Analysis of type of venues and average income in London
Capstone Project

1. Introduction

Income inequality has increased in the last decades, becoming a major economic a social issue in both developed and developing countries. According to OECD data, the average income of the richest 10% of the population is about nine times that of the poorest 10% across OECD countries, up from seven times 25 years ago.

Income inequality is a significant problem in the United Kingdom (UK). The UK is the fifth most unequal developed country and the fourth most unequal in Europe. In 2018, households in the bottom 20% of the population had an average income of £12,798, whilst the average income of the top 20% was more than five times higher (£69,126) (The Equality Trust, 2019).

Within the UK, London is one of the richest regions; however deep inequalities are observed across its inhabitants. The differences in income are much greater in London than in the rest of the UK. The richest ten per cent has around 9.7 times the income of the lowest income households (Mayor of London, 2020).

Against this background this project aims to explore whether the type of venues found in London's boroughs is associated to the average income of households located there. Policy makers and NGOs are the main potential audience of this report. However, the geospatial analysis of household incomes and venues may also be of interest to business for predicting demand and identifying main competitors.

2. Methodology

2.1 Data

The data used in this project comes from two main sources: Foursquare and the UK Office for National Statistics.

Venue data of London boroughs is obtained through Places API – Foursquare. The data extracted include the location (latitude and longitude) and the category of venue. The top 100 venues were obtained, within a 500m radius of the centroid of each Borough.

Annual income data is obtained from the 'income estimates for small areas, England and Wales' published by ONS. Income is measured in net terms after deducting housing costs. Data was obtained for middle layer super output areas (MSOA) and then grouped for London boroughs (local authorities).

2.2 Analysis

Income data was grouped into the 33 London boroughs using mean values of annual income and explored using descriptive statistics. Geocode library was used to obtain the coordinates of the boroughs an locate them on a map (Figure 1).

Figure 1. London boroughs
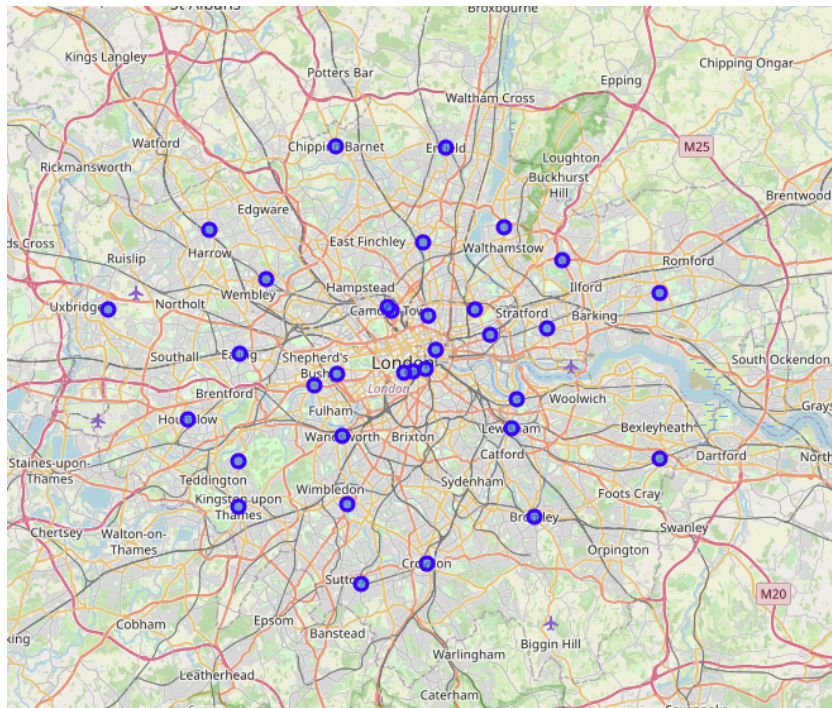
Source: Author's based on OpenStreetMap.

Foursquare data was used to identify the top 5 most common venues. These categories were used to group the boroughs in 4 clusters. Since types of venue are categorical variables, k-modes clustering algorithms were applied. The k-modes algorithm developed is an extension of the k-means algorithm used for numerical values. It groups data in k clusters such that the distance between the observations to the assigned cluster modes is minimised.

Annual income mean values were grouped in the four clusters, in order to determine if the type of venue was associated with the mean income the clusters of boroughs. Then analysis of variance (ANOVA) was used to determine if the differences observed in income were statistically significant.

3. Results

Large differences were observed in the average net annual income of London boroughs. As shown in Table 1, while in some boroughs, such as Barking and Dagenham and Enfield, the annual income is below £30K, in Richmond upon Thames and City of London, the annual income is over £40K.

Table 1. Net annual income of London boroughs

	Net annual income
count	33.000000
mean	32069.384795
std	3748.531588
min	25722.727273
25%	30000.000000
50%	31211.363636
75%	33486.666667
max	41800.000000

Source: Author's based on ONS data.

Differences in the type of most common venues across London boroughs were less evident. Since London is a very popular tourist destination, venues are very similar across boroughs. As shown in Table 2, some of the most common venues are pubs and coffee shops. Nonetheless, it was also possible to identify boroughs where the most common venues were more likely to serve locals rather than tourists, such as Enfield where optical shops and supermarkets were among the most common venues.

Mean values of the net annual income were computed for each of the clusters to identify whether the type of venues (clusters) were associated with the mean values of income (Table 3). Differences were observed particularly between clusters 0 and 2. However, these differences were not found statistically significant from the ANOVA performed.

Table 2. Five most common venues in London boroughs

	Borough	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Barking and Dagenham, London, UK	Bus Stop	Convenience Store	Liquor Store	Grocery Store	English Restaurant
1	Barnet, London, UK	Coffee Shop	Pub	Pharmacy	Convenience Store	Restaurant
2	Bexley, London, UK	Pub	Greek Restaurant	Breakfast Spot	Toy / Game Store	Train Station
3	Brent, London, UK	Coffee Shop	Hotel	Supermarket	Platform	Pedestrian Plaza
4	Bromley, London, UK	Clothing Store	Coffee Shop	Burger Joint	Gym / Fitness Center	Pub
5	Camden, London, UK	Pub	Coffee Shop	Burger Joint	Italian Restaurant	Café
6	City of London, London, UK	Coffee Shop	Gym / Fitness Center	French Restaurant	Seafood Restaurant	Clothing Store
7	Croydon, London, UK	Pub	Coffee Shop	Spanish Restaurant	Malay Restaurant	Gaming Cafe
8	Ealing, London, UK	Coffee Shop	Pub	Platform	Clothing Store	Bakery
9	Enfield, London, UK	Clothing Store	Coffee Shop	Optical Shop	Supermarket	Pub
10	Greenwich, London, UK	Pub	Boat or Ferry	Burger Joint	Pizza Place	History Museum
11	Hackney, London, UK	Pub	Coffee Shop	Café	Brewery	Supermarket
12	Hammersmith and Fulham, London, UK	Pub	Café	Coffee Shop	Clothing Store	Sandwich Place
13	Haringey, London, UK	Café	Park	Fast Food Restaurant	Coffee Shop	Bistro
14	Harrow, London, UK	Afghan Restaurant	Fast Food Restaurant	Grocery Store	Clothing Store	Bakery
15	Havering, London, UK	Café	Pub	Coffee Shop	Ice Cream Shop	Bar
16	Hillingdon, London, UK	Pub	Chinese Restaurant	Fast Food Restaurant	Park	Cupcake Shop
17	Hounslow, London, UK	Coffee Shop	Hotel	Clothing Store	Fast Food Restaurant	Indian Restaurant
18	Islington, London, UK	Pub	Mediterranean Restaurant	Burger Joint	Bakery	French Restaurant
19	Kensington and Chelsea, London, UK	Café	Italian Restaurant	Cupcake Shop	Supermarket	Pub
20	Kingston upon Thames, London, UK	Coffee Shop	Café	Pub	Clothing Store	Italian Restaurant
21	Lambeth, London, UK	Coffee Shop	Hotel	Bar	Sandwich Place	Café
22	Lewisham, London, UK	Clothing Store	Coffee Shop	Café	Fast Food Restaurant	Bus Stop
23	Merton, London, UK	Tram Station	Pub	Construction & Landscaping	Steakhouse	Sushi Restaurant
24	Newham, London, UK	Pub	Bus Station	Café	Electronics Store	Flea Market
25	Redbridge, London, UK	Hotel	Metro Station	Pizza Place	Eastern European Restaurant	Pub
26	Richmond upon Thames, London, UK	Pub	Construction & Landscaping	Home Service	Bus Station	Eastern European Restaurant
27	Southwark, London, UK	Hotel	Pub	Coffee Shop	Gym / Fitness Center	Bar
28	Sutton, London, UK	Grocery Store	Italian Restaurant	Train Station	Bakery	Park
29	Tower Hamlets, London, UK	Pub	Coffee Shop	Pizza Place	Bus Stop	Burger Joint
30	Waltham Forest, London, UK	Pub	Bus Stop	Art Gallery	Auto Garage	Lounge
31	Wandsworth, London, UK	Clothing Store	Pub	Coffee Shop	Burger Joint	Gym / Fitness Center
32	Westminster, London, UK	Coffee Shop	Outdoor Sculpture	Historic Site	Café	Sandwich Place

Source: Author's based on Places API -Fousquare data.

Table 3. Mean net annual income per cluster

| Cluster Labels	Net annual income
0	32483.268996
1	30484.280303
2	28472.222222
3	30845.161290

Source: Author's based on ONS data.

4. Discussion

Significant differences were observed across London boroughs in terms of household's net annual income. Although some differences were identified in the type of most common venues, more commonalities than differences were found. Food establishments, such as pubs, coffee shops and restaurants were the most common venues across the 33 London boroughs. This finding is related with the fact that London is a popular tourist attraction. Tourism is an important source of employment an income in London; however, this also means that locals may be paying higher prices when eating out. Further analysis would be needed to identify the impact of this issue on Londoners' lives.

Non-significant association was identified between the geospatial distribution of household's annual incomes and the most common type of venues available in the area. Since venues were similar across boroughs, cluster analysis generated uneven groups. However, as shown in Figure 2, the smaller clusters (purple, green and blue) tended to be located at the edges of London (region). These clusters group the boroughs that showed the main differences in the type of venues. An interesting finding is that some of these boroughs were Barking and Dagenham and Enfield, two of the poorest boroughs in London, as mentioned above. Further analysis would be needed to explore this relationship in more detail.

Figure 2. London boroughs grouped by type of venue
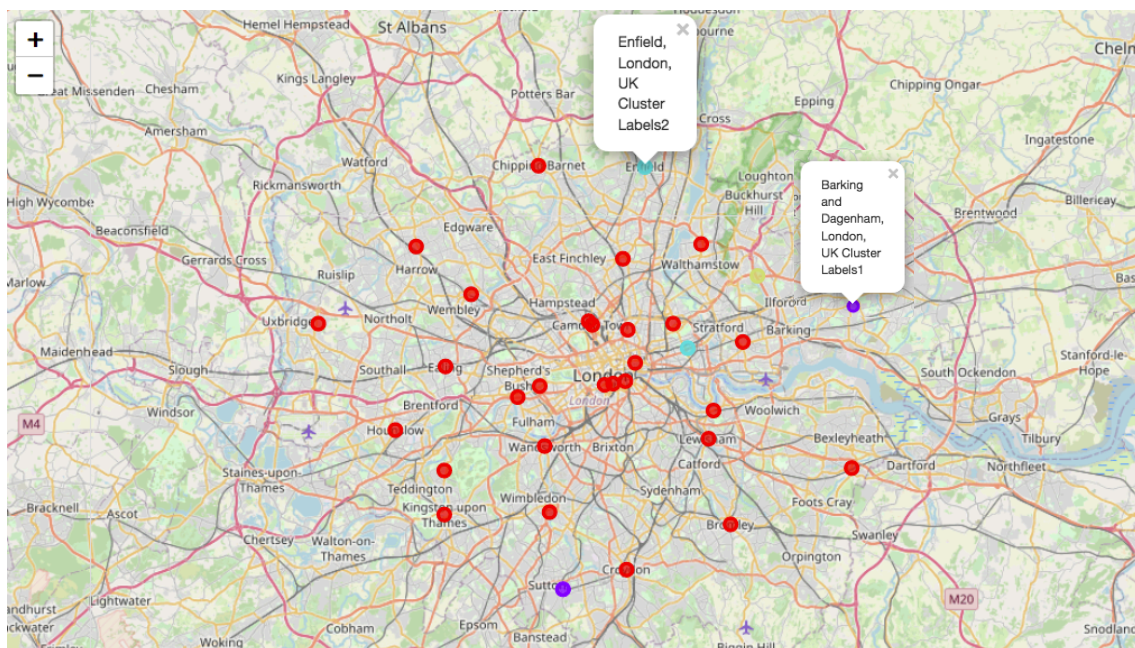
In light of these findings, the main recommendation of this report is to conduct further analysis to study the geospatial inequality across London boroughs and its implications. One potential direction of analysis is to reclassify the venues depending on the function performed, for example: food preparation, medical services, accommodation, etc.

5. Conclusions

This project explored whether the type of venues found in London's boroughs is associated to the average income of households located there. As a popular tourist destination, London is a challenging case to identify differences in the type of venues across boroughs. Nonetheless, some indication was found of poorer boroughs located at the edges of London (region) and showing venues different than in other areas. These venues included supermarkets and optical shops, while in other boroughs the most common venues were pubs, coffee shops and restaurants.

London is one of the richest regions in the UK, however large income and wealth inequalities are observed across households and these gaps are larger than in the rest of the UK. The analysis presented in this report provided exploratory evidence on how income inequality may be associated with the type of venues available in London boroughs. Businesses may use these findings to identify which are the best areas (at the edges) to locate shops targeting demand of London residents with lower incomes. Nonetheless, further analysis would be needed to provide specific policy recommendations.