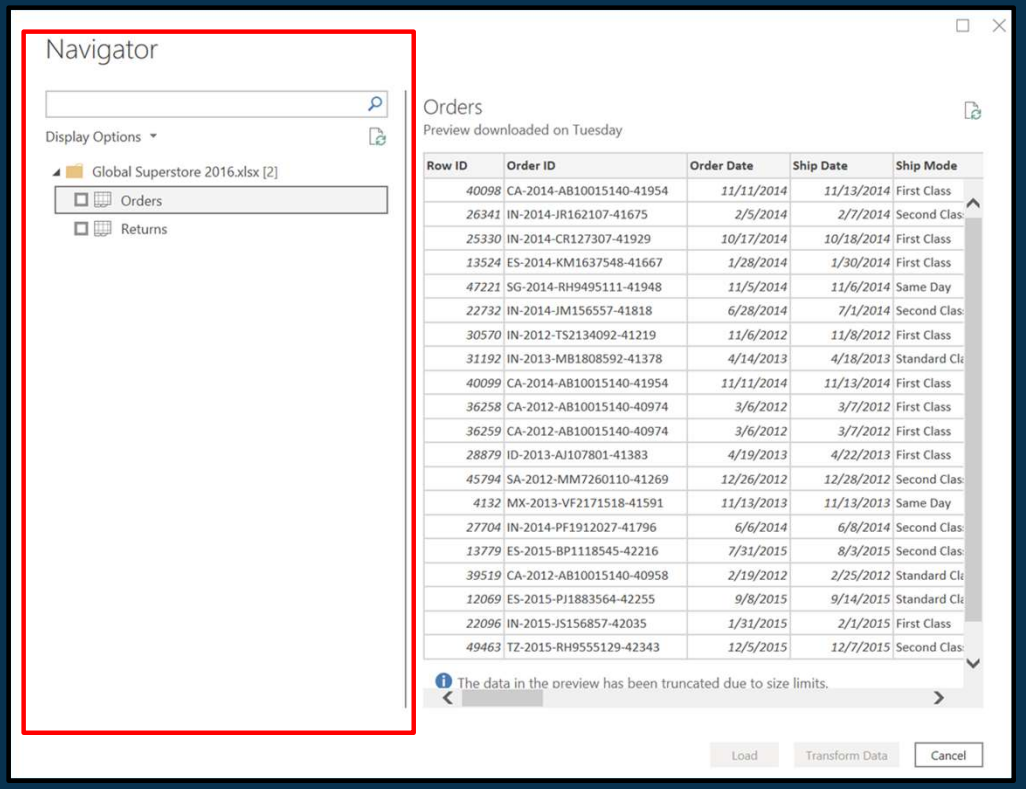
Data Connection - Example

- Select “Excel” and click on “Connect”
- Choose the file from your local machine



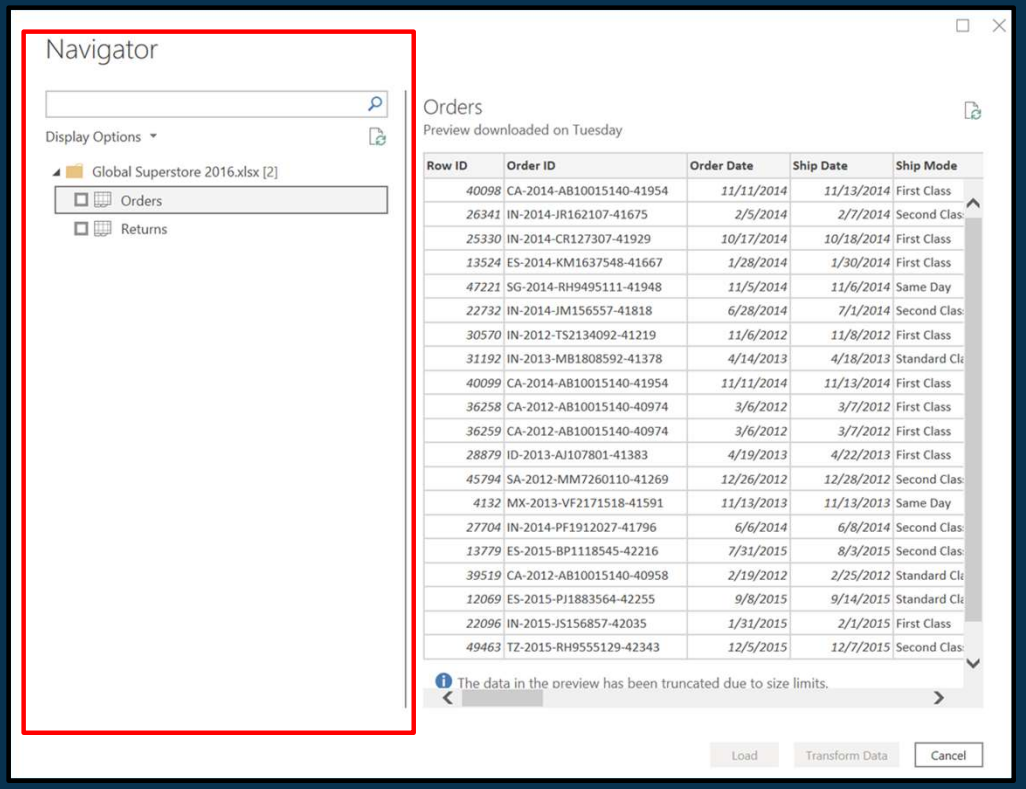
Data Connection - Example

- Navigator showing list of all the sheets in the data connection
- Select a field to view data preview on right side
- Click on the check box next to sheet name and click on “Load”



Data Connection - Example

- Navigator showing list of all the sheets in the data connection
- Select a field to view data preview on right side
- Click on the check box next to sheet name and click on “Load”



Build a Visualization

To create a basic bar chart -

1. Select the desired fields from the “Fields” panel

2. Select the chart type from “visualization” panel

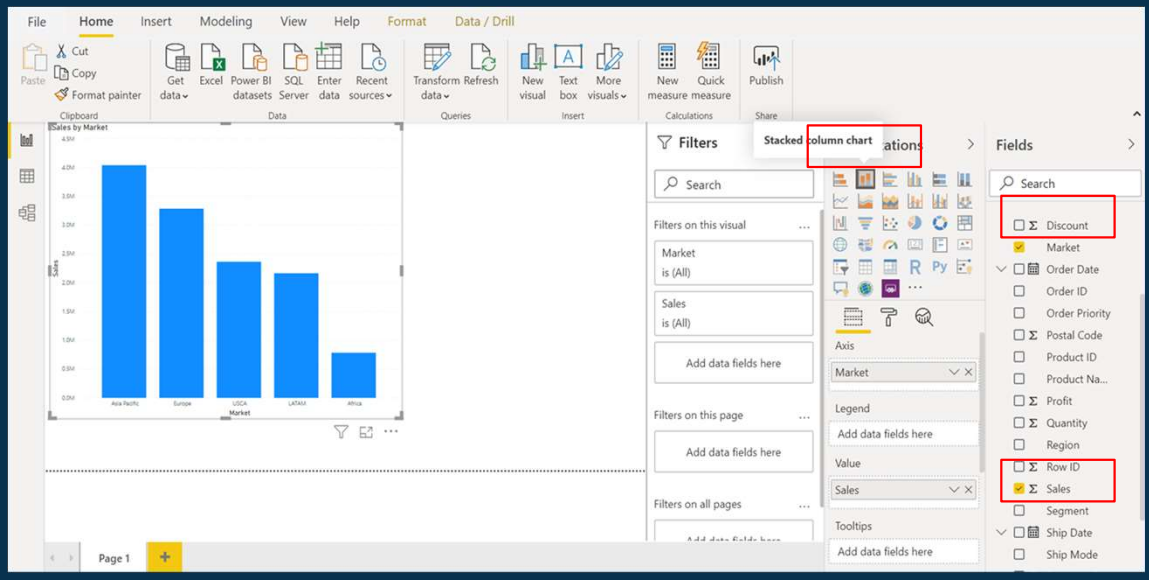
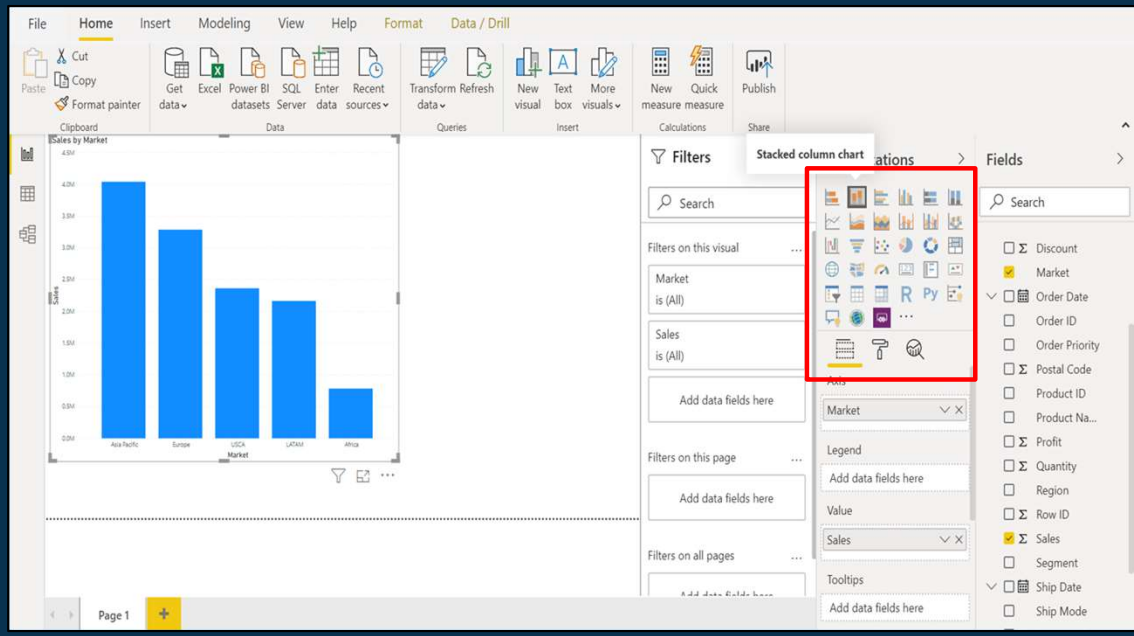
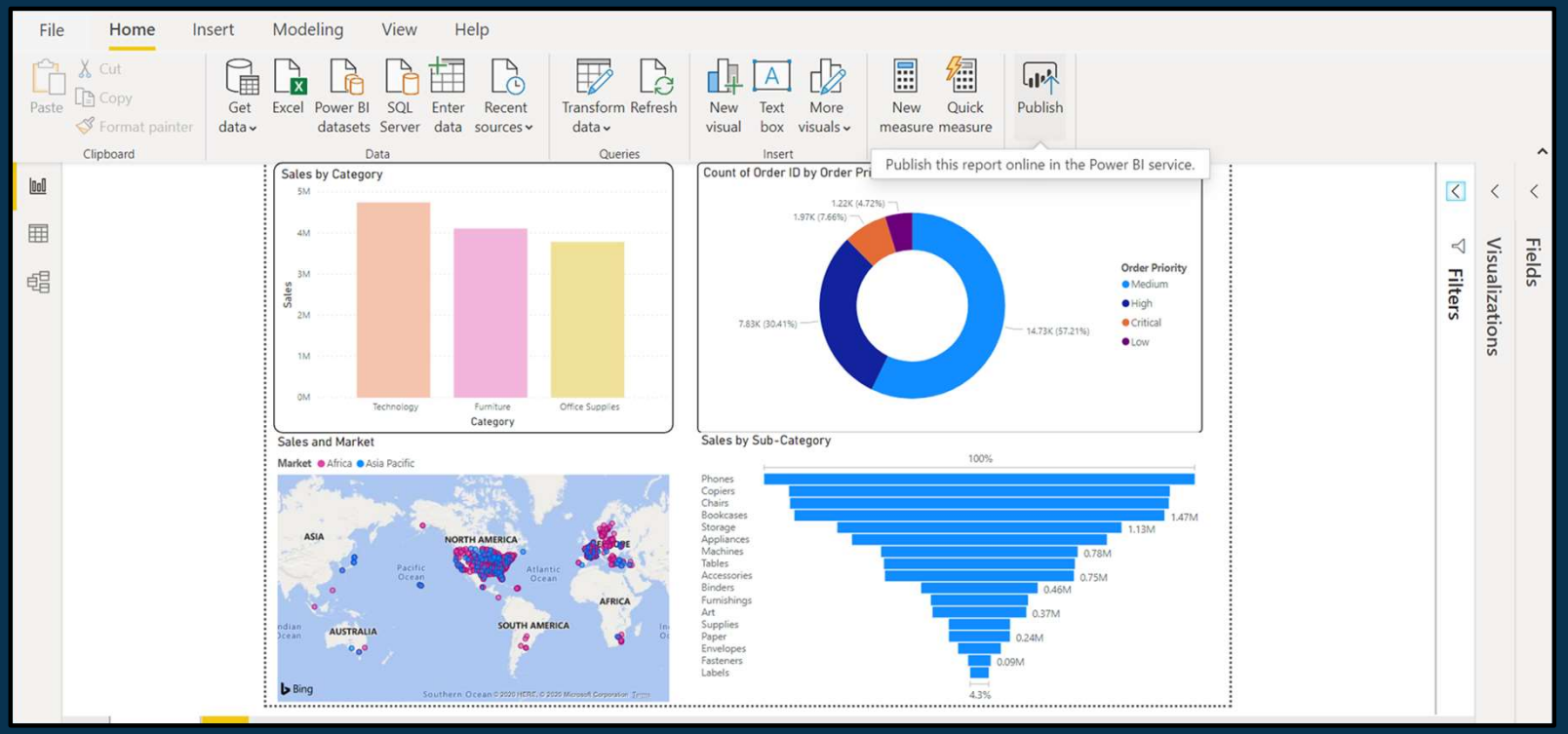


Chart options

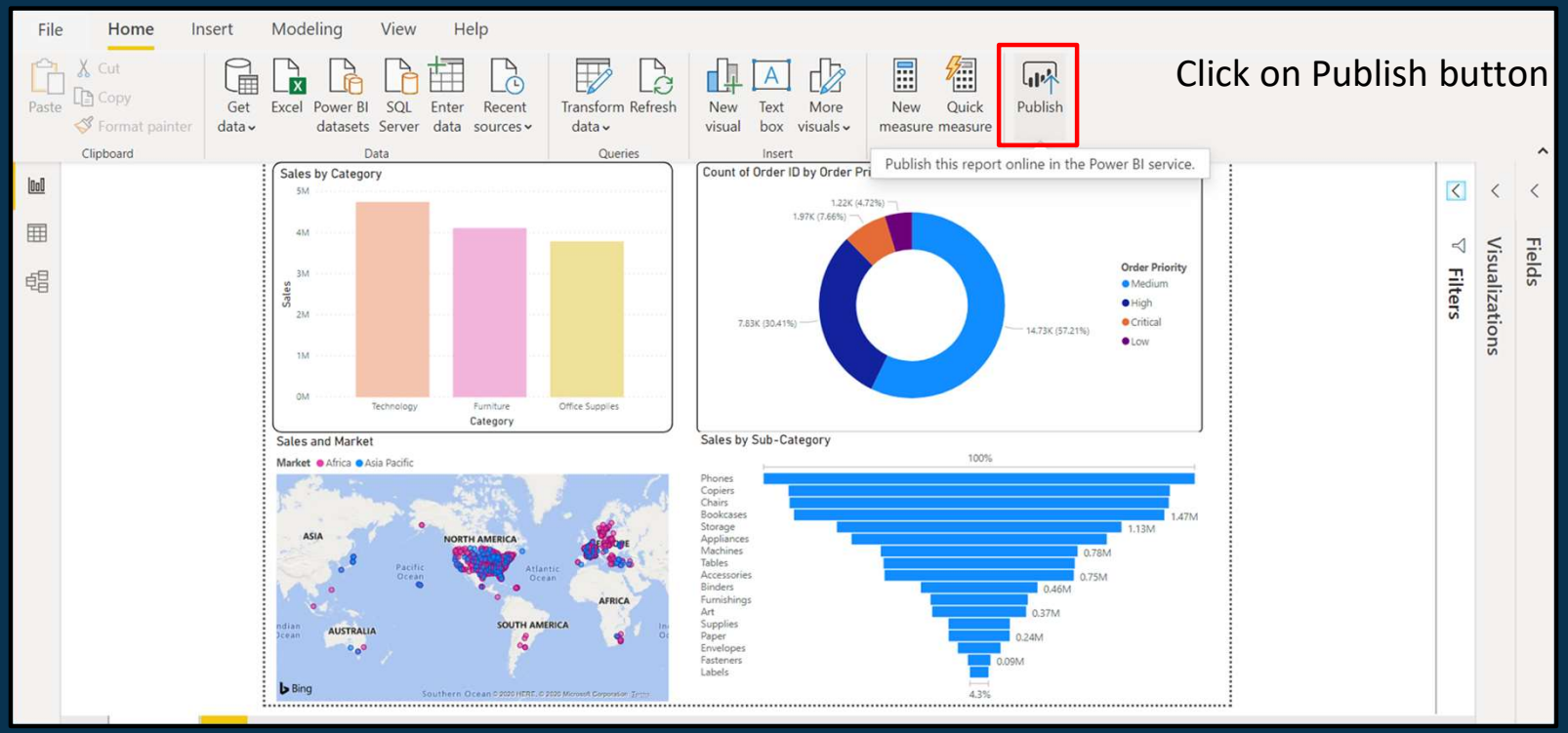
- Chart options helps to control chart types
- Power BI supports 30+ default charts and a library of 250 charts from marketplace
- 80 charts types are freely available and others can be purchased
- Trial versions are also available for the licensed chart types



Creating a dashboard



Publishing a dashboard



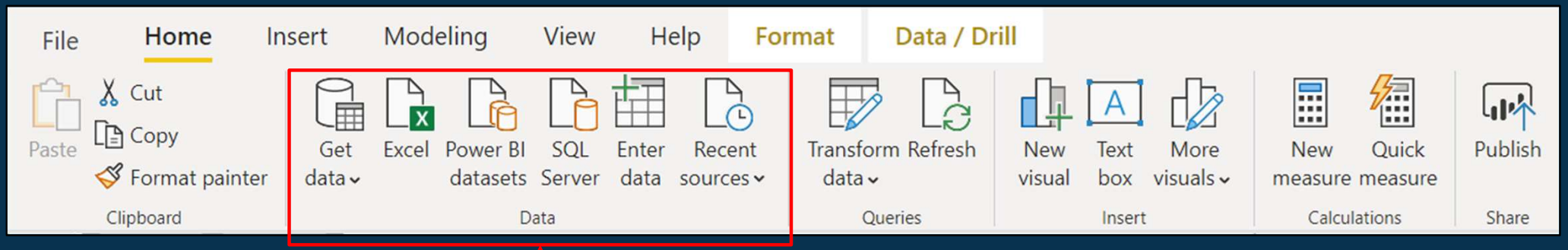
Power BI Interface and Basic Chart types

Learning Objective –

- Power BI Desktop Interface
- Report View – Data View – Model View
- Basic Chart Types – Bar chart, Line chart, Pie Chart, Table chart, Card
- Customize Workbook Theme

Power BI Interface

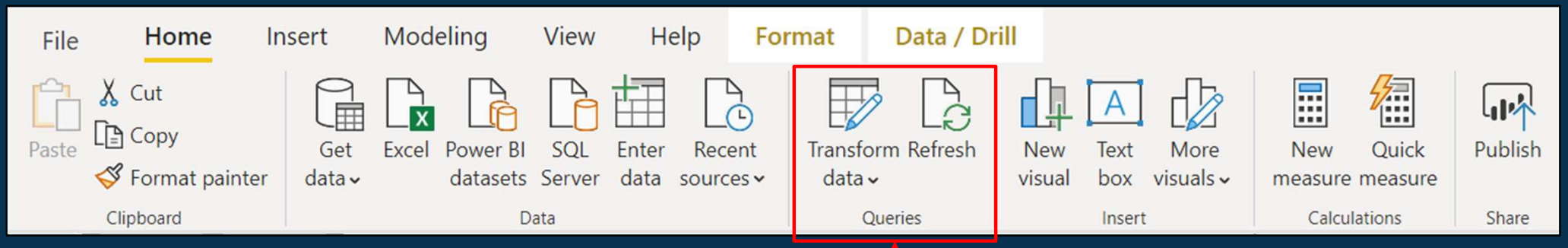
Home Tab - Data Ribbon



Data Connection options

Get Data window, allows to scroll through the list of All data sources

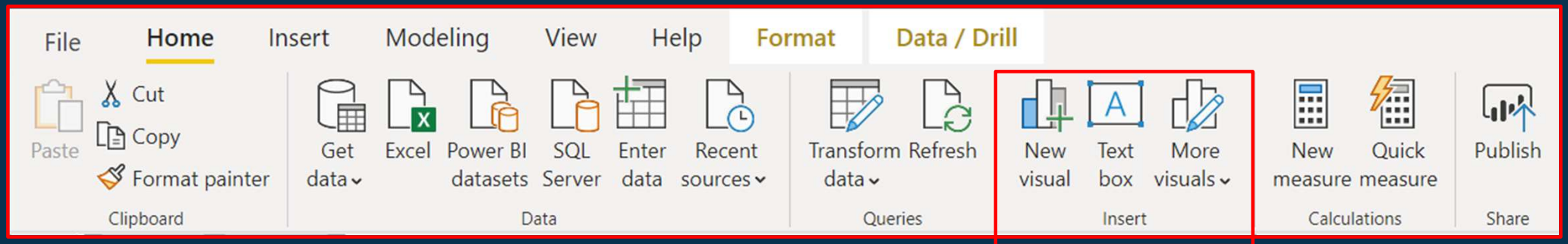
Home Tab - Queries Ribbon



Data Transformation and edit queries

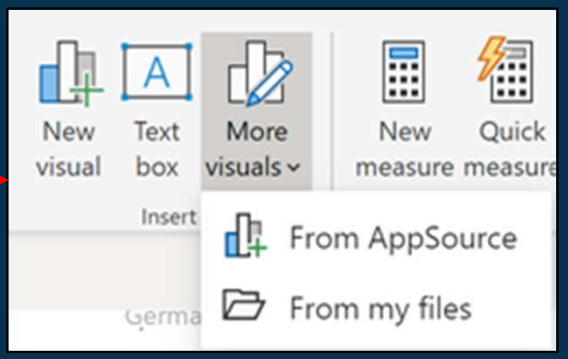
- Provides access to Power Query Editor.
- It helps shaping data.
- Shaping can mean transforming the data, such as renaming columns or tables, removing rows or columns, or changing data types

Home Tab - Insert Ribbon

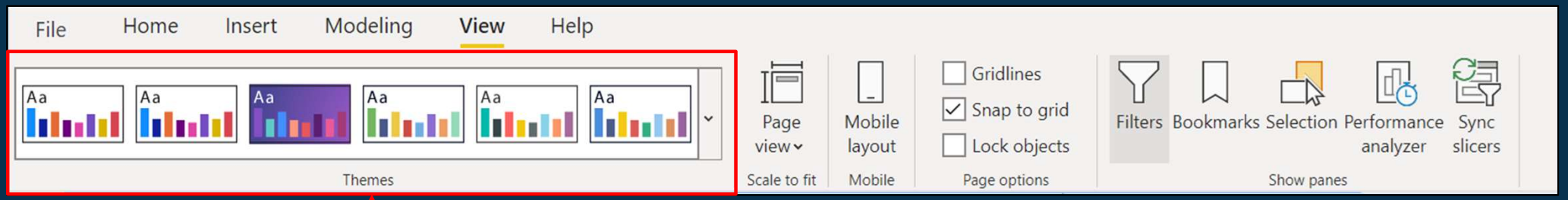


Helps add new visuals

Custom visuals are hosted in AppSource and can be accessed by pressing the “More visuals” dropdown and selecting “From AppSource”



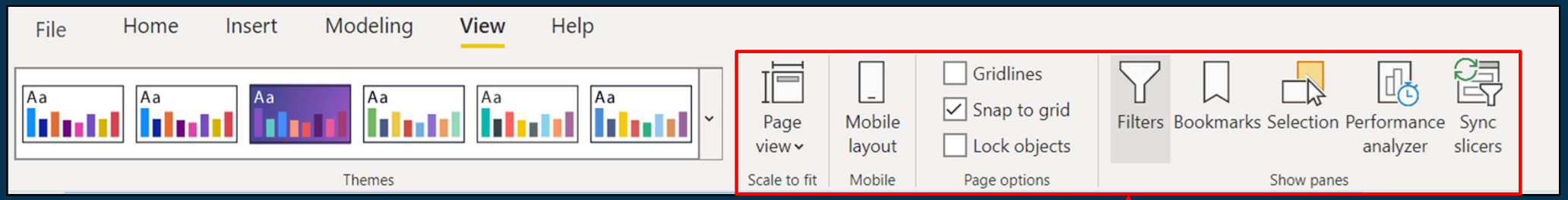
View Tab - Themes Ribbon



Themes ribbon displays themes in a gallery making it more easily to see what colours will be applied to report.

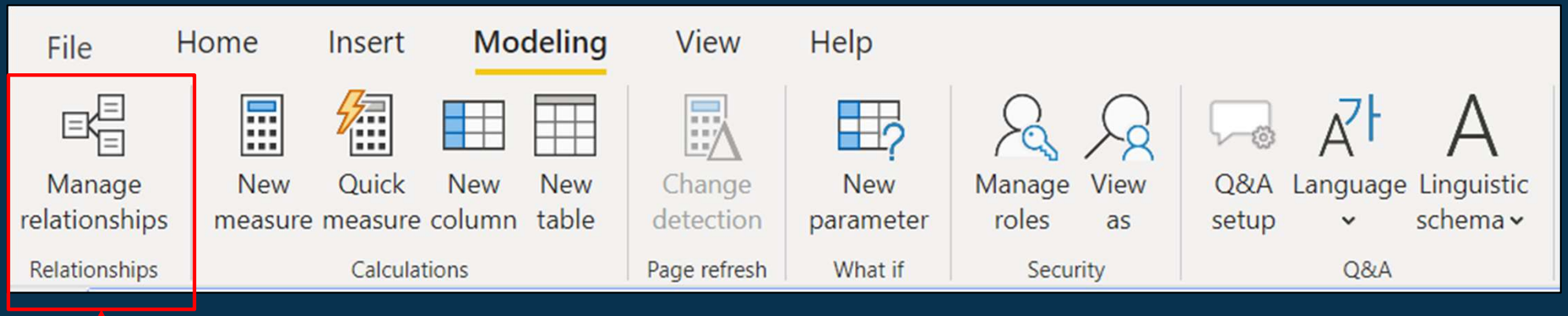
Also allows user to set custom themes

View Tab



Various Page visibility options and hide/show panes on report view tab

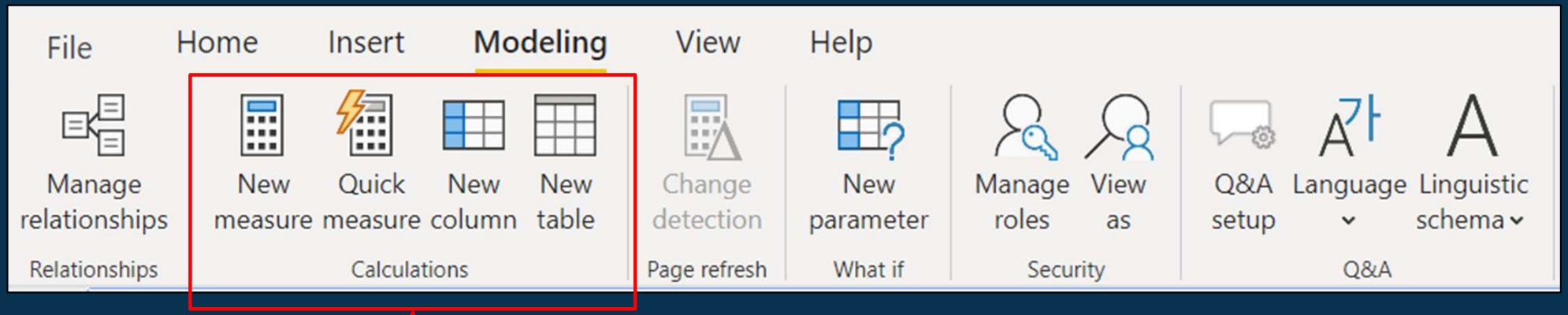
Modelling Tab - Relationships Ribbon



Opens Data Model view

Relationships between tables can be created and managed here.

Modelling Tab - Calculations Ribbon



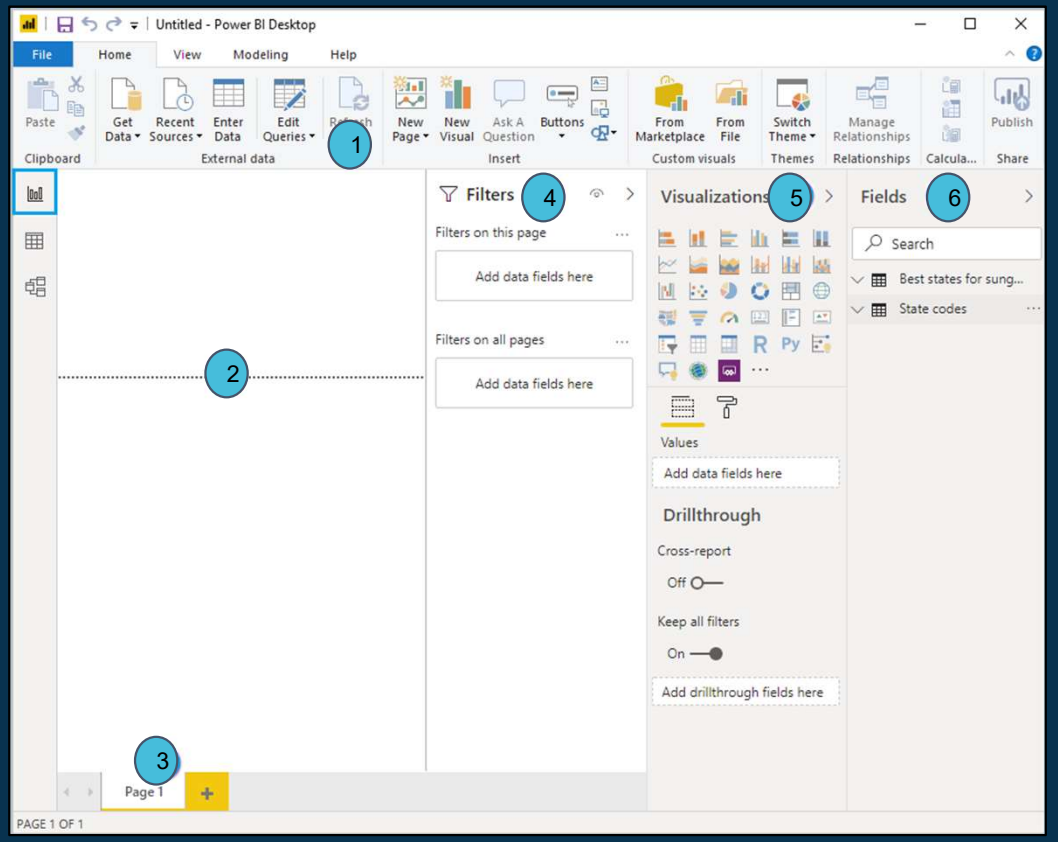
Allows user to create new fields by writing DAX expressions.

Report View – Data View – Model View

Report View

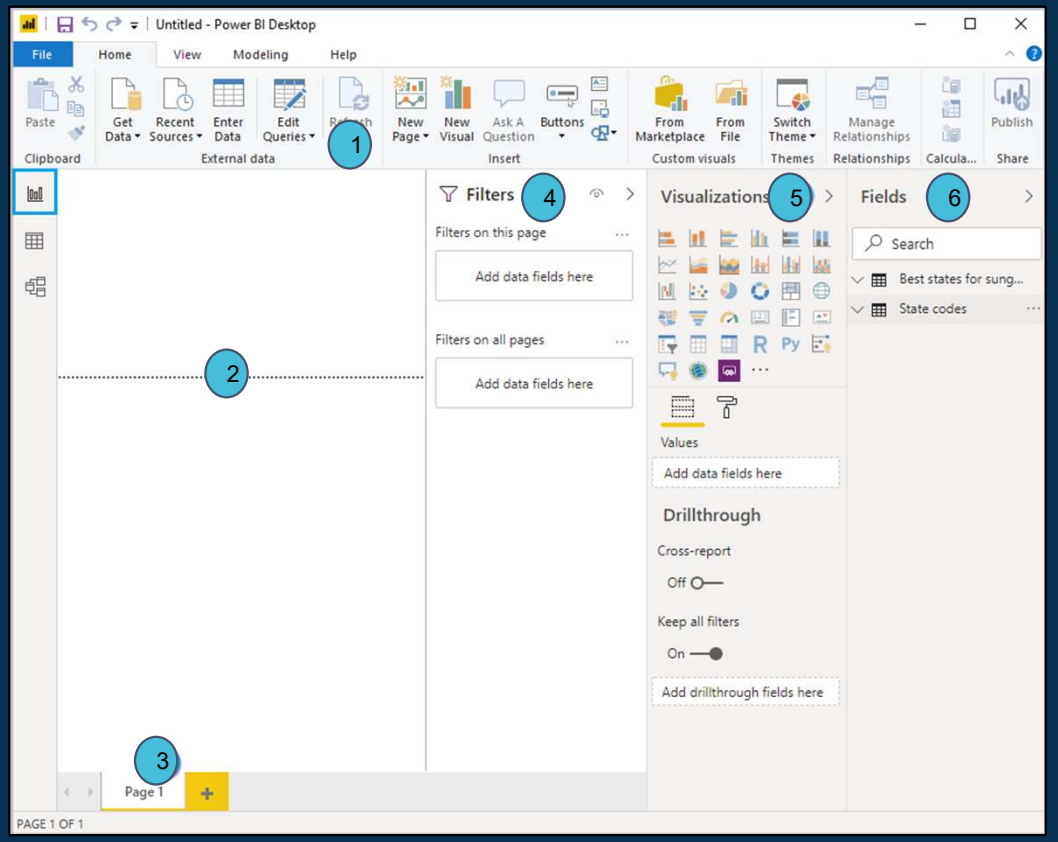
- In Power BI Desktop Report view, user can build visualizations and reports.

1. The ribbon at the top, which displays common tasks associated with reports and visualizations.
2. The canvas area in the middle, where visualizations are created and arranged.
3. The pages tab area at the bottom, which lets you select or add report pages.



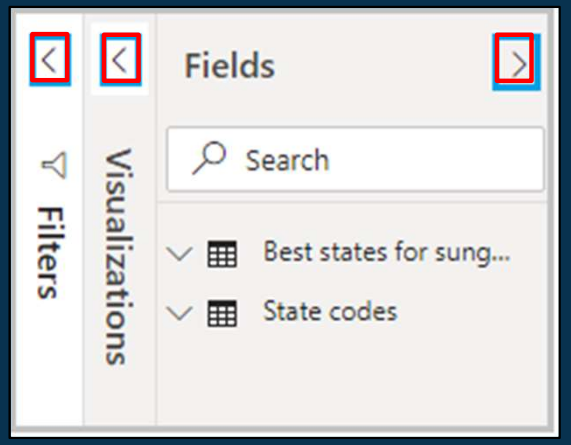
Report View

4. The Filters pane, where you can filter data visualizations.
5. The Visualizations pane, where you can add, change, or customize visualizations, and apply drill through.
6. The Fields pane, which shows the available fields in your queries. You can drag these fields onto the canvas, the Filters pane, or the Visualizations pane to create or modify visualizations.



Report View

- To expand and collapse the Filters, Visualizations, and Fields panes click the arrows at the tops of the panes.
- Collapsing the panes provides more space on the canvas to build cool visualizations.

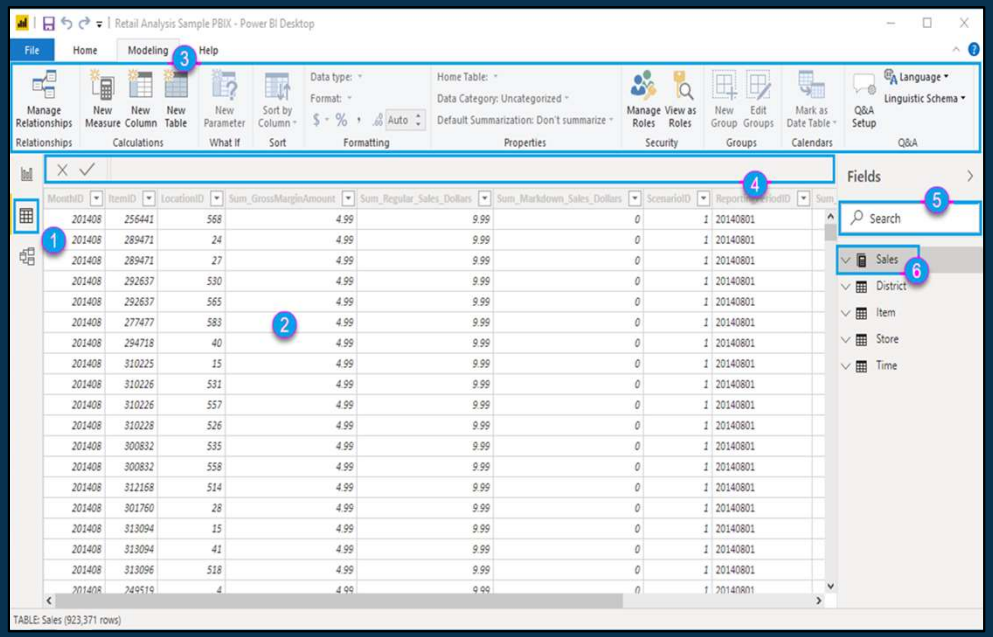


Data View

- Data view helps inspect, explore, and understand data in your Power BI Desktop model.
- It's different from tables, columns, and data in Power Query Editor.
- With Data view, user can look into data after it has been loaded into the model.
- This ability is especially useful when creating measures and calculated columns, or need to identify a data type or data category.

Data View

1. Data view icon - Select this icon to enter Data view.
2. Data Grid - This area shows the selected table and all columns and rows in it. Columns hidden from Report view are greyed out. Right-click on a column for options.
3. Modeling ribbon - Tab to manage relationships, create calculations, change data type, format, data category for a column.

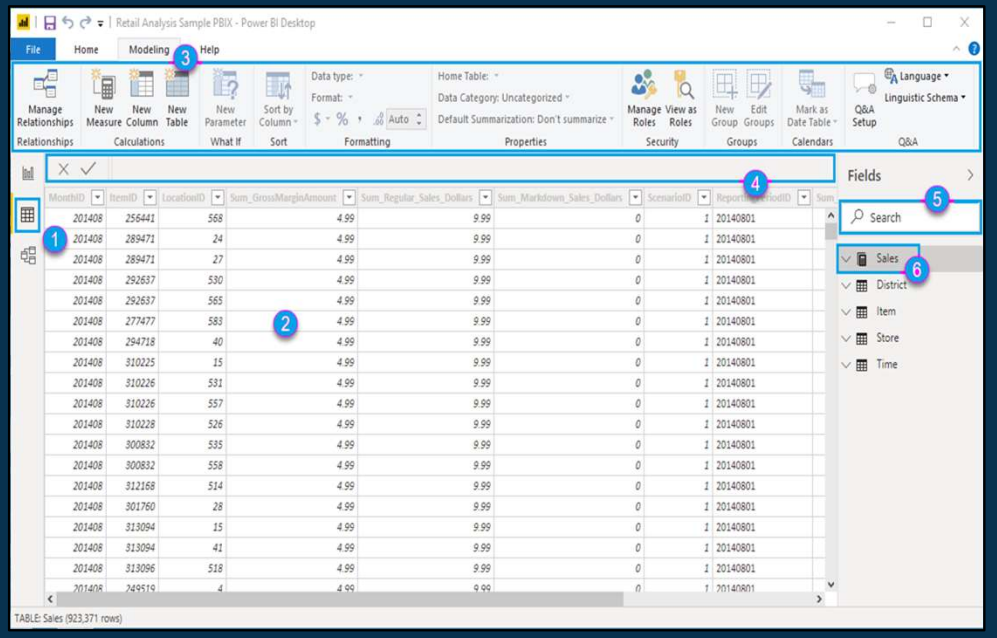


Data View

4. Formula bar - Enter Data Analysis Expression (DAX) formulas for Measures and Calculated columns.

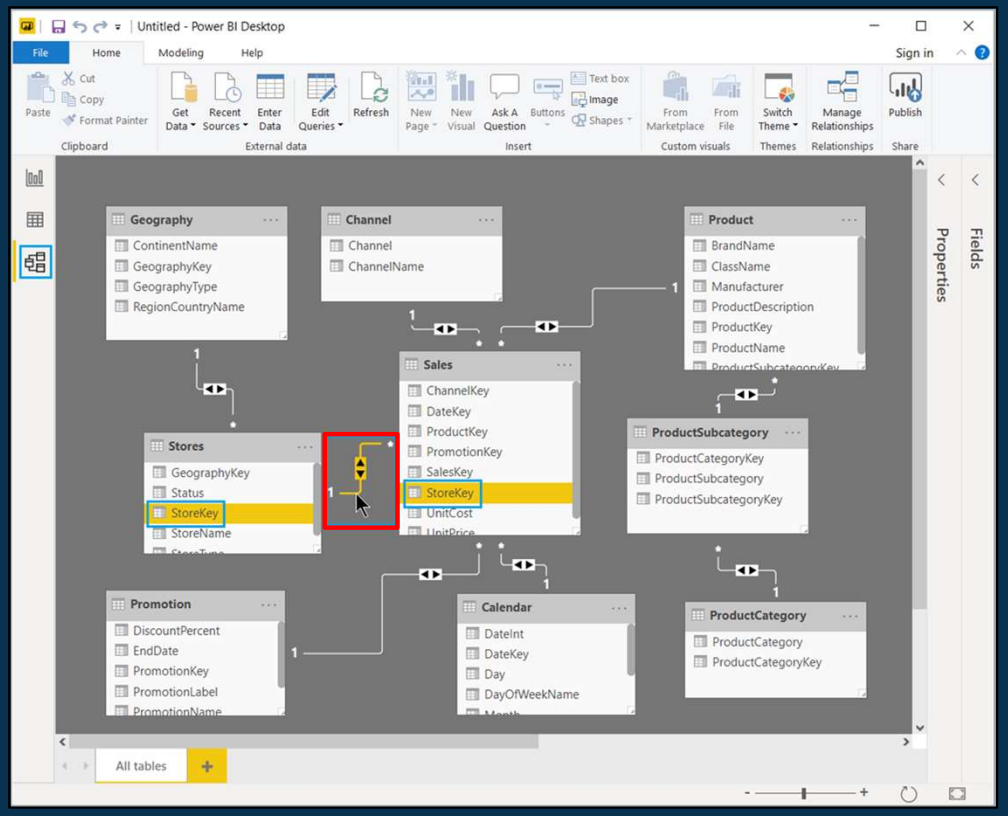
5. Search - Search for a table or column in model.

6. Fields list - Select a table or column to view in the data grid.



Model View

- Model view shows all of the tables, columns, and relationships in data model.
- This view can be especially helpful when model has complex relationships between many tables.
- Select the Model icon near the side of the window to see a view of the existing model.



Basic Chart Types

Bar and Column Charts

- Power BI Desktop has a variety of bar and column chart visualizations that present specific data across different categories in a stacked or clustered format.
- The stacked format will stack the information items on top of each other.

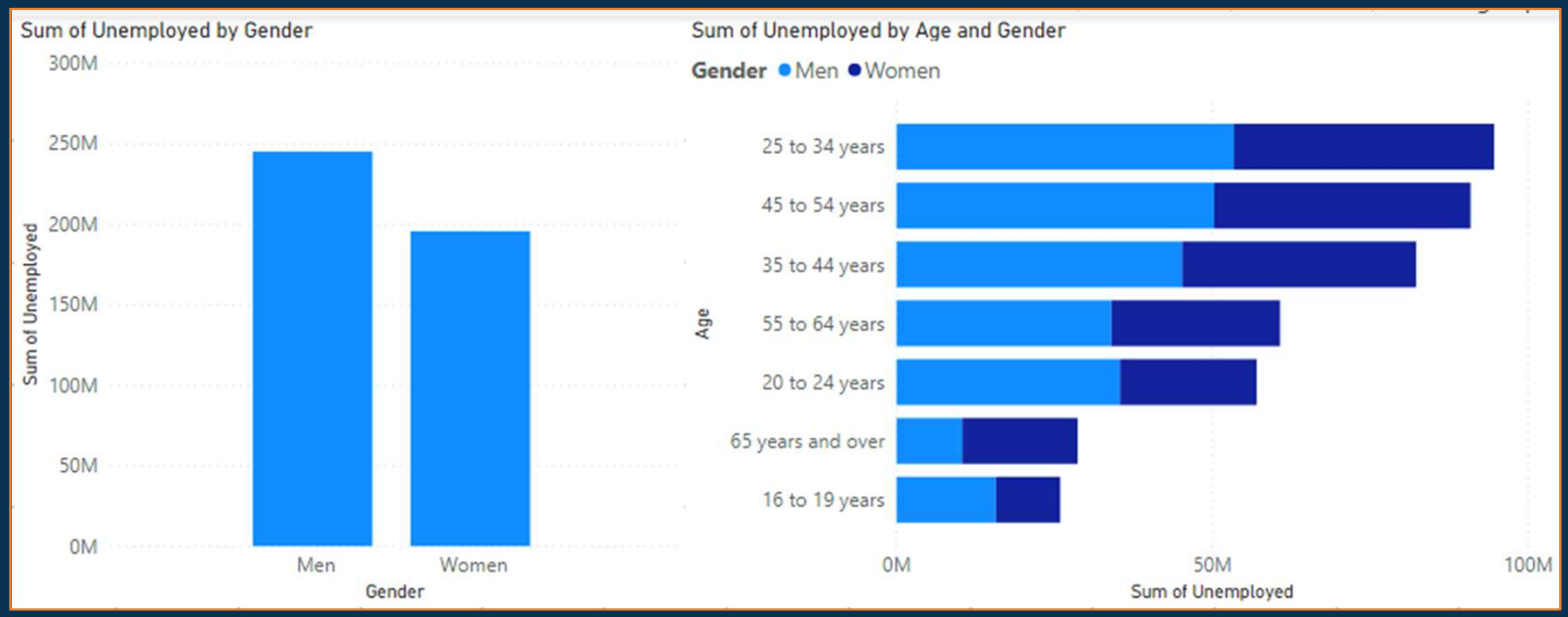


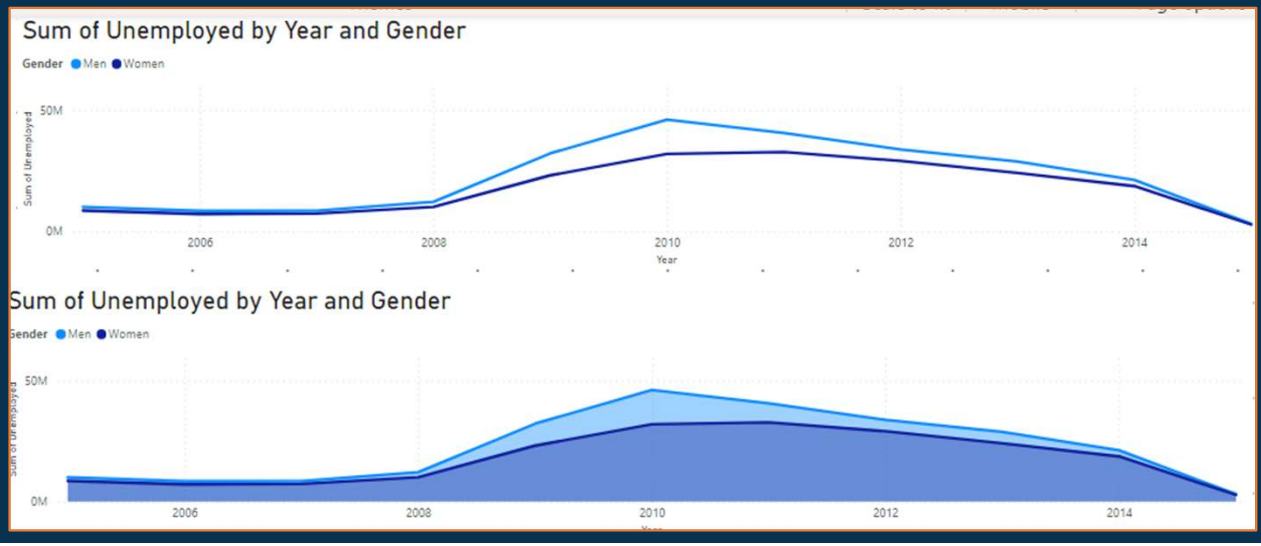
Table | Matrix | Card | Multivalued Card Chart

- The table is a grid that contains related data in a logical series of rows and columns.
- The table supports two dimensions and it can also contain headers and a row for totals.

Age	Gender	Sum of Unemployed
16 to 19 years	Men	15800000
16 to 19 years	Women	10170000
20 to 24 years	Men	35464000
20 to 24 years	Women	21613000
25 to 34 years	Men	53474000
25 to 34 years	Women	41191000
35 to 44 years	Men	45289000
35 to 44 years	Women	37032000
45 to 54 years	Men	50363000
45 to 54 years	Women	40574000
55 to 64 years	Men	34091000
55 to 64 years	Women	26676000
65 years and over	Men	10480000
65 years and over	Women	18246000
Total		**440463000**

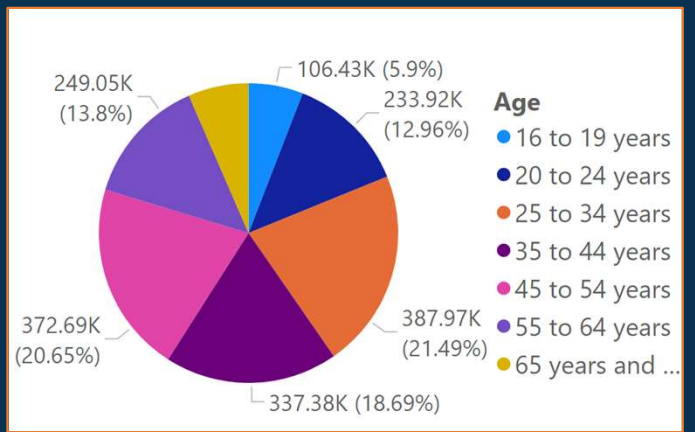
Line and Area Chart

- The line chart and area chart visualizations are beneficial in helping you present trends over time.
- The basic area chart is based on the line chart, with the area between axis and line filled in.
- The main difference between these two chart types is that the area chart highlights the magnitude of change over time.



Pie Chart

- Pie Chart shows the relationship of parts to the whole by dividing the data into segments.
- From a data analysis perspective, these charts are not useful because interpreting the data that they present can be difficult.
- However, these charts are often used for aesthetic reasons due to the colorful segments that they display.
- These charts are best suited for illustrating percentages.
- Pie charts are not suitable for representing measures with possibility of negative values.



Prepare Data for Analysis – Part 1

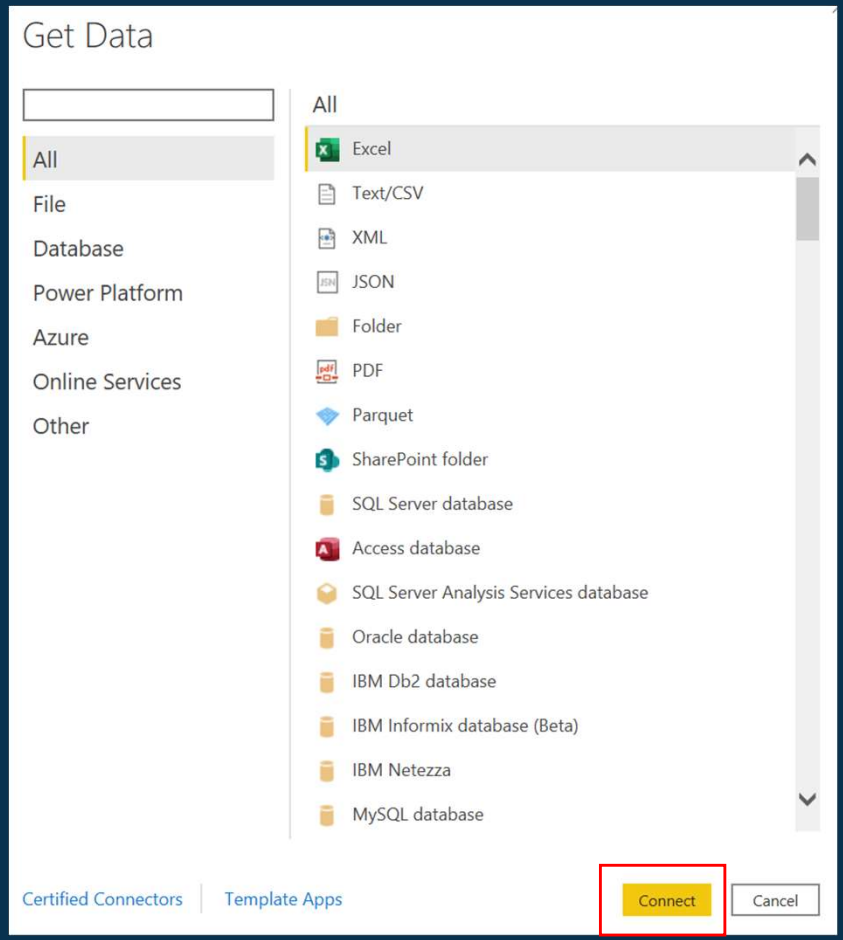
Learning Objective –

- Get Data in Power BI
- Introduction to Power Query Editor
- Column Profiling
- Append Queries
- Merge Queries

Get Data in Power BI

Connect to Data Source

1. *Home Tab > Data Ribbon> Get data*
2. *Select the appropriate data connection*
3. *Enter credentials (if required)*
4. *Load data*

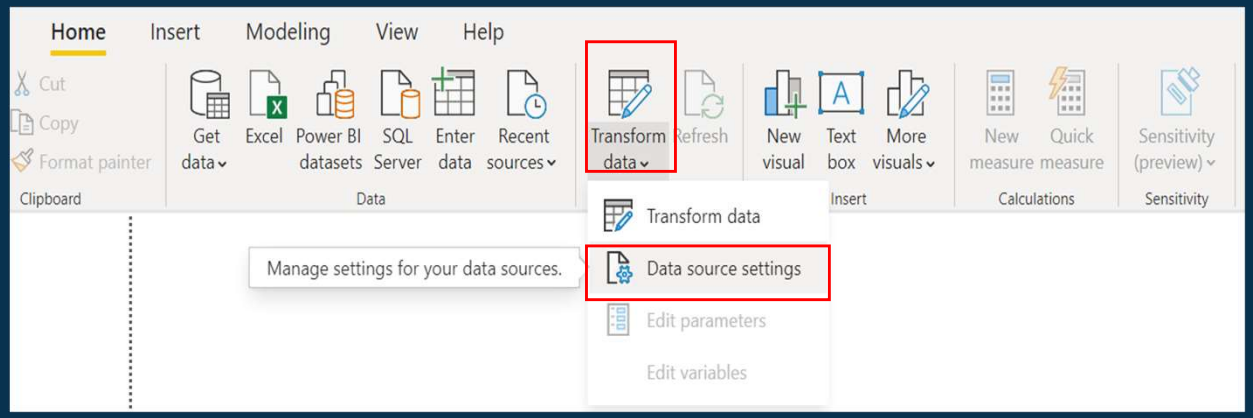


Change to Data Source Settings

1. *Home Tab > Queries Ribbon> Transform*

   Data > Data source settings

2. *Assign new data connection*



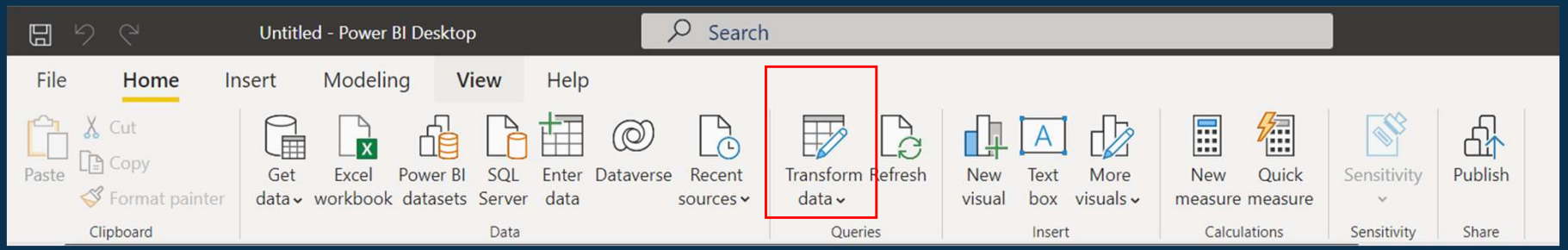
Introduction to Power Query Editor

Power Query Editor

- The Power Query Editor provides us with:
  - Import - You can make connection to a data source which can be located locally or over a network or even in the cloud
  - Transform - You can take up the raw data from its source and enrich the data through various transformations
  - Combine - You may also take data from multiple sources and combine them together in a meaningful way

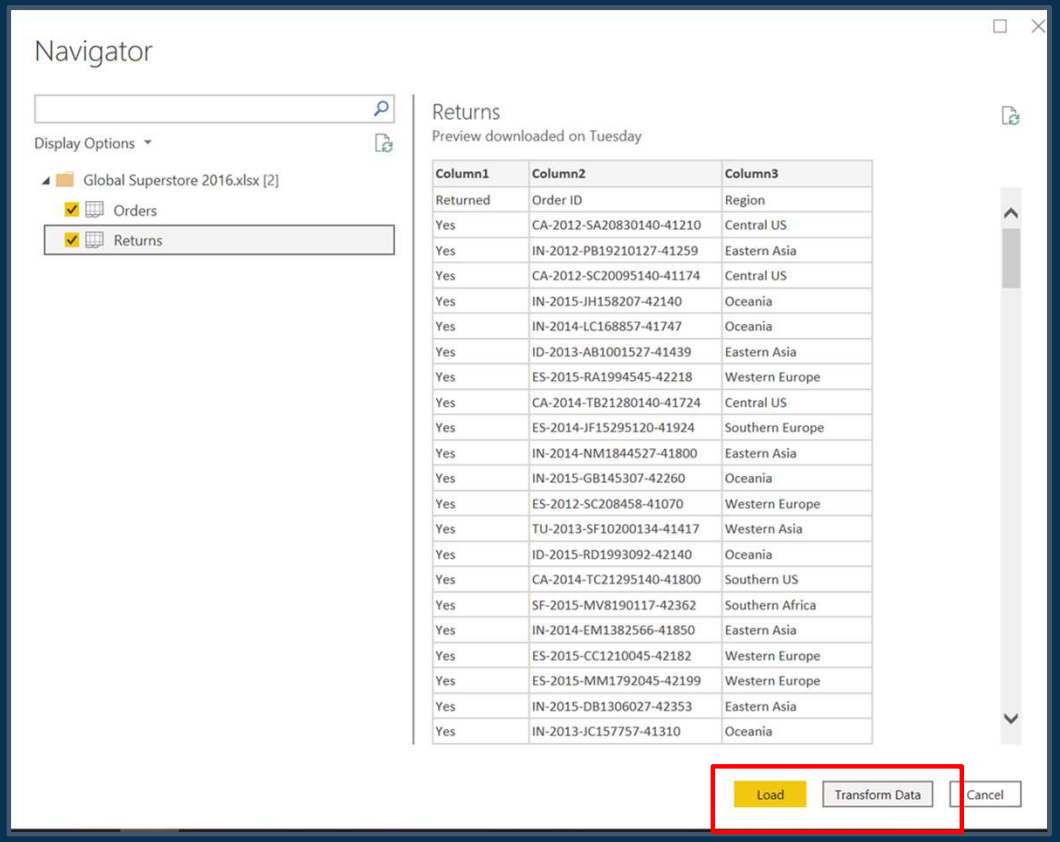
Power Query Editor

- Power Query Editor is a Power BI component. Its main use is to:
- Perform ETL (Extraction, Transformation, Loading)
- Data Transformation
- Power Query Editor uses the M language.
- It is available under the *[Home] > [Queries] > [Transform Data]*



Power Query Editor

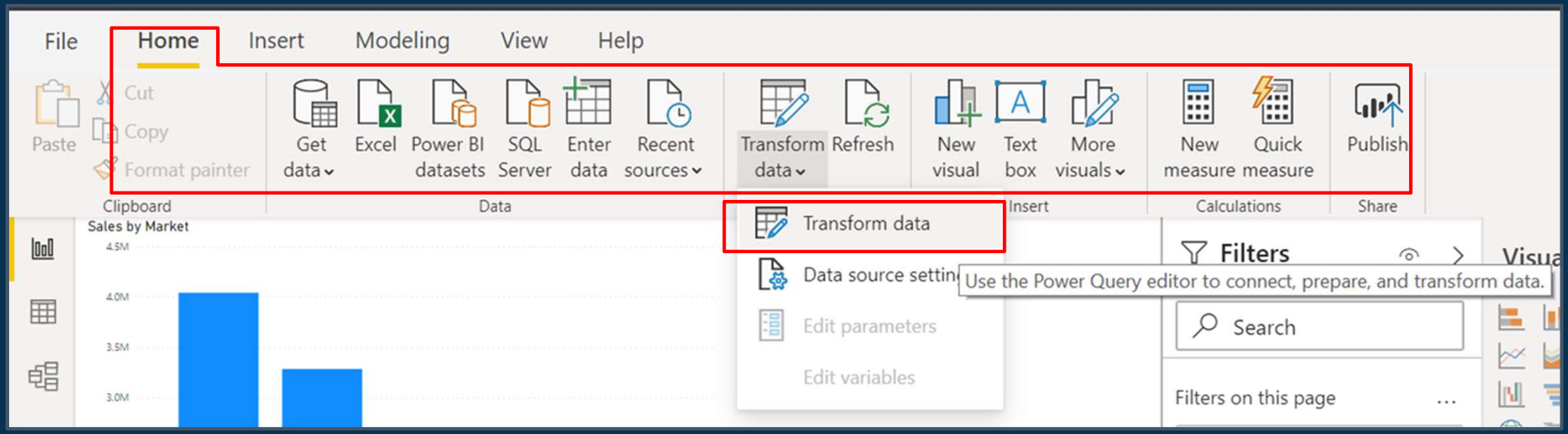
While loading data from the navigator select *Transform Data*



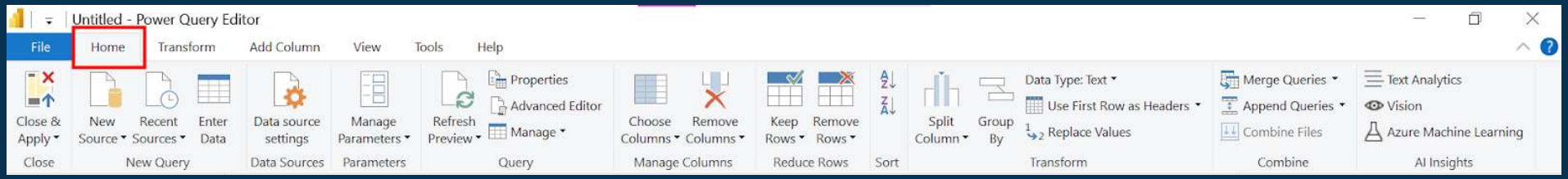
Column1	Column2	Column3
Returned	Order ID	Region
Yes	CA-2012-SA20830140-41210	Central US
Yes	IN-2012-PB19210127-41259	Eastern Asia
Yes	CA-2012-SC20095140-41174	Central US
Yes	IN-2015-JH158207-42140	Oceania
Yes	IN-2014-LC168857-41747	Oceania
Yes	ID-2013-AB1001527-41439	Eastern Asia
Yes	ES-2015-RA1994545-42218	Western Europe
Yes	CA-2014-TB21280140-41724	Central US
Yes	ES-2014-JF15295120-41924	Southern Europe
Yes	IN-2014-NM1844527-41800	Eastern Asia
Yes	IN-2015-GB145307-42260	Oceania
Yes	ES-2012-SC208458-41070	Western Europe
Yes	TU-2013-SF10200134-41417	Western Asia
Yes	ID-2015-RD1993092-42140	Oceania
Yes	CA-2014-TC21295140-41800	Southern US
Yes	SF-2015-MV8190117-42362	Southern Africa
Yes	IN-2014-EM1382566-41850	Eastern Asia
Yes	ES-2015-CC1210045-42182	Western Europe
Yes	ES-2015-MM1792045-42199	Western Europe
Yes	IN-2015-DB1306027-42353	Eastern Asia
Yes	IN-2013-JC157757-41310	Oceania

Power Query Editor

- *Home Tab > Queries Ribbon > Transform data*

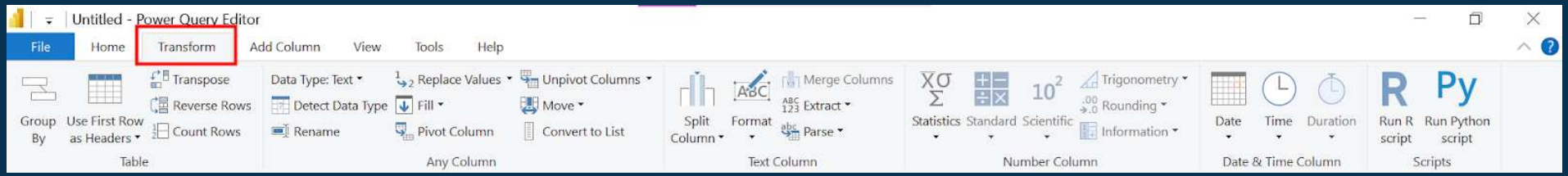


Power Query Editor – Interface



› *Home Tab* frequently used query tasks

Power Query Editor – Interface

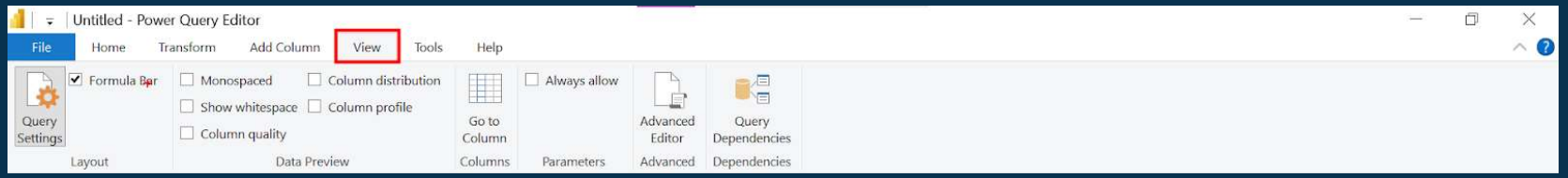


- The *Transform tab* helps perform various transformations on data such as : Adding or removing columns, changing data types, splitting columns, other data-driven tasks

Power Query Editor – Interface

- The *Add Column tab* adds a new column on performing data transformations. It also provides options to add custom columns and conditional column.

Power Query Editor – Interface

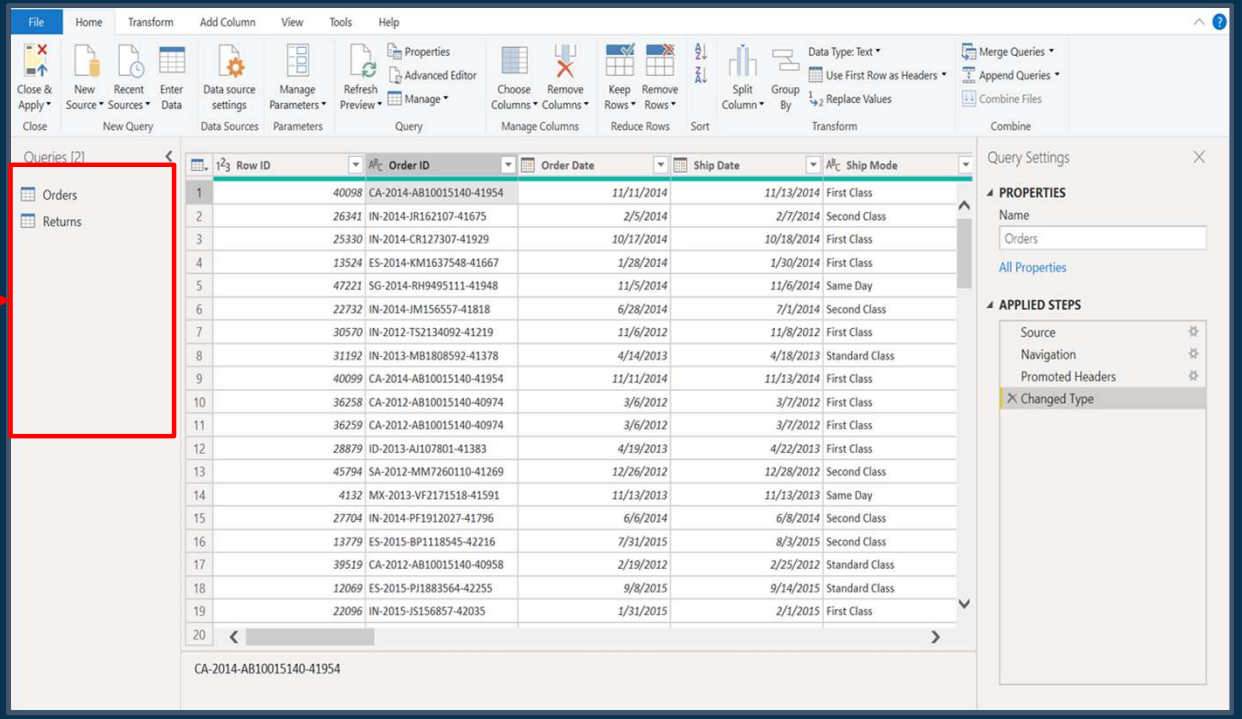


- The *View tab* adds is used to toggle whether certain panes or windows are displayed. It's also used to display the Advanced Editor.

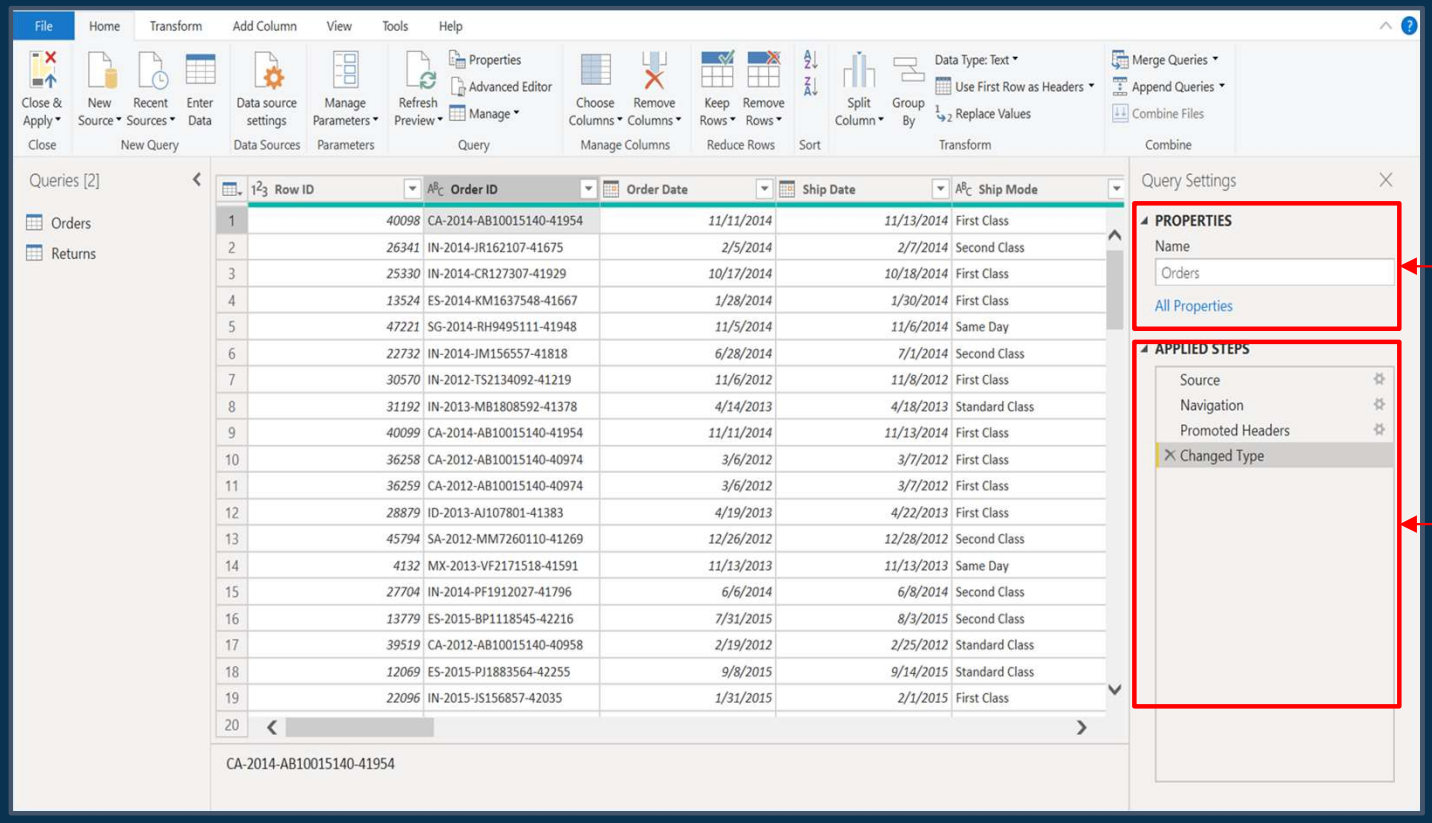
Power Query Editor – Interface

List of queries

Data preview of the selected sheet gets loaded



Power Query Editor – Interface

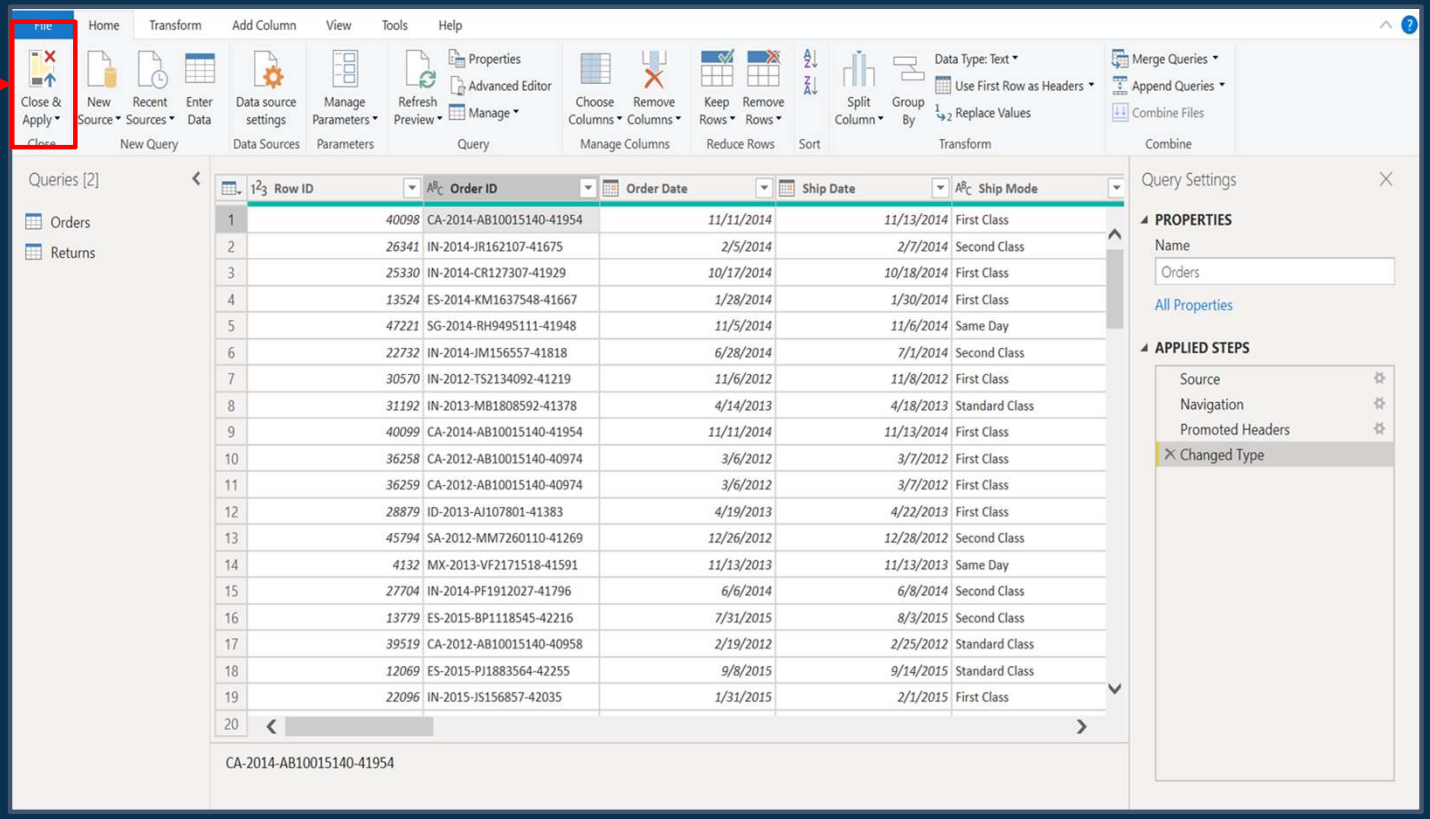


Renaming query

List of all the transformation steps applied on the data

Power Query Editor – Interface

Apply changes

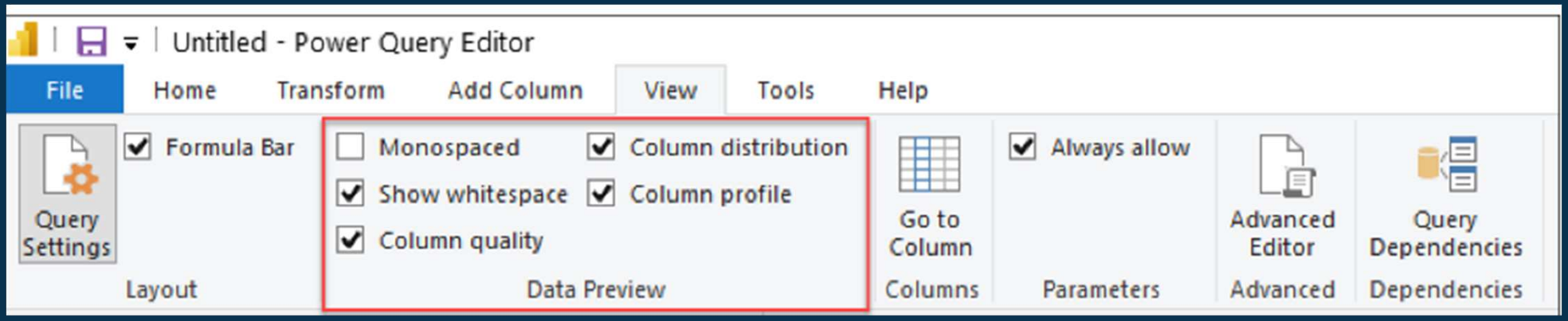


Profiling Data

- Profiling data is about understanding the patterns in the data such as : determining anomalies, examining underlying data structures, and querying data statistics such as row counts, value distributions, minimum and maximum values, averages, and so on.
- This process allows you to shape and organize the data so that interacting with the data and identifying the distribution of the data is simple.
- The data profiling tools include:
  - Column quality
  - Column distribution
  - Column profile

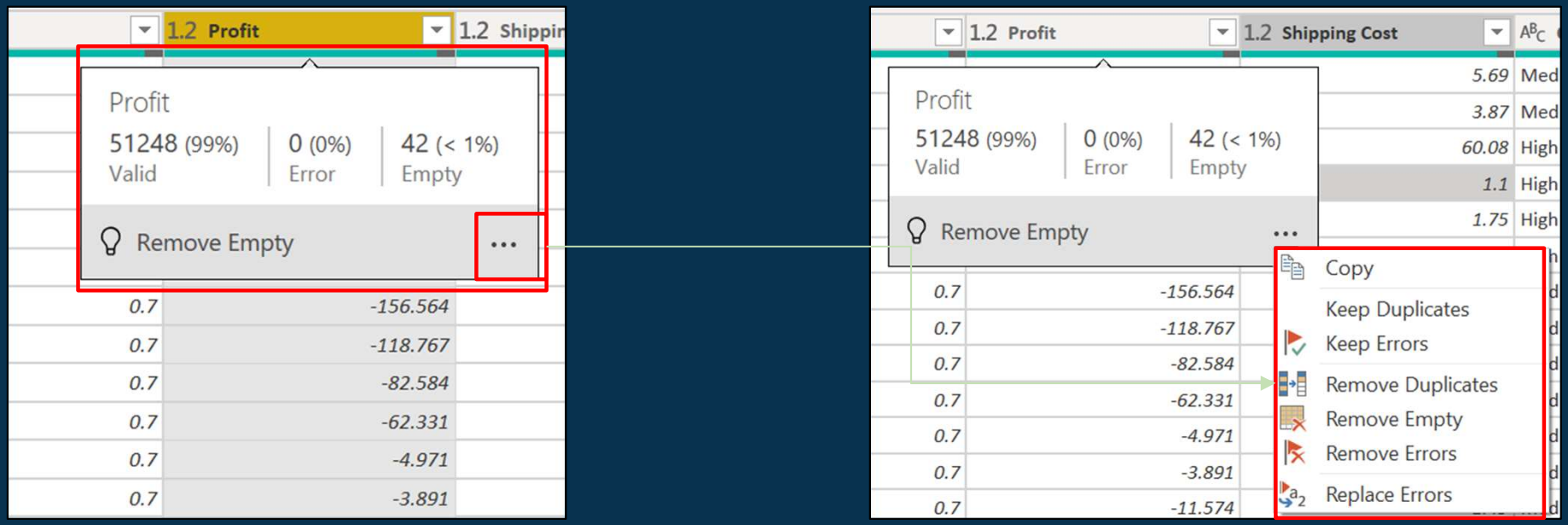
Using Data Profiling tools

- Click on *View tab* > *Data Preview* group, select one or more of the elements



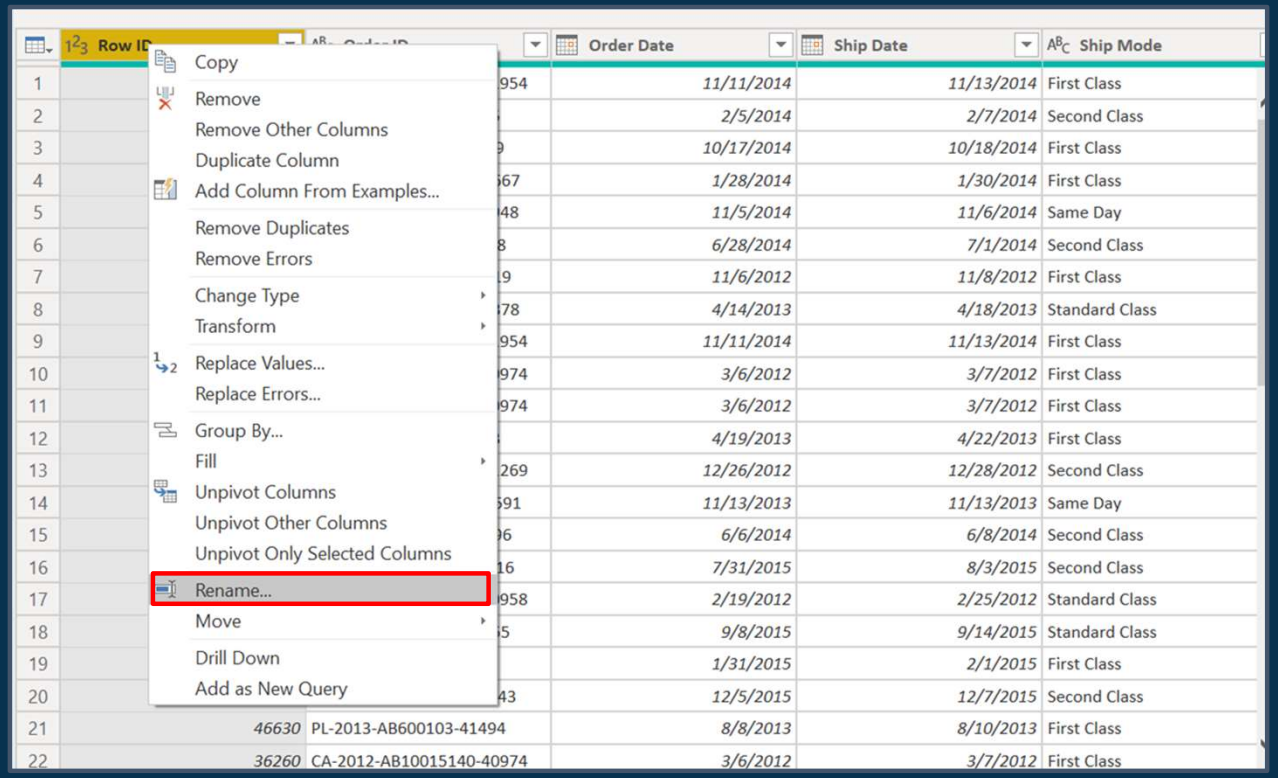
Column info

- Hovering over column bar tells the validity of data in the column.
- E.g. null or empty values, error values etc.



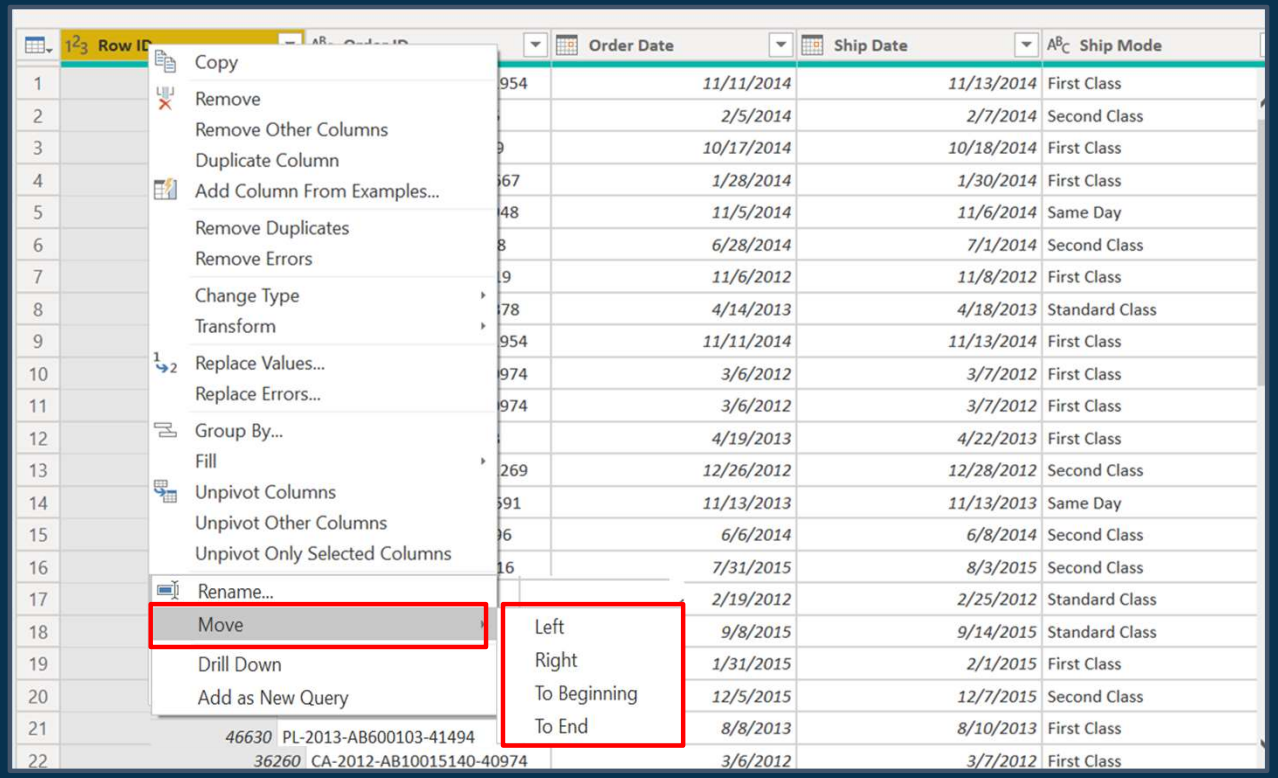
Renaming columns

- Right click on the column to be renamed
- Select Rename
- Assign a new name



Reordering columns

- Right click on the column to be reordered
- Select Move > desired position

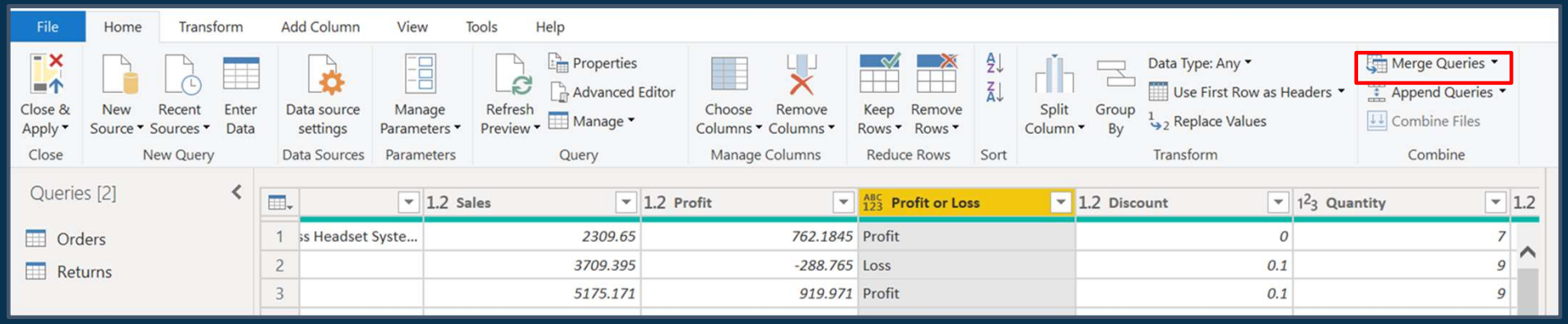


Merge queries

- Most commonly used operation
- Works similar to join operation in SQL
- Merging also includes appending data so that it can be stacked up one after another

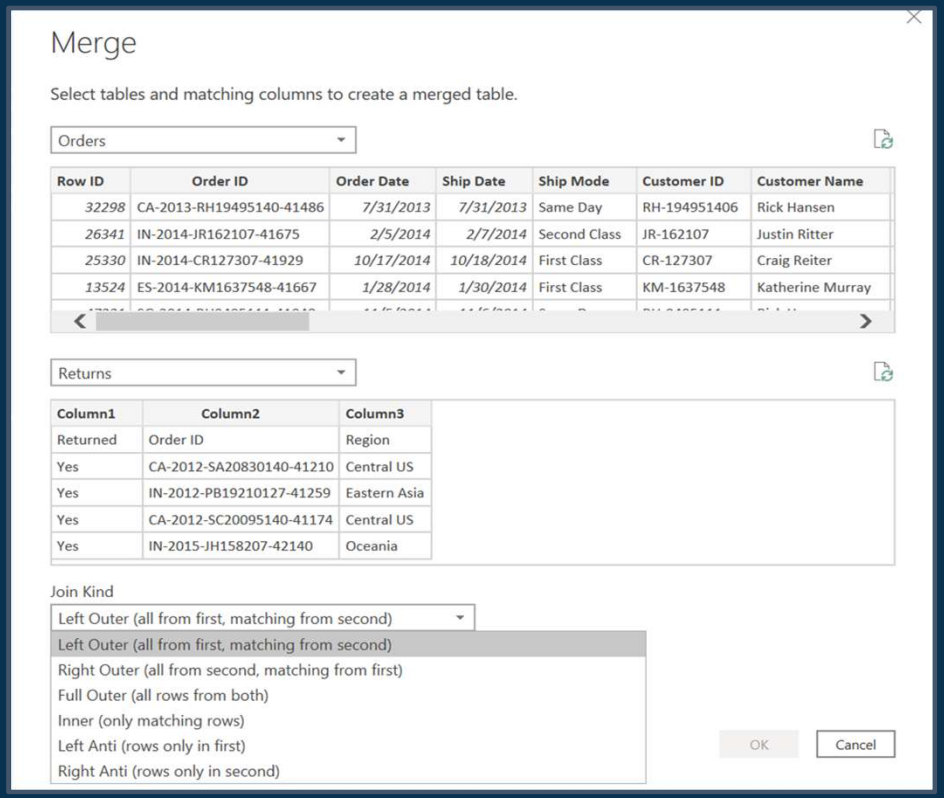
Merge query

- It works similar to join operations
- In Home Tab > Combine Ribbon > Merge Queries
- It can be either merged into same query or merged as a new query



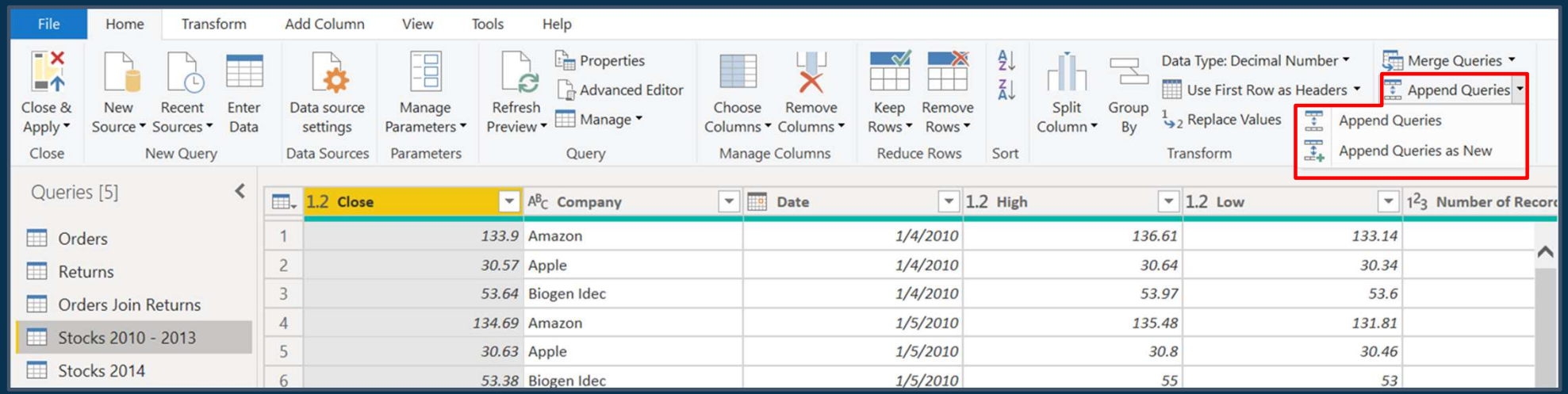
Merge query

- Select join type
- Select first query
- Select second query



Append query

- It works similar to union operation
- In Home Tab > Combine Ribbon > Append Queries
- It can be either append into same query or append as a new query



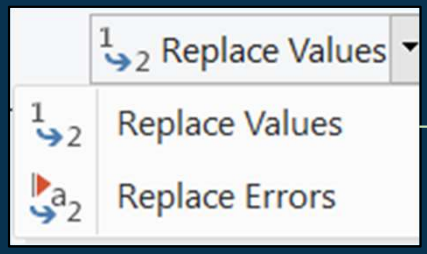
Prepare Data for Analysis – Part 2

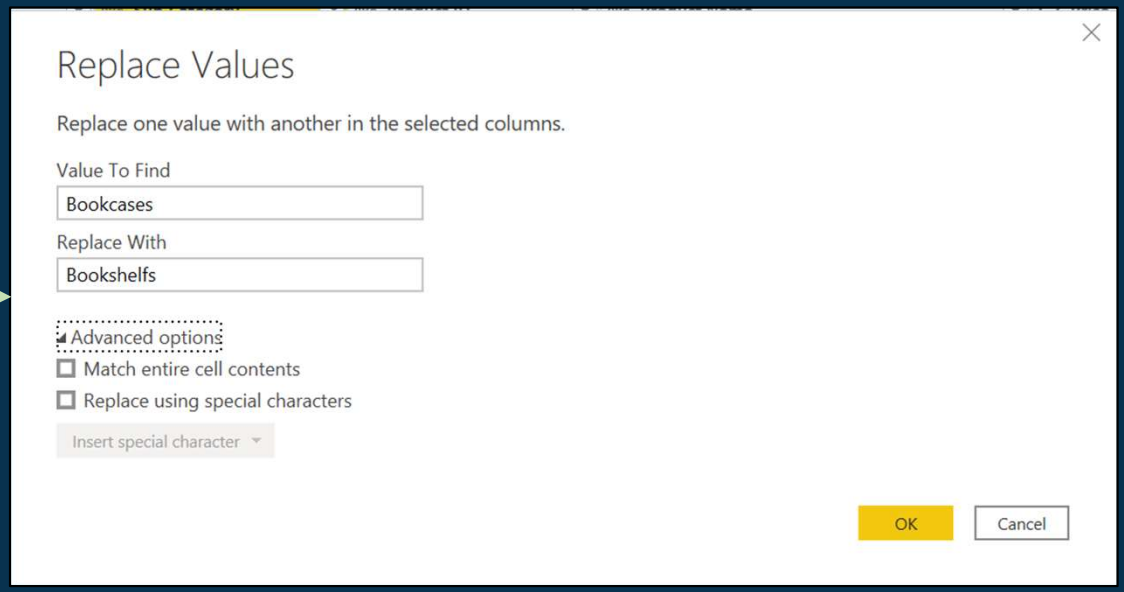
Learning Objective –

- Replace Values
- Splitting Values
- Date Field
- Sorting and Filtering
- Adding index column
- Conditional Column
- Pivot | Group By

Replacing values

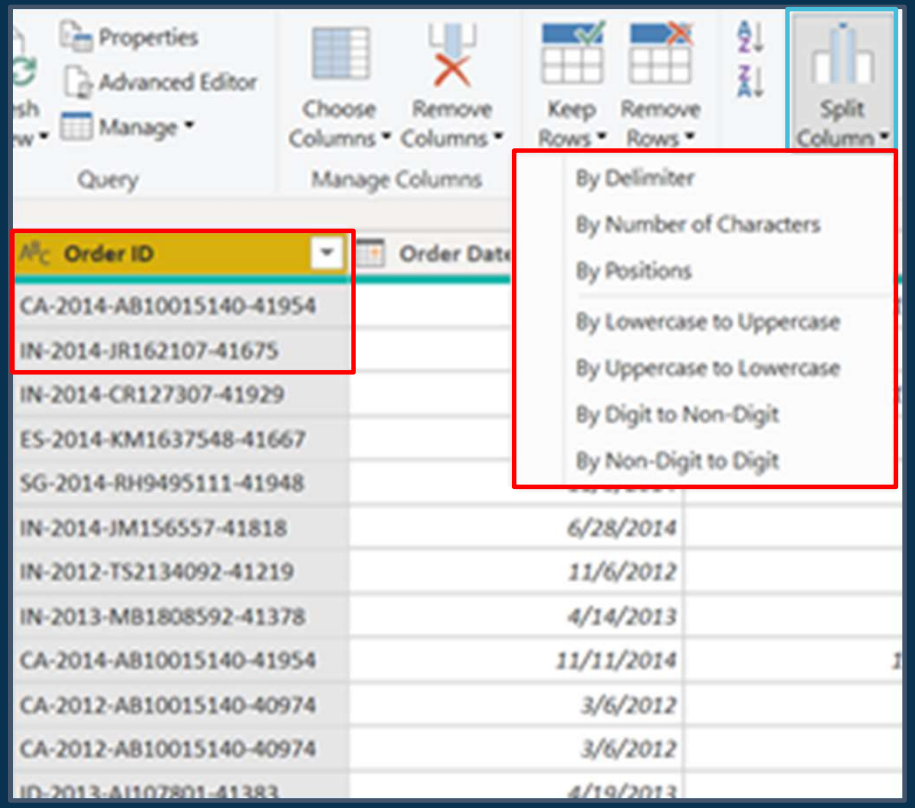
- Transform tab > Any Column ribbon > Replace Values





Splitting columns

- In the Home Tab > Transform ribbon
- Select Split Column > choose the desired option
- This will create new columns by splitting the text into multiple columns

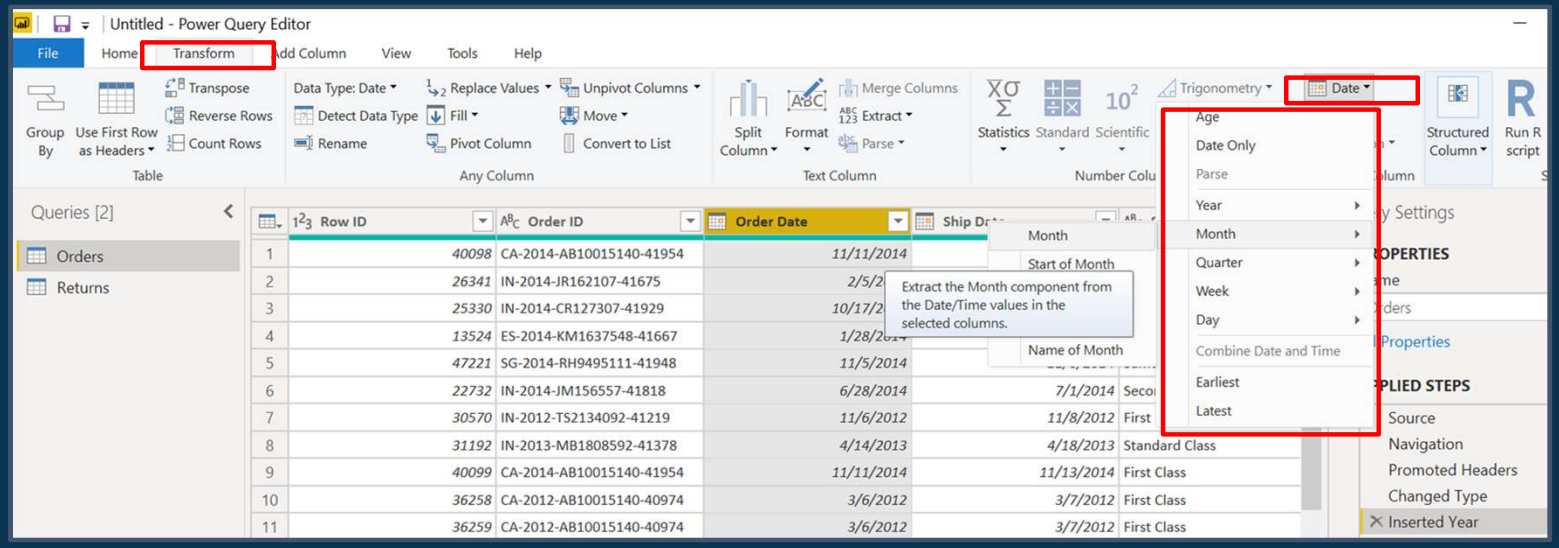


Date field

- Power BI works behind the scenes to automatically identify columns that represent dates
- It creates date hierarchies and other enabling metadata for models
- These built-in hierarchies can be used when creating report features like visuals, tables, quick measures, slicers, and so on

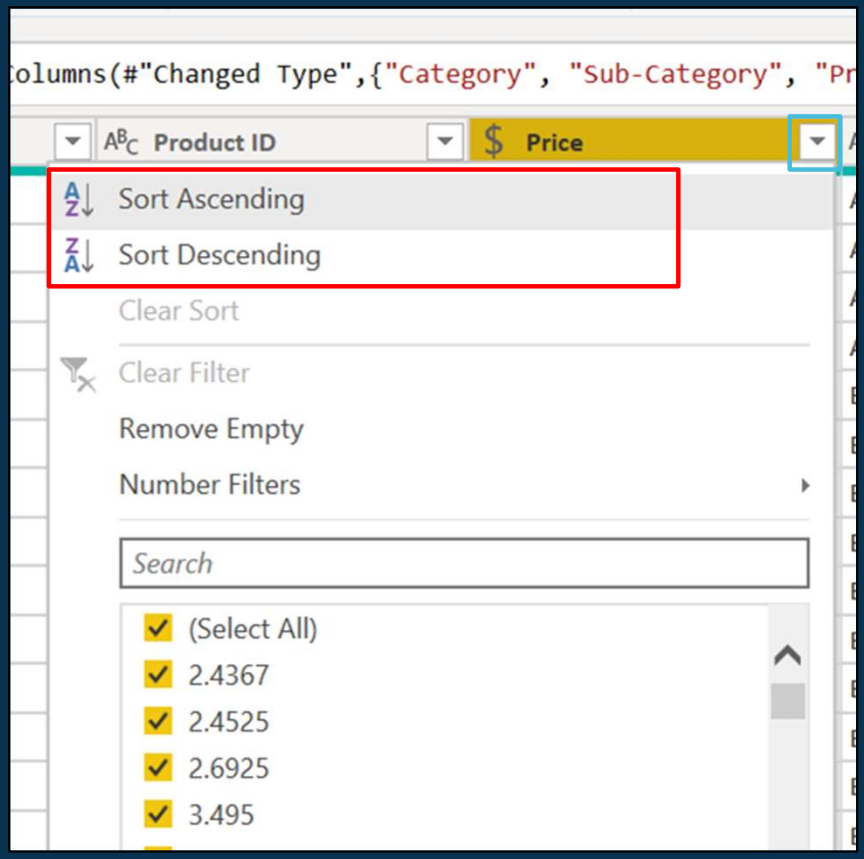
Extracting year, month, day from date field

- Select the date column to be modified
- Transform Tab > Date (drop-down) > select the part of the date to be extracted



Sorting values

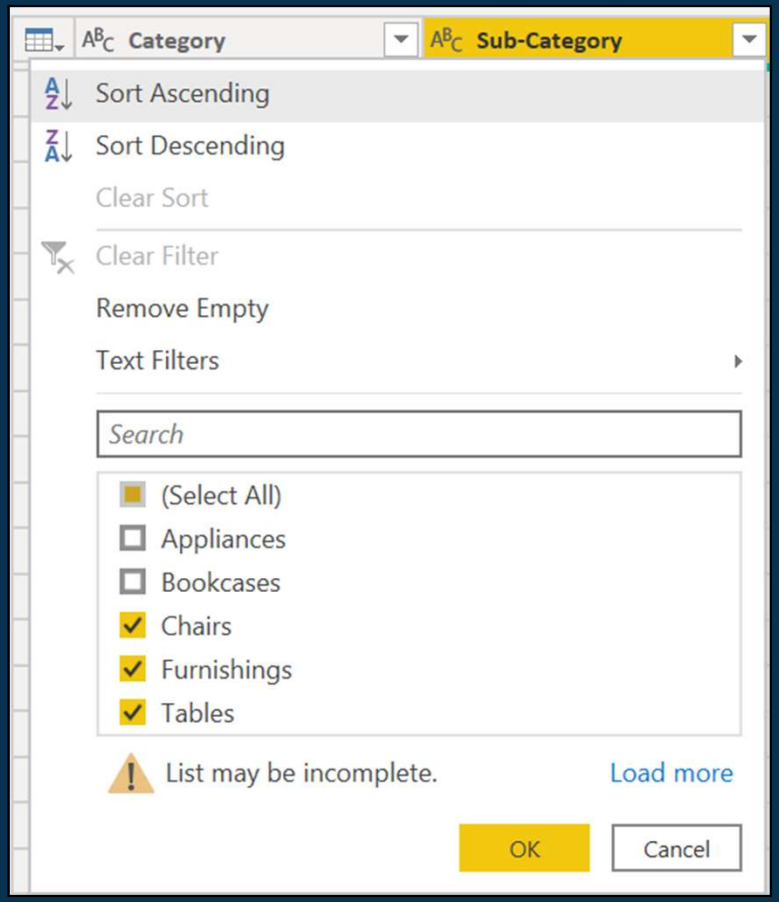
- Sort Ascending/ Descending will sort the data in the column.
- This sort is limited to the data view in Power Query Editor.



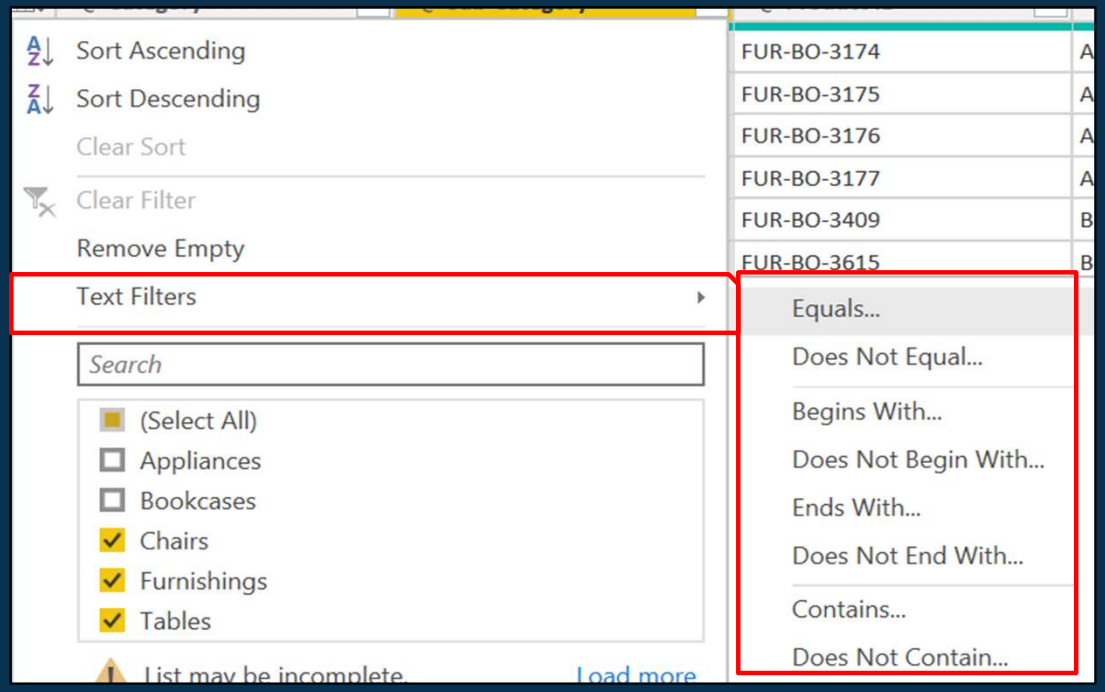
Filtering columns

- Deselect the nonessential values from the data.
- Filtering enables removing undesired values

  from the data model



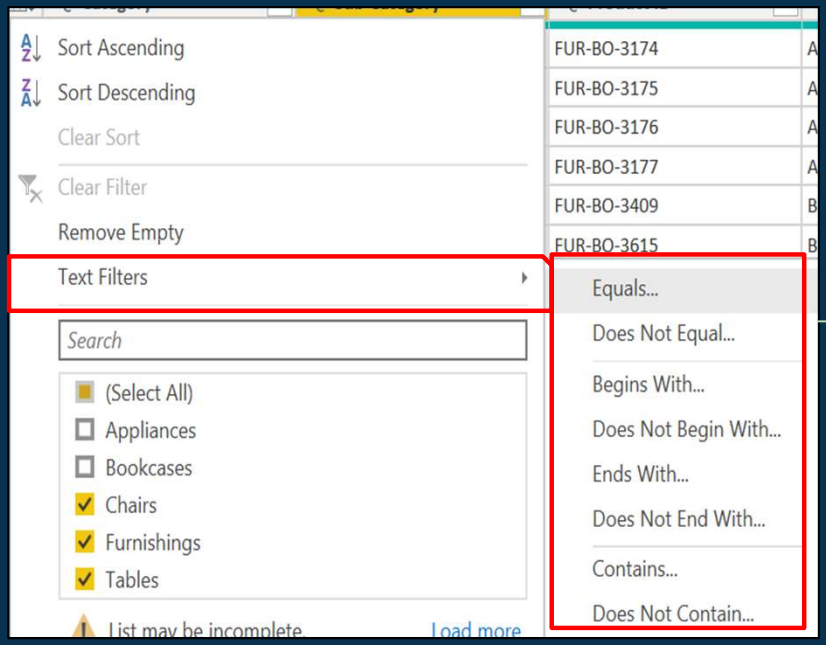
Filtering columns - Text Filters

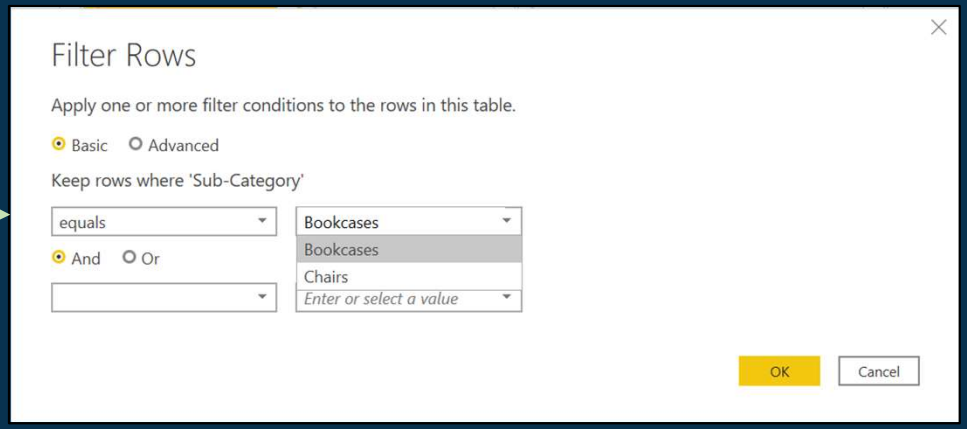


Custom condition can be

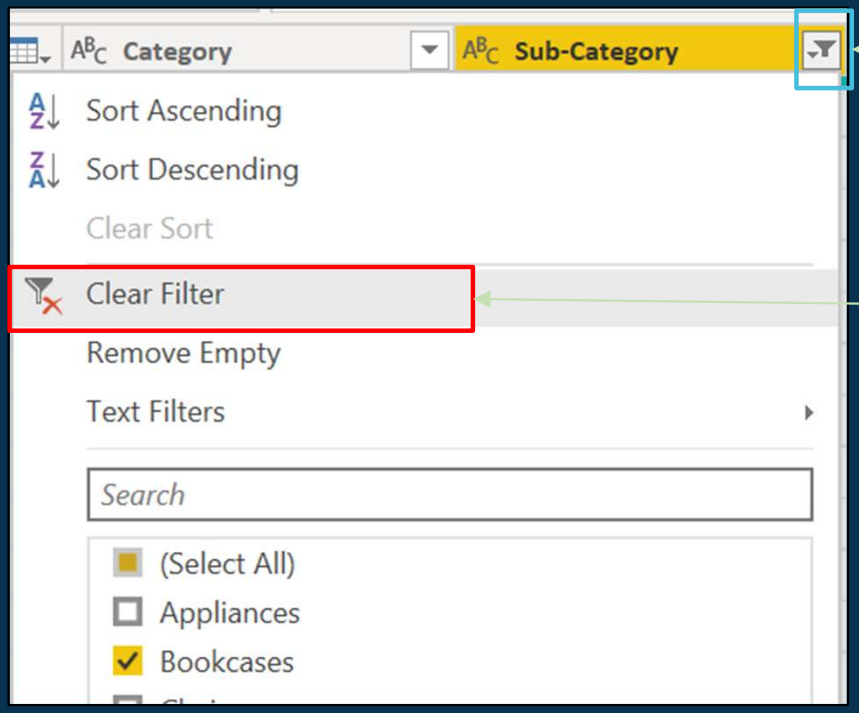
applied to filter text data

Filtering columns - Text Filters





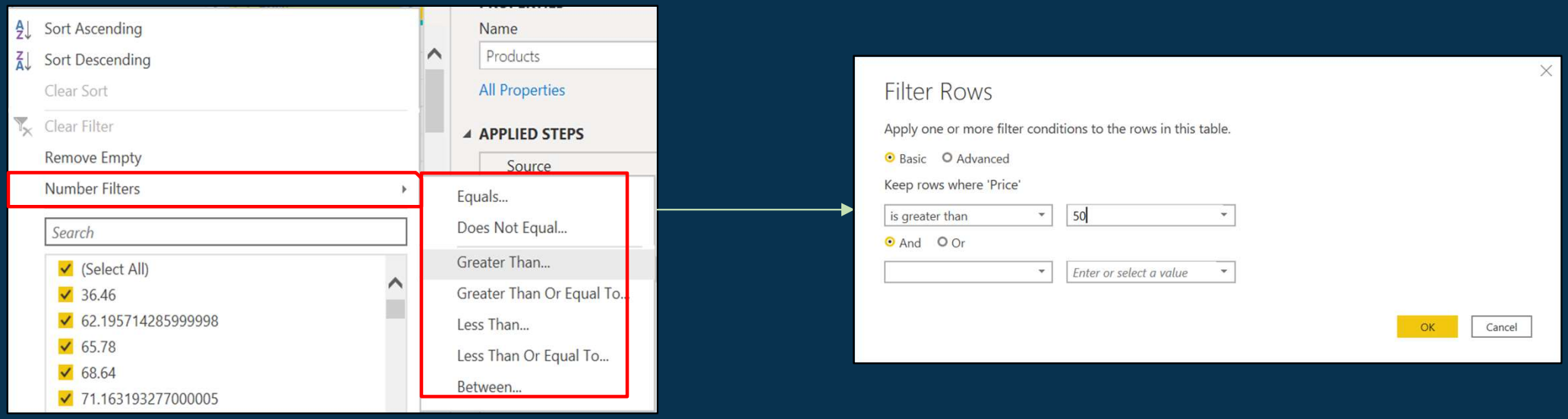
Filtering columns - Text Filters



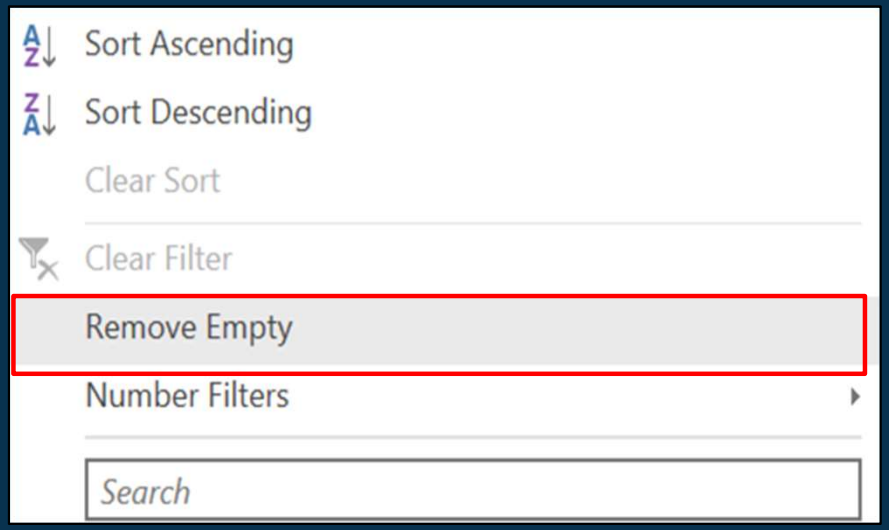
This symbol denotes, filter is applied to this column

Clears all the applied filters

Filtering columns - Numeric Filters



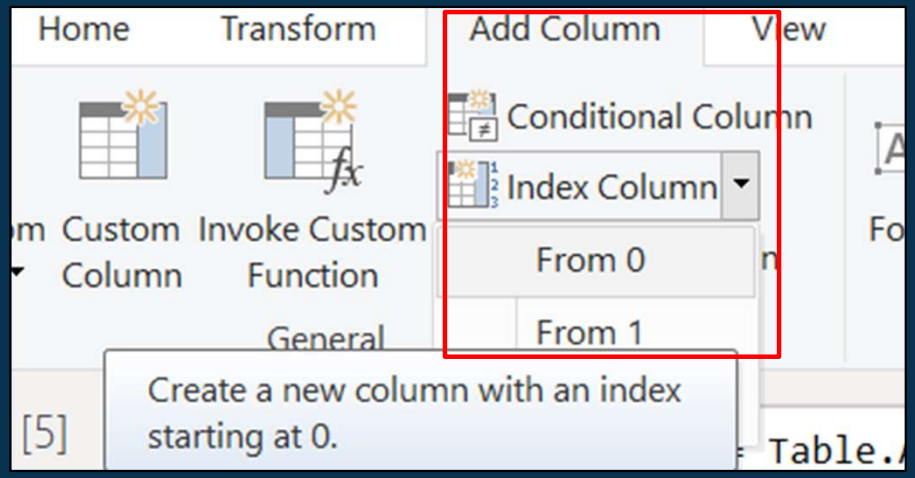
Filtering columns - Remove Empty



Removes blank cells from the column

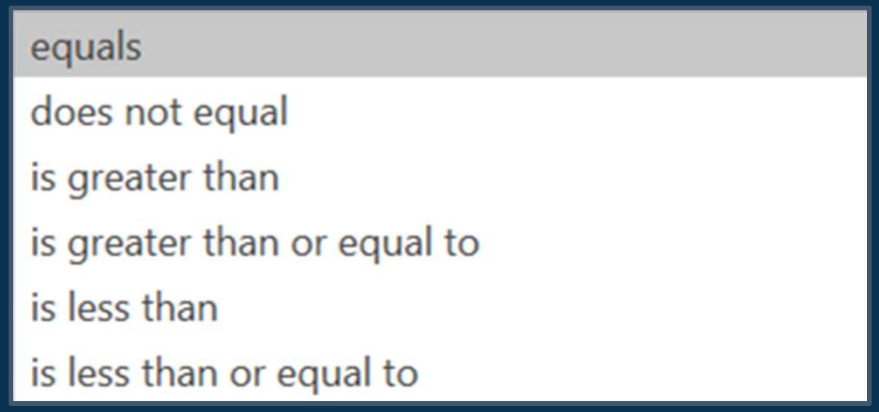
Adding Index Column

- Add Column tab > Index Column
- It lets you add an index column that serves as a row counter to your data.
- This is especially useful when you are filtering by row position or by a range or rows.



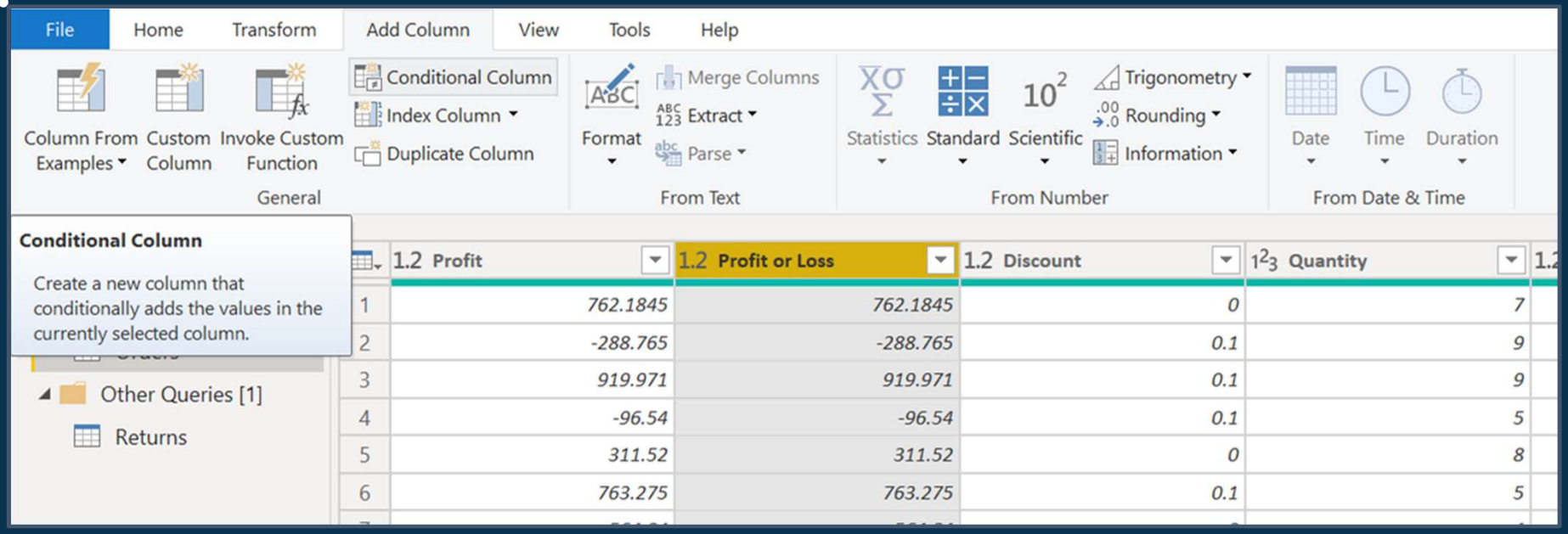
Conditional columns

- They are derived by applying logical conditions to an existing column
- Logical conditions can be as mentioned in the figure



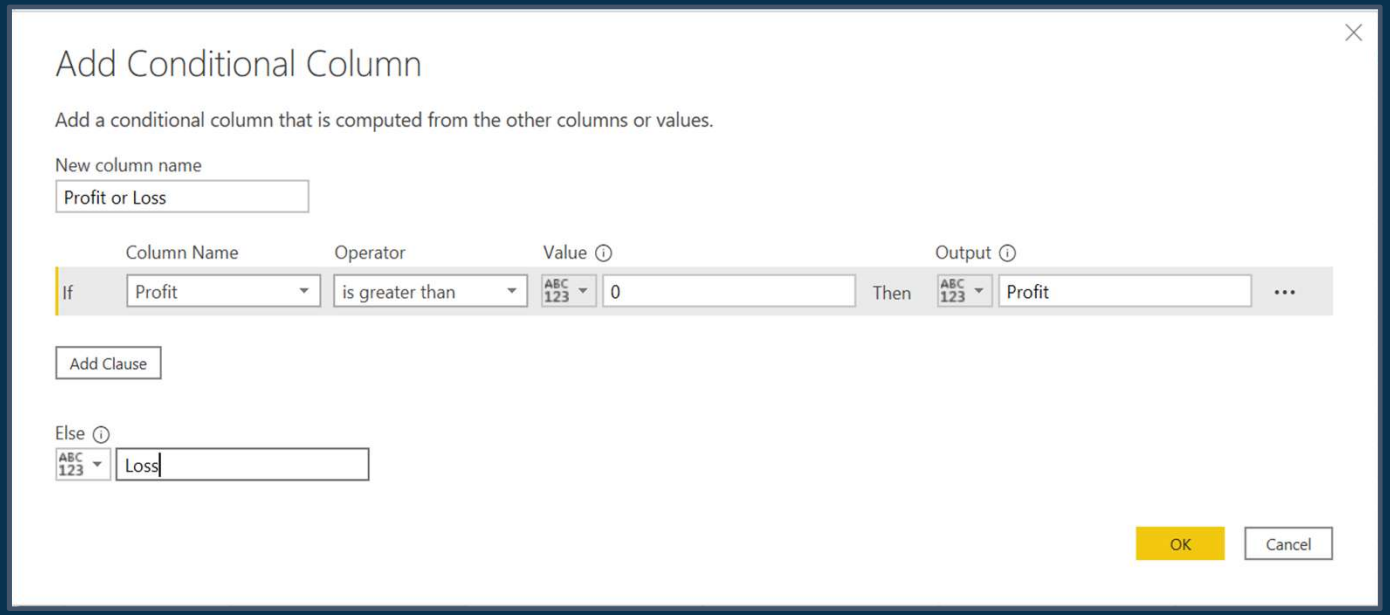
Creating conditional column

- Select the column whose the values are to be conditionally modified
- Add Column Tab > Conditional Column
-



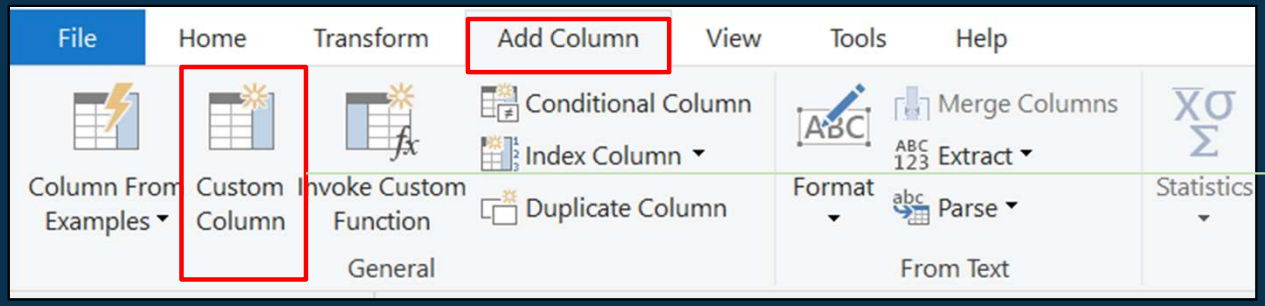
Assign column name

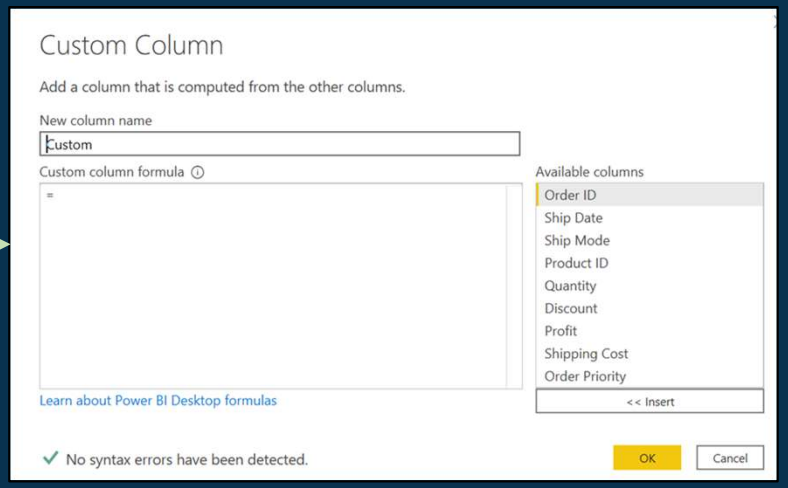
- Creating conditional column
- Set up the condition



Adding custom column

- Add Column tab > Custom Column





It lets you create a new column based on the existing

fields. DAX formula is used to defined the column

Pivot Table?

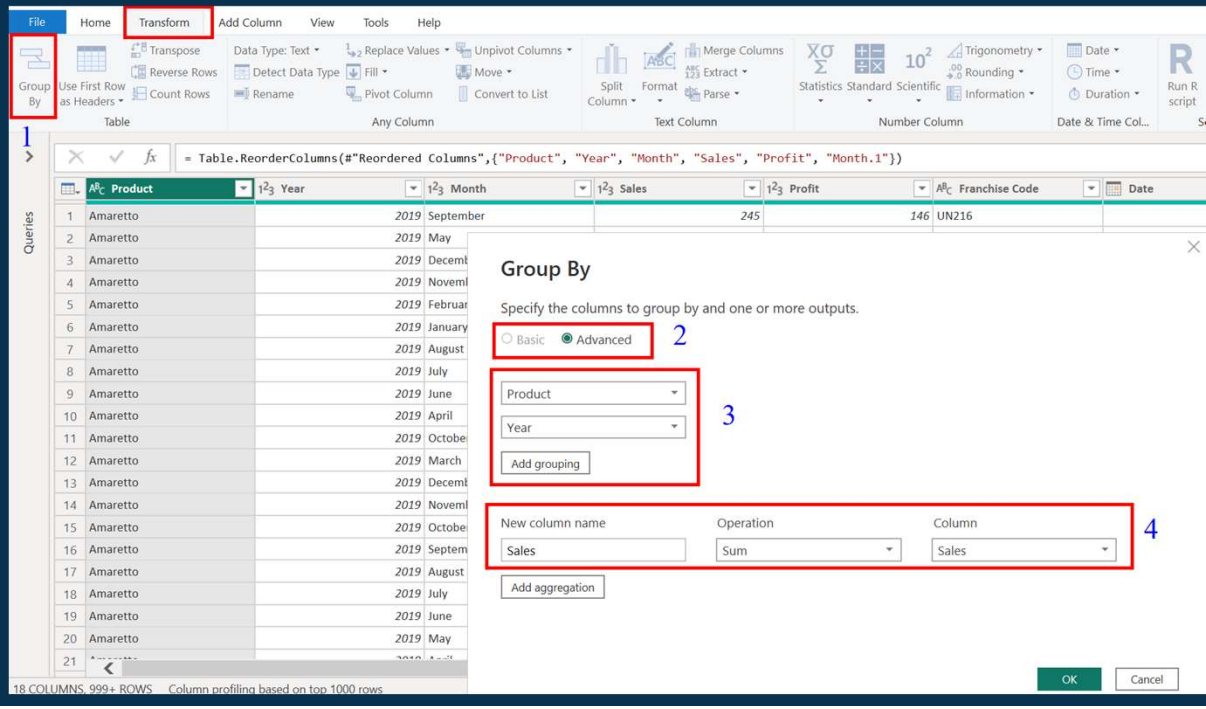
- A pivot table is a data summarization tool that is used in the context of data processing
- Pivot tables are used to summarize, sort, reorganize, group, count, total or average data stored in a database
- It allows its users to transform columns into rows and rows into columns. It allows grouping by any data field

When to use pivot tables

- Pivot tables are most commonly used in situations where data needs to be aggregated, and sliced and diced for analysis
- It’s particularly useful when you are looking to calculate and summarize data in order to make comparisons

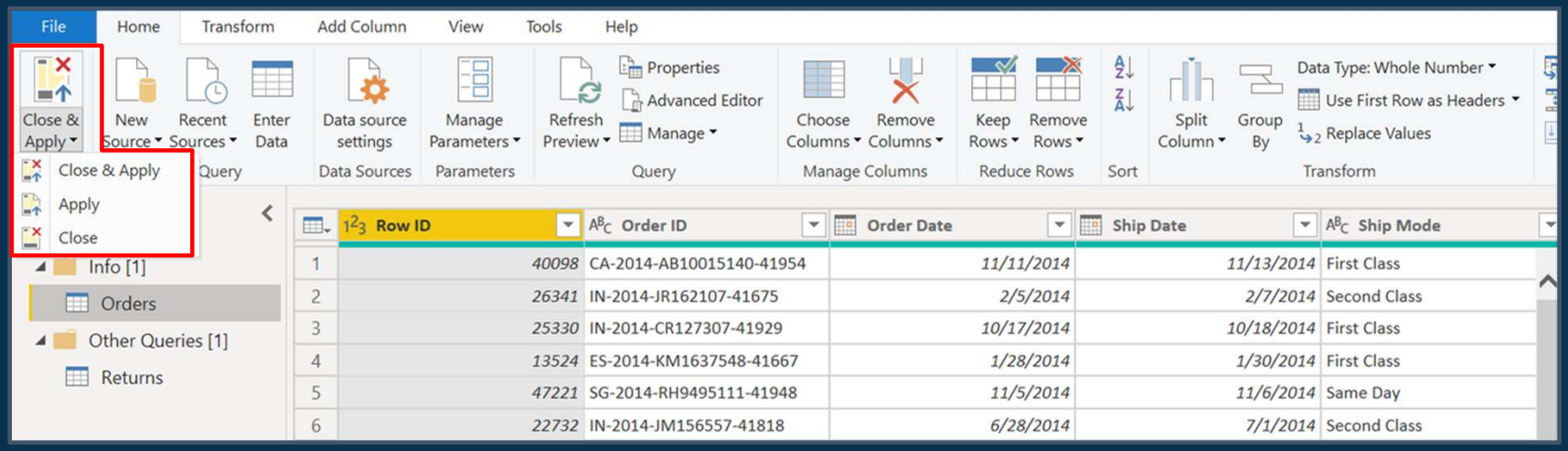
Group By

- *Transform tab > Group By*
- Select the categorial fields
- Assign new name to columns
- Select aggregation and column



Applying changes

- To apply all the enforced transformation changes, click Home Tab > Close & Apply



Data Modelling

Learning Objective –

- Design a data model in Power BI
- Managing relationships
- Difference between merge and relationship

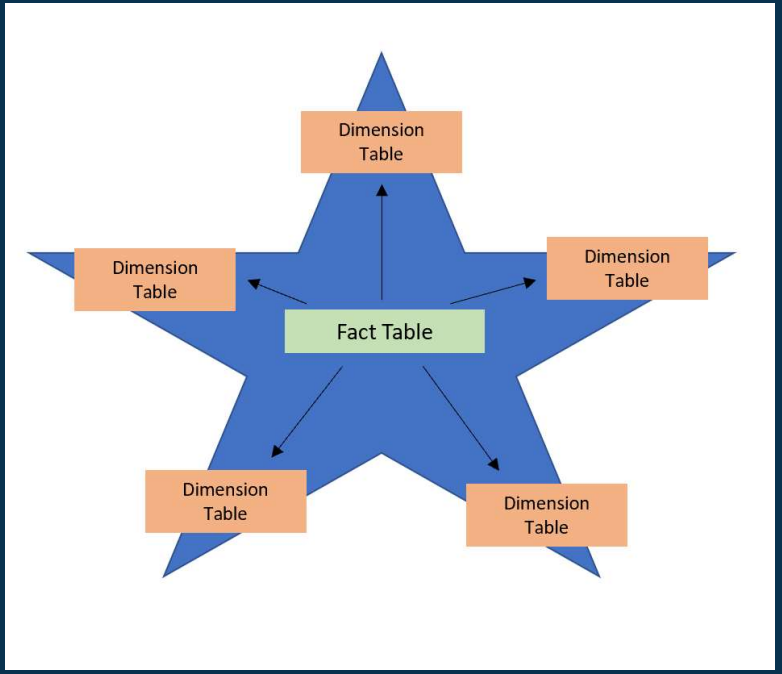
Introduction to Data Modelling

A good data model offers the following benefits:

- Data exploration is faster.
- Aggregations are simpler to build.
- Reports are more accurate.
- Writing reports takes less time.
- Reports are easier to maintain in the future.

Star Schemas

- Star schema is helpful to simplify data. Though it's not the only way to simplify data, but it is a popular method
- In a star schema, each table within dataset is defined as a dimension or a fact table

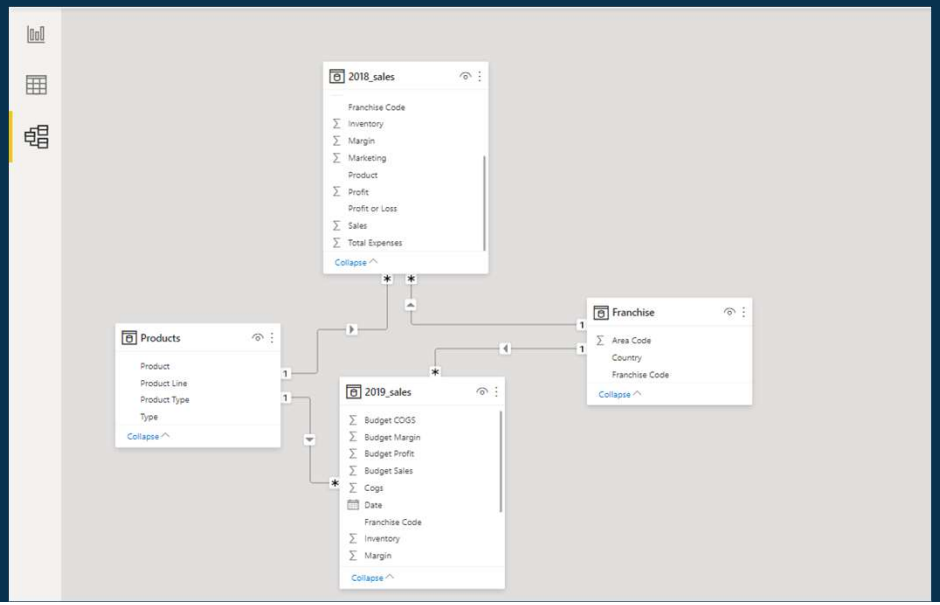


Star Schemas

- Fact tables contain observational or event data values: sales orders, product counts, prices, transactional dates and times, and quantities. Fact tables can contain several repeated values. For example, one product can appear multiple times in multiple rows, for different customers on different dates. These values can be aggregated to create visuals.
- Dimension tables contain the details about the data in fact tables: products, locations, employees, and order types. These tables are connected to the fact table through key columns. Dimension tables are used to filter and group the data in fact tables. The dimension tables, by contrast, contain unique values, for instance, one row for each product in the Products table and one row for each customer in the Customer table.

Model View

- Model view shows all of the tables, columns, and relationships in your model.
- This view can be especially helpful when your model has complex relationships between many tables.
- Select the Model icon near the side of the window to see a view of the existing model.

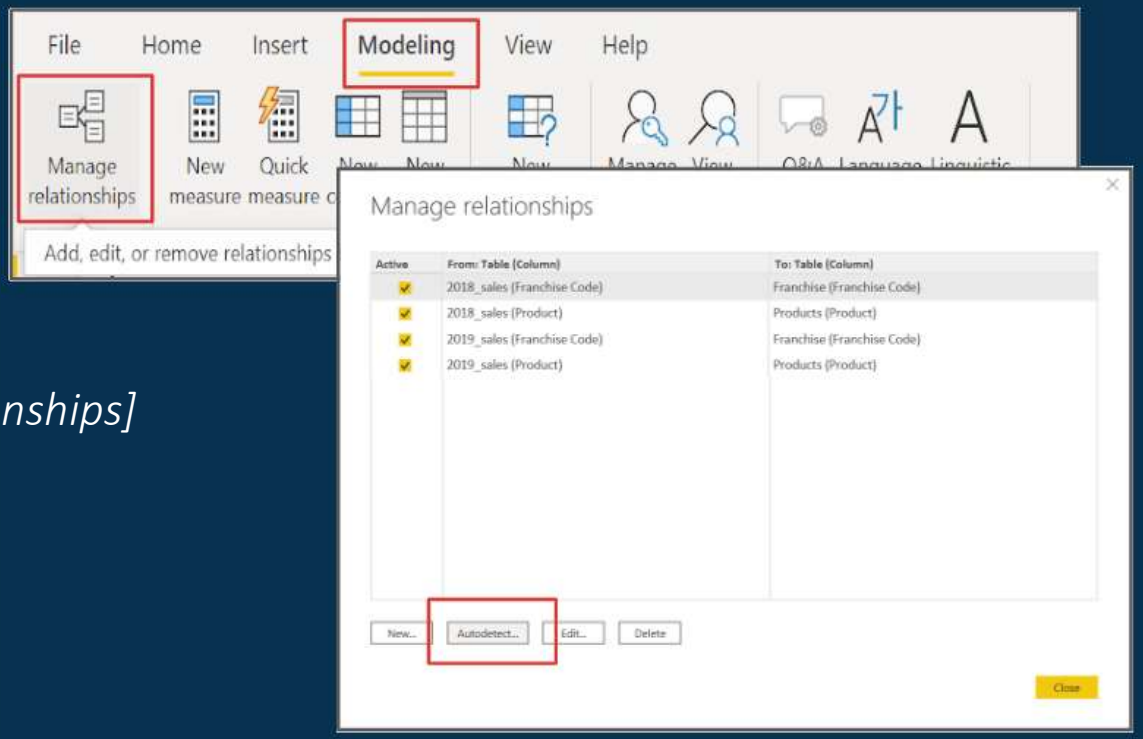


Managing relationships

- When multiple tables are imported, data in the tables is jointly is analyzed
- Relationships between those tables are necessary to accurately calculate results and display the correct information
- Power BI Desktop makes creating those relationships easy using the modelling tool
- In most cases auto-detection of relationship works great
- However, sometimes relationships can be edited manually

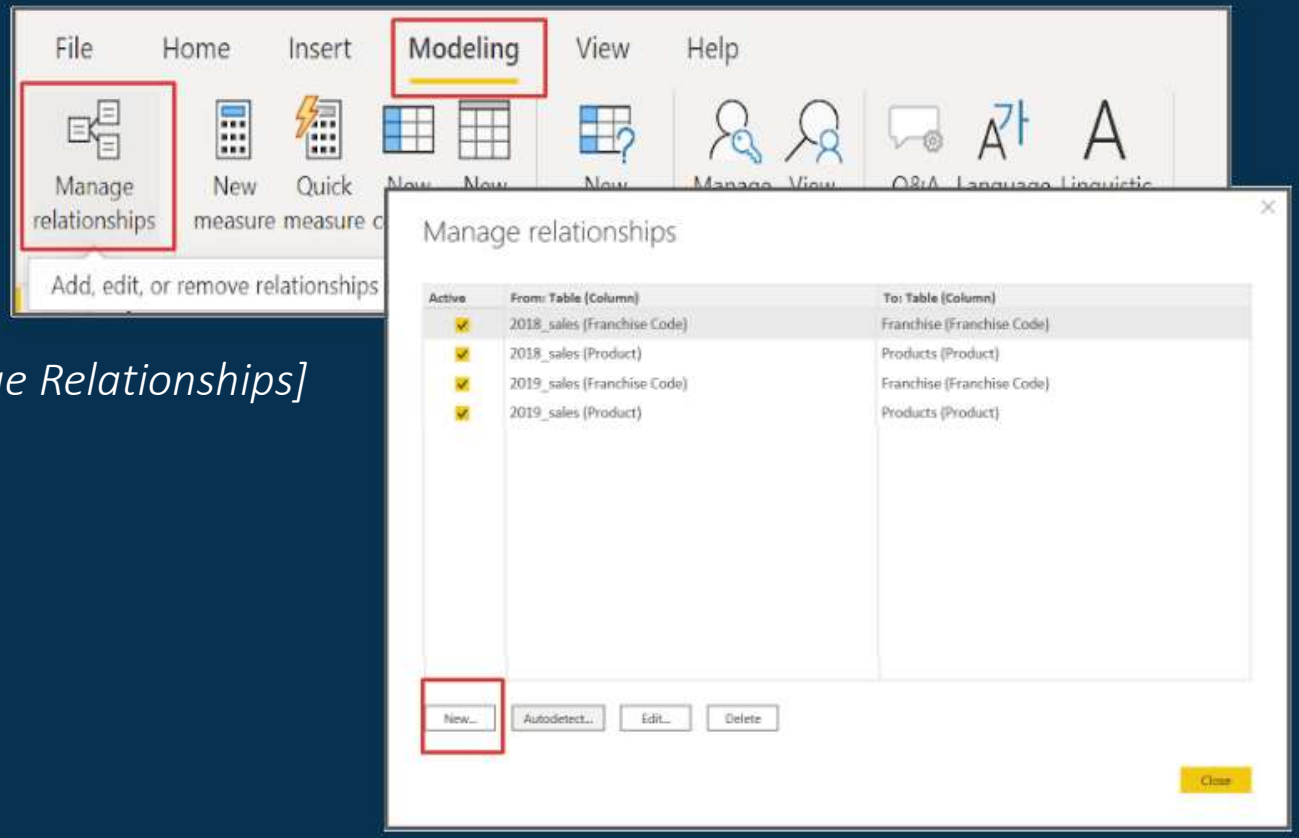
Create a Relationship with “Autodetect”

- In Model View
- Go to *[Modelling Tab] > [Manage Relationships]*
- Click on *Autodetect*



Create a Relationship Manually

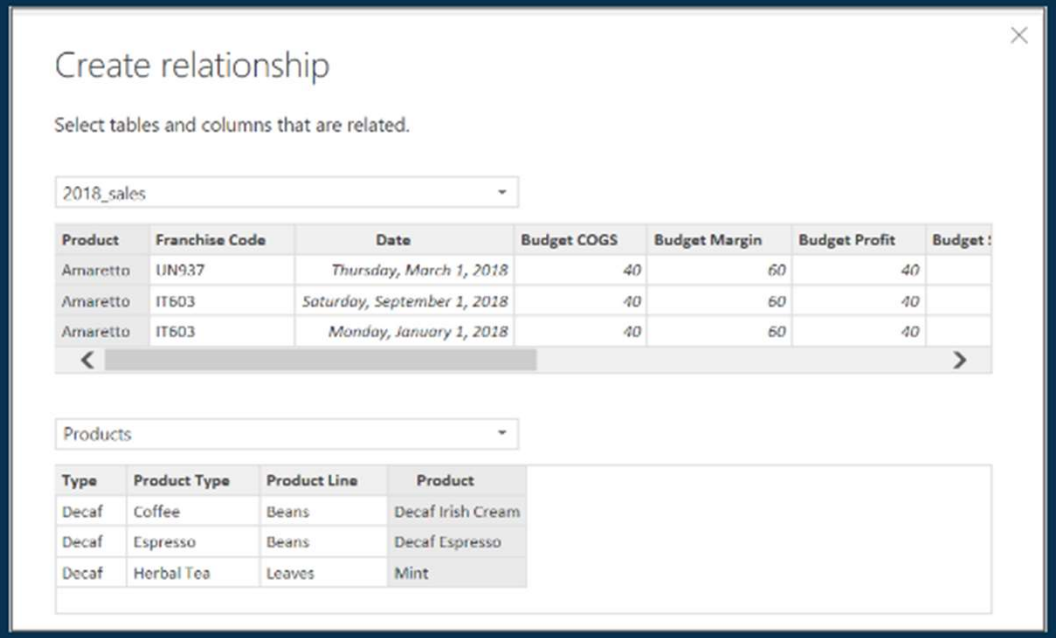
- In Model View
- Go to *[Modelling Tab] > [Manage Relationships]*
- Click on *New*



Create a Relationship Manually

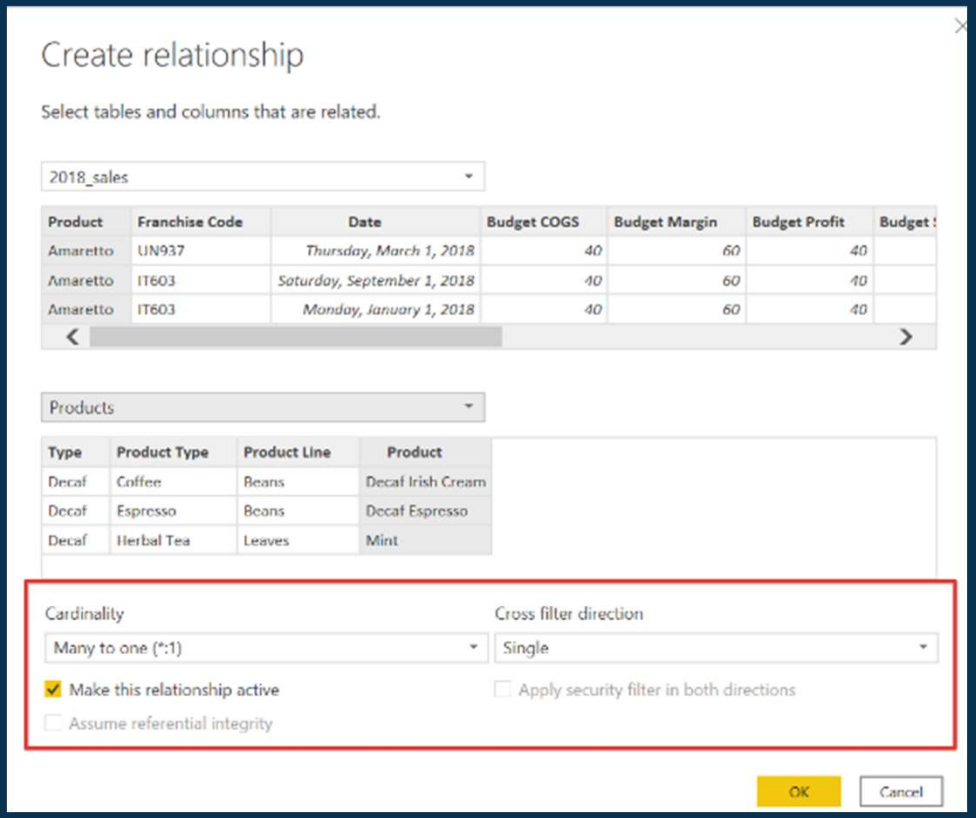
- In the Create relationship dialog box, in the first table drop-down list, select a table
- In the second table drop-down list, select the other table you want in the relationship

› Select the desired columns from both tables and click OK you want to use, and then elect OK.



Create a Relationship Manually

- By default, Power BI Desktop automatically configures the following options
  - Cardinality (direction),
  - Cross filter direction, and
  - Make this relationship active for new relationships



Configure additional relationship options

- The Cardinality option can have one of the following settings:
  - Many to one (*:1): The column in one table can have more than one instance of a value, and the other related table, often known as the lookup table, has only one instance of a value
  - One to one (1:1): The column in one table has only one instance of a particular value, and the other related table has only one instance of a particular value
  - One to many (1:*): The column in one table has only one instance of a particular value, and the other related table can have more than one instance of a value
  - Many to many (*:*): Removes requirements for unique values in tables while using composite models

Configure additional relationship options

- The Cross filter direction option can have one the following settings:
- Both:
  - For filtering purposes, both tables are treated as if they're a single tables
  - The Both setting works well with a single table that has a number of lookup tables that surround it
- Single:
  - The most common, default direction, which means filtering choices in connected tables work on the table where values are being aggregated.

Difference between relationship and merge

- Merge Queries (Data Transformation using power Query)
  - flattens all data into a single table
- Establishing relationships (Data Modelling using Power BI Desktop)
  - Data is stored in separate tables and only a relationship is defined between them based on a common column
  - This helps in building an efficient data model
- Modelled Relationship are preferred over merge queries unless operations like union or join are involved

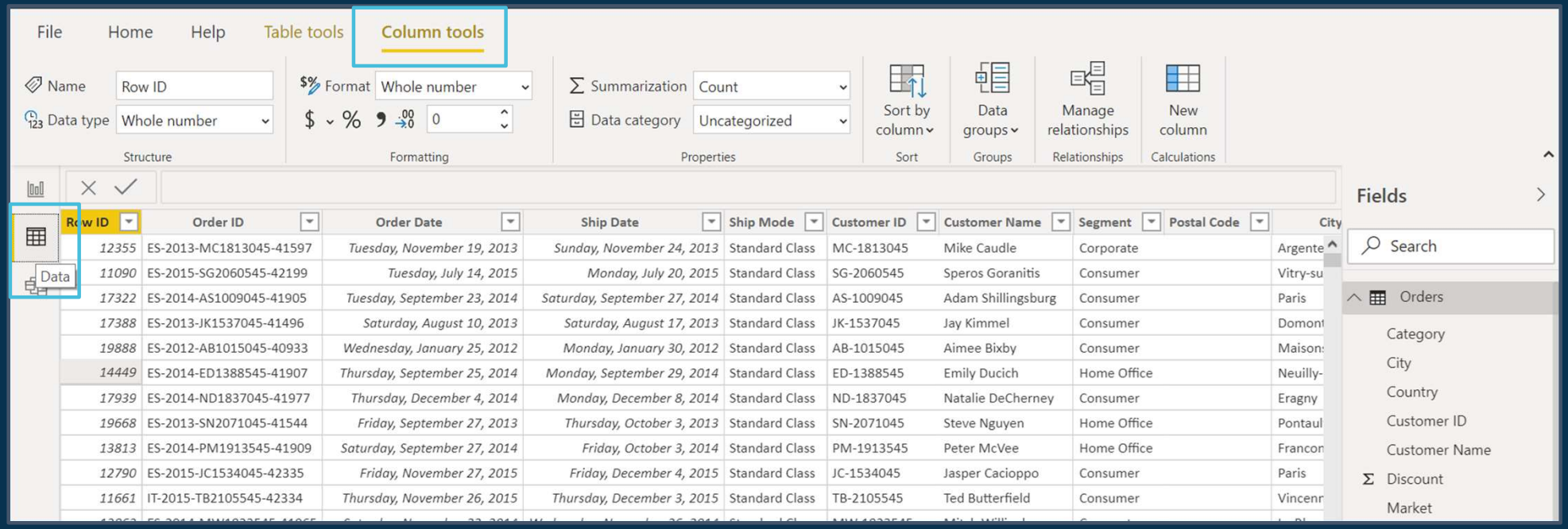
Data Tab

Learning Objective –

- Column Tools
- Sort by Column
- Data Groups | Groups | Bins
- Creating Hierarchies

Data Tab

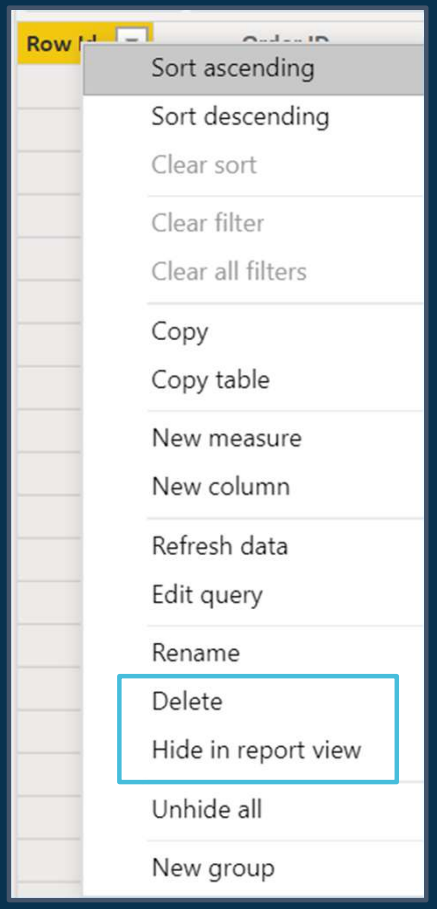
- *Data Tab* gives a preview of data and also provides *Column tools* to modify the default properties of a selected column



Row ID	Order ID	Order Date	Ship Date	Ship Mode	Customer ID	Customer Name	Segment	Postal Code	City
12355	ES-2013-MC1813045-41597	Tuesday, November 19, 2013	Sunday, November 24, 2013	Standard Class	MC-1813045	Mike Caudle	Corporate		Argente
11090	ES-2015-SG2060545-42199	Tuesday, July 14, 2015	Monday, July 20, 2015	Standard Class	SG-2060545	Speros Goranitis	Consumer		Vitry-su
17322	ES-2014-AS1009045-41905	Tuesday, September 23, 2014	Saturday, September 27, 2014	Standard Class	AS-1009045	Adam Shillingsburg	Consumer		Paris
17388	ES-2013-JK1537045-41496	Saturday, August 10, 2013	Saturday, August 17, 2013	Standard Class	JK-1537045	Jay Kimmel	Consumer		Domont
19888	ES-2012-AB1015045-40933	Wednesday, January 25, 2012	Monday, January 30, 2012	Standard Class	AB-1015045	Aimee Bixby	Consumer		Maisons
14449	ES-2014-ED1388545-41907	Thursday, September 25, 2014	Monday, September 29, 2014	Standard Class	ED-1388545	Emily Ducich	Home Office		Neuilly-
17939	ES-2014-ND1837045-41977	Thursday, December 4, 2014	Monday, December 8, 2014	Standard Class	ND-1837045	Natalie DeCherney	Consumer		Eragny
19668	ES-2013-SN2071045-41544	Friday, September 27, 2013	Thursday, October 3, 2013	Standard Class	SN-2071045	Steve Nguyen	Home Office		Pontaul
13813	ES-2014-PM1913545-41909	Saturday, September 27, 2014	Friday, October 3, 2014	Standard Class	PM-1913545	Peter McVee	Home Office		Francon
12790	ES-2015-JC1534045-42335	Friday, November 27, 2015	Friday, December 4, 2015	Standard Class	JC-1534045	Jasper Cacioppo	Consumer		Paris
11661	IT-2015-TB2105545-42334	Thursday, November 26, 2015	Thursday, December 3, 2015	Standard Class	TB-2105545	Ted Butterfield	Consumer		Vincenr

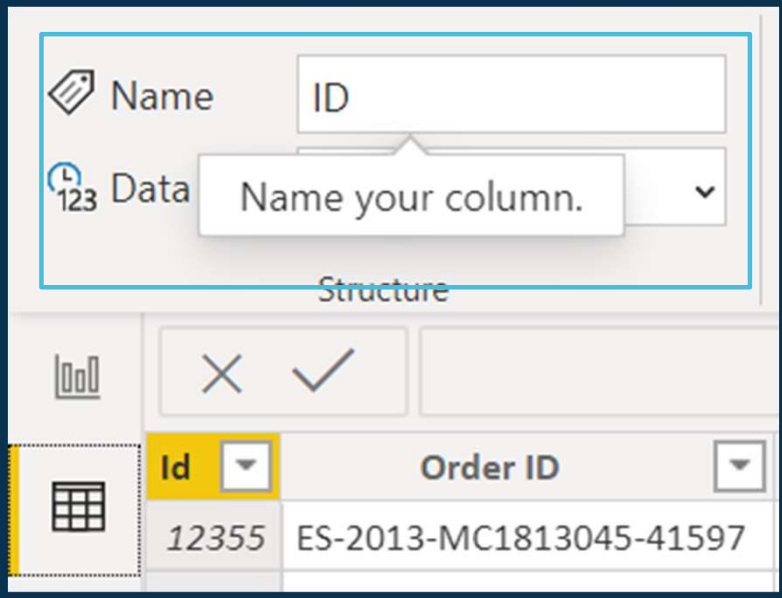
Hide or Delete a column

- Select the column
- Right click on the column header
- Select Delete or Hide in report view from the drop down



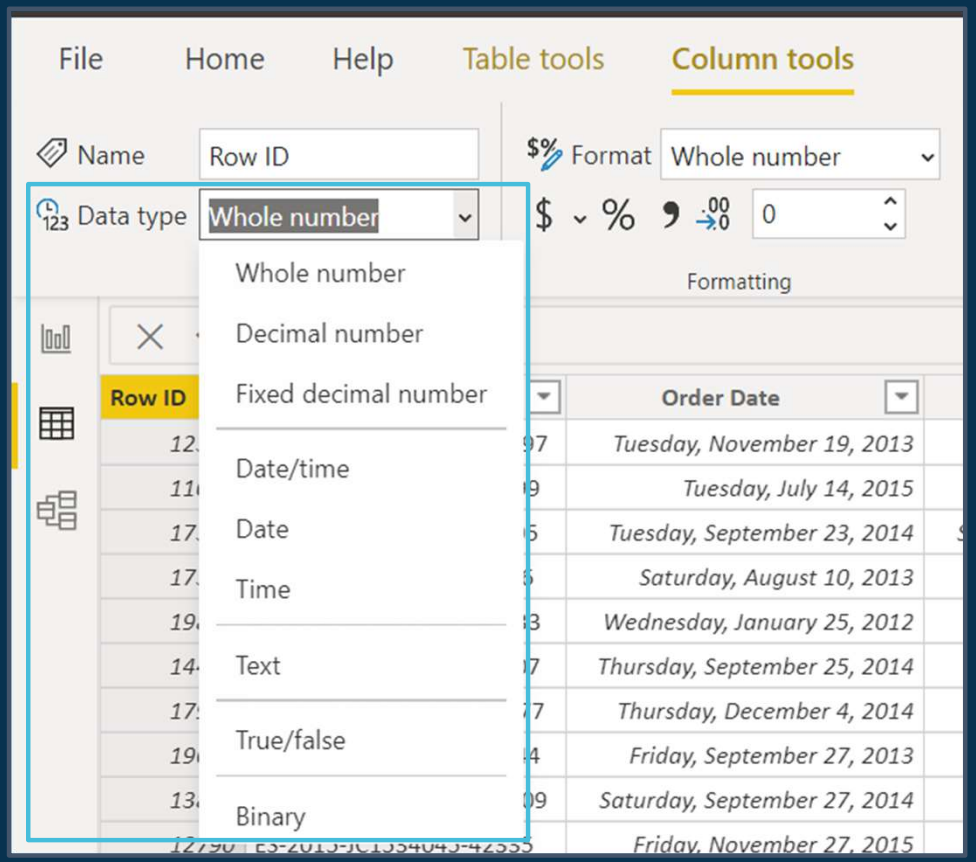
Renaming a column

- Select the column to be renamed
- In Structure Ribbon > Name assign new name to column



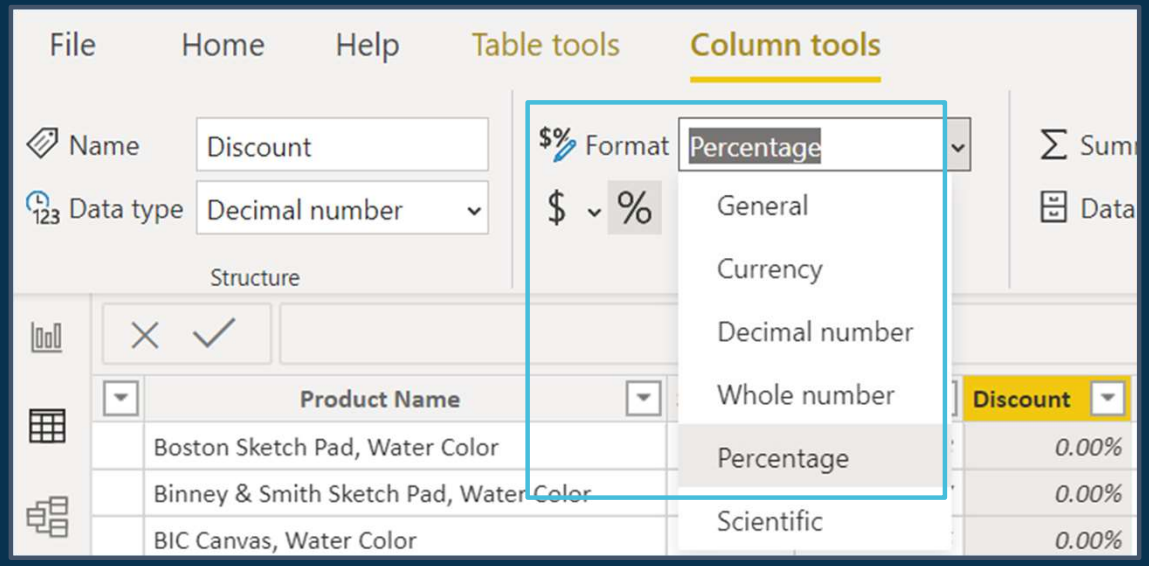
Changing data type for a column

- When data is imported a model, the model designer will automatically detect assign data types
- The data type of a column can be manually changed depending on how data is stored
- To change the data type
- Select the column
- select Structure Ribbon > data type to change data type



Changing data display format for a column

- To change the data display format for a column
  - Select the column
  - Select Formatting Ribbon > Format to change display format



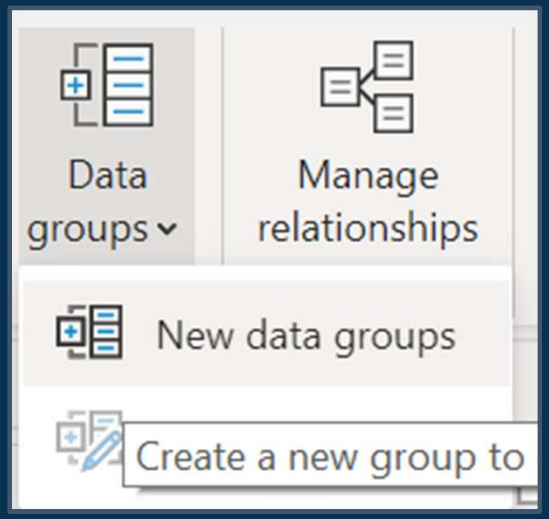
Data Groups

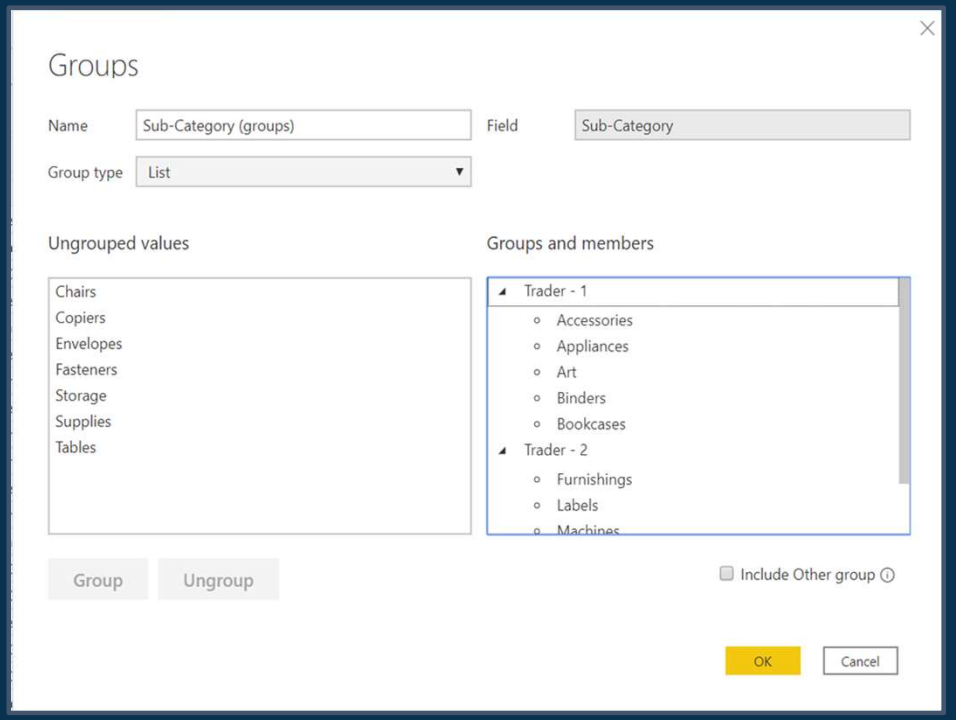
- Power BI Desktop aggregates data into chunks (or groups) based on values found in the underlying data.

- It allows to group data points to help analyze and explore data and trends in visuals.

Creating a grouped column

- *Data tools > New data group*



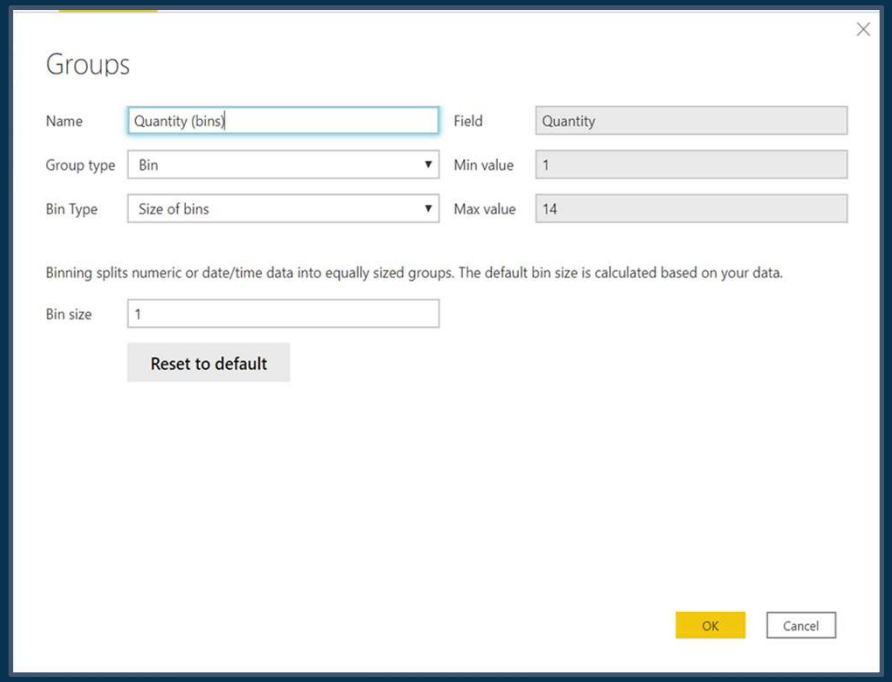


Creating a grouped column

- In the Groups dialog box
  - create new groups or modify existing groups.
  - rename any group
  - Add items from the Ungrouped values list into a new group or into one of the existing groupsTo create a new group, select two or more items (using Ctrl+click) from the Ungrouped values box, and then select the Group button below that box
  - Add an ungrouped value into an existing group - just select the one of the Ungrouped values, then select the existing group to add the value to, and select the Group button
  - To remove an item from a group - select it from the Groups and members box, and then select Ungroup
  - Move ungrouped categories into the Other group or leave them ungrouped

Binning

- A bin can be defined with a bin size to put values into equally sized groups that better enable to visualized data in ways that are meaningful
- This action is often called binning
- Binning works with numeric and date fields

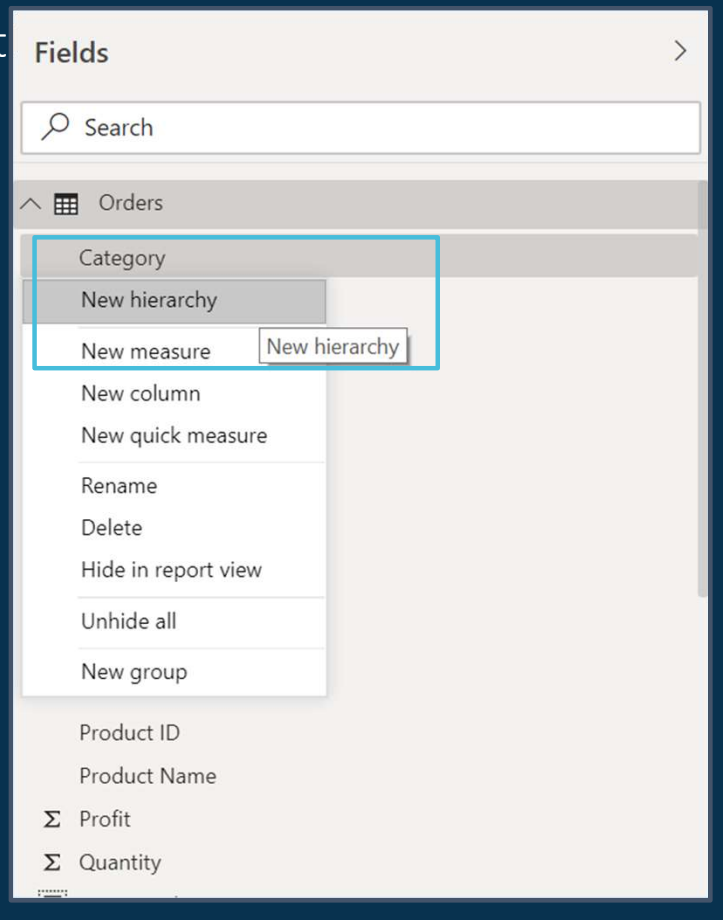


Creating hierarchies

- To help users navigate through the data, hierarchies can be created
- Hierarchies assembles the fields into the stages which helps understand the different levels in data
- Hierarchies also aids in representing the leveled data into interactive visuals

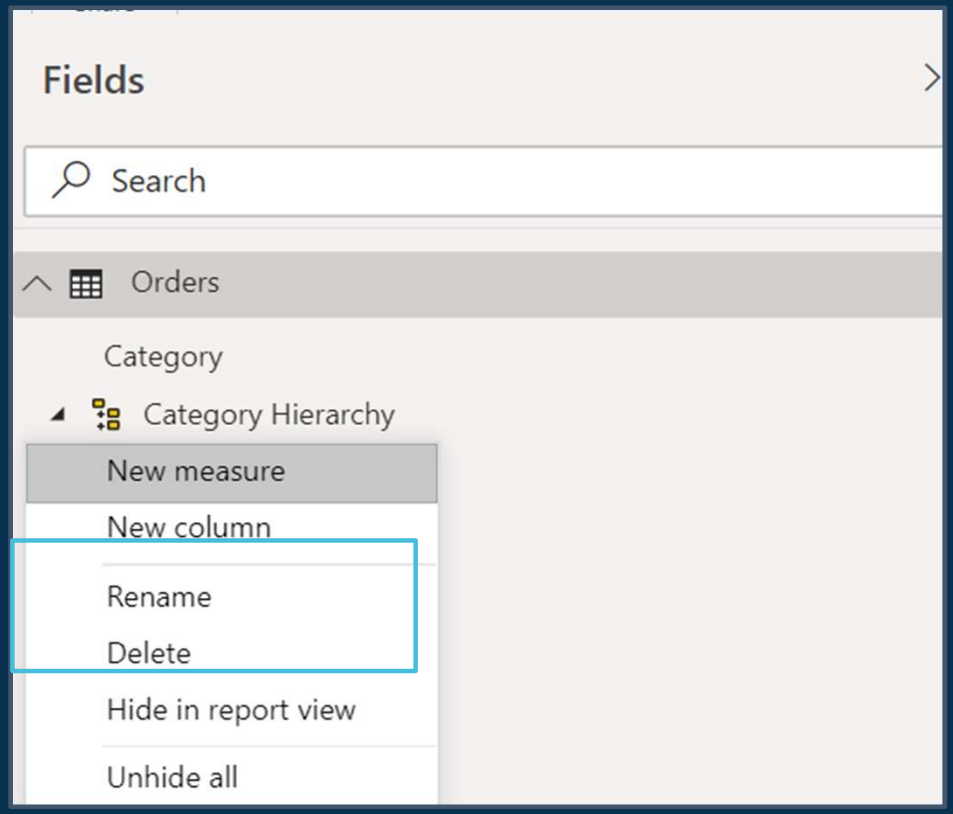
Create a new hierarchy

- From the Fields pane right click the field to be added into t
- Click New Hierarchy



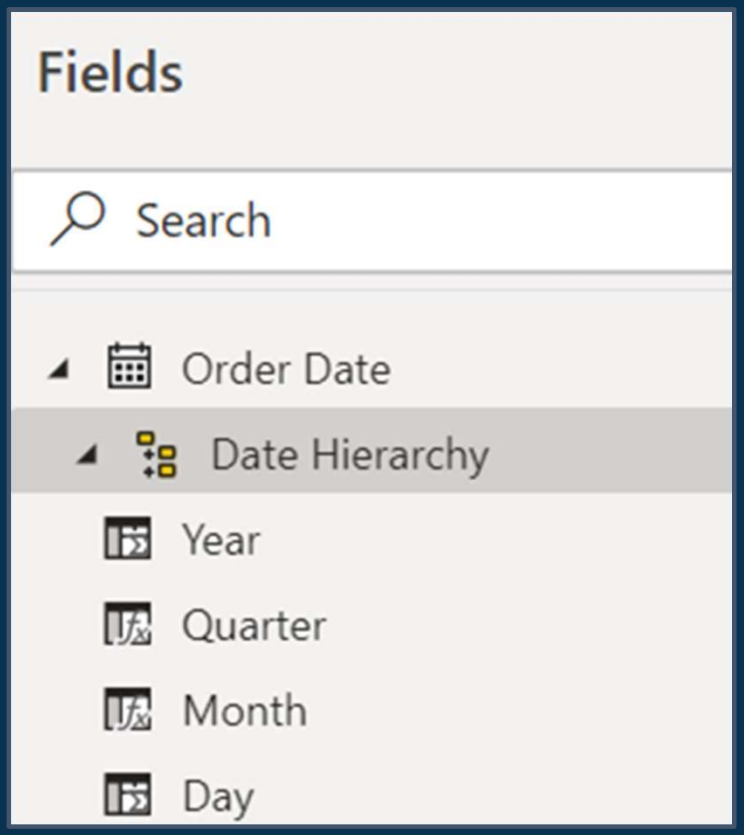
Rename/delete hierarchy

- From the Fields pane right click the hierarchy field to be modified
- Click Rename or Delete



Date hierarchies

- Date fields in Power BI come with an implicit hierarchy.
- A date field is split into Year, Quarter, Month, Day



DAX Expressions

Learning Objective –

- Creating Columns
- Creating Measures
- Date Table

Introduction to DAX

- DAX stands for Data Analysis Expressions.
- It is a formula expression language used in Analysis Services, Power BI Desktop, and Power Pivot in Excel.
- DAX formulas include functions, operators, and values to perform advanced calculations and queries on data in related tables and columns in tabular data models.
- DAX calculation formulas are used in measures, calculated columns, calculated tables, and row filters.

Introduction to DAX

- DAX can be used in –
  - calculated columns
  - measure
  - calculated tables
- It can also be used in defining filter for roles in Row-Level-Security

Introduction to DAX

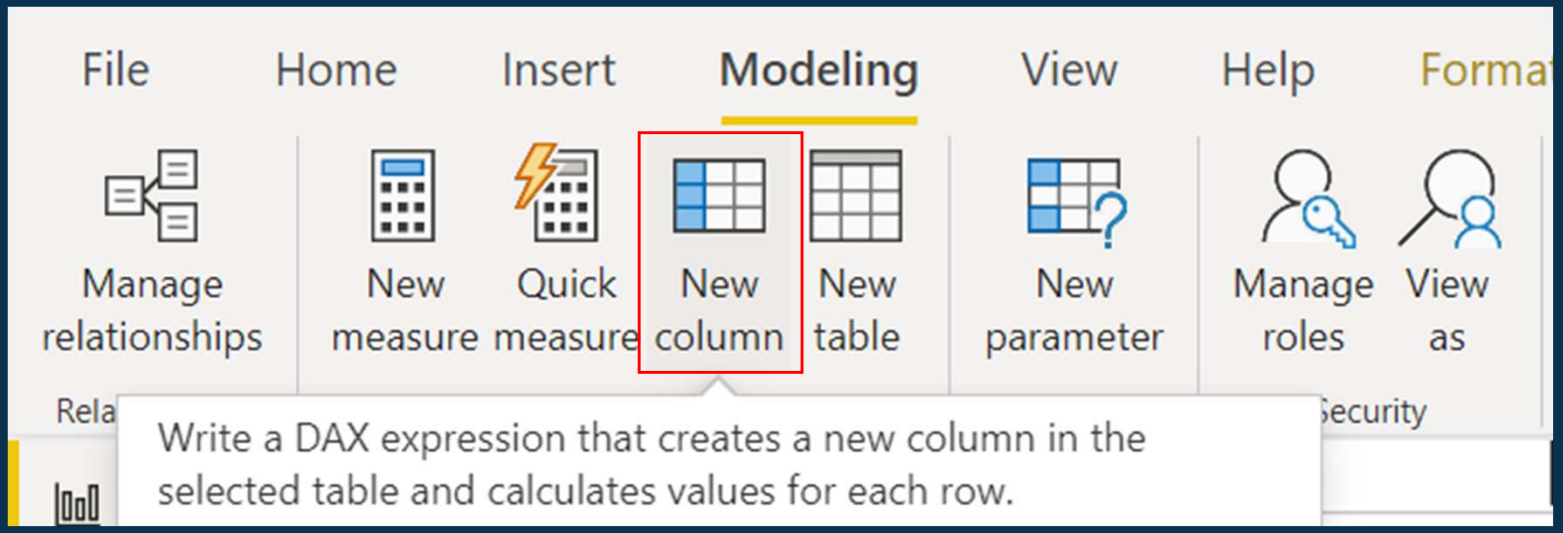
- DAX formula consist of expressions that return results either as a table object or a scalar
- DAX formula contain –
  - DAX functions
  - DAX operators
  - References to model objects
  - DAX and constant variables
  - Whitespaces

Calculated Columns

- A calculated column added to an existing table and then assigned a DAX formula that defines the column's values
- Since a calculated column is created in a table in the data model, it is not supported in models that retrieve data exclusively from a relational data source using DirectQuery mode
- When a calculated column contains a valid DAX formula, values are calculated for each row as soon as the formula is entered
- Values are then stored in the in-memory data model

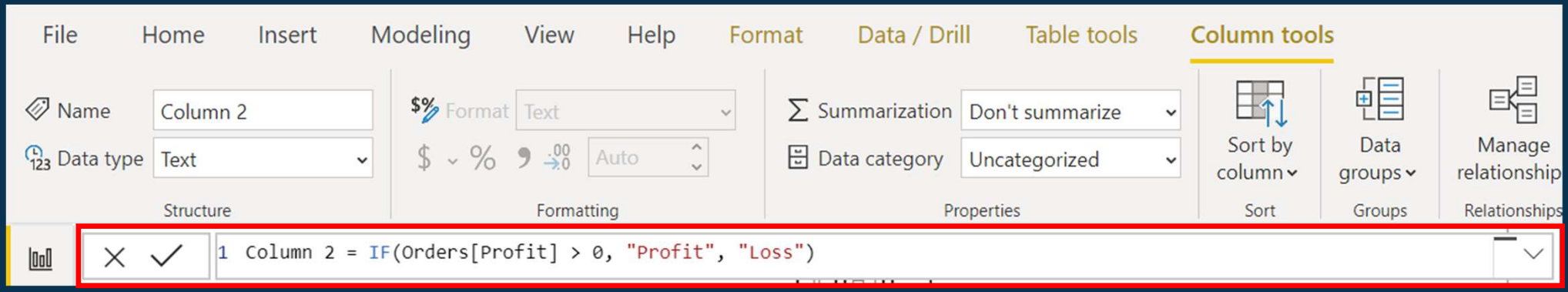
Creating Calculated Columns

- Click *Modeling tab > Calculations Ribbon > New Column*



Creating Calculated Columns

- The DAX formula for the calculated column can be entered in the formula bar

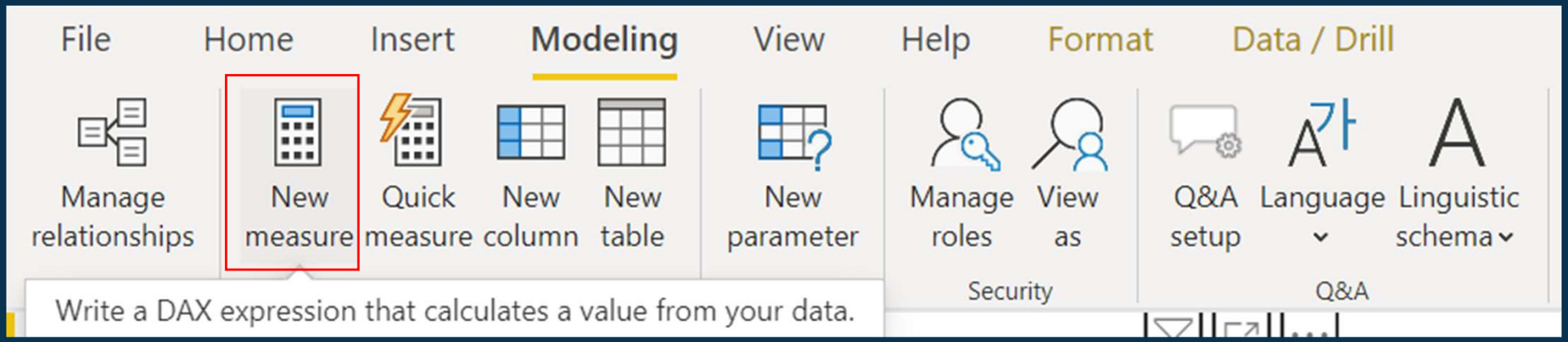


Measures

- Measures are dynamic calculation formulas where the results change depending on context
- A formula in a measure can use standard aggregation functions automatically created by using the Auto sum feature, such as COUNT or SUM, or can be defined by the user using the DAX formula bar
- Named measures can be passed as an argument to other measures.
- The syntax for a measure includes the measure's name preceding the formula

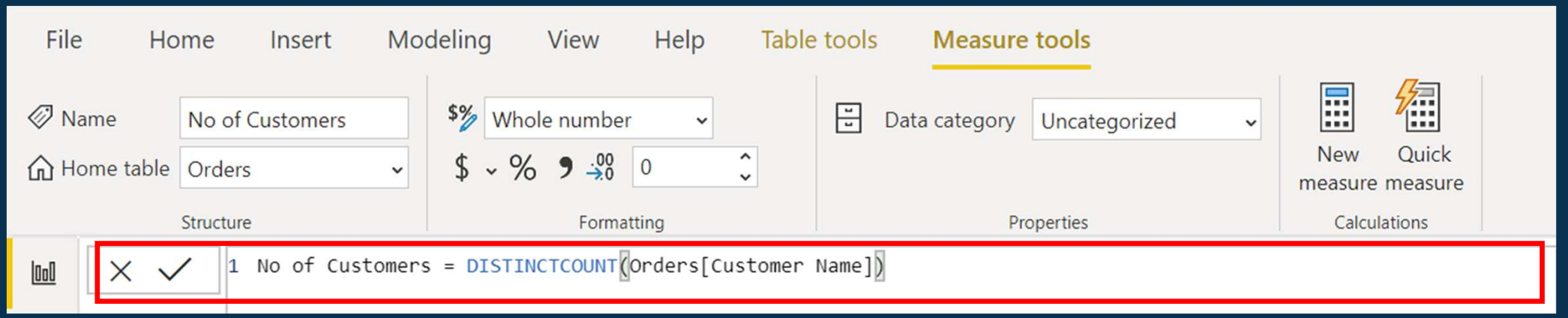
Creating Measure

- Click *Modeling tab > Calculations Ribbon > New Measure*



Creating Measure

- The DAX formula for the measure can be entered in the formula bar



Difference between Calculated Column and Measure

- The value of a calculated column is computed during data refresh and uses the current row as a context; it does not depend on user interaction in the report.

- A measure operates on aggregations of data defined by the current context, which depends on the filter applied in the report – such as slicer, rows, and columns selection in a pivot table, or axes and filters applied to a chart.

Date Table

- Date tables in Power BI are tables that only contain date-related data.
- It is a standard dimension table that can be used to reference dates in your model and analyze data based on these dates.
- They are also useful for time intelligence calculations and when creating reports that require precise date information.
- Date tables allow you to slice and dice your data by date attributes such as weekday, month, quarter, and year.
- They also allow you to use DAX time intelligence functions that can be used to generate Fiscal year calculations

Date Table

- A date table is a table that meets the following requirements:
  - It must have a column of data type date (or date/time)—known as the date column.
  - The date column must contain unique values.
  - The date column must not contain BLANKs.
  - The date column must not have any missing dates.
  - The date column must span full years. A year isn't necessarily a calendar year (January-December).
  - The date table must be marked as a date table.

Optimizing Data Model

Learning Objective –

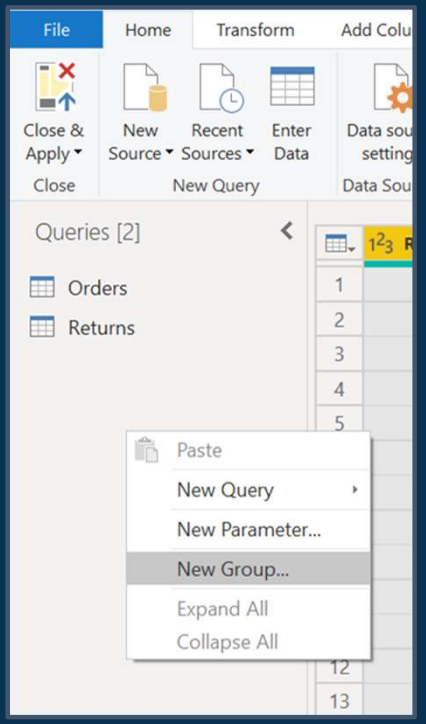
- Storage Modes
- Using Parameter in PQE
- Query Groups
- Query Folding
- Query Dependency
- Query Diagnostic

Query groups

- Query Groups help users better organize their queries within a given workbook, as well as perform bulk operations on all queries within a group (such as Refresh)
- The tables from the source are arranged in alphabetical order
- In real-time, these tables can be grouped based on the content

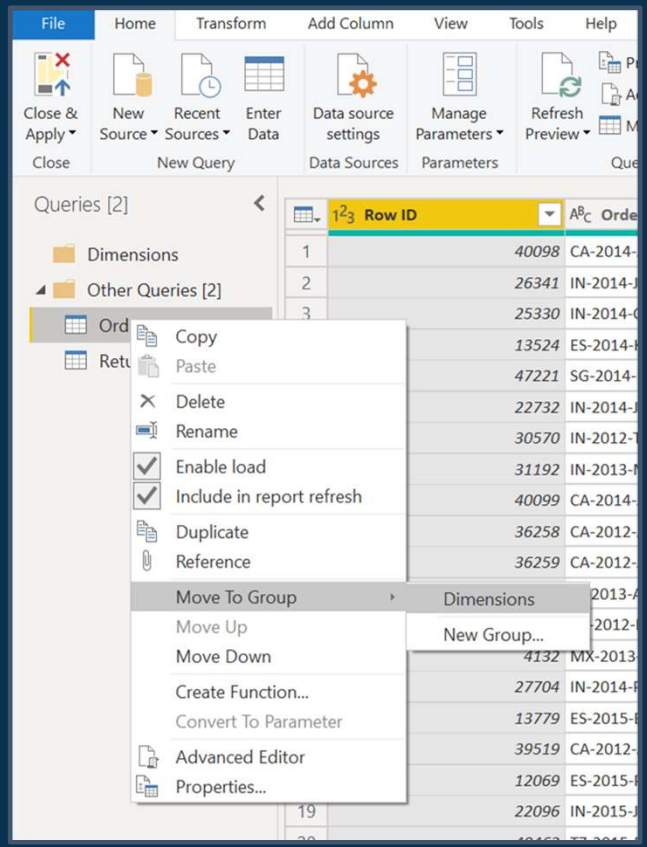
Create query group

- To create a new query group, right-click on the empty space under the queries section and select the New Group option from the context menu.
- Selecting the New Group option will open a new window
- Please specify the Group Name and the description. For now, we are assigning the Dimensions as the Group name.



Add tables to query group

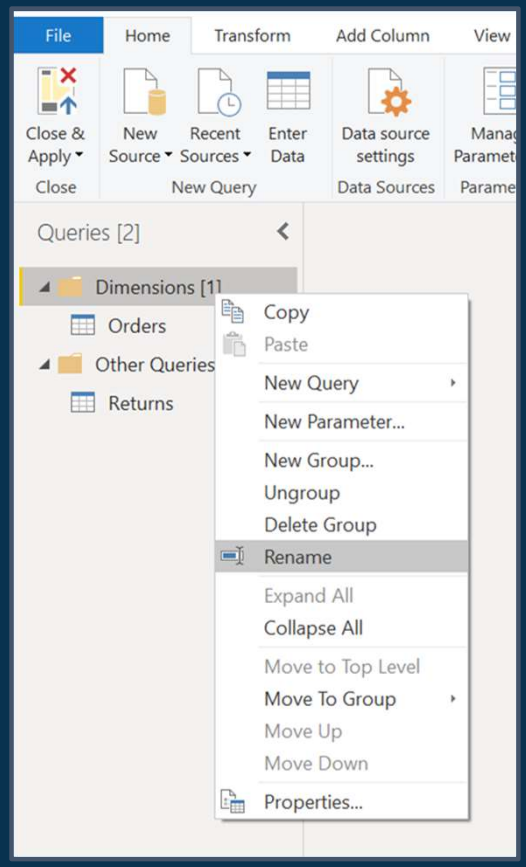
- Right-click on the table that you want to add opens the context menu.
- Please select the Move to Group option, and then select Group Name, i.e., Dimensions.



Rename query group

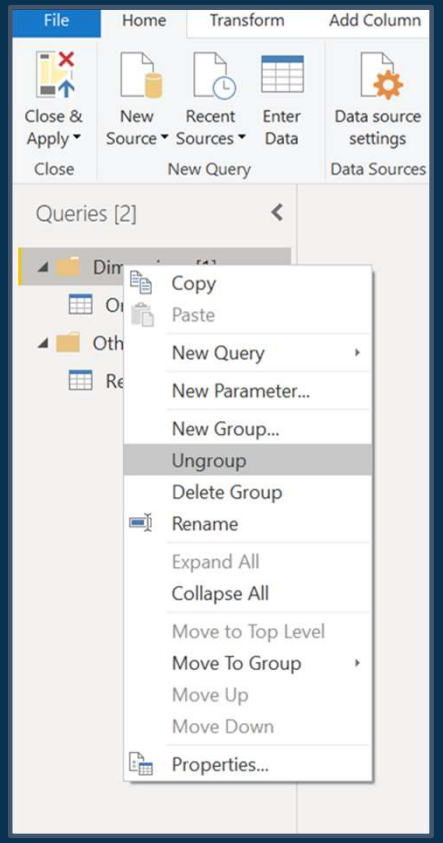
- Right-click on the query group you want to rename and select

  Rename option



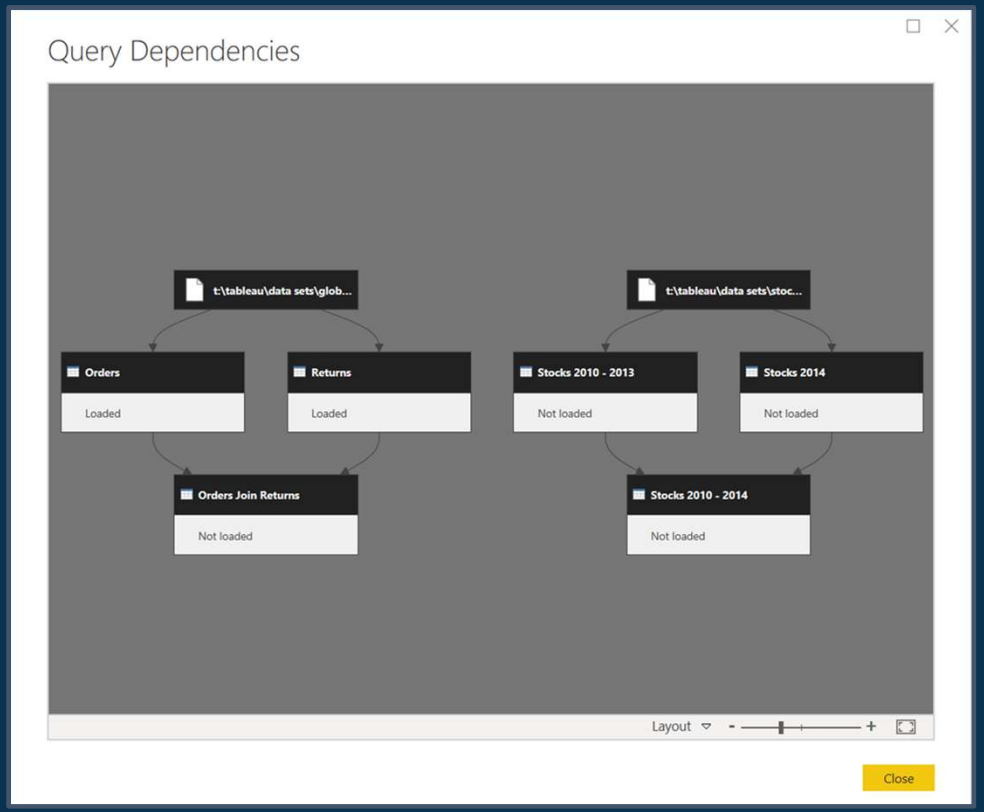
Ungroup query group

- Right-click on the query group and select Ungroup option
- This will ungroup all the tables from the query group



Query dependency view

- To identify various queries being used
- To make out the source of the queries.
- Go to View tab > Query Dependencies
- This opens adjoining chart



Using Parameter (in PQE) to store database information

- A parameter serves as a way to easily store and manage a value that can be reused.
- Parameters give you the flexibility to dynamically change the output of your queries depending on their value
- You can easily manage your parameters inside the Manage Parameters window.
- To get to the Manage Parameters window,
  - select the *Home tab > Manage Parameters* in Power Query Editor
- Example – Create a parameter to store database information

Points to Remember –

When using an SQL query to import data –

- Avoid using the wildcard character (*) in your query. If used in SELECT statement, it will import all columns may lead to redundant data in data model, which will cause performance issues and require additional steps to normalize your data for reporting.
- All queries should also have a WHERE clause. This clause will filter the rows to pick only filtered records.
- It is a best practice to avoid writing SQL query directly in Power BI. Instead, consider writing a query in SQL and import a view. If Power BI uses a view, when it retrieves data, it participates in query folding, a feature of Power Query.
- Refer this link for more information on query folding - https://learn.microsoft.com/en-us/power-query/power-query-folding

Storage Modes

Storage modes gives control caching of data in memory for report in Power BI Desktop

Power BI supports three different methods for connecting to data -

- Import Data
- DirectQuery
- Dual

Storage Modes - Import

- Import data is also referred as schedule refresh
- Data from the source will be loaded into Power BI, consuming memory and disk space
- During development phase data will be stored on local machine with Power BI Desktop, then it would be memory and disk space of local machine
- Once the report is published to the website, then it will be memory and disk space of Power BI cloud machine.

Storage Modes - DirectQuery

- DirectQuery is a direct connection to data source
- Data will NOT be stored in Power BI model
- Power BI will be a visualization layer, then query the data from data source every time
- Power BI will only store metadata of tables but not the data
- Power BI file size will be much smaller, without hitting size limitation
- It is suitable in cases where data changes frequently and near real-time reporting is required.
- It can handle large data without the need to pre-aggregate
- Data modeling and transform limitations

Storage Modes - Live Connection (Limited Use)

- Live Connection is very similar to DirectQuery in the way that it works with the data source
- It will not store data in Power BI, and it will query data source every time.
- Live Connection is only supported for data sources with inbuilt data engines, like -
  - SQL Server Analysis Services (SSAS) Tabular
  - SQL Server Analysis Services (SSAS) Multi-Dimensional
  - Power BI Service
- Power BI fetches all model metadata and modeling requirements are handled in the data source, and Power BI just surface that data through Visualization

Query Folding

- Query folding is the ability for a Power Query to generate a single query statement to retrieve and transform source data.
- The Power Query mashup engine strives to achieve query folding whenever possible for reasons of efficiency.
- Query folding may occur for an entire Power Query query, or for a subset of its steps.
- When query folding cannot be achieved—either partially or fully—the Power Query mashup engine must compensate by processing data transformations itself.
- Data sources such as relational databases, OData feeds (including SharePoint lists), Exchange, and Active Directory supports query folding. However, data sources like flat files, blobs, and web typically do not.

Optimize Performance in Power Query

- Use query folding
- The benefits to query folding include –
  - More efficiency in data refreshes and incremental refreshes. When you import data tables by using query folding, Power BI is better able to allocate resources and refresh the data faster because Power BI does not have to run through each transformation locally.
  - Automatic compatibility with DirectQuery and Dual storage modes
- Use Query diagnostics
- Process as much data as possible in the original data source. Separate date and time, if bound together
- Use native SQL queries. When using DirectQuery for SQL databases, avoid pulling data from stored procedures or common table expressions

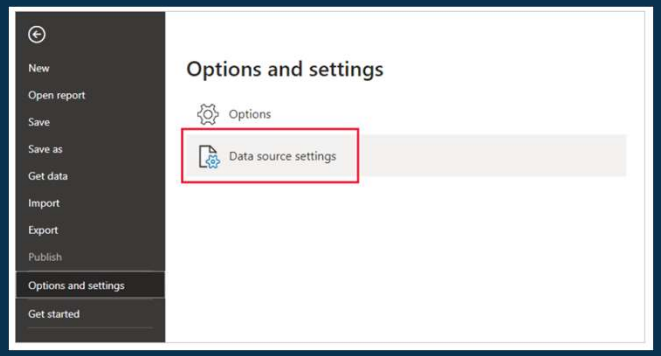
Use or Create PBIDS Files

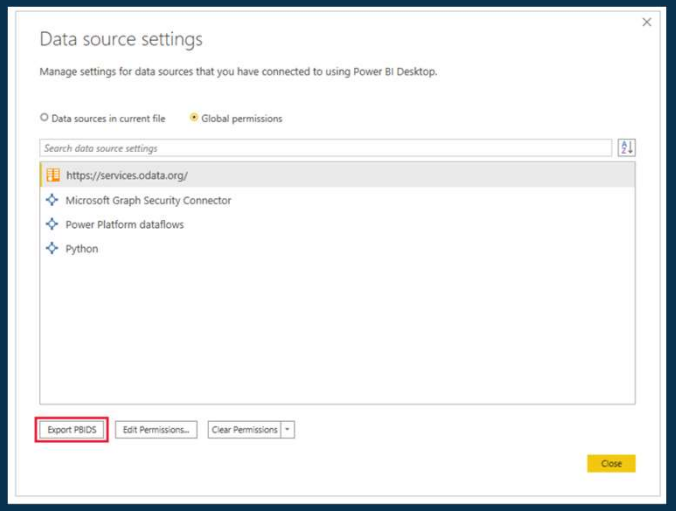
- PBIDS stands for Power BI Data Source file. It has .pbids extension.
- PBIDS file can be used to streamline the Get Data experience for new or beginner report creators in your organization. If you create the PBIDS file from existing reports, it's easier for beginning report authors to build new reports from the same data.
- PBIDS files store the data connection information.
- It will not store the authentication information.

Creating PBIDS Files

- To create the PBIDS file,

1. *File > Options and settings > Data source settings*
2. *In the dialog that appears, select the data source you want to export as a PBIDS, and then select Export PBIDS.*





Visualize data in Power BI

Learning Objective –

- Sorting | Filtering
- Format Pane
- Analytics Pane

Learning Objective

- Sorting
- Filtering
- Page Layout
- Analytics Pane

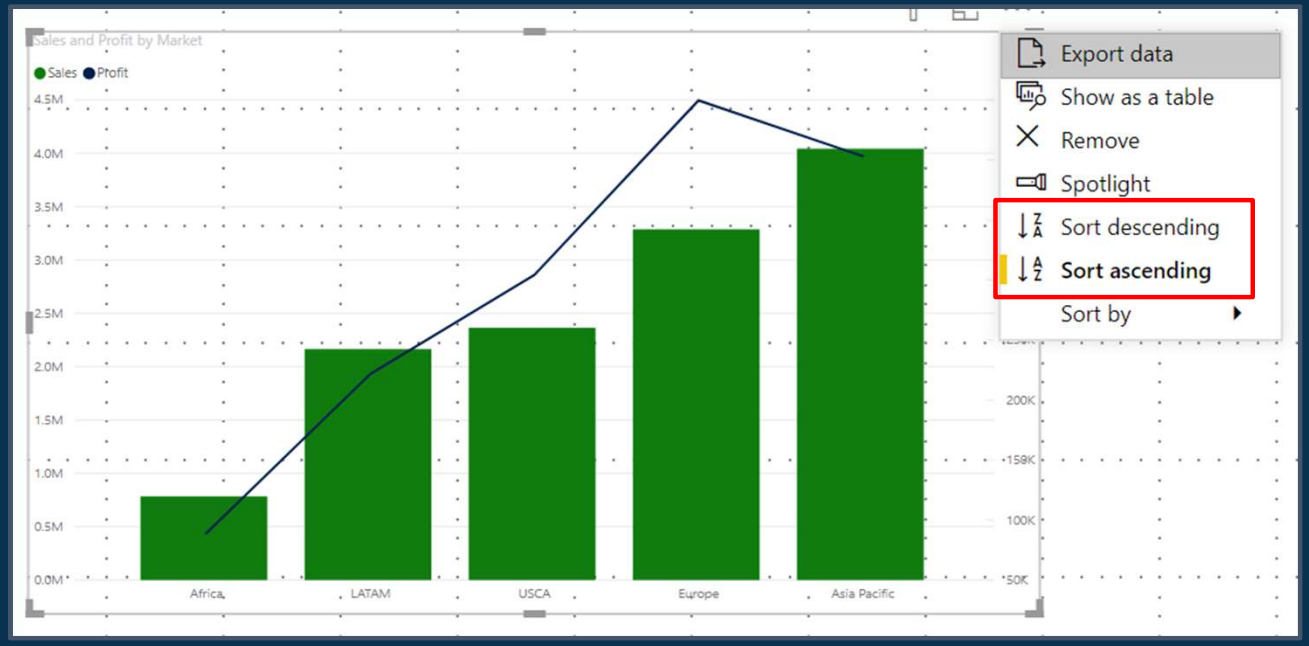
Sorting

Sorting by Columns

- In Power BI Desktop, one can change how a visual looks by sorting it by different data fields
- By changing how sorting method for a visual, the necessary information can be highlighted and conveyed rightly

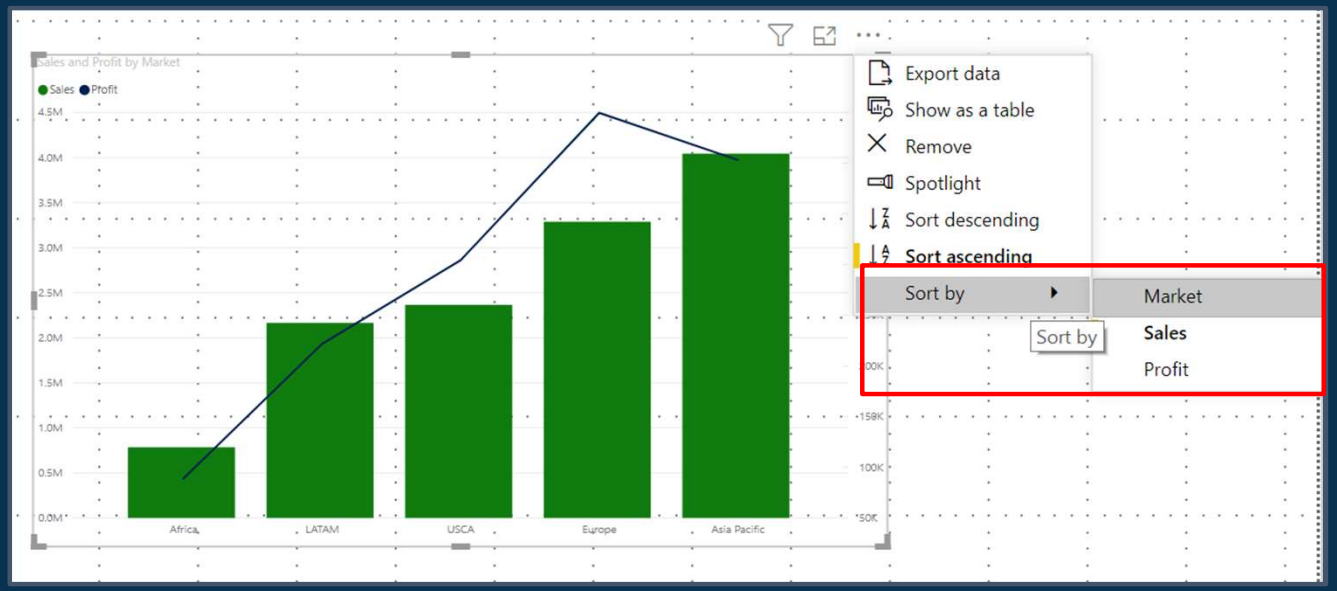
Sorting by Columns

- Once the visual is ready it can be sorted by clicking the more option and select the sort direction as ascending or descending



Sorting by Columns

- *Sort by* option lets us sort the column on a specific field



Filtering

Filtering

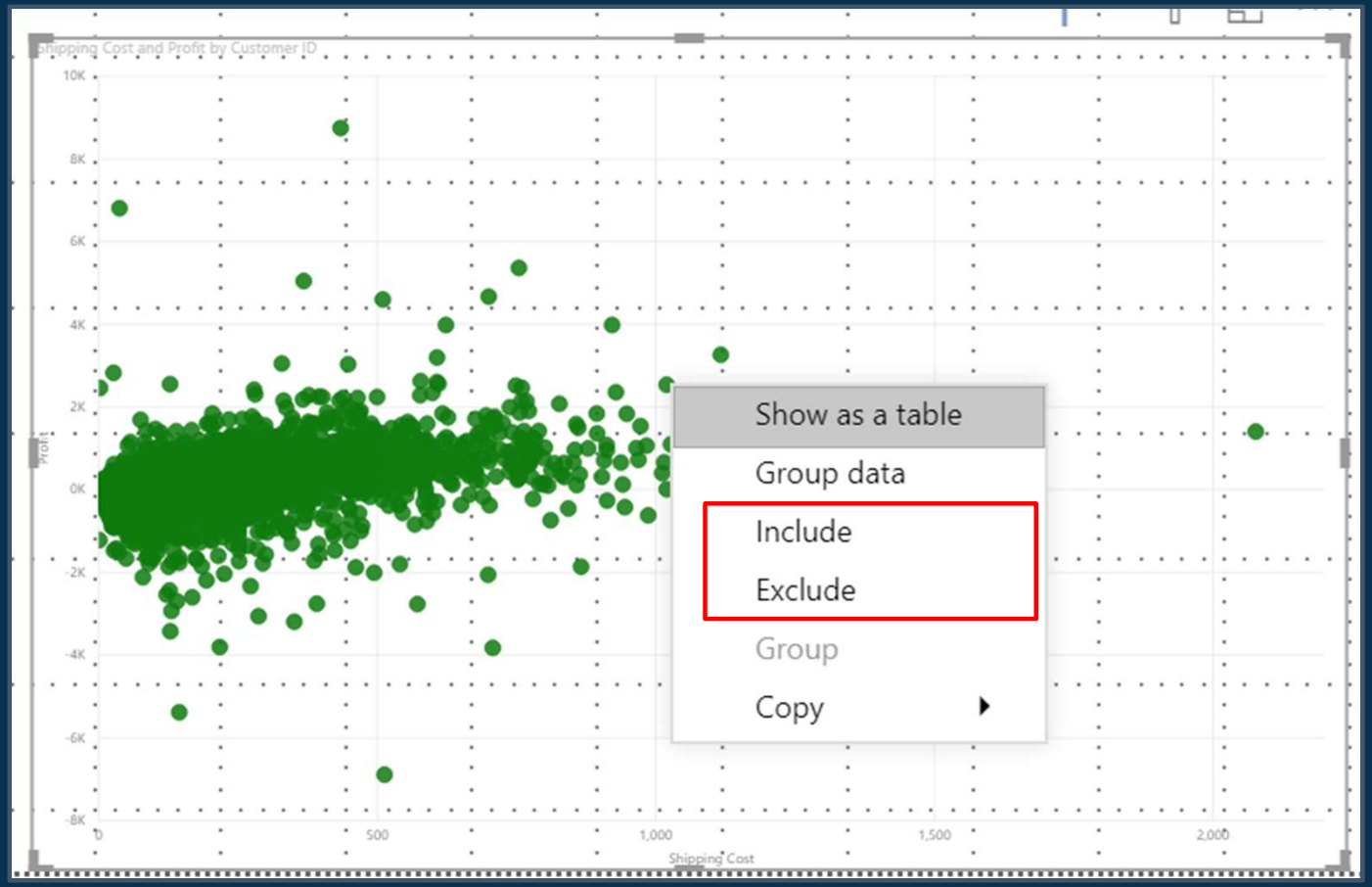
Types of filters -

- Include/exclude filter
- page filter applies to all the visuals on the report page
- visual filter applies to a single visual on a report page.
- report filter applies to all pages in the report
- Drill through filter applies to a single entity in a report
- Slicers

Include/Exclude Filter

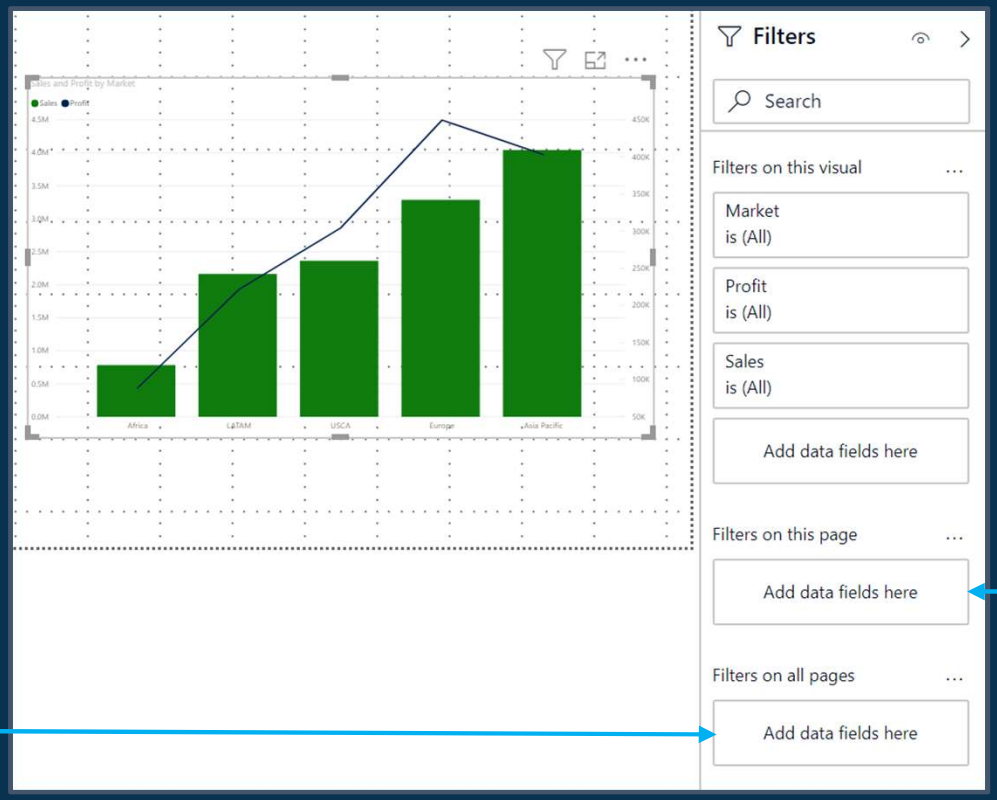
- When exploring data, it’s often useful to focus on a specific set of points
- In addition, you may want to remove specific points because you don’t want to focus on them
- You can now select one or more points and right click to include or exclude points in your visual
- Include will filter out all points except for those you’ve chosen to include. Exclude will filter out the points you’ve chosen to exclude.

Include/Exclude Filter

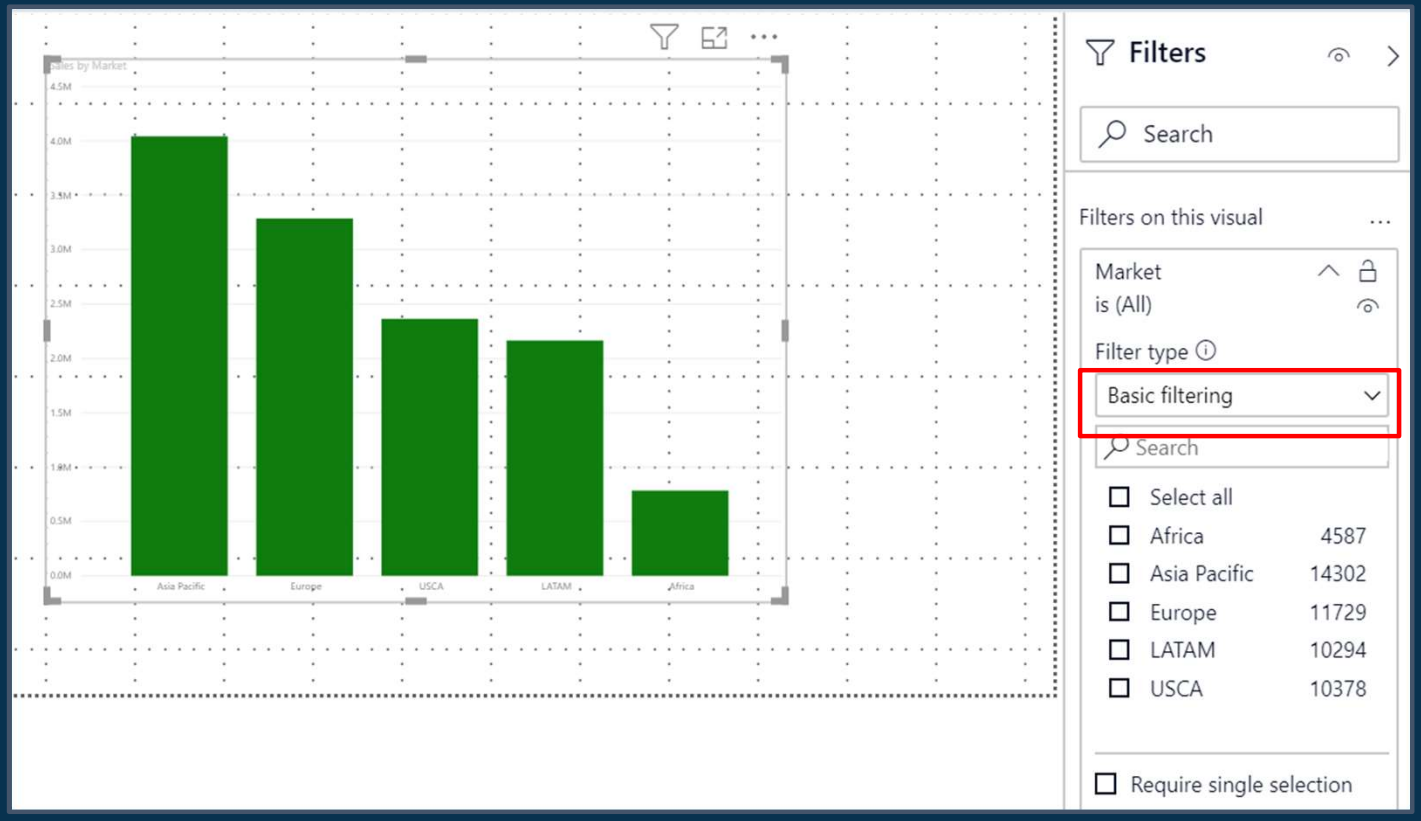


Filtering using Filter Pane

- Filter pane allows us to add fields which can referred to filter the data on the canvas

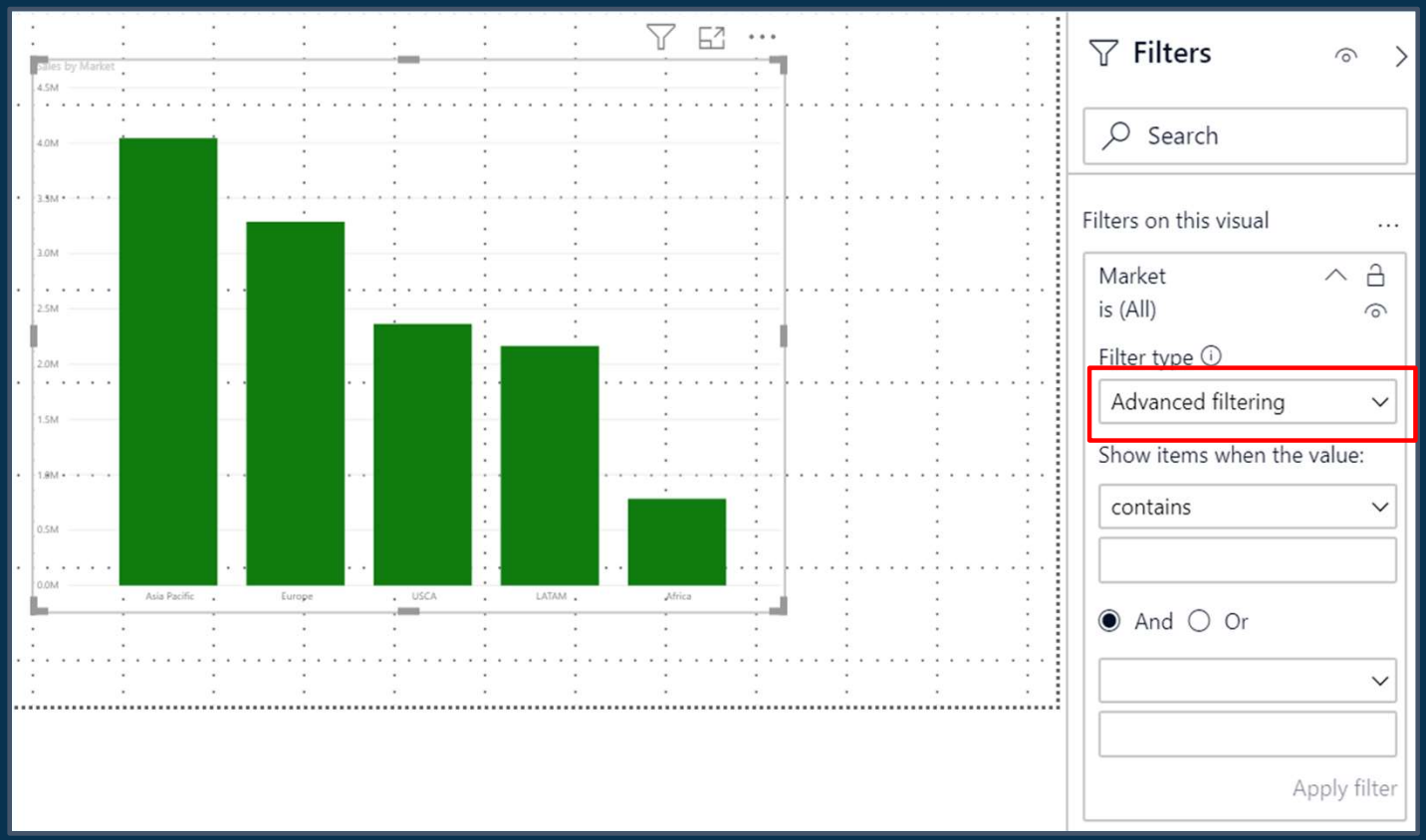


Filtering on Categorical Data



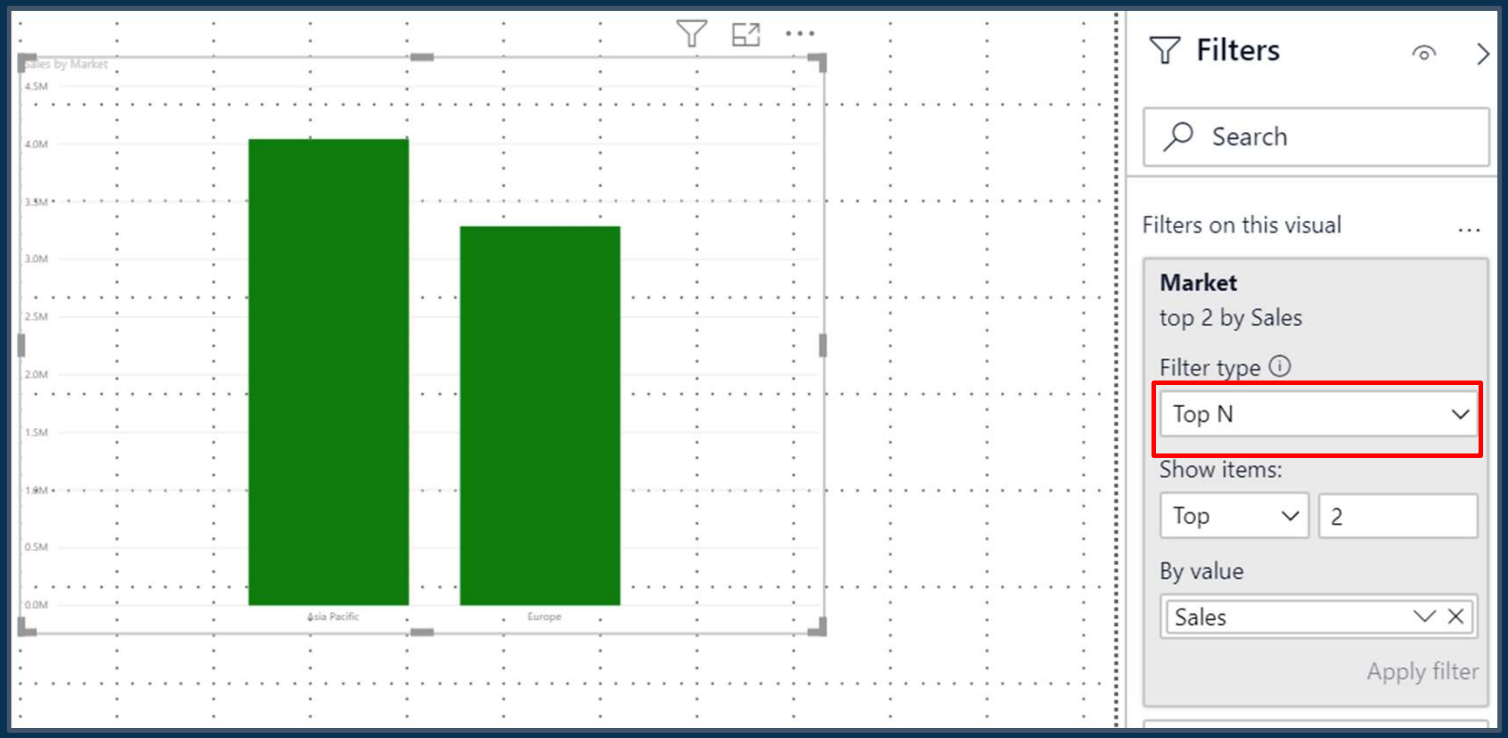
Select the elements to be filtered

Filtering on Categorical Data



Allows to filter on the basis of text values - contains, starts with, ends with etc.

Filtering on Categorical Data



Filters the Top or bottom N number of elements based on a field

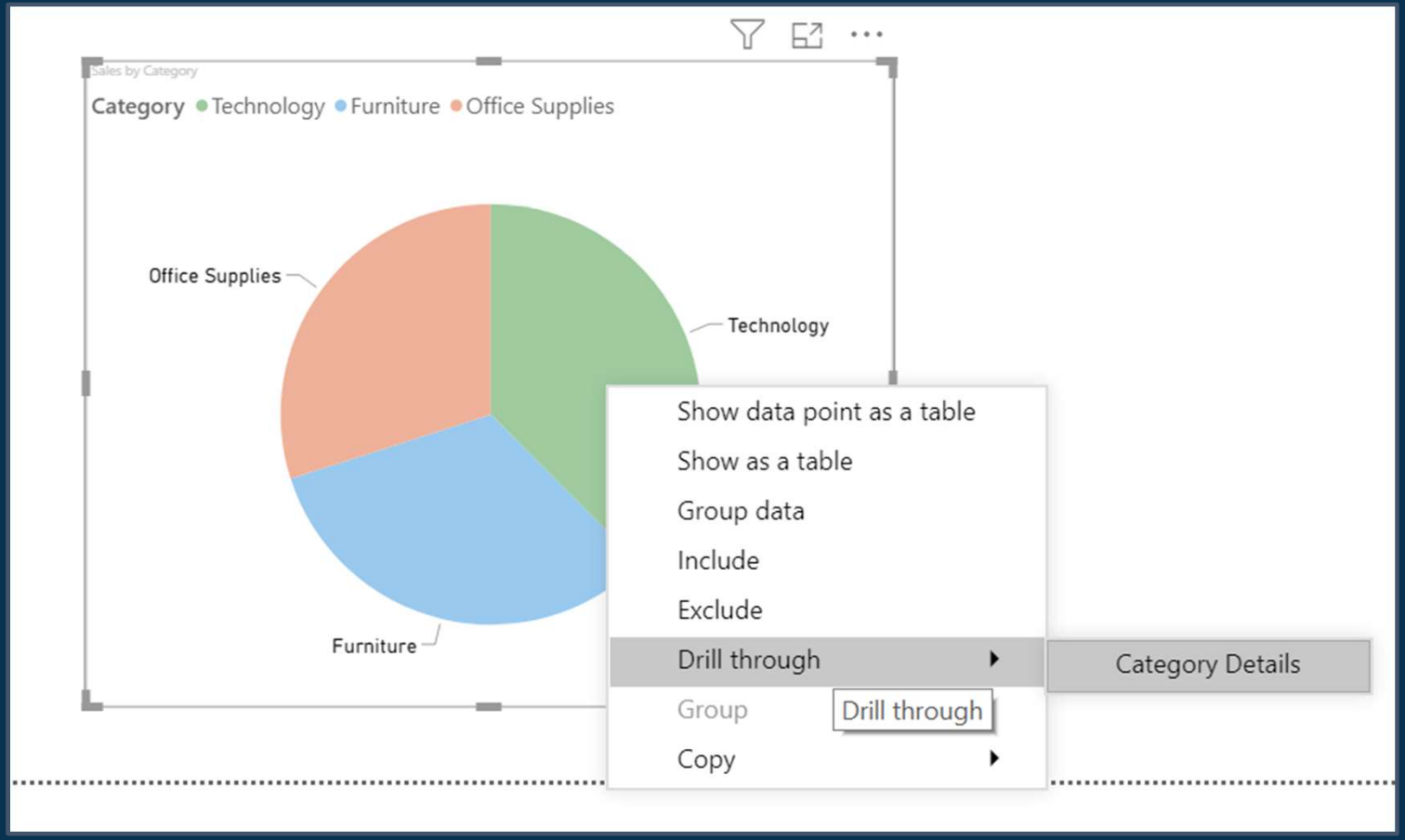
Filtering on Continuous Data

Filter on a continuous field can be applied on the basis of range of values

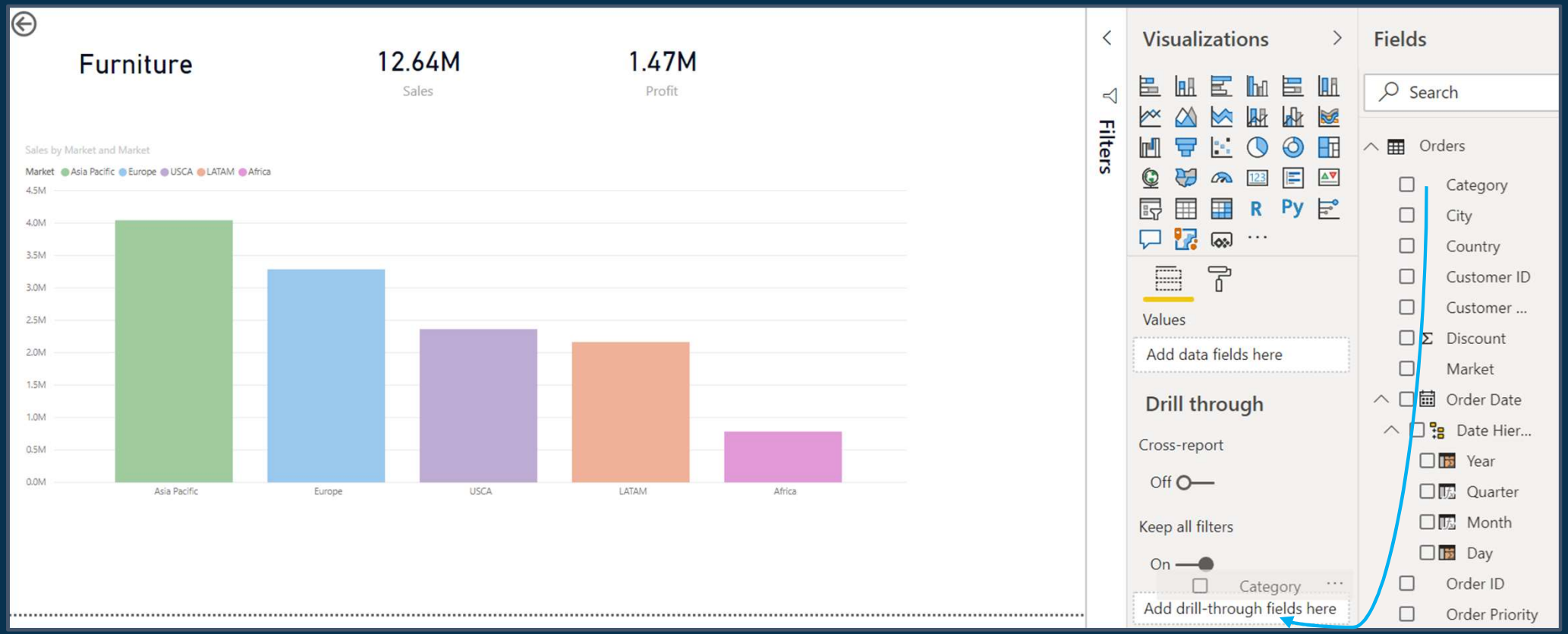
Drill-through Filter

- With drill through in Power BI reports, you can create a page in your report that focuses on a specific entity
- When your report readers use drill through, they right-click a data point in other report pages, and drill through to the focused page to get details that are filtered to that context

Set up drill-through Source Page

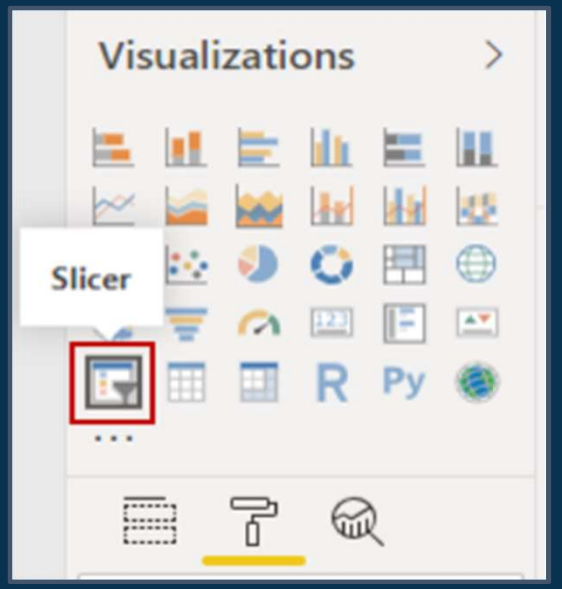


Set up drill-through Destination Page



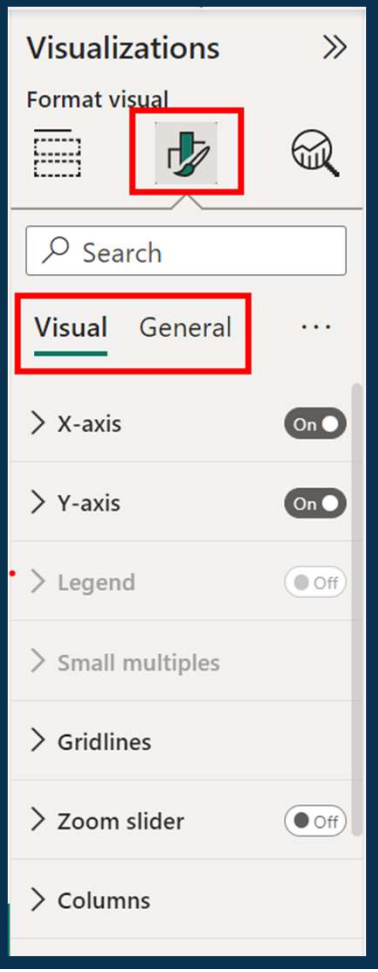
Slicers

- Slicers are one of the most powerful types of visualizations, particularly as part of a busy report.
- A slicer is an on-canvas visual filter that allows report users to segment the data by a specific value.
- To add a slicer to report, select *Slicer* from the *Visualizations* pane.



Format Pane

- Format tab in Power BI desktop can be used to change the formatting of a visual.
- It available in visualization pane second tab.
- Each visualization will have a different set of format options depending of the visual attributes.

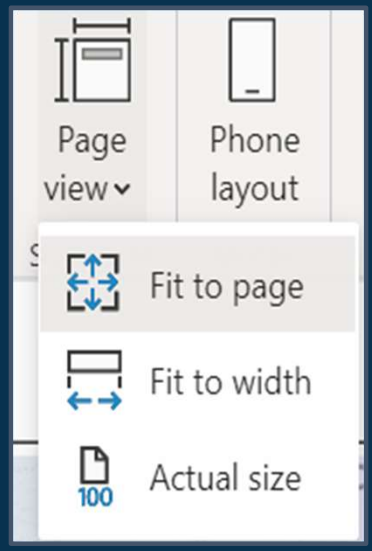


Page Layout

Page layout

- Power BI Desktop gives ability to control the layout and formatting of report pages, such as size and orientation
- View menu from the View tab to change the way that report pages scale
- The available options include Fit To Page (default),

  Fit To Width, and Actual Size

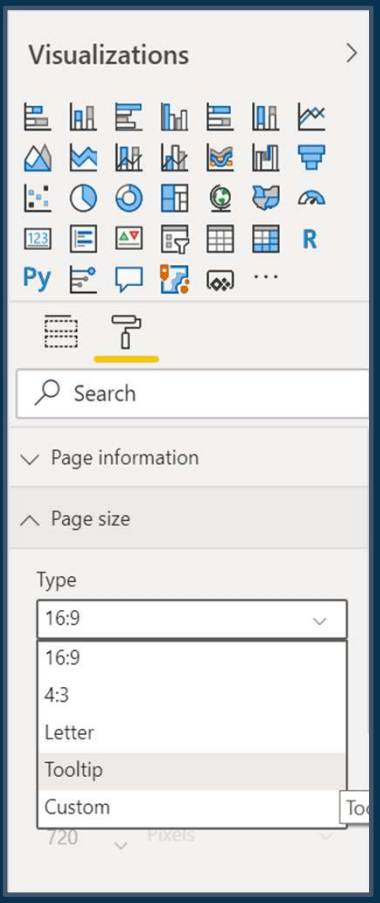


Page layout

- Page size can be modified from the format pane
- Default report page size is set to 16:9
- To change the page size, make sure that no visuals are selected,

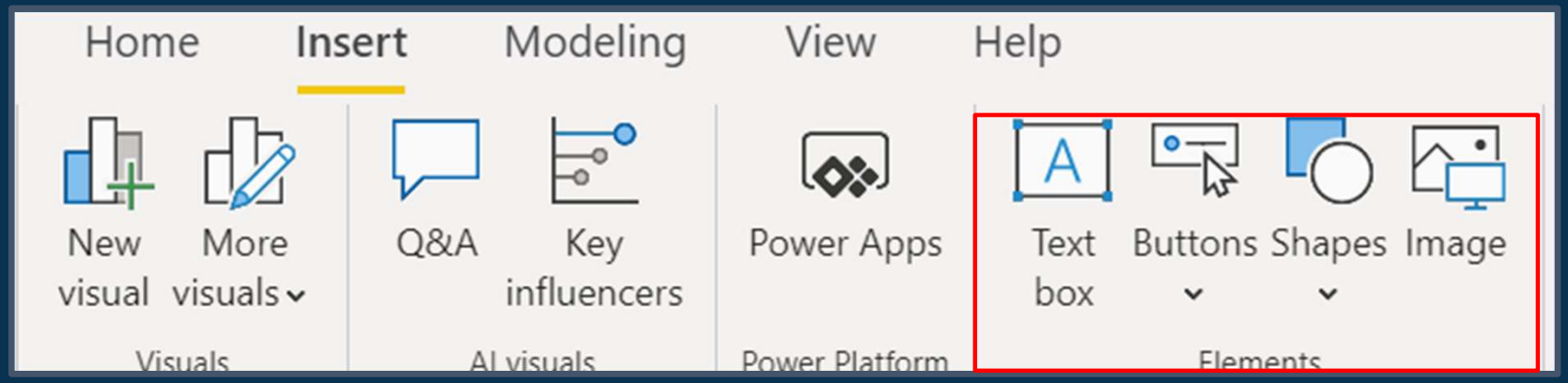
  select the paintbrush icon on the Visualizations pane,

  then select Page Size to expand that section



Add static elements

- Along with data-bound visuals, static elements such as text boxes, images, and shapes can be added to improve the visual design of reports.
- These elements can be added from *Insert tab* > *Elements* ribbon.



Analytics Pane

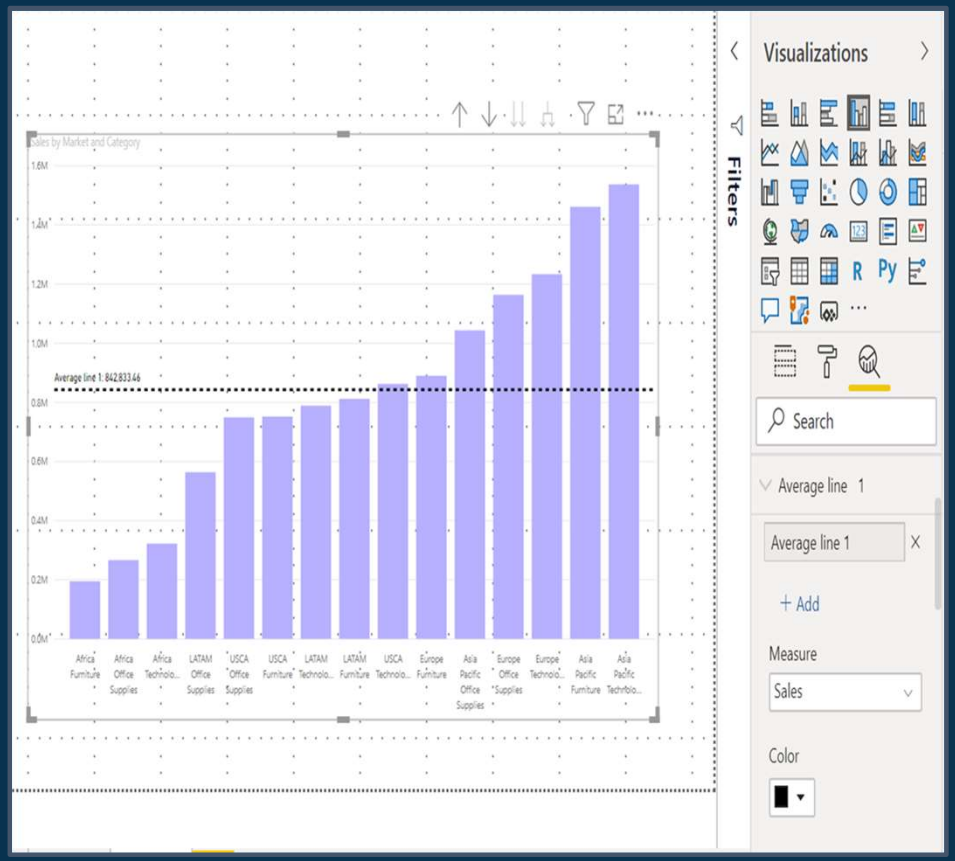
Analytics Pane

- With the Analytics pane, following types of dynamic reference lines can be added:
  - X-Axis constant line
  - Y-Axis constant line
  - Min line
  - Max line
  - Average line
  - Median line
  - Percentile line
  - Symmetry shading

Note - Not all lines are available for all visual types.

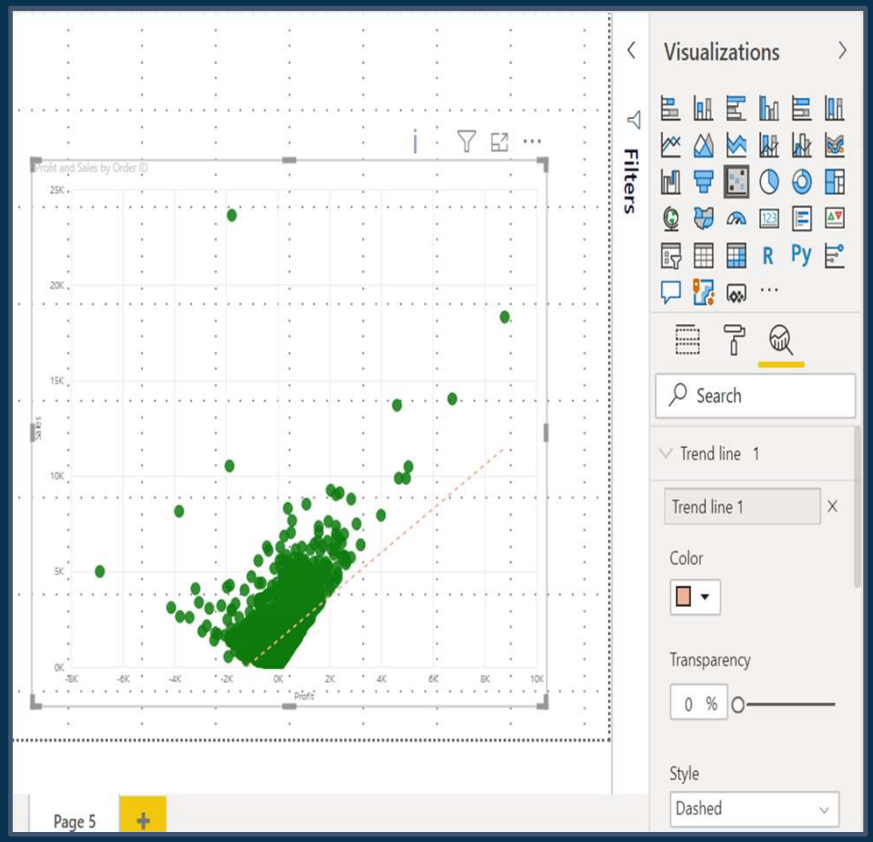
Adding a Reference Line

- Select or create a visual, then select the Analytics icon from the Visualizations section
- Select the type of line, it can be like - Average line, Min line, Max line, Constant line, Median line etc.
- To create a new line, select + Add
- To rename the line - Double-click the text box and enter name
- Its Color, Transparency percentage, Line style, and Position, Data label can be applied



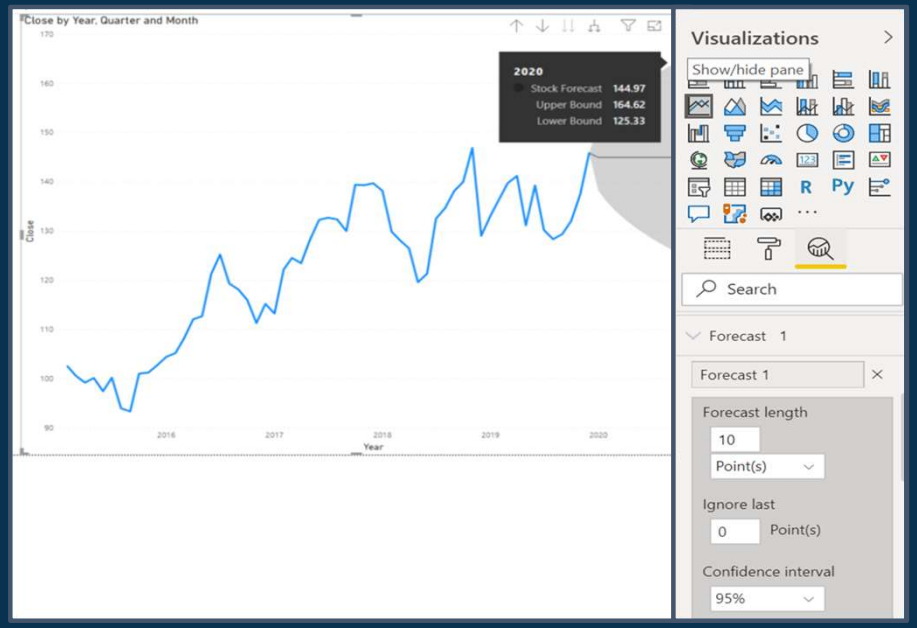
Adding a Trend Line

- Once the visual is ready, go to Analytics menu.
- Select the Trend line options and add it to the visual by clicking on the add link.
- Color and style of a trend line can be modified with the help of mentioned options in the pane.



Adding a Forecast

- Select a visual, then expand the Forecast section of the Analytics pane.
- Many inputs can be specified to modify the forecast, such as the Forecast length or the Confidence interval



Power BI Objects

- Power view supports various interactive objects such as images, shapes, text box, buttons and even navigation bar.
- These objects can be accessed using insert tab as shown in the image.