**As you design, document your design choices and rationales**. Document the story you want to tell and how the design choices you make help to strengthen it. If you processed the data, describe how you did it. Were there alternate ways of representing you considered? Did you encounter any trade-offs in showing what you wanted to show? What potential messages or areas of the data are sacrificed or hidden by the design choices you made?

- A 2-4 paragraph write-up in PDF form that:

- Outlines the story you want to tell with the data  
- Explains any processing you did of the data  
- Identifies the *visual encodings* you use in your image to link data to visual channels  
- Provides a rigorous rationale for your design choices and explain how they help to facilitate the communication of the story you want to tell.

**Story**

The story I want to convey is the scale of tree planting efforts. I wanted to highlight the small, yet increasing, effort for tree planting. While I don’t know the full complexity of the situation in San Francisco (cost, land, etc.), I wanted to show how tree planting effort has changed over the years by looking at the big picture and then breaking it down by most common species.

**Processing Data**

The data I used was the Filtered dataset because I wanted to be able to use the species data later on. To process the data, I added a new column that converted the “PlantDate” into just the year to be able to make a time series, with any NULL values 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. For the species graphs, I extracted the common name (after the “::”) from the “qSpecies” column for each of the trees.

The “newly” planted tree data does not include the year 2022, because only a few trees were planted then (the data only went up to January). I decided not to include it in order to avoid the perception that trees were planted less in 2022 because the data for the full year was not provided.

**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**

I initially had the idea of creating 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, since it might be too cluttered and difficult to interpret the number of trees.

I also had the idea of using one chart to show the different species of trees planted over time, but when I created it, it was difficult to interpret without adding some interactivity to distinguish the lines. Thus, I opted to make multiple smaller line charts that each house a unique species, so you are able to compare them without cluttering the view.