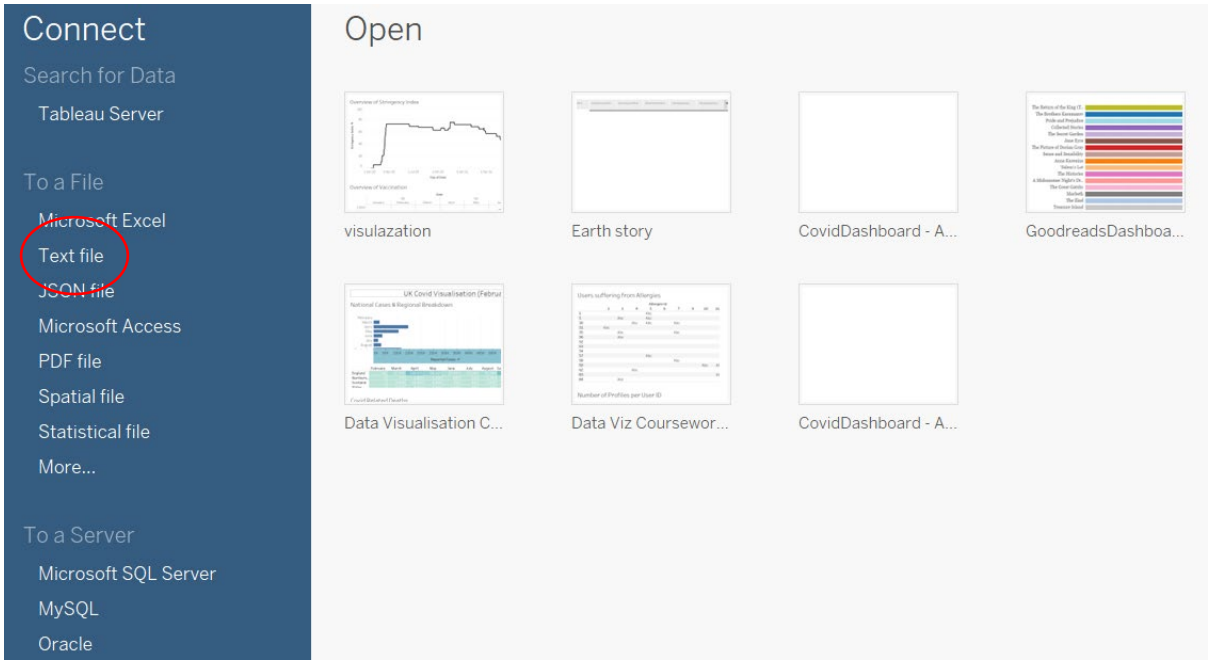
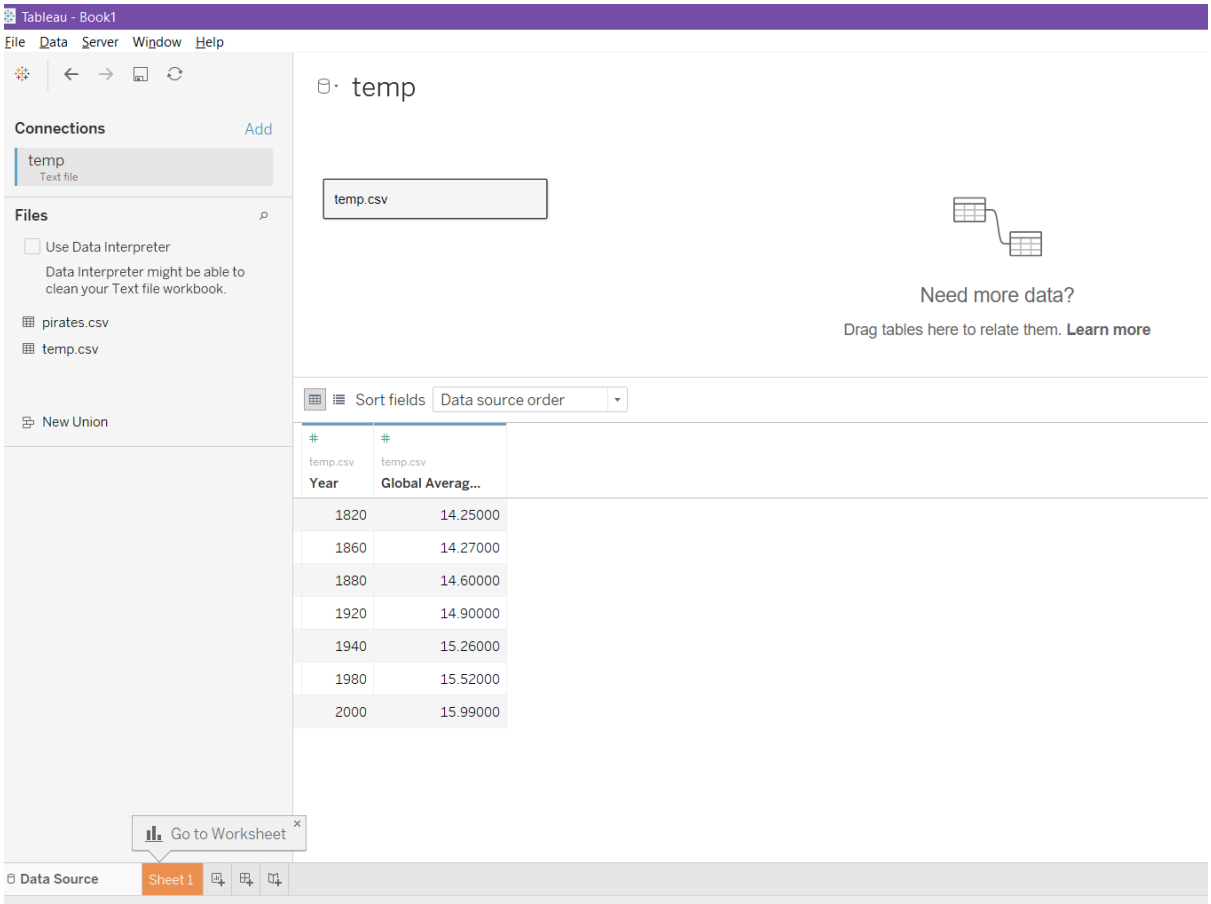
Topic 1 Lab Data Viz Tutorial (Tableau, 2023.1)

In this tutorial, you will learn the basics on how to load data and do simple plots using **Tableau**.

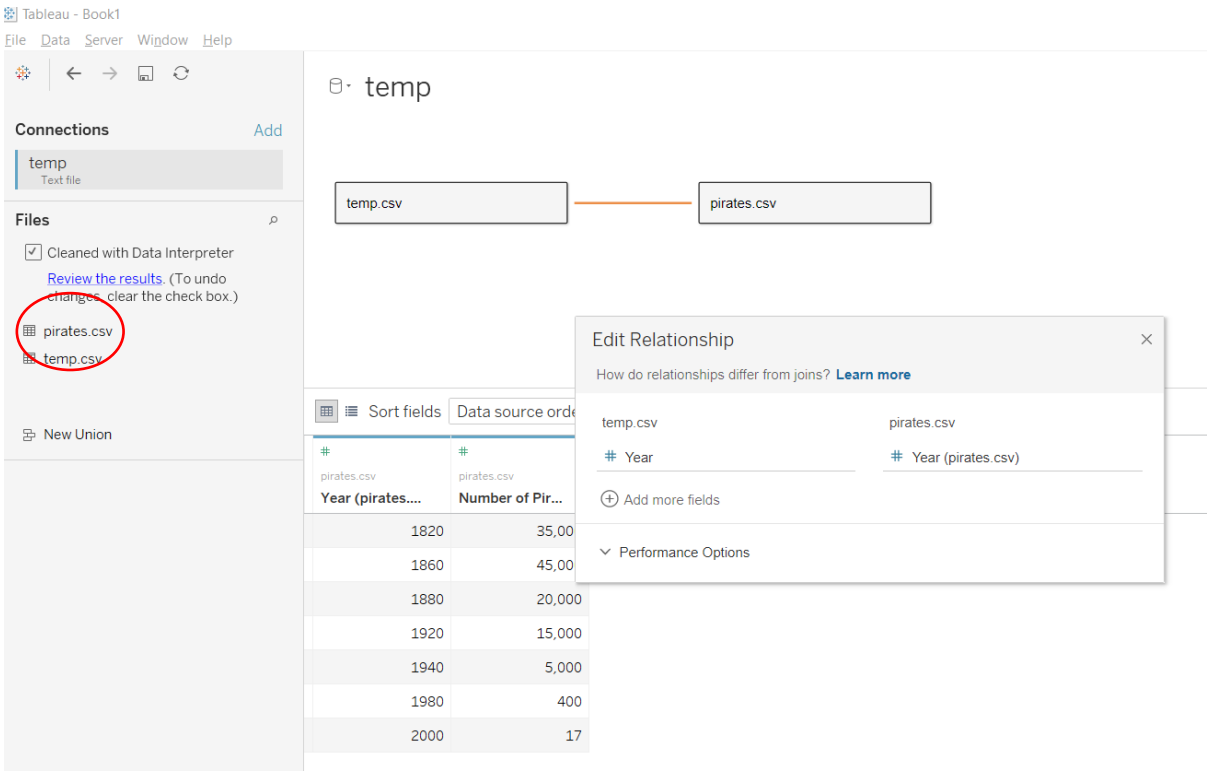
1. Open Tableau and **Connect → To a File** selecting **Text File**



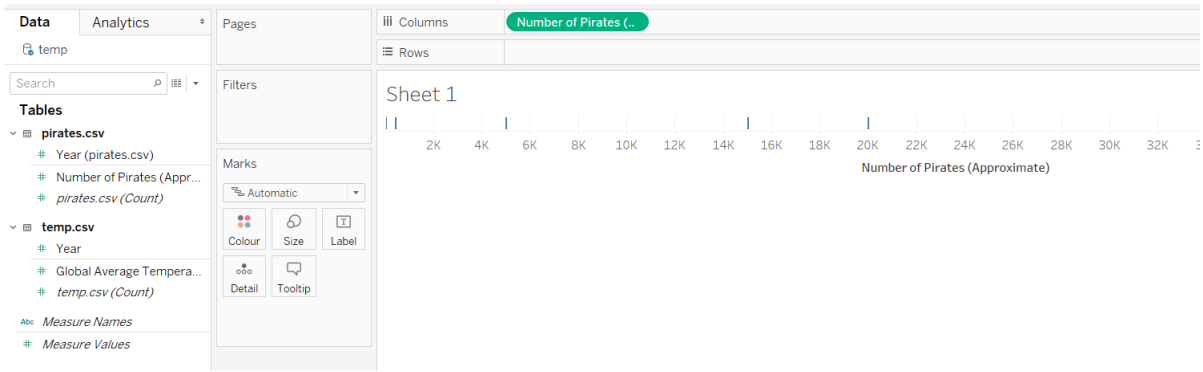
2. Select temp.csv. This will display a Data Source page (you can see the tab names on the bottom-left corner) where you can see a preview of the data.



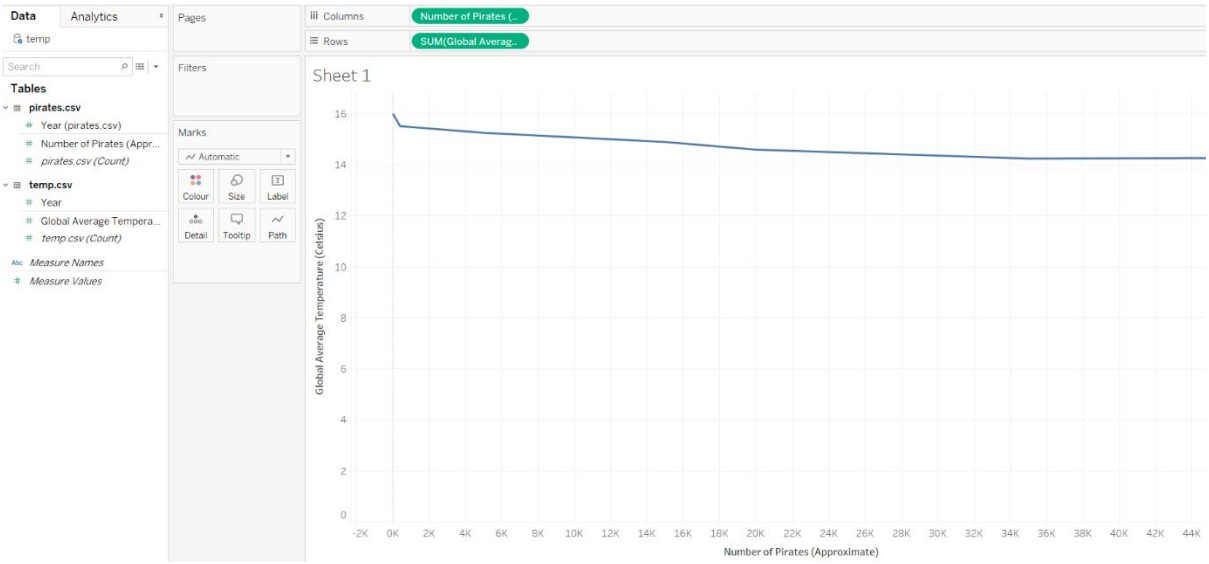
3. As you can see, there is a message in the centre of the screen asking you for more data! You can drag & drop pirates.csv in here. This will link the two files. Also, Tableau may have "seen" pirates.csv is in the same folder, so you can click on the file name in the left menu. This will open the following wizard, where you can see that the relation will be made by the year column:



4. Go to the worksheet by clicking on the orange tab in the bottom left corner. You will see a workspace where you can drag & drop variables from the tables to create visualisations on the fly. To do the first Pastafarian plot, we need to have the pirates on the x-axis and the global temperature on the y-axis; however, we also need to sort the pirate numbers by year! To do so, firstly drag & drop the *Number of Pirates (Approximate)* variable to the columns box (on top of the workspace).
NOTE: If the variable is changed to *SUM(Number of Pirates (Approximate))*, right click on it and change it from a *measure* to a *dimension* until SUM(disappears (it still needs to show it in green).

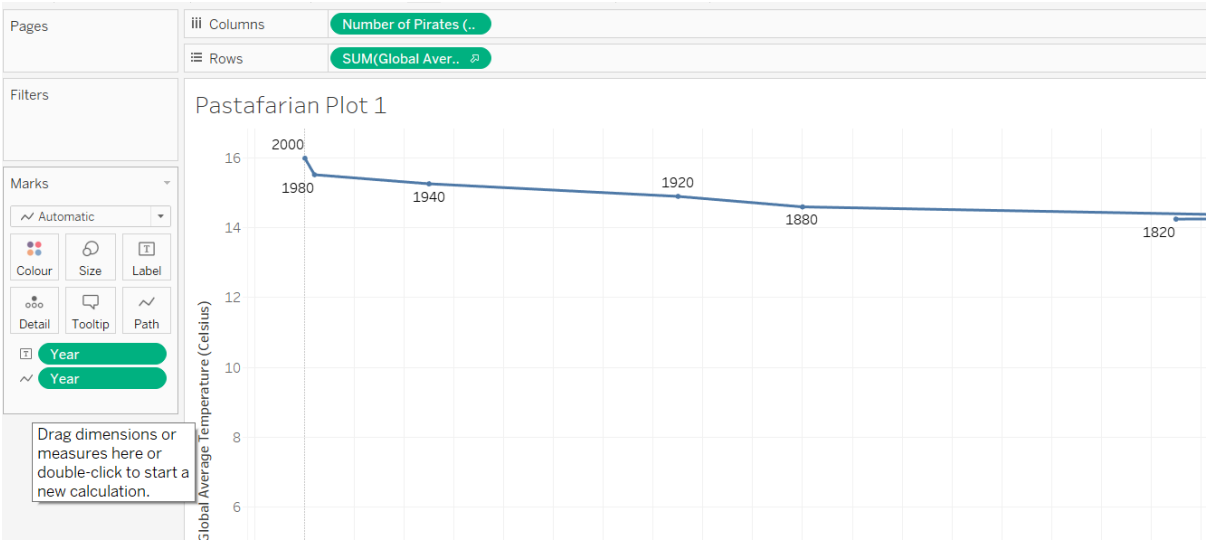


5. Now, drag & drop the *Global Average Temperature (Celsius)* variable into the rows section. By default, it will say that it is plotting the sum (other statistics can be used such as mean, median, etc). This will draw a line plot automatically.

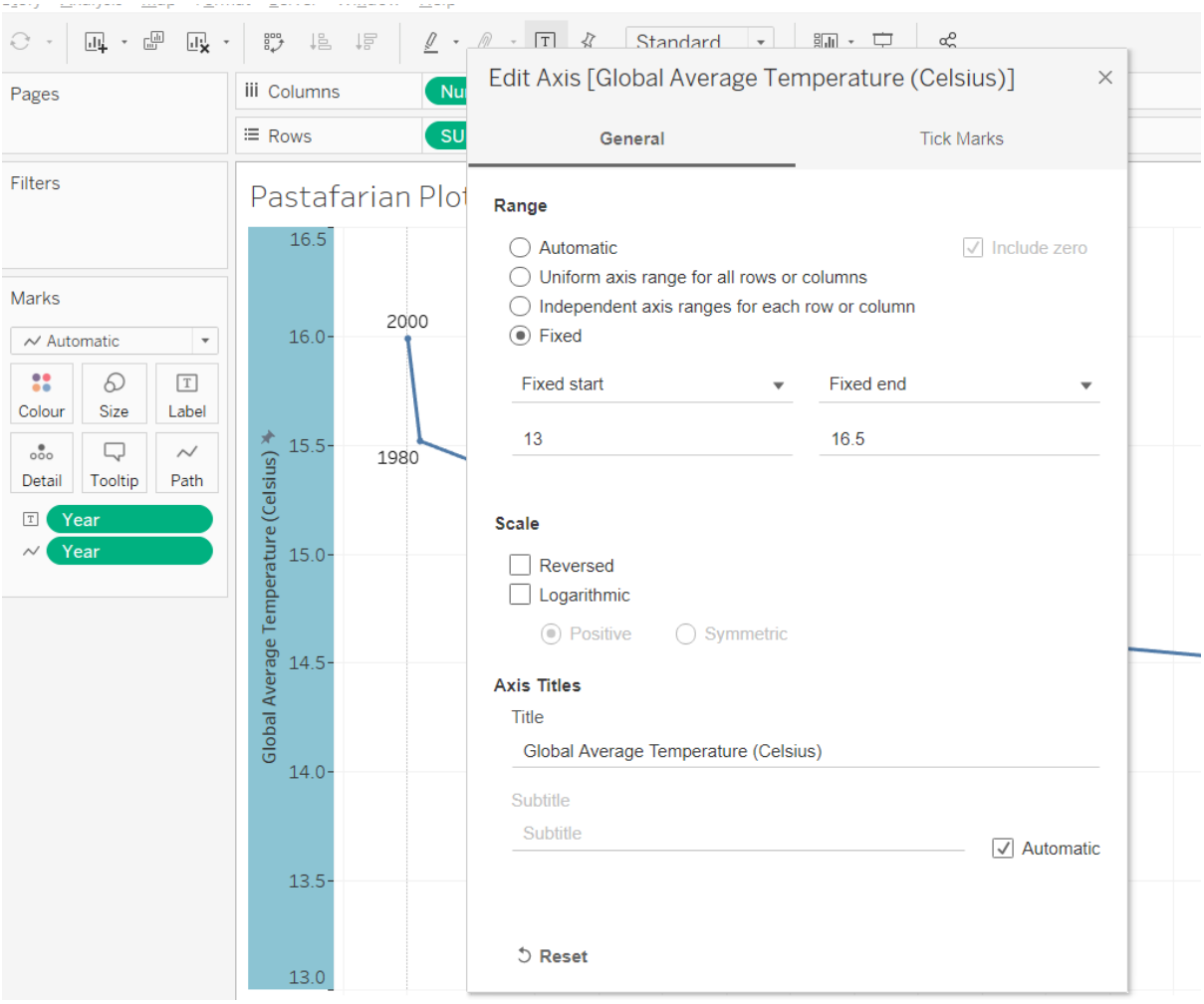


6. Drag & drop any of the *Year* variables (remember, they are connected, so any of them work equally!) to the *Marks* section (below the six options) and click on the four-circle symbol to the left of the *Year* variable, so that you can change it from **Detail** to **Label**. You will see that the years appear over each data point (and that the line disappears). Notice that markers are not sorted the way we want.

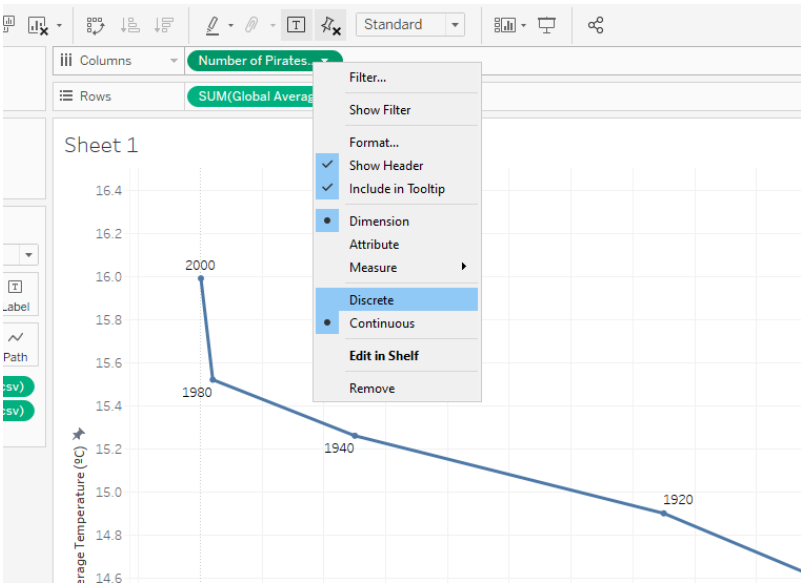
7. You can add the line again by adding *Year* in the marks section for a second time, and changing the marks type to **Path**. Again, this confirms that the plot is not in "chronological" order!

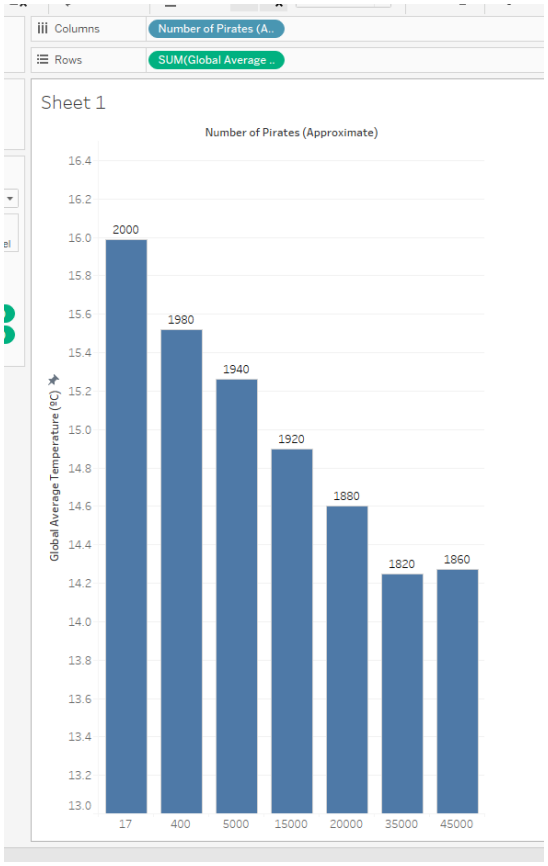


8. To mantain the aesthetics, right-click on the y-axis and select **Edit Axis**. Here, you can change the range to 13 – 16.5:

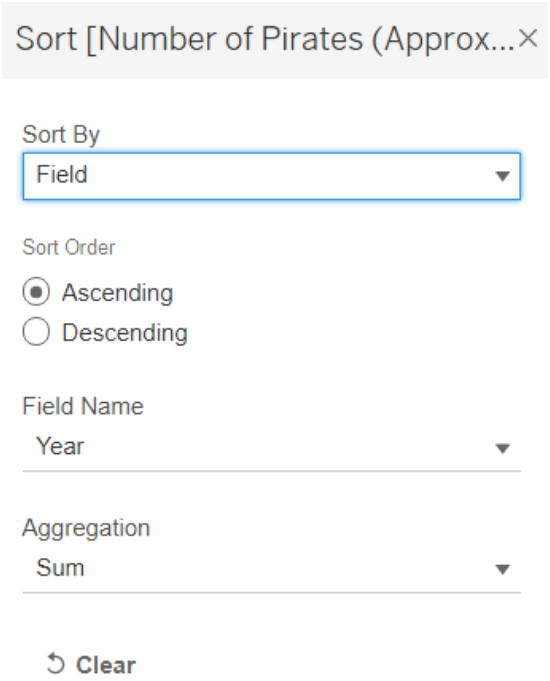


9. The data points are sorted since Tableau (and most data visualisation tools) treats the number of pirates as continuous values, which, in theory, should be sorted along the x-axis. One trick to change this is to convert this variable into a **discrete** one. This will convert your plot into a bar one! Don’t panic; this is normal…





10. To sort the values in chronological order, right click on the *Number of Pirates* variable (which you dragged to the columns area on step 4 and is now coloured blue) and select **Sort.** Here, you can indicate that you want to sort according to a **Field** (in this case, Year) in ascending order and using the **Sum** aggregation.

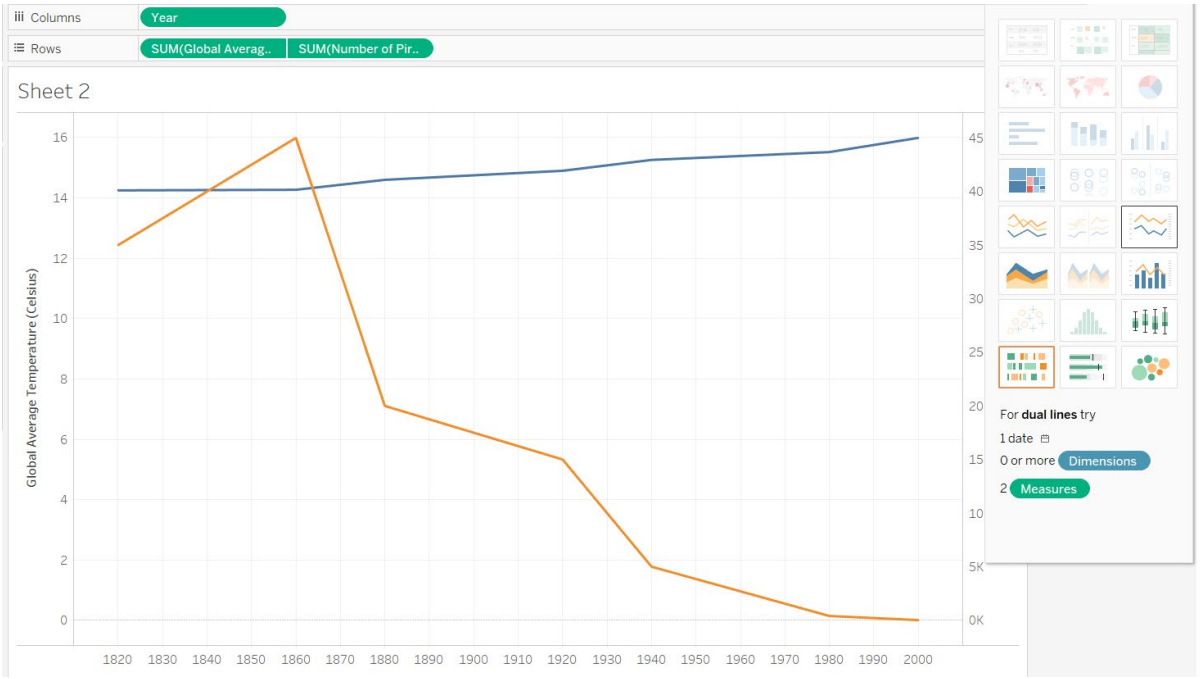


11. Finally, below the **Marks** section there is a drop-down menu which shows the “Automatic” option. This is the one that was converting our dots into bars! You can revert this by clicking on this menu and selecting **Line.**
    NOTE: If the line connecting the dots doesn’t appear, then have a go at it (I don’t remember how to do it…)

12. For the second Pastafarian plot, we will go to a new workspace. Just like in Excel, you can add a second sheet using the menu on the bottom-left corner. In here, we will first add *Year* as a column, since in this case we want two lines plotted with respect to this variable.

13. Add both *Number of Pirates* and *Global Temperature* as row data. Keep in mind that you need to change *Number of Pirates* back to a continuous numeric data type, and that both variables should appear as sum:

14. To put both lines in the same plotting space, click on **Show me** (top-right corner) and change the plot to a “dual lines” one:



15. Have a go at “beautifying” the plot!