Homework 06

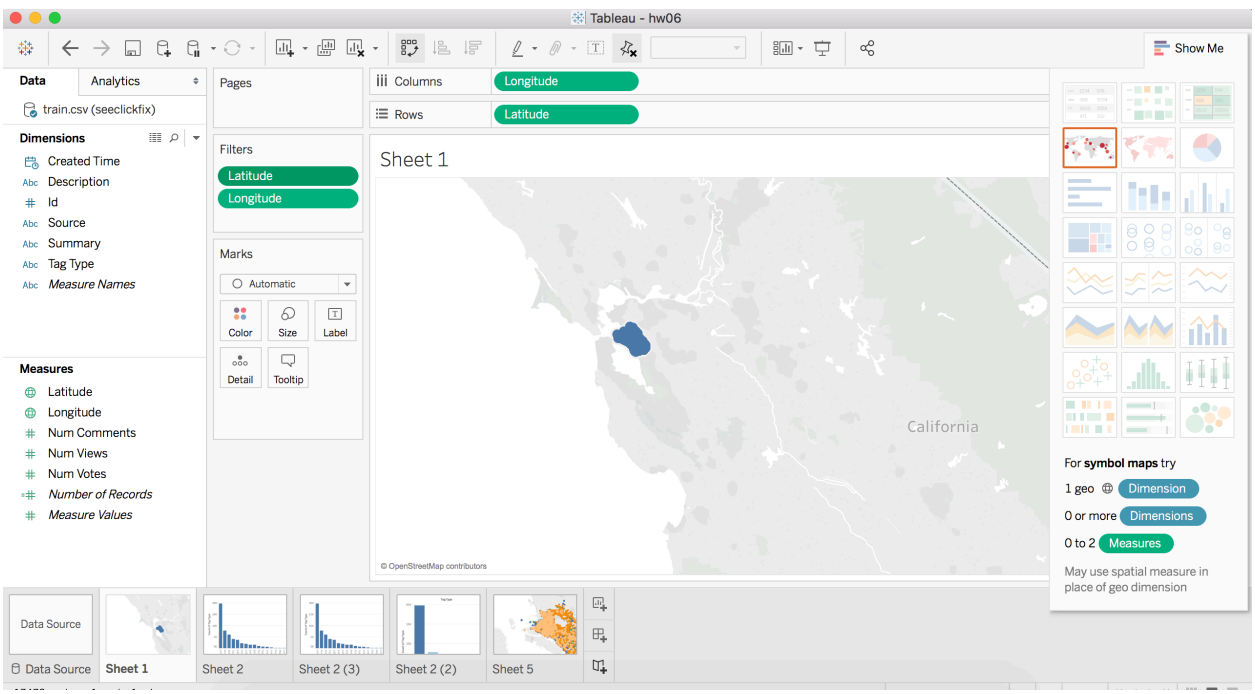
Xun Zhou 661966819

(1). IN YOUR OWN WORDS, describe what the data involves and key fields that are identified.

Q(1) Answer:
There are 11 variables involved in this data set: Created Time, Description, Id, Source, Summary, Tag Type, Latitude, Longitude, Num Comments, Num Views, Num Votes.

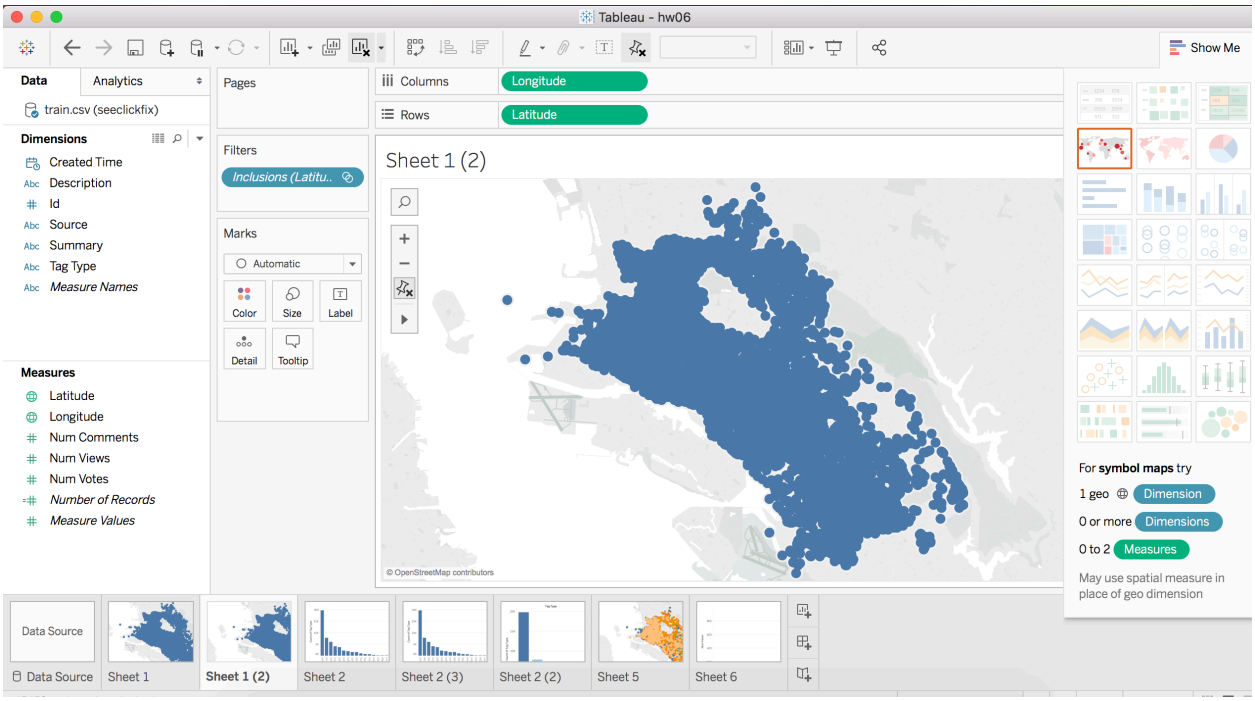
We have 14 data fields in Dimensions and Measures area. For example, Created Time, Description, Id, Source, Summary, Tag Type and Measure Names(generated by Tableau) are belong to Dimensions area, Latitude, Longitude, Num Comments, Num Views, Num Votes, Number of Records(generated by Tableau) and Measure Values(generated by Tableau) are from Measures area.

Since we drag the Longitude to the column and the Latitude to the rows, the key fields that are identified are the Longitude and Latitude.

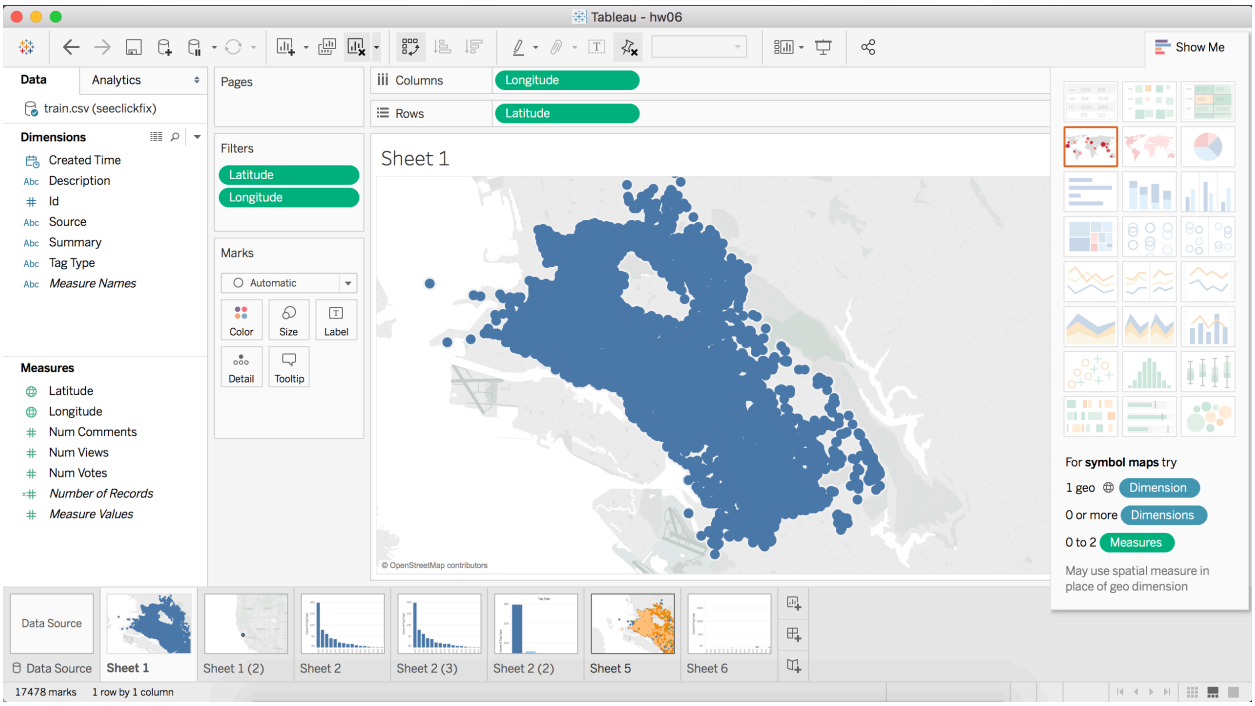


(2). IN YOUR OWN WORDS, describe the filter you created to select only one city.

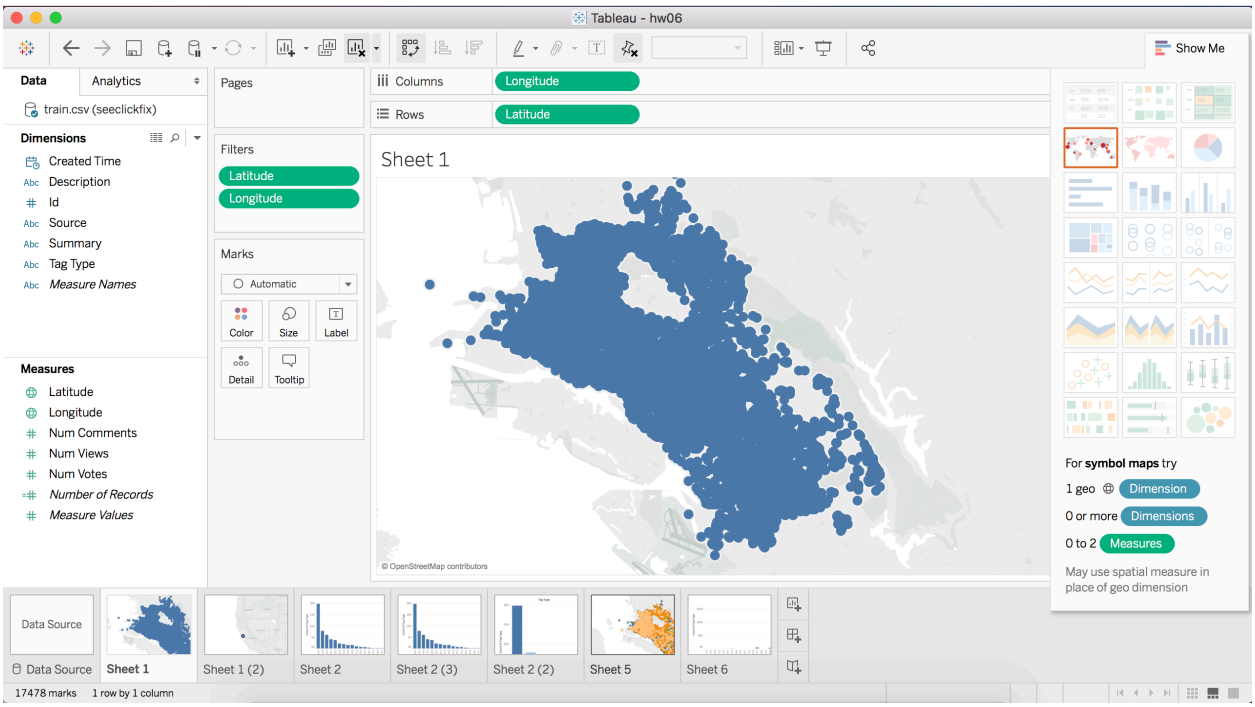
Q(2) Answer:
Since we only need data points from California, we can use two ways:
1. Based on the map, I click and drag in the view to select all the data points(marks) in California. And then Select **Keep Only** to keep only the selected marks in the view.



2. Since the Latitude and Longitude of California is 32° 30' N to 42° N and 114° 8' W to 124° 24' W, I drag Latitude to filters and choose the range from 32° to 42° , and then drag Longitude to filters and choose the range from -114° to -125° to show only relevant values.

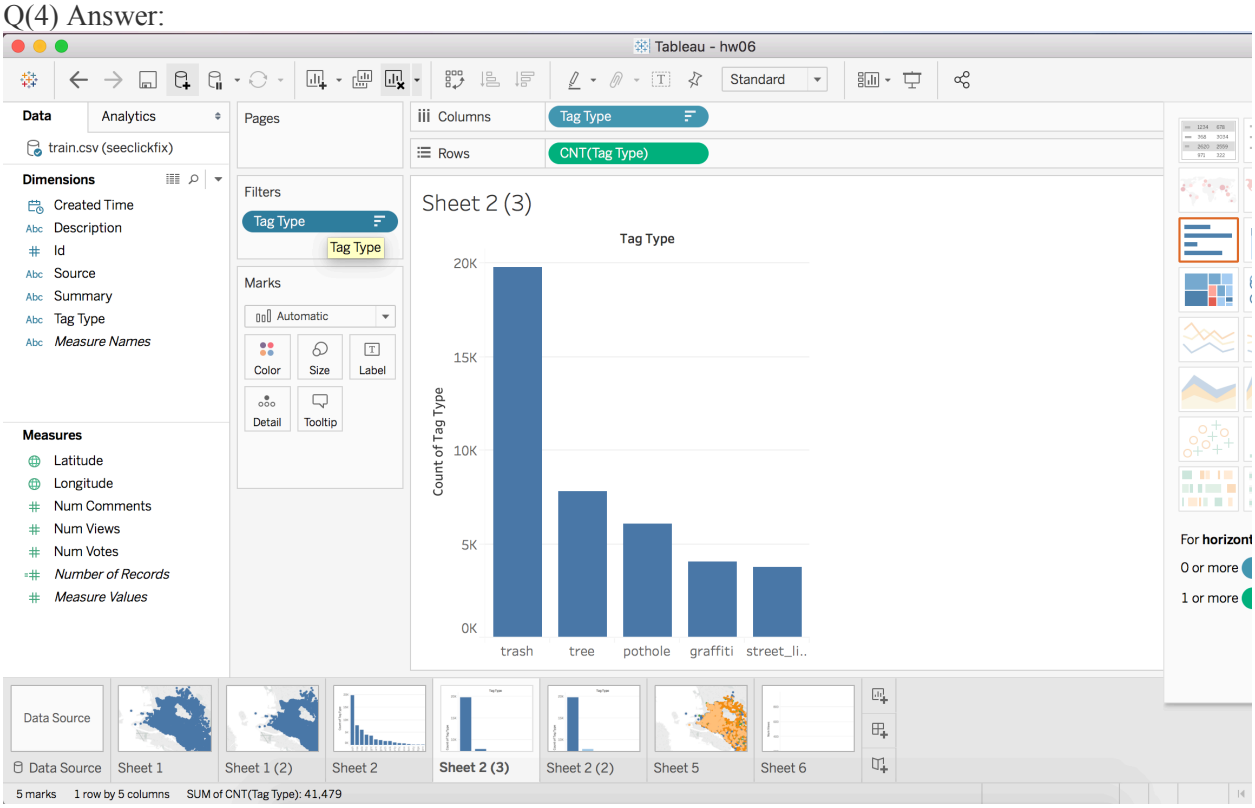


(3) *Change the type to a Map.* Select on the *Show Me* map to show your city on a map. (This actually happens automatically).



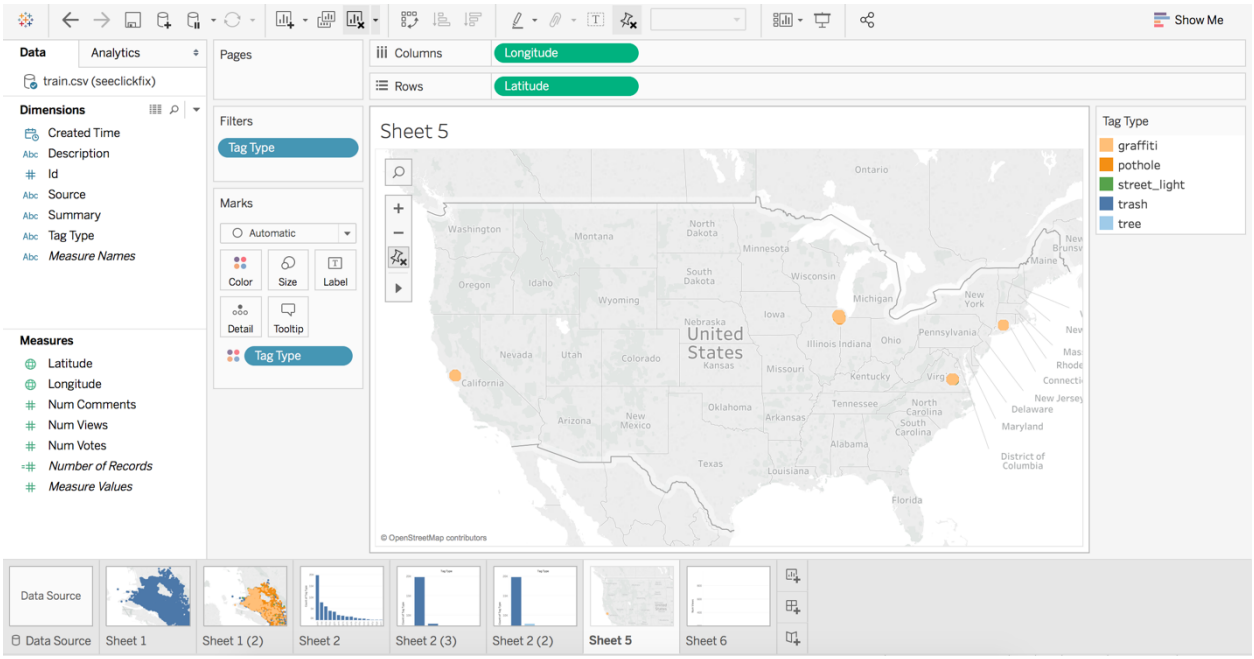
(4) *Understand Tags*. Create a new worksheet and drag Tag Type from the dimensions and create a visualization that is sorted by the count of the number of tags. Provide an image showing the 5 most frequent tag types (not including NAs).

Q(4) Answer:

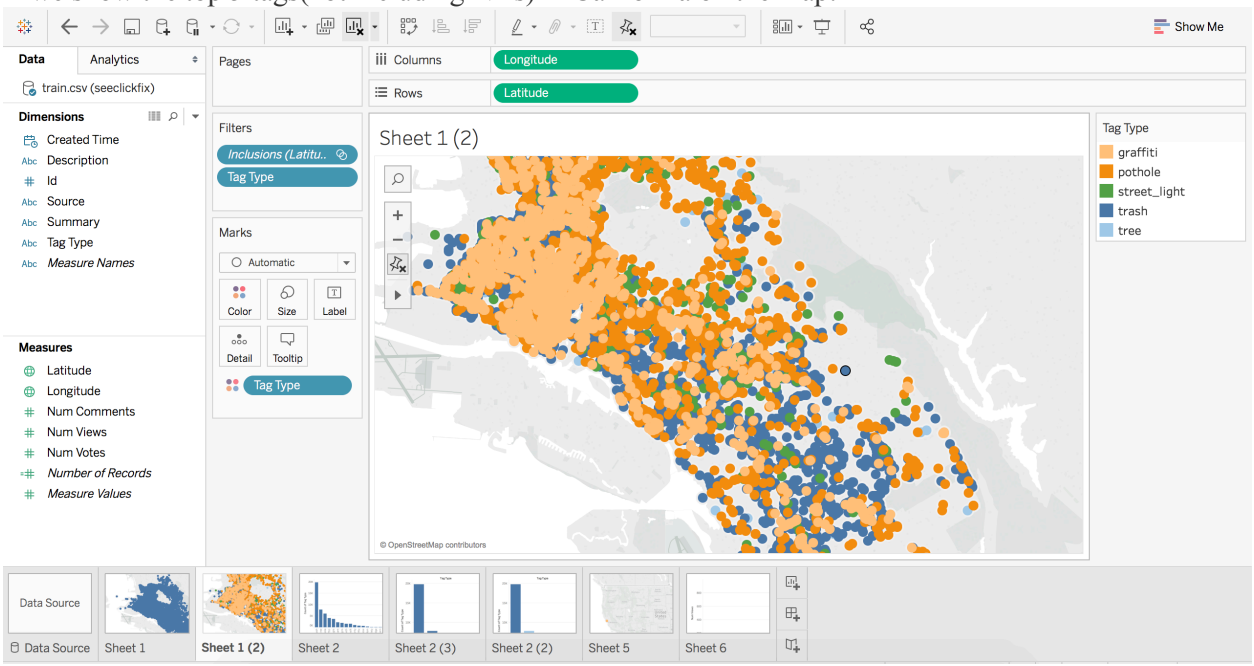


(5) *Update your Map in order to include Tag Type.* Drag the tags over to the colors and select filter and add. Select the top 5 tags and show them on the map.

Q(5) Answer:
If we show the top 5 tags(not including NAs) in all four cities on the map:

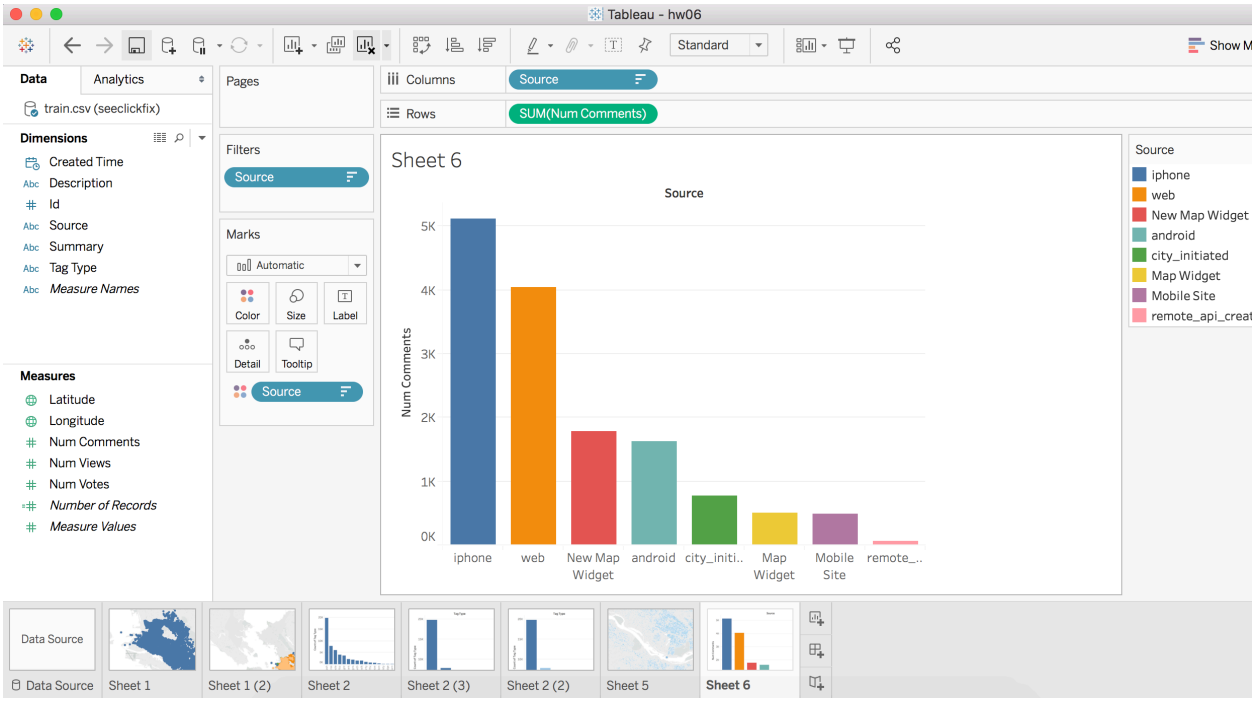


If we show the top 5 tags(not including NAs) in California on the map:

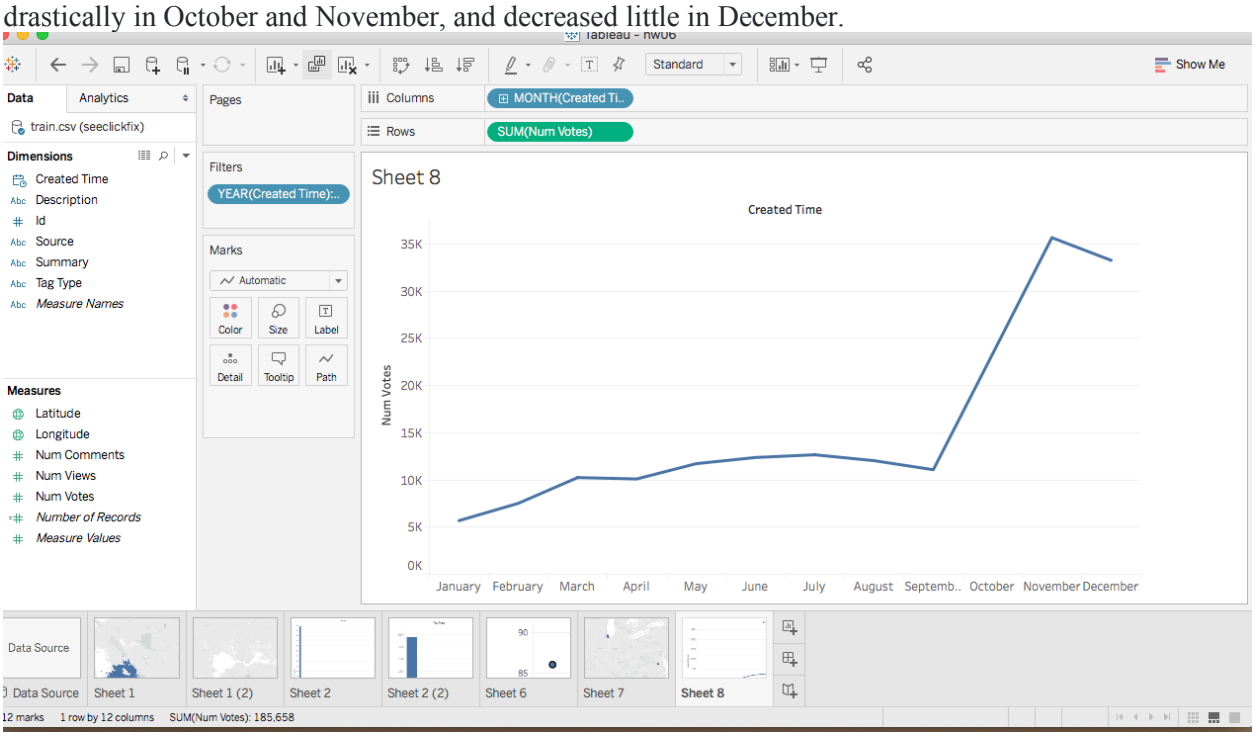


(6). Create 4 visualizations that help you to understand more details about the data or make a specific point. At least one should use the Map. Provide 1-3 sentences to describe each visualization.

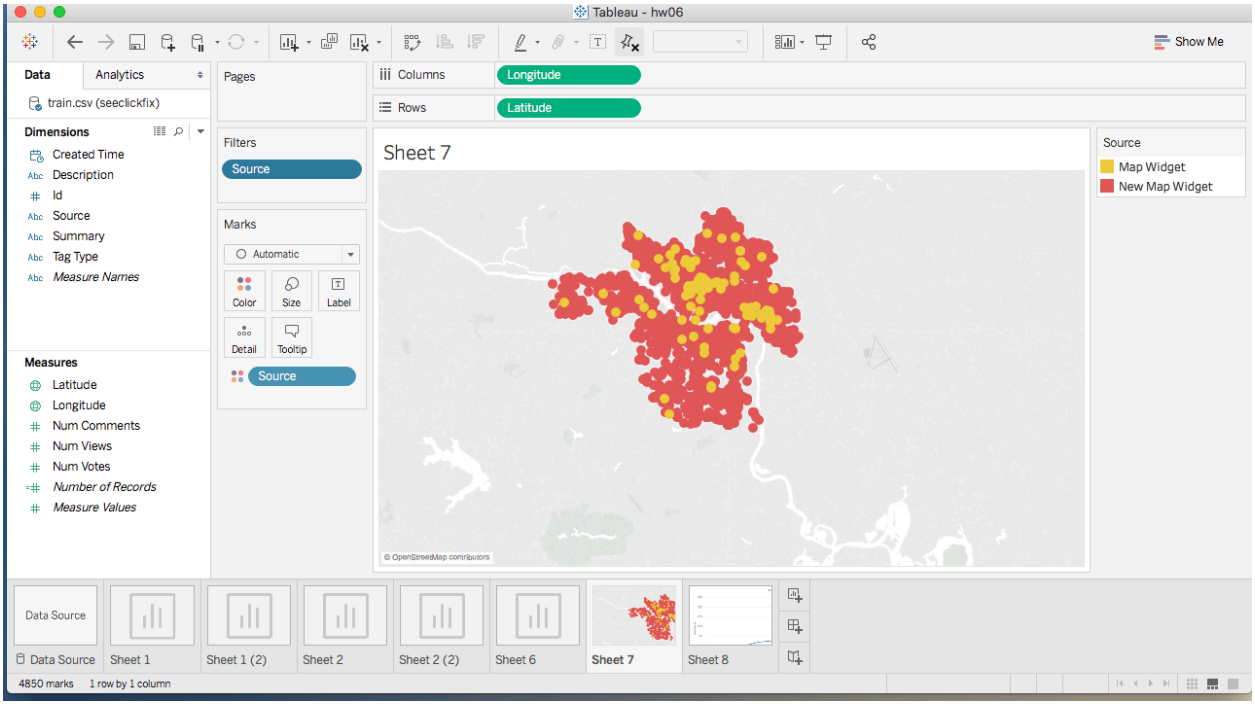
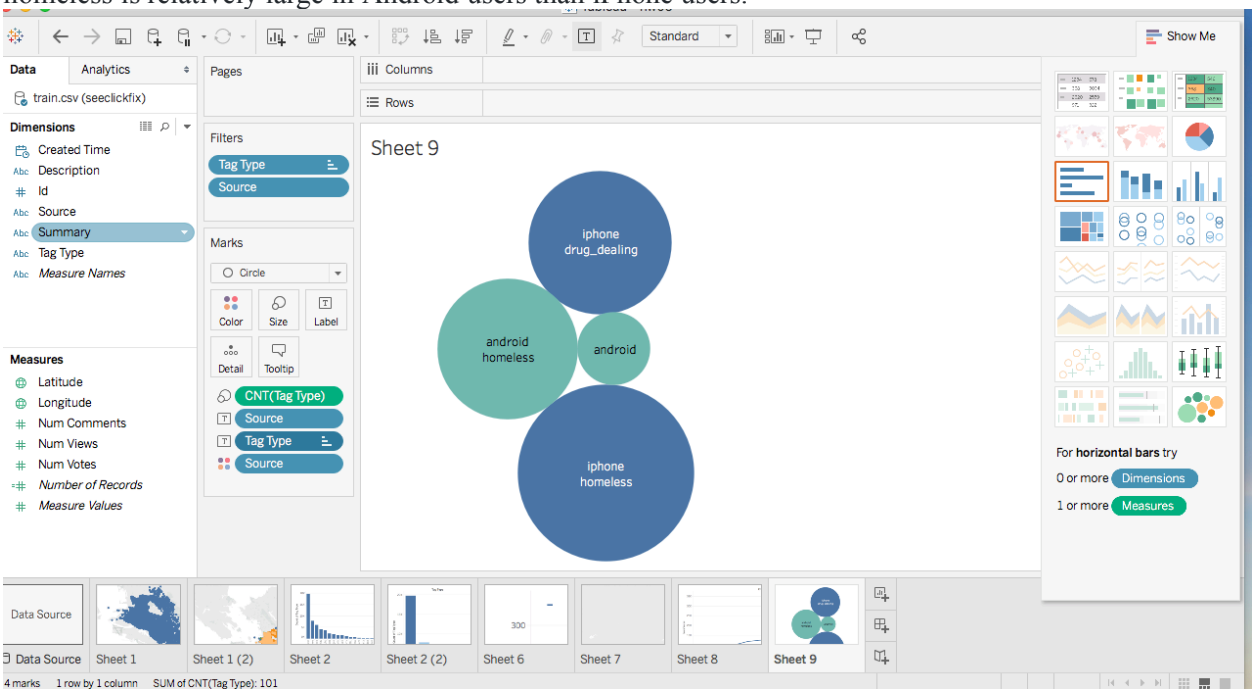
a. I create a bar chart to visualize the relationship between source types (not including NAs) and number of comments. As we can see from the graph below, most comments were sent by iphone, the second and third largest source is web and new map widget respectively. Android is the fourth largest source.



b. This image below gives us a broad understand that how the number of votes changes with time. In 2012, the number of votes increased slowly from January to September. The number of votes increased drastically in October and November, and decreased little in December.



c. This shows that in Virginia, more people use New Map Widget than Map Widget. And Map Widget users mostly located in the center of the city.



d. As the graph shown below, we can see that both Android and iPhone users care more about dealing with homeless than drug dealing. Specifically, the proportion of people who care about dealing with homeless is relatively large in Android users than iPhone users.