CS910 Project : Crime and Wealth: A Glimpse into Greater London

University ID : 2200579

January 4, 2022

Abstract-This paper will analyse data relating to the level of crime and wealth of an area throughout Greater London. The paper will then attempt to identify trends connecting the two to identify a pattern between crime and wealth. Through data analysis weak correlations between different crimes and wealth will be uncovered and classification tools will be used to discover if crimes can be used to predict the wealth of an area. Findings will highlight the links between the level of crime and wealth which would be useful for effective allocation of resources to combat crime in London in a COVID-19 ravaged economy.

1 Introduction

Research suggests that there exists a relationship between the wealth of the residents and the level of crime of an area. The purpose of this study is to examine the type of relationship between average wealth of the residents and the level of crime within an area, to examine which types of crimes by ward may be more sensitive to the average wealth of the residents and consider why it possible is this way. Existing research will be considered to understand the current landscape in the research on wealth and crime. From looking at the research, it is anticipated the wealthier the borough the lower the amount of crime in the ward. Therefore, different crime types in each London ward and the average wealth of residents are correlated, to observe the impact of wealth on crime levels. In order to test the hypothesis that these different crimes can predict the average income level of a ward, a classification model was built. The strength and reliability of these observations are then discussion, followed by recommendations for further work.

1.1 Crime in London

London has seen a significant reduction in the crimes reported in 2020 by a total of 15% from 921,211 the year prior to 790,123 in 2020[5], this is due in part to COVID-19. The subsequent lockdowns within the UK resulted in reduced movement to combat the virus, it also affected the level of criminal activity over the whole of the UK. The decrease had been the largest annual drop since 2010 with crimes like

robbery and theft dramatically falling in report levels while crimes such as drug offences increased[1], which could be a combination of more people using drugs to cope with stress during lockdown and police operations being much easier due to restrictions on leaving the house making operations against drug dealers easier since they are at home.

1.2 Wealth and Crime

It has been long questioned what the true causes of crime are, these discussions tend to split people into two groups, Individualists who tend to focus more on the individual referring to weakness and greed as the factors behind why anyone could commit a crime. Then there are those who are collectivist who believe an impoverished and dangerous environment[4] are the main causes behind criminal behaviour. It is now generally accepted that despite the causes of crime being very complex, factors such as poverty, alcohol and drug abuse and parental neglect have long term effects on crime[2]. Research has indicated there is a the link between crime and wealth as there is evidence that wealth is closely linked with crime on a broader societal level[7]. As shown in figure 1 there is evidence that show a link with the level of crime and the wealth of an area "80% more crimes were recorded in the most income-deprived areas"[6]. This is further compounded by research that people with household income of less than £10,000 are "62% more likely to experience personal crime as well as, 73% more violent crime" [3] than households of over £50,000.

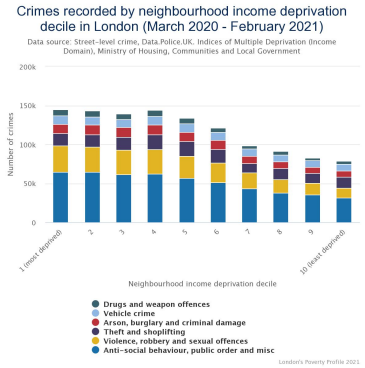

Figure 1 : Crimes by neighbourhood income [6]

2 The Data

For this project, the data sets were chosen to investigate the relationship between the crime in and wealth were from the primary source of data: with the total number of crimes within London and the mean taxpayer income of the boroughs from the London Data Store.

2.1 Crime Data

The London Wards recorded crime data by geographic breakdown, is a collection of CSV files which contains 66.14MB of data which breaks down the entire recorded crime per month within London by crime type which includes Arson, Burglary, Drug Offences, Possession of Weapons, Robbery, Theft, Sexual Offences. The data sets are broken down into three geographic levels within London by Borough, Ward, LSOA from December 2019 to November 2021. The CSV file chosen for this project contains 3MB of data, showcasing the recorded crimes reported by the wards and boroughs of London.

Figure 2 was created using excel and is used to visually represent the level of crime within each borough, highlighting the levels of crime throughout Greater London. It showcases that the level of crime varies across London but seems to increase towards the centre point, which gives premise for the investigation.

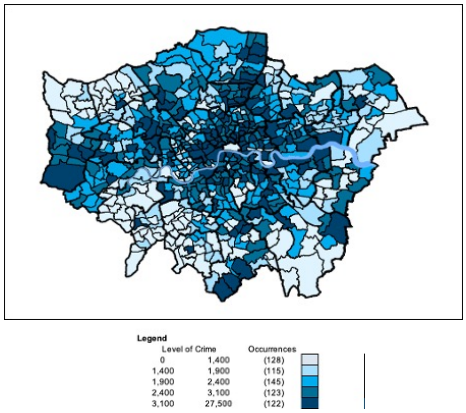

Figure 2 : level of crime in London

2.2 Wealth Data

The data set of 40KB contains the number of taxpayers and their Mean and Median Income by tax year by boroughs of London and regions of the UK, with data being retrieved from a survey of HMRC records, year on year from 1999 to 2019. However, this data set was already processed and gave the median and mean for each borough and this meant it was immediately accessible and available to be used for analysis.

3 Hypothesis

With the knowledge that wealth can be beneficial in reducing the level of crime in an area, it follows that there would exist a negative relationship between the wealth and the level of crime in an area. As such my current Hypotheses are:

1) Wards with higher mean taxpayer income will have lower levels of crime

2) The level of crime within a ward can be used to predict the average amount of wealth of taxpayers within the ward

4 Data Cleaning

4.1 Combining data set

The first data set contained the number of reported crimes such as Arson, Burglary, Drug Offences, Possession of Weapons, Robbery, Theft, Sexual Offences and more within each ward and borough of London, this data set broke down the crime types of each ward by month from Dec 2019 to Nov 2021 and did contain some missing values which will be addressed later. It also did not have a total amount of each crime for each ward over a 24 month period by each crime and this was one of the first actions undertaken. Then the Boroughs were removed with the wards being kept and the total of each crime over the 24 month period was listed next to the wards for each of the 11 crime types. The total reported crimes from the 24 month period were summed and then the data set was reformatted for the total number of crimes in each ward by crime type with missing data being screened for using python.

The second data set contained the processed data of the mean and median income of taxpayers within each borough, this data contained values for

individuals paying tax in the borough along with the median and mean income within the borough year on year from 1999 to 2019. As only the most recent year was necessary, the data from 1999 up until 2018 was removed along with the number of individuals and the median income for the borough and the mean income was kept. The mean income for each borough was then applied to all the wards within each of the boroughs for the analysis of how a ward's wealth impacts the amount of crime within it.

4.2 Missing Values

There were wards within the Crime data which did not include any quantitative information for one of the crime types, the reason for this could be because data was not collected for the wards Cranham and Sanderstead for the crime Possession of Weapons or the data could have been inputted wrongly by whoever was handling the data collection, regardless as there were no data, all the missing values were set to 0.

4.3 Groups

For the data to be classified the mean income of each borough was compared to the mean income of London, so that if the income of a borough of London was below the mean income of London the borough and the wards within them would be labelled as "Not Wealthy" and if the average income of the borough was above London it would be labelled as "Wealthy".

5 Results

5.1 Correlations

To observe the patterns within the cleaned reported crime data set, the reported crimes were plotted against the wealth of the ward. The linear equation of each correlation was calculated along with the coefficient of determination. The coefficient is used to determine if the variations in wealth may be explained by the different crime types; this was used to establish the strength of the relationship between wealth and each crime type.

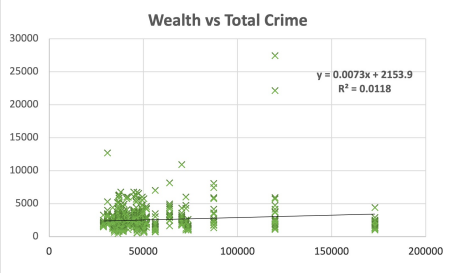

Figure 4 : Wealth and Total Crime per Ward

As shown by the figure 4 it was found that instead of a negative correlation between the wealth of a ward and the level of crime reported within, it was found that there is a positive correlation between the wealth of a ward and the level of crime within the ward. However, with the coefficient of determination being quite small as R^2 is closer to 0 than 1 it means we can infer the correlations between the two variables is weak.

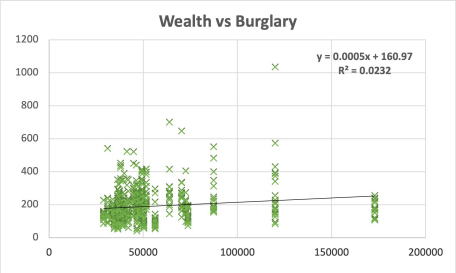

Figure 5 : Wealth and Level of Burglary per Ward

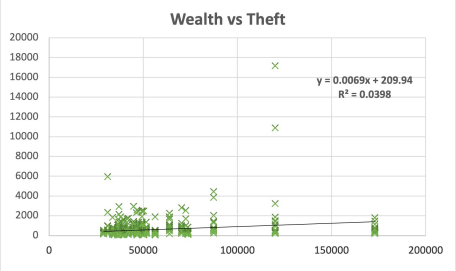

Figure 6 : Wealth and Level of Theft per Ward

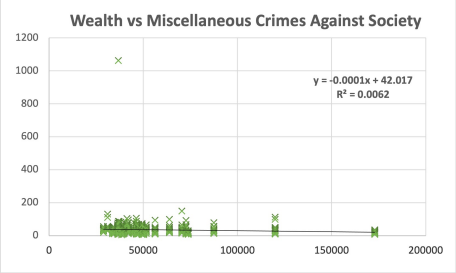

Figure 7 : Wealth and Level of MCAS per Ward

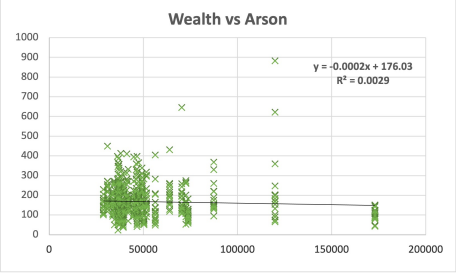

Figure 8 : Wealth and Level of Arson per Ward

Through examining the figures, it can be inferred that the level of crimes such as arson and miscellaneous crimes against society reduce as the borough average income increases. However, crimes such as burglary and thefts increase as the borough's average income increases. From this we can infer reported crimes related to stealing may increase the wealthier the borough is. Although as the coefficient of Determination for each of these plots are closer to 0 than 1, we must accept that the correlation between specific crimes and the wealth of a borough are weak

5.2 Building Classifiers

To build on my investigation of the relationship between wealth and crimes of boroughs of London, Weka software was used to show if the level of crime within wards of London can be determined by the wealth of the borough. To do this wealth classes were created for the average income of boroughs and assigned to the wards within the boroughs as described above. Along with wealth class were created from the crime types within the dataset 11 other nominal attributes: sexual offences, robberies, Drug Offences, Thefts, Arson and criminal damage, Possession of weapons, Vehicle offences, Burglary, Miscellaneous Crimes Against Society, Public Order Offences this totalled 614 data points. These datapoints where tested on 2 different classifiers, Naïve Bayes and a K-nearest neighbours' with K=1 to create a wealth class descriptor. For the Naïve Bayes, 63.5% of instances where correctly classified with Precision for Wealthy being 0.41

and Not Wealthy being 0.779 with a ROC of 0.651 and for the K-nearest neighbours' a summary of the results shown in figure 9.

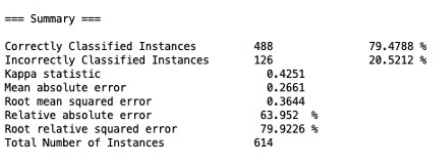

Figure 9 : KNN Summary

Therefore, with almost 80% of the instances correctly classified, and the precision of Wealthy being 0.796 and Not Wealthy being 0.795, along with ROC at 0.839 it can be determined that there exists some positive correlation between the wealth and level of crime in an area.

6 Analysis

As shown above there exists a weak correlation between the wealth and the several crime types, most notably theft, burglary and robbery in each ward. It could be suggested that as the strongest relationship was observed with burglary, that wealthier homes and areas may be more appealing to criminals who wish to undertake in crimes related to stealing and therefore it is reported more in wealthier wards. When considering the K- Nearest neighbour model, a significant percentage of data points were correctly assigned to their wealth class, a positive classification reinforced by the ROC values together shows that the findings have some significance. In effect, as wealth rises there seems to be a greater chance of Theft, Robbery and Burglary in London wards, this implies that wealth is one of the factors which influence the level of crime within an area. Therefore, it can be implied that the level of overall crime in an area can be used to predict the amount of wealth of its inhabitants, although with the motivations behind criminal activities being very complex, more factors, external and internal, need to be incorporated to improve the accuracy of the model.

7 Conclusion

It can be concluded that there is a relationship between wealth and the level of crime within a London ward. However, it was anticipated that the correlation would be negative, but instead it is positive, with the crimes most related to stealing such as burglary and theft being more of an indicator that the area is a wealthier borough. Furthermore, the level of crimes most closely associated with violence such as arson, possessing a weapon and violence against a person have the opposite relationship with these crimes an indicator for a less wealthy borough with the exception being sexual offences. This shows that it will be more beneficial for combating crime through policing and government funded programs in these areas to focus resources and budget in effective training that helps officers and charities combat crimes related to stealing in wealthier boroughs and in less wealthier boroughs resources and training should be more focused on tackling violent crimes. Despite the general trend, it should not be discounted that crimes unrelated to stealing such as sexual and drug offences also positively correlated with wealthy boroughs along with the overall level

of crime. While this was shown to be much more than random, a relative lack of strength in the reliability tests highlight the need for further investigation into other factors to discover important factors in crime.

7.1 Limitations

Because of the format and cleaning of the Wealth data set it was averaged to the boroughs and the issue being as my data was broken down by ward it resulted in the average mean income of taxpayers by borough to be used in the wards within them and the data is from 2019. This could have affected the sensitivity of the data as the wards average taxpayer income could have variation within the borough, and there could be a change in the mean income of boroughs in 2020-21. Also, another issue could be that fraud was not included in the crime data set, this was due to the data being transferred from police forces to the National Action Fraud in 2013 and this may have affected the accuracy of the analysis. Furthermore, the crime data set had no wards or crimes recorded within the City of London borough which may affect the analysis, while using reported crime disregards the crimes that are committed but not reported which could also effect the analysis.

7.2 Recommendations

As the correlations for each crime against wealth were weak, it suggests that where higher levels of income are observed, there is more of a chance a higher level of crime may be observed, rather than a direct link between the two. Therefore, further research is required to consider a large number of other factors such as the impact of Covid-19 and density of population in an areas effects on the levels of crime. Also, using the average income of wards or postcodes may produce a more insightful analysis of how wealth can impact the level of crime. It can be assumed with a large, more diverse data set there will be a more accurate conclusion to be drawn on how factors such as wealth effect crime in order to effectively combat crime in London.

References

[1] Dominic Casciani Ben Butcher. *COVID: What's happened to crime during the pandemic?* Mar. 2021. URL: `https://www.bbc.co.uk/news/56463680`.

[2] Francis T Cullen and Pamela Wilcox. *The Oxford handbook of criminological theory*. Oxford University Press, 2015.

[3] Peter Cuthbertson. *Poverty and Crime: Why a new war on criminals would help the poor most*. July 2018. URL: `https://www.civitas.org.uk/content/files/povertyandcrime.pdf`.

[4] Marijana M Kotlaja. "Cultural contexts of individualism vs. collectivism: Exploring the relationships between family bonding, supervision and deviance". In: *European Journal of Criminology* 17.3 (2020), pp. 288–305.

[5] regionadmin. *London Crime Rate 2021*. July 2021. URL: https://www.regionsecurityguarding.co.uk/security-company-london/london-crime-rate-2021/.

[6] unknown. *crime and income deprivation*. July 2021. URL: https://www.trustforlondon.org.uk/data/crime-and-income-deprivation/.

[7] CS Webster and Sarah Kingston. "Anti-poverty strategies for the UK: poverty and crime review". In: (2014).