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

Coursera Capstone Project Week 2

Introduction

London is a major tourist destination, with a great deal of cultural diversity and international profile. One result of this is that the city has a thriving restaurant trade.

In order to begin planning for a restaurant supply business to meet the high levels of demand, our clients want to develop some analysis of London's borough's and the restaurants they support.

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.

Methodology

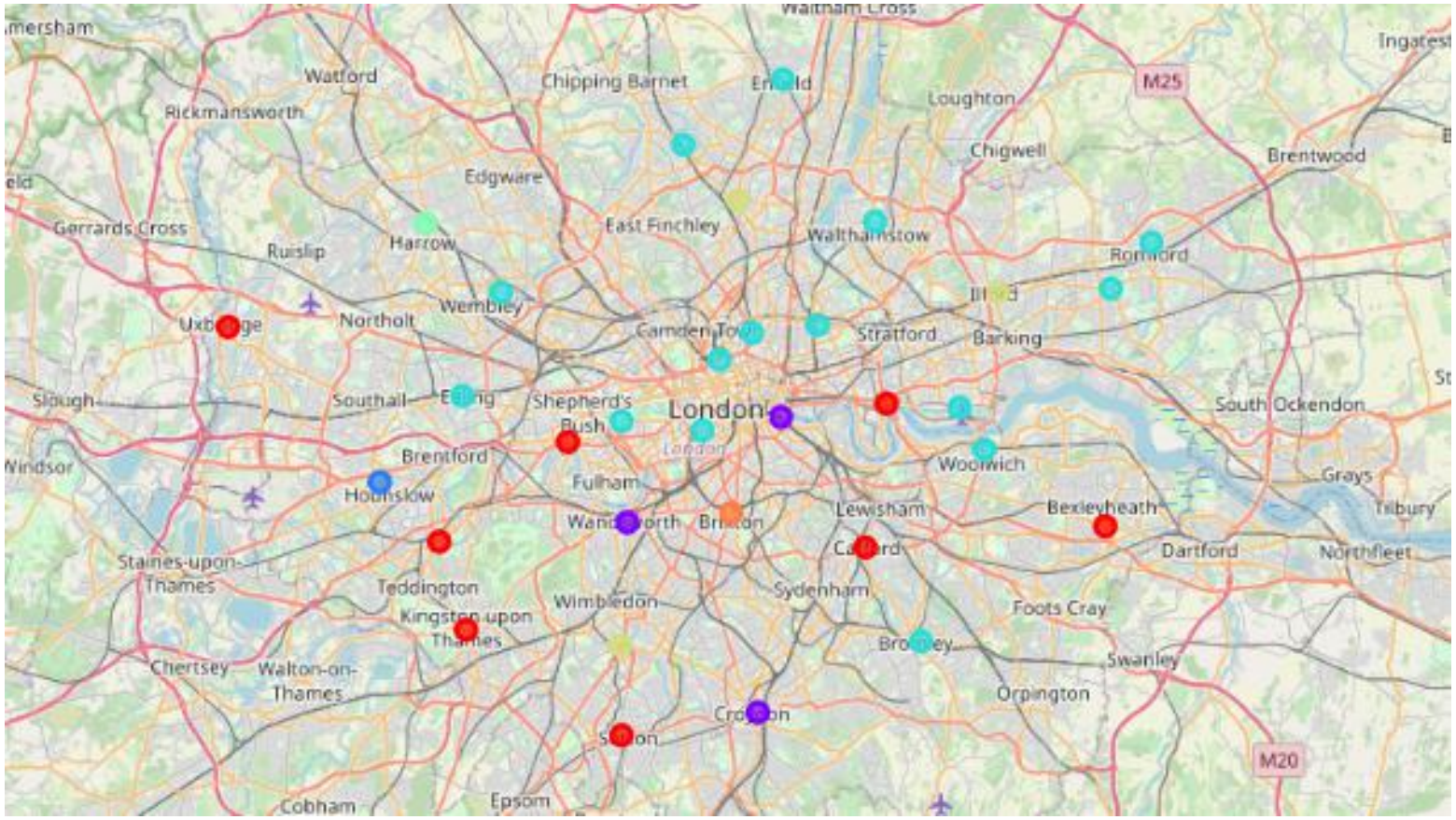
- The two data components require slightly different approaches. The borough data is obtained using beautiful soup to scrape data from wikipedia tables of the boroughs. After cleaning, this provides us with a dataframe including a borough name, longitudes and latitudes.
- We then use Foursquare to obtain 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 algorhythm to highlight which neighbourhoods share similar characteristics in terms of the restaurants therein. This means one-hot encoding the variables and then undertaing a clustering analysis which sorts the boroughs into clusters based on the restaurants they contain.

Results

We set k for 7, and have seven resulting clusters.

- 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.

Visualising the Cluster



Discussion

- We have now identified the clusters of restaurants in London.
- This is a good first step to identifying the best location for a food supply business.
- Next steps will include exploring the comparative advantage of individual clusters, i.e. restaurant density, and getting more granular neighbourhood data. We can also obtain information on the relative popularity of the restaurants.