Story and Topic

According to the dataset of trees in SF, each tree has a species type and site information (where the tree resides). The data visualisation tries to explore the distribution of species of all the trees that reside on the sidewalk. There are six types of sites under the sidewalk category. They are firstly divided by whether they are on the curbside of the sidewalk or the property side of the sidewalk. They are also divided by how they are planted: in the yard, in the pot or cutout.

Data Process

The first step in processing the data is to find different attributes I am interested in and figure out the types and number of each type in that column. After looking at several columns, I decide to explore the species and site information about the trees. However, there are over 400 species and around 20 types of site information. I then find that in the site information, the sidewalk is where most trees reside and there are six subtype site information in the sidewalk category. One question comes to me: do all the subtypes of sidewalk have the same species distribution? Hence, I process the data to figure out the species distribution in each subtype of sidewalk. I only choose the top 5 species in each category and drop the rest because they represent the main feature of each sidewalk type. Since I focus on species and site information only, I also delete irrelevant columns.

Design Rationale

For the map, I choose different colours of circles on the map to display types of site information. The six subtypes of sidewalks can be categorised into two general types: curb site and property side. Since the curtsied is more dangerous, I choose warm(red-related) colours to represent. On the contrary, the property side types use cold colours. Since the number of points of each type differs a lot, types with more numbers have lower opacity and smaller radius while types with fewer numbers have higher opacity, darker and brighter colour, and bigger radius, making it easier for viewers to find each type. The map can show the number difference of each type vividly and how trees with different site information are located.

For the six bar charts, I choose to use the same colour sets as are shown on the map. The visual channel is the difference in the length of the rectangle. And they have an aligned vertical axis. In each bar chart, there are the top 5 tree species with names and particular numbers and they are sorted from high to low. From the comparison of each row and column, we can find some common tree types displayed in multiple bar charts.

Outcome

Based on the map, we can see the cutout plant at the curb site is the most on the sidewalk. On the property side of the sidewalk, most trees are in the yard.

The six bar charts display the top 5 tree types in each category. In most cases, different site information has different tree species. However, we can still find that on the curbside, New Zealand Xmas Tree and London Plane are quite popular. On the property side, Juniper shows in both pot and yard types. For the cutout plant, New Zealand Xmas and Victorian Box appear on both sides.