Perfect Location and Venue to Start a Business

-Swapnil Rane

1.0 Introduction

Problem Description

Scarborough is a popular destination for new immigrants in Canada to reside. As a result, it is one of the most diverse and multicultural areas in the Greater Toronto Area, being home to various religious groups and places of worship. It includes a number of natural landmarks, including the Toronto Zoo, Rouge Park and the Scarborough Bluffs. As my client need specific information regarding what kind of business, they should start which can make the highest profit and what kind business is trending in a particular area. So, me as Data Scientist will create a list of popular places and venues in the Scarborough city and provide it to my client

Analytic Approach

To find the solution for the above-mentioned issue, I will be using the neighborhood information that is available on the internet and then will use the Foursquare API to find the venues in a particular area.

2.0 Data:

Data Requirements

For this implementation I will be using the Wikipedia to gather data regarding postal codes of Canada.

Data Collection

a. I will do a web scraping of the following Wikipedia link ' https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M' using Beautiful Soup python library to get the CANADA neighborhood mapped to Postal Code and Borough. I will use 'https://cocl.us/Geospatial_data' to get the mapping of latitude and longitude of each postal code.

b. I will filter only the Scarborough Data from the table of data collected in step 1 for this project.

c. I will use foursquare API calls to get the location details corresponding to each neighborhood captured in step 2, something as mentioned below:

Data Understanding

For understanding the data collected in the Data collection stage, I will explore the data and do some analysis like:

a. Explore the neighborhoods in Scarborough

b. Analyze each areas in Scarborough.

c. Explore the venues corresponding to a particular area.

Data Preparation

a. If I see any gap in the data, the I will go back to the data collection stage and collect any missing data.

b. If I see any erroneous data / unnecessary data, then I will filter and remove those data to prepare the final data.
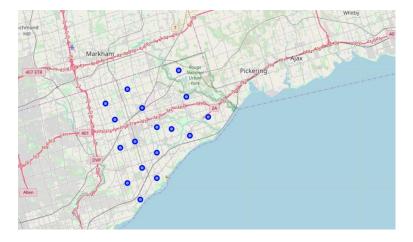
3.0 Data Analysis

Preparing neighborhood data

a. Importing all the libraries that are necessary for this project.

b. Extract the table with postal code, borough and neighborhood data from the Wikipedia link 'https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M' using Beautiful Soup library and save it in a csv file.

c. If a postal code has a borough but a Not assigned neighborhood (as shown below), then the neighborhood will be the same as the borough.

d. Group neighborhood based on postal code and borough.

e. Read the geo spatial data from 'https://cocl.us/Geospatial_data' and merge it with the data frame created in step d. and clean out every other Borough except the Scarborough

The resulting data frame looks something like:

	PostalCode	Borough	Neighborhood	Latitude	Longitude
0	M1B	Scarborough	Rouge, Malvern	43.806686	-79.194353
1	M1C	Scarborough	Highland Creek, Rouge Hill, Port Union	43.784535	-79.160497
2	M1E	Scarborough	Guildwood, Morningside, West Hill	43.763573	-79.188711
3	M1G	Scarborough	Woburn	43.770992	-79.216917
4	M1H	Scarborough	Cedarbrae	43.773136	-79.239476
5	M1J	Scarborough	Scarborough Village	43.744734	-79.239476
6	M1K	Scarborough	East Birchmount Park, Ionview, Kennedy Park	43.727929	-79.262029
7	M1L	Scarborough	Clairlea, Golden Mile, Oakridge	43.711112	-79.284577
8	M1M	Scarborough	Cliffcrest, Cliffside, Scarborough Village West	43.716316	-79.239476
9	M1N	Scarborough	Birch Cliff, Cliffside West	43.692657	-79.264848
10	M1P	Scarborough	Dorset Park, Scarborough Town Centre, Wexford ...	43.757410	-79.273304
11	M1R	Scarborough	Maryvale, Wexford	43.750072	-79.295849
12	M1S	Scarborough	Agincourt	43.794200	-79.262029
13	M1T	Scarborough	Clarks Corners, Sullivan, Tam O'Shanter	43.781638	-79.304302
14	M1V	Scarborough	Agincourt North, L'Amoreaux East, Milliken, St...	43.815252	-79.284577
15	M1W	Scarborough	L'Amoreaux West	43.799525	-79.318389
16	M1X	Scarborough	Upper Rouge	43.836125	-79.205636

Create a map showing the neighborhoods in Scarborough

Test the Foursquare API to get venues

The venues details regarding categories and location is collected using the neighborhood data and Foursquare API as shown below:

	name	categories	lat	lng
0	African Rainforest Pavilion	Zoo Exhibit	43.817725	-79.183433
1	Toronto Pan Am Sports Centre	Athletics & Sports	43.790623	-79.193869
2	Toronto Zoo	Zoo	43.820582	-79.181551
3	Canadiana exhibit	Zoo Exhibit	43.817962	-79.193374
4	Polar Bear Exhibit	Zoo	43.823372	-79.185145
5	penguin exhibit	Zoo Exhibit	43.819435	-79.185959
6	Australasia Pavillion	Zoo Exhibit	43.822563	-79.183286
7	Morningside Park	Park	43.786546	-79.205322
8	Lamanna's Bakery, Cafe & Fine Foods	Bakery	43.797971	-79.148432
9	Americas Pavillon	Zoo Exhibit	43.822083	-79.185665
10	Lion Exhibit	Zoo Exhibit	43.819228	-79.186977
11	Orangutan Exhibit	Zoo Exhibit	43.818413	-79.182548

Check how many venues are there for each neighborhood

Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
Agincourt	4	4	4	4	4	4
Agincourt North, L'Amoreaux East, Milliken, Steeles East	3	3	3	3	3	3
Birch Cliff, Cliffside West	4	4	4	4	4	4
Cedarbrae	8	8	8	8	8	8
Clairlea, Golden Mile, Oakridge	10	10	10	10	10	10
Clarks Corners, Sullivan, Tam O'Shanter	10	10	10	10	10	10
Cliffcrest, Cliffside, Scarborough Village West	3	3	3	3	3	3
Dorset Park, Scarborough Town Centre, Wexford Heights	6	6	6	6	6	6
East Birchmount Park, Ionview, Kennedy Park	4	4	4	4	4	4
Guildwood, Morningside, West Hill	9	9	9	9	9	9
Highland Creek, Rouge Hill, Port Union	2	2	2	2	2	2
L'Amoreaux West	12	12	12	12	12	12
Maryvale, Wexford	7	7	7	7	7	7
Rouge, Malvern	2	2	2	2	2	2
Scarborough Village	1	1	1	1	1	1
Woburn	4	4	4	4	4	4

Apply One hot encoding and group rows on neighborhood by taking mean

	Neighborhood	American Restaurant	Asian Restaurant	Athletics & Sports	Auto Garage	Bakery	Bank	Bar	Breakfast Spot	Bus Line	Bus Station	Café	Caribbean Restaurant	Chinese Restaurant	Coffee Shop	College Stadium	Construction & Landscaping	Convenience Store	Department Store	Discount Store	Electronics Store
0	Rouge, Malvern	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	Rouge, Malvern	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	Highland Creek, Rouge Hill, Port Union	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
3	Highland Creek, Rouge Hill, Port Union	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
4	Guildwood, Morningside, West Hill	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0

	Neighborhood	American Restaurant	Asian Restaurant	Athletics & Sports	Auto Garage	Bakery	Bank	Bar	Breakfast Spot	Bus Line	Bus Station	Café	Caribbean Restaurant	Chinese Restaurant	Coffee Shop	College Stadium	Construction & Landscaping	Convenience Store	Department Store	Discount Store	Electro...
0	Agincourt	0.000000	0.000000	0.000	0.000000	0.000000	0.000	0.0	0.250000	0.0	0.0	0.00	0.000	0.250000	0.000000	0.00	0.0	0.00	0.00	0.00	0.00
1	Agincourt North, L'Amoreaux East, Milliken, St...	0.000000	0.333333	0.000	0.000000	0.000000	0.000	0.0	0.000000	0.0	0.0	0.00	0.000	0.000000	0.000000	0.00	0.0	0.00	0.00	0.00	0.00
2	Birch Cliff, Cliffside West	0.000000	0.000000	0.000	0.000000	0.000000	0.000	0.0	0.000000	0.0	0.0	0.25	0.000	0.000000	0.000000	0.25	0.0	0.00	0.00	0.00	0.00
3	Cedarbrae	0.000000	0.000000	0.125	0.000000	0.125000	0.125	0.0	0.000000	0.0	0.0	0.00	0.125	0.000000	0.000000	0.00	0.0	0.00	0.00	0.00	0.00
4	Clairlea, Golden Mile, Oakridge	0.000000	0.000000	0.000	0.000000	0.200000	0.000	0.0	0.000000	0.2	0.1	0.00	0.000	0.000000	0.000000	0.00	0.0	0.00	0.00	0.00	0.00
5	Clarks Corners, Sullivan, Tam O'Shanter	0.000000	0.000000	0.000	0.000000	0.000000	0.100	0.0	0.000000	0.0	0.0	0.00	0.000	0.100000	0.000000	0.00	0.0	0.00	0.00	0.00	0.00
6	Cliffcrest, Cliffside, Scarborough Village West	0.333333	0.000000	0.000	0.000000	0.000000	0.000	0.0	0.000000	0.0	0.0	0.00	0.000	0.000000	0.000000	0.00	0.0	0.00	0.00	0.00	0.00
7	Dorset Park, Scarborough Town Centre, Wexford ...	0.000000	0.000000	0.000	0.000000	0.000000	0.000	0.0	0.000000	0.0	0.0	0.00	0.000	0.166667	0.000000	0.00	0.0	0.00	0.00	0.00	0.00

4.0 Results:

Below table shows how the data looks for all the neighborhoods

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Agincourt	Lounge	Sandwich Place	Breakfast Spot	Chinese Restaurant	Vietnamese Restaurant	Coffee Shop	General Entertainment	Fried Chicken Joint	Fast Food Restaurant	Electronics Store
1	Agincourt North, L'Amoreaux East, Milliken, St...	Park	Asian Restaurant	Playground	Vietnamese Restaurant	General Entertainment	Fried Chicken Joint	Fast Food Restaurant	Electronics Store	Discount Store	Department Store
2	Birch Cliff, Cliffside West	General Entertainment	Skating Rink	Café	College Stadium	Vietnamese Restaurant	Coffee Shop	Grocery Store	Fried Chicken Joint	Fast Food Restaurant	Electronics Store
3	Cedarbrae	Caribbean Restaurant	Thai Restaurant	Athletics & Sports	Fried Chicken Joint	Bakery	Bank	Lounge	Hakka Restaurant	College Stadium	General Entertainment
4	Clairlea, Golden Mile, Oakridge	Bakery	Bus Line	Intersection	Fast Food Restaurant	Soccer Field	Bus Station	Metro Station	Park	College Stadium	General Entertainment
5	Clarks Corners, Sullivan, Tam O'Shanter	Pizza Place	Noodle House	Thai Restaurant	Fried Chicken Joint	Fast Food Restaurant	Italian Restaurant	Bank	Chinese Restaurant	Pharmacy	Coffee Shop
6	Cliffcrest, Cliffside, Scarborough Village West	Motel	American Restaurant	Coffee Shop	Grocery Store	General Entertainment	Fried Chicken Joint	Fast Food Restaurant	Electronics Store	Discount Store	Department Store
7	Dorset Park, Scarborough Town Centre, Wexford ...	Indian Restaurant	Chinese Restaurant	Latin American Restaurant	Vietnamese Restaurant	Pet Store	Bakery	Grocery Store	General Entertainment	Fried Chicken Joint	Asian Restaurant
8	East Birchmount Park, Ionview, Kennedy Park	Discount Store	Department Store	Playground	Coffee Shop	Vietnamese Restaurant	Chinese Restaurant	Grocery Store	General Entertainment	Fried Chicken Joint	Fast Food Restaurant
9	Guildwood, Morningside, West Hill	Intersection	Breakfast Spot	Mexican Restaurant	Tech Startup	Spa	Electronics Store	Medical Center	Pizza Place	Rental Car Location	College Stadium
10	Highland Creek, Rouge Hill, Port Union	Bar	Construction & Landscaping	Vietnamese Restaurant	Coffee Shop	Grocery Store	General Entertainment	Fried Chicken Joint	Fast Food Restaurant	Electronics Store	Discount Store
11	L'Amoreaux West	Chinese Restaurant	Fast Food Restaurant	Coffee Shop	Grocery Store	Pharmacy	Pizza Place	Breakfast Spot	American Restaurant	Thrift / Vintage Store	Sandwich Place
12	Maryvale, Wexford	Middle Eastern Restaurant	Auto Garage	Bakery	Shopping Mall	Sandwich Place	Breakfast Spot	Vietnamese Restaurant	College Stadium	General Entertainment	Fried Chicken Joint
13	Rouge, Malvern	Fast Food Restaurant	Print Shop	Vietnamese Restaurant	Chinese Restaurant	Grocery Store	General Entertainment	Fried Chicken Joint	Electronics Store	Discount Store	Department Store

Now that I have the required top 10 data, my client has full access to the information regarding the venues and what kind of business is profitable in that area.

Here are some of the output Data on the frequency of venues in a particular area

```
----Agincourt----
            venue  freq
0          Lounge  0.25
1  Breakfast Spot  0.25
2   Sandwich Place  0.25
3  Chinese Restaurant  0.25
4  American Restaurant  0.00
```

```
----Agincourt North, L'Amoreaux East, Milliken, Steeles East----
              venue  freq
0   Asian Restaurant  0.33
1               Park  0.33
2         Playground  0.33
3           Pharmacy  0.00
4  Korean Restaurant  0.00
```

```
    ----Birch Cliff, Cliffside West----          ----Cedarbrae----
                      venue  freq                               venue  freq
0            College Stadium  0.25          0  Caribbean Restaurant  0.12
1      General Entertainment  0.25          1     Athletics & Sports  0.12
2              Skating Rink  0.25          2                 Bakery  0.12
3                      Café  0.25          3                   Bank  0.12
4                Pizza Place  0.00          4        Thai Restaurant  0.12
```

```
  ----Clairlea, Golden Mile, Oakridge----    ----Clarks Corners, Sullivan, Tam O'Shanter----
            venue  freq                               venue  freq
0          Bakery   0.2                      0          Pizza Place   0.2
1        Bus Line   0.2                      1            Pharmacy   0.1
2    Intersection   0.1                      2  Fried Chicken Joint   0.1
3   Metro Station   0.1                      3   Italian Restaurant   0.1
4    Soccer Field   0.1                      4      Thai Restaurant   0.1
```

```
 ----Cliffcrest, Cliffside, Scarborough Village West----
                          venue  freq
0                         Motel  0.67
1            American Restaurant  0.33
2            Italian Restaurant  0.00
3  Latin American Restaurant  0.00
4                        Lounge  0.00
```

5.0 Discussions & Observation:

While conducting the Data Analysis of this information, I observed that the frequency of the Motel in the Cliffside neighborhood is the highest amongst all the other venues and that this business is in highly demand compared to other businesses.

6.0 Conclusion:

Hence, I conclude my Data Analysis by providing all the information that is required by my client and advise them to build a Motel in the Cliffside Neighborhood as it has highest frequency and demand. I have also provided them with other business possibilities such venues and location, which can earn them profit in future.