

Capstone Project Week 5

Battle of the Neighbourhoods

1 Introduction

The client, Mr Roy, would like to start up an Indian Restaurant in one of the London Boroughs. He currently owns two successful restaurants in Birmingham but is unfamiliar with the London Boroughs and which one would be best suited to set up his third restaurant.

Mr Roy's restaurant would be opened 7 days a week in the evening only and ideally, he would like there to be passing trade from people enjoying an evening out. Mr Roy's restaurant would also provide a take-away service and he would offer free delivery within a 3-mile radius.

Mr Roy would like to understand the make-up of the London boroughs and would like a recommendation on which ones best meet his criteria. Areas with evening entertainment but also a good residential area to requiring a take-away service.

2 Data Description

2.1 Data Sources

The following wiki page will be used to obtain the London Boroughs, the geo data and populations.

https://en.wikipedia.org/wiki/List_of_London_boroughs

The following CSV document shows the business success for each of the London Boroughs. This can be merged with the wiki data to give a further factor for later analysis. The 5-year survival will be used.

<https://data.london.gov.uk/dataset/business-demographics-and-survival-rates-borough>

Downloaded data: Business Survival Rate (CSV)

The Foursquare API will be used to show venue categories for the London Boroughs. The London Borough central geo data will be used as the central search point. As the boroughs differ in size considerably, the search radius will be calculated as the square root of the borough area and converted the figure to metres.

Business survival rates were captured as a GSV from the London local government website. This was

	Borough	Area	Latitude	Longitude	Radius	5_year_survival_rate
0	Barking and Dagenham	13.93	51.5607	0.1557	3003	39.6
1	Barnet	33.49	51.6252	-0.1517	4656	40.9
2	Bexley	23.38	51.4549	0.1505	3890	44.1
3	Brent	16.70	51.5588	-0.2817	3288	39.2
4	Bromley	57.97	51.4039	0.0198	6126	44.6

Paraphrase	Latitude	Longitude	Radius	Verpus	Verpus Latitude	Verpus Longitude	Verpus Category
------------	----------	-----------	--------	--------	-----------------	------------------	-----------------

[illegible]

3 Design Methodology

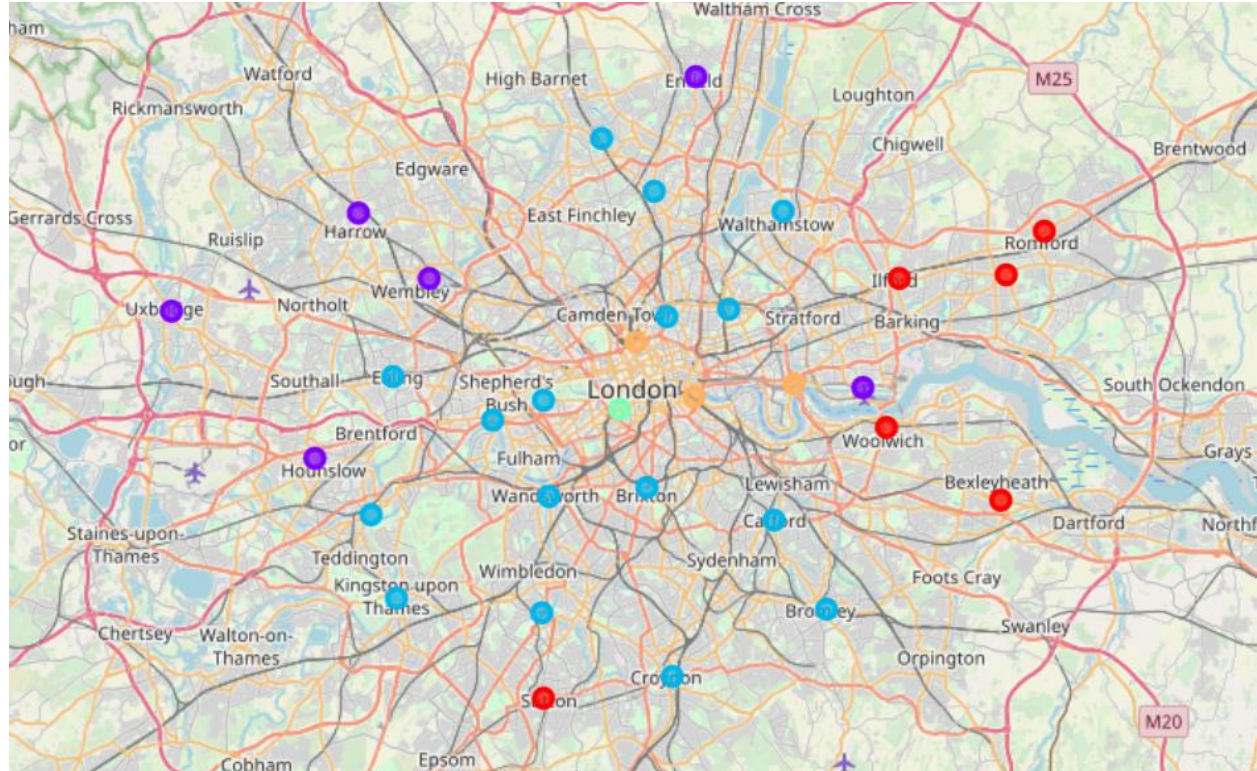
To gain insight into the different London boroughs, based on the Foursquare API data K-means clustering was used.

To achieve this the incidence of each venue category was found for each borough using one-hot coding. The average for each venue type was taken for each borough. This process gave a table of 32 boroughs by 263 venue categories. A sample is shown below.

	Borough	Afghan Restaurant	African Restaurant	Airport	American Restaurant	Aquarium	Argentinian Restaurant	Art Gallery	Art Museum	Arts & Crafts Store	...	Video Game Store	Vietnamese Restaurant
27	Sutton	0.0	0.0	0.0	0.0125	0.0	0.00	0.00	0.00	0.00	...	0.0125	0.0
28	Tower Hamlets	0.0	0.0	0.0	0.0000	0.0	0.01	0.01	0.00	0.00	...	0.0000	0.0
29	Waltham Forest	0.0	0.0	0.0	0.0000	0.0	0.00	0.02	0.00	0.01	...	0.0000	0.0
30	Wandsworth	0.0	0.0	0.0	0.0000	0.0	0.00	0.00	0.00	0.01	...	0.0000	0.0
31	Westminster	0.0	0.0	0.0	0.0100	0.0	0.01	0.03	0.03	0.00	...	0.0000	0.0

K-means clustering was used on the above data (K=5 assumed) and each borough was assigned to 1 of 5 clusters that had the closest similarity to other boroughs.

The borough clusters are shown below.



Colour Cluster Number

Red 1

Purple 2

Blue 3

Green 4

Orange 5

In order to understand the characteristics of the cluster based on the venues in the boroughs, the top 10 venue categories for each borough were found and this was merged with the cluster data. Each borough could then be evaluated.

Cluster 1

	Borough	Area	5_year_survival_rate	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	Barking and Dagenham	13.93	39.6	0	Grocery Store	Coffee Shop	Park	Supermarket	Furniture / Home Store	Restaurant	Pub	Clothing Store	Fast Food Restaurant	Multiplex
2	Bexley	23.38	44.1	0	Grocery Store	Pub	Supermarket	Train Station	Italian Restaurant	Fast Food Restaurant	Coffee Shop	Pizza Place	Clothing Store	Convenience Store
9	Greenwich	18.28	40.0	0	Pub	Grocery Store	Clothing Store	Coffee Shop	Hotel	Supermarket	Gym / Fitness Center	Park	Furniture / Home Store	Fast Food Restaurant
14	Havering	43.35	45.2	0	Coffee Shop	Pub	Grocery Store	Park	Furniture / Home Store	Supermarket	Italian Restaurant	Fast Food Restaurant	Café	Clothing Store
24	Redbridge	21.78	39.9	0	Park	Grocery Store	Indian Restaurant	Supermarket	Pub	Bakery	Coffee Shop	Fast Food Restaurant	Clothing Store	Train Station
27	Sutton	16.93	45.1	0	Pub	Grocery Store	Coffee Shop	Park	Supermarket	Italian Restaurant	Sandwich Place	Café	Train Station	Clothing Store

There is a high incidence of grocery stores and pubs in these boroughs. This would indicate these are residential areas. There is a good mix of venues, such as parks, restaurants, shops and coffee shops indicating these are busy areas during the day, but night life is limited to pubs and a few restaurants. These boroughs are on the outskirts of London, generally to the East.

Cluster name: Residential

Cluster 2

	Borough	Area	5_year_survival_rate	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	16.70	39.2	1	Indian Restaurant	Coffee Shop	Hotel	Clothing Store	Sandwich Place	Gym / Fitness Center	Grocery Store	Pizza Place	Café	Portuguese Restaurant
8	Enfield	31.74	43.7	1	Coffee Shop	Pub	Gym / Fitness Center	Garden Center	Supermarket	Turkish Restaurant	Hotel	Grocery Store	Greek Restaurant	Café
13	Harrow	19.49	43.4	1	Indian Restaurant	Coffee Shop	Pub	Gym / Fitness Center	Grocery Store	Supermarket	Park	Café	Italian Restaurant	Pizza Place
15	Hillingdon	44.67	44.3	1	Pub	Indian Restaurant	Coffee Shop	Gym / Fitness Center	Park	Golf Course	Hotel	Grocery Store	Supermarket	Bar
16	Hounslow	21.61	43.4	1	Pub	Indian Restaurant	Coffee Shop	Hotel	Park	Rugby Stadium	Pizza Place	Grocery Store	Gym	Gym / Fitness Center
23	Newham	13.98	36.8	1	Hotel	Coffee Shop	Pub	Park	Gym / Fitness Center	Clothing Store	Grocery Store	Supermarket	Thai Restaurant	Toy / Game Store

There is a high incidence of Indian restaurants, along with coffee shops, pubs and hotels. This also looks to be a residential area but livelier in the evening. The boroughs are on the outskirts of London, generally to the west and north.

Cluster name: Residential with love of Indian food

Cluster 3

	Borough	Area	5_year_survival_rate	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
1	Barnet	33.49	40.9	2	Café	Coffee Shop	Turkish Restaurant	Park	Pub	Grocery Store	Gym / Fitness Center	Greek Restaurant	Bakery	Italian Restaurant
4	Bromley	57.97	44.6	2	Pub	Coffee Shop	Park	Gym / Fitness Center	Pizza Place	Italian Restaurant	Indian Restaurant	Gastropub	Supermarket	Mediterranean Restaurant
6	Croydon	33.41	39.4	2	Pub	Coffee Shop	Grocery Store	Park	Indian Restaurant	Pizza Place	Café	Clothing Store	Mediterranean Restaurant	Caribbean Restaurant
7	Ealing	21.44	42.5	2	Pub	Park	Coffee Shop	Hotel	Grocery Store	Pizza Place	Gastropub	Gym / Fitness Center	Brewery	Italian Restaurant
10	Hackney	7.36	43.5	2	Pub	Café	Coffee Shop	Bakery	Yoga Studio	Brewery	Bookstore	Cocktail Bar	Pizza Place	Farmers Market
11	Hammersmith and Fulham	6.33	35.5	2	Pub	Café	Coffee Shop	Gastropub	Park	Indian Restaurant	Hotel	Thai Restaurant	French Restaurant	Pizza Place
12	Haringey	11.42	40.7	2	Café	Turkish Restaurant	Coffee Shop	Pub	Park	Bakery	Pizza Place	Mediterranean Restaurant	Indian Restaurant	Greek Restaurant
17	Islington	5.74	38.1	2	Pub	Café	Coffee Shop	Bakery	Mediterranean Restaurant	Indian Restaurant	Theater	Pizza Place	Park	Music Venue
18	Kensington and Chelsea	4.68	38.9	2	Restaurant	Pub	Italian Restaurant	Café	Gym / Fitness Center	Bakery	Hotel	English Restaurant	Garden	Indian Restaurant
19	Kingston upon Thames	14.38	50.0	2	Pub	Café	Coffee Shop	Park	Garden	Gastropub	Hotel	Italian Restaurant	Thai Restaurant	Sushi Restaurant
20	Lambeth	10.36	25.0	2	Coffee Shop	Pub	Park	Pizza Place	Market	Brewery	Café	Restaurant	Tapas Restaurant	Caribbean Restaurant
21	Lewisham	13.57	42.4	2	Pub	Coffee Shop	Café	Gastropub	Park	Italian Restaurant	Indian Restaurant	Supermarket	Turkish Restaurant	Food Truck
22	Merton	14.52	42.5	2	Park	Pub	Sushi Restaurant	Supermarket	Café	Theater	Fast Food Restaurant	Italian Restaurant	Coffee Shop	Lebanese Restaurant
25	Richmond upon Thames	22.17	45.8	2	Pub	Park	Italian Restaurant	Café	Garden	Coffee Shop	Rugby Stadium	Hotel	Bakery	Restaurant
29	Waltham Forest	14.99	37.5	2	Pub	Coffee Shop	Café	Pizza Place	Park	Brewery	Restaurant	Bakery	Turkish Restaurant	Supermarket
30	Wandsworth	13.23	39.9	2	Park	Pub	Coffee Shop	Pizza Place	Café	Bakery	Supermarket	French Restaurant	Thai Restaurant	Breakfast Spot

These boroughs really are dominated pubs and coffee shops and restaurants. There are few grocery stores, indicating the local people enjoy eating out during the day and night.

Cluster Name: Pub and café society

Cluster 4

	Borough	Area	5_year_survival_rate	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
31	Westminster	8.29	35.6	3	Hotel	Boutique	Theater	Clothing Store	Garden	Plaza	Park	Art Gallery	Art Museum	Burger Joint

This cluster is clearly unique, dominated by hotels and theatres, and cultural venues during the day. Central London.

Cluster name: Cultural

Cluster 5

	Area	5_year_survival_rate	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	8.40	39.5	4	Coffee Shop	Bookstore	Hotel	Pizza Place	Theater	Seafood Restaurant	Bakery	Steakhouse	Canal	Grocery Store
26	11.14	39.8	4	Coffee Shop	Hotel	Cocktail Bar	Italian Restaurant	Scenic Lookout	Theater	Street Food Gathering	Steakhouse	Café	Grocery Store
28	7.63	35.9	4	Coffee Shop	Burger Joint	Hotel	Bar	Plaza	Park	Lounge	Italian Restaurant	Pub	Gym / Fitness Center

These boroughs are dominated by coffee shops but would be quieter in the evening. Perhaps less residential but more for tourists with its plentiful hotels. This is central London.

Cluster name: Coffee shop and tourist.

4 Results

In the previous section, the London boroughs were clustered, based on venue category incidence, to other boroughs which were most similar, and the clusters were given label that best summed up the results based on the observations that could be taken from the data.

The question is, what is the best borough for Mr. Roy to open his new Indian restaurant?

Based on his requirements. From the data that is available cluster 2, residential with a love of Indian food, would be the most obvious suggestion. Clearly in these areas there are numerous Indian restaurants. Newham stands out as not having many Indian restaurants but has the other characteristics of areas that do. There are hotels in this area, so there could be an opportunity there. The downside would be that the 5-year survival rate for businesses is lower there. Being a more residential cluster there is also the opportunity for the take away business.

The second, riskier option would be cluster 1, residential. Mr. Roy would need to be careful to select an area of the borough with more night life. Perhaps near a pub. That way he could pick up passing trade, but also take advantage of the take away market. Sutton would be a suggested area, as it has a high business survival rate but many pubs.

5 Discussion

With clustering it is possible to interpret what the clusters represent in many ways. For instance, whether the area is residential is important in this case as it is required to support the take-away business. The issue is, the Foursquare API only gives the venue category, whether the area is residential has been assumed based whether there are grocery stores and supermarkets. Generally, grocery stores are near where people live.

The opening hours of venues was also a consideration. Generally coffee shops in the UK open during the day only, and pubs are open from late morning to late evening. Pubs don't always sell food, but many do. In areas where there are numerous coffee shops but no pubs or restaurants, it has been assumed there's not much happening in the evening.

For Mr. Roy, his business really relies on people wanting to be out and about in the evening. Indian restaurants in the UK, don't generally open at lunchtime, they rely on evening business but will stay open late, past pub closing time to pick up additional business. Also, many Indian restaurants will offer a take away service.

The two options suggested in the results section are based on these interpretations made about the areas.

6 Conclusion

The London boroughs have been split into 5 different clusters based on K-means. Each cluster has been interpreted and 2 options have been put forward as suitable boroughs.

Cluster 2 would be the safest option. Newham doesn't have many Indian restaurants but there are hotels and pubs in the area.

Cluster 1 would be a riskier option as there isn't much night life in these areas and the business may be more reliant on the take away business but in a borough such as Sutton, with lots of pubs it might do well.