Lab Sheet 07

Data Modelling and Exploration using Power BI Desktop

In this section, we will learn about the key parts of the Power BI desktop. We will model and explore the data and build visuals.

In this lab you will learn how to:
- create a range of different charts.
- highlight and cross-filter.
- add new measures to the model to do additional analysis.

Power BI Desktop – Data Modeling and Exploration

Power BI Desktop – Data Modeling

1. Click on the **Model** icon on the left panel of Power BI Desktop. You see the tables you have imported along with Relationships. **The Power BI Desktop automatically infers relationships between the tables.** Hover the mouse on the relationships, then you can see the keys related with the relationship.
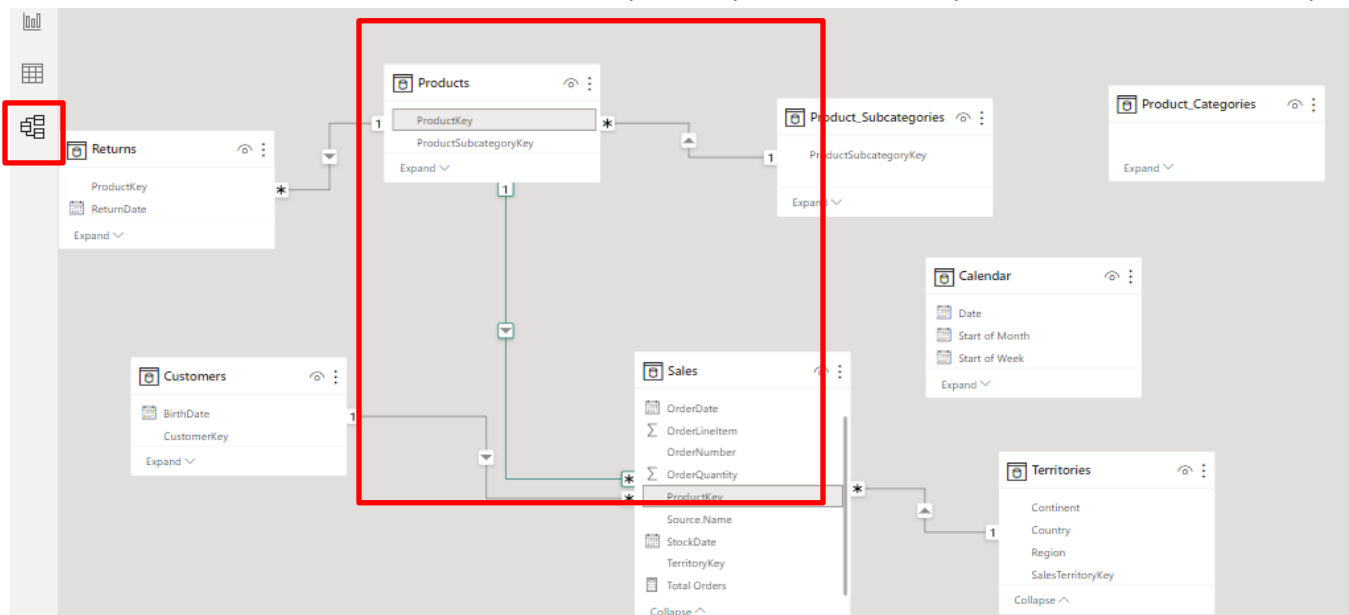
- A relationship is created between the Sales and Products tables using the **Productkey** column.

- A relationship is created between the Products and Returns tables using the **Productkey** column.

- A relationship is created between the Products and Product_Subcategories tables using the **ProductSubcategorykey** column.

- A relationship is created between the Product_Subcategories and Product_Categories tables using the **ProductCategorykey** column.

- A relationship is created between the Sales and Customers tables using the **Customerkey** column.

Power BI supports multiple types of relationships:

- 1 to many

- 1 to 1

- Many to many

In this lab, we will be using the 1 to many type of relationship, the most common type of relationship. This means one of the tables involved in the relationship should have a unique set of values.

When talking about data modeling in general, the term you will hear most often is the star schema. This is a widely adopted approach to designing data warehouses and relational databases and is the recommended approach to take in Power BI as well.

A star schema is made up of a central fact table with multiple dimension tables branching off of this fact table, much like the appearance of a star. The tables which are updated frequently called as fact table. A fact table is made up of values that can be summarized and aggregated, as well as one or more keys that link to the dimension tables. The purpose of the key is to make connections between tables.

The tables which are not updated frequently called as dimension tables. The dimension tables should contain unique values for the features they are describing, and their columns will be responsible for filtering or grouping the data in the Power BI report. Here Sales table is a fact table. All other tables are dimension tables.

Power BI Desktop – Creating a relationship manually

Power BI automatically detect all the possible relations between different sets of data if the primary key and the foreign key of the tables have the same name. Sometimes we need to create a relation between the data manually.

From Sales table, drag and drop the **TerritoryKey** column to the **SalesTerritoryKeycolumn** in the Territories table. It creates a line between Sales and Territories.

Now we need to connect Sales and Calendar tables. Sales table has two date type columns **OrderDate** and **StockDate**. Suppose you need to do the analysis using OrderDate (eg: how many orders place in the particular day, month etc.
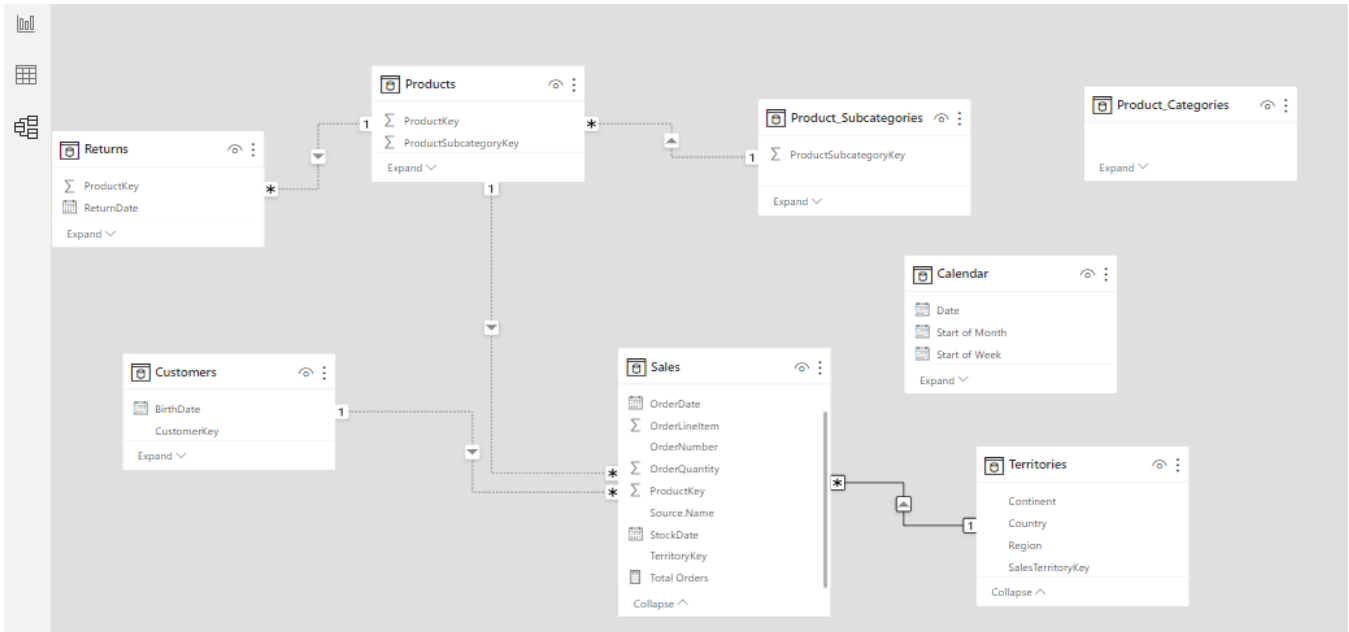
So drag and drop the **OrderDate** colum in the Sales table to the **Date** column in the Calendar table.

A line is created in between Sales and the Calendar tables. This model is called as **Star Schema**.

Note the filter directions (shown as arrows) in each relationships, these will always point from the 'one' side of the relationship (dimensions) to the 'many' side (fact) table.

Power BI Desktop – Deleting a relationship

1. Right click on the relationship that you need to delete, select **Delete** from the drop down list.

2. **Delete Relationship** dialog box opens.

3. Click **Yes.**
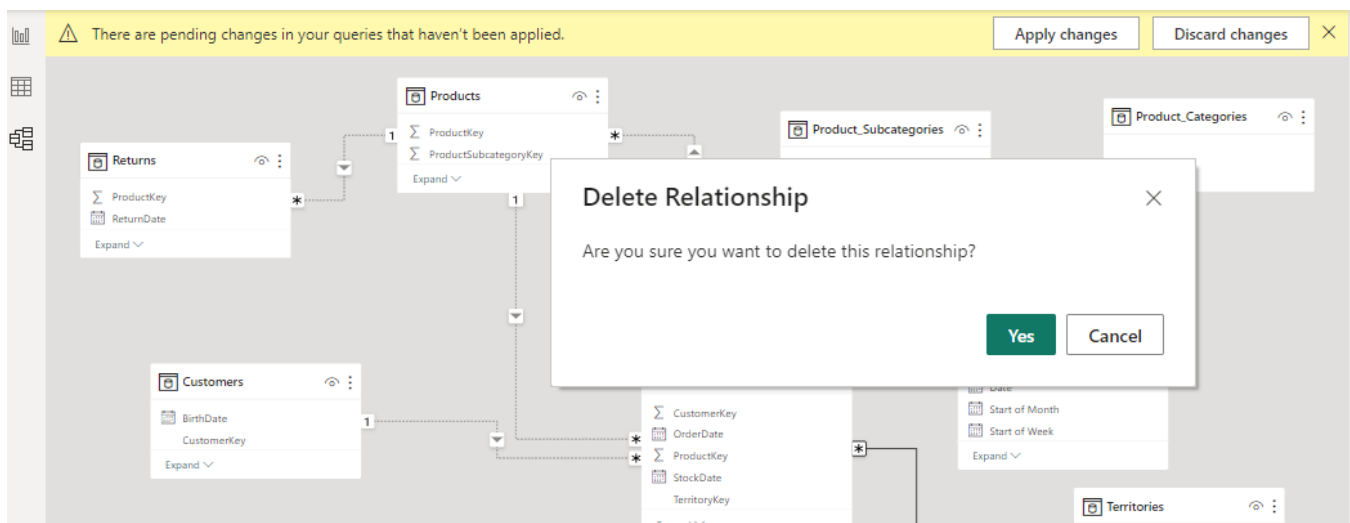
Power BI Desktop – Active/ Inactive a relationship

Double click on the relationship that need to be inactive.

1. **Edit Relationship** dialog box opens.
2. Deselect **Make this relationship active**.
3. Click **OK**.

Edit relationship

Select tables and columns that are related.

Sales

Source.Name	OrderDate	StockDate	OrderNumber	ProductKey	CustomerK
Sales_2016.csv	Sunday, January 3, 2016	Friday, September 6, 2002	SO48819	352	
Sales_2016.csv	Wednesday, January 13, 2016	Thursday, September 19, 2002	SO48901	352	
Sales_2016.csv	Thursday, January 14, 2016	Sunday, September 22, 2002	SO48909	352	

Products

ProductKey	ProductSubcategoryKey	ProductSKU	ProductName	ModelName	Produc
214	31	HL-U509-R	Sport-100 Helmet, Red	Sport-100	Universal fit, well-vent
215	31	HL-U509	Sport-100 Helmet, Black	Sport-100	Universal fit, well-vent
220	31	HL-U509-B	Sport-100 Helmet, Blue	Sport-100	Universal fit, well-vent

Cardinality

Many to one (*:1)

Cross filter direction

Single

☑ Make this relationship active

☐ Apply security filter in both directions

☐ Assume referential integrity

OK   Cancel

Data Analysis Expressions (DAX)

Data Analysis Expressions (DAX), is the formula language that drives power pivot. With DAX, you can,

Add **calculated columns** and **measures** to your model.

COMMON FUNCTION CATEGORIES

| MATH & STATS Functions	LOGICAL Functions	TEXT Functions	FILTER Functions	DATE & TIME Functions
Basic **aggregation** functions as well as "**iterators**" evaluated at the row-level	Functions for returning information about values in a given **conditional expression**	Functions to manipulate **text strings** or **control** formats for dates, times or numbers	**Lookup** functions based on related tables and **filtering** functions for dynamic calculations	
Common Examples:	**Common Examples:**	**Common Examples:**	**Common Examples:**	
• SUM	• IF	• CONCATENATE	• CALCULATE	
• AVERAGE	• IFERROR	• FORMAT	• FILTER	
• MAX/MIN	• AND	• LEFT/MID/RIGHT	• ALL	
• DIVIDE	• OR	• UPPER/LOWER	• ALLEXCEPT	
• COUNT/COUNTA	• NOT	• PROPER	• RELATED	
• COUNTROWS	• SWITCH	• LEN	• RELATEDTABLE	
• DISTINCTCOUNT	• TRUE	• SEARCH/FIND	• DISTINCT	
	• FALSE	• REPLACE	• VALUES	
Iterator Functions:		• REPT	• EARLIER/EARLIEST	
• SUMX		• SUBSTITUTE	• HASONEVALUE	
• AVERAGEX		• TRIM	• HASONEFILTER	
• MAXX/MINX		• UNICHAR	• ISFILTERED	
• RANKX			• USERELATIONSHIP	
• COUNTX				

*Copyright 2018, Excel Maven & Maven Analytics, LLC*

Adding Calculated columns

When you create a data model in Power BI Desktop, you can extend a table by creating new columns. The content of the columns is defined by a DAX expression evaluated row by row.

To add a calculated column,

1. Click on the table which needs to be added a new column. As an example click on the **Sales** table.
2. Click on the 3 dots appear in front of the **Sales** table name.
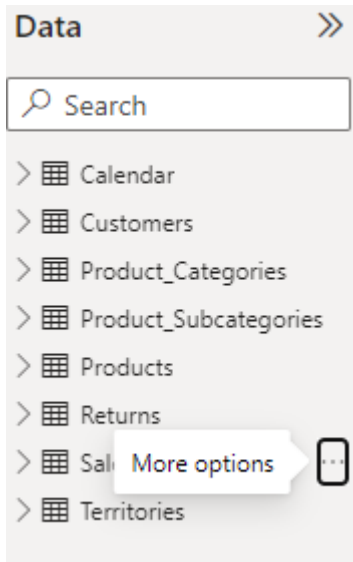
3. Then select **New Column**.
4. Type the following DAX formula.
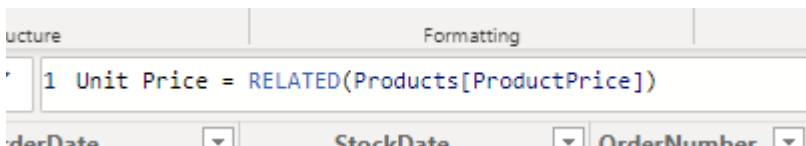
**Related()** will return the corresponding value from one table to another.

Adding Measures

There is another way of defining calculations in a DAX model, useful whenever you do not want to compute values for each row but, rather, you want to aggregate values such as sums, averages, minimum, maximum and counts from many rows in a table. These calculations are measures. First we need to calculate the **total orders**. To calculate the **total orders**,

1. Click on the **Report** icon on the left panel to navigate to the Report view.

2. Click 3 dots appear Infront of the Sales table in the right **Data** panel.
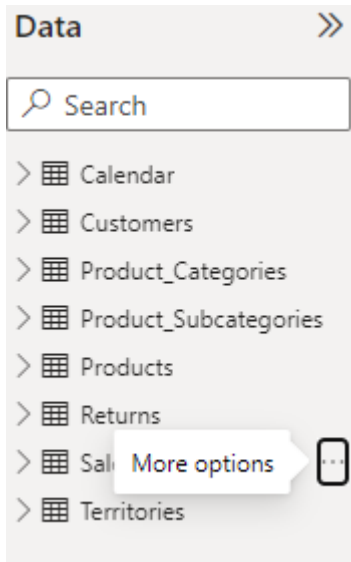
3. Select **New Measure** from the drop down menu.
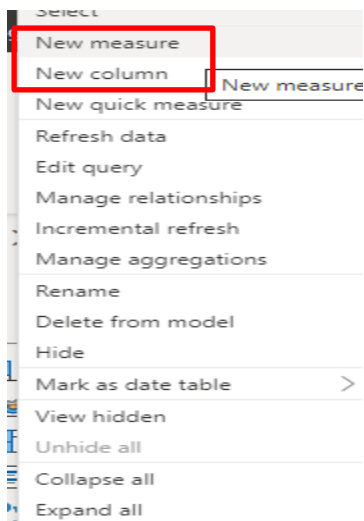
4. Type the following DAX formula.
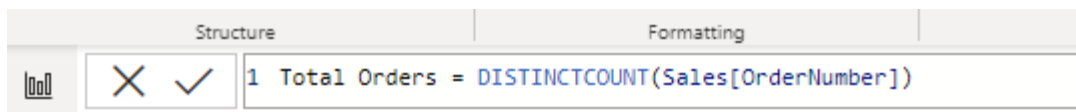
Now the Total Orders is calculated.
DISTINCTCOUNT() is used to calculate the number of distinct values in a column or a set of columns

Power BI Desktop – Data Exploration

Before move on to the data exploration in Power Bi, let's get an idea of **data types** which are used in analysis. Please refer to the **Handout**.
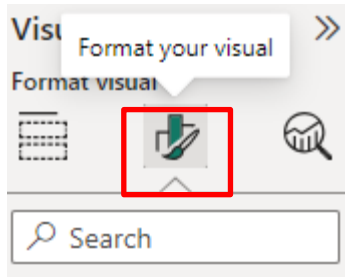
Adding donut or pie chart

Now that we have loaded data, let's start with analyzing **total orders** by **gender**. Click on **Donut** from the visualization.

5. Add **Gender** from the **Customers** table into the **legend**.
6. Add **Total Orders** from the **Sales** table into the **values**.
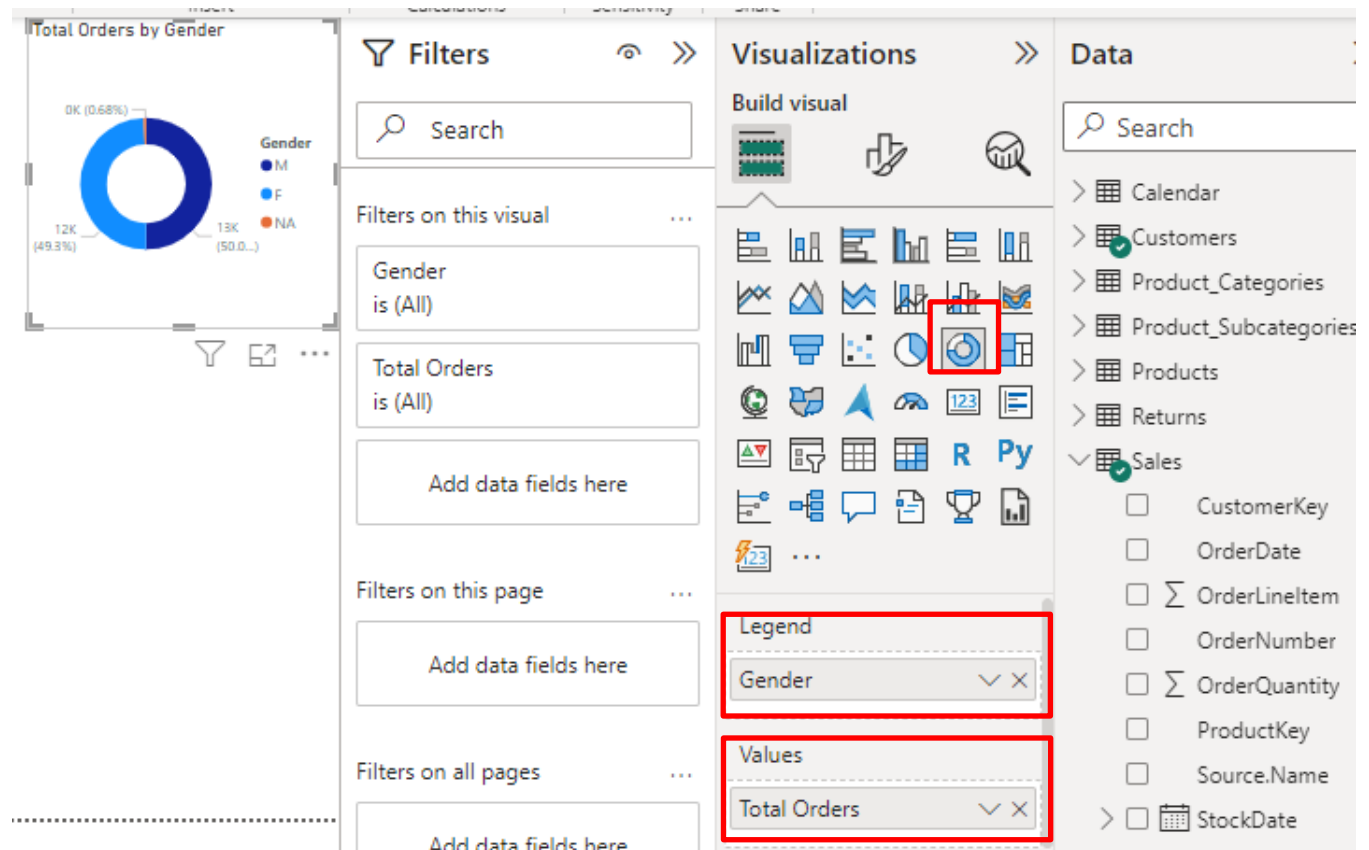
Do the formatting if necessary using the **Format your visual** icon in the **Visualization** pane.

Adding a Matrix

The matrix visual is a type of table visual (see Tables in this article) that supports a stepped layout. A table supports two dimensions, but a matrix makes it easier to display data meaningfully across multiple dimensions.

1. Calculate the **Total Revenue** by following formula.

```
1  Total Revenue = SUMX(sales,Sales[OrderQuantity]*RELATED(Products[ProductPrice]))
```

SUMX() is used to iterate over a table and calculate the sum of a specified expression or measure.

2. Click on the **Matrix** icon on the **Visualizations** pane**.**

3. As the rows add **ProductName**.

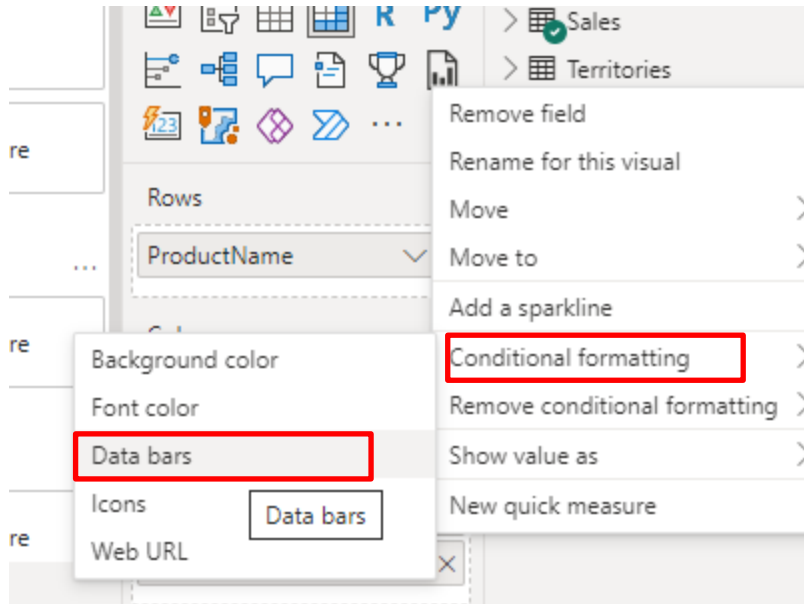4. **Add Total Orders** and **Total Revenue** as the values.
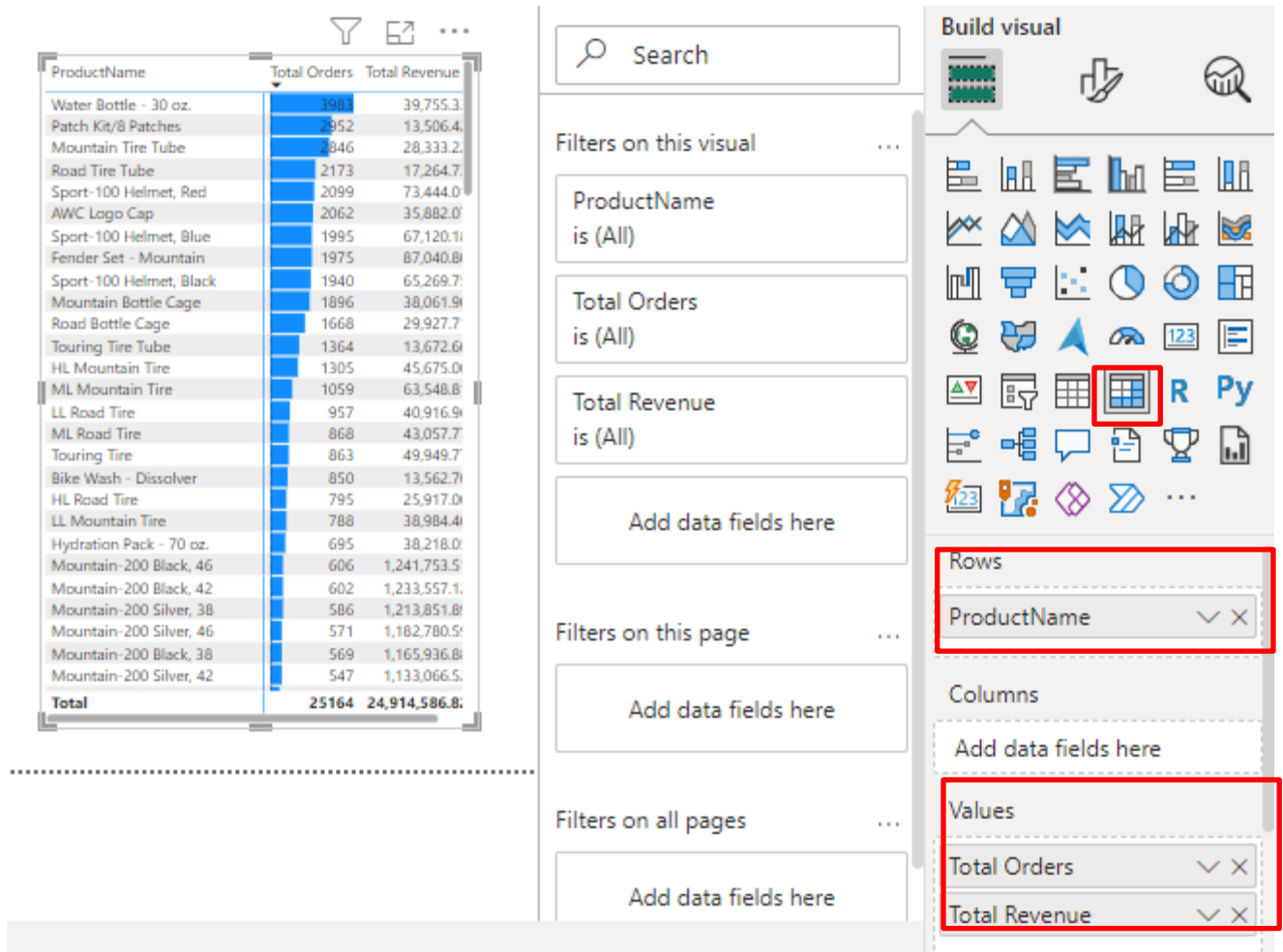
Adding bar chart to the columns, will enhance the visualization.

To add a **bar chart** to the **Total Orders** column,

1. Click on the drop down arrow in the **Total Orders** in Values field.

2. Then select **Conditional Formatting**.

3. Then **Data bars**.

4. **Data bars – Total Orders** dialog box opens
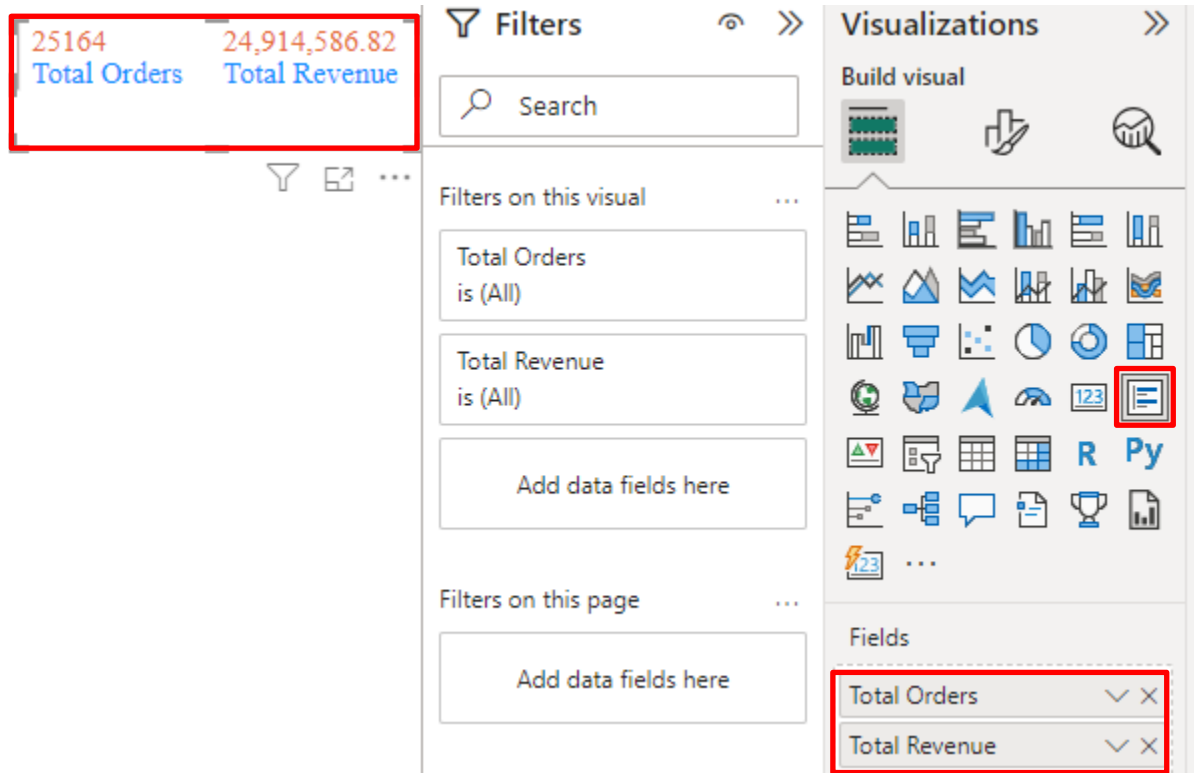
5. Click OK.
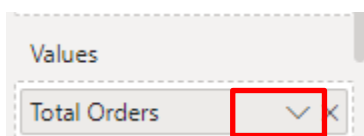
Adding a card/ multi-row card

A single number, such as total sales, market share year over year, or total opportunities, is sometimes the most important thing you want to track. A type of visualization in Power BI called a *card/ multi-row card* might be the best way to view that number.

1. Click on **multi-row card** icon.
2. Add **Total orders** and **Total Revenue** into the Fields.
3. Do the formatting according to your desire (Change font color, size, style etc.).

Adding a line chart

Line charts emphasize the overall shape of an entire series of values, usually over time.

To represent the data varies over the time period Line charts can be used in the Visualizations pane.

1. Click on the **Line and Clustered Column** Chart.
2. Add **Start of Month** into the **X-axis** from the **Calendar** table.
3. Add **Total Orders** into the **Column y-axis** from the **Sales** table.
4. Add **Total Revenue** into the **Line y-axis** from the **Sales** table.
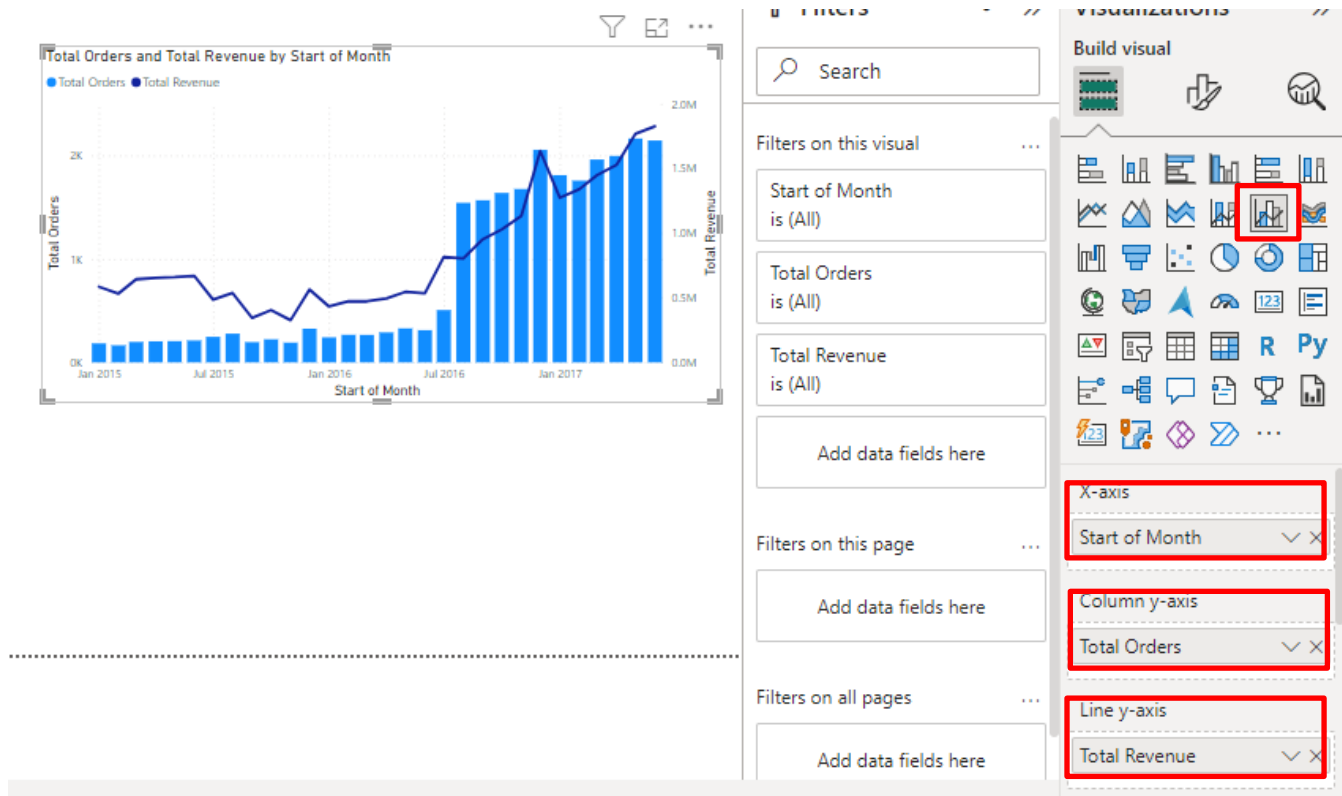
This shows the total order value and the total revenue value in each start of the month.

Adding a tree map

Treemaps are charts of colored rectangles, with size representing value. They can be hierarchical, with rectangles nested within the main rectangles. The space inside each rectangle is allocated based on the value being measured. And the rectangles are arranged in size from top left (largest) to bottom right (smallest).

We are going to look at how the total orders are varied based on the customer age. Add a calculated column and calculate the customer age using following formula.
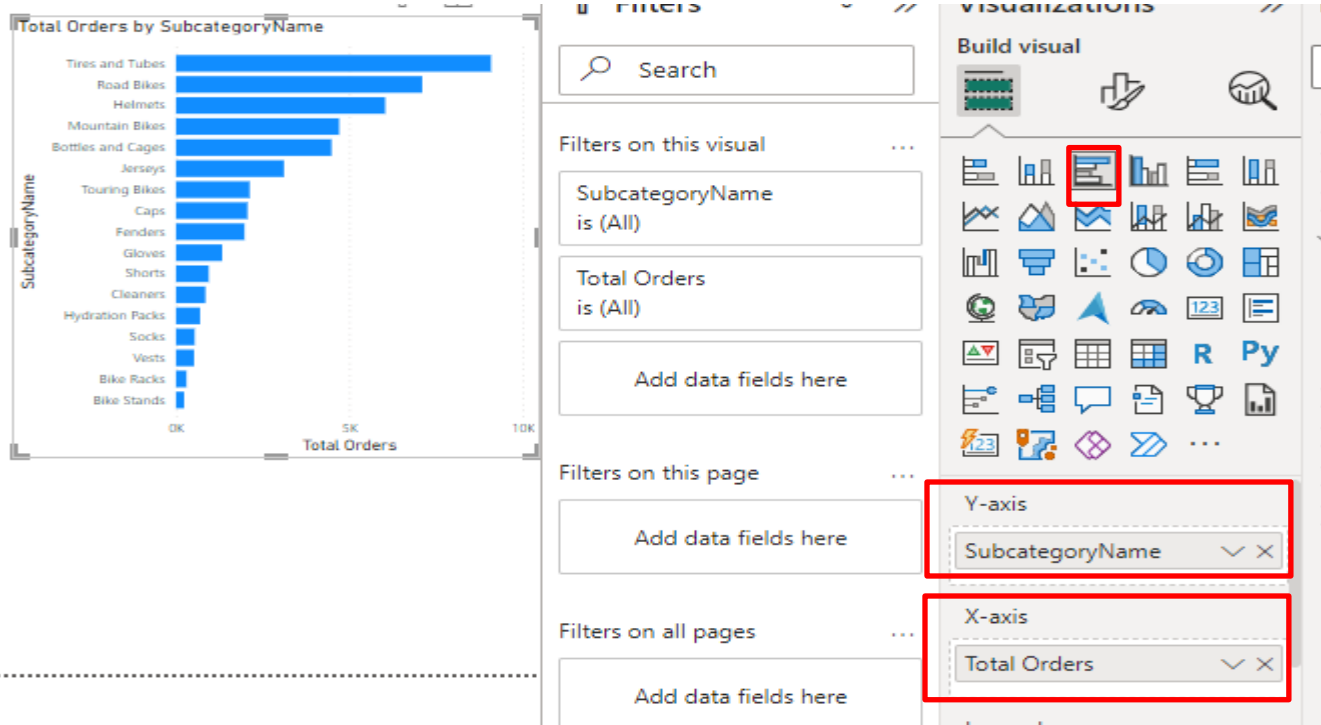
```
1 age = DATEDIFF(Customers[BirthDate],TODAY(),YEAR)
```

1. Click on the **Treemap** icon on the **Visualization** pane.
2. Add the **age** in to the **Category**.
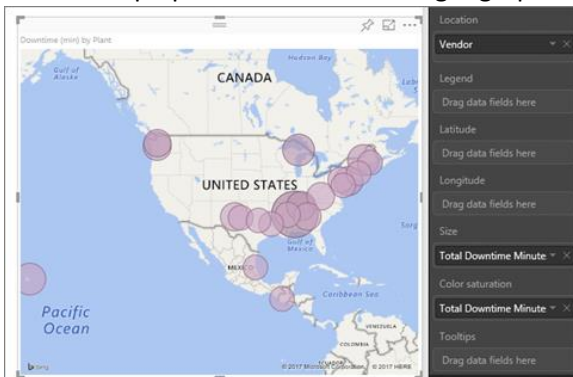3. Add the **Total Orders** in to the **Values** field.
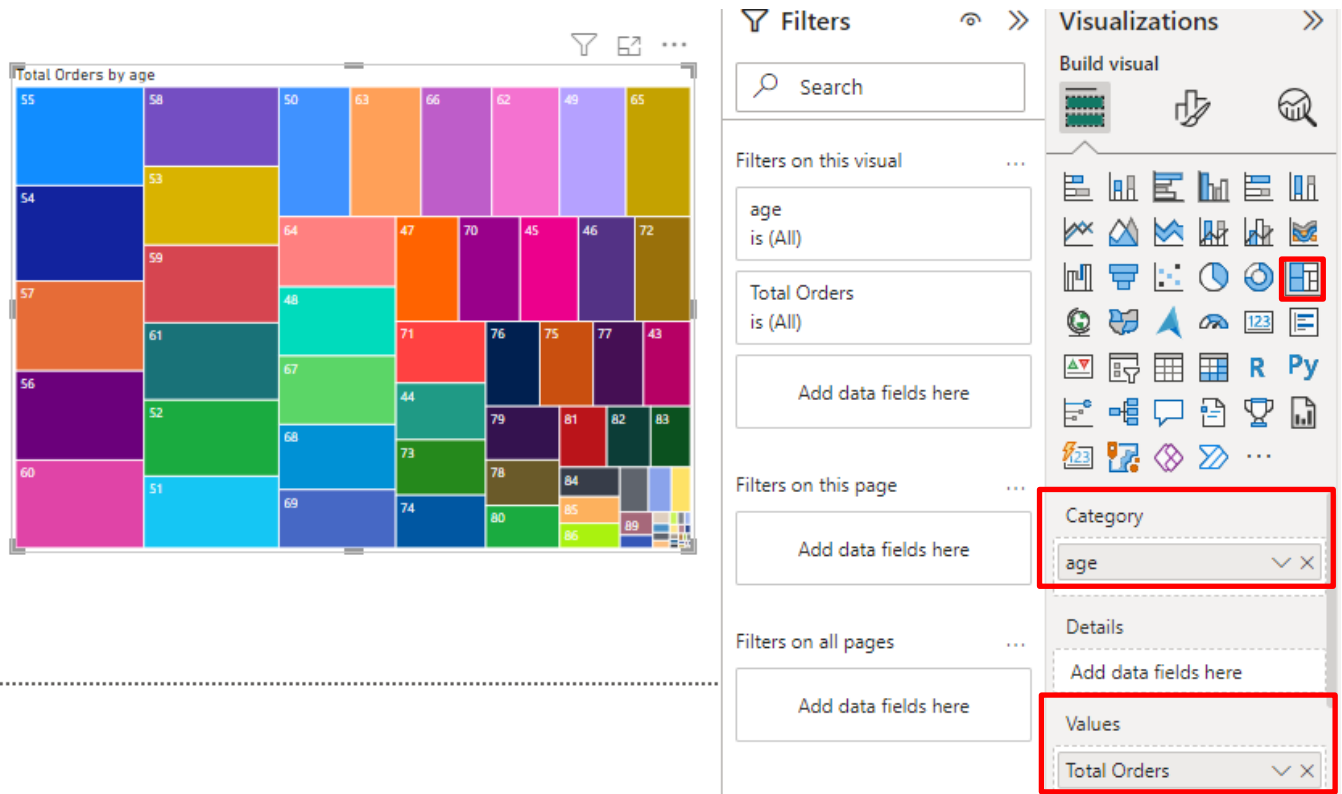
Adding a bar chart

1. Click on the **Clustered bar chart** on the **Visualization** pane.

2. Add **SubcategoryName** as the **Y-axis**.

3. Add **Total Orders** as the **X-axis**.

Adding a bubble/ filled map

Bubble Map - places bubbles over a geographic point

Filled Map – Shows the outline of area that you want to visualize

Power BI integrates with Bing Maps to provide default map coordinates (a process called geo-coding) so you can create maps. Together they use algorithms to identify the correct location, but sometimes it's a best guess. If Power BI tries, but can't create the map visualization on its own, it enlists the help of Bing Maps.

1. Click on the **map chart** on the **Visualization** pane.
2. Add **Country** to the **Location**.
3. Add **Total Orders** to the **Bubble size**.

   Bubble size is varied based on the total orders in each location.
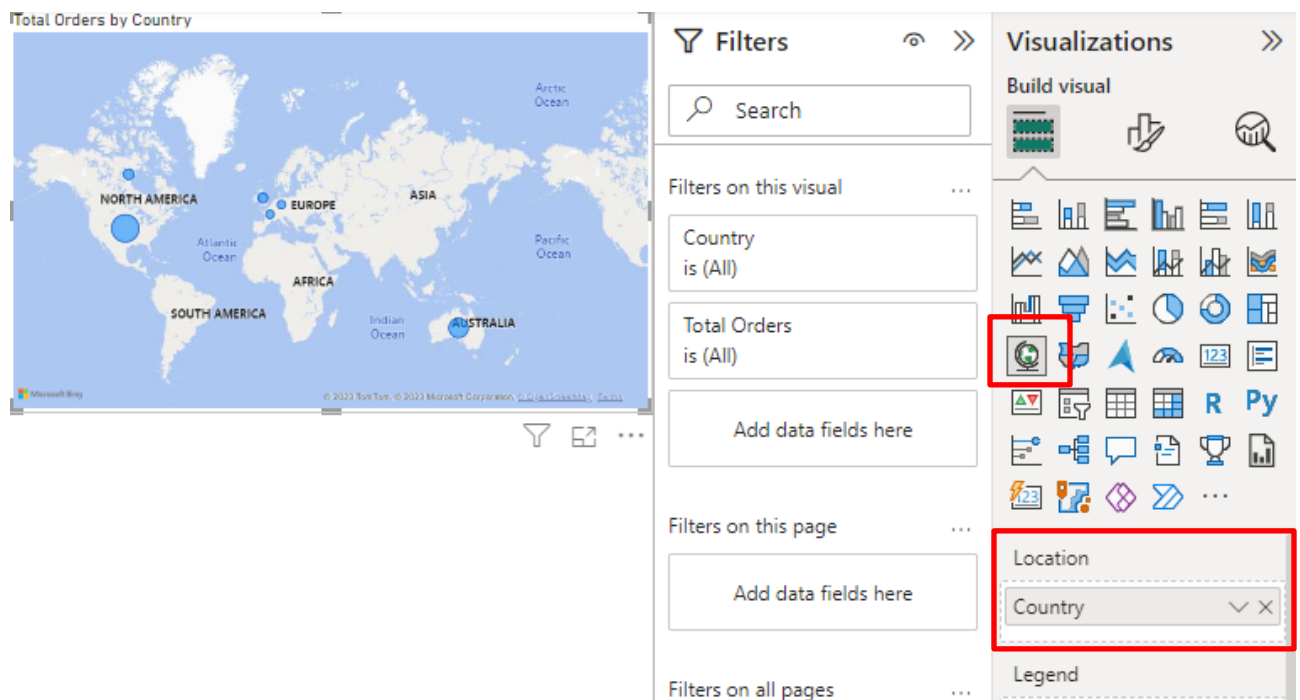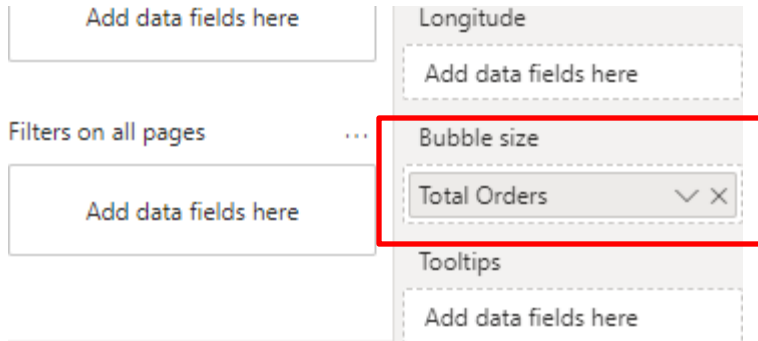
Adding a slicer

Slicers are a great choice when you want to:

- Display commonly used or important filters on the report canvas for easier access.
- Make it easier to see the current filtered state without having to open a drop-down list.
- Filter by columns that are unneeded and hidden in the data tables.
- Create more focused reports by putting slicers next to important visuals.

Here we are going to filter the **Total Orders** in each **Country** by the **Continent**. To add a slicer into the report,

1. Click on the **Slicer** icon on the **Visualization** pane.

2. Add the **Continent** in to the **Field**.
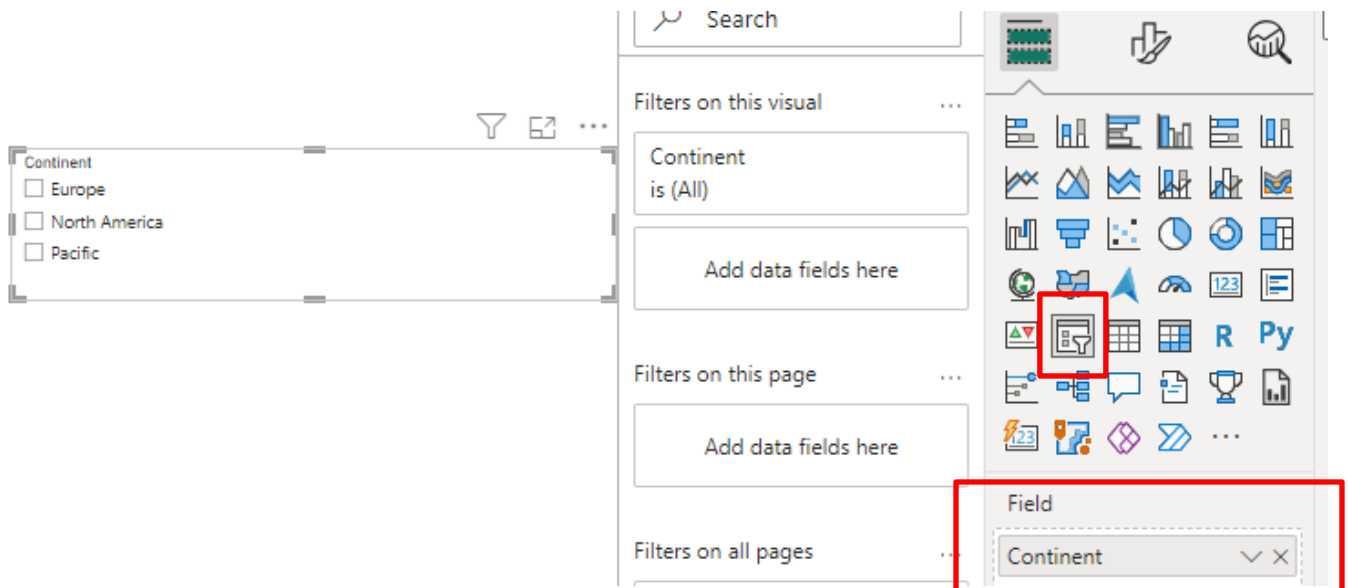
3. Continent names are displayed in the vertical order.

4. To change the appearance, Click on the **Format your visual**.

5. Then click on the Slicer settings.

6. Under the **Options**, **Style** field select **Tile**. The continent names are displayed in the Tiles as follows.
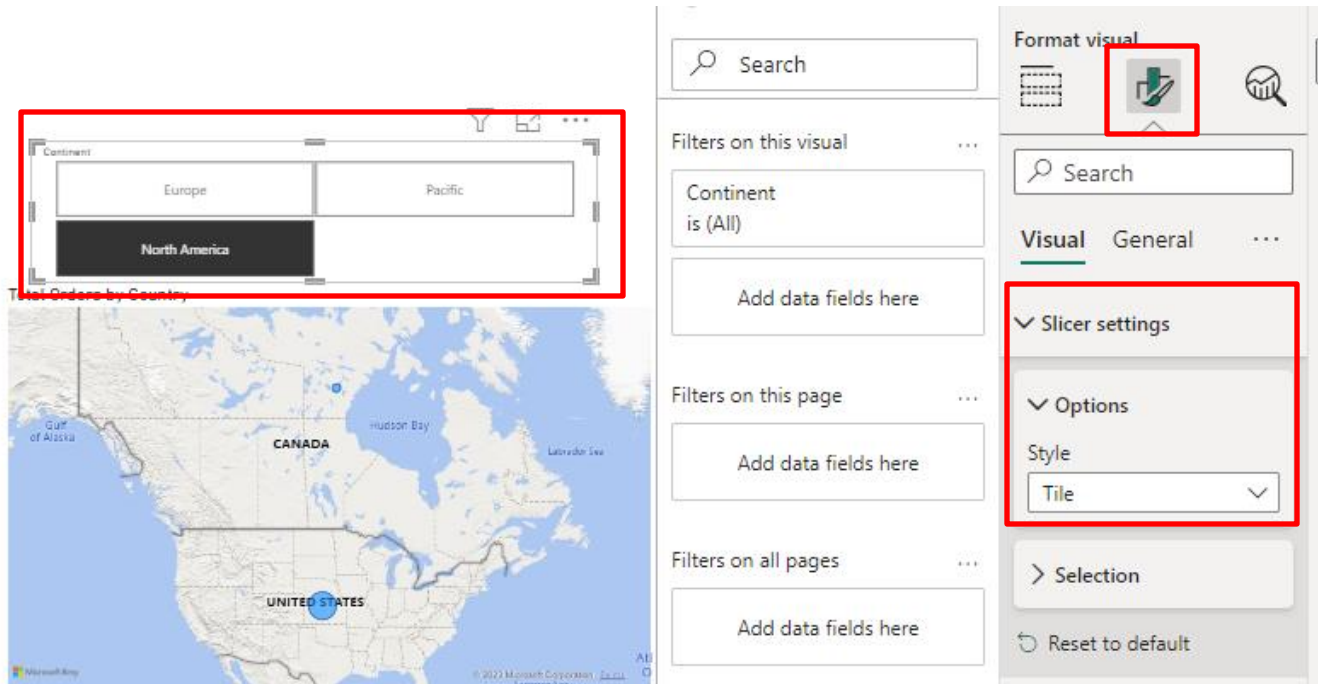
The Total orders in each Country will be filtered based on the Continent as above.

Adding a gauge chart

Gauge charts are used to show progress towards a particular goal. It helps understand to what extent a goal has been completed. The gauge line (or needle) represents the goal or target value. The shading represents the progress toward the goal. The value inside the arc represents the progress value. Power BI spreads all possible values evenly along the arc, from the minimum (left-most value) to the maximum (right-most value).

When to use a radial gauge

1. Radial gauges are a great choice in many scenarios:
2. Show progress toward a goal.
3. Represent a percentile measure, like a KPI.
4. Show the health of a single measure.
5. Display information that's easy to scan and understand.

In order to check the total orders in the current month reaches the order target or not, we can use the gauge charts in power bi as follows. Suppose the order target should increase 10% than the previous month orders. First we need to find the previous month total orders. Add a calculated column and calculate the previous month total orders using following formula.

```
1  previous month orders = CALCULATE([Total Orders],DATEADD('Calendar'[Date],-1,MONTH))
```

Add a calculated column and calculate the order target using following formula.


```
X  ✓  1 Order Target = [previous month orders]*1.1
```
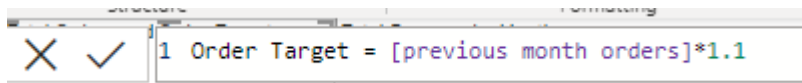
1. Click on the **gauge chart** on the **Visualization** pane.

2. Add **Total Orders** to the **Value** field.

3. We need to set the target for the current month. To do that, drag and drop the **Start of Month** to the **Filters on this visual**. From the **Filter type** select **Top N**. Under the **Show item** select **Top** and type the value as 1 in the text field.

4. Again drag and drop the **Start of Month** to the **By value** field. Select Latest from the drop down List.
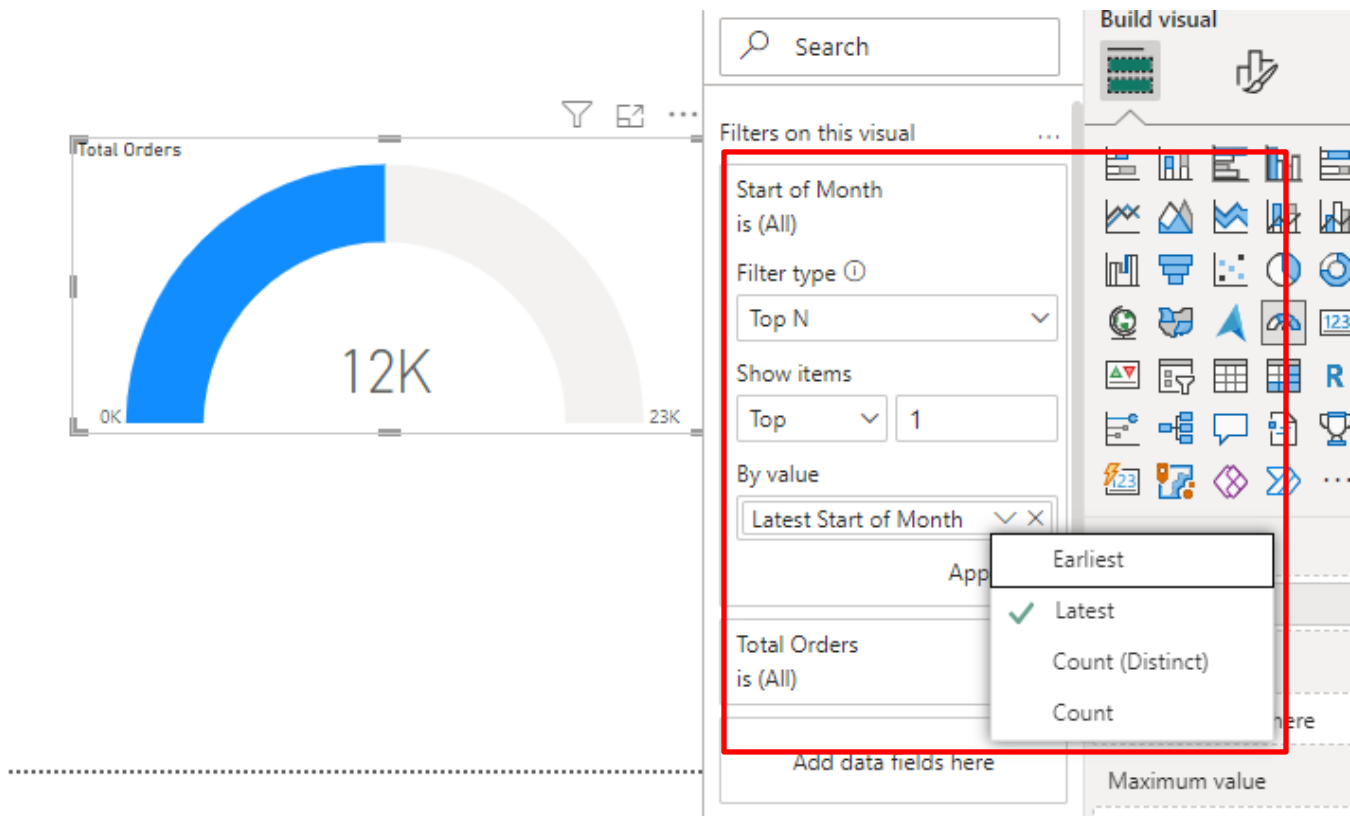
5. Click on the **Apply filter**.
6. To set the target, drag and drop the **Order Target** to the **Target value** field.
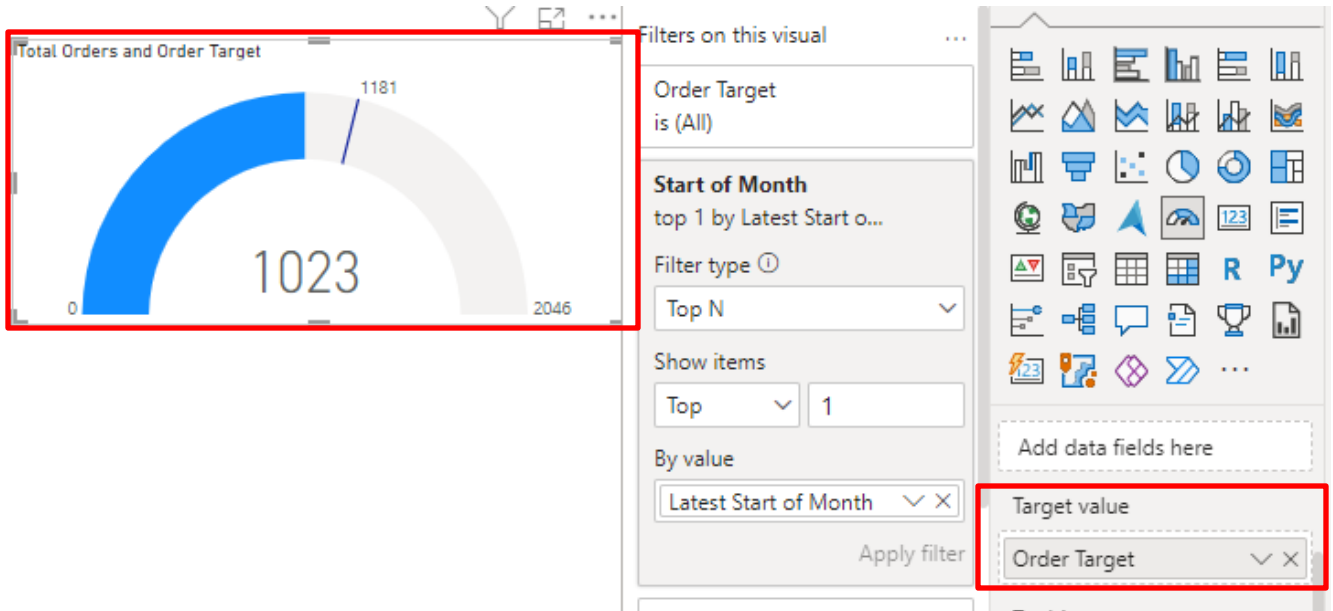
Adding Q&A

The Q&A visual in Power BI allows you to create visuals using Power BI's natural language processing engine to interpret your question and get an answer from your data. This new feature (introduced in 2019) means you can ask a question about your data in plain English and have Power BI build the visual to answer your question for you. It is similar to the smart narrative visual, another recent addition to Power BI that will automatically summaries your reports for you. Once you get used to it, this can be a really fast way to try out new reports and ways to look at your data.

State whether the given mathematical computations are possible to perform using the store sales data, and if possible, find the answers using Power Bi. (Use the Q&A in Power Bi to check the possibility)

1. Find the frequency of orders placed by each customer segment.

2. Create a frequency distribution for the order quantity.

3. Find the median of product category.

4. Find the median of shipping cost.

5. Find the total sales.

6. Create a frequency distribution for order priority.

7. Find the mean of order quantity.

8. Find the mean of the order priority.

9. Find the median of the order priority.

10. Find the standard deviation of order quantity.

11. Find the total number of orders placed each year.

12. Find the number of orders shipped for each province.

13. Find the total number of orders placed under each product category.

Activity 01

1. Load the given data set into the power BI.
2. Create a measure to calculate total number of transactions.
3. Create a measure to calculate the total manufacturing cost.
4. Create a measure to calculate average profit.
5. Create a measure to calculate the total unit price.
6. Create a measure to calculate average manufacturing price of units.

Activity 02

1. Download the sample sales data.
2. Explore the data available in the excel sheet.
3. Compare the customers with ship mode.
4. Compare customer segments and ship modes with related to customers.
5. Interpret and visualize data related to city, 'Los Angeles'.
6. Visualize the data related to product categories and sub categories.
7. What can you say about future sales on each category?

Activity 03

1. Create a new table with the data given below
2. Load them in to the power BI
3. Change the data type of the District to the proper format
4. Create a Bubble map to visually identify the number of patients in each district
5. Create a filled map to visually identify the number of patients in each district
6. Add a slicer in order to filter the map by district

DISTRICT	Count
COLOMBO	44
PUTTALAM	34
KALUTARA	26
GAMPAHA	16
KANDY	7

JAFFNA	7
RATNAPURA	3
KURUNEGALA	3
MATARA	2
GALLE	1
KEGALLE	1
BATTICOLOA	1
BADULLA	1