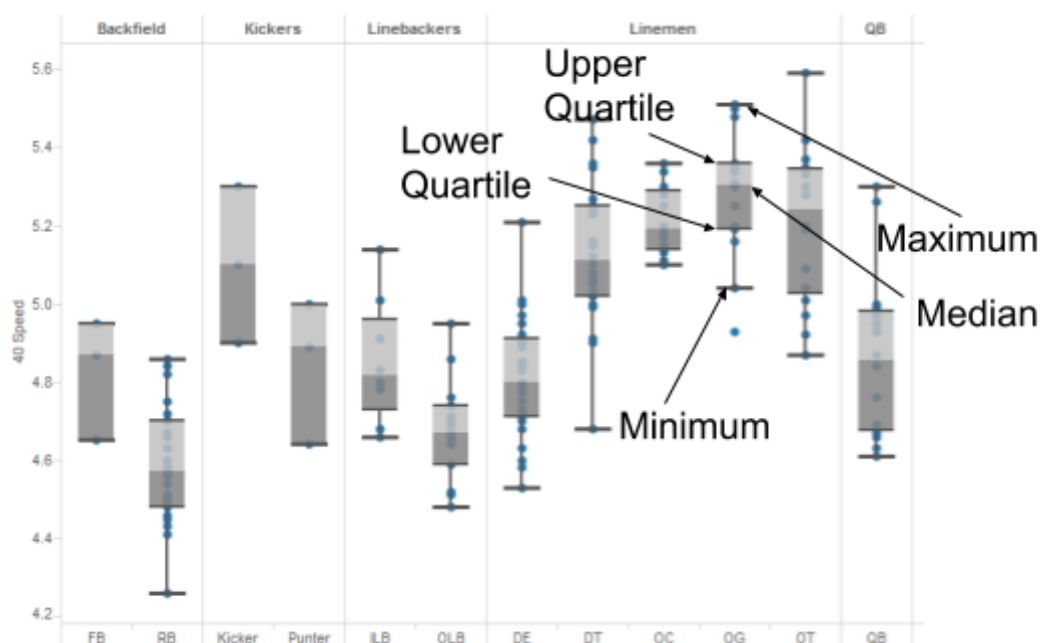


# TABLEAU

Day 12

## Box and Whisker Plot

Box and whisker plots are a great visualization chart if you want to infer important points by looking at the chart. You can easily find an outlier (if any) from box and whisker plot. A box and whisker plot is a way of summarizing a set of data measured on an interval scale. It is often used in Exploratory Data Analysis (EDA). In a box and whisker plot: the ends of the box are the upper and lower quartiles, so the box spans the interquartile range. the median is marked by a vertical line inside the box.



IQR stands for Interquartile Range, which are the data points between the first and third quartile.  $IQR = Q3 - Q1$  where Q3 and Q1 are the upper and lower quartile respectively.

### How to check if there is an outlier in our data points?

If the minimum is less than  $Q1 - 1.5 \cdot IQR$ , then it is an outlier or if the maximum is more than  $Q3 + 1.5 \cdot IQR$ , then also it is an outlier.

The most important line is the one right in the middle of each “box”, which represents the median. With median displayed, you can quickly look across the dimension members and compare medians, regardless of how big or small the range of values is within each column.

### Interpreting a Boxplot:

- The middle line in the box indicates 50% of the data i.e median value of the data. The upper edge of the box indicates 75th percentile of the data set, and lower hinge indicates 25th percentile.
- If the median line within the box is not equidistant from the hinges, then the data is skewed.
- The ends of the vertical lines or “whiskers” indicate the minimum and maximum data values, unless outliers are present in which case the whiskers extend to a maximum of 1.5 times the interquartile range.
- The points outside the ends of the whiskers are outliers or suspected outliers.

### How to create a box and whisker plot in Tableau?

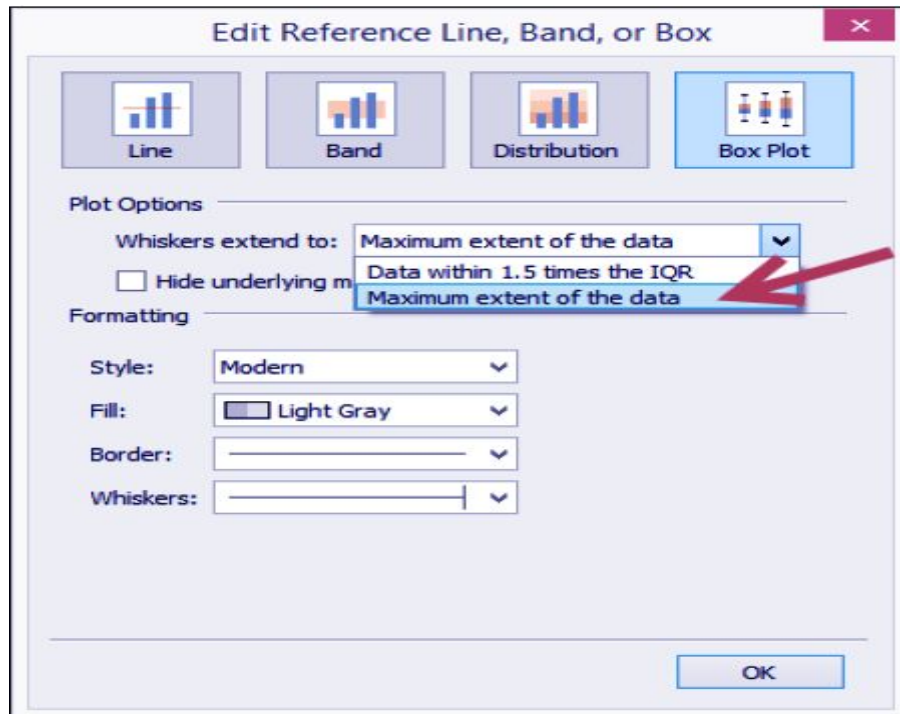
#### Using Show Me:

**Step 1:** Put dimension field in column section and measure field in rows section.

**Step 2:** In the show me option, choose the type of chart as box and whisker.



**Step 3:** A box will open up.



So the default options are telling Tableau to make all of the data points on the box-and-whisker plot fit into 1.5 times the IQR; anything outside of that range is an outlier.

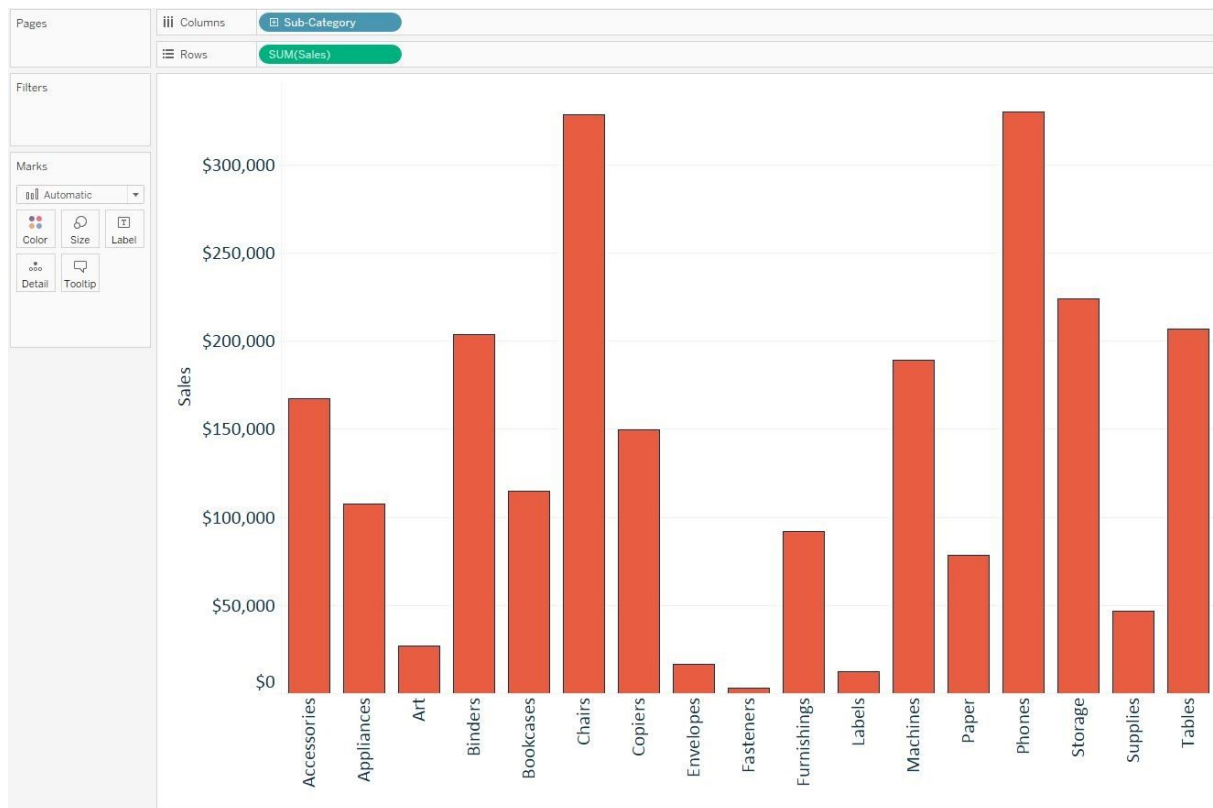
**Step 4:** Add more details to the chart by putting sub-category in the details section and you can also change the size of the circle by clicking on Size in the Marks card to make the Circle marks smaller.

That's it, your box and whisker plot is ready.

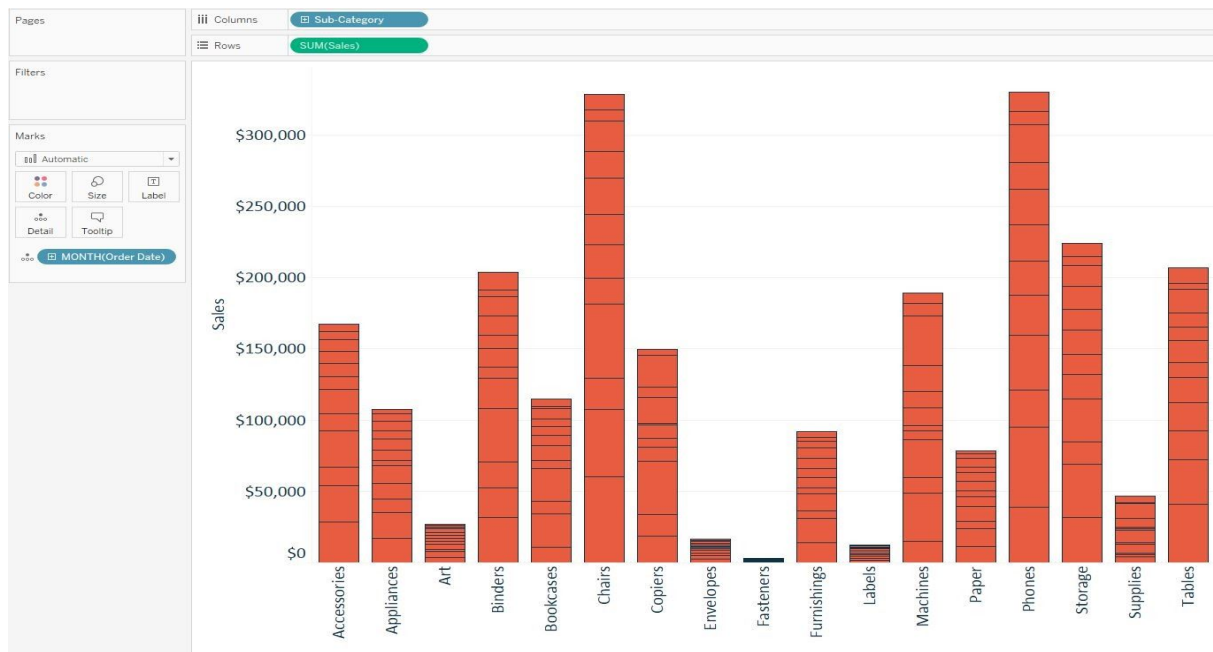
### **Without using Show Me:**

Box and whisker plot can be created using reference lines.

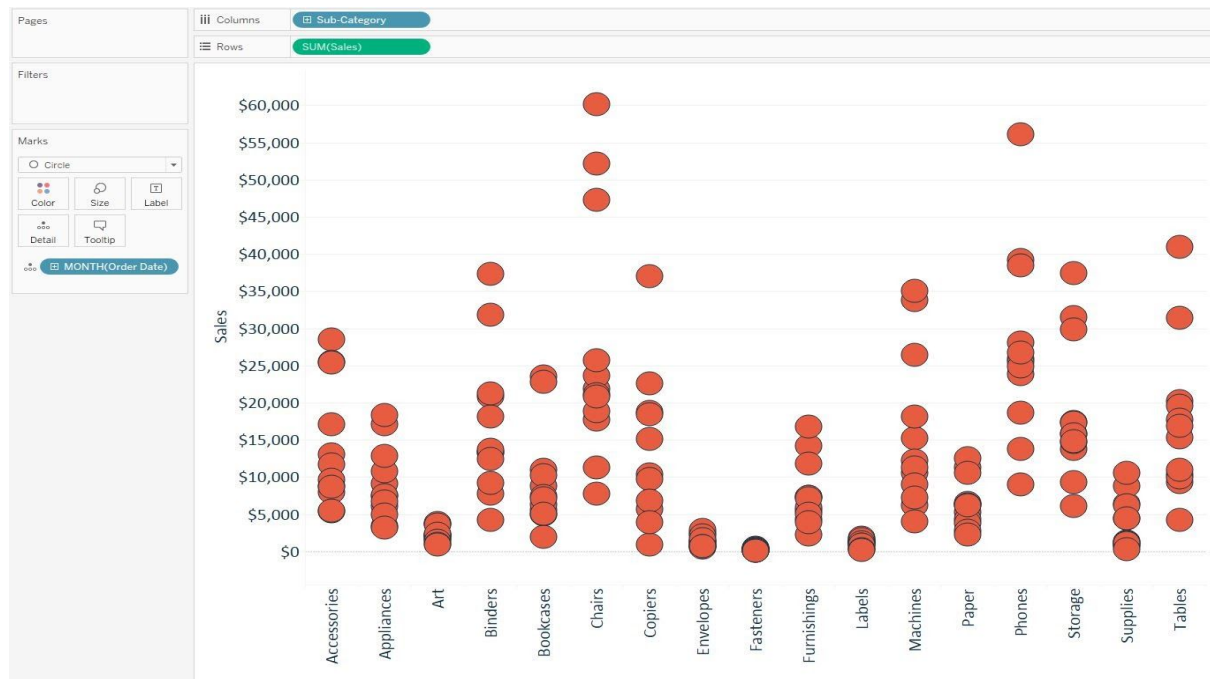
To create a box-and-whisker plot, start by creating a bar chart with the dimension and measure of interest. In the example above, we are looking at Sales by Sub-Category.



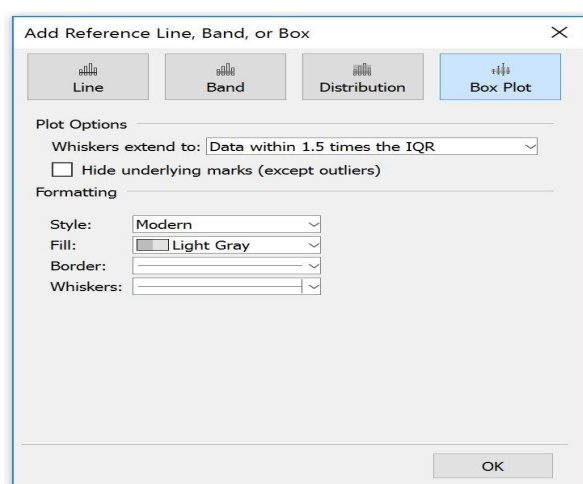
Next, add the distribution that you care about to the Detail Marks Card. In this case, we are looking at how Sales are distributed by Sub-Category, by *Month of Order Date*. So in this example, Month of Order Date is added to the Detail Marks Card.



By increasing the level of detail, a stacked bar chart is created, with each stack per sub-category representing one of the twelve months of the year. To set the foundation of the box-and-whisker plot, convert this stacked bar chart to a dot plot by changing the mark type from Automatic (Bar), to Circle.



Lastly, to create a box-and-whisker plot, right-click on the Y-Axis, and choose “Add Reference Line”. When the add reference line dialog box appears, click on the choice for Box Plot. There are some formatting options available, but the default settings are usually best:



*Happy learning, happy coding!*

