

Segmenting London boroughs and recommending the right place to open a restaurant

Evgeniy Kravchenko

25/12/2020

1. Introduction

a. Background

London is a huge city with lots of cultures blended in it. This city seems to be full of opportunities. London is also a popular tourist destination. Many businessmen see great opportunities (as well as risks) in this city.

b. Problem

The prices and the competition here are high, so it's very important to find the right place to open a new venue. It is also important to answer the following questions:

- Is the competition in the area already too high?
- Who are the competitors?
- Is there enough relevant audience in the area?
- What are the most popular types of cuisine in specific region?

c. Interest

Businessmen that are seeking for an opportunity, will find the information provided by such a research very valuable. This can also be helpful for locals or tourists who want to get a bigger picture of food culture in London.

2. Data

a. Data sources

We will scrape the information from the following Wikipedia webpage:

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

This page contains the table with list of boroughs in London. The important information in the table is: Borough name, Area and Population. We will also use *foursquare api* to receive local restaurants and food places in a specific radius of each borough.

b. Data cleaning

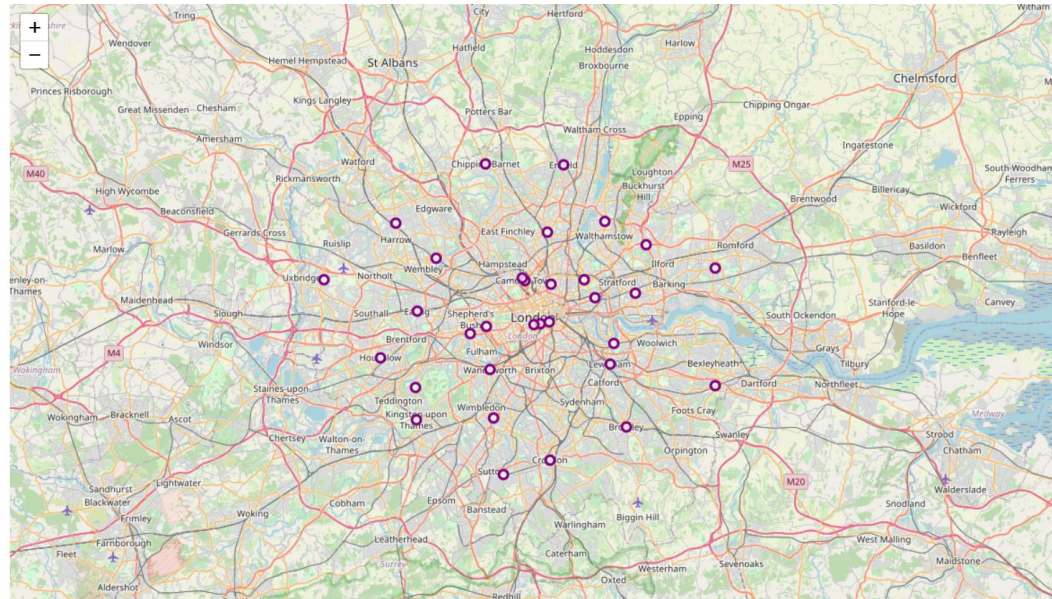
Although there was a column with coordinates of each borough, it had a link inside it and the format was hard to read. That's why I used *foursquare api* to receive the correct coordinates for each borough.

3. Methodology

a. Exploratory data analysis

Our final purpose is to understand where it would be a good idea to open a new restaurant, according to the information about already existing places.

First let's create a map with boroughs coordinates, using folium library:



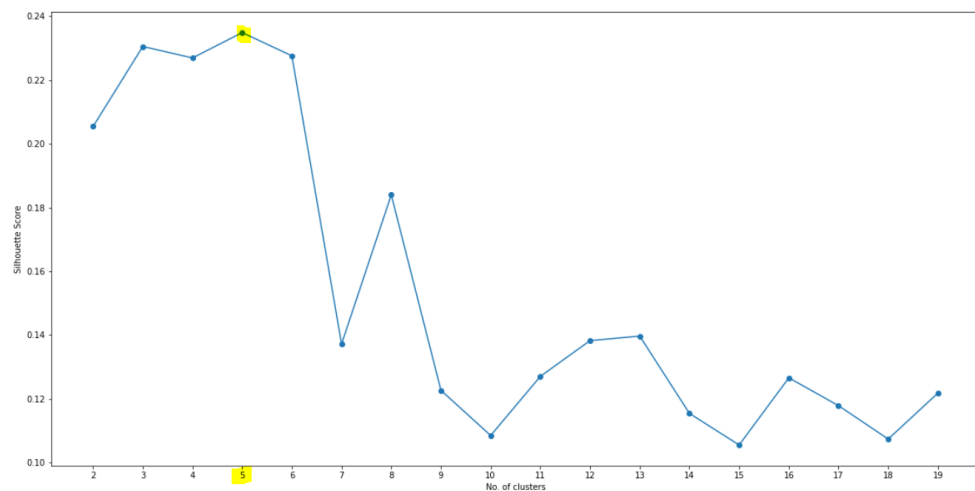
One of the difficulties as we can see in the map is overlapping boroughs in the city center. It can make our analysis less accurate.

b. Modeling using k-Means clustering algorithm

After running foursquare api and finding venues in the radius of 1,500 meter , we can see that there are 2,404 different venues in this radius (around each borough center coordinate) and 246 unique categories.

The second step is to filter restaurants, cuisine, cafe, snack and other food places. This brings us to 540 venues and 59 unique categories.

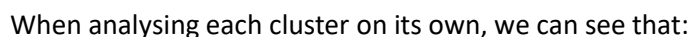
Using Silhouette Score the number of clusters giving highest score is k=5.



	Borough	Borough Latitude	Borough Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
1	Barking and Dagenham	51.554117	0.150504	Lara Grill	51.562445	0.147178	Turkish Restaurant	2	Restaurant	Turkish Restaurant	Chinese Restaurant	Vietnamese Restaurant	Indian Restaurant
9	Barking and Dagenham	51.554117	0.150504	The Pipe Major	51.545795	0.165834	Restaurant	2	Restaurant	Turkish Restaurant	Chinese Restaurant	Vietnamese Restaurant	Indian Restaurant
19	Barking and Dagenham	51.554117	0.150504	New China Gold	51.561269	0.141839	Chinese Restaurant	2	Restaurant	Turkish Restaurant	Chinese Restaurant	Vietnamese Restaurant	Indian Restaurant
32	Barnet	51.653090	-0.200226	Spizzico	51.656600	-0.201802	Italian Restaurant	1	Fast Food Restaurant	Sandwich Place	Seafood Restaurant	Indian Restaurant	Restaurant
37	Barnet	51.653090	-0.200226	Dudley's Pancake House	51.652965	-0.199625	Restaurant	1	Fast Food Restaurant	Sandwich Place	Seafood Restaurant	Indian Restaurant	Restaurant

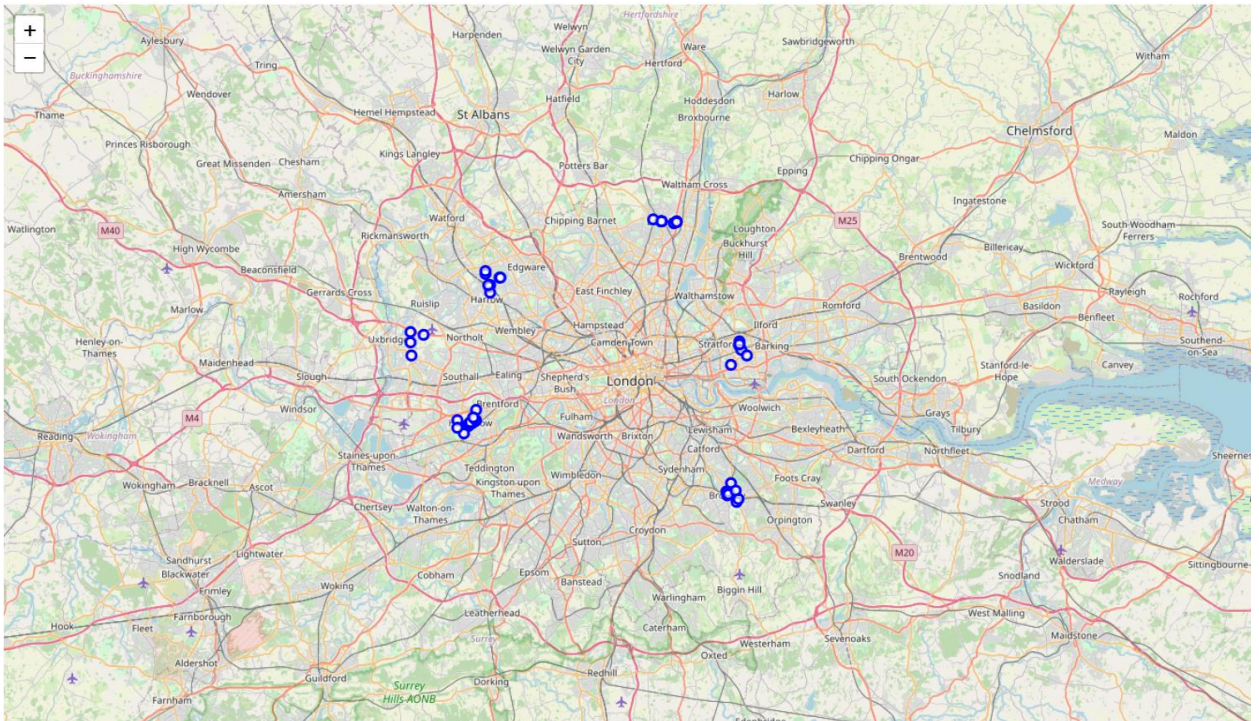
The results of clustering into k=5 different clusters:

The folium map shows the following results:



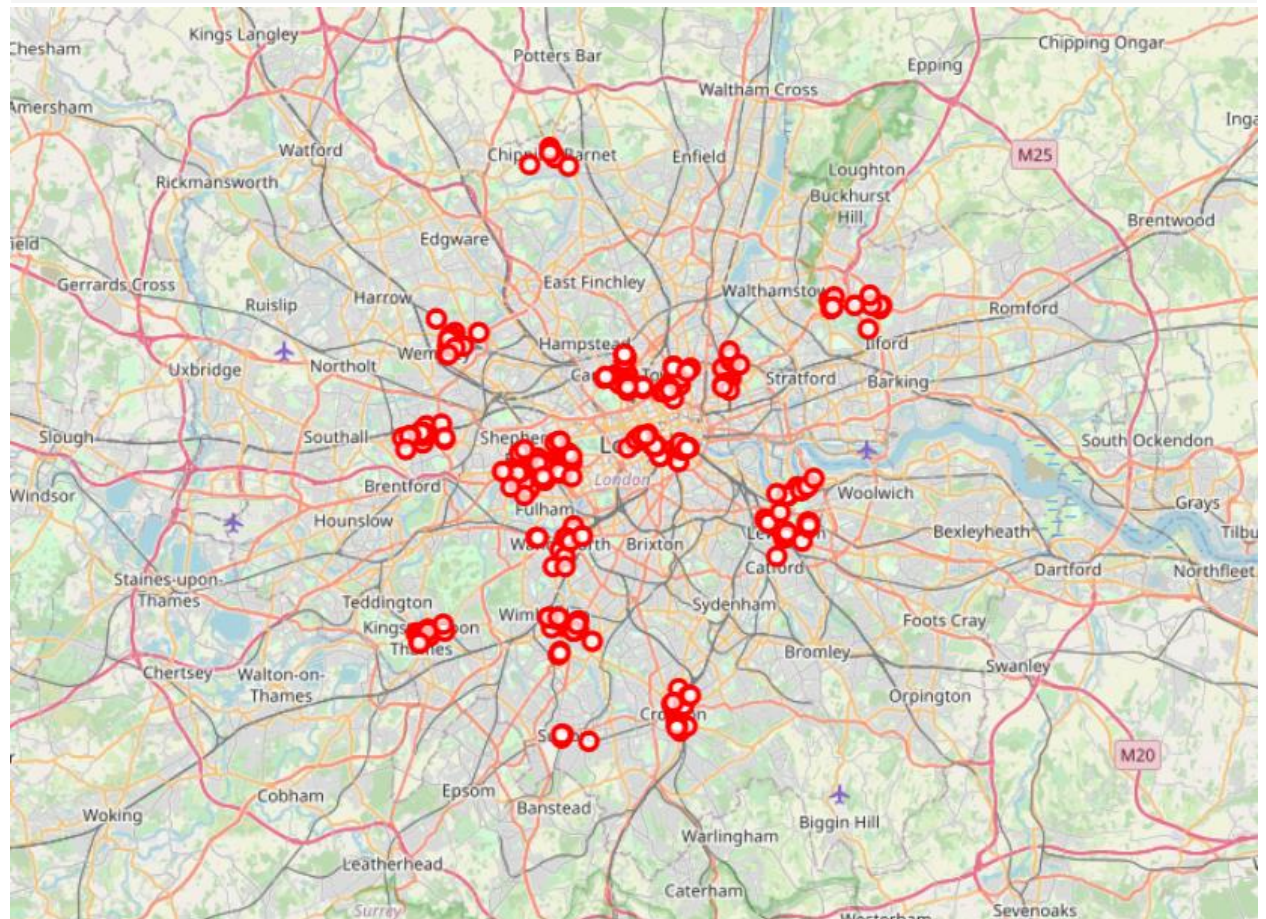
Cluster #1, Most common is Indian Restaurant:

	Borough	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
181	Bromley	Indian Restaurant	Asian Restaurant	Fast Food Restaurant	Portuguese Restaurant	Sushi Restaurant
536	Enfield	Fast Food Restaurant	Indian Restaurant	Sandwich Place	Turkish Restaurant	American Restaurant
999	Harrow	Indian Restaurant	Sandwich Place	Fast Food Restaurant	Thai Restaurant	Asian Restaurant
1125	Hillingdon	Chinese Restaurant	Fast Food Restaurant	Indian Restaurant	Vietnamese Restaurant	Italian Restaurant
1139	Hounslow	Indian Restaurant	Sandwich Place	Fast Food Restaurant	Restaurant	Portuguese Restaurant
1800	Newham	Indian Restaurant	Sandwich Place	Fast Food Restaurant	Asian Restaurant	Italian Restaurant



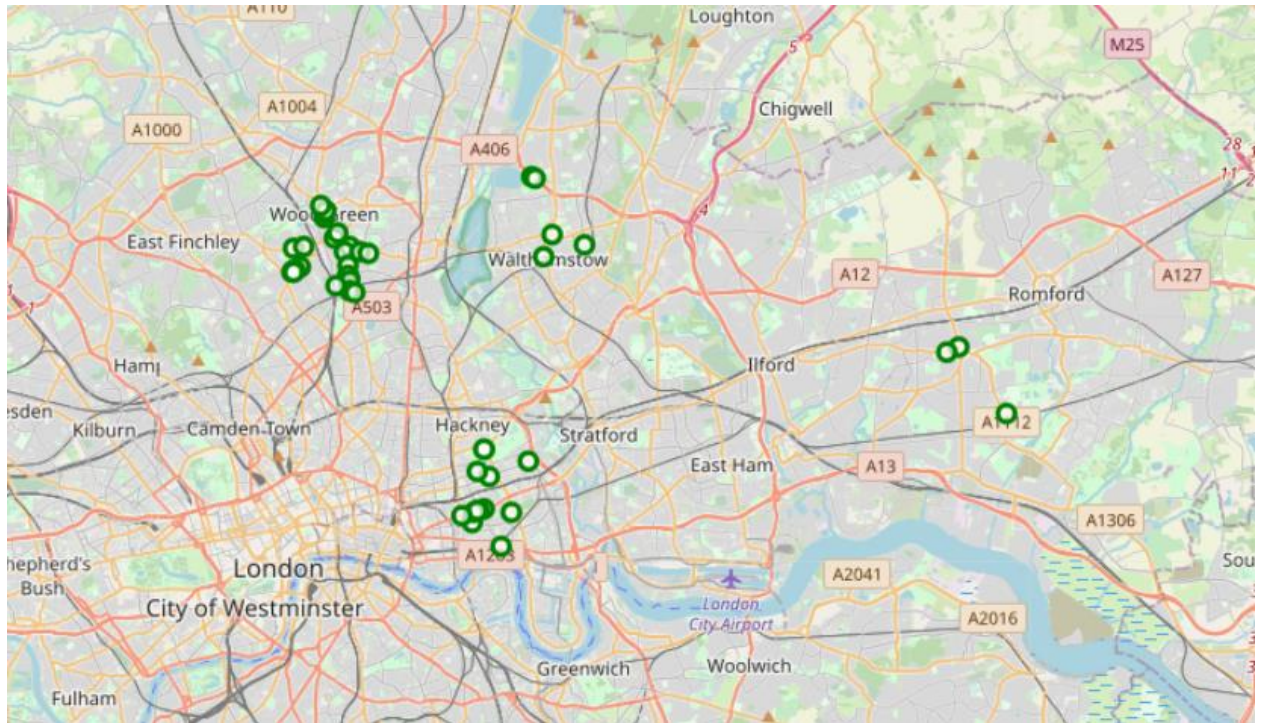
Cluster #2, Most common are Indian and Italian Restaurants:

	Borough	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
32	Barnet	Fast Food Restaurant	Sandwich Place	Seafood Restaurant	Indian Restaurant	Restaurant
84	Brent	Indian Restaurant	Sandwich Place	American Restaurant	Food Court	Fast Food Restaurant
235	Camden	Italian Restaurant	Middle Eastern Restaurant	French Restaurant	Falafel Restaurant	Greek Restaurant
334	Croydon	Indian Restaurant	Mediterranean Restaurant	Italian Restaurant	Portuguese Restaurant	Sushi Restaurant
433	Ealing	Italian Restaurant	Vietnamese Restaurant	Indian Restaurant	Sushi Restaurant	Persian Restaurant
603	Greenwich	Indian Restaurant	Turkish Restaurant	Mediterranean Restaurant	Italian Restaurant	Japanese Restaurant
706	Hackney	Vietnamese Restaurant	Seafood Restaurant	Italian Restaurant	Japanese Restaurant	Vegetarian / Vegan Restaurant
801	Hammersmith and Fulham	Thai Restaurant	Italian Restaurant	French Restaurant	Japanese Restaurant	Turkish Restaurant
1026	Havering	French Restaurant	Falafel Restaurant	Italian Restaurant	Greek Restaurant	Middle Eastern Restaurant
1198	Islington	Mediterranean Restaurant	Italian Restaurant	Japanese Restaurant	Vietnamese Restaurant	Seafood Restaurant
1300	Kensington and Chelsea	Restaurant	Italian Restaurant	Thai Restaurant	Indian Restaurant	Chinese Restaurant
1400	Kingston upon Thames	Thai Restaurant	Italian Restaurant	Sushi Restaurant	Japanese Restaurant	Sandwich Place
1503	Lambeth	Mexican Restaurant	Sushi Restaurant	Street Food Gathering	Restaurant	Food & Drink Shop
1597	Lewisham	Food Truck	Restaurant	Indian Restaurant	Turkish Restaurant	Fast Food Restaurant
1699	Merton	Italian Restaurant	Thai Restaurant	Sushi Restaurant	Lebanese Restaurant	Fast Food Restaurant
1838	Redbridge	Indian Restaurant	English Restaurant	Sushi Restaurant	Mediterranean Restaurant	Fast Food Restaurant
1928	Southwark	Italian Restaurant	Street Food Gathering	Seafood Restaurant	Restaurant	Asian Restaurant
2028	Sutton	Italian Restaurant	Portuguese Restaurant	French Restaurant	Sandwich Place	Asian Restaurant
2216	Wandsworth	Indian Restaurant	Thai Restaurant	Portuguese Restaurant	Restaurant	Lebanese Restaurant
2327	Westminster	Korean Restaurant	Seafood Restaurant	Japanese Restaurant	Greek Restaurant	French Restaurant



Cluster #3, Most common is Turkish Restaurant:

	Borough	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
1	Barking and Dagenham	Restaurant	Turkish Restaurant	Chinese Restaurant	Vietnamese Restaurant	Indian Restaurant
898	Haringey	Turkish Restaurant	Mediterranean Restaurant	Tapas Restaurant	Indian Restaurant	French Restaurant
2084	Tower Hamlets	Turkish Restaurant	Italian Restaurant	Chinese Restaurant	Food Truck	Middle Eastern Restaurant
2172	Waltham Forest	Turkish Restaurant	Vegetarian / Vegan Restaurant	Portuguese Restaurant	Vietnamese Restaurant	Fast Food Restaurant



Cluster #4 and 5, were not big enough.

The final table can be concluded as follows:

Cluster	1st most common	2nd most common
1	Indian	Sandwich Place
2	Italian	Indian
3	Turkish	Chinese

5. Conclusions and discussion

The analysis can be widened to other fields like: crime rate, kind of population in the specific borough etc. The number of clusters we received was eventually low. It would be a better analysis if we also had neighborhoods, and it would allow us to divide the information in more accurate way. Yet we can still see specific clustering according to the kind of restaurants.

Now that we have clusters for different boroughs in London, we can try and get decisions if we will open a new restaurant, in which borough, what type of food, etc. We can open a popular cuisine or we can do the other way around. It now depends on the business model we prefer.