**SUGGESTING THE BEST OFFICE LOCSTION FOR INTERNET SERVICE PROVIDING COMPANY**

**Nana Hamzat**

**April 9, 2020**

1. **Problem**

Starting a new business is one of the greatest challenges. The success of the business is also unpredictable. Data has improved the way we do business and has an improved guarantee of success.

A new internet service provider (TELNEL) wants to open its first office in the heart of Toronto alongside its High priority base station. Our target location is generally a busy spot where there are lots of active daily routines and business activities.

1. **Data Collection**

The Dataset will be scrapped from the Wikipedia page of Canada via, <https://en.wikipedia.org/w/index.php?title=List_of_postal_codes_of_Canada:_M&oldid=942851379>, in order to obtain the data that is in the table of postal codes and to transform the data into a pandas dataframe. Also, Foursquare API calls will be used to retrieve geolocation information about the venues in each Borough. And downloaded geographical coordinates will be used to find the latitude and longitude of each Neighborhood with unique postal codes.

The resulting dataframe has the following features, Postal Code, Borough, Neighborhood, Longitude and Latitude as shown in Table I.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Postcode | Borough | Neighbourhood | Latitude | Longitude |
| M1B | Scarborough | Rouge, Malvern | 43.806686 | -79.194353 |
| M1C | Scarborough | Highland Creek, Rouge Hill, Port Union | 43.784535 | -79.160497 |
| M1E | Scarborough | Guildwood, Morningside, West Hill | 43.763573 | -79.188711 |
| M1G | Scarborough | Woburn | 43.770992 | -79.216917 |
| M1H | Scarborough | Cedarbrae | 43.773136 | -79.239476 |

Table I

1. **Methodology**

In this project all venues will be explored as we are interested in the busy areas in Toronto. We will limit our analysis to area ~10km around all Districts in Toronto. We have collected and cleaned the required data and Foursquare API was used to get all venues around each district. And folium library was used to view the resulting data on a map.

The map below (Fig I) shows all the Boroughs with their neighborhoods popping up on click.

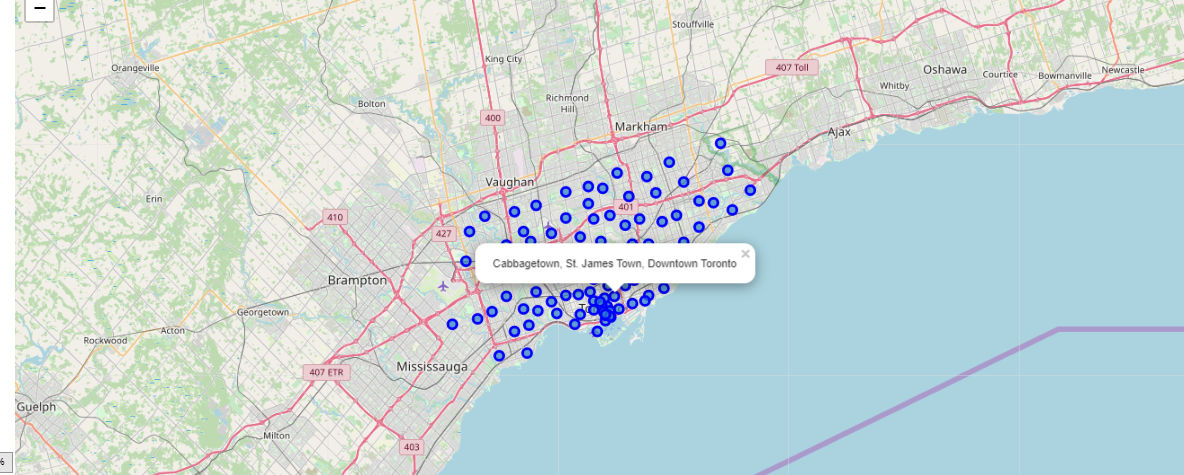


Fig I: Map of Toronto and its Neighborhoods

Scarborough was first analyzed using the foursquare API by setting the LIMIT as 100 and corresponding radius as 10000. As a result, 100 categories of venues was deduced. Table II below shows 10 out of the 100 categories of venues found around Scarborough with Rouge and Malvern as the neighborhood.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Name | Categories | Latitude | Longitude |
| 1 | Toronto Pan Am Sports Centre | Athletics & Sports | 43.790623 | -79.193869 |
| 2 | African Rainforest Pavilion | Zoo Exhibit | 43.817725 | -79.183433 |
| 3 | Polar Bear Exhibit | Zoo | 43.823372 | -79.185145 |
| 4 | Toronto Zoo | Zoo | 43.820582 | -79.181551 |
| 5 | Lamanna's Bakery, Cafe & Fine Foods | Bakery | 43.797971 | -79.148432 |
| 6 | Australasia Pavillion | Zoo Exhibit | 43.822563 | -79.183286 |
| 7 | Mona's Roti | Caribbean Restaurant | 43.791613 | -79.251015 |
| 8 | The Real McCoy Burgers And Pizza | Burger Joint | 43.774081 | -79.230496 |
| 9 | Port Union Waterfront Park | Park | 43.774967 | -79.134393 |
| 10 | Orangutan Exhibit | Zoo Exhibit | 43.818413 | -79.182548 |

Table II: Venues around Rouge and Malvern

The same analysis was repeated all other Boroughs with different Neighborhoods. As a result of this, 206 unique categories of venues were found in total. Then the top 10 most common venues were analyzed, and the results were shown as seen below.



**Clustering**

One of the clustering algorithms, K-Means was used to group the 206 unique categories of venues in all the Boroughs into 5 clusters. K-means was imported from sklearn library.

1. **Result.**

It was discovered that cluster 4 labelled green as seen in Fig II had two Borough with the highest number of venues. Downtown Toronto and Central Toronto. Still, Downtown Toronto had the highest number of venues between the two.

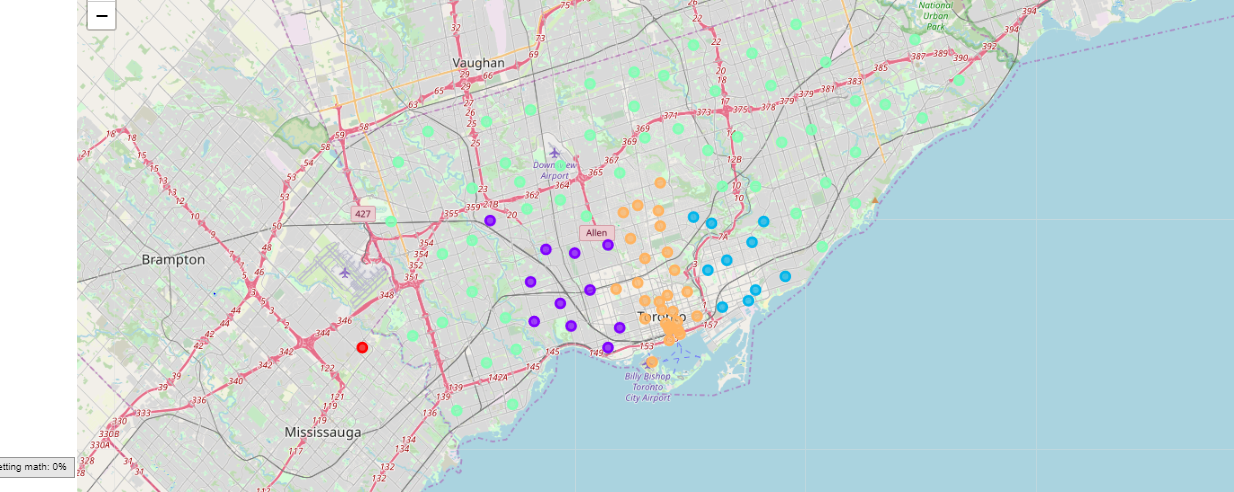


Fig II: Map of Toronto showing the 5 clusters

1. **Conclusion**.

It is suggested that TELNET office should be in the suburb of Downtown Toronto as it has the highest number of visitors thereby bringing more traffic.

1. **Recommendation.**

Although this project is built with focus on possible network traffic in a given area. Users can also be explored to determine the age category so as to deduce properly the possible number of internet users in the proposed area.