The Battle of Neighborhoods

JUNE 16

Authored by: Saurabh Kumar

Table of Contents

Introduction & Business Problem :

Problem Background:

The City of Toronto is the most populous city in the Canada. It is diverse and is the financial capital of Canada. It is multicultural. It provides lot of business opportunities and business friendly environment. It has attracted many different players into the market. It is a global hub of business and commerce. The city is a major center for banking and finance, retailing, world trade, transportation, tourism, real estate, new media, traditional media, advertising, legal services, accountancy, insurance, theater, fashion, and the arts in the Canada.
This also means that the market is highly competitive. As it is highly developed city so cost of doing business is also one of the highest. Thus, any new business venture or expansion needs to be analyzed carefully. The insights derived from analysis will give good understanding of the business environment which help in strategically targeting the market. This will help in reduction of risk. And the Return on Investment will be reasonable.

Problem Description:

A restaurant is a business which prepares and serves food and drink to customers in return for money, either paid before the meal, after the meal, or with an open account. The City of Toronto is famous for its excellent cuisine. It's food culture includes an array of international cuisines influenced by the city's immigrant history.
Central and Eastern European immigrants, especially Jewish immigrants - bagels, cheesecake, hot dogs, knishes, and delicatessens Jewish immigrants and Irish immigrants - pastrami and corned beef Chinese and other Asian restaurants, sandwich joints, trattorias, diners, and coffeehouses are ubiquitous throughout the city mobile food vendors.
So it is evident that to survive in such competitive market it is very important to strategically plan. Various factors need to be studied in order to decide on the Location such as :

1. **Population**
2. **Demographics**
3. **Competitor**
4. **Crime Rate profile**

Analyze the dataset to recommend a borough to standup Indian Restaurant.

Target Audience:

To recommend the correct location, ABC Company Ltd has appointed me to lead of the Data Science team. The objective is to locate and recommend to the management which neighborhood of Toronto

city will be best choice to start a restaurant. The Management also expects to understand the rationale of the recommendations made.

This would interest anyone who wants to start a new Indian restaurant in Toronto.

Success Criteria:

The success criteria of the project will be a good recommendation of borough/Neighborhood to stand up an Indian Restaurant based on lack of such restaurants in that location and low in crime rate

Data

To create a recommendation to open an Indian Restaurant in the Toronto city neighborhood, below sets of data which are required.

Dataset 1 : - Toronto City neighborhood and Borough Postal Code

In order to segment the neighborhoods and explore them, we will essentially need a dataset that contains the boroughs and the neighborhoods that exist in each borough as well as the latitude and longitude coordinates of each neighborhood.

This dataset exists for free on the web. Link to the dataset is :
https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

Dataset 2 : Demographic and population

Demographic and population analysis of the Toronto city to find out what is the population structure and where are the South Asian population lives in.
The dataset exists at below location

https://en.wikipedia.org/wiki/Demographics_of_Toronto http://www12.statcan.ca/census-recensement/2016/dp-pd/hlt-fst/imm/Table.cfm?Lang=E&T=22&Geo=535

Dataset 3 : Toronto City Crime Profile

Analysis of the crime profile of the Toronto City. This can be found at
http://data.torontopolice.on.ca/datasets/neighbourhood-crime-rates-boundary-file-/data?orderBy=Neighbourhood

Dataset 4 : Venue Information

Toronto city geographical coordinates data will be utilized as input for the Foursquare API, that will be leveraged to provision venues information for each neighborhood. We will use the Foursquare API to explore neighborhoods in Toronto City.

Methodology

To explore the options of opening Indian Restaurant in Toronto City, the socio-economic data profile of the Toronto City needs to be analyzed. After that depending on the analysis, recommendation needs to be put forward to the ABC company.

Methods of Data Collection

Methods which has been opted for the analysis is by using the data available

- by Canadian government in the public forum
- by the World organization available over the internet.
- Wikipedia
- FourSquare API

Methods of Analysis

Toronto City is having 6 borough and 140 neighborhood having people belonging to multiple ethnicity. Analytical approach along with exploratory analysis has been used to analyze these borough and neighborhood of the Toronto City.

Dataset 1 :- Toronto City borough and neighborhood

To analyze the Toronto City, below are the steps which have been carried out

- all the postal code of the city has been retrieved by reading the **https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M**

- Transform the data and store the dataset in Python Pandas Data Frame which has the field namely postal code, borough and neighborhood

- Cleaning up the data set in the Data Frame like where borough is no assigned then clean up the data.

- Get the latitude and longitude of each of the postal code in the Data Frame by calling “geopy” and “pgeocode” API’s

	PostalCode	Borough	Neighborhood	latitude	longitude
0	M3A	North York	Parkwood	43.7545	-79.3300
1	M4A	North York	Victoria Villag	43.7276	-79.3148
2	M5A	Downtown Toronto	Regent Park, Harbourfron	43.6555	-79.3626
3	M6A	North York	Lawrence Manor, Lawrence Height	43.7223	-79.4504
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Governmen	43.6641	-79.3889

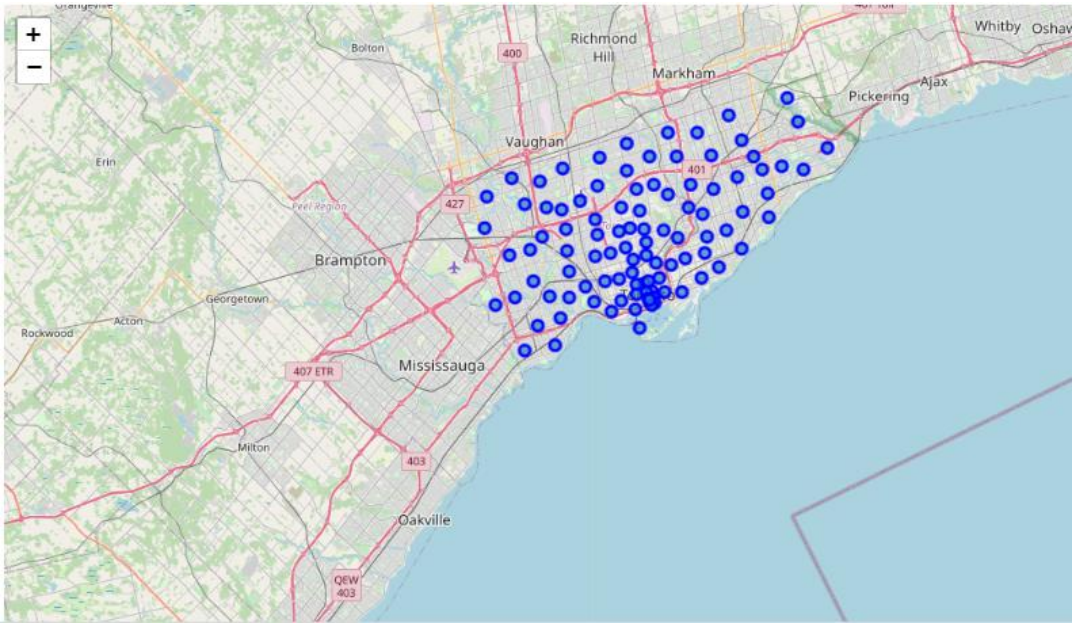
- Get the venues related information by calling Venues explore of the FourSquare which is https://api.foursquare.com/v2/venues/explore

	PostalCode	Borough	Neighborhood	latitude	longitude	Accessories_Store	Afghan_Restaurant	Airport	American_Restaurant	Art_Gallery	...	Vegetarian_/_
0	M3A	North York	Parkwood	43.7545	-79.3300	NaN	NaN	NaN	NaN	NaN	...	
1	M4A	North York	Victoria Villag	43.7276	-79.3148	NaN	NaN	NaN	NaN	NaN	...	
2	M5A	Downtown Toronto	Regent Park, Harbourfron	43.6555	-79.3626	NaN	NaN	NaN	NaN	NaN	...	
3	M6A	North York	Lawrence Manor, Lawrence Height	43.7223	-79.4504	NaN	NaN	NaN	1.0	NaN	...	
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Governmen	43.6641	-79.3889	NaN	NaN	NaN	NaN	NaN	...	

- Get the number of Indian Restaurant in each borough of Toronto City

Borough	Indian_Restaurant
East York	4.0
Downtown Toronto	3.0
Central Toronto	2.0
East Toronto	1.0
North York	1.0
West Toronto	1.0

- Using the Folium API, the neighborhood and postal code of the Toronto City.



With the above dataset, it will result in getting what are the venues like Restaurant, market, supermarket, hotel located in each neighborhood. One of the venues that we are looking as part of the analysis is “Restaurant” and its type.

Dataset 2 :- Demographic and Population

In the dataset 1 analysis, though the profile of the neighborhood in terms what type of venues are present have been analyzed. However, to open an Indian Restaurant, it is recommended and required to analyze the demographic and population of the neighborhood.

So, in order analyze the demographic and population of the Toronto City, data available on the wiki page and Canada Census data for metropolitan city have been analyzed.

Below are the steps taken to analyze the dataset.

- Read the tables present in the wiki page using pandas read_html function and get the dataset into data frame.
- Get the Visible Minorities as % of Population (2016 Census) & Top Ethnic Origins for each neighborhood and borough

	Riding (2013 Redistribution)	Percentage	Most Populous Ethnic Origin	Percentage.1	Top Riding for Ethnic Origin (Percentage)	Most Populous Visible Minority	Percentage.2	Top Riding for Visible Minority (Percentage)
0	Parkdale-High Park	26.2	English	22.3	Irish (20.0); German (9.8); French (8.9)	Black	5.3	NaN
1	Etobicoke Centre	27.1	Italian	15.1	Italian (15.1); Ukrainian (8.1)	South Asian	5.9	NaN
2	Etobicoke-Lakeshore	27.7	English	17.1	NaN	South Asian	5.5	NaN
3	Toronto-St. Paul's	28.3	English	18.5	NaN	Black	5.1	NaN
4	Eglinton-Lawrence	31.8	Canadian	14.7	Polish (12.0)	Filipino	10.7	NaN

- Get the age group diversification of the Toronto City by the age group

	Age Groups	Total	Male	Female
0	0 to 14 years	398135	204190	193945
1	0 to 4 years	136000	69895	66105
2	5 to 9 years	135025	69345	65680
3	10 to 14 years	127105	64940	62165
4	15 to 64 years	1906495	925070	981420
5	15 to 19 years	145525	74240	71280
6	20 to 24 years	194750	97415	97330
7	25 to 29 years	232945	113905	119035
8	30 to 34 years	224580	108895	115680
9	35 to 39 years	196310	94065	102240
10	40 to 44 years	182390	86535	95860
11	45 to 49 years	190925	90860	100065
12	50 to 54 years	202405	98735	103670
13	55 to 59 years	182805	88145	94655
14	60 to 64 years	153865	72265	81600
15	65 to 85 years	426945	184325	242615
16	65 to 69 years	130540	60360	70185

- Get the highest concentration of each ethnic group in Toronto City . This results in establishing that Etobicoke borough has the highest concentration of Indian.

	Top 20 Ethnic origins in the City of Toronto	Population in 2016	Percentage	Riding with Highest Concentration	Percentage
0	Chinese	332830	12.5	Scarborough—Agincourt	47.0
1	English	331890	12.3	Beaches—East York	24.2
2	Canadian	323175	12.0	Beaches-East York	19.7
3	Irish	262965	9.8	Parkdale—High Park	20.0
4	Scottish	256250	9.5	Beaches—East York	18.9
5	East Indian	202675	7.5	Etobicoke North	22.2
6	Italian	182495	6.8	Etobicoke Centre	15.1
7	Filipino	162605	6.0	York Centre	17.0
8	German	130900	4.9	Parkdale—High Park	9.8
9	French	122615	4.6	Parkdale—High Park	8.9
10	Polish	114530	4.3	Eglinton—Lawrence	12.0
11	Portuguese	100420	3.7	Davenport	22.7
12	Jamaican	90065	3.3	Humber River—Black Creek	8.5
13	Russian	74465	2.8	York Centre	9.5
14	Ukrainian	72340	2.7	Etobicoke Centre	8.1
15	Sri Lankan	58180	2.2	Scarborough—Rouge Park	11.1

- Assumption that we are making now is it is better and profitable to open Indian Restaurant in the borough where Indian or South Asian population is higher. So, lets find out which borough has Indian population higher. Below is the snapshot which puts the percentage of each ethnicity by borough.

Out[28]:

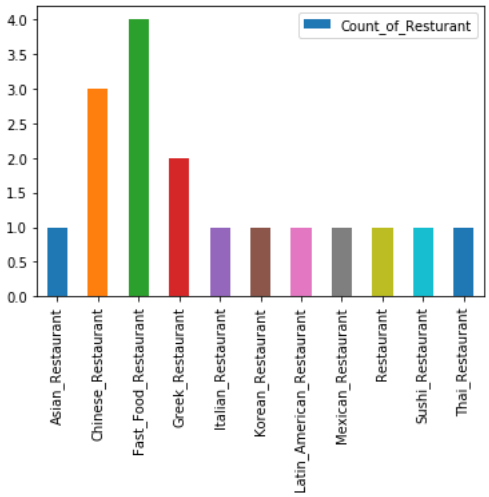
	Borough	Ethnic Group Position 1	Ethnic Group Position 2	Ethnic Group Position 3	Ethnic Group Position 4	Ethnic Group Position 5	Ethnic Group Position 6	Group 1 Percentage	Group 2 Percentage	Group 3 Percentage	Group 4 Percentage	Group 5 Percentage	Group 6 Percentage	Popu
0	TORONTO & EAST YORK	White	Chinese	South Asian	Black	NaN	NaN	56.3	14.8	8.3	5.1	NaN	NaN	1
1	TORONTO & EAST YORK	White	South Asian	Black	Chinese	NaN	NaN	64.5	10.9	6.6	5.7	NaN	NaN	1
2	TORONTO & EAST YORK	White	Black	Chinese	Latin American	NaN	NaN	66.9	6.4	5.9	5.4	NaN	NaN	1
3	TORONTO & EAST YORK	White	Black	NaN	NaN	NaN	NaN	72.4	5.3	NaN	NaN	NaN	NaN	1
4	TORONTO & EAST YORK	White	Chinese	South Asian	Black	NaN	NaN	65.5	12.3	5.4	5.0	NaN	NaN	1
5	TORONTO & EAST YORK	White	Black	NaN	NaN	NaN	NaN	70.8	5.1	NaN	NaN	NaN	NaN	1
6	TORONTO & EAST YORK	White	Chinese	NaN	NaN	NaN	NaN	66.5	14.0	NaN	NaN	NaN	NaN	1
7	TORONTO & EAST YORK	White	South Asian	Chinese	Black	NaN	NaN	48.8	11.8	11.1	9.1	NaN	NaN	
8	NORTH YORK	White	Chinese	West Asian	Korean	South Asian	Filipino	33.1	25.3	10.9	10.3	5.9	5.4	1
9	NORTH YORK	White	Filipino	Black	NaN	NaN	NaN	67.7	10.7	5.5	NaN	NaN	NaN	1
10	NORTH YORK	Chinese	White	South Asian	West Asian	NaN	NaN	31.3	29.4	10.2	7.6	NaN	NaN	1
11	NORTH YORK	White	Black	Latin American	Southeast Asian	Filipino	NaN	25.4	22.8	9.5	8.9	5.5	NaN	1
12	NORTH YORK	White	Filipino	Black	Latin American	NaN	NaN	53.1	16.5	7.9	5.1	NaN	NaN	1
13	NORTH YORK	White	South Asian	Chinese	NaN	NaN	NaN	57.9	13.3	10.6	NaN	NaN	NaN	1
14	NORTH YORK	White	South Asian	Black	Chinese	Filipino	West Asian	40.9	17.1	9.3	7.5	7.4	5.5	
15	SCARBOROUGH	White	South Asian	Filipino	Black	Chinese	NaN	29.4	25.6	12.5	9.6	9.3	NaN	1
16	SCARBOROUGH	White	South Asian	Black	Filipino	Chinese	NaN	42.0	21.6	11.2	9.0	5.8	NaN	1
17	SCARBOROUGH	Chinese	White	South Asian	Black	Filipino	NaN	45.8	19.1	14.0	6.3	5.4	NaN	1
18	SCARBOROUGH	South Asian	White	Black	Filipino	NaN	NaN	32.6	26.8	15.9	8.7	NaN	NaN	1
19	SCARBOROUGH	South Asian	White	Black	Filipino	Chinese	NaN	33.2	27.6	14.3	7.9	5.4	NaN	1
20	SCARBOROUGH	Chinese	South Asian	Black	White	Filipino	NaN	45.0	26.1	7.6	7.6	6.4	NaN	
21	ETOBICOKE & YORK	White	South Asian	Black	NaN	NaN	NaN	71.3	5.5	5.0	NaN	NaN	NaN	1
22	ETOBICOKE & YORK	South Asian	White	Black	NaN	NaN	NaN	28.9	23.8	23.4	NaN	NaN	NaN	1
23	ETOBICOKE & YORK	White	South Asian	Black	NaN	NaN	NaN	72.3	5.9	5.9	NaN	NaN	NaN	1
24	ETOBICOKE & YORK	White	Black	Latin American	Filipino	South Asian	NaN	44.2	23.2	8.5	5.9	5.7	NaN	1

This results in establishing that “ETOBICOKE” and “SCARBOROUGH” has a sizable population of South Asian in that area. So, it is better to open the Indian Restaurant there. However, let’s check factor like what is the crime rate in these areas and type of restaurant in these areas

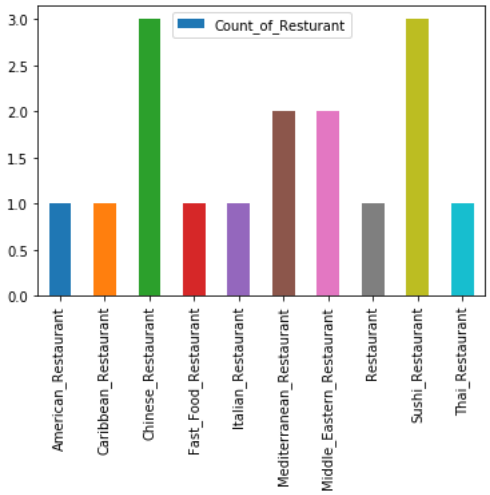
- Restaurant type in “ETOBICOKE” and “SCARBOROUGH”

  Below bar graph shows type of restaurant.

  - “SCARBOROUGH”



- ETOBICOKE

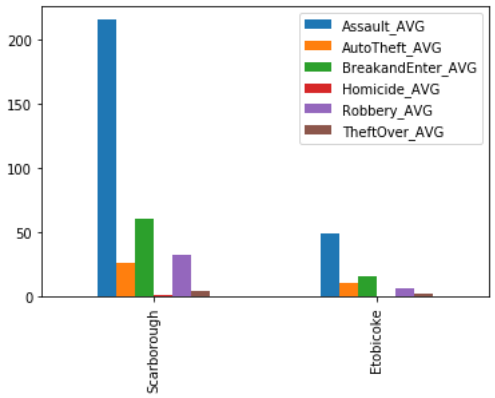


Dataset 3 :- Toronto City Crime Profile

With the dataset – 2, “ETOBICOKE” and “SCARBOROUGH” are the two borough where there is sizable population of the South Asian and therefore, it will be profitable to open the Indian Restaurant in these area. However, crime profile needs to be checked in these areas. To do that below are the steps carried out.

- Read the data in CSV format from **http://data.torontopolice.on.ca/datasets/neighbourhood-crime-rates-boundary-file-/data?orderBy=Neighbourhood** and load into Pandas DataFrame
- After processing the dataset, get the crime rate in the borough “ETOBICOKE” and “SCARBOROUGH”

	Neighbourhood	Assault_AVG	AutoTheft_AVG	BreakandEnter_AVG	Homicide_AVG	Robbery_AVG	TheftOver_AVG
0	Scarborough	215	26.7	60.5	1.6	32.5	4.7
1	Etobicoke	49.3	11.2	16	0.7	6.7	2.7



Results

1. The analysis resulted in finding the neighborhood having low competition and less crime to open an Indian Restaurant

Discussion

1. There will be competition from the Chinese Restaurant as there is presence of them more than other in the Etobicoke area.

Conclusion

Various factors need to be studied in order to decide on the Location . Below tables describes the factor that have been considered to reach out to the conclusion.

Borough	South Asian Population	Competitor	Crime Rate profile
Scarborough	Position 1	0	High
Etobicoke	Position 2	0	Low

Recommendation

Depending on the south Asian population and crime profile it is better to choose for Borough "Etobicoke" as compared to "Scarborough" to open the Indian Restaurant.