Christopher Sun

INFO 4310 - HW 1

While looking at the data and the different columns of the filtered version of the data, I was first curious to see how a map of the tree data points would change over the years. Since I was scrolling a bit through the data, I found that many of the "PlantDate" entries were left empty for the trees, meaning that many of these trees were planted before the year 1955. Since, in much more recent years, our environment has been slowly crumbling and more and more people are becoming aware of the detrimental effects humans have on the environment, I was curious to see whether people have been actually trying to aid the world, which, in this example, is planting trees. After getting a basic map of all the trees, I had to decide how to represent the age of the trees planted on the map. At first I was thinking if it would be possible to create a color scale and make a legend to show how old each tree is, but after looking at the first map and how small each data point is, as well as the fact that we can't add any interaction, I decided this wouldn't be the best idea. Since there's a large density of trees in many spots of the map, a color scale might make it very hard to point out differences in age in these clusters and might end up cluttering up the map more, making it visually hard to look at. I then decided it would be best to measure the year based on the opacity of each point. Since I mainly wanted to focus on the recent years after 1955, any points that were planted before that could essentially be ignored.

In order to actually get dates for points without a plant date, I assigned an arbitrary value of January 1st, 1950, since the goal was to make older trees completely transparent. Using d3.timeParse and then converting the given PlantDate data to a datetime object, I was able to extract the year planted for each tree using d3.timeFormat("%Y"), since year was what was most important and not the months or days themselves. After seeing both maps next to each other on

the same page, I had to choose what color I wanted the maps to be as well as what color the points would be, as I originally defaulted to black points with a gray background. Without the ability to interact, mainly pan and zoom closer into the map, I needed to make sure that the opacity would be noticeable enough to be seen from afar. In general, the user can definitely see the drastic change from the first map to the second, with a lot of areas missing dots and the density of trees getting smaller in a lot of areas. I experimented with a lot of colors, testing to see which would not only be pleasing to the eye and have a noticeable opacity effect. Since the map was a representation of San Francisco, I decided on making the color representative of that, making the points, which were small circles, a shade of red. Each point would feel more like an obtainable "thing" on the map, representing an item of San Francisco as well as accomplish my goal of having a noticeable opacity effect. The background color was kept as gray, for it didn't make the map too flashy and it paired nicely with the red.

After this, I doubled back to the original goal, which was seeing if people have been trying to aid the world. In order to connect the beginning maps to this, I used the "qCaretaker" column, which identifies whether the association that takes care of the tree is private or the DPW. The base idea for the next graph was a graph that could show the number of trees at each age, which I could then divide into separate counts, one for private caretakers and another for DPW caretakers. In order to do this, I had to go through the data in JavaScript. First I added a new variable to each point, an age variable, measuring the difference between the year 2022 and the year planted to find the age. In order to collect the count of trees for each age division, I created several arrays with the different possible ages as indices in the arrays, namely 0 to 67. Since trees planted before 1955 were given a date planted of 1950, any tree planted in 1955 and before were put in the 67 age category, which was really the 67 and above category. I then went through all

the data in a for loop that went from 0 to 67, and when the age of the point matched with a certain i value in the for loop, 1 would be added to the index i of the array. After the count was made for each age, the graph was created. The simplest and neatest way to visualize these counts, I thought, would be through a bar graph. Not only that, but dividing a single bar in a bar graph into separate colors to show the caretaker division would be easily readable. Each age would be on the x axis with the count of trees on the y axis. I decided that putting a tick for each age would be adding too much clutter to the x-axis and left it in divisions of 5, as the bars themselves were very clear and evenly spaced. A user could easily count to tell what age each bar is at based on the ticks.

In order to make the divisions of each bar, I used the previous arrays I had made. The main array was the total counts of trees at each age, but I also made another array which contained the count of just the private caretaker trees at each age. With this, I was able to replicate the bar graph making process for the private caretaker trees. I put this process after the original one, so that the private caretaker colored bar graph would overlap the original. When deciding on colors I decided that the two easiest contrasting colors would be blue and red, to easily tell the difference. I thought that it would be too confusing to have any other colors as the categories, as even though the caretakers aren't "opposites" blue and red are very clear in distinction and are easily considered representative of different groups. The original bar graph was made with blue, so that when the red overlapped, the blue part of the original graph that was left would represent the DPW caretakers, with the new red bar being the private caretakers. With the new divisions, the user can clearly see in specific years or age of the tree which caretaker held the majority.

In the end, what was seen is that the DPW rarely held the majority of trees in any age, except for trees of age 14. I decided to put in an after-note describing this in the data, as it would've been clear, but maybe the reason wouldn't have been. Since these are trees of age 14, this means that this was during the wildfires of California in 2008, which were incredibly severe and burned hundreds of acres of land. In this sense, it makes sense that the government, or the DPW, decided to step in and help the environment by supplying California with an abundance of new trees, showing their care for the people and its environment. The problem is, though, the consistent pattern of private caretakers being the majority in the other ages. Although the goal was to see whether people were stepping up to take care of our environment, which, as seen in the bar graph, they are, it seems that the DPW isn't doing as much as it could be doing when compared to regular citizens or agencies.

I created another final graph, displaying the final age of 67 and above, since this was a majority of the data. Because including this with the rest of the data would've ruined the scale of the graph and made it impossible to read, these had to have a separate graph. I also decided that instead of making it resemble the graph of data for the other ages, I essentially made the bar horizontal on its own. This is because I thought this was the most effective use of space without adding too many unnecessary graphics such as more axes or gridlines. Unlike the first two maps, for these two bar graphs I used color-coding legends placed near each bar graph to specify what meant what. For the first two maps I simply added small annotations after the title of each map to show the number of points on each map as well as describe what was being put on the map. With the order of both the maps and bar graphs, the story is communicated to the user in an orderly fashion. First, with the maps, it's seen how in recent years San Francisco has definitely planted more trees, but even so a handful of these trees are old and even more are older, having been

planted before 1955. After, with the bar graphs, we can see who is responsible for the planting and taking care of each of these trees, young and old. Through both of these we can see how much we're trying to help the environment and we can make a judgement on how much more either, or both, parties need to increase the trees planted in order to protect our environment.