“Spatial Analysis of Crime Patterns in London”

CEGE0097: Spatial Analysis and Geocomputation

Ng Jia Wen & Junju Ng

Contents

1 Introduction **3**

2 Data description **5**

3 Exploratory Spatial Data Analysis **6**

4 Methodology **10**
4.1 Methods in Analysing Violence and Sexual Offences 10
4.2 Methods in Analysing Vehicle Crime 11

5 Results and discussion **13**
5.1 Analysing Violence and Sexual Offences 13
5.1.1 Analysing Violence and Sexual Offences in City of London 001F 14
5.2 Analysing Vehicle Crime 21
5.2.1 Spatial autocorrelation of vehicle crime 21
5.2.2 Reasons for vehicle crime rates 23
5.2.3 Discussion 29

6 Overall discussion and conclusions **31**

7 References **32**

1 Introduction

Crime (and its prevention and responses) is associated with significant social and economic costs. In England and Wales, the total cost of crimes against individuals and businesses is £50.1 billion and £8.7 billion respectively in 2015/16 alone (Brand and Price 2014). A significant portion of these costs are incurred in London, given that about 20% of crimes in England and Wales occurs in London (Mayor of London 2016).

Consequently, much work has been done to understand the occurrence of crime, and to predict crime. To date, the spatial concentration of crime in cities is well-established in the literature (Malleson and Andresen 2016). (Sherman, Gartin, and Buerger 1989)'s seminal analysis of predatory crime in Minneapolis revealed that 50% of police calls come from 3% of street segments, indicating that crime hotspots, down to the street level, are present within cities. Recent works have thus focussed on trying to improve spatial analysis techniques for crime. Specifically, (Chainey, Tompson, and Uhlig 2008) examined the use of different hotspot mapping techniques to predict spatial patterns of crime in Camden and Islington local authority district areas in London. Other researchers have been working on developing techniques to examine spatial-temporal patterns of crime (Cheng and Williams 2012), in order to understand how crime develops and evolves to improve crime prediction.

Given that a significant portion of crimes occurs in London and the occurrence of crime is non-random, this study will therefore examine spatial crime patterns in London. Specifically, this study aims to answer two main questions: 1) where do crime types occur? 2) what factors are associated with the occurrence of different types of crime?

Study area

London is a thriving metropolis in the United Kingdom with a population of 8.8 million. Beyond the large residential population, London also attracts a huge number of visitors, with more than 56 million overnight stays from tourists in 2016 (Mayor of London 2016).

London consists of 33 local government districts (32 London Boroughs and the City of London). The City of London Police is the police force responsible for law enforcement within the City of London while the Metropolitan Police Service is reponsible for policing the Greater London region, excluding the City of London. Greater London can be split into 4835 geographic regions known as lower layer super output areas(LSOAs) for the purposes of reporting small-area statistics. These LSOAs will form the basis of the analysis.


Enfield
Barnet
Harrow
Haringey
Waltham Forest
Redbridge
Havering
Hillingdon
Brent
Camden
Hackney
Barking and Dagenham
Newham
Ealing
Hammersmith and Fulham
Westminster
Tower Hamlets
Kensington and Chelsea
Southwark
Greenwich
Hounslow
Richmond upon Thames
Wandsworth
Lambeth
Bexley
Lewisham
Merton
Kingston upon Thames
Bromley
Sutton
Croydon

2 Data description

Official geocoded crime data from London is required to analyse spatial patterns of crime. Crime data in London for a one-year period (1 January 2017 to 31 December 2017) from the Metropolitan Police and the City of London Police was used. These data are taken from the Metropolitan Police and the City of London Police because they are the two police forces covering the London and Greater London area. The data was obtained via the police.uk website as a CSV file. Whilst most crimes are geocoded, some reported crimes do not have a location and are therefore excluded from the analysis. The effect on excluding crimes without locations from the analysis is anticipated to be minimal, as they only form a small proportion of reported crimes (5.3% for the City of London Police, and 1.2% for the Metropolitan Police).

The geocoded crime data are categorised into 14 crime types: anti-social behaviour, bicycle theft, burglary, criminal damage and arson, drugs, other crime, other theft, possession of weapons, public order, robbery, shoplifting, theft from a person, vehicle crime, and violence and sexual offences. Geomasking techniques are applied to the data to reduce their spatial accuracy for privacy purposes. Specifically, each crime is mapped to its nearest map point on the master list of map points kept by the Home Office (Home Office 2019). However, analysis by (Tompson et al. 2015) reveals that at the lower super output area (LSOA) level, 85% of areas exhibited no statistically significant difference between the masked and raw crime data. Hence, analysis of patterns in crime will predominantly be conducted at the LSOA level in this study.

2011 census data at the LSOA level was also downloaded from the London Datastore. This is to obtain information about the total resident population in each London LSOA to calculate crime rates in each LSOA.

In addition, industrial landuse areas are obtained from OpenStreetMap using the overpass API as a polygon dataset. This is to observe if trends between the occurrence of crimes and industrial areas exist. The information from OpenStreetMap is crowd-sourced, and not created by geospatial experts. Thus, some inaccuracies may be present in the dataset (Basiri et al. 2016). However, corroboration of the industrial landuse polygons with aerial imagery and street imagery from Google Maps shows that the data is generally accurate, and therefore suitable for broader-scale exploratory analysis.

A point dataset of the location of entertainment outlets (defined as pubs, bars, and restaurants) in London is also obtained from OpenStreetMap. Further verification of the location of entertainment outlets also suggests that the dataset is broadly up-to-date.

3 Exploratory Spatial Data Analysis

An exploratory analysis of the crime data will be conducted in this section.

A total of 1,048,712 crimes were reported in 2017. Of these crimes, 1,035,826 (98.7%) have a location. As discussed previously, only the geocoded crimes will be used in the analysis.

Variations in Crime by Type

The barplot below depicts the frequency of crimes by crime type in London in 2017 (Figure 1). Over 40% of reported crimes in London come from 2 categories - anti-social behaviour (22.0%) and violence and sexual offences (20.7%). The next two most significant crime types making up 20% of crimes are other theft (10.5%) and vehicle crime (10.0%). 36.9% of crimes fall under the remaining 10 crime types. As such, different types of crimes are reported at different frequencies, with antisocial behaviour and violence and sexual offences being the most commonly reported crimes, followed by vehicle crime and other theft.

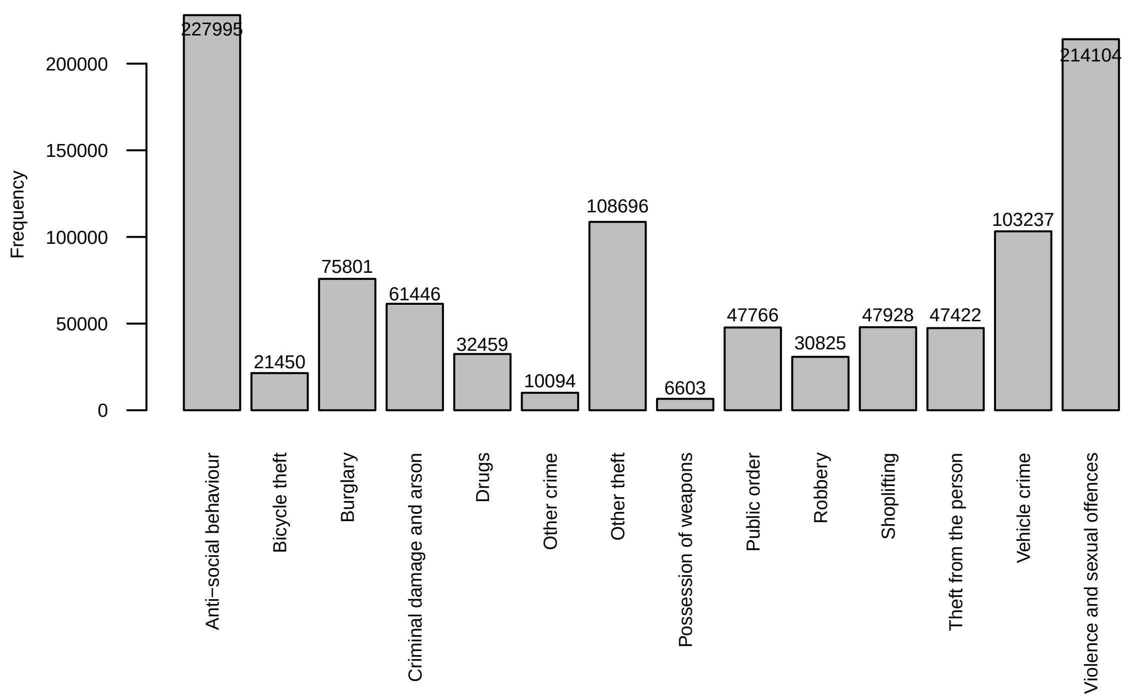


Figure 1: 2017 London crime counts by crime type

Variations in Crime by Type and Month

A line graph of crime count by type and month (Figure 2) is plotted to examine temporal patterns in crime in London across the year. From Figure 2, it is evident that the frequency of some crime types varies throughout the year. For instance, reports of anti-social behaviour appear to peak in July and August. Burglary reports appear to increase between October and January, while bicycle theft reports appear to increase between May and October).

Table 1: Crime Count Statistics by LSOA

	Value
Minimum	1.0
1st Quartile	64.0
Median	127.0
Mean	175.3
3rd Quartile	211.0
Maximum	7329.0

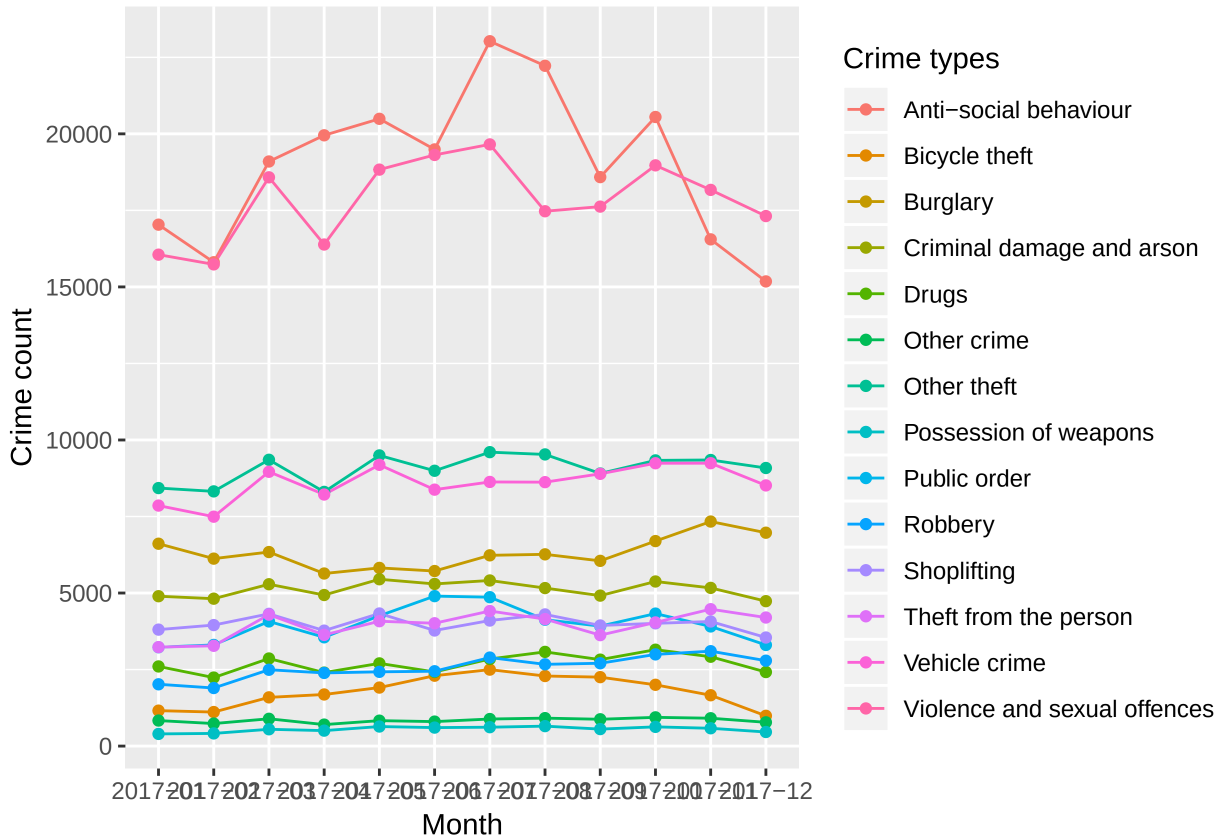


Figure 2: 2017 London crime count by type and month

Variations in Crime by LSOA

Crime counts are also examined by LSOA, in order to study the distribution of crime in London. From Table 1, it is evident that crime counts vary significantly between LSOAs. Few LSOAs have extremely high crime counts, while 50% of LSOAs have crime counts between 64 and 211. Consequently, the jenks classification scheme will be used to map the distribution of crime counts across London, to minimise the variation within each class but highlight differences between classes.

Crime count maps are plotted to visualise the spatial distribution of crime across London. The crime count map (Figure 3) illustrates the total number of reported crimes in each LSOA in 2017. By visual inspection, crime counts are highest in central London, although some regions with higher crime counts are present along the river Thames in East London, the Lea Valley and in Hillingdon.

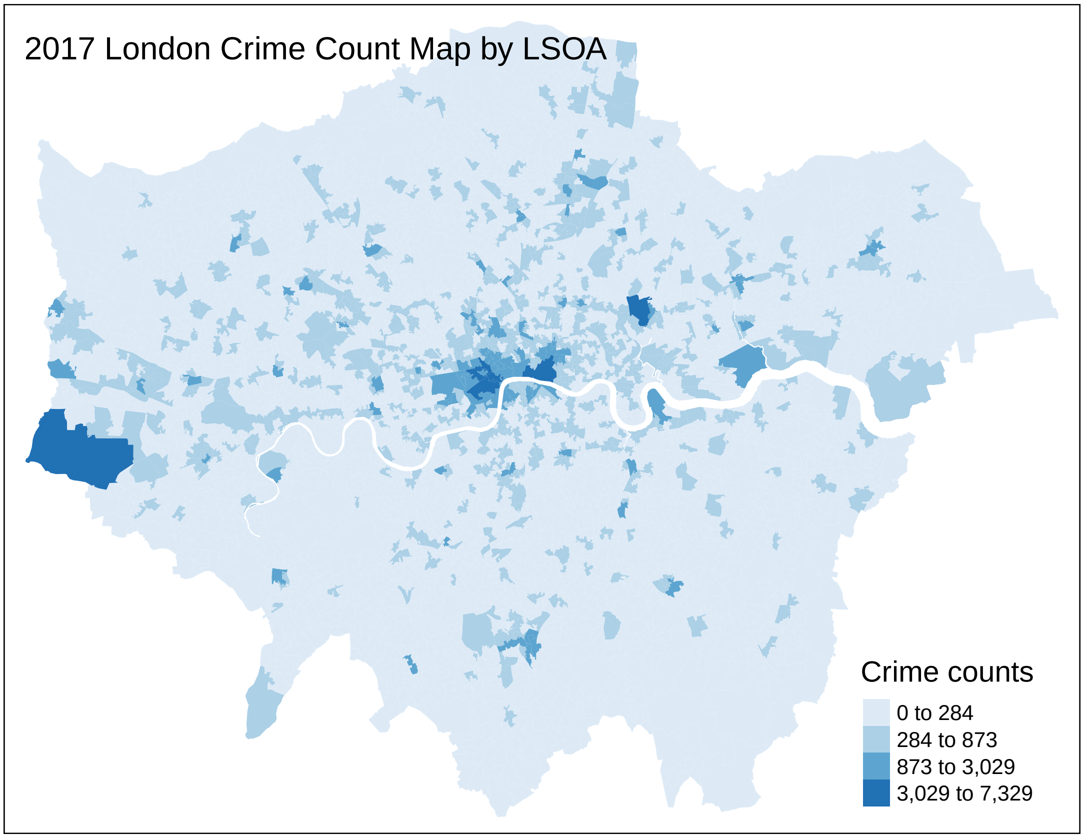
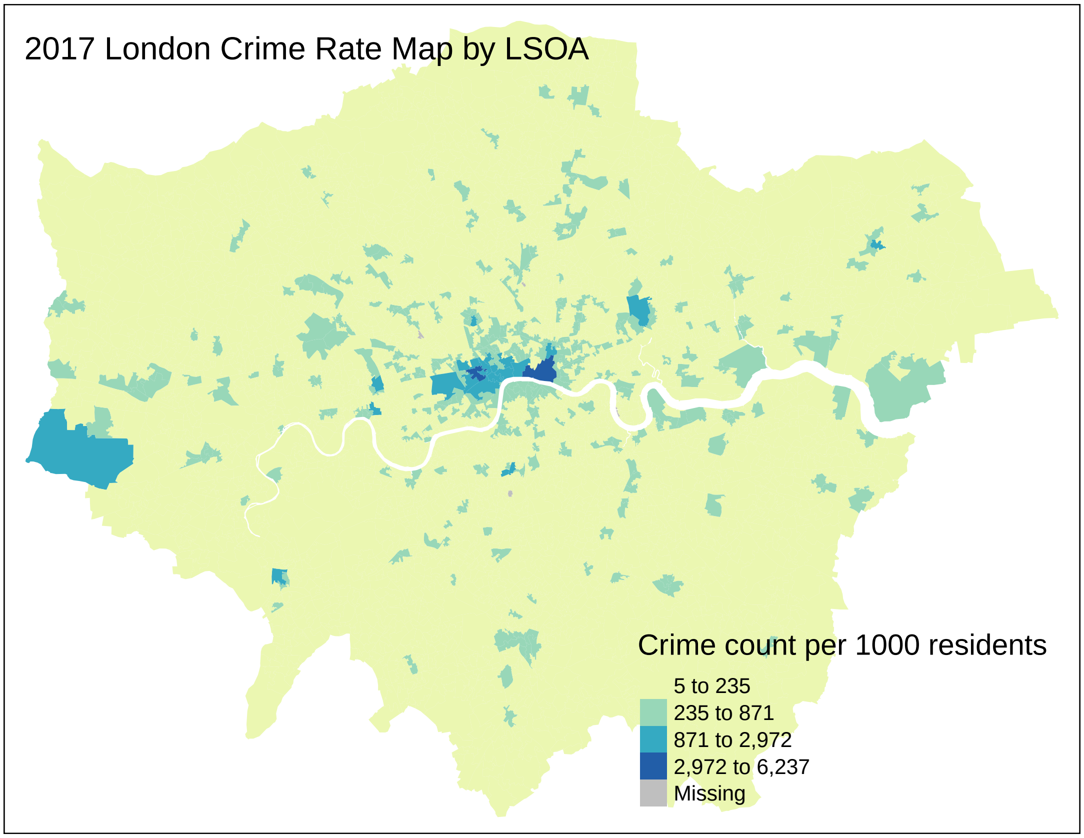


Figure 3: 2017 London crime count map by LSOA

A crime rate map (Figure **??**) is also plotted to take into account population size in each LSOA. This is derived by taking the crime count divided by the resident population in the corresponding LSOA and then multipled by 1000 (i.e. crime count per 1000 residents). These crime rate maps allow for the assessment of the risk of crime (P. L. Brantingham and Brantingham 1997). However, it is to be noted that resident populations are used to calculate crime rate. Thus, regions with a low resident population but high pedestrian population (e.g. commercial areas) would have an inflated crime rate.

Visually, although there are some differences between the crime rate and crime count maps, the regions with the highest crime counts are also the regions with the highest crime rates. However, a number of regions outside of central London with slightly higher crime counts (235 to 871) have lower-than-expected crime rates. This indicates that residential population does not have a significant influence on crime counts, except when crime counts are slightly elevated and occur in LSOAs outside central London (presumably with a higher residential population).


2017 London Crime Rate Map by LSOA
Crime count per 1000 residents
5 to 235
235 to 871
871 to 2,972
2,972 to 6,237
Missing

4 Methodology

Based on our exploratory data analysis, different crime types have different counts and spatial variations. As such, we have decided to analyse two different crime types in isolation: Violence and Sexual Offences (VSO), and Vehicle crime.

VSO is one of the most commonly reported, as well as the more severe crimes that bring about damage to people and property. Vehicle crime is chosen due to its unique nature in that it can be committed without direct contact with the victim, unlike violence and sexual offences which often requires direct contact with the victim, thus making it an interesting constrast to study.

As the two crime types have its own spatial pattern and are likely to be caused by different factors, the methodology that we undertake will be different for each of them.

4.1 Methods in Analysing Violence and Sexual Offences

A spatial point pattern analysis was done on the City of London 001F (COL001F) LSOA.

Two hypotheses were made in the process of analysing VSO crime patterns in the LSOA:

`H1. VSO crimes are randomly distributed in COL001F`

`H2. VSO crime locations positively correlate with the locations of entertainment outlets`

The reason for these hypotheses will be made clearer in section 5.1.

Here is the methodology employed to investigate the hypotheses:

1. Test H1 with Quadrat Counting
   - If homogenous, find out if there is clustering with Ripley's K Function.
     - If there is clustering, H1 is rejected.
     - If not, H1 is accepted.
   - If non-homogenous, use Kernel Density Estimation to identify areas with high intensity.
     - H1 is rejected.
2. Test H2 with Inhomogenous K cross function
   - If Inhomogenous K cross function reveals positive correlation, H2 is accepted.
   - If Inhomogenous K cross function reveals no correlation, H2 is rejected.
   - If Inhomogenous K cross function reveals negative correlation, H2 is rejected.

The following statistical methods were used in the process:

1. Quadrat Counting

   To check for homogeneity is to determine whether regions of equal area contain roughly equal number of points.

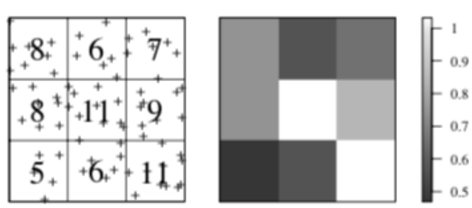


Figure 4: Picture 1 shows Quadrat counting for Swedish Pines data. Left: Quadrat counts. Right intensity. (Baddeley, Rubak, and Turner 2015)

   In quadrat counting, the observation window W is divided into subregions $B_1, ..., B_m$ called quadrats. The numbers of points falling in each quadrat is counted. If the spatial point process is homogeneous, the intensity for each quadrat should be equal 'on average'– the expected count in each quadrat is proportional to the area of the quadrat. Otherwise, the spatial point process is inhomogeneous. (Baddeley, Rubak, and Turner 2015)

2. Chi-sq test

   The Chi-sq test is used to determine whether there is a significance difference between the observed number points and the expected number of points in each quadrats, where chi-squared can be computed via:

   $\chi^2 = \sum_{i=1}^{m} \frac{(k_i - \bar{\lambda})^2}{\bar{\lambda}}$

   If there is a significant difference, then we can reject the null hypothesis that the spatial point process is homogeneous as it is not behaving in the way it is expected to behave.

3. Kernel Density Estimation

   Kernel Density Estimation (KDE) is a method for estimating the probability density function of a random variable. It can be used to extract the relative likelihood of obtaining a particular value on a random drawn. In spatial analysis, KDE can be used to estimate how the density of a point process varies across space. (Haworth 2018)

   A KDE can be written as: $\hat{f}(x, y) = \frac{1}{nh^2} \sum_{i=1}^{n} k_s(\frac{x - x_i}{h_s} \frac{y - y_i}{h_s})$

   where n is the number of points; x_i and y_i are the coordinates of the point i, i = 1,2, … , n; h_s is the bandwidth of a spatial kernel k_s.

4. Inhomogenous K cross function

   The inhomogenous K cross function measures whether 2 spatial point patterns are correlated, with the assumption that the spatial point patterns are inhomogenous.

   $K_{ij}(r) = \frac{1}{\lambda_j} E[t(u, r, X^j | u \epsilon X^i)]$

   For any pair of types i and j, the multitype K-function $K_{ij}(r)$, also called the bivariate or cross-type K-function, is the expected number of points of type j lying within a distance r of a typical point of type i, standardised by dividing by the intensity of points of type j, adjusted for spatially varying intensity. (Baddeley, Rubak, and Turner 2015)

5. Monte-Carlo Simulation

   The monte-carlo simulation is used to generate 100 simulations of a point pattern according to a specified probability distribution, in order to test whether the result from a test is statistically significant.

4.2 Methods in Analysing Vehicle Crime

Visual inspection of the maps suggests that vehicle crime rates appear to concentrate in regions. Thus, the following hypothesis was generated:

```
H3. Vehicle crimes are spatially correlated
```

```
H4. Vehicle crime rates are correlated with industrial areas
```

The reason for these hypotheses will be made clearer in section 5.2.

The Moran's I (Moran 1950) was used to test for spatial autocorrelation. It is a weighted correlation coefficient that measures if the spatial data sampled at adjacent locations are correlated with one another. It measures a single average spatial autocorrelation across the entire area. The Moran's I is calculated using the following equation:

$$I = \frac{n}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_i (x_i - \bar{x})^2}$$

Where n is the number of LSOAs, w_{ij} is the element of a spatial weight matrix W giving the spatial weight between LSOAs i and j,
x_i and x_j are the vehicle crime rates measured at LSOAs i and j respectively and $i, j = 1, 2, ..., n$,
$\bar{x}$ is the mean of x (i.e. the mean LSOA vehicle crime rate)

A expectation of I under a null hypothesis is $\frac{-1}{n-1}$. Values greater than that indicate a positive spatial autocorrelation, while values smaller than that indicate a negative spatial autocorrelation.

A local Moran's I (Anselin 1995), I_i, was subsequently calculated for each LSOA, to determine the spatial autocorrelation between a given LSOA and its adjacent LSOAs. This is to observe if variations in spatial autocorrelation exist. The local Moran's I can be calculated using the following equation:

$$I_i = \frac{z_i}{m_2} \sum_j w_{ij} z_j$$

Where $z_i = x_i - \bar{x}$, $z_j = x_j - \bar{x}$, $m_2 = \frac{1}{n} \sum_i z_i^2$, $\bar{x}$ is the mean vehicle crime rate, x, at n LSOAs and $i = 1, 2, ..., n$

The expectation of I_i is $\frac{-w_i}{(n-1)}$, where $-w_i = \sum_j w_{ij}$. Local hotspots of positive or negative autocorrelation will be present if vehicle crime rates are spatially heterogeneous. A bonferroni adjustment was then applied to the results, to identify Moran's I values that are statistically significant at the 95% confidence interval.

The areas with high vehicle crime rates and which are statistically significant high-high clusters are then analysed using Google Map imagery. Further analysis is also conducted with industrial landuse data from OpenStreetMap.

5 Results and discussion

Two crime types, violence and sexual offences, and vehicle crime have been used for further analysis. Vehicle crime refers to any theft, tampering, or interference with motor vehicles (Home Office 2019).

As the nature of violence and sexual offences, and vehicle crime differs, examination of the spatial distribution of the two crime types in London will allow for better understanding of the factors influencing the occurrence of these two crimes.

5.1 Analysing Violence and Sexual Offences

By plotting the occurrence of violence and sexual offences per 1000 people on a chloropeth, we get the following map below (Figure 5):

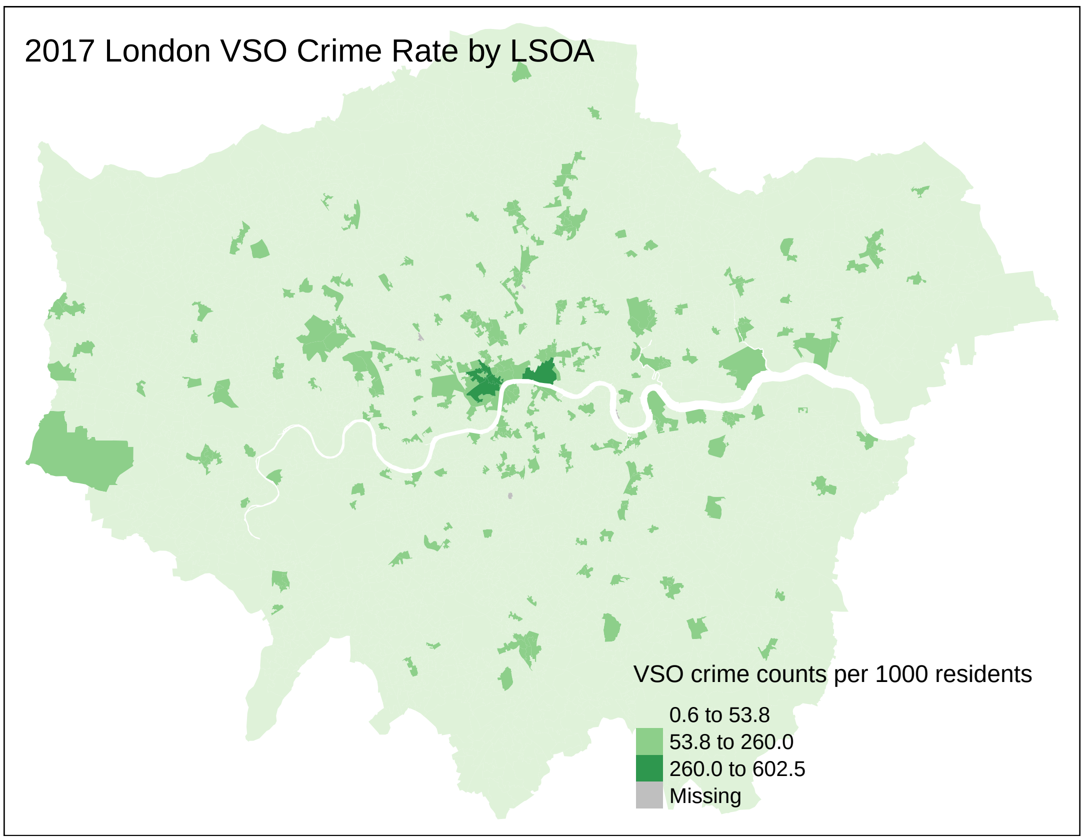


Figure 5: 2017 London crime rate map by LSOA

This map shows us that VSO crimes are mainly concentrated in central london, with a few LSOAs having mid-range crime density dispersed across london.

Ranking the LSOAs by crime density, we obtain the following rank table (Table 2), showing the top 20 LSOAs with the highest crime rates (crime count per 1000 people):

We select the City of London 001F (COL001F) to do our analysis as it has the highest crime density. We are also interested to investigate why this is so.

Table 2: 20 LSOAs with the highest crime rates

LSOA name	Crime Rate
City of London 001F	602.4904
Westminster 013E	594.0426
Westminster 018A	520.2840
Westminster 018C	342.8899
Westminster 013B	304.1843
Camden 021A	259.9595
Hackney 027G	254.5341
Lambeth 011B	244.0074
Kingston upon Thames 009C	214.7472
Hillingdon 031A	195.8538
Havering 013C	193.6538
Islington 004B	186.2427
Newham 013G	184.6803
Croydon 027B	183.4756
Sutton 012D	183.0226
Greenwich 004E	173.0580
Westminster 013F	165.9919
Westminster 018B	164.5708
Hammersmith and Fulham 004A	162.2313
Greenwich 036B	161.7647

5.1.1 Analysing Violence and Sexual Offences in City of London 001F

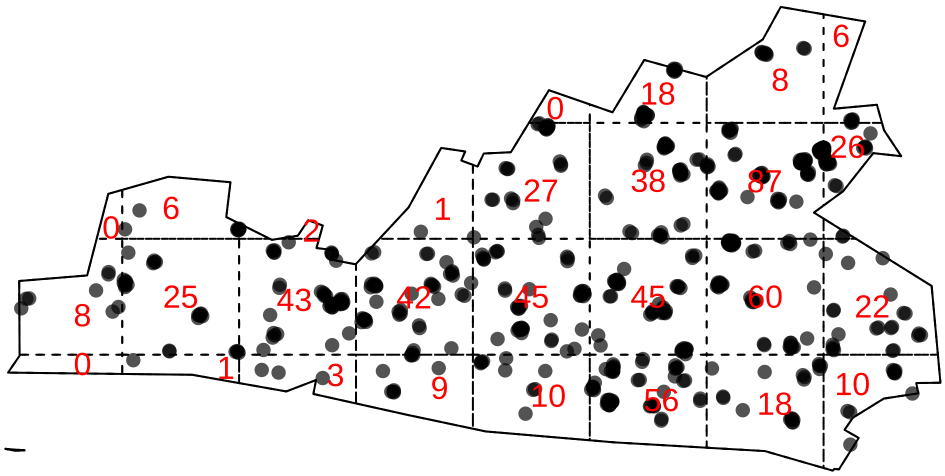
```
## OGR data source with driver: ESRI Shapefile
## Source: "C:\Users\Junju\Desktop\Masters\Term 1\Spatial Analysis and Geocomputation\Geocomputation\bou
## with 4303 features
## It has 8 fields
```

By plotting the spatial distribution of the crimes on a map (Figure ??), we can make our first hypothesis:

H1. VSO crimes are randomly distributed in the City of London 001F LSOA

To test for H1, we employ the use of quadrat counting. If the spatial pattern is homogeneous, each quadrat should have roughly the same amount of intensity (points per area) (Figure ??).

COL001F_ppp

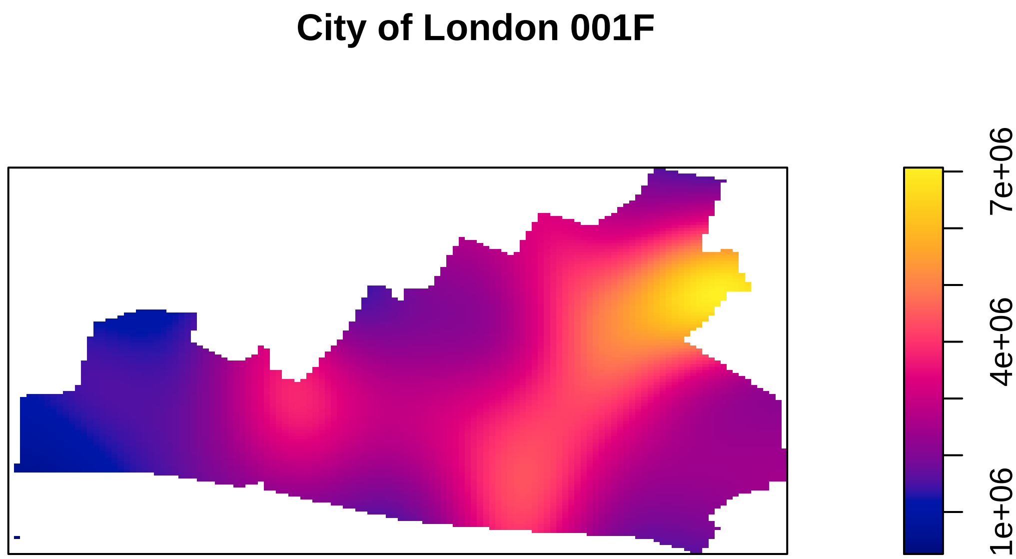


A chi-sq test gives the following results:

```
##
##  Chi-squared test of CSR using quadrat counts
##  Pearson X2 statistic
##
## data:  COL001F_ppp
## X2 = 246.52, df = 27, p-value < 2.2e-16
## alternative hypothesis: two.sided
##
## Quadrats: 28 tiles (irregular windows)
```

As the p-value is less than 0.05, it rejects the null hypothesis that the crimes are homogeneous.

Next, we use the Kernel Density Estimation to visualise the different intensities of the VSO crime on the map.



From the map, clusters of VSO crime seem to occur North-East of City of London 001F. To understand why this is so, we can study the area in more detail (red box).

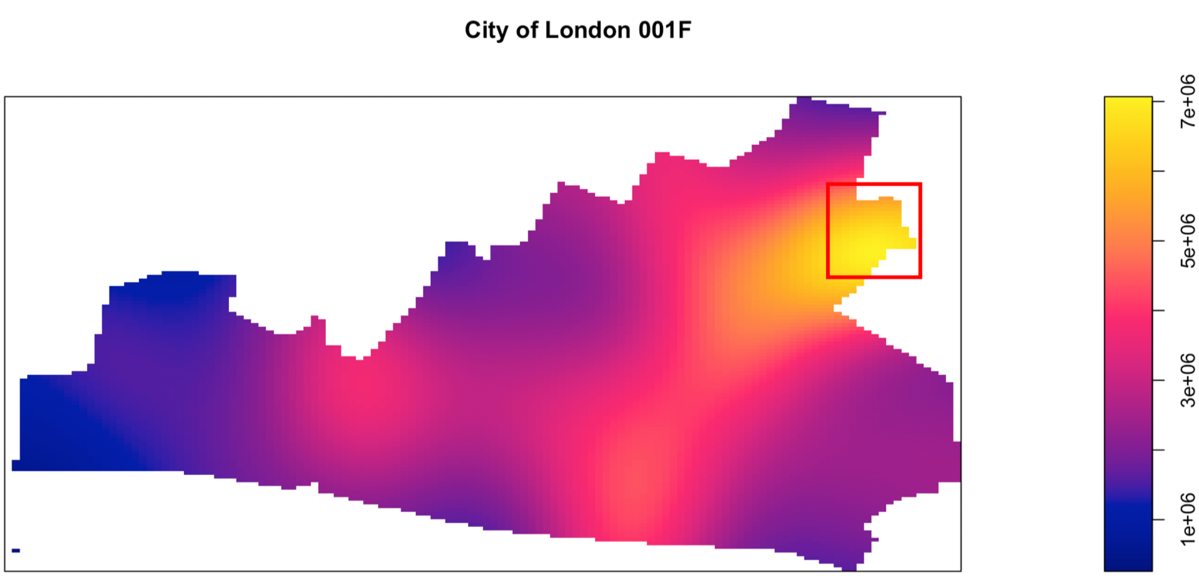


Figure 6:

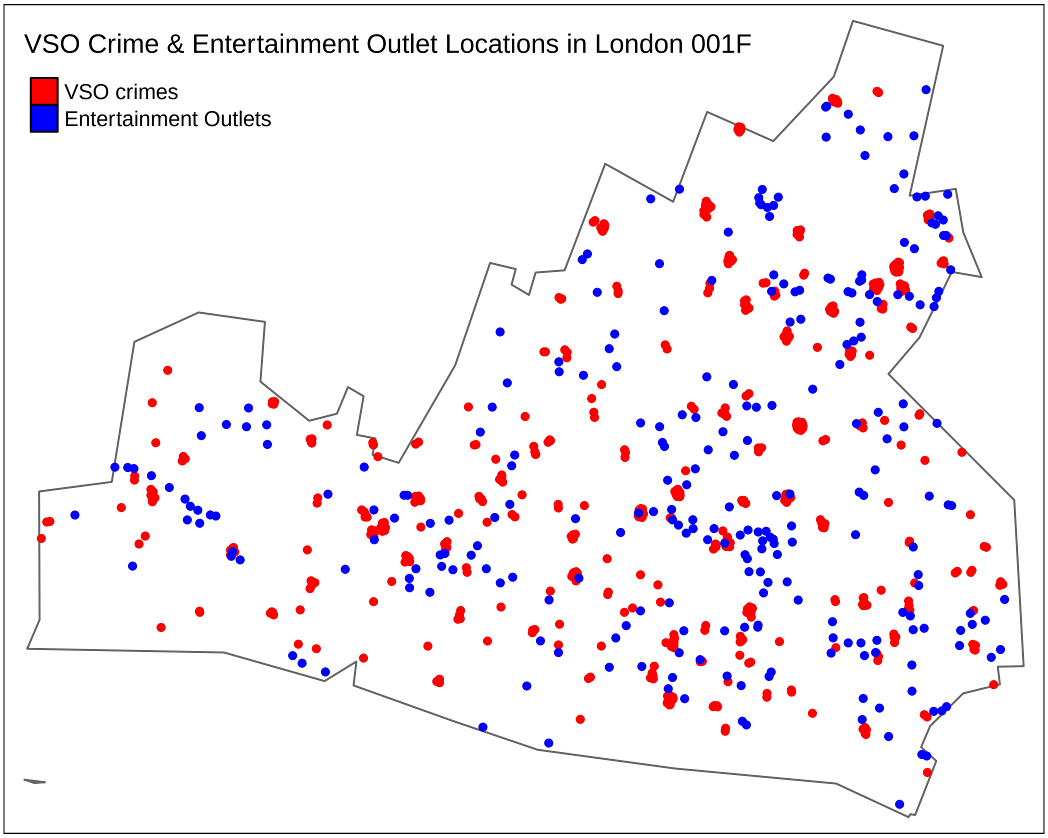
The street level view reveals a hotspot of 67 VSO crimes in the region.

Figure 7:

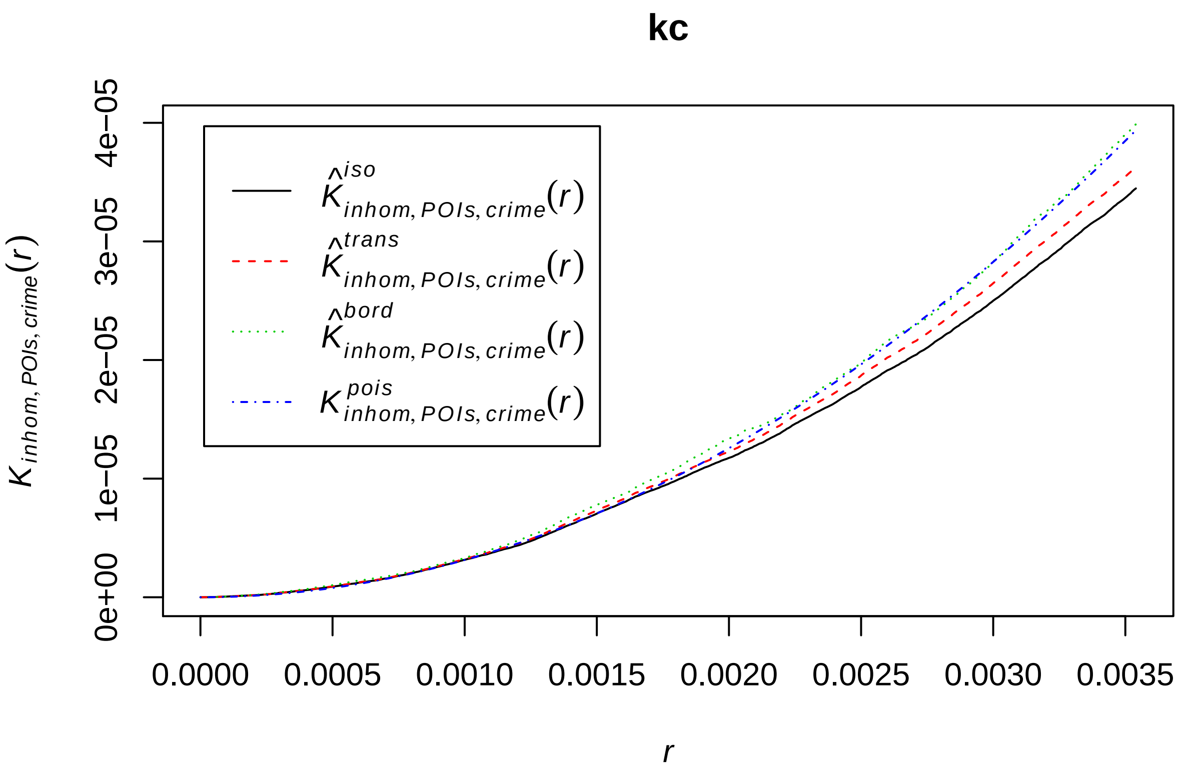
Based on visual inspection alone, the offences seem to congregate around public entertainment outlets (orange markers above). Here, we construct a second hypothesis.

H2: The location of VSO crimes correlate with the location of entertainment outlets.

We define entertainment outlets here to be: restaurants, bars and pubs. With open street map, we can extract the locations of these outlets and visualise them on a map, alongside the VSO crimes:



Visually, the map seems to agree with our hypothesis. VSO crime locations seem to correlate with locations of entertainment outlets. We can perform an inhomogeneous kcross function to determine their spatial relationship.



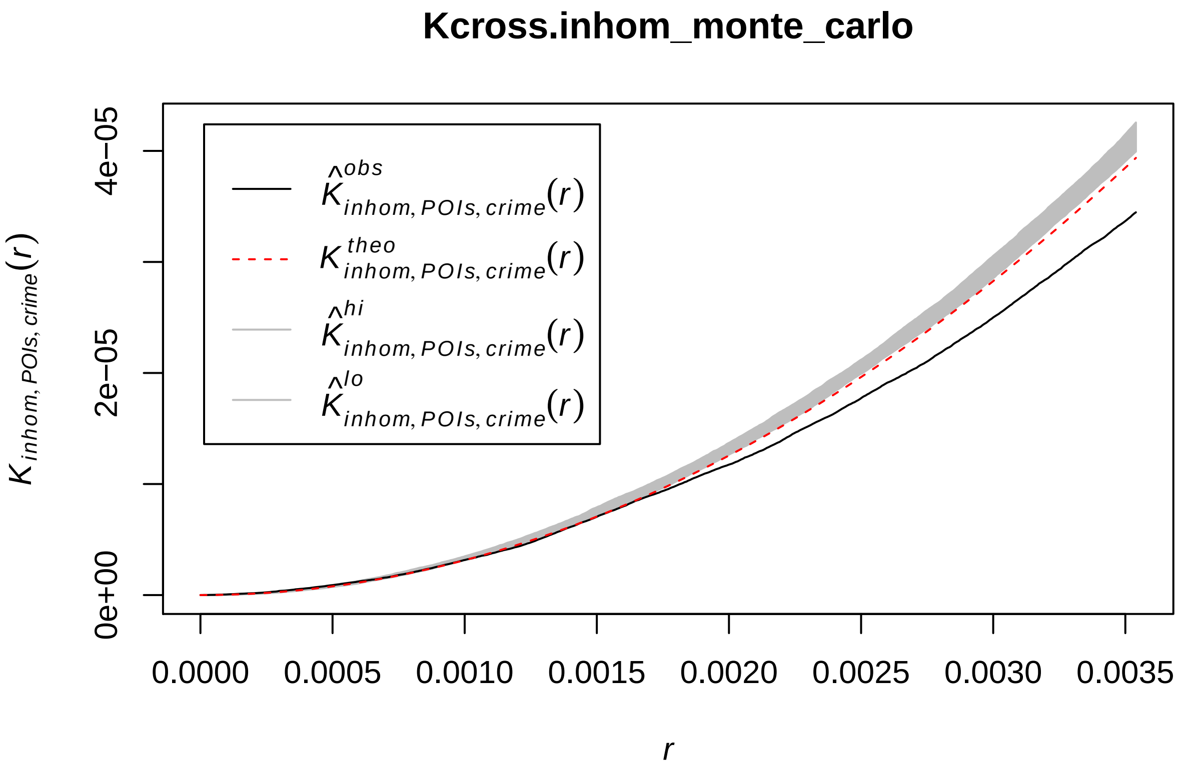
The inhomogeneous kcross function tells us that there is in fact, a repulsive spatial relationship between the location of VSO crimes and the location of entertainment outlets (lines below the blue line) – VSO crimes do not cluster around entertainment outlets.

Running a monte-carlo simulation tells us whether this observation is statistically significant:

```
## Generating 100 simulations of CSR  ...
## 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 2
## 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 
## 77, 78, 79, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99,  100.
##
## Done.
```

Table 3: Points on the streets

points_on_roads	Freq
FALSE	235
TRUE	381



The observed *Kcrossinhom* line lies below and outside of the acceptance region. This means that the null hypothesis that the 2 variables are independent is rejected. The results also indicate a dispersive relationship between the 2 variables.

This finding is surprising because we expected the 2 variables to be attractive. We may be able to explain this finding though, that the inhomogeneous Kcross function accounted for the the inhomogeneity in the spatial point pattern, brought about by private buildings and parks (random empty spaces on the map).

So although the 2 variables seem to be occurring in the 'same place', they are in fact brought together because of the spatial structure of the environment. The presence of parks and buildings 'forces' the points to go along in a particular direction. This means if we plot other variables on the map as well (for example: other crimes or locations of supermarkets, etc.), their spatial locations may appear similar too – avoiding the spaces that can't be occupied (parks, private spaces and buildings) and occupying the spaces that can be.
Just out of curiosity, we can do simple tabulation of the number of VSO crimes that happen on the streets.

The results show that **62%** of VSO crimes happen on the streets. This may support our hypothesis that the spatial pattern of the VSO crimes is heavily influenced by the spatial structure of the environment.

The repulsive relationship between VSO crimes and entertainment outlets (bars, pubs and restaurants) is still surprising though and we are uncertain why this is so. Perhaps these are places people usually go to with a group of friends of family and hence it is more difficult for perpatrators to commit VSOs? These are also places with high visibility therefore perpatrators do not dare to commit VSOs? VSOs at bars and pubs are also likely to go unreported, because of the nature of their environments? Or perhaps it could be because of a related variable that causes this repulsive relationship? Clearly, more analysis and research is needed in this area.

5.2 Analysing Vehicle Crime

Figure 8 shows the spatial distribution of vehicle crime rates by LSOA. Several generalised observations are described below:
1. Majority of LSOAs with the highest crime rates are located north of the river Thames
2. LSOAs with the highest crime rates tend to be found along the river Thames
3. LSOAs in (2) can be grouped into 4 main areas: East London (e.g. Newham, Barking, Greenwich), Central London (e.g. City of London, Westminster), Southwest London (e.g. Kensington, Knightsbridge), West London (e.g. Heathrow, Acton, Hammersmith, Chiswick)

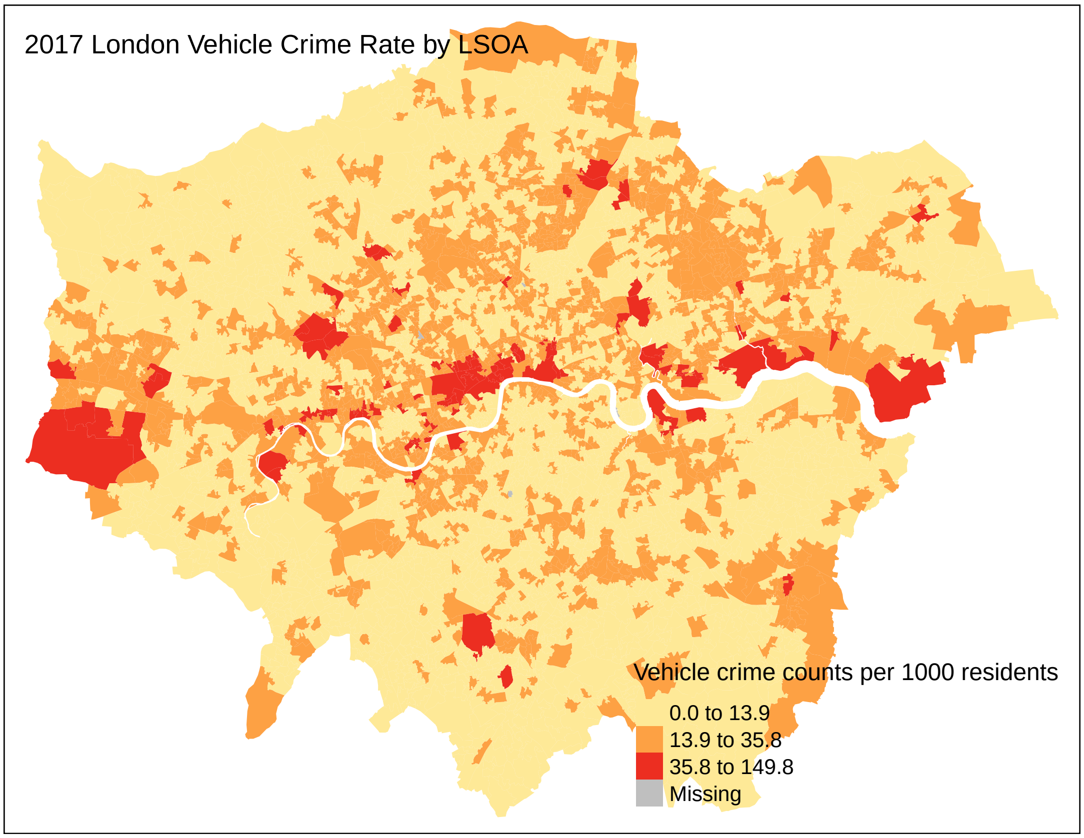


Figure 8: 2017 London vehicle crime rate map by LSOA

5.2.1 Spatial autocorrelation of vehicle crime

Earlier observations suggests that LSOAs with high vehicle crime rates occur in regions. Thus, this section will investigate this further, by examining spatial autocorrelation of vehicle crime at the LSOA level based on regions with contiguous boundaries.

A global Moran's I calculation was ran for all the London LSOAs. The resultant I value of 0.36013 suggests that a positive autocorrelation in vehicle crime counts exist on average across the entire area. Results for the

Moran's I test using randomisation and the Monte-Carlo simulation using 999 simulations both produce small p-values of 2.2e-16 and 0.001 respectively, suggesting that the calculated Moran's I value is highly statistically significant. Thus, at the LSOA level, frequency of vehicle crimes display, on average, a significant positive autocorrelation in London.

Local spatial autocorrelation

Local variations in spatial autocorrelation are also examined, to determine if levels of autocorrelation vary due to spatial heterogeneity. This is done by calculating local Moran's I values for each LSOA.

LSOAs with local Moran's I values that are statistically significant at the 95% confidence interval were then mapped. Figure 9 shows that spatial autocorrelation is non-significant across most of London. It is however noteworthy that the LSOAs with the highest crime rates also tend to form statistically significant high-high clusters. Specifically, high-high clusters are present in West London, central London and East London along the river Thames. This confirms the earlier hypothesis that there is a concentration of crimes on a local level in different regions of London. It is also notable that several other LSOAs (that do not have relatively high vehicle crime rates) also have high-high clusters, indicating that spatial autocorrelation is not only restricted to regions with high vehicle crime rates. This may be related to characteristics of the region or criminal behaviour. Interestingly, a single low-low cluster is observed in South London, in the Bexley region.

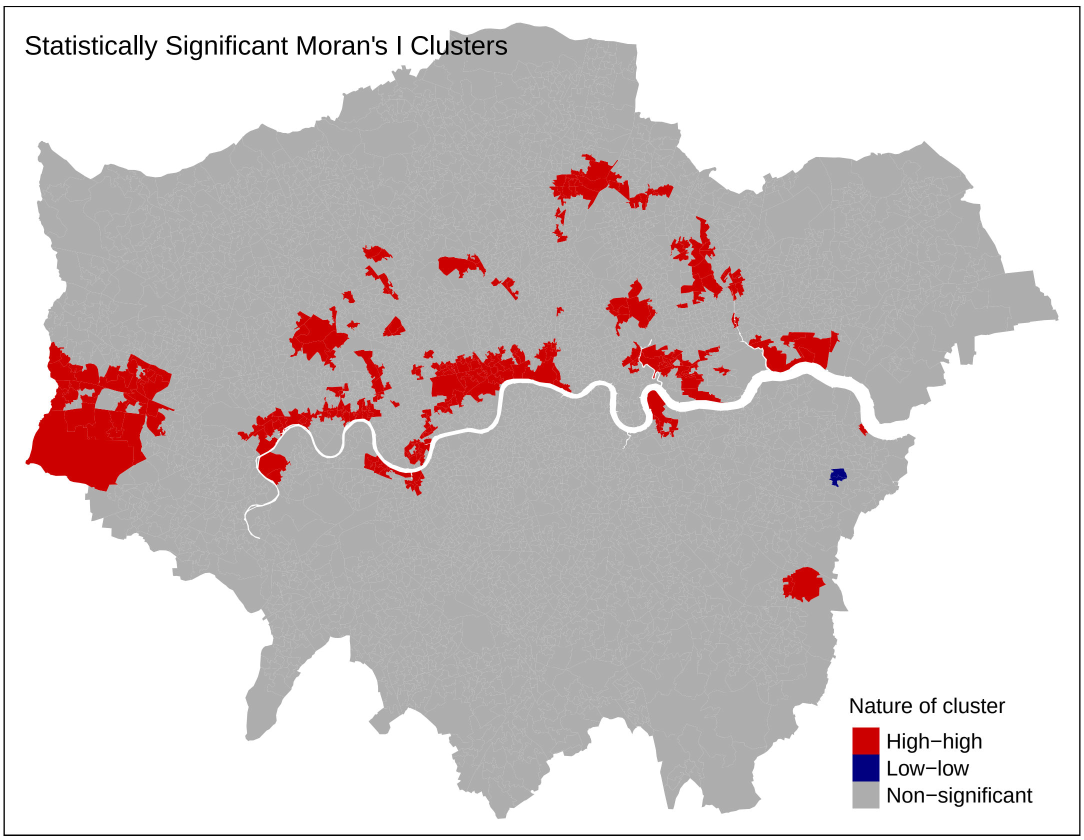


Figure 9: LSOAs with statistically significant local Moran's clusters for vehicle crime

These high-high clusters can be used to identify buffer regions with high vehicle crime rates - a high-high clusters are only identified when an LSOA with high crime rates is surrounded by LSOAs with high crime rates. This can aid the allocation of police resources and outreach efforts in the event of a resource shortage.

It is also a strategic move, as the spatial autocorrelation in vehicle crimes may mean that a reduction of crime in one LSOA has effects on crime rates in neighbouring LSOAs.

5.2.2 Reasons for vehicle crime rates

Earlier analysis reveals that elevated crime rates tend to occur in specific regions, and there is spatial autocorrelation at the LSOA level. Therefore, this section will examine regions with high vehicle crime rates to identify variables that may contribute to high crime rates.

East London

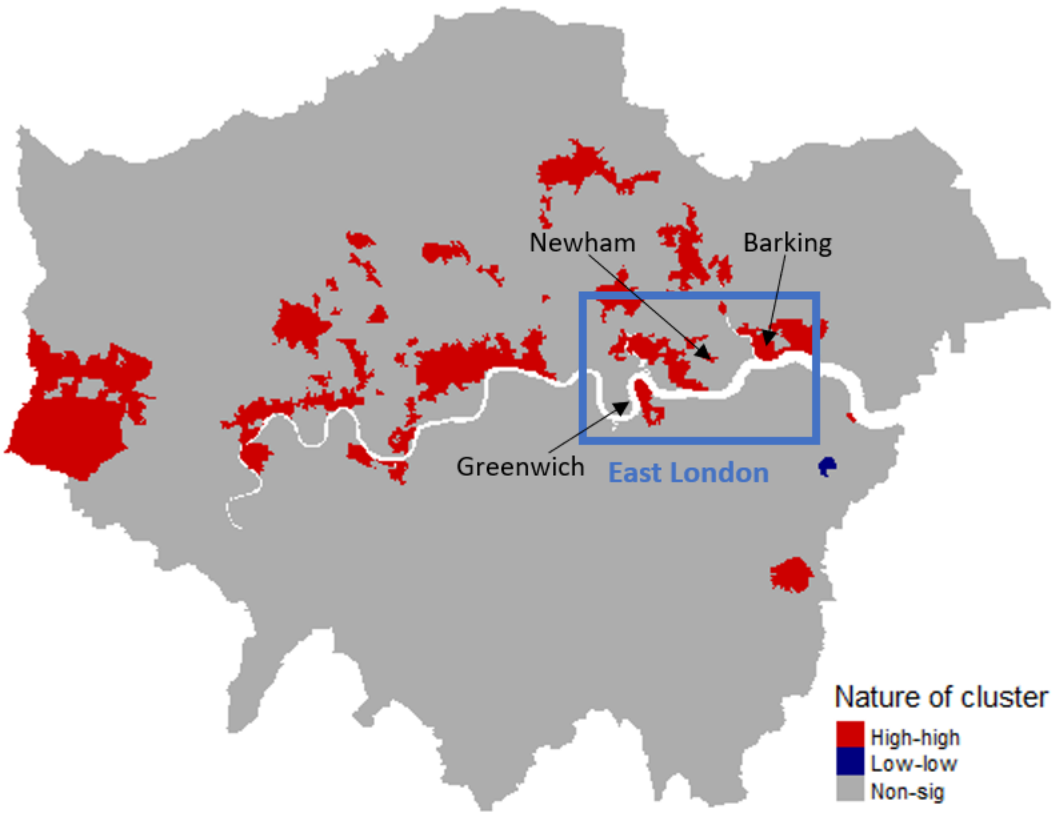


Figure 10: East London

In East London, the LSOAs with high crime rates either have 1) industrial parks/scrapyards/depots (e.g. Barking LSOAs, Newham LSOAs) and/or 2) retail parks/leisure parks (e.g. Newham LSOAs, Greenwich LSOAs). For instance, the agglomeration of retail parks in Newham (and the accompanying parking lots), such as the Gallions Reach shopping park, Beckton Triangle retail park and Gateway retail park may contribute to the elevated vehicle crime rates in the LSOA. Likewise, in Greenwich, the presence of leisure parks and the O2 stadium may contribute to the high vehicle crime rates. The high vehicle crime rates at retail parks/leisure parks suggests that thieves are likely opportunistic, and target these areas due to the high volume of vehicles parked in these areas.

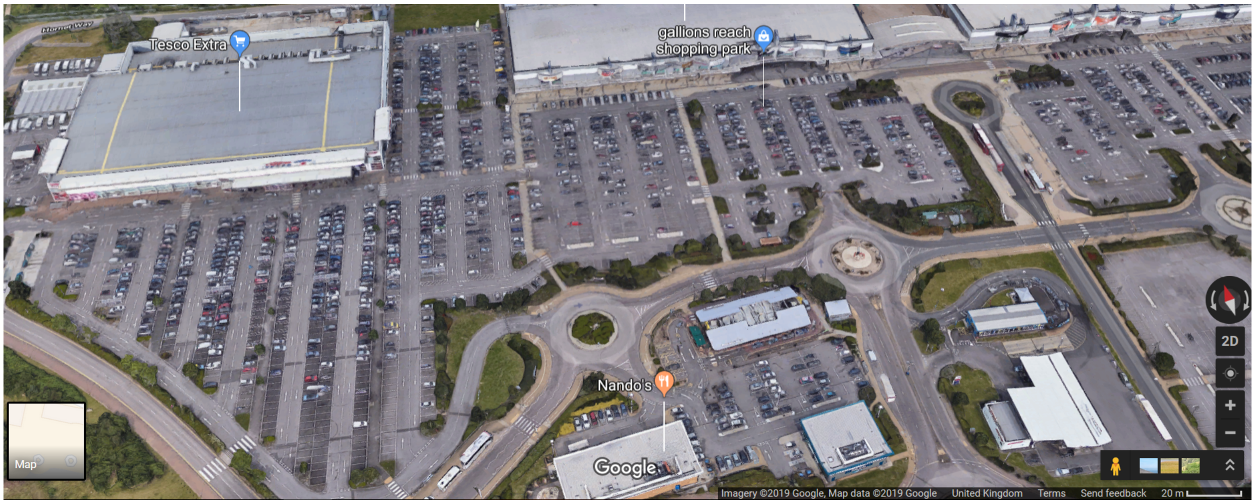


Figure 11: Parking lots at retail parks at Newham

On the other hand, in parts of Barking, the high vehicle crime rates may be attributed to industrial parks and scrapyards. The high rates of vehicle crime at industrial areas is interesting, as the vehicles are often old and/or damaged. This suggests that thieves may steal vehicle parts from these areas. Industrial areas are also relatively quiet and empty at night, which may mean that the risk of getting caught is lower. The perceived risk may be a contributory factor to the higher crime rates.

Figure 12: Vehicles at industrial area in Barking

West London

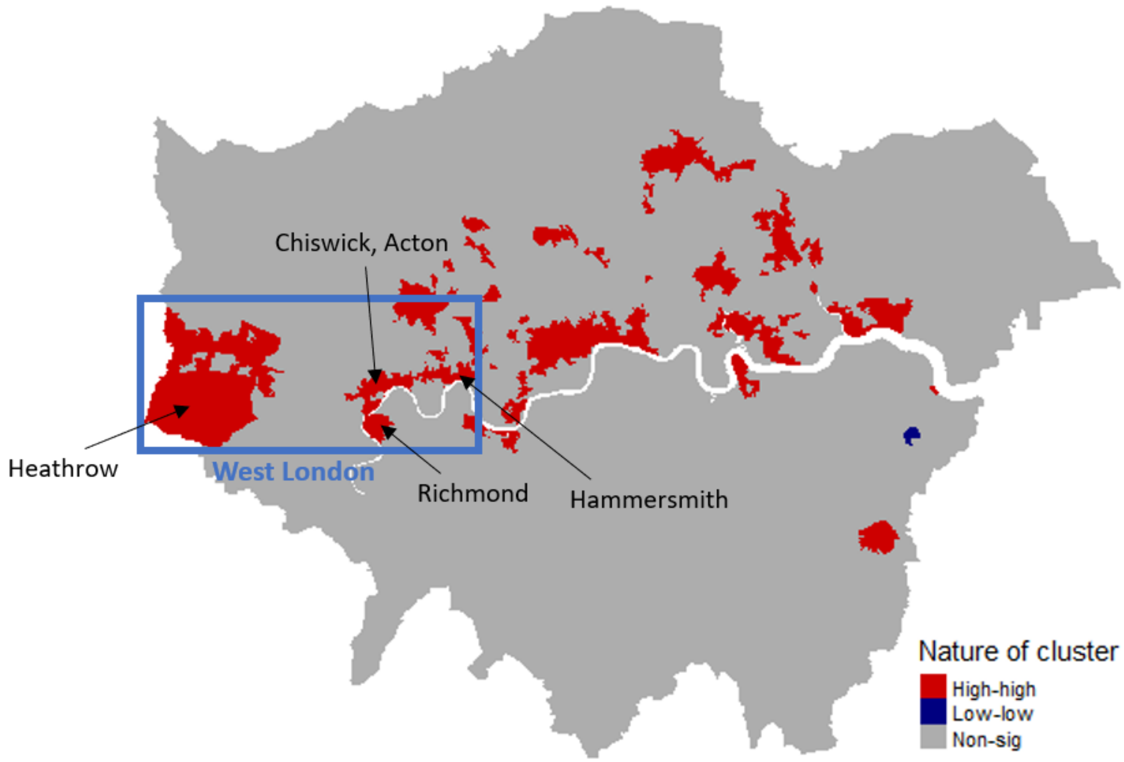


Figure 13: West London

Interestingly, crime patterns are different in West London and can generally be linked to either carparks or car dealers. High rates of vehicle crime near Heathrow airport may be attributed to the high volumes of parked vehicles at short-stay and long-stay carparks. Criminals may therefore choose to operate near the airport, particularly at the long-stay carparks which are likely used by individuals travelling abroad.

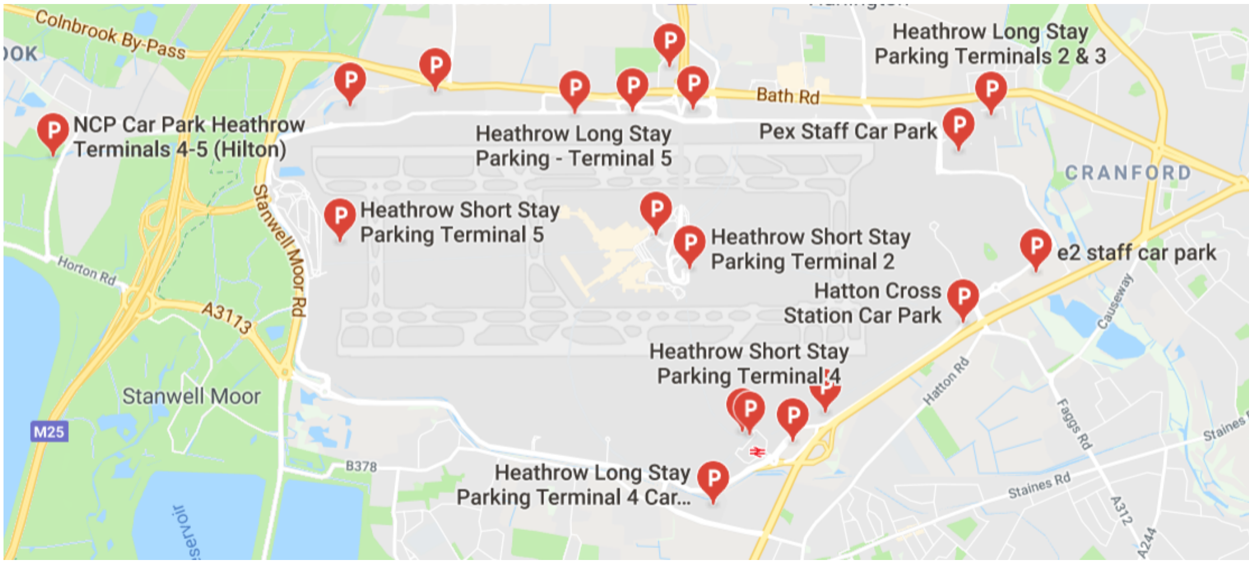


Figure 14: Locations of carparks around Heathrow airport

The high crime rates in Chiswick may also be linked to the presence of large numbers of parked cars, as the LSOAs with the highest crime rates tend to have either carparks associated with supermarkets or car dealers. These observations are consistent with the hypothesis that thieves are opportunistic, and pick areas with more parked vehicles.

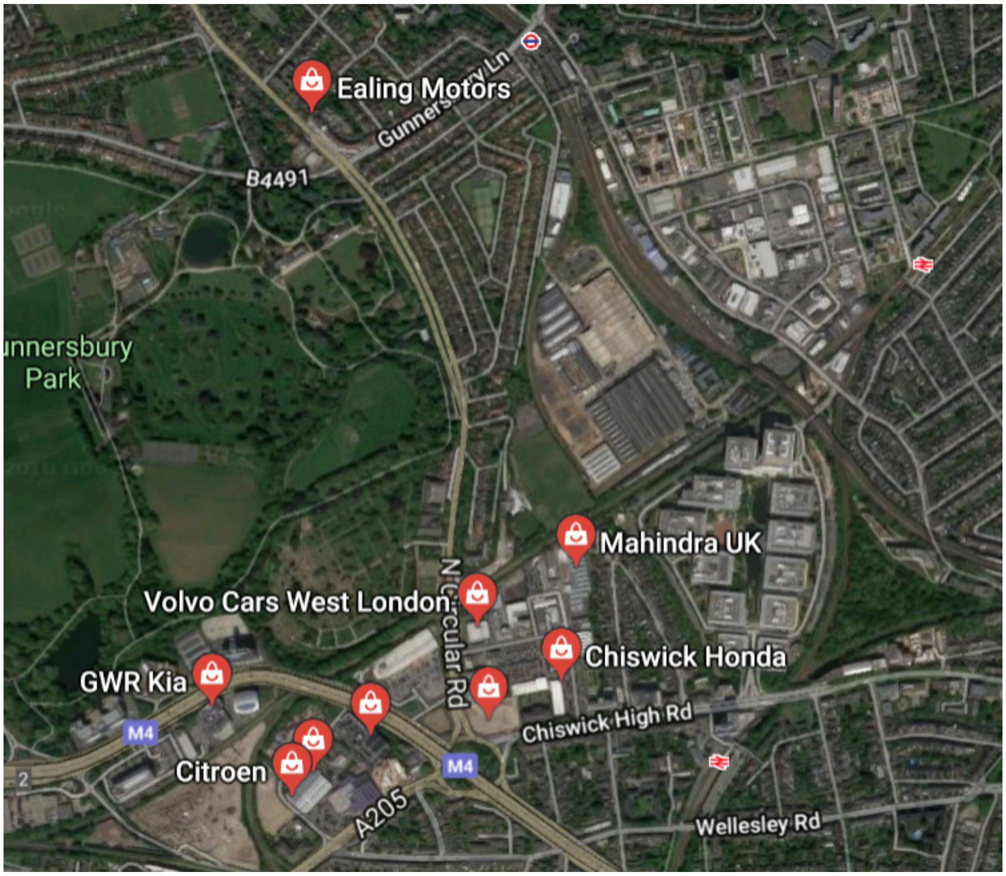


Figure 15: Agglomeration of car dealers in Chiswick

However, these hypotheses do not apply to Hammersmith and parts of Hounslow, both of which appear to be predominantly residential. It may be possible that these LSOAs have high crime rates as they are in the vicinity of other hotspots, such as those in Chiswick.

Vehicle crime rates are also high in Richmond, and may be linked to the presence of multiple carparks and the affluent demographic of the area.

Southwest London

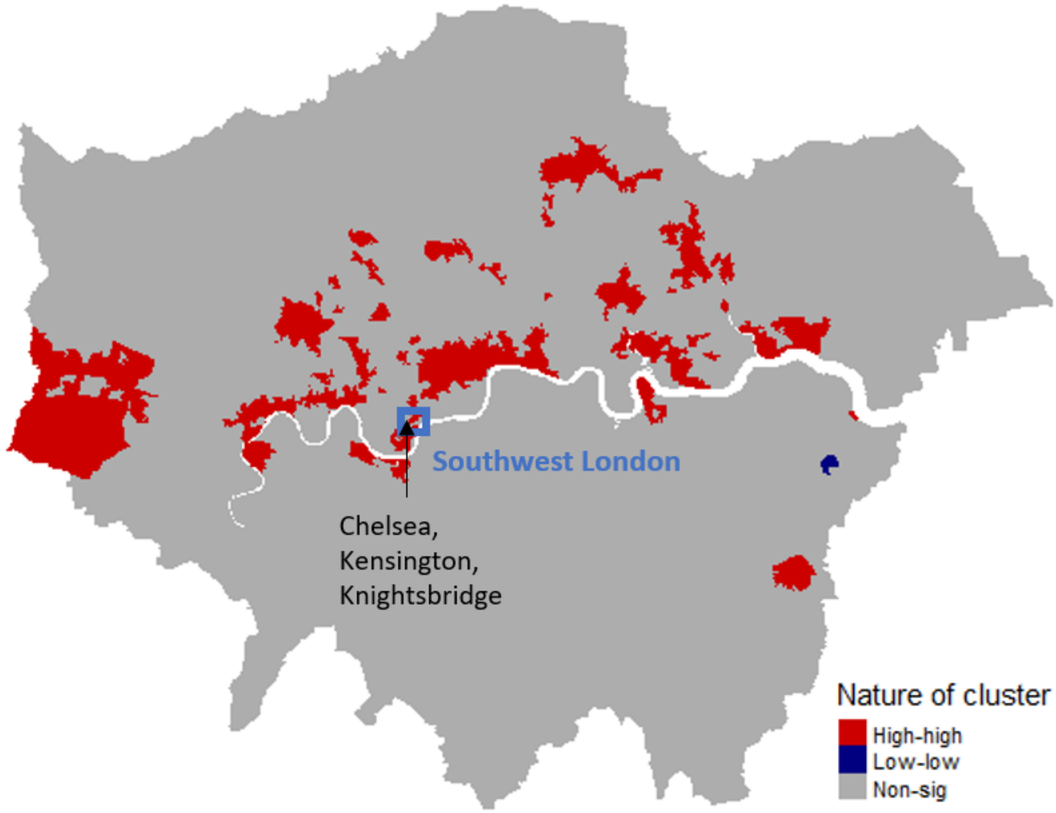


Figure 16: Southwest London

In contrast to other regions, where high crime rates are associated with industrial and commercial areas, in Southwest London, high vehicle crime rates occur in residential areas such as Kensington, Chelsea and Knightsbridge. Unlike other areas, the volume of parked vehicles is not exceptionally high - there are no carparks/industrial areas/car dealers. However, there are numerous luxury vehicles parked in these areas, as they are affluent neighbourhoods. Hence, it is likely that vehicle crime rates are high in these region, as thieves operating in these areas specifically target high-value vehicles.

Figure 17: Luxury vehicles in Knightsbridge

Central London

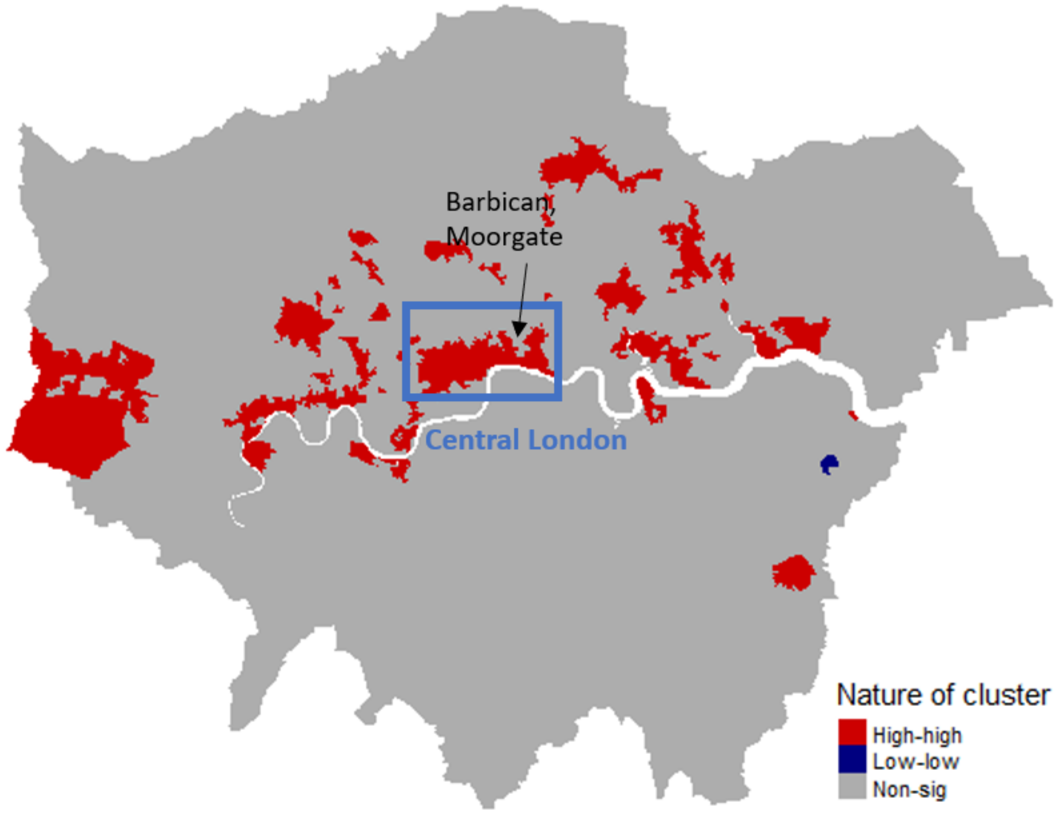


Figure 18: Central London

It is not clear why vehicle crime rates are high in central London, particularly in the City of London and Westminster. However, it is also interesting that some parts of the City of London (e.g. Moorgate, Barbican) have low vehicle crime rates - this may be due to the presence of secure parking in underground carparks.

5.2.3 Discussion

More detailed examination of regions with high vehicle crime rates suggests two main modes of operation for criminals:
1) Operating in areas with large numbers of parked vehicles (i.e. quantity over quality)
2) Operating specifically in affluent areas to target high-value vehicles (i.e. quality over quantity)

Based on hypothesis (1), further analysis has identified industrial areas and large carparks to be significant variables associated with elevated vehicle crime rates. A map of industrial areas and the vehicle crime rates by LSOAs is therefore generated to visualise the relationship between vehicle crime rates and industrial areas (Figure 19). There appears to be a correlation between industrial areas and vehicle crime rates. This may be because these areas tend to be quiet at night thus, the risk of getting caught is lower. Future work to model vehicle crime patterns can therefore consider including distance from industrial areas or the proportion of land allocated to industrial uses in each LSOA as variables to be modelled. Distances to the nearest police station can also be considered as a proxy for perceived risk of getting caught.

Ideally, the relationship between large carparks (as a proxy for the number of parked vehicles) and vehicle crime rates will also be visualised. However, datasets for carparks are not fully complete - for instance, the data on OpenStreetMap mainly limited to central London. Moreover, the dataset provides information on general parking spaces (which could be anywhere), and not the location of large carparks. Future work can consider examining the relationship between large car parks and vehicle crime rates.

Based on hypothesis (2), future work can consider incorporating income data in residential areas as a variable contributing to vehicle crime rates.

That said, high crime rates in central London and other residential (but not particularly affluent) parts of West London such as Hammersmith cannot be explained using the above hypotheses. Further studies can look towards understanding criminal behaviour in these regions.

Lastly, analysis of the vehicle crime dataset is conducted in the form of areal analysis at the LSOAs level. Different levels of analysis may produce different results, and a higher level of aggregation (e.g. at the borough level) may produce different trends. Whilst the observations at the broader areal level between LSOAs suggests that a relationship between industrial areas, car parks with vehicle crime, as well as a relationship between affluent areas and high-value vehicle crime exists, more detailed point analyses at a local level may suggest otherwise. Thus, further work can be conducted within LSOAs, to examine the interactions between these variables at a local level.

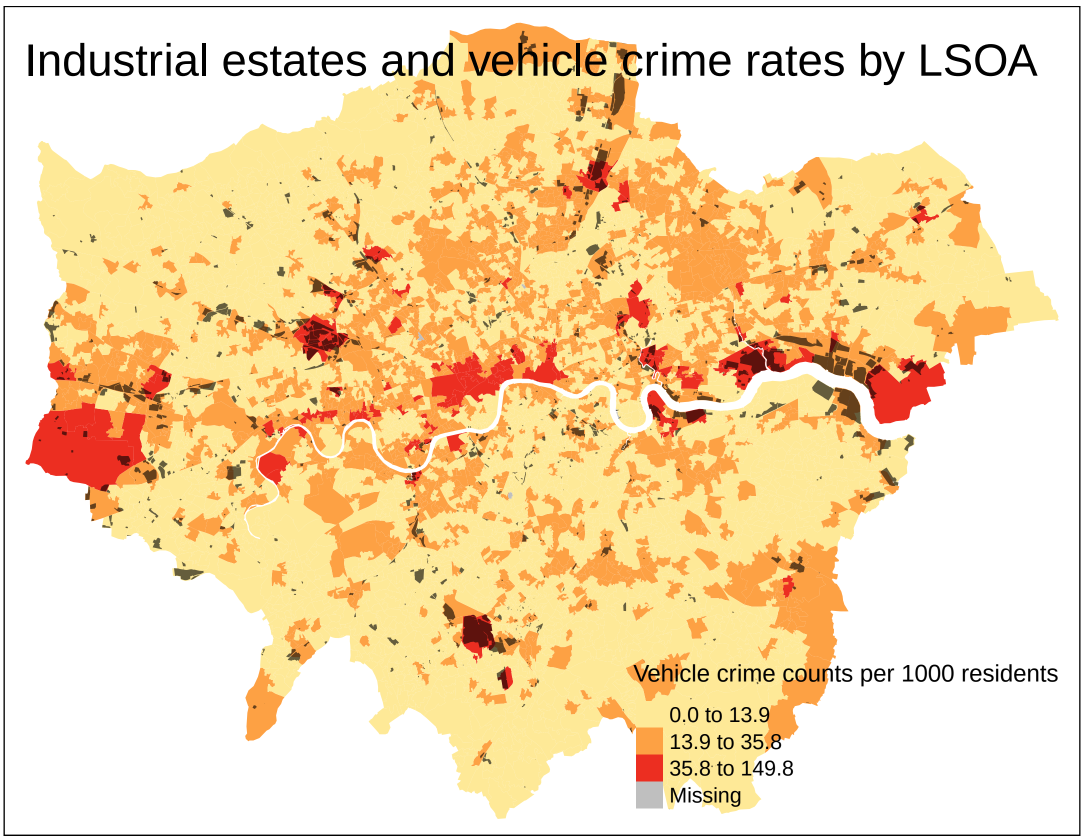


Figure 19: Location of carparks and industrial areas on top of vehicle crime rates

6 Overall discussion and conclusions

compare between crime type Different crime type – different patterns (so employ different method)

comparing between analysing local scale vs LSOA level - prove that results obtained by analysing a single LSOA cannot be generalised to other LSOA - hence shows the limitations of a local scale analysis — that can only be achieved in a LSOA-level analysis - Insert heathrow VSO map – show that theres no pubs etc. – show that spatial structure of the environment is vastly different

Future work

Spatial point pattern vs Areal analysis

In conclusion, visual analysis of the location of entertainment outlets and VSO crimes suggest that they are positively correlated. However, further statistical analysis indicates that a repulsive relationship exists between the two. This scenario underpins the importance of rigorous statistical analysis in testing observations. The reasons for the repulsive relationship are unknown, and further research can be conducted in this area. This can include the incorporation of pedestrian footfall data.

Areal analysis of vehicle crime rates at the LSOAs level indicates cluters of high vehicle crime in parts of West London, East London, Central London and Southwest London. Several hypotheses for these cluster have been formulated. Specifically, these clusters may be attributed to 1) large numbers of parked vehicles, or 2) a concentration of high-end vehicles. Future work can looking into quantifying the effect of variables such as the distance from industrial areas, carparks, and the nearest police station.

Analysis at the LSOA level is subject to the modifiable areal unit problem, and further investigation of the individual points can be considered to better understand the dynamics at the local level. The limitations of a localised point-based analysis can thus be partially addressed using a broader areal-level analysis, and vice versa.

As such, an integrated approach will be the best approach.

7 References

Anselin, Luc. 1995. "Local indicators of spatial analysis–LISA." *Geographical Analysis* 27 (2): 93–115. doi:10.1111/j.1538-4632.1995.tb00338.x.

Baddeley, A, E Rubak, and R Turner. 2015. *Spatial Point Patterns: Methodology and Applications with R*. doi:10.18637/jss.v075.b02.

Basiri, Anahid, Mike Jackson, Pouria Amirian, Amir Pourabdollah, Monika Sester, Adam Winstanley, Terry Moore, and Lijuan Zhang. 2016. "Quality assessment of OpenStreetMap data using trajectory mining." *Geo-Spatial Information Science* 19 (1). Taylor & Francis: 56–68. doi:10.1080/10095020.2016.1151213.

Brand, Sam, and Richard Price. 2014. "The Economic and Social Costs of Crime." *Home Office Research Study 217*, no. July.

Brantingham, P L, and P J Brantingham. 1997. "Mapping Crime for Analytic Purposes: Location Quotients, Counts and Rates." *Mapping Crime for Analytic Purposes: Location Quotients, Counts and Rates*, no. May: 263–88. doi:10.1016/j.amc.2008.06.050.

Chainey, Spencer, Lisa Tompson, and Sebastian Uhlig. 2008. "The Utility of Hotspot Mapping for Predicting Spatial Patterns of Crime." *Security Journal* 21 (4): 291–92. doi:10.1057/sj.2008.6.

Cheng, T, and D Williams. 2012. "Space-time analysis of crime patterns in central London." *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives* 39 (September): 47–52. doi:10.5194/isprsarchives-XXXIX-B2-47-2012.

Haworth, James. 2018. *Spatial Point Pattern Analysis*.

Home Office. 2019. "About data.police.uk." https://data.police.uk/about/.

Malleson, Nick, and Martin A. Andresen. 2016. "Exploring the impact of ambient population measures on London crime hotspots." *Journal of Criminal Justice* 46. The Authors: 52–63. doi:10.1016/j.jcrimjus.2016.03.002.

Mayor of London. 2016. "A City for all Londoners," 1–84.

Moran, P A P. 1950. "Notes on continuous stochastic processes." *Biometrika* 37 (3): 17–23.

Sherman, Lawrence W., Patrick R. Gartin, and Michael E. Buerger. 1989. "Hot spots of predatory crime: Routine activities and the criminology of place." *Criminology* 27 (1): 27–56. doi:10.1111/j.1745-9125.1989.tb00862.x.

Tompson, Lisa, Shane Johnson, Matthew Ashby, Chloe Perkins, and Phillip Edwards. 2015. "UK open source crime data: Accuracy and possibilities for research." *Cartography and Geographic Information Science* 42 (2). Taylor & Francis: 97–111. doi:10.1080/15230406.2014.972456.