Capstone Project: The Battle of Neighborhoods

Naiara Marcos Gonzalez

April 20, 2020

1 Introduction

1.1 Description of the problem

Previous weeks, we studied two cities, New York and Toronto, which are two big financial cities from North America and their population is made up of diverse nationalities. In this project, we will focus on comparing London to these other two cities and discuss if it is more similar to New York or Toronto or none.

1.2 Discussion of the background

Finding similarities or dissimilarities between cities can be very interesting from different points of views. For example, expanding business, that is to say, if there is a business or infrastructure which is working well in one city, maybe it would be a good idea reproduce it in the city which looks more alike. Another possible example, it might be finding accommodation for people who move from another city and they want to live in similar area than they used to, and so on.

In this study, we chose London because it shares the same characteristics as New York and Toronto but in Europe, it is a big financial centre with high and heterogenous population.

1.3 Target audience

The potential audience of this project is a person or company which already has a business in London and wants to move/expand to North America, but he has doubts about which of the two cities (New York and Toronto) is a better fit. For that reason,

we should study the similarity between these three cities: London, New York and Toronto.

2 Data

In order to compare these three cities, we will use the studies of New York and Toronto which we made the previous weeks in Coursera. We add new data from the city of London. The data sources are:

New York:`https://geo.nyu.edu/catalog/nyu_2451_34572`
Toronto:`https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M`
London:`https://www.doogal.co.uk/london_postcodes.php`

This data provides us the Postcode, Borough's Name, Neighbourhood, the respective Latitude and Longitude for each Neighbourhood that every city contains. From this data, we can use the Foursquare API to get information about restaurants, coffee shops, parks, theatres, and so on for each neighbourhood. Then, with this data we plan to apply the k-means algorithm to create a cluster analysis of the neighbourhoods from each cities. With these analysis we plan to tell which city is London more alike and to find a correspondence between each city neighbourhood.