Capstone Project: Business expansion

Opening a new branch in South London

# Introduction

There is a successful business operating in Bristol, who want to open another branch in South London. The business currently has a branch in the Bedminster neighbourhood of Bristol, and wants to find a similar neighbourhood in London to open a store in. The business has already decided that south London (SW and SE postal codes) are the optimum overall area to base themselves but need to refine the exact neighbourhood to expand into.

The business believes that the type of people who frequent other highly rated venues in the Bedminster area are the key customers who make their business successful, therefore they are interested in finding a neighbourhood in South London which is similar to Bedminster in terms of the types of venues which are frequently found in that area.

# Data Requirements

The data which will be required to solve this problem includes:

* Data about “nearby venues” in Bedminster
* A list of SW and SE postcode areas and associated neighbourhood names
* Latitude and longitude data for the SE and SW postcodes
* Data about “nearby venues” around the centre of each of the SE and SW postcodes

The “nearby venues” data will be accessed through Foursquare, within the notebook. The location data about South London Postcodes has been accessed through the government website for London (see references, section 7). The list of postcode areas and associated neighbourhood names was found through doogal.co.uk (see references, section 7), which is a website which contains various data about UK maps and post codes.

# Methodology

The .csv files containing the London postcodes were imported from a github repository then inspected and cleaned up. This included renaming, sorting and dropping any data which was not useful to this project. The detailed postcode data was reduced down using the “groupby” function, to get an average latitude and longitude for each post code area, by taking the average latitude and longitude of all the individual postcodes in that area.

The two dataframes were then joined together to create a single dataframe with the latitude and longitude values and neighbourhood names for all the postcode areas in South London. Using this dataframe, a map was created of the South London postcode areas to display where each neighbourhood was located, see Figure 3.1 Map of the South London neighbourhoods.



Figure . Map of the South London neighbourhoods

The South London postcode areas were then explored using Foursquare, and a dataframe was created showing up to 100 venues within 1km of the centre of each postcode areas. Onehot encoding was then used to normalise the data in the dataframe and allow comparison between neighbourhoods. A similar exercise was then conducted for Bedminster, exploring nearby venues, and normalising the data to enable comparison.

The normalised data for South London and Bedminster was then combined into a single dataframe, and this was used for K-Means clustering, to group each neigbourhood into a cluster of neighbourhoods with similar attributes. The cluster containing Bedminster was then extracted and printed on a map, see Figure 3.2.

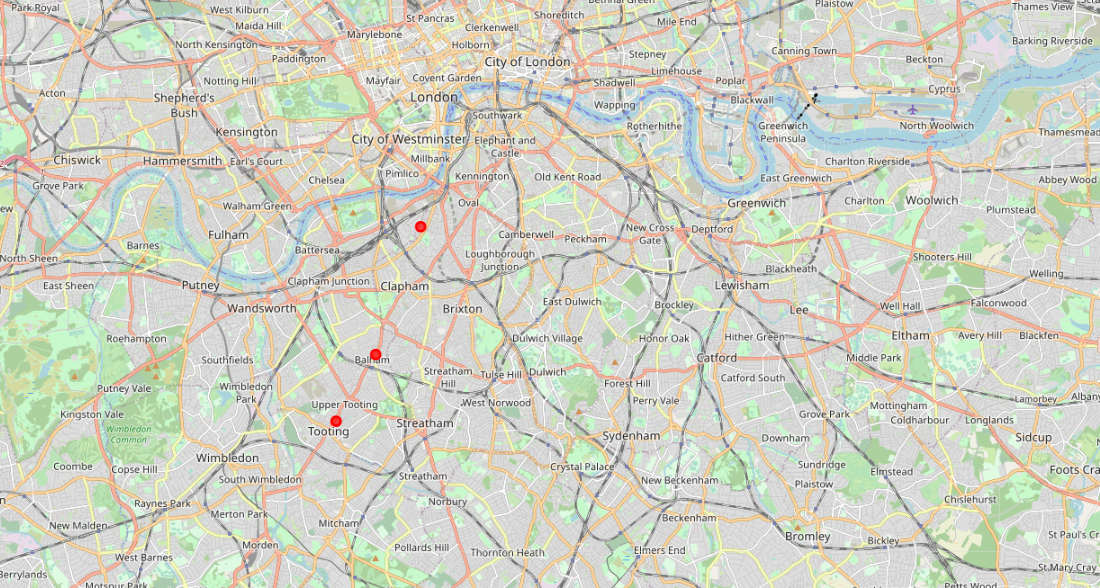


Figure . Map of the South London neighbourhoods which were clustered with Bedminster

# Results

Three neighbourhoods were identified by this exercise as being similar to Bedminster, as follows:

* Balham
* Tooting
* Nine Elms, South Lambeth

These three neighbourhoods, along with Bedminster had pubs and coffee shops in their top three most common venues, along with bars, cafes, grocery stores/super markets, and some international cuisine restaurants being within the top ten across all four neighbourhoods.

# Discussion

There have been some challenges in conducting this project, in that there were a very low number of venues which were returned by the Foursquare search for the Bedminster area, specifically only 46 venues within the 1km radius specified in the query, compared with an average of 66 venues across each of the London postcodes. There are significantly more venues than this in the area than suggested by Foursquare, which points to the fact that the Foursquare database for this area is not well populated or updated. Foursquare is not such a commonly used application in the UK, and outside of major cities, it does not appear to have good coverage. The majority of venue types appeared only once within the Bedminster area – there were only five venues which appeared more than once, which means that a top ten or top 15 list of venues contains these five, and subsequently lists other venue types in alphabetical order.

Despite these limitations in the data, it does appear that the model has worked well in grouping Bedminster with Balham, Tooting and South Lambeth, as reviewing the most common venue types in these areas does fit with not only the Foursquare-reported venues in Bedminster, but also venues types and frequency as known by the author.

The recommendation based on the results is that the most similar areas to Bedminster within South London, and therefore the best areas to explore for opening a new branch of the Bedminster business, are Balham, Tooting and Nine Elms, South Lambeth, on account of their propensity to be populated with bars, pubs, cafes and international restaurants.

# Conclusion

In conclusion, despite the relative lack of Foursquare data available for Bedminster, it has been possible to successfully compare neighbourhoods with South London to Bedminster and present a recommendation to the business owner that the best neighbourhoods to open another branch of the business are Balham, Tooting or Nine Elms, South Lambeth. These neighbourhoods are all in a relatively common geographic location and may now be further considered by the business owner to ascertain any other limitations around these locations.

# References

List of postcode areas

<https://www.doogal.co.uk/london_postcodes.php>

Latitude and longitude data for London postcodes

<https://data.london.gov.uk/dataset/postcode-directory-for-london>