

Analysing Crime and Venue Data in London Boroughs

Table of Contents

- Analysing Crime and Venue Data in London Boroughs1*
- 1.1 Introduction2**
- 1.2 Data.....3**
- 1.3 Methodology.....4**
 - Downloading and exploring the dataset4
 - Exploring neighborhoods in London.....4
 - Analysing each neighborhood4
 - Clustering neighborhoods.....4
 - Examining clusters4
- 1.4 Results and discussion.....5**
- 1.5 Conclusion.....6**

1.1 Introduction

Over the last five years, crime rates in London have increased by over 23%. A total of 15,590 knife crimes were reported across London from 2019-2020. The concentration of crimes across different London boroughs should be considered by commercial and residential real estate agents, as well as anybody looking to buy or rent a home in the city. Hence, this report is of interest to anyone that fulfils such criteria.

The business problem we are faced with is: which parts of London are safe and attractive for residential buyers/renters across various demographic groups? To solve this problem, we will assess the crime rates in each of the London boroughs, and cluster neighborhoods to assess the venues on offer, such as pubs, cafes and parks.

1.2 Data

The data that will be used to assess the best residential areas in London is: (1) the Metropolitan Police Service (MPS) Borough Level Crime from August 2018 to July 2020 (<https://data.london.gov.uk>), and (2) Foursquare venue data.

The headings from the MPS Borough Level Crime data are: MajorText, MinorText, LookUp_Borough and 24 separate columns for the 24 months of crime data. The major and minor text describe the accused crime, for example “violence against the person” and “violence with injury”. In order to analyse this data, we must sum the 24 columns to find the total crime levels in each borough. We must then group the data by Borough, which we will refer to as the ‘Address’.

1.3 Methodology

The methodology section will consist of five sections covering the exploratory data analysis, statistical testing and machine learning involved in the project.

- I. Downloading and exploring the dataset
- II. Exploring neighborhoods in London
- III. Analysing each neighborhood
- IV. Clustering neighborhoods
- V. Examining clusters

Downloading and exploring the dataset

The first step was collecting the MPS Borough Level Crime data and transforming it into a pandas dataframe for data exploration, preparation and pre-processing. After examining the data, it was essential to assign meaningful column names and drop the non-essential columns. It was also necessary to sum the 24 months of data to get a total crime level, and group the crime data by borough.

Exploring neighborhoods in London

The second step involved exploring neighborhoods in London. We were able to pull the coordinates of each borough using the geopy library, and explore the venues in each borough using the Foursquare API.

Analysing each neighborhood

Analysing each neighborhood involved using the Foursquare venue data to explore the ten most common venues in each borough.

Clustering neighborhoods

It was essential to use machine learning to further analyse the data. We used K-means clustering to group similar neighborhoods together into five clusters and to discover underlying patterns in the data. We then successfully visualised the clusters on a map of London.

Examining clusters

Finally, we were able to examine each cluster by looking at both the crime data and venue data, in order to produce final recommendations for agents or individuals looking at residential properties in London.

1.4 Results and discussion

We analysed two sources of data in this report. Hence, we may discuss our results in two main parts.

In terms of the Metropolitan Police Service Borough Level Crime data, we found that the boroughs with the least crime are: (1) Kingston upon Thames, (2) Richmond upon Thames and (3) Sutton. The boroughs with the most crime are: (1) Westminster, (2) Newham, and (3) Southwark. The quartiles for the crime data are: 893.4, 1106.6, 1271.3, 2868.6 crimes over the course of 24 months. The mean number of crimes is 1118.9.

Clusters 2, 4 and 5 all showed crime levels below average. However, clusters 1 and 3 showed mixed crime levels across each borough.

The most common venues were as follows:

- Cluster 1 – Pubs, fast food restaurants, parks.
- Cluster 2 – Bus stops, grocery stores, convenience stores.
- Cluster 3 – Coffee shops, pubs and clothing stores.
- Cluster 4 – Bakeries, train stations, parks.
- Cluster 5 – Sports clubs, home services, pubs.

The clusters that would appeal most to families and older people, with fewer pubs and more active spaces and independent stores, are Cluster 4 and Cluster 5 as they offer suitable facilities and low crime rates. Hence, the boroughs of Sutton (4) and Richmond (5) are attractive for family homes and older people.

Cluster 1 seems to be appropriate for younger residential buyers/renters with more pubs, fast food restaurants and parks. Within Cluster 1, the borough with the least crime is Merton, followed by Bexley.

Cluster 3 would be fitting for either group, especially families with slightly older children, with a variety of coffee shops, pubs and clothing stores. Within this cluster, the boroughs with the least crime are Kingston upon Thames and Harrow.

Finally, Cluster 2 appears to have low crime rates, but has much fewer desirable venues such as pubs or cafes. Hence, Barking and Dagenham may be safe but less attractive for buyers/renters.

1.5 Conclusion

The business question that we have answered in this report is: which parts of London are safe and attractive for residential buyers/renters across various demographic groups? To solve this problem, we have assessed the crime rates in each of the London boroughs, and clustered neighborhoods to assess the venues on offer, such as pubs, cafes and parks.

Overall, the ideal borough for a residential buyer/renter depends on the age and venue preference of the individual. However, it seems that the Sutton and Richmond are most suitable for older people or families with young children. Kingston upon Thames and Harrow are ideal for families with older children. Merton and Bexley are appropriate for young people looking for more pubs and fast food restaurants. Barking and Dagenham is a safe option but with few desirable venues.

Whilst London crime rates have been growing over the last few years, there are many boroughs which are both relatively safe and offer venues which are appealing to all ages. This report acts as a guide to all commercial and residential real estate agents, as well as anybody looking to buy or rent a home in the city, and should inform individuals across all ages.