Coursera Capstone – The Battle of Neighbourhoods

Data Description

Data description

Data URL - https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

We will use the Scarborough dataset that was scraped from Wikipedia on Week 3.
This dataset should provide us with details such as latitude and longitude values and zip codes which is paramount to helping us pick the right location.

We will need to use Foursquare API locational information in order to gain data regarding different venues in the neighbourhoods of each borough. Foursquare API will also provide us with various categories associated with that neighbourhood such as population which is a key factor in picking the best location for high demand.

The information gathered from each neighbourhood will be as follows:

1. Neighbourhood
2. Neighbourhood Latitude
3. Neighbourhood Longitude
4. Venue Name
5. Venue Category
6. Venue Latitude
7. Venue Longitude

How will this data be used

With these features, along with the techniques such as K-clustering to segment and cluster the neighbourhoods, we will be able to provide IndyRest with the best possible neighbourhood location to start off the beginning of their journey. By analysing the most common venues and their categories in each neighbourhood, this will give us an idea of the competition levels as well as if there is a gap in the market for this particular venture. We will want to find a location which is densely populated and preferably with the least amount of restaurants in the vicinity which lowers competition levels making the restaurant a market leader in this location. Other additional features which could improve demand is to analyse the most common language spoken in the area which can improve demand with people who are accustomed to this type of cuisine.