Deciding where to open a new venue in London, Outer Boroughs

Introduction

London is a great city. It is diverse, multicultural and full of opportunities. But there is also another side to its popularity and appeal. London is an expensive place to live and thrive. Many businesses want to open a venue here and many pay top price to have their window on Piccadilly or Oxford Street. But what about those, who want to start their own business and cannot really afford to open in the City yet? Where is it best to open a new place? Where will it be cheapest and will have enough people living around to be popular? Where the competition is not too overwhelming?

If we consider all of these questions, it might be a good idea to turn to data on the outer London boroughs, to look into the numbers with a bit more scrutiny. In my analysis I have turned to:

1. Wikipedia list of London Boroughs with coordinates.
2. Foresquare data on the most popular venues in the respective boroughs
3. Online based data on rent in London boroughs

Methodology

But first things first – let`s take a look at the boroughs we are to investigate!

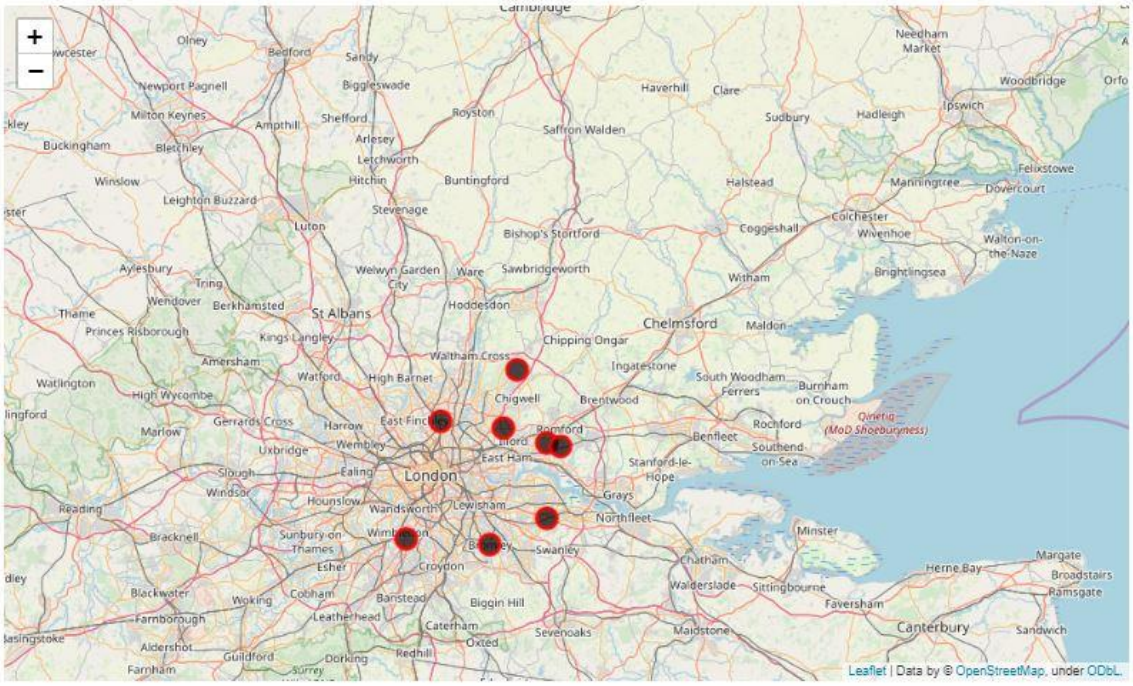
	Borough	Inner	Status	Local authority	Political control	Headquarters	Area (sq mi)	Population (2013 est)[1]	Co-ordinates	Nr. in map
0	Barking and Dagenham [note 1]	NaN	NaN	Barking and Dagenham London Borough Council	Labour	Town Hall, 1 Town Square	13.93	194352	51°33′39″N 0°09′21″E / 51.5607°N 0.1557°E	25
1	Barnet	NaN	NaN	Barnet London Borough Council	Conservative	North London Business Park, Oakleigh Road South	33.49	369088	51°37′31″N 0°09′06″W / 51.6252°N 0.1517°W	31
2	Bexley	NaN	NaN	Bexley London Borough Council	Conservative	Civic Offices, 2 Watling Street	23.38	236687	51°27′18″N 0°09′02″E / 51.4549°N 0.1505°E	23
3	Brent	NaN	NaN	Brent London Borough Council	Labour	Brent Civic Centre, Engineers Way	16.70	317264	51°33′32″N 0°16′54″W / 51.5588°N 0.2817°W	12
4	Bromley	NaN	NaN	Bromley London Borough Council	Conservative	Civic Centre, Stockwell Close	57.97	317899	51°24′14″N 0°01′11″E / 51.4039°N 0.0198°E	20
5	Camden	NaN	NaN	Camden London Borough Council	Labour	Camden Town Hall, Judd Street	8.40	229719	51°31′44″N 0°07′32″W / 51.5290°N 0.1255°W	11
6	Croydon	NaN	NaN	Croydon London Borough Council	Labour	Bernard Weatherill House, Mint Walk	33.41	372752	51°22′17″N 0°05′52″W / 51.3714°N 0.0977°W	19
7	Ealing	NaN	NaN	Ealing London Borough Council	Labour	Perceval House, 14-16 Uxbridge Road	21.44	342494	51°30′47″N 0°18′32″W / 51.5130°N 0.3089°W	13
8	Enfield	NaN	NaN	Enfield London Borough Council	Labour	Civic Centre, Silver Street	31.74	320524	51°39′14″N 0°04′48″W / 51.6538°N 0.0799°W	30

After downloading the entire page from Wiki, it is rather obvious, that we will not need all of these columns – and they will distract us going forward. So we will trim the dataframe a little.

	Borough	Area	Population	Latitude	Max_Rent	Longitude
0	Barking and Dagenham	13.93	194352	51.554117	102.25	0.150504
2	Bexley	23.38	236687	51.441679	97.00	0.150488
4	Bromley	57.97	317899	51.402805	118.50	0.014814
8	Enfield	31.74	320524	51.663600	102.25	0.079900
12	Haringey	11.42	263386	51.587930	107.75	-0.105410
14	Havering	43.35	242080	51.549900	86.00	0.183700
22	Merton	14.52	203223	51.410803	123.75	-0.188099
24	Redbridge	21.78	288272	51.576320	118.50	0.045410

Much better! Now we will use the Folium map to see where the outer boroughs are spread around London

Out[148]:



First thing that springs to mind is how far apart they are, except maybe the top right corner ones. Due to the different size boroughs, quite a drastic difference in cultural and ethnic diversity in each borough, it will be more than interesting clashing them against each other!

To get a feel for what is the most and least favored venues for the people who live in the boroughs in question, we will use Foursquare data for venue exploration around the borough coordinates. Since in London you have everything pretty much everywhere, we will need to limit the amount of reviewed venues. For that, we will need to filter each borough`s preferences for their top 5 places of interest in the format below.

```
brgh_venues_sorted.head(8)
```

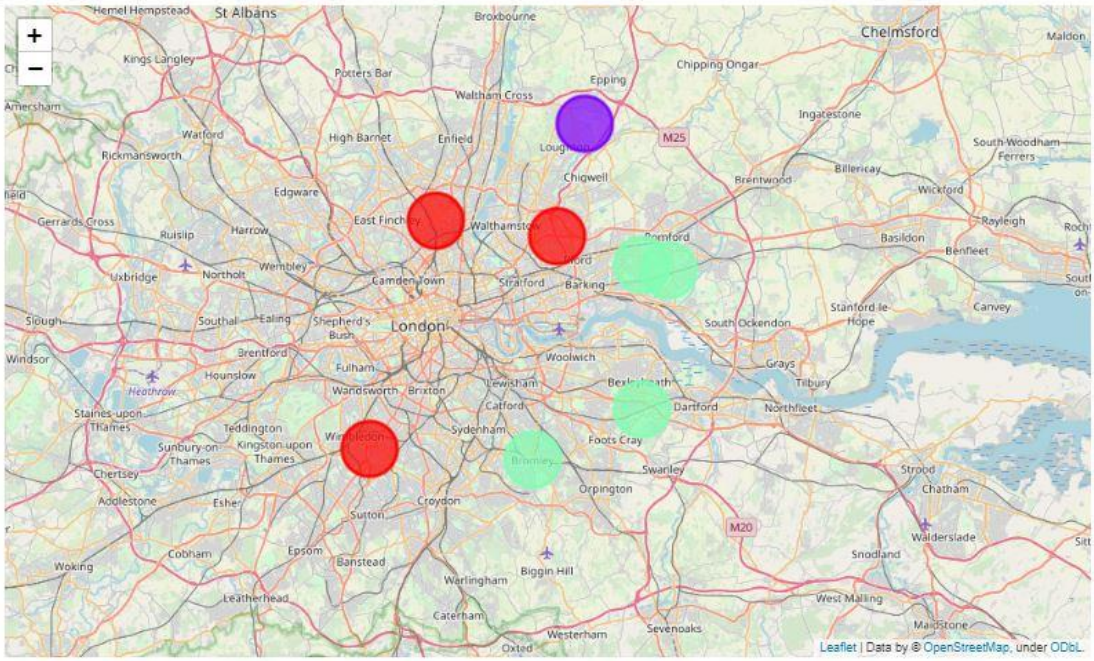
:

	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	Supermarket	Grocery Store	Coffee Shop	Pub	Park
1	Bexley	Grocery Store	Pub	Coffee Shop	Supermarket	Park
2	Bromley	Coffee Shop	Pub	Grocery Store	Park	Gym / Fitness Center
3	Enfield	Pub	Coffee Shop	Café	Italian Restaurant	Grocery Store
4	Haringey	Pub	Coffee Shop	Park	Café	Turkish Restaurant
5	Havering	Pub	Coffee Shop	Supermarket	Grocery Store	Italian Restaurant
6	Merton	Pub	Park	Italian Restaurant	Coffee Shop	Sushi Restaurant
7	Redbridge	Pub	Park	Coffee Shop	Pizza Place	Restaurant

Now that we have figured out the top places for every borough – it`s time to cluster. Because the amount of boroughs is so little (on earlier stages it was decided to slash some of them due to high rent price compared to the finalists) the amount of clusters will be small.

```
map_clusters
```

Out[159]:



Although the picture, that seemingly is worth a thousand words, is rather divisive. Until we look at the clusters individually, and try to determine patterns.

```
Borough_merged.loc[Borough_merged['Cluster Label'] == 0, Borough_merged.columns[[0,1,2,4] + list(range(6, Borough_merged.shape[1]))]]
```

:

	Borough	Area	Population	Max_Rent	Cluster Label	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
4	Haringey	11.42	263386	107.75	0	Pub	Coffee Shop	Park	Café	Turkish Restaurant
6	Merton	14.52	203223	123.75	0	Pub	Park	Italian Restaurant	Coffee Shop	Sushi Restaurant
7	Redbridge	21.78	288272	118.50	0	Pub	Park	Coffee Shop	Pizza Place	Restaurant

The first pattern and a very British picture: Pub as a most popular venue, and for those, who don`t feel like a visit to a pub can enjoy the nature. This cluster is hardly of interest, because it will be problematic to open a new Pub. Not to mention a Park... Amongst the most popular venues are also restaurants, which can be of interest to foodies and the chefs living close by.

```
Borough_merged.loc[Borough_merged['Cluster Label'] == 1, Borough_merged.columns[[0,1,2,4] + list(range(6, Borough_merged.shape[1]))]]
```

:

	Borough	Area	Population	Max_Rent	Cluster Label	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
3	Enfield	31.74	320524	102.25	1	Pub	Coffee Shop	Café	Italian Restaurant	Grocery Store

Second cluster, and the most controversial. Enfield is one of the bigger outer London boroughs, and the data, that we get can differ massively within the same borough: those living closer to M25 in the suburban houses tend to stay indoors and spend more time with their families and cook at home. On the other hand, those living closer to the inner borough limits may be more outgoing, and because of this contribute the most to our dataset. Regardless, with not too high rent, big and a very diverse population, Enfield is definitely worth considering

```
Borough_merged.loc[Borough_merged['Cluster Label'] == 2, Borough_merged.columns[[0,1,2,4] + list(range(6, Borough_merged.shape[1]))]]
```

]:

	Borough	Area	Population	Max_Rent	Cluster Label	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Barking and Dagenham	13.93	194352	102.25	2	Supermarket	Grocery Store	Coffee Shop	Pub	Park
1	Bexley	23.38	236687	97.00	2	Grocery Store	Pub	Coffee Shop	Supermarket	Park
2	Bromley	57.97	317899	118.50	2	Coffee Shop	Pub	Grocery Store	Park	Gym / Fitness Center
5	Havering	43.35	242080	86.00	2	Pub	Coffee Shop	Supermarket	Grocery Store	Italian Restaurant

Biggest and the most diverse cluster. Most of the boroughs are not heavy drinkers, but neither are they very outgoing. Because I would not call visit to the grocery store as a social route! However, Havering and Bromley are good fits to our target boroughs: both have entertainment venues as most common, both rather big and quite populated. In this situation perhaps Havering has a slight financial edge - cheaper rent a more incline to pubs and coffee shops in the area.

Results

Results of the above analysis and clustering cam be summarized:

1. The most popular social venues, ouside of Inner London boroughs are Pubs and Coffee shops
2. Northern boroughs are more prone to visiting pubs, whereas southern boroughs are most likely to shop and have the social life from home
3. Within top 5 places of interest in every borough is an ethnic restaurant
4. Rent price is not so much a factor for going out - the demand is not affected by difference in costs

Discussion

Looking at the data, Havering, Bromley and Enfield are the best places outside of Central London where a new venue is worth opening. However, a lot of information is not taken into account, and cannot be obtained from Foursquare Developer:

1. Higher ethnic presence in a given borough can and will influence the popularity of a given cuisine.
2. Closer proximity to Inner boroughs and better transport links allows people to travel to the neighbouring borough and impact the measurements
3. Many small venues are not registered in Foursquare and are marketed via word-of-mouth, and are not taken into account

Regardless, the analysis provided an insight into what people like and opt for, when it comes to going out in their own neighbourhoods.

Conclusion

Finally to conclude this project, I have had a good trial run at solving a real-life problem, using available data to find a business solution - choosing to open a venue in London .I have made use of some frequently used python librairies to manipilate data, use Foursquare API to explore the information on the Boroughs I looked into and managed to make a map of results, that allowed me to ilustrate my point graphicly and quite clearly to someone, not familiar with data manipulation and who only wants to know one thing - where will my venue be flourishing??