Discovering the driving factors behind the high crime rate in Baltimore

Industry/Civic/Academic Question

The high crime rate of Baltimore has long been a concerning issue for both the city government and local residents. So, to better ensure our personal safety, it is beneficial for us to understand the driving factors behind the crime rate and hopefully decrease the chance of being the victim. After analyzing the data, I found out that the median house hold income, high school dropouts and population density are all significant factors affecting crime rate in Baltimore. Here are some useful articles related to my topic.

[1] Inter-Provincial Spatial Linkages of Crime Pattern in Indonesia: Looking at Education and Economic Inequality Effects on Crime

https://jurnal.ugm.ac.id/ijg/article/view/34026

[2] Property Crimes and Violence in United States: An Analysis of the influence of Population density   http://www.sascv.org/ijcjs/harries.pdf

[3] Multivariate temporal modeling of crime with dynamic linear models

https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0218375

Data Questions

For my data source, I used both the data related to crime and demographic characteristics from Baltimore City Open Data and SimplyAnalytics (from year 2010 to 2018). I want to find what's the relationship between demographic features and the crime amount, and what's the influence factors for each feature. My initial hypothesis is that the number of crimes is roughly affected by 5 factors: the number of high
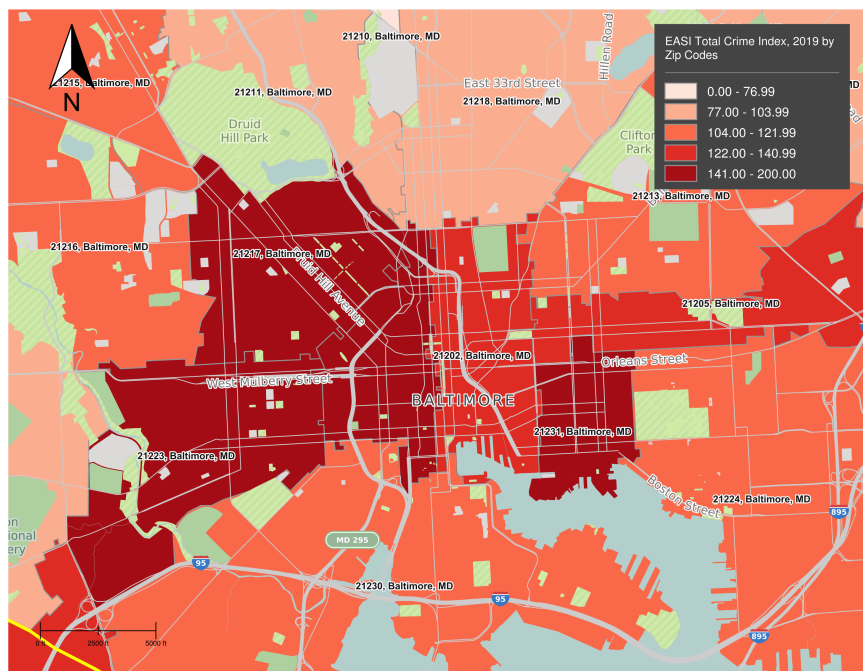
school dropout, median household income, total poverty, total population and the unemployment. To discover the correlation between independent variables and the outcome, multiple regression analysis is used. In my case, the crime amount is the outcome variable and the five hypnotized factors are predictors. And I will try to determine how each type of crime can be influenced by the predictors. From the p-value in coefficient table, I can determine the significant level of each factors and pick out the most influential one. Moreover, the regression equation will show the overall relationships, and how the crime amount will change with respect to the change in other independent variables. This will also allow me to roughly predict the rate amount in the future. Here is my data source:

[1] https://data.baltimorecity.gov/Crime/Crime-by-Neighborhood/6ayg-3z5z

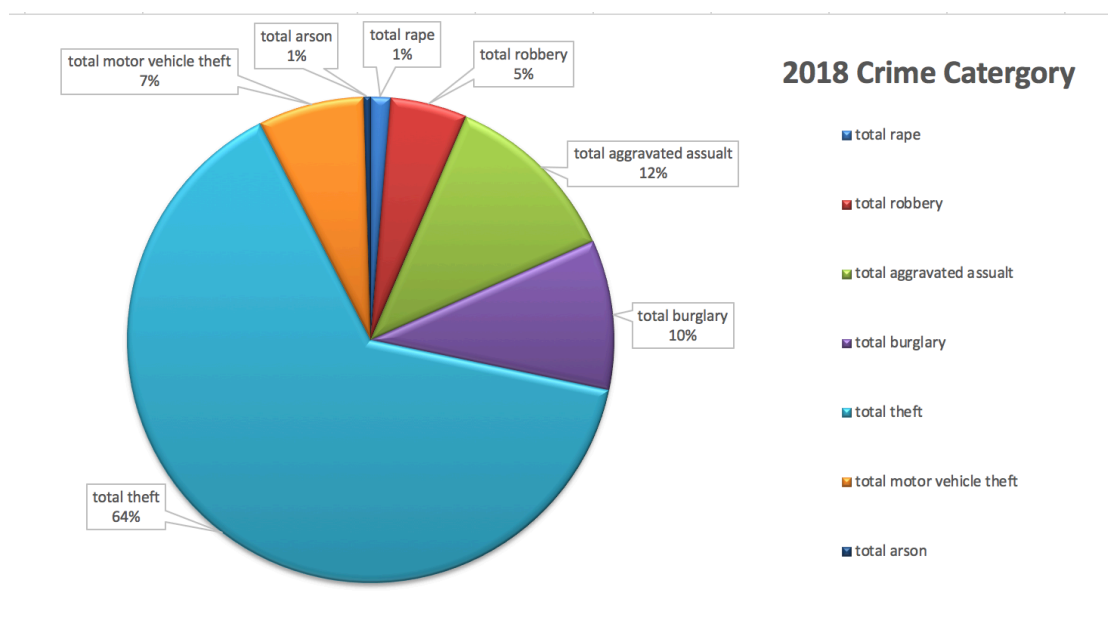[2] https://data.baltimorecity.gov/Crime/canton/f42y-mbfm

[3] https://simplyanalytics.com/

Data Answers

In order to explain crime occurrence in Baltimore, one need to first look at the problem in spatial context. The heat map was generated using SimplyAnalytics, which shows the crime occurrence in terms of location. The data was from the reported crimes to the FBI in year 2019. From the crime map, we can see that most of the region in Baltimore have a crime index number higher than the national average. Moreover, the regions with high crime index tend to aggregate in space, and neighborhoods next to each other tend to share similar level of crime index. These evidences might indicate that there is inter-relationship between geographical distribution and crime rate.

Next, lets dive deeper into the categories of crime. The data was from Baltimore police department, showing the overall crime types in 2018. The pie chart clearly shows the percentage of each category with respect to the overall crime amount. We can see that theft is the number one common crime, which takes up nearly 64% of the overall amount. Next followed by aggravated assault which occupy 12% and burglary 10%.

Regression Equation

Property Crime Count = −57142 + 0.14080 Education: High School Dropouts
+ 0.3229 Median Household Income + 0.09627 Population

From the regression equation, we can see that the median household income influenced the property crime most, by a factor of 0.32. and the positive sign suggests that they are positively correlated. This makes sense as region with high income inequality will lead to frustration for the people in low income category; therefore, triggering them to commit crime. Also, the high school dropout rates positively affect the crime rate. The result suggests that low education level will decrease people's chance of getting a decent job, therefore drive them toward committing crime. Moreover, the population density has a positive association with property crimes as well. This can be explained by the fact that higher population density will offer more opportunities and higher returns for crime.

Regression Equation

Violent Crime Count = 61536 + 0.08172 Education: High School Dropouts − 0.09484 Population

In the case of violent crime, the result shows that crime rate is in direct proportional to high school dropout rate. However, it is interesting to notice that the crime rate is negatively influenced by the population density, it is a little bit counter-intuitive. One of my guesses for this is that witnesses are more abundant, and events are more likely to be reported to police in densely populated areas. So, the high populated area offers natural surveillance that has the effect of inhibiting violent crimes.

Industry/Civic/Academic Answer(s)

In both models, the high school dropout rate and population are significant variables affecting crime rate in Baltimore. In the case of property crime, the median household income is the most significant factor, indicating that economic inequality may trigger crime occurrence. And the high population density offers more opportunities for property crimes, since the opportunity cost is relatively low. In the case of violent

crime, it is interesting to find out that the high density can have a suppression effect on the crime rate.

So, in order to reduce the crime rate in Baltimore, it is up to the government to shorten the economic inequality between the neighborhood areas and increase the level of schooling. And some suggestions for individuals, if one wants to avoid being attacked or robed, they may consider go to places where there are lots of people around. And if one wants to avoid property loss, they may avoid living in rich and highly-populated neighborhood.