The Battle of Neighbourhoods

Harsh Shah

2nd August 2020

1. Introduction

1.1 Background

In London, there are numerous boroughs. The city is famous for its many touristic spots and is one of the most visited cities on the planet. In addition to tourists, many students come to London in order to study at university. As well as this, many people come for short term stays for e.g. internships, as well as for graduate programs.

1.2 Problem

Finding a good place to stay is very important, somewhere that has local access to shops and supermarkets, as well as restaurants, bars and a number of other facilities.

HS is moving to London in order to study at University College London. In addition to university life, HS also wants to get a feel of the city's culture, food etc. He's looking to live in a neighbourhood that is culturally vibrant but not too touristy, as he feels that he can visit tourist sites in his spare time, therefore the proximity of tourist attractions is not something that matters to him.

From the perspective of students, many factors are involved when searching for the best accommodation, which includes distance and rent. However, this project will focus on the general atmosphere of different areas of London as well as safety.

The research carried out is expected to be of benefit to international students looking to live in London, because despite the opportunity to explore the local culture, they would very likely want to feel at home at the same time.

2. Data Acquisition and Preprocessing

In the project, I will be using the following datasets to produce an outcome - Borough Level Crime, List of London boroughs and Foursquare API. After using legitimate resources to acquire them, they will be cleansed into more useful forms so that they can be analysed further.

2.1 Borough Level Crime

	MajorText	MinorText	LookUp_BoroughName	201807	201808	201809	201810	201811	201812	201901	...	201909	201910	201911
0	Arson and Criminal Damage	Arson	Barking and Dagenham	6	5	3	8	5	1	5	...	6	9	8
1	Arson and Criminal Damage	Criminal Damage	Barking and Dagenham	127	101	107	132	105	88	97	...	109	109	97
2	Burglary	Burglary - Business and Community	Barking and Dagenham	30	18	33	32	39	33	45	...	37	30	30
3	Burglary	Burglary - Residential	Barking and Dagenham	94	84	99	94	106	164	114	...	80	97	114
4	Drug Offences	Drug Trafficking	Barking and Dagenham	8	7	10	7	7	4	5	...	7	8	13

The table above shows crime numbers per crime type at borough level over the last 24 months. The data consists of 1569 observations and 27 columns. It was obtained from the London Datastore.

To go into further detail, the number of crimes were used to calculate monthly averages, without taking into consideration crime categories, for simplification purposes.

	BoroughName	MonthlyAverage
0	Barking and Dagenham	1616.166667
1	Barnet	2467.208333
2	Bexley	1403.833333
3	Brent	2496.708333
4	Bromley	1983.833333
5	Camden	3031.833333
6	Croydon	2754.875000
7	Ealing	2521.750000
8	Enfield	2443.666667
9	Greenwich	2285.458333
10	Hackney	2705.958333
11	Hammersmith and Fulham	1853.208333
12	Haringey	2577.875000
13	Harrow	1369.000000
14	Havering	1527.958333

2.2 List of London Boroughs

The second dataset that I used provides information on the different London boroughs, obtained from Wikipedia

Borough	Inner	Status	Local authority	Political control	Headquarters	Area (sq mi)	Population (2013 est)[1]	Co-ordinates	No. in map
Barking and Dagenham [note 1]			Barking and Dagenham London Borough Council	Labour	Town Hall, 1 Town Square	13.93	194,352	51.5607°N 0.1557°E	25
Barnet			Barnet London Borough Council	Conservative	Barnet House, 2 Bristol Avenue, Colindale	33.49	369,088	51.6252°N 0.1517°W	31
Bexley			Bexley London Borough Council	Conservative	Civic Offices, 2 Watling Street	23.38	236,687	51.4549°N 0.1505°E	23
Brent			Brent London Borough Council	Labour	Brent Civic Centre, Engineers Way	16.70	317,264	51.5588°N 0.2817°W	12
Bromley			Bromley London Borough Council	Conservative	Civic Centre, Stockwell Close	57.97	317,899	51.4039°N 0.0198°E	20
Camden	✓		Camden London Borough Council	Labour	Camden Town Hall, Judd Street	8.40	229,719	51.5290°N 0.1255°W	11
Croydon			Croydon London Borough Council	Labour	Bernard Weatherill House, Mint Walk	33.41	372,752	51.3714°N 0.0977°W	19
Ealing			Ealing London Borough Council	Labour	Perceval House, 14-16 Uxbridge Road	21.44	342,494	51.5130°N 0.3089°W	13
Enfield			Enfield London Borough Council	Labour	Civic Centre, Silver Street	31.74	320,524	51.6538°N 0.0799°W	30
Greenwich [note 2]	✓ [note 3]	Royal	Greenwich London Borough Council	Labour	Woolwich Town Hall, Wellington Street	18.28	264,008	51.4892°N 0.0648°E	22

From this data, we will only be using the population data and coordinates. The population data can be used to calculate the ratio of crime to population in order to generate a better comparison, while the coordinates can be used to obtain geographical data from Foursquare. The simplified data looks like this:

	BoroughName	Population	Latitude	Longitude
0	Barking and Dagenham	194352	51.5607	0.1557
1	Barnet	369088	51.6252	-0.1517
2	Bexley	236687	51.4549	0.1505
3	Brent	317264	51.5588	-0.2817
4	Bromley	317899	51.4039	0.0198

2.3 Foursquare API

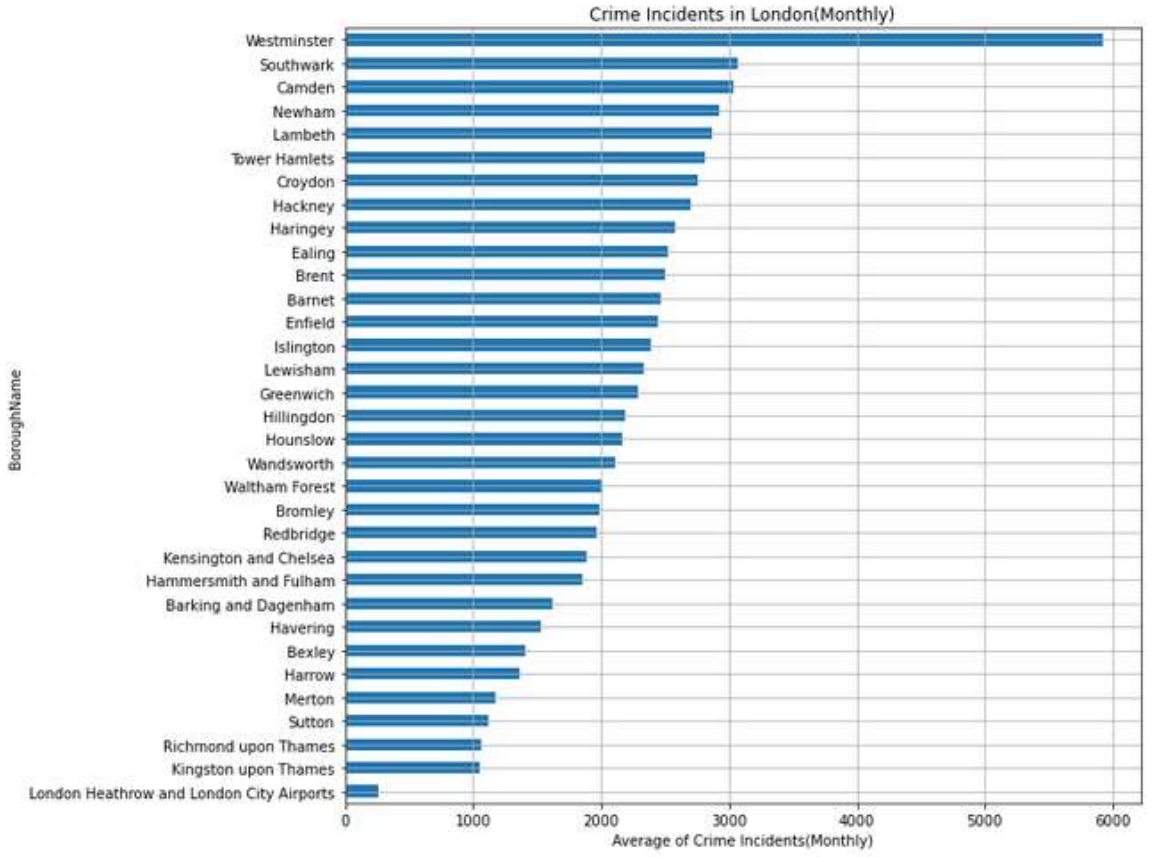
Foursquare API was used to obtain the 50 most popular venues in each borough, which was done using the 'explore' function to request the URL. As a result, I was able to get data which looks like this:

	BoroughName	Borough Latitude	Borough Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Barking and Dagenham	51.5607	0.1557	Central Park	51.559560	0.161981	Park
1	Barking and Dagenham	51.5607	0.1557	Crowlands Heath Golf Course	51.562457	0.155818	Golf Course
2	Barking and Dagenham	51.5607	0.1557	Robert Clack Leisure Centre	51.560808	0.152704	Martial Arts Dojo
3	Barking and Dagenham	51.5607	0.1557	Beacontree Heath Leisure Centre	51.560997	0.148932	Gym / Fitness Center
4	Barking and Dagenham	51.5607	0.1557	Becontree Heath Bus Station	51.561065	0.150998	Bus Station

3. Methodology

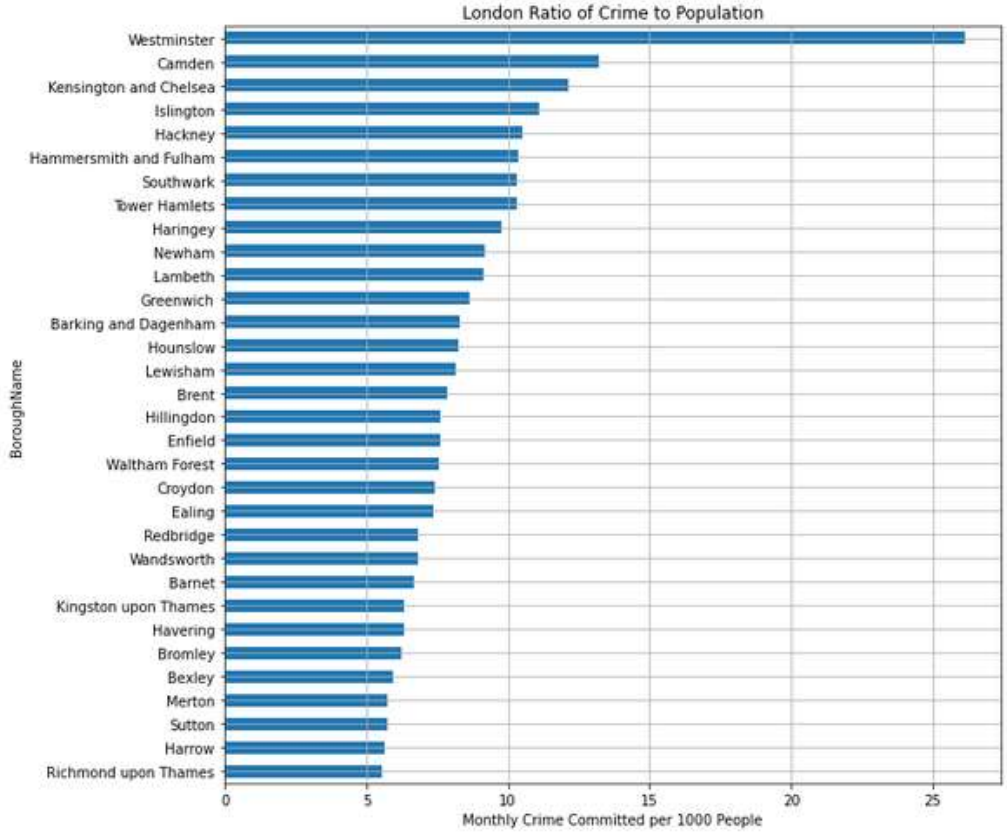
3.1 Exploratory Analysis

Using the cleansed data, I created visualisations to better represent the data



This bar chart shows the boroughs in descending order of monthly crime incidents. As you can see, Westminster has the highest number of recorded crimes, followed by Southwark, Camden and Newham.

Despite this, I felt that it wouldn't be sensible to directly compare the number of recorded crimes due to differences in population. Therefore, I used the population to calculate the number of crimes per 1000 people in each borough:



It can still be seen that Westminster and Camden still remain high up the graph, although Kensington and Chelsea, Islington and Hackney have also moved towards the top of the chart.

Also, before conducting further analysis, I decided that it would be a good idea to take a look at the different locations to get a better idea of London as a region.

3.2 Cluster Analysis

Next, K-means clustering was performed in order to group the boroughs based on the boroughs they had using the Foursquare data, to get a feel of the general atmosphere of each of the boroughs.

The first step that was taken was to carry out 'one hot encoding', in order to create binary values for each of the clusters.

	BoroughName	African Restaurant	Airport	Airport Lounge	Airport Service	American Restaurant	Argentinian Restaurant	Art Gallery	Art Museum	Arts & Crafts Store	...	Turkish Restaurant	Used Bookstore
0	Barking and Dagenham	0	0	0	0	0	0	0	0	0	...	0	0
1	Barking and Dagenham	0	0	0	0	0	0	0	0	0	...	0	0
2	Barking and Dagenham	0	0	0	0	0	0	0	0	0	...	0	0
3	Barking and Dagenham	0	0	0	0	0	0	0	0	0	...	0	0
4	Barking and Dagenham	0	0	0	0	0	0	0	0	0	...	0	0

Next, the data was grouped according to borough name in order to find out the number of venues of each category that exist within the 50 most popular venues. However, due to the lack of data from Foursquare, some boroughs displayed less than 50 venues, therefore the category count was altered to the frequency of how often each category appears amongst others. Based on this frequency, we could obtain a list of the most common venue categories within each borough as follows:

	BoroughName	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Barking and Dagenham	Pool	Golf Course	Bus Station	Supermarket	Park	Gym / Fitness Center	Martial Arts Dojo	Film Studio	Fast Food Restaurant	Farmers Market
1	Barnet	Café	Bus Stop	Yoga Studio	English Restaurant	Food Court	Flea Market	Fish Market	Film Studio	Fast Food Restaurant	Farmers Market
2	Bexley	Clothing Store	Coffee Shop	Pub	Supermarket	Fast Food Restaurant	Pharmacy	Warehouse Store	Portuguese Restaurant	Hotel	Department Store
3	Brent	Coffee Shop	Hotel	Clothing Store	Sporting Goods Shop	Grocery Store	American Restaurant	Sandwich Place	Bar	Pizza Place	Chocolate Shop
4	Bromley	Coffee Shop	Clothing Store	Pizza Place	Burger Joint	Gym / Fitness Center	Bar	Irish Pub	English Restaurant	Furniture / Home Store	Burrito Place

Using these categories, K-means clustering was then carried out in order to group the boroughs into 5 different clusters based on similarity. The different coloured dots show different clusters:

Based on the characteristics that each of the clusters possess, we have named each of the clusters according to how their characteristics are best depicted:

- Cluster 0 – Healthy area (Gyms, pools)

BoroughName	CrimePer1000	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
Barking and Dagenham	8.315668	0	Pool	Golf Course	Bus Station	Supermarket	Park	Gym / Fitness Center	Martial Arts Dojo	Film Studio	Fast Food Restaurant	Farmers Market

- Cluster 1 – Lively area (Pubs, Restaurants)

BoroughName	CrimePer1000	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
Brent	7.869498	1	Coffee Shop	Hotel	Clothing Store	Sporting Goods Shop	Grocery Store	American Restaurant	Sandwich Place	Bar	Pizza Place
Camden	13.198009	1	Hotel	Café	Coffee Shop	Bakery	Breakfast Spot	Burger Joint	Italian Restaurant	Garden	Train Station
Kensington and Chelsea	12.148069	1	Café	Juice Bar	Restaurant	Clothing Store	Gym / Fitness Center	Burger Joint	Bakery	Sporting Goods Shop	Lebanese Restaurant
Newham	9.189195	1	Hotel	Pharmacy	Light Rail Station	Rafting	Sandwich Place	Chinese Restaurant	Airport Service	Airport Lounge	Airport
Southwark	10.292419	1	Coffee Shop	Hotel	Pub	Hotel Bar	Cocktail Bar	Pizza Place	Art Gallery	English Restaurant	Theater
Tower Hamlets	10.288181	1	Italian Restaurant	Sandwich Place	Hotel	Coffee Shop	Gym / Fitness Center	Outdoor Sculpture	Chinese Restaurant	Grocery Store	Light Rail Station
Westminster	26.100514	1	Hotel	Coffee Shop	Theater	Sushi Restaurant	Juice Bar	Hotel Bar	Sandwich Place	Sporting Goods Shop	Korean Restaurant

- Cluster 2 – Shopping area (Coffee shops, clothing stores)

BoroughName	CrimePer1000	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue
Bexley	5.931181	2	Clothing Store	Coffee Shop	Pub	Supermarket	Fast Food Restaurant	Pharmacy	Warehouse Store
Bromley	6.240452	2	Coffee Shop	Clothing Store	Pizza Place	Burger Joint	Gym / Fitness Center	Bar	Irish Pub
Croydon	7.390638	2	Pub	Coffee Shop	Portuguese Restaurant	Asian Restaurant	Clothing Store	Park	Spanish Restaurant
Ealing	7.362903	2	Coffee Shop	Italian Restaurant	Pizza Place	Hotel	Burger Joint	Vietnamese Restaurant	Clothing Store
Enfield	7.623974	2	Coffee Shop	Clothing Store	Pub	Supermarket	Stationery Store	Bookstore	Optical Shop
Greenwich	8.656777	2	Clothing Store	Pub	Fast Food Restaurant	Supermarket	Coffee Shop	Hotel	Sandwich Place
Hackney	10.513516	2	Pub	Café	Brewery	Bakery	Coffee Shop	Hotel	Grocery Store
Hammersmith and Fulham	10.371370	2	Pub	Indian Restaurant	Italian Restaurant	Café	Gastropub	Clothing Store	Vietnamese Restaurant

- Cluster 3 - Market area

BoroughName	CrimePer1000	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
Barnet	6.684607	3	Café	Bus Stop	Yoga Studio	English Restaurant	Food Court	Flea Market	Fish Market	Film Studio	Fast Food Restaurant	Farmers Market

- Cluster 4 – Traveller area

BoroughName	CrimePer1000	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue
Hounslow	8.259275	4	Metro Station	Park	Bed & Breakfast	Café	Yoga Studio	English Restaurant	Food Court	Flea Market	Fish Market

4. Results

Using different analysis, I was able to discover the best areas to live based on the criteria of general atmosphere and safety. Now, we will look back at all the analysis performed in the project before making a conclusion on which area would be the best for HS to live in.

4.1 Safety

For safety, I decided to normalise the crime to population video and reversed the safety score so that the closes value to 1 represents the area that has the lowest crime per head.

4.2 Atmosphere

With regards to atmosphere, I decided to give an arbitrary value to each cluster on the basis of personal preference, since preference is difficult to quantify without being subjective. The highest value was given to the Shopping Area (Cluster 2) while the lowest was given to the Traveller Area (Cluster 4).

BoroughName	CrimePer1000	Cluster Labels	Safety	Atmosphere
Barking and Dagenham	8.315668	0.8	0.864607	0.0
Barnet	6.684607	3.0	0.943901	0.7
Bexley	5.931181	2.0	0.980529	1.0
Brent	7.869498	1.0	0.886298	0.9
Bromley	6.240452	2.0	0.965494	1.0

4.3 Final Score

Finally, by adding the Safety score to the Atmosphere value, I was able to find the best area to live in, which was Richmond upon Thames, scoring 2.0.

BoroughName	Safety	Atmosphere	Score
Richmond upon Thames	1.000000	1.0	2.000000
Harrow	0.995407	1.0	1.995407
Sutton	0.990868	1.0	1.990868
Merton	0.988647	1.0	1.988647
Bexley	0.980529	1.0	1.980529

5. Conclusion

5.1 Final result

Based on the analysis, I have found that the 5 boroughs below would be the best places for HS to live according on the criteria that he set out, based on general atmosphere and safety. They all belong to the Shopping Area cluster, which contains many clothing shops and coffee shops, as well as a number of restaurants that serve a range of cuisines. The only thing that sets them apart is the safety score.

5.2 Limitations

From the map, we can see that all of the top 5 boroughs are located far out from the city, which is due to the limitations held by the research carried out.

As well as taking into account crime rates and the types of stores that exist in each area, I could've perhaps also taken into account the number of each type of store that exists in each borough, as well as other factors such as house/rental prices.

Nevertheless, performing this research was still incredible enjoyable as I was able to explore each borough in a lot of depth.

References

- **"**Borough Level Crime", London Datastore
- "List of London Boroughs"
- Foursquare API
- IBM Data Science Professional Certificate notes, Coursera