Capstone Project -Battle of Neighbourhood.

Recommender system for finding suitable location to start Pizza places in the City of Toronto.

Table of Contents:

1. Introduction: Business Problem

1. Background:

Toronto is the capital of the province of Ontario in southeastern Canada. It is the most populous city in Canada, a multicultural city, and the country's financial and commercial centre. Being the largest city in Canada with an estimated population of over 6 million, there is no doubt about the diversity of the population. Multiculturalism is seen through the various neighbourhoods including; Chinatown, Corso Italia, Little India, Kensington Market, Little Italy, Koreatown and many more.

Downtown Toronto being the hub of interactions between ethnicities brings many opportunities for entrepreneurs to start or grow their business. It is a place where people can try the best of each culture, either while they work or just passing through. Toronto is well known for its great food.

2. Problem:

The objective of this project is to help a prospective investor to find the best location for the opening of a Pizza restaurant/takeout for locals living in Toronto, Canada.

3. Interest:

Pizza is one of the most popular dishes in any city , kind of business easy to start up with a moderate investment, great chance to expand in futre, there are numerous opportunities to open a new Pizza place. Through this project, we will find the most suitable location for an entrepreneur to open a new Pizza restaurant/take out in Toronto, Canada**.**

2. Data acquisition and cleaning:

1. Data sources:

We require geographical location data for Toronto . Postal codes in each city serve as a starting point. Using Postal codes we can find out the neighborhoods, boroughs, venues and their most popular venue categories.

For the Toronto neighborhood data, a Wikipedia page exists that has all the information we need to explore and cluster the neighborhoods in Toronto.
Data Link: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
From week 3- Dataset consisting of latitude and longitude, zip codes.

Two of the neighbouring boroughs chosen for comparison are North York and Downtown Toronto.

-Foursquare API Data:

We will need data about different venues in different neighbourhoods of that specific borough. In order to gain that information we will use "Foursquare" locational information. Foursquare is a location data provider with information about all manner of venues and events within an area of interest. Such information includes venue names, locations, menus and even photos. As such, the foursquare location platform will be used as the sole data source since all the stated required information can be obtained through the API.

2. Data Cleaning:

Once the web scraping is done is done, data frame is created with features of interest, unwanted rows removed, combined rows contains similar data. Combined geographical information to the columns of data frame.

After finding the list of neighbourhoods, we then connect to the Foursquare API to gather information about venues inside each and every neighbourhood. For each neighbourhood, we have chosen the radius to be 500 meters.

3. Feature Selection:

The data retrieved from Foursquare contained information of venues within a specified distance of the longitude and latitude of the postcodes. The information obtained per venue as follows:

Neighbourhood : Name of the Neighbourhood
Neighbourhood Latitude : Latitude of the Neighbourhood
Neighbourhood Longitude : Longitude of the Neighbourhood
Venue : Name of the Venue
Venue Latitude : Latitude of Venue
Venue Longitude : Longitude of Venue
Venue Category : Category of Venue
Based on all the information collected for both boroughs in Tornto , we have sufficient data to build our model. We cluster the neighbourhoods together based on similar venue categories. We then present our observations and findings. Using this data, our stakeholders can take the necessary decision.

Postcode	Borough	Neighbourhood	Latitude	Longitude
M3A	North York	Parkwoods	43.7545	-79.3300
M4A	North York	Victoria Village	43.7276	-79.3148
M5A	Downtown Toronto	Harbourfront	43.6555	-79.3626
M6A	North York	Lawrence Heights, Lawrence Manor	43.7223	-79.4504
M7A	Downtown Toronto	Queen's Park	43.6641	-79.3889

The resulting data frame looks as above:

3. Modelling

After data wrangling is performed next step is to analyse the data.
First we analyse the borough North York for number of pizza resaurants and common venues in the neighbourhood, and the same analysis performed for Downtown Toronto. The data exploration, analysis and visualization for both boroughs are done in the same way but separately.

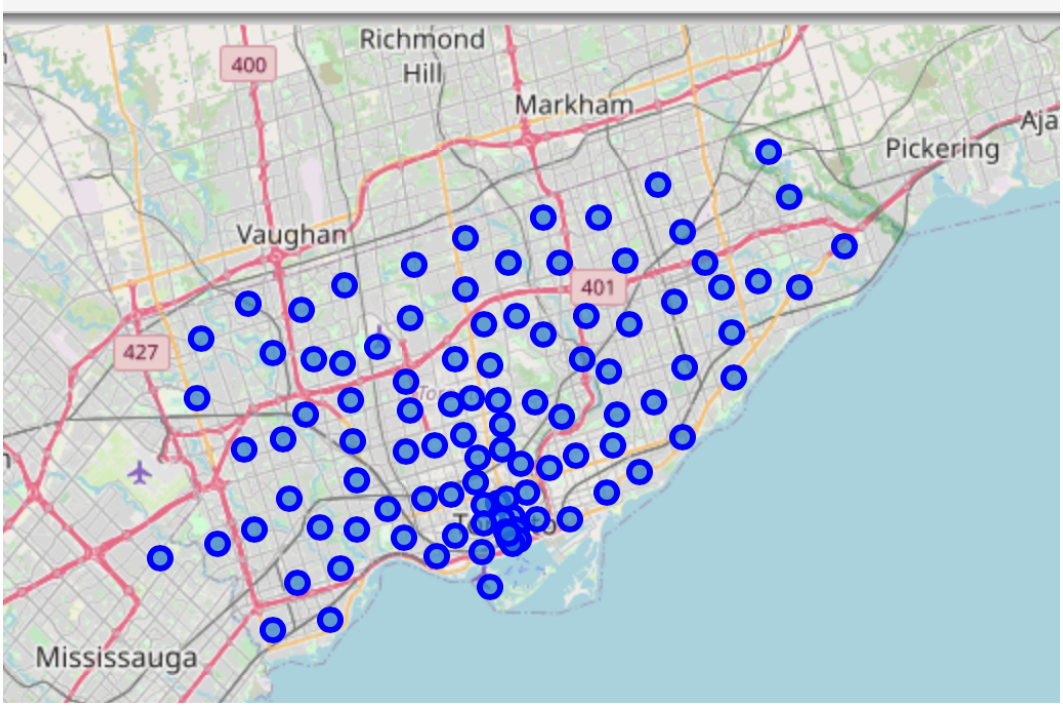
1. Business understanding

Main objective is to compare two boroughs North York and Downtown Toronto and find the optimum location for opening Pizza place business.

2. Exploratory Data Analysis

Using Foursquare API the number of Pizza places in North York and Downtown Toronto are explored and also the most common venues in the neighbourhood of each them analysed for the popularity by comparing the frequency of visit to each of them.

Now let's visualise all neighborhoods in a map using Folium and colour-coded each.



Toronto Neighbourhood

3. Problem Approach and K-means clustering

We analyse both boroughs neighbourhoods through one hot encoding (giving '1' if a venue category is there, and '0' in case of venue category is not there). On the basis of one hot encoding, we calculate mean of the frequency of occurrence of each category and picked top ten venues on that basis for each neighbourhood. It means the top venues are showing the foot traffic or the more visited places.

There is around 50 pizza places in Downtown Toronto and only 5 in North York,

```
0                              Pizza Pizza
1                              Amato Pizza
2                            Mamma's Pizza
3                           Domino's Pizza
4                               Pizza 2 Go
5                            Express Pizza
6                              Blaze Pizza
7                St. Lawrence Pizza and Pasta
8                     Pizzaiolo Gourmet Pizza
9                                241 Pizza
10                               JZs Pizza
11                            Pizza Shab
12                           Pizza Pizza
13                          Boston Pizza
14                        Colombo's Pizza
15                   Vinnie's Pizza & Pasta
16                           Pizza Nova
17                            Pizza Hut
18                           Mamma's Pizza
19                             Pizza Pizza
20                         Fantastico Pizza
21                             Pizza Studio
22                           Domino’s Pizza
23                            Pizza Rustica
24    Classic italian style pizza food truck
25                         Bocconcini Pizza
```

Pizza places in Downtown Toronto

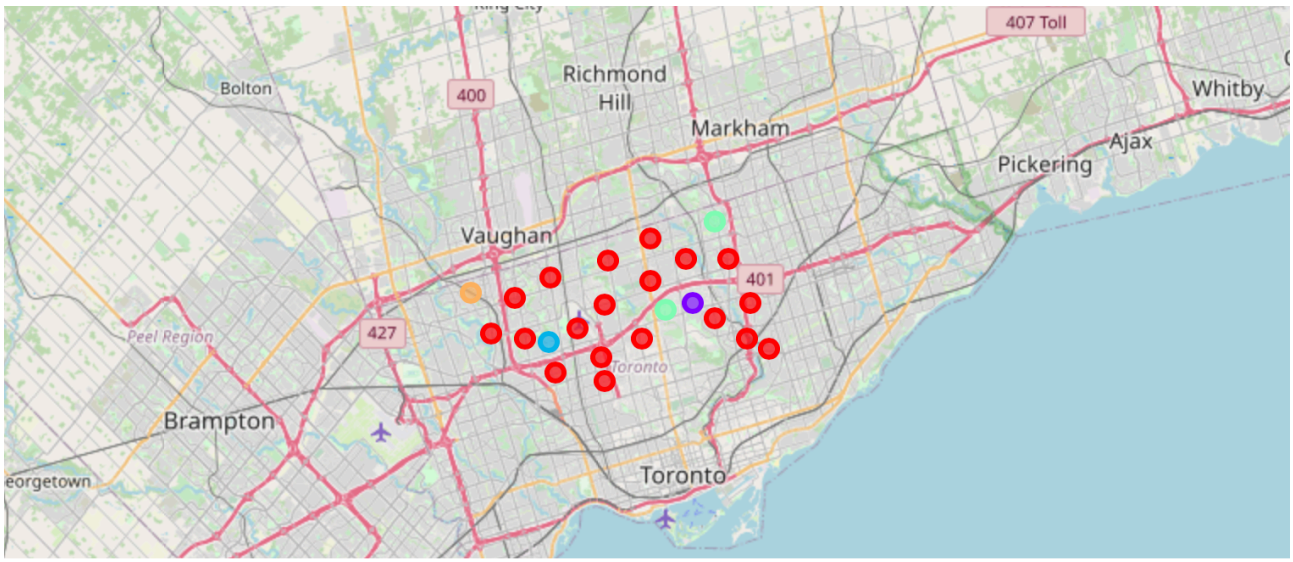
```
                                Domino's Pizza
                                 Pizza e Pazzi
            Double Double Pizza and Chicken
                                    Pizza Nova
                                   Pizza Pizza
ıme: name, dtype: object
```

Pizza places in North York

Above two results show that Downtown Toronto has many more number of Pizza places to compete than North York.

Now we need to check for the most common venues in each borough. For that we use k-means clustering

Cluster of Venues in North York



Cluster of Venues in Downtown Toronto

Clustering Venues in North York

Cluster 1

```
In [39]: NorthYork_merged.loc[NorthYork_merged['Cluster Labels'] == 0, NorthYork_merged.columns[[1] + list(range(5, NorthYork
```

Out[39]:

	Borough	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 Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
1	North York	0	Pizza Place	Intersection	Portuguese Restaurant	Coffee Shop	Hockey Arena	Park	Accessories Store	Middle Eastern Restaurant	Pharmacy	Pastry Shop
2	North York	0	Clothing Store	Coffee Shop	Shoe Store	Women's Store	Cosmetics Shop	Bakery	Restaurant	Food Court	Sushi Restaurant	Toy / Game Store
4	North York	0	Grocery Store	Gas Station	Latin American Restaurant	Pizza Place	Asian Restaurant	Ice Cream Shop	Fast Food Restaurant	Pharmacy	Pastry Shop	Park
7	North York	0	Pizza Place	Coffee Shop	Mediterranean Restaurant	Grocery Store	Fried Chicken Joint	Accessories Store	Middle Eastern Restaurant	Pharmacy	Pastry Shop	Park
8	North York	0	Clothing Store	Fast Food Restaurant	Coffee Shop	Baseball Field	Food Court	Lingerie Store	Electronics Store	Shoe Store	Japanese Restaurant	Bakery
9	North York	0	Pizza Place	Metro Station	Middle Eastern Restaurant	Sports Bar	Sandwich Place	Pharmacy	Pastry Shop	Park	Nightclub	Moving Target
11	North York	0	Airport	Coffee Shop	Food Court	Shoe Store	Accessories Store	Middle Eastern Restaurant	Pizza Place	Pharmacy	Pastry Shop	Park

Clustering Venues in Downtown Toronto:

	Borough	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 Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
1	Downtown Toronto	0	Sushi Restaurant	Burrito Place	Gym	Gym / Fitness Center	Escape Room	Ethiopian Restaurant	Coffee Shop	Restaurant	College Cafeteria	Ramen Restaurant
	click to unscroll output; double click to hide			Café	Italian Restaurant	Cocktail Bar	Restaurant	Clothing Store	Department Store	Beer Bar	American Restaurant	Japanese Restaurant
4	Downtown Toronto	0	Sandwich Place	Cocktail Bar	Coffee Shop	Hotel	Japanese Restaurant	Café	Beer Bar	Bakery	Bank	Sporting Goods Shop
12	Downtown Toronto	0	Café	Vegetarian / Vegan Restaurant	Gaming Cafe	Burger Joint	Art Gallery	Arts & Crafts Store	Mexican Restaurant	Coffee Shop	Bakery	Caribbean Restaurant
13	Downtown Toronto	0	Coffee Shop	Italian Restaurant	Gym / Fitness Center	Bar	Speakeasy	Sandwich Place	Restaurant	Pub	Pizza Place	Park
16	Downtown Toronto	0	Coffee Shop	Pizza Place	Italian Restaurant	Restaurant	Bakery	Café	Gym / Fitness Center	Snack Place	Pharmacy	Plaza
18	Downtown Toronto	0	Sushi Restaurant	Japanese Restaurant	Gay Bar	Restaurant	Dance Studio	Coffee Shop	Mediterranean Restaurant	Burrito Place	Indian Restaurant	Fast Food Restaurant

There are 1108 different venues in Downtown Toronto and and 287 venues in North York.

Clustering the venues in North York neighbourhood shows that there are more than 5 pizza places coming under first 5 most common venues.
Its seen that pizza places are one of the most common venue in North York neighbourhood, shows the popularity of Pizza places in the borough.

Cluster analysis of Downtown Toronto shows that there's only 1 pizza place coming under first 5 most common venues, which indicate that pizza places are less popular in the locality.

4. Discussion

North York is borough in Toronto with 289 venues and only 5 Pizza Places, with frequency of Pizza places occur in 1st, 3rd and 4th most common venues, and having good number of restaurants around for a healthy competition.
Downtown Toronto at the same time has 1108 venues with 50 pizza places with frequency of occurrence is hardly 2 in first five most common venue.

So with the help of k-means clustering its concluded that North York is the better choice for opening up a new pizza place as it is identified to be one of the most popular venue in the borough.

Project could further be used for analysing the opportunity open up any kind of business anywhere in the world, also could be extended to analyse based on other parameters depending on the investor.

5. Conclusion

The final decision on optimal restaurant location will be made by the future investor based on the recommendations given in this analysis and also on specific characteristics of neighbourhoods in every recommended area, taking into consideration additional factors like real estate availability, prices, etc.

6. References

https://en.wikipedia.org
www.google.com
developer.foursquare.com
IBM Skills Network Lab