Best Place to start a new Firm

1. Introduction:

The problem statement is that if a person is willing to start a new business such as a restaurant in a locality, I should be able to show him the prospective location in which he can establish his business and gain more profits based on the location advantages.

To start a new restaurant there should be many factors that should be considered before establishing the restaurant. The main focus point is, one should look for the number of people in the locality and their spending habits. If a restaurant is started in a location in which there are less number of people and they do not spend much amount, then that restaurant will not give returns.

Also the new restaurant firm and its services should be available and in reach to the customers. So, if one establishes his business in far remote location where people cannot reach, it doesn't give visibility to the new restaurant and it results in the loss of the organization.

If there are similar kind of already established restaurants and if this new restaurant is established in that location, then there will be huge competition and there will not be much profits. So, whenever a person is starting a new business all the above mentioned points must be taken into consideration to gain more from the business.

So one should have the previously established firm's data and the neighbourhood data so that he can get the insight of the locality.

Now the question is where one can find that information? The possible answer is the usage of Foursquare location data services which provides ample information on how the data can be retrieved and how that can be used to get the insights from the data to solve the problem at hand.

Foursquare provides the information on the venues that are available in a certain locality by passing the necessary inputs to the API, and in return it provides the list of places that are in and around the location.

The target audience for this will be those whoever is interested in starting a new restaurant in Toronto and wants to get the overview of the location possibilities through which one can improve their business.

2. Data:

2.1 Data Acquisition:

I took the data required for this project by using the Wikipedia page https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M.

We have the dataset which is having four columns which contains the postal code, borough, neighbourhood, latitude and longitude of the neighbourhoods. We can get the venues of that neighbourhood by passing the latitude and longitude of the neighbourhood to the four square API.

We can take the count of the venues available in the neighbourhood and based on the count and the distance metrics we can predict the location in which one can make the better use of the location. We can also get the frequency of the venues available and then make a suggestion on which place is best suitable for establishing the new business.

We can also get the most common venue at a given location and to which venue most people are willing to go. We can then analyse the neighbourhood and suggest the place which is best suitable for establishing the new business. We can take the features like the distribution of the businesses, density of the locality, living habits, popular venues.

Fig 1: Data before cleaning:

	Postcode	Borough	Neighbourhood
0	M1A	Not assigned	Not assigned
1	M2A	Not assigned	Not assigned
2	M3A	North York	Parkwoods
3	M4A	North York	Victoria Village
4	M5A	Downtown Toronto	Harbourfront
5	M5A	Downtown Toronto	Regent Park
6	M6A	North York	Lawrence Heights
7	M6A	North York	Lawrence Manor
8	M7A	Queen's Park	Not assigned
9	M8A	Not assigned	Not assigned
10	M9A	Etobicoke	Islington Avenue

2.2 Data Cleaning:

After scraping the data from the Wikipedia page using the beautiful soup, I then cleaned the data for which Postal codes, the boroughs are not assigned. After that I combined the neighbourhoods of similar boroughs and postal codes. The number of these rows are less, thereby removing them will not have much effect on the overall result.

Fig 2: Data after removing the Not Assigned places

	Postcode	Borough	Neighbourhood
0	M3A	North York	Parkwoods
1	M4A	North York	Victoria Village
2	M5A	Downtown Toronto	Harbourfront
3	M5A	Downtown Toronto	Regent Park
4	M6A	North York	Lawrence Heights
5	M6A	North York	Lawrence Manor
6	M7A	Queen's Park	Not assigned
7	M9A	Etobicoke	Islington Avenue
8	M1B	Scarborough	Rouge
9	M1B	Scarborough	Malvern
10	M3B	North York	Don Mills North

Fig 3: Data after combining the neighbourhoods

	Postcode	Borough	Neighbourhood
0	M1B	Scarborough	Rouge,Malvern
1	M1C	Scarborough	Highland Creek,Rouge Hill,Port Union
2	M1E	Scarborough	Guildwood,Morningside,West Hill
3	M1G	Scarborough	Woburn
4	M1H	Scarborough	Cedarbrae
5	M1J	Scarborough	Scarborough Village
6	M1K	Scarborough	East Birchmount Park,Ionview,Kennedy Park
7	M1L	Scarborough	Clairlea,Golden Mile,Oakridge
8	M1M	Scarborough	Cliffcrest,Cliffside,Scarborough Village West
9	M1N	Scarborough	Birch Cliff,Cliffside West
10	M1P	Scarborough	Dorset Park,Scarborough Town Centre,Wexford He...

Fig 4: Data which shows the latitude and longitude of the places:

	Postcode	Latitude	Longitude
0	M1B	43.806686	-79.194353
1	M1C	43.784535	-79.160497
2	M1E	43.763573	-79.188711
3	M1G	43.770992	-79.216917
4	M1H	43.773136	-79.239476

Fig 5: Data after combining the places and their coordinates:

	Postcode	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

Fig 6: Map of Toronto showing the places.

3. Methodology:

The main focus of this project is to identify the suitable place in which one can start a new restaurant in the city of Toronto. The steps included in order to achieve our task are listed below.

Task 1: Retrieve the data of the Toronto city venues using the Foursquare API.

Task 2: Combine the venues of each neighbourhood using K Means clustering algorithm.

Task 3: Get the top most venues of each location.

Task 4: Take the least venue where there are less number of restaurants available, so that one can get more profits because of less competition.

Task 5: The obtained Results can then be used to identify the suitable location for staring a new restaurant.

Fig 7: Data which shows the popular venues at a given location.

	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	Adelaide,King,Richmond	Coffee Shop	Bar	Thai Restaurant	Steakhouse	Café	American Restaurant	Bakery	Hotel	Sushi Restaurant	Asian Restaurant
1	Agincourt	Lounge	Sandwich Place	Skating Rink	Breakfast Spot	Donut Shop	Diner	Discount Store	Dog Run	Doner Restaurant	Dumpling Restaurant
2	Agincourt North,L'Amoreaux East,Milliken,Steel...	Playground	Park	Women's Store	Donut Shop	Dessert Shop	Dim Sum Restaurant	Diner	Discount Store	Dog Run	Doner Restaurant
3	Albion Gardens,Beaumond Heights,Humbergate,Jam...	Grocery Store	Pharmacy	Coffee Shop	Beer Store	Fast Food Restaurant	Sandwich Place	Fried Chicken Joint	Liquor Store	Pizza Place	Dim Sum Restaurant
4	Alderwood,Long Branch	Pizza Place	Pharmacy	Coffee Shop	Gym	Skating Rink	Sandwich Place	Dance Studio	Pool	Pub	Diner

Fig 8: Data after clustering the places using K Means Clustering Algorithm.

	Postcode	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	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 Cor
0	M1B	Scarborough	Rouge,Malvern	43.806686	-79.194353	0.0	Fast Food Restaurant	Drugstore	Dim Sum Restaurant	Diner	Discount Store	Dog Run	Resta
1	M1C	Scarborough	Highland Creek,Rouge Hill,Port Union	43.784535	-79.160497	3.0	Bar	Dumpling Restaurant	Diner	Discount Store	Dog Run	Doner Restaurant	
2	M1E	Scarborough	Guildwood,Morningside,West Hill	43.763573	-79.188711	0.0	Medical Center	Electronics Store	Pizza Place	Spa	Mexican Restaurant	Rental Car Location	Bre
3	M1G	Scarborough	Woburn	43.770992	-79.216917	0.0	Coffee Shop	Korean Restaurant	Women's Store	Drugstore	Diner	Discount Store	Do
4	M1H	Scarborough	Cedarbrae	43.773136	-79.239476	0.0	Athletics & Sports	Bank	Bakery	Thai Restaurant	Caribbean Restaurant	Fried Chicken Joint	Resta

4. Results:

After doing all our analysis on exploring the Toronto city, we got the places where there are less number of restaurants. These places can be used by the investors who are planning to start a new business in the city of Toronto thereby increasing their profits.

I used the less common places where the restaurants are available, so that one can use these places for starting the new restaurants.

Fig 9: Places where there are less number of restaurants:

	Boroughs
0	Scarborough
1	North York
2	Central Toronto
3	Downtown Toronto
4	York
5	West Toronto
6	East Toronto
7	Etobicoke

5. Discussion:

So by looking at these results we can say that the places we got in the results section are best suited for our needs.

6. Conclusion:

The main purpose of this project is to find the places where one can start their new business such as a restaurant. Not only the restaurant, but we can also predict the places such as spa, parks, Coffee shops and many other places. These can be further used to explore the city to find the prospective place for starting a new business.