Finding Best Venue for new Business in outer London Boroughs

Hassan Ashiq

August, 1, 2020

1. Introduction :

1.1 Background :

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. With such a competitive market and high revenues (London GDP accounts for 25% of the whole country`s) it is challenging for newcomers to find a great place to make their first steps and competition if fiercer than ever. In this situation, there is no such thing as too much data to help make that choice.

1.2 Problem

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 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.

1.3 Interest

The analysis will be of interest to the following groups:

1. First time entrepreneurs, who want to start their first business. Below dataset will give a comprehensive insight into where best to open a new venue, to maximise the value for money.
2. People who already run a business and want to branch out. Given the extra information, it may provide valuable information before decision making.

2. Data acquisition and cleaning

2.1 Data sources

As data sources the below were chosen:

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

2.2 Data cleaning

Once the dataset of London boroughs has been downloaded, we must edit the dataset provided to only have information, necessary for our problem. Wikipedia provided information on political situation, headquarters of the borough council etc. that will not be required. After cleaning, we will only be left with Name, Area, Population, Coordinates and rent prices for each borough. Since the dataset will only include outer boroughs, all the inner ones will be omitted as well. Foursquare provides a dataset of venues around the specific coordinates or venues, if we use the "Explore" function in the Developer tab. Once requested, we get a full breakdown of all recorder venues around the boroughs of interest.

2.3 Feature selection

For convenience, we transform the dataset only to show top 5 places to work with. After that, we merge the data frames together for a comprehensive set of values, worth analysing.

	Borough	Area	Population	Max_Rent	Latitude	Longitude	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	51.554117	0.150504	1	Supermarket	Park	Grocery Store	Coffee Shop	Pub
1	Bexley	23.38	236687	97.00	39.969238	-82.936864	2	Pizza Place	Coffee Shop	Ice Cream Shop	Discount Store	Chinese Restaurant
2	Bromley	57.97	317899	118.50	51.402805	0.014814	1	Coffee Shop	Pub	Grocery Store	Park	Pizza Place
3	Enfield	31.74	320524	102.25	51.652085	-0.081018	1	Pub	Coffee Shop	Turkish Restaurant	Park	Supermarket
4	Haringey	11.42	263386	107.75	51.587930	-0.105410	1	Café	Pub	Turkish Restaurant	Park	Coffee Shop
5	Havering	43.35	242080	86.00	51.004361	-2.337475	0	Event Service	IT Services	Indoor Play Area	Hotel	Electronics Store
6	Merton	14.52	203223	123.75	51.410803	-0.188099	1	Pub	Park	Coffee Shop	Café	Bar
7	Redbridge	21.78	288272	118.50	51.576320	0.045410	1	Pub	Park	Café	Coffee Shop	Restaurant

3. Data Analysis

3.1 Clustering

Once the boroughs are selected, we will place them onto the map and cluster the boroughs to analyse the similar ones against each other and within the clusters themselves. Each of the clusters will be compared against the popularity of the venues, worth considering for a new business venture, as well as the hospitability of the venture climate. Depending on the area, population and rent, each of the clusters offer different advantages and disadvantages in terms of venue choices.
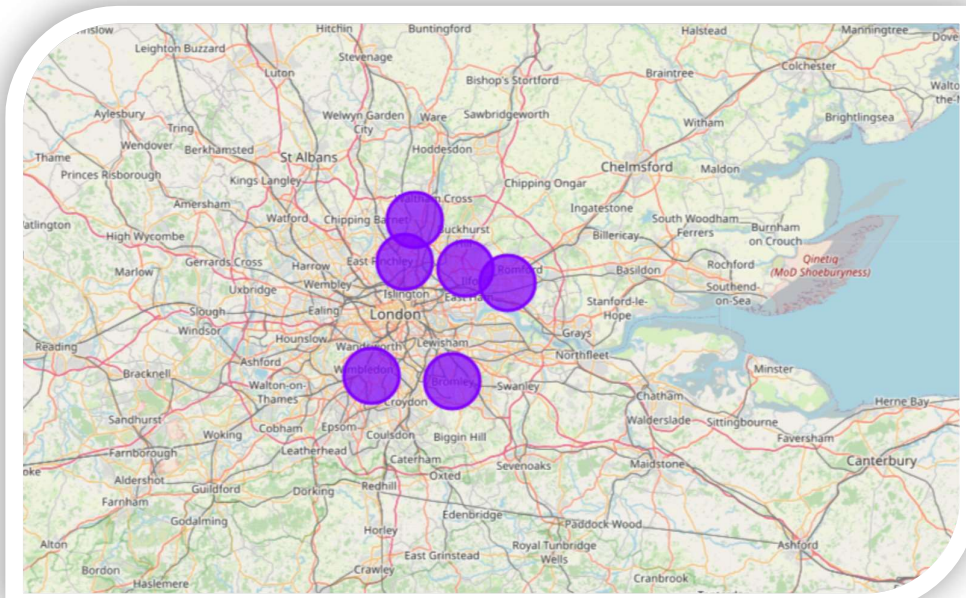
4. Results

Reviewing each cluster as per method above, we learn:

1. First Cluster that is Cluster with label 0, have a very Electric Type Picture. Event Service is a most popular venue. This cluster is of no interest since most the venues are related to IT or Electric Appliances.

2. Second cluster, that is Cluster with Label 1 : 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, Haringey 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 Haringey has a slight financial edge - cheaper rent a more incline to pubs and coffee shops in the area.

3. Third Cluster, that is Cluster with Label 2: Not that much good. With little Area there are a lot of things for eating only. No park, no entertainment etc. So, on should not go for it

4. Within top 5 places of interest in every borough is an ethnic restaurant. Because of the different ethnicities in the boroughs, some choices will be more favourable amongst the specific group in the area.

5. Rent price is not so much a factor for going out - the demand is not affected by difference in costs. There is a spike in rent price going into London, but further away the cost is not too much of an issue.

5. 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 considered, and cannot be obtained from Foursquare Developer:

- Higher ethnic presence in each borough can and will influence the popularity of a given cuisine.
- Closer proximity to Inner boroughs and better transport links allows people to travel to the neighbouring borough and impact the measurements
- Many small venues are not registered in Foursquare and are marketed via word-of-mouth, and are not considered.
- Regardless of all that, our analysis provided an insight into what people like and opt for, when it comes to going out in their own neighbourhoods.

6. Conclusion

Finally to conclude this project, I have had a good experience of a trial run at solving a real-life problem, using available data to find a business solution. Which in our case was to find a best venue for new Business Venture in London.

- Besides that this project made me to :
- Use of some frequently used python libraries to manipulate 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 illustrate 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??
- Overall, this project gives me a really good platform to practice my all previous courses in this specialization.

Thank You Coursera and IBM for changing lives.

End Our Journey at IBM Data Science Specialization

Project by : Hassan Ashiq

https://www.linkedin.com/in/hassan-ashiq