Applied data science capstone project ‘The battle of neighbourhoods’

Opening an Italian resturant in London

Francesco Bianconi
bianco@ieee.org

Wednesday 24th June, 2020

Background

London is the capital and largest city of England and the United Kingdom with a population of $\approx$9M

It is one of the most influential, most visited and most culturally active place in the world

Offers countless opportunities to the visitor and the residents: '*When a man is tired of London, he is tired of life*' (Samuel Jackson)

Objectives

We want to open an Italian restaurant in London; the problem is to decide the most convenient place

Methodology I

We considered the town in London divided into its boroughs

The tentative location for the new restaurant would be a generic place within one mile from the centre of the borough

Methodology II

We compared the boroughs based on the following features:

- Number of **Italian restaurants** already present (the fewer the better)
- Average **income** of each borough (the higher the better)
- **Population** of the borough (the more populated the better)
- Presence of amenities around like **cinemas** and **theatres** (the more the better)
- Presence of **universities** around (the more the better)

We ranked the boroughs by each of the above features and obtained the overall rank by averaging the results (we assumed that each feature weighed the same in determining the overall ranking)

Data sources

We retrieved the data from the following following sites:

- `https://en.wikipedia.org/wiki/London_boroughs` (list of London boroughs)
- `https://data.london.gov.uk/dataset/earnings-place-residence-borough` (average income by borough, 2018)
- `https://www.citypopulation.de/en/uk/greaterlondon/` (population by borough, 2018)

Data collection and analysis

We used Python 3.7.4 and the following packages:

- Folium for map generation
- Foursquare to get the points of interest around the target locations
- GeoPy to retrieve the location of the centre of each borough
- Pandas for data collection and manipulation

All the source code is available in the notebook

Results I

The following table shows the data arranged in a Pandas dataframe (first five rows):

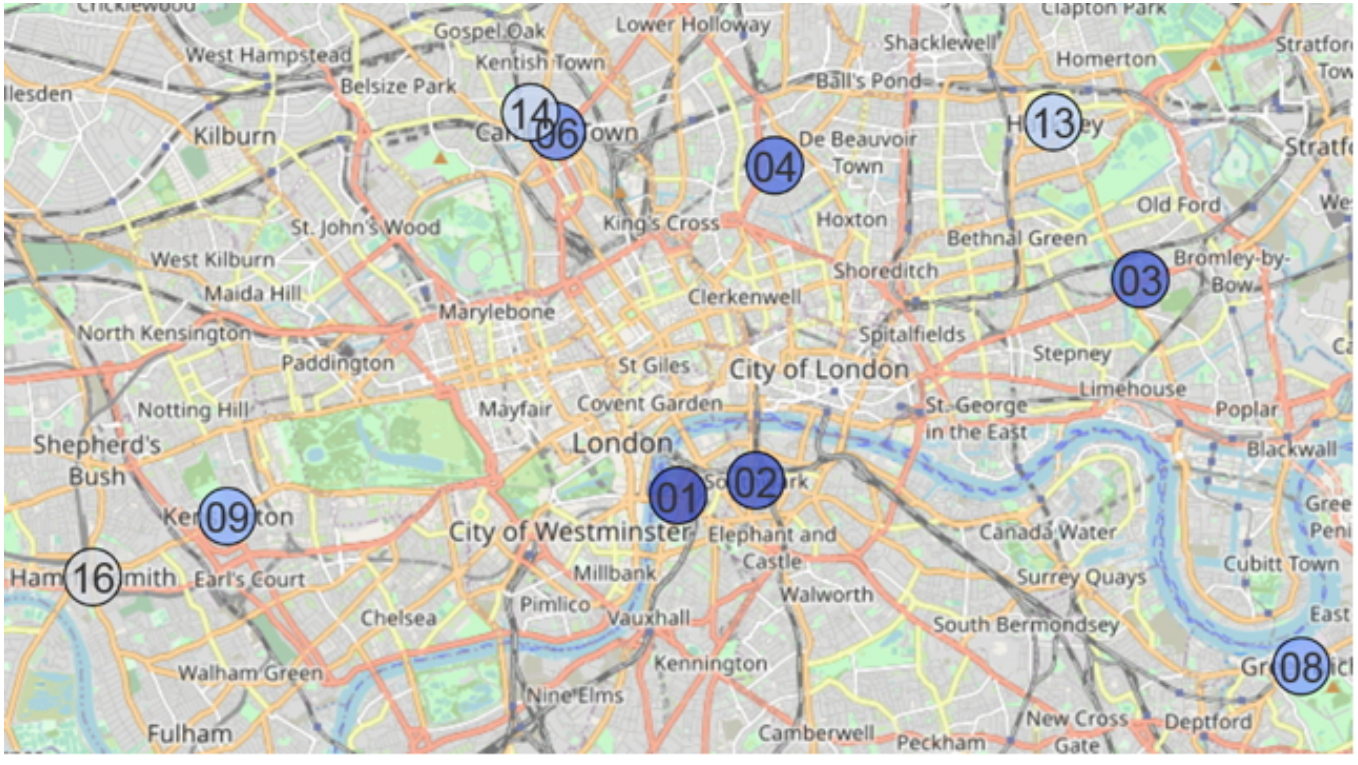
	Cinema	Income	Italian Restaurant	Latitude	London borough	Longitude	Population	Theatre	University
0	5	634.7	50	51.542305	Camden	-0.139560	262226.0	25	10
1	3	573.7	29	51.482084	Greenwich	-0.004542	286186.0	4	25
2	4	555.6	34	51.543240	Hackney	-0.049362	279665.0	4	5
3	4	681.3	50	51.492038	Hammersmith and Fulham	-0.223640	185426.0	10	1
4	9	687.6	50	51.538429	Islington	-0.099905	239142.0	28	50

And here are the data after ranking (top five results):

	London borough	Population_rnk	Income_rnk	Italian Restaurant_rnk	Cinema_rnk	Theatre_rnk	University_rnk	Overall_rnk
6	Lambeth	23.0	22.0	4.0	31.0	30.5	28.0	23.083333
8	Southwark	21.0	20.0	4.0	30.0	30.5	27.0	22.083333
9	Tower Hamlets	22.0	24.0	9.0	27.0	26.0	23.0	21.833333
4	Islington	8.0	30.0	4.0	29.0	29.0	31.0	21.833333
10	Wandsworth	24.0	31.0	22.5	10.0	19.5	13.5	20.083333

Results II

The picture shows the results on a map (interactive version available here)



Conclusions

Our study indicates that the best boroughs to open an Italian restaurant in London are (in top-down order):

- Lambeth (ranked 1st)
- Southwark
- Tower Hamlets
- Islington
- Wandsworth

Please visit the notebook for the complete placings