My visualization aims to tell the story of tree planting and forestation in SF in terms of the two main owners - the Department of Public Works (DPW) and private owners and companies. I wanted to show how a mass number of tree planting all across SF is done by private owners. Through the map, one can see that the DPW mostly plants trees on streets, which is represented by the straight rows of green. Private owners take up the majority of the map, and especially the border of Outer Sunset. Through further online research, my findings show that SF Public Works is only planting trees on SF streets. Furthermore, I traced the Outer Sunset phenomenon to a community event done in 2019 where anyone could plant a street at the neighborhood border through an event done by Friends of the Urban Forest, a non-profit group not associated with DPW. As a result, these trends are not a flaw in data but rather happening in real life. Through my other graphs it seems that most tree planting in recency is done by private groups. In the site order graph, one can notice that the newer trees (indicated by low DBH) and the trees planted in bulk (indicated by a high Site Order number) all are private. The older trees with high DBH are mostly owned by DPW - showing that most of the urban forest initiative is done by citizens. That trends is further seen by the Site Info graph, as it seems that most DPW trees were sidewalk cut out trees with large diameters at breast height while private trees are more varied. Also, it shows that many of the private trees that have been planted in recency are done in pots or the backyard. Finally, I wanted to note how the DPW trees are concentrated mostly in the 3-6 feet plot size range. It appears that most private planting initiatives are either not done through plots or are done in plots of great size. As a result, my visualization shows that more recent tree planting efforts are done by private owners in great bulks in non side-walk areas. Throughout my data story, I make the assumption that the age of a tree is dictated by the size of the DBH. Trees with a large diameter are older while trees with a small diameter are younger.

To process the data, I modified the script into a new one called post-process-all.py. For this script, I wanted to clean up the data to make sure there weren't extraneous variables. I limited the caretaker to only two types - "DPW" and "Private" because I noticed when looking through the data that it was very rare to see other types such as "Police Department" and "SFUSD". Similarly, because I wanted to see plot size, I filtered and reformatted the data. The plot size data was originally inconsistent with empty strings and strange formats like "60M". I wanted to limit the formats to one of two kinds - "width x height", and "Width:". That way, I was sure that I was consistently looking at the width of the plot. I reformatted the plot size to just be a single integer which would indicate the width in feet. Then I changed the long strings of text unde Site Info to be shorter and more legible - for example Back Yard: Yard just became Back Yard. I did not filter out the trees in other locations because that would limit the extensiveness of the map as I would not be able to show the locations of all the trees.

For the visual channels, I chose to go with a limited color palette of blue and orange, to show the maximum amount of contrast between Private and DPW and make trends easily findable. I made the background of the map gray to increase the contrast between the datapoints and the background for more easy distinction. Furthermore, I chose two visual forms, a map and

scatterplot to fully represent the data as the map represents coverage while the scatter represents both frequency and spread. The tradeoffs of the map was that there were too many trees to see clear area based trends, although the one benefit was that it became very clear there was a pattern to the DPW planted trees. The tradeoffs of the scatterplot was that it was difficult to distinguish between the orange and blue when they were layered on top of each other. As a result, the frequency itself is hard to clearly distinguish, however, the spread is very clear. Because there aren't many old trees, it's clear that the ones that do exist were planted by DPW at cut outs. What's hidden from the graphs is the true distribution of DPW versus private planted trees, as it's difficult to tell from the three scatter plots what the exact ratio is between the two. At the same time, the format does make it more clear who has been mostly planting the older, thicker, trees. Although it does not distinguish between similar frequencies, it does make the distinguish between vast differences in frequencies between the two kinds of care takers for older trees. Even though I could have done a stacked bar graph to examine the different distributions, it would have been difficult to do it at every single DBH. Finally, one last trade-off I made was deciding between representing age of the trees from its plant date versus through its DBH. I ultimately decided on DBH because many large trees were missing their plant date because this data had just started being collected recently. However, this does lead to the question of whether my observations are accurate - it could be that private trees were not mandated to be recorded until recently. The accuracy of the trends ultimately depends on the assumption that private and government trees are all recorded and accounted for. All in all, my data visualization shows the story of tree planting efforts in SF and how citizen led initiatives have made a lasting impact in the composition of trees in SF.