

# HW 5

Malik Hassan Qayyum

## Tasks:

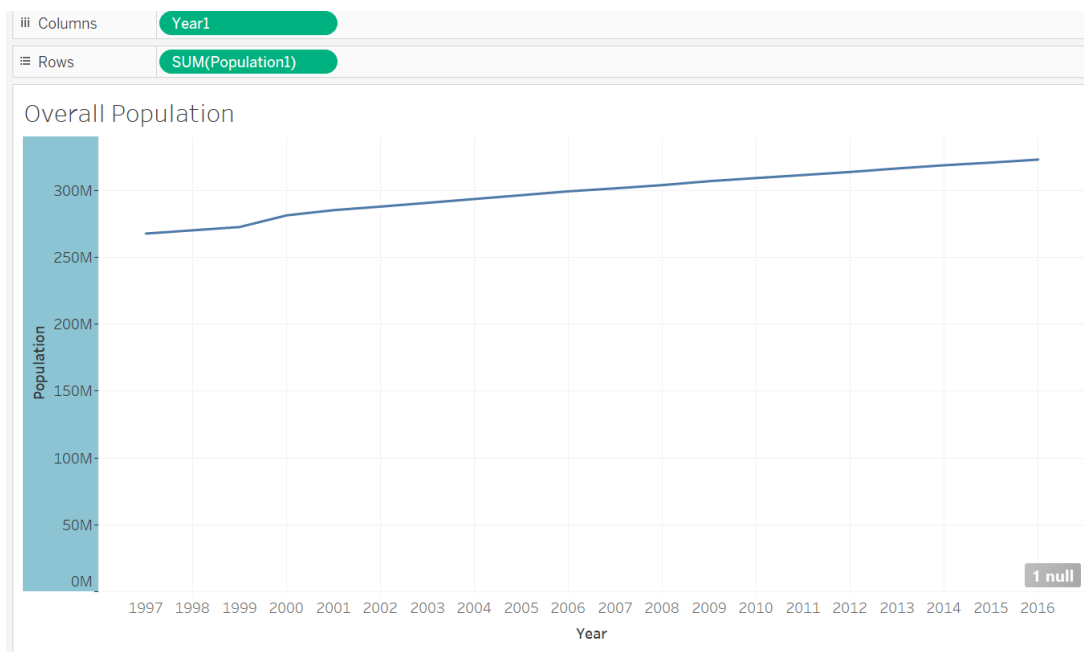
251

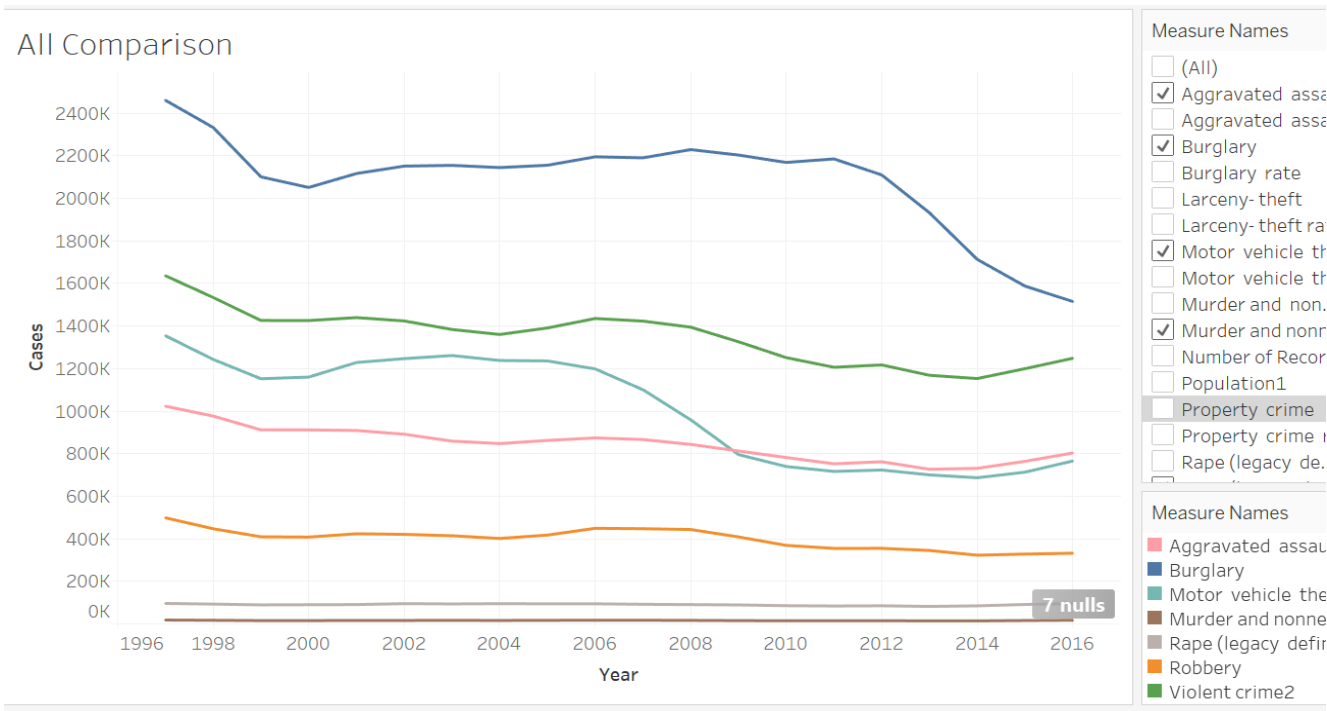
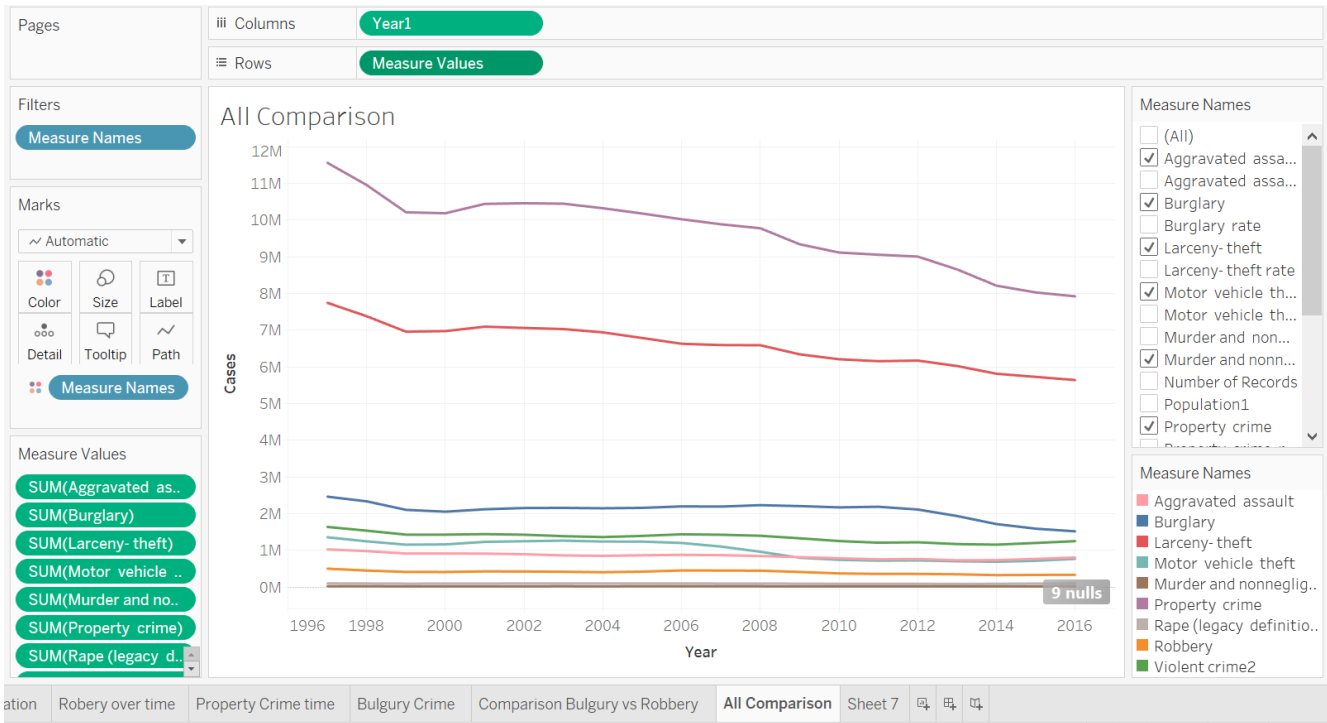
1. Plot population vs time from FBI data in either **Tableau** or R, each plot is expected to have an x-axis, y-axis, and title labels.
2. Now having explored the population data, write down a paragraph detailing your thoughts about how your observations of this plot expand your interpretation of the data from the last two assignments.
3. Choose 1 time series or distribution comparison graph (boxplot) from HW3 or HW4 that you're proud of and improve the visualization and communication it provides. Provide a snapshot of the original graph alongside the new graph to demonstrate the progress. Use only Tableau or R, NO EXCEL. This part of the assignment will be graded based on application of material learned in class, put in your best effort.
4. Give me an interesting analysis visual with a short supporting explanation for how you came about this idea and why is it useful in the understanding of the data.

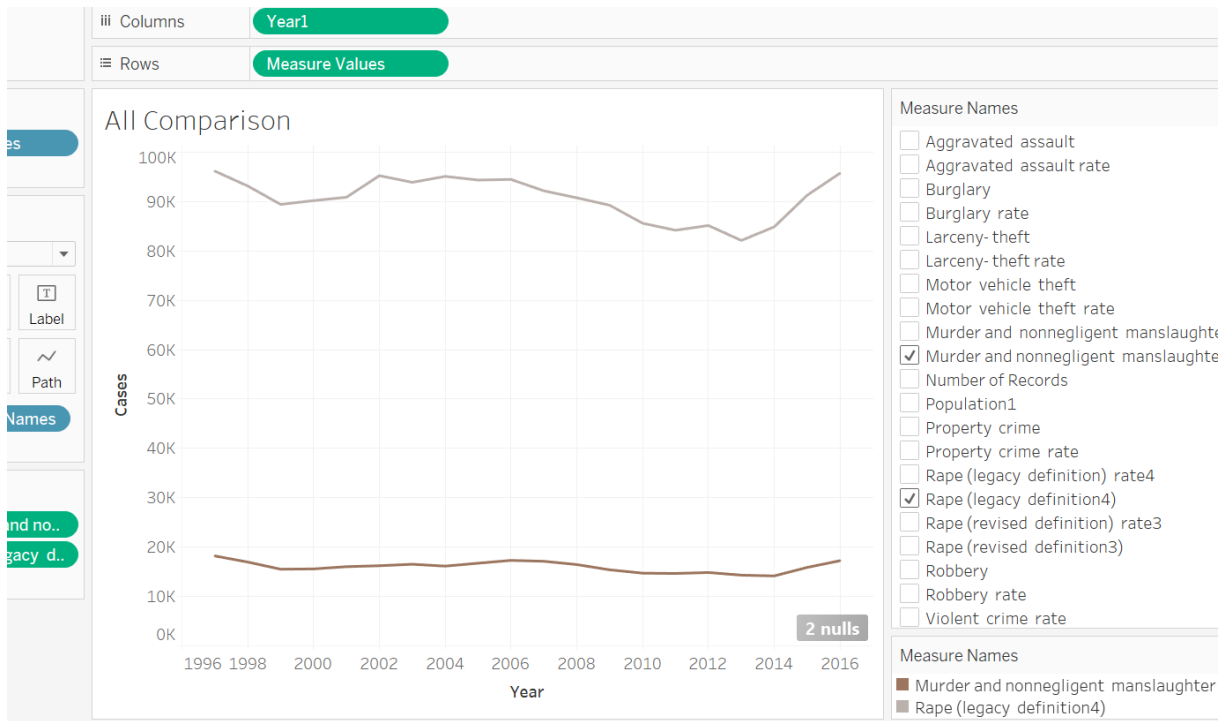
502

5. Same as above, but now setup the data so that your geographical map in tableau also has variable filters.

## 1. Population vs Time



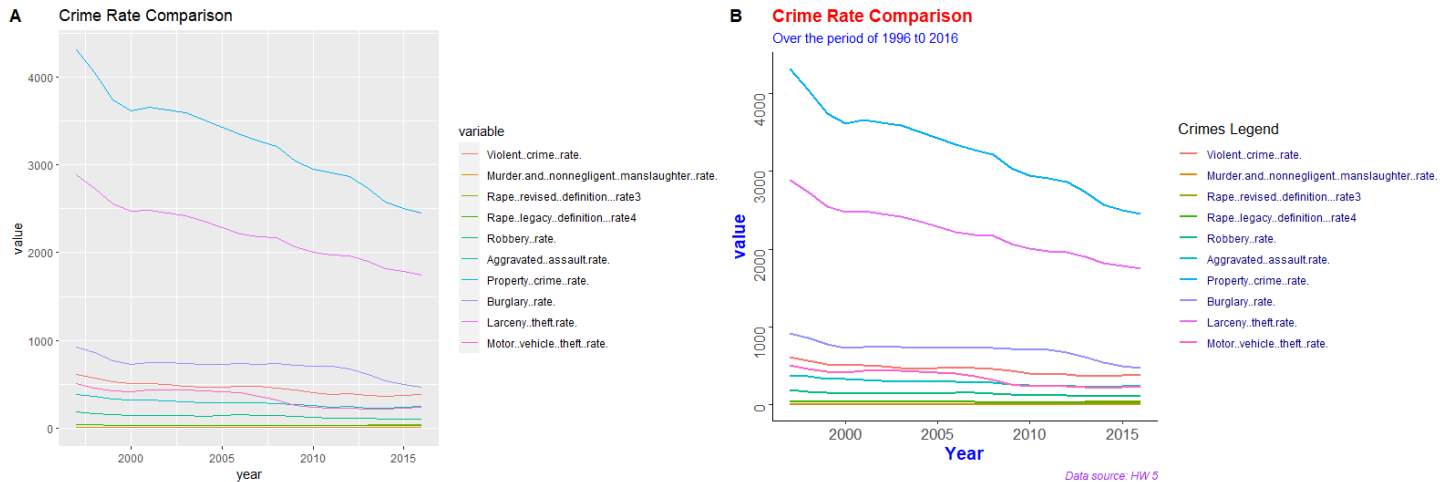




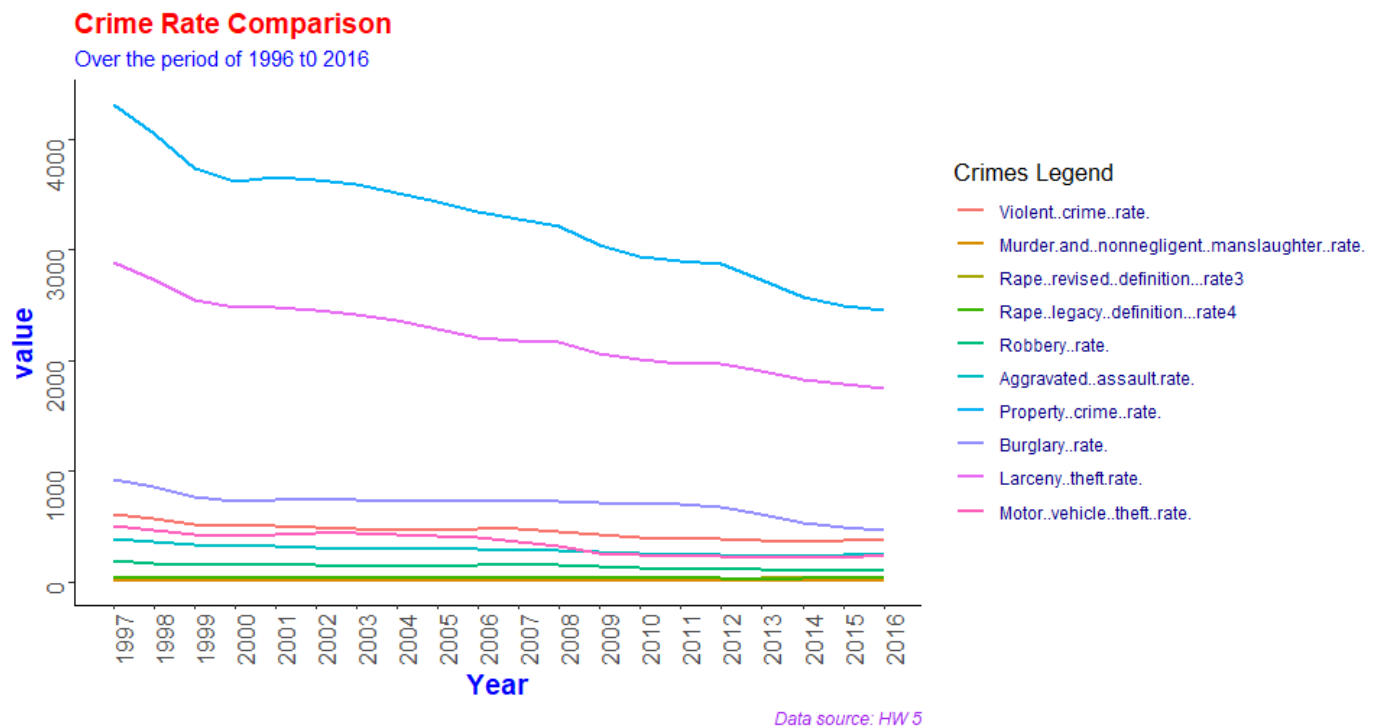
## 2. Description

From 1997 to 2016 Population has increased from roughly 265M to 210M in the USA. We can see an upward population growth. And from previous assignments, we can observe that despite the increase in population there is a decline in most of the crime rate. Overall, both the violent and property crimes both are decreasing in number over the given period. However, rapes and murders have shown spike up from the year 2014 to 2016. Murder numbers remain almost the same and the number of rapes showed a few fluctuations in between.

### 3. Improved plot



One More Change: Added and Rotated all xticks. Hence, all can be seen on screen without overlapping.



Code:

```
#####

crime_df <-read.csv("C:/Users/Malik/Documents/GitHub/Data-Visualization-
Data502/Dataset/DATA_FBI Crime Rate.csv")
crime_df_cases = crime_df[c(1:20),c(1,2,3,5,9,11,13,15,17,19,21)]
crime_df_rates = crime_df[c(1:20),c(1,2,4,6,8,10,12,14,16,18,20,22)]
names(crime_df_rates)
#####
```

```

crime_df_rates_melted = reshape2::melt(crime_df_rates,id.vars="Year", measure.vars=c(3:ncol(c
rime_df_rates)), value.name="value")
write.csv(crime_df_rates_melted,"C:/Users/Malik/Documents/GitHub/Data-Visualization-
Data502/Dataset/DATA_FBI Crime Rate_melted_r.csv")

# Line Plot
p = ggplot(crime_df_rates_melted, aes(Year,value,color=variable)) + geom_line() + xlab("year"
) + ylab("value") + ggtitle("Crime Rate Comparison")
p

# improving line graph

plt = ggplot(crime_df_rates_melted, aes(Year,value,color=variable)) + geom_line(lwd=1) + lab
s(title = "Crime Rate Comparison",
  subtitle = "Over the period of 1996 t0 2016",
  caption = "Data source: HW 5",
  color = "Crimes Legend") + scale_x_continuous(breaks=seq(1997, 2016, 1))

plt = plt+ theme_bw() + theme(panel.grid.major = element_blank(), panel.grid.minor = element_
blank(),
  panel.background = element_blank(),
  axis.line = element_line(colour = "Black"),
  panel.border = element_blank())

plt = plt+theme(axis.text.y = element_text(angle = 90,size = 13),
  axis.text.x = element_text(size = 13,angle = 90))

plt = plt + theme(plot.title = element_text(color = "red", size = 14, face = "bold"),
  plot.subtitle = element_text(color = "blue"),
  plot.caption = element_text(color = "purple", face = "italic"))

plt = plt+theme(axis.title.y = element_text(angle = 90, color = "Blue", size = 14, face = "bo
ld"),
  axis.title.x = element_text( color = "Blue", size = 14, face = "bold"))

plt = plt+ theme(legend.title = element_text(color = "black", size = 13),
  legend.text = element_text(color = "navy"))

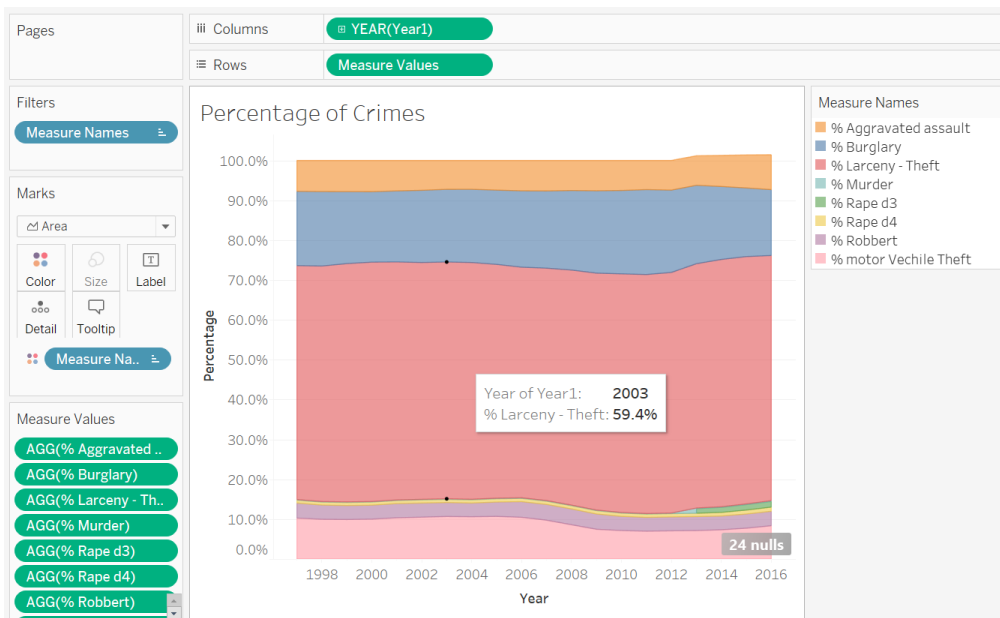
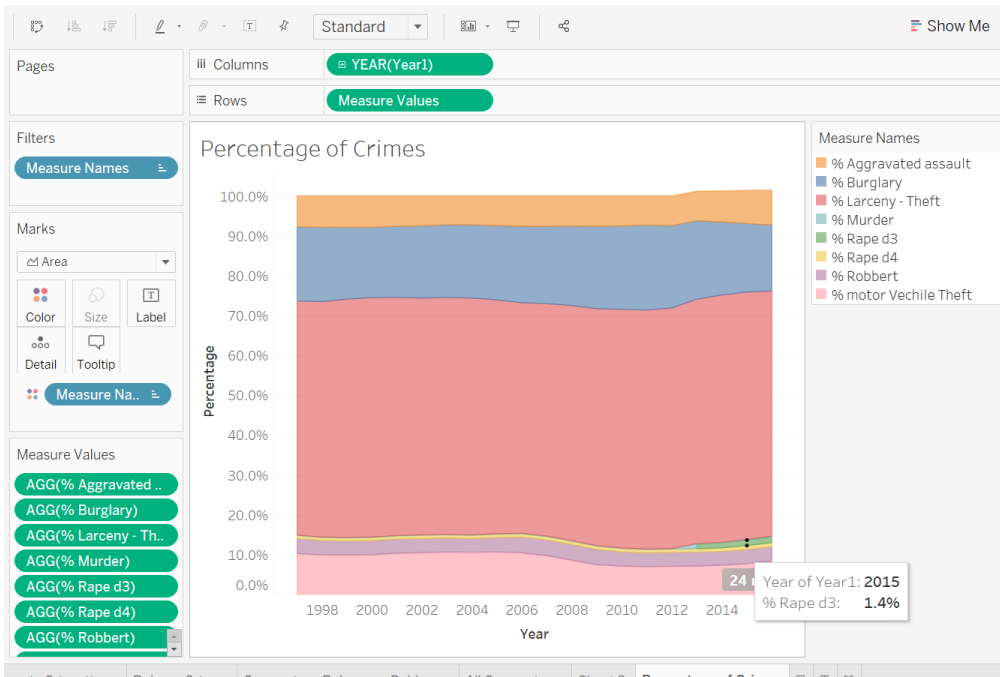
plt

# Plting together
plot_grid(p, plt, labels = "AUTO")

```

---

## 4. A New Visual



How I Designed it:

I Created this Area plot to observe what percentage of a particular crime has been throughout the period.

A formula that I used for *Calculated fields*:

Percentage of crime = Sum (That Crime) / (Sum (Property Crime) + Sum (Violent Crime)

This gives the proportion of crimes from overall crimes.

Observation:

Over the time proportion of the crimes almost remains the same. Murder and Rape are less than the 5% of overall crimes throughout the period.

Moreover, larceny theft has the dominant proportion around 55% to 65% of overall crimes. Rest of the crimes has around 10 to 20 proportion for each of them.

5. Geographical Crimes with filter

