

Capstone Project - The Battle of Neighbourhoods

Final Report

Living in Toronto after quarantine days.

by Muharrem Sener Yuce

April / 2020

I	-	Instruction
II	-	Data
III	-	Methodology
IV	-	Results
V	-	Discussion
VI	-	Conclusion

I - Introduction

Toronto:

Toronto demographics show that it is a large and ethnically diverse metropolis.

With around half of its population born outside the country, Toronto is often referred to as 'the most multicultural city in the world.

Toronto boasts 200 ethnic groups with over 140 languages spoken.

2019 World Ranking for Toronto

(<https://www.toronto.ca/city-government/data-research-maps/toronto-progress-portal/world-rankings-for-toronto/>)

Toronto's Ranking	# of Cities	Report	Description
–	–	2019 National Geographic Traveler Magazine	Toronto is among the 28 must-see destinations.
15	1,000	StartupBlink Startup Ecosystem Rankings 2019	Toronto in the top 20 amongst 1,000 cities according to their startup ecosystem strength
12	35	CBRE Global Living 2019	Toronto ranks 12th of 35 in terms of average property prices
7	10	Newsweek The 10 Best Hospitals in the World	Toronto is home to one of the top 10 hospitals in the world
9	50	Making the Most of a Layover	Toronto's Pearson airport ranks as one of the world's best airports to enjoy a layover.
7	140	Economist Intelligence Unit – Global Liveability Index 2019	Toronto ranks as one of the most liveable cities in the world.
3	50	CBRE Research 2019 Scoring Tech Talent (North America Only)	CBRE Research's 2019 Scoring Tech Talent ranked Toronto 3rd of top 50 tech talent markets in North America.
6	60	2019 Safe Cities Index	Economist Intelligence Unit released its 2019 Safe Cities Index, ranked Toronto at 6th out of 60 cities.

Covid-19,

Pandemic

The complicated problem for World.

Problem

Covid-19 caused;

People were caught unprepared for this rapidly spreading problem.

Losses are increasing rapidly.

People have to live in quarantine.

This negatively affects the psychology of people in cities with social and economic life.

At the end of these difficult days, it is predicted that people will prefer to stay away from crowded places to maintain social distance.

Activity venues that give people the chance to maintain this social distance will be preferred.

One of the most suitable activities that provides the conditions suitable for this prediction is Yoga Studios.

Question (for Toronto local) :

In which area is it appropriate to open a Yoga Studio in Toronto?

Who would be interested in this project?:

by Governments, Municipalities, Entrepreneurs to all age citizens and customers.

II - Data

For this project we need the following data:

Toronto data that contains list Boroughs, Neighborhoods along with their latitude and longitude.

Data source	: https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
Description	: This data set contains the required information. And we will use this data set to explore various neighborhoods of Toronto

Yoga Studios in each neighborhood of Toronto.

Data source	: Foursquare API
Description	: By using this API we will get all the venues in each neighborhood. We can filter these venues to get only Yoga Studios.

GeoSpace data

Data source	: http://cocl.us/Geospatial_data
Description	: By using this geo space data we will get the Toronto Borough boundaries that will help us visualization.

III - Methodology

The list of neighborhoods in Toronto, Canada preparing by extracting the list of neighborhoods from

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

The web scraping realized by utilizing the pandas HTML table scraping method.

This is given to us only a list of neighborhood names and postal codes.

For their coordinates to utilize Foursquare to pull the list of venues near these neighborhoods.

To get the coordinates, used CSV file provided by IBM team to match the coordinates of Toronto neighborhoods.

http://cocl.us/Geospatial_data

Coordinates gathered.

Visualizing the map of Toronto used the Folium package to verify whether these are correct coordinates.

Foursquare API to pull the list of top 100 venues within 500 meters radius.

The data pulled the names, categories, latitude, and longitude of the venues.

Checked how many unique categories that we can get from these venues.

Analyzed each neighborhood by grouping the rows by neighborhood and taking the mean on the frequency of occurrence of each venue category.

Here, for made a justification to specifically look for “Yoga Studio”.

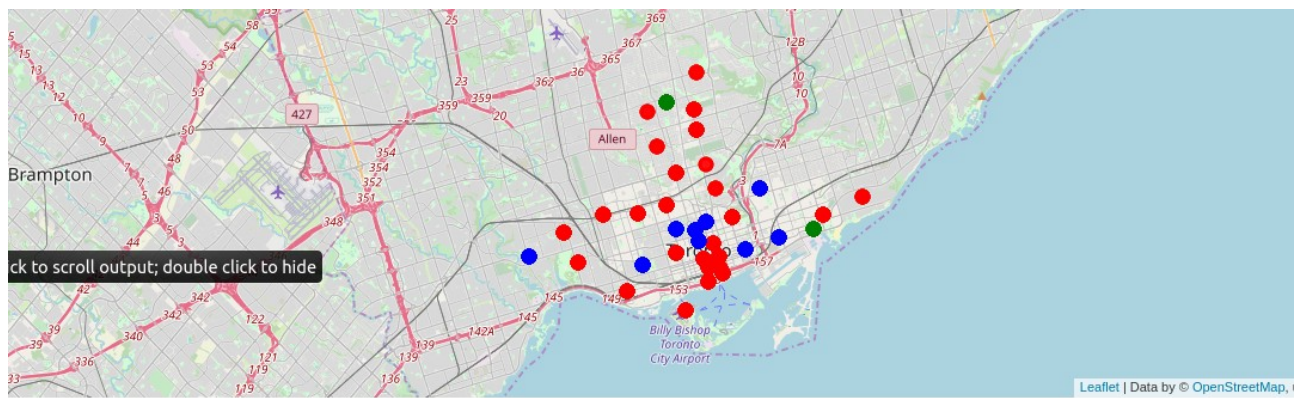
Performed the clustering method by using k-means clustering.

K-means clustering algorithm identifies k number of centroids, and then allocates every data point to the nearest cluster while keeping the centroids as small as possible.

It is one of the simplest and popular unsupervised machine learning algorithms and it is highly suited for this project as well.

Clustered the neighborhoods in Toronto into 3 clusters based on their frequency of occurrence for “Yoga Studio”.

Based on the results (the concentration of clusters), for will be able to recommend the ideal location to open the restaurant.



```
In [48]: #Cluster 0 (marking color red)
to_merged.loc[(to_merged['Cluster Labels'] == 0) & (to_merged['Venue Category'] == 'Yoga Studio') ]
```

```
Out[48]:
```

	Neighborhood	Yoga Studio	Cluster Labels	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
--	--------------	-------------	----------------	-----------------------	------------------------	-------	----------------	-----------------	----------------

```
In [49]: #Cluster 1 (marking color blue)
to_merged.loc[(to_merged['Cluster Labels'] == 1) & (to_merged['Venue Category'] == 'Yoga Studio') ]
```

```
Out[49]:
```

	Neighborhood	Yoga Studio	Cluster Labels	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
4	Central Bay Street	0.012987	1	43.657952	-79.387383	The Yoga Sanctuary	43.661499	-79.383636	Yoga Studio
32	Studio District	0.023810	1	43.659526	-79.340923	Spirit Loft Yoga	43.663548	-79.341333	Yoga Studio
38	University of Toronto / Harbord	0.028571	1	43.662096	-79.400049	Sivananda Yoga Centre	43.662754	-79.402951	Yoga Studio
36	The Danforth West / Riverdale	0.023810	1	43.679557	-79.352188	Moksha Yoga Danforth	43.677622	-79.352116	Yoga Studio
23	Queen's Park / Ontario Provincial Government	0.031250	1	43.662301	-79.389494	The Yoga Sanctuary	43.661499	-79.383636	Yoga Studio
24	Regent Park / Harbourfront	0.022222	1	43.654260	-79.360636	The Yoga Lounge	43.655515	-79.364955	Yoga Studio
6	Church and Wellesley	0.025316	1	43.665860	-79.383160	Bikram Yoga Yonge	43.668205	-79.385780	Yoga Studio
6	Church and Wellesley	0.025316	1	43.665860	-79.383160	The Yoga Sanctuary	43.661499	-79.383636	Yoga Studio
19	Little Portugal / Trinity	0.020000	1	43.647927	-79.419750	YogaSpace	43.647607	-79.420133	Yoga Studio
28	Runnymede / Swansea	0.025000	1	43.651571	-79.484450	(The New) Moksha Yoga Bloor West	43.648058	-79.485242	Yoga Studio

```
In [50]: #Cluster 2 (marking color green)
to_merged.loc[(to_merged['Cluster Labels'] == 2) & (to_merged['Venue Category'] == 'Yoga Studio') ]
```

```
Out[50]:
```

	Neighborhood	Yoga Studio	Cluster Labels	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
2	Business reply mail Processing Centre	0.062500	2	43.662744	-79.321558	Toronto Yoga Mamas	43.664824	-79.324335	Yoga Studio
21	North Toronto West	0.058824	2	43.715383	-79.405078	Barreworks	43.714070	-79.400109	Yoga Studio

IV - Result

This study was completed on April 4, 2020.

According to this date, there are yoga studios in the region defined as Cluster 1.

There is no Yoga studio at Cluster 0.

There are only 2 Yoga Studios in the region defined as Cluster 2.

For this reason, more detailed research can be done in the areas defined as Cluster 0 and Cluster 2 for the installation of Yoga studios.

V - Discussion

This study is a course study in which 20 hours of working time is envisaged.

Made only for the purpose of testing basic research methods and sources of information are limited.

In order to make a commercial investment decision on this subject, it is necessary to participate in the analysis in other variables such as rent, transportation convenience.

VI- Conclusion

With the Covid 19 outbreak, humanity is going through a huge challenge.

This Covid 19 pandemic; causing irreversible and irreversible losses.

At the end of this period, physical and mental difficulties await us in a world where everything will not be the same anymore.

For this reason, activities such as yoga will be continued and mental meditation activities will increase.

I think that this solution proposal should be taken into consideration by companies and governments as both business investment and social responsibility.

P.S. In the Coursera - Applied Data Science course, many educational work has been done about Toronto. For this reason, the city of Toronto was preferred in this study.

I wish healthy days.