

FINDING THE BEST LOCATION FOR A START-UP COFFEE BUSINESS



BACKGROUND - LONDON

1. Fastgrowing city, it has plenty of talent and it is a prime location for securing funding and investments.
2. Densely populated.
3. High competition among restaurants.
4. Location such as Piccadilly, Mayfair and Soho in Central London, are popular for start ups although most don't have the funds available to locate there.

BACKGROUND - THE CLIENT

1. The client is running a variety of bars in other parts of the world and would like to open their first coffee bar in London.
2. Opportunity to expand into the English market.
3. Significant funds to back up their start up so money will be less of a limitation in this case.

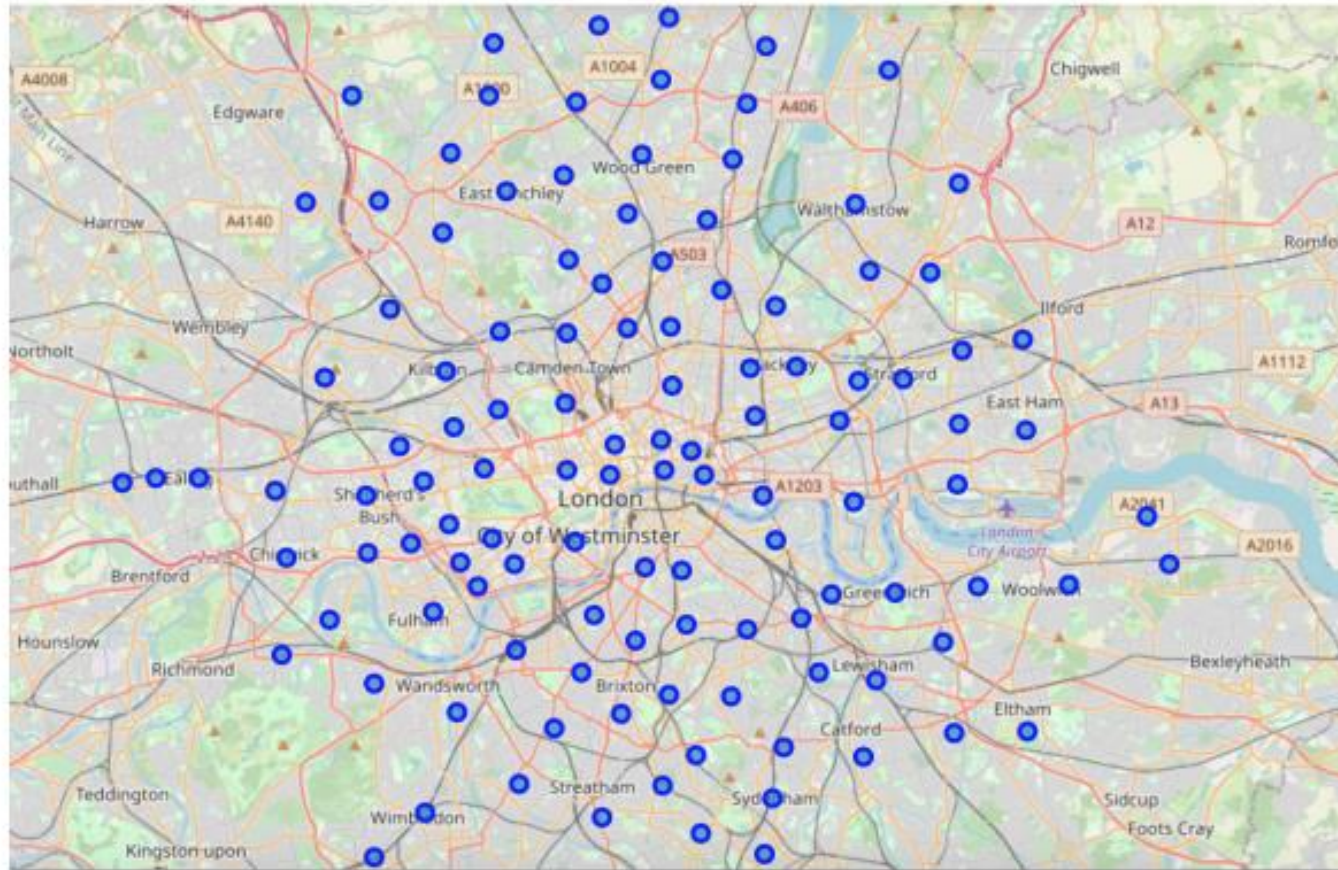
DATA ACQUISITION AND CLEANING

1. Data acquired through various online sources.
2. The two data set containing postal code information was transformed and merged into a panda's data frame.
3. Unneeded columns resulting in a data frame with only two columns 'Postcode' and 'District'.

METHODOLOGY — EXPLORATORY ANALYSIS

1. Using exploratory analysis, we first take a look at some initial parameters of the dataset.
2. We want to be able to dive deeper and look at the distribution of the dataset across London and thus, using the merged data we input the latitude, longitude coordinates along with the names of each district and plotted using Folium on a map of London.

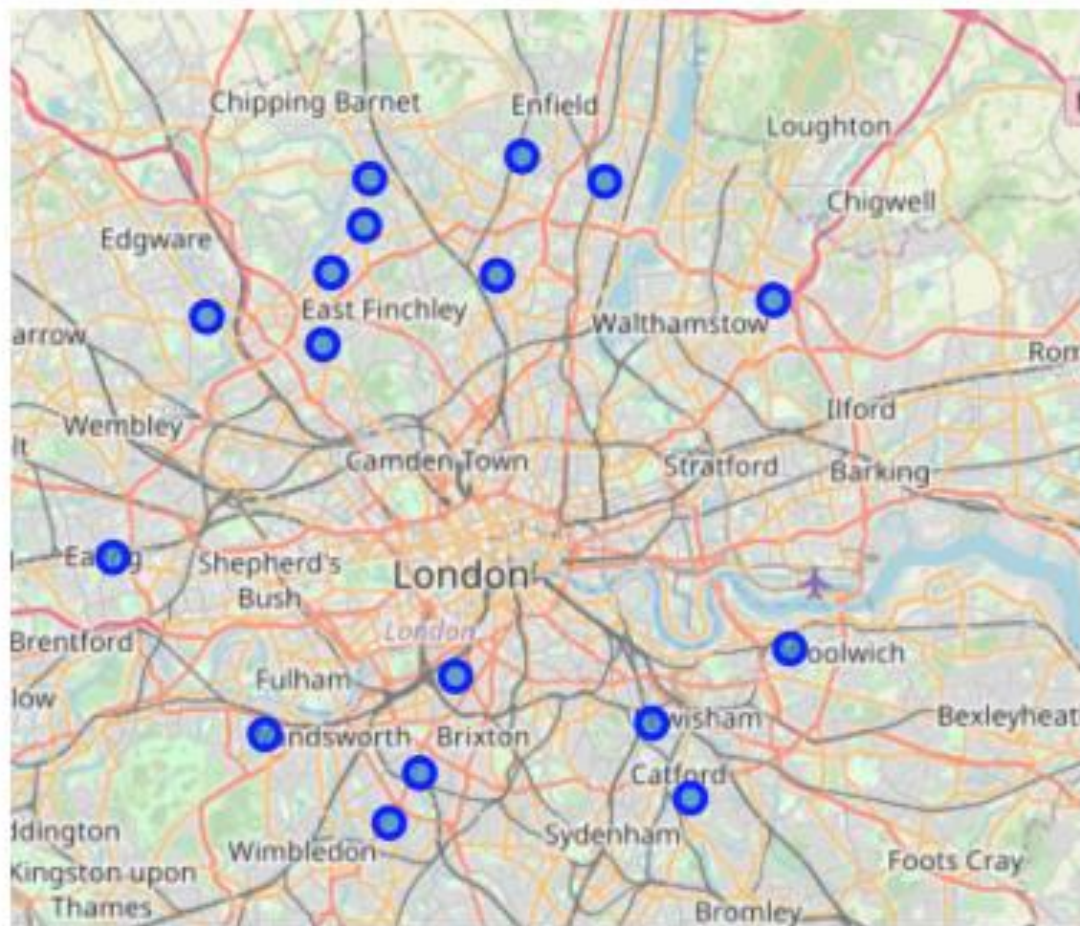
DISTRICTS OF LONDON PLOTTED USING FOLIUM



K-MEANS CLUSTERING — EXPLORE

1. The objective here using k-means clustering is to find similarities between districts based on the most popular venues around its area.
2. The number of bars in the area is assumed to be directly correlated with the popularity of that type of venue.
3. However, this information is not enough and we want to be able analyse further and look into cluster one for more information. We can also visualise the spread of the cluster using Folium.

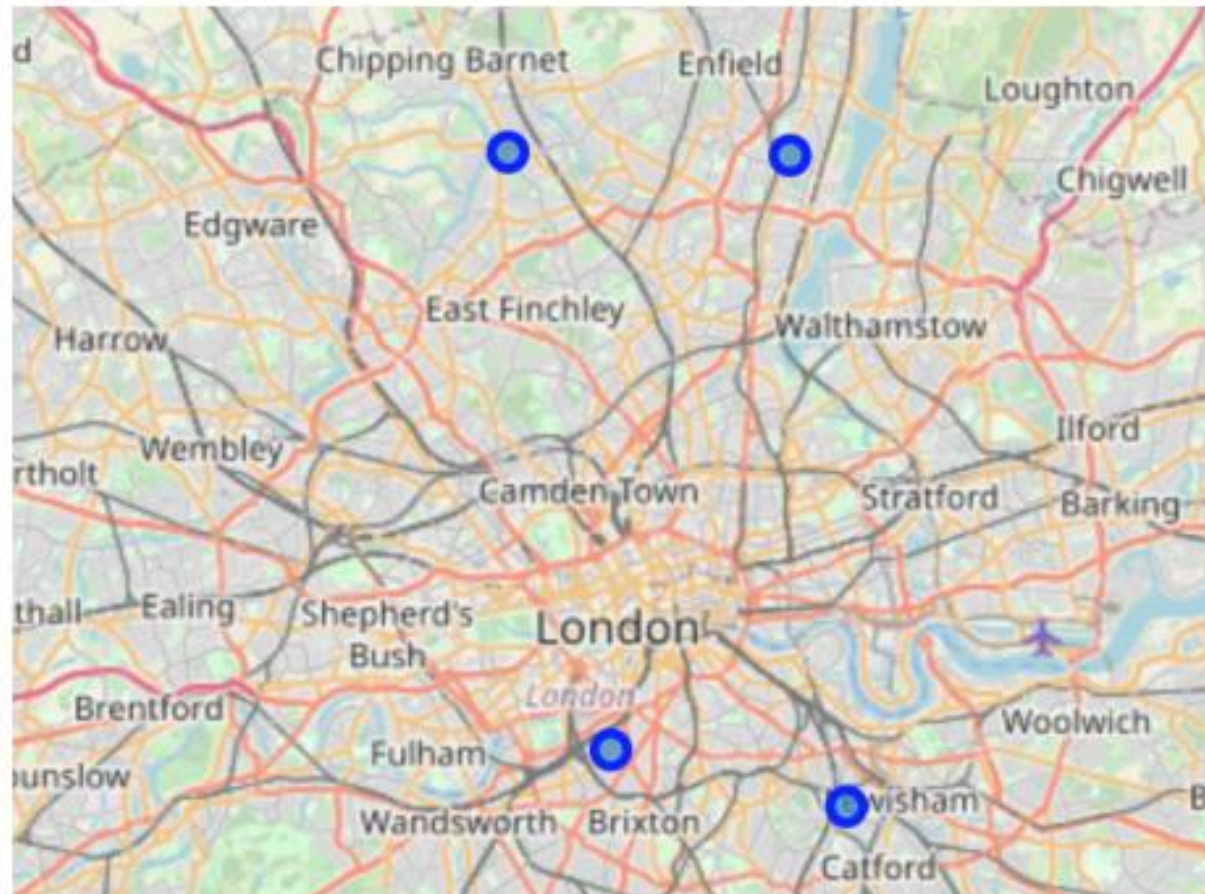
FOURSQUARE - EXPLORE



FOURSQUARE SEARCH

1. The goal here is to find, quantitatively, the number of coffee shops present within each district of Cluster 0.
2. To do this, we use the search end point along with a query. We then merge and take the segment of the data frame containing only rows from cluster 0 with the district's corresponding latitude and longitude coordinates.

FOURSQUARE SEARCH



RESULTS

1. As shown, most points are focused on the outskirts of London, even moving into other towns.
2. However, many factors other than location goes into play when opening a business such as budget, target customers, marketing, competition and business plan. The client can now define these parameters themselves and with the help of this assignment find the best location.
3. Points closer to the city or central London may be a popular attraction for bars and thus, in theory be a good location, however as this study suggests, the huge competition within those areas due to a larger number of already established bars may cause high barrier to entry.

DISCUSSION

1. Through exploring the dataset, we've established a list of locations that are good locations for a start-up coffee bar business. The client can now review these locations and based on their criteria, preferences and defined parameters, conclude on which location they think will best fit all their requirements.
2. However, there are still many limitations to this study such as we have not considered the actual demand for bars in each of the districts. This parameter is difficult to measure and thus, not considered for our study.
3. The other limitation is that, many areas within Central London may not have bars as the most common venue as a result of restaurant and cafes being more prominent. However, this does not mean this isn't significant demand for bars in these areas especially in the centre of the city. This shows why demand is an important parameter to consider.

CONCLUSION

1. In this report, we looked at different datasets, conducted exploratory analysis and using kmeans clustering, we were able to identify good locations for a start-up coffee bar business.
2. We interpret the results and their meaning, and applied real world logic to the problem in order to form an argument backed by evidence produced in this study.
3. In conclusion, the results shown in this study gives an indication, however are subjected to a few limitations which can be overcome using further research.