INFO 4310 HW 1

Amy Ouyang (aco49)

Website: [*https://info4310-hw1-pwwg.onrender.com*](https://info4310-hw1-pwwg.onrender.com)

This site tells the story of the different types of species of trees in San Francisco. The viewer can first get an overview of the most populated 12 tree species in San Francisco on the map. One can see where different the top 12 tree species occupy neighborhoods. The top 12 tree species in San Francisco represent 17,167 trees out of the 37,119. These 12 species of 404 species represent 46% of the trees in San Francisco. The legend is right below the map and denotes the name of each species with a color hue to differentiate each species. The legend also includes the number of trees in that species. Below this legend, a bar chart is shown which allows the viewer to see and understand the number of tree species in a different representation that is simpler. This change in visualization draws out the quantitative side of the data. Following the bar chart, there are 3 line charts which show when the 3 tree species with the most trees are planted over time. These line charts are to show the progression of various trees over time and elaborate on when these trees have been planted so one can see trends per species. These different visualizations tell the story of tree species throughout physical space and passing time.

I utilized Street_Tree_List-2022-01-30_FILTERED.csv as my data for the map and bar chart. To process the data, for the column 'qSpecies', I kept the common name of the tree species by splitting the string on "::" and taking the second element. I then kept a dictionary of the number of trees in each species and only kept the top 12 dictionary items with the highest values. For the line charts, I used the same data set but processed the data differently. I cleaned the column 'PlantDate' by not accounting for those with null or empty string values. The column 'PlantDate' had the date and a time of '00:00' that I spliced and then formatted to become just the year. While many points do not have 'PlantDate' with a year, it is a trend worth noting.

The marks in the map of San Francisco are circles that denote a tree. Visual channels in the map are the color hue and position on a common scale. The color hue of the circles, which represent what species the tree is. The position is on the scale of the earth of Longitude and Latitude on the map shown by horizontal alignment and vertical alignment. On the bar chart, the marks are lines/bars. The visual channels are vertical length and horizontal positions. For the line chart, the marks were the lines and the visual channels were color hue, and horizontal and vertical alignment in position.

I made certain design choices to aid in understanding the story of tree species. For the map, I made the design choice to keep the background of the states black as this contrasts with the circle marks that have various color hues. With a lighter background, the marks get drowned in the visual channel and are difficult to differentiate. The legend and the color hues were chosen such that each hue is noticeably different. The legend denotes the number of trees with each species to help understand the groupings of trees in the map. These color choices communicate the location of various species. In the bar chart, the x axis labels are skewed slightly to allow the visualization to still be compact and easily comparable. If the labels were not skewed, the graph would be quite wide and spread too thin making it hard to see all the data. The same colors as the previous legend are used to identify groups quicker than just the x axis labels. The total number of trees are in text above the bar for the related tree species. This number allows the user to compare bars similar in height and identify gaps or closeness in numbers of trees. In the series of line charts, I chose to only show the Top 3 tree species planted over time as showing too many lines on the same graph would make the visual be too cluttered. Placing the 3 species in one line graph with the same x and y-axis allows one to see trends in planting in one species or multiple species. I then chose to show each line graph per species planted over time separately so that one can then see the time trends individually. These line charts tell the viewer the growth and expansion of the top tree species over time.