In-class assignment #5

CMSE 402, Data Visualization Principles and Techniques
Spring 2022

Purpose: The purpose of this assignment is to for you to learn to manipulate plots, including modifying axes, annotations, and creating subplots. We will do all of this using the same dataset that you used for the pre-class assignment, `congress-terms.csv`, which is included with this assignment for your convenience.

Part 1: Using the Pyplot 2D histogram method to make a 2D histogram of ages of incoming congresspeople as a function of the number of the congress (a 2D histogram is also referred to as a "heat map"). Use the `bins` argument to choose an appropriate number of bins for the x and y-dimensions separately. Now, do the following:

1. Label the x and y axes and give the plot a title. Use the matplotlib text properties arguments to change the size, color, weight (i.e., boldness), and position of the text. Note the coordinates used when you set positions!

2. Try changing the color map, using this page of color maps as a guide.

3. Use the pyplot `gca` method to get a copy of the current Axes class for the figure you have created. This can be used to modify the plot axes in a variety of interesting ways - consult this section of the Axes API description for more information. Try setting the range of the axis (`set_xlim` and `set_ylim`), the location of the ticks (`set_xticks` and `set_yticks`), and manually set the tick labels (`set_xticklabels` and `set_yticklabels`). Try to create tick marks that are unevenly spaced, and to change the size of the tick labels! You can also have no label for a tick mark by giving an empty entry in a list. Note that if you know how to use LaTeX, you can get matplotlib to give you any LaTeXsymbol you want using this syntax: `r'$\alpha$'`, with the letter 'r' before the quotes telling it to use TeX formatting (demo here).

4. Finally, try putting text and arrows inside the plot using Pyplot's annotation features. Note that the coordinate system for these annotations uses the data values, not the plot coordinate system (i.e., use numbers that make sense given the congress and ages, not in the 0-1 range).

Part 2: Using the same dataset, create a `subplot` with a single row and two panels – the left panel is a 1D histogram of ages of all congresspeople, and the right panel is the 2D histogram from above. Use the techniques discussed in the previous section to modify the two subplots as you see fit (but make sure to make separate modifications to each subplot, and give each subplot its own title). Use the `tight_layout` method to ensure reasonable spacing between the subplots so that you can read all of the labels, and use `suptitle()` to give the entire figure a title.

Handing in the assignment: Turn in all this assignment via the GitHub classroom repository where you found these instructions. You can do this by committing your assignment and pushing it – i.e., in the directory that contains your assignment repository type "`git add (name of file containing writeup)`", then "`git commit (filename)`", and then "`git push`". The last step is critical, since that pushes your changes to GitHub where I can see and grade them!