

Crime vs Poverty Data Analytics

May 17, 2020

Law & Order Analytics: Pooja Nagrecha, Ryan Thomas, Ana Gill & Chris Krokus

Overview:

The purpose of this analysis is to carefully examine and evaluate the impact of poverty on crime. Washington D.C. was chosen as a model city for the study. The city has a prevalent amount of crime and poverty, being the city in the nation with the widest income inequality graph. The bottom fifth of D.C. households shows just two percent of the total D.C. income during the year 2017, while the top fifth shows 51.5 percent (Gregory, 2017). The crime rates during 2017 are 102% higher than the national average (DC Crime, 2018). The goal of analyzing these datasets was to see if there was any correlation between the amount of crime that occurred and the amount of poverty in a designated area during a specific period of time.

Data and Model:

The data was collected from two datasets that show the amount of crime and the level of poverty during the year 2017. For this data, Washington D.C. is broken up into 179 Census Tracts. A Census Tract is a geographic region defined for the purpose of taking a census. This was determined to be the unifying piece between the two datasets of poverty status and instances of crime. From the poverty dataset, we analyzed the total amount of poverty, the amount of child poverty, and percentages of poverty. From the crime dataset: offenses, number of crimes, longitude, latitude_____

Prior to merging the two datasets, we took some time to explore them individually. The poverty dataset had many columns breaking the Total Population up into 'above' or 'below' the poverty line and then into age segments to include < 6, 6-11, 12-17, 18-59, 60-74, 75-84, and 85+. Since we were under tight time constraints, we elected to simplify this down to Population Below the Poverty Line, Child Population Below the Poverty Line, and the Total Population; all of these were broken down into rows of the 179 Census Tracts.

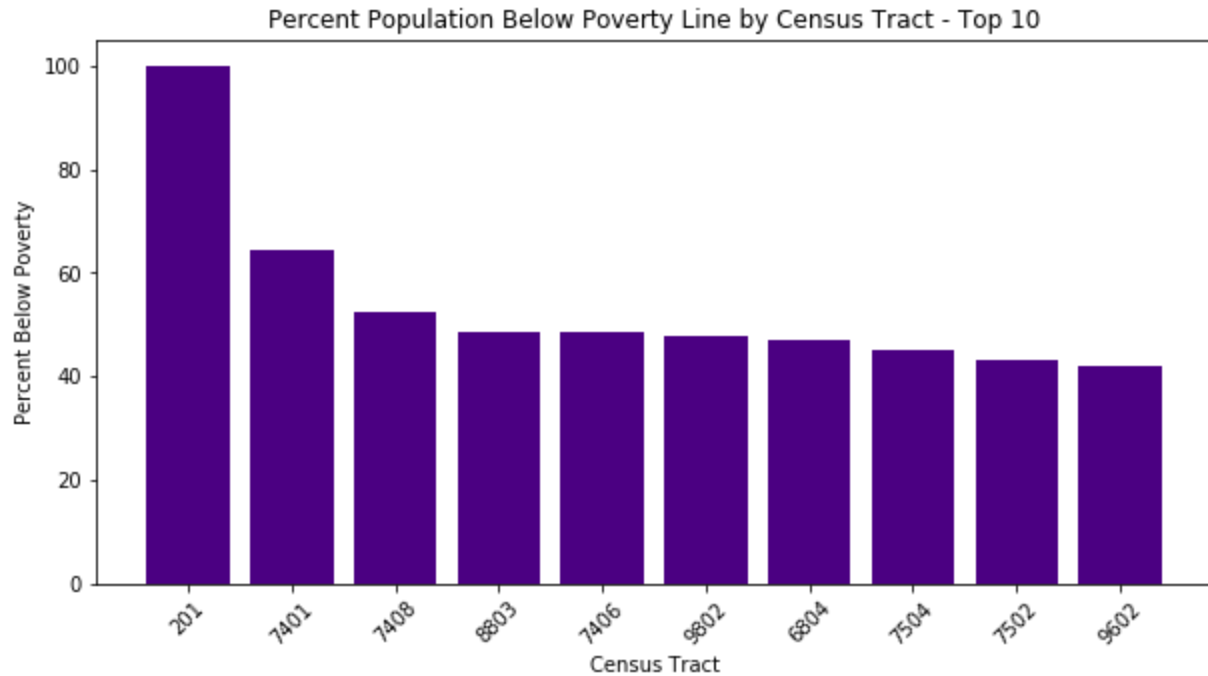


Figure 1.

Figure 1. depicts the top ten poorest census tracts in Washington DC. Of the 179 census tracts in our dataset, 176 of them had more than 1000 people. Census Tract 201, the left-most bar in Figure 1. is displaying a 100% poverty rate; well it turns out there are only 24 people accounted for in that particular census tract and all of them are below the poverty line. The remainder of the top ten range from about 45-65% of people below the poverty line.

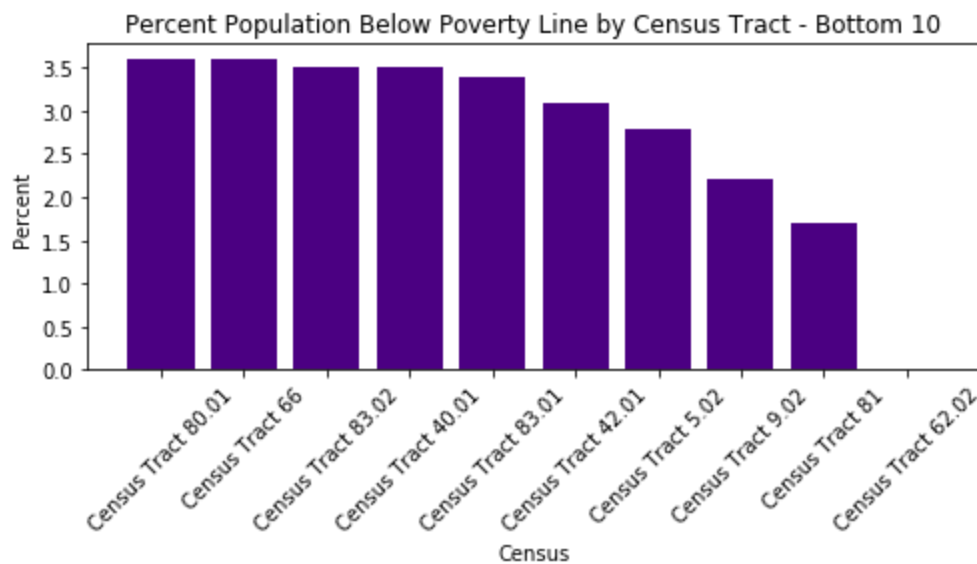


Figure 2.

At the opposite end of the spectrum from Figure 1., we have Figure 2. - the top ten wealthiest census tracts in Washington DC. There were zero residents in Census Tract 6202 below the poverty line and the remainder of the top 10 all fell at or below 3.5%.

Establishing whether poverty has an impact on crime

Although the link between poverty and crime is complex. There is a widespread impression showing crime is prevalent among poor social classes. In the regression analysis, the aim is to discover whether a strong link exists or not between income poverty and crime in Washington D.C.

The map below shows the top 5 most dangerous areas in D.C. during the year 2017. Crimes include homicide, sexual assault, robbery, and battery. In 2017 there are an estimated 33,082 offenses committed in Washington, D.C. The area known as Georgetown tops the list with 603 criminal offenses. Only Georgetown and Ivy City made up the list of the top 10 areas with the highest percentage of people below the poverty line.

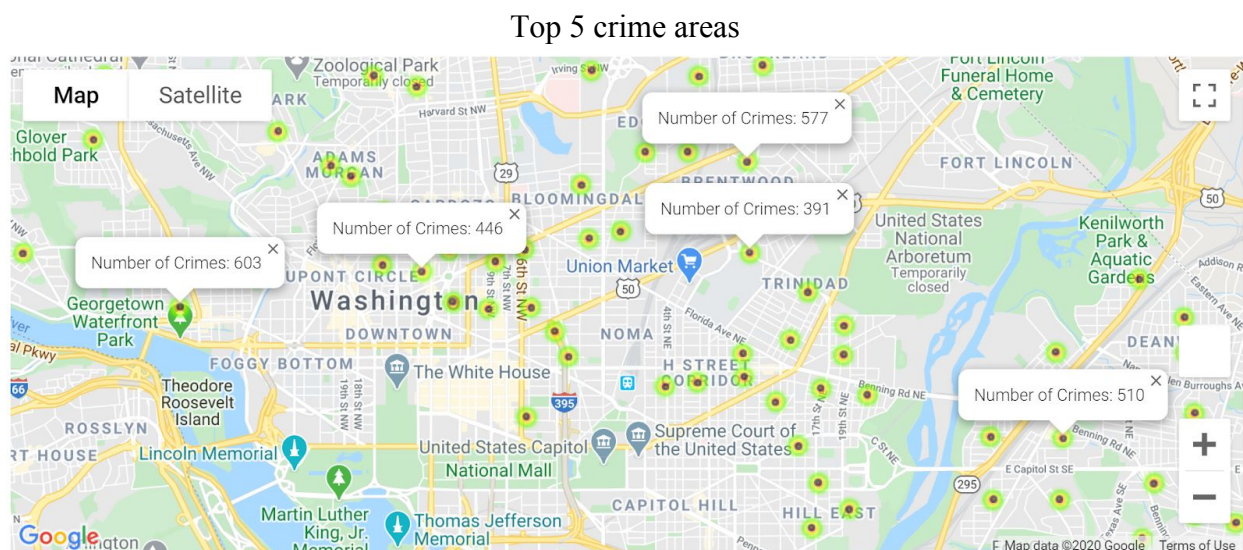


Figure 3.

Figure 3. Displays the most dangerous areas in D.C. based on data from incidents reported in the ASAP (Analytical Services Application) crime report database by the District of Columbia Metropolitan Police Department (MPD). <https://dcatlas.dcgis.dc.gov/crimecards/>

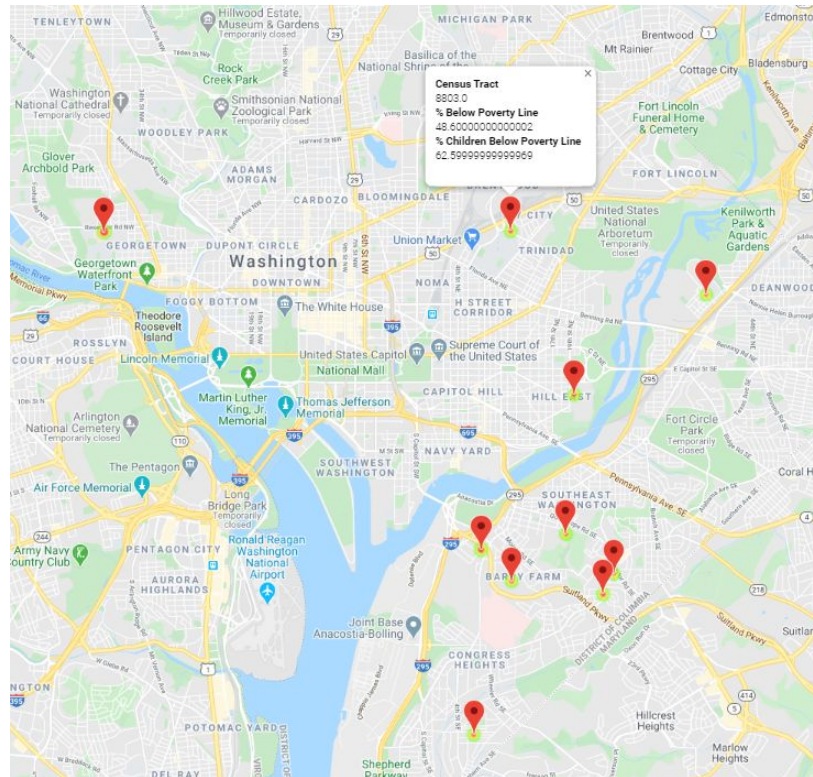


Figure 4.

Figure 4. Displays the 10 poorest neighborhoods by percentage. Washington DC is divided up into four quadrants with the Capitol Building at the center. We have seven markers in the Southeast quadrant, two in the Northeast and 1 in the Northwest. It is important to note that the one in the Northwest is Census Tract 201, which only accounted for 24 people, all of whom were below the poverty line. This Tract is an outlier for two reasons; one, because the size of the sample was so small in comparison (most other Tracts exceeded 1000 people) and two, it led the percentage of people below the poverty line (100% vs. 64.6%, which was 2nd poorest). In the future, we would probably add the data from the three census tracts that fell below 1000 people in with a neighboring tract.

Given that these points only correspond with the poorest neighborhoods, this data may be of use to someone involved with planning where a Food Bank should be established.

Correlation Analysis:

What is the strength of the relationship between poverty and crime?

A Statsmodels evaluation of the Total Population vs Number of Crimes was assessed. The following are key takeaways:

a) A regression analysis was used to quantify the relationship between the total population of poverty to the number of crimes. The accuracy of the model in Figure 10. is mildly relevant. The data shows a weak uphill positive linear relationship with an R squared value of 0.82.

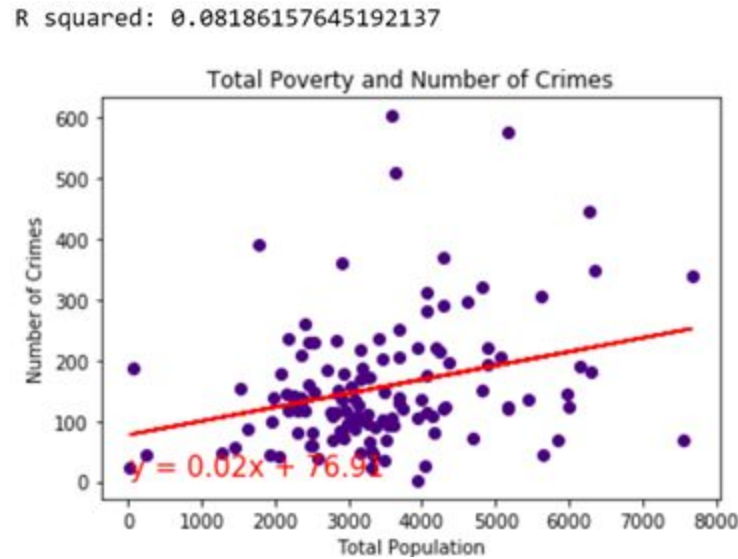


Figure 5.

b) Although there is a correlation between the Total Population and Number of Crimes, this does not imply that there is a causal relationship between the two. For example, a person who becomes poor does not necessarily have an impact on whether the person will commit crimes.

	Unnamed: 0	Census	Number of Crimes	Total Population	Percentage of Population in Poverty
Unnamed: 0	1.000000	0.969295	0.065111	-0.306494	0.251593
Census	0.969295	1.000000	0.096092	-0.317483	0.298728
Number of Crimes	0.065111	0.096092	1.000000	0.286115	0.014407
Total Population	-0.306494	-0.317483	0.286115	1.000000	-0.207819
Percentage of Population in Poverty	0.251593	0.298728	0.014407	-0.207819	1.000000
Percentage of Children in Poverty	0.370181	0.419992	0.121176	-0.082997	0.764240
Latitude	-0.491519	-0.577254	-0.046053	0.188685	-0.541075
Longitude	0.754286	0.796605	0.147368	-0.266772	0.338487

Figure 6.

Figure 6. The Dataset Correlation table shows a small to moderate correlation between Total Population and Number of Crimes, $r=.286$.

OLS Regression Results

Dep. Variable:	Number of Crimes	R-squared:	0.082
Model:	OLS	Adj. R-squared:	0.075
Method:	Least Squares	F-statistic:	11.15
Date:	Sun, 10 May 2020	Prob (F-statistic):	0.00111
Time:	15:24:26	Log-Likelihood:	-766.87
No. Observations:	127	AIC:	1538.
Df Residuals:	125	BIC:	1543.
Df Model:	1		
Covariance Type:	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
const	76.9122	25.251	3.046	0.003	26.937	126.888
Total Population	0.0228	0.007	3.338	0.001	0.009	0.036

Omnibus:	53.126	Durbin-Watson:	1.632
Prob(Omnibus):	0.000	Jarque-Bera (JB):	144.582
Skew:	1.634	Prob(JB):	4.02e-32
Kurtosis:	7.079	Cond. No.	1.03e+04

Figure 7.

c) Figure 7. displays the Stats model table with a Low R-squared value of 0.082 and a low p-value of 0.001. These values do not explain much of the variation of the data but indicate that these relationships are somewhat significant and predicting.

d) The following figure shows the correlation matrix on a heatmap.

We can derive the following insights from this correlation matrix:

- There is a moderate correlation between Total Population and Number of Crimes.
- Longitude and Percentage of Children in Poverty also shows a high moderate correlation.

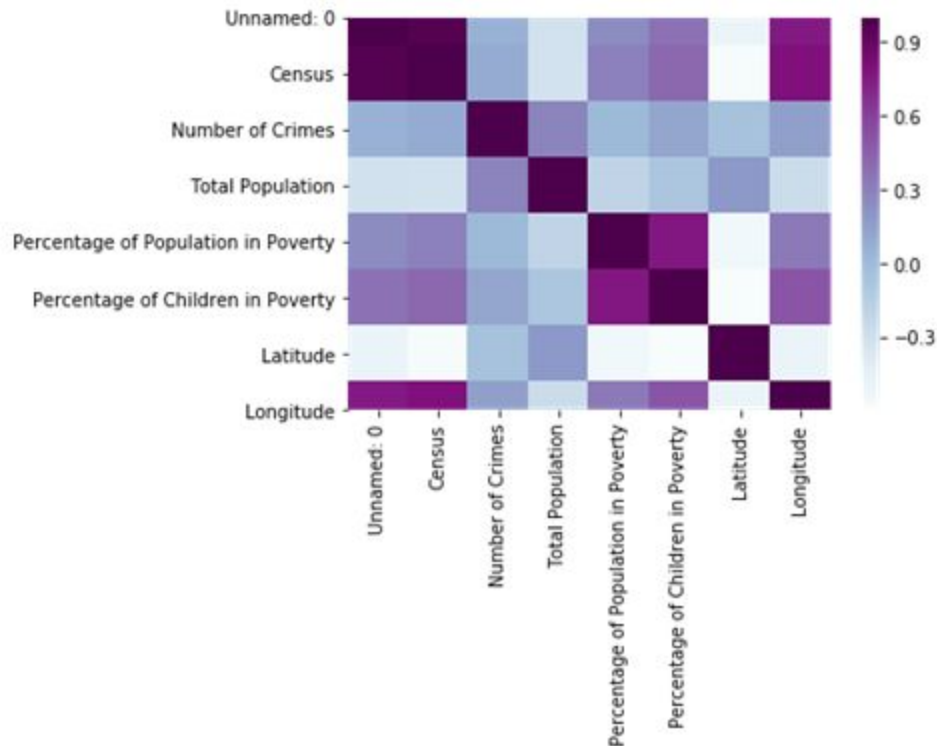


Figure 8. Heatmap Matrix

e) The Predicted vs Actual plot gives a visualization of the relationship between Total Population and Number of Crimes. From this scatter plot you can see that the model greatly underestimates the actual values with 250 predicted crimes and actual value of 600.

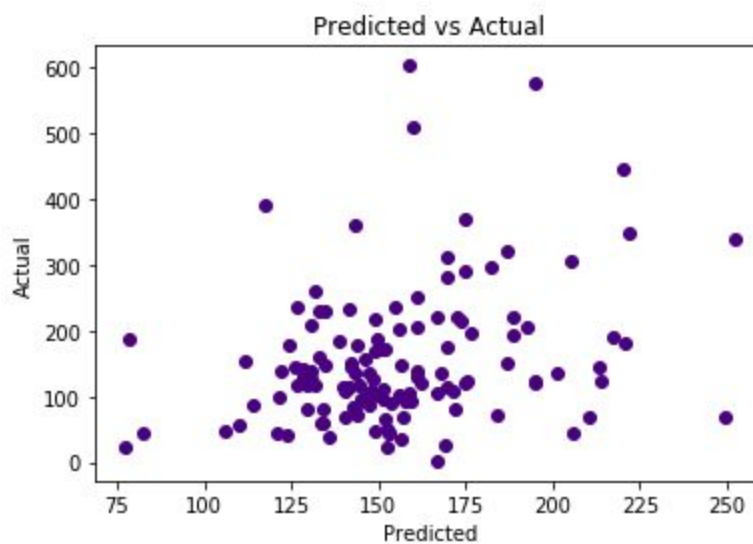


Figure 9.

f) The residual plot does not show any randomness. The positive values for the residual mean the prediction was too low. A clear pattern or trend in the residuals can be detected meaning the model has room for improvement.

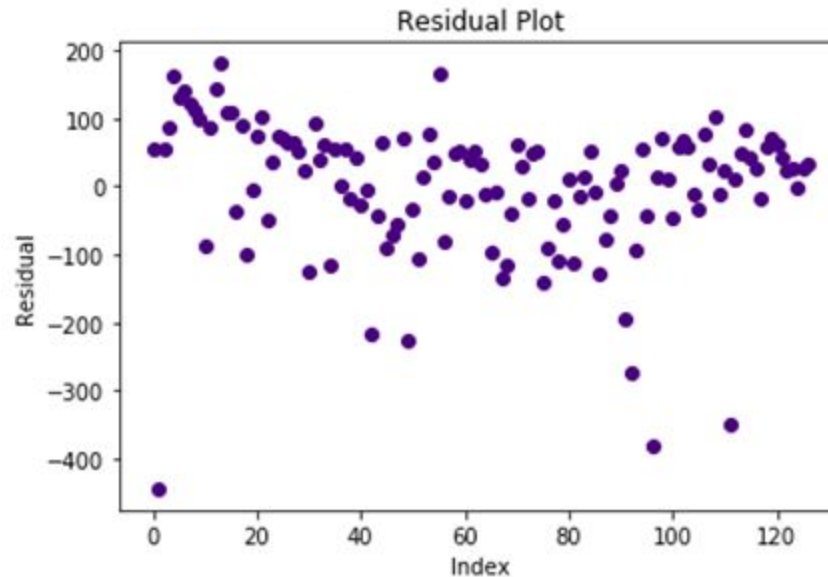


Figure 10.

Conclusions:

The influence of poverty on crime is not easily captured by the connotation that poverty alone leads to crime. While the research shows that there is a link between poverty and crime, the link can be elusive. The variables that make up this link are many and complex. Crime is connected to people of poverty if in part because poor people may be more likely to be victimized by it. While most poor people are lawful, most criminals are poor. Poverty may be a necessary means for crimes to be committed but not necessarily a sufficient reason or condition that leads to crime.

References:

Gregory, G(2017, March 21). Income inequality and economic mobility in D.C.

Retrieved from:

https://www.dcpolicycenter.org/publications/income-inequality-and-economic-mobility-in-d-c/#_ftn3

DC crime (2018), Washington, DC Crime

Retrieved from:

<https://www.areavibes.com/washington-dc/crime/>