

Story

I wanted to convey the fluctuating, yet overall increasing, scale of tree planting efforts in San Francisco. There were more trees from before 1955 than have been recently planted (although the visuals don't cover if those older trees were natural or from earlier tree planting efforts). I wanted to highlight how the tree planting efforts have changed from no new trees being planted, to multiple spikes in trees scattered throughout the years. Additionally, I wanted to show the most common species of trees are in San Francisco, and when they were popular: some species had constant numbers planted over the years, while others had large spikes in certain years. The latest year shown in the data, 2021, had a greater number of trees planted overall than 1956.

Processing Data

I used the filtered dataset because I wanted to be able to utilize the species attribute in my visuals. To process the data, I added a new column that extracted the year from "PlantDate" to be able to make the line charts, with any NULL values to be set as 1955 (according to the data dictionary provided). To acquire the counts of trees for each year, I used the function rollups which provided a mapping of the year to the number of trees planted in that year. Additionally, I removed any data from 2022 since the provided dataset only had 3 trees planted that year (all in January). To avoid the perception that trees were planted less in 2022 (I cannot confirm or deny this is the case), I filtered 2022 out and only focused on the years 1956 to 2021 for all the charts.

Visual Encodings

For the **"Trees Planted Before and After 1955"** bar chart, the marks are the bars that represent the number of trees planted during the given period. The channels include the length of the bar on a horizontal, aligned scale that corresponds to the magnitude of the trees planted. For the **"Number of Trees Planted Each Year"** line chart, the mark is the line that indicates the number of trees planted and its continuity over time. The channels include horizontal and vertical positions on an aligned scale and the length of the lines to show changes during the provided

duration. For the **“Top 6 Most Frequent Species Among New Trees”** line charts, the marks are the lines that indicate the number of trees planted and its continuity over time. The channels include horizontal and vertical positions on an aligned scale, the length of the lines to show changes during the provided duration, and the color hue to differentiate the tree species.

Design Choices

For the Most Frequent Species chart, my initial idea was to have on chart to show the different species of trees planted overtime. However, it was difficult to interpret without adding some interactivity to distinguish the lines and compare them adequately, even with unique colors. Thus, I opted to make multiple smaller line charts that each house a unique species, allowing the user to compare them without cluttering the view. I kept the colors to help differentiate them in the multiple charts. Moreover, I ensured that the species charts all used the same scale to make comparing the lines more intuitive. Based on feedback from the critique session, I decided to decrease the size of the charts so 3 could fit in a row compared to the initial 2 per row. This makes it easier to notice and compare all of them simultaneously without having to scroll down too far. Additionally, I decided to include a bar chart comparing the number of trees before and after 1955 to ease the user into the focus on after 1955. Despite “newly” planted trees being a quarter of all trees in the city, the line charts give perspective on the rate at which new trees are being planted. While my charts show the number of trees planted each year and even broken down by species, there are other directions that the visuals are not showing, like potential differences each year in a geographical lens. The way I approached the visuals aggregates counts and lacks the dimension of location. An alternative visual I had in mind, but decided not to implement, was a network with caretakers and species, mapping each caretaker to the tree species they cared for with the link’s thickness dependent on the number of trees. However, I thought this would be difficult to implement as a static visual and felt like it did not contribute too much to the tree-planting story I had in mind.