THE BATTLE OF NEIGHBORHOODS

Capstone Project

Abstract

To find the best place to live in London, considering the crime occurrence and venues density

Barbora Husáková

20.04.2020

Table of Contents

1. Introduction: Business Problem

London, the capital of the UK, is considered to be one of the most important global and multicultural cities. It has been ranked by many to be the most visited, expensive, innovative, popular for working, studying, and simply living in. Whether one is moving to London because of studies or work opportunities, one needs to find a nice neighborhood to live in that provides him/her with all the safety, high quality services and standard of life.

Business Problem

Each London borough has its own characteristics, depending on what one's preferences are, one needs to decide where to settle down when relocating to London. Some districts offer higher safety, others more local amenities and nightlife activities while somebody may seek for quite neighborhood with park areas and local markets. Of course, there is a budget limitation and time requirements for commuting to school or work that are important aspects when selecting a location of the accommodation, however, these are not to be included in our analysis.

As there are many factors to be taken into account when looking for a borough where one is going to spend several months or even years of life, we have decided to analyze London's neighborhoods from safety and venues perspective in order to help potential future Londoners to find a proper and suitable borough. Moreover, these analytics can aid real estate agencies as well, when categorizing the boroughs and giving recommendations to the people searching for a real estate.

Main Goal

The main objective of this project is to find the best place to live in London, taking into account the crime occurrence and venues density meaning to classify boroughs in London based on the safety indicator (reflecting the current status of the crime incidents) and topmost common venues in each of them.

Target Audience

By analyzing two main attributes (safety and amenities availability), we aim to help people relocating to London to give a better overview of the current living situation in London and recommend them the areas as per their interests and expectations so they can better orient themselves in an unknown city and narrow down their search. This report will be focusing mainly on the stakeholders of general audience ("newcomers" to London) while summarizing the safety and venues categories so they can more conveniently choose the final location as per their preferences.

Success Criteria

The success criteria of the project will be a good recommendation of a borough to people based on the above-mentioned factors i.e. safety and venues types.

Problem Solving

Firstly, we analyze the crime dataset in order to better understand the current distribution of crimes across the respective boroughs. In the next step, after obtaining the data related to various venues available in the respective boroughs, we cluster the boroughs using the k-means algorithm to get the groups of similar boroughs that can be recommended to the prospects searching for the best area as per their needs/preferences. All the results will be supported with the visualizations in order to more effectively present the outputs of the analysis and underlying relations.

2. Data Acquisition and Cleaning

In order to proceed with the analytics part and to achieve the set objectives, the below data has been collected and processed in order to give us better insight of the safety and venues availability:

- Total number of crimes in the last two years per borough and crime category
- The most common venues in each borough that are reported at the time of analysis

The datasets have been acquired from the following sources:

- London crime dataset:
  - London Datastore: https://data.london.gov.uk/dataset/recorded_crime_summary
- London boroughs dataset:
  - Wikipedia: https://en.wikipedia.org/wiki/List_of_London_boroughs
- London boroughs coordinates:
  - Geopy API: https://pypi.org/project/geopy/
- London venues in the respective boroughs:
  - Foursquare API: https://api.foursquare.com

London Crime Dataset

London Crime dataset obtained from the London Datastore consists of all the crimes recorded in the last two years while broken down by the respective boroughs, crime categories and months. The below table shows an excerpt from this dataset as it originally looks like after downloading from the London Datastore. The data has 1575 records and 27 fields.

	MajorText	MinorText	BoroughName	201803	201804	201805	201806	201807	201808	201809	...	201905	201906	201907	201908	201909	201910	201911
0	Arson and Criminal Damage	Arson	Barking and Dagenham	6	3	4	12	6	5	3	...	11	3	5	3	6	9	8
1	Arson and Criminal Damage	Criminal Damage	Barking and Dagenham	115	122	126	123	127	101	107	...	138	113	134	118	109	109	99
2	Burglary	Burglary - Business and Community	Barking and Dagenham	38	36	24	33	30	18	33	...	22	27	31	35	37	30	30
3	Burglary	Burglary - Residential	Barking and Dagenham	122	75	93	77	94	84	99	...	114	96	71	67	80	97	114
4	Drug Offences	Drug Trafficking	Barking and Dagenham	7	3	8	6	9	7	10	...	8	6	8	6	6	9	10

Figure 2-1 London Crime Dataset

The dataset has been transformed to the below format in order to calculate the Total number of crimes in the last two years and Monthly Average crime incidents, afterwards, we are able to aggregate the data on the needed borough and major crime category level.

	BoroughName	CrimeCategory	Total	MonthlyAverage
0	Barking and Dagenham	Arson and Criminal Damage	2904	121.000000
1	Barking and Dagenham	Burglary	3156	131.500000
2	Barking and Dagenham	Drug Offences	2142	89.250000
3	Barking and Dagenham	Miscellaneous Crimes Against Society	577	24.041667
4	Barking and Dagenham	Possession of Weapons	370	15.416667

Figure 2-2 Transformed London Crime Dataset

London Boroughs Dataset

Another dataset used in the analysis is the list of all London boroughs that have been scraped from Wikipedia website using Beautiful Soup library. From this dataset, only borough name and population fields have been used for further analytics as coordinates have been acquired from a Geopy API.

Borough	Inner	Status	Local authority	Political control	Headquarters	Area (sq mi)	Population (2013 est)[1]	Co-ordinates	Nr. 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	North London Business Park, Oakleigh Road South	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
Hackney	✓		Hackney London Borough Council	Labour	Hackney Town Hall, Mare Street	7.36	257,379	51.5450°N 0.0553°W	9
Hammersmith and Fulham [note 4]	✓		Hammersmith and Fulham London Borough Council	Labour	Town Hall, King Street	6.33	178,685	51.4927°N 0.2339°W	4
Haringey	[note 3]		Haringey London Borough Council	Labour	Civic Centre, High Road	11.42	263,386	51.6000°N 0.1119°W	29
Harrow			Harrow London Borough Council	Labour	Civic Centre, Station Road	19.49	243,372	51.5898°N 0.3346°W	32
Havering			Havering London Borough Council	Conservative (council NOC)	Town Hall, Main Road	43.35	242,080	51.5812°N 0.1837°E	24
Hillingdon			Hillingdon London Borough Council	Conservative	Civic Centre, High Street	44.67	286,806	51.5441°N 0.4760°W	33
Hounslow			Hounslow London Borough Council	Labour	Hounslow House, 7 Bath Road	21.61	262,407	51.4746°N 0.3680°W	14
Islington	✓		Islington London Borough Council	Labour	Municipal Offices, 222 Upper Street	5.74	215,667	51.5416°N 0.1022°W	10
Kensington and Chelsea	✓	Royal	Kensington and Chelsea London Borough Council	Conservative	The Town Hall, Hornton Street	4.68	155,594	51.5020°N 0.1947°W	3
Kingston upon Thames		Royal	Kingston upon Thames London Borough Council	Liberal Democrat	Guildhall, High Street	14.38	166,793	51.4085°N 0.3064°W	16
Lambeth	✓		Lambeth London Borough Council	Labour	Lambeth Town Hall, Brixton Hill	10.36	314,242	51.4607°N 0.1163°W	6
Lewisham	✓		Lewisham London Borough Council	Labour	Town Hall, 1 Catford Road	13.57	286,180	51.4452°N 0.0209°W	21
Merton			Merton London Borough Council	Labour	Civic Centre, London Road	14.52	203,223	51.4014°N 0.1958°W	17
Newham	[note 3]		Newham London Borough Council	Labour	Newham Dockside, 1000 Dockside Road	13.98	318,227	51.5077°N 0.0469°E	27
Redbridge			Redbridge London Borough Council	Labour	Town Hall, 128-142 High Road	21.78	288,272	51.5590°N 0.0741°E	26
Richmond upon Thames			Richmond upon Thames London Borough Council	Liberal Democrat	Civic Centre, 44 York Street	22.17	191,365	51.4479°N 0.3260°W	15
Southwark	✓		Southwark London Borough Council	Labour	160 Tooley Street	11.14	298,464	51.5035°N 0.0804°W	7
Sutton			Sutton London Borough Council	Liberal Democrat	Civic Offices, St Nicholas Way	16.93	195,914	51.3618°N 0.1945°W	18
Tower Hamlets	✓		Tower Hamlets London Borough Council	Labour	Town Hall, Mulberry Place, 5 Clove Crescent	7.63	272,890	51.5099°N 0.0059°W	8
Waltham Forest			Waltham Forest London Borough Council	Labour	Waltham Forest Town Hall, Forest Road	14.99	265,797	51.5908°N 0.0134°W	28
Wandsworth	✓		Wandsworth London Borough Council	Conservative	The Town Hall, Wandsworth High Street	13.23	310,516	51.4567°N 0.1910°W	5
Westminster	✓	City	Westminster City Council	Conservative	Westminster City Hall, 64 Victoria Street	8.29	226,841	51.4973°N 0.1372°W	2

Figure 2-3 London Boroughs Dataset (view from Wikipedia)

By leveraging Beautiful Soup, the table format dataset could have been extracted. Once data has been scraped, it has been turned into a below dataframe (an excerpt of the first 5 records) while applying pre-processing techniques in order to keep consistency across all the datasets.

	Borough	Population
0	Barking and Dagenham [note 1]	194352
1	Barnet	369088
2	Bexley	236687
3	Brent	317264
4	Bromley	317899

Figure 2-4 London Boroughs Dataset (view from Pandas dataframe)

London Boroughs Coordinates

Geopy API has been leveraged for obtaining London boroughs' coordinates and merged with the London boroughs dataset.

	Borough	Population	Latitude	Longitude
0	Barking and Dagenham	194352	51.554117	0.150504
1	Barnet	369088	51.653090	-0.200226
2	Bexley	236687	51.441679	0.150488
3	Brent	317264	51.563826	-0.275760
4	Bromley	317899	51.402805	0.014814

Figure 2-5 London Boroughs Coordinates

London Venues in the Respective Boroughs

Foursquare API has been utilized to acquire data on the most popular venues in the respective London boroughs, using explore API endpoint.

	BoroughName	Borough Latitude	Borough Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Barking and Dagenham	51.554117	0.150504	Central Park	51.559560	0.161981	Park
1	Barking and Dagenham	51.554117	0.150504	Lara Grill	51.562445	0.147178	Turkish Restaurant
2	Barking and Dagenham	51.554117	0.150504	Iceland	51.560578	0.147685	Grocery Store
3	Barking and Dagenham	51.554117	0.150504	Wilko	51.541002	0.148898	Furniture / Home Store
4	Barking and Dagenham	51.554117	0.150504	Shell	51.560415	0.148364	Gas Station

Figure 2-6 London Venues in the Respective Boroughs

3. Methodology

The Methodology section consists of 3 main parts:

- Exploratory data analysis serves us as the source of better understanding the distribution of the crimes across the boroughs as well as map's visualizations will help us fully grasp the geographic layout of the boroughs
- Modelling part focuses on the unsupervised learning algorithm, namely k-means through which we will cluster the similar boroughs into 5 clusters in order to be able to analyze these from venues perspective/density/type
- Cluster analysis concentrates on the overview of each group of boroughs with the venues in order to better comprehend which venues are most typical for the respective clusters and perform the classification of these into more informative categories

Exploratory Data Analysis

The London crime dataset has been processed to an aggregated form on the level of borough and crime category respectively in order to perform an analysis of the highest monthly crime incidents per borough and crime category as can be seen on the graphs below. For visualization creation, Matplotlib and Seaborn libraries have been utilized.

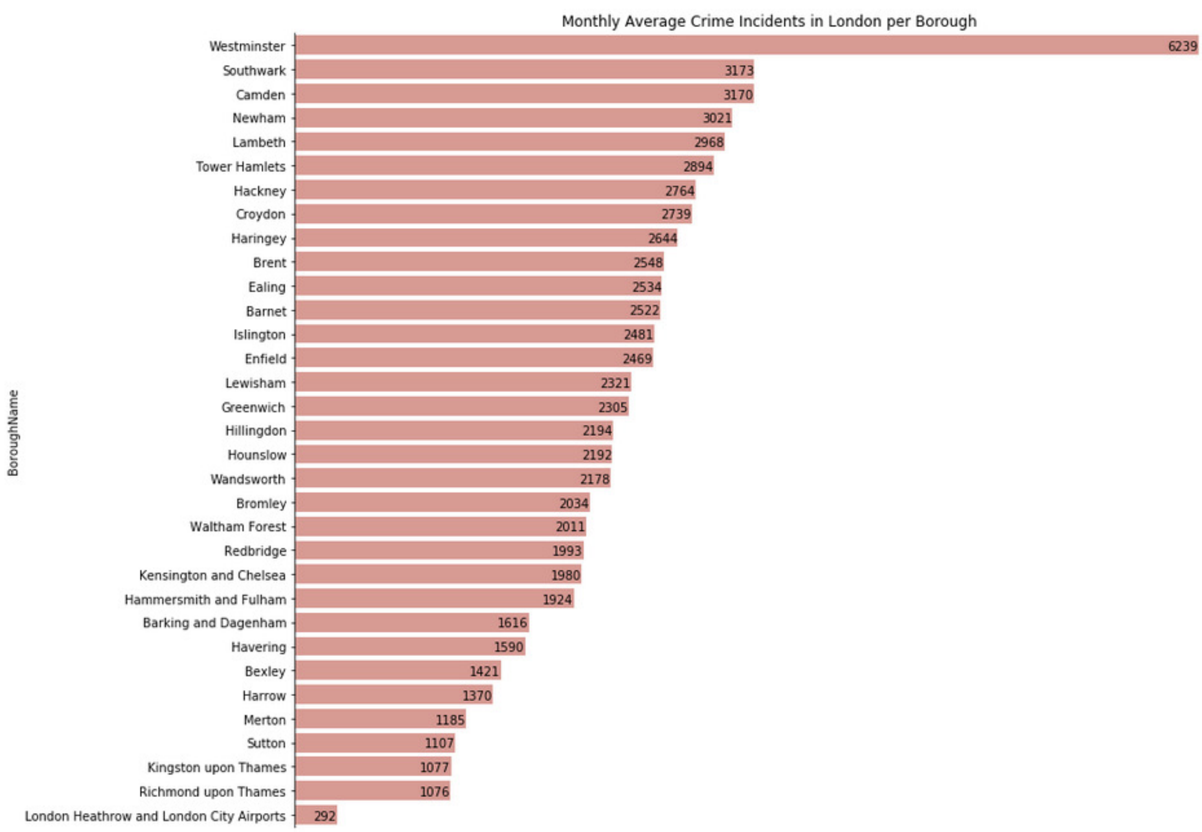

Figure 3-1 Monthly Average Crime Incidents in London per Borough

On the Figure 3-1, Westminster has the biggest number of reported crimes, followed by Southwark, Camden and Newham. When looking at the crime category breakdown, the Figure 3-2 displays the distribution of the crimes where theft and violence against person are two most occurring crime categories, followed by vehicle offences and burglary.

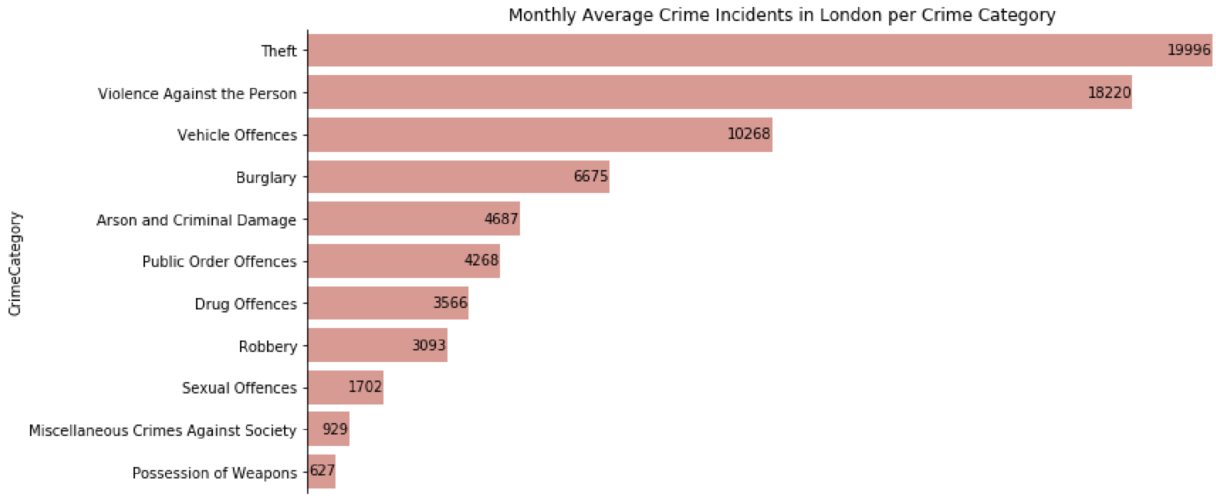

Figure 3-2 Monthly Average Crime Incidents in London per Crime Category

As each borough is different in terms of expanse or population, the crime indicator has been calculated in order to wisely reflect the population size and its effect on the crime occurrence.

	BoroughName	Total	MonthlyAverage	Population	Latitude	Longitude	CrimeIndicator
0	Westminster	149734	6238.916667	226841	51.501356	-0.124930	27.503479
1	Southwark	76149	3172.875000	298464	51.502922	-0.103458	10.630679
2	Camden	76081	3170.041667	229719	51.542305	-0.139560	13.799649
3	Newham	72506	3021.083333	318227	51.530000	0.029318	9.493485
4	Lambeth	71240	2968.333333	314242	51.501301	-0.117287	9.446011

Figure 3-3 Crime Indicator Dataframe

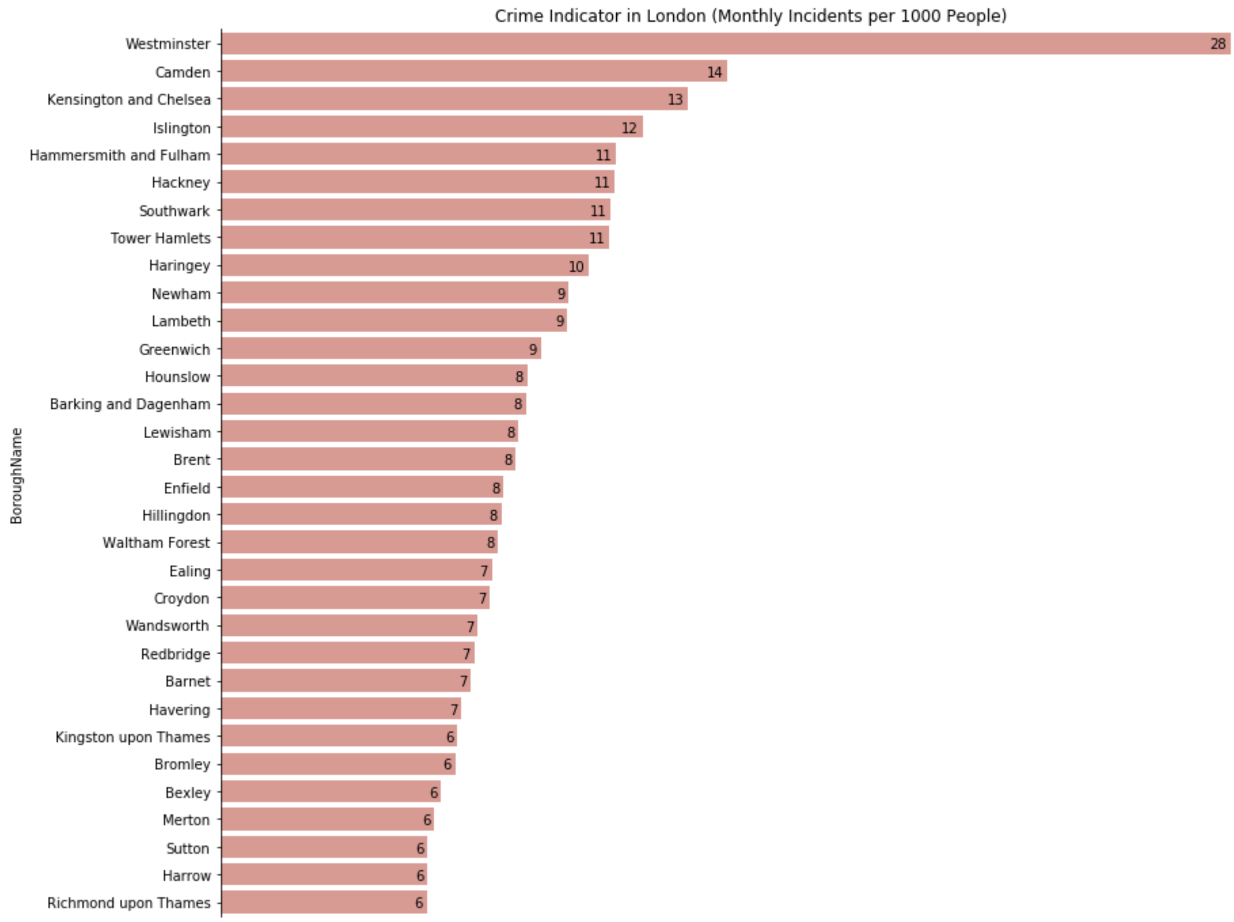

Figure 3-4 Crime Indicator in London (Monthly Incidents per 1000 People)

Westminster and Camden remain top 2 dangerous areas even after taking population into consideration. However, there is a change on the 3rd and 4th position as Kensington and Chelsea, and Islington appeared in the list.

In order to visualize the boroughs from a geographical perspective, Folium library has been leveraged to display a map where each borough is represented by a blue circle with the pop-up markers indicating its name.

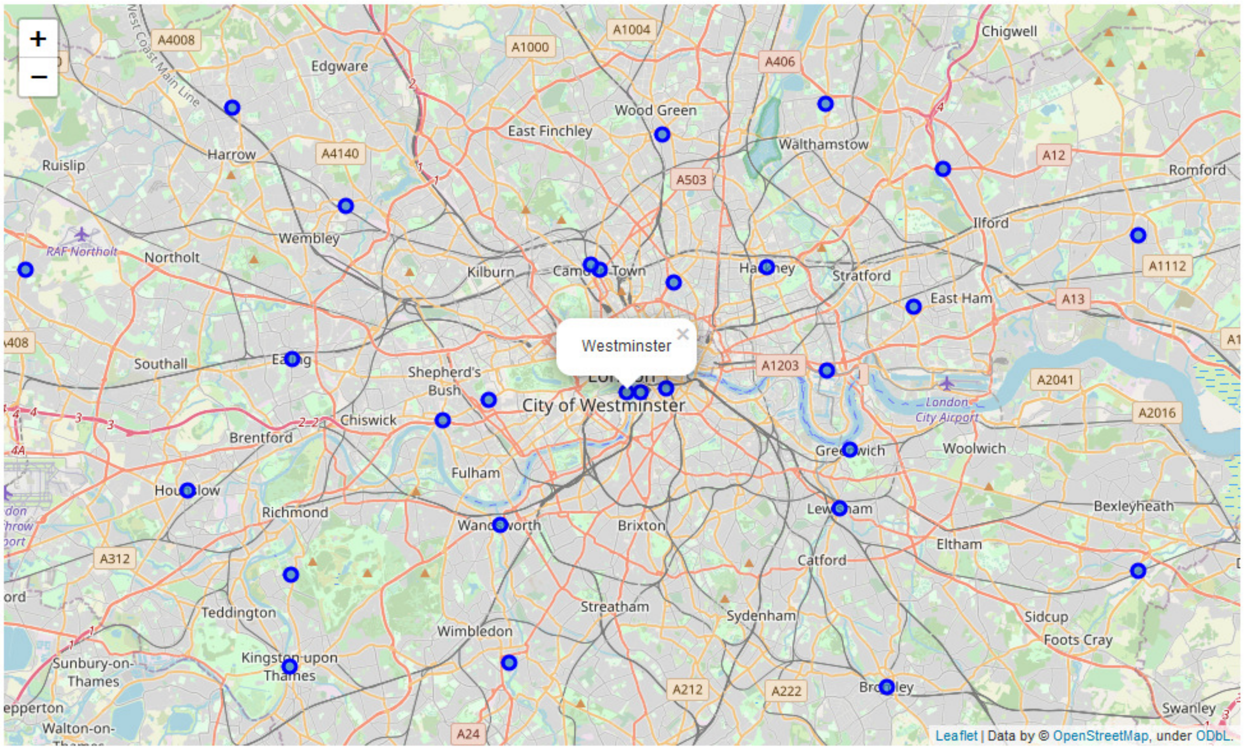

Figure 3-5 London Boroughs Map

Modelling

In this section, we will perform k-means clustering to group the boroughs according to what amenities they have, using Foursquare data. In order to utilize the k-means algorithm, one hot encoding has been used to transform the categorical data into a numerical data. Afterwards, the data is grouped by borough and the mean of frequency of each venue category is considered in the next steps when modelling.

	BoroughName	Afghan Restaurant	African Restaurant	American Restaurant	Argentinian Restaurant	Art Gallery	Art Museum	Arts & Crafts Store	Asian Restaurant	Athletics & Sports	...	Vegetarian / Vegan Restaurant	Veterinarian	Video Game Store	Video Store
0	Barking and Dagenham	0	0	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	0	0
2	Barking and Dagenham	0	0	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	0	0
4	Barking and Dagenham	0	0	0	0	0	0	0	0	0	...	0	0	0	0

Figure 3-6 One Hot Encoded Venues Dataset

	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	Supermarket	Grocery Store	Bus Station	Gas Station	Martial Arts Dojo	Bus Stop	Convenience Store	Restaurant	Park	Golf Course
1	Barnet	Coffee Shop	Pub	Park	Grocery Store	Italian Restaurant	Golf Course	Bookstore	Café	Restaurant	Pizza Place
2	Bexley	Pub	Greek Restaurant	Playground	Fast Food Restaurant	Chinese Restaurant	Fruit & Vegetable Store	Steakhouse	Garden	Gastropub	Museum
3	Brent	Coffee Shop	Hotel	Grocery Store	Clothing Store	Indian Restaurant	Sporting Goods Shop	Sandwich Place	Burger Joint	Stadium	Pedestrian Plaza
4	Bromley	Pub	Clothing Store	Gym / Fitness Center	Coffee Shop	Burger Joint	Pizza Place	Indian Restaurant	Supermarket	Stationery Store	Burrito Place

Figure 3-7 Ten Most Common Venues per Borough

As mentioned above, the k-means unsupervised algorithm has been applied to the data. It is an iterative method that partitions data into non-overlapping groups/clusters while user has to define the number of clusters upfront i.e. before modelling takes place. In this project, k (the number of clusters) has been set to 5. Boroughs with similar venues will be grouped together so people searching for the area with some specific type of venues can shortlist these and select the one that fits their preferences the best.

Cluster Analysis

Now, we look at each cluster to see what the main attributes are and create naming representing cluster's characteristics.

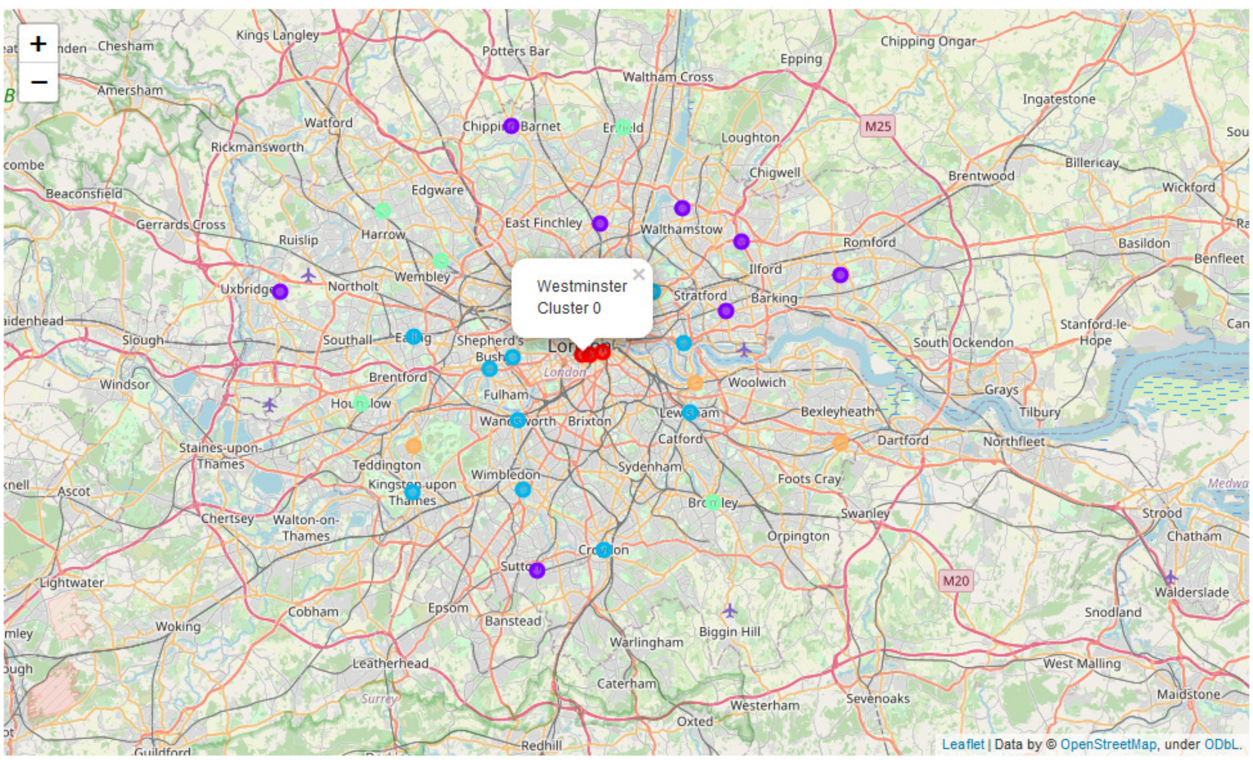

Figure 3-8 London Boroughs Clustered by Using k-means

Each cluster is color coded for the ease of presentation. We can see that majority of the neighborhood falls in the blue cluster which is the cluster No 2. Other clusters' size ranges from 3 - 8 boroughs.

Cluster 0: Cultural Area

The cluster No 0 is located in the city center so it can be considered to be a cultural and traveler area which consists of venues such as theaters, hotels, parks, restaurants, museums and galleries. This borough might be more touristic, however, it also has its own magic and atmosphere that makes this area unique.

	BoroughName	CrimeIndicator	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
0	Westminster	27.503479	0	Hotel	Coffee Shop	Café	Theater	Park	Event Space	Plaza	Art Gallery	Bar	Skate Park
1	Southwark	10.630679	0	Theater	Scenic Lookout	Coffee Shop	Hotel	Grocery Store	Italian Restaurant	Art Museum	Street Food Gathering	Seafood Restaurant	Pub
4	Lambeth	9.446011	0	Theater	Hotel	Scenic Lookout	Park	Art Gallery	Plaza	Cupcake Shop	Sandwich Place	Cocktail Bar	Coffee Shop

Figure 3-9 Cluster 0: Cultural Area

Cluster 1: Vibrant Area

In the cluster 1, we can see that the most common venues are pubs, restaurants, cafes, supermarkets, grocery stores and parks. So, everybody can find here all one needs, places to go for shopping as well as entertainment, and with gardens and parks to do some sports or some outside activities.

	BoroughName	CrimeIndicator	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
8	Haringey	10.036600	1	Turkish Restaurant	Café	Pub	Mediterranean Restaurant	Bakery	Park	Coffee Shop	Polish Restaurant	Supermarket
3	Newham	9.493485	1	Grocery Store	Pub	Indian Restaurant	Park	Café	Bus Stop	Gym / Fitness Center	Breakfast Spot	Coffee Shop
24	Barking and Dagenham	8.317383	1	Supermarket	Grocery Store	Bus Station	Gas Station	Martial Arts Dojo	Bus Stop	Convenience Store	Restaurant	Park
16	Hillingdon	7.651368	1	Pub	Grocery Store	Chinese Restaurant	Pizza Place	Park	Fast Food Restaurant	Sports Club	Gas Station	Breakfast Spot
20	Waltham Forest	7.566081	1	Pub	Turkish Restaurant	Park	Gym	Grocery Store	Café	Supermarket	Pizza Place	Bus Stop
21	Redbridge	6.912453	1	Pizza Place	Pub	Park	Bakery	Grocery Store	Italian Restaurant	English Restaurant	Indian Restaurant	Café
11	Barnet	6.833736	1	Coffee Shop	Pub	Park	Grocery Store	Italian Restaurant	Golf Course	Bookstore	Café	Restaurant
29	Sutton	5.651927	1	Pub	Grocery Store	Park	Café	Italian Restaurant	Coffee Shop	Bar	Pizza Place	Bakery

Figure 3-10 Cluster 1: Vibrant Area

Cluster 2: Cosmopolitan Area

The second cluster is the biggest one in terms of number of boroughs (13 out of 32 boroughs) that have been classified into this group. One can see that majority of venues are pubs, restaurants, cafes and gyms.

As there are many boroughs to choose from, this offers us a wide range of possibilities to search for a place to live while having all the main services available in the walking distance.

	BoroughName	CrimeIndicator	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	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
28	Merton	5.829393	2	Park	Pub	Italian Restaurant	Café	Coffee Shop	Sushi Restaurant	Thai Restaurant	Theater	Portuguese Restaurant
30	Kingston upon Thames	6.459604	2	Café	Pub	Coffee Shop	Burger Joint	Thai Restaurant	Sushi Restaurant	Cajun / Creole Restaurant	German Restaurant	Latin American Restaurant
25	Havering	6.566183	2	Coffee Shop	Pub	Market	Yoga Studio	Vegetarian / Vegan Restaurant	French Restaurant	Beer Store	Music Venue	Fish Market
18	Wandsworth	7.013863	2	Pub	Coffee Shop	Park	Gastropub	Pizza Place	Supermarket	Café	Gym	Bar
7	Croydon	7.348273	2	Pub	Coffee Shop	Clothing Store	Mediterranean Restaurant	Bookstore	Portuguese Restaurant	Hotel	Park	Indian Restaurant
10	Ealing	7.398061	2	Pub	Coffee Shop	Café	Park	Wine Bar	Hotel	Italian Restaurant	Burger Joint	Pizza Place
14	Lewisham	8.110717	2	Pub	Café	Park	Food Truck	Gastropub	Gym	Turkish Restaurant	Farmers Market	Restaurant
5	Tower Hamlets	10.604395	2	Gym / Fitness Center	Hotel	Coffee Shop	Plaza	Bakery	Street Food Gathering	Park	English Restaurant	Italian Restaurant
6	Hackney	10.737084	2	Pub	Bakery	Coffee Shop	Butcher	Cocktail Bar	Café	Gastropub	Park	Wine Shop
23	Hammersmith and Fulham	10.770350	2	Pub	Coffee Shop	Park	Japanese Restaurant	French Restaurant	Café	Pizza Place	Grocery Store	Cocktail Bar
12	Islington	11.503073	2	Pub	Bakery	Japanese Restaurant	Gastropub	Cocktail Bar	Trail	Arts & Crafts Store	Theater	Movie Theater
22	Kensington and Chelsea	12.724623	2	Pub	Gym / Fitness Center	Restaurant	Hotel	Italian Restaurant	Café	Pizza Place	Japanese Restaurant	Bakery
2	Camden	13.799649	2	Pub	Coffee Shop	Café	Market	Music Venue	Vegetarian / Vegan Restaurant	Italian Restaurant	Pizza Place	Portuguese Restaurant

Figure 3-11 Cluster 2: Cosmopolitan Area

Cluster 3: Busy Area

The cluster No 3 has many coffee shops, clothing stores, restaurants and pubs so if one enjoys lively area this is the correct choice for him.

	BoroughName	CrimeIndicator	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
17	Hounslow	8.352959	3	Clothing Store	Coffee Shop	Indian Restaurant	Pub	Sandwich Place	Fast Food Restaurant	Discount Store	Hotel	Supermarket
9	Brent	8.031429	3	Coffee Shop	Hotel	Grocery Store	Clothing Store	Indian Restaurant	Sporting Goods Shop	Sandwich Place	Burger Joint	Stadium
13	Enfield	7.703921	3	Coffee Shop	Pub	Clothing Store	Sandwich Place	Turkish Restaurant	Indian Restaurant	Supermarket	Department Store	Grocery Store
19	Bromley	6.397472	3	Pub	Clothing Store	Gym / Fitness Center	Coffee Shop	Burger Joint	Pizza Place	Indian Restaurant	Supermarket	Stationery Store
27	Harrow	5.629242	3	Indian Restaurant	Coffee Shop	Park	Grocery Store	Afghan Restaurant	Café	Sports Bar	Chinese Restaurant	Fish & Chips Shop

Figure 3-12 Cluster 3: Busy Area

Cluster 4: Quiet Area

The last cluster has only 3 boroughs where one can find mainly pubs, parks, gardens, playgrounds and restaurants. Therefore, this is a borough that is more peaceful, however, one can find here all the family needs for living.

	BoroughName	CrimeIndicator	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
15	Greenwich	8.732374	4	Pub	Garden	Park	Café	Market	Turkish Restaurant	Grocery Store	History Museum	Historic Site
26	Bexley	6.002829	4	Pub	Greek Restaurant	Playground	Fast Food Restaurant	Chinese Restaurant	Fruit & Vegetable Store	Steakhouse	Garden	Gastropub
31	Richmond upon Thames	5.621674	4	Pub	Café	Park	Garden	Grocery Store	Playground	Hotel	Historic Site	Scenic Lookout

Figure 3-13 Cluster 4: Quiet Area

4. Results and Discussion

The aim of this project is to select the safest area, assembled by the venue's similarity. Therefore, we have sorted the dataframe based on the Crime Indicator field, assigning each Borough the Cluster from k-means modelling section. Now, per one's preferences, newcomers to London can shortlist the areas of their interest.

It highly depends what people are looking for so if somebody prefers vibrant or busy areas, these can be found in the boroughs clustered as No 1 and 3. On the other side, if somebody is seeking a quiet area with many parks and playgrounds, then cluster No 4 is ideal for him/her. Cluster 0 is cultural one, however, as we can see it has high criminal indicator. Thus, the factor of safety plays an important role when deciding which borough to select.

In the below dataframe, one can find both aspects (crime indicator along with cluster specifics) in one place. In the top 5 safest boroughs, 4 types of clusters (Quiet, Busy, Cosmopolitan and Vibrant) are listed, so everybody can find a borough differentiating itself based on the venues' types while the crime indicator remains low.

	BoroughName	CrimeIndicator	Total	MonthlyAverage	Population	Latitude	Longitude	Cluster Labels	Cluster Name	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
0	Richmond upon Thames	5.621674	25819	1075.791667	191365	51.440461	-0.305519	4	Quiet Borough	Pub	Café	Park
1	Harrow	5.629242	32880	1370.000000	243372	51.596769	-0.337275	3	Busy Borough	Indian Restaurant	Coffee Shop	Park
2	Sutton	5.651927	26575	1107.291667	195914	51.357511	-0.173640	1	Vibrant Borough	Pub	Grocery Store	Park
3	Merton	5.829393	28432	1184.666667	203223	51.410803	-0.188099	2	Cosmopolitan Borough	Park	Pub	Italian Restaurant
4	Bexley	6.002829	34099	1420.791667	236687	51.441679	0.150488	4	Quiet Borough	Pub	Greek Restaurant	Playground

Figure 4-1 Top 5 Safest Boroughs with Cluster Area

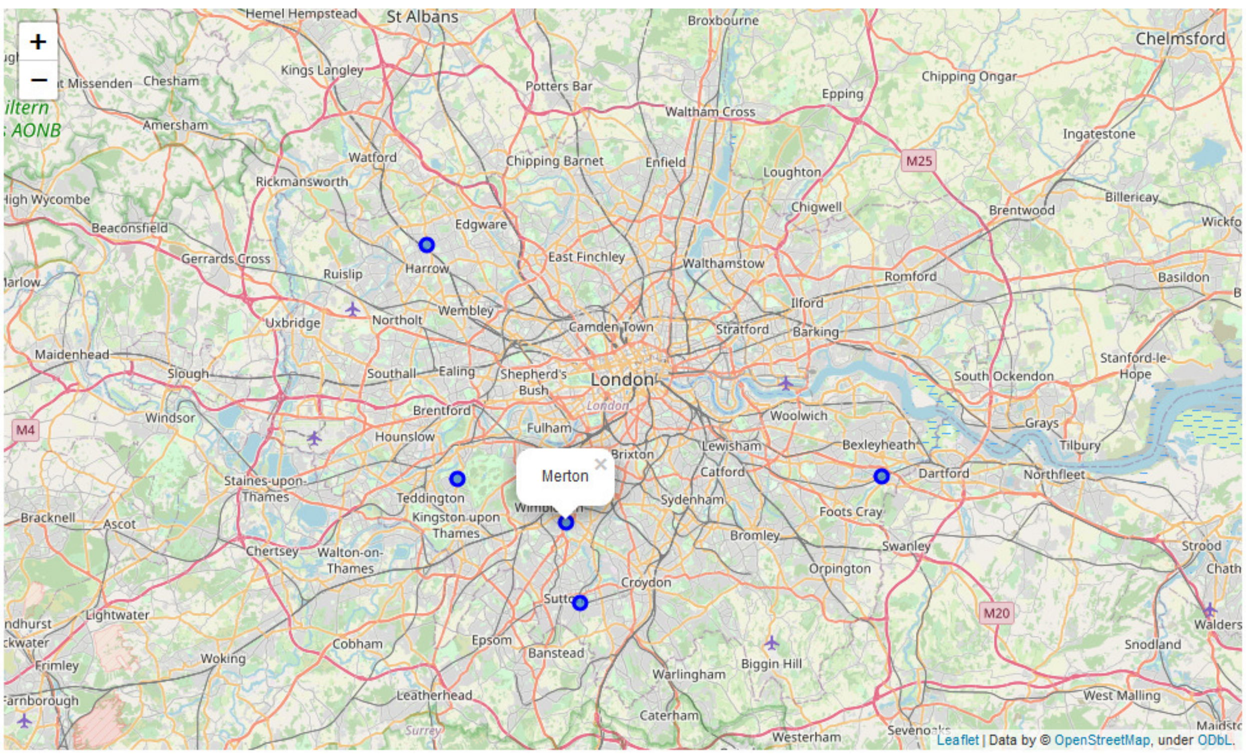

Figure 4-2 Top 5 Safest Boroughs Map

5. Conclusion

From our data-driven analysis, one can find useful information on the overview of the current safety status along with the most common venues. As one can see, all the safest areas are suburban ones, in south-west, east and north London. The analytics comprises of two perspectives of the respective boroughs, so it would be helpful to incorporate also other factors, such as the real estate prices, time to commute to work/school, cost of living, budget limitations, etc. The crime dataset provides the details on the type of the crime so this could be also included into analyzing safety in order to distinguish the seriousness of the committed crime and this way create weights that could increase the accuracy of the safety indicator. However, the current research successfully performed the high-level analysis and provides people searching for a place to live in London a recommendation based on the most popular venues types and safety. This project enables/helps future Londoners to make more informed decision when looking for a safe borough in London.