|  |  |
| --- | --- |
|  |  |

|  |
| --- |
| Box and Whisker Features Document |
|  |

# Short Description

Displays data distribution quartiles in a box plot. A line (whiskers) through the box indicates variability outside the upper and lower quartiles.

# Overview

Box and whisker charts are useful for quickly comparing distributions between several sets of data, for example comparing a series of medical trial results or students’ test scores. The chart consists of two parts: the main body, called the box, and the vertical lines coming out of the box, called whiskers. The first quartile forms the bottom of the box, and the third quartile forms the top. The whiskers connect the minimum and the maximum data values to the box.

In addition to showing the median, first and third quartiles, and the maximum and minimum values, Box and Whisker by MAQ Software displays the mean, standard deviation, and quartile deviation.

While other Power BI box plots display some of these statistics, Box and Whisker by MAQ Software provides additional user controls and capabilities such as the ability to add a parent axis and flip the visual horizontally or vertically based on the reporting requirements.

Other features include:

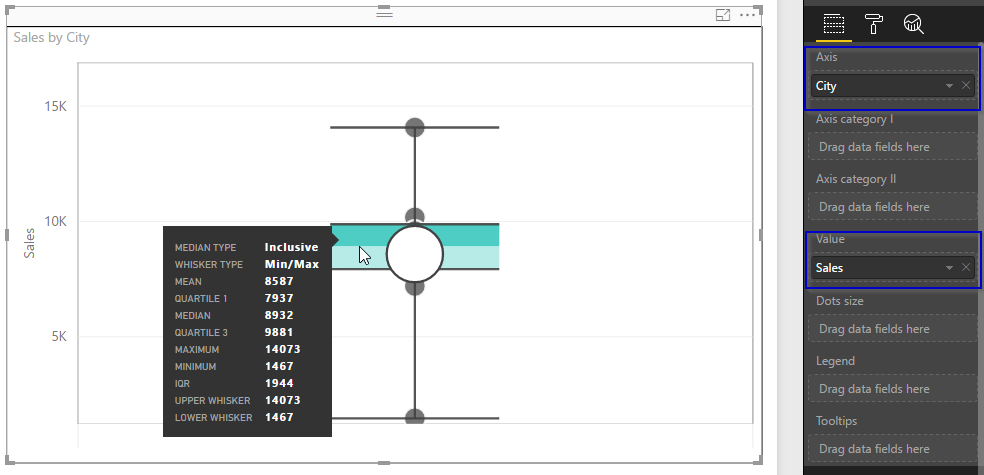
* Option to enable or disable mean dot and median line.
* Option to choose the type of quartile: inclusive or exclusive.
* Option to choose type of whiskers: Min/Max, < 1.5 IQR, > 1.5 IQR, or custom.
* Option to choose shapes of mean and data points: circle, square, or rectangle.
* Option to hide or show all data points.
* Option to hide or show outliers.
* Option to control the opacity of boxes.

Box and Whisker by MAQ Software allows you to see a complete picture of your data distribution at a glance.

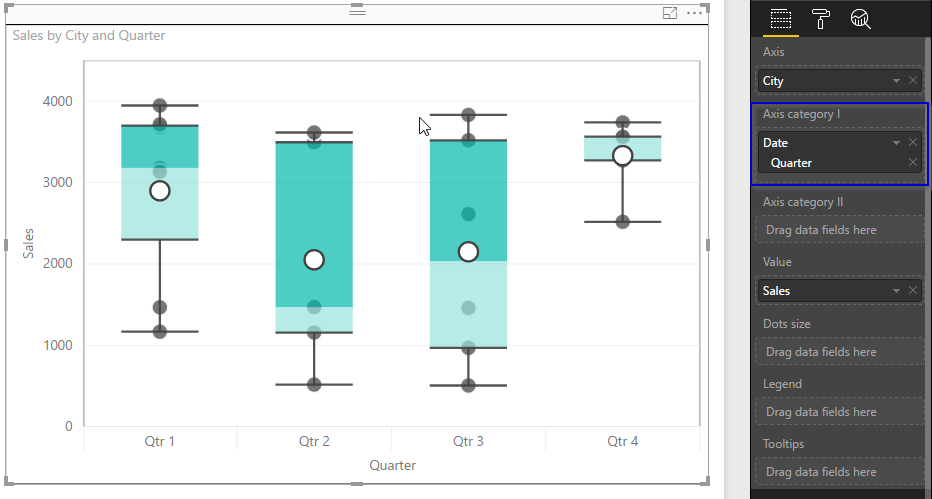
# Example

Imagine you are the CEO of a company and you would like to know total sales for all cities where your business is located.

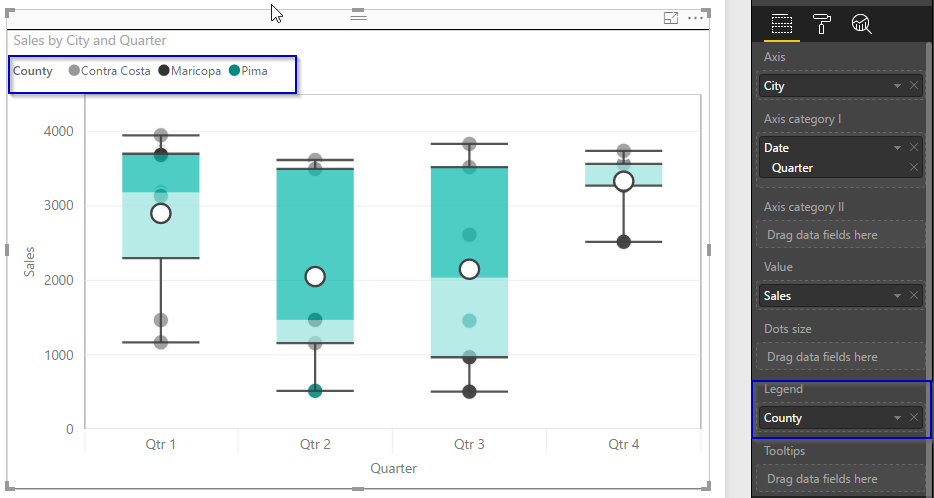
To get started, you enter 'City' in the 'Axis' field and 'Sales' in the 'Value' field. You will then see dots that represent the cities and their respective sales. Plotted on top of the dots, you will see a box showing the quartiles, mean, and median, along with whiskers indicating the minimum and maximum values.



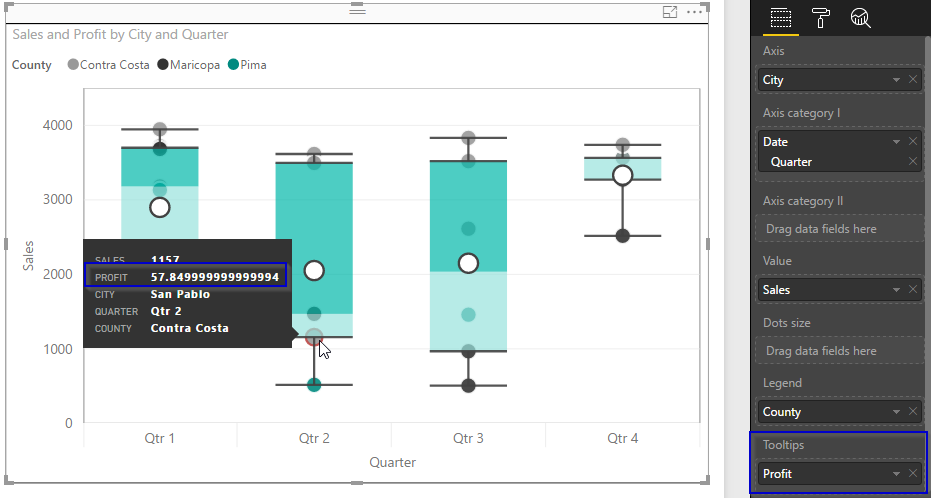
Next, you would like to compare quarterly sales. You can determine this by dragging 'Quarter' into the 'Axis Category I' field.



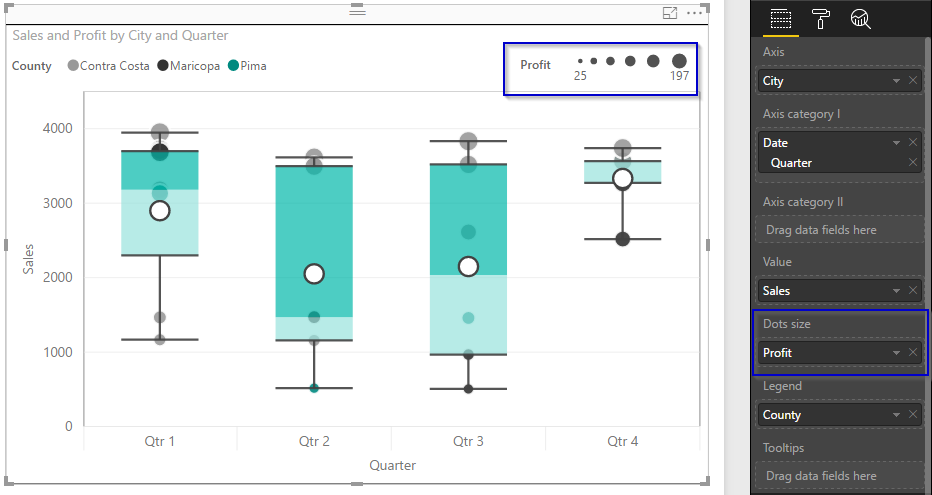
Now, you'd like to take your analysis even deeper and discover which sales are attributed to each county. To discover the answer, select 'County' in the 'Legend' field.



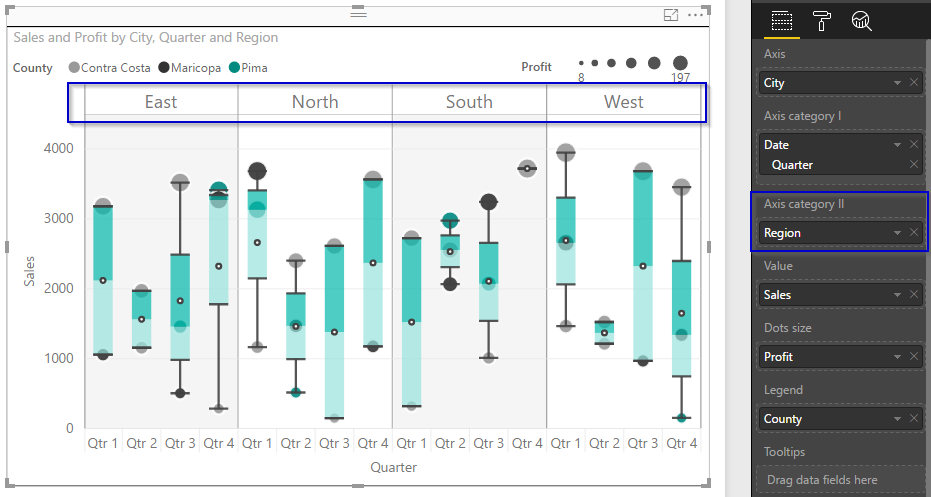
You can add tooltips to analyze values not displayed on the graph. For example, if you want to know the profit associated with every sales value, you can add 'Profit' in the 'Tooltips' field.



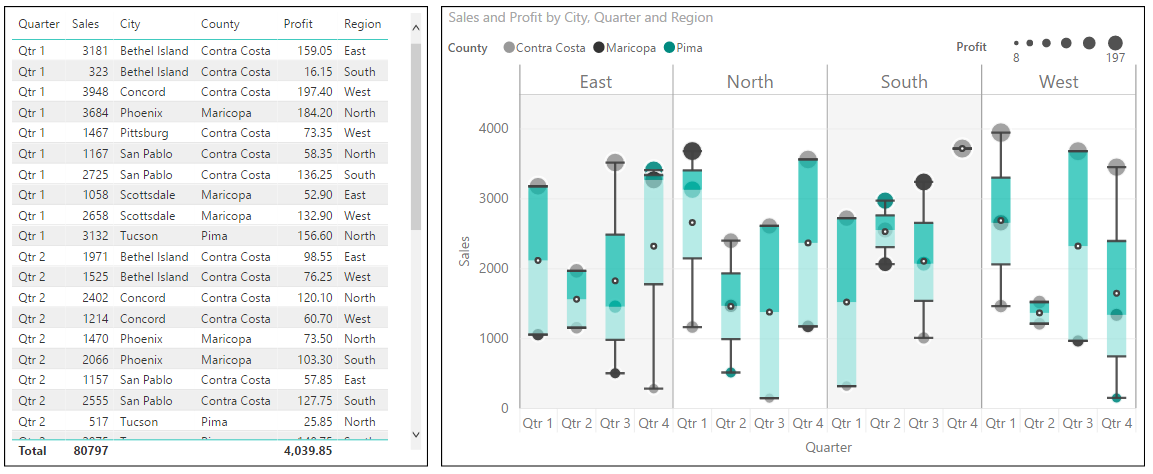
If you want to see the profit associated with each city, you can drag the 'Profit' value into the 'Dots size' field. Now, the size of the dot indicates profits; large dots indicate higher profits, and small dots indicate lower profits.



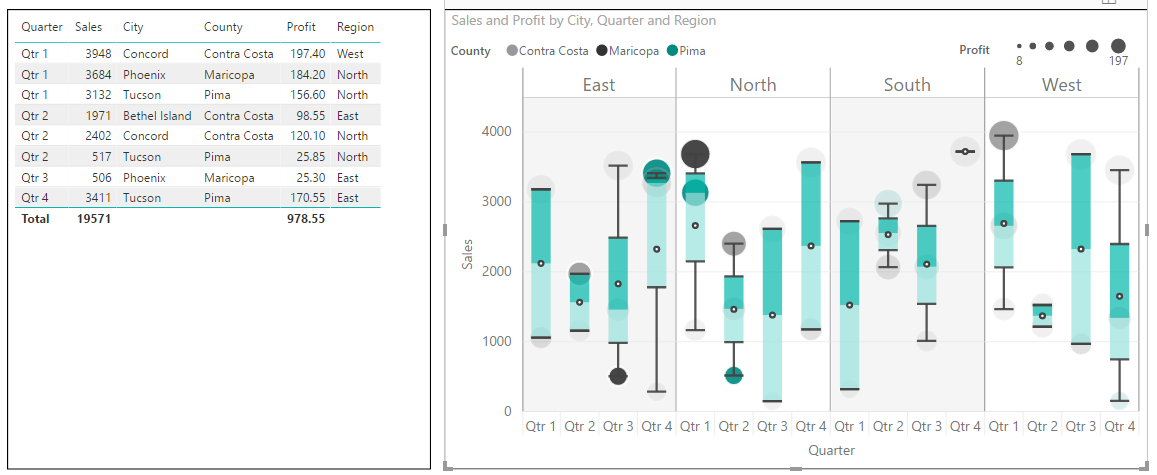
You can add 'Region' in 'Axis Category II' to classify your data based on region and quarter.



If you choose to include a table with the Box and Whisker visual, the page will appear as follows:



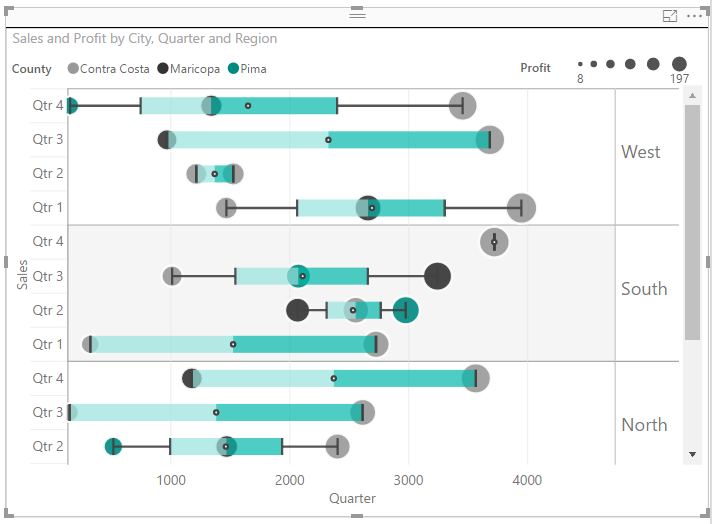
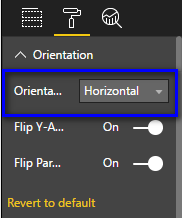
In the table view, you can filter data by clicking on the bubbles or the table.



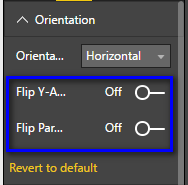
# Formatting

Users can format Box and Whisker by MAQ Software with the following options in the Format panel:

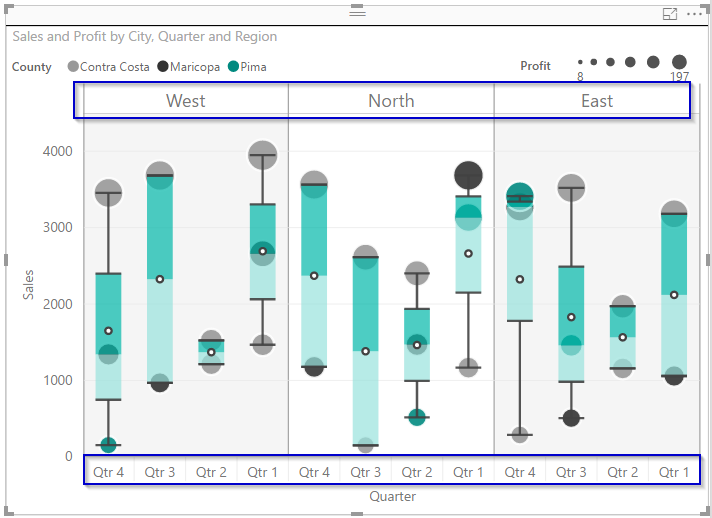
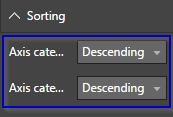
1. Use 'Orientation' to switch the visual between vertical and horizontal views.



When you select the horizontal view, you can optionally flip the orientation of the Y-axis labels and parent labels.

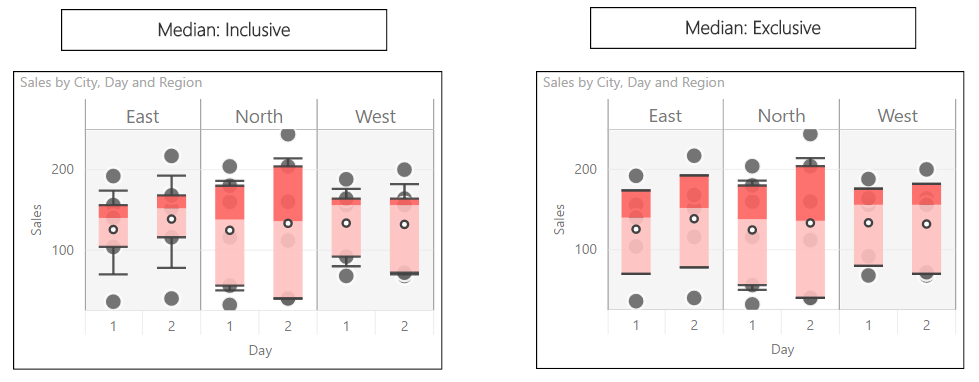
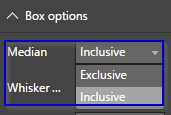


1. You can sort the visual based on the 'Axis category I' and 'Axis category II' fields, in ascending or descending order.

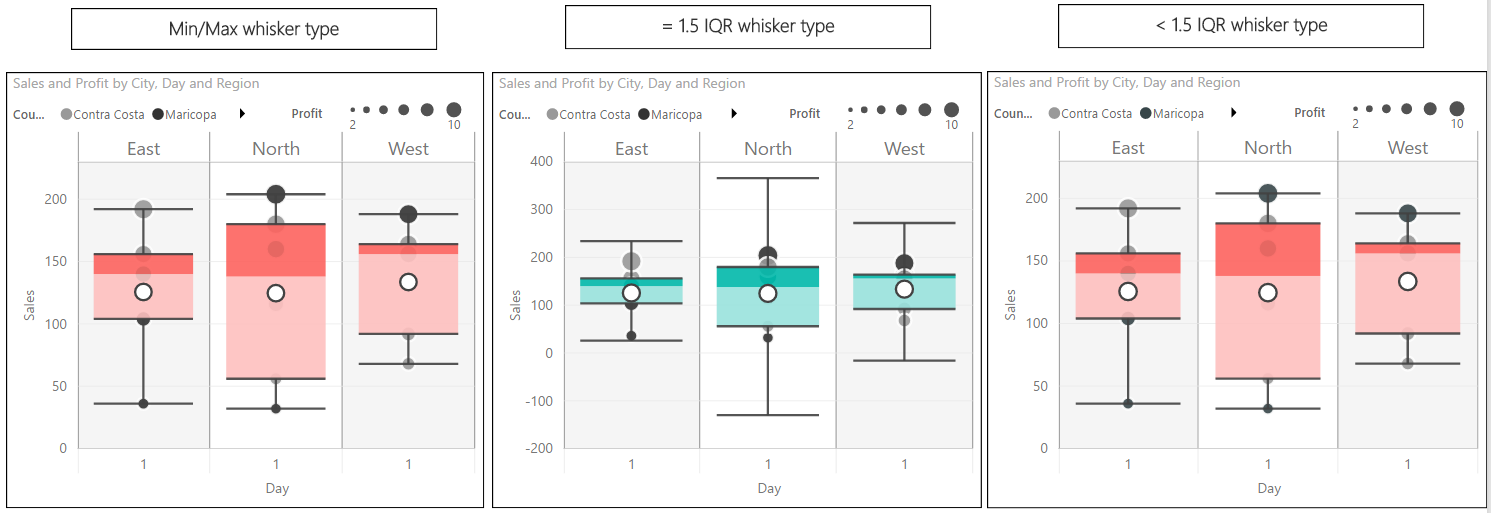
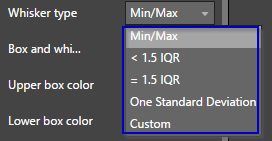


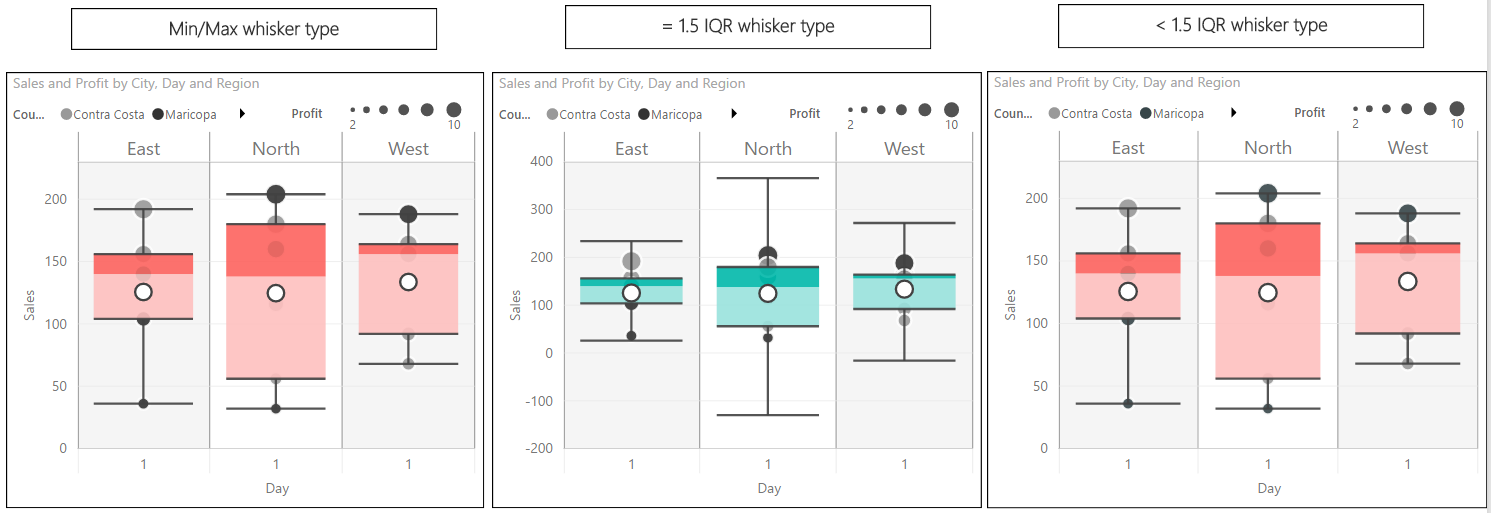
1. 'Box options' include the following choices:
   1. Median:
      1. Inclusive: The median is included while calculating quartile values Q1 and Q3.
      2. Exclusive: The median is excluded while calculating quartile values Q1 and Q3.

If the number of data points is even, both options will give the same quartile values. However, an odd number of data points returns different results because the median value is included in both the upper and lower half of the dataset.



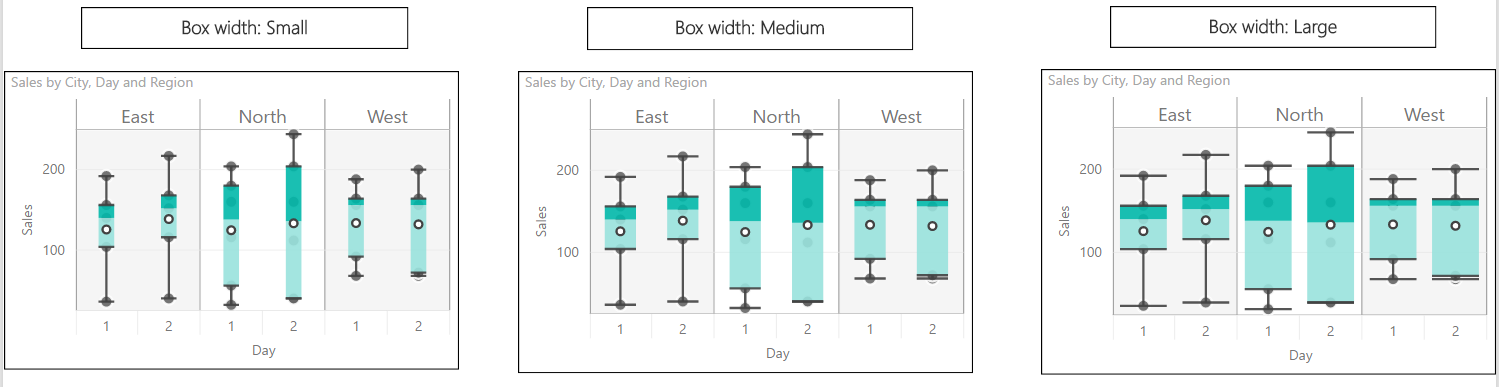
* 1. Whisker type:
     1. Min/Max
     2. < 1.5 IQR
     3. = 1.5 IQR
     4. One Standard Deviation
     5. Custom



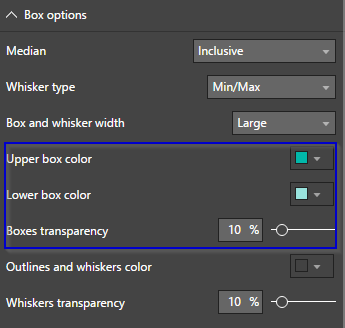


* 1. Box and whisker width:

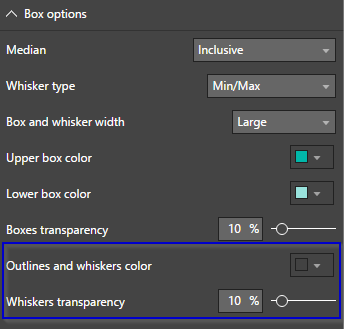




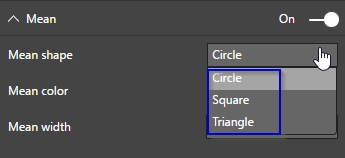
* 1. Box colors and transparency:

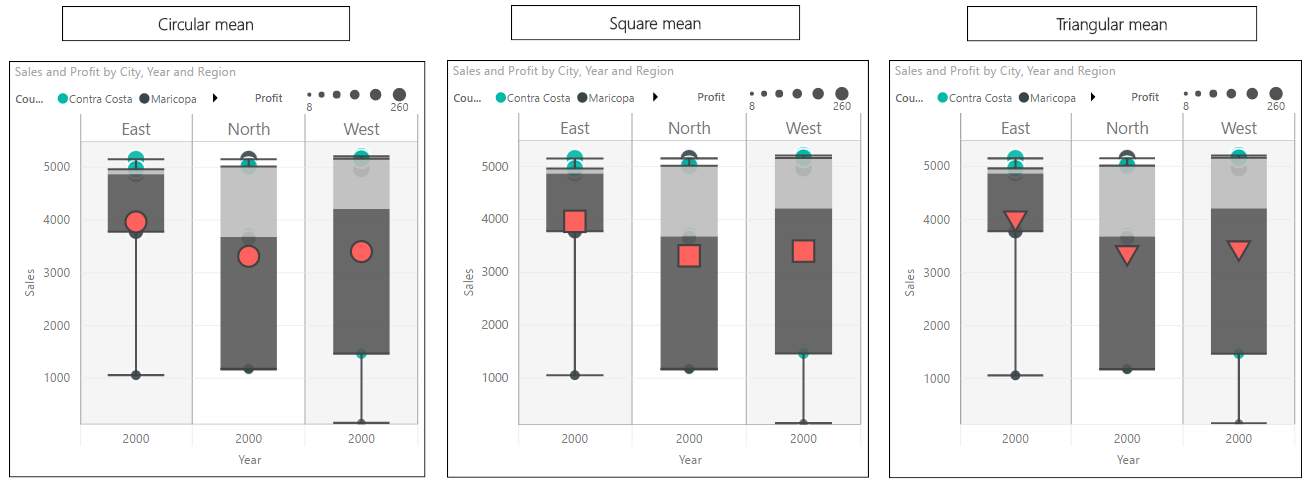


* 1. Outlines, whisker color, and whisker transparency:

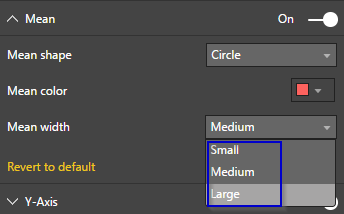


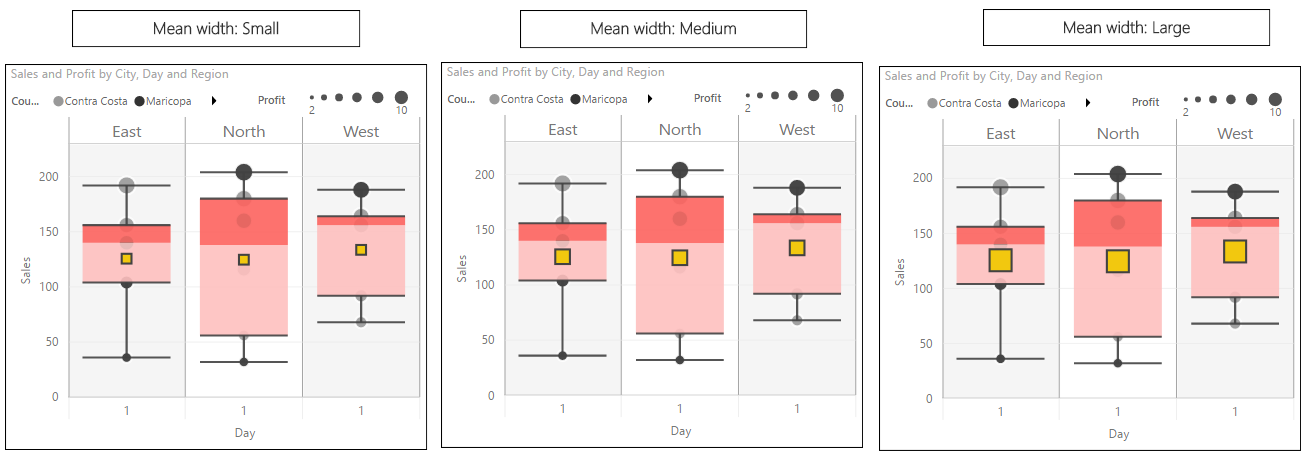
1. Formatting options for 'Mean' include:
   1. Mean shape:



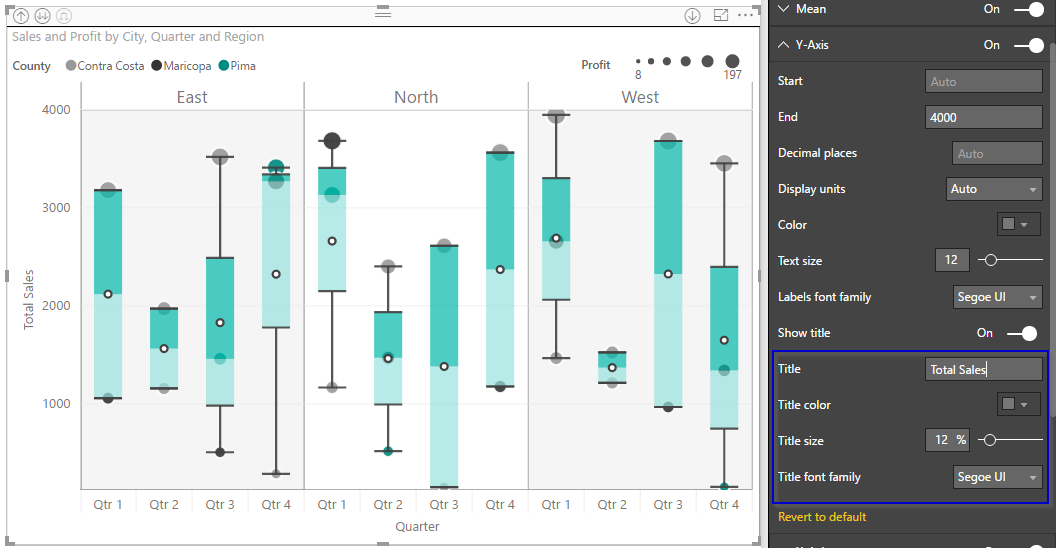


* 1. Mean color and width:

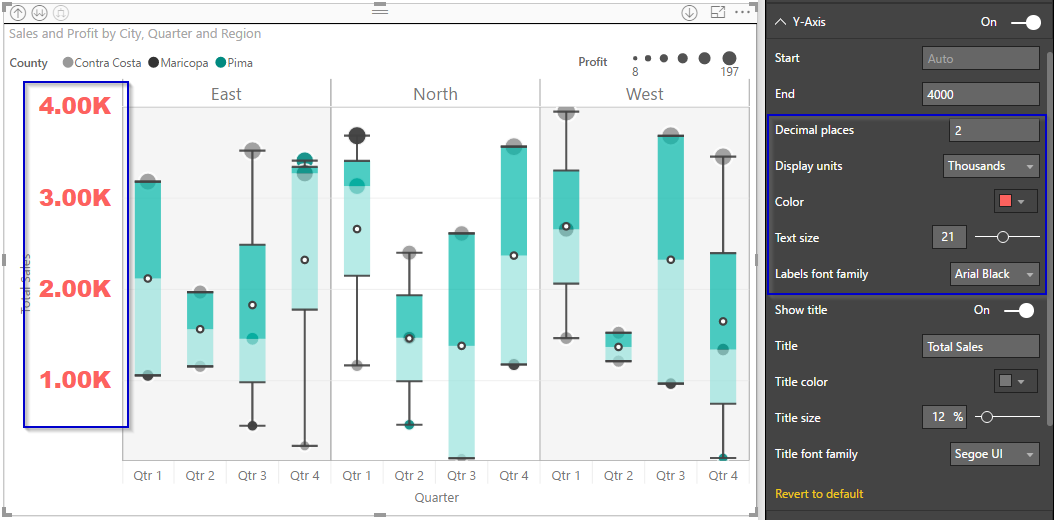




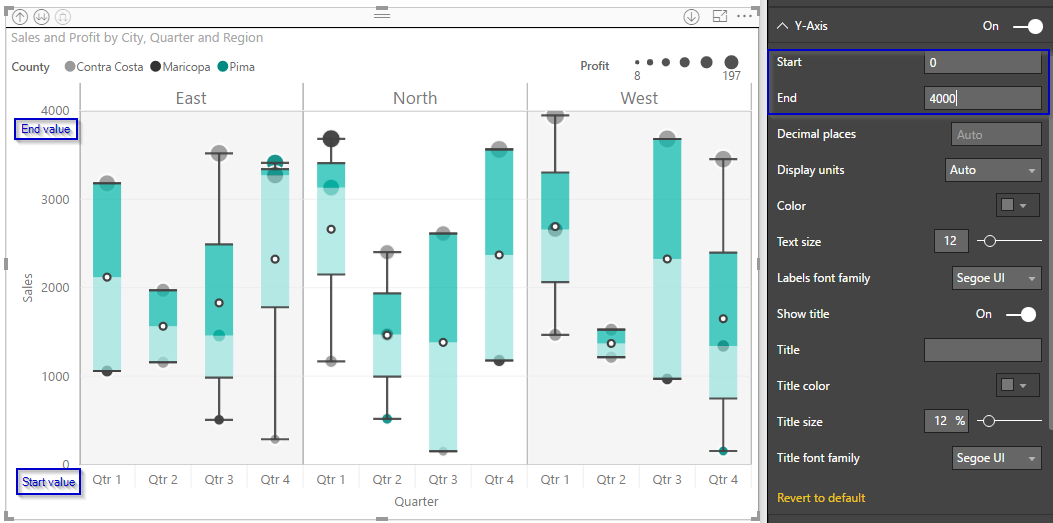
1. Formatting options for the 'Y-Axis':
   1. Users can enter a custom title, along with its color, size, and font.



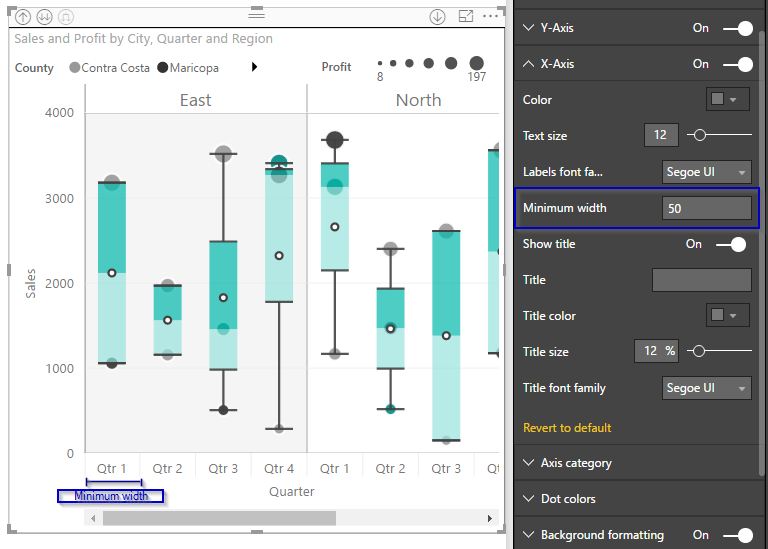
* 1. Users can update decimal places and display units, along with their color, size, and font.



* 1. Users can change the start and end values for the Y-axis.



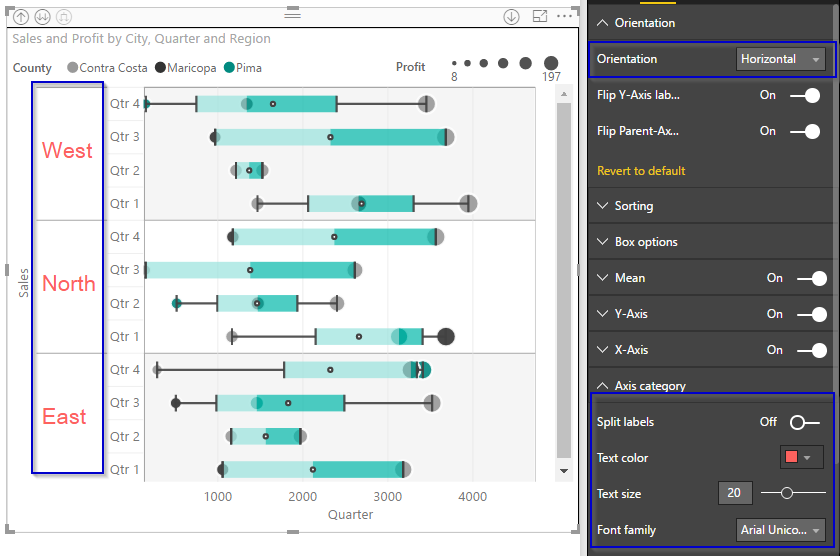
1. Similar options are available for the X-axis. Additionally, users can configure the minimum width of each X-axis label. After updating the width, a horizontal scrollbar appears if the labels cannot be accommodated.



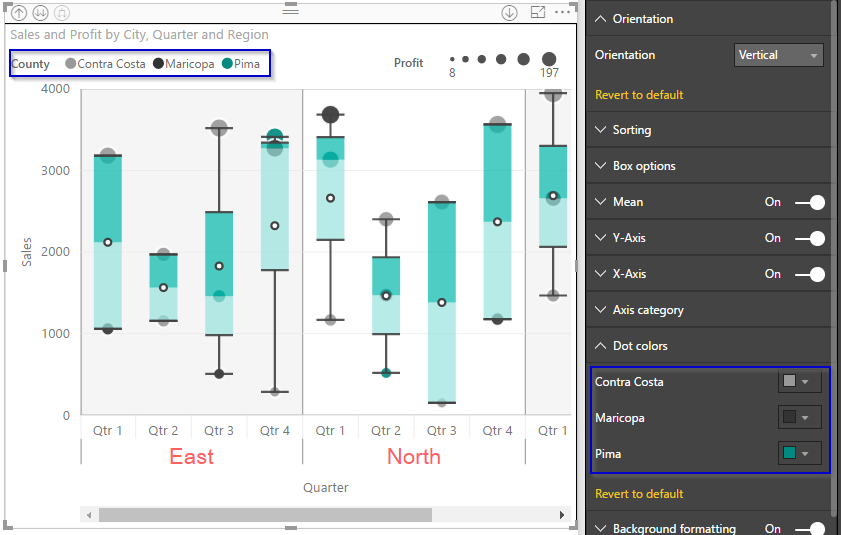
1. Users can flip the 'Axis category II' labels from the top to the bottom using the 'Axis category' settings. The color, size, and font of the labels can also be updated.



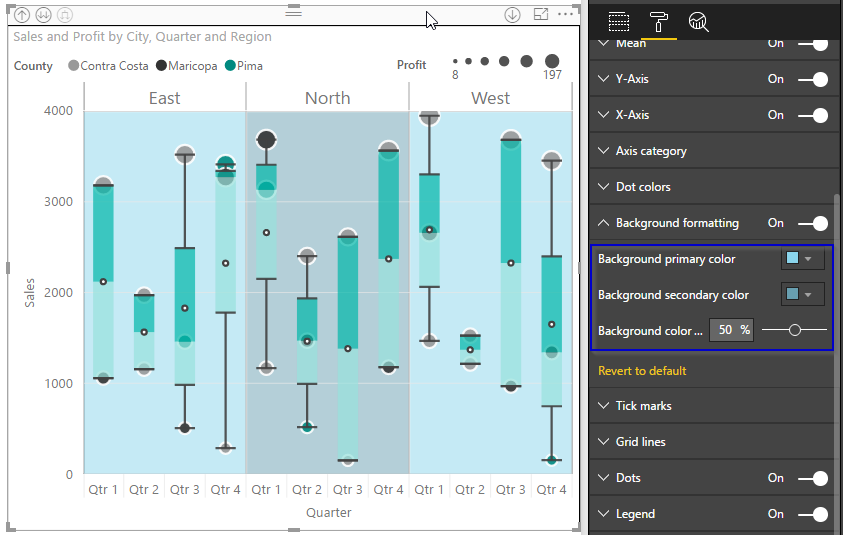
In the 'Horizontal' orientation, users can flip the 'Axis category II' labels from right to left.



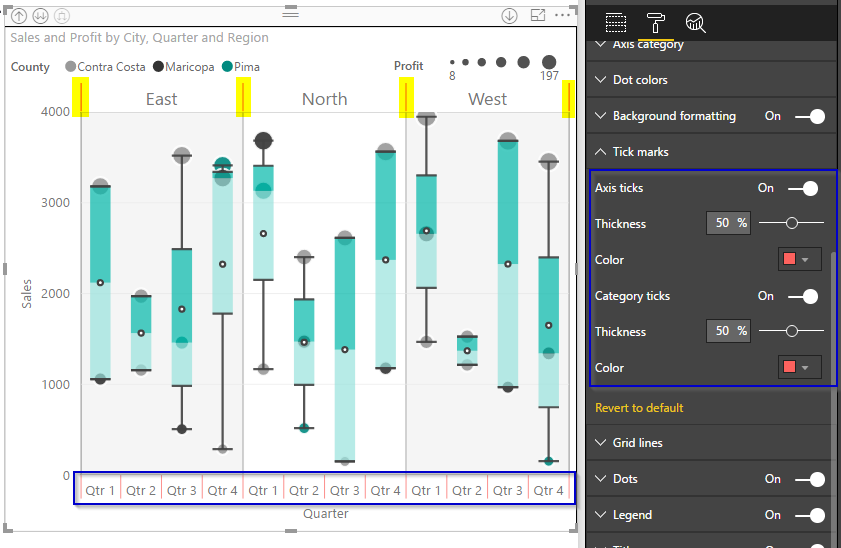
1. When the legend is present, users can modify the color of the various legend items and dots using 'Dot Colors.'

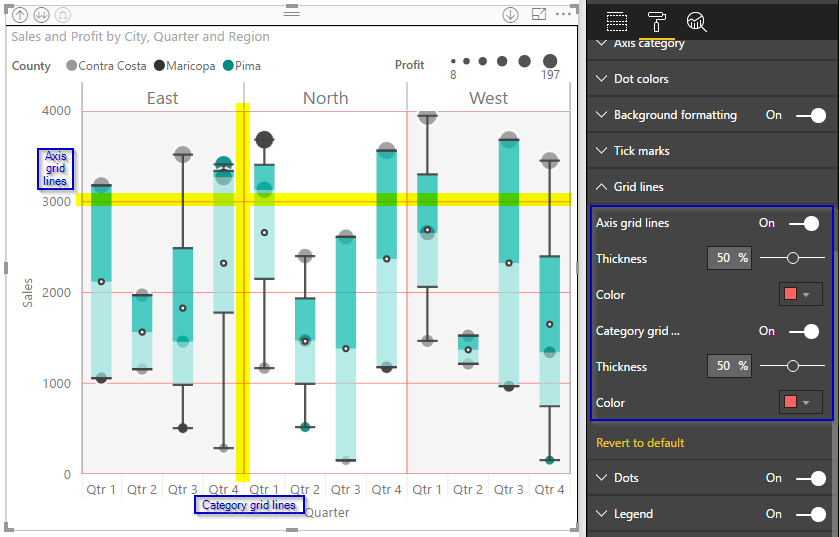


1. The primary and secondary colors of the graph can be modified, along with the transparency.

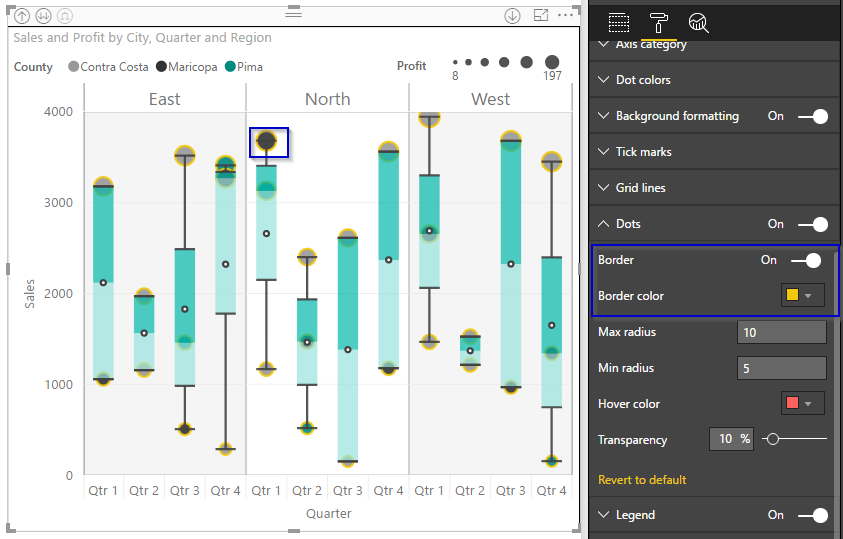


1. The axis and category tick marks can be shown or hidden. Users can also update the color and thickness of the axis and category tick marks. Similar options are available for the grid lines.

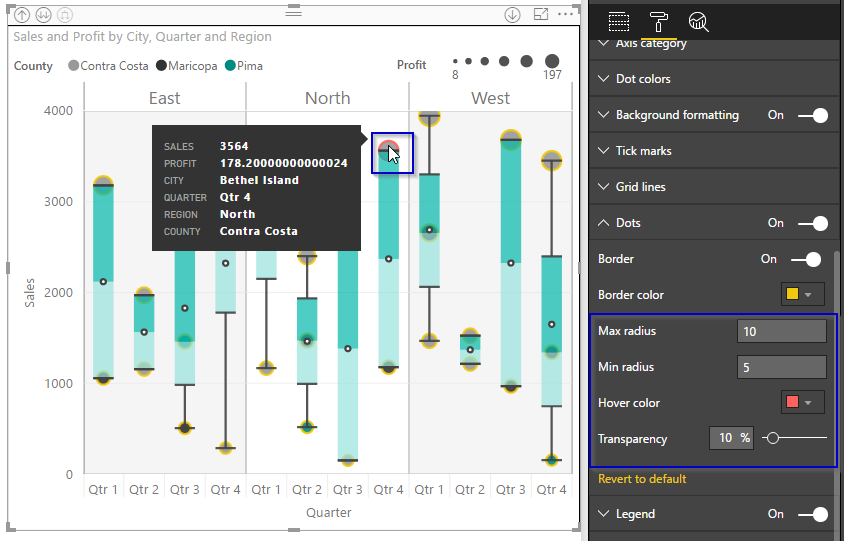




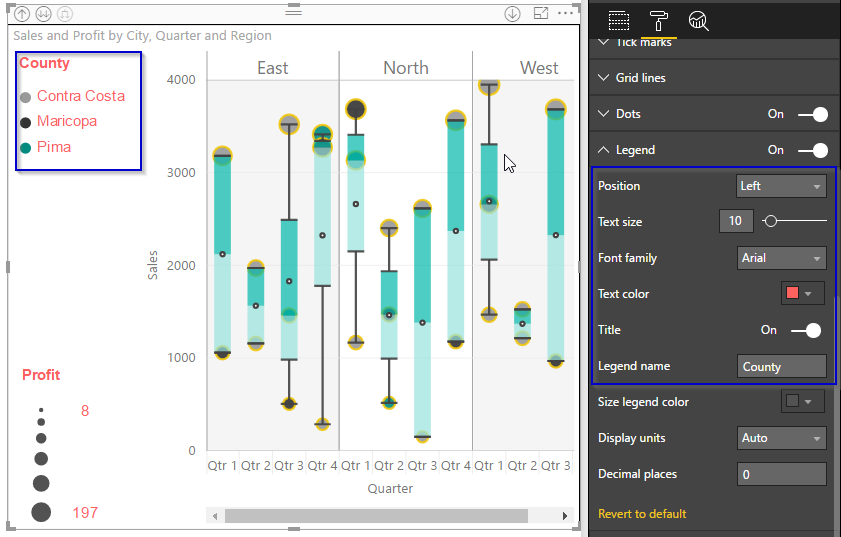
1. Formatting option for 'Dots':
   1. Users can enable or disable a border, with the option to change its color.



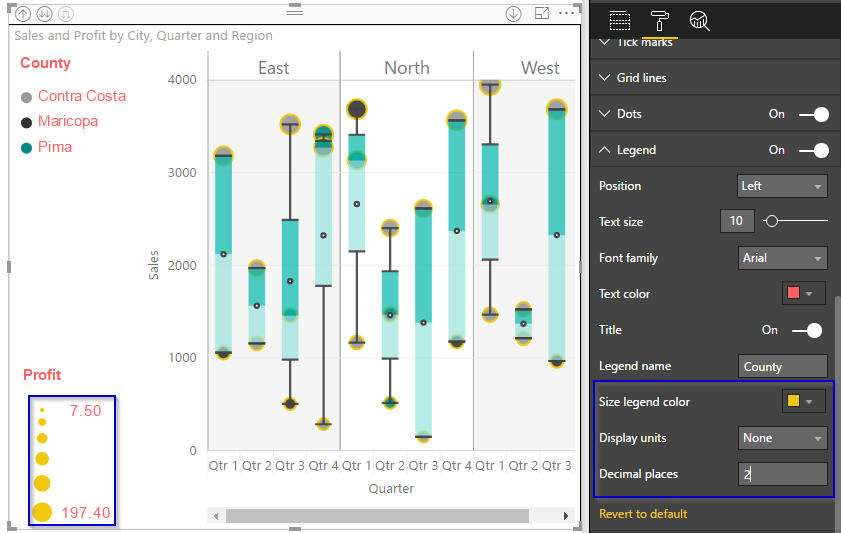
* 1. Users can update the radius range of the dots, as well as the hover color and transparency.



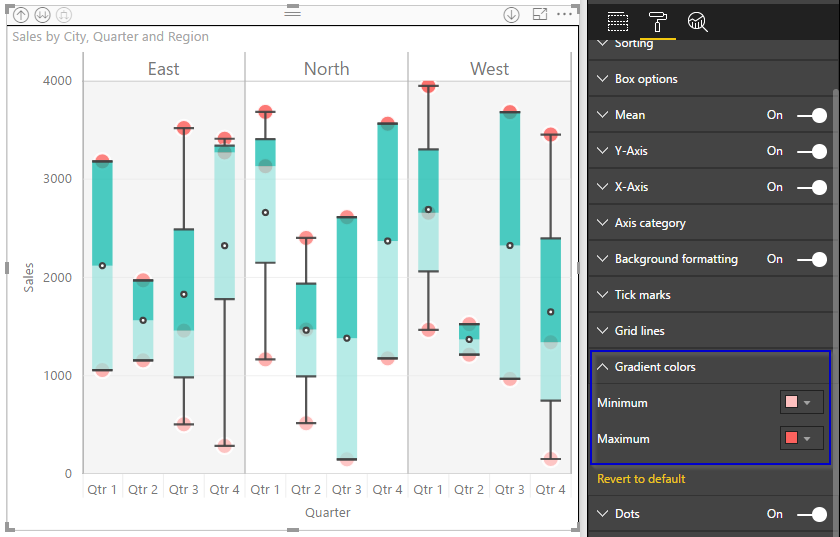
1. Formatting options for the 'Legend':
   1. The position of the 'Legend' can be altered, as well as the 'Title' and its color, size, and font. An option is provided to toggle the title on or off.



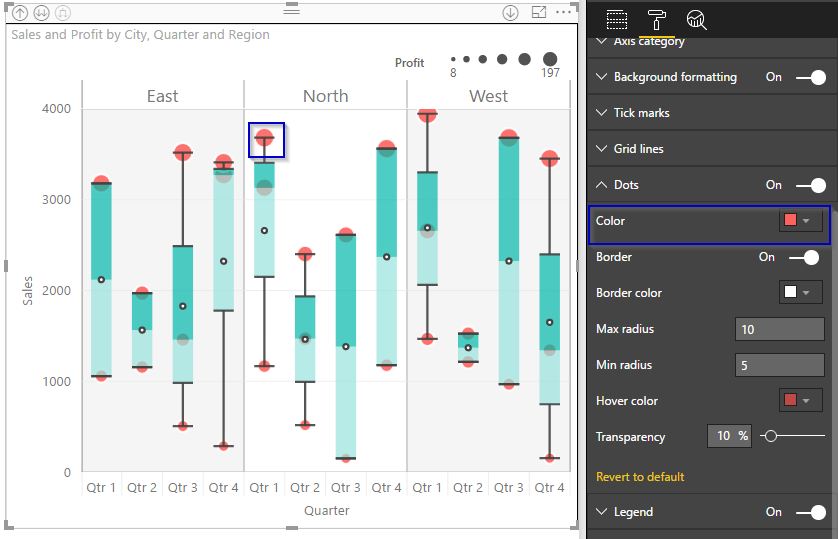
* 1. Color, display units, and decimal places for the 'Size legend' can be updated.



1. If a measure is selected instead of a category in the 'Legend' field, options appear for 'Gradient colors.' In this view, the color of each dot appears on a gradient scale, ranging from the minimum to maximum value.



1. If the legend is absent, the 'Dot Colors' option disappears. In this case, users can set the colors of the dots using the 'Dots' formatting option.



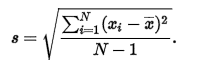
# Calculations Used

1. **Order** data in numerical order before calculation
2. **Median** = (n+1)/2th element; //if n is odd

**Median** = ((n/2th element) + (n/2 + 1th element)) / 2; //if n is even

1. Inter Quartile Range(IQR) = Q3 - Q1
2. Find median of whole data (**Q2**), divides data into two halves
3. Find medians of these two halves, (**Q1<Q3**), divides the data into quarters
4. **Whisker Choices**:
   1. Min/Max – Whiskers on minimum and maximum value
   2. = 1.5 IQR – Whiskers on Q3 + 1.5 IQR and Q1 – 1.5 IQR
   3. < 1.5 IQR – Whisker on dots <= Q3 + 1.5 IQR and >= Q1 – 1.5 IQR
   4. One Standard Deviation
      1. Upper whisker – Max(Mean + sigma, Q3)
      2. Lower whisker – Min(Mean – sigma, Q1)

Sigma,



* 1. Custom percentile calculation, example below for Pth percentile:
     1. The first step is to compute the rank (R) of the Pth percentile. This is done using the following formula:

R = P/100 x (N + 1)

where P is the desired percentile and N is the number of values.

* + 1. If R is an integer, the Pth percentile is the number with rank R. When R is not an integer, we compute the Pth percentile by interpolation as follows:

1. Define IR as the integer portion of R (the number to the left of the decimal point).
2. Define FR as the fractional portion of R.
3. Find the scores with Rank IR and with Rank IR + 1.
4. Interpolate by multiplying the difference between the scores by FR and add the result to the lower score.