TABLEAU

Day 22-23

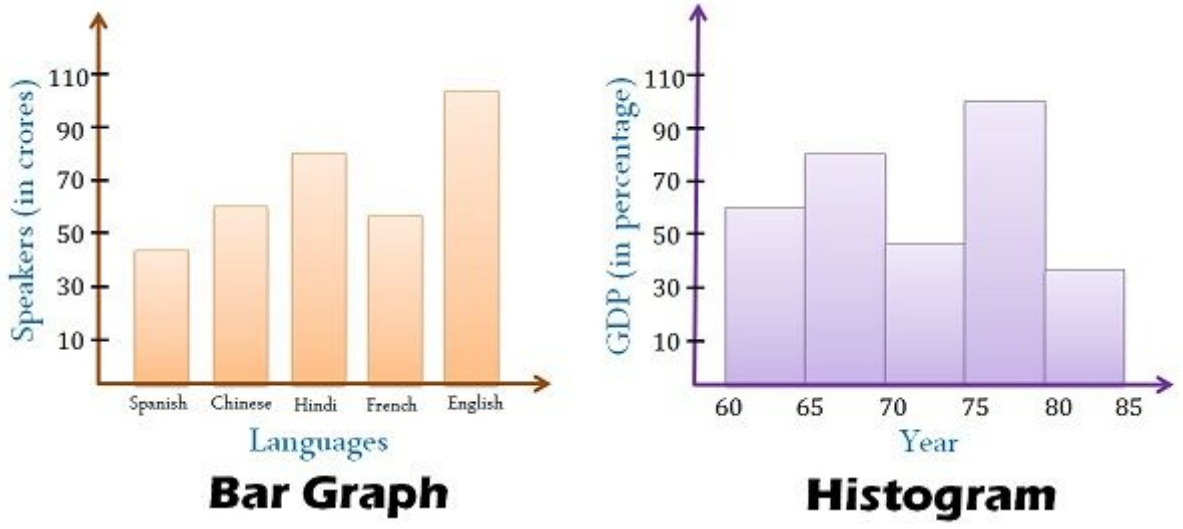
Histogram

The histogram is a diagram consisting of rectangles whose area is proportional to the frequency of a variable and whose width is equal to the class interval. They show the distribution of your data by dividing it into bins of equal size and assigning each bin a bar that is proportional to the number of data points that fit within that bin.

Histograms can be summarized roughly as an inventory of what "kinds of items" you have and "how many of each kind" you have.
In a histogram, the total range of the data set (i.e from minimum value to maximum value) is divided into 8 to 15 equal parts. These equal parts are known as bins or class intervals. *When we have **huge datasets** it can be easily visualized using a histogram.*

It often happens that we confuse a bar graph with a histogram.

How is histogram different from a bar graph?



- There are gaps between bars in a bar graph but in the histogram, the bars are adjacent to each other.
- Histogram refers to a graphical representation, that displays data by way of bars to show the frequency of numerical data whereas, Bar graph is a pictorial representation of data that uses bars to compare different categories of data.
- Histograms are distribution of non-discrete, continuous values whereas bar graphs are distribution of discrete, categorical values.

- Width of the bars in histogram need not be the same whereas in the bar graph the width needs to be the same.
- In histogram, bars cannot be reordered whereas in bar graphs it can be reordered.

Histogram is used to represent statistical information by way of bars to show the frequency distribution of continuous data. It indicates the number of observations which lie in-between the range of values, known as class or bin.

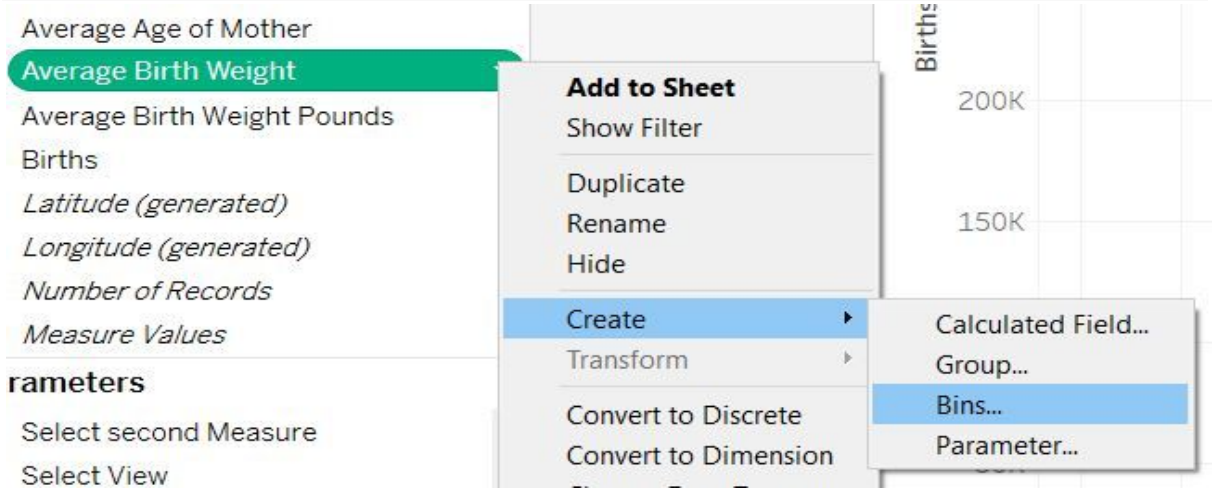
When should you use a histogram?

Histogram is only used to plot the frequency of score occurrences in a continuous data set that has been divided into classes, called bins.

How to create histograms in Tableau?

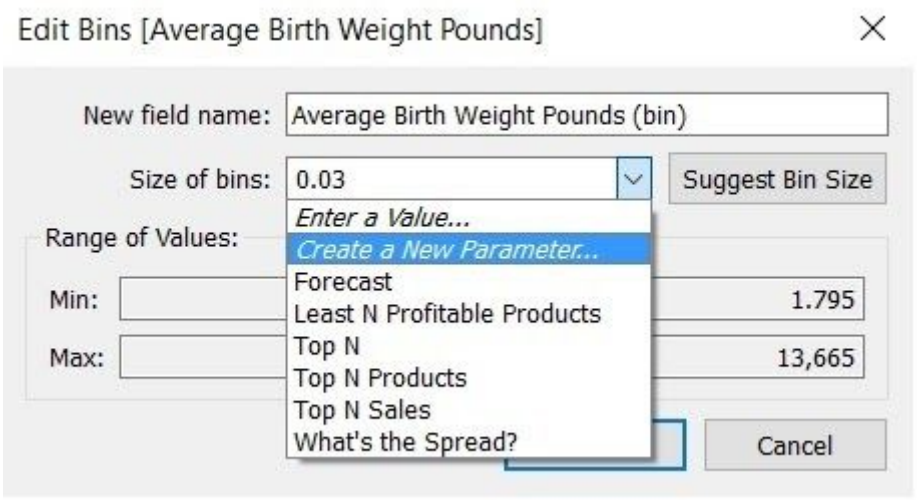
Step 1: Building your bins

In order to build our histogram, our first step is going to build our bins within which our data points will then be distributed. Click on the little arrow at right side of your *Measure* of interest in your data pane – scroll down to the *Create* option, and select *Bins.*

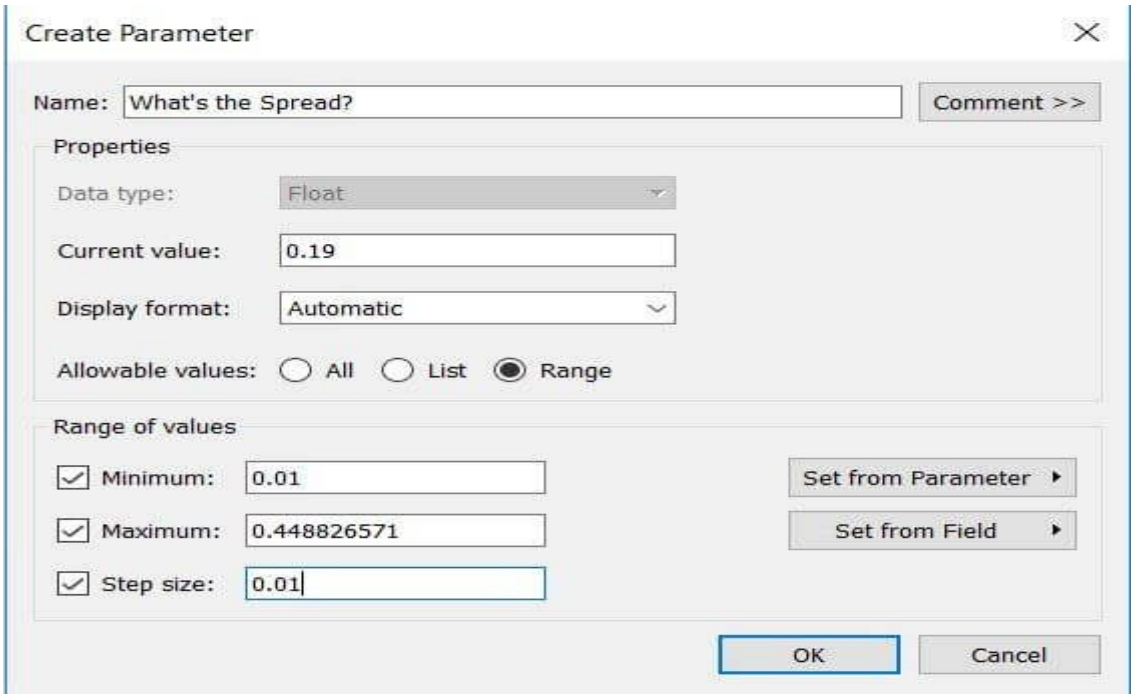


Step 2: Setting up your parameter control

- After you've hit Bins, you'll be prompted with a configuration window. All you're going to want to do is click on the Size of Bins menu and select Create a New Parameter.



- Once you've hit Create a New Parameter, you'll be prompted with another configuration window. You can leave all the prepopulated values – Tableau will estimate what the minimum, maximum, and step size should be – or adjust them to your own liking. I am going to change the minimum value and step size to 0.01 and leave the maximum value as is. This means that the user of my dashboard will be able to use the parameter control to adjust the size bins in increments of 0.01, with a minimum size of 0.01 and a maximum size of 0.449.
- Remember to give your parameter a name so your future user can easily understand what he is adjusting when he plays around with the slider!

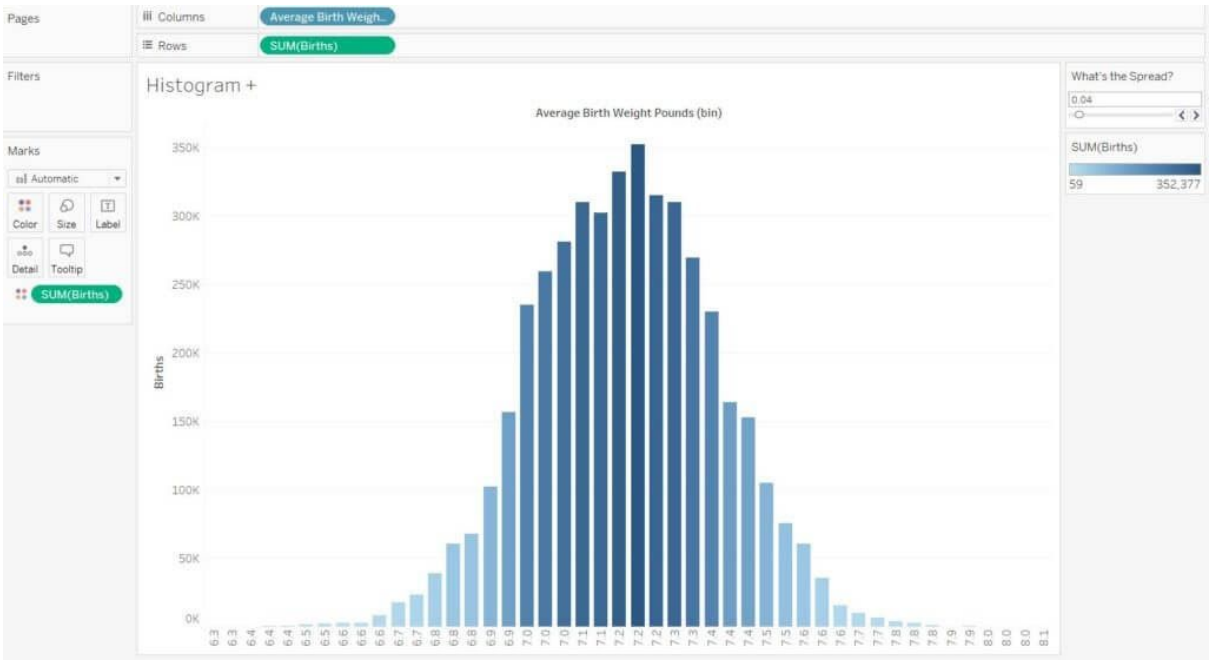


- Hit okay, and you should see your parameter control appear on the top-right of your screen. You can hover over it, hit the little caret, and change the display option to Slider for extra usability and clarity. It should end up looking like this:

Step 3: Putting it all together!

Now that you've set up your bins and configured your parameter control, you're ready to build your histogram!

- Drag your key measure into Rows –and the bins Dimension into columns.
- You can drag your key measure unto Color in the marks card to add some additional visual information and show your viewer where the majority of your data points are concentrated.
- Congo! You have learned to create an awesome histogram chart!

- 

Disadvantages of histogram:

- Cannot read exact values because data is grouped into categories.
- More difficult to compare two data sets.
- Use only with continuous data.

Bonus tip:

How to update your worksheet changes with a timestamp when changes are made to it?

1. Open the workbook that contains the dashboard in Tableau Desktop, and then navigate to the sheet for which you want to display the time of the last data refresh.
2. Select **Worksheet** > **Show Title**.
3. Double-click the title.
4. In the Edit Title dialog box, select **Insert** > **Data Update Time**, and then click **OK**.

The Data Update Time value depends on the type of connection:

- Extract-based connections display the last time the extract was refreshed.
- Live database connections display the last time the data source was refreshed.
- Live file-based connections display the last time the file was updated.