Identifying the Best Locations in the City of Toronto for a New Restaurant

# Dataset

The project uses Foursquare location data. It also uses a Wikipedia web page that contains table of Toronto neighbourhoods and their postal codes. The Wikipedia web page is <https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M>. In addition to that, one more dataset is used which contains latitudes and longitudes of Toronto Neighbourhoods. This data is provided by IBM in a data science course they launched through Coursera.com (coursera.com, 2020).

Foursquare is “a location technology platform dedicated to improving how people move through the real world” (foursquare.com, 2020). This project utilises Foursquare location data and other type of data like ratings and comments of users.

The table that exists in the Wikipedia page contains three features: postal code, borough and neighbourhood (Table 1).

Table 1 Sample of Postal Codes of Toronto Neighbourhoods

|  |  |  |
| --- | --- | --- |
| **Postal Code** | **Borough** | **Neighbourhood** |
| **M1A** | Not assigned | Not assigned |
| **M1B** | Scarborough | Malvern, Rouge |
| **M1C** | Scarborough | Rouge Hill, Port Union, Highland Creek |
| **M1E** | Scarborough | Guildwood, Morningside, West Hill |
| **M1G** | Scarborough | Woburn |

A sample of the latitude and longitude dataset is shown in Table 2.

Table 2 Sample of Latitude and Longitude Data

|  |  |  |
| --- | --- | --- |
| Postal Code | Latitude | Longitude |
| M1B | 43.80669 | -79.1944 |
| M1C | 43.78454 | -79.1605 |
| M1E | 43.76357 | -79.1887 |
| M1G | 43.77099 | -79.2169 |
| M1H | 43.77314 | -79.2395 |

# Methodology

The project will retrieve data of all neighbourhoods from Foursquare and then analyse this data. The analysis will include finding all businesses in the neighbourhoods, finding all restaurants in each and every neighbourhood, clustering and grouping all restaurants within 100 metres from each successful business (based on user ratings), and then selecting the location that has a successful business with the least number of restaurants.

This approach is selected based on the fact that successful businesses attracts more customers. So, starting a new restaurant near a successful business is bound to succeed. However, if there are too many restaurants near that business, then starting a new restaurant in that area is not the best idea.

# References

Coursera.com, 2020. “Applied Data Science Specialization”. Internet source, retrieved on 25-10-2020 from: <https://www.coursera.org/specializations/applied-data-science?skipBrowseRedirect=true>

Foursquare.com, 2020. “About Us”. Internet source, retrieved on 25-10-2020 from: <https://foursquare.com/about>