**ECON GROUP PROJECT**

**I. EXPLORATORY DATA ANALYSIS**

We are working on a balanced panel dataset on 50 US states, plus the District of Columbia over a 23-year period of time. The total number of observations is 1173 and there's no null values in the dataset.

Before building our hypotheses and models, we will do descriptive analysis and some exploratory analysis to have an overview, as well as better insights about our data such as relationship between variables, trend, ... Besides 3 dependent variables (vio, rob, mur), we will also focus on 2 important explanatory variables - shall and incarc\_rate to answer our main the question "Do more guns reduce crime?"

**1. Descriptive statistics:**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Variable** | ***vio*** | ***mur*** | ***rob*** | ***incarc\_rate*** | ***shall*** |
| Mean | 503.07 | 7.67 | 161.82 | 226.58 | 0.24 |
| Standard Error | 9.76 | 0.22 | 4.98 | 5.22 | 0.01 |
| Median | 443.00 | 6.40 | 124.10 | 187.00 | - |
| Mode | 256.80 | 3.60 | 111.60 | 98.00 | - |
| Standard Deviation | 334.28 | 7.52 | 170.51 | 178.89 | 0.43 |
| Sample Variance | 111,741.24 | 56.59 | 29,073.65 | 32,000.95 | 0.18 |
| Skewness | 2.54 | 5.79 | 3.89 | 3.89 | 1.20 |
| Range | 2,874.80 | 80.40 | 1,628.70 | 1,894.00 | 1.00 |
| Minimum | 47.00 | 0.20 | 6.40 | 19.00 | - |
| Maximum | 2,921.80 | 80.60 | 1,635.10 | 1,913.00 | 1.00 |
| Sum | 590,106.60 | 8,991.20 | 189,815.10 | 265,778.00 | 285.00 |
| Count | 1,173.00 | 1,173.00 | 1,173.00 | 1,173.00 | 1,173.00 |
| Confidence Level (95.0%) | 19.15 | 0.43 | 9.77 | 10.25 | 0.02 |

For these 5 main variables (measured per 100,000 people), on average over 23 years and 51 states, violence rate is much higher compared to robbery and murder. However, the standard deviation of violence is smaller than standard deviation of murder and robbery compared to theirs mean. Violence is also less skewed than murder and robbery.

Although incarceration rate is also high, it's still much lower compared to total crime rate (violence and robbery and murder).

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Variable** | ***pb1064*** | ***pw1064*** | ***pm1029*** | ***pop*** | ***avginc*** | ***density*** |
| Mean | 5.34 | 62.95 | 16.08 | 4.82 | 13.72 | 0.35 |
| Standard Error | 0.14 | 0.29 | 0.05 | 0.15 | 0.07 | 0.04 |
| Median | 4.03 | 65.06 | 15.90 | 3.27 | 13.40 | 0.08 |
| Mode | #N/A | #N/A | #N/A | #N/A | 11.66 | #N/A |
| Standard Deviation | 4.89 | 9.76 | 1.73 | 5.25 | 2.55 | 1.36 |
| Sample Variance | 23.87 | 95.29 | 3.00 | 27.58 | 6.53 | 1.84 |
| Skewness | 2.35 | (2.23) | 0.27 | 2.43 | 0.74 | 6.70 |
| Range | 26.73 | 54.75 | 10.14 | 32.74 | 15.09 | 11.10 |
| Minimum | 0.25 | 21.78 | 12.21 | 0.40 | 8.55 | 0.00 |
| Maximum | 26.98 | 76.53 | 22.35 | 33.15 | 23.65 | 11.10 |
| Sum | 6,259.38 | 73,834.99 | 18,863.16 | 5,649.57 | 16,099.19 | 412.94 |
| Count | 1,173.00 | 1,173.00 | 1,173.00 | 1,173.00 | 1,173.00 | 1,173.00 |
| Confidence Level (95.0%) | 0.28 | 0.56 | 0.10 | 0.30 | 0.15 | 0.08 |

**2. Correlation between variables:**

Below is the heatmap for correlation indexes between our variables. With blue is strong positive correlation, red is strong negative correlation.   
 **A picture containing shoji, crossword puzzle

Description automatically generated**

One of the important points we should notice is violence, murder and robbery are highly correlated (positive correlation), especially violence and robbery. Incarceration rate is highly correlated with violence and murder, but not that high with robbery. However, the shall-carry law is not correlated with crime rate, according to this overall analysis. This may not be accurate as the number of shall-carry law record is much lower than no shall-carry law in this dataset. We also use R to do pair plot for correlation of each variable - mur, vio, rob with the rest variables and they seem to have the same pattern. We will examine this further in our later parts.

There's also a positive correlation between density and crime rate, which makes sense to us. The percentage of white in the population is strongly negatively correlated with the percentage of black in the population. And, as time gone by, the percentage of young male in the population decrease.

**3. Variables studying in detail:**

**3.1. Dependent variable - Crime rate:**

Below is the violent, robbery, murder crime rate respectively by year and by shall-carry law. These 3 variables have the same pattern in each case (shall-carry law and no shall-carry law). Additionally, shall-carry law had utility in reducing the violent crime up till 1989. (Since last decade the effect of Shall Carry Law is more or less similar to its counterpart.)

**A close up of a map

Description automatically generated**

We have also known from the previous part that violence, murder and robbery are highly correlated. We will consider all of them as one, which is crime rate, by summing these 3 variables. From 1977 to 1999, the average crime rate has an upward trend, with shall-carry law has lower crime rate than none shall-carry law.

A close up of a map

Description automatically generated

We can see from the below graph, the states with highest crime rate (regardless of applying shall-carry law or not) are District of Columbia, Florida, California, Illinois, and Maryland.

A close up of a map

Description automatically generated

*We see a strong positive correlation between the variables- mur,vio,rob which is justifiable since murder, violence and robbery rate are interrelated.*

*A screenshot of a cell phone

Description automatically generatedA screenshot of a cell phone

Description automatically generatedA screenshot of a cell phone

Description automatically generated*

*Hence we create another variable* ***all\_crime*** *that sums up all 3 variables. There is a moderate to strong correlation between incarceration rate and these 3 variables individually. So we plot all\_crime against incarc\_rate to check whether our new variable captures this effect or not.*

*A screenshot of a cell phone

Description automatically generated*

***Histogram plots of dependent and explanatory variables.***

*Below we see the distribution of variables like mur,vio,rob,density,incarc\_rate,all\_crime,pm1029,avginc through the histogram plots.*

A close up of a map

Description automatically generated

*All the above variables have a positively skewed histogram. Thus we use the ln() in order to take logarithmic values of these variables as these values will assist in dealing with issues like heteroscedasticity or non normal residuals if encountered further thus making our model estimates more efficient and improving the precision of the model.*

*A close up of a map

Description automatically generated*

*The percentage of males in the age group 10-29 and the average seem to have a fairly symmetrical distribution i.e no skewness observed and hence don’t need any transformation.*

*Now in order to study the effect of the implementation of the Shall-Carry Law, we perform the following explanatory analysis.*

*A screenshot of a video game

Description automatically generated*

*In the above boxplot, we can observe that the overall crime rate is less in states where the Shall-Carry Law is implemented as compared to the states without the Shall-Carry Law. Thus we can say here that the implementation of the Shall-Carry Law has a considerable effect on the overall crime rates.*

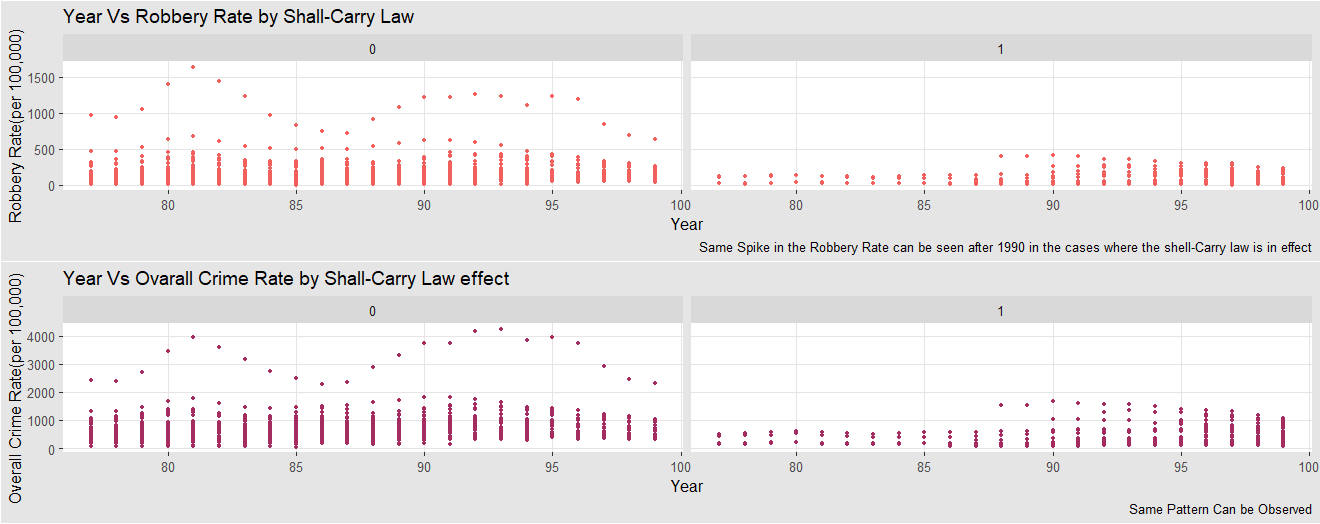
*Since the data is captured over a period, we can further investigate if the effect of Shall-Carry Law implementation varies with time.*

*A screenshot of a cell phone

Description automatically generated*

*We observe that over the period of 1977-1999, the violent crime rate is less throughout in states with Shall-Carry Law in comparison with the states without the Shall-Carry Law. However we can note a trend in the states with Shall-Carry Law that the violent crime rate has increased in the last decade (after 1989) and thus the law seems less effective in later years.*

*A similar trend is observed with murder rate from 1989 to 1999.*

**

*There exist a similar trend with Robbery rate over the period of time. Since all these variables are highly correlated, we can observe a similar trend in the overall crime rate.*

*Thus we can deduce that with time, the effect of Shall-Carry Law has not been effective in reducing the Crime Rates.*

*A screenshot of a cell phone

Description automatically generated*

*We see the trend between incarceration rate and over all crime rate with respect to Shall-Carry Law. There is a difference in the slope where the law is implemented and where its not. The states where the Law is implemented, the overall crime rate is less and the incarceration rate*

*Our data segregates the population by the percentage of blacks and whites in a region. Their effect on the Crime rates and the effect of Shall Law on this segregation can be viewed below.*

*A screenshot of a social media post

Description automatically generated*

*There exist a trend between the percentage of blacks in a region and the overall crime rates. There is an increase in the overall crime along with an increase in the percentage of Blacks lying in the age group of 10-64. Thus there is a moderate-strong positive correlation between these two groups.*

*We can further see a bifurcation based on the Shall-Carry Law implementation. However there is no evidence that suffices to say that there exist a definite trend or pattern.*

*A screenshot of a social media post

Description automatically generated*

*There is descending trend observed between the percentage of whites and the overall crime rate. The overall crime rate is on a decline as the percentage of whites in the state reduces. There is a negative correlation between these two groups.*

*As observed with Percentage of Blacks, there is no significant effect of implementation of Shall-Carry Law.*

*Intuitively, we can also justify the difference in the trends as the number of blacks in a region will be less where the population of whites is more and vice versa.*

*The data provides a bifurcation of percentage of young males(age group 10-29) of the entire population.*

*A screenshot of a social media post

Description automatically generated*

*The overall crime rate doesn’t fluctuate much with an increase in concentration of Young males in a state. The overall crime is almost constant for all Percentages of males (age 10-29). Thus, there is no correlation observed between the two groups.*

*The implementation of Shall-Carry law doesn’t provide any significant effect on the relation between the overall crime rate and the Young Male population. We can see that the implementation of Shall-Carry law has led to a decrease in the overall crime rate for states with high Percentage of young male population. Also we can see no implementation of Shall-Carry law where the Percentage of Male population. However this effect is not remarkable.*

*We can study the relation of population and density with the overall crime rate.*

*A close up of a map

Description automatically generated*

*We see positive trend between Population and the Overall Crime rates. Similar trend can be seen between Density and the Overall Crime rate. Intuitively highly populated regions tend to have high crime rate as denser the region, higher scope for robbery, murders or other crimes to occur. Thus there is a strong positive correlation between these groups.*

*We can see that the implementation of Shall-Carry Law doesn’t seem to have any significant effect. There is no remarkable trend with respect to low or highly dense states.*

*A screenshot of a social media post

Description automatically generated*

*From the scatterplot above, we can observe that there is not strong linear relation between the Average Income of the population and the over all income. Thus we cannot say there is no correlation between these groups.*

*Also the implementation of the Shall-Carry Law doesn’t provide any significant trend in the relation between the Overall crime rate and the average income of the state.*

*Further we study the effect of various variables on the incarceration rate.*

*A picture containing photo, sky, bird, wall

Description automatically generated*

*We can observe that there is an increase in the incarceration rate from 1977 to 1999 irrespective of the implementation of Shall Carry Law. However in states where the Law had been implemented, we can see that there has been a spike in the incarceration rate in the last decade i.e after 1990.*

*Similar effect was observed previously with overall crimes. There was a spike in the overall crime rate in the last decade. Intuitively, we can say that since there was an increase in the crime rate post 1990, with the implementation of the Shall-Carry Law, there is an increase in the incarceration rate as well.*

*A screenshot of a social media post

Description automatically generated*

*As discussed above, we can see the trend in states where the population of blacks is higher, the population of whites is lower and vice versa. There is an inverse relationship between the Population of Blacks and Whites of a state.*

*From the plot above we can see that states with higher percentage of white have lower incarceration rates and states having comparatively more percentage of Blacks have higher incarceration rate.*

*A screenshot of a computer

Description automatically generated*

*We can observe from these plots that states with low population blacks which also implies higher population of whites, the overall crime rate is lower which also explains the low incarceration rate.*

*States with high Population of Blacks i.e low Population of Whites have higher overall crime rate along with higher incarceration rates.*

*A screenshot of a map

Description automatically generated*

*As observed earlier, there is no strong correlation between the Percentage of young males and the overall crime rates. However we can see that where there is low Percentage of Young Males with Higher Overall crime rates, the incarceration rate is also higher as compared to other scenarios.*

*A screenshot of a computer

Description automatically generated*

*We know that population and density are correlated with the overall crime rate. The above plot shows us the variation of incarceration rate with these groups. No remarkable trend is observed however we can say that states with higher density of population, having higher Overall Crime rates tend to have higher Incarceration rates.*

*A picture containing photo, indoor

Description automatically generatedFrom the above plot we can infer that there is a decline in the Percentage of the Young Male Population over the years from 1977-1999. Time has an effect on the overall crime rates and the incarceration rates. They increase over the time and substantially increased in the last decade. Although descending, but we can see a similar effect or trend in the above plot.*

*A picture containing screenshot

Description automatically generated*

*The Average per Capita Income have been seen to rise over the time period from 1977-1999. Thus a positive trend is observed.*

*A picture containing photo, showing

Description automatically generated*

*There are no significant inferences or trends that can be drawn from the plot above. However we see that there is a slight decline in the incarceration rate as the Population of Young Males increases.*

*A screenshot of a cell phone

Description automatically generated*

*The Average per Capita Income is observed to be on a decline with increase in the Population of Young Males in the state. This can be justified intuitively. A person tends to have a lower income as compared to in the later years. Thus if the Population of Young Males in the state is higher, the average per capita income is lower. We observe a descending or negative trend here.*

*A close up of a map

Description automatically generated*

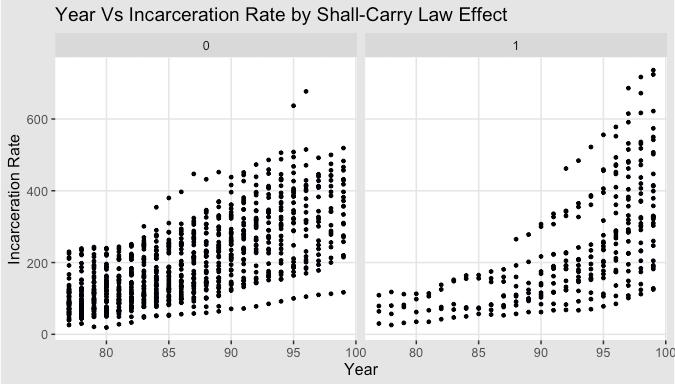
**3.2. Incarceration rate:**

By analyzing incarceration rate and overall crime rate, we can see that there is a positive correlation between crime rate and incarceration rate. It seems that states with shall law in place have higher incarceration rates.

**A screenshot of a map

Description automatically generated**

Same pattern of the incarceration rate spiking up in after 1990 where the shall-carry law is in effect can be seen.



**3.3. Shall-issues law:**

Regarding shall-issues law, we can see that over 23 years, the total number of times these states adopted shall-carry law is much less than none shall-carry law (285 compared to 1,173). Also, there are some states that always applied shall-carry law throughout 23 years (Washington, New Hampshire, Vermont, ...), some states have never applied shall-carry law and some states (New Mexico, Colorado, Kansas, ...), and some states switch shall-carry law on and off. We will group these states and analyze in the next parts.

A close up of a map

Description automatically generated A close up of a map

Description automatically generated

We also want to see if shall law implemented, will there be more incarceration? As we can see from the boxplot, states having shall-carry law in effect tend to have higher incarceration rates and less overall crime rates.

A close up of a map

Description automatically generated

**3.4. Density and Population:**

Does higher density mean more crime? There is a positive correlation between overall crime and population as well as between density and overall crime rate. Also, densely populated states do not have shall law in effect.

A close up of a map

Description automatically generated

How about incarceration rate? We can see that states having higher population have more overall crime and hence have higher incarceration rates. Also, cities having high densityalso have more crime, hence higher incarceration rates.

A close up of a map

Description automatically generated

**3.5. Average Income (avginc):**

People who live in states with higher density seems to have higher income. It is also obvious that average per capita income increases over year.

A close up of a map

Description automatically generatedA screenshot of a computer

Description automatically generated