Capstone Project – The Battle of Neighbourhoods (Week 1)

Data/Methodology

We will scrap the web for the information needed, and in particular:

- https://en.wikipedia.org/wiki/London_boroughs for the list of London boroughs;
- https://data.london.gov.uk/dataset/earnings-place-residence-borough for the average income in each borough;
- https://www.citypopulation.de/en/uk/greaterlondon/ for the population of each borough.

We will use 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 will be carried out on Pandas.

The final objective is to determine a ranking of boroughs for the new restaurant, from the most to the least suitable one. The ranking will be based on the features listed above (1—5). We make the simplifying assumption that each feature will weigh the same in determining the global ranking. Therefore, we will compute a rank for each feature separately and average the results to obtain the overall ranking. Finally, we will show the results on a map using Folium.