Emir Polat
Professor Rzeszotarski
INFO 3300
8 February 2024

INFO 3300 — Homework 1 Write-up

Part 1: Story

Using the dataset of municipal trees in San Francisco, my visualization unfolds a narrative of growth and collaboration, emphasizing the evolving dynamics of tree ownership and maintenance on San Francisco's city sidewalks. This exploration delves into species diversity, ownership dynamics, and the concerted efforts to enhance the city's green infrastructure. My goal is to shed light on the crucial roles both public and private sectors play in fostering and enhancing San Francisco's urban forest, and underline the importance of partnership in urban environmental management.

Part 2: Data Processing

I processed the dataset "Street_Tree_List-2022-01-30_RAW.csv" and generated two distinct datasets: "annual_summary.csv" and "caretaker_and_species.csv." To generate the "annual_summary.csv," I excluded the rows lacking "PlantDate" data or not marked under "qCaretaker" as either "Private" or "DPW." Subsequently, the "PlantDate" field was transformed into datetime format to isolate the year component. The data was then classified by year, leveraging aggregation functions to count and categorize trees separately based on private and DPW ownership. This preprocessing stage helped me to capture the distribution and ownership of newly planted trees up to the year 2021.

To generate the "caretaker_and_species.csv," the preprocessing began with the exclusion of entries missing species information ("qSpecies"). The dataset was further refined to include only trees under the care of "Private" or "DPW" entities. The "qSpecies" field underwent a transformation to extract and clean the species name by splitting on ' :: ' and selecting the last element. Additionally, any entries ambiguously categorized under "Tree(s) ::" were excluded. The cleaned dataset was then grouped by caretaker and species, with a count of occurrences computed for each group. This aggregation made it easier to study the distribution of trees owned by different caretakers.

Part 3: Design Rationale

The stacked bar chart illustrates the number of trees planted each year from 1969 to 2022, categorizing them based on ownership. It differentiates between trees maintained by the Department of Public Works (DPW), represented in blue, and those under private care, depicted in green. This chart effectively showcases the annual efforts to enhance San Francisco's urban

forest, emphasizing the joint contributions of public and private entities in these greening initiatives. When I read about San Francisco reclaiming the ownership and maintenance of nearly 200,000 trees along city sidewalks in 2017, I became curious about the ownership of newly planted trees over multiple years. Although I initially considered using a map visualization, I ultimately decided on a stacked bar chart because it allows me to display multiple aspects of the data in one visualization.

The pie chart provides a comprehensive breakdown of tree species within San Francisco's urban forest, classifying them by their prevalence across different caretakers, namely the Department of Public Works (DPW) and private owners. This visual representation underscores the biodiversity of the city's green spaces. Each segment of the chart is color-coded to differentiate species, offering an intuitive understanding of species distribution.

Part 4: Visual Encodings

The stacked bar chart utilizes position, length, and color visual channels to convey complex data in an accessible manner. The x-axis represents year, while the y-axis quantifies the number of trees planted each year, with the height of each bar indicating the total count and the colors distinguishing between DPW and private ownership. This visual encoding allows for a direct comparison across years, highlighting trends in planting efforts and the evolving landscape of tree ownership. The consistent width of the bars across the chart ensures readability and uniform comparison, making it easy for viewers to trace the growth of San Francisco's urban forest over time.

The pie charts utilize angle, color, and area visual channels together to illustrate the distribution of the most common five tree species under DPW ownership and private ownership.The angle and area of each pie slice indicate the proportion of each species relative to the total, providing a clear visual representation of their prevalence in San Francisco's tree population. Distinct colors for each species enhance the chart's readability, enabling quick identification and comparison of the species. This visual strategy effectively communicates the diversity of the trees within different ownerships.