Visualizing crime patterns

This script is just the start of analysis; I am working on this problem from last week. In addition, I Will update any new patterns observed, I am just a starter please correct me if I am wrong with any approach.

Data set:

* **Dates** – The date and time of the crime happened
* **Category** – Variable which we are going to predict has 39 categories of crimes in the data set, among these

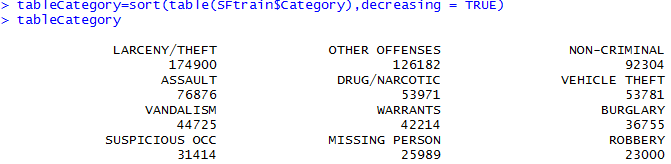
Script

Let us see how many crimes are from top 10 crime categories in the data

* **Descript** - detailed description of the crime incident (only in train.csv)
* **DayOfWeek** - the day of the week
* **PdDistrict** - name of the Police Department District
* **Resolution** - how the crime incident was resolved (only in train.csv)
* **Address** - the approximate street address of the crime incident
* **X** - Longitude
* **Y** – Latitude

Visualization:

We can observe from the table below that top 10 crimes constitute to 83.5% of total crimes which happened in San Francisco.



When we see the percentage of top 10 crimes in all the crimes.



# we are plotting the location with the categody variable to seggregate crimes

#by coloring them

#There is not much we can derive from the below plot, But if you observe closely

#there are dark shaded blue points and light shaded blu points in most of the occurences

#From the Hour meter it is said that light blue dots are after 8'o pm and dark blue dots

#are after 12' am in night.

#I believe most of the crimes occur between this intervel.





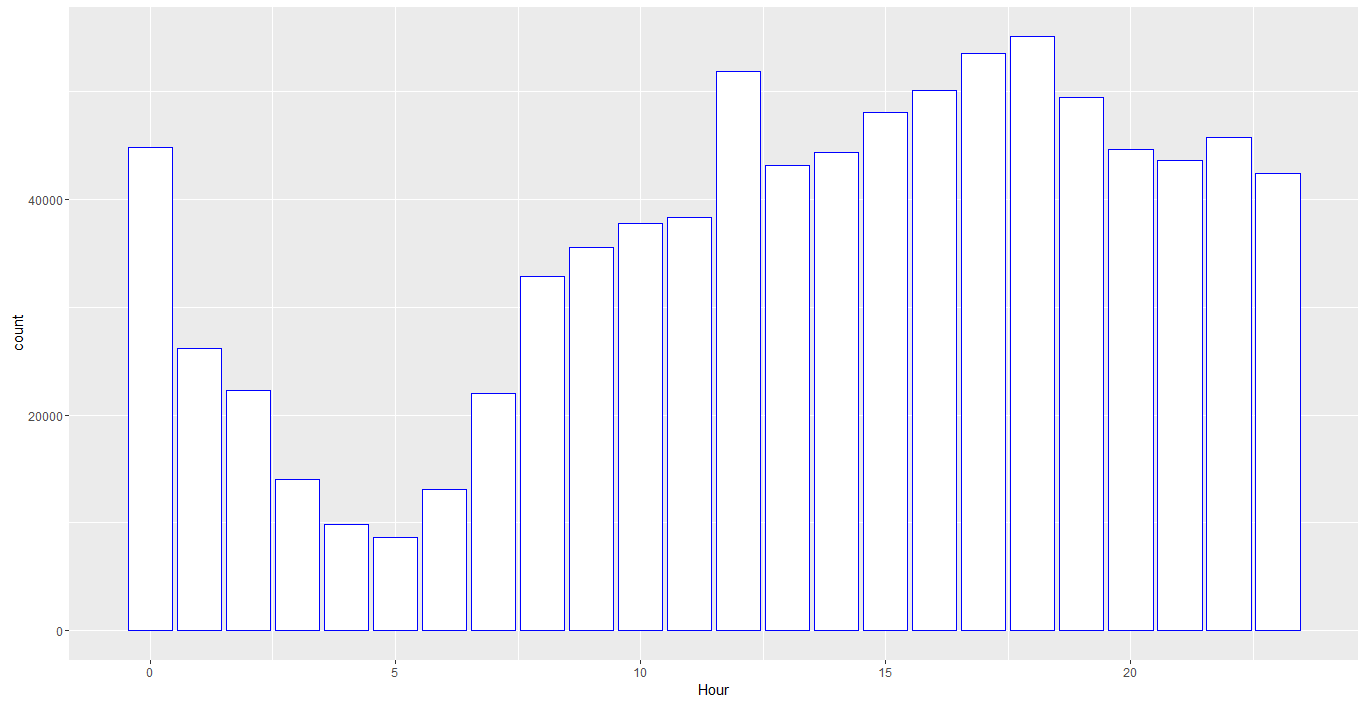
#we try plotting number of crimes with respect to Hour of the crime Occured

#in this initial analysis we can observe that more crimes occure in the afternoon,

#where many people go out for lunch etc and night time where people go from

#offices, dinner etc.



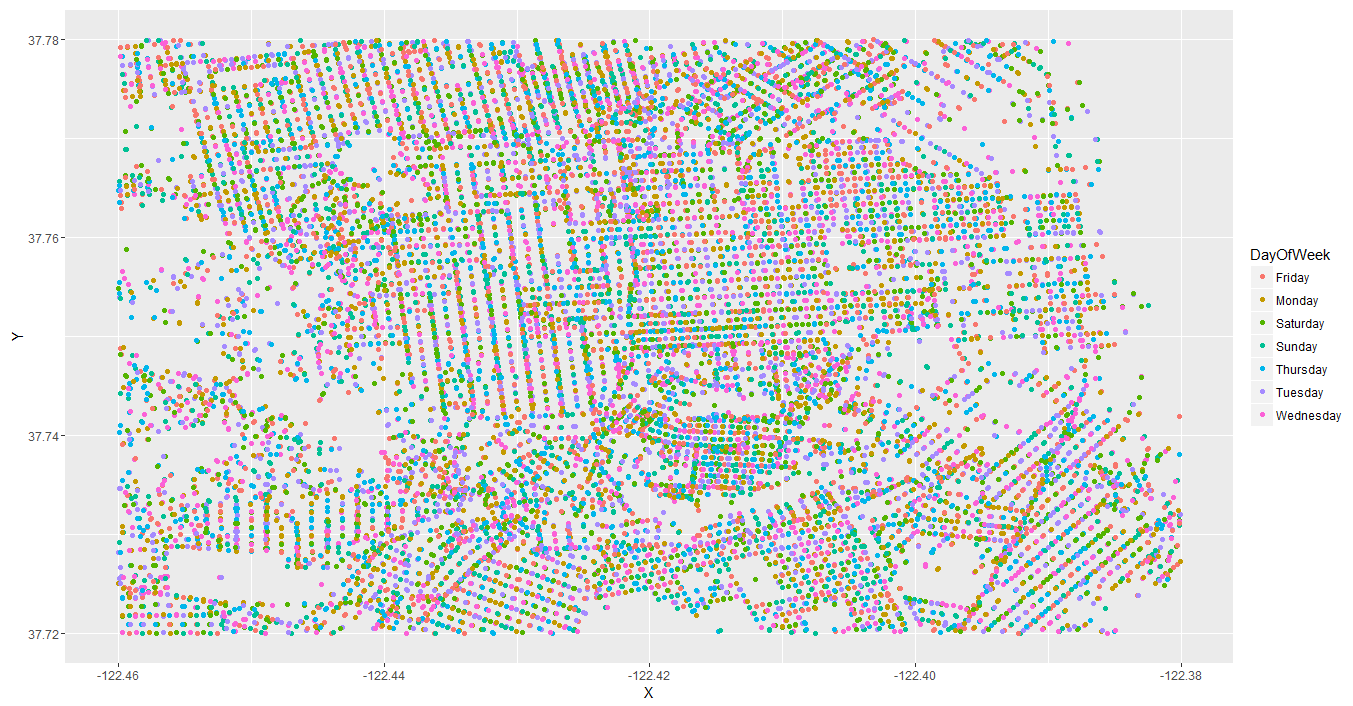


# setting blue colored triangular structured points for the observation

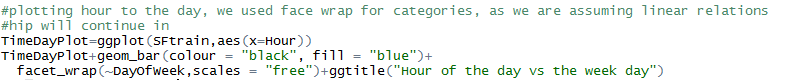
# we can see that we did not achieve anything from this map as the crimes are spread through

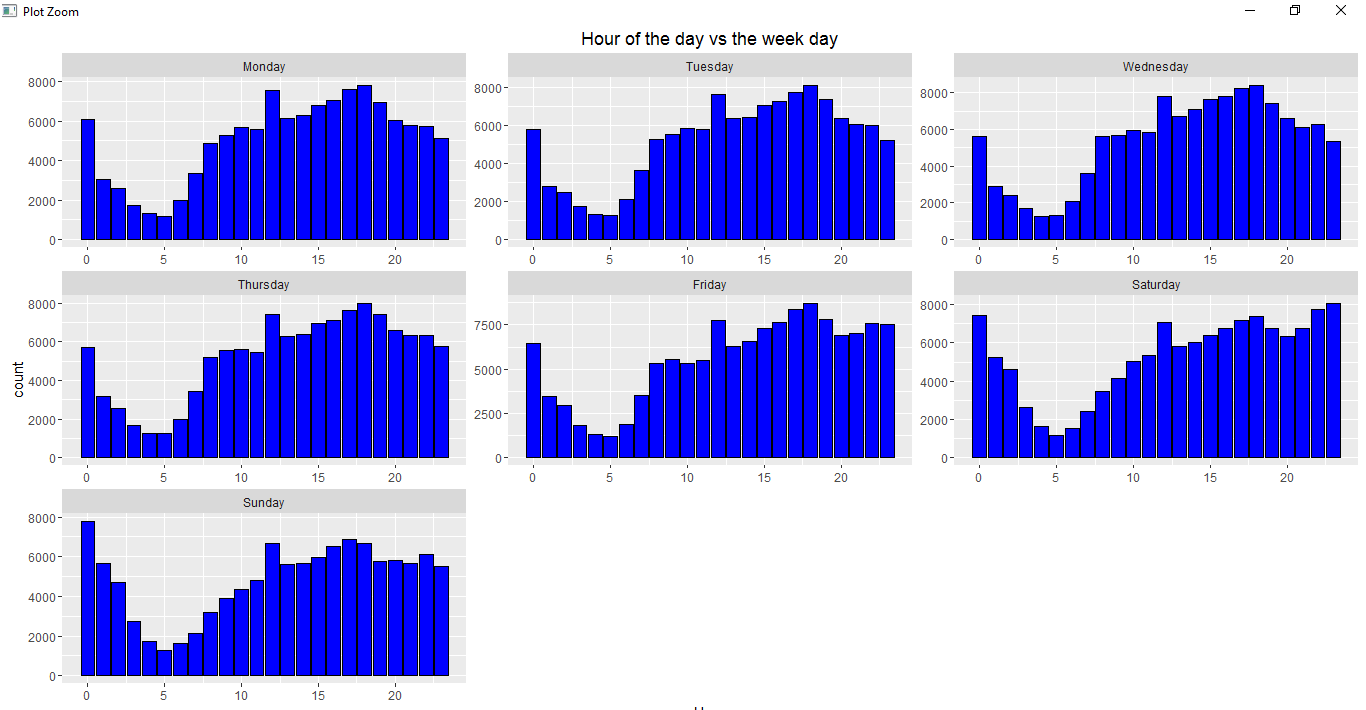
# the region, mostly we can observe pink dots which say crime occurred on Friday.



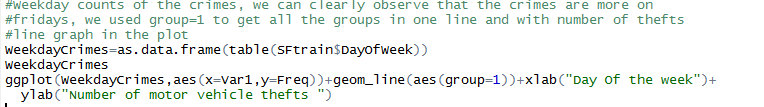


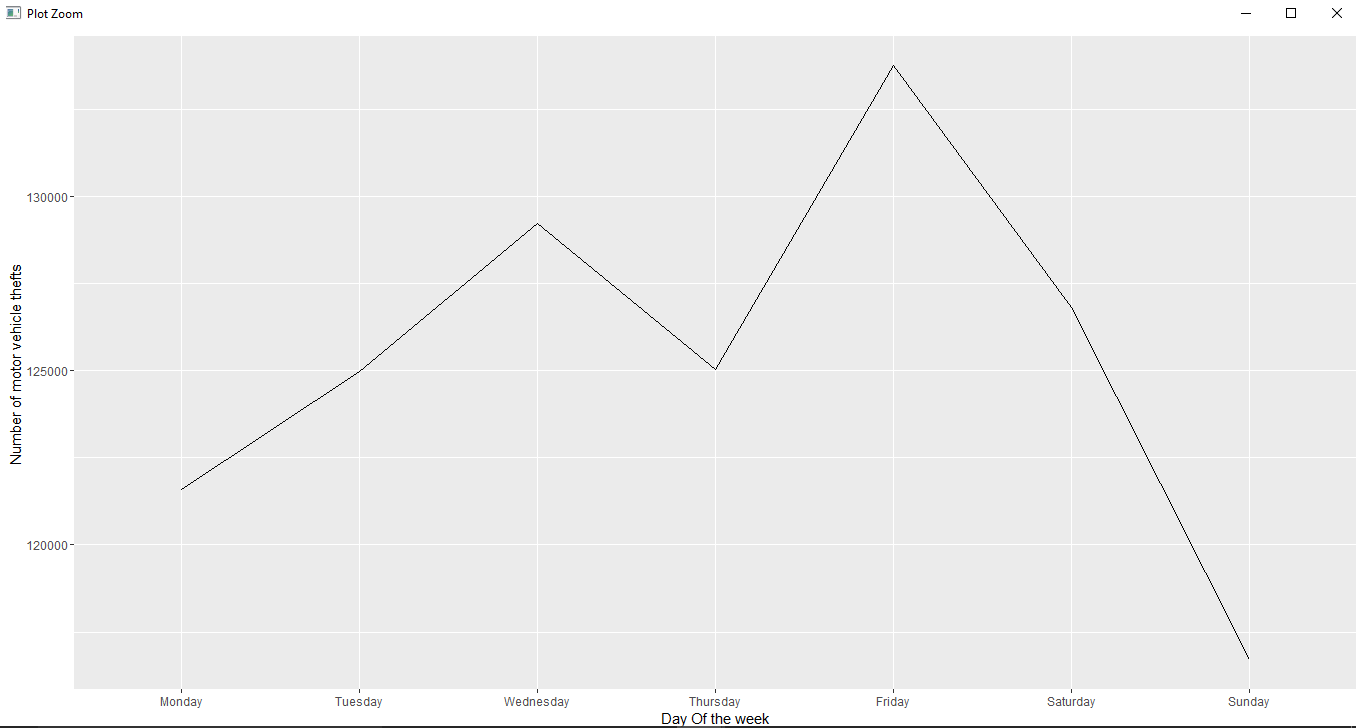
Plotting crimes based on the hour and day, we can see that on Saturdays after 10 there are large number of crimes happening , Which states that many people are celebrating as it is a normal weekends.



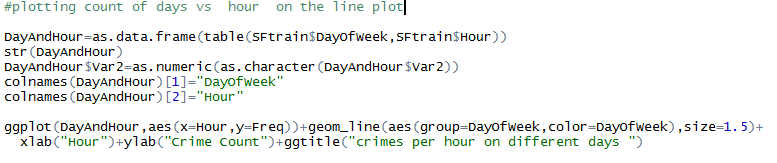


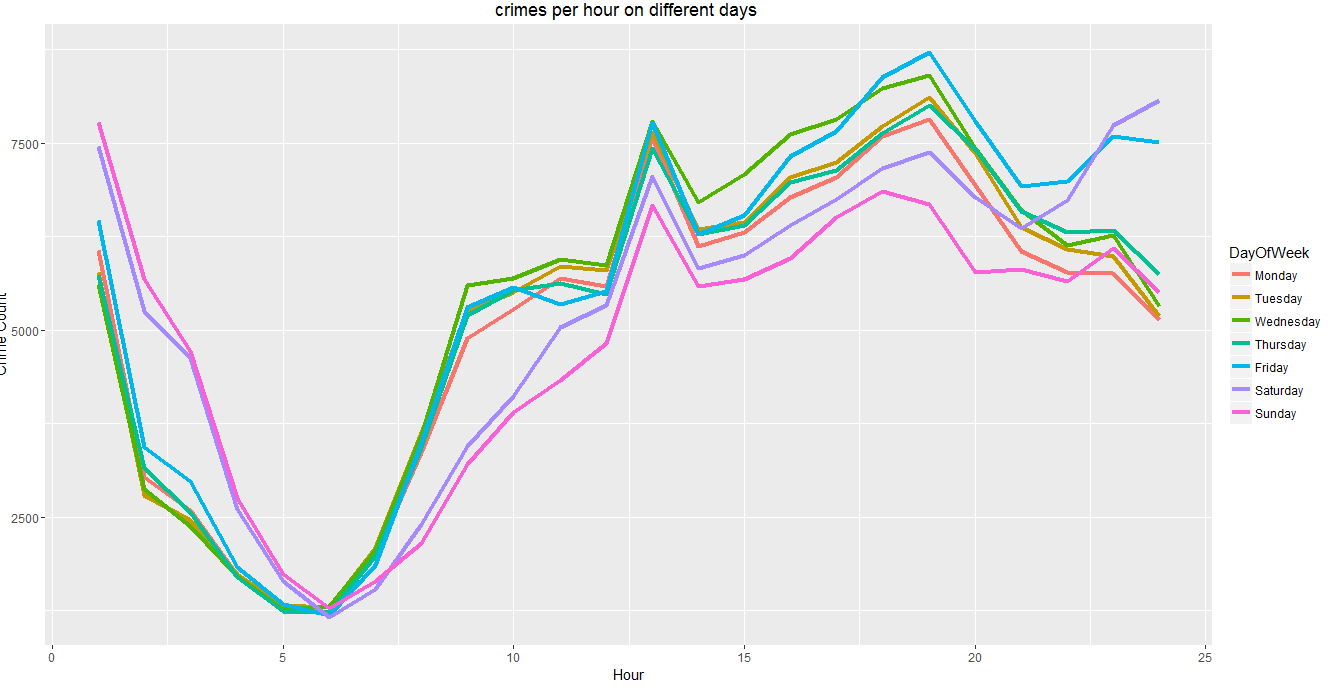
We just plotted a line plot to check trends in crimes in our data with respect to day, We can observe that Friday has most number of crimes.





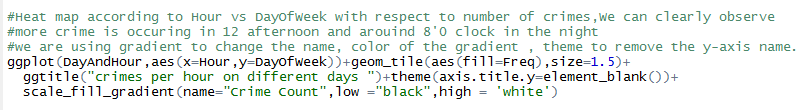
We will try putting the line plot of days vs hour the crime occurred and check for any trends in the data, we can observe

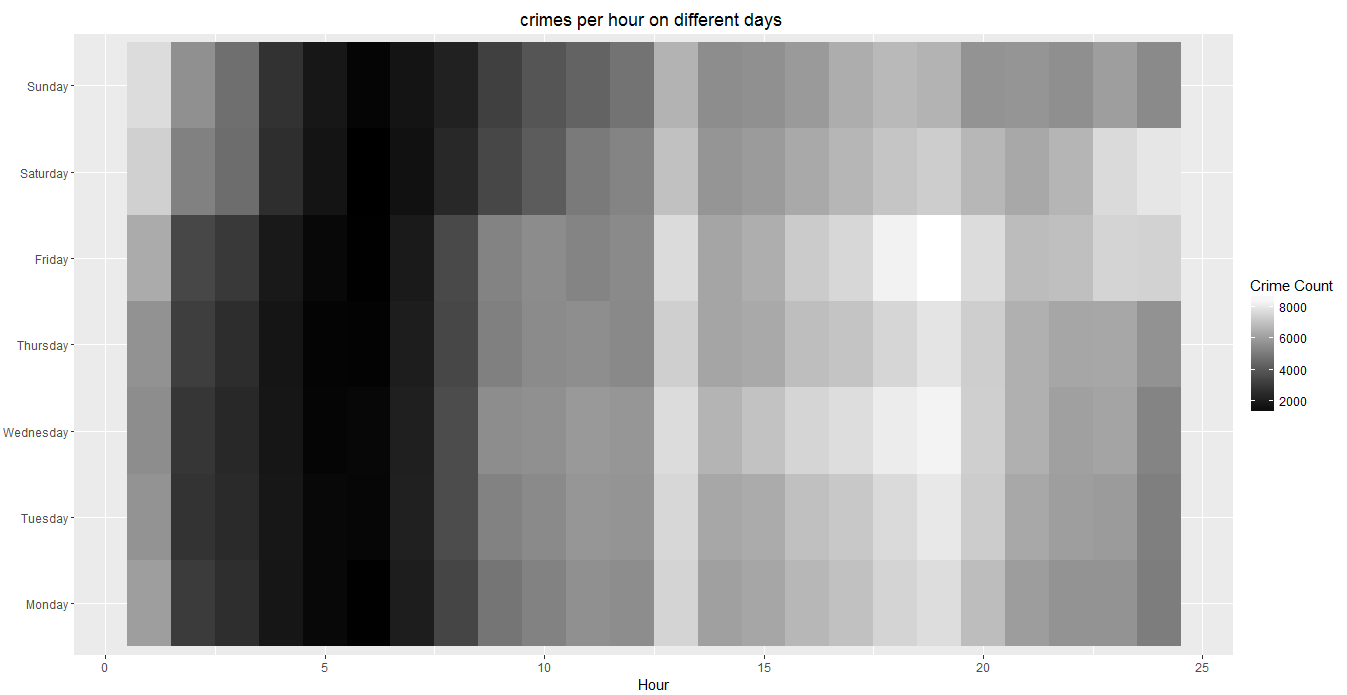




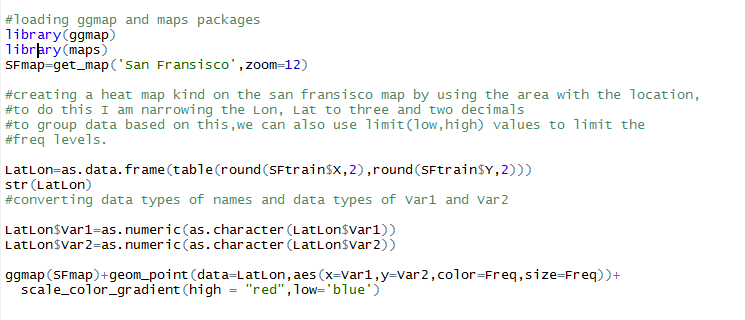
#Heat map according to day vs date with respect to number of crimes, we can clearly observe

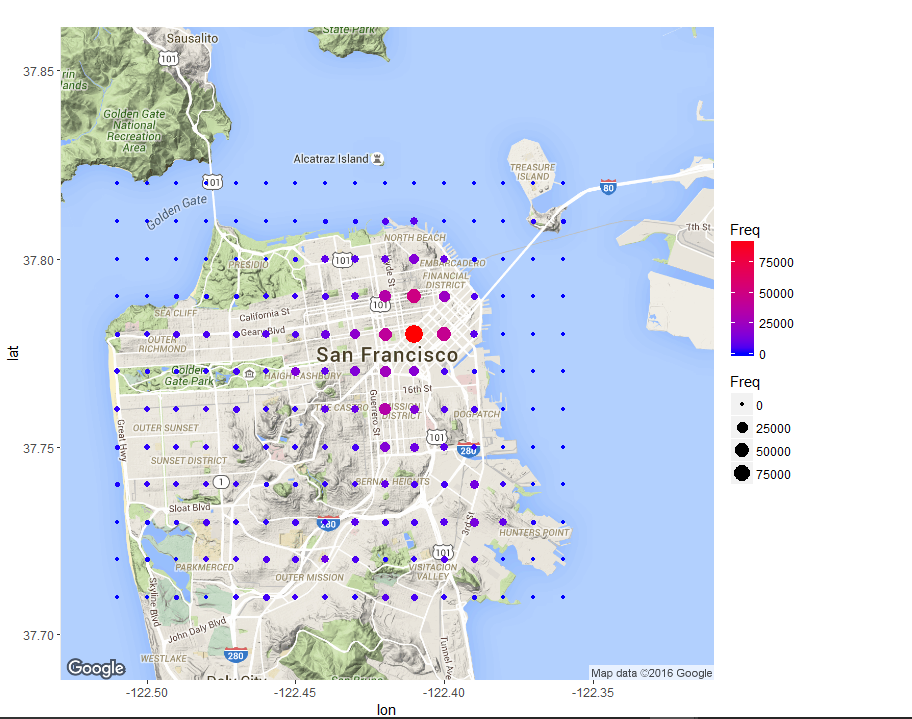
#more crime is occurring in 12 after noon and around 8'0 clock in the night



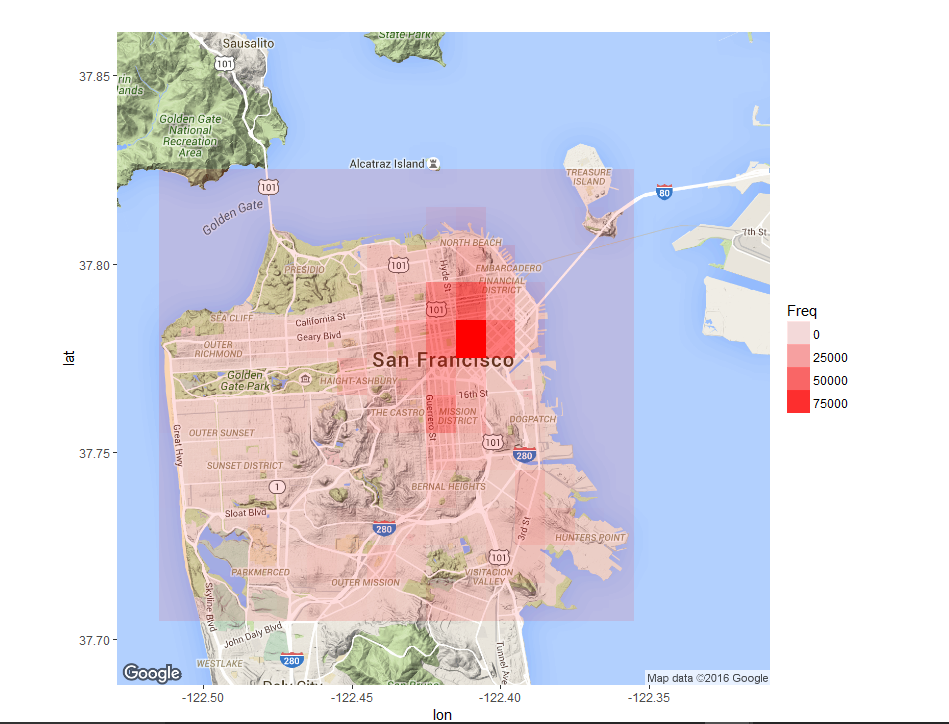


Big dots say more number of crims in a particular location, We can observe this in Freq meter.





I have plotted the above plot using the tile function which created a heat map, It is a better solution to visualization than points in my opinion. We have Freq to observe the number of times a crime occurred.



Finally, I have created heat maps with respect to each category, these are better solutions for visualization, I have added few of them here. However, when you run the code all the individual heat maps will be saved on your local for better visualization. Below way I have plotted all the categories and interestingly most of the crimes occur in financial district of San Francisco which is highlighted by dark red tile. This is obvious because there are lot of people working.