Comparing the demographics and facilities of London neighbourhoods

Coursera Capstone Project

Introduction

London is one of the busiest cities in the world with over 8 million people calling it home. But do the demographics of the people who live in London, where they live and the facilities available to them, have distinct patterns?

This will be investigated in two parts:

1. Exploring the London boroughs by their demographic data
2. Using the Foursquare API to investigate the types of venues found in each area

Aim of the project

The types of venues in an area influence and are influenced by the people who live there.

Analysing this could be used to:

- Identify areas which could benefit from certain facilities based on their demographics
- Lead to suggestions for government funding in certain regions
- Give an indication to people who want to open a business in London of where their audience lives and the other facilities in the area

Results: demographic data

We cluster the London boroughs into 5 groups taking into account the following characteristics:
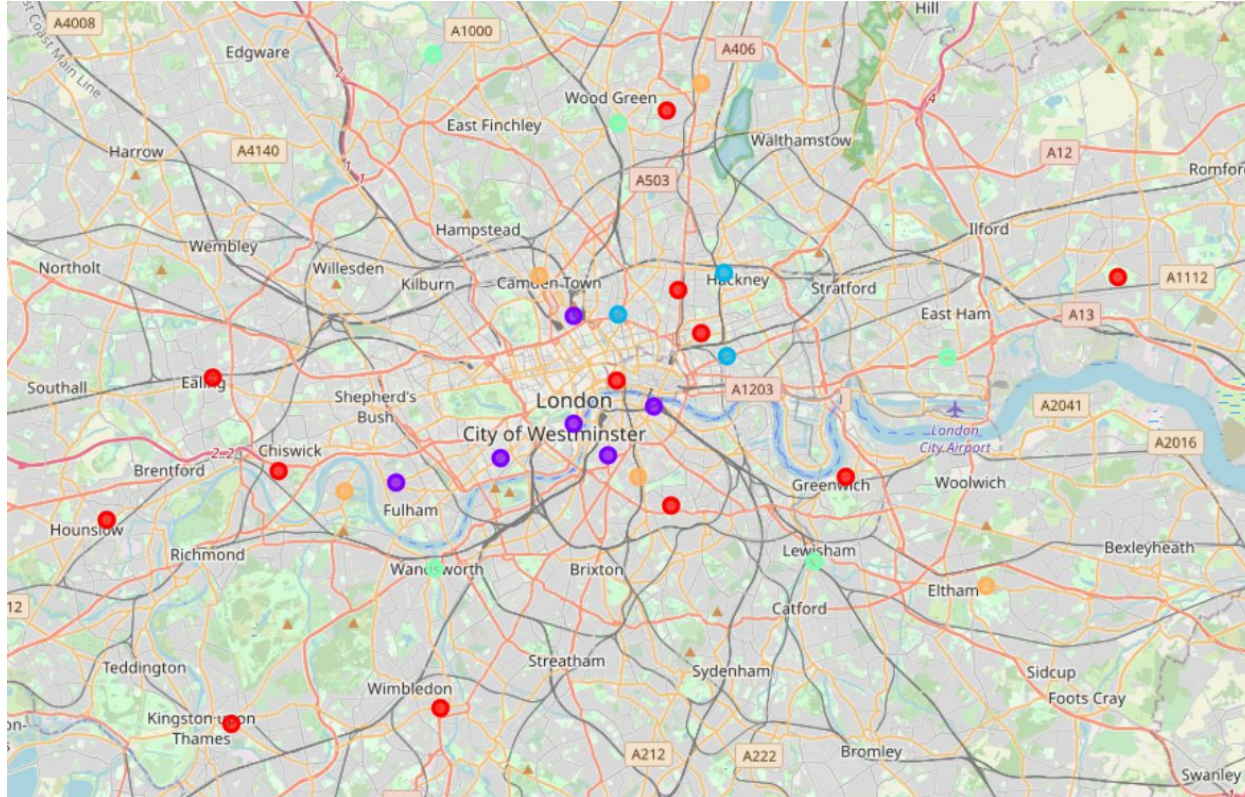
- Population density (per hectare) 2017
- Average Age, 2017
- % of resident population born abroad (2015)
- Happiness score 2011-14 (out of 10)

Summary of each demographic cluster

Cluster	Population density (per hectare) 2017	Average Age, 2017	% of resident population born abroad (2015)	Happiness score 2011-14 (out of 10)
0	51.1	36.4	37.2	7.3
1	116.6	36.3	42.8	7.2
2	151.1	33.1	37.0	7.1
3	86.1	34.7	42.1	7.2
4	29.0	38.9	20.3	7.3

Cluster 2 has the highest population density and lowest average age and happiness score. Cluster 4 has the lowest population density as well as the highest average age and happiness score.
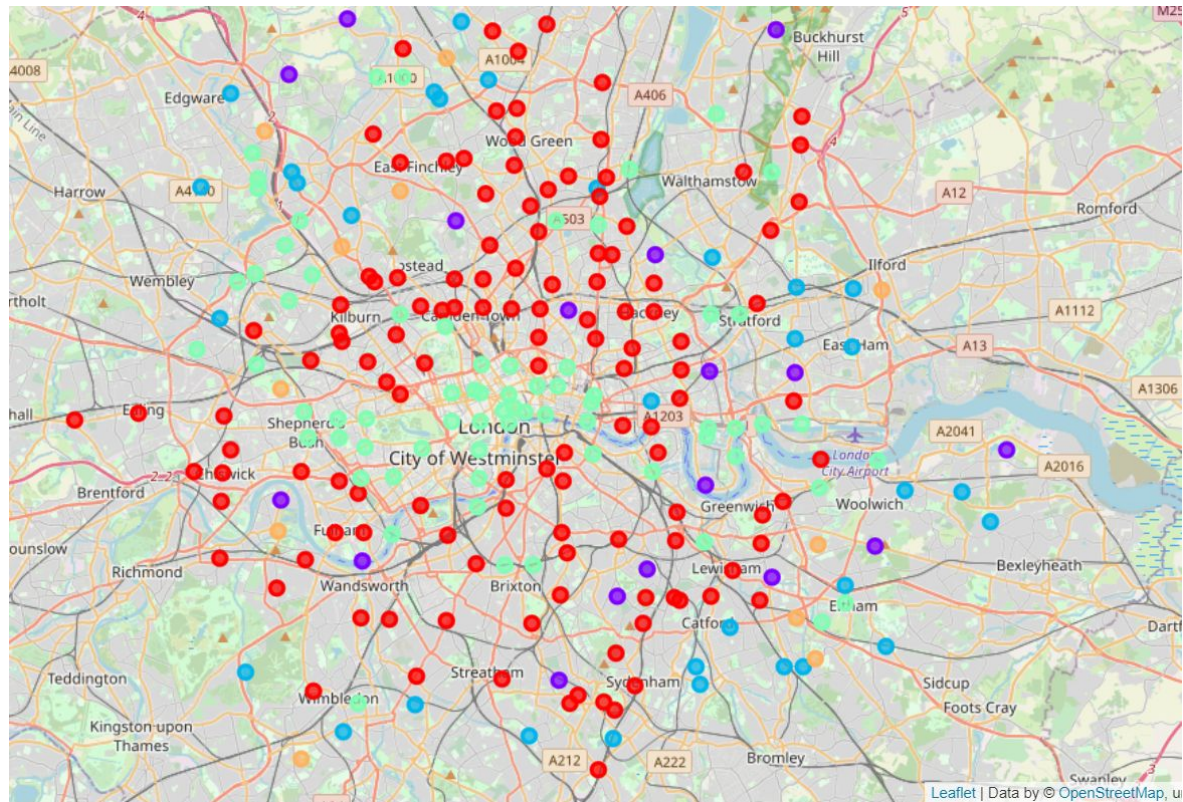
Map of the demographic clusters

Legend
Cluster 0
Cluster 1
Cluster 2
Cluster 3
Cluster 4

Cluster 2 boroughs are located in inner London, in the East. Cluster 4 boroughs are in the suburbs of London.

Results: venues data

We now look at the types of venues in different areas of London. For each area we find the proportion of each type of venue and then group the areas into 5 clusters. We then plot the clusters on a map of London.

Map of the venues clusters

Legend
Cluster 0
Cluster 1
Cluster 2
Cluster 3
Cluster 4

Summary of venues clusters

From looking at the location of each area, and the most common venues of each cluster, we make the following generalisations about each cluster:

- **Cluster 0** is the most common cluster with areas being found in a circle around central London. They stretch from inner London out to the suburbs. The most common venues are pubs, cafes and parks.
- **Cluster 1** areas are found amongst cluster 0 areas but are further out from central London. Pubs and cafes are common as well as outdoor venues such as zoos and golf courses.
- **Cluster 2** is found mostly on the edges of Greater London, representing suburban areas. The most common venue is a grocery store, as well as various nationality restaurants.
- **Cluster 3** is found in central London, along the river and in the North West of London. The areas in this cluster have a range of types of most common venues, from hotels to shopping to restaurants.
- **Cluster 4** is found also in the suburbs of London. The most common venues are parks and gyms.

Conclusions

We can see patterns in the demographic data of the people in London as well as in the types of venues present in each area. To summarise three geographic areas:

- **Central London**
  - Population density is high, higher than average happiness score and average age.
  - There are a lot of different types of venues including shops, restaurants, zoos, hotels and ice cream shops.
  - It may be the case that the venues are catered towards tourists and the areas are missing basic amenities such as supermarkets.
- **Inner London**
  - Some areas have the highest population density and lowest happiness scores, but there is a lot of variation.
  - The most popular venues are pubs, cafes and parks which are expected for residential areas.
- **Outer London**
  - These areas have the highest average age, lowest population density and highest happiness score.
  - The types of venues are as you would expect for residential areas, grocery stores, restaurantes, as well as a lot of outdoor facilities like parks.

Future directions of research

1. Analysis of the frequency of different venues in each area would lead to more insight on how well facilitated each area is in London. This could then be compared to the demographics of the population. Eg. group the supermarkets, sports facilities or restaurants and compare the density of them in each area. This may lead to suggestions for where areas are lacking facilities.
2. Broader investigation into the demographics would also be insightful as this report only looks at four characteristics. This would add more detail to describe each area of London, and would surely lead to patterns of similarities between different areas in London.