Design Rationale

Using the data, I wanted to convey the differences in tree density or population amongst various San Francisco neighborhoods. This was done using a choropleth map to visually compare the total number of trees in each neighborhood. Preprocessing of the dataset was done in JavaScript in the original index.html file. Data processing involved importing the original trees CSV dataset, looping through each neighborhood object and checking if each tree falls inside it using d3.geoContains(), and then creating a new JSON dataset that details the counts of trees in each neighborhood with the structure {neighborhood name : number of trees}. This dataset was named num_trees.JSON and later imported back into the index.html file for further analysis.

The choropleth map employs the San Francisco neighborhood shapes as marks that visually and spatially represent each neighborhood. A white mesh or white outlines for each neighborhood polygon were used to provide a clean, minimal feel for the map that puts the focus on the neighborhood shapes and their colors. Visual channels of color hue and saturation are used to denote the total number of trees in each neighborhood. Both channels were used in conjunction to provide more visual contrast between the quantile bins. Moreover, 6 quantiles (sextiles) were chosen for the bins as this number of bins provides for enough granular detail while also not introducing too many bins that could confuse the viewer. D3.scaleQuantile() was used instead of d3.scaleQuantize() to provide a more even distribution of colors on the map and data points in each bin. Colors range from a light green to a dark blue, with darker colors representing a higher volume of trees. Natural or earthy colors like green were chosen to match the subject matter of trees, and to serve as a visual reminder that the dataset aims to convey information on neighborhood greenery. These colors are also displayed as a legend in a horizontal bar/scale format at the top of the SVG to allow the viewer to identify the numerical quantity of trees in each neighborhood. The threshold values for each bin are placed at the breakpoints in the legend scale to denote the numerical cutoffs.

Using a choropleth map that employs color as a visual channel provides viewers with a simple basis for comparison. Initially, I considered conveying this information by creating a

bubble map that plots each tree as a point on the map using their longitude and latitude. However, points are often small and overlap with this method, making it hard for the viewer to count the total number of trees in a neighborhood. Moreover, many points sit on the border between two neighborhoods, making it difficult to distinguish which point belongs to which neighborhood. A choropleth map solves these issues by doing the counting for the viewer, using color to compare values for two neighborhoods. However, since there are only 6 bins and no hover interactivity, viewers are unable to determine the exact number of trees that sit in each neighborhood and where they are located. This is a design trade-off, where precision was sacrificed for simplicity and ease of comparison.