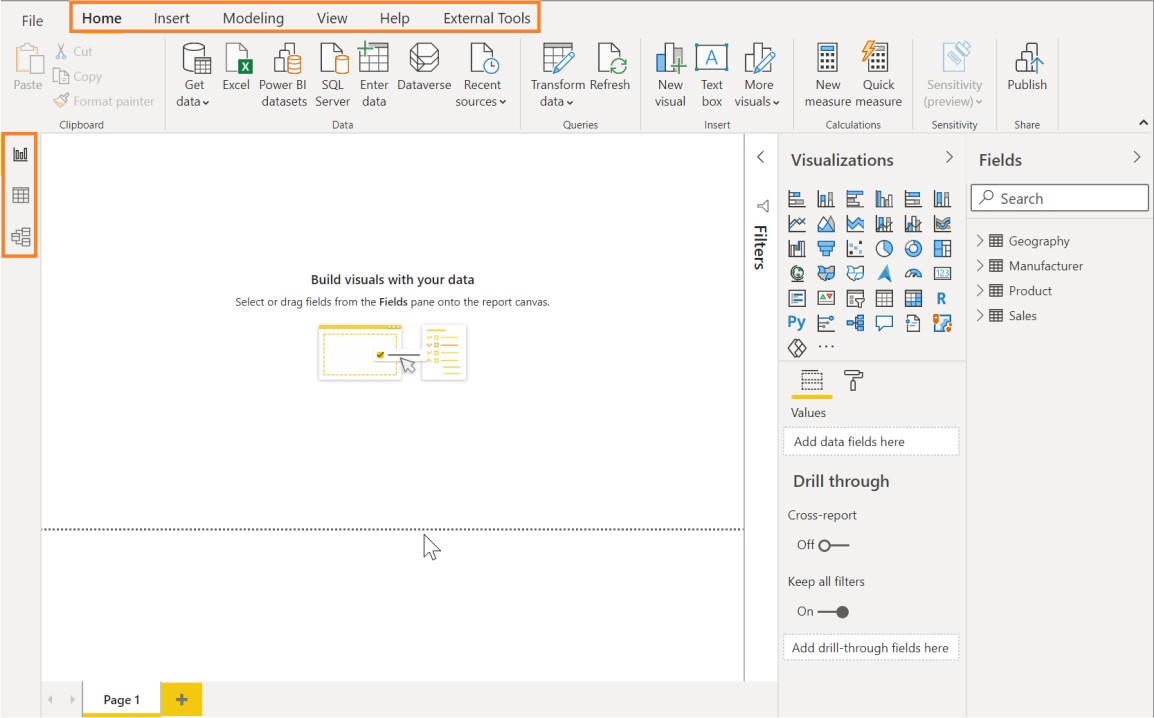
# Lab 3 - Data Modeling

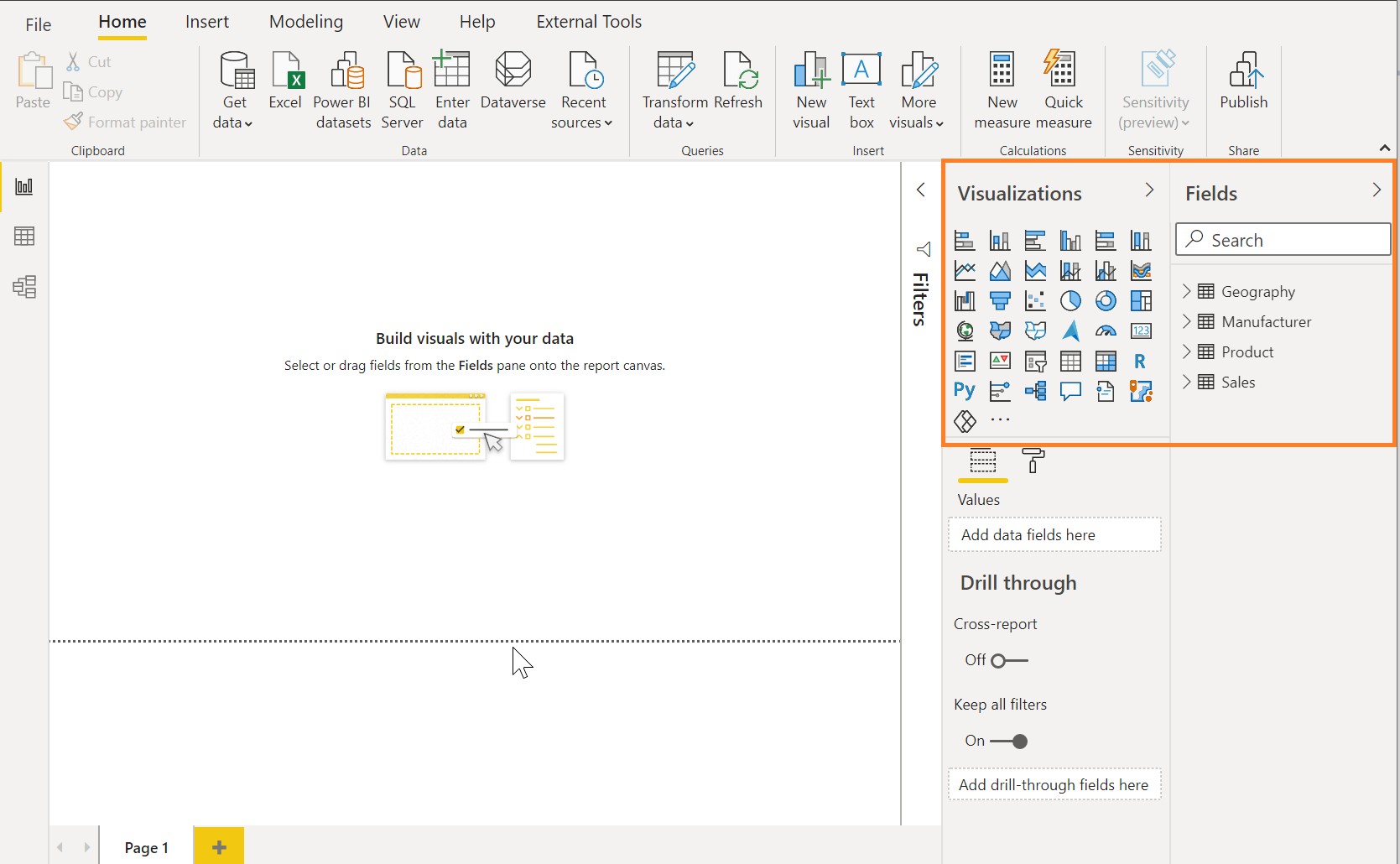
In this section, we will learn about the [key parts of the Power BI desktop](https://powerbi.microsoft.com/en-us/documentation/powerbi-desktop-query-overview/). We will model and explore the data and build visuals.

## Power BI Desktop - Layout

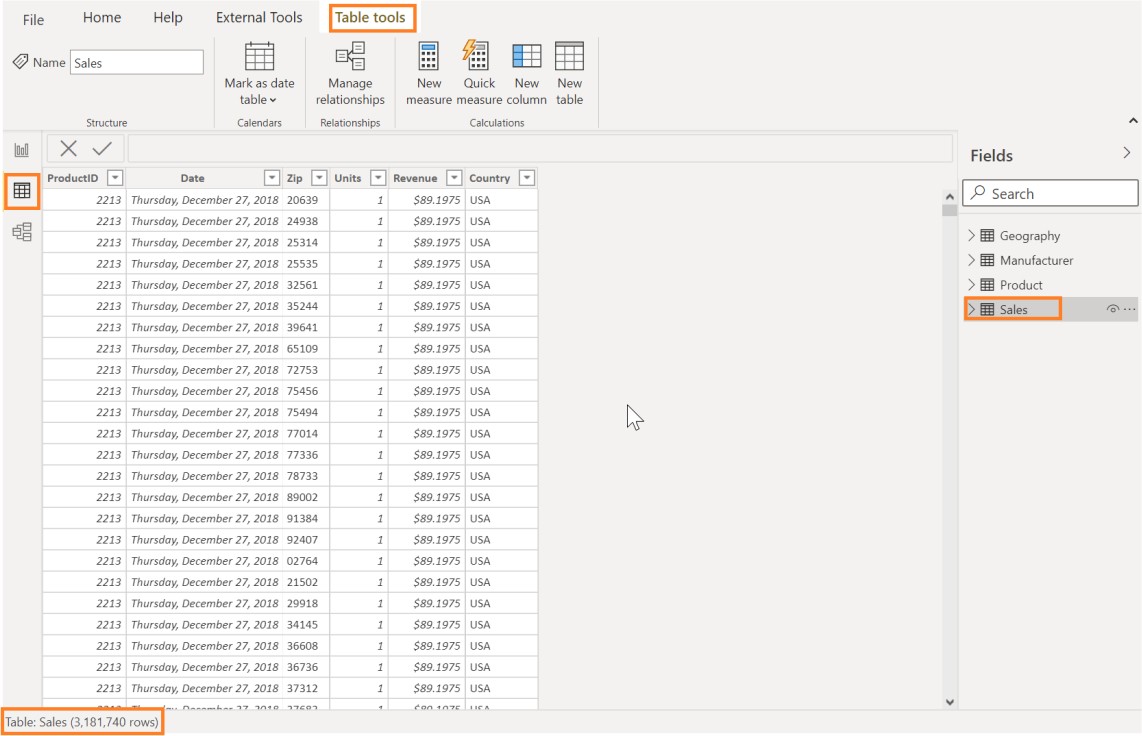
Let’s start with the main **Power BI Desktop** window and become familiar with the distinct sections available.



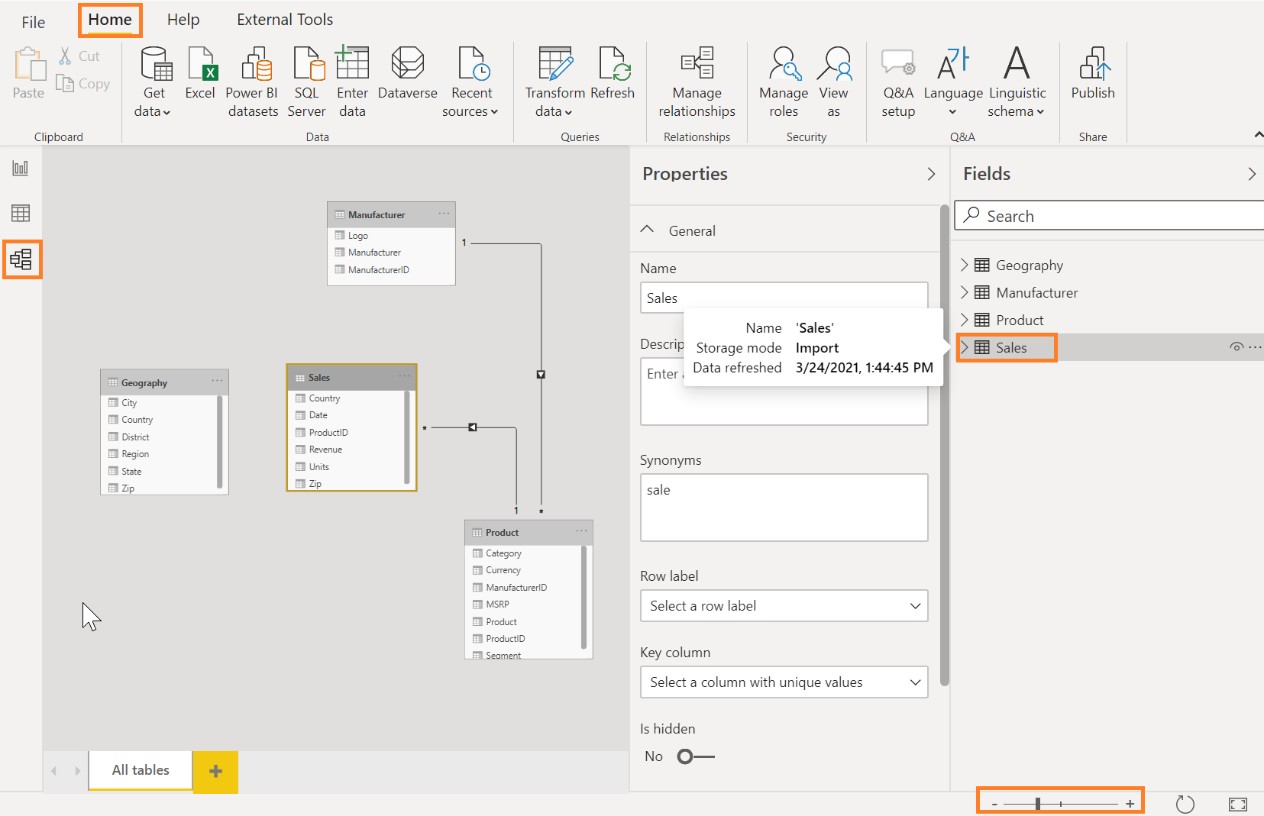
1. On the top of the window, you see the **Home** tab where the most common operations you perform are available.
2. The **Insert** tab in the ribbon allows you to insert shapes, a text box or new visuals
3. The **Modeling** tab in the ribbon enables additional data modelling capabilities like adding custom columns and calculating measures.
4. The **View** tab has options to format the page layout.
5. The **Help** tab provides self-help options like guided learning, training videos and links to online communities, partner showcase and consulting services.
6. On the left side of the window, you have three icons, **Report, Data and Model**. If you hover over the icons, you can see the tooltips. Switching between these allows you to see the data and the relationships between the tables.
7. The center **white space** is the canvas where you will be creating visuals.
8. The **Visualizations** panel on the right allows you to select visualizations, add values to the visuals, and add columns to the axis or filters.



1. The **Fields** window on the right panel is where you see the list of tables which were generated from the queries. Click the icon (downward facing triangle) next to a table name to expand the field list for that table.



1. Click on the **Data** icon on the left side. Expand the **Sales** table in the **Fields** pane as shown in the figure above. Scroll up and down to notice how fast you can navigate through over three million rows.



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.
   * A relationship is created between the Sales and Product tables using the **ProductID** column.
   * A relationship is created between the Product and Manufacturer tables using the

**ManufacturerID** column.

Power BI supports multiple types of relationships:

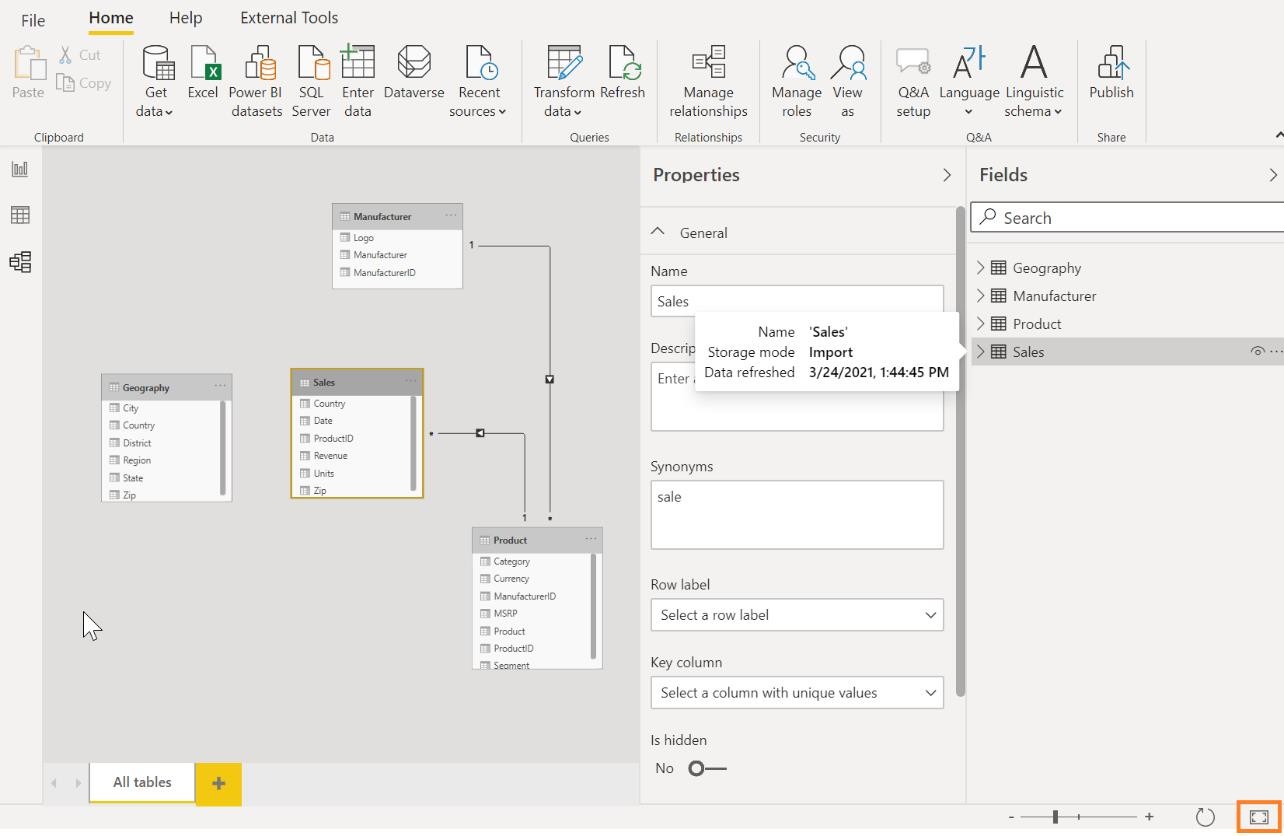
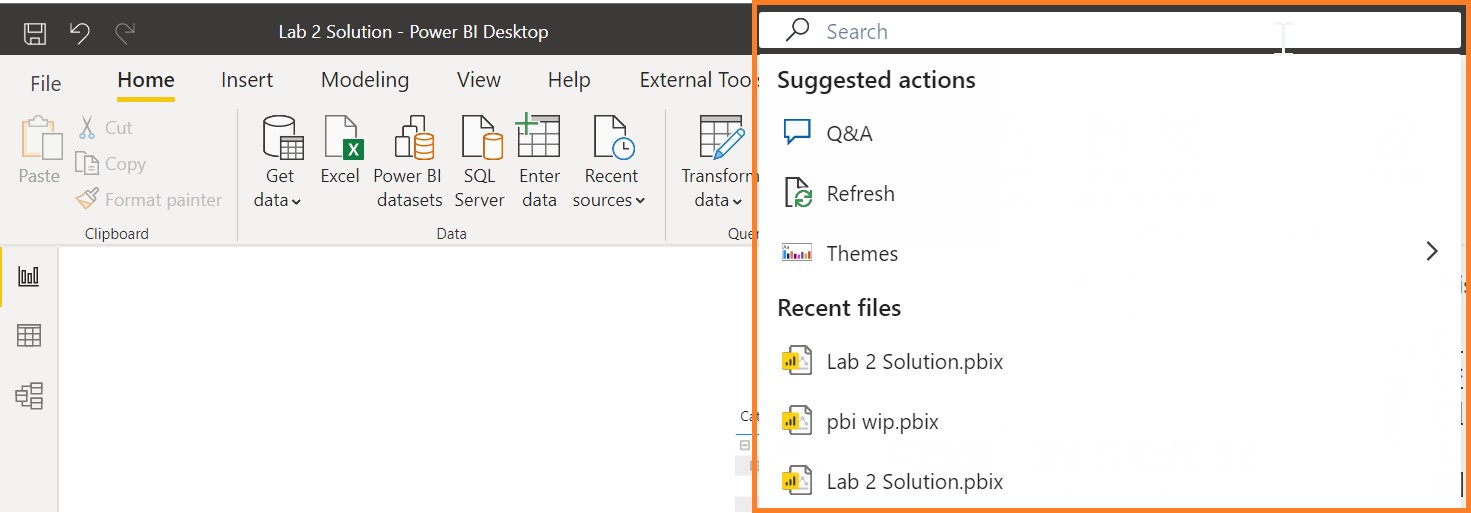
* + 1 to many
  + 1 to 1
  + Many to many

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. We will create additional relationships later in this lab.

**Note**: Tables may not appear as shown in the figure. You can zoom in and out of the **Relationships**

Fit to screen button in relationships tabpage by dragging the zoom slider in the bottom right corner of the window. Also, if you want to ensure

you are seeing all the tables, use the fit to screen icon: . Drag and move the tables to appear as shown in the figure:

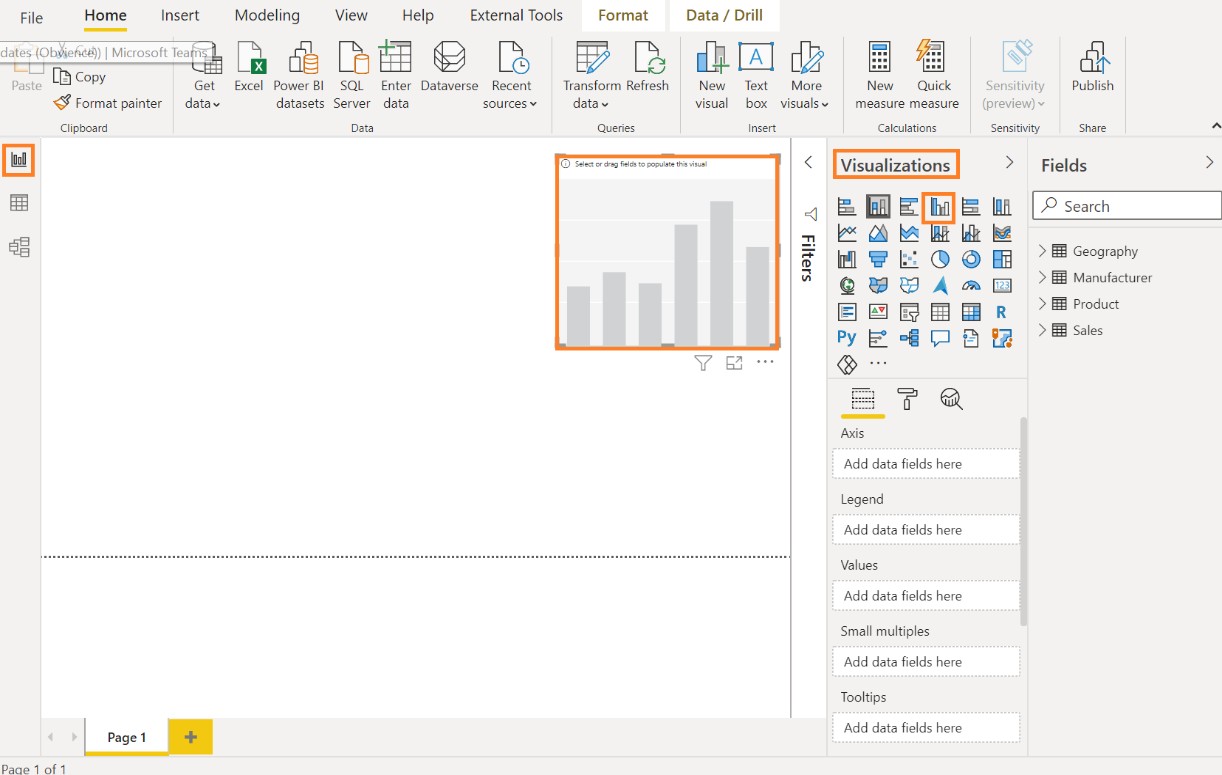


1. Click on the Search box and notice the options available. The options change based on what you are clicked on in canvas

## Power BI Desktop – Data Exploration

Now that we have loaded data, let’s start with analyzing sales by country.

1. Click on the **Report** icon on the left panel to navigate to the Report view.
2. Click the **Clustered column chart** visual in **Visualizations** as shown in the screenshot.



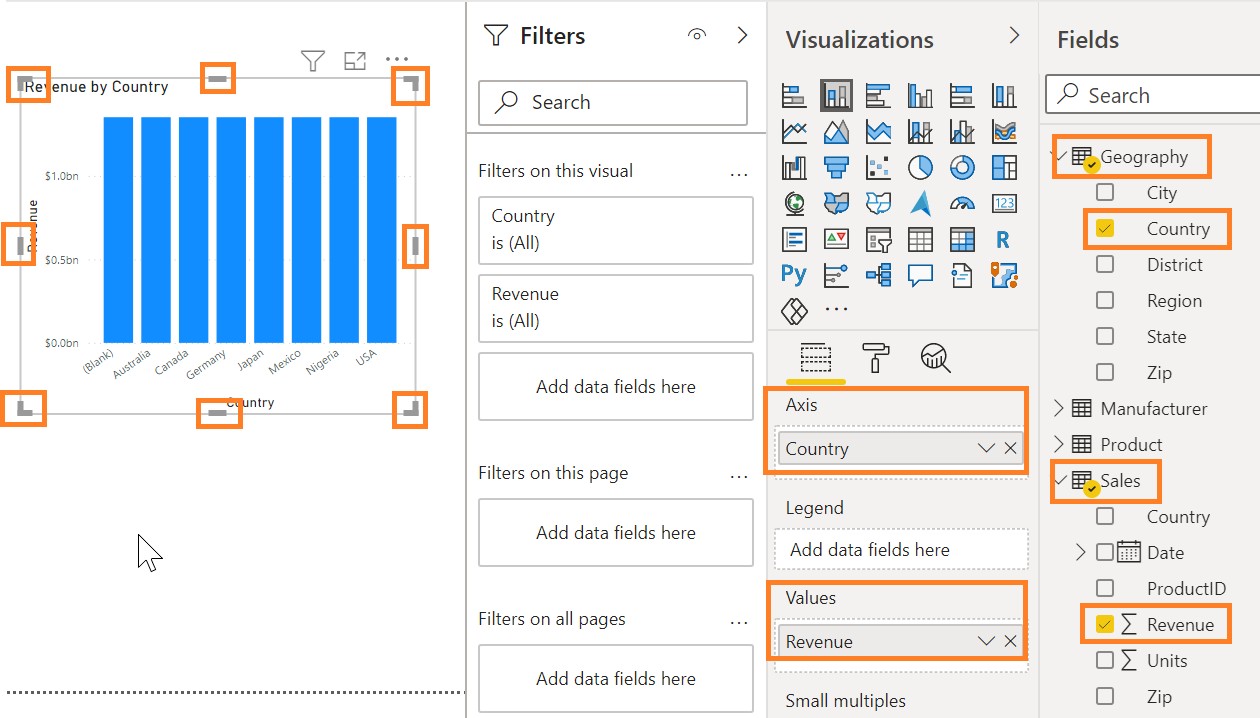
1. From the **Fields** section, expand the **Geography** table and then click the checkbox next to the

**Country** field.

1. From the **Fields** section, expand the **Sales** table and then click the checkbox next to the **Revenue**

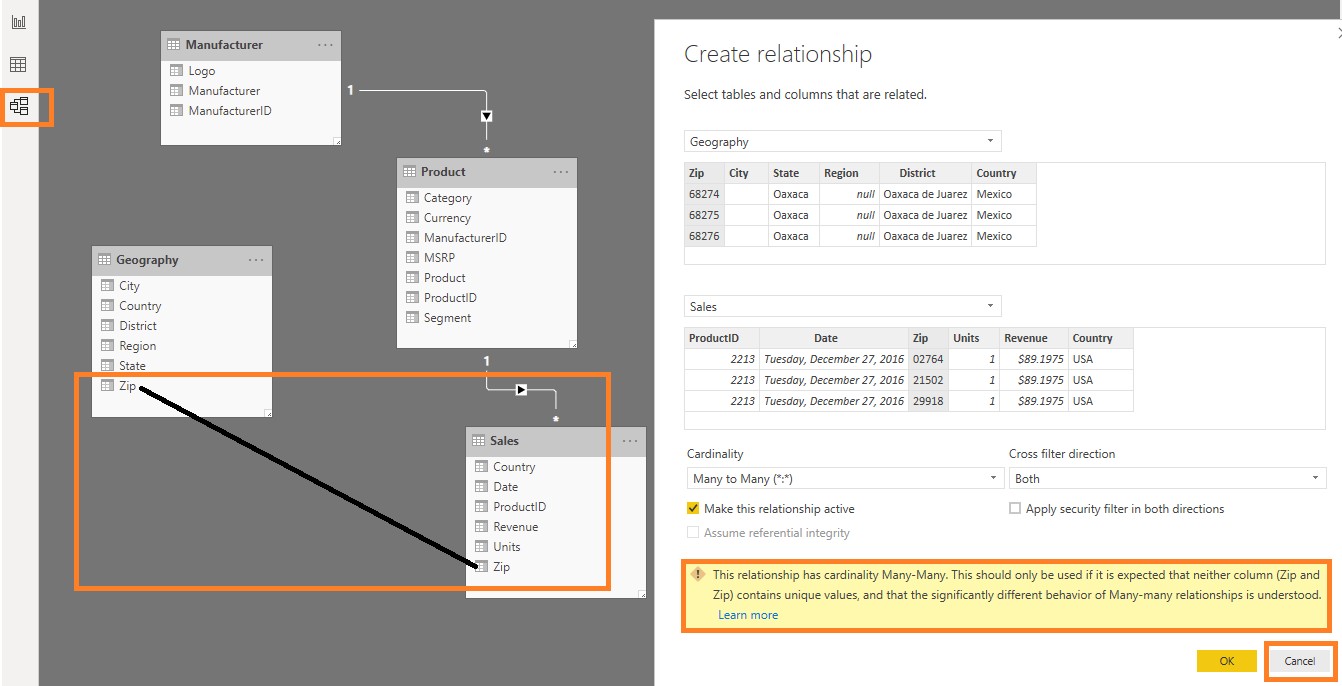
field.

1. **Resize** the visual as needed by dragging the edges.



Notice that the revenue of each country is the same. Now we need to create a relationship between the Sales and Geography tables.

1. Click on the **Model** icon on the left panel to navigate to the Relationship view.
2. Our sales data is by Zip code, so we need to connect the **Zip** column from the **Sales** table with **Zip** column in the **Geography** table. You can do this by dragging the **Zip** field in the **Sales** table to connect the line with the **Zip** field in the **Geography** table.



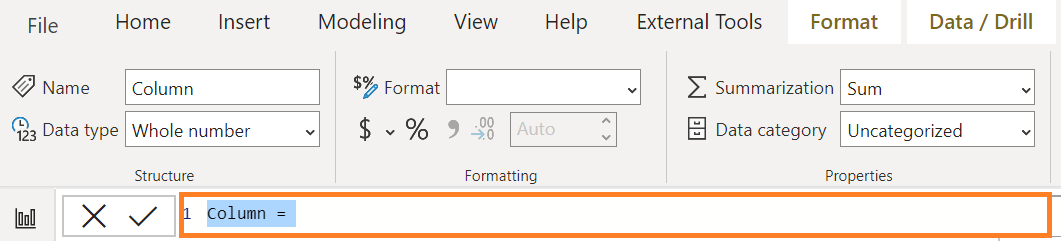
You will notice the **Create relationship** dialog opens with a warning message at the bottom stating the relationship has a many-many cardinality. The reason for the warning is that we don’t have unique Zip values in the **Geography** table. This is because multiple countries could have the same Zip code. Let’s concatenate the **Zip** and **Country** columns to create a unique value field.

1. Click **Cancel** in the **Create relationship** dialog box.

We need to create a new column in both the Geography table and the Sales table that combines the

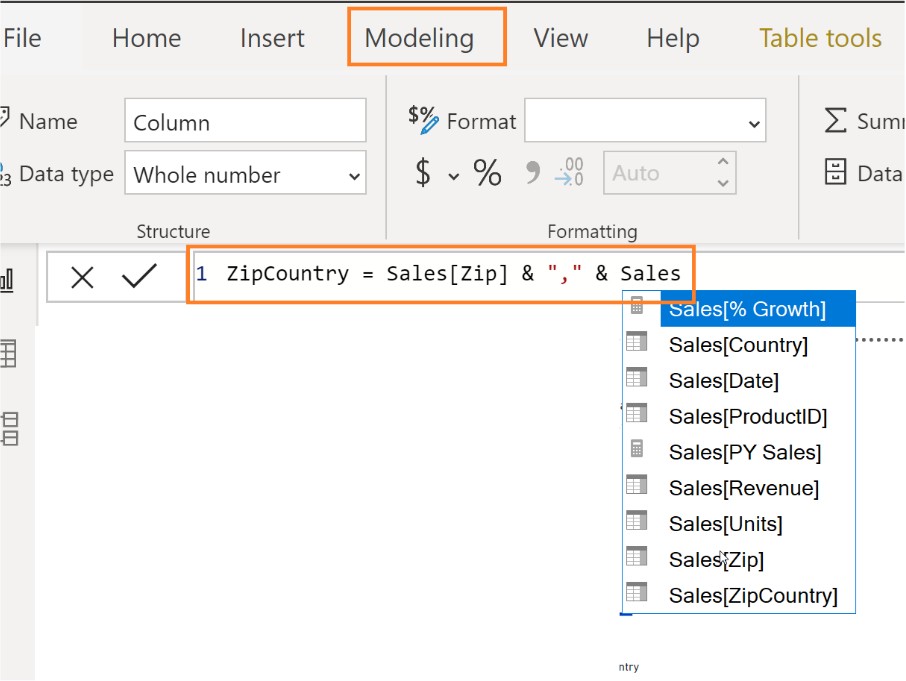
**Zip** and **Country** columns. Let’s start by creating a new column in the Sales table.

1. Click on the **Report** icon on the left panel to navigate to the **Report** view.
2. In the **Fields** section, click on the ellipse next to the **Sales** table. Click the **New Column** as shown in the figure. You will see a formula bar appear, as shown in the screenshot, to help create this new column.



1. Now we are ready to combine the Zip and Country columns into a new column called ZipCountry, separated by a comma. To create this column called **ZipCountry**, type the following calculation in the editor.

#### ZipCountry = Sales[Zip] & "," & Sales[Country]



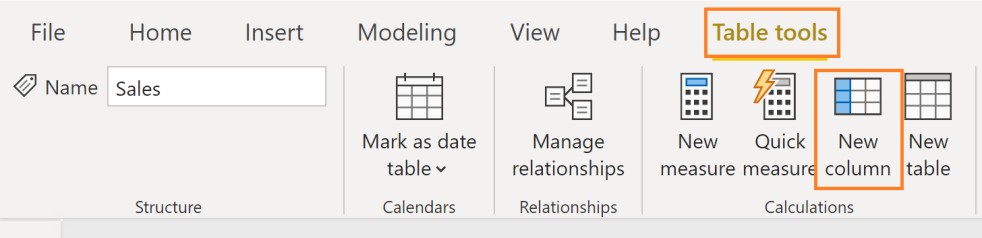
1. Once you are done entering the formula press **Enter** or click the checkmark on the left side of the formula bar.

**IMPORTANT!**

If you get an error creating a new column, make sure your Zip column is the Text Data Type.

You will notice that IntelliSense appears guiding you to choose the correct column. The language you used to create this new column is called Data Analysis Expression (DAX). We are connecting columns (Zip and Country) in each row by using the “&” symbol. The icon with an (fx), near the new column ZipCountry, indicates that you have a column containing an expression, also referred to as a calculated column.

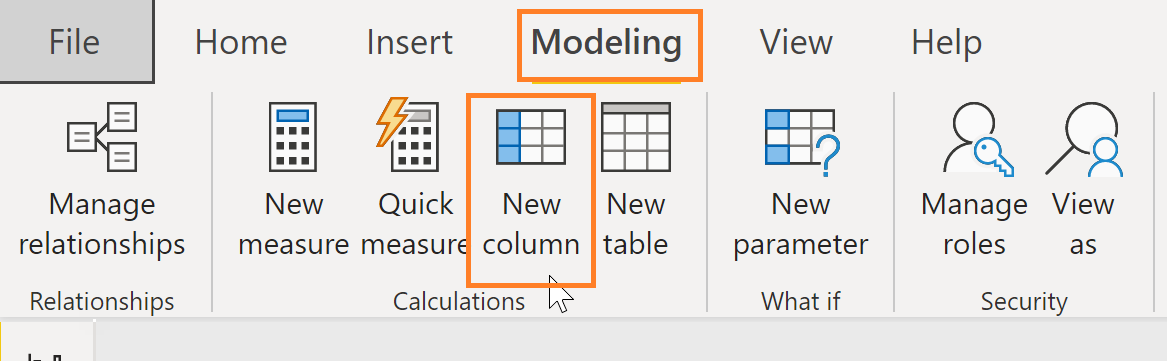
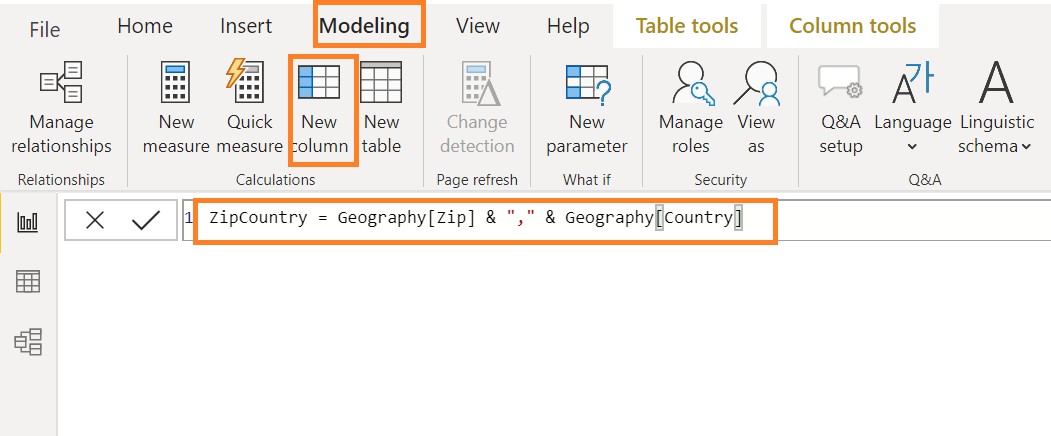
**Note**: An alternative way to add a new column is by selecting the table, click **Table Tools,** click **New Column** or **Modeling**, and then click **New Column** from the ribbon.



Let us use this method to create a **ZipCountry** column in the **Geography** table.

1. From the **Fields** section, click the **Geography** table, from the ribbon click **Modeling**, and then click

**New Column** as shown in the figure.



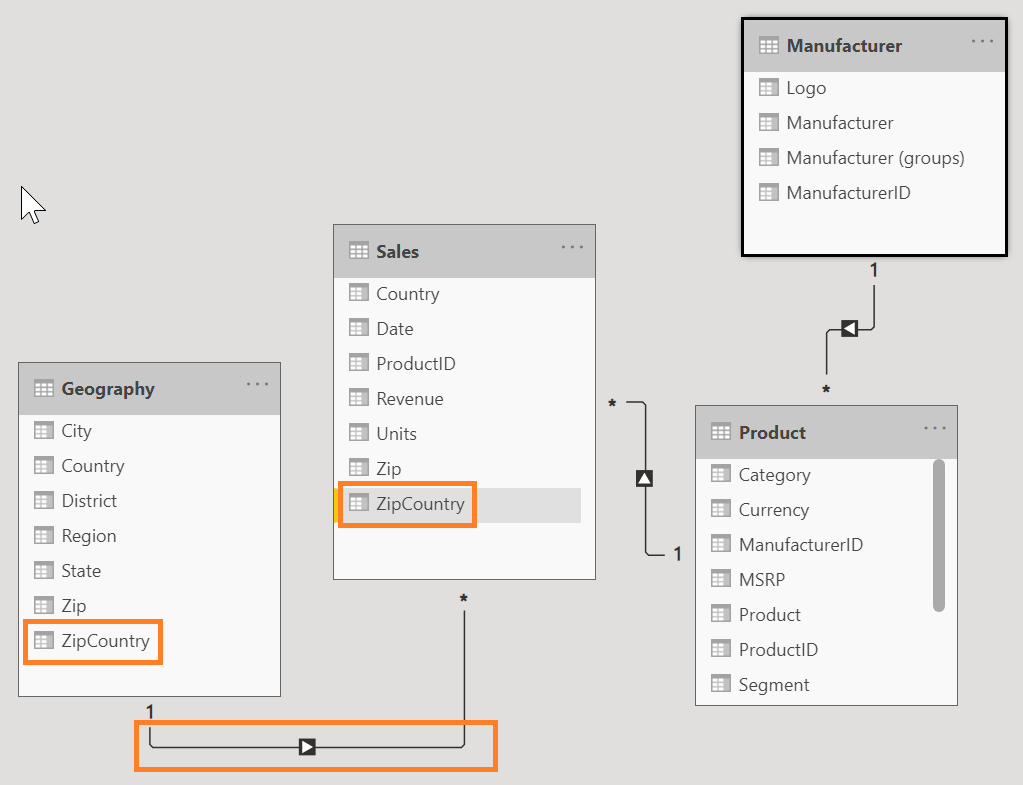
1. A formula bar now appears. Enter the following DAX expression in the formula bar:

#### ZipCountry = Geography[Zip] & "," & Geography[Country]

You will see a new column, **ZipCountry**, in the **Geography** table. The final step is to set up the relationship between the two tables using the newly created **ZipCountry** columns in each of these tables.

1. Click on the **Model** icon on the left panel to navigate to the **Relationship** view.
2. Drag the **ZipCountry** field from the **Sales** table and connect it to the **ZipCountry** field in the

**Geography** table.



Now we have successfully created a relationship. The number “1” next to Geography indicates it is on the one side of the relationship and the “\*” next to Sales indicates it is on the many side of the relationship.

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

Notice the clustered column chart that we created earlier. It shows different sales for each country or region. USA has the most sales, followed by Australia and Japan. By default, the chart is sorted by **Revenue**.

1. Click on the **ellipse** on the top right corner of the visual (alternatively, the ellipse may be at the bottom of the chart). Notice there is an option to Sort by **Country** as well.

