**Introduction**

**Background**

London is a large city, a real metropolis where you can find many things and do many activities. There is a large amount of venues i.e. restaurants, pubs, sports centres etc. Many people would like to open their business in London. The right place to open a business is not an easy task. For this reason a thorough analysis is needed in order to take a well-informed decision.

London includes many boroughs where some of them belong to inner London and some of them in the outer area. Different boroughs have different preferences with respect to the different venues that exist in their area. Furthermore the rent price is an important factor in choosing the right place to open a business.

**Problem**

Deciding the right place of opening a business is not an easy task. Many parameters or variables could be taking into account. The preferences of the citizens of a borough is quite important while the price of rent plays an important role.

**Interest**

The interest could appeal to somebody that wants to open a business or somebody who wants to visit these places. The parameters that could be taken into account is the price, the popularity of a venue etc.

## **Data Section**

### Data Sources

**The used data sources are the following:**

**1. List of London Boroughs with their coordinates based on the following Wikipedia url:** <https://en.wikipedia.org/wiki/List_of_London_boroughs>

**2. Data from the Foursquare API which will enrich the above data with the respective venues and their geographical coordinates.**

**3. Online based data on the average rent price in London boroughs**

**https://www.theweek.co.uk/99093/london-house-prices-which-boroughs-are-on-the-rise**

## **Data Cleaning**

The main data source for our problem is the above Wikipedia url. Wikipedia provides information about London boroughs and local authorities. This includes political control information and headquarter information that will not be used in our analysis. Thus we will filter out these data and keep the rest of them which include Name, Area size, Population, coordinates and rent prices for each borough. Furthermore in our study we will exclude the London area and thus we will focus on boroughs in the outer area of London.

Foursquare provides a dataset of venues around the specific coordinates or venues, if we  
use the “Explore” function in the Developer tab. Once requested, we get a full breakdown of  
all recorder venues around the boroughs of interest.

**Feature selection**For convenience, we transform the dataset only to show top 10 places to work with. After that, we merge the dataframes together for a comprehensive set of values, worth analysing.

**Data Analysis**

**Clustering**

Once the boroughs are selected, we will place them onto the map and cluster the boroughs  
to analyse the similar ones against each other and within the clusters themselves. Each of the clusters will be compared against the popularity of the venues, worth considering for a new business venture, as well as the hospitability of the venture climate. Depending on the area, population and rent, each of the clusters offer different advantages and disadvantages in terms of venue choices.

**Feature selection**

For convenience, we transform the dataset only to show top 10 places to work with. After that, we merge the dataframes together for a comprehensive set of values, worth analysing.

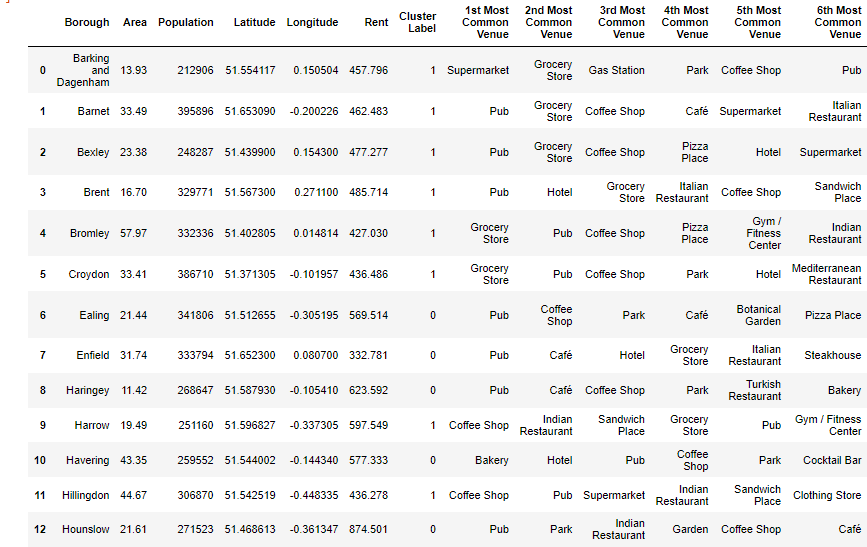


Figure Boroughs & Venues

**Data Analysis**

**Clustering**

Once the boroughs are selected, we will place them onto the map and cluster the boroughs to analyse the similar ones against each other and within the clusters themselves. Each of the clusters will be compared against the popularity of the venues, worth considering for a new business venture, as well as the hospitability of the venture climate. Depending on the area, population and rent, each of the clusters offer different advantages and disadvantages in terms of venue choices.

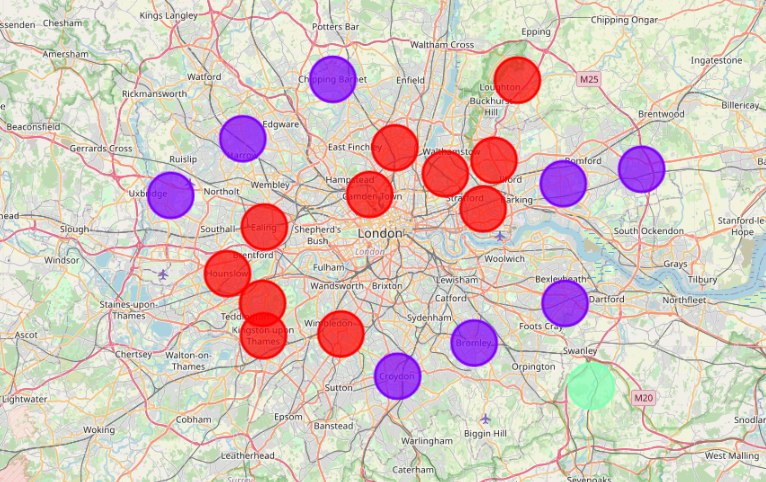


Figure Map of the Clusters

**Results**

Reviewing each cluster as per method above, we see that:

1. In the first cluster, the most popular place is a pub. This is a common habit in Britain which is also understood by the data analysis. The second most popular mainly is a park or a café. Then, the coffee shops or the bars follow. The ethnic restaurants like Indian or Italian restaurants follow in the popularity rating. Also houses tend to have higher rents in this cluster since it covers places closer to London city centre.

2. In the second cluster, grocery stores receive a higher rating. This could be explained since this is cluster, farther from London where people do not consider pubs as their first priority. Rents are also lower in this area.

3. The third cluster concerns just one borough which is located quite far from London. Still in this case, the pubs are the most popular venues but other places like golf or sports activities follow in the ratings. Shops or supermarkets are not in the first ten of the choices.

• **Conclusion**

This study helps the candidate owners as well as the already owners of a business in the outer area of London to understand where it is the best place to open or change the location of their business. Different types of venues have been analysed based on their location, their popularity and the price of rent in the area they belong to. Persons interested in opening a new business or change the location of their current business can take advantage of this analysis.