# Static Visualization HW 1 : Design Document

EXPLORATION

I began by importing the data into RStudio to get an initial look at the two datasets. As I explored the columns by looking at their "summary()" data, I was intrigued by the "PlantDate" variable. I was curious about whether the distribution of new trees was different across the years. A quick histogram (with faulty year translations before 1950) indicated that there was something interesting going on.

The research question then became: Does the existence of new versus old trees differ in the various neighborhoods of San Francisco?
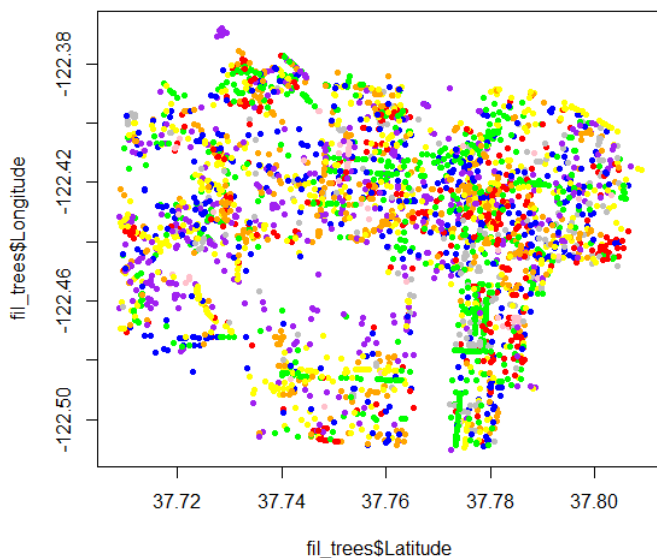
Histogram of fil_trees$PlantDate
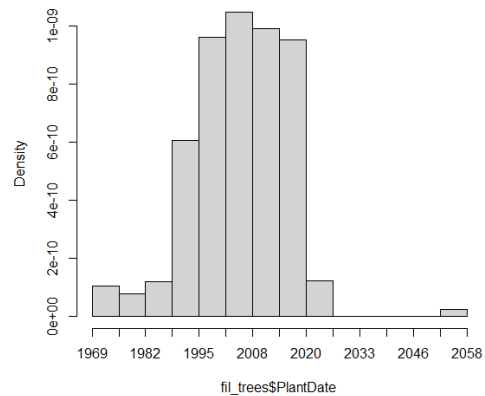
After removing all the rows without a PlantDate, I assigned different years to different colors. I then generated a scatterplot as a quick proof of concept. I could tell that there were sections more dominated by some colors than others.

From there, I moved into recreating the visualization with D3, but with a clearer vision of how to display the age of the trees depicted.

DATA CLEANING

I started by cleaning the data and preparing it for the visualization. I removed all the data points that didn't have a valid "PlantDate," since including them (even if grayed out) would muddle the visualization. Next, I converted the "PlantDate" values into Date format and grabbed just the year from output.

COLOR ENCODINGS

After setting up the map of the San Francisco and its neighborhoods, it was time to decide what each tree mark should look like.

The most natural way we understand new growth is by examining what trees look like in the spring. The pre-existing growth from previous year is a darker, deeper, and bluer shade of green, whereas the new growth from the present year's spring is a brighter, lighter, and yellower color green.

The image on the right shows the inspiration for the color scheme I chose. I excluded a couple of greens that I felt would be difficult to distinguish from the others if they were in different parts of the map.
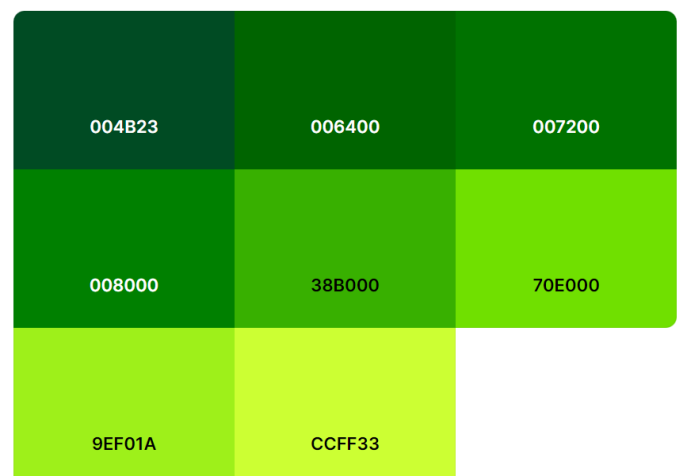
DESIGN AND LEGEND

I decided to make the dots for each tree large enough to distinguish the color, but not so big as to obscure other tree points.

I included a legend that showed the beginning year and the most recent year of recorded tree planting dates, as well as the colors that spanned the years in between.

THE STORY

Although I could have indicated the specific years associated with the other colors in the legend, the point of the visualization was not to be able to pick out when each individual tree was planted. Rather, the viewer naturally looks for clusters of newer or older trees. These clusters tell a story about the various neighborhoods across San Francisco.

Lines of darker trees at the southern part of the map indicate older, more established neighborhoods. Moreover, the clusters of bright green on the western part of map suggest newer neighborhoods that may have been something else before that wouldn't have had trees (e.g. industrial zones). Another straight. diagonal line of medium-darkness trees in the northeast corner along a coastline suggest a boardwalk that was put together around the turn of the century.

1956   2022