**Homework 3\_502**

**Name: Malik Hassan Qayyum**

**Data: FBI Crime Data, FBI Crime Data Geographically**

**Task:**

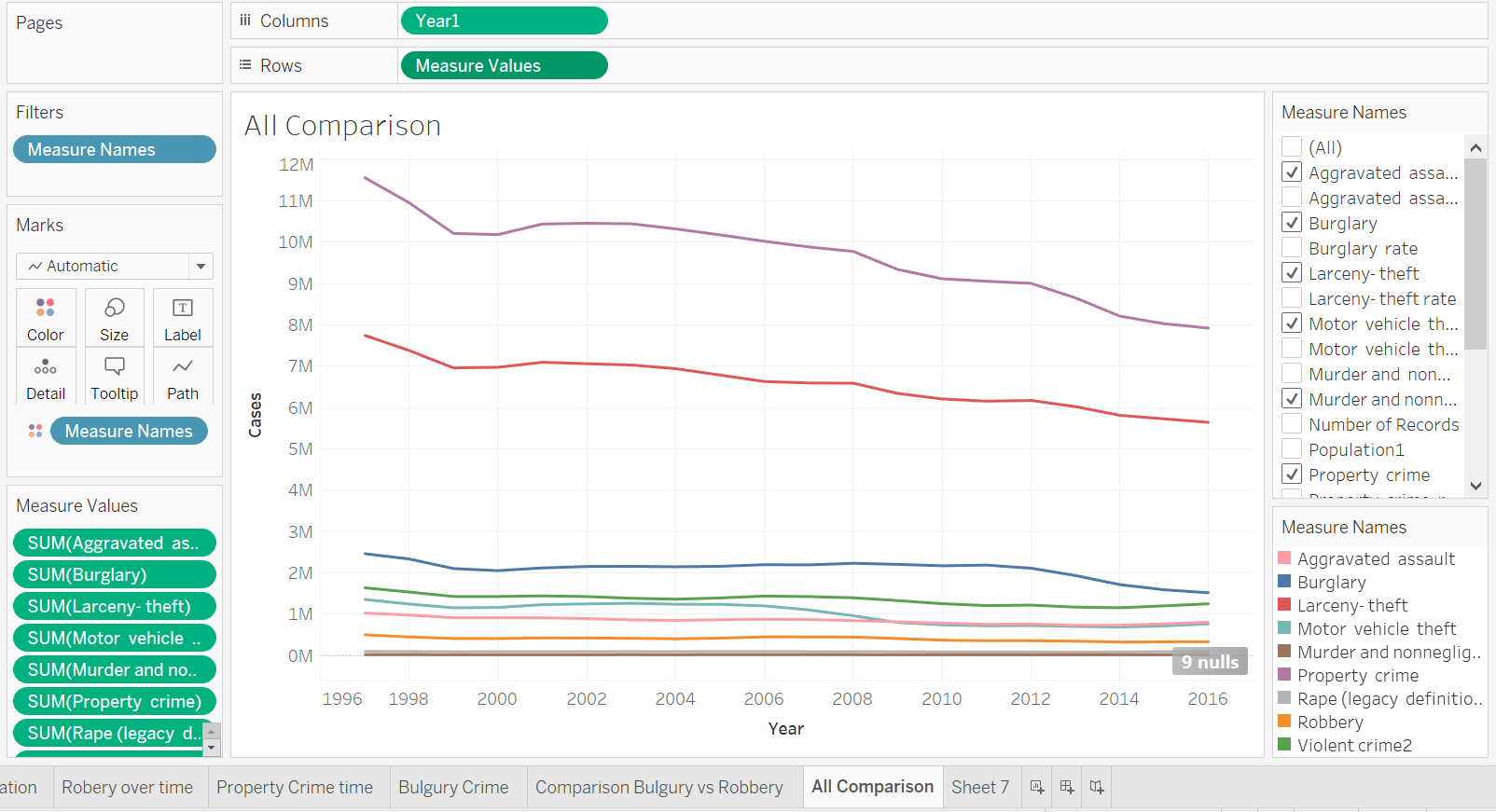
Using all three software (Excel, Tableau, R) create a times series plot for the data as well as a visual demonstrating the categorical nature of the data (ie comparison/distribution)

Label each of the x and y axis, along with a title for all plots

Briefly explain what you discover from your visuals as if you were presenting to an audience.

Create an additional plot in Tableau using the Geographical data file and justify your reasoning for the use of this plot.

**Tableau:**



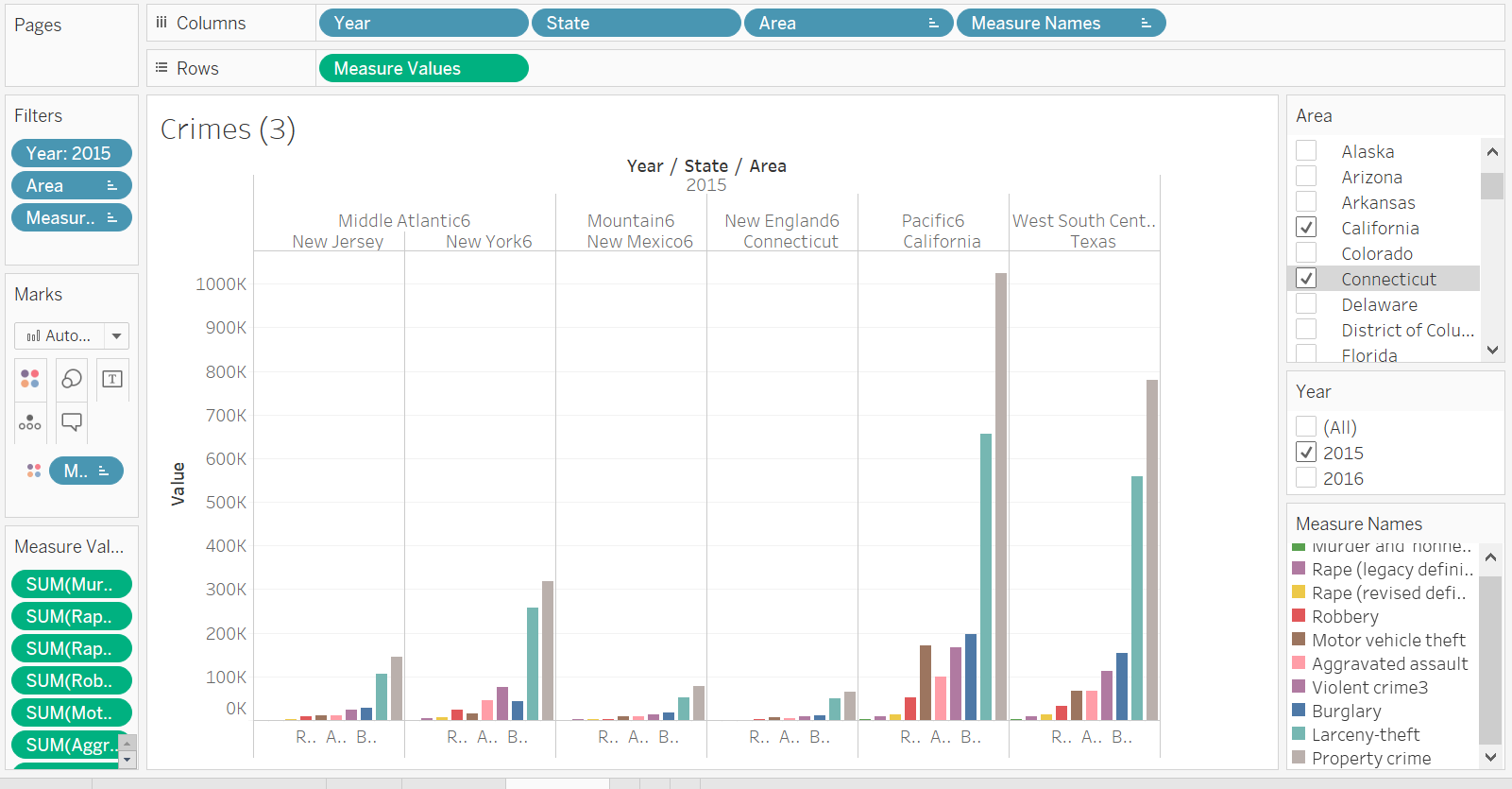
We can see that Property crime and larceny- left are the highest ones in numbers. Murders and Rapes are least in numbers as compared to other crimes throughout the period of 1997 to 2016. There is decrease in numbers of Property crime and larceny- left cases over the given period.

Note: Crime rate can also be used. It gives same comparison.



From 1997 to 2016 Population has increased from roughly 265M to 210M in the USA.

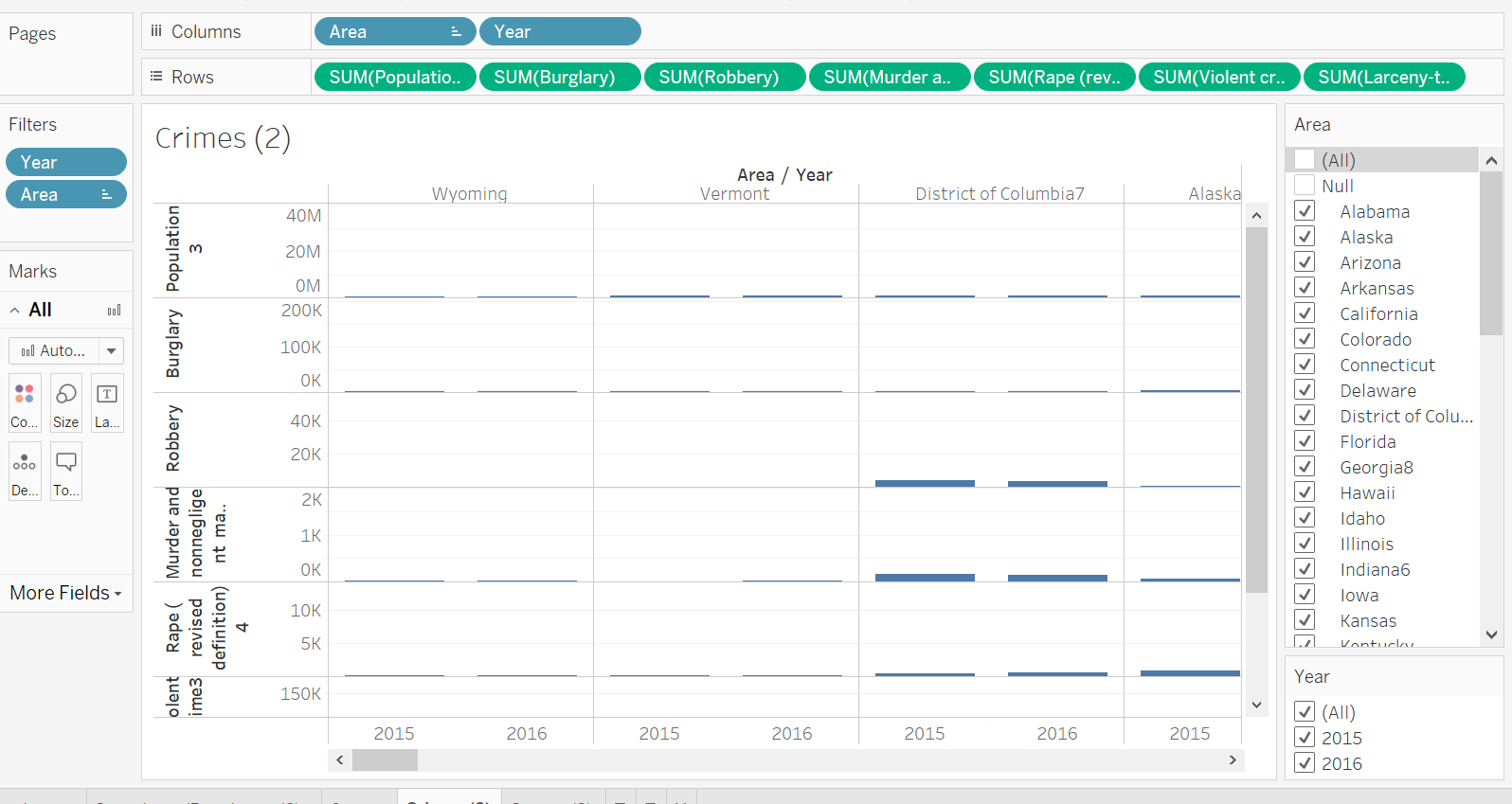
Geographical Data:



**Note:** I have selected only a few states and regions so the graph can fit on screen. States and year can be selected form the filters pane on the right side.

The same pattern can be observed for every state. Each state has Property crime and larceny- left as top crimes.





California, Taxes, Florida, New York has the highest crime rates. [ That could be biased we need to consider the population for further conclusions]

These states have a high population and high crime rates.

Similarly, Wyoming, Vermont, and the District of Columbia have the lowest crime but the population is also low.

**Excel:**

**R - Studio**

Code:

# install.packages("scales")

library(ggplot2)

library(scales)

#################################################################################

crime\_df <-read.csv("C:/Users/Malik/Documents/GitHub/Data-Visualization-Data502/Dataset/DATA\_FBI Crime Rate.csv")

#################################################################################

p = ggplot(data=crime\_df, aes( x=Year,y=Population1))

scale\_it = scale\_y\_continuous(labels = function(x) format(x, scientific = FALSE))

p + geom\_line() +

  scale\_it +

  ggtitle("Poplutaion over years")

#--------------------------------------------------------------------------#

# All Crimes

ggplot(crime\_df, aes( x=Year)) + scale\_it +

  geom\_line(aes(y =Violent.crime2, colour = "Violent.crime2")) +

  geom\_line(aes(y =Rape..revised..definition3., colour = "Violent.crime2")) +

  geom\_line(aes(y =Rape..legacy..definition4., colour = "Violent.crime2")) +

  geom\_line(aes(y = Murder.and.nonnegligent..manslaughter , colour = "Murder.and.nonnegligent..manslaughter")) +

  geom\_line(aes(y =Robbery, colour = "Robbery")) +

  geom\_line(aes(y =Aggravated..assault, colour = "Aggravated..assault")) +

  geom\_line(aes(y =Robbery, colour = "Robbery")) +

  geom\_line(aes(y =Property..crime, colour = "Property..crime")) +

  geom\_line(aes(y =Burglary, colour = "Burglary")) +

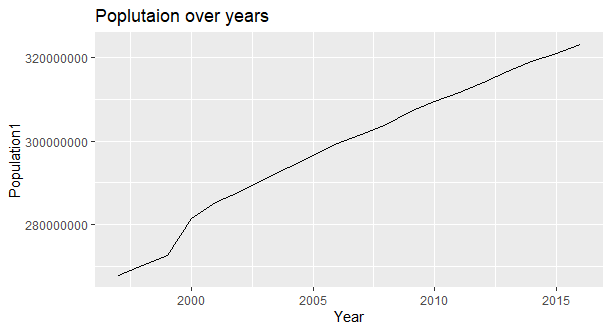
  geom\_line(aes(y =Larceny..theft, colour = "Larceny..theft")) +

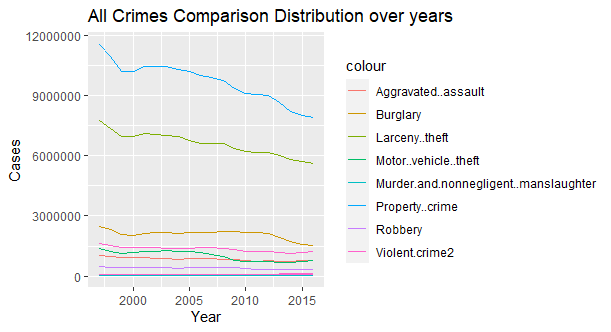
  geom\_line(aes(y =Motor..vehicle..theft, colour = "Motor..vehicle..theft"))+

  ggtitle("All Crimes Comparison Distribution over years") +

  ylab("Cases")

Plots:





----- END -----