CAPSTONE PROJECT

TOPIC: Opening a hotel in Toronto, Canada

By Jayanti Bhattacharya

Date – 5th January 2021

INTRODUCTION:

Toronto is one of Canada's leading tourism destinations. In 2017, the Toronto-area received 43.7 million tourists, of which 10.4 million were domestic visitors and 2.97 million were from the United States, spending a total of $8.84 billion. Toronto has an array of tourist attractions and a rich cultural life.

Hotels are definitely one of the fastest-growing sectors in the tourism sector and it is truly justified as accommodation is the key part in the development of any country or region's tourism. Tourism and Hotel Industry always go hand in hand and the presence of enough hotels also adds value and quite a lot of factors and punches it within the region's economy. The Existence of a Hotel isn't enough to single-handedly boost a region's tourism but they also give out a symptom of health tourism.

BUSINESS PROBLEM:

- *Location of a hotel is very important for determining the success or profit of a hotel*
- *The objective of this project is to find profitable and great locations for building hotels in Toronto*
- *The target audience are property developers who want to build a hotel in Toronto*
- *Business Problem:*

If a property developer is planning to build a hotel in Toronto where is a great location to build it and where are the recommended locations?

DATA:

- *Data Required:*
    1. *Lists of neighbourhoods in Toronto, Canada*
    2. *Coordinates of the neighbourhoods*
    3. *Venue data of the neighbourhoods*

- *Sources of Data:*
    1. *Wikipedia page for postal codes in Toronto*

    https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M

    2. *Excel spreadsheet of coordinates of the neighbourhoods*
    3. *Foursquare API for venue data*

Methodology:

- *First, we get the information about neighbourhoods by web scraping and the coordinates by python geocoder package and put into a dataframe.*
- *Use Foursquare API to get the venue data.*
- *Visualize the data using Folium.*

Exploring the Data

```
df
```

	Neighbourhood	Latitude	Longitude
0	Parkwoods	43.753259	-79.329656
1	Victoria Village	43.725882	-79.315572
2	Regent Park, Harbourfront	43.654260	-79.360636
3	Lawrence Manor, Lawrence Heights	43.718518	-79.464763
4	Queen's Park, Ontario Provincial Government	43.662301	-79.389494
5	Islington Avenue, Humber Valley Village	43.667856	-79.532242
6	Malvern, Rouge	43.806686	-79.194353
7	Don Mills	43.745906	-79.352188
8	Parkview Hill, Woodbine Gardens	43.706397	-79.309937

- *Group data by neighbourhood and taking mean of frequency of each venue.*
- *Filter the venue by category 'Hotel'*

Let's see number of hotels present

```
[15]: len(t_grouped[t_grouped["Hotel"] > 0])
```

```
[15]: 47
```

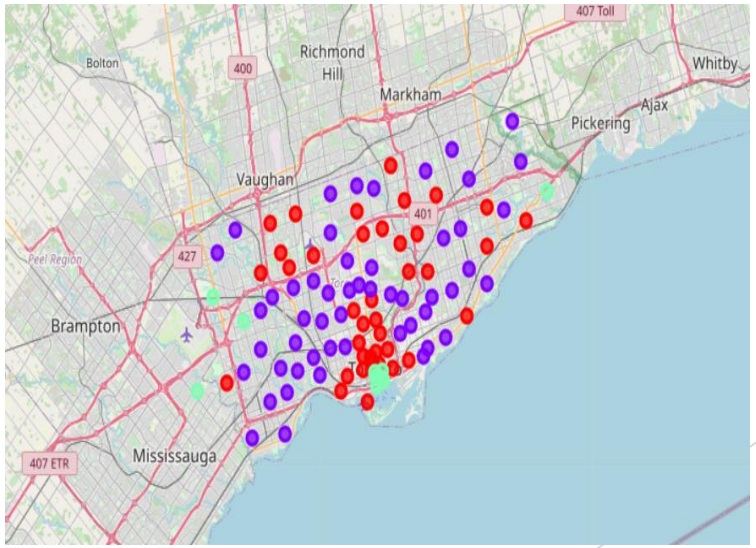
Let's see the number of hotels in all Neighbourhood

```
[16]: t_sh = t_grouped[["Neighbourhoods","Hotel"]]
      t_sh
```

[16]:

	Neighbourhoods	Hotel
0	Agincourt	0.000000
1	Alderwood, Long Branch	0.000000
2	Bathurst Manor, Wilson Heights, Downsview North	0.000000
3	Bayview Village	0.000000
4	Bedford Park, Lawrence Manor East	0.000000

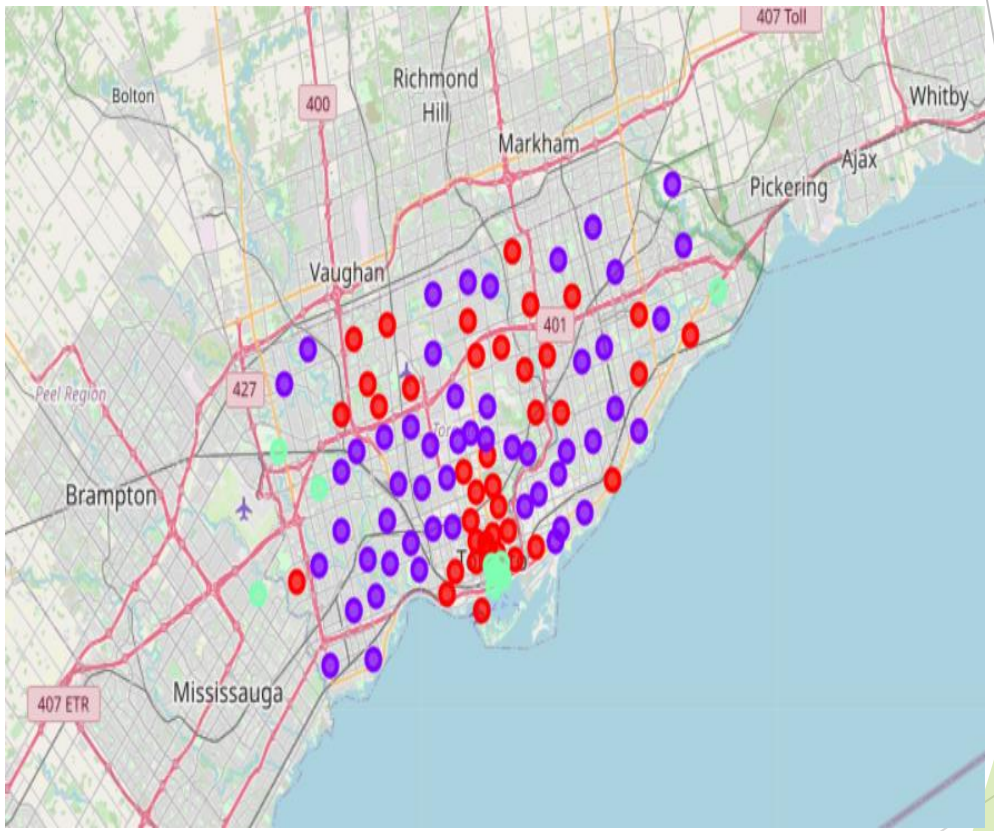
- *Perform clustering on the data by the technique of k-clustering.*
- *Finally, visualise the formed clusters with Folium.*



RESULTS:

The neighbourhoods of Toronto are Categorised into 3 parts:

- *Cluster 1 : Neighbourhoods with moderate number of hotels*
- *Cluster 2 : Neighbourhoods with very small number of hotels*
- *Cluster 3 : Neighbourhoods with large number of hotels*



DISCUSSION:

- *Most of the hotels are concentrated in the center part of the city which is cluster 3.*
- *Highest number of hotels are in cluster 3*
- *Lowest number of hotels are in cluster 2 and moderate number in cluster 1.*
- *There is high competition rate in Cluster 3, while there are less in Cluster 1 and minimum in Cluster 2.*

RECOMMENDATIONS:

- *It is better to avoid high competition places like Cluster 3 and it is preferrable to not build a hotel there.*
- *Cluster 1 or 2 can be a suitable option for building a hotel*
- *Cluster 1 with moderate competition which might be a profitable area if the place stands out from the others*
- *Cluster 2 is a safe position will almost no competition and a highly profitable area*

CONCLUSION:

- *Answer to the business problem:*

  The neighbourhoods in Cluster 2 are most preferred to open a hotel in followed by Cluster 1

- *Cluster 3 is overcrowded and has high competition and thus should not be preferred over the others to maximize profit*

- *The findings of this project might help project developers and future hotel owners to maximise their profit by selecting a better location*

THAT'S THE END OF THIS PRESENTATION