FINAL YEAR REPORT

SM 4290 - RESEARCH PROJECT (STATISTICS)

The Statistical Investigation on The Possible Factors of Crime In London boroughs

By

SITI NURFARHANA BTE FERRY ALFIAN
(16B2153)

Faculty of Science

Universiti Brunei Darussalam

Acknowledgement

First of all, I would like to show my most profound gratitude to Almighty Allah for providing myself the opportunity , courage and patience throughout the semester in completing this paper.

Here I would like to take the opportunity to express my gratitude to my supervisor Dr Haji Mohammad Haziq Bin Haji Mohammad Jamil for providing a continuous guidance on my research and writing, positive and supportive attitude and encouragement given without which it would have not been possible to complete this Final Year Project.

I would also like to thank the university, University Brunei Darussalam for providing an opportunity to be able to conduct a research work and providing learning ground and access to this project.

Finally I would like to express my profound gratitude to my family for the support and encouragement, my parents Ferry Alfian Bin Hj Mohd Djauhari and Siti Azizah Bte Hj Abdul Rashid, my brother Mohammad Shafi'ei Bin Ferry Alfian throughout my years of study. This accomplishment would not have been possible without them.

TABLE OF CONTENTS

ABSTRACT

Crime is an act of immorality that is triggered upon various reasons; it can be committed by any individual without any sense of moral respect. Human beings, recently, have stood up for a legitimate jurisdiction against such a circumstance to prevent any loss to any other parties. Crime is visible in various places not only in the suburb, rural, and dense areas such as public malls or schools, but sometimes, it can also be seen in holy places such as mosques, churches, and temples. It is evident that crime does not only depend upon ethnicity, race, location, and time; thus, this makes it difficult for the police and other parties within the Justice Department to control and act on the crime rates. However, it is possible to understand the origin, root, and potential reason for the crime by learning the pattern within a certain area. This study addresses the data of crime occurrences from April 2015 to March 2016 in the 33 boroughs of London, including the City of London. This study entails three positive correlation findings. Firstly, the findings include burglary crime with the unemployment rate, education, and police officers. Next, they denote theft crime with population density, median house price, education, tax charge, and police officers. Lastly, the findings entail robbery crime with population density, unemployment rate, education, tax charge, and police officers.

INTRODUCTION

This report aims to statistically investigate the possible factors of crime rates in London boroughs, the United Kingdom, from April 2015 to March 2016. In the United Kingdom, the crime rates are often high and as stated by Machin (2000), the United Kingdom reached a crime-record level in the early 1990s as a result of its rapid increase throughout the 1970s and 1980s, which led to the emergence of public policy issue and subsequently the issue of crime origins and reasons. It was further concluded by multiple studies that crime is mainly caused by economic, demographic, and deterrent factors (Gaviria and Pagés, 2002; Mathur, 1977; Stevans, 1983; Meera and Jayakumar, 1995; and Masih and Masih, 1996).

There are different types of crimes that took place in London. While some are unique and unorthodox, others are generally common and can be categorized under the same category. The focus of this investigation will only be directed on 3 specific types of crime: Burglary, Theft, and Robbery.

Burglary refers to theft, or attempted theft, from a premise without any authorised access. Burglary also takes into account the damage to a premise, which appears to have been caused by a person who attempted an entry to commit a burglary. There are three available types of theft; however, this analysis will focus on the data set of theft from the person. In general, "theft from the person" is deemed committed by a person if he or she steals a property while the victim holds or carries the property. Meanwhile, robbery is deemed committed by a person if he or she uses force or the threat of force while stealing or attempting to steal from someone, which colloquially refers to "mugging".

It is undeniable that people need money to fulfill their necessities in life, such as for food, shelter, education, or general responsibility for a country. Hence, it is an ideal option to have a permanent job to fulfill these necessities because the job will secure an individual's source of income besides covering the individual's needs for daily consumption for survival, purchasing a house for shelter, and paying annual tax to the country.

Besides, it is also undeniable that the number of crime rates can increase with the greater number of people living in an area. The reason is that people compete for limited resources in a highly dense area; thus, when they are unable to acquire the resources, they tend to achieve them in the wrong way.

OBJECTIVES

- Crime occurrences in the 33 boroughs of London, including the City of London.
- The different socio-economic variable rates in the 33 boroughs of London, including the City of London.
- The correlation between the crime rates and the socioeconomic variables in the 33 boroughs of London, including the City of London.

RESEARCH

The importance of these research questions is to mainly improve the overall safety of the areas affected by learning the patterns and contributing factors. With the focused research, it would enable the public citizen to have a greater awareness of the areas that has high crime occurrences and the type of crime, hence allowing the people to be more careful and much more vigilant. This knowledge would also benefit the tourist and tourism agency, as the tourist activity does contribute in boosting the petty-crime rates, this can be seen from Westminster as being the Europe's Largest Nighttime economy, but also has the capital's highest crime rate, where 37,315 theft incidents were reported in back in 2008. In conclusion, the knowledge obtained from these research questions would benefit the people and environment to be much more aware and alert of areas categorized as dangerous zone according to the statistic of the crime occurrence.

1. Which of the London Boroughs has the highest crime rate occurrences?
2. Does the crime rate correlate with the independent variables?

3. Which variable has the most influence on the number of crime occurrences in the boroughs?

LITERATURE REVIEW

Crime originates from various possible factors. In this report, the information will only focus on Gross Annual Pay, Population Density, Unemployment Rate, Median House Price, Job Density, Education, Tax Charge, and the Number of Police Officers.

Gross Annual Pay is defined as the yearly total of a particular resident's wage, whereas Population Density refers to the total number of residents living within the 33 boroughs of London, including the City of London. Meanwhile, the unemployment rate occurs when an economy is unable to generate employment for the individuals who intend to work but are not doing so regardless of their employment availability and active work-seeking. Median House Price is referred to as the required amount of money to be paid for the shelter, while Job density is known as the total number of jobs available per total population in the 33 boroughs of London, including the City of London. Additionally, education is included by the percentage of pupils obtaining A* to C in English and Mathematics GCSE examination; hence, more educated individuals are produced with the potential to develop a legal income, thus simultaneously reducing criminal activities. While Tax Charge specifically refers to housing tax, this report also analyzes the number of officers based on the logical assumption that the higher number of available officers would minimize crime activities in the area.

In an attempt to reduce crime and further enhance the citizens' standard of living, it is crucial to understand how these factors might affect the crime rates as well as controlling the crime level and frequency because it influences not only the overall standard of the society but also the country. The act of misconduct by offenders is one of the top major significant social problems that have an impact on public safety, youth development, and socioeconomic status, especially the adults (Wang, Hongjian, & Kifer; Daniel & Graif; Corina & Li; Zhenhui, 2016).

Many studies and investigations have been conducted to establish the correlation between crime rates and their factors. As such, we can emphasize and narrow down the possible factors to observe the correlation between crime rates and their underlying factors:

1. Based on Becker's (1968) framework, the opportunity of cost crime (foregone wages while incarcerated) increases as the wage increases; hence, this will reduce an individual's tendency to commit criminal activities.

2. The studies by Carmichael and Ward (2000) reported the following findings:
    a. A systematic positive relationship between burglary rates and male unemployment regardless of age.
    b. A consistent and positive association between youth unemployment to criminal damage and robbery rates.
    c. A positive relationship between adult unemployment and theft.

3. Gumus (2003) mentioned that crime factors are crucial determinants of crime rates by which these factors are the effects of the per capita income, income inequality, population, and the presence of the black population on crime rates in the US.

4. Based from the empirical study by Marvel and Moody (1996) a positive relationship was found between crime rates and police numbers. This was further supported by 36 reviewed papers that examined the association between police numbers and crime rates, where 15 papers found a positive correlation.

DATA

The data concerning various socioeconomic variables in the whole UK is publicly available and obtainable from the Office of National Statistics (ONS), UK. This data includes a variety of possible factors in the 33 boroughs of London; however, this report will only focus on 8 possible factors related to crime rates.

The possible factors in this report include Gross Annual Pay (GAP), Population Density (PD), Job Density (JD), Unemployment Rate (UR), Median House Price (HP), Job Density (JD), Education (Edu), House Tax (Tax), and Average Number of Police Officers (PO). Next, the data will be entered and the analysis will be carried out via a mathematical software R.

Besides, the crime sets data is also publicly available and obtainable from data.police.uk and met.police.uk. There are 14 types of crimes listed in the Police department, which include

Theft from the Person, Robbery, Violence and Sexual Offences, Burglary, Criminal Damage and Arson, Other Theft, Shoplifting, Vehicle Crime, Bicycle Theft, Possession of Weapons, Anti-Social Behaviour, Drugs, Public Order, and Other Crimes.

METHODOLOGY

The data concerning various socioeconomic variables in the whole UK is publicly available and obtainable from the Office of National Statistics (ONS), UK. The data includes 10 possible factors in the 33 boroughs of London, which will then be entered and analyzed using a mathematical software R.

The analysis aims to examine the correlation between crime occurrences and their possible factors.

The method of visualizing spatial data will be employed to achieve the first two objectives of this research, which are to observe crime occurrences and the possible factors of crime in the boroughs of London. By layering the data over the map of the London boroughs, this method will assist in observing the intensity of crime occurrences including the possible factors of crime in the 33 boroughs of London.

The method of visualizing spatial data will be employed to achieve the first two objectives of this research, which are to observe crime occurrences and the possible factors of crime in the boroughs of London. A template of heat map in Microsoft Excel was obtained from data.london.gov.uk where the data was inserted into the template, while the legend was classified into 5 number of ranges where each range can be seen by the intensity of colour. By layering the data over the map of the London boroughs, this method will assist in observing the intensity of crime occurrences including the possible factors of crime across the 33 boroughs of London.

The multiple linear regression model will be used to see the correlation between the number of crime occurrences and the possible factors because the factors comprise more than one variable. Besides able to observe the strength of these variables' effect on the dependent variable, this model can also calculate the amount of impact on the dependent variable if the independent variables are changed as well as predicting trends and future values.

To create a model, the First Data will be inserted into the R software; however, this model is not reliable due to the presence of a possible error. As such, the backward elimination method will be applied to solve this issue whereby any variable with a p-value lower than the significant level (SL = 0.05) will not be included for further analysis; thus, a model with all

significant independent variables will be created. Lastly, the stepwise regression in the R software will determine and reassure that the model built in this method is the best model.

RESULTS AND DISCUSSION

VISUAL SPATIAL DATA – HEAT MAP METHOD

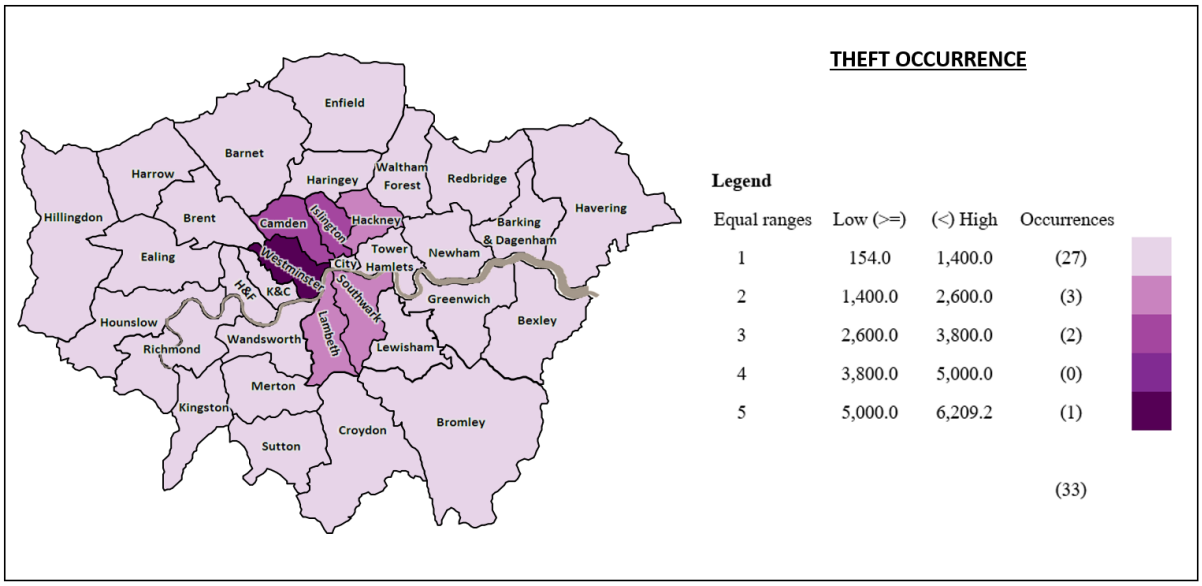

Figure 1: Theft Occurrences

Figure 1 shows the number of Theft occurrences in 33 London Boroughs, including the City of London. It shows Westminster has the highest occurrences of Burglary of 6203 cases.

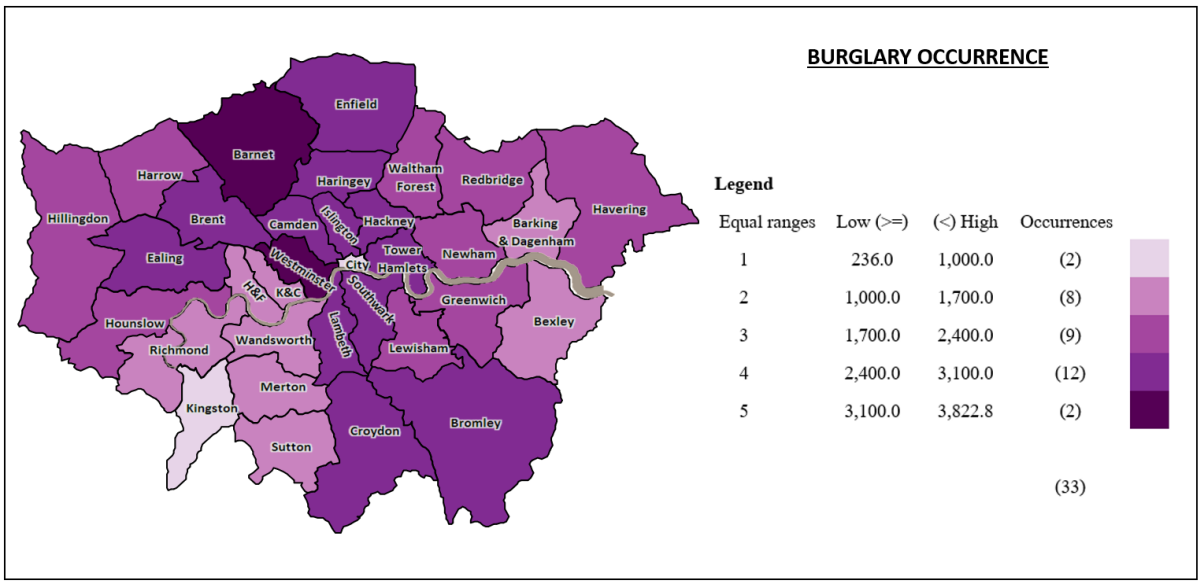

Figure 2: Burglary Occurrences

Figure 2 shows the number of Burglary occurrences in 33 London Boroughs, including the City of London. It shows Westminster has the highest occurrences of Burglary of Barnet with 3819 cases.

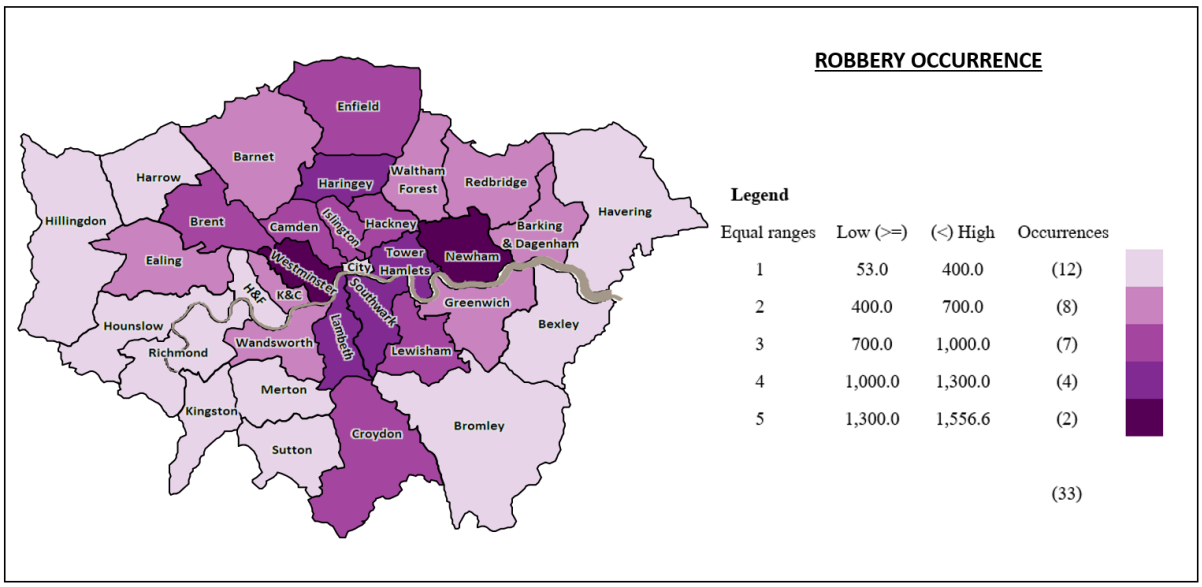

Figure 3: Robbery Occurrences

Figure 3 shows the number of Robbery occurrences in 33 London Boroughs, including the City of London. It shows Westminster has the highest occurrences of Burglary of Barnet with 1555 cases.

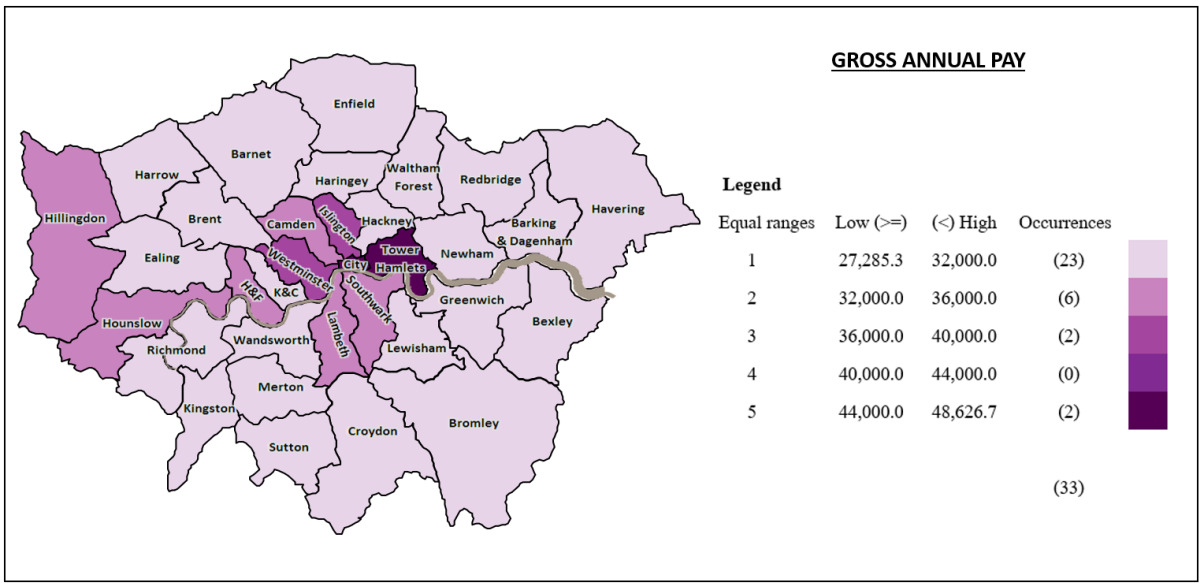

Figure 4: Gross Annual Pay

Figure 4 shows the Gross Annual Pay in 33 London Boroughs, including the City of London. It shows Tower Hamlets has the highest annual salary of £48578.

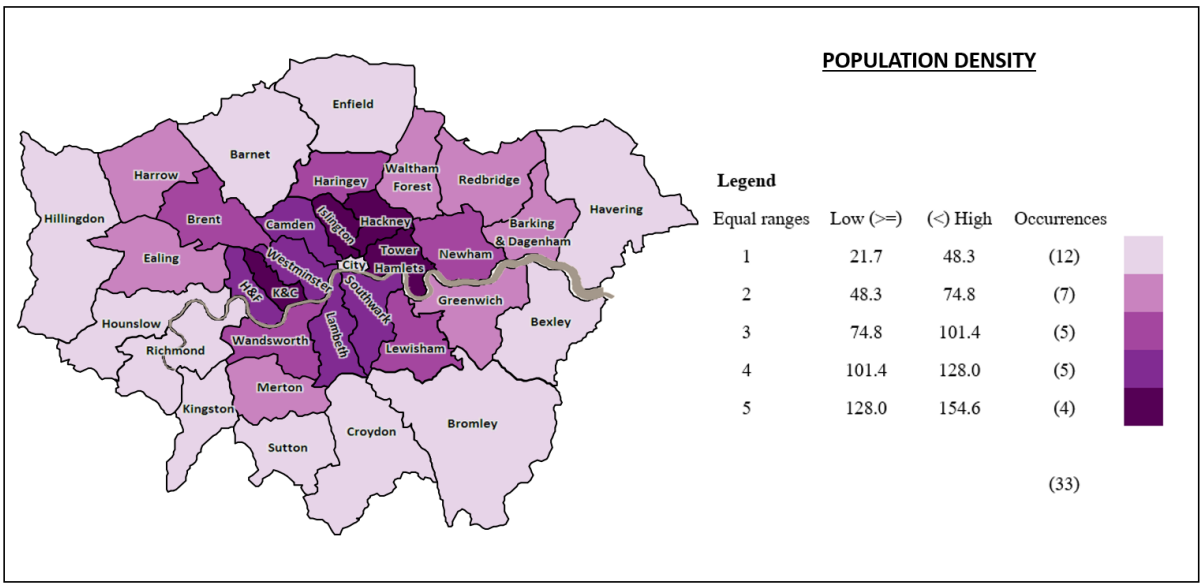

Figure 5: Population Density

Figure 5 shows the Population Density in 33 London Boroughs, including the City of London. It shows Islington has the highest population density of 154.4 people per hectare.

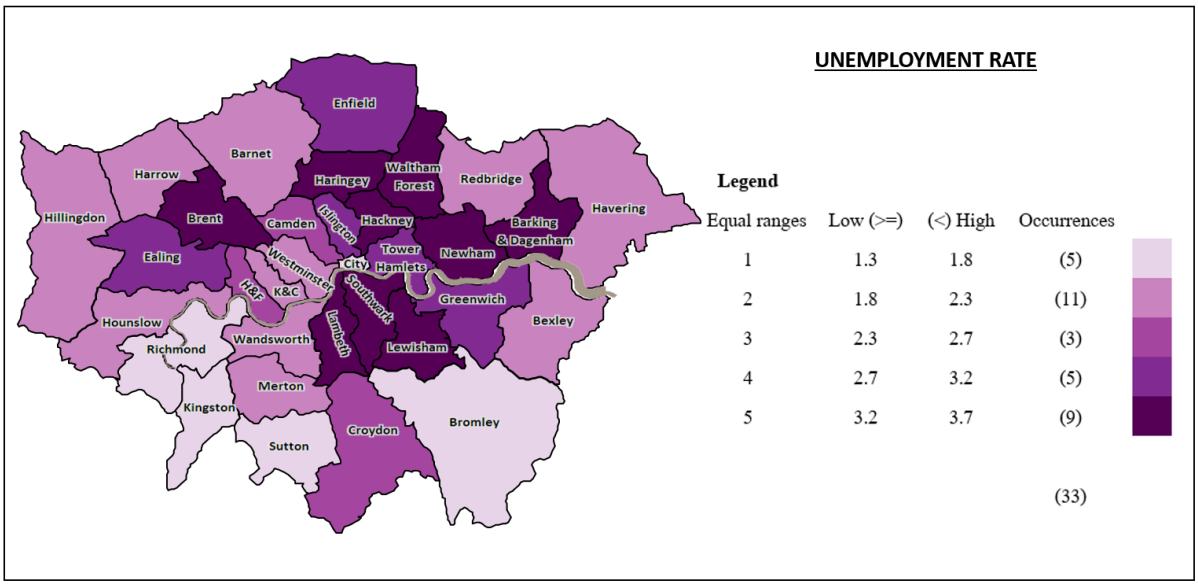

Figure 6: Unemployment Rate

Figure 6 shows the Unemployment Rate in 33 London Boroughs, including the City of London. It shows Hackney has the highest rate of unemployment of 3.7.

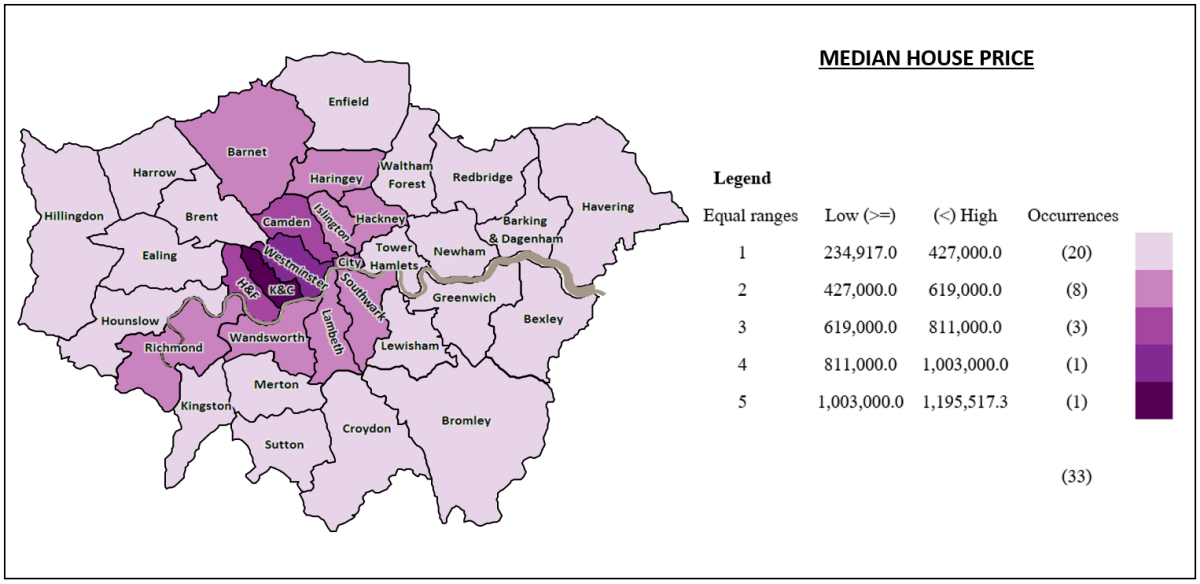

Figure 7: Median House Price

Figure 7 shows the Median House Price in 33 London Boroughs, including the City of London. It shows Kensington and Chelsea (K&C) has the highest median house price of £1,194,323.

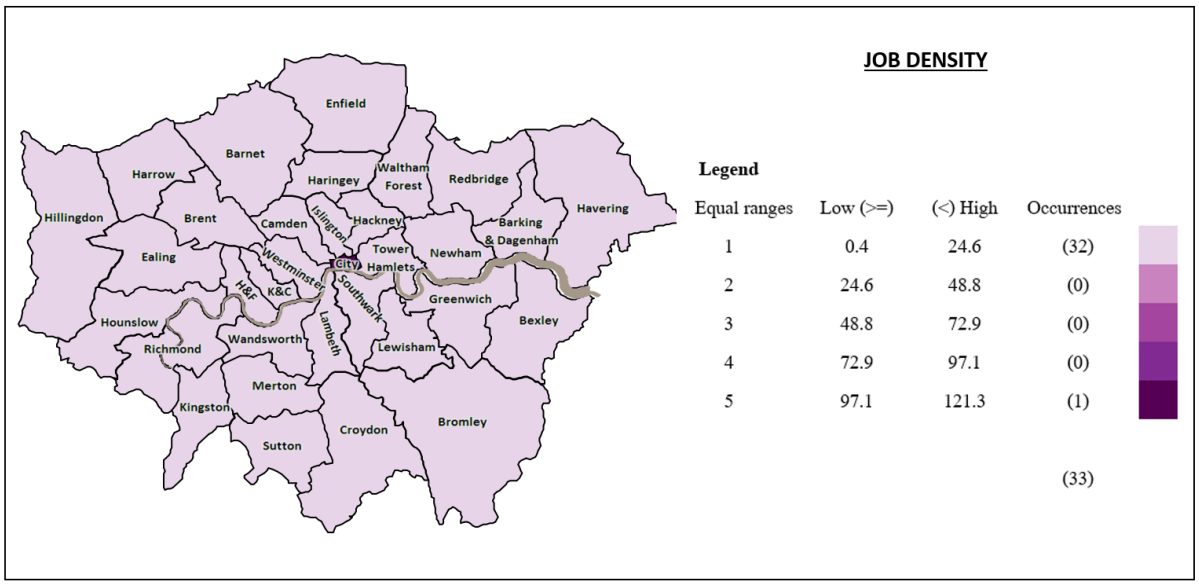

Figure 8: Job Density

Figure 8 shows the Job Density in 33 London Boroughs, including the City of London. It shows that the City of London has the highest Job density of 154.4 number of job per resident of working age.

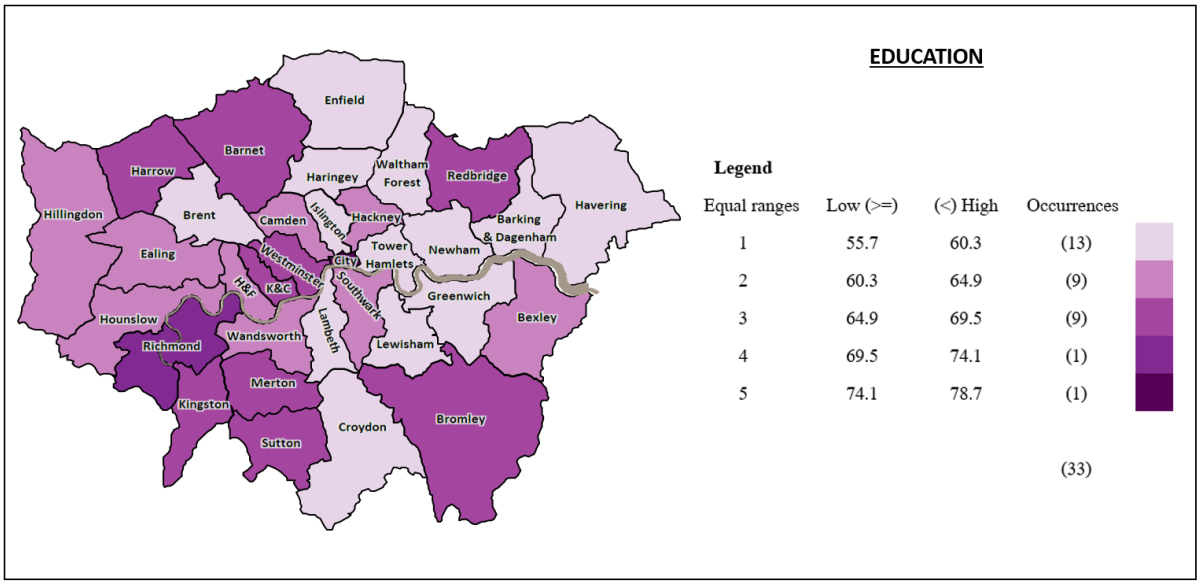

Figure 9: Education

Figure 9 shows the Education of in in 33 London Boroughs, including the City of London. It shows the City of London has the highest percentage of education with 78.7% .

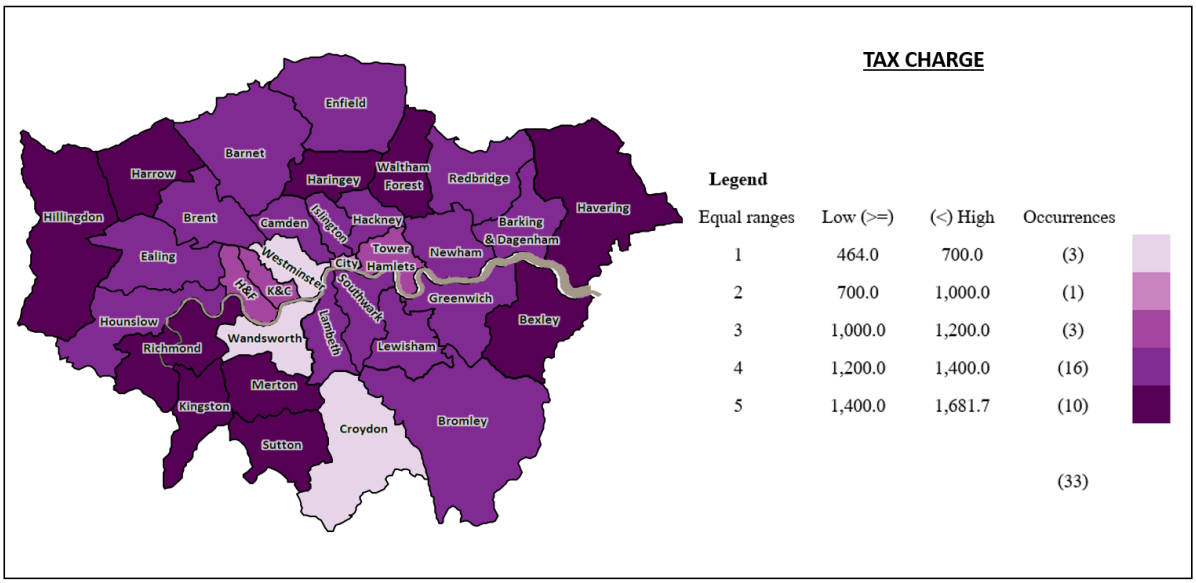

Figure 10: Tax Charge

Figure 10 shows the House Tax Charge in 33 London Boroughs, including the City of London. It shows Kingston has the highest Median House Tax Charge of £1680.

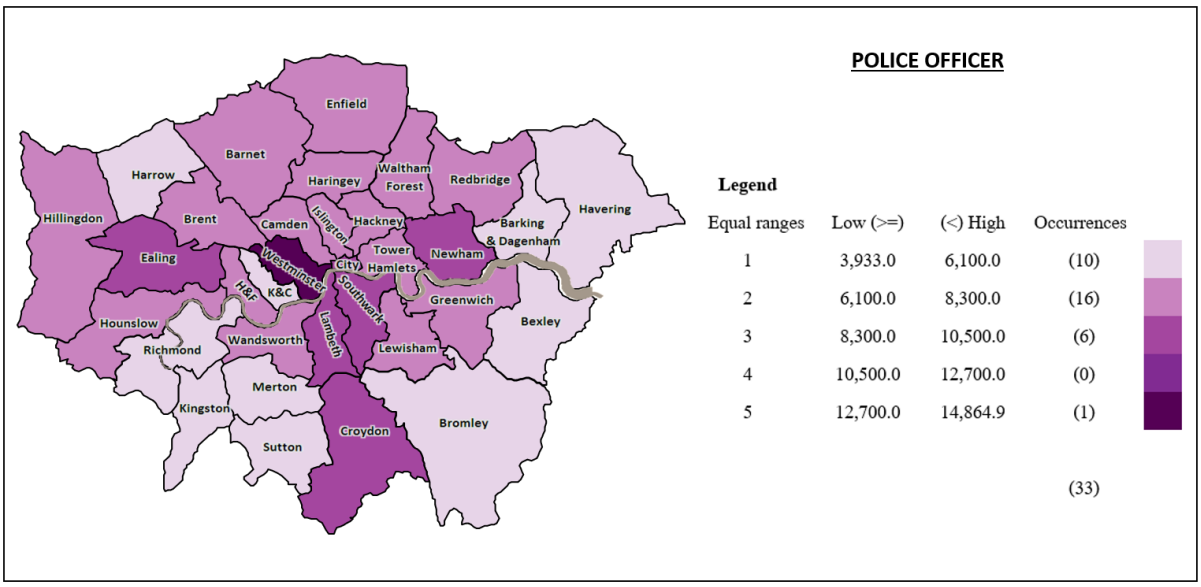

Figure 11: Police Officers

Figure 11 shows the number of Police Officers in 33 London Boroughs, including the City of London. It shows Westminster has the highest number of Police Officers of 14850 officers.

Table 1 Summary Statistics

Variable	Sample size	Mininum	Q1	Median	Mean	Q3	Maximum
Burglary	396	14.00	133.80	178.00	175.00	217.00	434.00
Theft	396	6.00	30.75	46.50	85.85	104.00	609.00
Robbery	396	1.00	27.00	46.50	53.54	77.50	169.00
GAP	396	2201.00	2409.00	2528.00	2650.00	2755.00	4152.00
PD	396	21.70	43.80	58.10	73.35	107.58	157.00
UR	396	0.90	1.80	2.30	2.44	3.20	4.00
HP	396	220000.00	327000.00	400000.00	456822.00	490000.00	1,210,000.00
JD	396	0.39	0.57	0.68	4.56	1.06	121.73
Education	396	55.70	59.70	61.20	62.84	67.30	78.60
Tax	396	56.00	103.00	113.00	108.90	120.00	140.00
PO	396	309.00	496.50	579.00	587.00	680.80	1275.00

Table 1 shows the basic statistic for each both dependent and independent variables in the regression model analyzed in this paper.

MULTIPLE LINEAR REGRESSION

For Burglary Occurrences

```
Call:
lm(formula = Burglary ~ GAP + PD + UnemploymentRate + HP + JD +
    Education + Tax + PO, data = regmodeldata)

Residuals:
    Min      1Q  Median      3Q     Max
-82.474 -30.917  -4.455  24.814 252.302

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)      -2.204e+02  8.007e+01  -2.752  0.00620 **
GAP               1.096e-02  1.075e-02   1.020  0.30854
PD               -5.874e-02  1.294e-01  -0.454  0.65001
UnemploymentRate  1.278e+01  6.034e+00   2.118  0.03480 *
HP               -2.594e-05  2.176e-05  -1.192  0.23385
JD               -1.937e+00  2.242e-01  -8.640  < 2e-16 ***
Education         3.320e+00  1.030e+00   3.223  0.00138 **
Tax               1.712e-01  2.082e-01   0.822  0.41138
PO                2.264e-01  1.962e-02  11.539  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 45.88 on 387 degrees of freedom
Multiple R-squared:  0.5008,    Adjusted R-squared:  0.4905
F-statistic: 48.53 on 8 and 387 DF,  p-value: < 2.2e-16
```

Figure 12: Input Data with 8 Variables

```
Call:
lm(formula = Burglary ~ GAP + UnemploymentRate + HP + JD + Education +
    Tax + PO, data = regmodeldata)

Residuals:
    Min      1Q  Median      3Q     Max
-82.486 -31.202  -4.782  23.831 252.427

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)      -2.160e+02  7.942e+01  -2.720  0.00682 **
GAP               8.354e-03  9.080e-03   0.920  0.35813
UnemploymentRate  1.132e+01  5.106e+00   2.218  0.02714 *
HP               -3.177e-05  1.754e-05  -1.811  0.07088 .
JD               -1.881e+00  1.868e-01 -10.072  < 2e-16 ***
Education         3.347e+00  1.028e+00   3.257  0.00122 **
Tax               1.877e-01  2.048e-01   0.917  0.35980
PO                2.276e-01  1.941e-02  11.729  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 45.83 on 388 degrees of freedom
Multiple R-squared:  0.5005,    Adjusted R-squared:  0.4915
F-statistic: 55.54 on 7 and 388 DF,  p-value: < 2.2e-16
```

Figure 13: Input Data with 7 Data (with PD removed)

```
Call:
lm(formula = Burglary ~ GAP + UnemploymentRate + HP + JD + Education +
    PO, data = regmodeldata)

Residuals:
   Min     1Q Median     3Q    Max
-83.08 -31.41  -4.74  23.21 253.16

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)      -1.972e+02  7.671e+01  -2.571 0.010515 *
GAP               7.304e-03  9.006e-03   0.811 0.417850
UnemploymentRate  1.222e+01  5.011e+00   2.439 0.015181 *
HP               -3.991e-05  1.513e-05  -2.638 0.008678 **
JD               -1.896e+00  1.861e-01 -10.188  < 2e-16 ***
Education         3.521e+00  1.010e+00   3.487 0.000543 ***
PO                2.193e-01  1.711e-02  12.813  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 45.82 on 389 degrees of freedom
Multiple R-squared:  0.4994,    Adjusted R-squared:  0.4917
F-statistic: 64.69 on 6 and 389 DF,  p-value: < 2.2e-16
```

Figure 14: Input Data with 6 Variable (with Tax removed)

```
Call:
lm(formula = Burglary ~ UnemploymentRate + HP + JD + Education +
    PO, data = regmodeldata)

Residuals:
    Min      1Q  Median      3Q     Max
-85.556 -31.033  -4.595  23.494 253.182

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)      -1.746e+02  7.141e+01  -2.444 0.014952 *
UnemploymentRate  1.214e+01  5.008e+00   2.425 0.015780 *
HP               -3.733e-05  1.479e-05  -2.525 0.011967 *
JD               -1.805e+00  1.491e-01 -12.107  < 2e-16 ***
Education         3.415e+00  1.001e+00   3.412 0.000711 ***
PO                2.227e-01  1.657e-02  13.435  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 45.8 on 390 degrees of freedom
Multiple R-squared:  0.4986,    Adjusted R-squared:  0.4922
F-statistic: 77.56 on 5 and 390 DF,  p-value: < 2.2e-16
```

Figure 15: Input Data with 5 significant Variables

$$\boldsymbol{Y(Burglary\ Occurrences)} = \boldsymbol{-174.6 + 12.14(UR) - 0.00003733(HP) - 1.805\,(JD) + 3.415(Edu) + 0.2227(PO)}$$

For Theft Occurrences

```
Call:
lm(formula = Theft ~ GAP + PD + UnemploymentRate + HP + JD +
    Education + Tax + PO, data = regmodeldata)

Residuals:
     Min       1Q   Median       3Q      Max
-119.931  -31.662   -3.158   21.536  292.750

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)      -5.419e+02  9.334e+01  -5.805 1.34e-08 ***
GAP              -5.314e-04  1.253e-02  -0.042  0.96618
PD                1.298e+00  1.508e-01   8.611  < 2e-16 ***
UnemploymentRate -2.972e+01  7.034e+00  -4.225 2.99e-05 ***
HP                5.104e-05  2.536e-05   2.012  0.04487 *
JD               -1.086e+00  2.614e-01  -4.153 4.04e-05 ***
Education         3.684e+00  1.201e+00   3.067  0.00231 **
Tax               9.598e-01  2.427e-01   3.955 9.10e-05 ***
PO                4.294e-01  2.287e-02  18.773  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 53.48 on 387 degrees of freedom
Multiple R-squared:  0.7374,    Adjusted R-squared:  0.7319
F-statistic: 135.8 on 8 and 387 DF,  p-value: < 2.2e-16
```

Figure 16: Input Data with 8 Variables

```
Call:
lm(formula = Theft ~ PD + UnemploymentRate + HP + JD + Education +
    Tax + PO, data = regmodeldata)

Residuals:
     Min       1Q   Median       3Q      Max
-120.258  -31.500   -3.162   21.581  292.766

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)      -5.434e+02  8.608e+01  -6.313 7.48e-10 ***
PD                1.295e+00  1.274e-01  10.166  < 2e-16 ***
UnemploymentRate -2.963e+01  6.744e+00  -4.394 1.44e-05 ***
HP                5.129e-05  2.461e-05   2.084  0.03784 *
JD               -1.093e+00  1.855e-01  -5.894 8.19e-09 ***
Education         3.687e+00  1.198e+00   3.079  0.00223 **
Tax               9.600e-01  2.424e-01   3.961 8.89e-05 ***
PO                4.292e-01  2.235e-02  19.199  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 53.41 on 388 degrees of freedom
Multiple R-squared:  0.7374,    Adjusted R-squared:  0.7326
F-statistic: 155.6 on 7 and 388 DF,  p-value: < 2.2e-16
```

Figure 17: Input Data with 7 significant Variables (with GAP removed)

Now all variables are significant with a p-value less than 0.05.

$$Y(\boldsymbol{Theft\ Occurrences}) = -543.4 + 1.295(\boldsymbol{PD}) - 29.63\ (\boldsymbol{UR}) + 0.00005129\ (\boldsymbol{HP}) - 1.093\ (\boldsymbol{JD}) + 3.687\ (\boldsymbol{Ed}) + 0.96(\boldsymbol{Tax}) + 0.4292(\boldsymbol{PO})$$

For Robbery Occurrences

```
Call:
lm(formula = Robbery ~ GAP + PD + UnemploymentRate + HP + JD +
    Education + Tax + PO, data = regmodeldata)

Residuals:
    Min      1Q  Median      3Q     Max
-37.590  -9.181  -0.835   7.682  57.096

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)      -1.643e+02  2.643e+01  -6.214 1.33e-09 ***
GAP              -5.081e-03  3.547e-03  -1.433 0.152796
PD                3.623e-01  4.270e-02   8.485 4.63e-16 ***
UnemploymentRate  8.388e+00  1.992e+00   4.212 3.16e-05 ***
HP               -3.425e-05  7.181e-06  -4.769 2.62e-06 ***
JD               -2.799e-01  7.402e-02  -3.781 0.000181 ***
Education         1.351e+00  3.401e-01   3.974 8.45e-05 ***
Tax               2.996e-01  6.872e-02   4.359 1.68e-05 ***
PO                1.425e-01  6.476e-03  21.998  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 15.14 on 387 degrees of freedom
Multiple R-squared:  0.8067,    Adjusted R-squared:  0.8027
F-statistic: 201.9 on 8 and 387 DF,  p-value: < 2.2e-16
```

Figure 18: Input Data with 8 Variables

```
Call:
lm(formula = Robbery ~ PD + UnemploymentRate + HP + JD + Education +
    Tax + PO, data = regmodeldata)

Residuals:
    Min      1Q  Median      3Q     Max
-37.089  -9.209  -0.568   7.784  57.625

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)
(Intercept)      -1.788e+02  2.444e+01  -7.316 1.48e-12 ***
PD                3.297e-01  3.617e-02   9.116  < 2e-16 ***
UnemploymentRate  9.186e+00  1.915e+00   4.797 2.30e-06 ***
HP               -3.182e-05  6.988e-06  -4.554 7.06e-06 ***
JD               -3.545e-01  5.267e-02  -6.730 6.12e-11 ***
Education         1.379e+00  3.400e-01   4.057 6.02e-05 ***
Tax               3.007e-01  6.881e-02   4.370 1.59e-05 ***
PO                1.406e-01  6.346e-03  22.147  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 15.16 on 388 degrees of freedom
Multiple R-squared:  0.8057,    Adjusted R-squared:  0.8021
F-statistic: 229.8 on 7 and 388 DF,  p-value: < 2.2e-16
```

Figure 19: Input Data with 7 significant Variables (with GAP removed)

$$\boldsymbol{Y(Robbery\ Occurrences)} = -\mathbf{178.8} + \mathbf{0.3297}\,(\mathbf{PD}) + \mathbf{9.186}\,(\mathbf{UR}) - \mathbf{0.00003182}\,(\mathbf{HP}) - \mathbf{0.3545}\,(\mathbf{JD}) + \mathbf{1.379}\,(\mathbf{Edu}) + \mathbf{0.3007}\,(\mathbf{Tax}) + \mathbf{0.1406}\,(\mathbf{PO})$$

STEPWISE REGRESSION

For Burglary

```
Step:  AIC=3034.77
Burglary ~ UnemploymentRate + HP + JD + Education + PO

                  Df Sum of Sq     RSS    AIC
<none>                         818041 3034.8
- UnemploymentRate  1     12331  830371 3038.7
- HP                1     13373  831413 3039.2
- Education         1     24424  842465 3044.4
- JD                1    307450 1125491 3159.1
- PO                1    378628 1196668 3183.4

Call:
lm(formula = Burglary ~ UnemploymentRate + HP + JD + Education +
    PO, data = regmodeldata)

Coefficients:
     (Intercept)  UnemploymentRate                HP                JD         Education
      -1.746e+02         1.214e+01        -3.733e-05        -1.805e+00         3.415e+00
              PO
       2.227e-01
```

Figure 20: Input Data and Results – Burglary

$$\boldsymbol{Y(Burglary\ Occurrences) = -174.6 + 12.14(UR) - 0.00003733(HP) - 1.805\,(JD) + 3.415(Edu) + 0.2227(PO)}$$

A unit rate person increase in the Unemployment rate will increase the number of Burglary occurrence by 12.14, for every one pound (£) increase in Median House Price it will decrease the number of Burglary occurrence by 0.00003733, an increase in Job Density will decrease the number of Burglary occurrence by 1.805, a one percent increase in Education will increase the number of Burglary occurrence by 3.415 and for every one Police Officer increase will increase the number of Burglary occurrence by 0.2227.

These results can be related to the studies by Carmichael and Ward (2000)1 that shows a positive relationship between burglary rates and male unemployment regardless of age. While according to the empirical study by Marvell and Moody (1996)2 along with 15 out of 36 previous papers examined there is a positive relationship between crime rates and police numbers.

For Robbery

```
Start:  AIC=2161.19
Robbery ~ GAP + PD + UnemploymentRate + HP + JD + Education +
    Tax + PO

                  Df Sum of Sq    RSS    AIC
<none>                         88743 2161.2
- GAP              1       471  89214 2161.3
- JD               1      3278  92021 2173.6
- Education        1      3621  92364 2175.0
- UnemploymentRate 1      4067  92810 2176.9
- Tax              1      4357  93101 2178.2
- HP               1      5216  93959 2181.8
- PD               1     16510 105253 2226.8
- PO               1    110966 199709 2480.4

Call:
lm(formula = Robbery ~ GAP + PD + UnemploymentRate + HP + JD +
    Education + Tax + PO, data = regmodeldata)

Coefficients:
     (Intercept)              GAP               PD  UnemploymentRate
      -1.643e+02       -5.081e-03        3.623e-01         8.388e+00
              HP               JD        Education              Tax
      -3.425e-05       -2.799e-01        1.351e+00        2.996e-01
              PO
       1.425e-01
```

Figure 21: Input Data and Results - Robbery

$$\boldsymbol{Y(Robbery\ Occurrences)} = -\mathbf{164.3} - \mathbf{0,005081}\ (\boldsymbol{GAP}) + \mathbf{0.3623}\ (\boldsymbol{PD}) + \mathbf{8.388}\ (\boldsymbol{UR}) - \mathbf{0.00003425}\ (\boldsymbol{HP}) - \mathbf{0.2799}\ (\boldsymbol{JD}) + \mathbf{0.2799}\ (\boldsymbol{JD}) + \mathbf{1.351}\ (\boldsymbol{Edu}) + \mathbf{0.2996}\ (\boldsymbol{Tax}) + \mathbf{0.1425}\ (\boldsymbol{PO})$$

For every one pound (£) increase in Gross Annual Pay will result in to decrease in the Robbery Occurrence by 0.005081. An increase in Population Density will increase the Robbery occurrence by 0.3623. A unit rate person increase in the Unemployment rate will increase the occurrence Robbery by 8.388, for every one pound (£) increase in Median House Price will decrease the occurrence of Robbery by 0.00003425, an increase in Job Density will decrease the occurrence of Robbery by 0.2799, a one percent increase in Education will increase the occurrence of Robbery by 1.351, for every one pound (£) increase in House Tax will increase the occurrence on Robbery by 0.2996 and for every one Police Officer increase will increase the occurrence of Robbery by 0.1425.

According to Gumus (2003) per capita income is one of the crucial determinants of crime rates which is not aligned with the results above. According to Gumus (2003) as well, population one of the crucial determinants of crime rate which can be related to the finding above. There is a positive relationship between the youth unemployment to criminal damages and robbery rates which can be related to the finding found in the Robbery regression. Similar to the Burglary model, it is found a positive relationship between crime rates and police numbers that can be related to the 15 out of 36 examined papers along with Marvell and Moondy (1996) empirical study.

For Theft

```
Start:  AIC=3160.49
Theft ~ GAP + PD + UnemploymentRate + HP + JD + Education + Tax +
    PO

                   Df Sum of Sq     RSS    AIC
- GAP               1         5 1106818 3158.5
<none>                         1106813 3160.5
- HP                1     11582 1118395 3162.6
- Education         1     26908 1133721 3168.0
- Tax               1     44735 1151549 3174.2
- JD                1     49331 1156144 3175.8
- UnemploymentRate  1     51048 1157861 3176.3
- PD                1    212044 1318857 3227.9
- PO                1   1007900 2114714 3414.9

Step:  AIC=3158.49
Theft ~ PD + UnemploymentRate + HP + JD + Education + Tax + PO

                   Df Sum of Sq     RSS    AIC
<none>                         1106818 3158.5
- HP                1     12386 1119205 3160.9
- Education         1     27039 1133857 3166.0
- Tax               1     44753 1151571 3172.2
- UnemploymentRate  1     55071 1161889 3175.7
- JD                1     99098 1205916 3190.4
- PD                1    294802 1401620 3250.0
- PO                1   1051432 2158250 3420.9

Call:
lm(formula = Theft ~ PD + UnemploymentRate + HP + JD + Education +
    Tax + PO, data = regmodeldata)

Coefficients:
     (Intercept)                PD  UnemploymentRate                HP                JD
      -5.434e+02         1.295e+00        -2.963e+01         5.129e-05        -1.093e+00
       Education               Tax                PO
       3.687e+00         9.600e-01         4.292e-01
```

Figure 22: Input Data and Results - Theft

$$\boldsymbol{Y(Theft\ Occurrences)} = -543.4 + 1.295(PD) - 29.63\ (UR) + 0.00005129\ (HP) - 1.093\ (JD) + 3.687\ (Ed) + 0.96(Tax) + 0.4292(PO)$$

An increase in Population Density will increase the occurrence of Theft by 1.295, a unit rate person increase in the Unemployment rate will decrease the occurrence of Theft by 29.63, for every one pound (£) increase in Median House Price will increase the occurrence of Theft by 0.00005129, an increase in Job Density will decrease the occurrence of Theft by 1.093, a one

percent increase in Education will increase the occurrence of Theft by 3.687, for every one pound (£) increase in House Tax will increase the occurrence of Theft by 0.96 and for every one Police Officer increase will increase the occurrence of Theft by 0.4292.

According to findings in Carmichael and Ward (2002) studies, shows a positive relationship between adult unemployment and theft. The results from the finding also show population density can be related to the study by Gumus that claimed that population is one of the crucial determinants of crime rates (2003). Lastly, from the empirical study by Marvell and moody (1996) claims that there is a positive relationship between the crime rate and police numbers.

CONCLUSION – SUMMARY

The research questions have been validated based on visualizing spatial data from the heat maps. In summary, the highest Burglary occurrences are in Barnet with 3,819 cases, whereas in Westminster, the highest occurrences for both Theft and Robbery are recorded with 6,203 and 1,555 cases respectively.

Based on the modelling analysis, Unemployment Rate, Education, and Police Officers have a positive correlation with the number of Burglary occurrences, whereas Median House Price and Job Density also correlate with the number of Burglary occurrences but with very little impact. Meanwhile, a positive correlation is observed between the number of Robbery occurrences with Population Density, Unemployment Rate, Education, House Tax, and Police Officers, but very little impact is seen with Gross Annual Pay, Median House Price, and Job Density. In another context, there is a positive correlation between the number of Theft occurrences with Population Density, Median House Price, Education, House Tax, and Police Officers, but a minimal impact is observed with Unemployment Rate and Job Density. Out of all variables, Unemployment Rate is the most influential factor of Burglary and Theft occurrences with coefficients of 12.14 and 8.39 respectively, whereas Education is most influential to Theft occurrences with a coefficient of 3.69.

REFERENCES

Becker, G. S. (1968). Crime and punishment: An economic approach. Journal of Political Economy 76(2), 169–217.

Carmichael, Fiona & Ward, Robert. (2000). Youth Unemployment and Crime in the English Regions and Wales.. Applied Economics. 32. 559-71. 10.1080/000368400322462.

GAVİRİA, ALEJANDRO and CARMEN PAGÉS. (2002). Patterns of Crime Victimization in Latin American Cities, Journal of Development Economics, 67, 181-203,

MATHUR, VİJAY K. (1977). Economics of crime: An investigation of the deterrent hypothesis for urban areas, The Review of Economics and Statistics.

STEVANS, L. K. (1983). Urban Property Crime and the distribution of Income, Dissertation, Oklahoma State University, May.

MEERA, AHAMED KAMEEL and JAYAKUMAR, MALİYAKAL D. (1995). Determinants of crime in a developing country: a regression model, Applied Economics, 27, 455-460.

MASİH, ABUL M.M. and MASİH RUMİ. (1996). Temporal causality and the dynamics of different categories of crime and their socioeconomic determinants: evidence from Australia, Applied Economics, 28, 1093-1104.

Machin, S. (n.d.). Stephen Machin. Retrieved from http://jhr.uwpress.org/lookup/doi/10.3368/jhr.XXXIX.4.958.

GAVİRİA, ALEJANDRO and CARMEN PAGÉS. (2002). Patterns of Crime Victimization in Latin American Cities, Journal of Development Economics, 67, 181-203,

Ogburn, W. (1935). Factors in the Variation of Crime Among Cities. Journal of the American Statistical Association, 30(189), 12-34. doi:10.2307/2278526

GUMUS, ERDAL, Crime in Urban Areas: An Empirical Investigation. Akdeniz I.I.B.F. Dergisi, Vol. 4, No. 7, pp. 98-109, May 2004

Wang, Hongjian & Kifer, Daniel & Graif, Corina & Li, Zhenhui. (2016). Crime Rate Inference with Big Data. 10.1145/2939672.2939736.

Marvell, T.B. and Moody, C.E. (1996) Specification Problems, Police Levels, and Crime Rates, Criminology 34(4): 609-646.