NAME: Awojobi Oluwaseyitan

Student ID: 11087659

Course: INFO 5709

PROJECT 3

INTRODUCTION:

The goal of our project is to derive insights from the trees dataset by designing an interactive visualization system using various interactive techniques within tableau. We are using a dataset called tree which has in it 9831 rows and 8 rows. To achieve this goal, using the tree file and the SVG map, I am using Tableau to create our interactive design and to specify locations or data points within our visualization.

STEPS

In solving this, our first step is to perform exploratory data analysis to understand the dataset we are working with. We observe a few missing values. We can however proceed with our visualization as the missing values plays little or no impact.

Next, the dataset is imported into tableau as a Text file given our file is in the CSV file format. Upon completion, the SVG map for **San Francisco** is imported which serves as our base map as seen below.

For this visualization, we are using 6 out of the 7 interactive techniques which are; Select, Explore, Reconfigure, Encode, Abstract/Elaborate, and Filter.
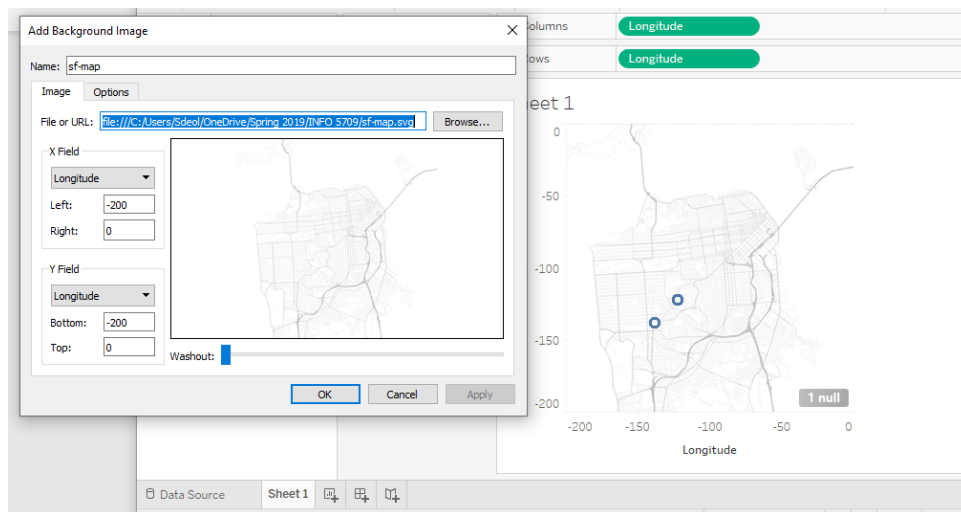

Once our base map and dataset is imported, we zoom(explore) into the data point to get a better view of the San Francisco trees which we see below.
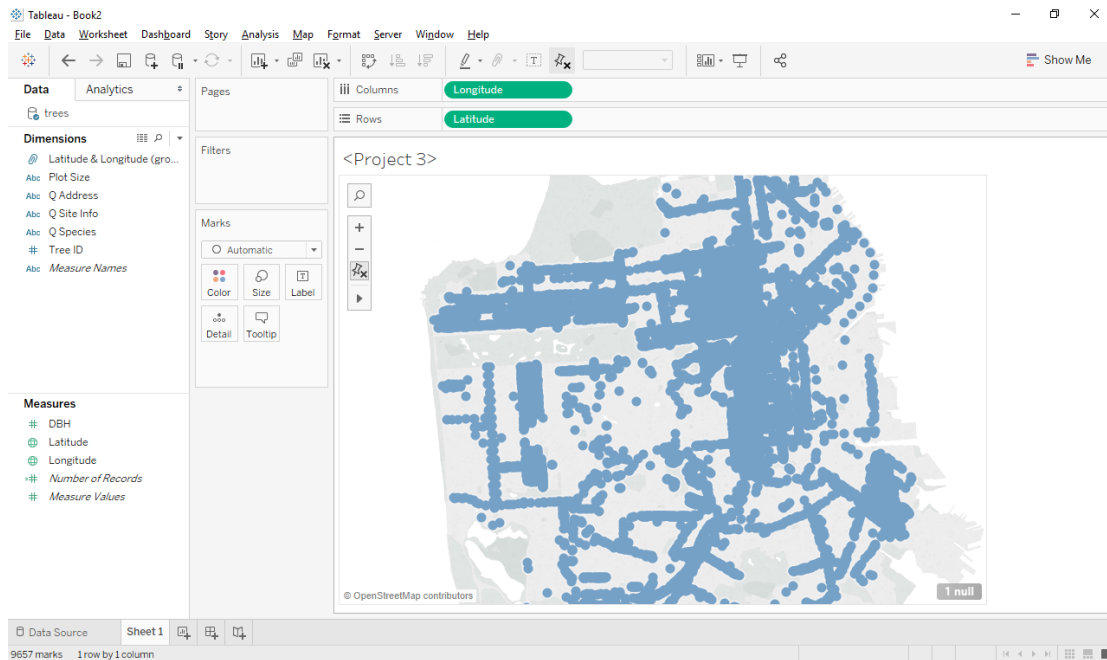
RADIAL SELECTION

Using the radial selection tool, specifically the circle, our goal is to specify two data points A and B as well as a radius and then apply the filter interaction functionality for the trees which lie within the intersection of the circles. Visual representations of each of the functionality is further described below.
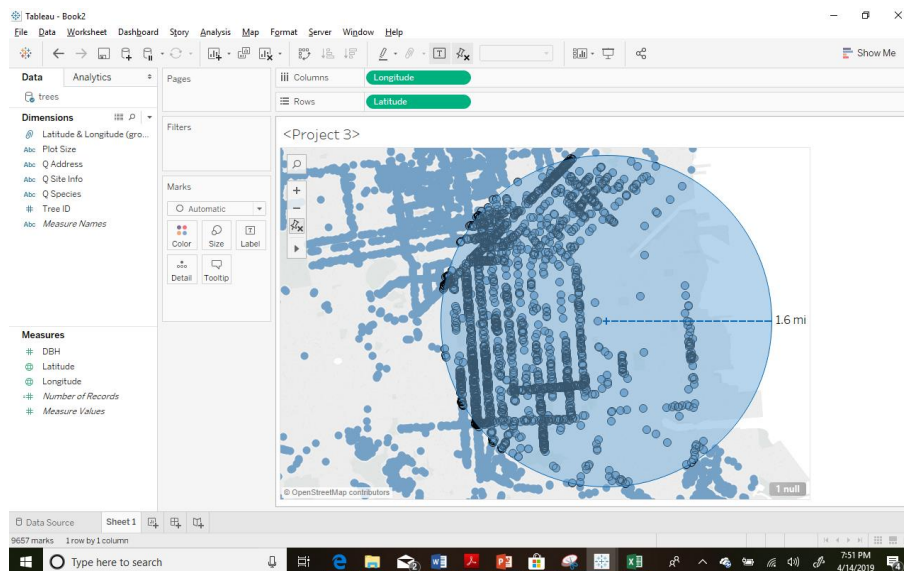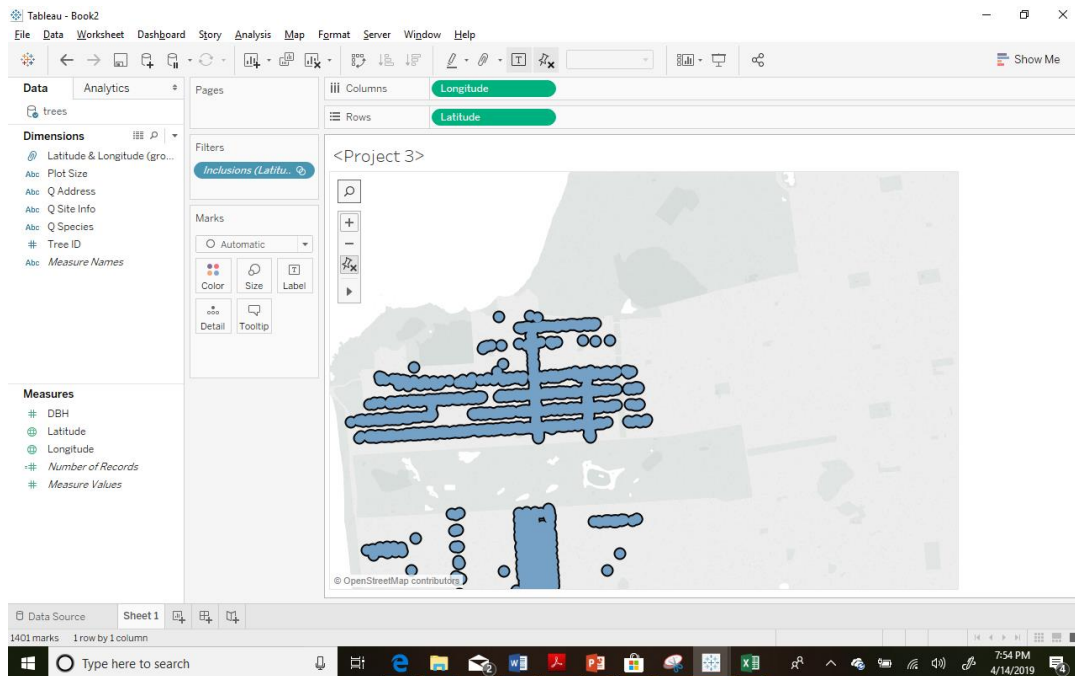

POINT A with 1.6mile radius
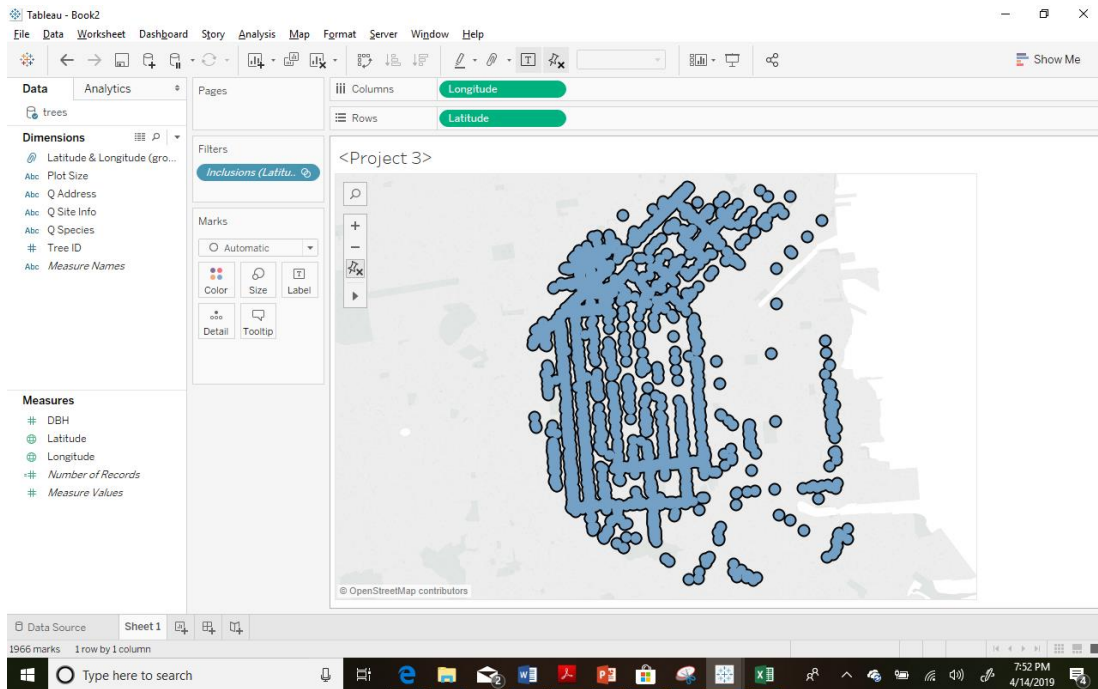
**POINT A: A selection of data points within 1.6-mile radius using the Keep Only function.**
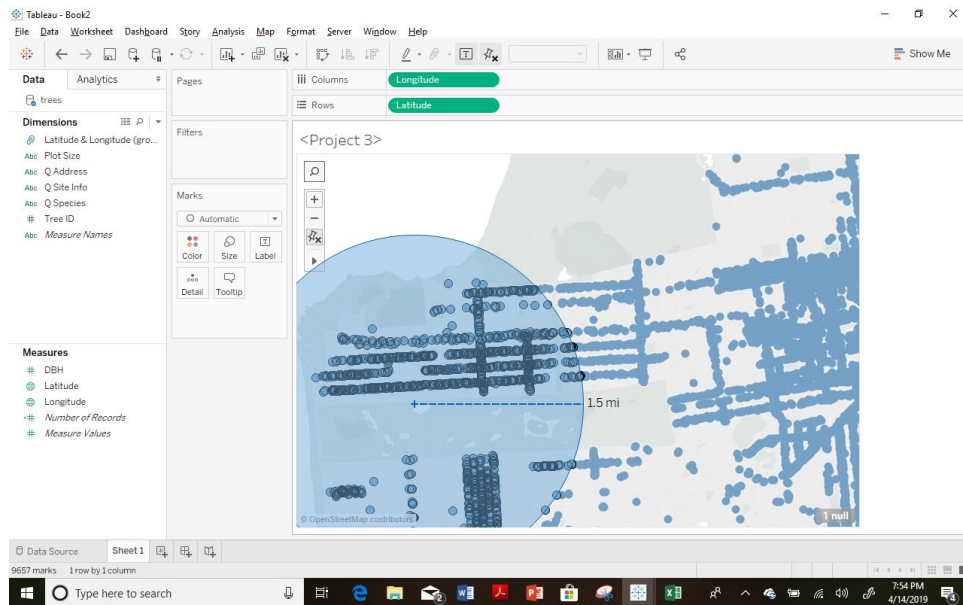

POINT B with 1.5mile radius

POINT B: A selection of data points within 1.5 miles radius using the Keep Only function.

FILTER BY TREE SPECIE TYPE:

As visualized below, we select only trees from a specific specie. In this case, I have selected trees in the **Ulmus family** to further streamline the trees in San Francisco which are visualized.
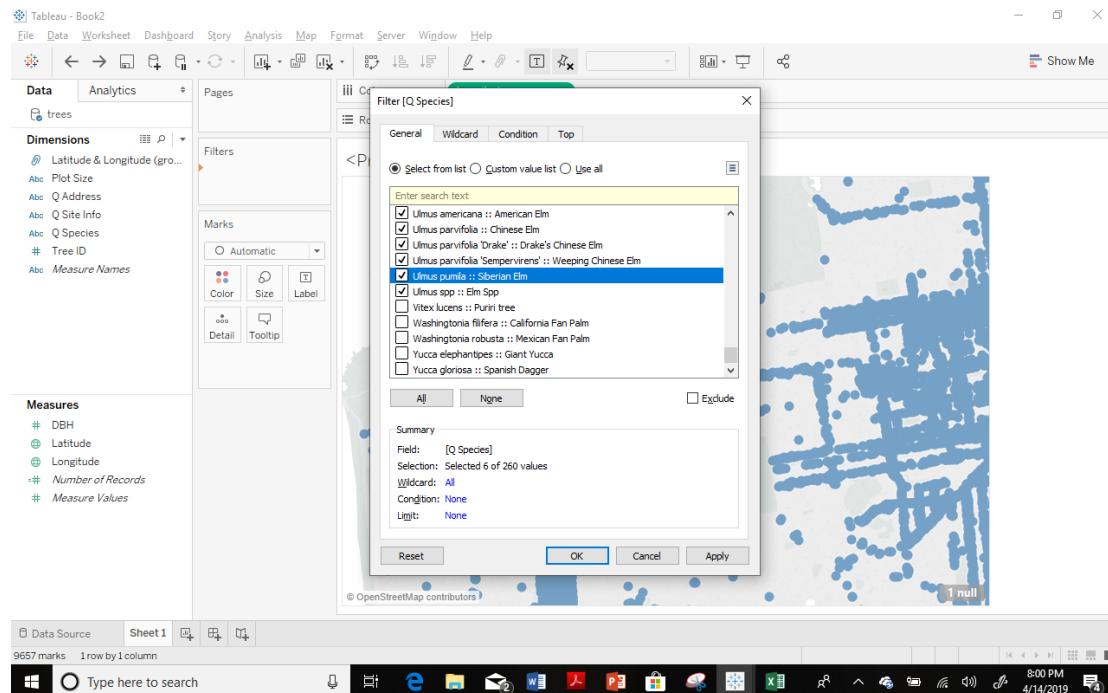
This filter gives us a better refined visualization which is displayed below.
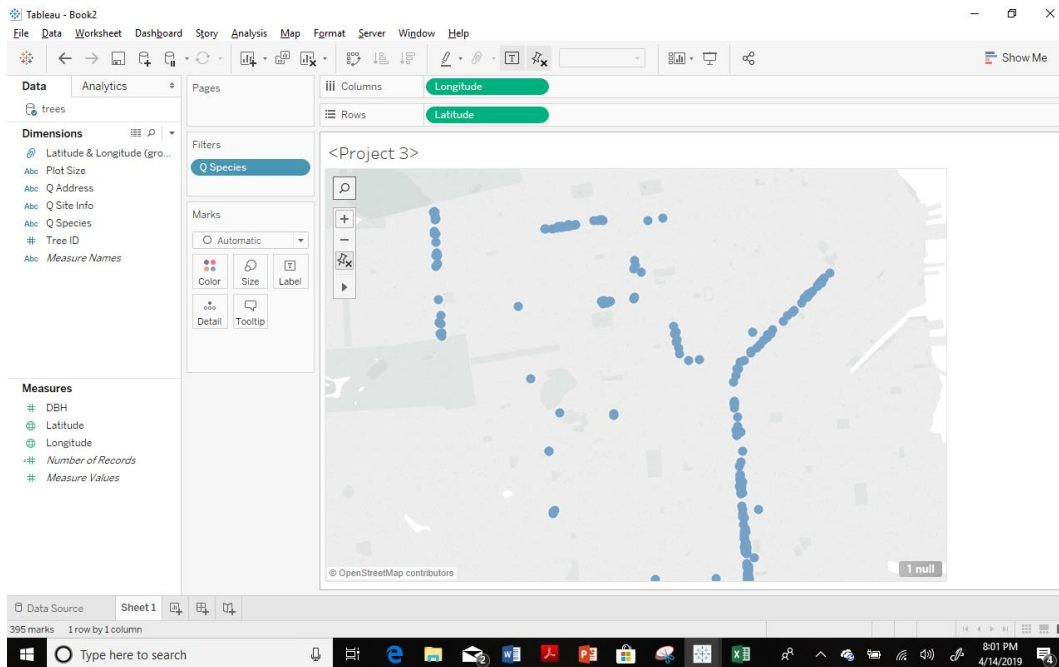

FILTER BY SPECIES

Next, we would be filtering by DBH (Diameter at breast height)

FILTER BY DBH:

Here, we filter by the diameter at breast height. In doing this, we see that the DBH ranges from 2 to 135. I applied a filter of at least 50. This means that we are keeping only those trees that have a minimum of 50 units DBH or more. Here is a representation of our filtering process.
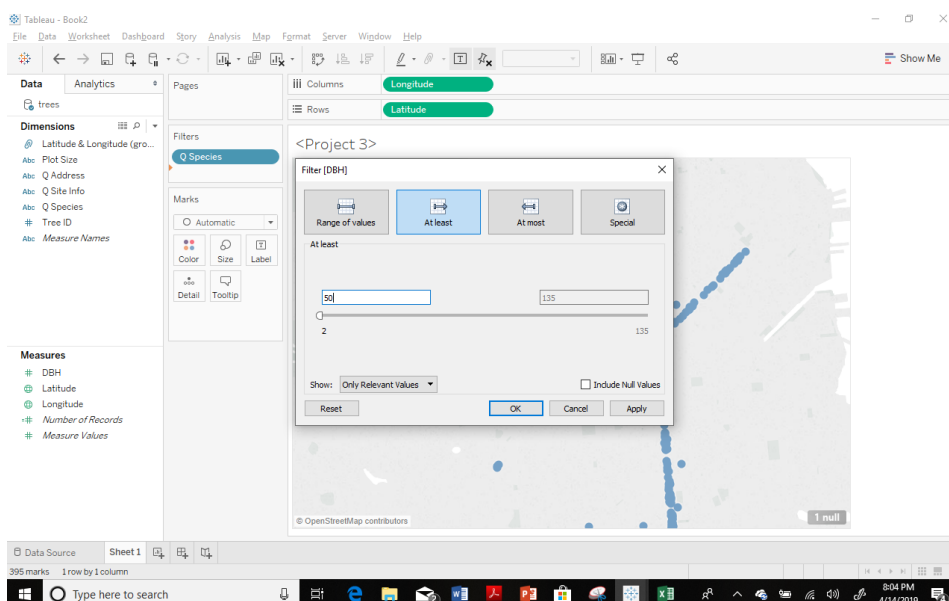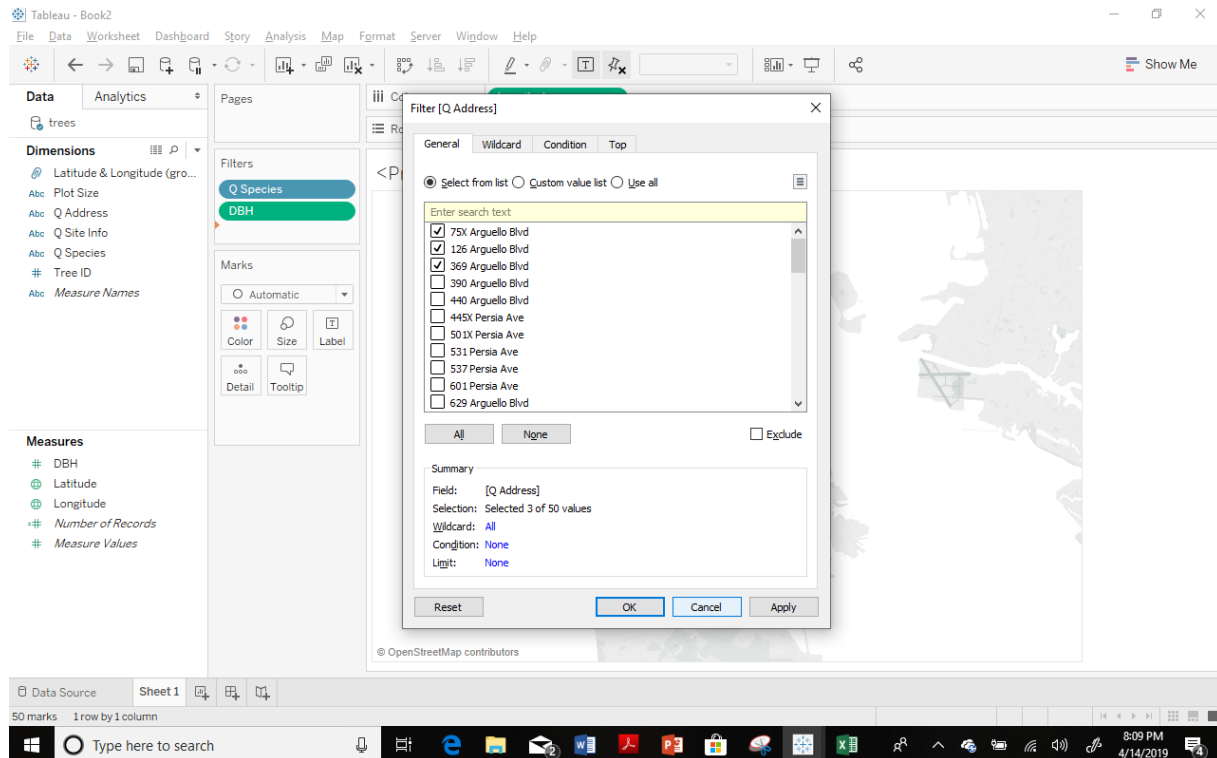
Upon application of the filter, we see below the visual representation of the filter applied.
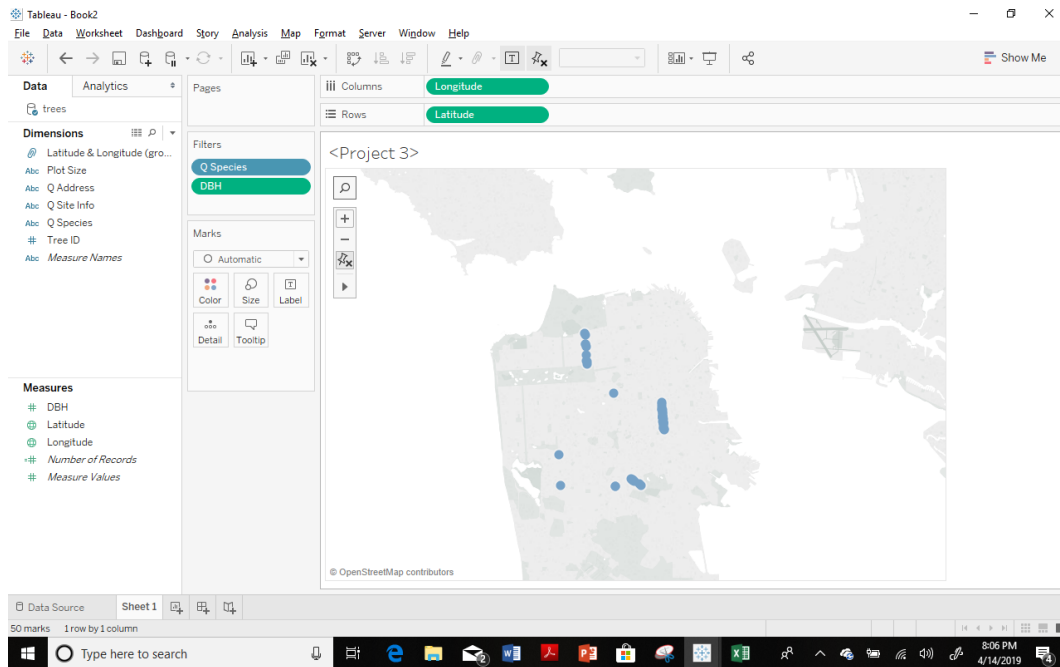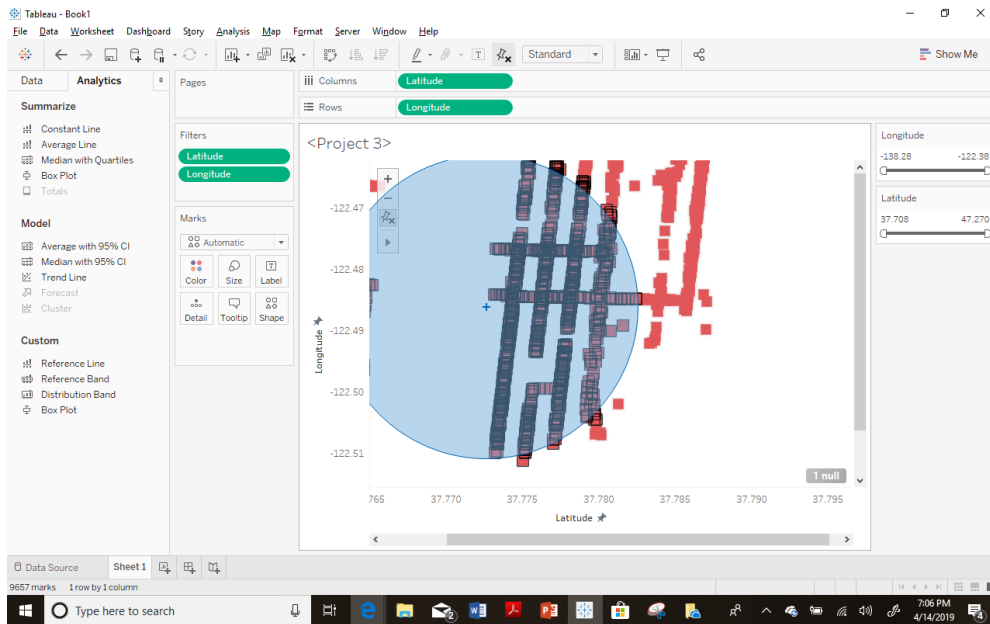

**FILTER BY DBH (DIAMETER AT BREAST HEIGHT)**

Next we filter by Address:

As we see below, there is a smaller amount of points shown. Therefore, the more components/filters we apply, the lesser the datapoints available.

**DISPLAYING LABELS/VALUES FOR EACH LONGITUDE.**

ABSTRACT/ELABORATE:

This interactive method allows for user to adjust details from a general overview into specific details. In the below image, we have filtered our visualization into specifics in respect to the Tree ID, the plot size, the DBH and the address of the trees planted.

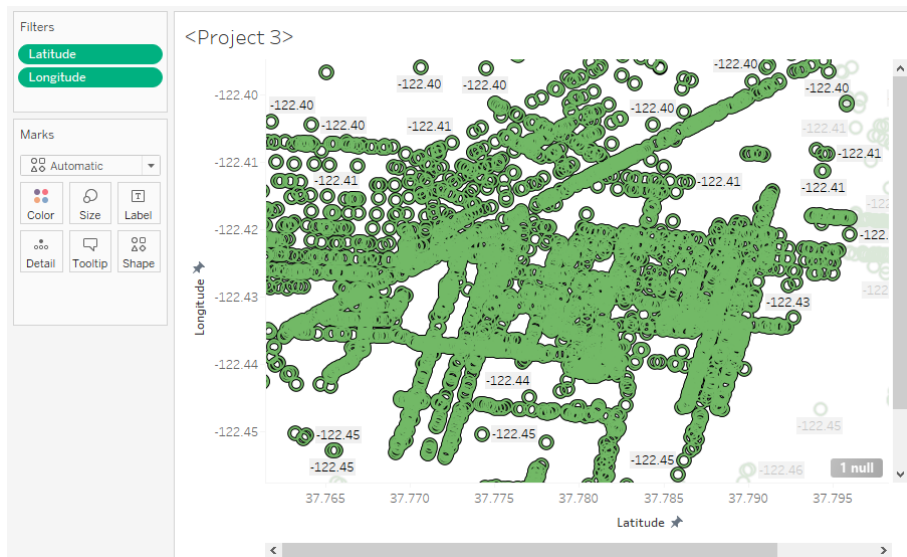This has therefore resulted in a smaller view than the initial visualization displayed.
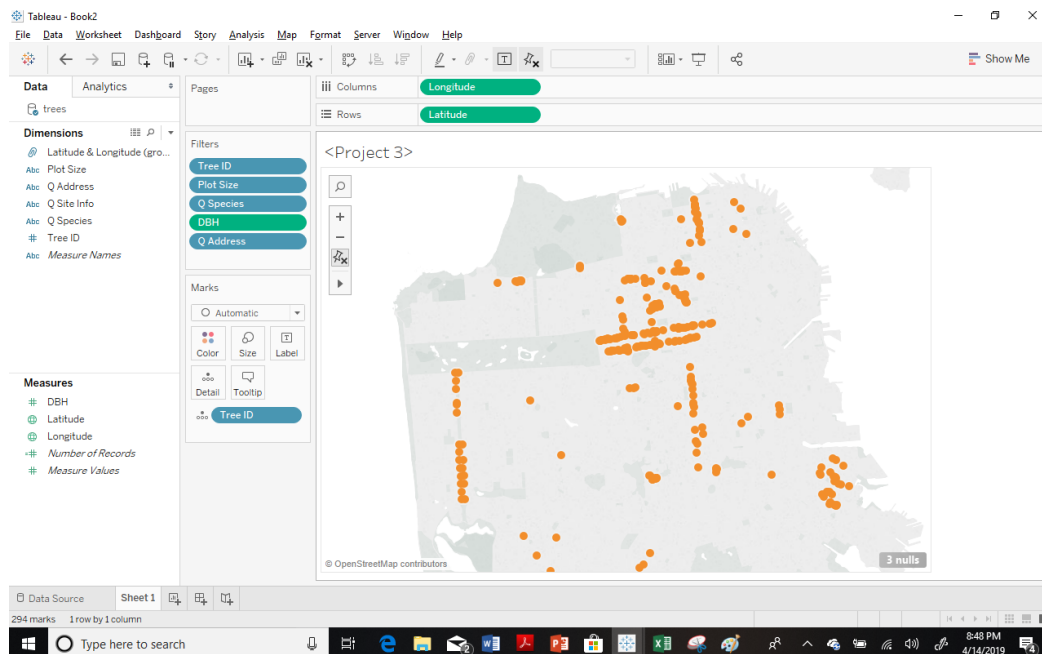
AFTER APPLYING ABSTRACT FILTERS
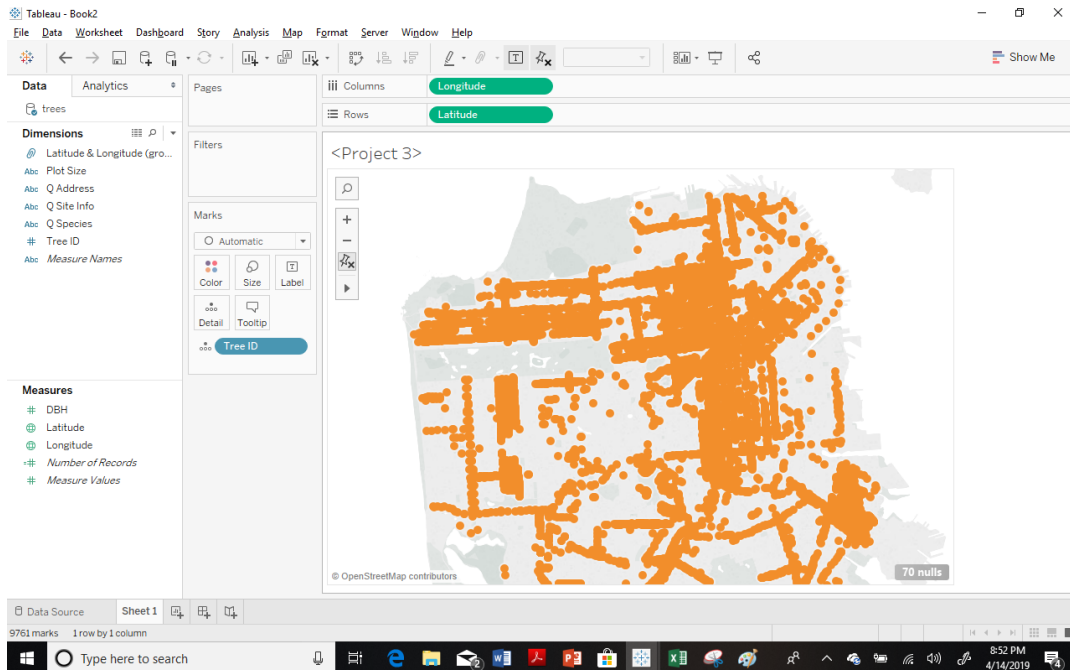
BEFORE APPLYING FILTERS.

SELECT:

Below, we see a high cluster of variables in one this location. It would be interesting to know what happens more frequently around this area and help make business decisions.
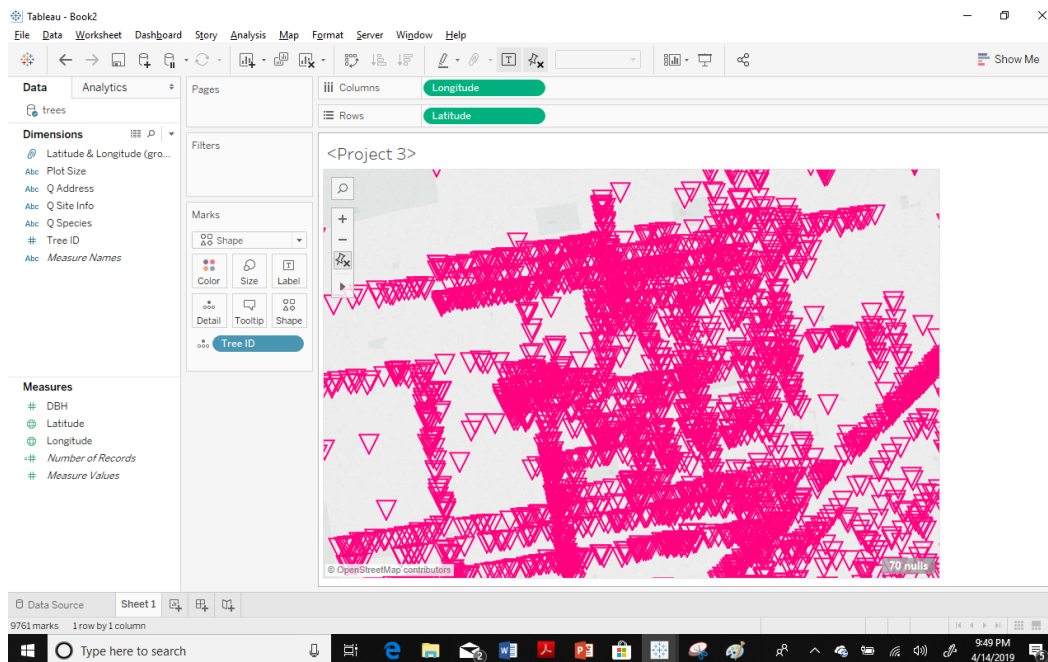
ENCODE:

Below, the shape of the identifier was changed from circle as seen in previous visualizations into a triangle. We also performed color encoding as I changed the color was from blue/green to orange. Alongside the changes mentioned above, the size of the shape was made larger. This is visualized below.
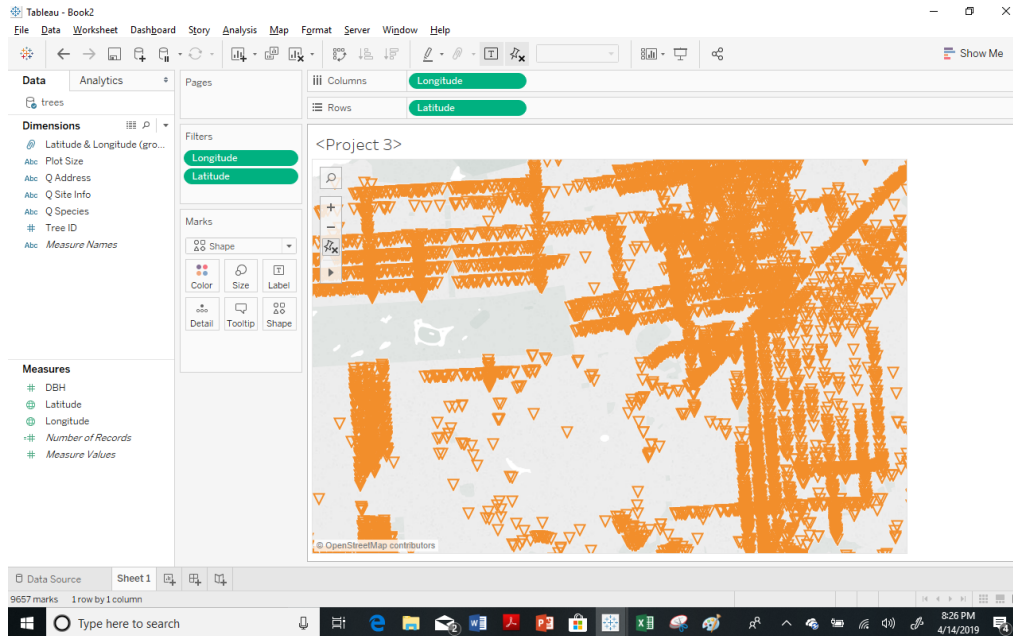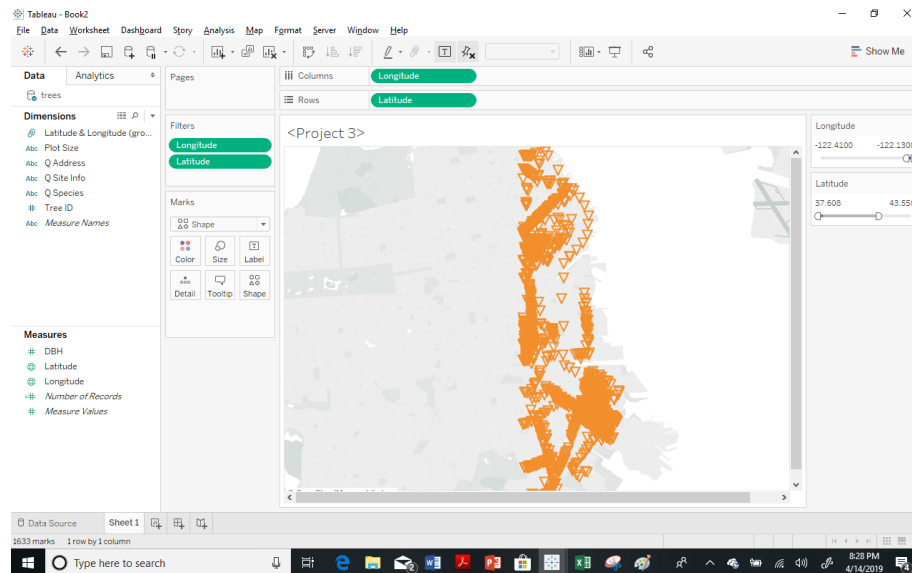

**COLOR AND SHAPE ENCODING**

SIZE ENCODING

FILTER:

Below, we have applied filters to both the longitude and the latitude. The visualized data shows longitude between -122.41 to -122.13 while the latitude displays data between 37.608 and 43.550 as shown on the right.
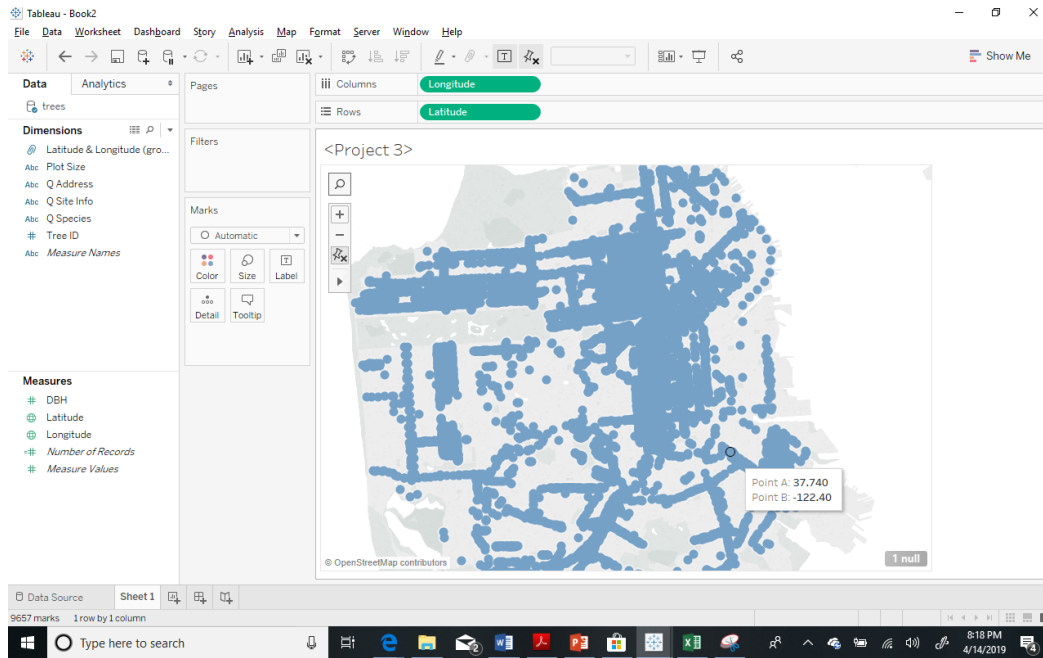

EXPLORE:

Using the panning tool and zoom in tool, we explored different parts of the dataset. Below, I have taken snapshots of different parts of the clusters and differentiated them with colors for easy understanding.

We begin our data visualization task with two distinct data points (clusters). As we explore the datapoint on California by zooming in and panning, we can see a larger view of the different trees planted in this location and we are thereby able to focus on a specific location. This is visualized below.
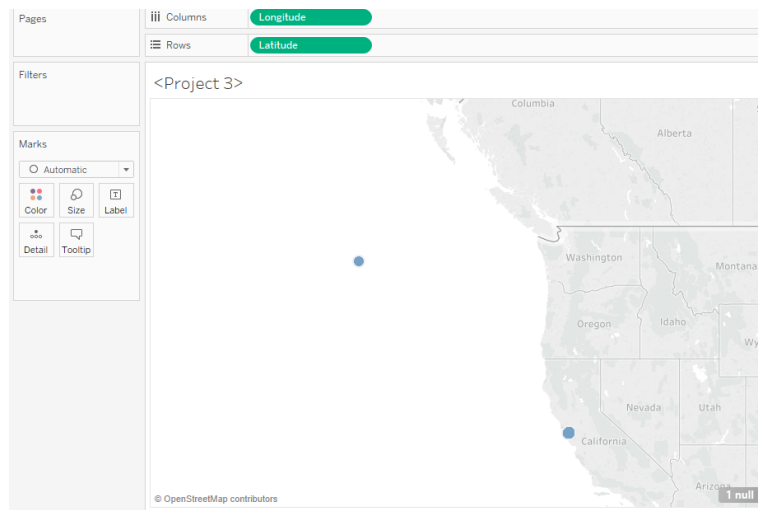
**BEFORE EXPLORATION**

AFTER EXPLORATION

From the other datapoint located away from California and upon exploration, we find no adequate data for analysis. This is shown below.
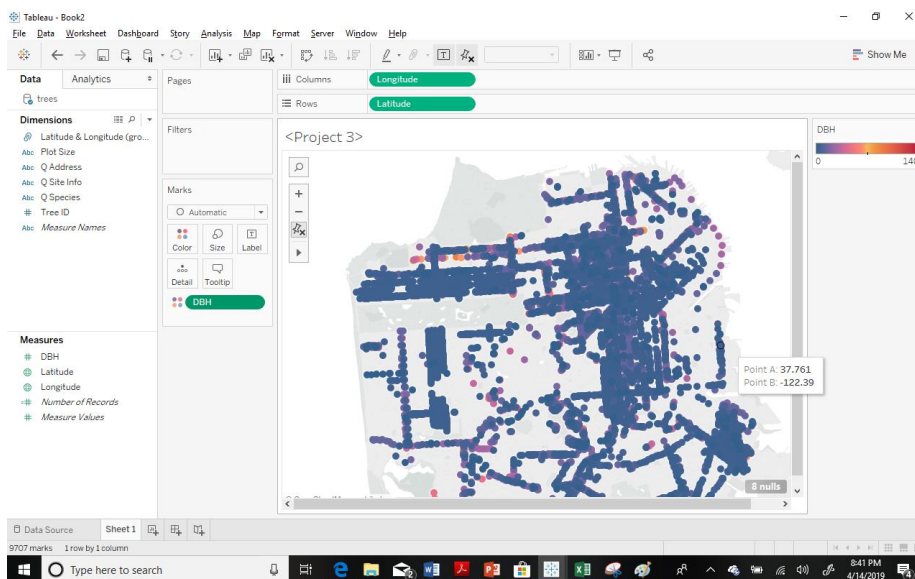
EXPLORING CLUSTER

EXPLORE BY SPECIES:

Here, we are filtering by the tree species and trying to understand through our visualization which tree specie is planted in an area. As we see below, there are various tree species represented by various colors.
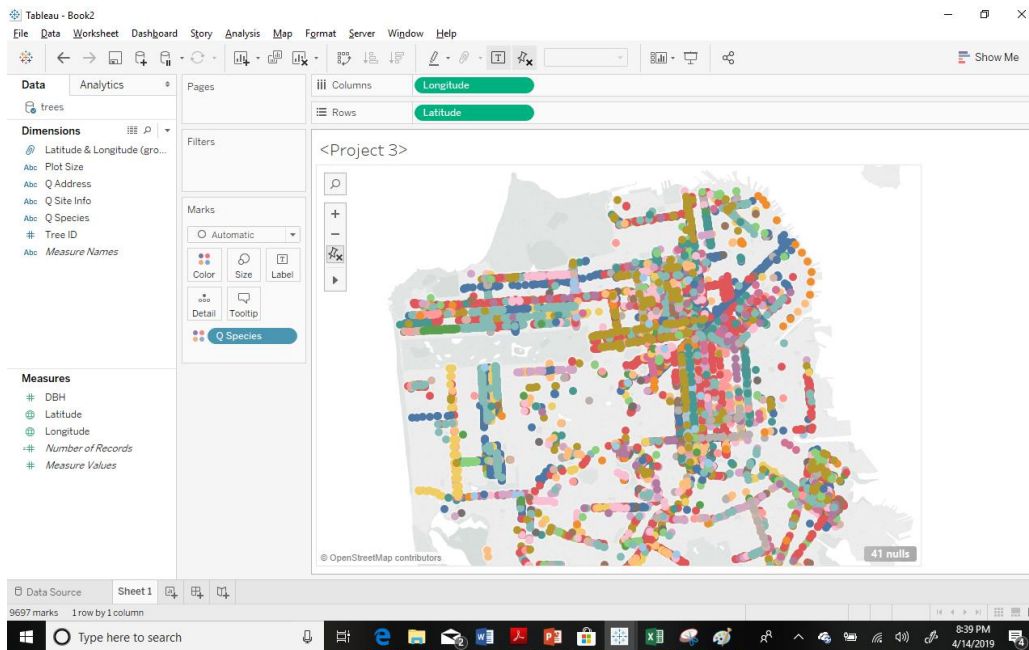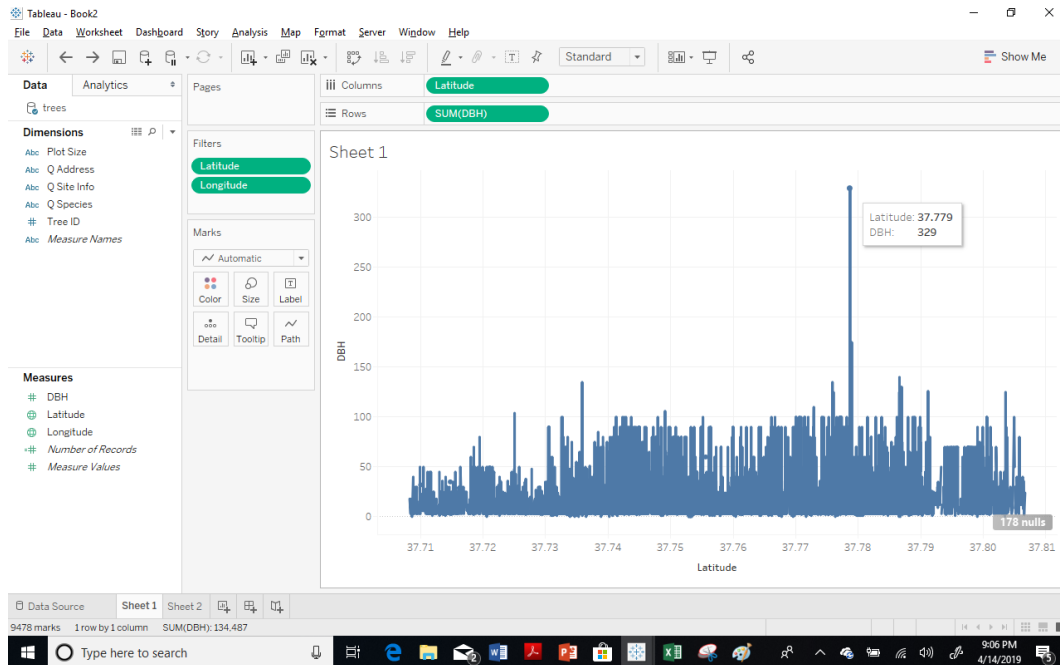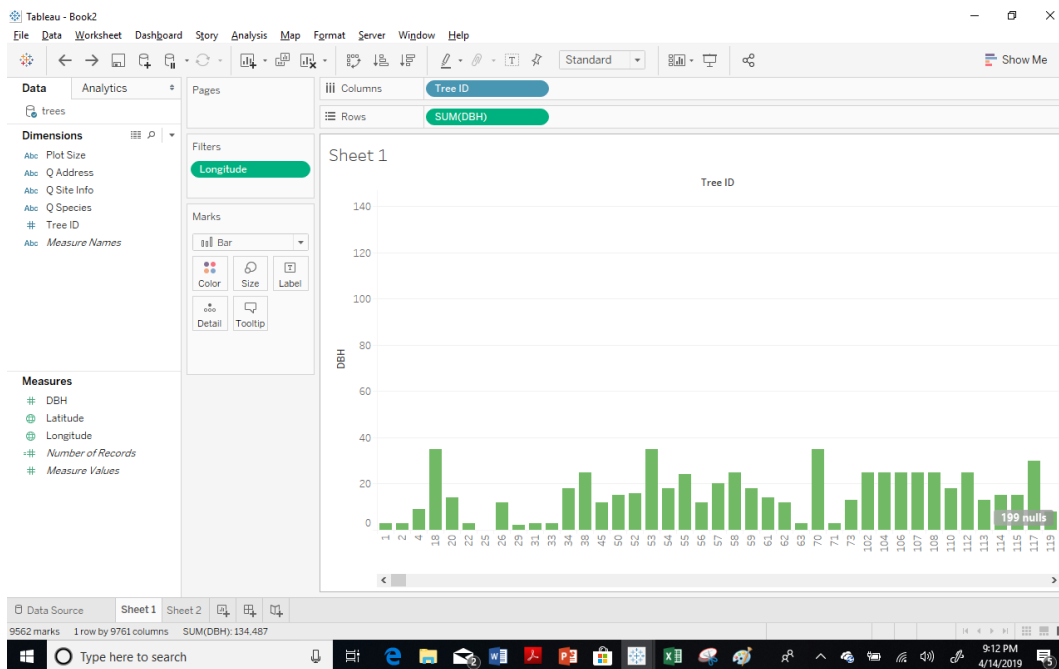

EXPLORE BY DBH:

RECONFIGURE:

To give us a different view of our dataset, we would be trying out some of the data variables we have.
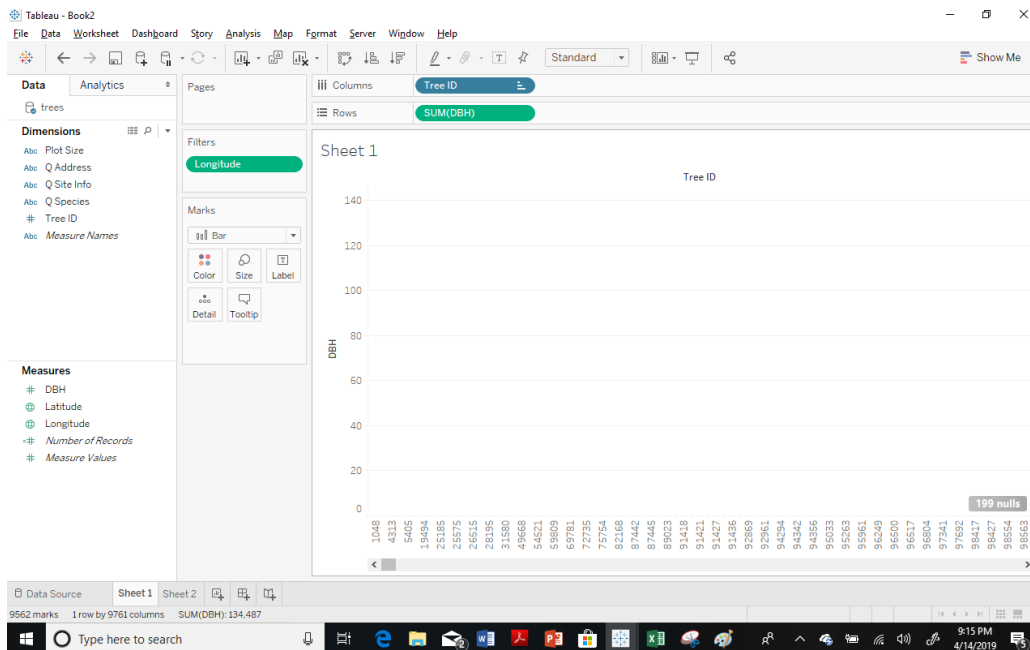

Here, using the Latitude and the DBH parameters, we see an unusually high trees planted in latitude 37.779 with DBH being 329
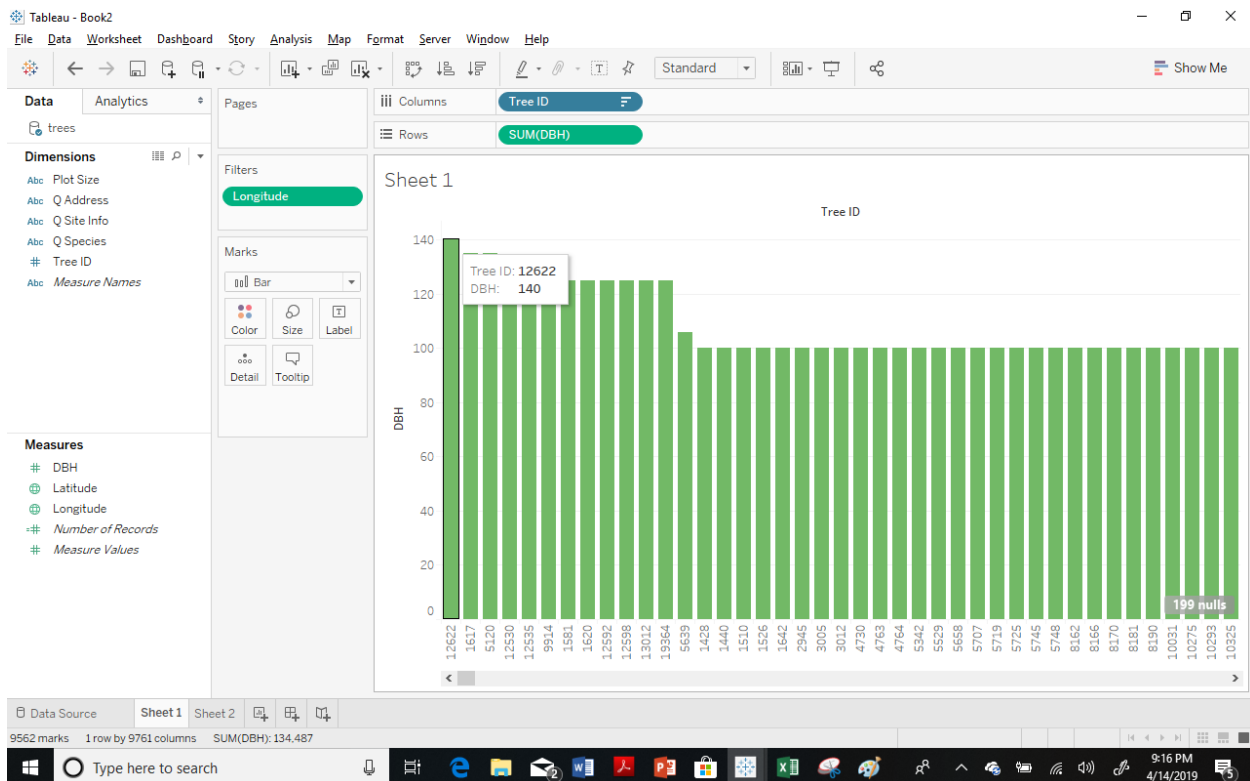
Next, we reconfigure using the Tree ID and DBH. Below is the first visualization displayed.

Subsequently, we apply the sort function on the Tree ID to determine which has the least usage.

USING ASCENDING SORT FUNCTION.

USING THE DESCENDING SORT FUNCTION ON THE DBH FIELD NAME.

The above image shows us there is high concentration of trees with ID 12622 and DBH 140.

To sort, click on the bar chat like icon on the Column (Tree ID) and specify the sort criteria's as seen below.
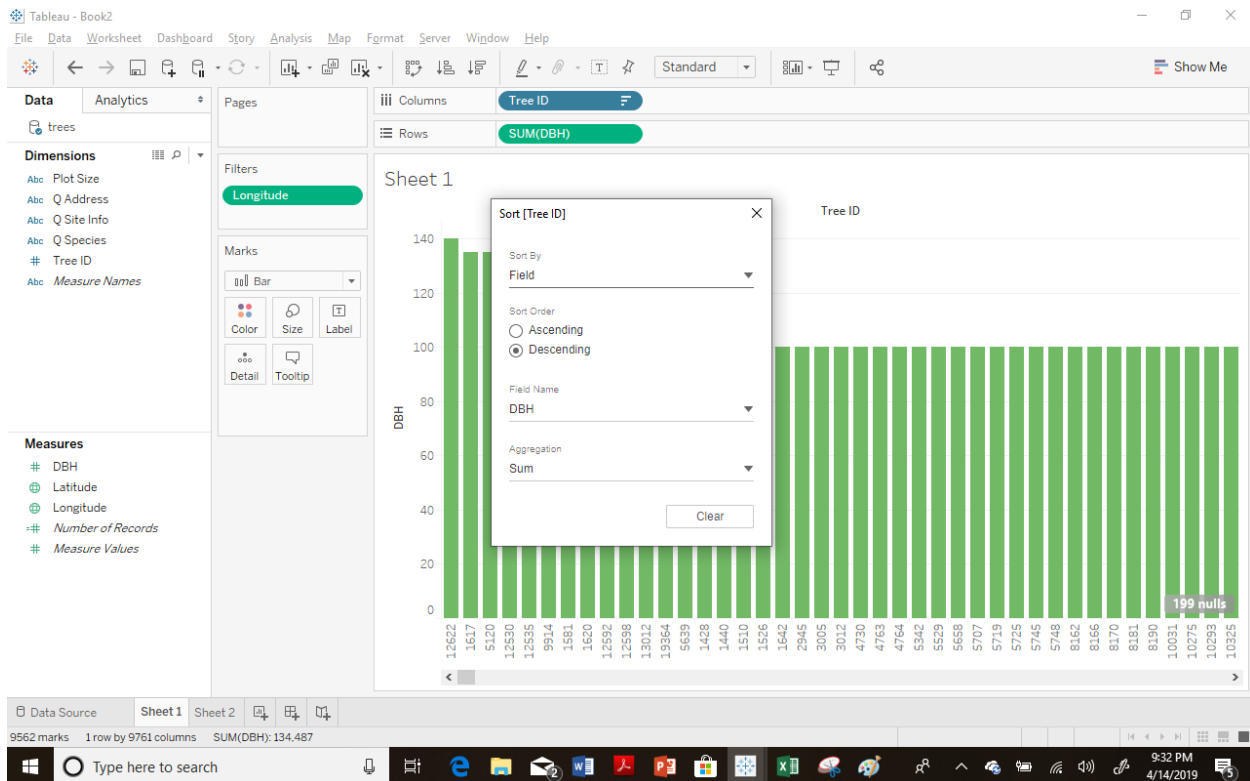
CONCLUSION

From the above visualizations and explanations, we can see the use of some of the most important interactive techniques used in visualizing data and deriving insights from a dataset.