Casey Campbell

Dr. Thomas Craig

MG-212: Data Analysis

9 December 2022

Final Project

Violent Crime Analysis

Section 1: Introduction/Overview

Violent Crime occurs all over the United States multiple times a day and millions of Americans fall victim to them every year. Violent crimes include: Murder/Manslaughter, Rape, Robbery, Aggravated Assault, Property Crime, Burglary, Larceny-Theft, Motor Vehicle Theft, and Arson. To get a better understanding, murder is when the defendant kill another person with full intent to do so, while manslaughter is the killing of another person with the absence of intent. For example, point blank gunshot to the head would be murder while if a person was driving and accidentally hit someone walking across the street and killed them it would be manslaughter. Rape is the unlawful penetration of another without consent. Aggravated Assault is when the defendant caused or intended to cause serious bodily injury to the victim, and had an indifference to the value of human life. Property crime is when the victim has their belongings stolen or damaged when threat and force is not involved. Burglary is simply the illegal entry into a building with full intent to commit a crime, which is usually theft. Larceny-Theft is the act of taking, carrying, etc. of property from another person's possession without the threat or use of force. While Robbery is just larceny-theft with the use or threat of force. Motor Vehicle Theft is

fairly self explanatory, it's the attempt or taking of a motor vehicle. Finally, Arson is the attempt or act of setting on fire of a structure. Violent Crimes are the most common serious crimes in our country and if data can help us understand where certain crimes happen more and what crimes are correlated.

Section 2: Statement of Hypotheses

When examining this data of violent crime supplied by the FBI's Uniform Crime Report, I have gathered a few assumptions and questions about the data. I believe there will be a correlation between robbery and larceny-theft, aggravated assault and murder/manslaughter, property crime and burglary, burglary and arson, then larceny-theft and motor-vehicle theft. Furthermore, I am curious to see which variable is a better predictor of violent crime, murder/manslaughter or aggravated assault. Assuming the answer is not murder/manslaughter, from analysis I am looking to see what is the second best predictor of violent crime in Pennsylvania.

Hypotheses:

Null Hypotheses (H_0): There is a positive correlation between larceny-theft and robbery.

Alternate Hypotheses (H_1): There is not a positive correlation between larceny-theft and robbery.

Null Hypotheses (H_0)*:* There is a positive correlation between aggravated assault and murder/manslaughter.

Alternate Hypotheses (H_1): There is not a positive correlation between aggravated assault and murder/manslaughter.

Null Hypotheses (H_0): There is a positive correlation between property crime and burglary

Alternate Hypotheses (H_1): There is not a positive correlation between property crime and burglary.

Null Hypotheses (H_0): There is a positive correlation between burglary and arson.

Alternate Hypotheses (H_1): There is not a positive correlation between burglary and arson.

Null Hypotheses (H_0): There is a positive correlation between larceny-theft and motor-vehicle theft

Alternate Hypotheses (H_1)*:* There is not a positive correlation between larceny-theft and motor-vehicle theft.

Null Hypotheses (H_0)*:* Murder/Manslaughter is better predictor of violent crime than aggravated assault.

Alternate Hypotheses (H_1): Aggravated Assault is better predictor of violent crime than murder/manslaughter.

Null Hypotheses (H_0): Murder/Manslaughter is second best predictor of violent crime.

Alternate Hypotheses (H_1): Murder/Manslaughter is not the second best predictor of violent crime.

Section 3: Data and Analysis Methodology

This data was gathered by the Federal Bureau of Investigation (FBI) through their yearly Uniform Crime Report (UCR) Program. With this program the FBI institutes the Hierarchy Rule which takes the most serious offense within a multiple offense criminal incident. Meaning, if a person was found guilty of murder and larceny the UCR only tallies the murder from that case and larceny theft isn't counted. Moreover, cities are not required to report their data for each year. Although most cities do indeed report, some don't for various reasons being, they didn't have time, had a bad criminal year and don't want to scare the public, or they simply do not want to report because it is more work on the city's behalf.

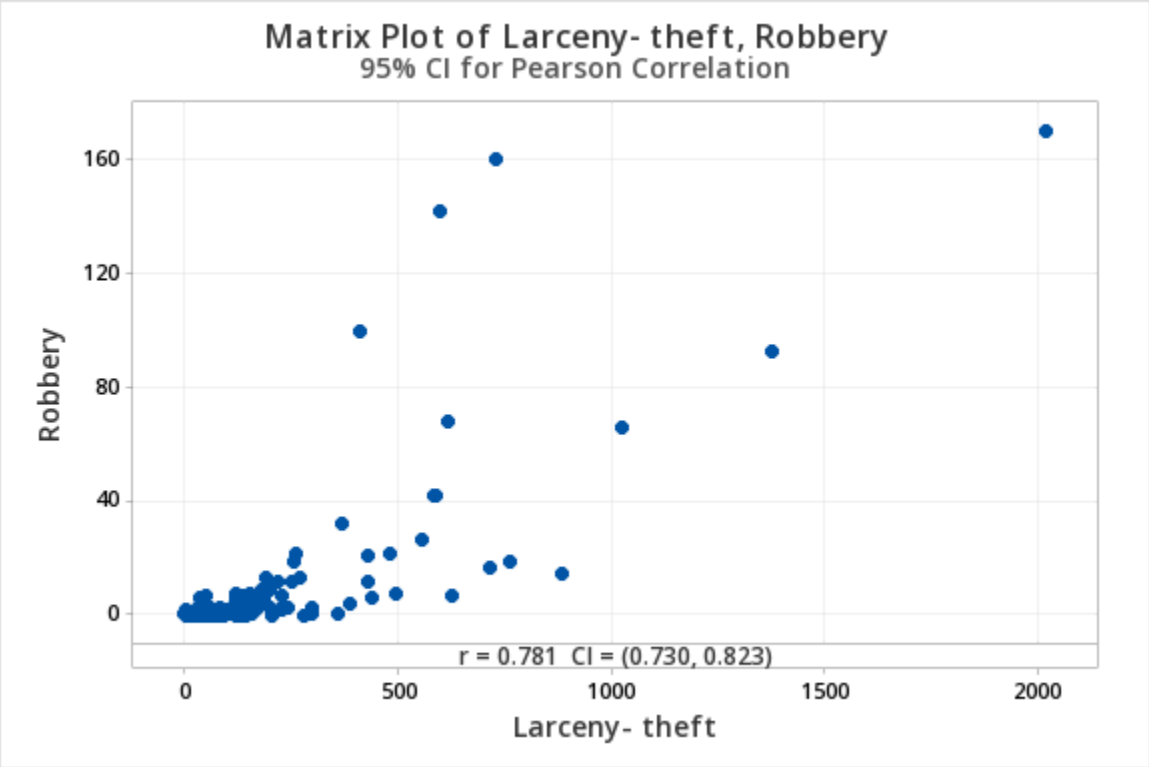
To test my hypotheses of the correlation between two variables I first found the correlation coefficient, also known as the r-value. This analysis gives a number from 1.0 to -1.0 which tells us how strong the correlation is, and whether it is positive or negative. An r-value of 0.0 means that there is no correlation between the variables. If the r-value is closest to .25 it means that the correlation is weak. A moderate correlation is when the r-value is nearest to .5. Two variables are strongly correlated if the r-value's nearest neighbor is .75. Lastly, a perfect correlation is when the r-value is 1.0, however this is very rare. The correlation is determined negative or positive by the r-value, if the r-value is negative then so is the correlation, and if it's positive then the correlation is positive. Furthermore, the scale for the strength of the correlation is the same on both sides of the zero. Finally, when it comes to analyzing the predictability of violent crime I will be using a regression analysis which will give me an R-squared (R-sq) value which is measured by a percentage ranging from 0%-100%. R-squared will tell us how well that equation and variable explains and predicts violent crime.

Section 4: Results

A. Is there a correlation between Larceny and Robbery?

Null Hypotheses (H_0): There is a positive correlation between larceny-theft and robbery.

Alternate Hypotheses (H_1): There is not a positive correlation between larceny-theft and robbery.

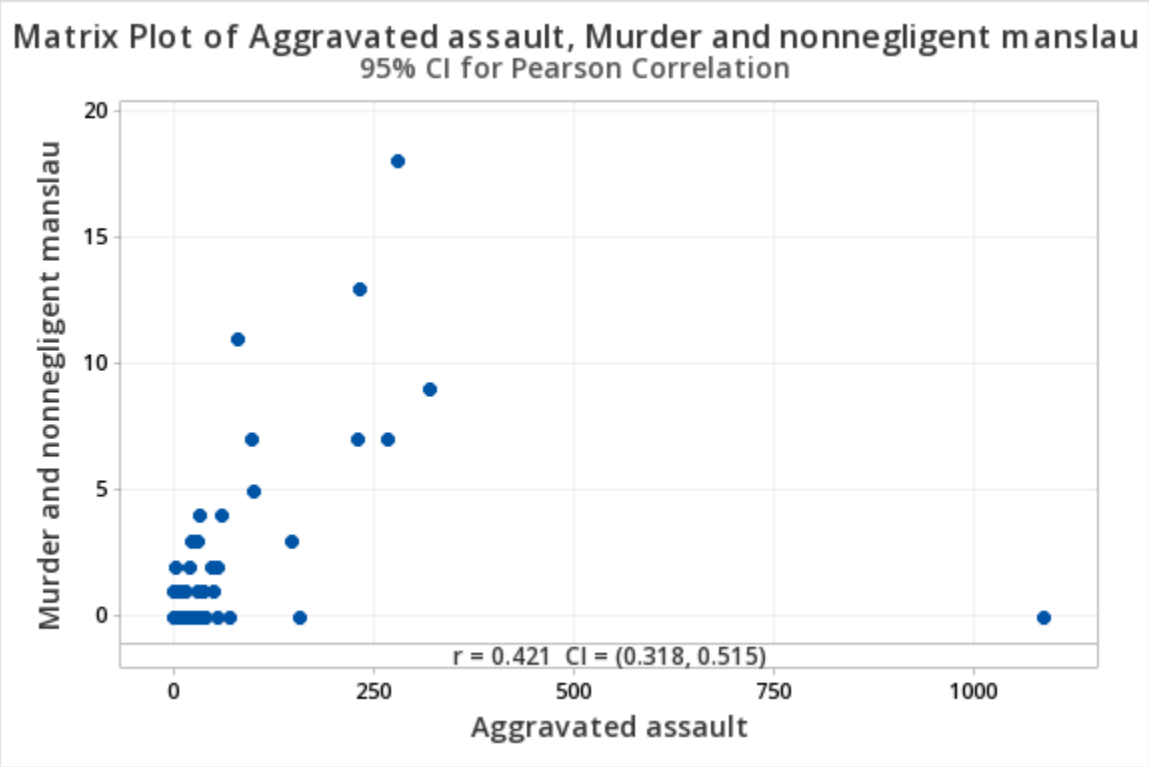


From this information we can accept the Null Hypotheses. We see that the r-value equals .78 which means that we are comfortable with saying there is strong positive correlation between larceny-theft and robbery.

B. Is there a correlation between Aggravated Assault and Murder/Manslaughter?

Null Hypotheses (H_0): There is a positive correlation between aggravated assault and murder/manslaughter.

Alternate Hypotheses (H_1): There is not a positive correlation between aggravated assault and murder/manslaughter.

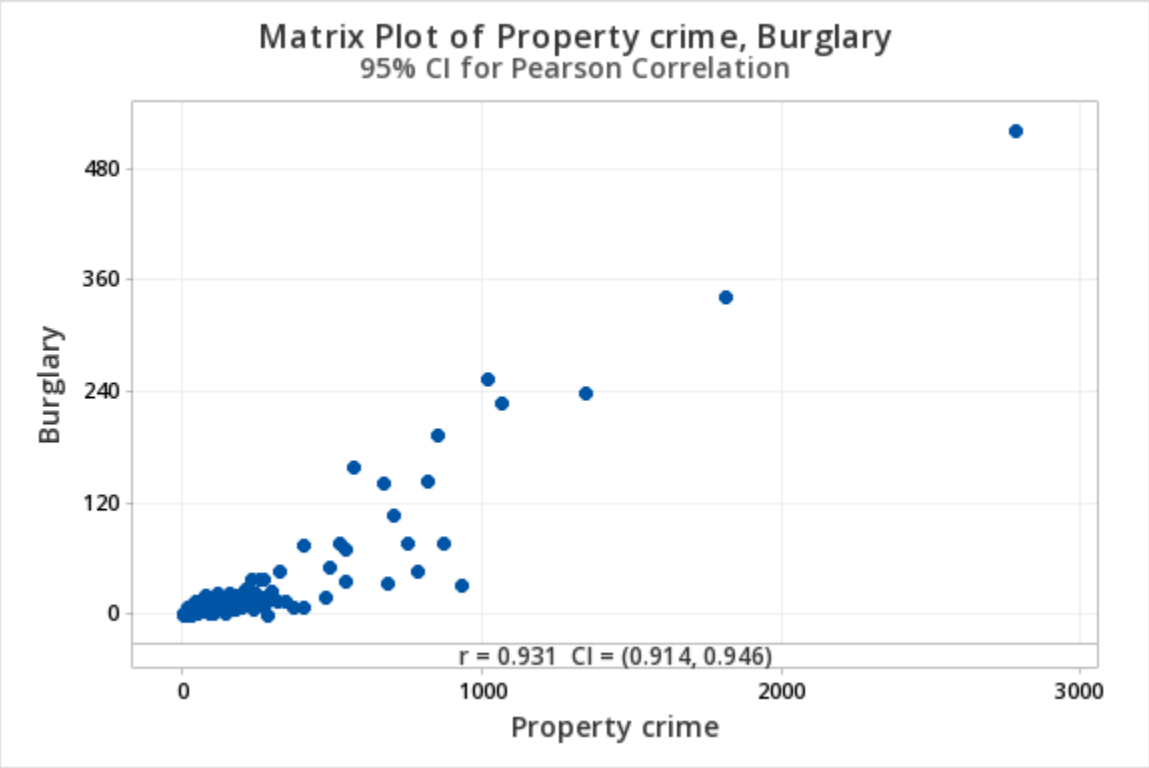


This correlation graph shows us the association between the two variables aggravated assault and murder/manslaughter. The correlation however is positive and moderate since it sits at .421 which is neighbors with .5. With this analysis we can confirm and accept our null hypotheses that there is positive correlation between aggravated assault and murder/manslaughter.

C. Is there a correlation between Property Crime and Burglary?

Null Hypotheses (H_0): There is a positive correlation between property crime and burglary

Alternate Hypotheses (H_1): There is not a positive correlation between property crime and burglary.

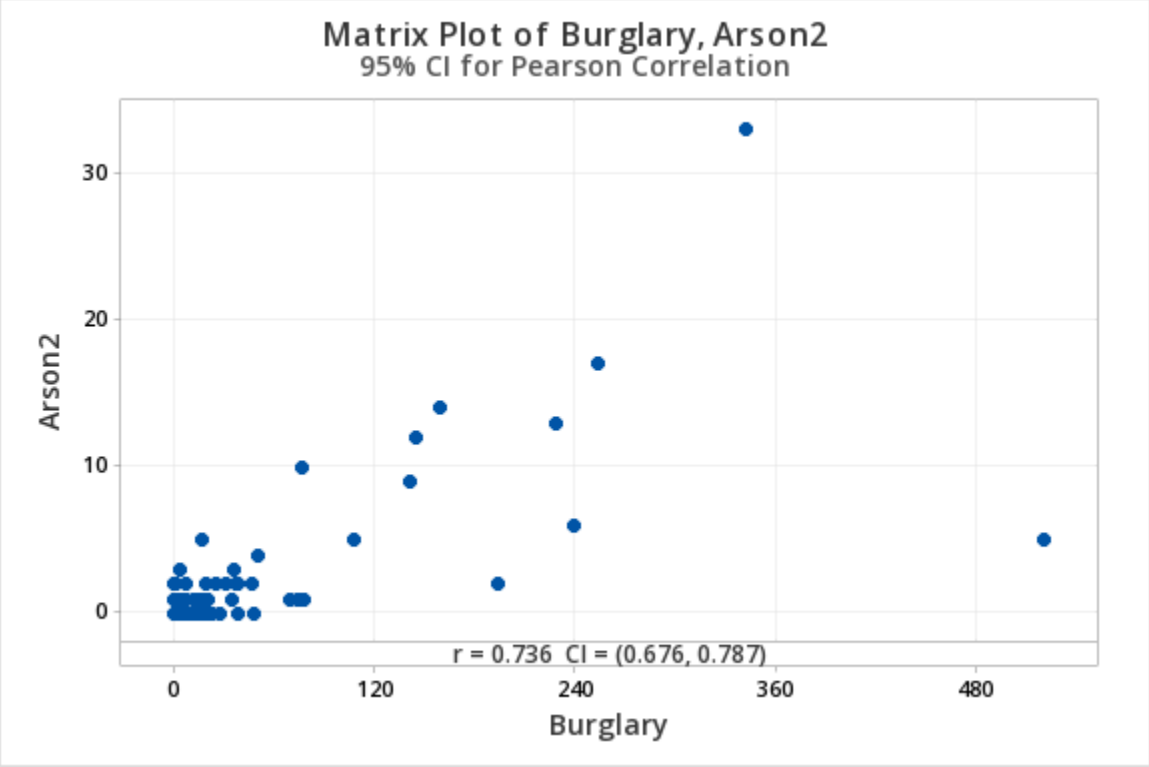


Analyzing these two variables we can clearly see that there is a very strong and nearly perfect correlation between property crime and burglary with the r-value being .931. From this graph and correlation analysis we can confirm and accept the null hypothesis and conclude that there is positive correlation between property crime and burglary.

D. Is there a correlation between Burglary and Arson?

Null Hypotheses (H_0): There is a positive correlation between burglary and arson.

Alternate Hypotheses (H_1): There is not a positive correlation between burglary and arson.

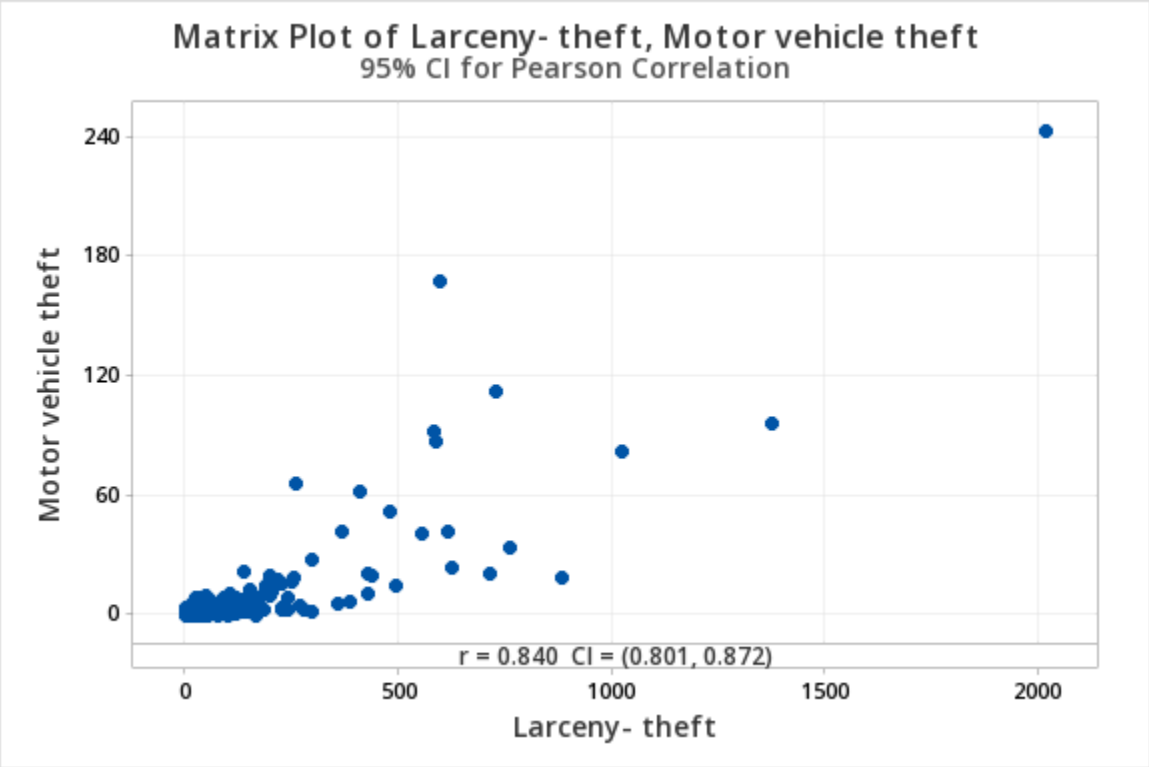


Here we are looking at the correlation between burglary and arson. It can be observed that both of these crimes record very low counts except for a few outliers. However, the correlation between these two variables is strong and positive resting at .736. From this we can conclude that the null hypothesis is correct and that there is a positive correlation between burglary and arson.

E. Is there a correlation between Larceny and Motor Vehicle Theft?

Null Hypotheses (H_0): There is a positive correlation between larceny-theft and motor vehicle theft

Alternate Hypotheses (H_1)*:* There is not a positive correlation between larceny-theft and motor-vehicle theft.

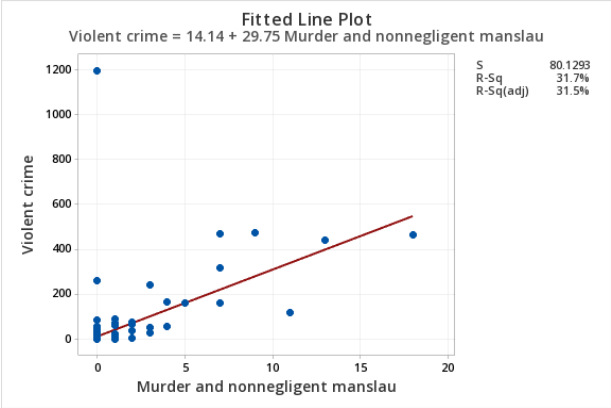


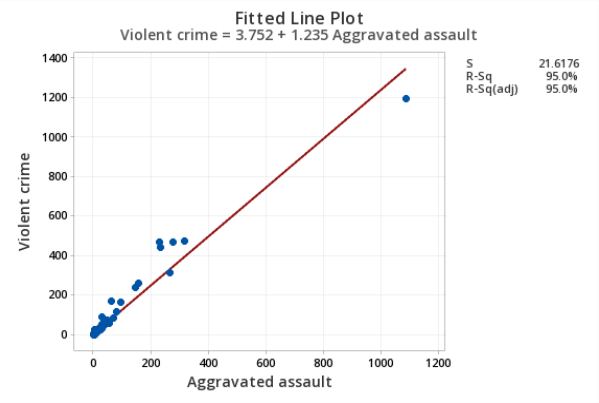
This correlation analysis is between larceny-theft and motor vehicle theft. This graph shows us that the r-value is .84 and consequently shows a strong and positive association between the two variables. Thus, we can accept the null hypothesis and confidently say there is a correlation between larceny-theft and motor vehicle theft.

F. Which variable is the best predictor of violent crime in Pennsylvania?

Null Hypotheses (H_0): Murder/Manslaughter is better predictor of violent crime than aggravated assault.

Alternate Hypotheses (H_1): Aggravated Assault is better predictor of violent crime than murder/manslaughter.



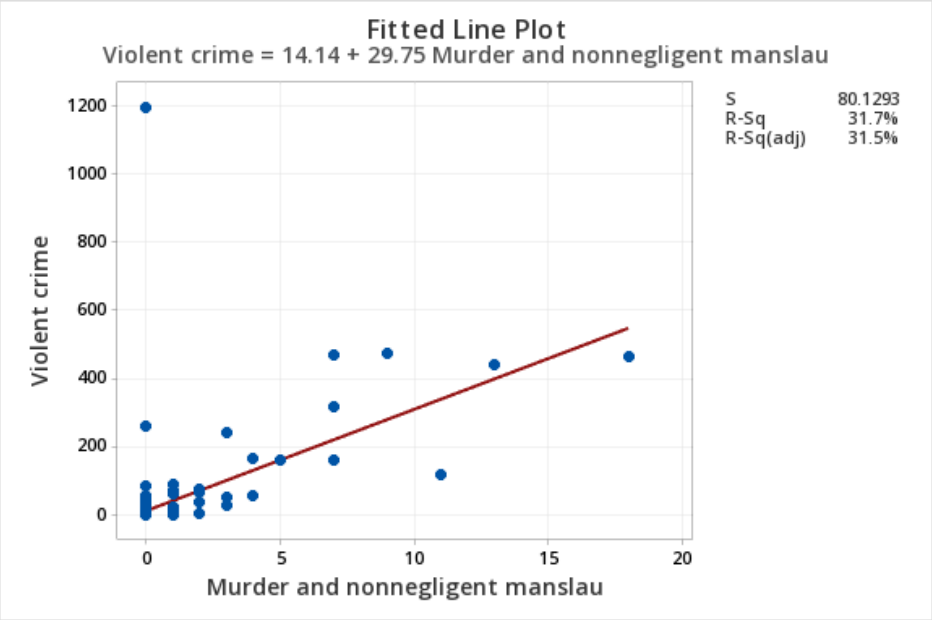


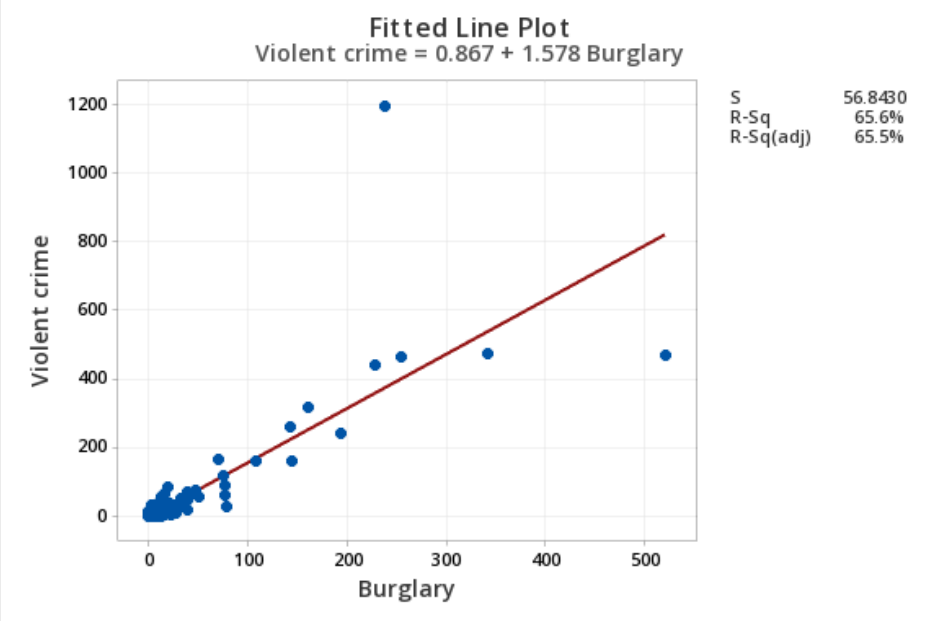
Regression analysis shows us that aggravated assault is clearly a better predictor of violent crime than murder/manslaughter. The R-Sq value shows us the reliability of the regression analysis of that variable and equation, so the higher the percentage the more reliable the variable. Murder/manslaughters R-Sq. value rests at 31.7% while aggravated assaults R-Sq. value sits at a comfortable 95%. Proving without a doubt we can reject the null hypothesis and conclude that aggravated assault is better predictor of violent crime than murder/manslaughter.

G. Is Murder/Manslaughter the second best predictor of violent crime in Pennsylvania?

Null Hypotheses (H_0): Murder/Manslaughter is second best predictor of violent crime.

Alternate Hypotheses (H_1): Murder/Manslaughter is not the second best predictor of violent crime.





To test my hypothesis I chose to do another regression analysis which produced these two graphs. A result that I did not anticipate was burglary being the second best predictive variable of violent crime in Pennsylvania. As discussed in the previous analysis the murder/manslaughter R-Sq. value is only at 31.7% but burglary achieved an R-Sq. value of 65.6%. Due to these results we can reject the null hypothesis and proudly state that burglary is the second best predictor of violent crime.

Section 5: Discussion and Conclusion

Violent Crime occurs all over the nation however it is important to remember that this data is only from the cities in Pennsylvania that decided to report to the FBI's Uniform Crime Report in 2019. Therefore, this is a good portion of Pennsylvania but not the whole so we cannot be entirely certain that Pennsylvania's population resembles these exact same numbers. However, with the data we were able to access and gather these numbers are extremely eye opening. Going into this project I expected at most half of the correlation hypotheses' to be accepted, not all of them. The correlation tests performed indicate that all of the crimes that are related, in action and differentiated by use or threat of force, are most likely correlated. However we cannot say definitively that they are always correlated because we do not have the whole population's data. The data we possess is just a large sample that represents about 75% of the state. Furthermore, with our first regression analysis testing which variable is a better predictor of violent crime, it was shocking to see the separation in results between the two variables. Aggravated assault didn't just turn out to be the best predictor, it was the best by far and almost a perfect predictor of violent crime in Pennsylvania with an R-Sq of 95%. On the other hand, looking at murder/manslaughter's R-Sq of 31.7%, it was the worst predictor out of all the variables in this dataset. Burglary came out to be the second strongest predictor of violent crime in Pennsylvania at 65.6% which was shocking. Nonetheless, that is what this data and project is for, so that we can properly analyze the data and not rely on assumption, intuition, and history. Recent data is so crucial to help law enforcement better mitigate, understand, and prevent crime in the United States.

References

"Violent Crime." *FBI*, FBI, 13 Sept. 2019, https://ucr.fbi.gov/crime-in-the-u.s/2019/crime-in-the-u.s.-2019/topic-pages/violent-crime.