Re-locating in London: insights into London Borough similarities

Matthew Lawrence

8th May, 2020

1.1 Introduction

London is the largest city, capital of the UK and home to over 8.9 million people. Private renters in London account for over 30% of this population and a high proportion of tenants move frequently from one property to the next, to accommodate changes in lifestyle and working environments. Finding a suitable locale is a top priority for tenants in addition to the other challenges of renting. It would be of great benefit if tenants were able to gain insights on similar boroughs to facilitate informed decision making in where to re-locate in London.

1.2 Problem

Renting in London is expensive, fast-paced and often confusing. The rental market is challenging, primarily due to the fact that properties are advertised and let in such a short period of time. This makes it challenging for tenants and in many cases tenants have to make quick decisions on whether they would like to view a property.

London is diverse, international and multicultural city with many different communities and neighbourhoods. Distinct differences can also be seen between the various boroughs in the city. Prior to 1985, the Greater London Council acted as a single administrative body for the London area. However after it was abolished in 1985, additional power was devolved to the London boroughs. As a result, the London boroughs have significant character variation and this can be evidenced in the number, spread and type of venues in each.

For individuals living in London and looking to change locale, the borough in which they live is an important feature. In many cases, knowledge of these boroughs is poor and individuals run the risk of moving to a borough where they are not satisfied with local amenities and services.

Therefore, the problem statement can be summed up as 'tenants in London lack knowledge of how other boroughs compare to the one in which they live, which makes re-locating across borough boundaries difficult'.

1.3 Interest

Individuals currently renting in London and looking to move (due to changes in workplace location, family or other situation) would be interested in London borough similarity insights to allow them to make informed decisions with regards to moving. Borough similarity insights would allow renters to locate a Borough similar to their current one or indeed choose a borough that was different and more aligned with their preferences.

2. 1 Data Acquisition and Cleaning

2.3 Data Sources

To provide some findings to shed some light on the problem described, data will be needed. This will be formed of two key sources.

Firstly, London borough data will be needed to provide the names of the boroughs and their centre-point coordinates. This will form the basis of the data frame used to undertake clustering for the boroughs and is essential as we will need a list of borough names to identify boroughs in the resulting clusters.

Secondly, venue data from Foursquare will be required. The explore function will be used with the Foursquare API to obtain a list of venues within a certain radius of the boroughs centre-point coordinates. It is by obtaining this venue data (for each borough) that I will be able to build up a picture of how similar the boroughs are based on the venues within them.

Prior to clustering and analysis, data on the London boroughs and the Foursquare location data will have to be merged.

2. 3 Data Cleaning

The London borough data was cleaned to allow clustering analysis to be successfully undertaken. Wikipedia has a tendency to include references and indexes within tables and text to allow readers to see the source origin of the information. However this was not needed for the project and so these were removed using the stripping technique.

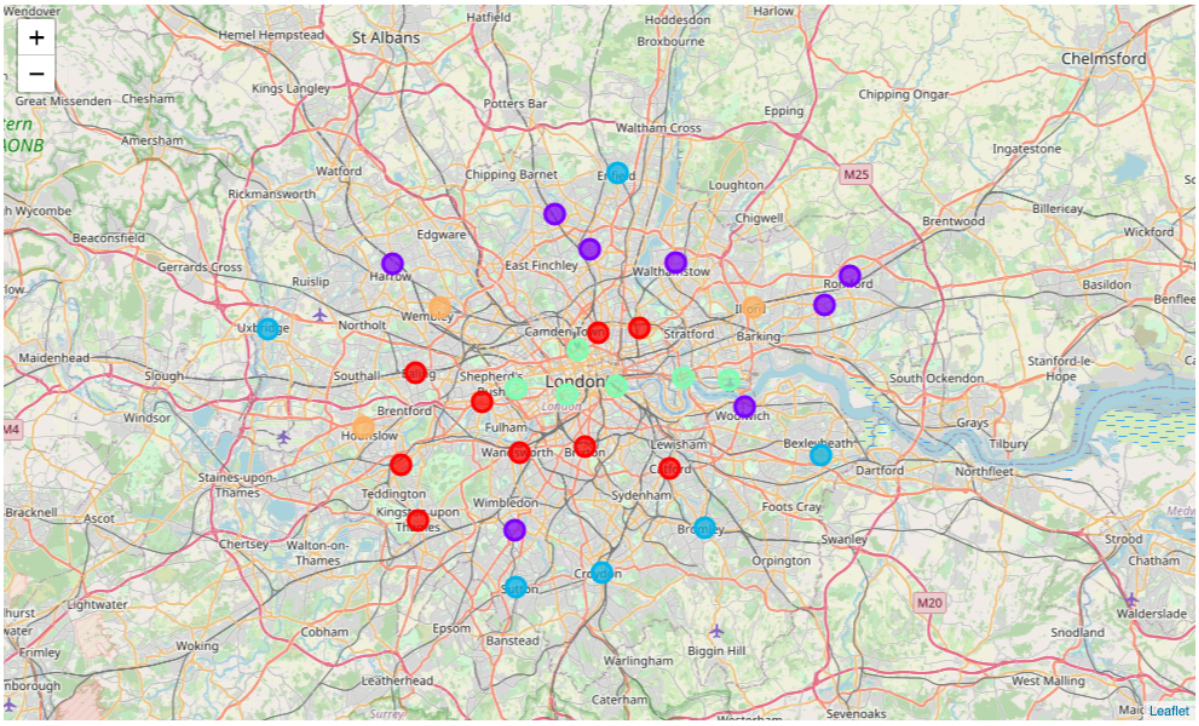
2.4 Feature Selection

Not all of the information included on the London boroughs source was relevant and useful for the project and so some of these columns were removed as they were redundant for the project.

3.1 Methodology

3.2 Clustering

To analyse the data and determine similarities between clusters in terms of the venues, amenities and services most prominent within them, clustering data analysis was used with k-means clustering. The output of this data analysis would be five clusters of London boroughs, based on the type of venues within them. The figure below shows a map of London with the different borough clusters depicted with corresponding colours.



4.1 Results

The results of the clustering can be depicted in a map like the one above. Furthermore, the groupings of the clusters are as follows.

Cluster 1

	Borough_Name	Borough_Name	Population	Coordinates	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
7	Ealing	Ealing	342494	51°30′47″N 0°18′32″W / 51.5130°N 0.3089°W /...	51.5130	-0.3089	0	Pub	Coffee Shop	Park	Hotel	Pizza Place
10	Hackney	Hackney	257379	51°32′42″N 0°03′19″W / 51.5450°N 0.0553°W /...	51.5450	-0.0553	0	Pub	Coffee Shop	Café	Bakery	Cocktail Bar
11	Hammersmith and Fulham	Hammersmith and Fulham	178685	51°29′34″N 0°14′02″W / 51.4927°N 0.2339°W /...	51.4927	-0.2339	0	Pub	Café	Coffee Shop	Park	Gastropub
17	Islington	Islington	215667	51°32′30″N 0°06′08″W / 51.5416°N 0.1022°W /...	51.5416	-0.1022	0	Pub	Park	Café	Gastropub	Theater
19	Kingston upon Thames	Kingston upon Thames	166793	51°24′31″N 0°18′23″W / 51.4085°N 0.3064°W /...	51.4085	-0.3064	0	Pub	Café	Coffee Shop	Thai Restaurant	Burger Joint

Cluster 2

	Borough_Name	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 Co
0	Barking and Dagenham	1	Grocery Store	Supermarket	Park	Platform	Gas Station	Bus Stop	History Museum	Metro Station	Restaurant	Breakf Spot
1	Barnet	1	Coffee Shop	Pub	Café	Grocery Store	Pharmacy	Italian Restaurant	Park	Hotel	Turkish Restaurant	Superr
9	Greenwich	1	Pub	Grocery Store	Coffee Shop	Park	Supermarket	Hotel	Fast Food Restaurant	Pharmacy	Bakery	Clothir Store
12	Haringey	1	Café	Pub	Park	Grocery Store	Mediterranean Restaurant	Turkish Restaurant	Greek Restaurant	Bakery	Gym / Fitness Center	Coffee
13	Harrow	1	Coffee Shop	Indian Restaurant	Pub	Park	Grocery Store	Fast Food Restaurant	Sandwich Place	Café	Gym / Fitness Center	Clothir Store

Cluster 3

	Borough_Name	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 Co
2	Bexley	2	Pub	Supermarket	Clothing Store	Coffee Shop	Italian Restaurant	Fast Food Restaurant	Hotel	Grocery Store	Bakery	Pharma
4	Bromley	2	Pub	Gym / Fitness Center	Clothing Store	Coffee Shop	Pizza Place	Indian Restaurant	Park	Indie Movie Theater	Department Store	Sandwi Place
6	Croydon	2	Pub	Coffee Shop	Clothing Store	Hotel	Park	Supermarket	Café	Indian Restaurant	Furniture / Home Store	Mediter Restaur
8	Enfield	2	Pub	Supermarket	Coffee Shop	Grocery Store	Pizza Place	Clothing Store	Train Station	Gym / Fitness Center	Pharmacy	Fast Fo Restaur
15	Hillingdon	2	Coffee Shop	Pub	Italian Restaurant	Sandwich Place	Pharmacy	Gym / Fitness Center	Supermarket	Restaurant	Gym	Clothin
27	Sutton	2	Pub	Grocery Store	Coffee Shop	Italian Restaurant	Clothing Store	Park	Café	Pizza Place	Supermarket	Indian Restaur

Cluster 4

	Borough_Name	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
5	Camden	3	Coffee Shop	Hotel	Pizza Place	Cocktail Bar	Gym / Fitness Center	History Museum	Wine Bar	Sushi Restaurant	Café	Bookstore
18	Kensington and Chelsea	3	Pub	Hotel	Garden	Café	Indian Restaurant	Park	Italian Restaurant	Gym / Fitness Center	Restaurant	Science Museum
23	Newham	3	Hotel	Coffee Shop	Gym / Fitness Center	Café	Grocery Store	Pub	Sandwich Place	Light Rail Station	Harbor / Marina	Gas Station
26	Southwark	3	Hotel	Coffee Shop	Cocktail Bar	Scenic Lookout	Theater	Art Museum	Pub	Brewery	Gym / Fitness Center	Grocery Store
28	Tower Hamlets	3	Coffee Shop	Hotel	Burger Joint	Bar	Pub	Gym / Fitness Center	Park	Italian Restaurant	Plaza	Lounge
31	Westminster	3	Hotel	Plaza	Cocktail Bar	Park	Garden	Art Gallery	Café	Art Museum	Bakery	Lounge

Cluster 5

	Borough_Name	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
3	Brent	4	Coffee Shop	Indian Restaurant	Hotel	Clothing Store	Grocery Store	Sandwich Place	Pizza Place	Sporting Goods Shop	Bar	Gym / Fitness Center
16	Hounslow	4	Indian Restaurant	Coffee Shop	Clothing Store	Hotel	Convenience Store	Grocery Store	Pub	Supermarket	Metro Station	Park
24	Redbridge	4	Grocery Store	Supermarket	Indian Restaurant	Coffee Shop	Fast Food Restaurant	Clothing Store	Irish Pub	Department Store	Sandwich Place	Pizza Place

5.1 Discussion

Based on the clusters depicted in the map above, I have deduced borough cluster characteristics as follows:

Cluster 1 - Residential with lots of pubs, cafes, coffee shops and parks This cluster is characterised by a high number of Pubs, Coffee Shops, Cafes and Parks. This is logical since

these boroughs are not directly in the centre of London and therefore are more likely to have venues suited to the primarily residential communities that surround them. I can interpret this borough to be primarily residential and to offer lots of amenities such as pubs, cafes, coffee shops and parks.

Cluster 2 - Residential with large number of supermarkets and quiet nightlife. This cluster of boroughs is characterised by high numbers of coffee shops, grocery stores and supermarkets. This is again logical, as due to these clusters being further out from the high-price of land in central London, there is financial opportunity to build large supermarkets for shopping. These boroughs are also primarily residential, however, in comparison to boroughs in cluster 1, there are large supermarkets for shopping. Finally, less pubs suggests that these boroughs have a quieter night life.

Cluster 3 - Lots of shopping opportunities and nightlife. Typically boroughs in cluster 3 are the furthest out from the centre of London. The cluster is characterised by a large number of pubs and coffee shops, however, there are also a large number of clothing stores and restaurants. This suggests that for shopping trips, people living in these boroughs are not willing to make the trip into central London to shop and instead prefer to shop locally.

Cluster 4 - Tourist areas with hotels and restaurants. The boroughs in cluster 4 have a distinct prominence of hotels. This suggests that these boroughs are the prime location for tourists and business people to stay when in London. There are a large number of restaurants, cafes and gyms which suggest that the majority of amenities cater to business people working in the vicinity. There are a number of pubs and cocktail bars also

Cluster 5 - Residential with large number of Indian restaurants. Cluster 5 is interesting as there is a significant prominence of Indian restaurants. This suggests potentially a large number of residents in the local area with Indian heritage. Other than the Indian restaurants, these boroughs appear to be very similar to the boroughs identified in Cluster 2 and so we can deduce therefore that these boroughs are primarily residential and offer similar services and venues as the boroughs in cluster 2 (with added Indian restaurants).

6.1 Conclusions

It is certainly a useful exercise to conduct London borough similarity analysis and indeed this provides useful information for those currently living in London to make decisions with regards to re-locating across borough boundaries. Whilst the clustering is accurate to some extent, to further enhance the accuracy of the results, the proportion of venues relative to population could be calculated as this would provide a more accurate picture of the relative number of venues as opposed to the absolute number which my analysis focuses on.