Report 1

Analysis of the correlation with the crime rate

December 9, 2023
Seogyeong Hwang

1. Introduction

Crime is always happening around us. There are various criminalities in the world, and there may or may not be reasons why crimes occur. Therefore, we tried to visualize and analyze crime-related data from Statis.org through graphs to see which items affect the crime rate. This has been studied using aggregate data on 47 states of the USA for 1960. The data set contains the following items: ‘M: percentage of males aged 14-24 in total state populations’, ‘Ed: mean years of schooling of the population aged 25 years or over’, ‘Po1: per capita expenditure on police protection in 1960’, ‘Po2: per capita expenditure on police protection in 1959’, ‘LF: labour force participation rate of civilian urban males in the age-group 14-24’, ‘M.F: number of males per 100 females’, ‘Pop: state population in 1960 in hundred thousands’, ‘NW: percentage of nonwhite in the population’, ‘U1: unemployment rate of urban males 14-24’, ‘U2: unemployment rate or urban males 35-39’, ‘Wealth: median value of transferable assets or family income’, ‘Ineq: income equality; percentage of families earning below half the median income’, ‘Prob: probability of imprisonment; ratio of number of commitments to number of offenses’, ‘Time: average time in months served by offenders in state prisons before their first release’.

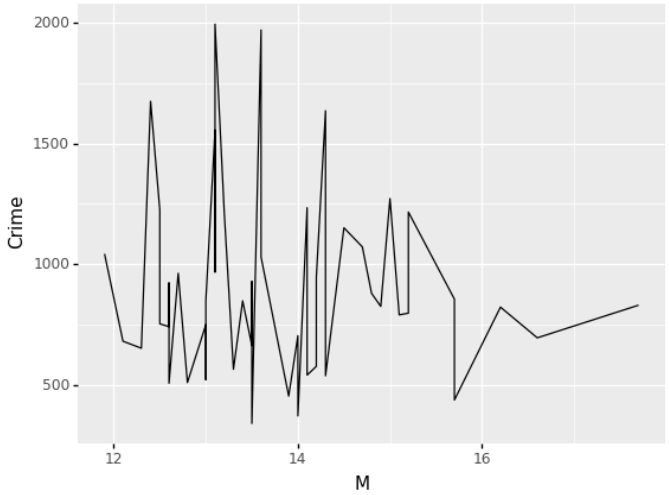
In this report, we checked how much each item affected the crime rate and which one affected the crime rate the most.

2. Method

Three methods were used in this code.

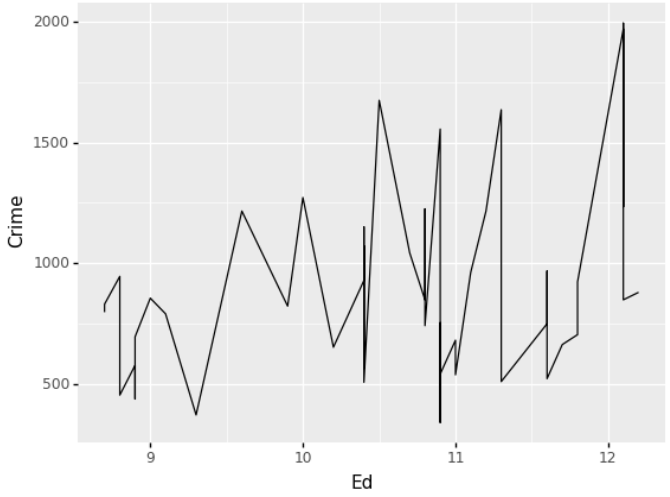
- pd.read_csv() : To read csv file into dataframe
  However, when I was reading the file, the comma delimiter did not work, so I changed the delimiter to ‘\t’.
- to_csv() : To make dataframe as a csv file
- ggplot() : To plot data
  for the graph type, I added ‘geom_line()’.

3. Result



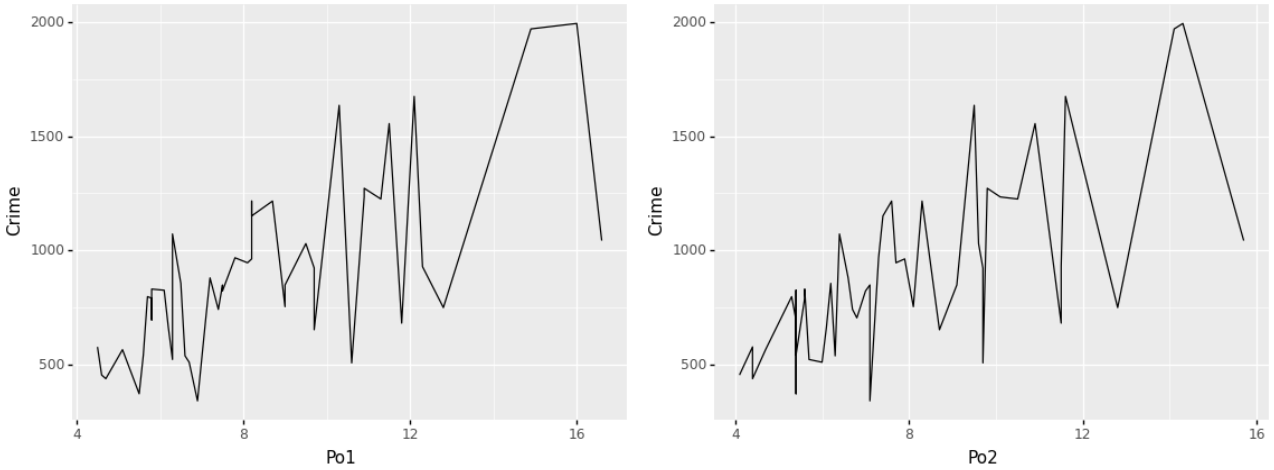
< rate of change with Percentage of 14-24 aged men >

: The highest crime rate occurs when the percentage of 14-24 aged men is 12 ~14%.
: If it is bigger than 14%, the crime rate shows a downward trend.



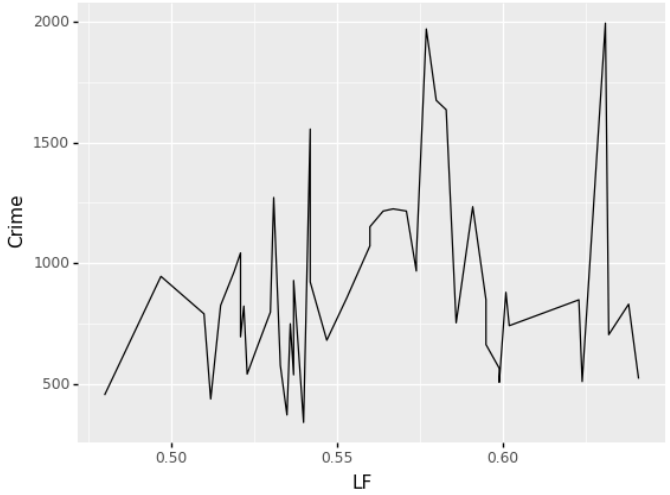
< rate of change with Means years of Education aged 25 years or over >

: Overall, The graph shows a trend where the crime rate increases, as the mean years increase.



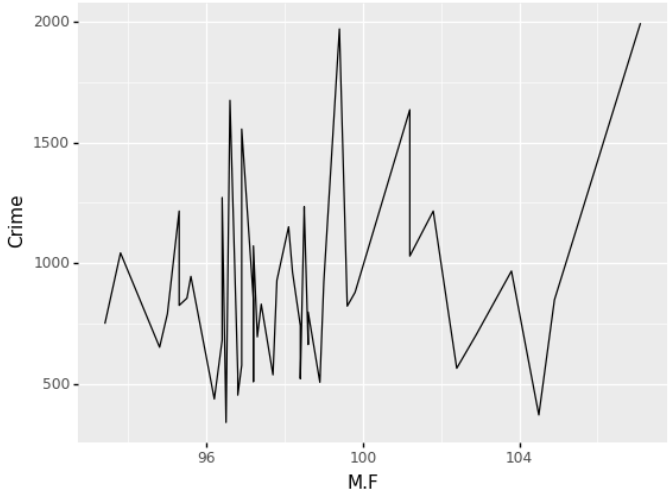
< rate of change with per capita expenditure on Police Protection >

: Left graph is in 1960, the right graph is in 1959. However, these have almost the same trend.
: Both of them show that the larger the expenditure on police protection is, the crime rate increases.



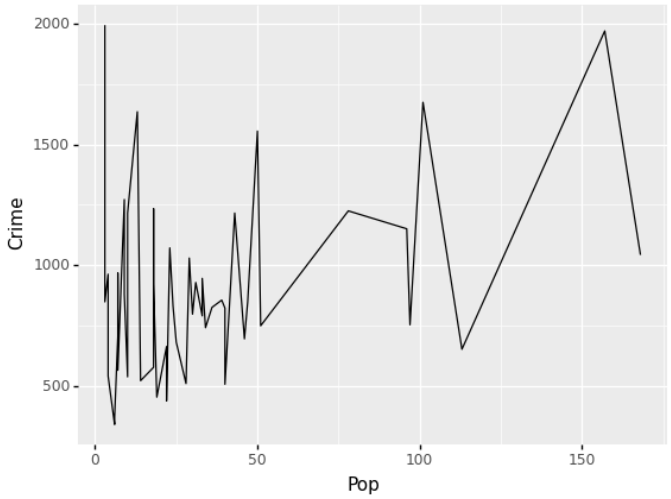
< rate of change with Labour Force participation rate of 14-24 aged males >

: In this graph, when X data is 0.55~0.60, the average crime rate is higher than other sections.



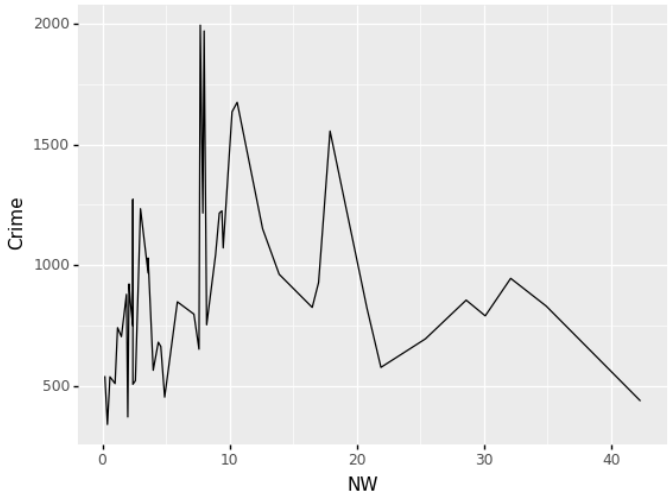
< rate of change with number of Males per 100 Females >

: The middle x value has the highest crime rate.
: At the end, when the number of males are going up, crime rate is also going up.



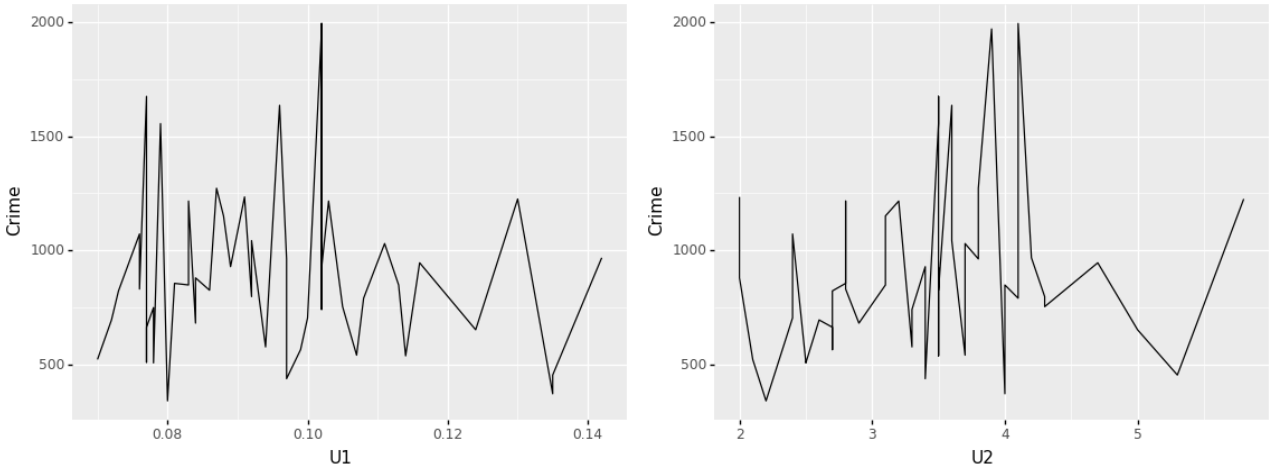
< rate of change with state Population in 1960 >

: The unit is hundred thousands.
: It goes up and down in large amounts like a frequency.

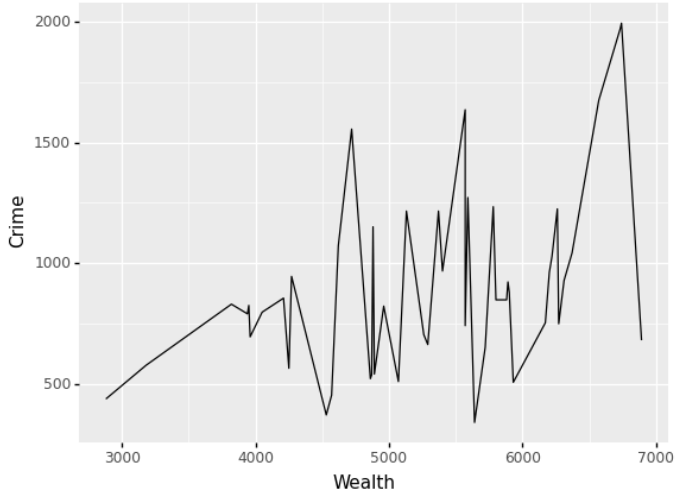


< rate of change with percentage of Nonwhites in the population >

: Before reaching 10%, the rate is increasing rapidly, but it starts to decrease after reaching 10%.

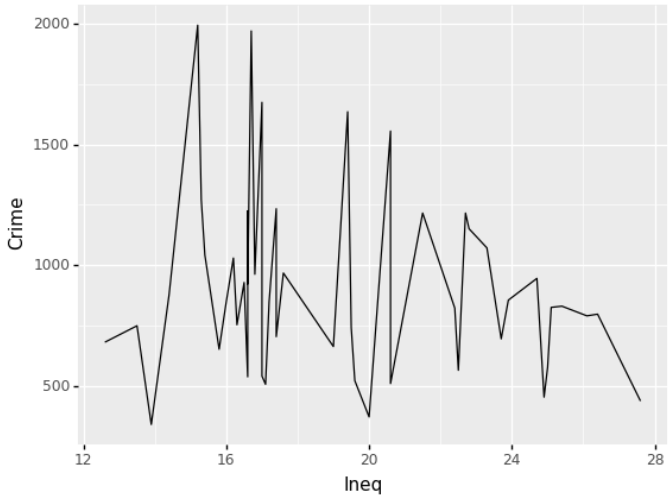


< rate of change with Unemployment rate or urban males >
: Left graph has 14-24 age group, and the right graph has 35-39 age group.
: For the 14-24 age group, the lower the unemployment rate, the higher the crime rate.
: For the 35-39 age group, this graph goes from low to high, and then low again, with the center of the unemployment rate value having the highest crime rate.



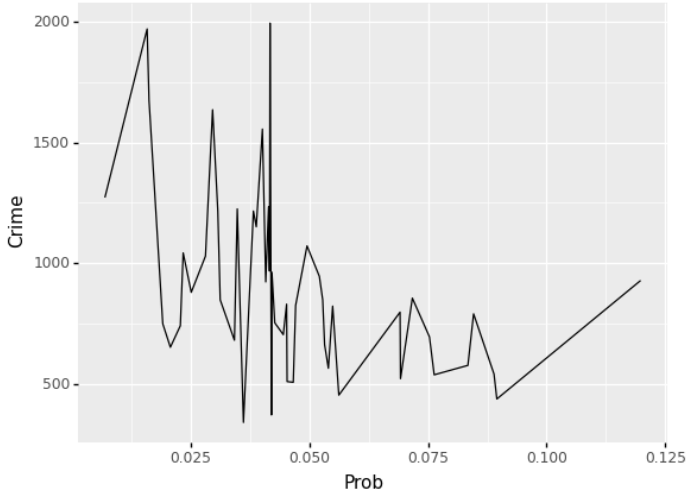
< rate of change with Wealth >

: This graph is showing a trend that when the wealth value goes up, the crime rate value also goes up.



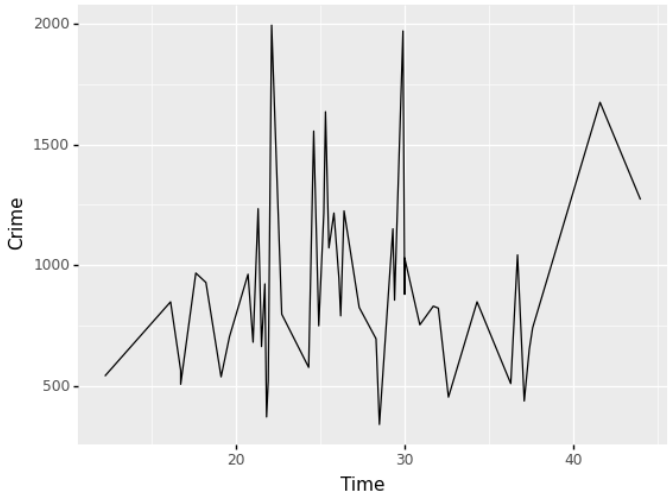
< rate of change with income Inequality >

: It has a gradually lowering shape. The percentage of families earning below half the median income is around 16%, it has the highest crime rate value, but when the x value goes down, the crime rate also goes down.



< rate of change with Probability of imprisonment >

: If the ratio of probability of imprisonment is less than 0.05, the crime rate is the highest.
: After 0.05, it maintains a relatively low crime rate, but after 0.1 it starts go up again.



< ratio of change with average Time served by offenders in state prisons before their first release >

: In parts 20~30 months, have a high median crime rate and low values elsewhere.
: After 40 months, the crime rate increases again.

4. Conclusion

From the above results, we can draw the following conclusions.
In the beginning, I expected that these following items would affect the crime rate a lot: mean education time, expenditure on police protection, wealth, family income, income inequality, probability of imprisonment, average time served by offenders in state prison. I cannot say that these conclusions are exact statistics, because the graph went up and down repeatedly in one graph, instead of going up or down clearly. Therefore, if we just look at the reformulation of the graph and draw the conclusion, the result was totally different from what I expected except for the average time served by offenders in state prison.
As we can see from the graphs, the crime rate increases, when…

- there are less 14-24 aged men in the state.
- mean years of education aged 25 years or over increases.
- per capita expenditure on police protection is high.
- the number of males per 100 females is 96 ~ 100 or more than 105.
- the percentage of nonwhites is less.

- 14-24 aged people have less unemployment rate and 35-39 aged people have around 4% unemployment rate.
- people are wealthier.
- the percentage of families earning below half the median income is less.
- the probability of imprisonment has less value.
- the average time served by offenders in state prison is longer.