

Battle of the Neighbourhoods: Exploring the best places in London to develop a restaurant supply business

Coursera Capstone Project Week 2

Joseph Ritchie

Introduction: The Business Problem

London is a vast metropolitan area, with a large restaurant trade. The city has a population of around 8.9 million.

My client wants to set up a business supplying restaurants with ingredients. In the first instance, they would like to develop their understanding of London's restaurants and how they are distributed in the various boroughs.

My task is to provide insights on the ways in which restaurant types are distributed in London boroughs.

Data

To provide our initial recommendations, we want to provide a map of London that shows the geographical distribution of restaurants.

Initially, we want to focus this analysis on a borough level.

To do this, we will need:

- London borough geo-spatial data.
- Data about the location of venues in London, which we can obtain from Foursquare API.

This data is obtained from Foursquare and wikipedia, and appropriately cleaned.

Methodology

The two data components require slightly different approaches. The borough data is obtained using beautiful soup to scrape data from wikipedia tables of the boroughs. After cleaning, this provides us with a dataframe including a borough name, longitudes and latitudes.

ut[404]:

	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

We then use Foursquare to obtain the co-ordinates of the various restaurants in London, first using a call to get all venues and then refining the resulting dataframe to simply use the columns that contain restaurant data.

Once we have obtained the data, the key method we are going to use is using a clustering algorithm to highlight which neighbourhoods share similar characteristics in terms of the restaurants therein. This means one-hot encoding the variables and then undertaking a clustering analysis which sorts the boroughs into clusters based on the restaurants they contain.

ut[501]:

	BoroughName	Population	Latitude	Longitude	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	
0	Barking and Dagenham	194352	51.5607	0.1557	1	Vietnamese Restaurant	Dim Sum Restaurant	Italian Restaurant	Indian Chinese Restaurant	Indian Restaurant	Fast Food Restaurant	Falafel Restaurant	Dumpling Restaurant	Chinese Restaurant	R
1	Barnet	369088	51.6252	-0.1517	1	Vietnamese Restaurant	Dim Sum Restaurant	Italian Restaurant	Indian Chinese Restaurant	Indian Restaurant	Fast Food Restaurant	Falafel Restaurant	Dumpling Restaurant	Chinese Restaurant	R
2	Bexley	236687	51.4549	0.1505	0	Italian Restaurant	Fast Food Restaurant	American Restaurant	Portuguese Restaurant	Chinese Restaurant	Dim Sum Restaurant	Indian Chinese Restaurant	Indian Restaurant	Falafel Restaurant	R
3	Brent	317264	51.5588	-0.2817	1	American Restaurant	Indian Restaurant	Asian Restaurant	Indian Restaurant	Brazilian Restaurant	Latin American Restaurant	Argentinian Restaurant	Caribbean Restaurant	Japanese Restaurant	R

Results

Based on the clustering algorithm, for which seven clusters were created, we can identify the following primary clusters of restaurants:

Cluster 1 appears to primarily contain Italian restaurants, with a high number of fast food restaurants and Portuguese restaurants.

Cluster 2 has many Indian and Asian restaurants

Cluster 3 contains has many Oriental and Chinese restaurants.

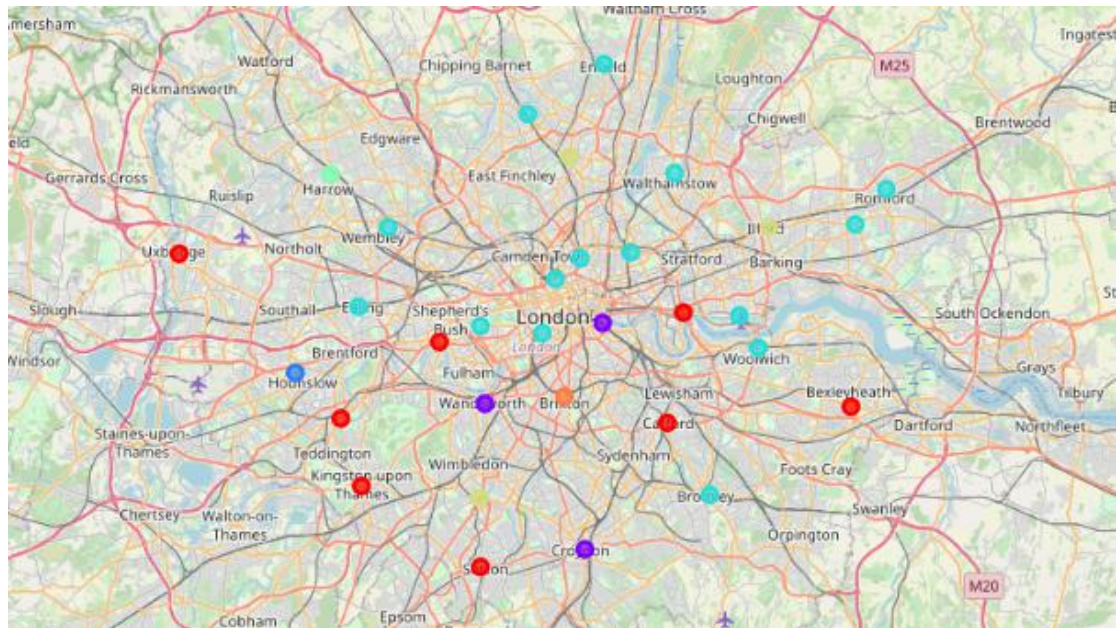
Cluster 4 contains many Vietnamese restaurants.

Cluster 5 has a particularly diverse and unique combination of restaurants.

Cluster 6 contains many fast food restaurants.

Cluster 7 also offers a specific combination of restaurants, like cluster 5.

We can visualise these clusters using a map of London boroughs, using folium.



Conclusion

To be effective, a food supply business has to develop effective relationships and effective access to customers. Boroughs offer an intuitive scale for thinking about a new business, given the challenges of transporting goods and maintaining market relations.

With this cluster analysis, we can identify which boroughs offer which opportunities to suppliers, at least insofar as suppliers may wish to attempt to supply similar restaurant businesses, who may have similar needs in terms of ingredient types.

Discussion

This cluster analysis provides an excellent starting point for developing the context of a restaurant supply business in London. The next steps are to cross reference this data with more data about business rental rates and a greater level of neighbourhood granularity.