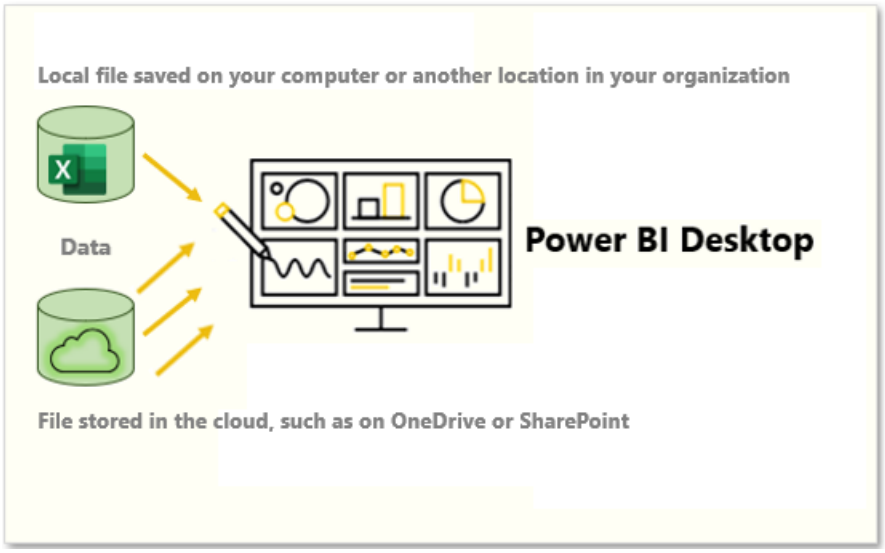
1. Creating Simple Report

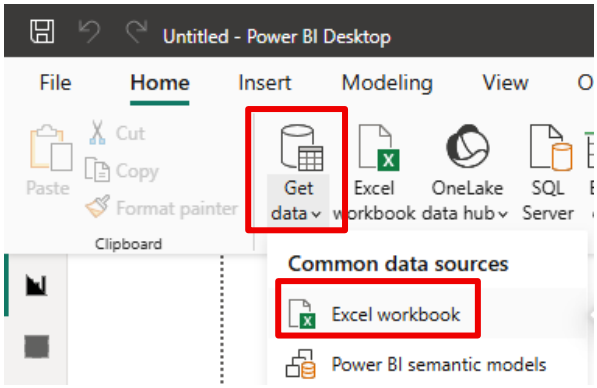
Organizations often export and store data in files. One possible file format is a flat file. A flat file is a type of file that has only one data table and every row of data is in the same structure. The file does not contain hierarchies. Likely, you're familiar with the most common types of flat files, which are comma-separated values files, delimited text files, and fixed width files. Another type of file would be the output files from different applications, like Excel workbooks.

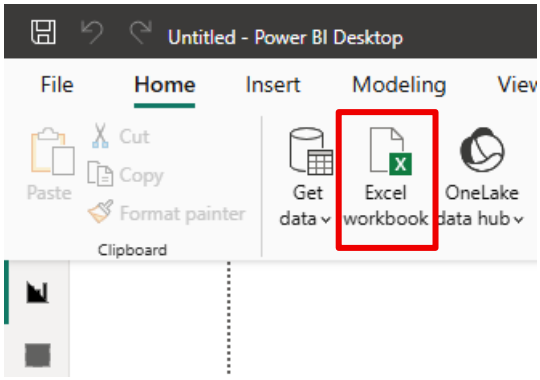
Power BI Desktop allows you to get data from many types of files. You can find a list of the available options when you use the Get data feature in Power BI Desktop.



1.1 Connect to Excel workbook

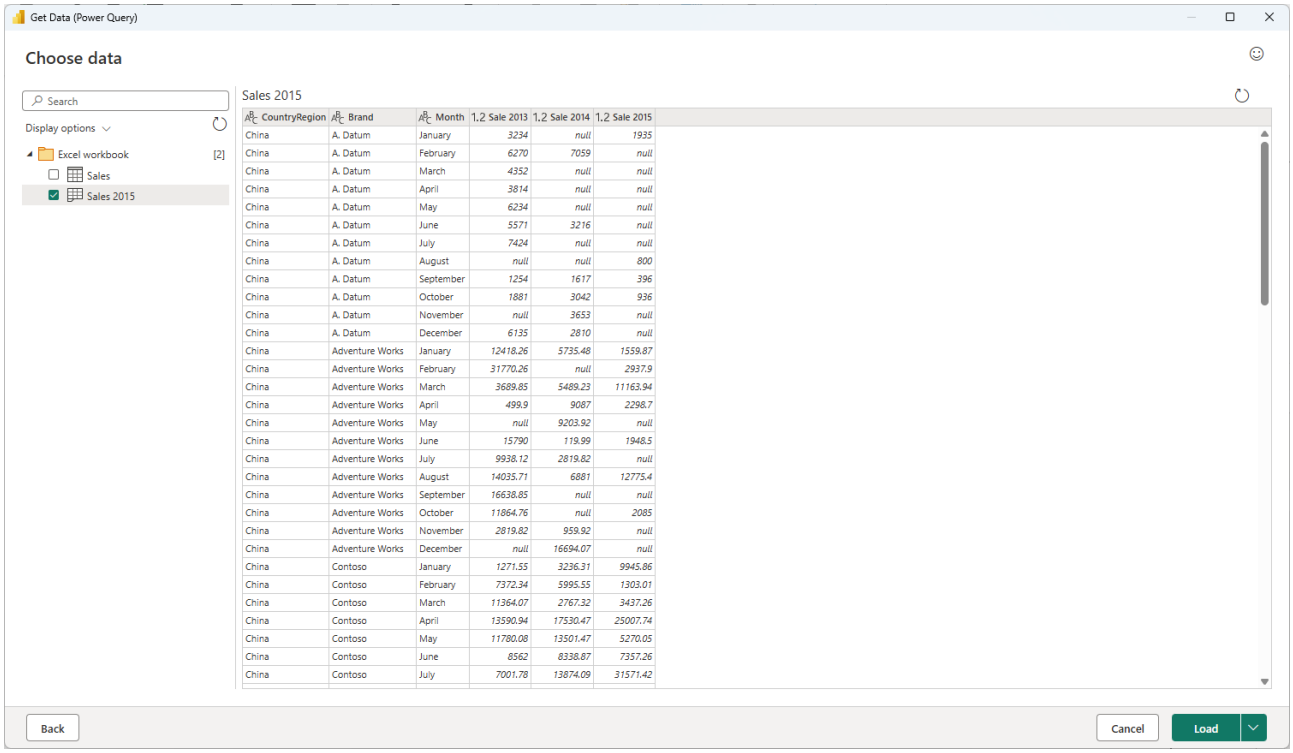
1. In Power BI, on the **Home** tab, select **Get data**. In the list that displays, select the option that you require, such as Text/CSV or XML. For this example, you will select **Excel Workbook**.



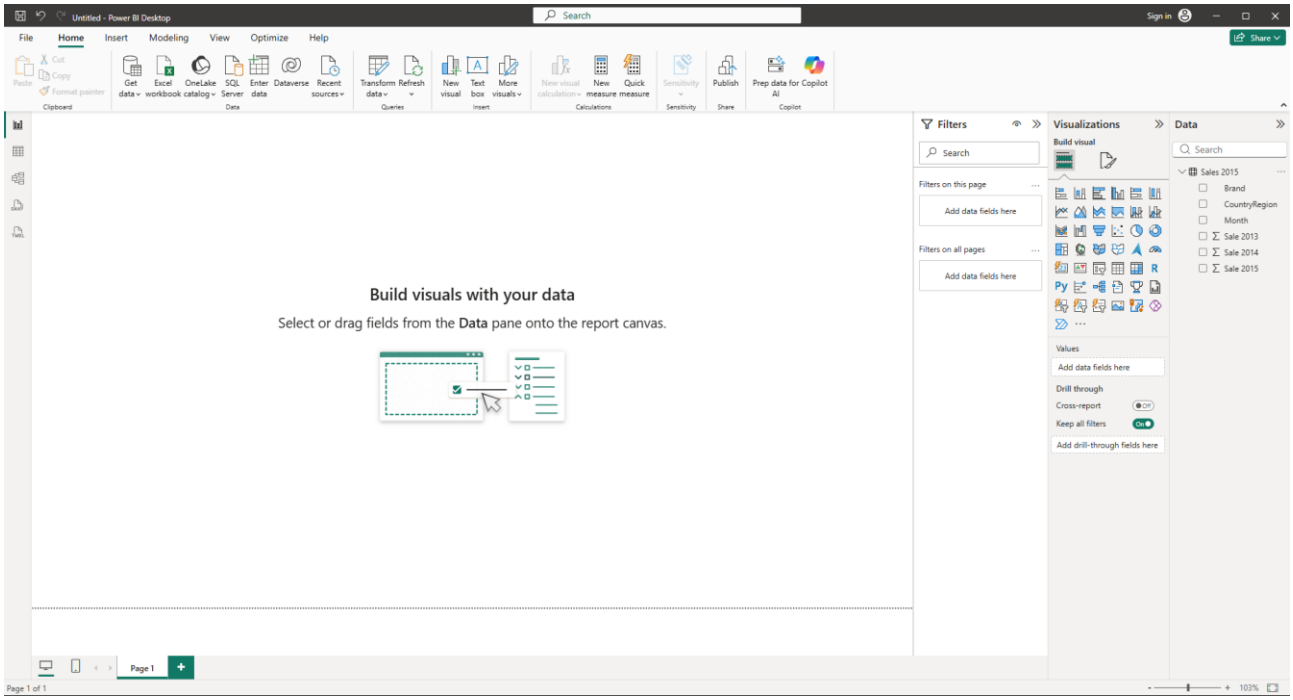


2. Depending on your selection, you need to find and open your data source. You might be prompted to sign into a service, such as OneDrive, to authenticate your request. In this example, you will open the *2015 Sales.xlsx* Excel workbook.
3. After the file has connected to Power BI Desktop, the Navigator window opens. This window shows you the data that is available in your data source. You can select a table or entity to preview its contents, to ensure that the correct data is loaded into the Power BI model.

4. Select the check box of the table that you want to bring in to Power BI. This selection activates the Load and Transform Data buttons as shown in the following image.



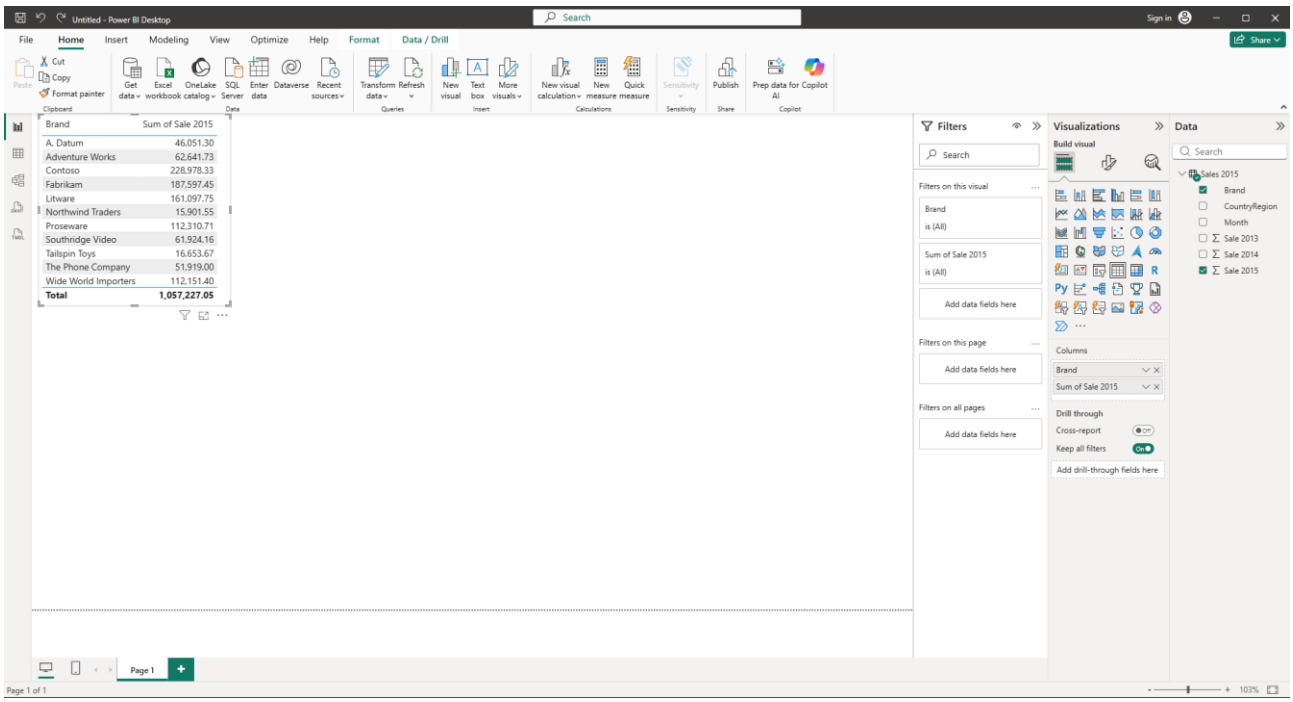
5. The file should be loaded successfully to the PowerBI, and you are able to create your report.



1.2 Creating Report

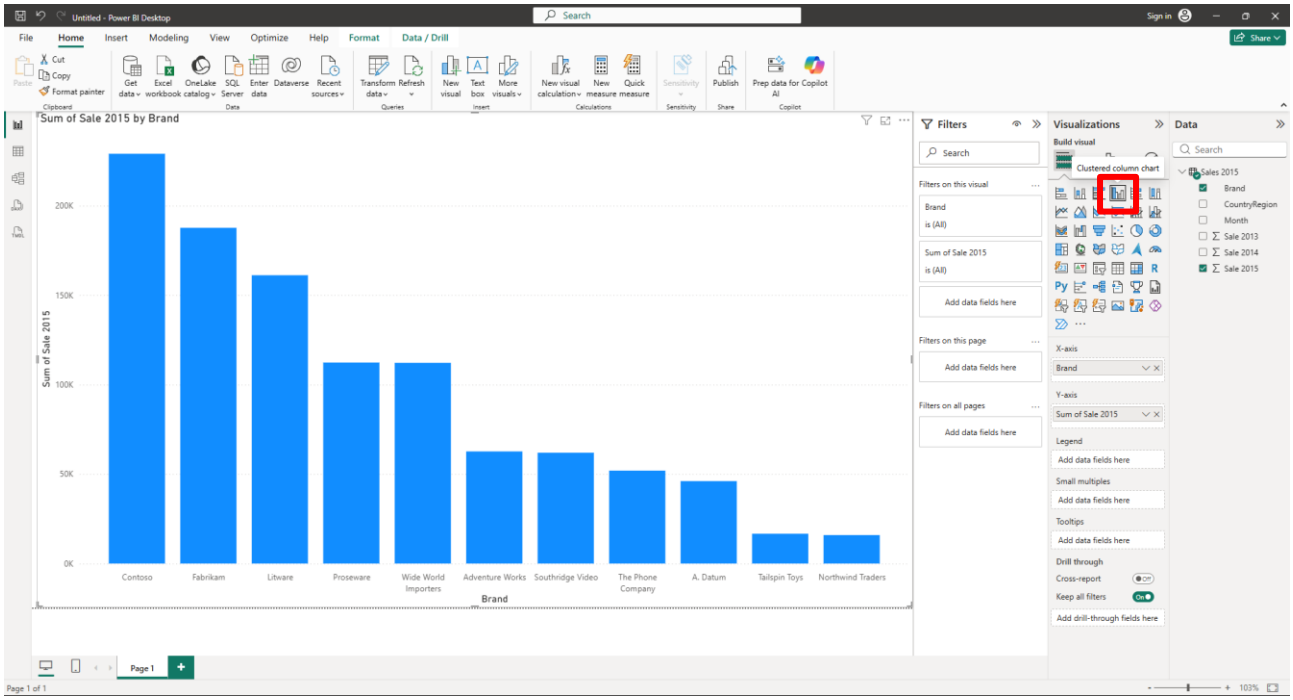
1.2.1 Create Table

Select *Brand* and *Sales 2015* in the **Data** panel, a summary table will be created in the report area.



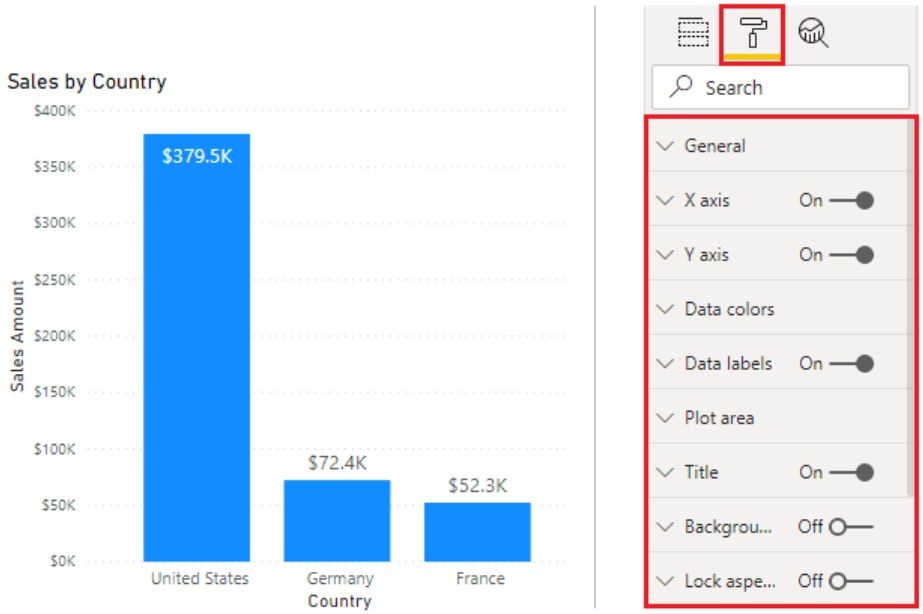
1.2.2 Create Report

Select **Cluster Column Chart** in **Visualization** panel to change the report type. The default sort order is "Sort Descending".



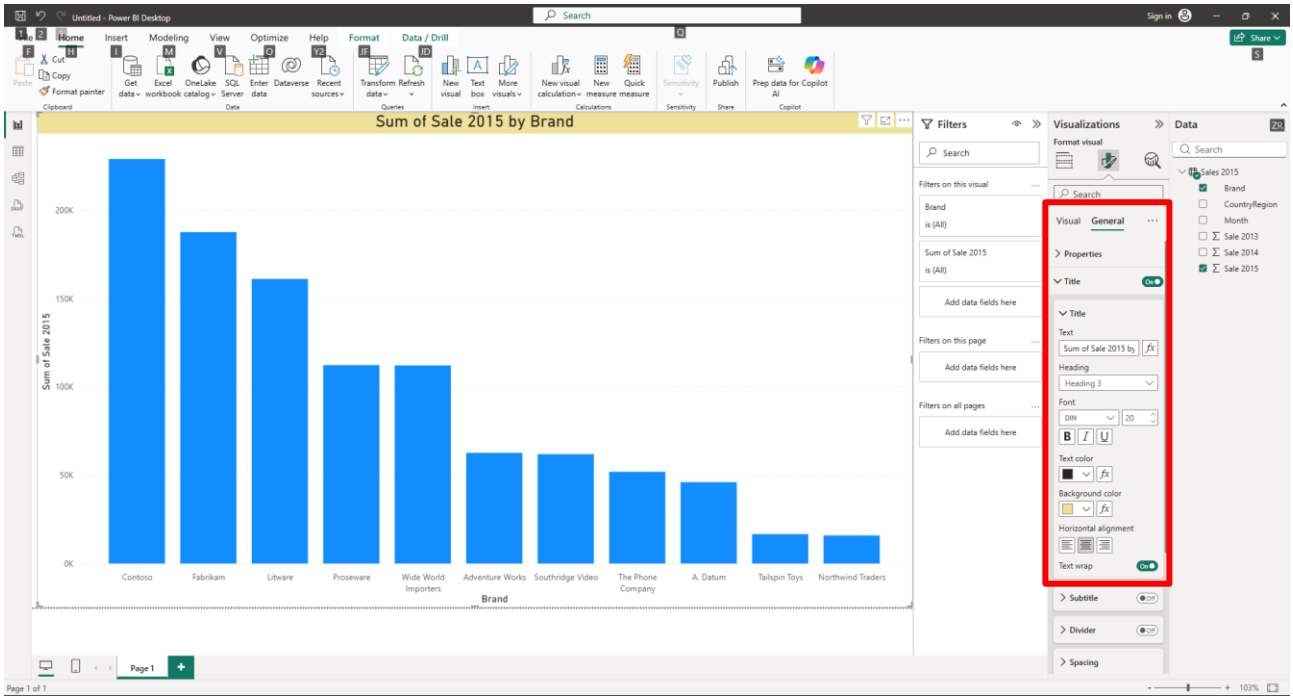
1.3 Custom Format

Power BI Desktop gives you a variety of options for customizing how your selected visualizations look, such as the colors and format of the text that they contain. You should take time to explore the options to determine what impact they each have on a visual. Start by selecting the visualization on the canvas, and then select the Format button (paint roller icon) to display the Format pane.



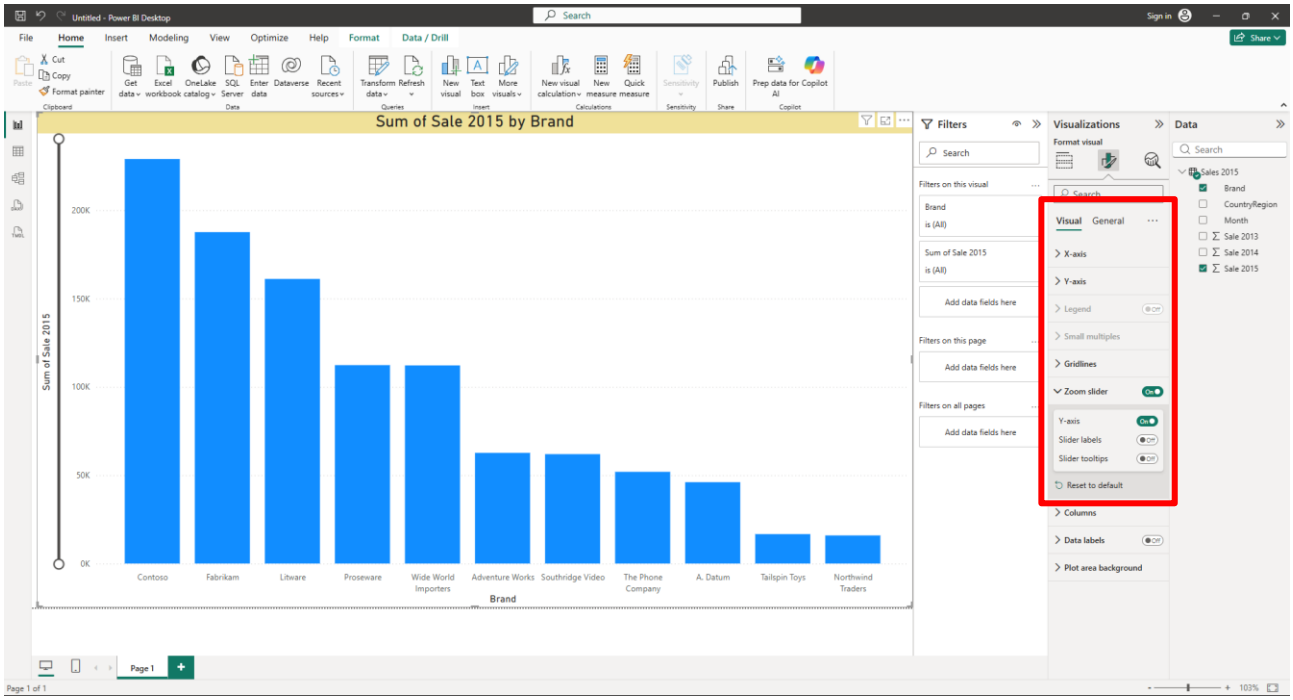
1.3.1 Title

You can edit a default title and add a title, if you don't have one. In this example, you will select the column chart visualization and then, in the Format pane, scroll down and expand the Title section. Change the ***Title Text*** (Annual Sale by Brand - 2015), ***Background*** (#F0E199), ***Alignment*** (Center), and ***Text Size*** (20) in **General** tab.



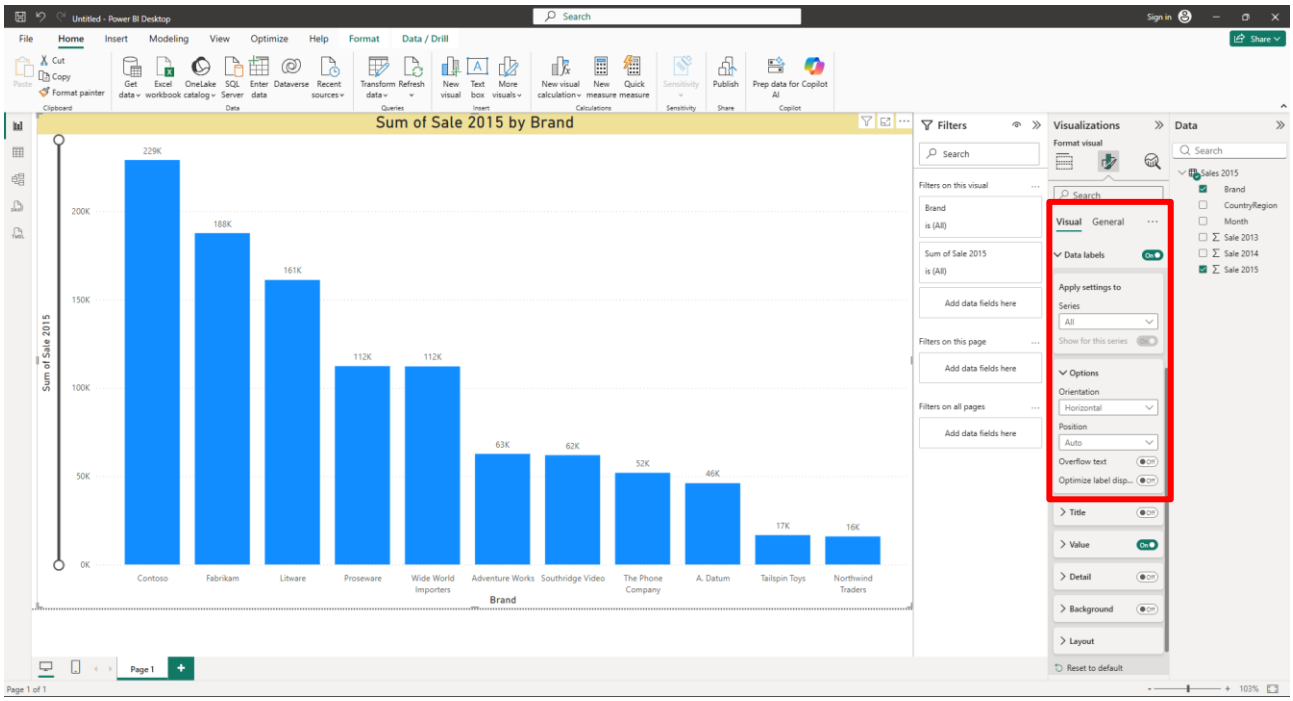
1.3.2 Zoom Slider

Enabling a **Zoom Sliders** feature on chart allows you to easily examine a smaller range of the data in a chart without having to use a filter.



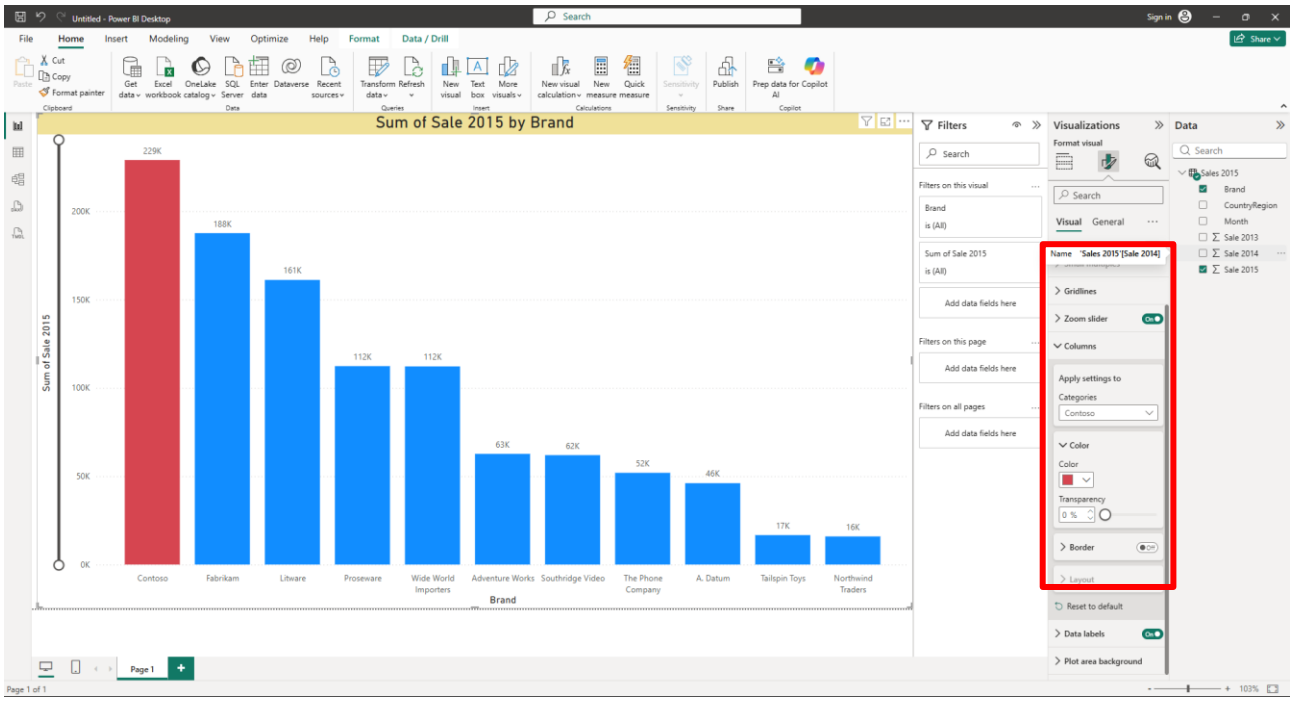
1.3.3 Data Label

Turn on the Data label if you want to show all the values in the Chart. By adjusting the **Display units**, the figure can be shown in thousand, million, etc.



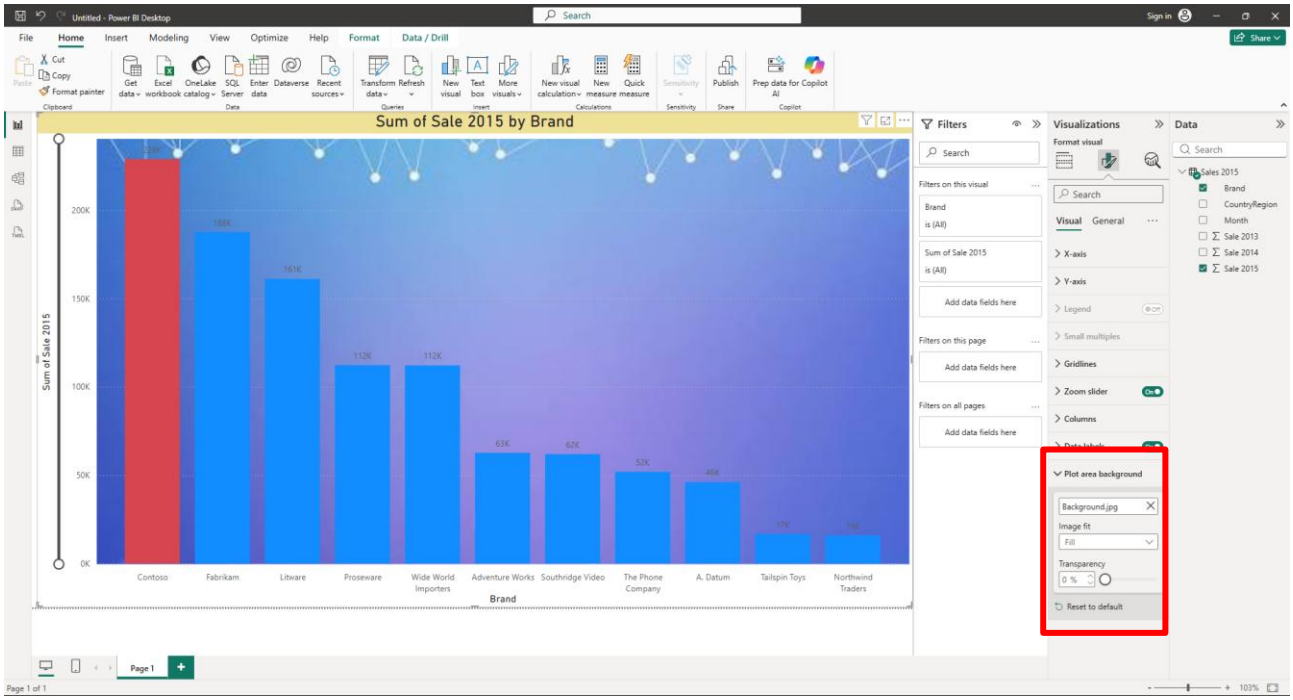
1.3.4 Column Color

By selecting the categories, the color and transparency percentage can be defined by columns.



1.3.5 Plot Area Background

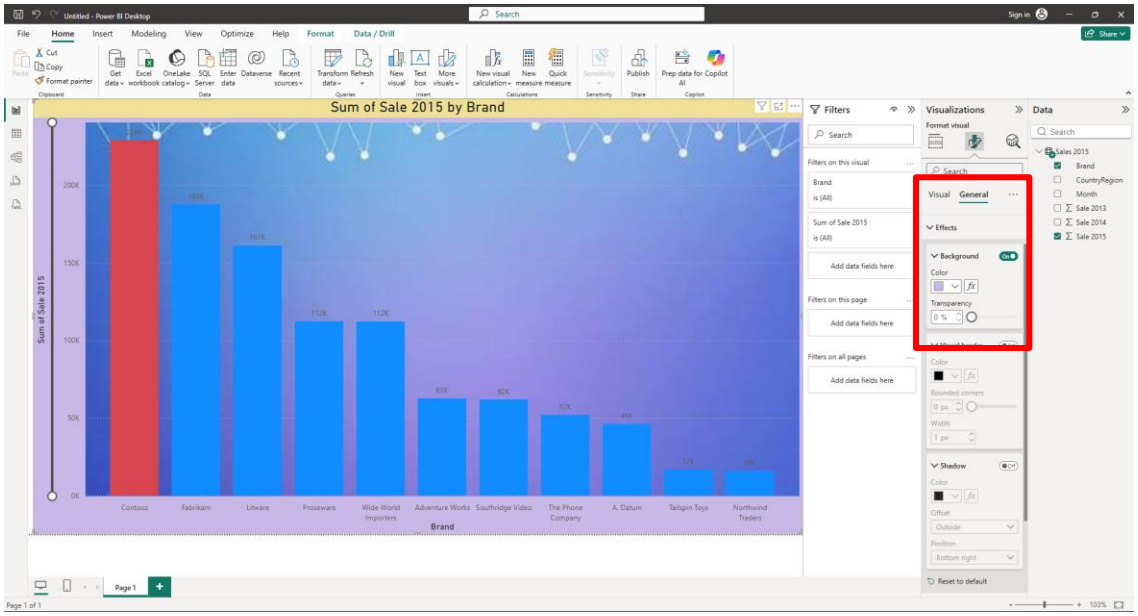
1. Use the **[Add image]** button to add the background picture, then change the Image Fit to "**Fill**".



1.3.6 Background

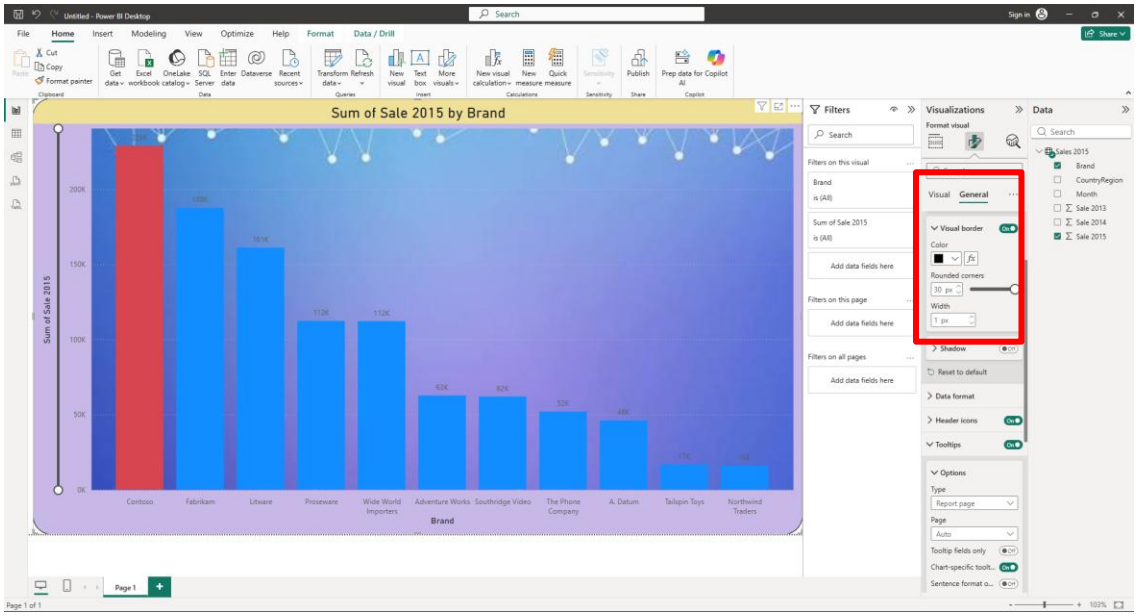
In the Background section, you can set any color or image as the background for the visual. If you plan to use an image as a background, try to select an image that won't have lines or shapes that would make it difficult for the user to read the data. It is best to keep a white background so the presented data can be clearly seen.

It is best practice to keep the default white background so the presented data can be clearly seen. However, you can change the default background color to make a visualization more colorful and easier to read or to match a particular color scheme.



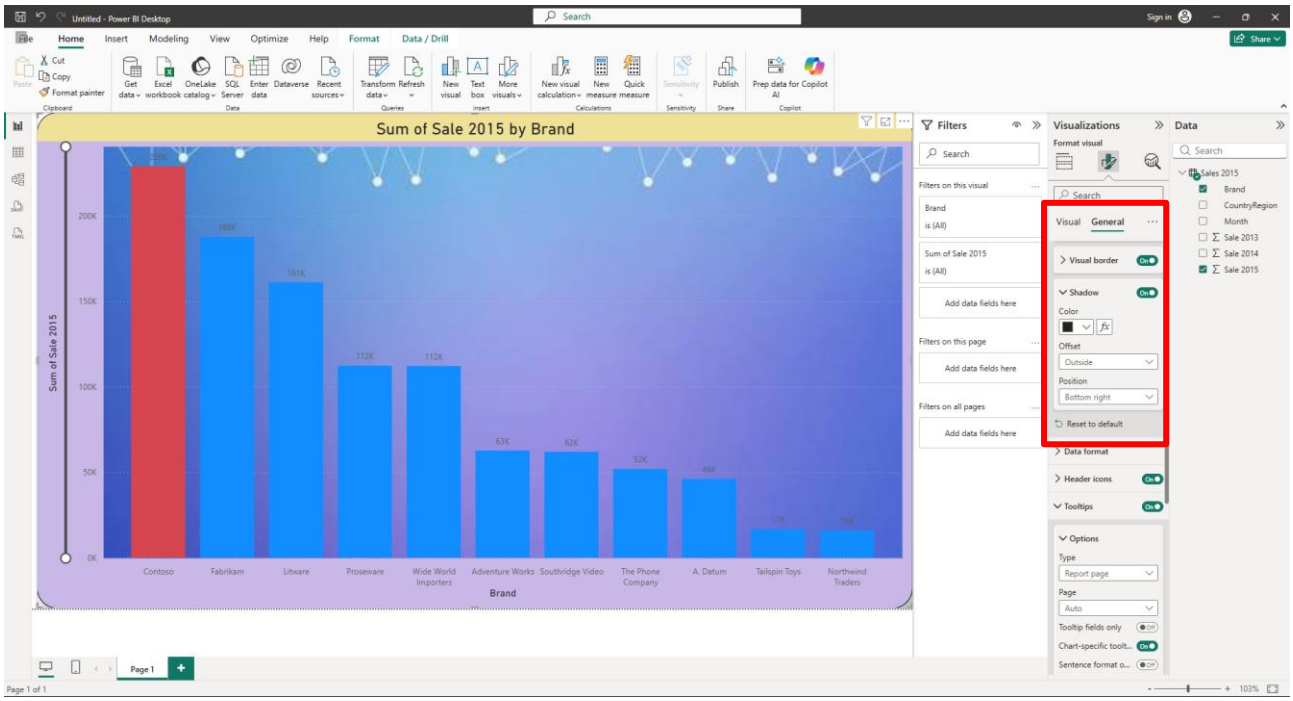
1.3.7 Visual Border

In the Border section, you can set a border around the visual to isolate the visual from other elements on the canvas, which helps make it easier for the user to read and understand the data. You can change the border color and radius to be consistent with your color scheme.



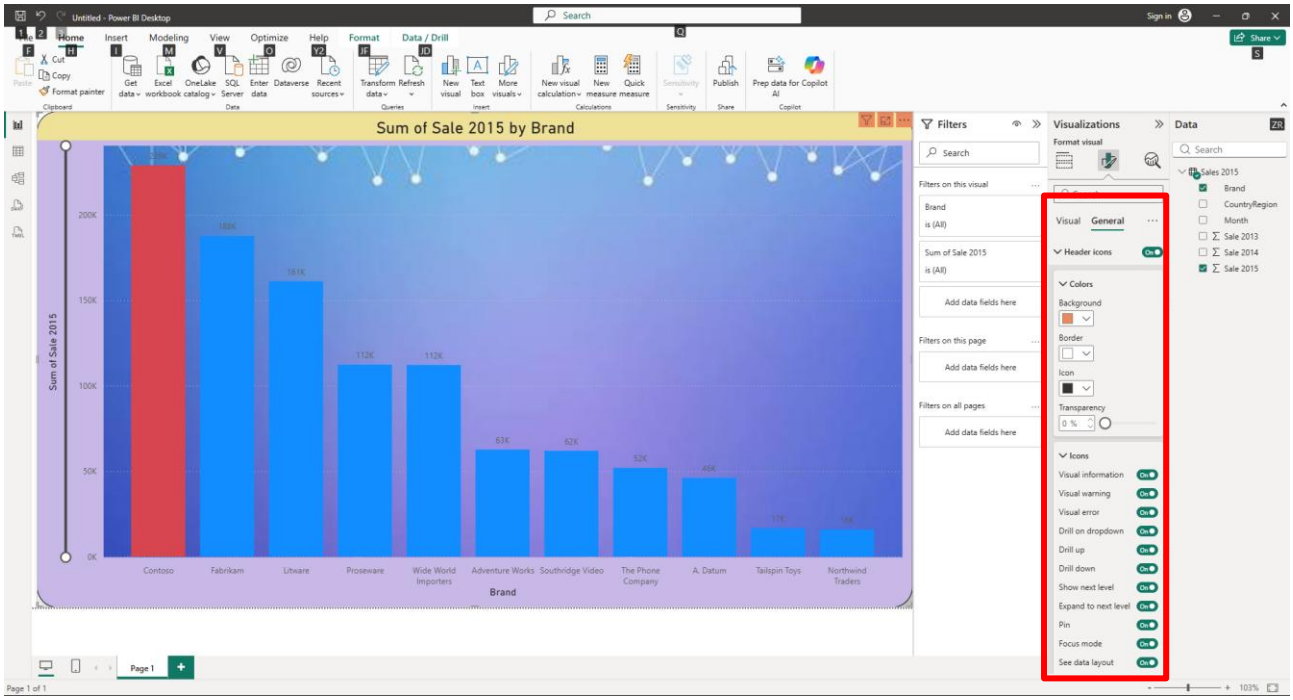
1.3.8 Shadow

By turning on the Shadow, the shadow color and position can be adjusted.



1.3.9 Header Icon

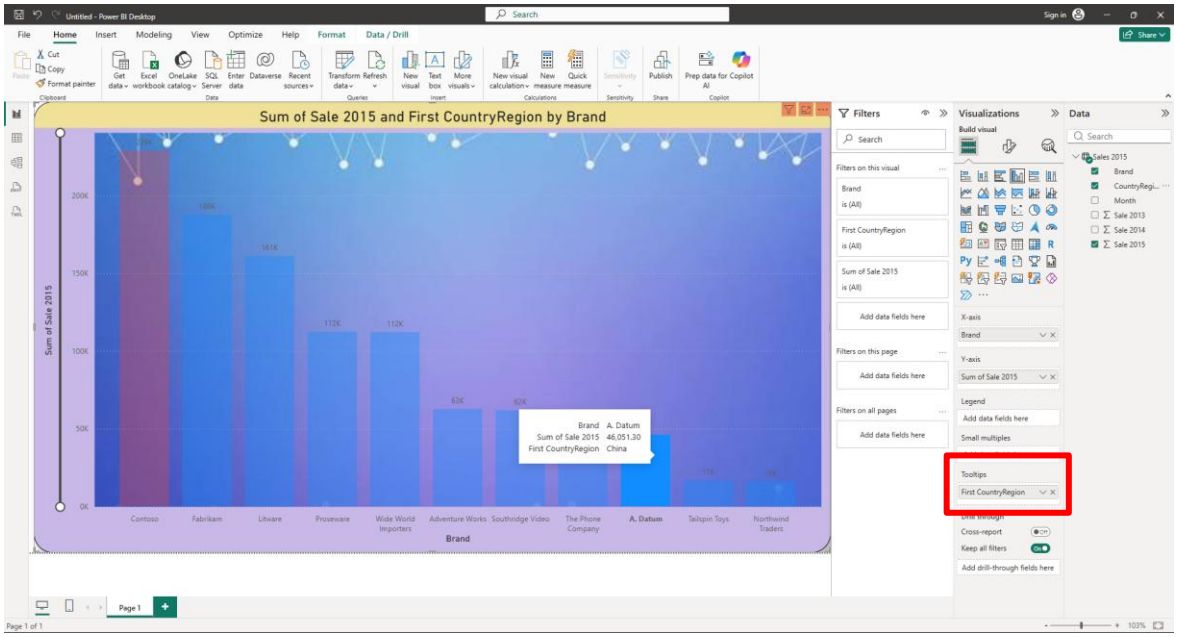
Each visual also has a card in the Formatting section of the Visualizations pane called header icon. In that card you can adjust all sorts of characteristics of the header icon



1.3.10 Tooltip

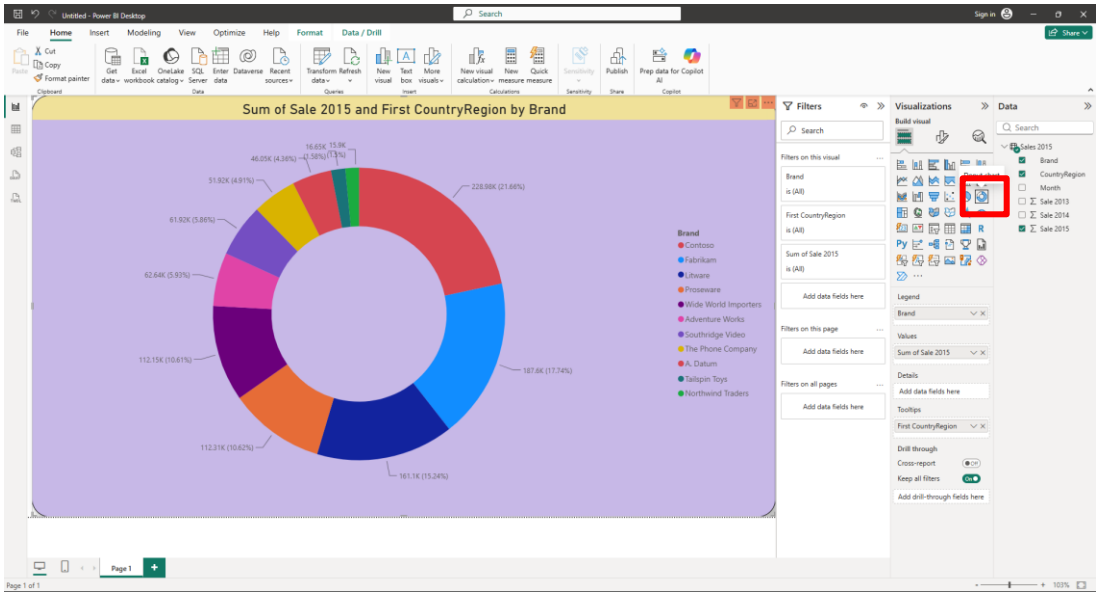
Using tooltips is a clever way of providing more contextual information and detail to data points on a visual. When you add a visual, the default tooltip displays the data point's value and category, but you can customize this information to suit your needs. For example, you might want to provide your report users with additional context and information or specify additional data points that you want users to see when they hover over the visual.

To expand on the data points that are displayed in the default tooltip, you can drag a field from the Fields panel into the Tooltips bucket. However, you should not add many more fields to the tooltips because adding too many fields can introduce performance issues and slow down your visuals.



1.4 Change Visualization

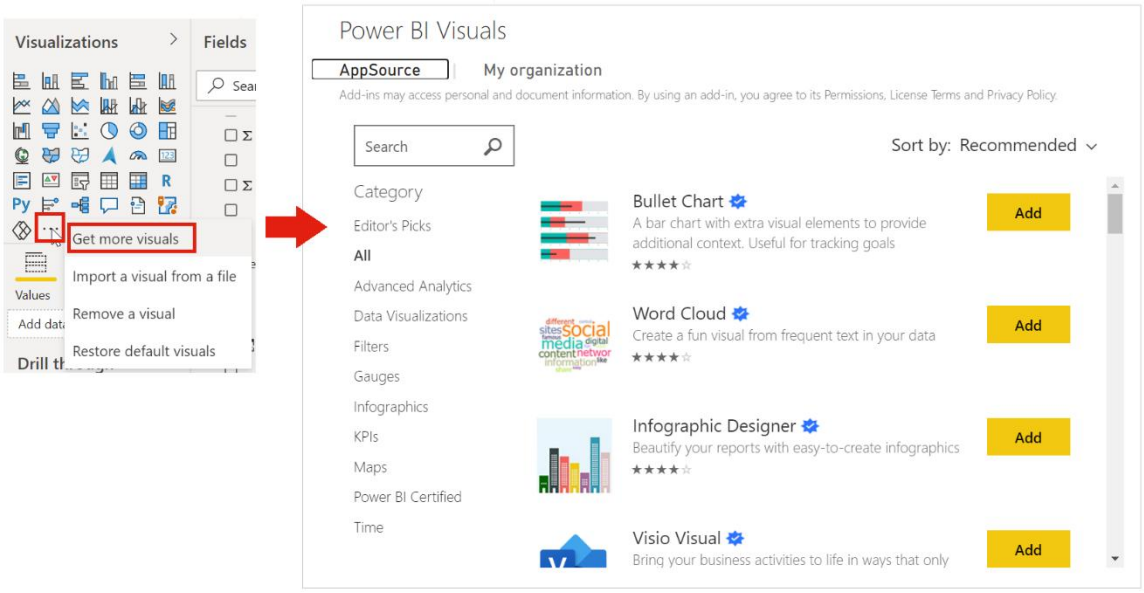
In order to change the visualization, you can select another report type (such as **Donut Report**) in **Visualization** panel.



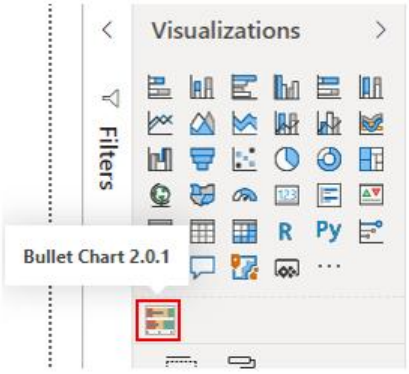
1.5 Import Custom Visual

In addition to the out-of-the-box visualizations in Power BI Desktop, hundreds of other developers have created a multitude of visuals for you to choose from. If you have a specific visual in mind, you can likely find it in the marketplace. If you can't find it, Power BI makes it possible for you to build your own.

In the **Visualizations** pane, select the **Get more visuals** icon and then select **Get more visuals**. On the window that displays, locate and select the visual that you want to import and then select **[Add]**.



The new visual will appear under the other visuals in the Visualizations pane. To add the visual to your report, select its icon. You can then add fields to the visual and customize its formatting, just like you would for any other visual.

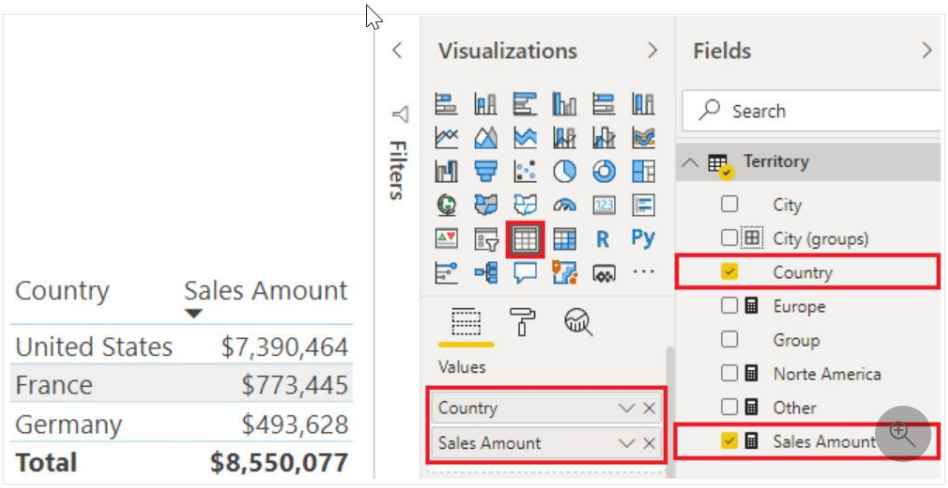


2. Introduction to Power BI Visual

Power BI visuals are attractive charts and graphics that you can use to revitalize your data. Visuals allow you to share data insights more effectively and increase comprehension, retention, and appeal. Visuals are a fundamental part of your report because they help your report audience connect and interact with the information to make informed decisions quickly.

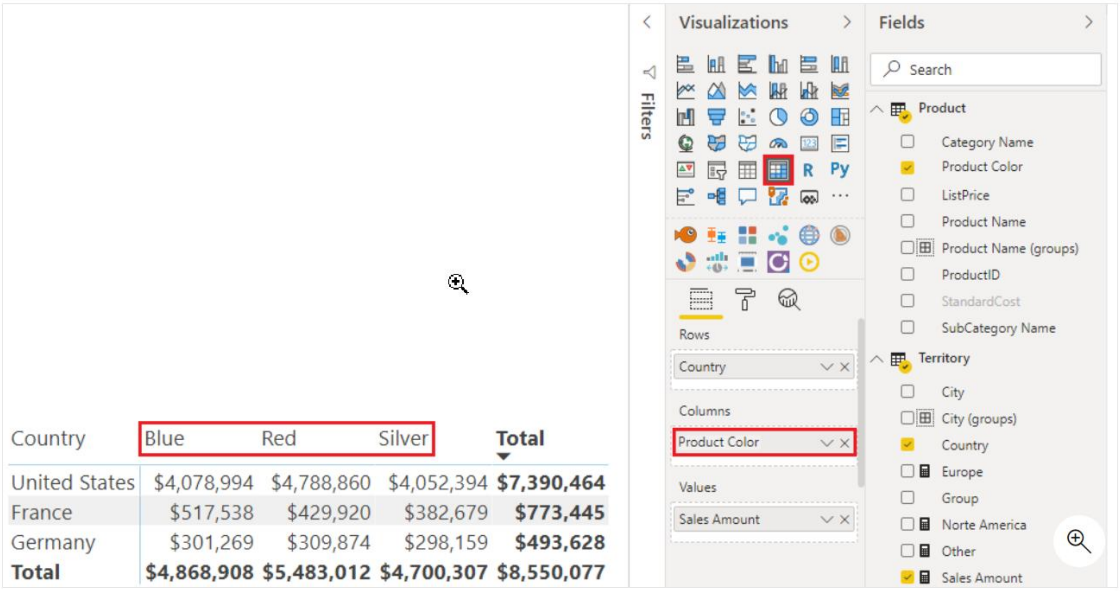
2.1 Table

The **Table** (default visualization) is a grid that contains related data in a logical series of rows and columns. The table supports two dimensions and the data is flat, which means that duplicate values are displayed and not aggregated. It can also contain headers and a row for totals.



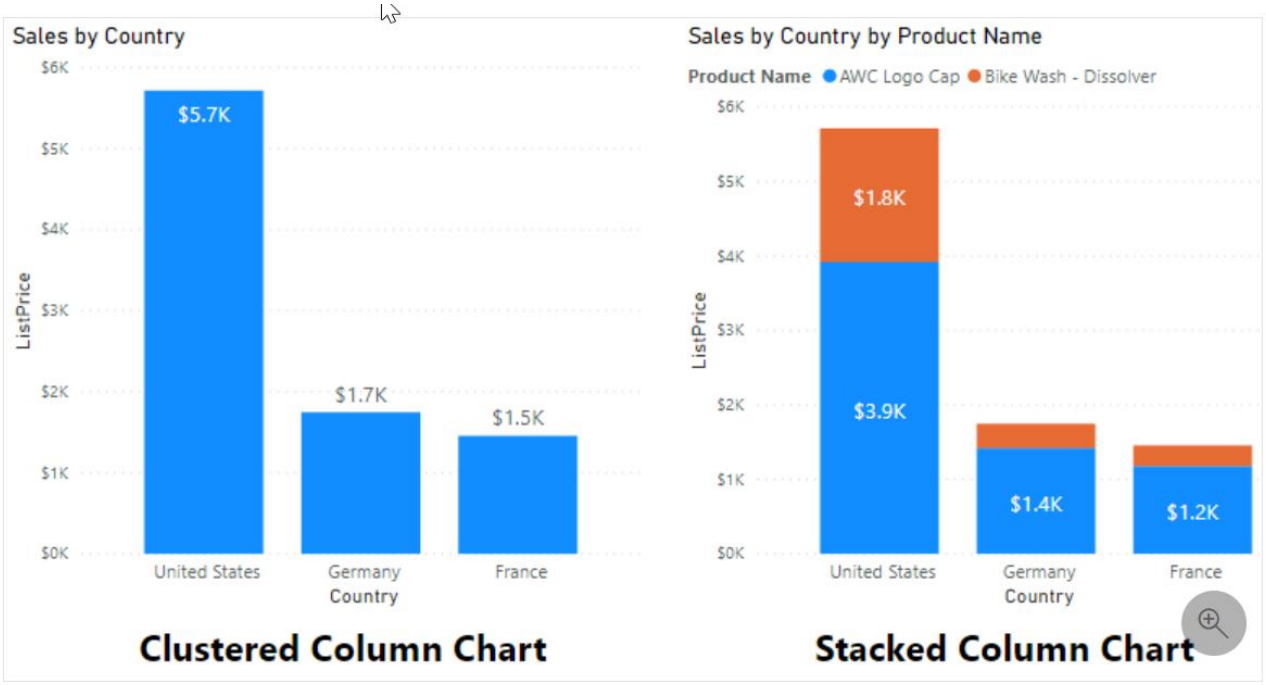
2.2 Matrix

The Matrix visualization looks similar to the table visualization; however, it allows you to select one or more elements (rows, columns, values) in the matrix to cross-highlight other visuals on the report page.



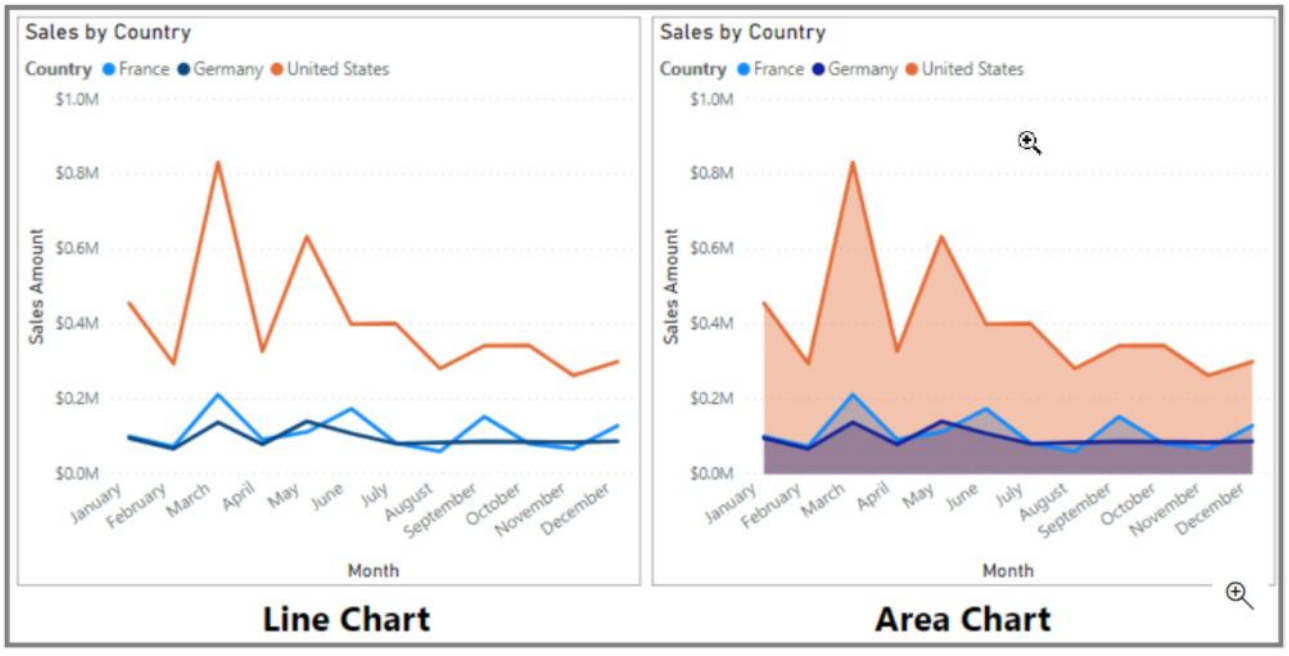
2.3 Bar and Column Charts

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.



2.4 Line and Area Charts

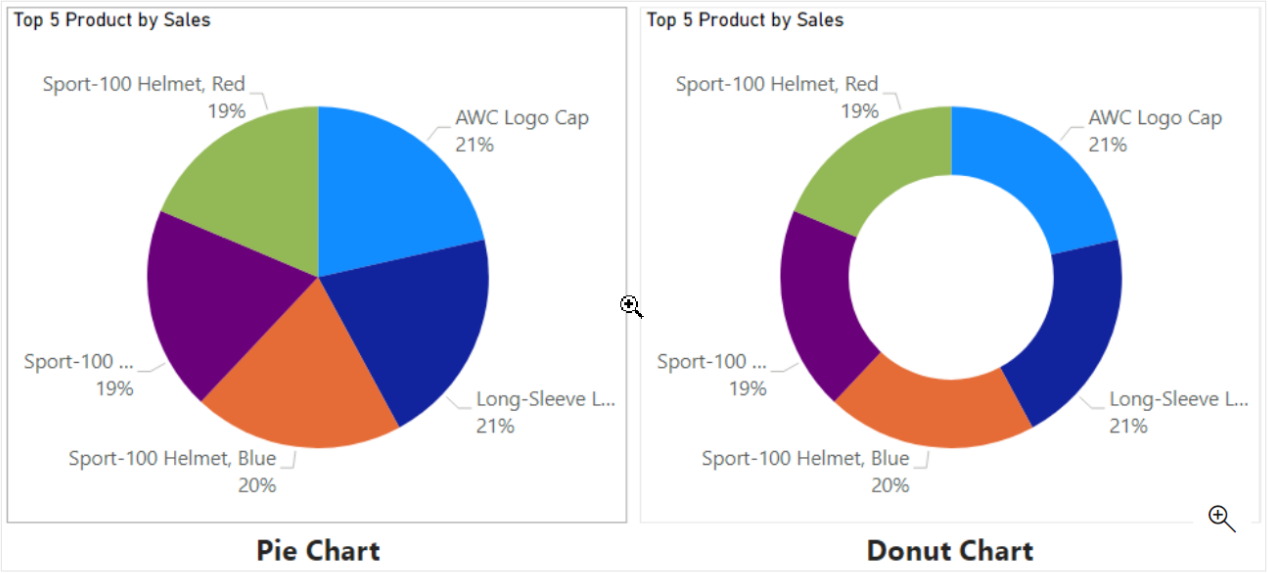
The line chart and area chart visualizations are beneficial in helping to present trends over time. The basic area chart is based on the line chart, with the area between axis and line filled in. The main difference between these two chart types is that the area chart highlights the magnitude of change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



2.5 Pie Chart, Donut Chart, and Treemaps

The pie chart, donut chart, and Treemap visualizations show you 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, such as the top five sales by product or country, or any other available categories.

The pie chart is a solid circle, whereas the donut chart has a center that is blank and allows space for a label or icon.

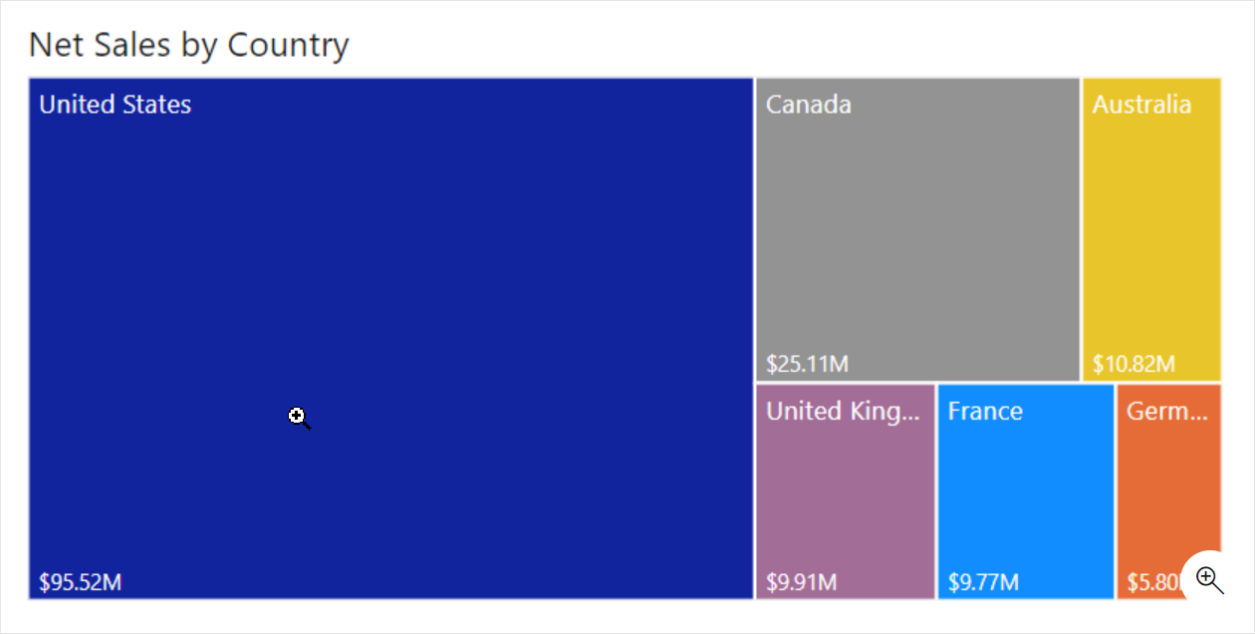


When using pie charts, donut charts, and Treemaps, try to avoid presenting too many categories because it results in thin slices (or rectangles) that provide no added value to the user. If you do need to present all categories in your dataset, it's better to use another type of visual, such as a column chart.

Pie charts and donut charts present data by dividing it into slices, while the Treemap visualization displays data as a set of nested rectangles. Each level of the hierarchy is represented by a colored rectangle (branch) containing smaller rectangles (leaves). The space inside each rectangle is allocated based on the value that is being measured. The rectangles are arranged in size from top left (largest) to bottom right (smallest).

A Treemap is ideal to visualize:

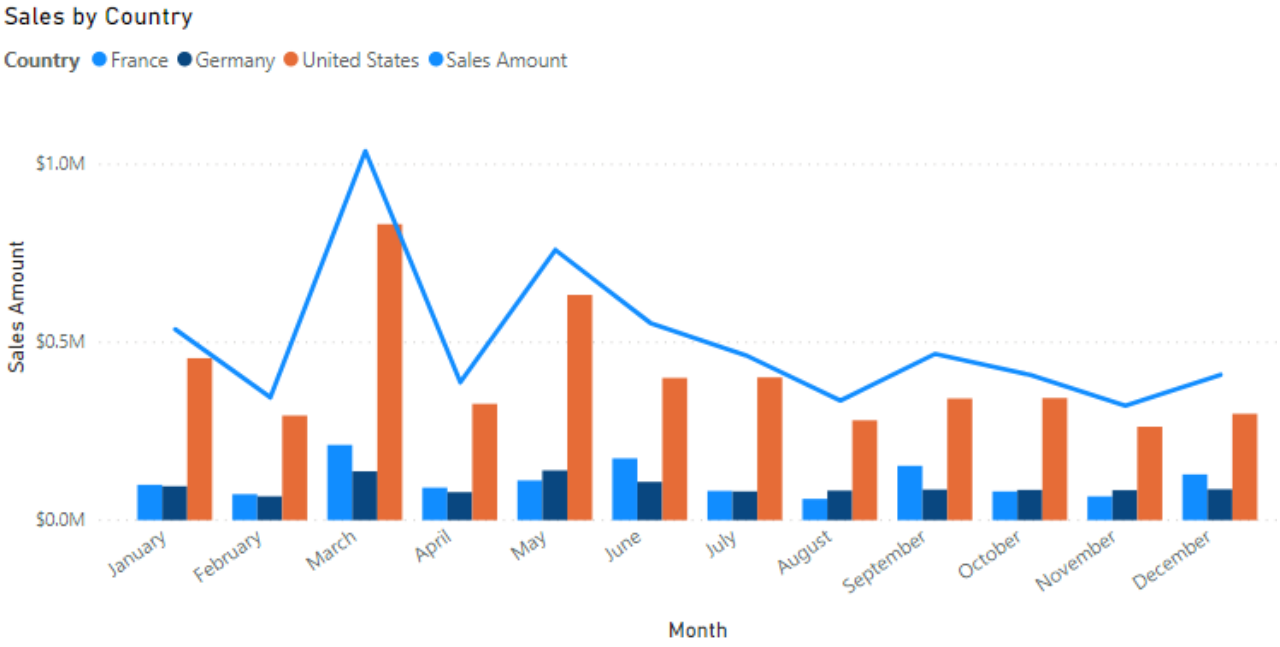
- Large amounts of hierarchical data when a bar chart can't effectively handle the large number of values.
- Proportions between each part and the whole.
- The distribution pattern of the measure across each level of categories in the hierarchy.
- Attributes, by using size and color coding.
- Spot patterns, outliers, most-important contributors, and exceptions.



2.6 Combo Charts

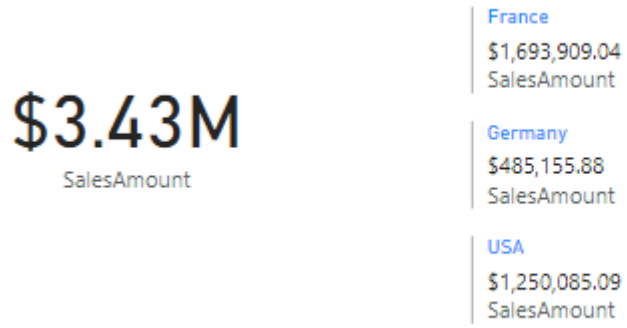
The combo chart visualization is a combination of a column chart and a line chart that can have one or two Y axes. The combination of the two charts into one lets you:

- Compare multiple measures with different value ranges.
- Illustrate the correlation between two measures in one visual.
- Identify whether one measure meets the target that is defined by another measure.
- Conserve space on your report page.



2.7 Card Visualization

The card visualization displays a single value: a single data point. This type of visualization is ideal for visualizing important statistics that you want to track on your Power BI dashboard or report, such as total value, YTD sales, or year-over-year change. The multi-row card visualization displays one or more data points, with one data point for each row.

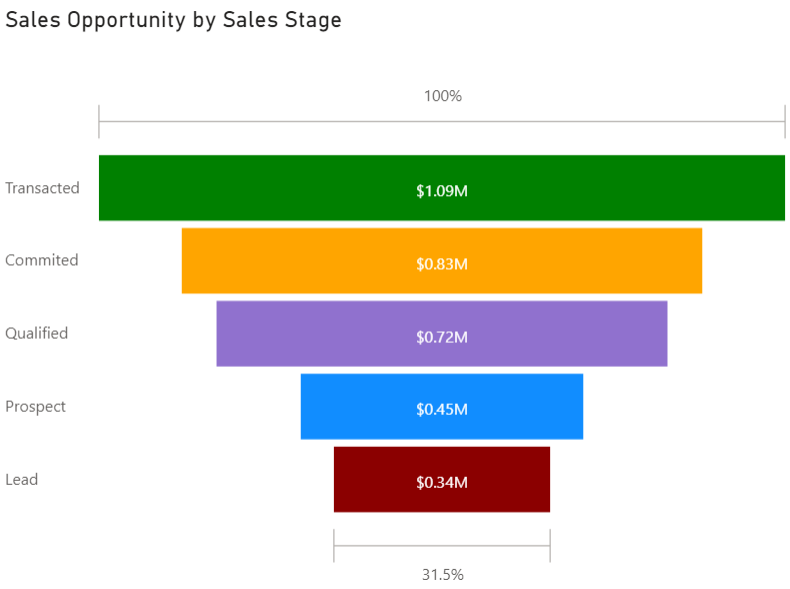


2.8 Funnel visualization

The funnel visualization displays a linear process that has sequential connected stages, where items flow sequentially from one stage to the next. Funnel charts are most often seen in business or sales contexts. For example, they are useful for representing a workflow, such as moving from a sales lead to a prospect, through to a proposal and sale.

Funnel charts are great options in the following contexts:

- When the data is sequential and moves through at least four stages.
- When the number of items in the first stage is expected to be greater than the number of items in the final stage.
- To calculate a potential outcome (revenue, sales, deals, and so on) by stages.
- To calculate and track conversion and retention rates.
- To reveal bottlenecks in a linear process.



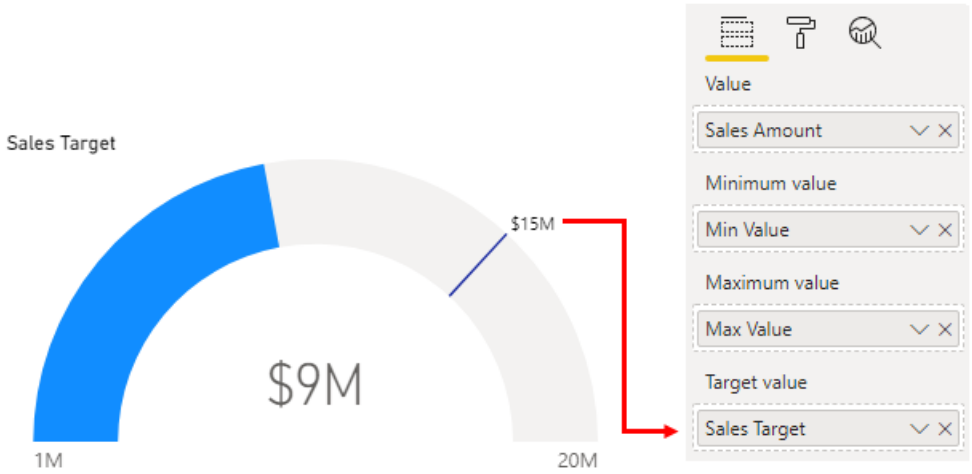
2.9 Gauge Chart

A radial gauge chart has a circular arc and displays a single value that measures progress toward a goal or target.

The value at the end of the arc represents the defaulted maximum value, which will always be double the actual value. To create a realistic visual, you should always specify each of the values. You can accomplish this task by dropping the correct field that contains an amount into the Target value, Minimum value, and Maximum value fields on the Visualization pane.

The shading in the arc represents the progress toward that target. 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).

Radial gauges can be used to show the progress that is being made toward a goal or target, or they can show the health of a single measure. However, radial gauges do take up a lot of space in comparison to the insights that they provide. It is more effective to use a pair of gauges with a spark line so users can see the trend and know what to do about it.

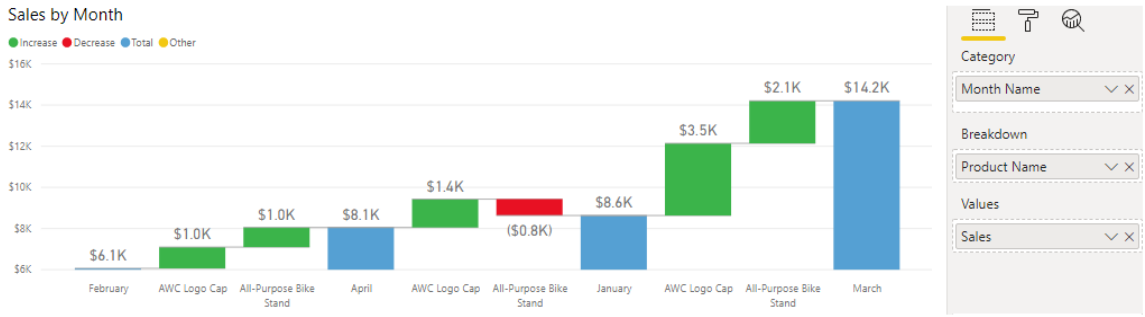


2.10 Waterfall Visualization

The waterfall visualization (also known as a bridge chart) shows a running total as values are added or subtracted, which is useful in displaying a series of positive and negative changes. The chart consists of color-coded columns, so you can quickly identify increases and decreases. The initial and the final value columns often start on the horizontal axis, while the intermediate values are floating columns.

Waterfall charts can be used to:

- Visualize changes over time or across different categories.
- Audit the major changes that contribute to the total value.
- Plot your organization's annual profit by showing various sources of revenue to help determine the total profit (or loss).
- Illustrate the beginning and ending headcount for your organization in a year.
- Visualize how much money you earn and spend each month and the running balance for your account.



2.11 Scatter Chart

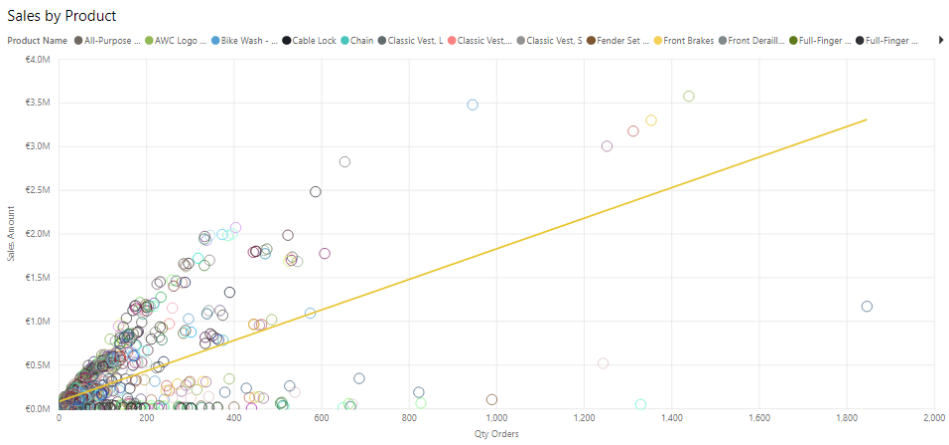
The scatter chart visualization is effective when you are comparing large numbers of data points without regard to time. The scatter chart has two value axes to show: one set of numerical data along a horizontal axis and another set of numerical values along a vertical axis. The chart displays points at the intersection of an X and Y numerical value, combining these values into single data points. These data points might be distributed evenly or unevenly across the horizontal axis, depending on the data. You can set the number of data points, up to a maximum of 10,000.

You might want to use a scatter chart instead of a line chart because it allows you to change the scale of the horizontal axis. Scatter charts also allow you to:

- Show relationships between two numerical values.
- Plot two groups of numbers as one series of x and y coordinates.
- Turn the horizontal axis into a logarithmic scale.
- Display worksheet data that includes pairs or grouped sets of values.
- Show patterns in large sets of data, for example, by showing linear or non-linear trends, clusters, and outliers.

Compare large numbers of data points without regard to time. The more data that you include in a scatter chart, the better the comparisons that you can make.

The following example shows a scatter chart that displays outliers (anomalies) with a trendline going up. The chart clearly shows that most products were sold at the same quantity, and only some products were sold in larger quantities. By identifying those outliers, you can run further analysis and break them down by country and region, which can help to improve logistics, decrease costs, and increase customer satisfaction.



2.12 Maps

A basic map (bubble map) is used to associate categorical and quantitative information with spatial locations. This type of map visual displays precise geographical locations of data points on a map, as illustrated in the following image. A fill map uses shading, tinting, or patterns to display how a value differs in proportion across a geographical region. Similarly, shape maps use colors to display relative comparisons of geographical regions. You can also use an ArcGIS map to display graphical information in a more interactive way.



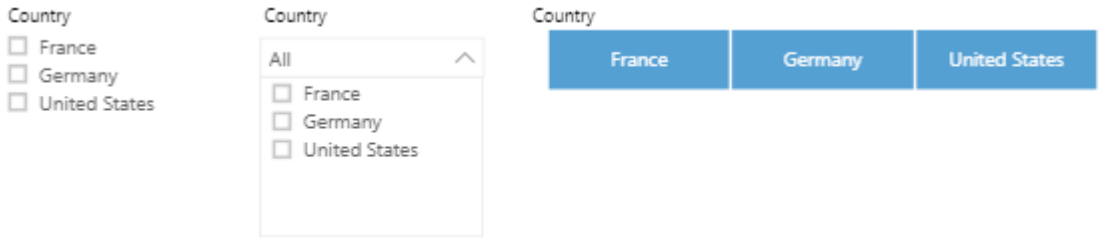
2.13 Slicer Visualization

The slicer visualization is a standalone chart that can be used to filter the other visuals on the page. Slicers provide a more advanced and customized way of filtering, in comparison to the Filters pane, which is suited to more basic filtering operations. You can learn more about these two filtering options in another module.

Slicers come in many different formats, including list, drop-down, and buttons, and they can be formatted to allow the selection of only one, many, or all available values.

Slicers are ideal to:

- Visualize commonly used or important filters on the report canvas for easier access.
- Simplify your ability 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.

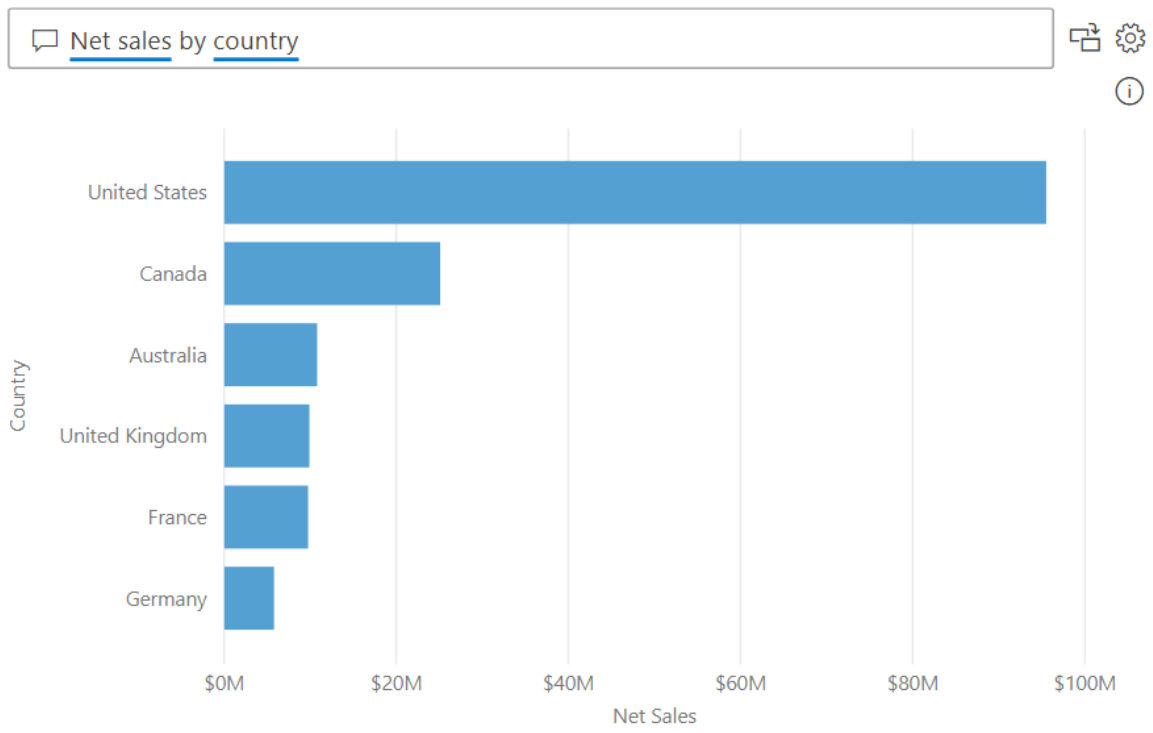


2.14 Q&A Visualization

The Q&A visualization allows you to ask natural language questions and get answers in the form of a visual. This ability to ask questions is valuable to consumers and to you, the report author. This visualization type can help you create visuals in the report, and it can also be used as a tool for consumers to get answers quickly.

The Q&A visualization consists of the following four core components:

- The question box, where users enter their question and are shown suggestions to help them complete the question.
- A pre-populated list of suggested questions.
- An icon that users can select to convert the Q&A visual into a standard visual.
- An icon that users can select to open Q&A tooling, which allows designers to configure the underlying natural language engine. When entering natural language queries with Power BI Q&A, you can specify the visual type in your query. The following example illustrates how to implement Net sales by country.

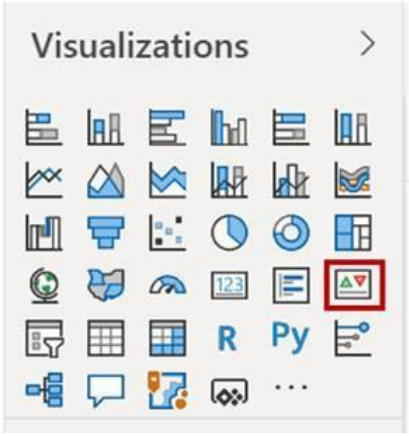


2.15 Key Performance Indicators

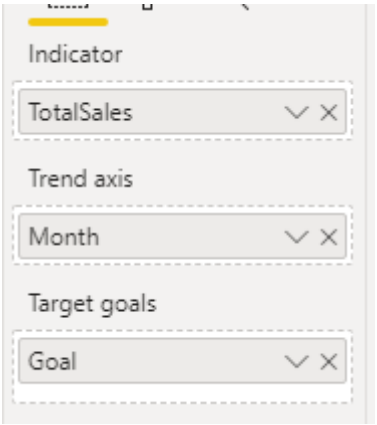
Key Performance Indicators (KPI) are excellent in helping you track progress toward a specific goal over time. To use a KPI, you need three pieces of information:

- A unit of measurement that you want to track, for instance total sales, number of employee hires, number of loans serviced, or number of students enrolled.
- A goal for the measurement so that you can compare your progress with that goal.
- A time series, for instance daily, monthly, or yearly.

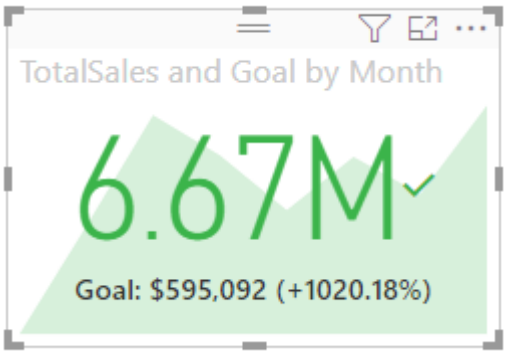
Start by adding the KPI visual to the design service. The following screenshot shows the KPI icon in the Visualizations pane.



When configuring the KPI visual, enter the unit of measurement that you are tracking in the **Indicator** prompt. Then, enter the goal under **Target goals** and select the time series from **the Trend axis** drop-down list.



This action will produce a KPI that looks similar to the following screenshot. KPIs work best in a series, for instance, showing the daily, monthly, and yearly goals in the section of a Power BI report.

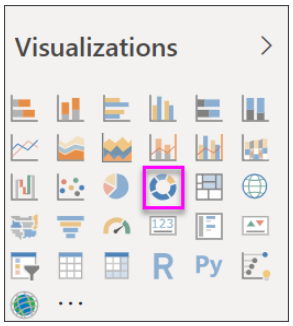


3. Working with Power BI Visual

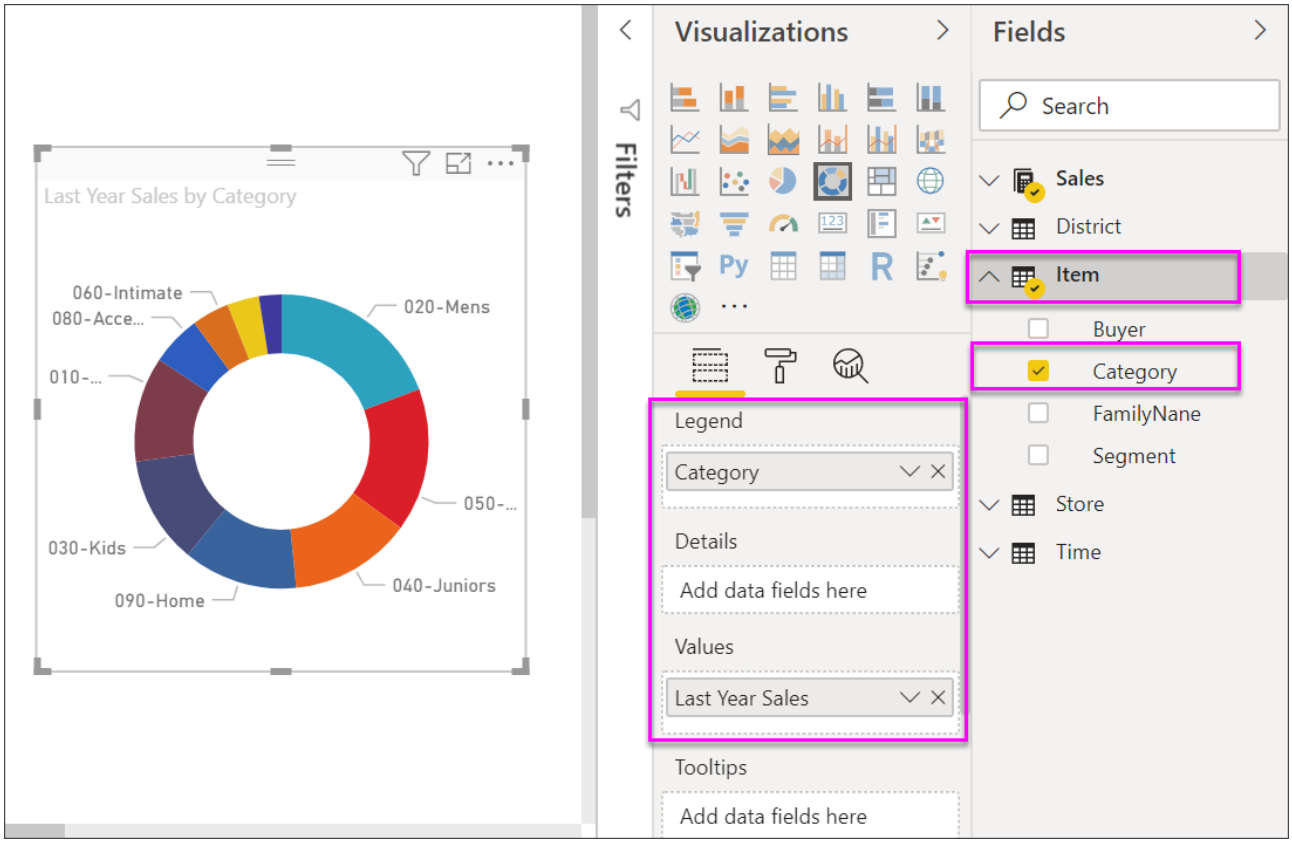
3.1 Doughnut Chart

Doughnut charts are best used to compare a particular section to the whole, rather than comparing individual sections with each other. The sum of the doughnut chart values must add up to 100%. However, too many categories make it difficult to read and interpret.

1. Select **File ➔ Open Report**, and open *Retail Analysis Sample.pbix*.
2. Start on a blank report page and from the **Fields** pane, select *Sales ➔ Last Year Sales*.
3. From the Visualizations pane, select the icon for doughnut chart to convert your bar chart to a doughnut chart. If Last Year Sales is not in the Values area, drag it there.



4. Select *Item ➔ Category* to add it to the Legend area.



3.2 Table

A table is a grid that contains related data in a logical series of rows and columns. It may also contain headers and a row for totals. Tables work well with quantitative comparisons where you're looking at many values for a single category.

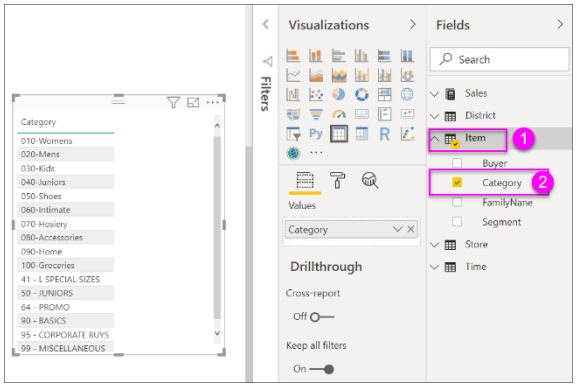
Create tables in reports and cross-highlight elements within the table with other visuals on the same report page. You can select rows, columns, and even individual cells and cross-highlight. You can also copy and paste individual cells and multiple cell selections into other applications.

Tables are a great choice:

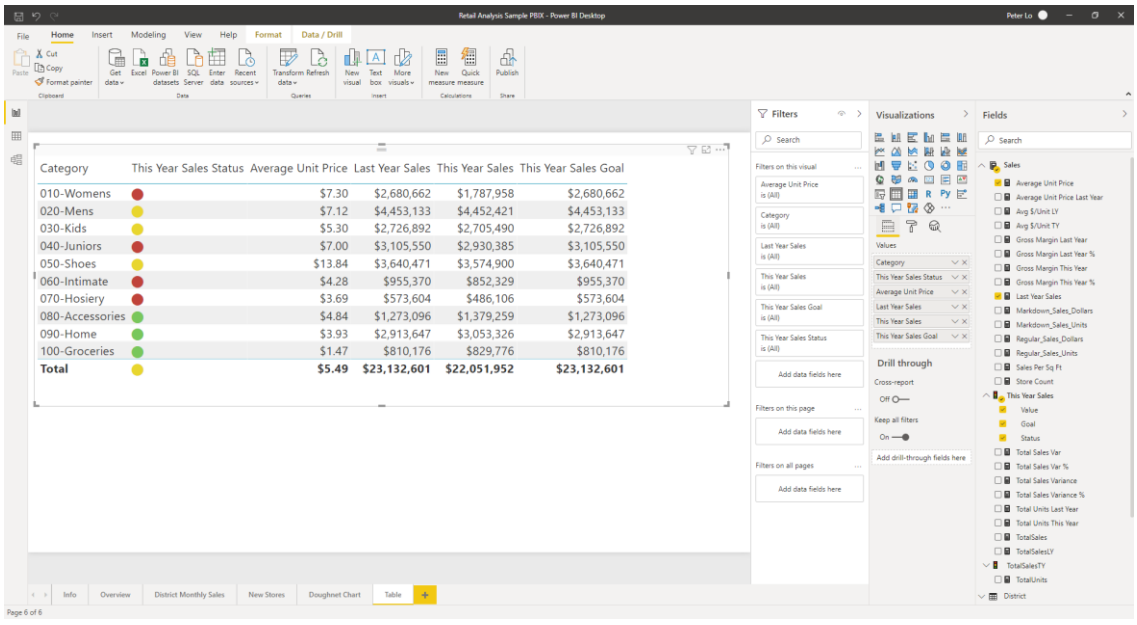
- To see and compare detailed data and exact values (instead of visual representations).
- To display data in a tabular format.
- To display numerical data by categories.

In order to create the table to display sales values by item category.

1. Select **File ➔ Open Report**, and open *Retail Analysis Sample.pbix*.
2. Start on a blank report page and from the **Fields** pane, select *Item ➔ Category*. Power BI automatically creates a table that lists all the categories.



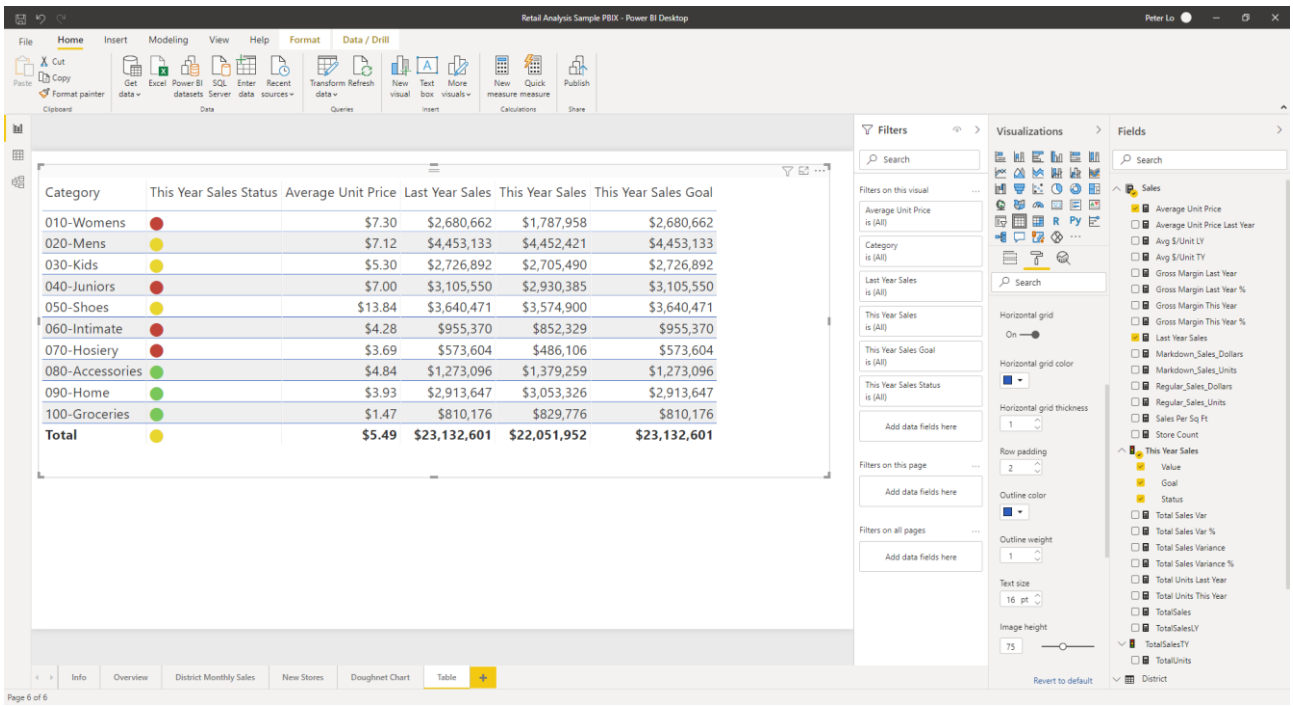
3. Select *Sales ➔ Average Unit Price* and *Sales ➔ Last Year Sales*.
4. Then select *Sales ➔ This Year Sales* and select all three options: *Value*, *Goal*, and *Status*.
5. In the **Visualizations** pane, locate the **Values** well and select the values until the order of your chart columns matches the first image on this page. Drag the values in the well if needed.



Category	This Year Sales Status	Average Unit Price	Last Year Sales	This Year Sales	This Year Sales Goal
010-Womens		$7.30	$2,680,662	$1,787,958	$2,680,662
020-Mens		$7.12	$4,453,133	$4,452,421	$4,453,133
030-Kids		$5.30	$2,726,892	$2,705,490	$2,726,892
040-Juniors		$7.00	$3,105,550	$2,930,385	$3,105,550
050-Shoes		$13.84	$3,640,471	$3,574,900	$3,640,471
060-Intimate		$4.28	$955,370	$852,329	$955,370
070-Hosiery		$3.69	$573,604	$486,106	$573,604
080-Accessories		$4.84	$1,273,096	$1,379,259	$1,273,096
090-Home		$3.93	$2,913,647	$3,053,326	$2,913,647
100-Groceries		$1.47	$810,176	$829,776	$810,176
Total		**$5.49**	**$23,132,601**	**$22,051,952**	**$23,132,601**

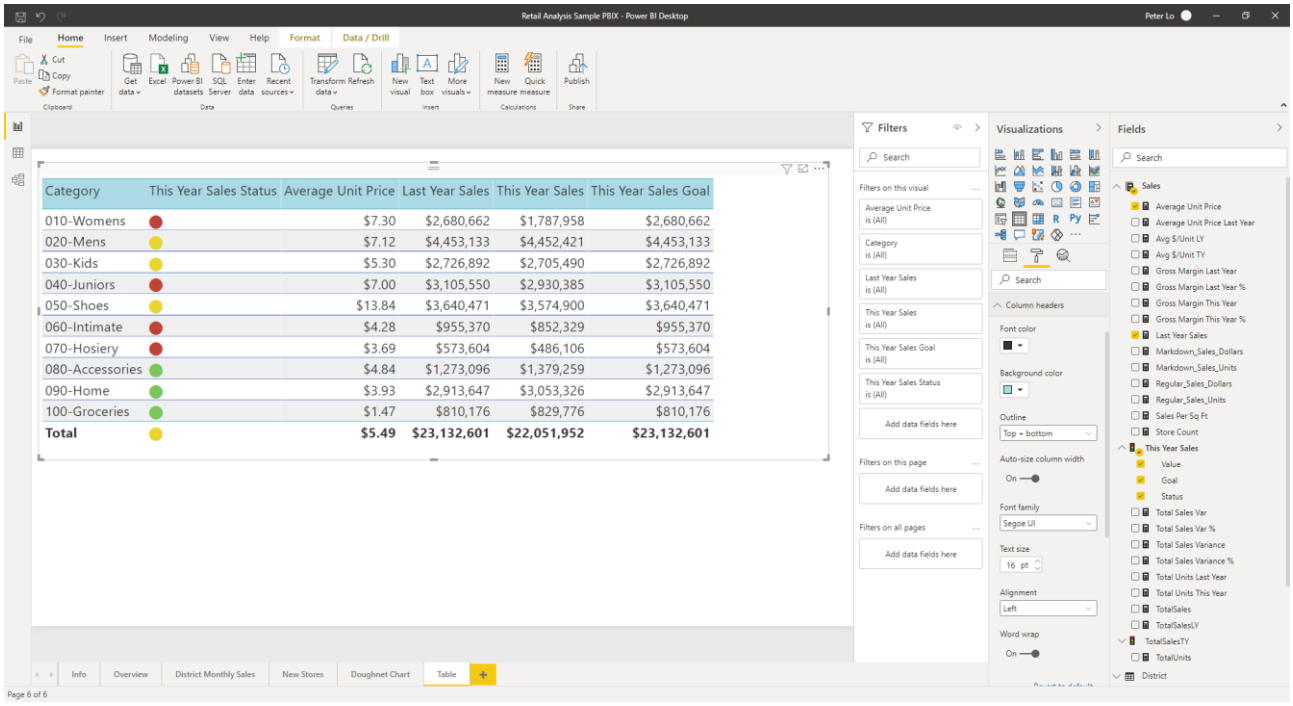
3.2.1 Format the Table

1. Try formatting the table grid. Here you'll add a blue vertical grid, add space to the rows, and increase the outline and text size.



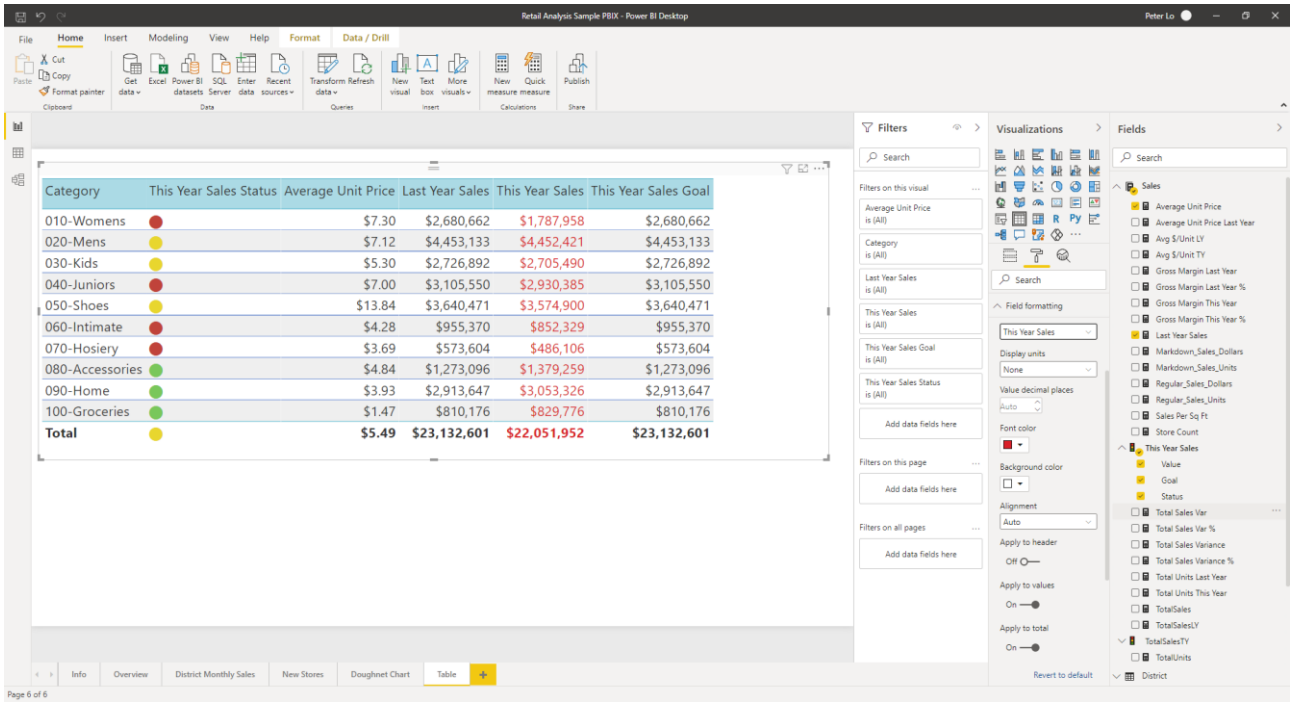
Category	This Year Sales Status	Average Unit Price	Last Year Sales	This Year Sales	This Year Sales Goal
010-Womens		$7.30	$2,680,662	$1,787,958	$2,680,662
020-Mens		$7.12	$4,453,133	$4,452,421	$4,453,133
030-Kids		$5.30	$2,726,892	$2,705,490	$2,726,892
040-Juniors		$7.00	$3,105,550	$2,930,385	$3,105,550
050-Shoes		$13.84	$3,640,471	$3,574,900	$3,640,471
060-Intimate		$4.28	$955,370	$852,329	$955,370
070-Hosiery		$3.69	$573,604	$486,106	$573,604
080-Accessories		$4.84	$1,273,096	$1,379,259	$1,273,096
090-Home		$3.93	$2,913,647	$3,053,326	$2,913,647
100-Groceries		$1.47	$810,176	$829,776	$810,176
Total		**$5.49**	**$23,132,601**	**$22,051,952**	**$23,132,601**

2. For the column headers, change the background color, add an outline, and increase the font size.



Category	This Year Sales Status	Average Unit Price	Last Year Sales	This Year Sales	This Year Sales Goal
010-Womens		$7.30	$2,680,662	$1,787,958	$2,680,662
020-Mens		$7.12	$4,453,133	$4,452,421	$4,453,133
030-Kids		$5.30	$2,726,892	$2,705,490	$2,726,892
040-Juniors		$7.00	$3,105,550	$2,930,385	$3,105,550
050-Shoes		$13.84	$3,640,471	$3,574,900	$3,640,471
060-Intimate		$4.28	$955,370	$852,329	$955,370
070-Hosiery		$3.69	$573,604	$486,106	$573,604
080-Accessories		$4.84	$1,273,096	$1,379,259	$1,273,096
090-Home		$3.93	$2,913,647	$3,053,326	$2,913,647
100-Groceries		$1.47	$810,176	$829,776	$810,176
Total		**$5.49**	**$23,132,601**	**$22,051,952**	**$23,132,601**

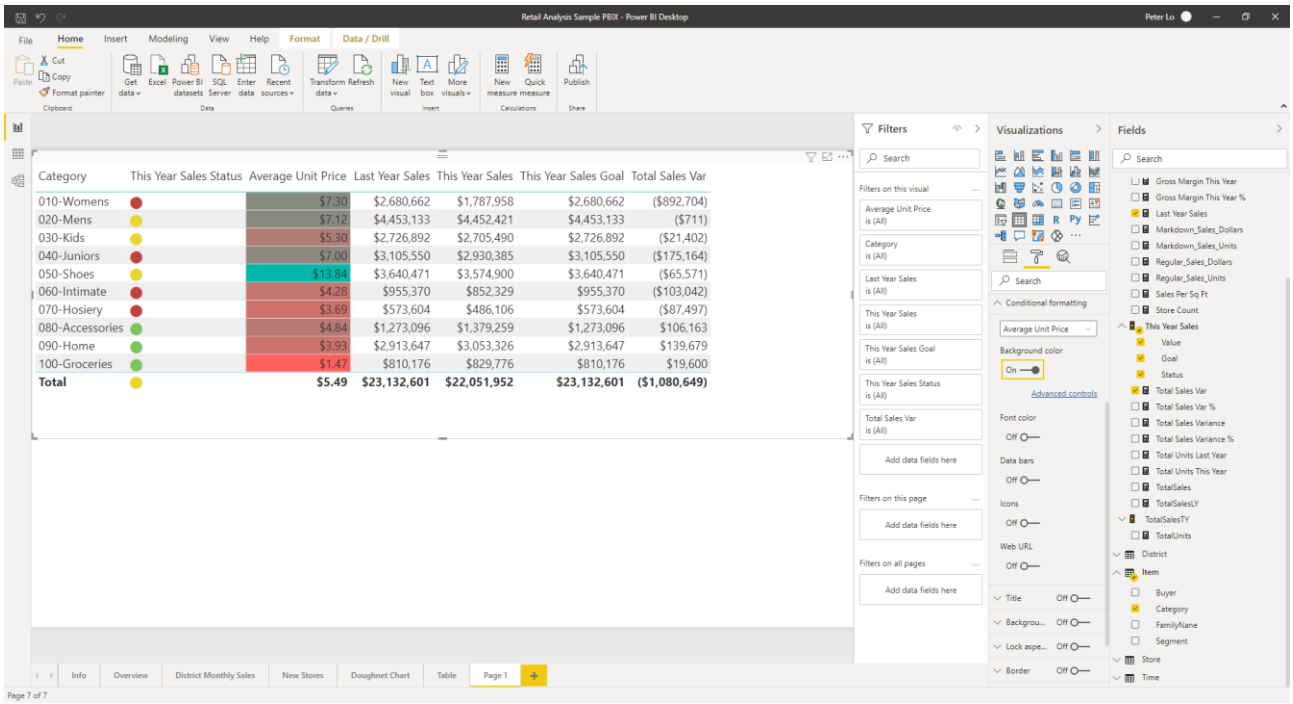
3. You can even apply formatting to individual columns and column headers. Start by expanding **Field formatting** and selecting the column to format from the drop-down. Depending on the column values, Field formatting lets you set things like: display units, font color, number of decimal places, background, alignment, and more. Once you've adjusted the settings, decide whether to apply those settings to the header and totals row as well.



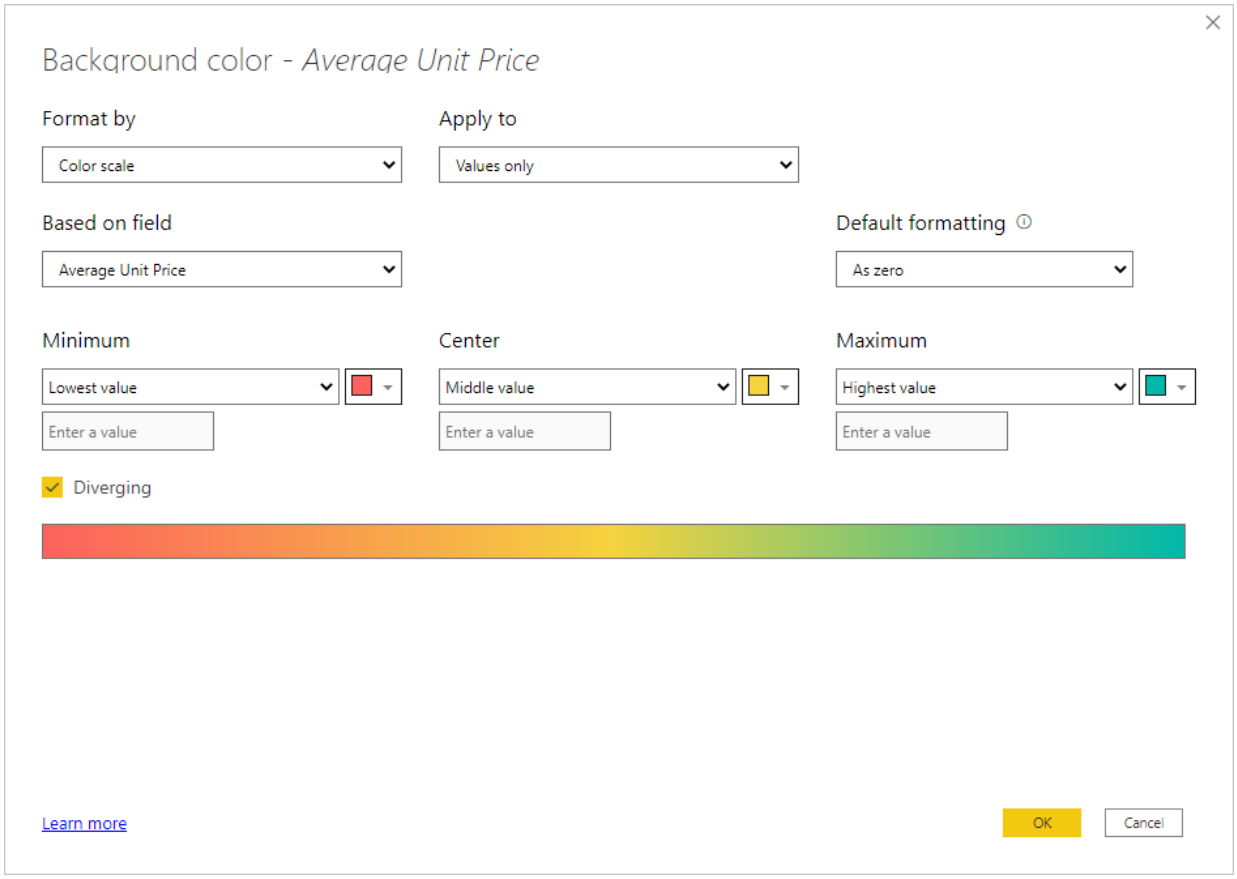
3.2.2 Conditional Formatting

Power BI can apply conditional formatting to any of the fields that you added to the **Values** well of the **Visualizations** pane. With conditional formatting for tables, you can specify icons, URLs, cell background colors, and font colors based on cell values, including using gradient colors.

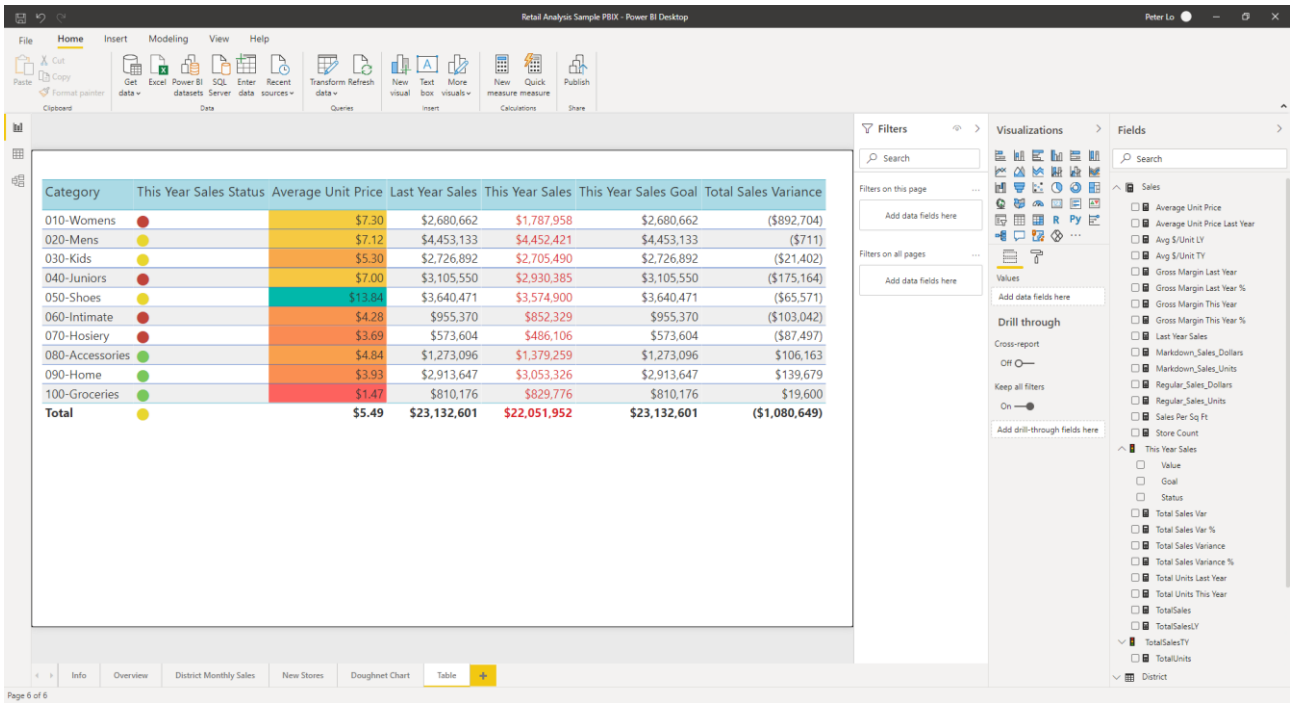
1. In the **Format** pane, open the **Conditional formatting** card.
2. Select the field *Average Unit Price* to format, and turn the slider for **Background color** to On. Power BI applies a gradient based on the values in the column. To change the default colors, select [**Advanced controls**].



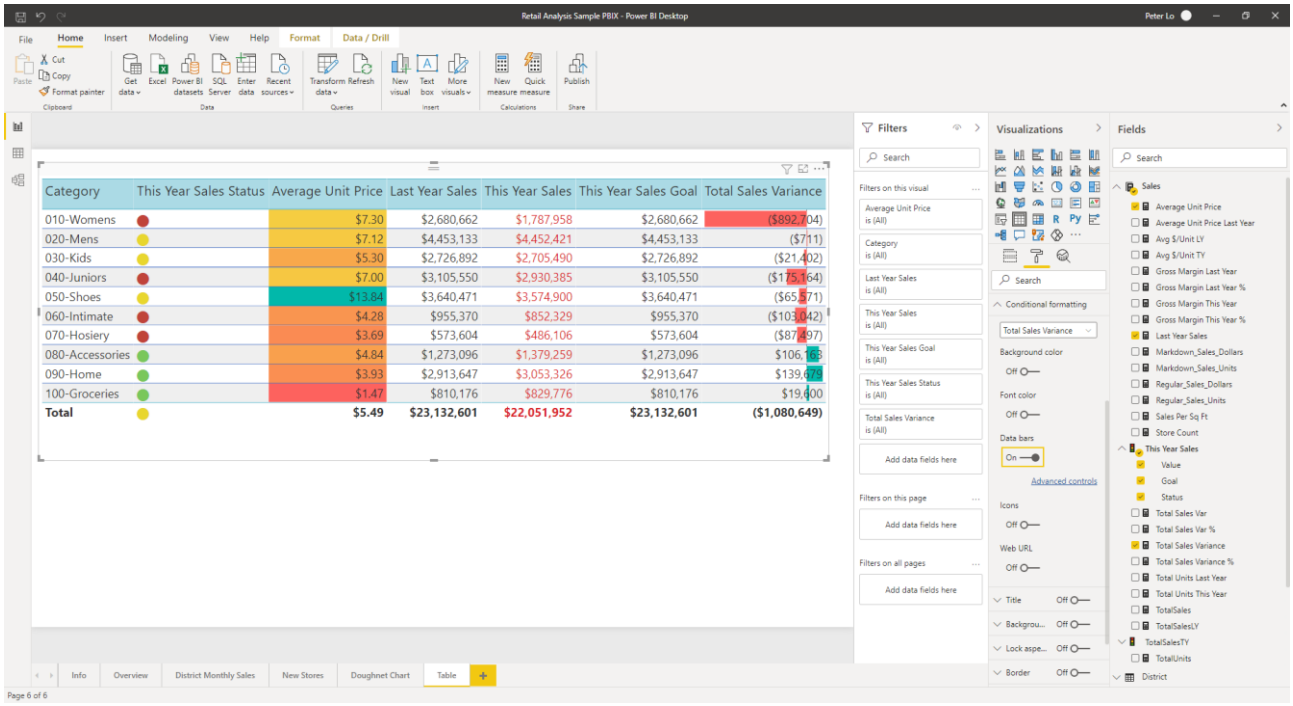
3. If you select the **Diverging** option, you can configure an optional **Center** value as well. Select **[OK]** after add some colors.



4. Add a new field to the table that has both positive and negative values. Select *Sales* ➔ *Total Sales Variance*.

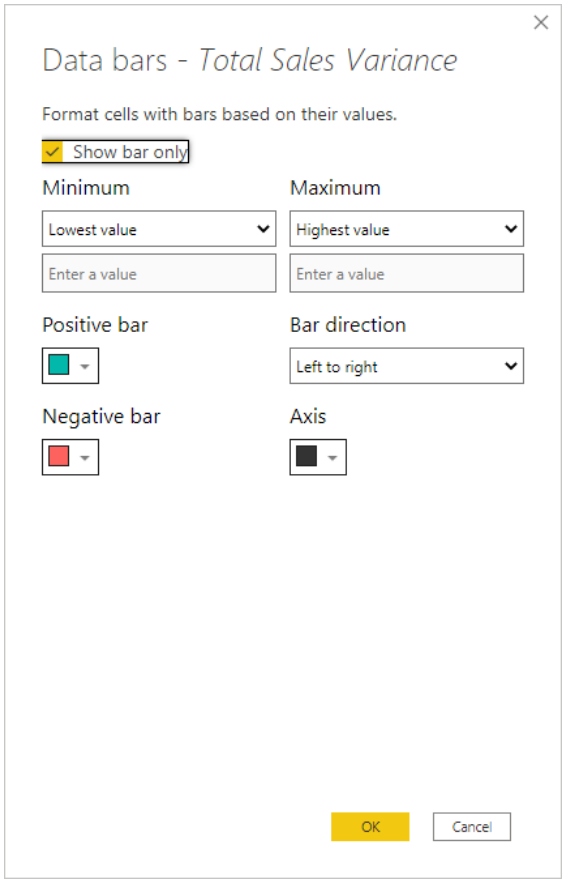


5. Add data bar conditional formatting by turning the **Data bars** slider to **On**.

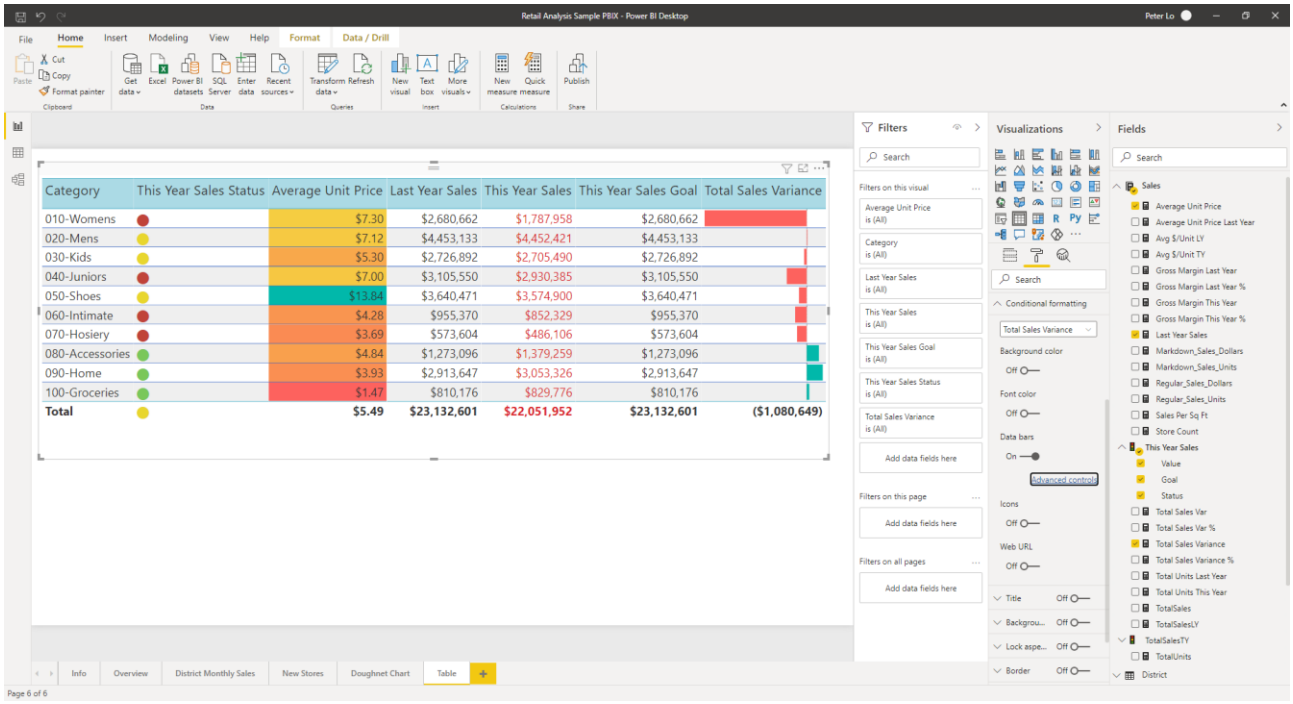


Category	This Year Sales Status	Average Unit Price	Last Year Sales	This Year Sales	This Year Sales Goal	Total Sales Variance
010-Womens		$7.30	$2,680,662	$1,787,958	$2,680,662	($892,704)
020-Mens		$7.12	$4,453,133	$4,452,421	$4,453,133	($711)
030-Kids		$5.30	$2,726,892	$2,705,490	$2,726,892	($21,402)
040-Juniors		$7.00	$3,105,550	$2,930,385	$3,105,550	($175,164)
050-Shoes		$13.84	$3,640,471	$3,574,900	$3,640,471	($65,571)
060-Intimate		$4.28	$955,370	$852,329	$955,370	($103,042)
070-Hosiery		$3.69	$573,604	$486,106	$573,604	($87,497)
080-Accessories		$4.84	$1,273,096	$1,379,259	$1,273,096	$106,163
090-Home		$3.93	$2,913,647	$3,053,326	$2,913,647	$139,679
100-Groceries		$1.47	$810,176	$829,776	$810,176	$19,600
Total		**$5.49**	**$23,132,601**	**$22,051,952**	**$23,132,601**	**($1,080,649)**

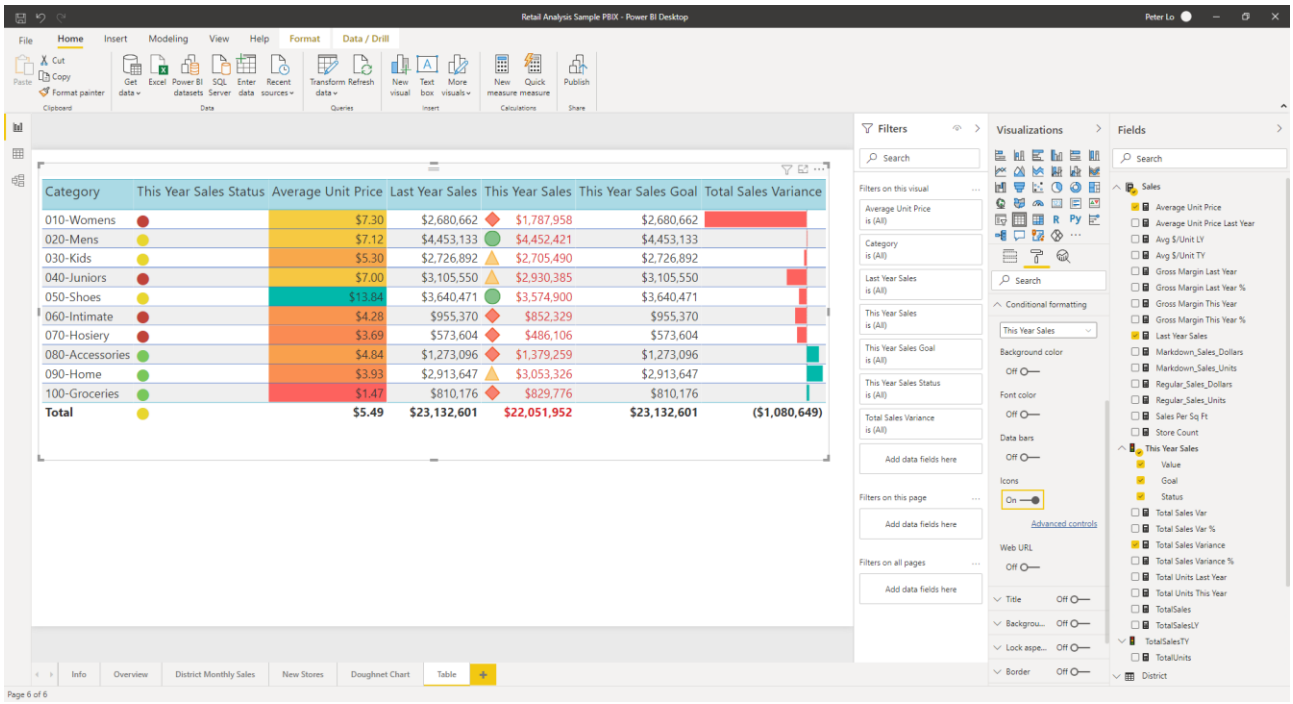
6. To customize the data bars, select **Advanced controls**. In the dialog that appears, set colors for **Positive bar** and **Negative bar**, select the **Show bar only** option, and make any other changes you'd like.



7. Once you select **[OK]**, **Data bars** replace the numerical values in the table, making it easier to scan.



8. Add visual cues to your table with conditional icons. In the **Conditional formatting** card, select *This year sales* from the dropdown. Turn the **Icons** slider to **On**. To customize the icons, select **Advanced controls**.

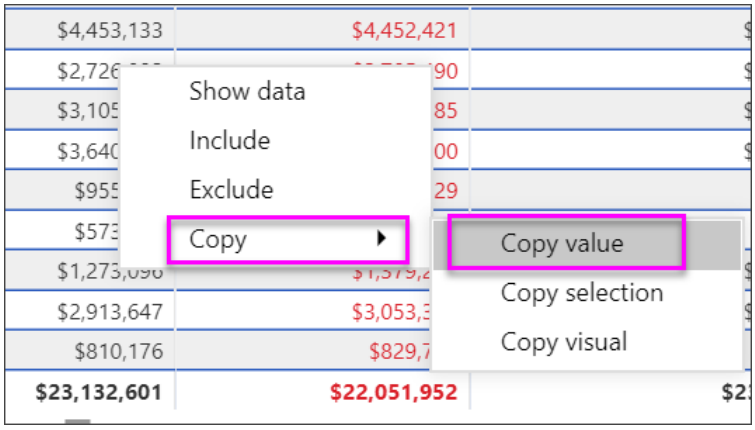


3.2.3 Copy Values from Power BI Tables to other Applications

Your table or matrix may have content that you'd like to use in other applications, like Excel. When you right-click inside a cell, you can copy the data in a single cell or a selection of cells onto your clipboard, and paste it into the other applications.

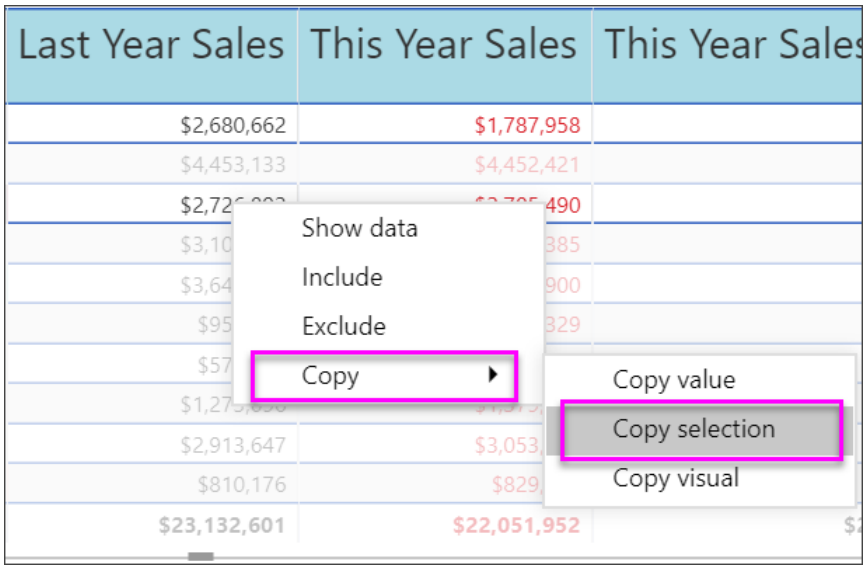
To copy the value of a single cell:

1. Select the cell you want to copy.
2. Right-click inside the cell.
3. Select **Copy ➔ Copy value**.
4. With the unformatted cell value on your clipboard, you can paste it into another application.



To copy more than a single cell:

1. Select a range of cells or use **[Ctrl]** key to select multiple cells.
2. Right-click inside one of the cells you selected.
3. Select **Copy ➔ Copy selection**.



3.3 KPI Visual

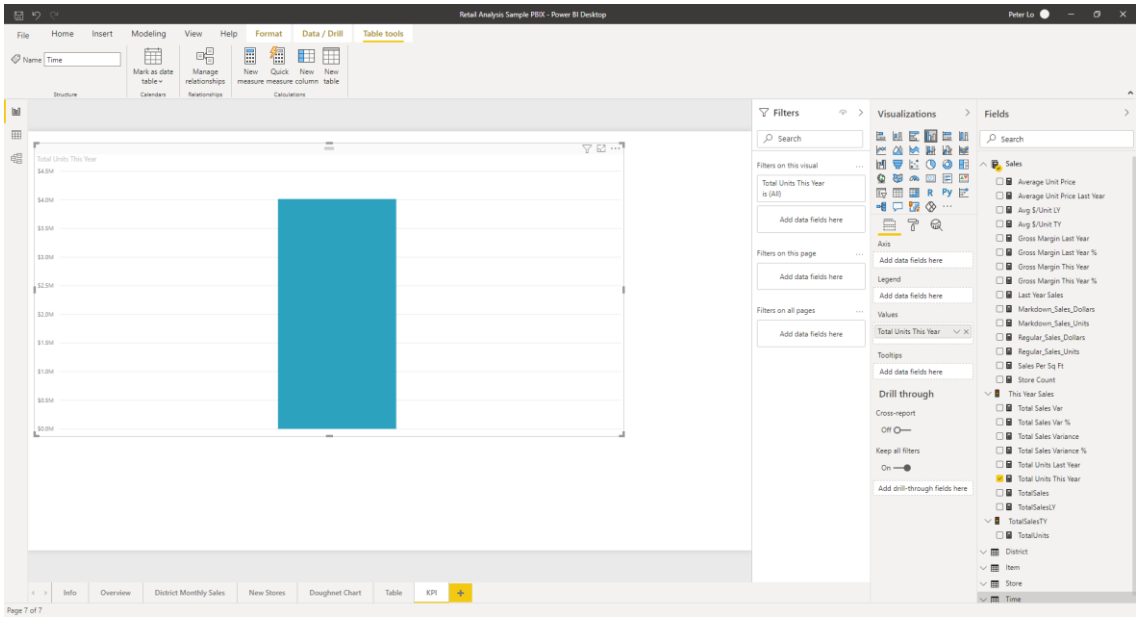
A Key Performance Indicator (KPI) is a visual cue that communicates the amount of progress made toward a measurable goal.

KPIs are a great choice:

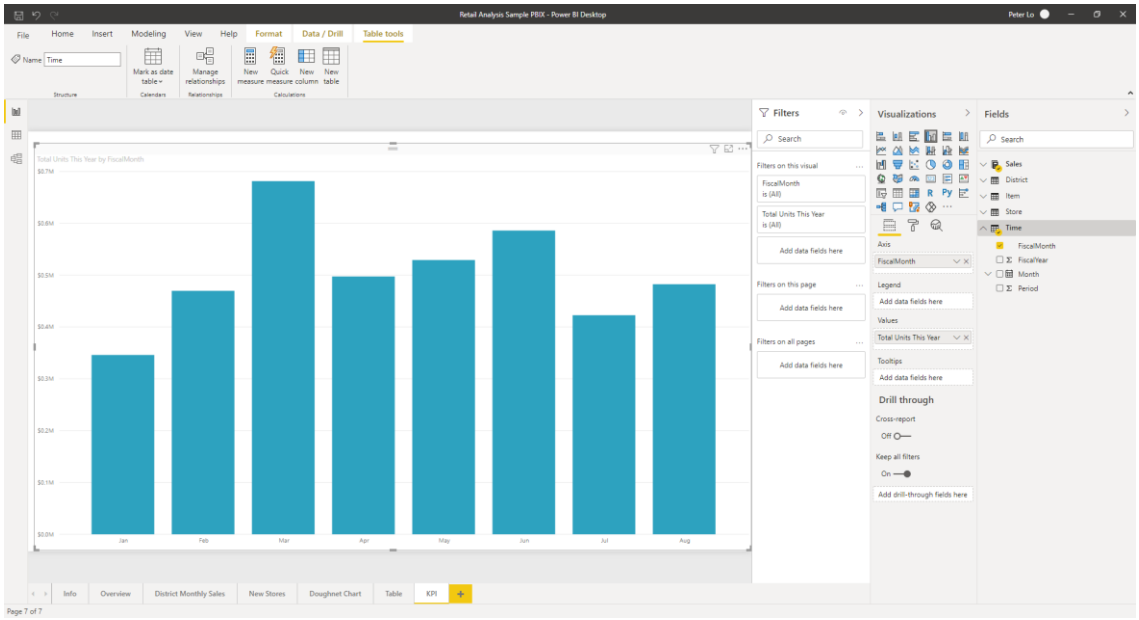
- To measure progress. Answers the question, "What am I ahead or behind on?"
- To measure distance to a goal. Answers the question, "How far ahead or behind am I?"

A designer bases a KPI visual on a specific measure. The intention of the KPI is to help you evaluate the current value and status of a metric against a defined target. A KPI visual requires a base measure that evaluates to a value, a target measure or value, and a threshold or goal.

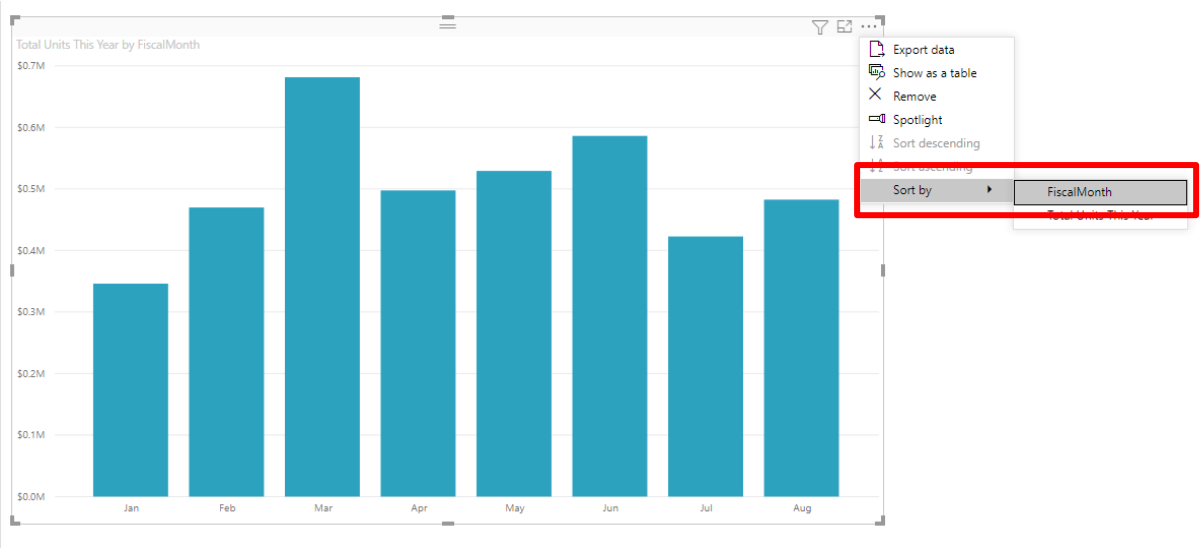
1. Select **File ➔ Open Report**, and open *Retail Analysis Sample.pbix*.
2. Start on a blank report page and from the **Fields** pane, select *Sales ➔ Total Units This Year*. This value will be the indicator.



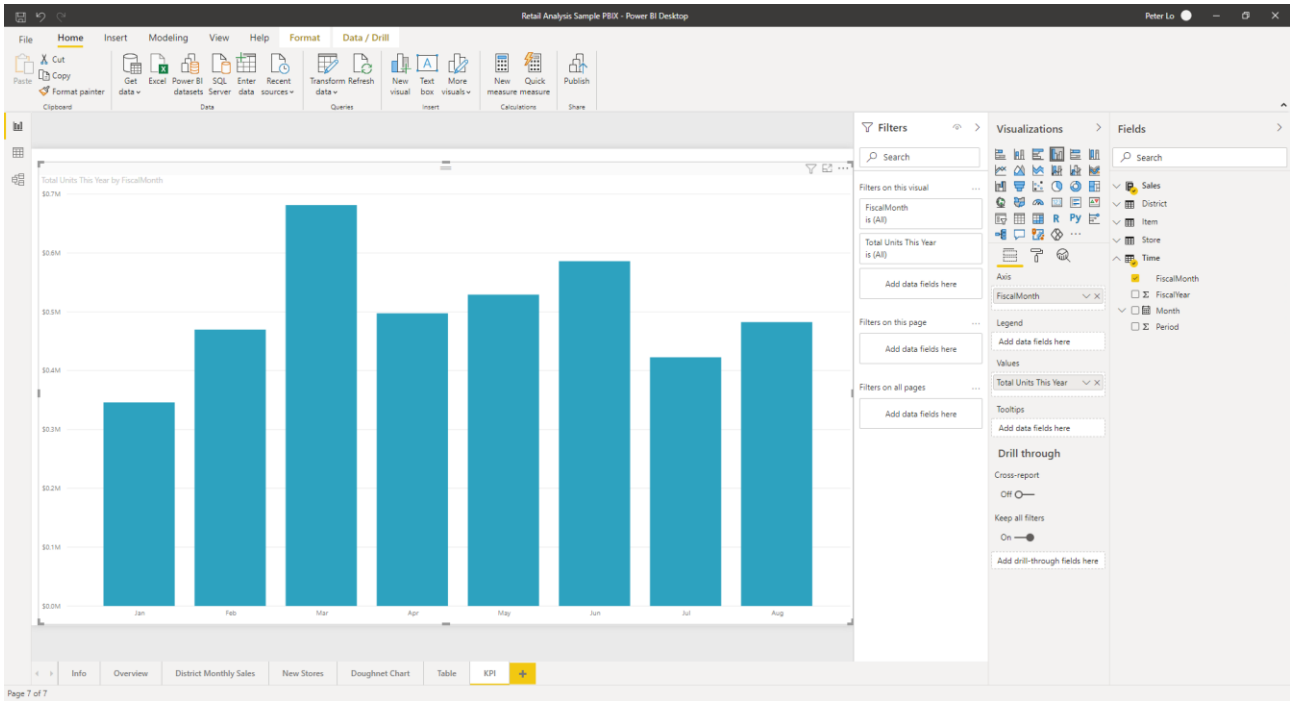
3. Add *Time ➔ FiscalMonth*. This value will represent the trend.



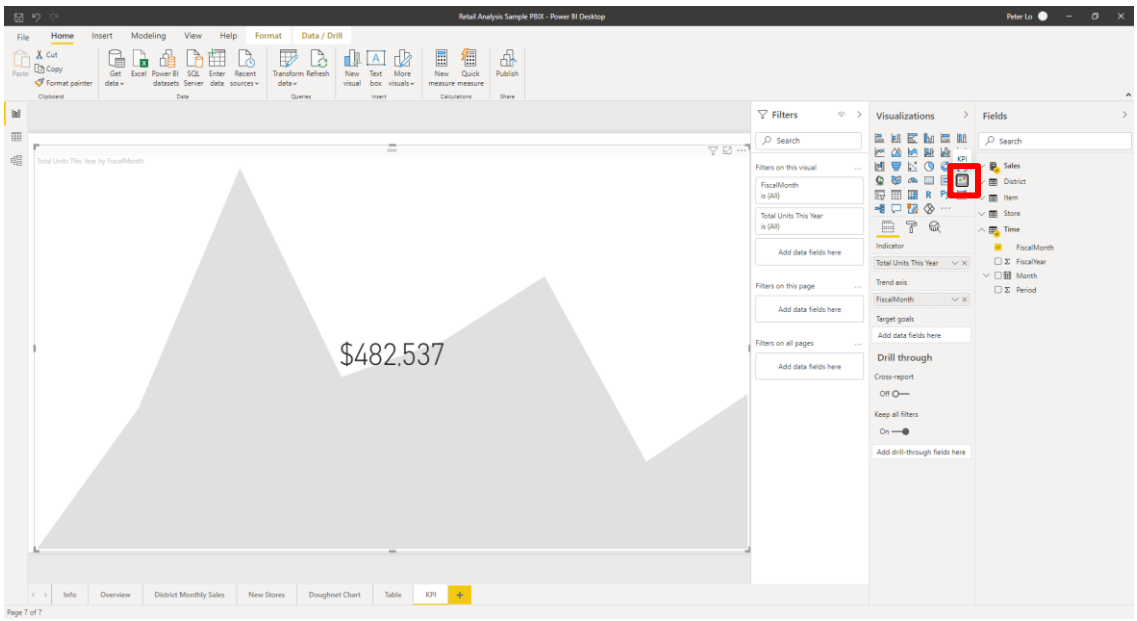
4. In the upper-right corner of the visual, select the ellipsis and check that Power BI sorted the columns in ascending order by *FiscalMonth*. Once you convert the visualization to a KPI, there's no option to sort.



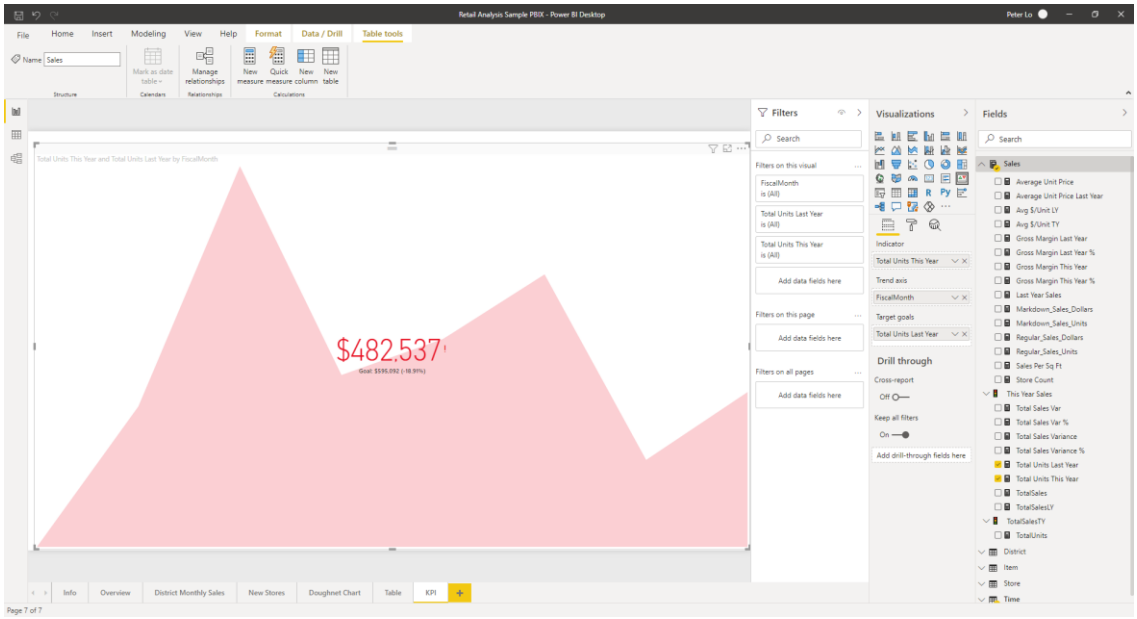
5. Once sorted correctly, your visual will look like this:



6. Convert the visual to a KPI by selecting the **KPI** icon from the **Visualization** pane.



7. To add a goal, drag *Sales* ➔ *Total Units Last Year* to the **Target goals** field.



8. Optionally, format the KPI by selecting the paint roller icon to open the Formatting pane.
   - **Indicator** – Controls the indicator's display units and decimal places.
   - **Trend axis** – When set to On, the visual shows the trend axis as the background of the KPI visual.
   - **Goals** – When set to On, the visual shows the goal and the distance from the goal as a percentage.
   - **Color coding ➔ Direction** – People consider some KPIs better for higher values and consider some better for lower values. For example, earnings versus wait time. Typically, a higher value of earnings is better versus a higher value of wait time. Select high is good and, optionally, change the color settings.

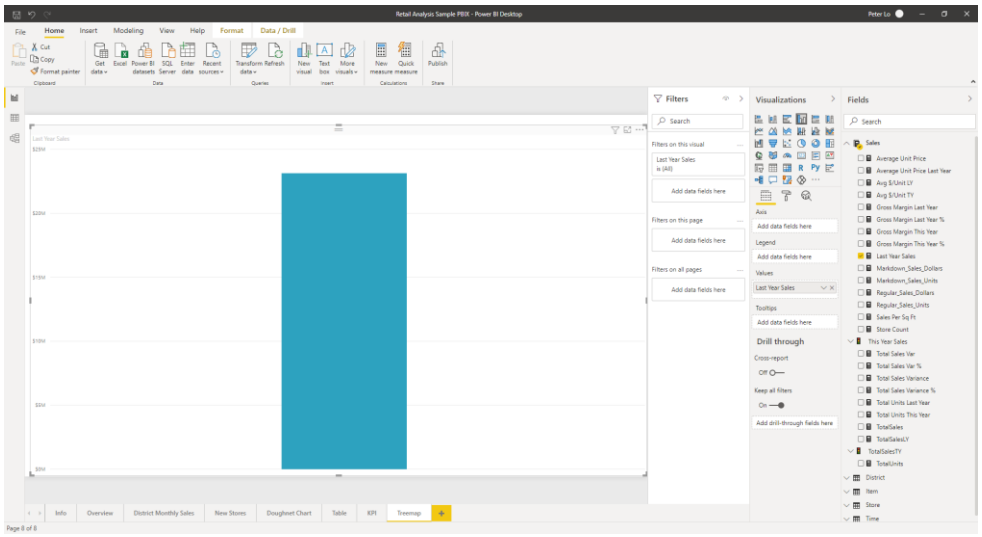
3.4 Treemap

Treemaps display hierarchical data as a set of nested rectangles. Each level of the hierarchy is represented by a colored rectangle (branch) containing smaller rectangles (leaves). Power BI bases the size of the space inside each rectangle on the measured value. The rectangles are arranged in size from top left (largest) to bottom right (smallest). Treemaps are a great choice:

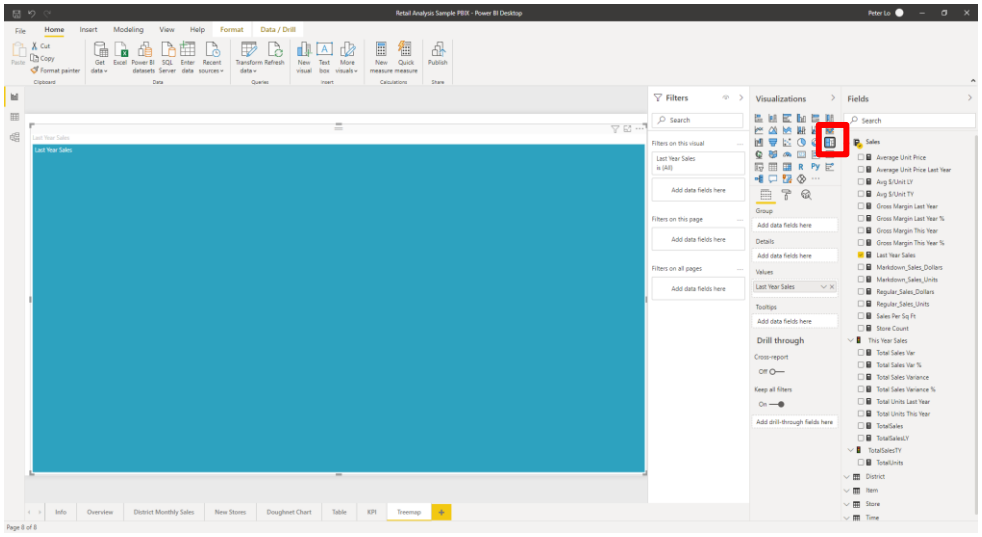
- To display large amounts of hierarchical data.
- When a bar chart can't effectively handle the large number of values.
- To show the proportions between each part and the whole.
- To show the pattern of the distribution of the measure across each level of categories in the hierarchy.
- To show attributes using size and color coding.
- To spot patterns, outliers, most-important contributors, and exceptions.

In order to create a Treemap,

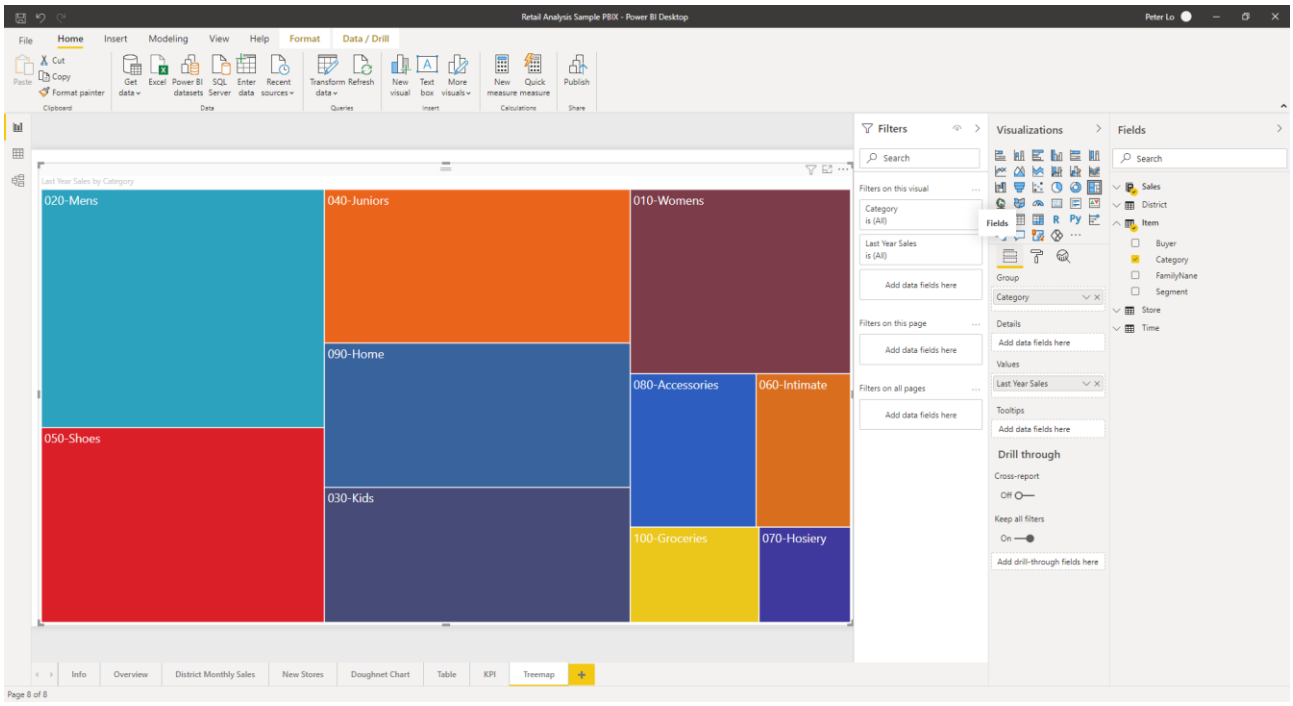
1. Select **File ➔ Open Report**, and open *Retail Analysis Sample.pbix*.
2. Start on a blank report page and from the **Fields** pane, select the *Sales ➔ Last Year Sales* measure.



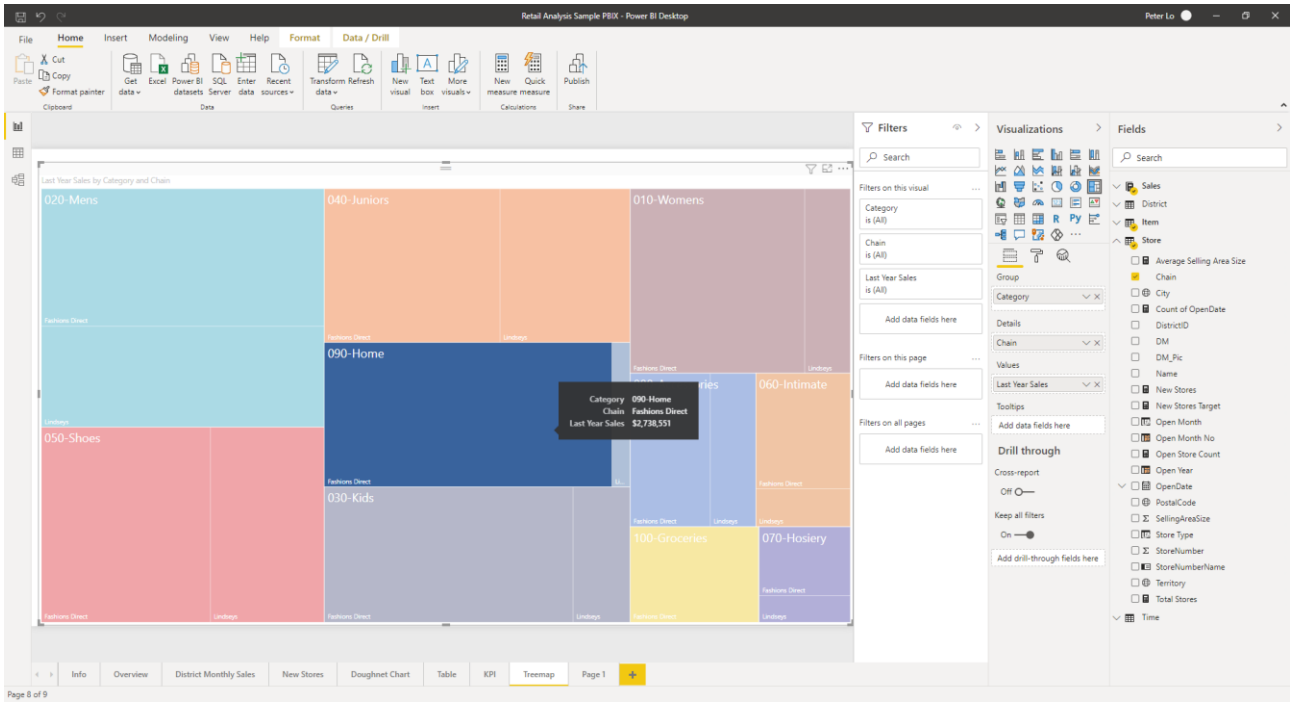
3. Select the treemap icon to convert the chart to a treemap.



4. Select *Item* ➔ *Category* which will add Category to the Group well. Power BI creates a treemap where the size of the rectangles is based on total sales and the color represents the category. In essence you've created a hierarchy that visually describes the relative size of total sales by category. The Men's category has the highest sales and the Hosiery category has the lowest.



5. Select *Store* ➔ *Chain* which will add Chain to the Details well to complete your treemap. You can now compare last year's sales by category and chain.
6. Hover over a Chain area to reveal the tooltip for that portion of the Category. For example, hovering over Fashions Direct in the 090-Home rectangle reveals the tooltip for Fashion Direct's portion of the Home category.



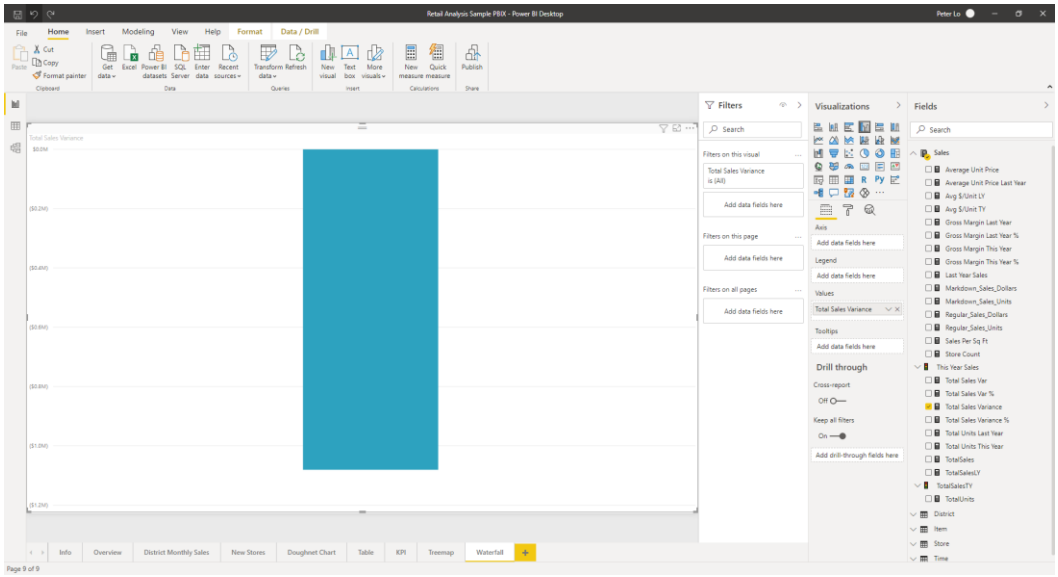
3.5 Waterfall Chart

Waterfall charts show a running total as Power BI adds and subtracts values. They're useful for understanding how an initial value (like net income) is affected by a series of positive and negative changes. Waterfall charts are a great choice:

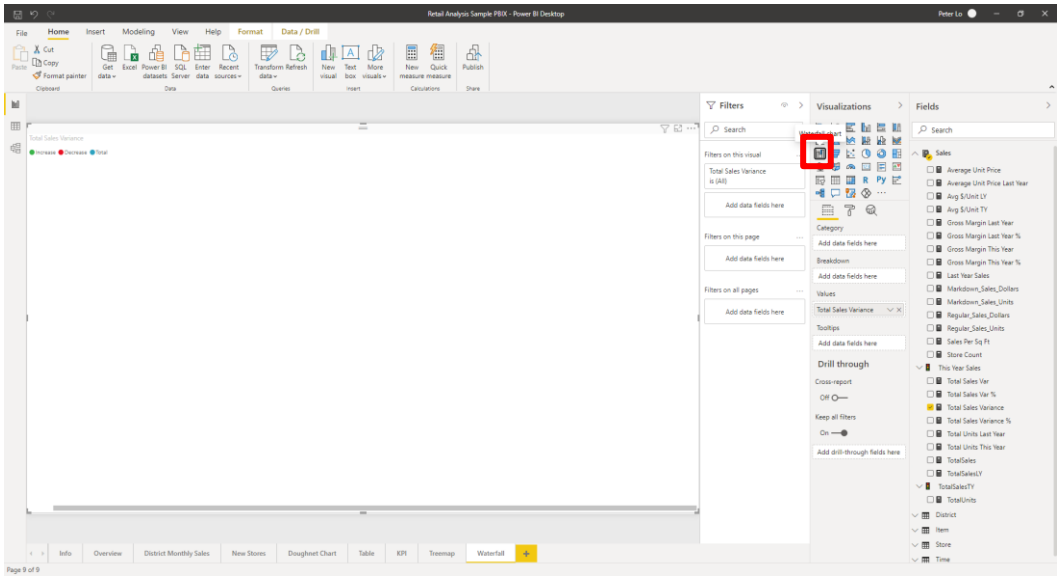
- When you have changes for the measure across time, a series, or different categories.
- To audit the major changes contributing to the total value.
- To plot your company's annual profit by showing various sources of revenue and arrive at the total profit (or loss).
- To illustrate the beginning and the ending headcount for your company in a year.
- To visualize how much money you make and spend each month, and the running balance for your account.

In order to create a waterfall chart that displays sales variance (estimated sales versus actual sales) by month:

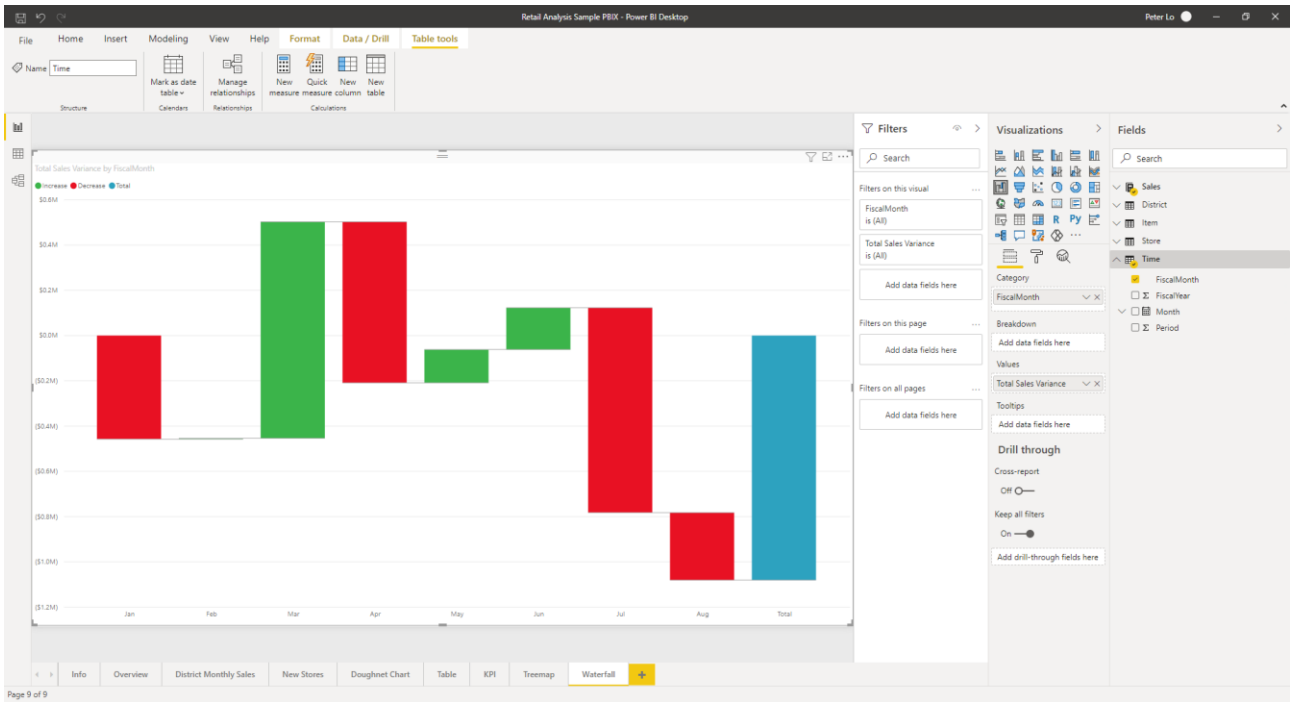
1. Select **File ➔ Open Report**, and open *Retail Analysis Sample.pbix*.
2. Start on a blank report page and from the **Fields** pane, select the *Sales ➔ Total Sales Variance*.



3. Select the waterfall icon

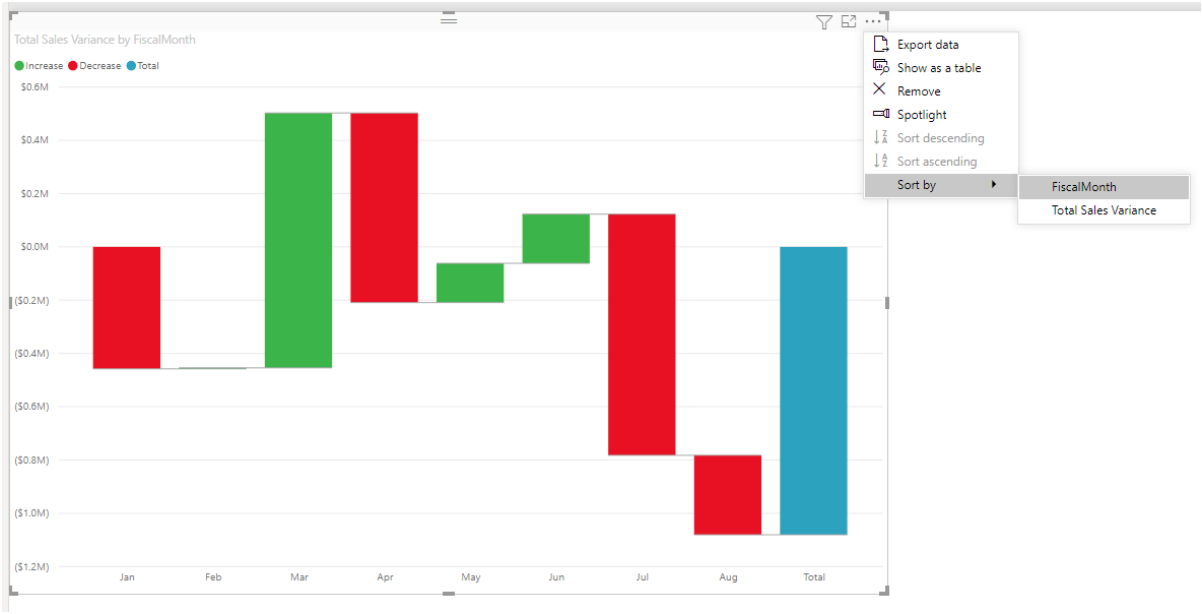


4. Select *Time* ➔ *FiscalMonth* to add it to the **Category** well.



3.5.1 Sort the Waterfall Chart

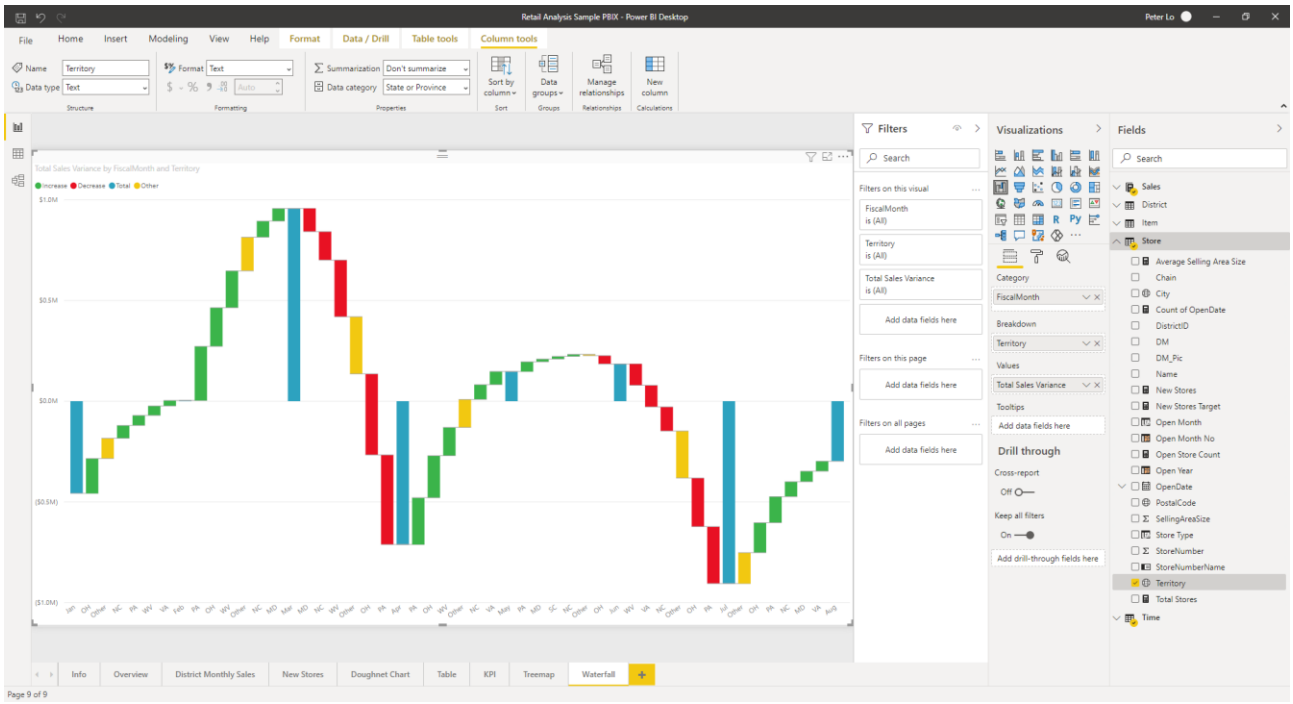
1. Make sure Power BI sorts the waterfall chart chronologically by month. From the top-right corner of the chart, select **More options (...)**. Select **Sort by** and choose **FiscalMonth**. A yellow indicator next to your selection indicates when your selection option is being applied.
2. To display the months in chronological order, select **Sort ascending**. A yellow indicator next to the left of Sort ascending, this indicates that your selected option is being applied.



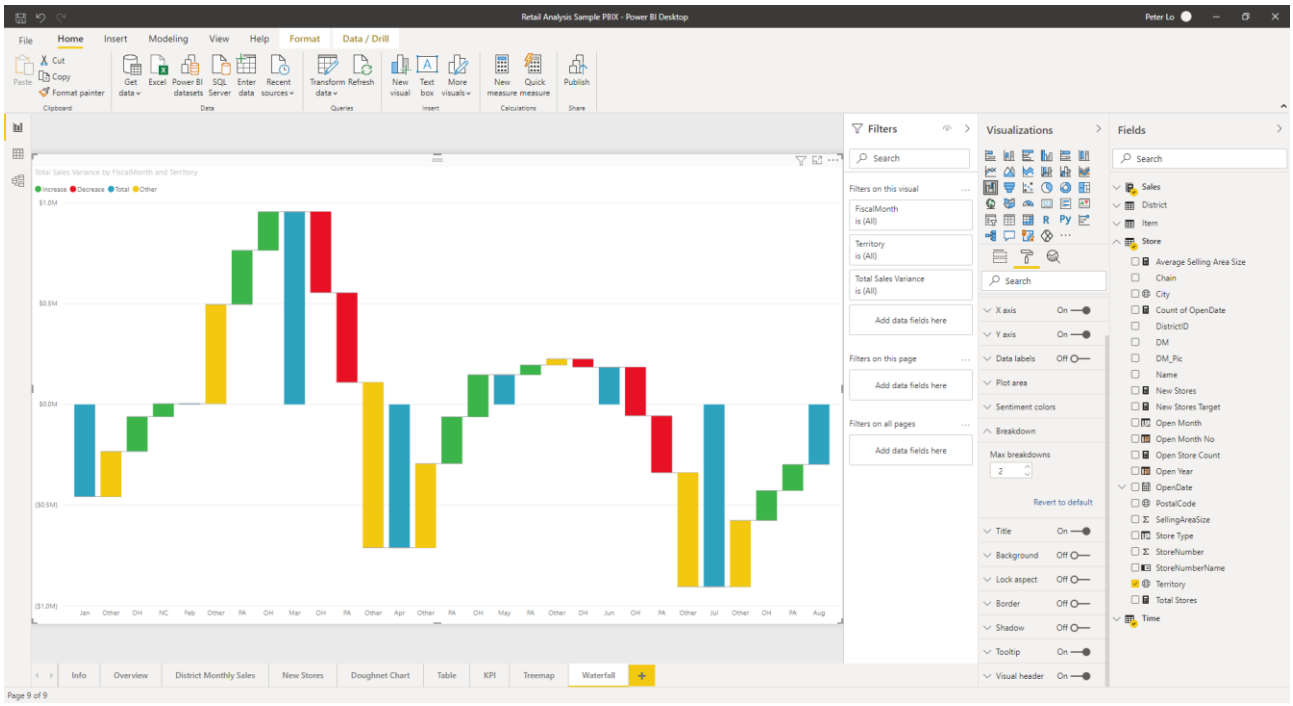
3.5.2 Explore the Waterfall Chart

Dig in a little more to see what's contributing most to the changes month to month.

1. Select *Store* ➔ *Territory*, which will add Territory to the **Breakdown** bucket. Power BI uses the value in Breakdown to add additional data to the visualization. It adds the top five contributors to increases or decreases for each fiscal month. This means that February, for example, now has six data points instead of just one.



2. Let's say that you're only interested in the top two contributors. In the **Format** pane, select **Breakdown** and set **Max breakdowns** to **2**. A quick review reveals that the territories of Ohio and Pennsylvania are the biggest contributors to movement, both negative and positive, in your waterfall chart.



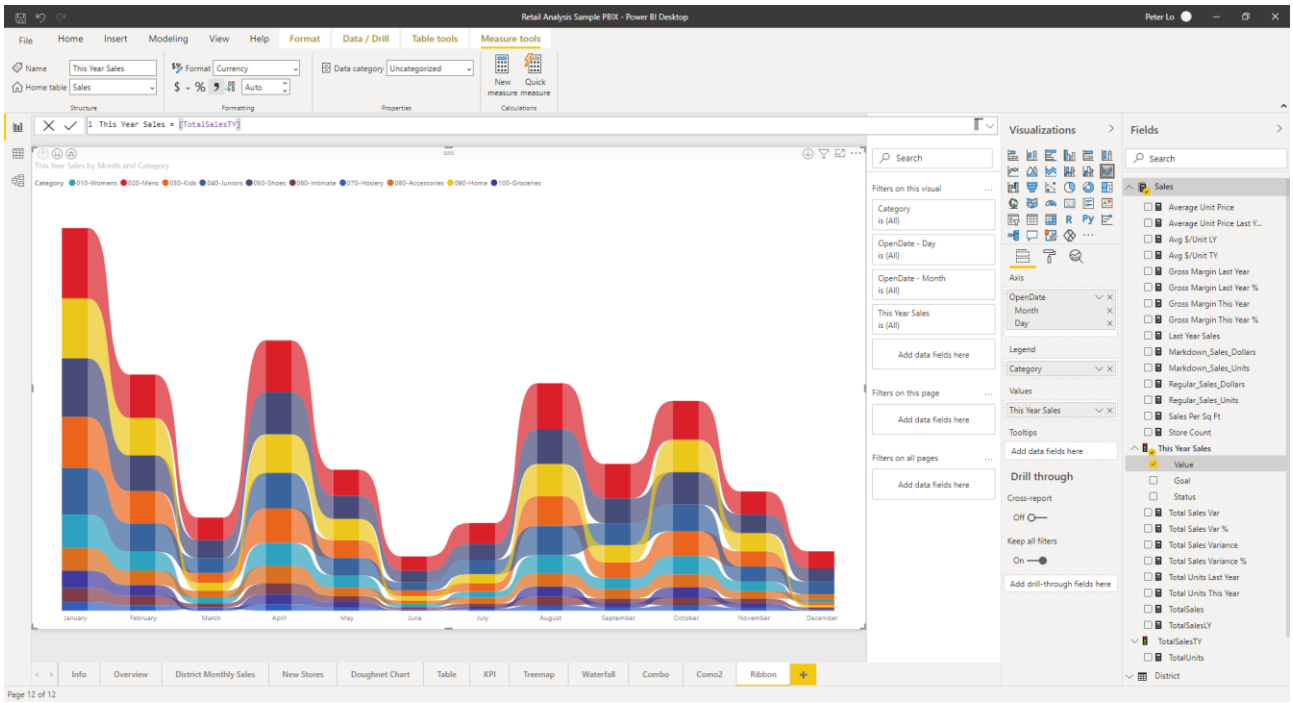
3.6 Ribbon Charts

You can create ribbon charts to visualize data, and quickly discover which data category has the highest rank (largest value). Ribbon charts are effective at showing rank change, with the highest range (value) always displayed on top for each time period.

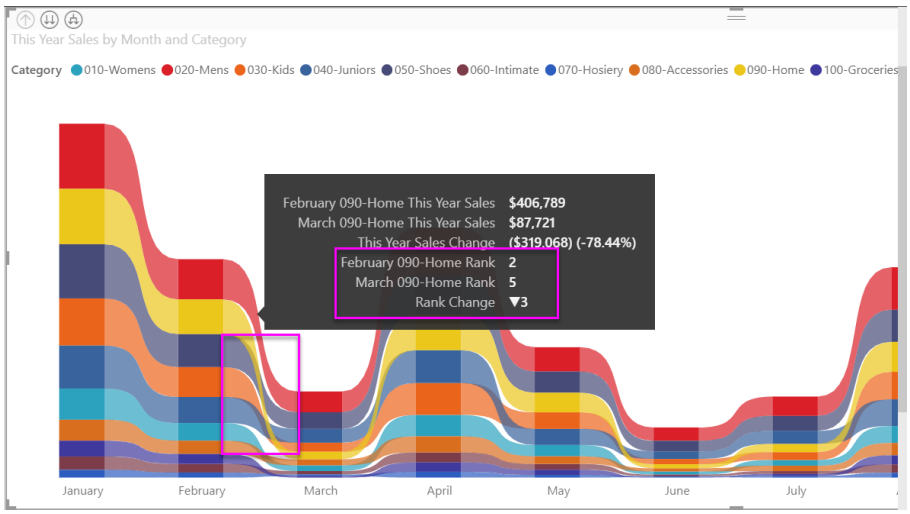
Ribbon charts connect a category of data over the visualized time continuum using ribbons, enabling you to see how a given category ranks throughout the span of the chart's x-axis (usually the timeline).

To create a ribbon chart, select Ribbon chart from the Visualizations panel.

1. Select **File ➔ Open Report**, and open *Retail Analysis Sample.pbix*.
2. Start on a blank report page and select **Ribbon** chart from the **Visualizations** panel
3. Select *Store* ➔ *OpenDate* (*Month* and *Day* only) as **Axis**, *Item* ➔ *Category* as **Legend**, and *Sales* ➔ *This year sales* ➔ *Value* as **Value**.



4. The ribbon chart shows rank for every month. Notice how rank changes across time. For example, the Home category moves from second to fifth from February to March.



3.6.1 Format a Ribbon Chart

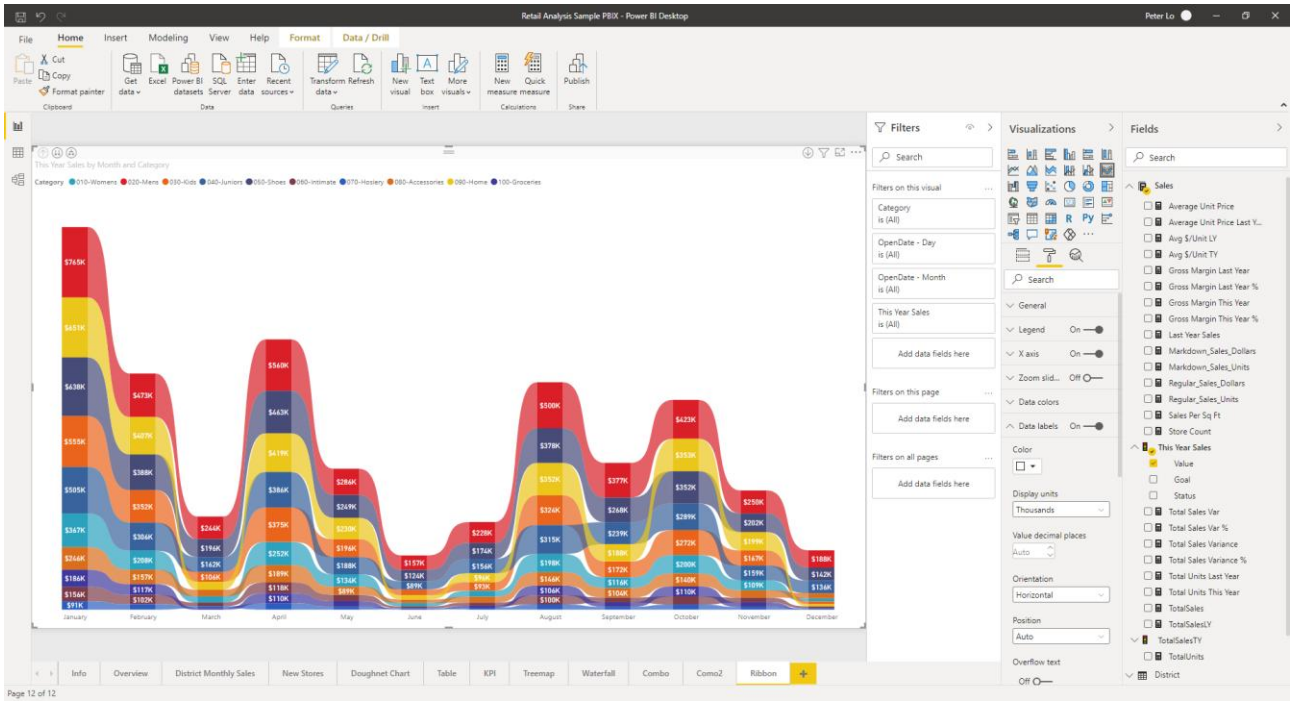
When you create a ribbon chart, you have formatting options available in the Format section of the Visualizations pane. The formatting options for ribbon charts are similar to those for a stacked column chart, with additional formatting options that are specific to the ribbons.

These formatting options for ribbon charts let you make adjustments.

- **Spacing** lets you adjust how much space appears between ribbons. The number is the percentage of the column's maximum height.
- **Match series color** allows you to match the color of the ribbons with the series color. When set to off, ribbons are gray.
- **Transparency** specifies how transparent the ribbons are, with the default set to 30.
- **Border** lets you place a dark border on the top and bottom of the ribbons. By default, borders are off.

Since the ribbon chart does not have y-axis labels, you may want to add data labels. From the **Formatting** pane, select **Data labels**.

- Set the **Text color** to **White**
- Set **Display units** to **Thousands**



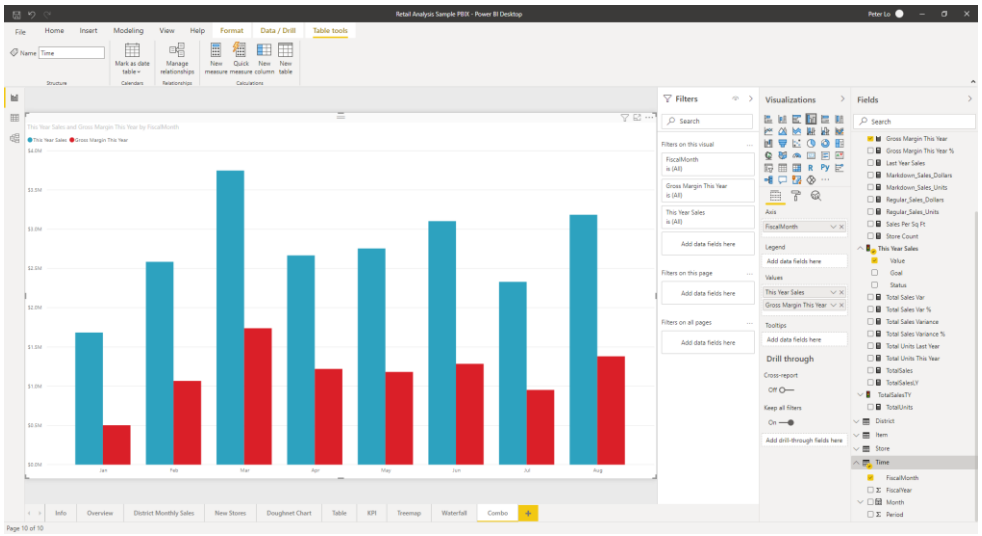
3.7 Combo Chart

A combo chart is a single visualization that combines a line chart and a column chart. Combining the 2 charts into one lets you make a quicker comparison of the data. Combo charts can have one or two Y axes. Combo charts are a great choice:

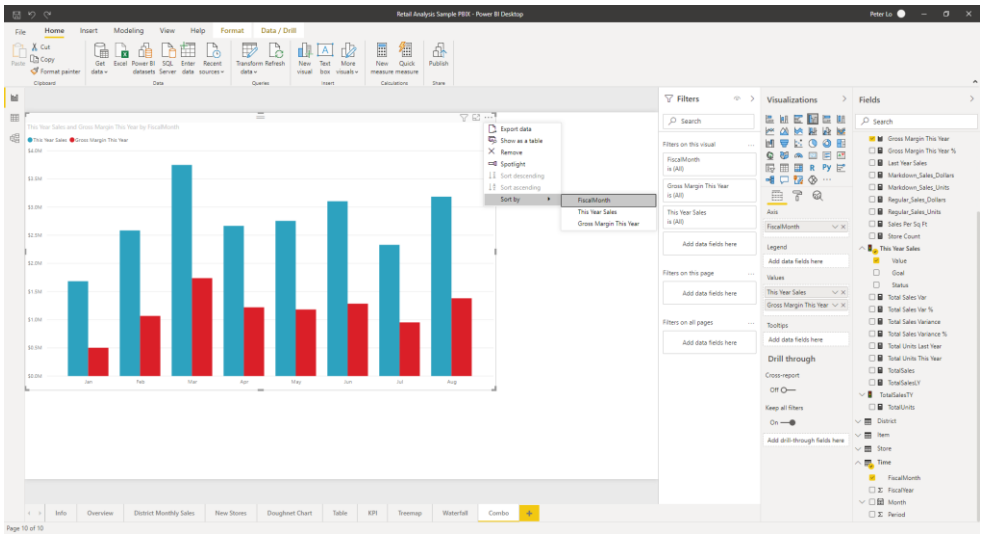
- When you have a line chart and a column chart with the same X axis.
- To compare multiple measures with different value ranges.
- To illustrate the correlation between two measures in one visualization.
- To check whether one measure meet the target which is defined by another measure
- To conserve canvas space.

In order to create a column chart that displays this year's sales and gross margin by month.

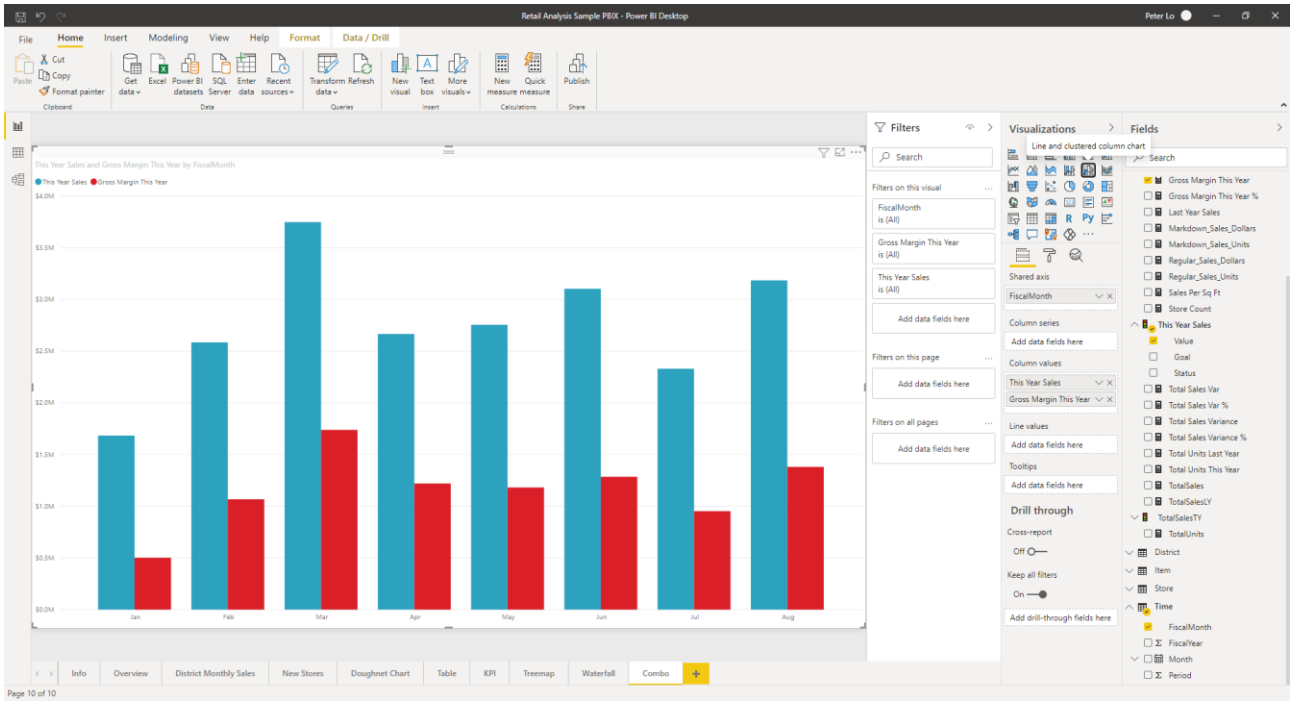
1. Select **File ➔ Open Report**, and open *Retail Analysis Sample.pbix*.
2. Start on a blank report page and from the **Fields** pane, select the *Sales ➔ This Year Sales ➔ Value*.
3. Drag *Sales ➔ Gross Margin This Year* to the **Value** well.
4. Select *Time ➔ FiscalMonth* to add it to the **Axis** well.



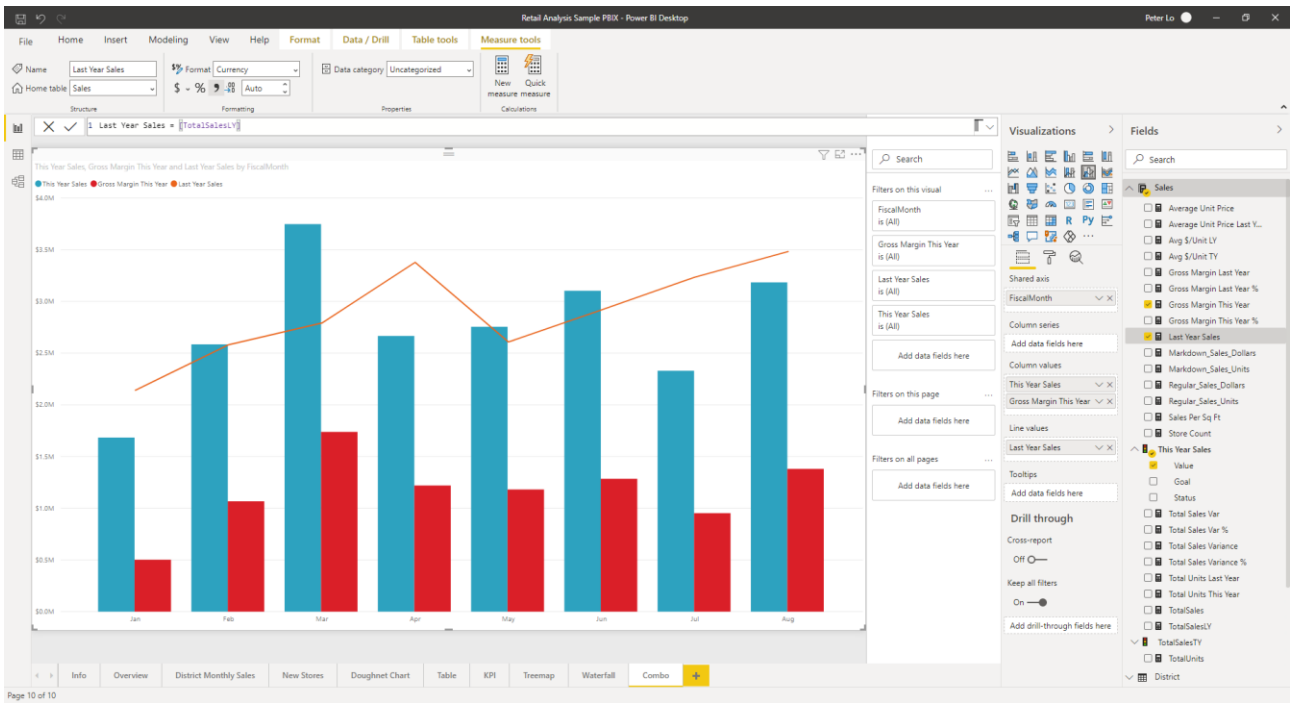
5. Select **More options (...)** in the upper-right corner of the visualization, and select **Sort by ➔ FiscalMonth**. To change the sort order, select the ellipsis again and choose either **Sort ascending** or **Sort descending**.



6. Convert the column chart to a combo chart. With the column chart selected, from the Visualizations pane select the **Line and clustered column chart**.



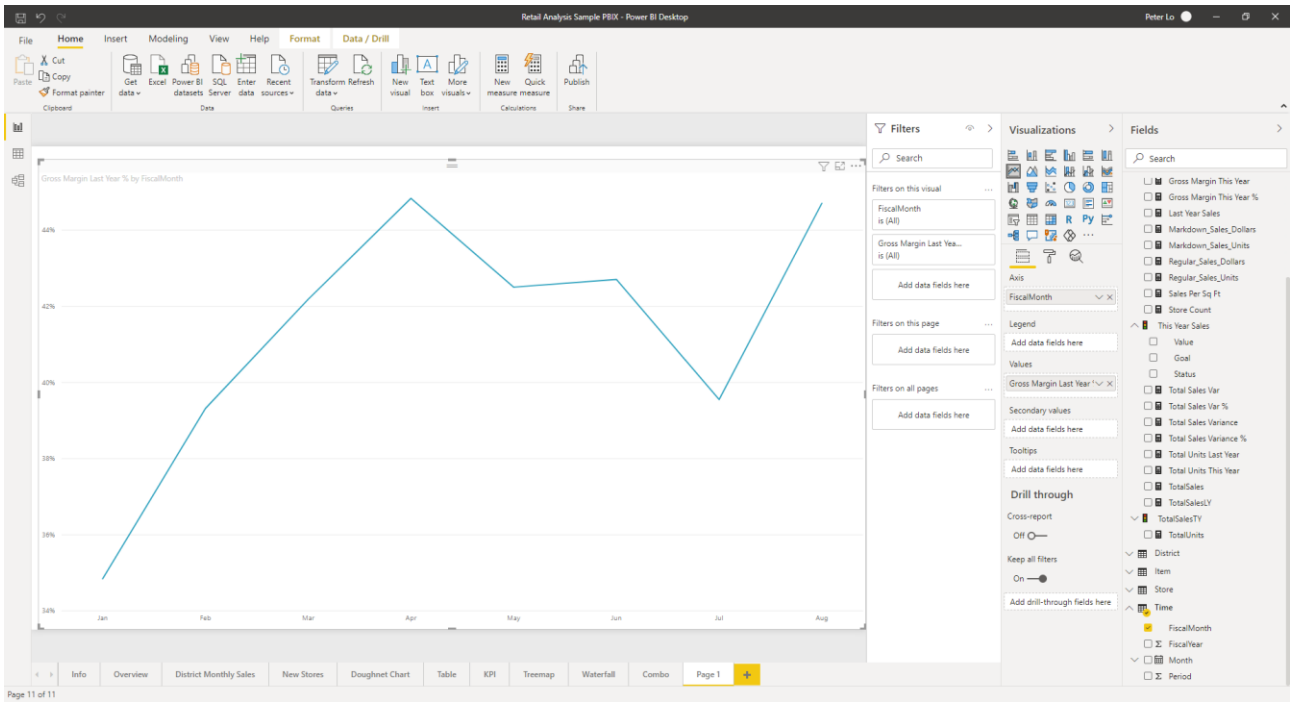
7. From the Fields pane, drag *Sales* ➔ *Last Year Sales* to the **Line Values** bucket.



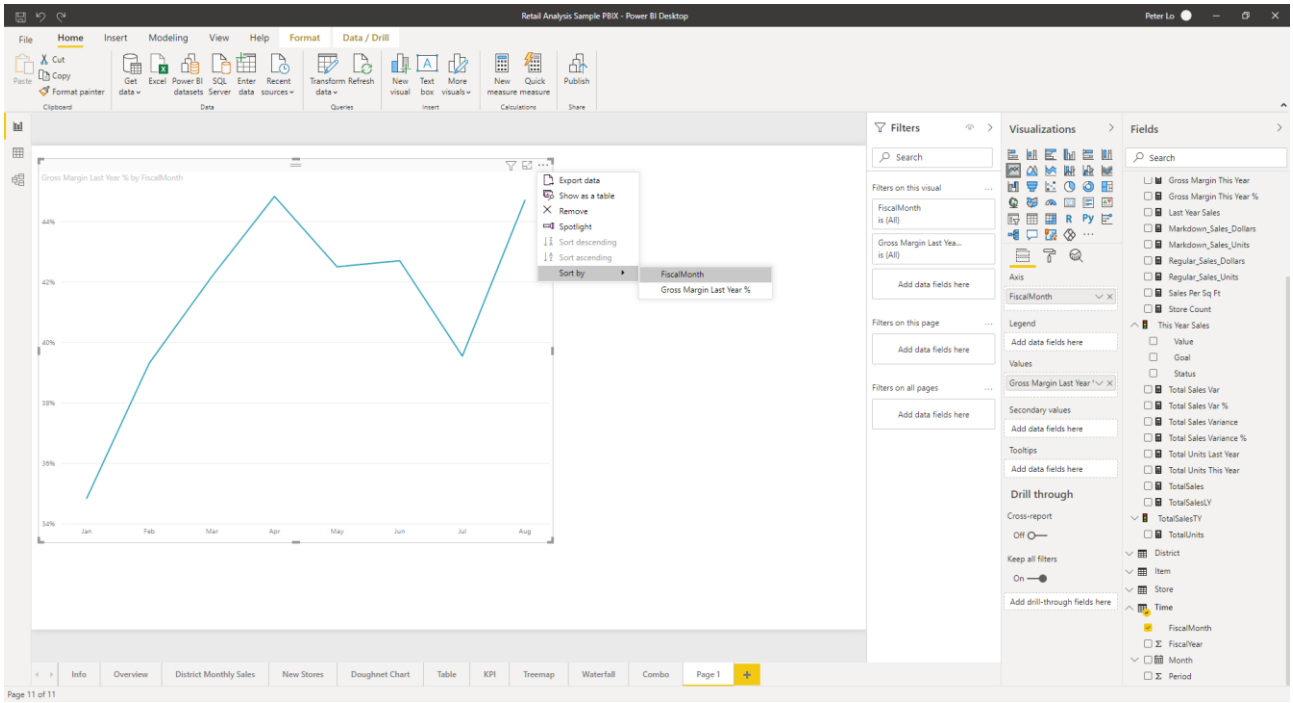
3.7.1 Create a Combo Chart with Two Axes

Create a new line chart that tracks Gross Margin last year % by Fiscal Month. In January GM% was 35%, peaked at 45% in April, dropped in July and peaked again in August. Will we see a similar pattern in sales last year and this year?

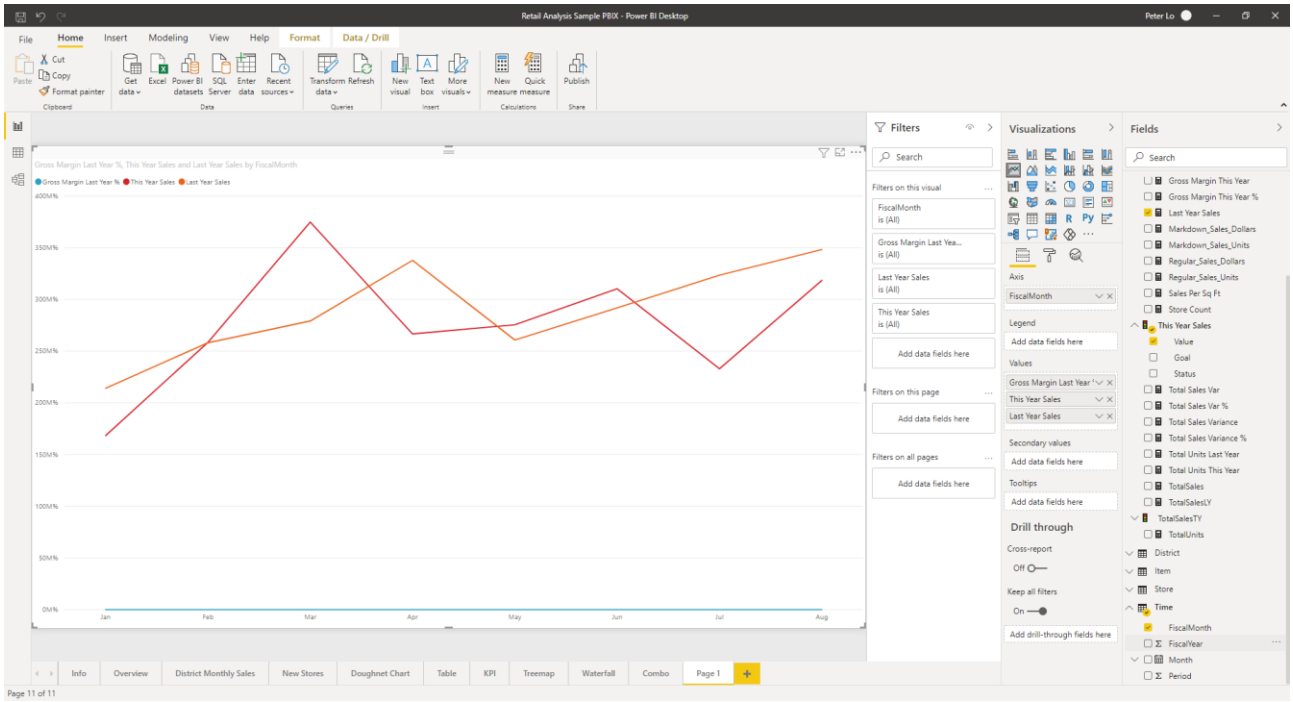
1. Start on a blank report page and select Line Chart. Then from the **Fields** pane, select *Sales* ➔ *Gross Margin last year %* and *Time* ➔ *FiscalMonth*.



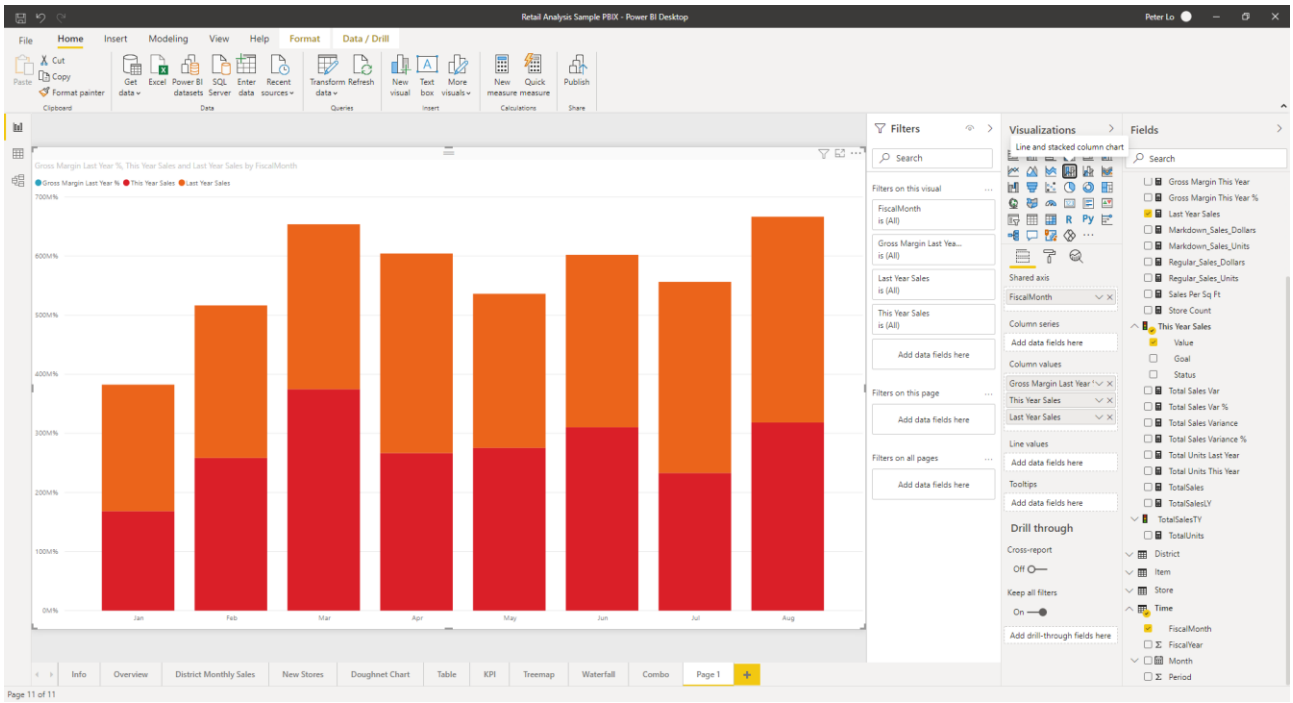
2. Select the ellipsis to sort it by Month and Ascending.



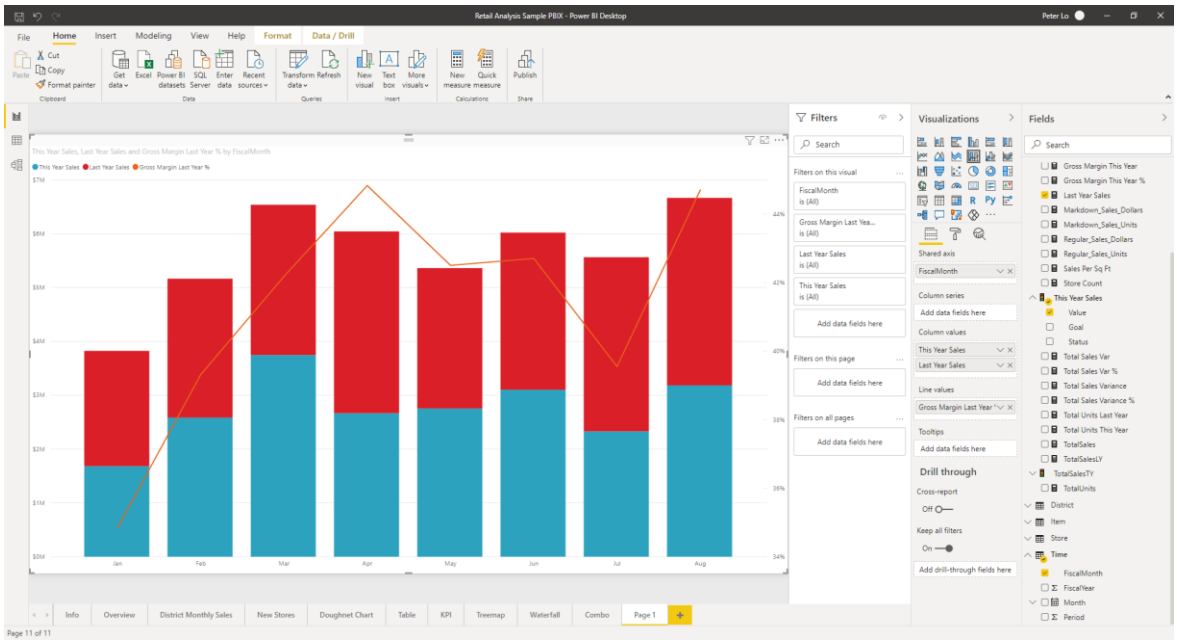
3. Add *Sales* ➔ *This Year Sales* ➔ *Value* and *Sales* ➔ *Last Year Sales* to the line chart. The scale of Gross Margin Last Year % is much smaller than the scale of Sales which makes it difficult to compare.



4. To make the visual easier to read and interpret, convert the line chart to a **Line and Stacked Column chart**.

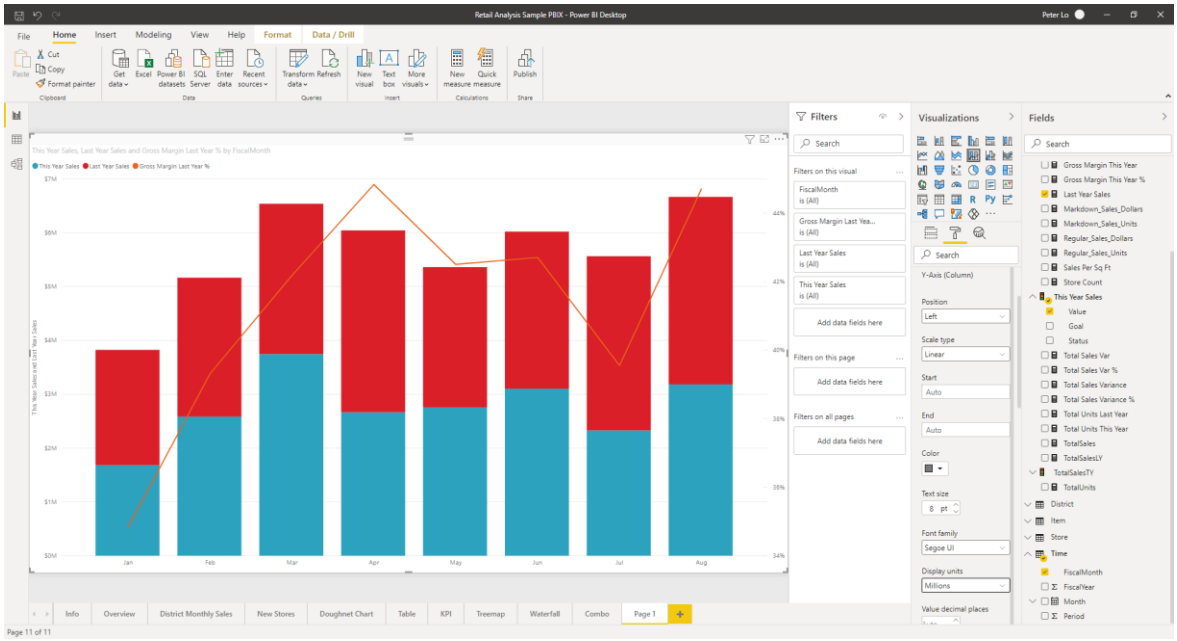


5. Drag *Gross Margin Last Year %* from **Column Values** into **Line Values**. Power BI creates two axes, thus allowing the datasets to be scaled differently; the left measures sales dollars and the right measures percentage. Now we see the answer to our question; yes, we do see a similar pattern.

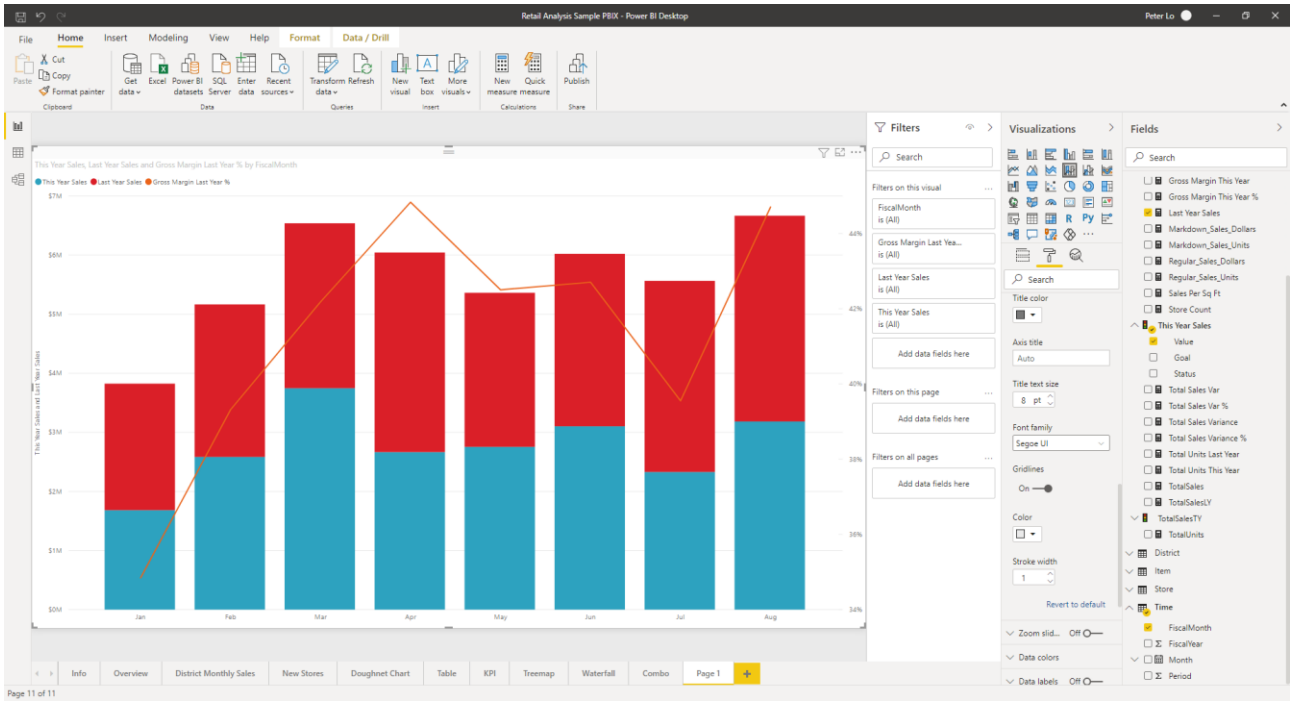


3.7.2 Add Titles to the Axes

1. Select the paint roller icon to open the **Formatting** pane, Select the down arrow to expand the **Y-axis** options.
    - Set **Position** to **Left**
    - Set **Title** to **On**
    - Set **Display units** as **Millions**
    - Set **Style** to **Show title only**

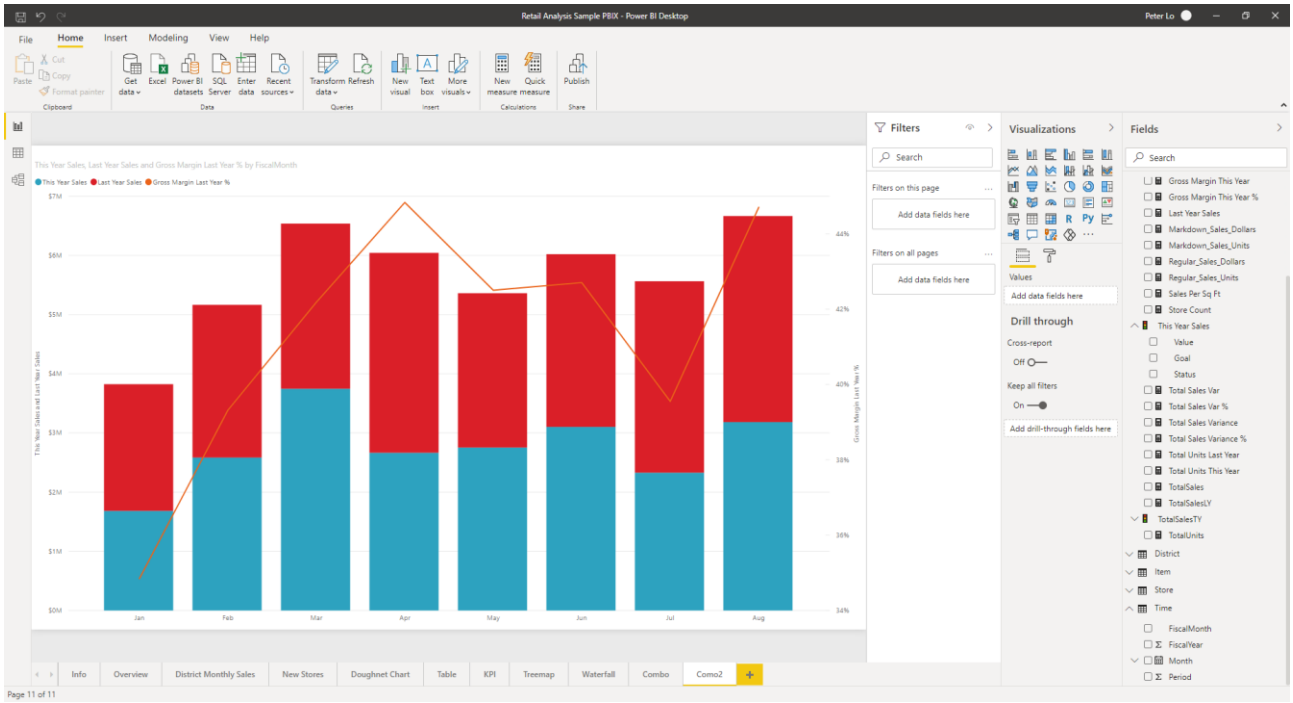


2. Under **Y-Axis (Column)**, scroll down until you see **Show secondary**. The Show secondary section displays options for formatting the line chart portion of the combo chart.



3. For **Y-Axis (Line)**,
    - Leave **Position** as **Right**
    - Turn **Title** to **On**
    - Set **Style** to **Show title only**

Your combo chart now displays dual axes, both with titles.



3.8 Funnel Chart

A funnel chart helps you visualize a linear process that has sequential connected stages. For example, a sales funnel that tracks customers through stages: Lead ➔ Qualified Lead ➔ Prospect ➔ Contract ➔ Close. At a glance, the shape of the funnel conveys the health of the process you're tracking.

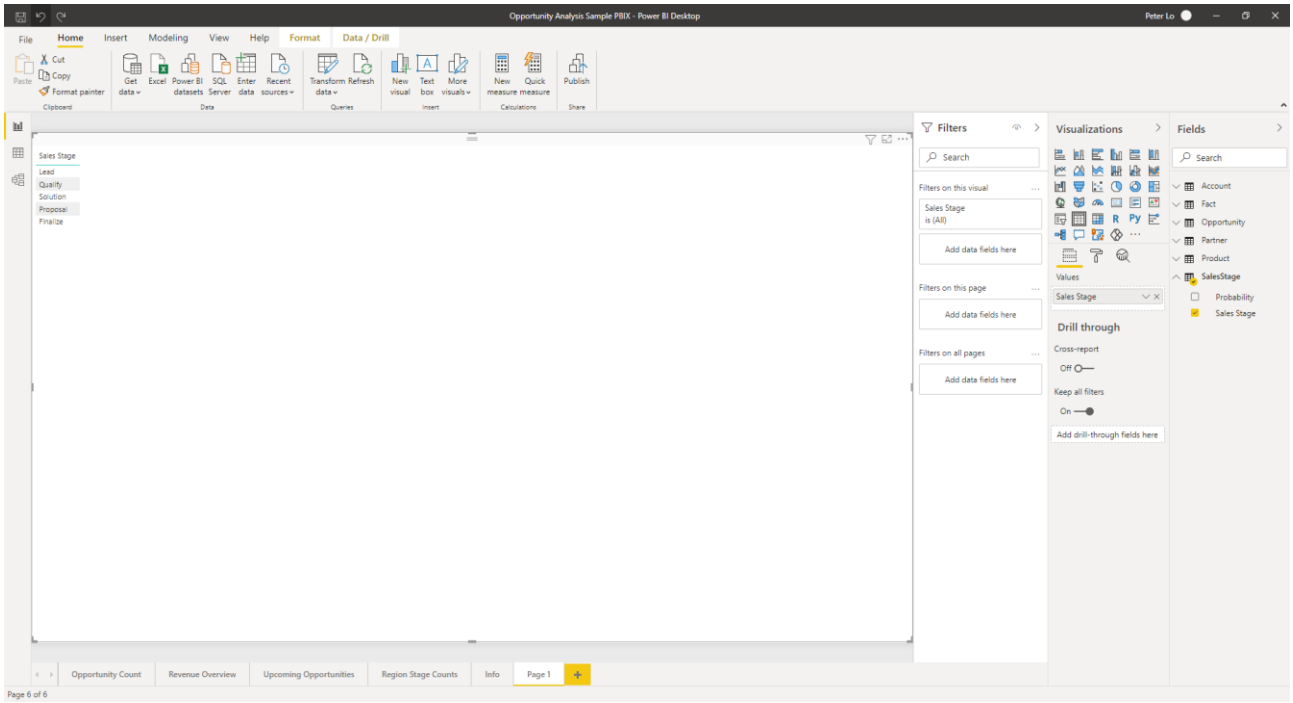
Each funnel stage represents a percentage of the total. So, in most cases, a funnel chart is shaped like a funnel – with the first stage being the largest, and each subsequent stage smaller than its predecessor. A pear-shaped funnel is also useful – it can identify a problem in the process. But typically, the first stage, the "intake" stage, is the largest.

Funnel charts are a great choice:

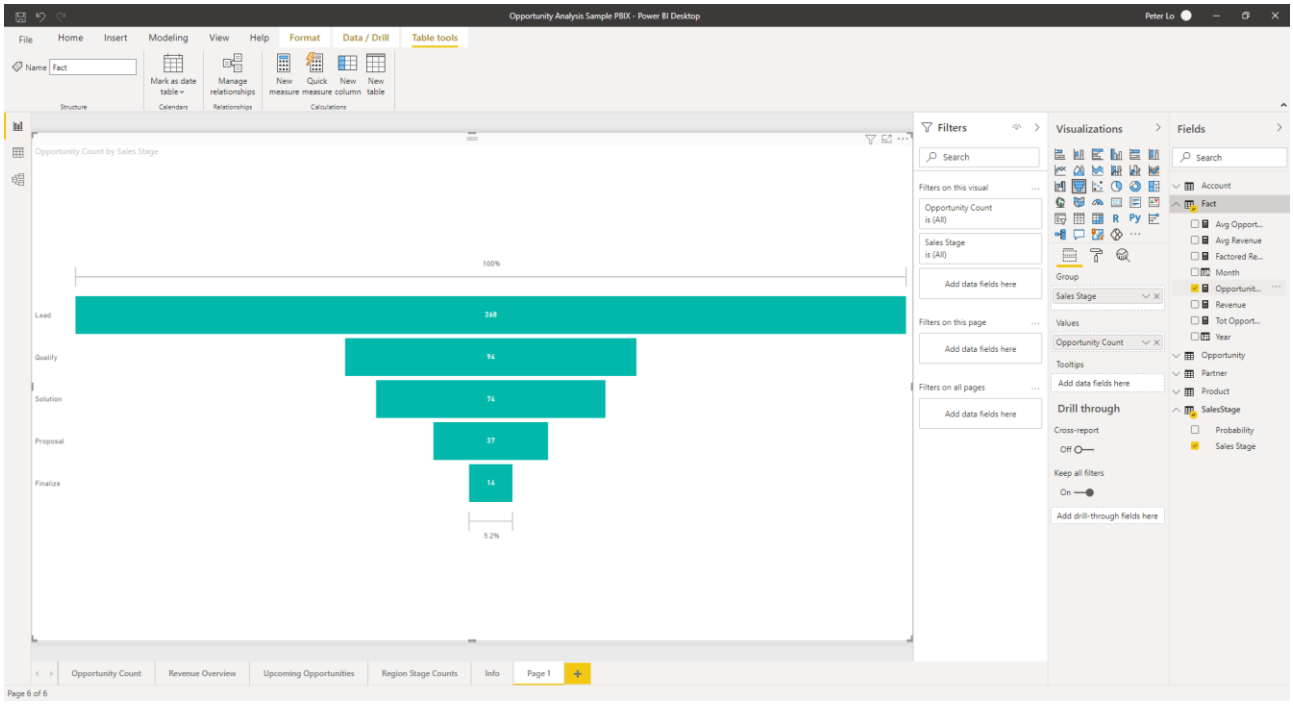
- When the data is sequential and moves through at least 4 stages.
- When the number of "items" in the first stage is expected to be greater than the number in the final stage.
- To calculate potential (revenue/sales/deals/etc.) by stages.
- To calculate and track conversion and retention rates.
- To reveal bottlenecks in a linear process.
- To track a shopping cart workflow.
- To track the progress and success of click-through advertising/marketing campaigns.

In order to create a funnel chart that shows the number of opportunities we have in each of our sales stages

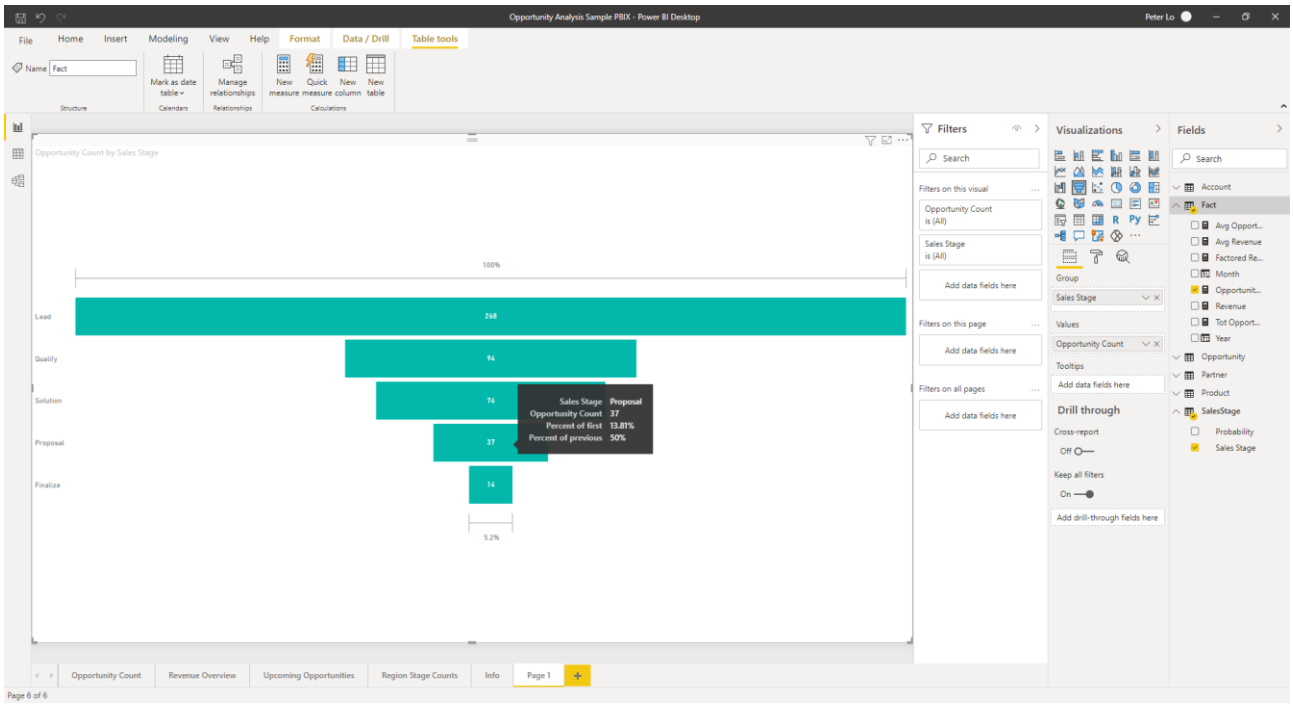
1. Select **File ➔ Open Report**, and open *Opportunity Analysis Sample.pbix*.
2. Start on a blank report page and from the **Fields** pane, select the *SalesStage ➔ Sales Stage*.



3. Select the funnel icon to convert the column chart to a funnel chart. From the Fields pane, select *Fact* ➔ *Opportunity Count*.



4. Hovering over a bar displays a wealth of information.
   - The name of the stage
   - Number of opportunities currently in this stage
   - Overall conversion rate (% of Lead)
   - Stage-to-stage (also known as Drop Rate) which is the % of the previous stage (in this case, Proposal Stage/Solution Stage)



3.9 Scatter Charts and Bubble Charts

A scatter chart always has two value axes to show: one set of numerical data along a horizontal axis and another set of numerical values along a vertical axis. The chart displays points at the intersection of an x and y numerical value, combining these values into single data points. Power BI may distribute these data points evenly or unevenly across the horizontal axis. It depends on the data the chart represents. You can set the number of data points, up to a maximum of 10,000.

Scatter charts are a great choice:

- To show relationships between two numerical values.
- To plot two groups of numbers as one series of x and y coordinates.
- To use instead of a line chart when you want to change the scale of the horizontal axis.
- To turn the horizontal axis into a logarithmic scale.
- To display worksheet data that includes pairs or grouped sets of values.
- To show patterns in large sets of data, for example by showing linear or non-linear trends, clusters, and outliers.
- To compare large numbers of data points without regard to time. The more data that you include in a Scatter chart, the better the comparisons that you can make.

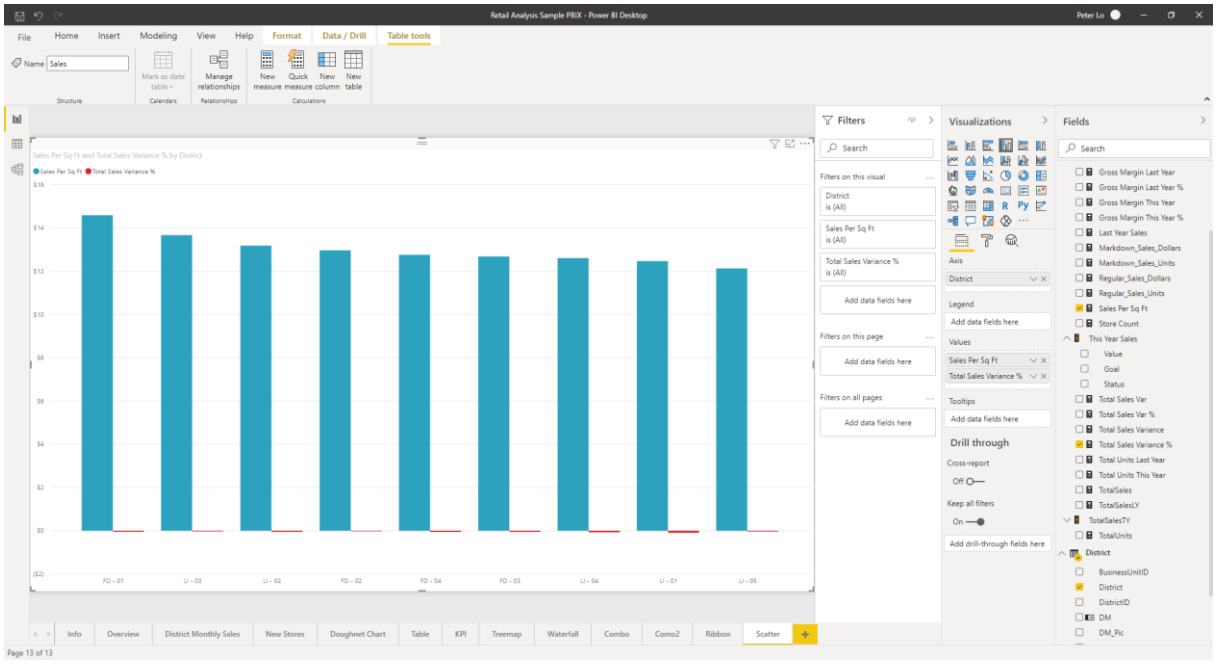
In addition to what Scatter charts can do for you, bubble charts are a great choice:

- If your data has three data series that each contains a set of values.
- To present financial data. Different bubble sizes are useful to visually emphasize specific values.
- To use with quadrants.

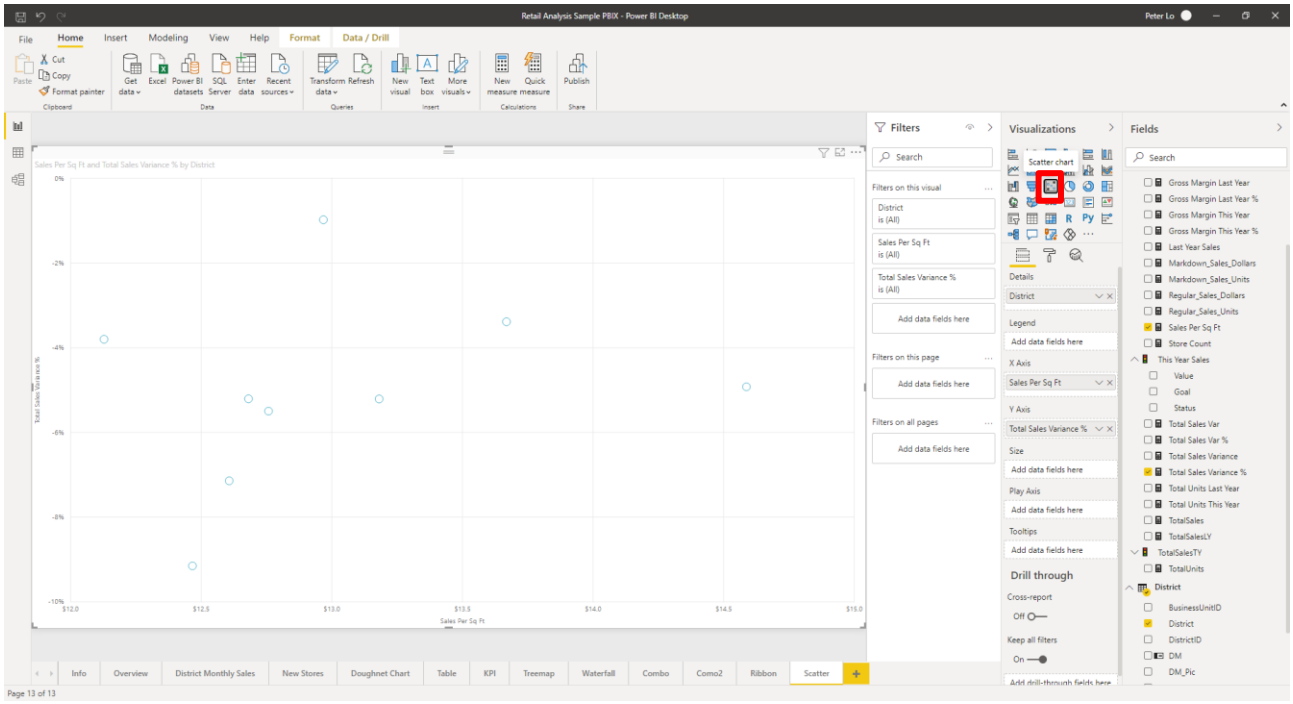
3.9.1 Scatter Chart

In order to create a scatter chart:

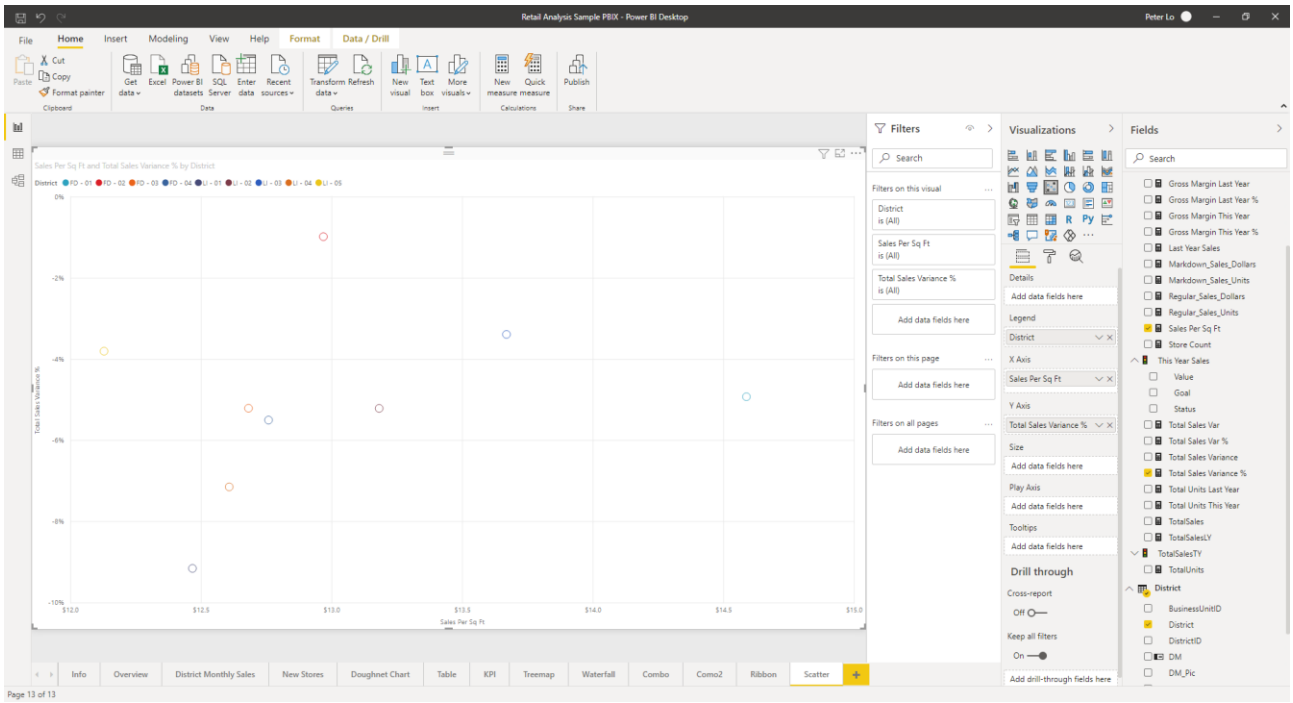
1. Select **File ➔ Open Report**, and open *Retail Analysis Sample.pbix*.
2. Start on a blank report page, select *Sales ➔ Sales Per Sq Ft, Sales ➔ Total Sales Variance %, and District ➔ District* from the **Field** pane.



3. In the **Visualization** pane, select **Scatter Chart** to convert the cluster column chart to a scatter chart.

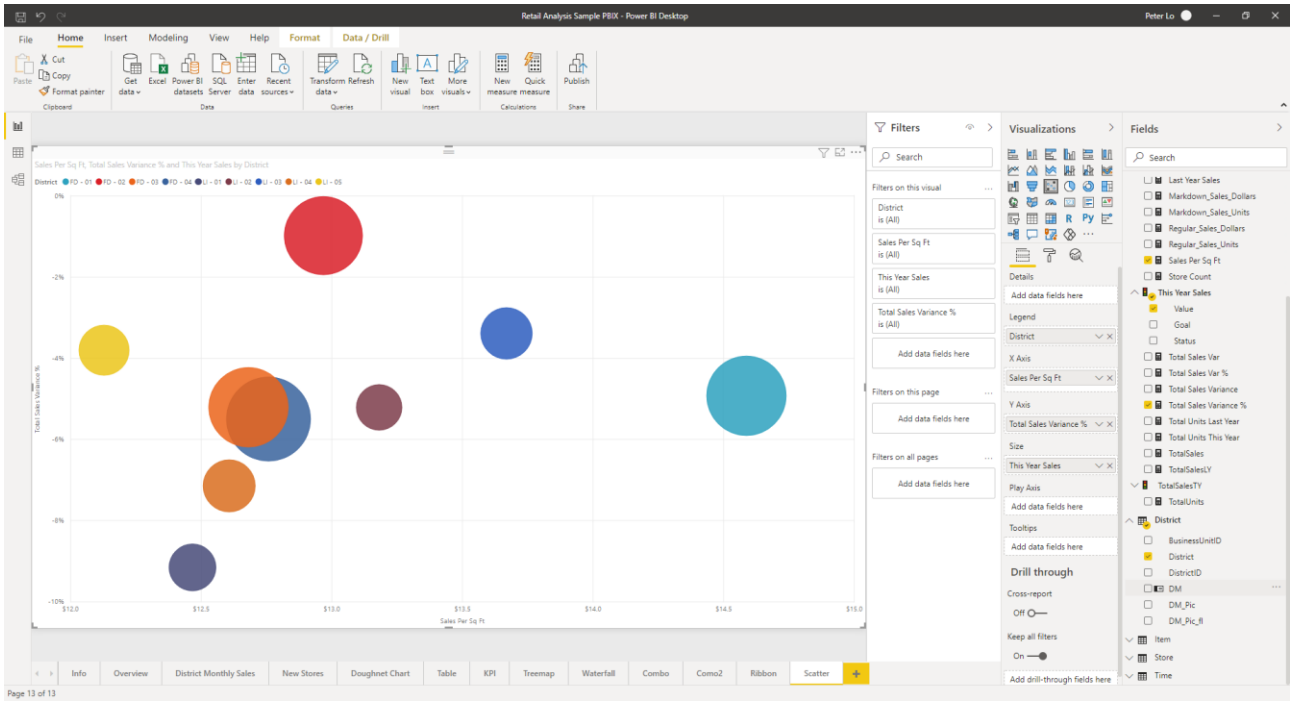


4. Drag *District* from **Details** to **Legend**. Power BI displays a scatter chart that plots Total Sales Variance % along the Y-Axis, and plots Sales Per Square Feet along the X-Axis. The data point colors represent districts:

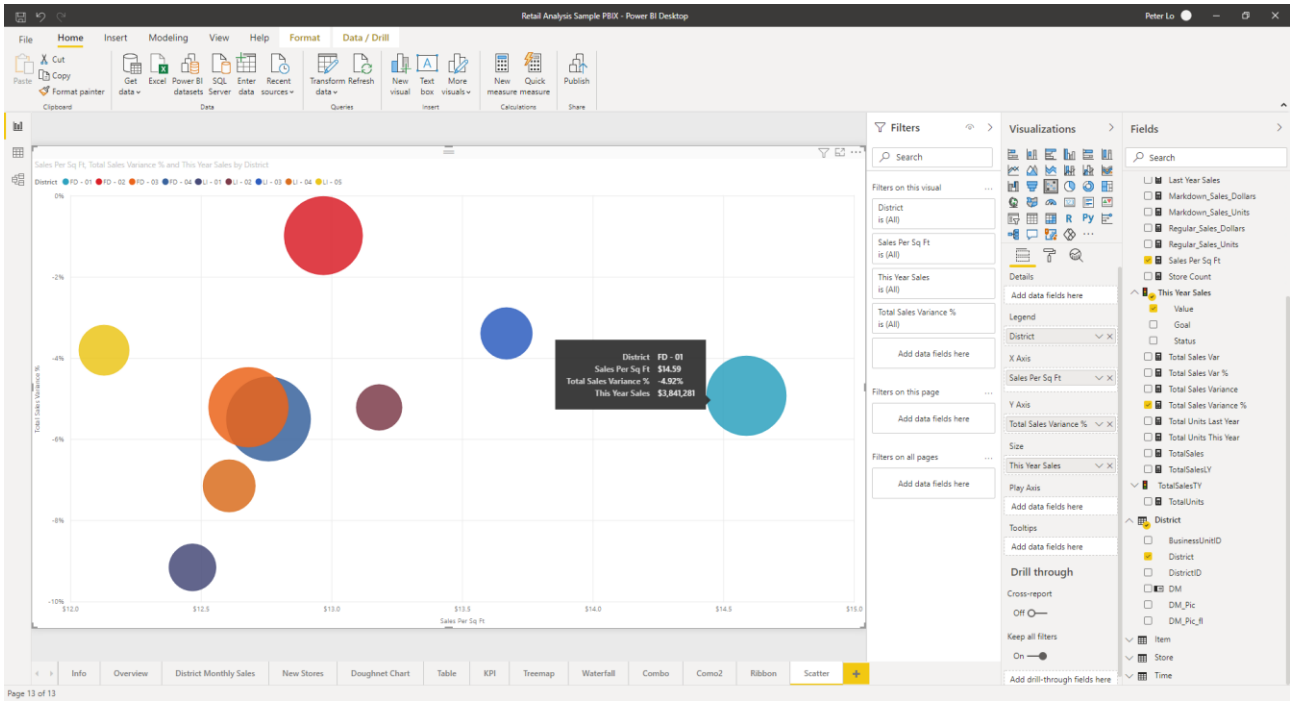


3.9.2 Bubble Charts

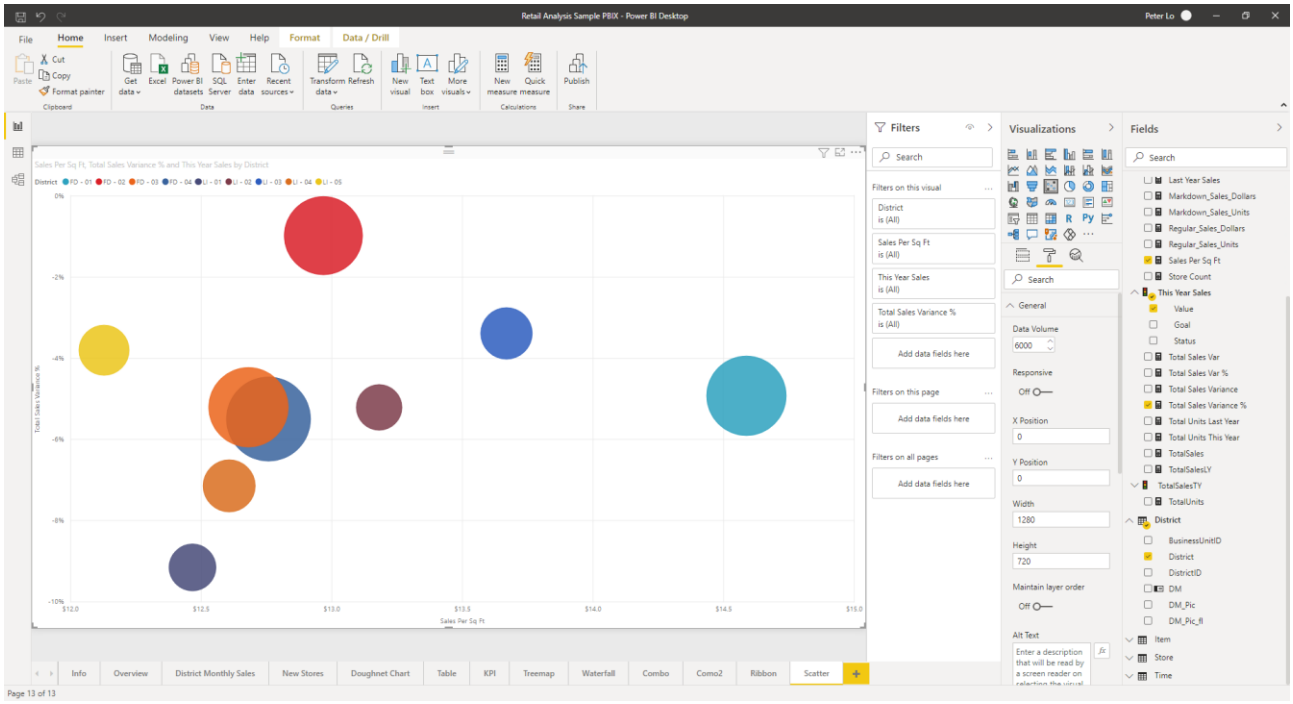
1. From the **Fields** pane, drag *Sales* ➔ *This Year Sales* ➔ *Value* to the **Size** well. The data points expand to volumes proportionate with the sales value.



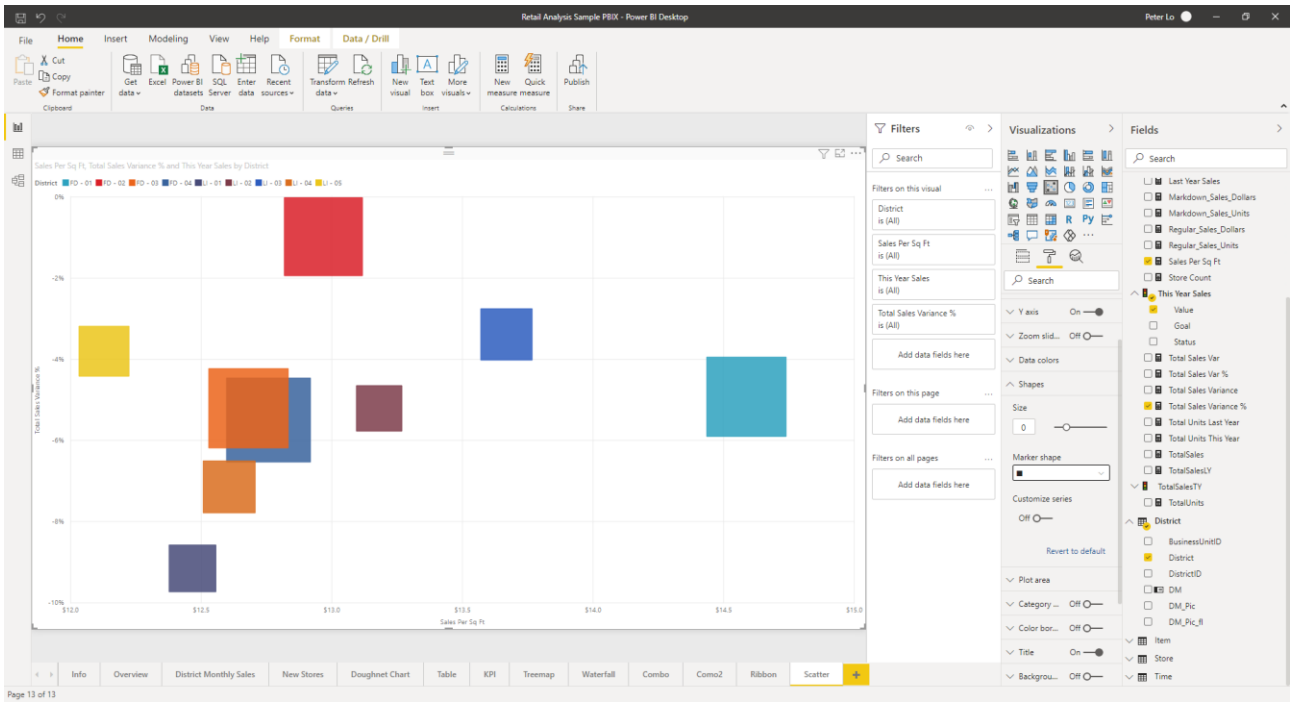
2. Hover over a bubble. The size of the bubble reflects the value of This Year Sales.



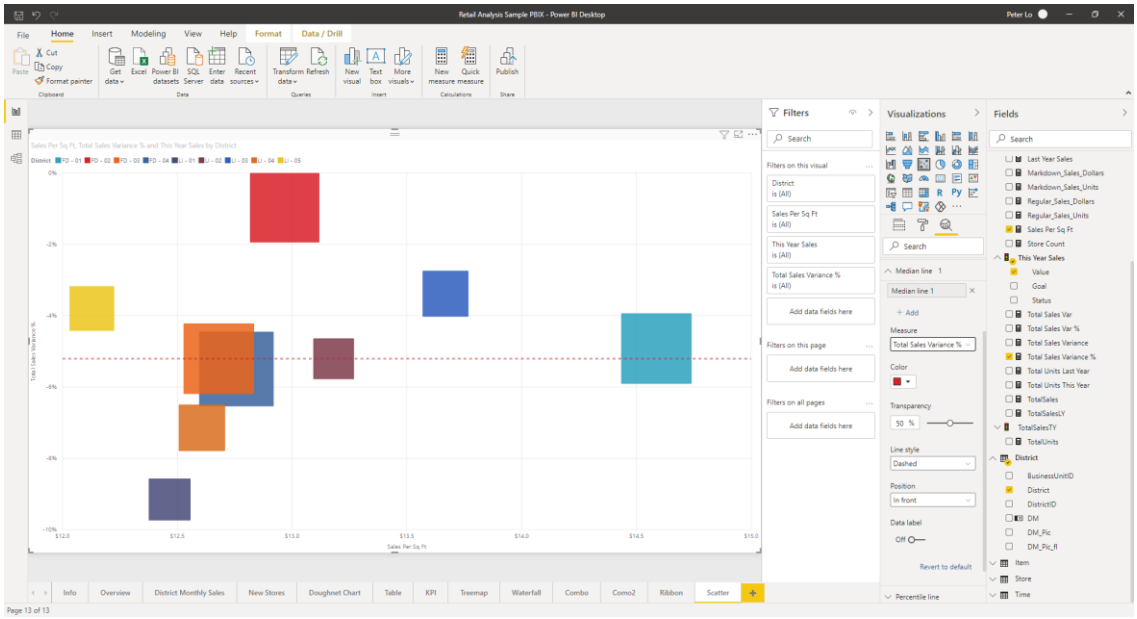
3. To set the number of data points to show in your bubble chart, in the Format section of the **Visualizations** pane, expand **General**, and adjust the **Data Volume** to **6000**.



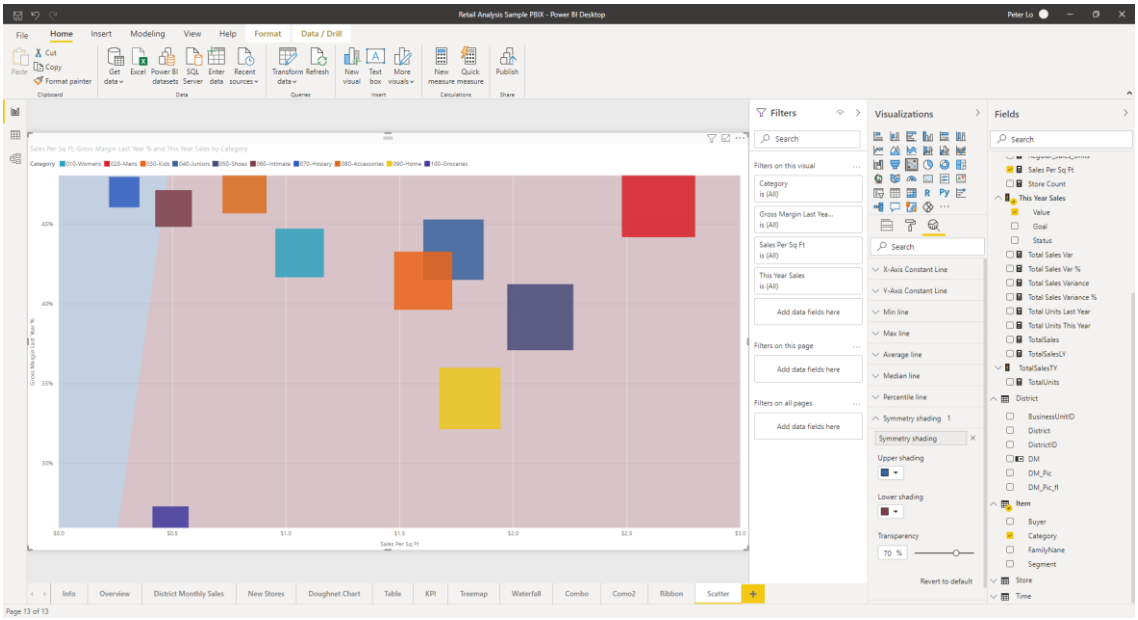
4. Continue formatting the visualization colors, labels, titles, background, and more. To improve accessibility, consider adding marker shapes to each line. To select the marker shape, expand **Shapes**, select **Marker shape**, and select a shape.



5. Open the **Analytics** pane, and select **Median line** and press **[+ Add]**. By default, Power BI adds a median line for *Sales per sq ft*. This isn't very helpful since we can see that there are 10 data points and know that the median will be created with five data points on each side. Instead, switch the Measure to *Total sales variance %*.



6. When you turn symmetry shading on in the Analytics pane, Power BI shows you the background of your scatter chart symmetrically based on your current axis upper and lower boundaries. This is a very quick way to identify which axis measure a data point favors, especially when you have a different axis range for your x- and y-axis.
    - Change **Legend** to *Item* ➔ *Category*
    - Change **Y Axis** to *Sales* ➔ *Gross margin last year %*
    - From the **Analytics** pane, add **Symmetry shading**. We can see from the shading that Hosiery (the green bubble in the pink shaded area) is the only category that favors gross margin rather than its sales per store square footage.



3.10 Radial Gauge Charts

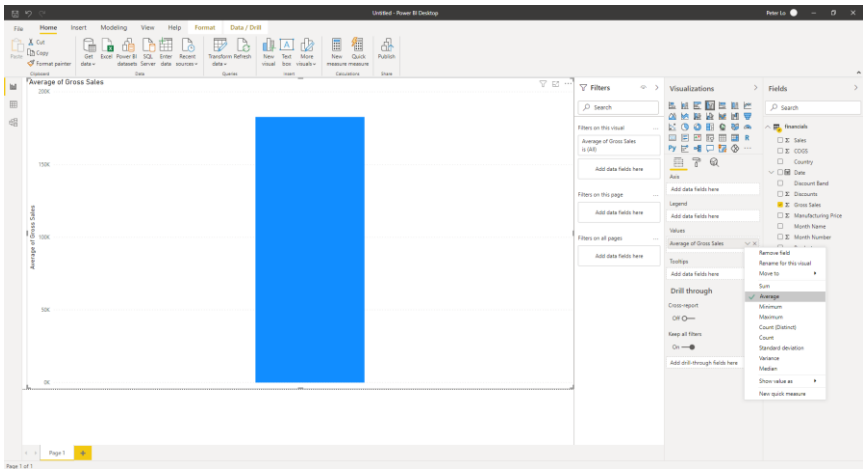
A radial gauge chart has a circular arc and shows a single value that measures progress toward a goal or a Key Performance Indicator (KPI). The line (or needle) represents the goal or target value. The shading represents the progress toward that goal. The value inside the arc represents the progress value. Power BI spreads all possible values evenly along the arc, from the minimum (left-most value) to the maximum (right-most value). Radial gauges are a great choice to:

- Show progress toward a goal.
- Represent a percentile measure, like a KPI.
- Show the health of a single measure.
- Display information you can quickly scan and understand.

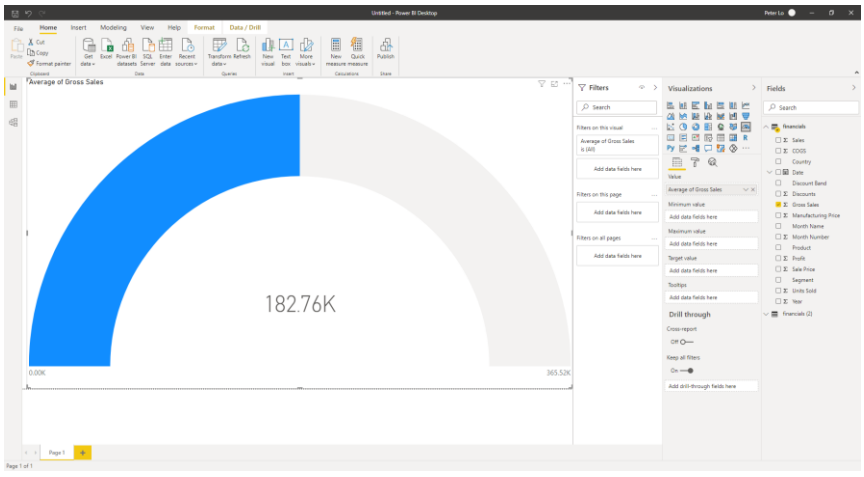
Assume that you're a car retailer tracking the sales team's average sales per month. The needle represents a 140 cars sales goal. The minimum possible average sales is 0 and the maximum is 200. The blue shading shows that the team is averaging approximately 120 sales this month.

3.10.1 Create a Gauge to Track Gross Sales

1. Select **Get Data ➔ Excel** in the **Home** ribbon tab, and open *Financial Sample.xlsx* workbook.
2. Select **financials** and click **[Load]**.
3. From the **Fields** pane, select *Gross Sales*. Change the **aggregation** to **Average**.

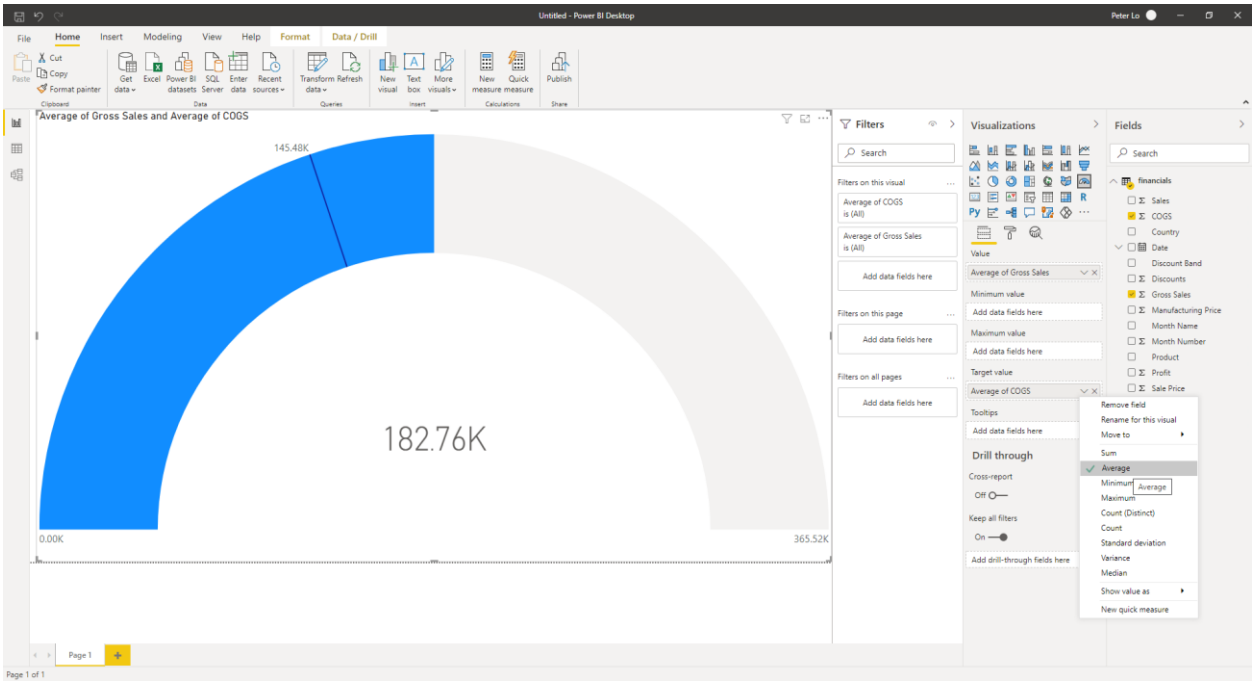


4. Select the **Gauge** icon to convert the column chart to a gauge chart.



3.10.2 Set a Target Value

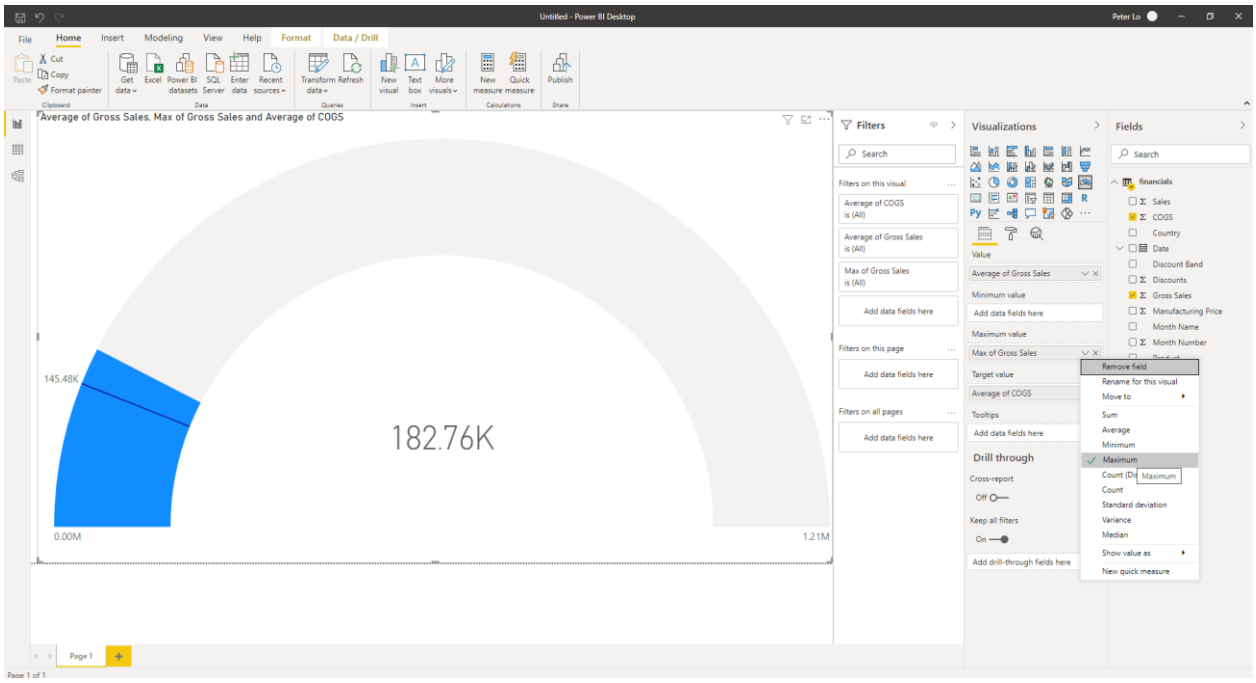
1. Drag *COGS* from the **Fields** pane to the **Target value** well.
2. Change the **aggregation** to **Average**. Power BI adds a needle to represent our target value of $145.48K.



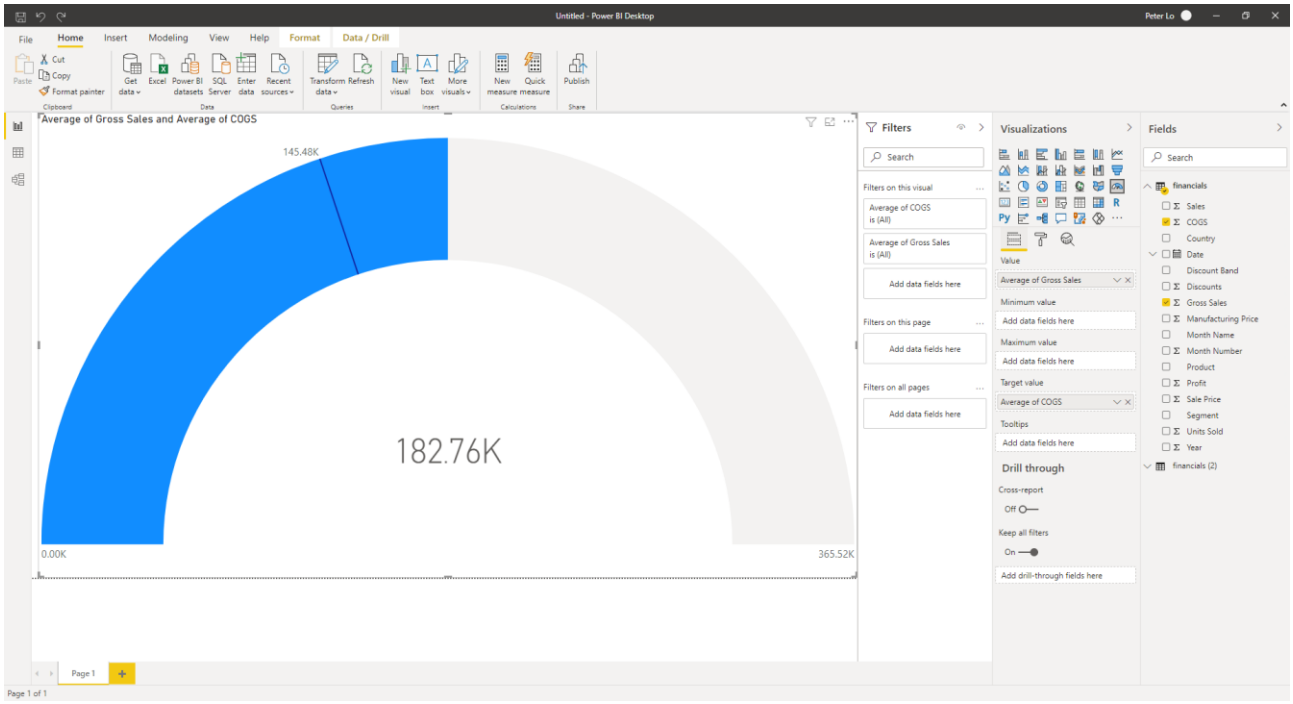
3.10.3 Set a Maximum Value

Power BI used the Value field to automatically set minimum and maximum values. What if you want to set your own maximum value? Let's say that, instead of using double the current value as the maximum possible value, you want to set it to the highest Gross Sales number in your dataset.

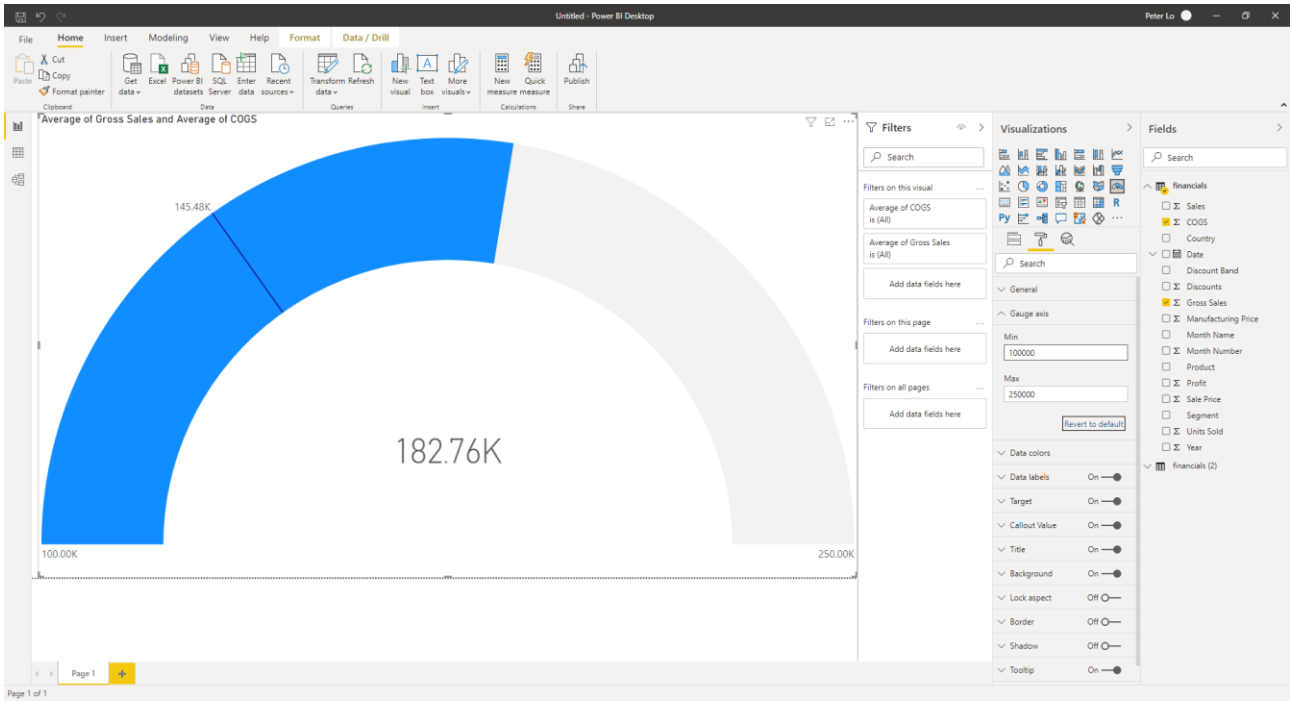
1. Drag *Gross Sales* from the **Fields** pane to the **Maximum value** well.
2. Change the **aggregation** to **Maximum**.



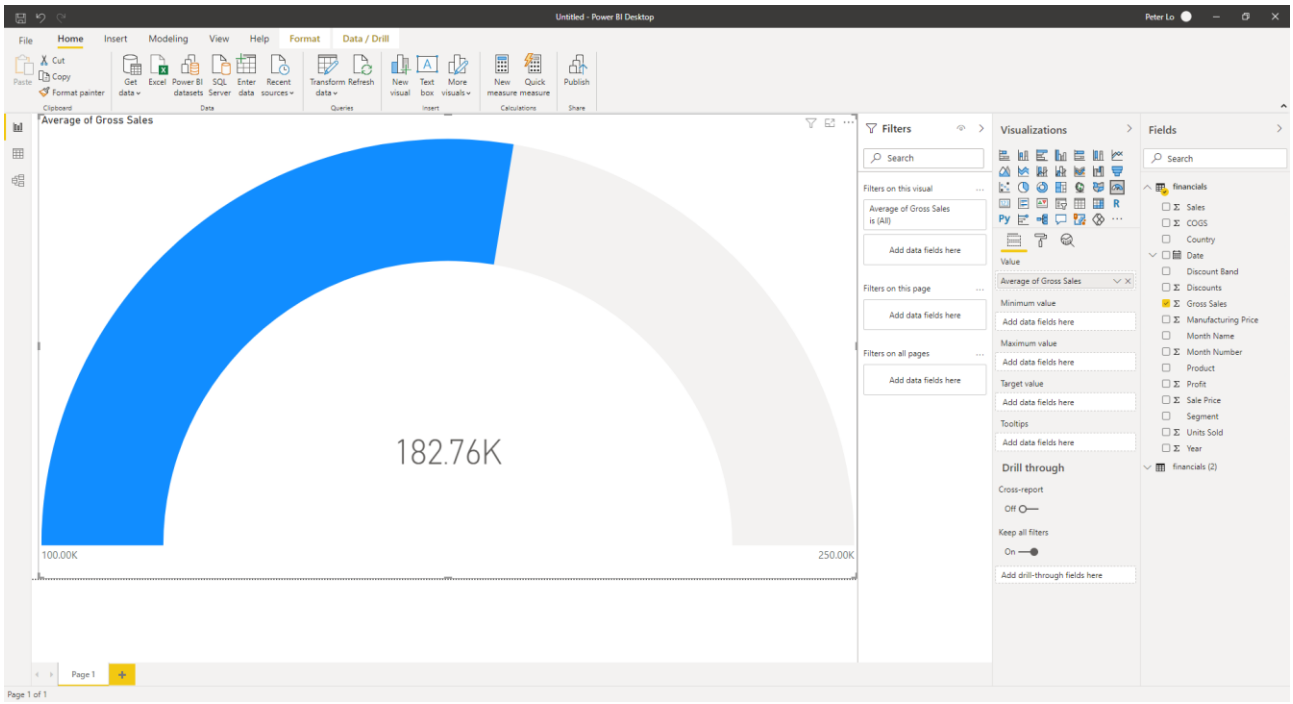
3. If you want to use manual format options to set Minimum, Maximum, and Target values, remove **Max of Gross Sales** from the **Maximum value** well.



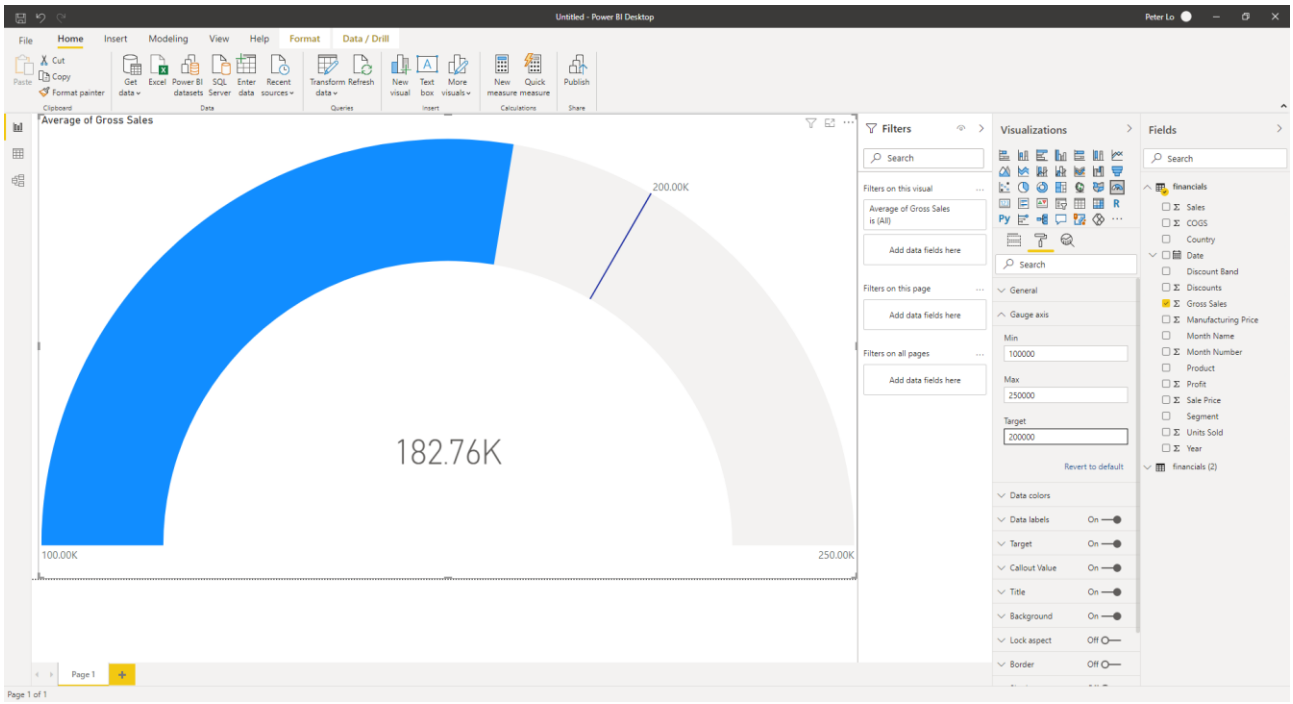
4. Select the paint roller icon to open the **Format** pane. Expand **Gauge axis** and enter values for Min and Max.



5. Remove *Average of COGS* from **Target Value**.



6. When the Target field appears under Gauge axis, enter a value.



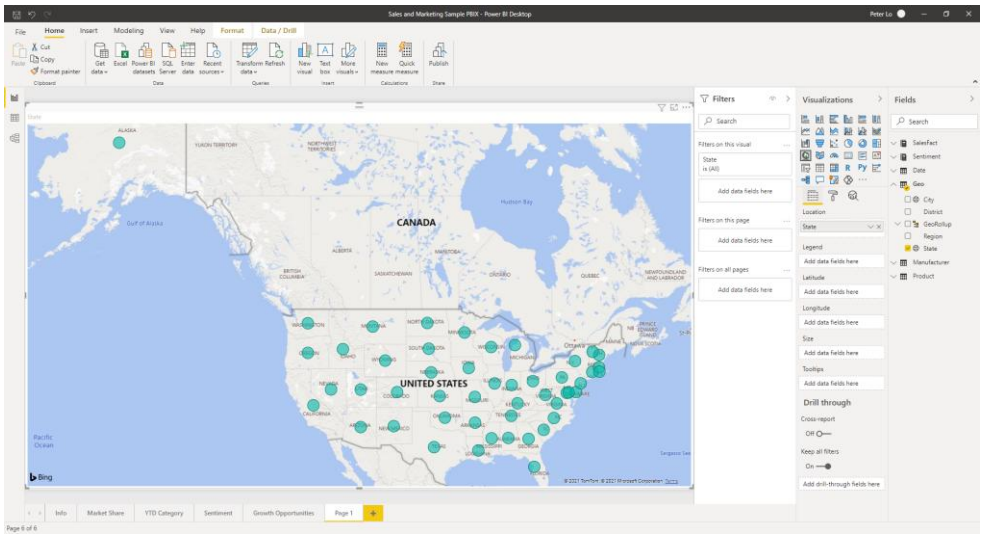
3.11 Filled Map

A filled map uses shading or tinting or patterns to display how a value differs in proportion across a geography or region. Quickly display these relative differences with shading that ranges from light (less-frequent/lower) to dark (more-frequent/more). Filled maps are a great choice:

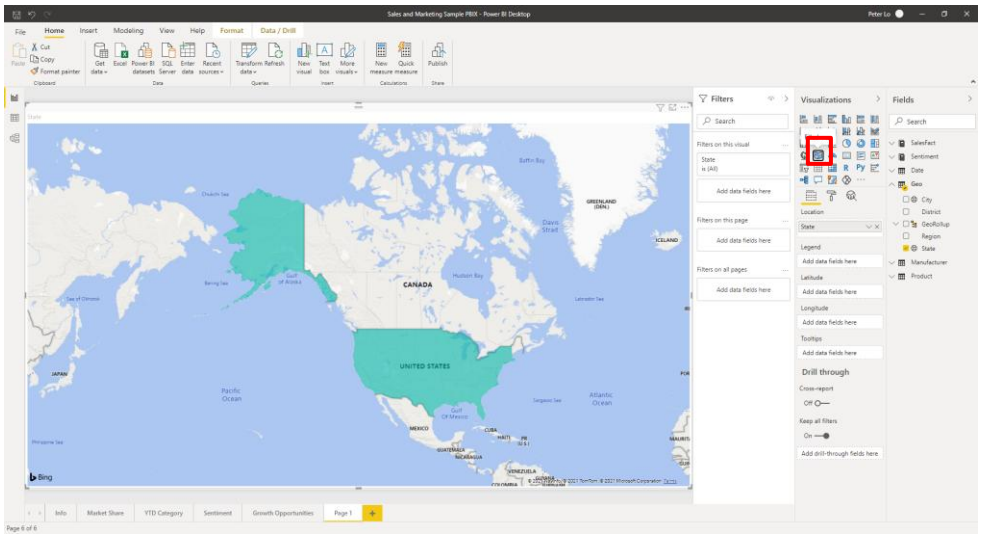
- To display quantitative information on a map
- To show spatial patterns and relationships
- When your data is standardized
- When working with socioeconomic data
- When defined regions are important
- To get an overview of the distribution across the geographic locations

In order to create a filled map:

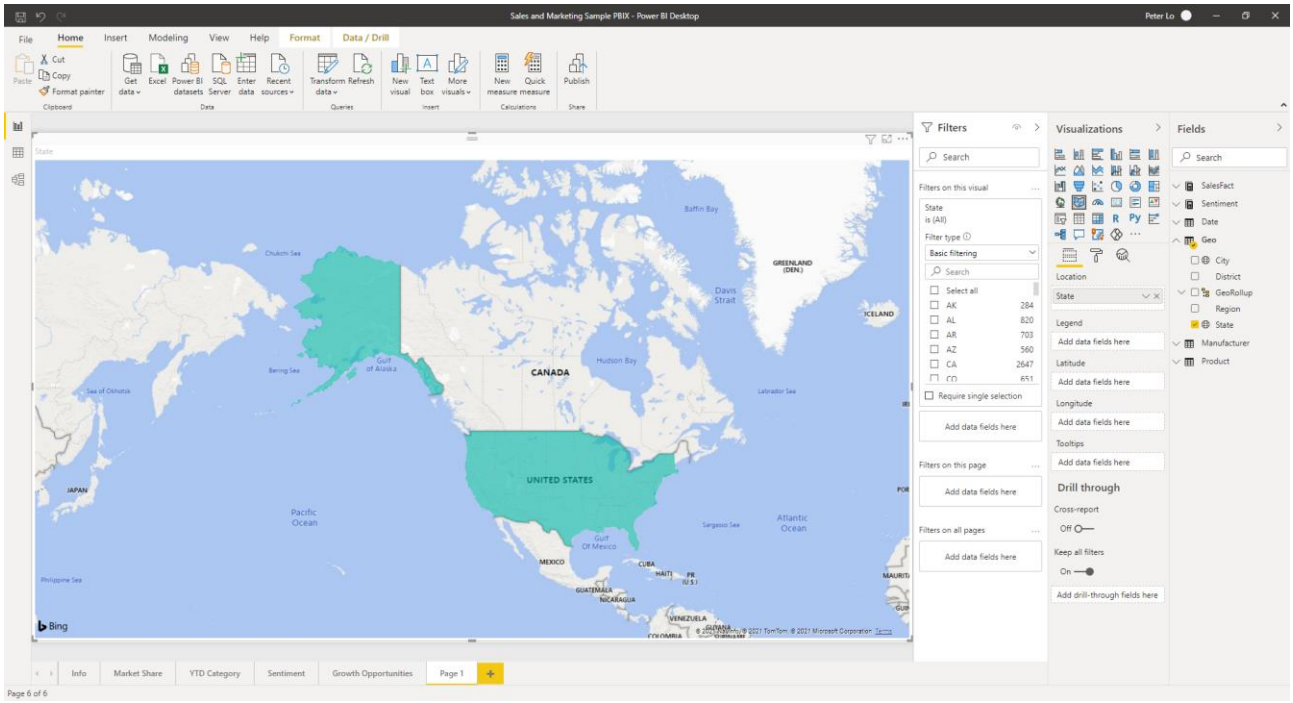
1. Select **File ➔ Open Report**, and open *Sales and Marketing Sample.pbix*.
2. Start on a blank report page and from the **Fields** pane, select *Geo ➔ State* field.



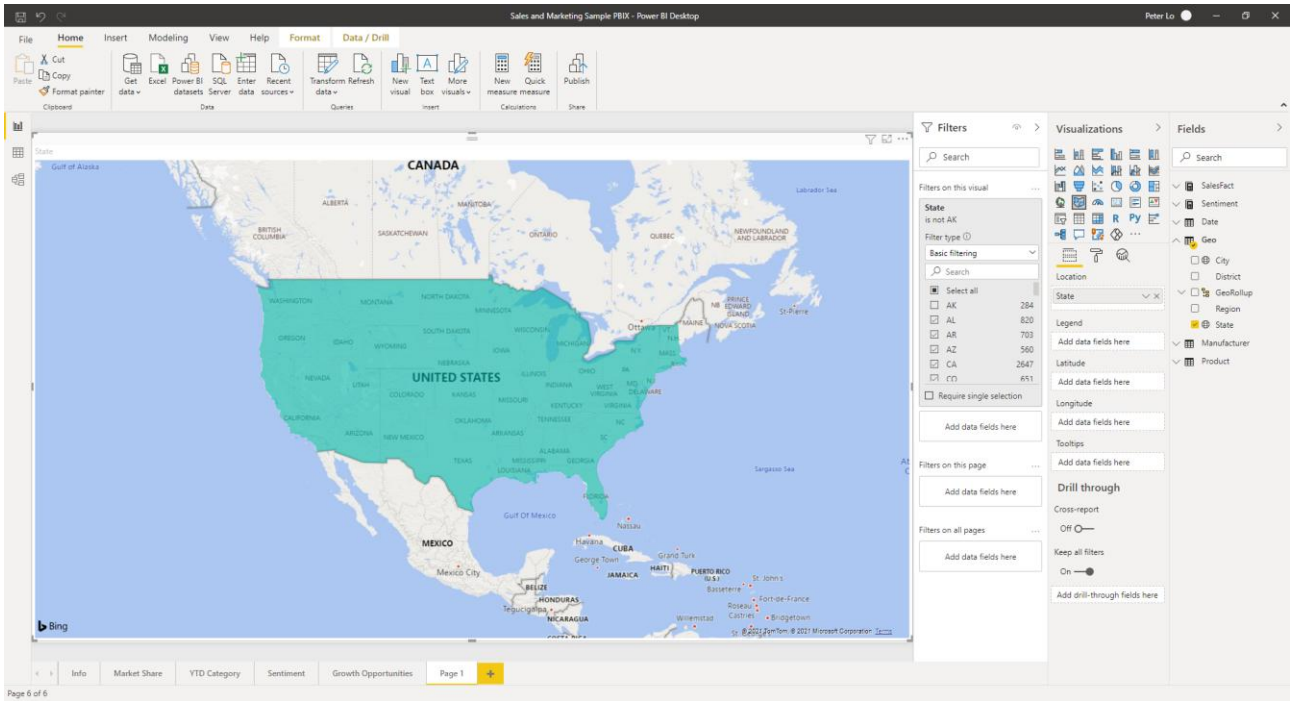
3. Convert the chart to a filled map. Notice that State is now in the Location well. Bing Maps uses the field in the Location well to create the map. The location can be a variety of valid locations: countries, states, counties, cities, zip codes, or other postal codes etc. Bing Maps provides filled map shapes for locations around the world. Without a valid entry in the Location well, Power BI cannot create the filled map.



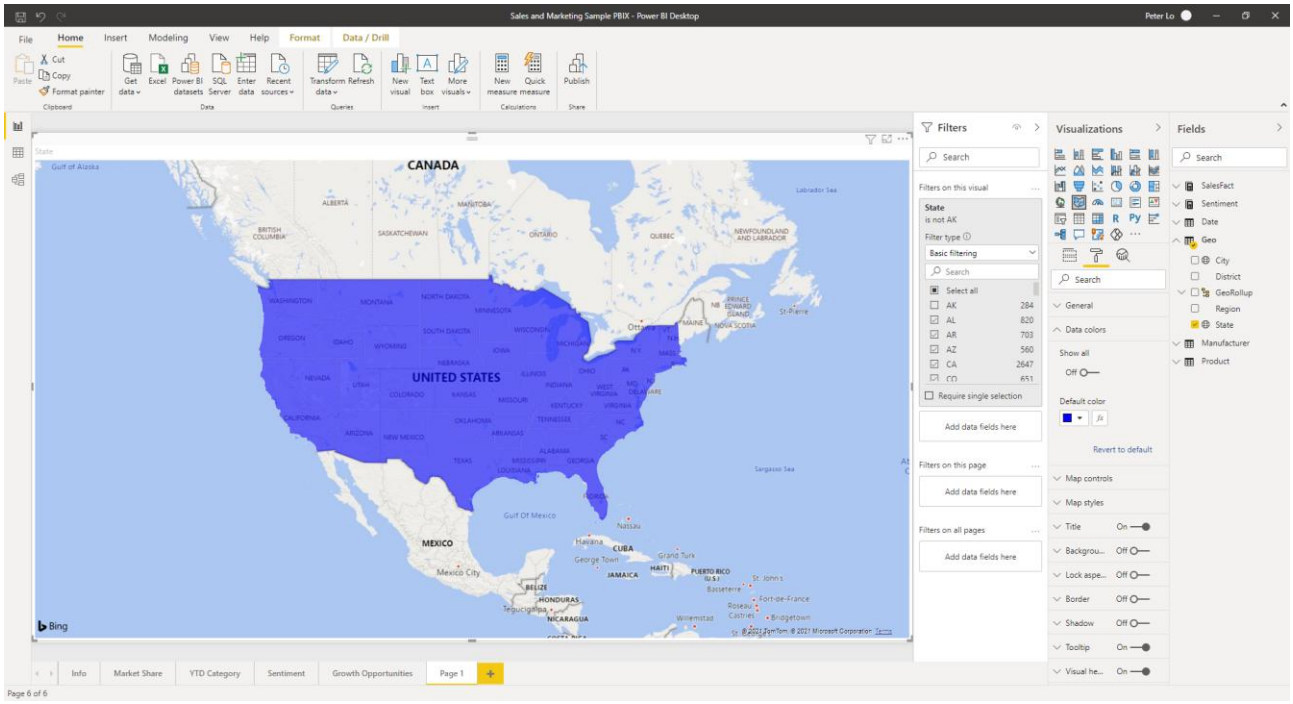
4. Filter the map to display only the continental United States. To the left of the **Visualizations** pane, look for the **Filters** pane. Expand it if it is minimized. Hover over **State** and select the expand chevron.



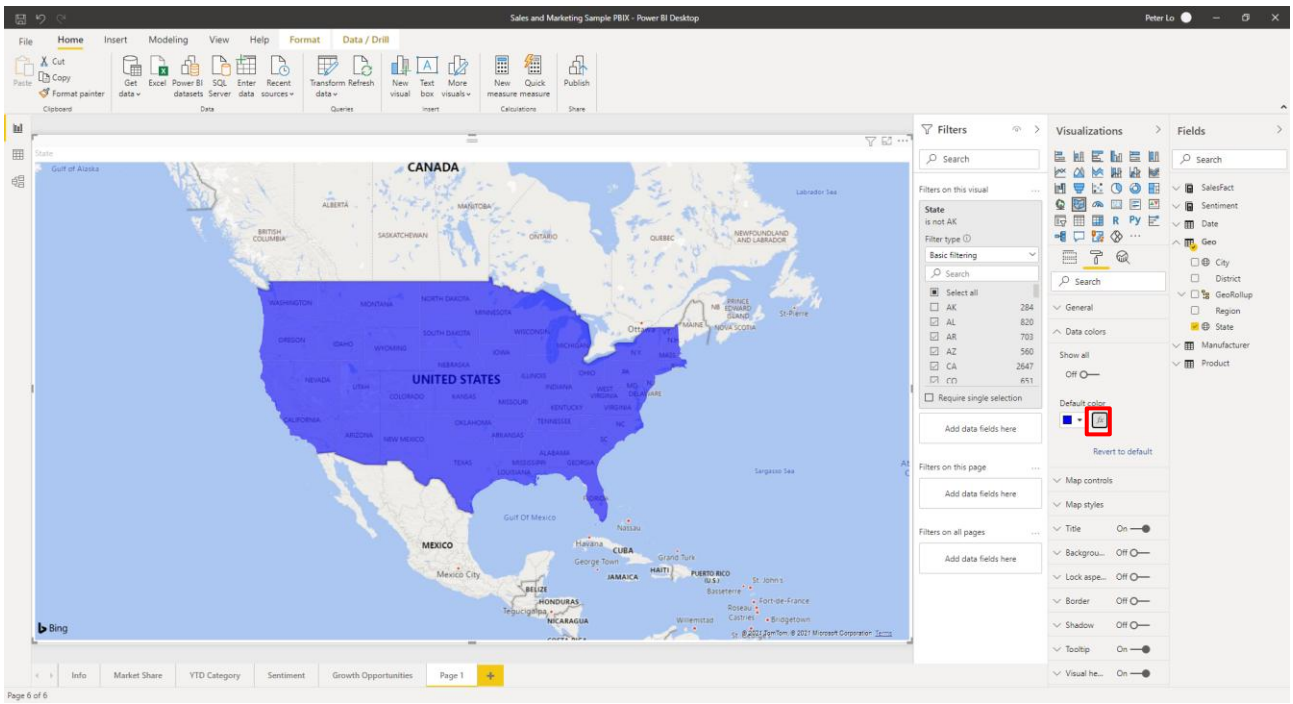
5. Place a check mark next to **All** and remove the check mark next to AK.



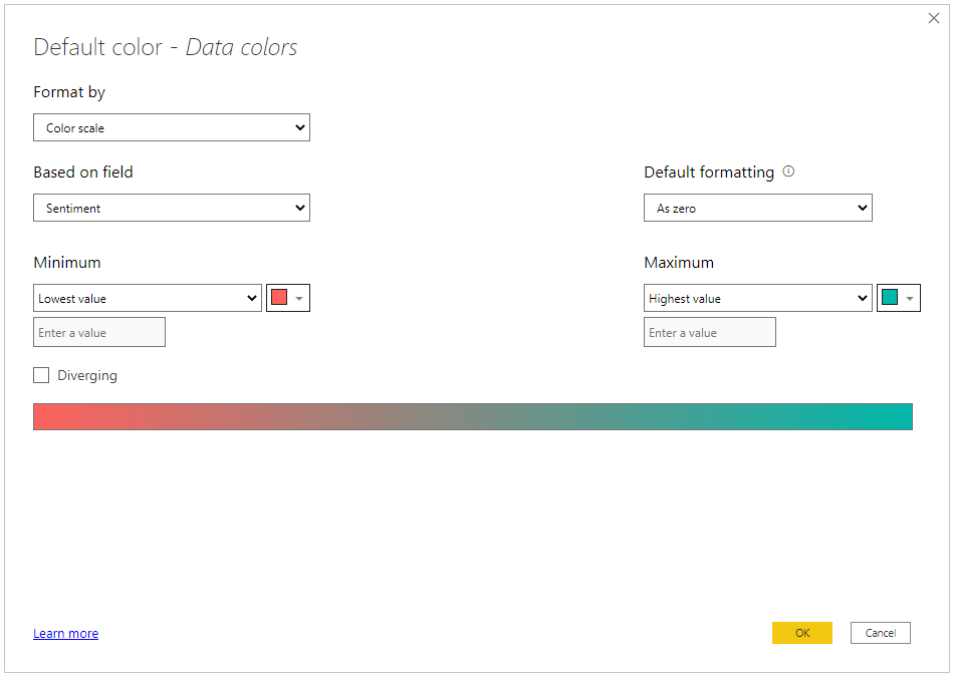
6. Select the paint roller icon to open the **Formatting** pane, and choose **Data colors**. Then change it to **Blue** (#0000FF).



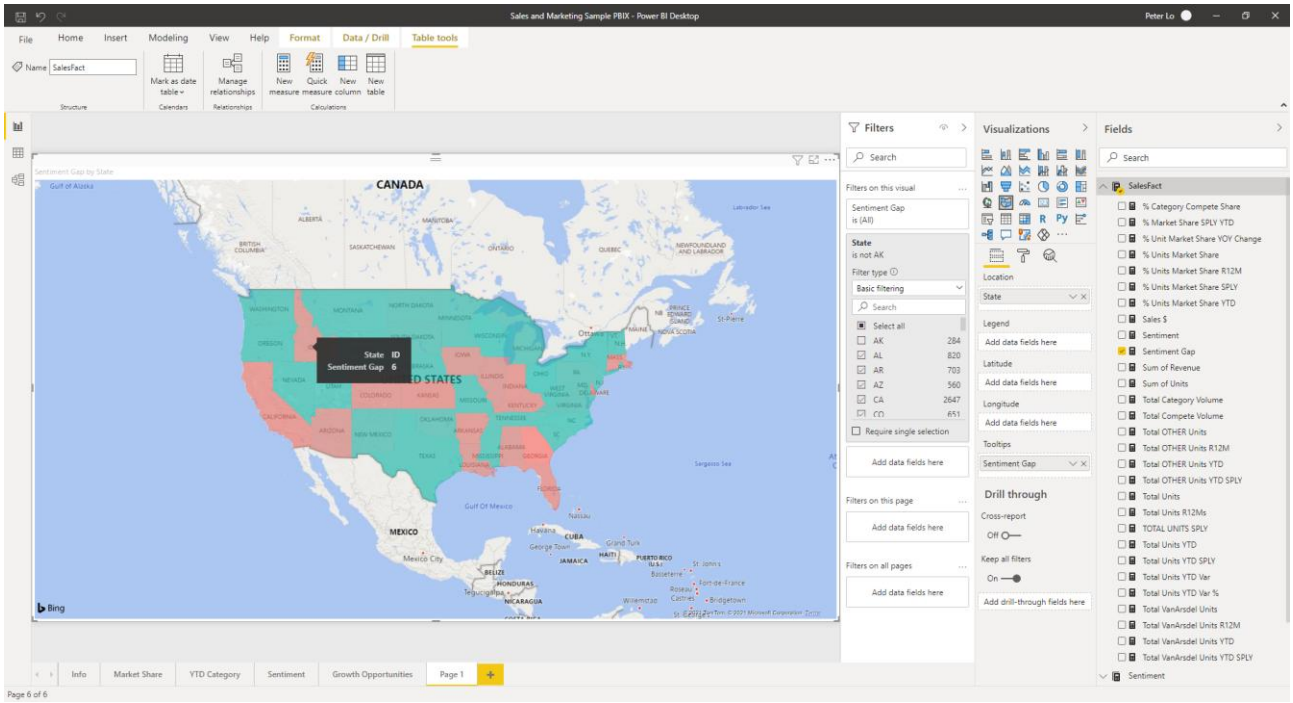
7. Click the **[*fx*]** button to choose Conditional formatting.



8. Use the **Default color – Data colors** screen to determine how your filled map will be shaded. The options available to you include which field to base the shading, and how to apply the shading. In this example we're using the field *SalesFact* ➔ *Sentiment*, and setting the lowest value for sentiment as orange and the highest value as blue. Values that fall between the maximum and minimum will be shades of orange and blue. The illustration at the bottom of the screen shows the range of colors that will be used.



9. The filled map is shaded in green and red, with red representing the lower sentiment numbers and green representing the higher, more-positive sentiment. To display additional detail, drag a field into the Tooltips well. Here we've added *SalesFact* ➔ *Sentiment gap*. Highlighting the state of Idaho (ID) shows us that sentiment gap is low, at 6.



4. Key Influencers

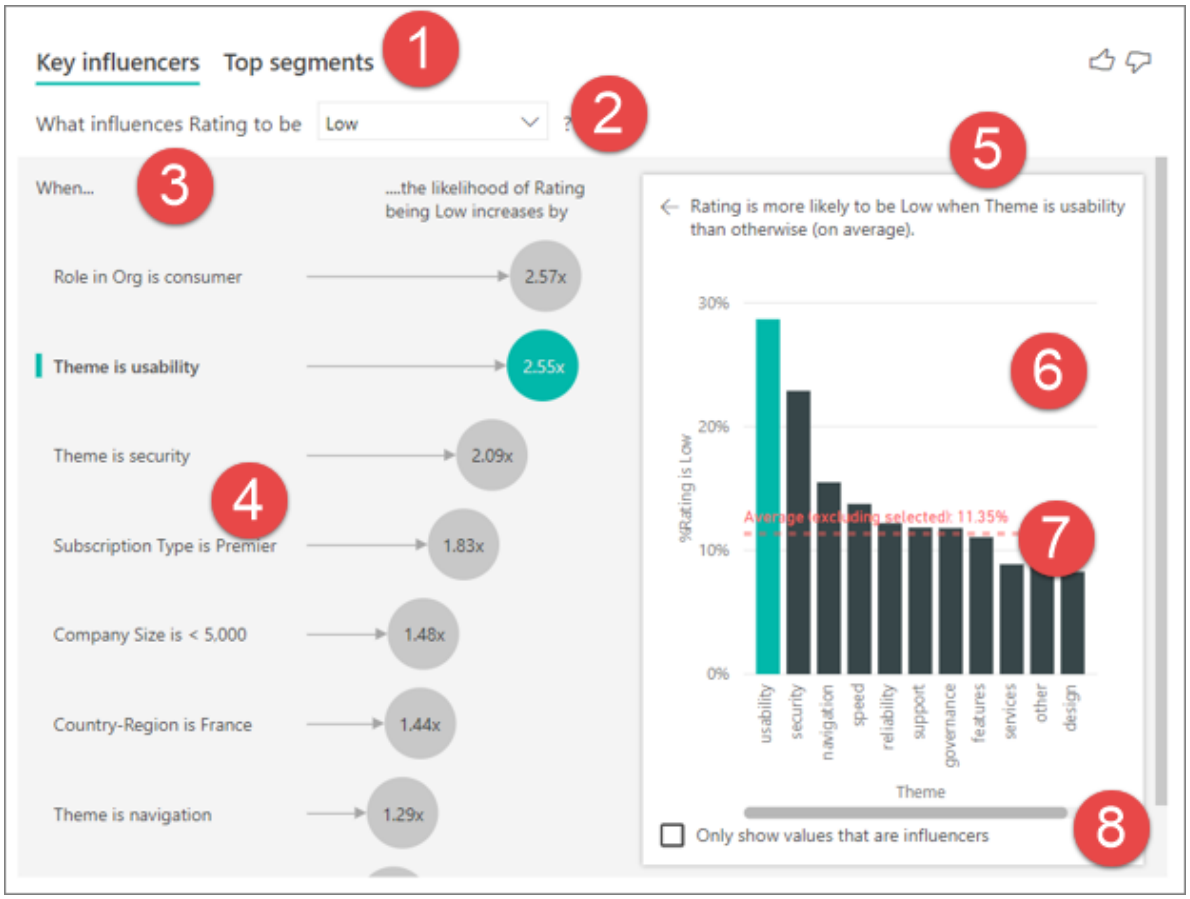
The key influencers visual helps you understand the factors that drive a metric you interested in. It analyzes your data, ranks the factors that matter, and displays them as key influencers. For example, suppose you want to figure out what influences employee turnover, which is also known as churn. One factor might be employment contract length, and another factor might be employee age.

4.1 When to use Key Influencers?

The key influencers visual is a great choice if you want to:

- See which factors affect the metric being analyzed.
- Contrast the relative importance of these factors. For example, do short-term contracts have more impact on churn than long-term contracts?

4.2 Features of the Key Influencers Visual



(1) **Tabs:** Select a tab to switch between views. Key influencers shows you the top contributors to the selected metric value. Top segments shows you the top segments that contribute to the selected metric value. A segment is made up of a combination of values. For example, one segment might be consumers who have been customers for at least 20 years and live in the west region.

(2) **Drop-down box:** The value of the metric under investigation. In this example, look at the metric Rating. The selected value is Low.

(3) **Restatement:** It helps you interpret the visual in the left pane.

(4) **Left pane:** The left pane contains one visual. In this case, the left pane shows a list of the top key influencers.

(5) **Restatement:** It helps you interpret the visual in the right pane.

(6) **Right pane:** The right pane contains one visual. In this case, the column chart displays all the values for the key influencer Theme that was selected in the left pane. The specific value of usability from the left pane is shown in green. All the other values for Theme are shown in black.

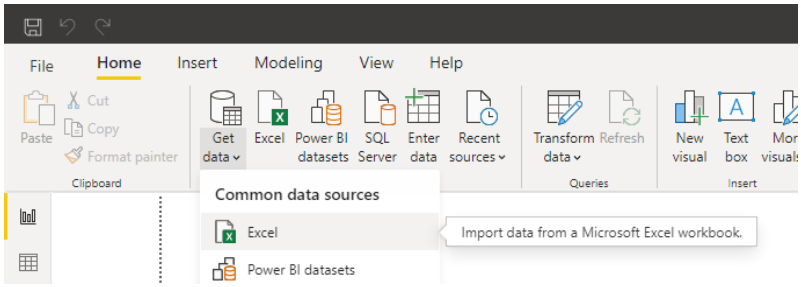
(7) **Average line:** The average is calculated for all possible values for Theme except usability (which is the selected influencer). So the calculation applies to all the values in black. It tells you what percentage of the other Themes had a low rating. In this case 11.35% had a low rating (shown by the dotted line).

(8) **Check box:** Filters out the visual in the right pane to only show values that are influencers for that field. In this example, this would filter the visual to usability, security and navigation.

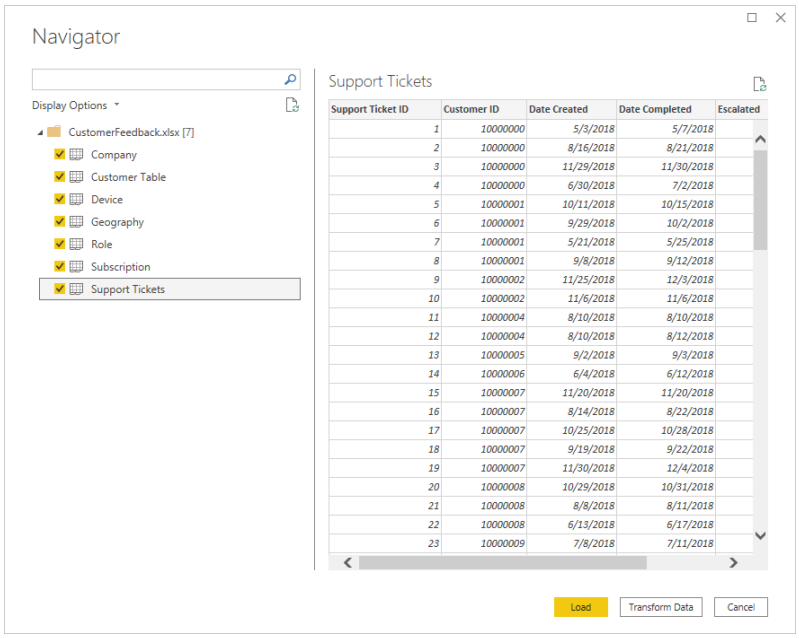
4.3 Analyze a Metric that is Categorical

Your Product Manager wants you to figure out which factors lead customers to leave negative reviews about your cloud service.

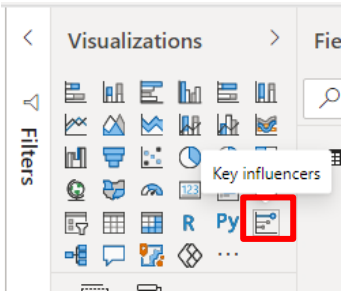
1. Select **Get Data ➔ Excel** in the **Home** ribbon tab, and open *CustomerFeedback.xlsx* workbook.



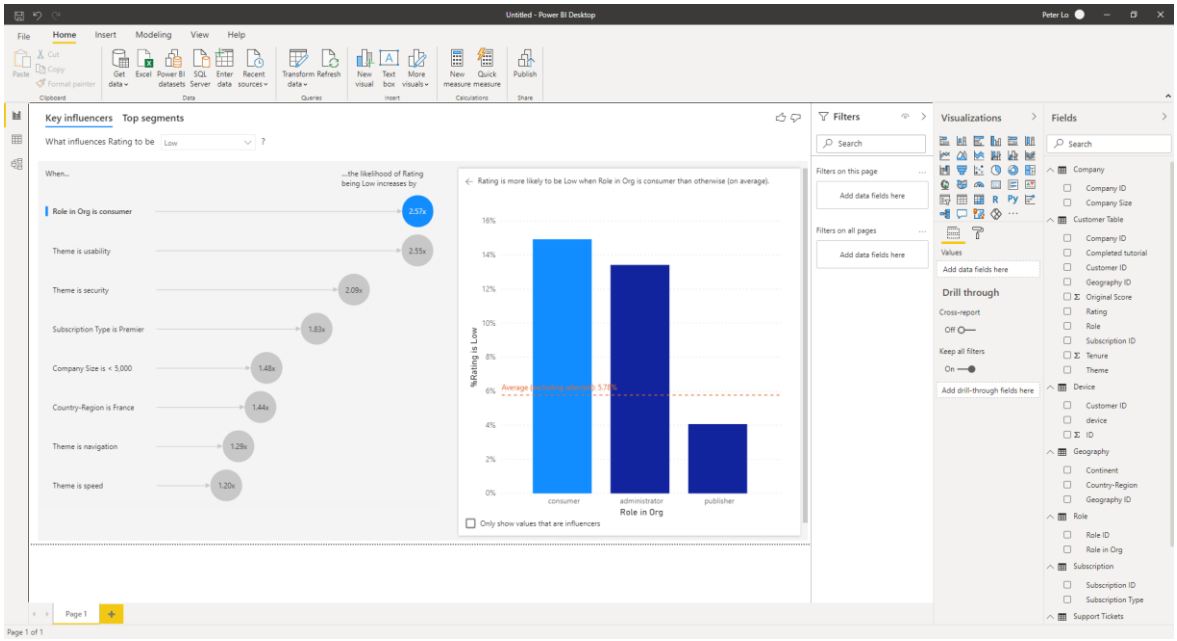
2. Select all and press **[Load]**.



3. Select the **Key influencers** icon.



4. Move the metric you want to investigate into the Analyze field. To see what drives a customer rating of the service to be low, select *Customer Table* ➔ *Rating*.
5. Move fields that you think might influence Rating into the Explain by field.
   - *Geography* ➔ *Country-Region*
   - *Role* ➔ *Role in Org*
   - *Subscription* ➔ *Subscription Type*
   - *Company* ➔ *Company Size*
   - *Customer Table* ➔ *Theme*
6. Leave the **Expand by** field empty. This field is only used when analyzing a measure or summarized field.
7. To focus on the negative ratings, select **Low** in the **What influences Rating to be** drop-down box.



The analysis runs on the table level of the field that's being analyzed. In this case, it's the Rating metric. This metric is defined at a customer level. Each customer has given either a high score or a low score. All the explanatory factors must be defined at the customer level for the visual to make use of them.

In this example, all of the explanatory factors have either a one-to-one or a many-to-one relationship with the metric. In this case, each customer assigned a single theme to their rating. Similarly, customers come from one country, have one membership type, and perform one role in

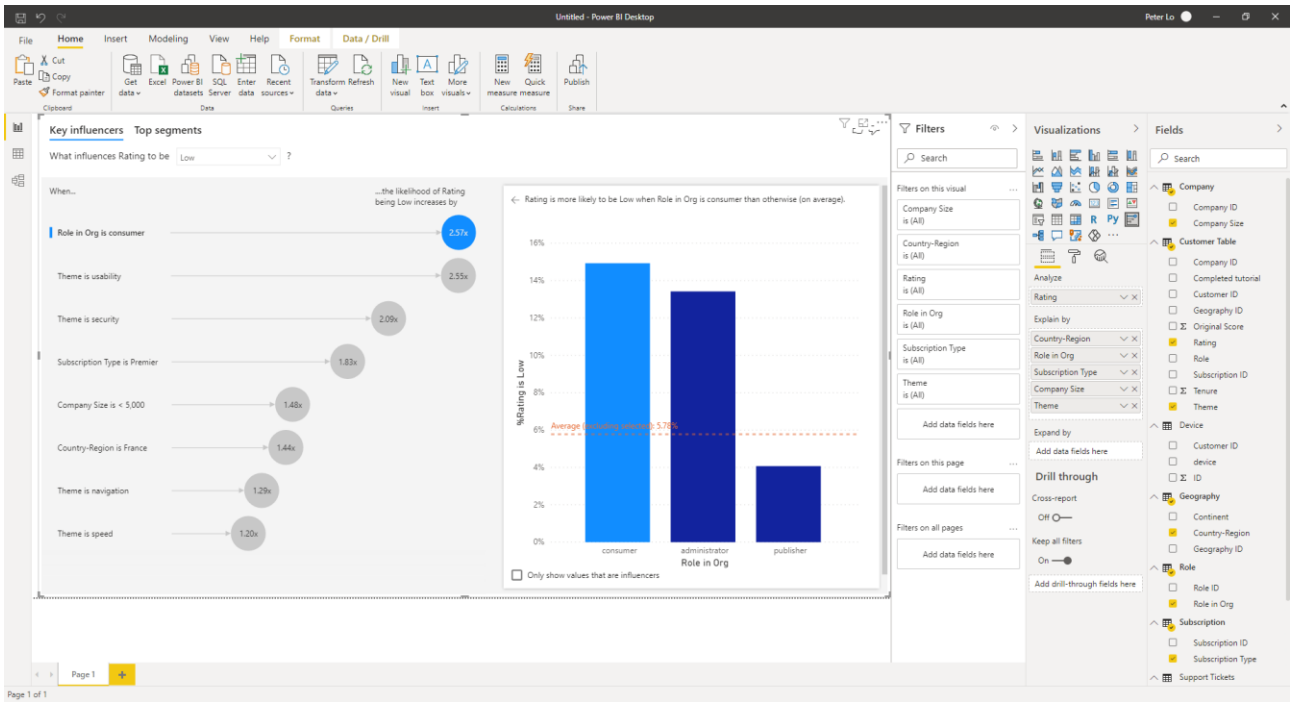
their organization. The explanatory factors are already attributes of a customer, and no transformations are needed. The visual can make immediate use of them.

Measures and aggregates used as explanatory factors are also evaluated at the table level of the Analyze metric.

4.4 Interpret Categorical Key Influencers

4.4.1 Top Single Factor that Influences the Likelihood of a Low Rating

The customer in this example can have three roles: consumer, administrator, and publisher. Being a consumer is the top factor that contributes to a low rating.

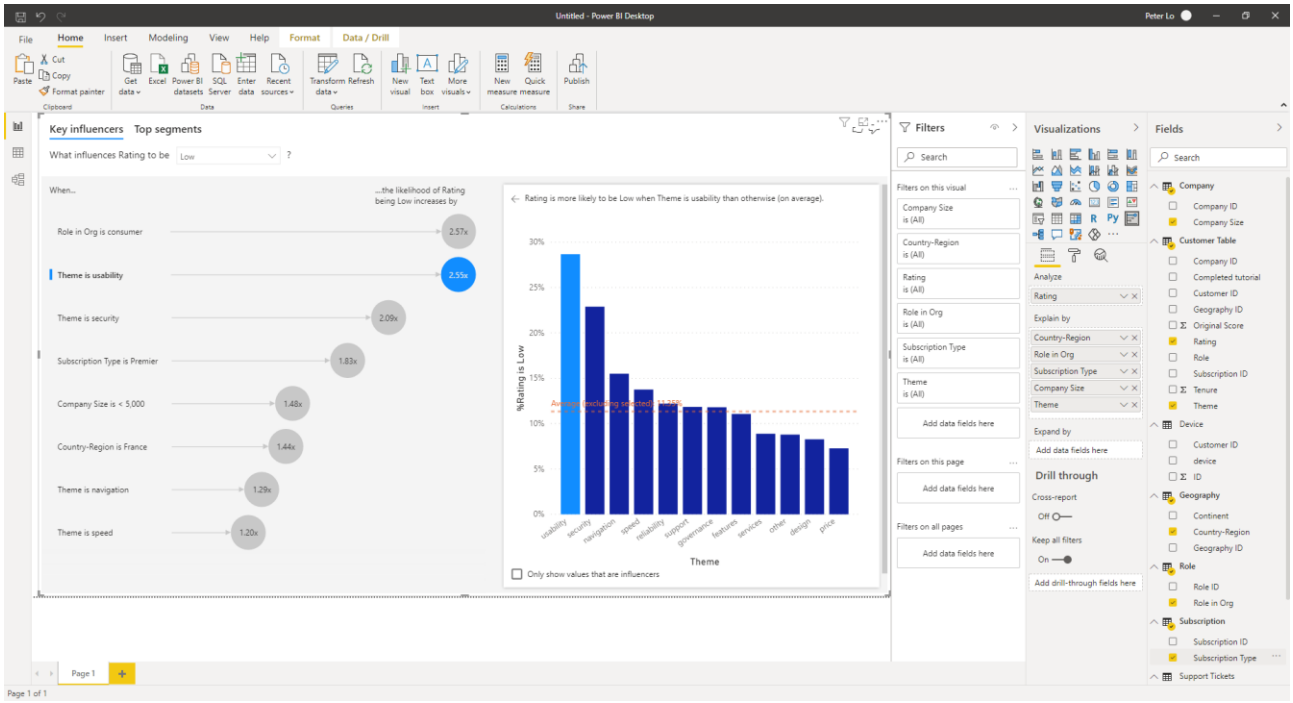


More precisely, your consumers are 2.57 times more likely to give your service a negative score. The key influencers chart lists Role in Org is consumer first in the list on the left. By selecting Role in Org is consumer, Power BI shows additional details in the right pane. The comparative effect of each role on the likelihood of a low rating is shown.

- 14.93% of consumers give a low score.
- On average, all other roles give a low score 5.78% of the time.
- Consumers are 2.57 times more likely to give a low score compared to all other roles. You can determine this by dividing the green bar by the red dotted line.

4.4.2 Second Single Factor that Influences the Likelihood of a Low Rating

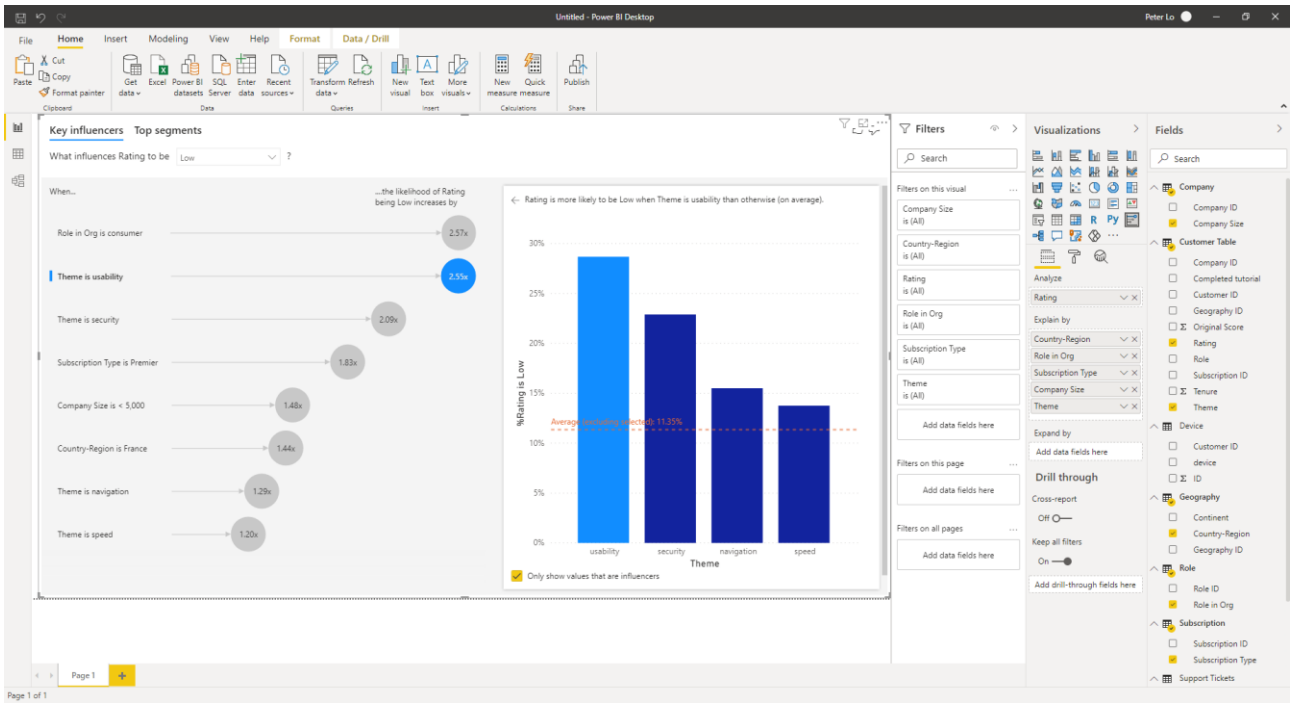
The key influencers visual compares and ranks factors from many different variables. The second influencer has nothing to do with *Role in Org*. Select the second influencer in the list, which is *Theme is usability*.



The second most important factor is related to the theme of the customer's review. Customers who commented about the usability of the product were 2.55 times more likely to give a low score compared to customers who commented on other themes, such as reliability, design, or speed.

Between the visuals, the average, which is shown by the red dotted line, changed from 5.78% to 11.34%. The average is dynamic because it's based on the average of all other values. For the first influencer, the average excluded the customer role. For the second influencer, it excluded the usability theme.

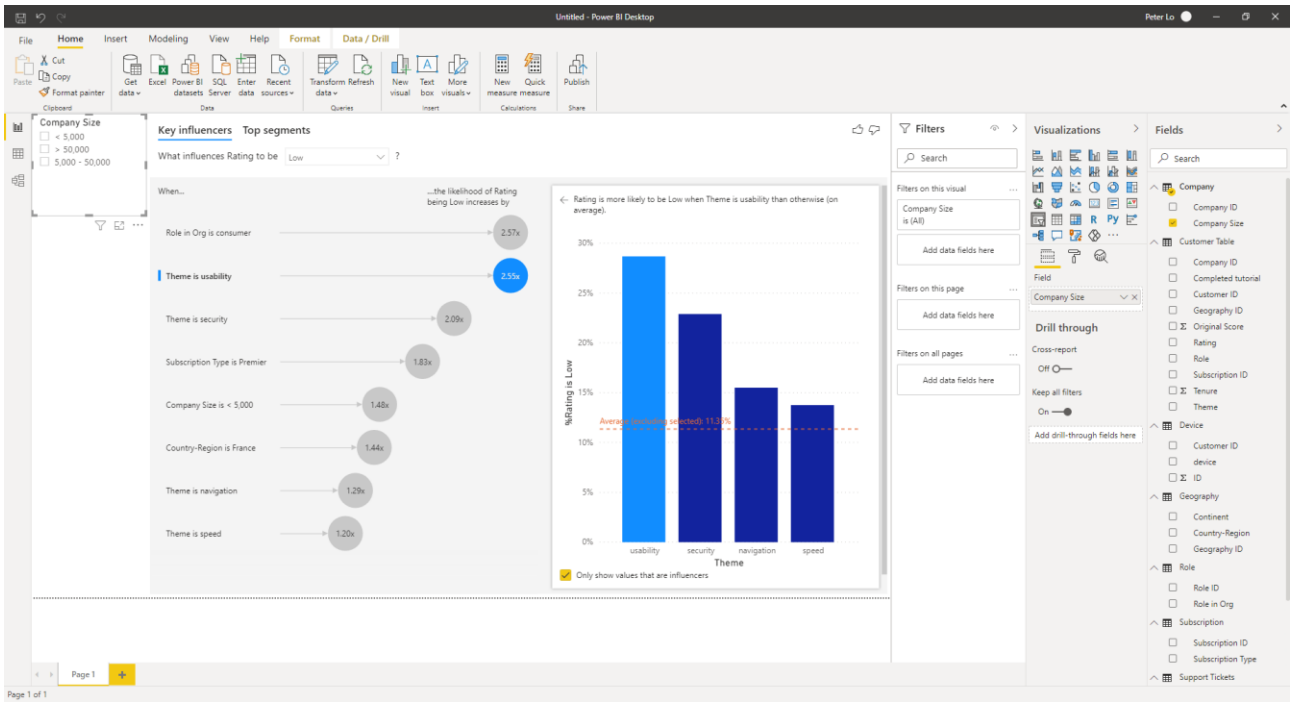
Select the **Only show values that are influencers** check box to filter by using only the influential values. In this case, they're the roles that drive a low score. Twelve themes are reduced to the four that Power BI identified as the themes that drive low ratings.



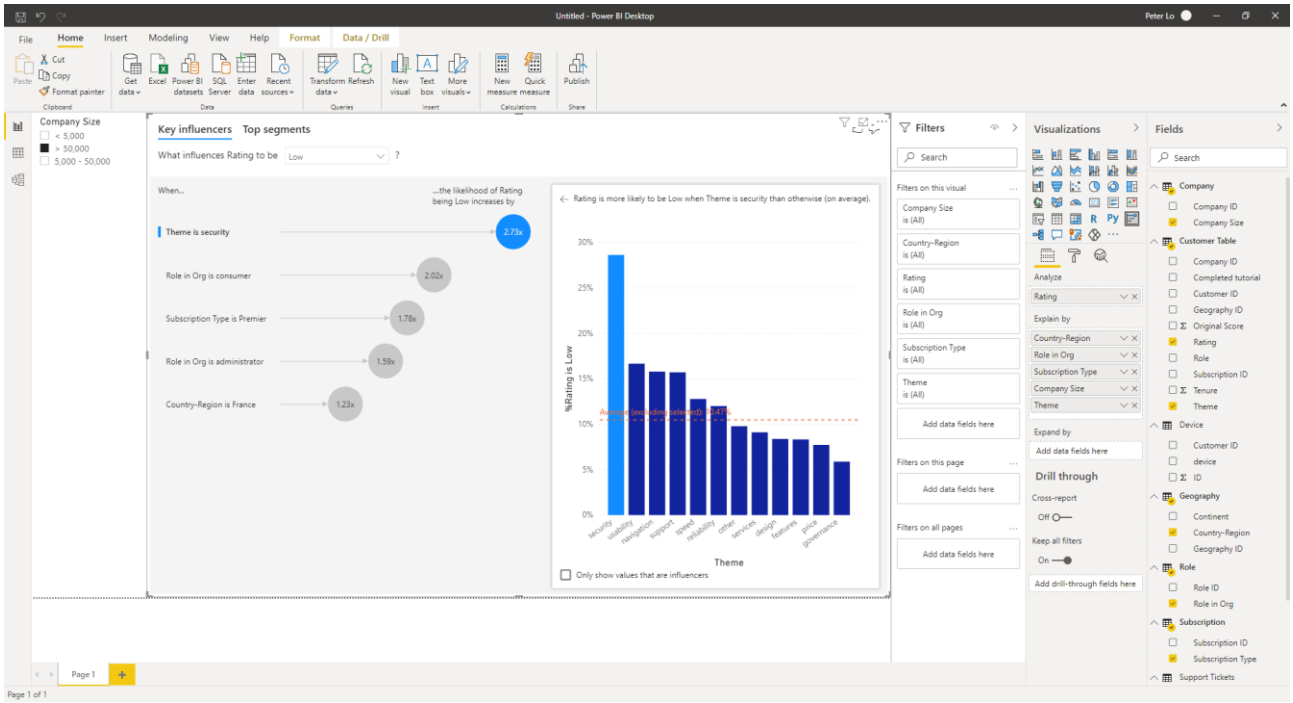
4.4.3 Interact with Other Visuals

Every time you select a slicer, filter, or other visual on the canvas, the key influencers visual reruns its analysis on the new portion of data.

1. Move Company Size into the report and use it as a slicer.



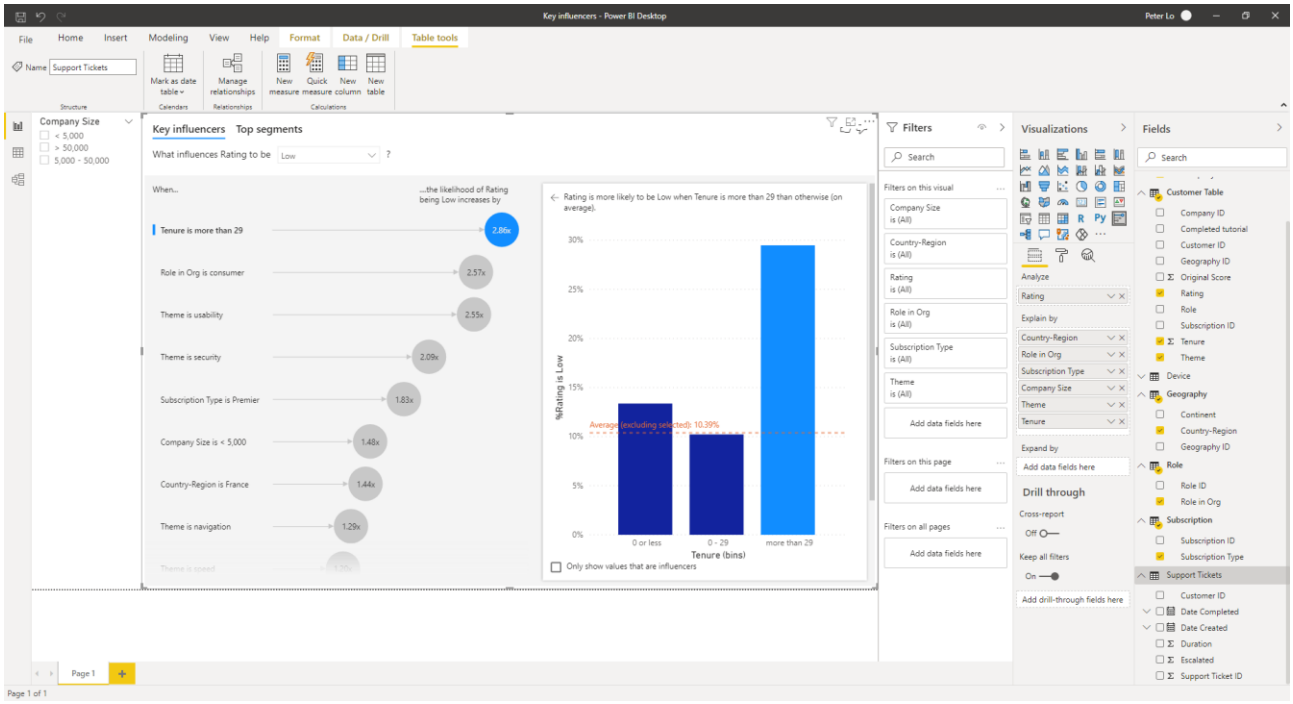
2. Use it to see if the key influencers for your enterprise customers are different than the general population. An enterprise company size is larger than 50,000 employees. Selecting >50,000 reruns the analysis, and you can see that the influencers changed. For large enterprise customers, the top influencer for low ratings has a theme related to security. You might want to investigate further to see if there are specific security features your large customers are unhappy about.



4.4.4 Interpret Continuous Key Influencers

So far, you've seen how to use the visual to explore how different categorical fields influence low ratings. It's also possible to have continuous factors such as age, height, and price in the **Explain by** field. Let's look at what happens when *Customer Table* ➔ *Tenure* is moved into **Explain by**. Tenure depicts how long a customer has used the service.

As tenure increases, the likelihood of receiving a lower rating also increases. This trend suggests that the longer-term customers are more likely to give a negative score. This insight is interesting, and one that you might want to follow up on later.

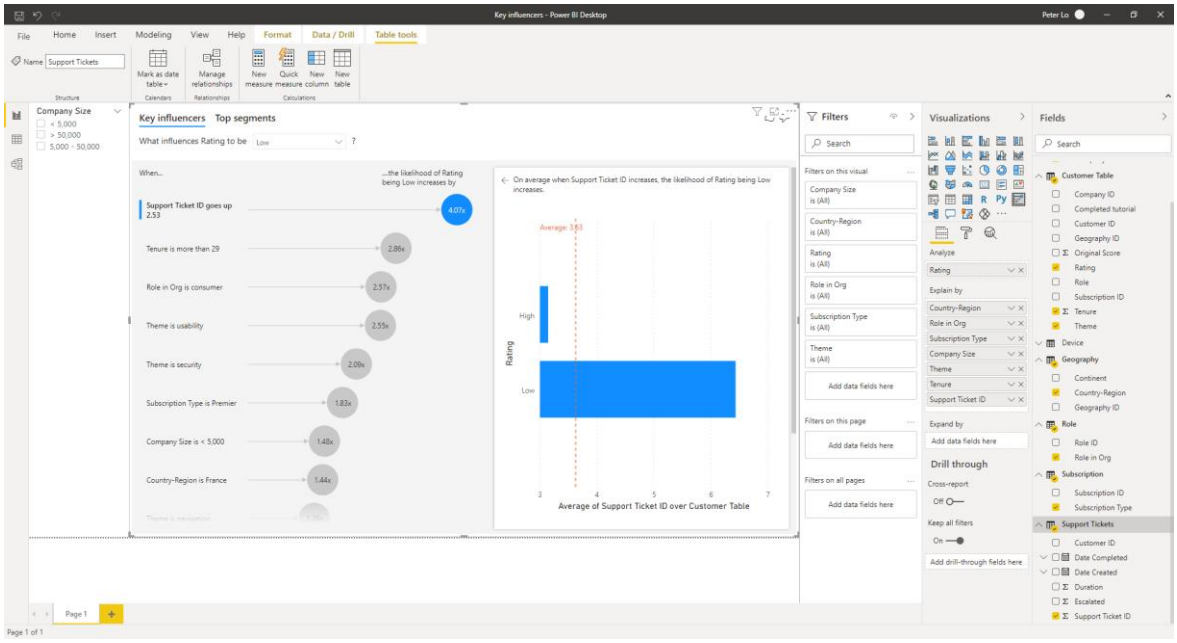


4.4.5 Interpret Measures and Aggregates as Key Influencers

You can use measures and aggregates as explanatory factors inside your analysis. For example, you might want to see what effect the count of customer support tickets or the average duration of an open ticket has on the score you receive.

You want to see if the number of support tickets that a customer has influences the score they give by bring in *Support Ticket* ➔ *Support Ticket ID*. Because a customer can have multiple support tickets, you aggregate the ID to the customer level. Aggregation is important because the analysis runs on the customer level, so all drivers must be defined at that level of granularity.

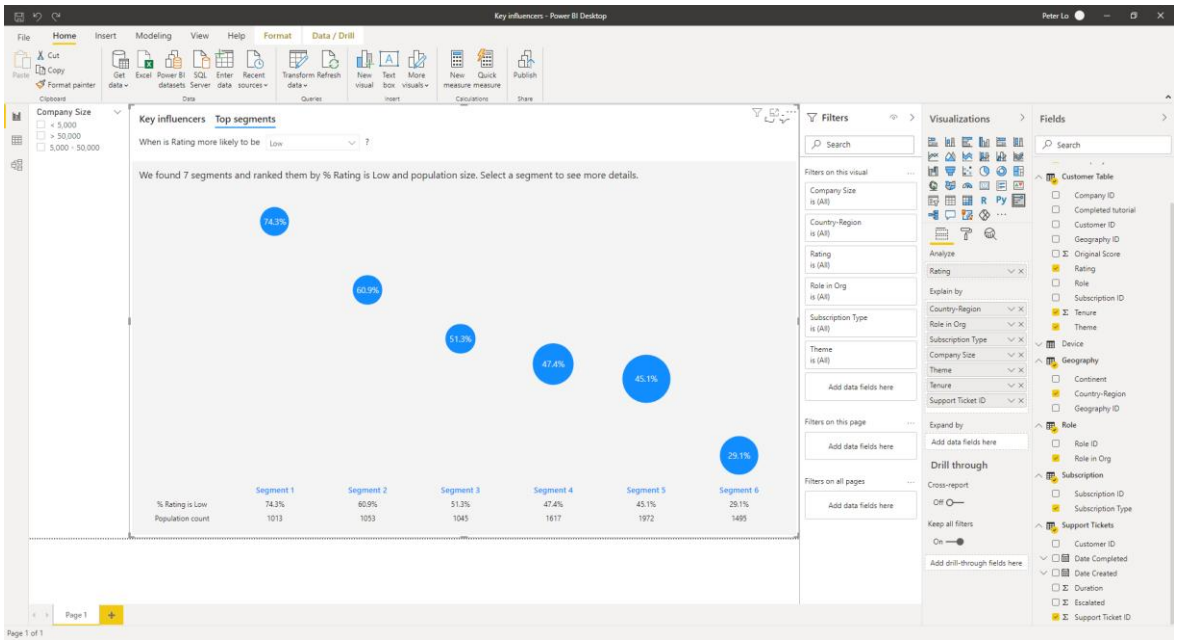
Let's look at the count of IDs. Each customer row has a count of support tickets associated with it. In this case, as the count of support tickets increases, the likelihood of the rating being low goes up 4.07 times. The visual on the right shows the average number of support tickets by different Rating values evaluated at the customer level.



4.4.6 Interpret the Results: Top Segments

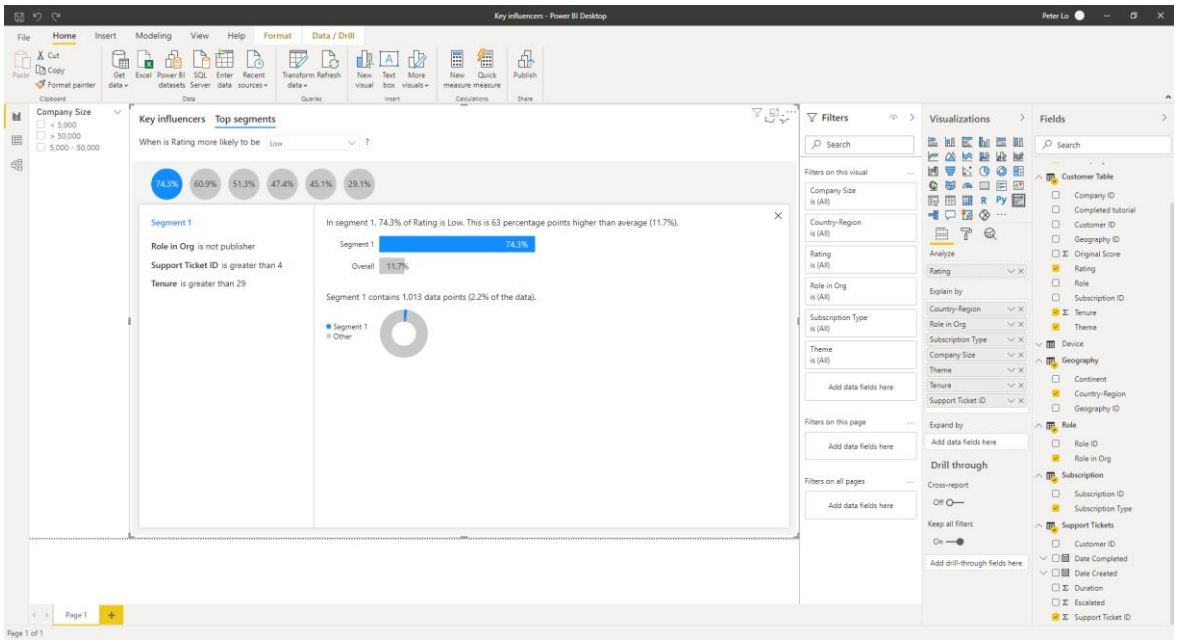
You can use the Key influencers tab to assess each factor individually. You also can use the Top segments tab to see how a combination of factors affects the metric that you're analyzing.

Top segments initially show an overview of all the segments that Power BI discovered. These segments are ranked by the percentage of low ratings within the segment. Segment 1, for example, has 74.3% customer ratings that are low. The higher the bubble, the higher the proportion of low ratings. The size of the bubble represents how many customers are within the segment.



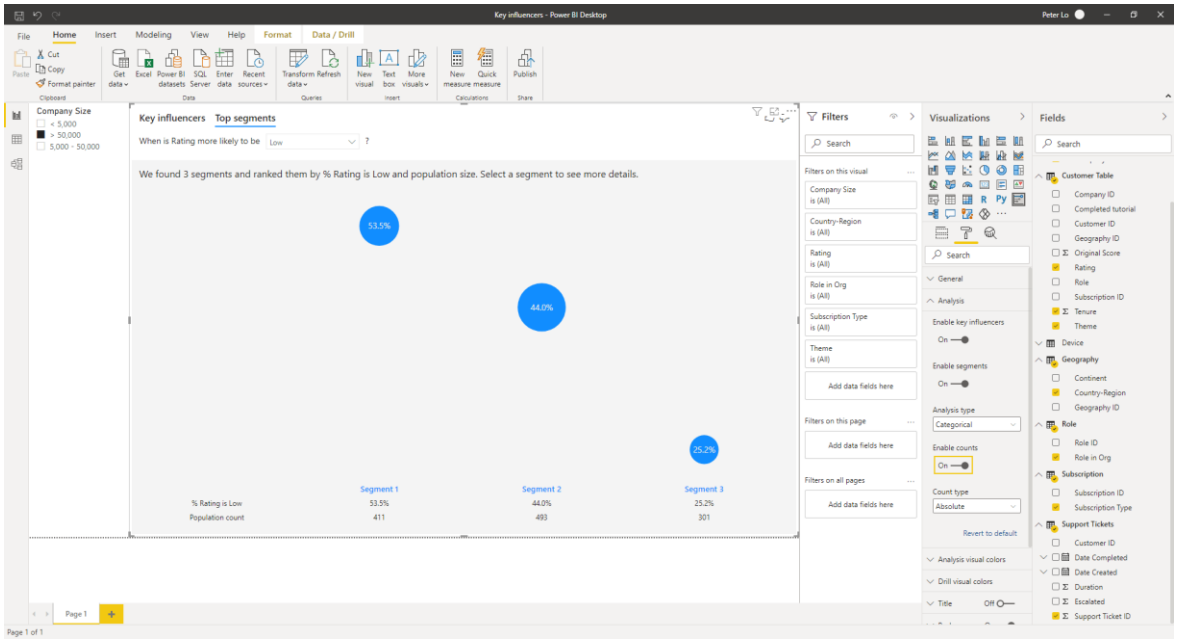
Selecting a bubble displays the details of that segment. If you select Segment 1, for example, you find that it's made up of relatively established customers. They've been customers for over 29 months and have more than four support tickets. Finally, they're not publishers, so they're either consumers or administrators.

In this group, 74.3% of the customers gave a low rating. The average customer gave a low rating 11.7% of the time, so this segment has a larger proportion of low ratings. It's 63 percentage points higher. Segment 1 also contains approximately 2.2% of the data, so it represents an addressable portion of the population.

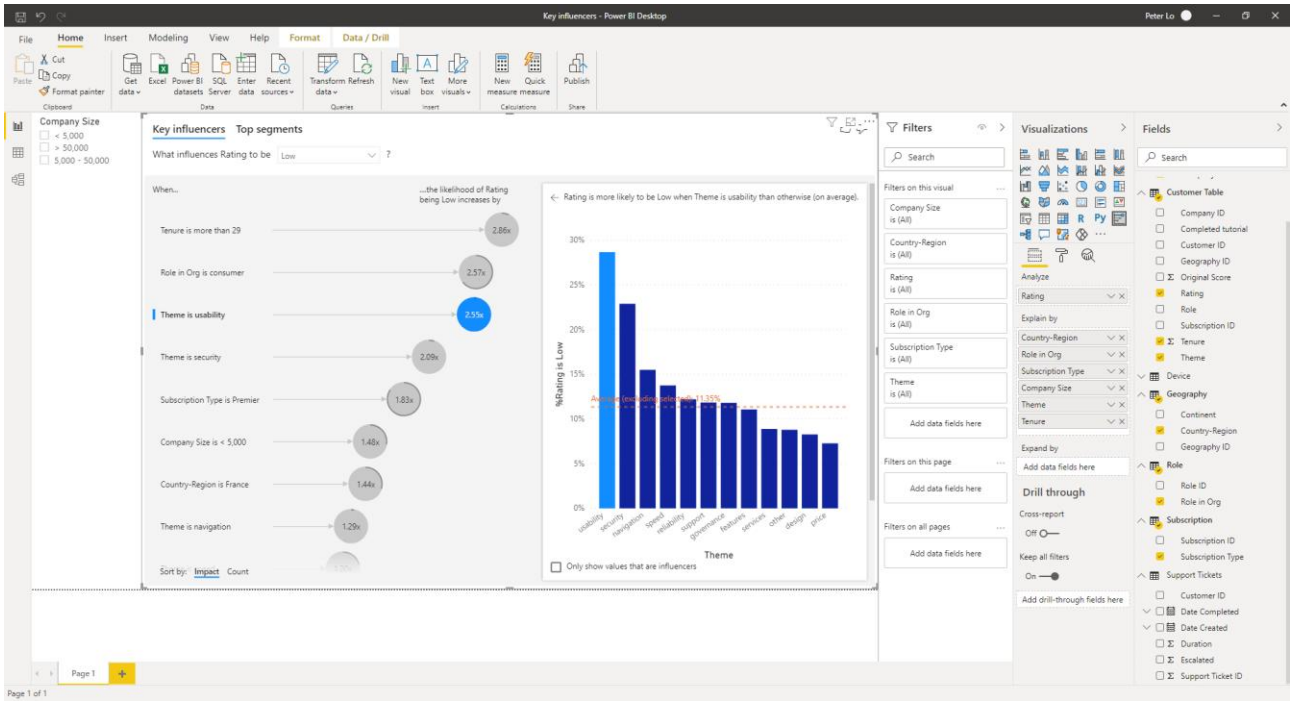


4.5 Adding Counts

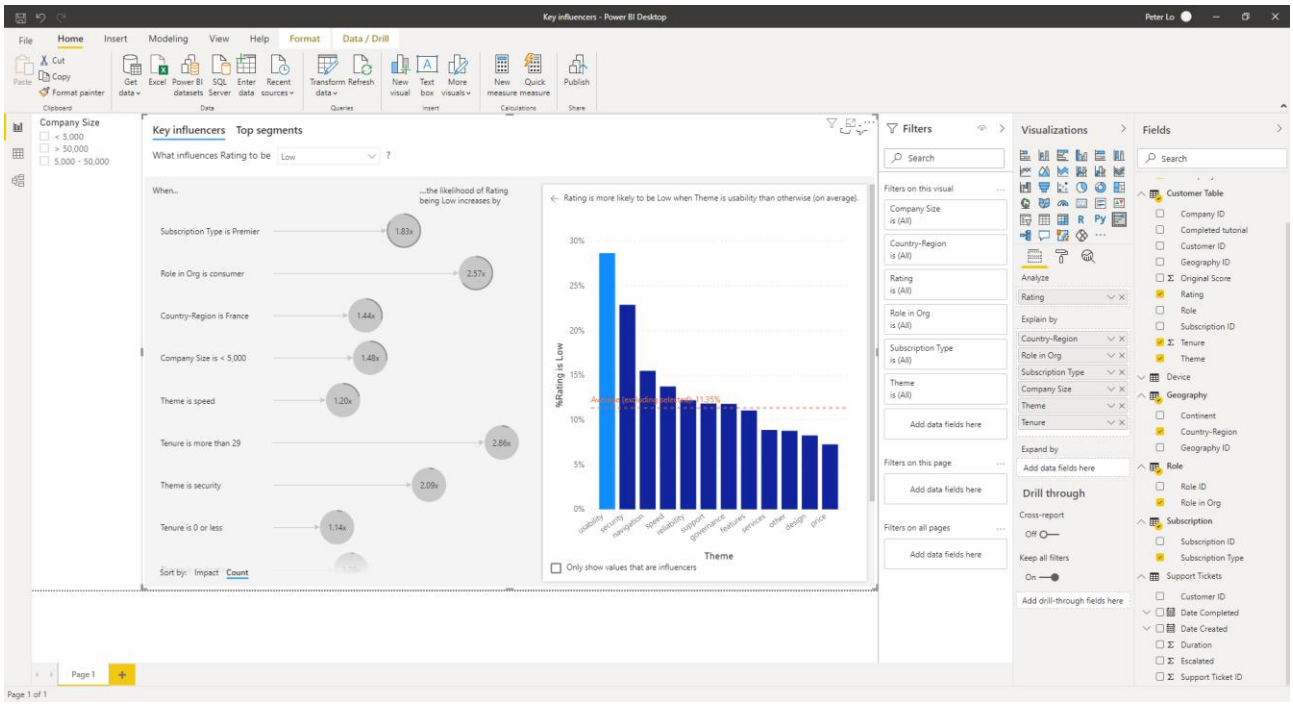
Sometimes an influencer can have a big impact but represent very little of the data. For example, Theme is usability is the second biggest influencer for low ratings. However, there might have only been a handful of customers who complained about usability. Counts can help you prioritize which influencers you want to focus on. You can turn counts on through the Analysis card of the formatting pane.



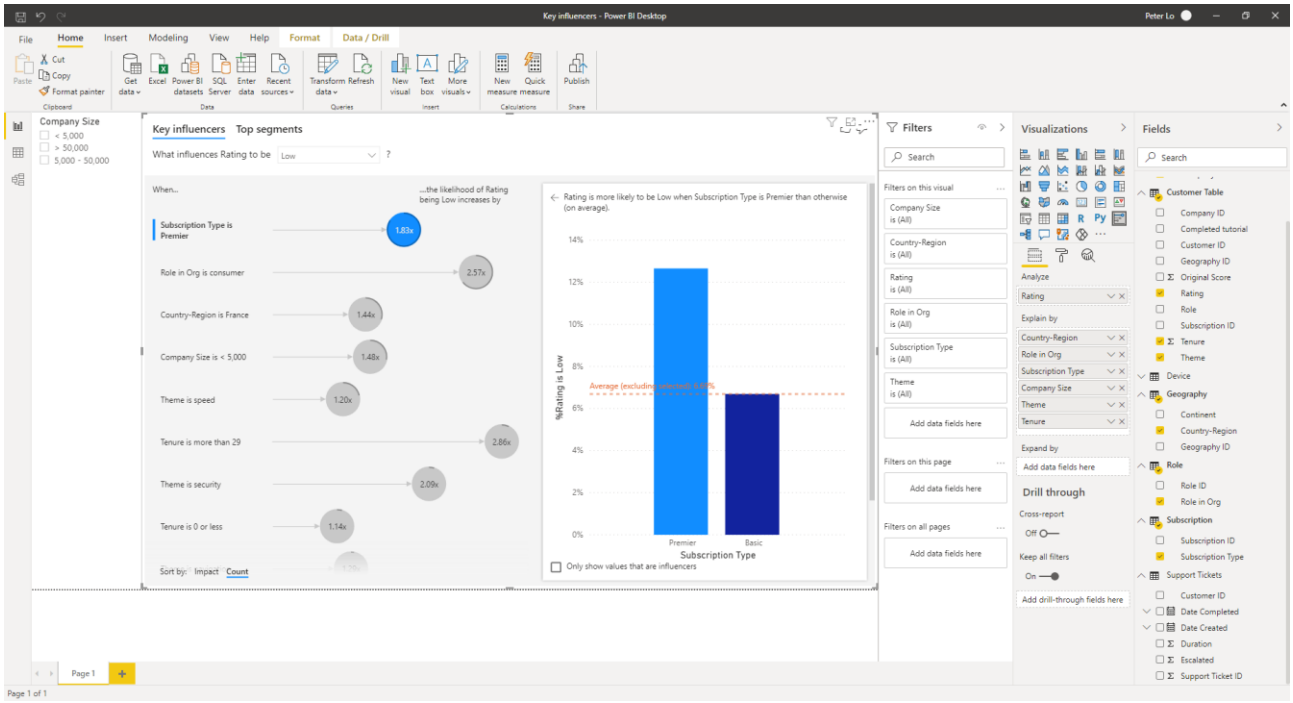
Once counts are turned on, you'll see a ring around each influencer's bubble, which represents the approximate percentage of data that influencer contains. The more of the bubble the ring circles, the more data it contains. We can see that Theme is usability contains a very small proportion of data.



You can also use the Sort by toggle in the bottom left of the visual to sort the bubbles by count first instead of impact. *Subscription Type* is *Premier* is the top influencer based on count.

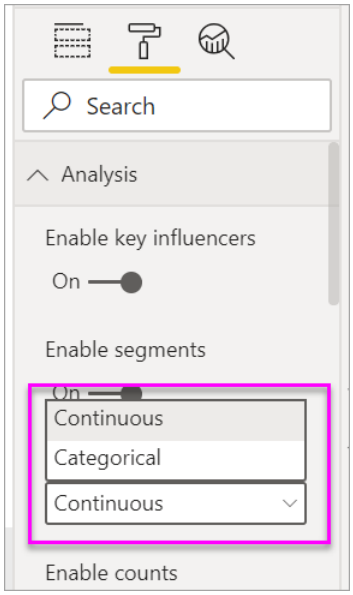


Having a full ring around the circle means the influencer contains 100% of the data. You can change the count type to be relative to the maximum influencer using the Count type dropdown in the Analysis card of the formatting pane. Now the influencer with the most amount of data will be represented by a full ring and all other counts will be relative to it.



4.5.1 Analyze a Metric that is Numeric

If you move an un-summarized numerical field into the **Analyze** field, you have a choice how to handle that scenario. You can change the behavior of the visual by going into the **Formatting Pane** and switching between **Categorical Analysis Type** and **Continuous Analysis Type**.

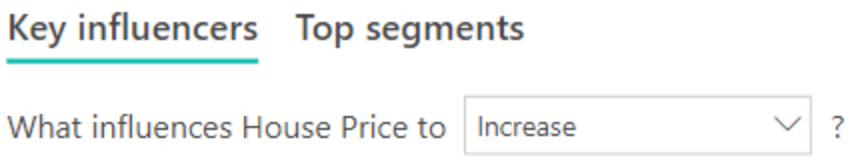


- A **Categorical Analysis Type** behaves as described above. For instance, if you were looking at survey scores ranging from 1 to 10, you could ask 'What influences Survey Scores to be 1?'
- A **Continuous Analysis Type** changes the question to a continuous one. In the example above, our new question would be 'What influences Survey Scores to increase/decrease?'

This distinction is very helpful when you have lots of unique values in the field you are analyzing. In the example below we look at house prices. It is not very meaningful to ask 'What influences House Price to be 156,214?' as that is very specific and we are likely not to have enough data to

infer a pattern.

Instead we may want to ask, 'What influences House Price to increase'? which allows us to treat house prices as a range rather than distinct values.

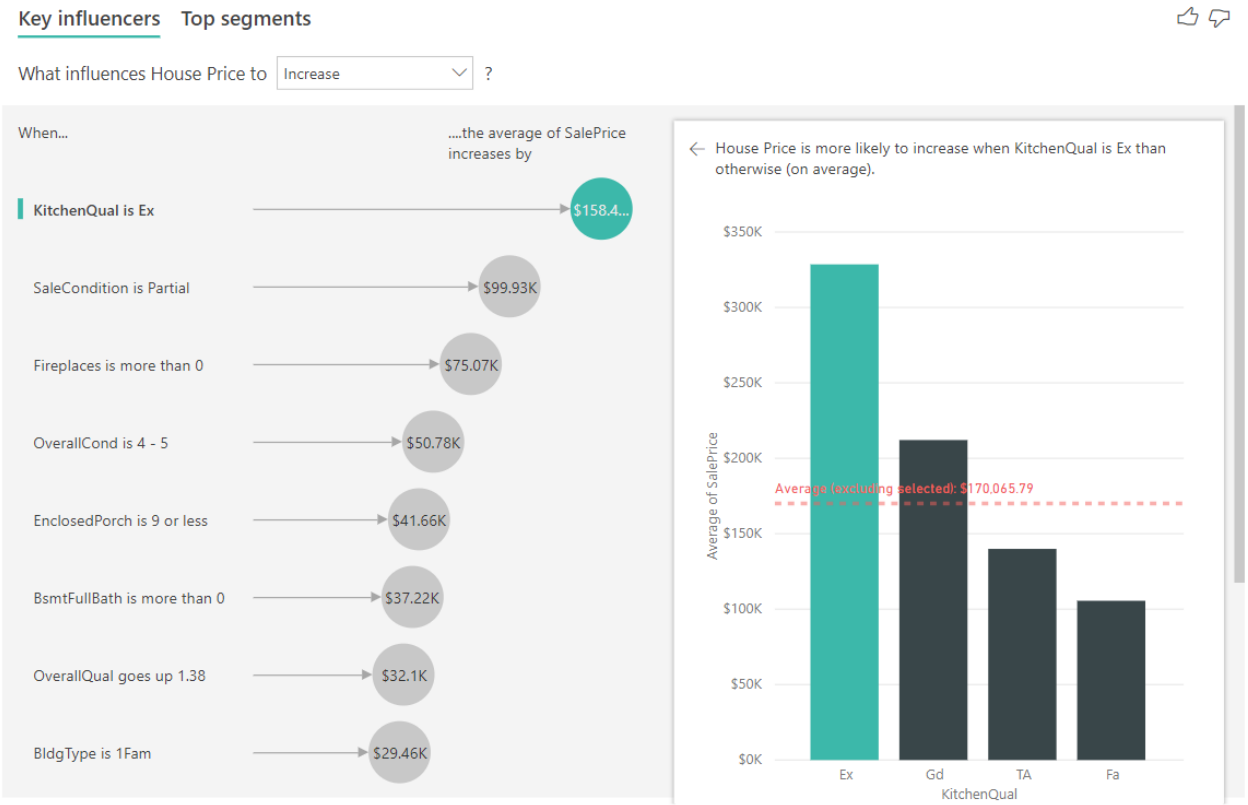


4.5.2 Interpret the Results: Key Influencers

In this scenario we look at 'What influences House Price to increase'. We are looking at a number of explanatory factors that could impact a house price like Year Built (year the house was built), KitchenQual (kitchen quality) and YearRemodAdd (year the house was remodeled).

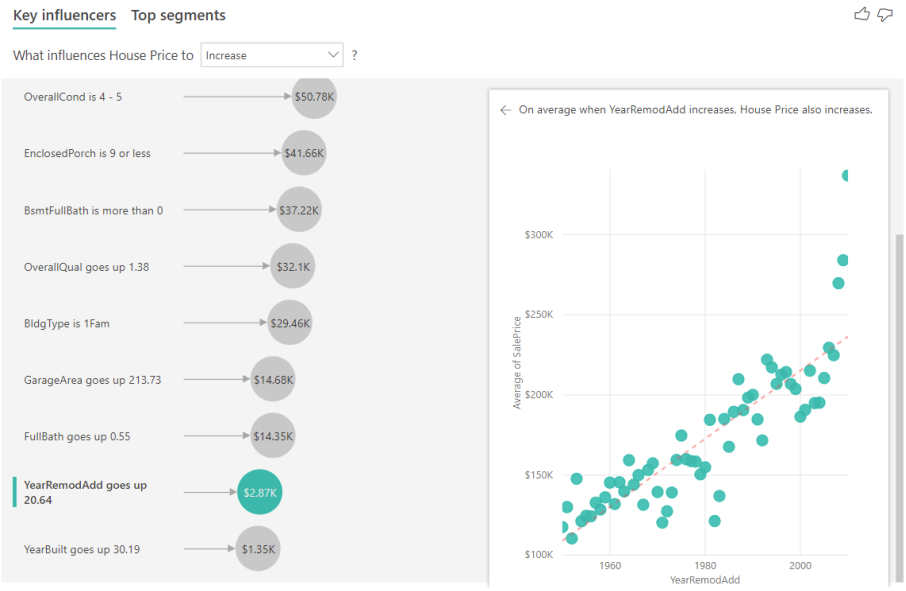
In the example below we look at our top influencer which is kitchen quality being Excellent. The results are very similar to the ones we saw when we were analyzing categorical metrics with a few important differences:

- The column chart on the right is looking at the averages rather than percentages. It therefore shows us what the average house price of a house with an excellent kitchen is (green bar) compared to the average house price of a house without an excellent kitchen (dotted line)
- The number in the bubble is still the difference between the red dotted line and green bar but it's expressed as a number ($158.49K) rather than a likelihood (1.93x). So on average, houses with excellent kitchens are almost $160K more expensive than houses without excellent kitchens.



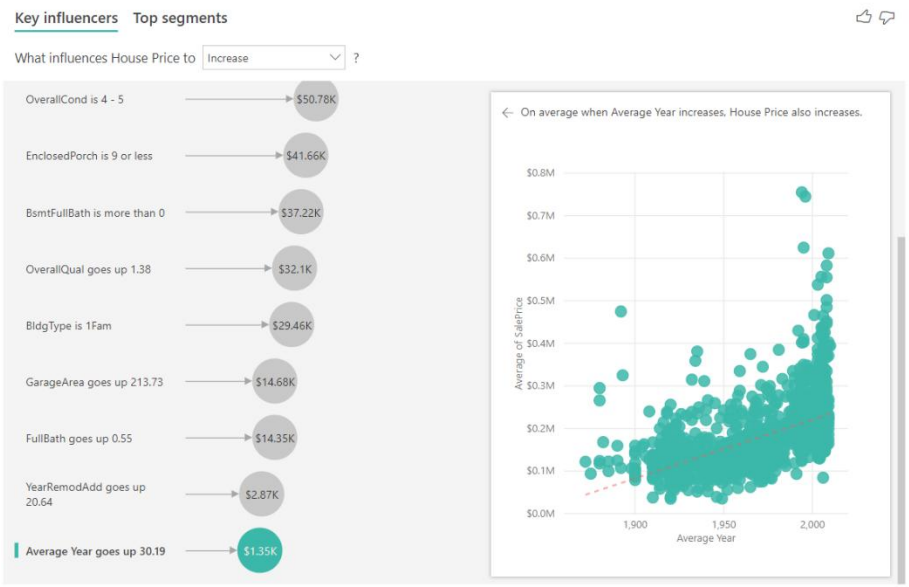
We are looking at the impact a continuous factor (year house was remodeled) has on house price. The differences compared to how we analyze continuous influencers for categorical metrics are as follows:

- The scatter plot in the right pane plots the average house price for each distinct value of year remodeled.
- The value in the bubble shows by how much the average house price increases (in this case $2.87k) when the year the house was remodeled increases by its standard deviation (in this case 20 years)



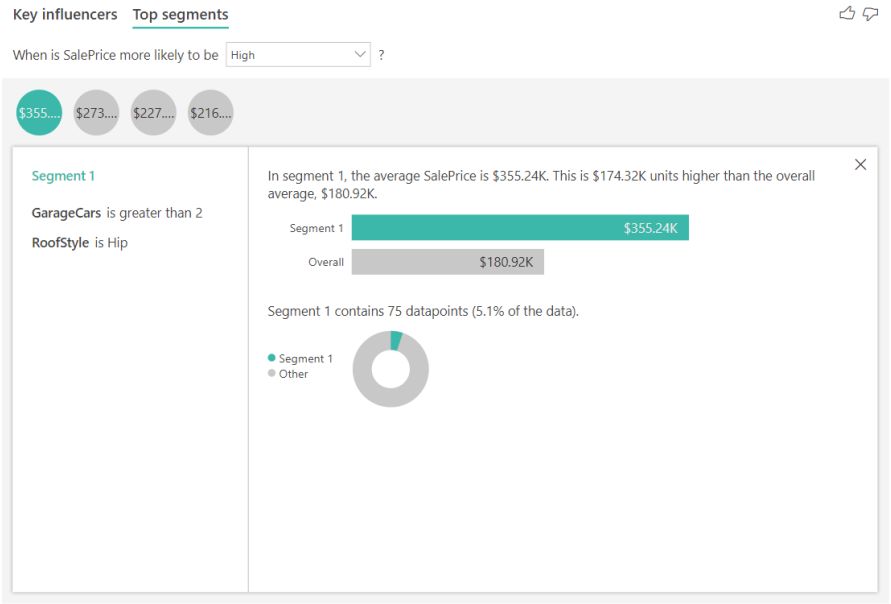
Finally, in the case of measures we are looking at the average year a house was built. The analysis here is as follows:

- The scatterplot in the right pane plots the average house price for each distinct value in the table
- The value in the bubble shows by how much the average house price increases (in this case $1.35K) when the average year increases by its standard deviation (in this case 30 years)



4.5.3 Interpret the Results: Top Segments

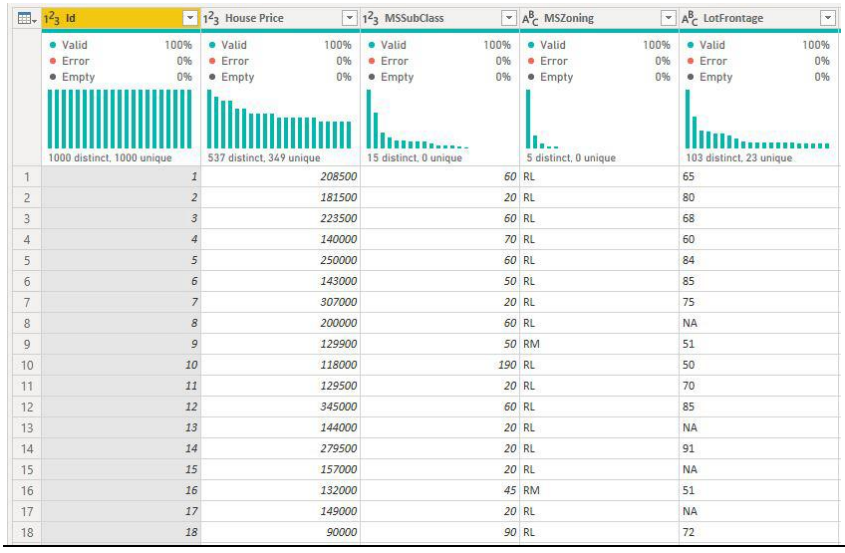
Top segments for numerical targets show groups where the house prices on average are higher than in the overall dataset. For example, below we can see that Segment 1 is made up of houses where GarageCars (number of cars the garage can fit) is greater than 2 and the RoofStyle is Hip. Houses with those characteristics have an average price of $355K compared to the overall average in the data which is $180K.



4.5.4 Analyze a Metric that is a Measure or a Summarized Column

In the case of a measure or summarized column the analysis defaults to the **Continuous Analysis Type** described above. This cannot be changed. The biggest difference between analyzing a measure/summarized column and an unsummarized numeric column is the level at which the analysis runs.

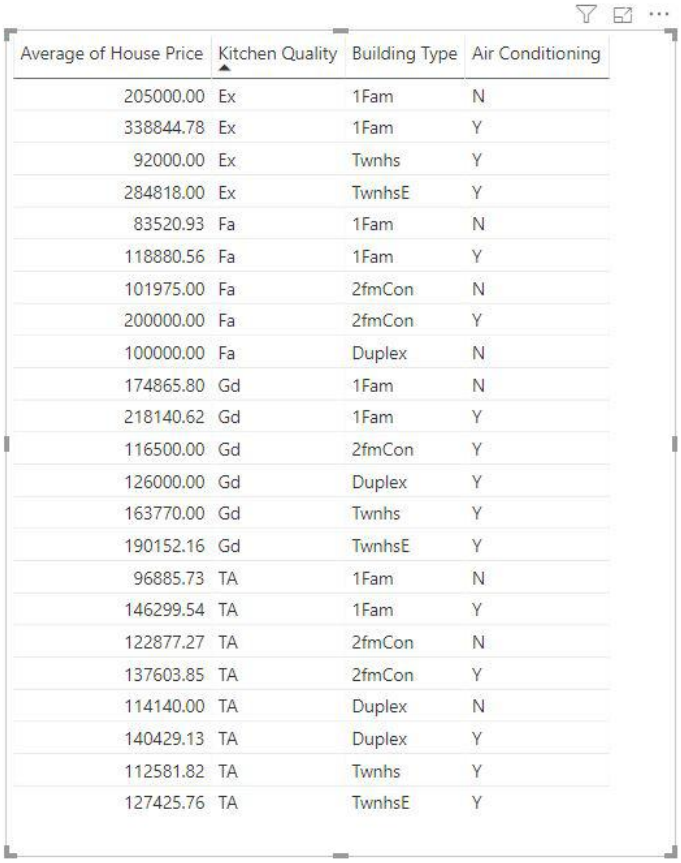
In the case of unsummarized columns, the analysis always runs at the table level. In the house price example above, we analyzed the *House Price* metric to see what influences a house price to increase/decrease. The analysis automatically runs on the table level. Our table has a unique ID for each house so the analysis runs at a house level.



	Id	House Price	MSSubClass	MSZoning	LotFrontage
	Valid 100% Error 0% Empty 0%	Valid 100% Error 0% Empty 0%	Valid 100% Error 0% Empty 0%	Valid 100% Error 0% Empty 0%	Valid 100% Error 0% Empty 0%
	1000 distinct, 1000 unique	537 distinct, 349 unique	15 distinct, 0 unique	5 distinct, 0 unique	103 distinct, 23 unique
1	1	208500	60	RL	65
2	2	181500	20	RL	80
3	3	223500	60	RL	68
4	4	140000	70	RL	60
5	5	250000	60	RL	84
6	6	143000	50	RL	85
7	7	307000	20	RL	75
8	8	200000	60	RL	NA
9	9	129900	50	RM	51
10	10	118000	190	RL	50
11	11	129500	20	RL	70
12	12	345000	60	RL	85
13	13	144000	20	RL	NA
14	14	279500	20	RL	91
15	15	157000	20	RL	NA
16	16	132000	45	RM	51
17	17	149000	20	RL	NA
18	18	90000	90	RL	72

For measures and summarized columns, we don't immediately know what level to analyze them at. If *House Price* was summarized as an **Average**, we would need to consider what level we would like this average house price calculated. Is it the average house price at a neighborhood level? Or perhaps a regional level?

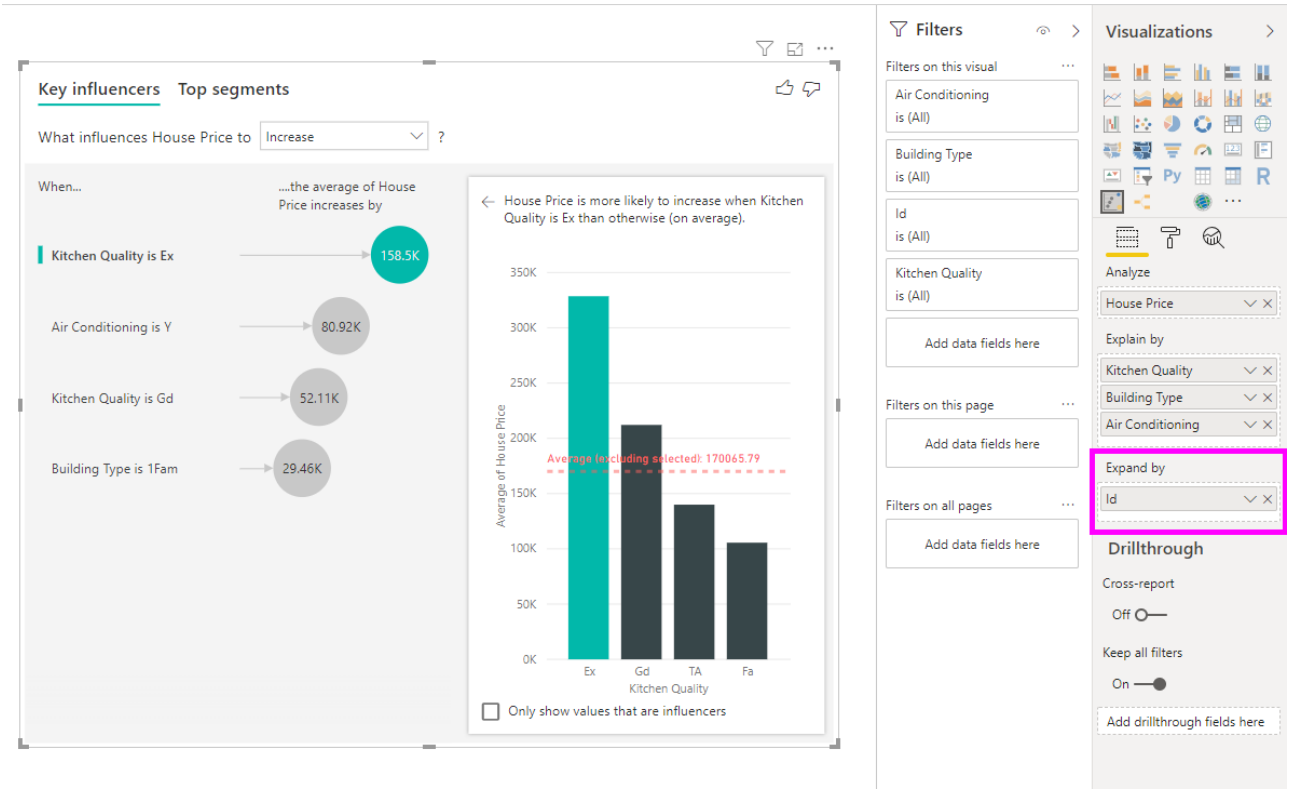
Measures and summarized columns are automatically analyzed at the level of the **Explain by** fields used. Imagine we have three fields in **Explain By** we are interested in: *Kitchen Quality*, *Building Type* and *Air Conditioning*. Average House Price would be calculated for each unique combination of those three fields. It is often helpful to switch to a table view to take a look at what the data being evaluated looks like



Average of House Price	Kitchen Quality	Building Type	Air Conditioning
205000.00	Ex	1Fam	N
338844.78	Ex	1Fam	Y
92000.00	Ex	Twnhs	Y
284818.00	Ex	TwnhsE	Y
83520.93	Fa	1Fam	N
118880.56	Fa	1Fam	Y
101975.00	Fa	2fmCon	N
200000.00	Fa	2fmCon	Y
100000.00	Fa	Duplex	N
174865.80	Gd	1Fam	N
218140.62	Gd	1Fam	Y
116500.00	Gd	2fmCon	Y
126000.00	Gd	Duplex	Y
163770.00	Gd	Twnhs	Y
190152.16	Gd	TwnhsE	Y
96885.73	TA	1Fam	N
146299.54	TA	1Fam	Y
122877.27	TA	2fmCon	N
137603.85	TA	2fmCon	Y
114140.00	TA	Duplex	N
140429.13	TA	Duplex	Y
112581.82	TA	Twnhs	Y
127425.76	TA	TwnhsE	Y

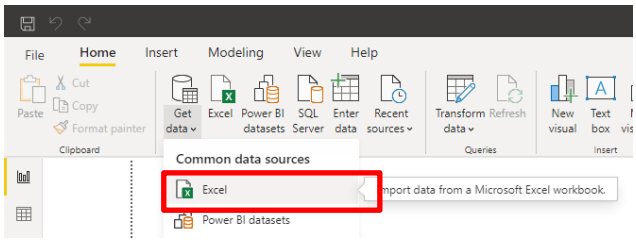
This analysis is very summarized and so it will be hard for the regression model to find any patterns in the data it can learn from. We should run the analysis at a more detailed level to get better results. If we wanted to analyze the house price at the house level we would need to explicitly add the ID field to the analysis. Nevertheless, we don't want the house ID to be considered an influencer. It is not helpful to learn that as house ID increases, the price of a house increase. This is where the Expand By field well option comes in handy. You can use **Expand By** to add fields you want to use for setting the level of the analysis without looking for new influencers.

Take a look at what the visualization looks like once we add ID to **Expand By**. Once you have defined the level at which you want your measure evaluated, interpreting influencers is exactly the same as for unsummarized numeric columns.

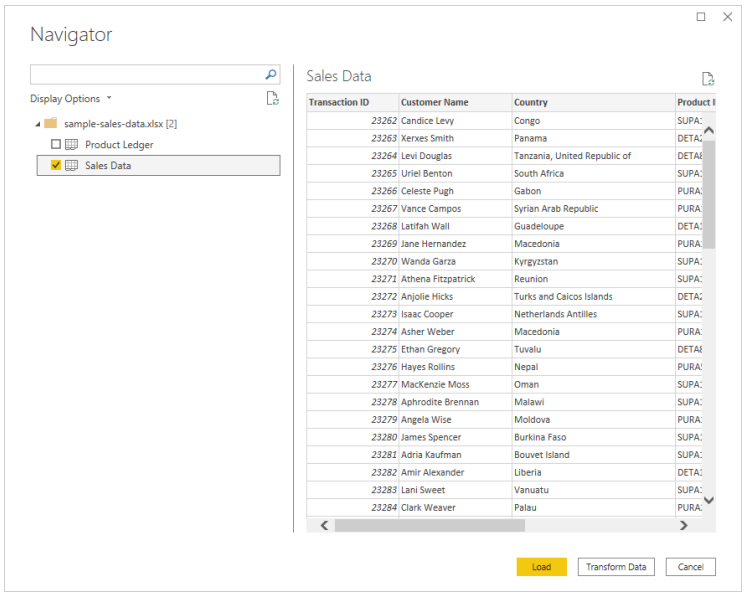

Key influencers
Top segments
What influences House Price to
Increase
?
When...
....the average of House Price increases by
Kitchen Quality is Ex
158.5K
Air Conditioning is Y
80.92K
Kitchen Quality is Gd
52.11K
Building Type is 1Fam
29.46K
House Price is more likely to increase when Kitchen Quality is Ex than otherwise (on average).
350K
300K
250K
200K
150K
100K
50K
0K
Average of House Price
Average (excluding selected): 170065.79
Ex
Gd
TA
Fa
Kitchen Quality
Only show values that are influencers
Filters
Filters on this visual
Air Conditioning is (All)
Building Type is (All)
Id is (All)
Kitchen Quality is (All)
Add data fields here
Filters on this page
Add data fields here
Filters on all pages
Add data fields here
Visualizations
Analyze
House Price
Explain by
Kitchen Quality
Building Type
Air Conditioning
Expand by
Id
Drillthrough
Cross-report
Off
Keep all filters
On
Add drillthrough fields here

5. Drill Down

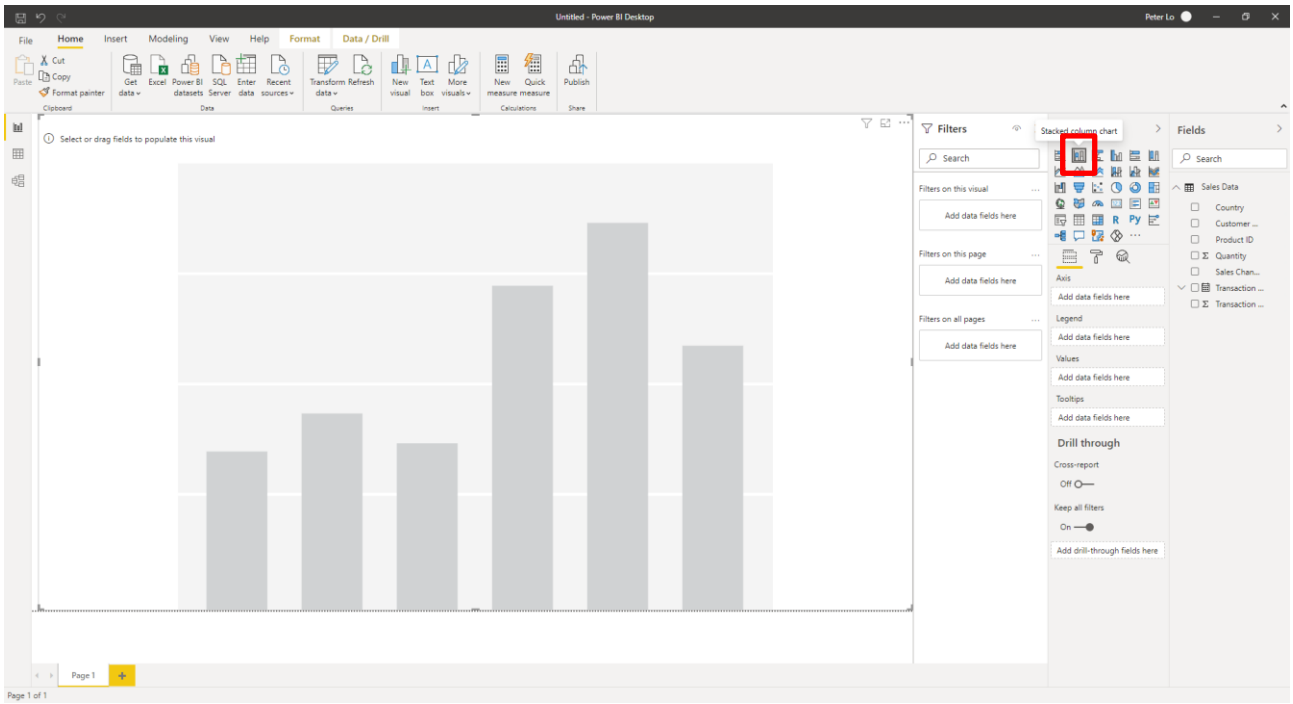
1. Select **Get Data ➔ Excel** in the **Home** ribbon tab.



2. In the Open File dialog box, select the *sample-sales-data.xls* file. Then in the Navigator pane, select the *Sales Data* table and then select **[Load]**.

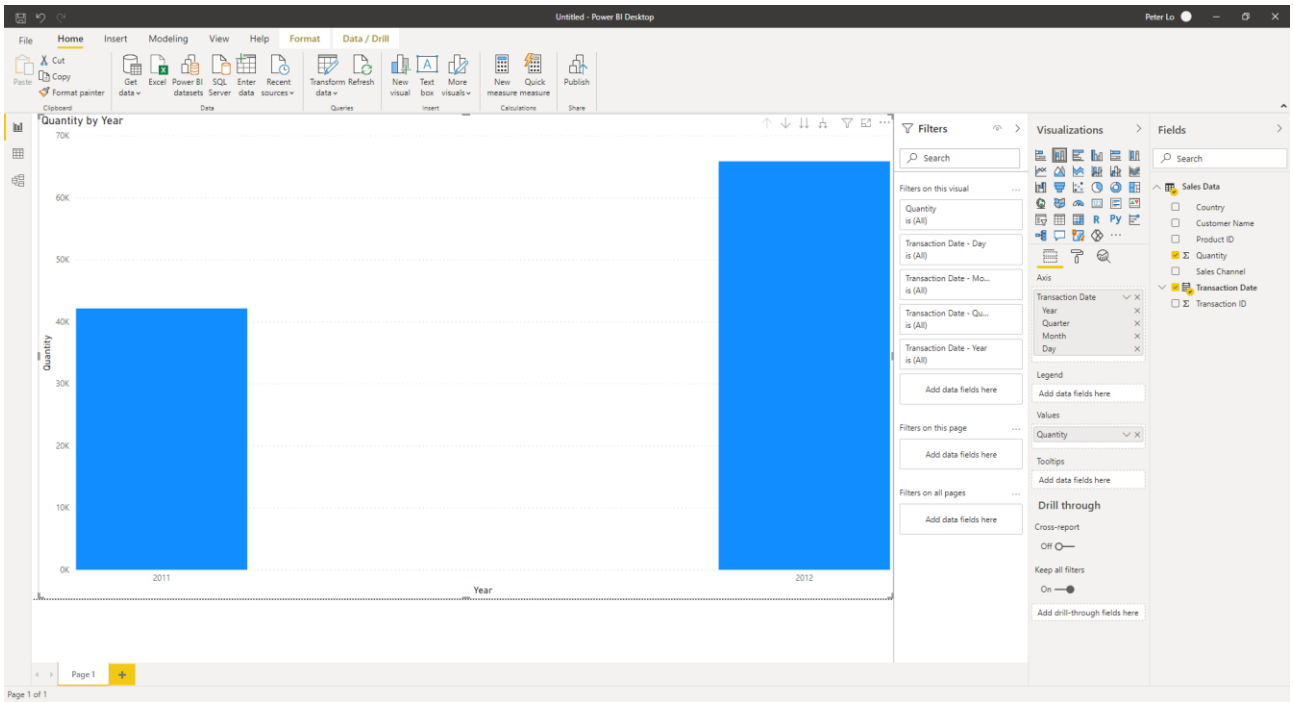


3. Add a **Stacked Column Chart** visual.

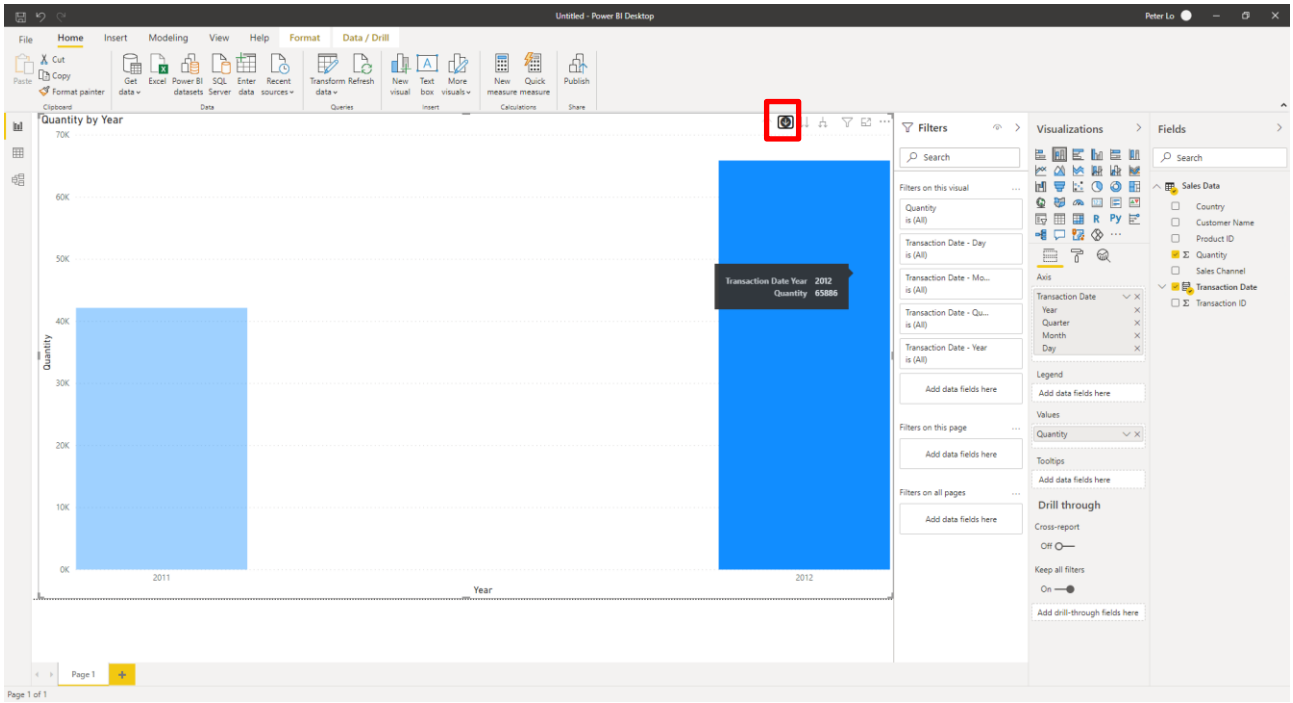


4. Add the following fields to the visual wells:
   - Assign *Transaction Date* to **X-Axis**
   - Assign *Quantity* to **Y-Axis**

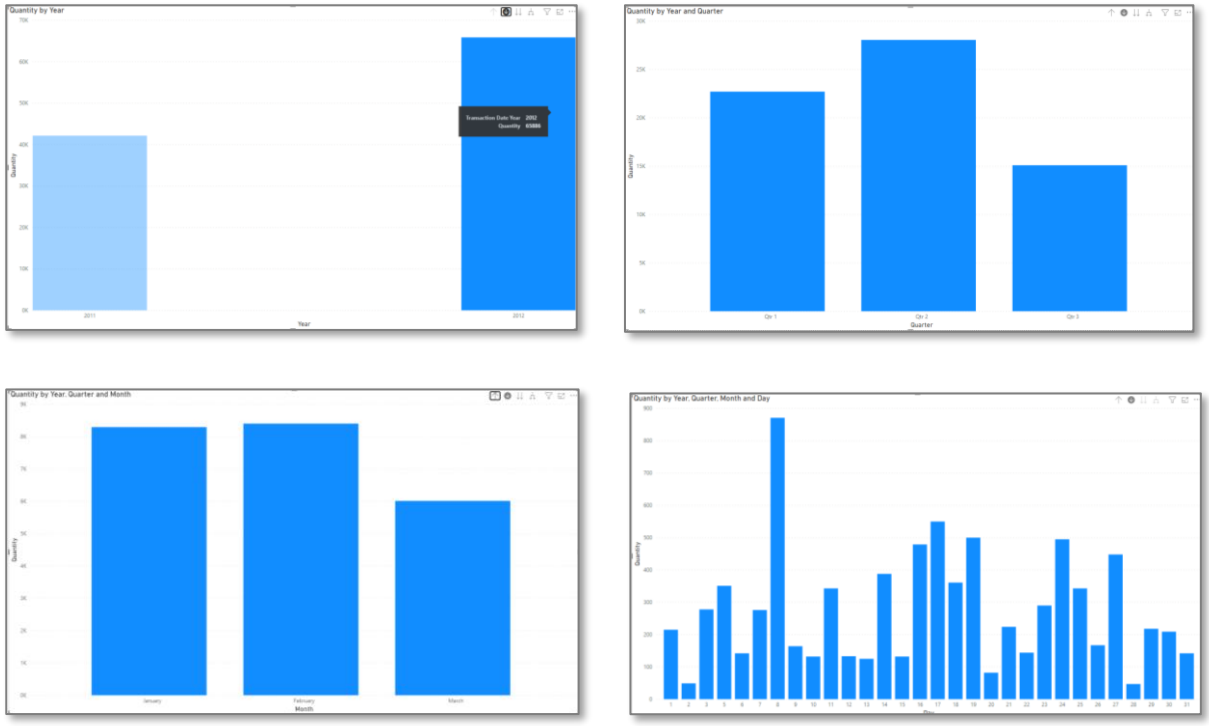
   When you add a date field to a visual in the Axis field bucket, Power BI automatically adds a time hierarchy that includes Year, Quarter, Month and Day. By doing this, Power BI allows your visuals to have time-based interaction with those viewing your reports, by letting users drill-down through those different time levels



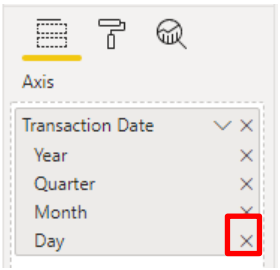
5. With a hierarchy in place, you can begin drilling down through the time hierarchy. For example, clicking a year in the chart drills down to the next level in the hierarchy.



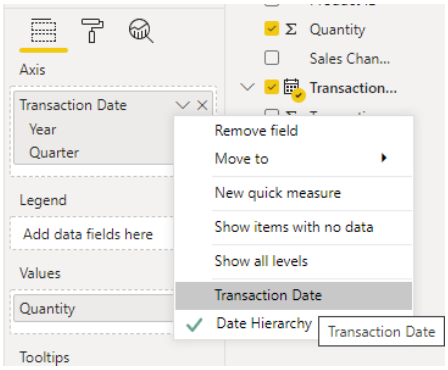
6. Now you can drill down to next level by double click the column bar.



7. In that automatically created hierarchy, you can also manage to which level your shared report allows people to drill. To do this, in the Visualizations pane, simply click the **[X]** beside the hierarchy that you want to remove. The deleted level is removed from the report, and drilling no longer displays that level.



8. If you need to get that level of the hierarchy back, just remove the date field, and then add it again from the Fields pane, and the hierarchy is once again created for you automatically. There may be times when you don't want the hierarchy to be used for a visual. You can control that by selecting the down-arrow button beside the Date field, and select **Date** rather than **Date Hierarchy**. That prompts Power BI to show the raw date values in the visual.



9. You can also expand all data elements currently visible at once, rather than selecting a single quarter, or a single year. To do that, select the Drill all icon in the top left of the visual, which is a double-down arrow icon

