**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 patter. 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

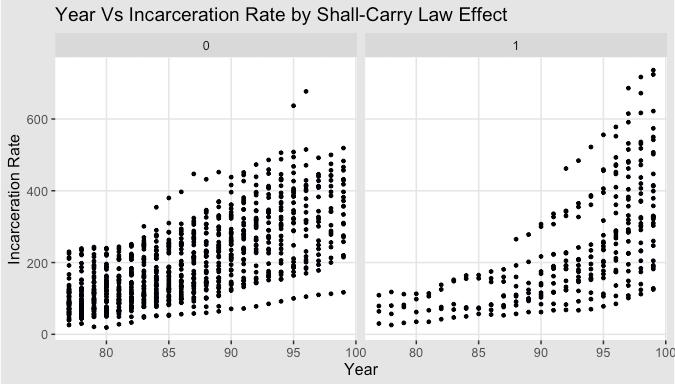
**3.2. Incarceration rate:**

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

**4. Table of expectation:**

From all the exploratory analysis above. We have the table of expectation for the variables as below.

|  |  |  |
| --- | --- | --- |
| **Variable** | ***Expected sign*** | ***Explain*** |
| *shall* | (-) | States having shall-carry law in effect tend to have less overall crime rates. |
| *incarc\_rate* | (+) | States with higher crime rate also have higher incarceration rate |
| *density* | (+) | Higher population density increases overall crime rate |
| *avginc* | (-) | Higher average income reduces overall crime rate |
| *pop* | (+) | More people increase overall crime rate |
| *pm1029* |  |  |
| *pw1064* |  |  |
| *pb1064* |  |  |
| *stateid* |  |  |
| *year* |  |  |