Applied Data Science Capstone Project – The Battle of Neighbourhoods
Opening an Italian restaurant in London

1. Background
London is the capital and largest city of England and the United Kingdom. With a population of about nine million it is one of the most influential, most visited and most culturally active place in the world. London offers countless opportunities to the visitor and the residents: as Samuel Jackson put it, '*when a man is tired of London, he is tired of life; for there is in London all that life can afford*'. As an Italian and former London resident I therefore decided that London could be the best place to set my project.

2. Business problem and methodology
The business plan consists in opening an Italian restaurant in London, and the problem is to decide the most convenient place. We consider the town of London divided into its boroughs. The tentative location for the new restaurant will be a generic place within one mile from the centre of the borough. The problem therefore reduces to identifying the most convenient borough. To this end we consider the following parameters:
1. Number of Italian restaurants already present (the fewer the better);
2. Average income of each borough (the higher the better);
3. Population of the borough (the more populated the better);
4. Presence of amenities around like cinemas and theatres (the more the better);
5. Presence of universities around (the more the better).

3. Data
We looked up the web for the information needed and in particular the following sites:
6. https://en.wikipedia.org/wiki/London_boroughs for the list of London boroughs;
7. https://data.london.gov.uk/dataset/earnings-place-residence-borough for the average income in each borough;
8. https://www.citypopulation.de/en/uk/greaterlondon/ for the population of each borough.

We used GeoPy to retrieve the location of the centre of each borough and Foursquare API to determine the points of interest (cinemas, theatres and universities) around the centre. Data analysis was carried out on Pandas.

4. Results and discussion
Table 1 summarises the data retrieved from the web. From Tab. 1 we obtained a rank table (Tab. 2) by sorting the following columns in the order shown in parenthesis: 'Income' (descending – i.e., highest value = highest rank), 'Italian Restaurant' (ascending – i.e., lowest value = highest rank), 'Population' (descending), 'Cinema' (descending), 'Theatre' (descending) and 'University' (descending). Finally, we obtained the overall rank (last column) by averaging over the above columns.

Table 1 Data retrieved from the web – head of the dataframe

	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

Table 2 Head of the dataframe after ranking

	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

The results indicate that Lambeth is the most suitable borough for opening an Italian restaurant, followed by Southwark and Tower Hamlets (complete results available in the notebook). Figure 1 shows the results on a map.

Figure 1 Most suitable boroughs for opening an Italian restaurant in London (interactive version of the map available *here*).