Abinaya Janakan
A20376287
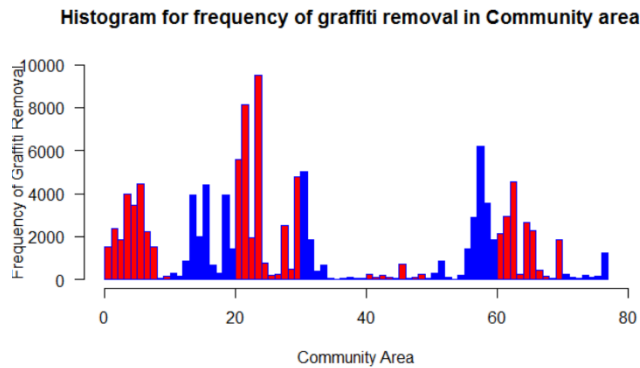
Homework 3

Histogram for 2016 incidents based on community area:

Histogram for frequency of graffiti removal in Community areas
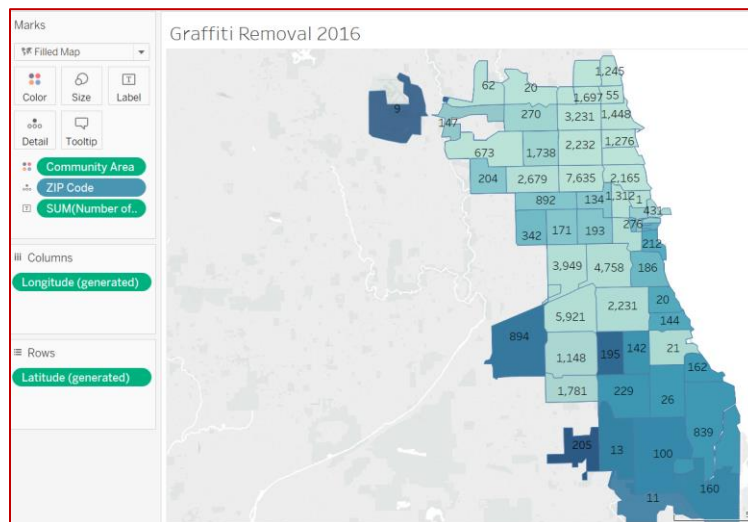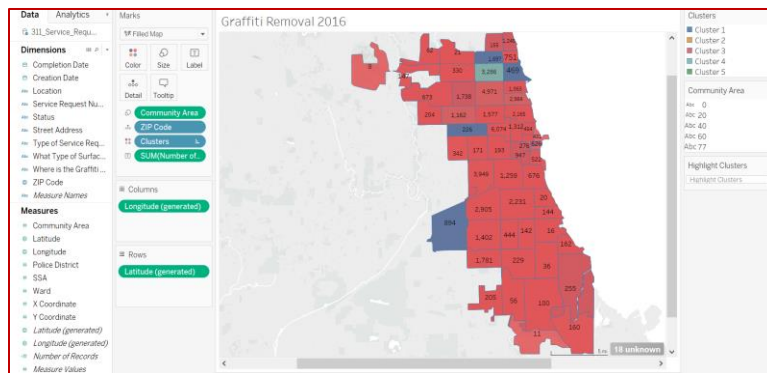

Histogram based on color coding:
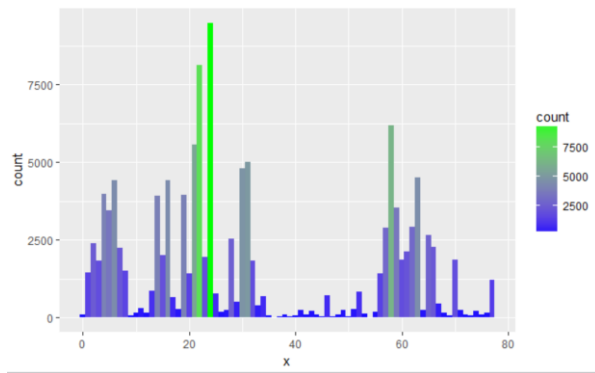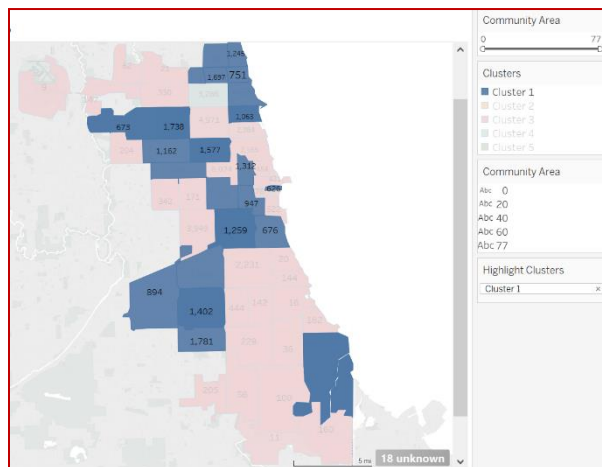

Map Based visualisation:
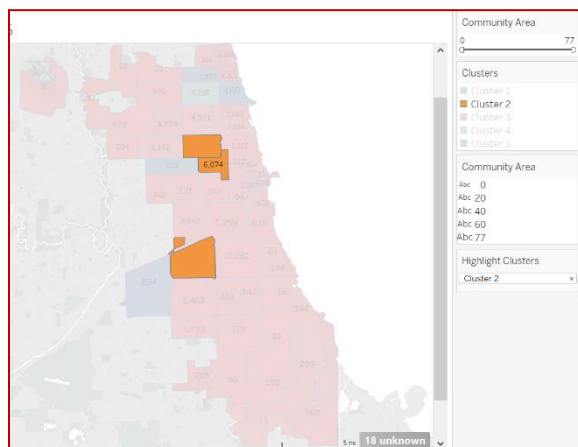
Clustering the map based on its values

Cluster 1 :

Range: 500 to 2000

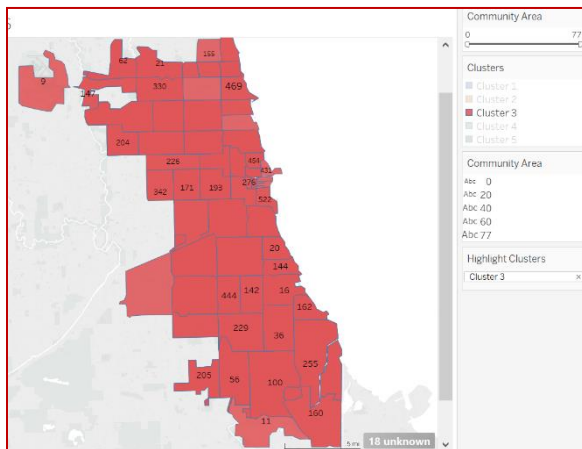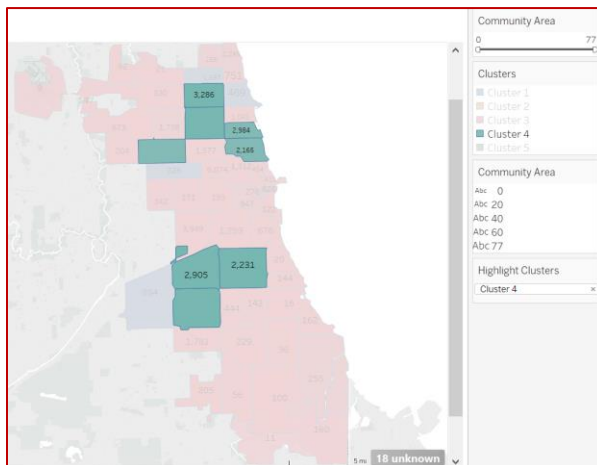

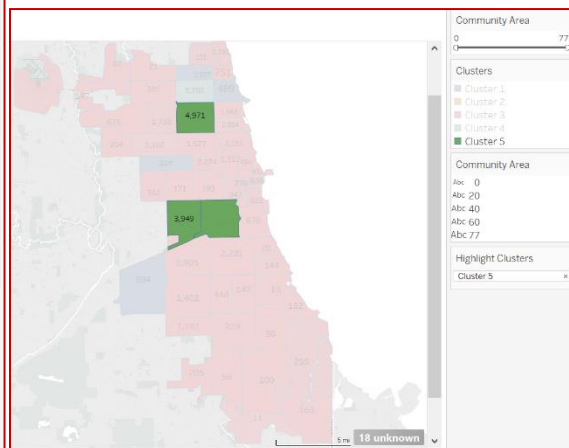Cluster 2: Range>5000

Cluster 3:

Range: 0 to 500

Cluster 4:

Range 2001to 4000

Cluster 5:

Range 3001 to 5000

Comparsion of Map based visualization vs histogram:

- The histograms takes just the count /frequency of incidents for each community area whereas the Map based visualization takes the longitude and latitude data along with the community area so it produces a much better visualization with a detailed information.
- Though we can color code the histogram data based on the range of values it doesn't give location based visualization as map does. Which helps in better analysis of data, Moreover in maps we can cluster the data based on the incidents along with the geographic location which could possibly help in finding the relation between the location and the graffiti incidents.
- Finding the highest and lowest values is quite straight forward in the histogram based visualization and which is quite tedious in map based visualization

Histogram for frequency of Street Light requests in 2016
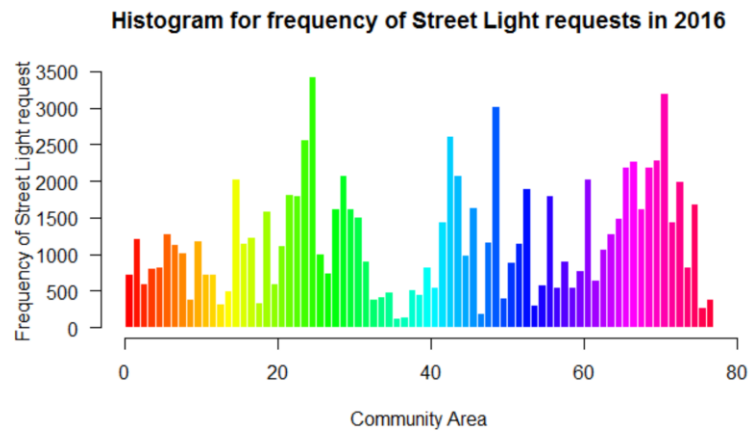
Color based coding:
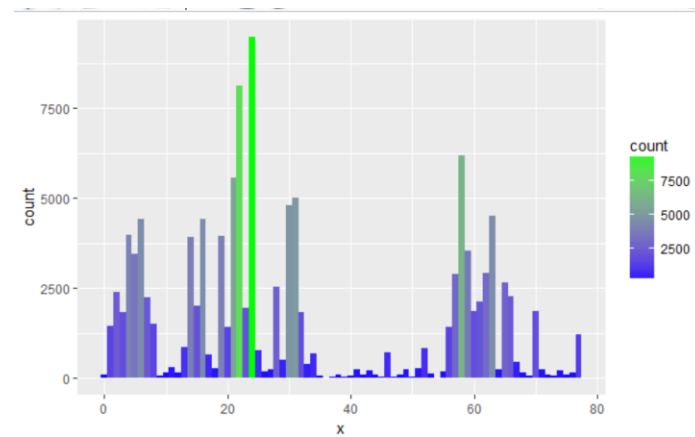
Combining Graffiti and street light data

graffiti Vs Street Light

Frequency

community area
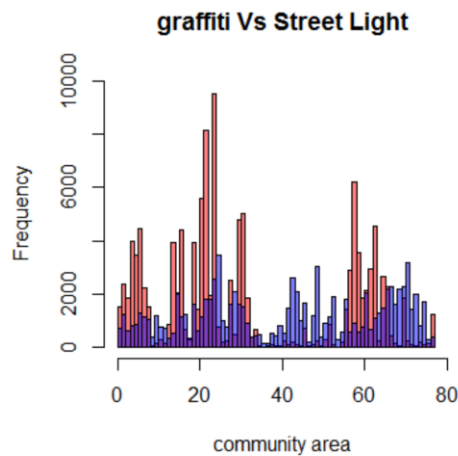
Relation between street light breakage and graffiti incidents:

The above graph shows the combined histogram of the graffiti and the street light incidents. The overlap between these two histograms is shown in violet color. Though there are some overlap between these two incidents there are lots of mismatch , some regions which has less graffiti incidents have higher street light incidents and vice versa.

Since we are comparing the data based on just the community area and frequency of the incidents the analysis may not be accurate because the date when the street light service request was received and the date when the graffiti incidents happened also play a huge role and more over chicago's broken street light data contains details of all the complains pertaining to street light.There may be other reasons for the complaint such as replacing worn out bulb or improving existing lights or they may be complains even due to the lack of proper lighting in particular area.Thus there are many factors influencing the broken street light incidents so we cant possibly say that graffiti incidents influence the street light complains.

Coding:

install.packages("readr")

library(readr)

install.packages("ggplot2", dependencies=TRUE)

library(ggplot2)

install.packages("COUNT")

#reading 2016 graffiti removal data

graffitidata <- read_csv("C:/Users/Abinaya/Desktop/graffitidata.csv")

graffitidata

View(graffitidata)

#Histogram for community area Vs Graffiti Data

colors = rainbow(length(unique(graffitidata$`Community Area`)))

colors = col<-c(rep("red",10),rep("blue",10))

x1 <- hist(graffitidata$`Community Area`,

    main="Histogram for frequency of graffiti removal in Community areas",

    xlab="Community Area",

    ylab = " Frequency of Graffiti Removal ",

    ylim=c(0, 10000),

    border="blue",

    col = colors,

    las=1,

    legend = rownames(graffitidata$`Community Area`),

    breaks=78)

Histogram using color

qplot(graffitidata$`Community Area`, geom="histogram",binwidth=1)

histo<-qplot(x =x, fill=..count.., geom="histogram",binwidth=1)

histo

histo+scale_fill_gradient(low="blue", high="green")

#reading 2016 street light request data

streetlightdata <- read_csv("C:/Users/Abinaya/Desktop/streetdata.csv")

streetlightdata

View(streetlightdata)

```
colors = rainbow(length(unique(graffitidata$`Community Area`)))

x2 <-hist(streetlightdata$`Community Area`,

    main="Histogram for frequency of Street Light requests in 2016",

    xlab="Community Area",

    ylab = " Frequency of Street Light request ",

    border="white",

    col = colors,

    las=1,

    breaks=78)

# Histogram using color

qplot(streetlightdata$`Community Area`, geom="histogram",binwidth=1)

histo_1<-qplot(x =x, fill=..count.., geom="histogram",binwidth=1)

histo_1

histo_1+scale_fill_gradient(low="blue", high="green")
```

#plotting graffiti and street light incidents

```
plot(x1, col=rgb(1,0,0,0.5),xlim=c(0,80), ylim=c(0,10000),,main='graffiti Vs Street Light',xlab='community
area')

plot(x2, col=rgb(0,0,1,0.5), add=T)
```