

Determining Suitable Locations for Launching Yoga Studios in Toronto

1. Introduction

Toronto is the provincial capital of Ontario and the most populous city in Canada, with a population of 2,731,571 as of 2016. It is Canada's most populous CMA and the fourth most populous city in North America. The city is the anchor of the Golden Horseshoe, an urban agglomeration of 9,245,438 people (as of 2016) surrounding the western end of Lake Ontario. Toronto is an international centre of business, finance, arts, and culture, and is recognized as one of the most multicultural and cosmopolitan cities in the world. Fast becoming the most desired location for migrants and natives alike, the city has developed into a buzzing centre of social and economic development.

As the corporate culture spreads, with more companies springing up, and an increasing share of the population switching to desk jobs, the concern to develop and sustain levels of personal fitness is also very real. In a fast paced life, where stresses - both mental and physical - are on a steady rise, it is more important than ever to have ideal solutions to enable the citizens to cope up. One among the many well known solutions, and arguably, the most ideal solution that caters to the mental, physical and metaphysical wellness is the ancient Vedic Indian art of Yoga.

Yoga is a group of physical, mental, and spiritual practices or disciplines which originated in ancient India. Yoga is one of the six Āstika (orthodox) schools of Hindu philosophical traditions. The practice of yoga has been thought to date back to pre-vedic Indian traditions; possibly in the Indus valley civilization around 3000 BCE. Yoga is mentioned in the Rigveda, and also referenced in the Upanishads. Yoga as exercise is a physical activity consisting largely of asanas, often connected by flowing sequences called vinyasas, sometimes accompanied by the breathing exercises of pranayama, and usually ending with a period of relaxation or meditation. It is often known simply as yoga, despite the existence of multiple older traditions of yoga within Hinduism where asanas played little or no part, some dating back to the Yoga Sutras, and despite the fact that in no tradition was the practice of asanas central. In addition to boosting emotional health and relieving stresses, regular practice of Yoga helps you in your fight against cardiac problems, back aches, arthritis, asthma, PTSD etc, or even something as simple as getting a good night's sleep.

Stakeholder: A1 Yoga Studios, a leading chain of Fitness Centres, is planning an entry into the city of Toronto. After having established their bases across the United States of America, after humble beginnings in Kerala, India. The branded has entrusted us with the responsibility of conducting a market survey prior to their launch in the city.

Mission: Our mission is to identify suitable locations across Toronto to launch the Yoga Studio chain.

Aim: The aim of conducting the analytical survey is to determine top locations within Toronto where a fair share of the population is inclined towards fitness activities, but the top 5 most common fitness centres do not include a Yoga Studio. A strong inclination towards fitness, plus lack of a great Yoga Studio in the proximity, presents an ideal case for the establishment of a renowned brand offering top class Yoga sessions.

2. Data

Primarily, three data sets have been used to conduct the analysis, and arrive at a possible solution. They are as enumerated below:

- a. The Wikipedia webpage:

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

has been used to create a data frame that classifies Neighborhoods in Toronto with respect to their Boroughs and Postal Codes. This is primarily required to determine the locations and addresses of various neighborhoods.

- b. The Geospatial Data from

https://cocl.us/Geospatial_data

has been used to get the latitude and longitude co-ordinates of the neighbourhoods in Toronto based on their postal codes. A data frame was created, which was later merged with data frame mentioned in point 1 above to form comprehensive location data.

- c. Location Information from Foursquare was obtained with an API call:

<https://api.foursquare.com/v2/venues/>,

with venue category as '4bf58dd8d48988d175941735', which represents Gyms and Fitness Centres, and various sub categories.

3. Methodology

The methodology adopted in determining the most suitable locations for establishing the fitness centres for A1 Yoga Studios are described in details in the ensuing paragraphs. it consisted of the following steps:

- a. Downloading, Exploring and Cleaning Data:

The data downloaded from the Wikipedia Webpage consisted of information in the form of a table containing Postal Code, Boroughs and Neighborhoods. The rows were cleaned to remove any unassigned or missing values in both Borough and Neighbourhood columns. It was noticed that the same Postal Code was being repeated multiple times, due to different neighbourhoods in the vicinity if each other sharing the same Postal Code. To bring uniqueness to the Postal Code, the neighbourhoods sharing the same postal code were grouped into a single row, with commas separating the values. The values were stored in a Pandas Dataframe.

- b. Downloading Geospatial Data and Merging with the Original DF:

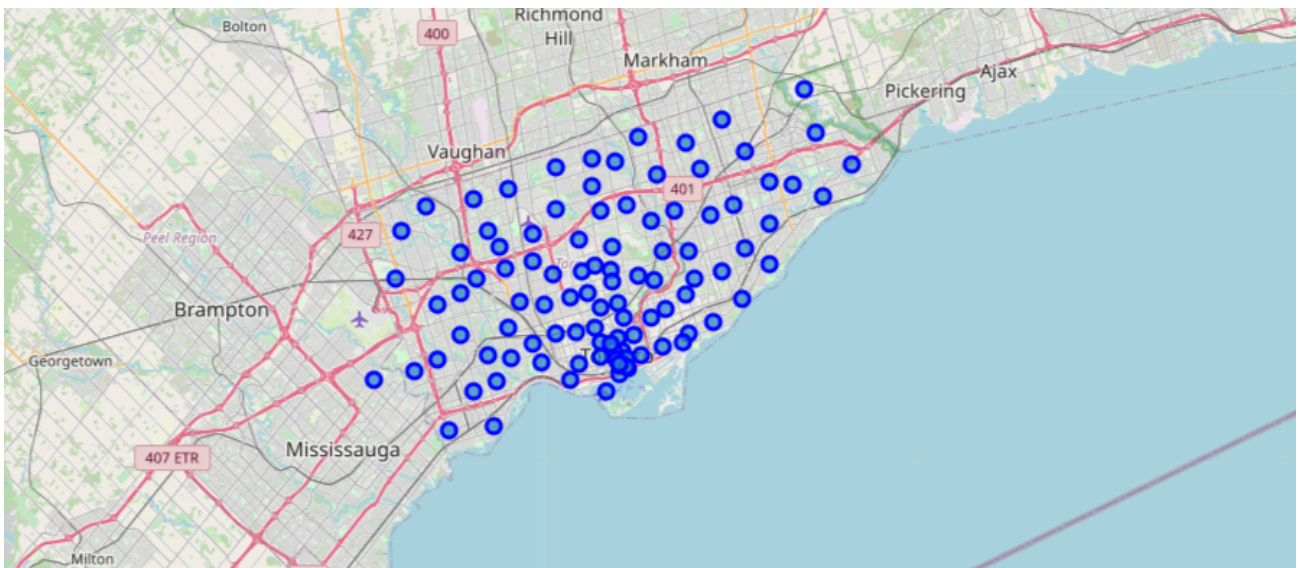
To explore location data, geographical co-ordinates, namely Latitudes and Longitudes of the various locations are required. Geospatial data is required to accurately reference and pinpoint locations while making API calls to Foursquare data, and also to superimpose location markers on maps generated using the Folium library.

In order to achieve this, the CSV file downloaded from https://cocl.us/Geospatial_data, containing geospatial data representing Postal Codes and Lat-Long co-ordinates of neighbourhoods in Toronto was copied to a data frame. The data frame containing geospatial data was then merged with the original data frame, ensuring that the Postal Codes match with the Lat-Long co-ordinates. The final table looked as shown below:

	Postal Code	Borough	Neighborhood	Latitude	Longitude
0	M1B	Scarborough	Malvern, Rouge	43.806686	-79.194353
1	M1C	Scarborough	Rouge Hill, Port Union, Highland Creek	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

c. Analysing Neighbourhoods in Toronto with respect to Fitness Centres:

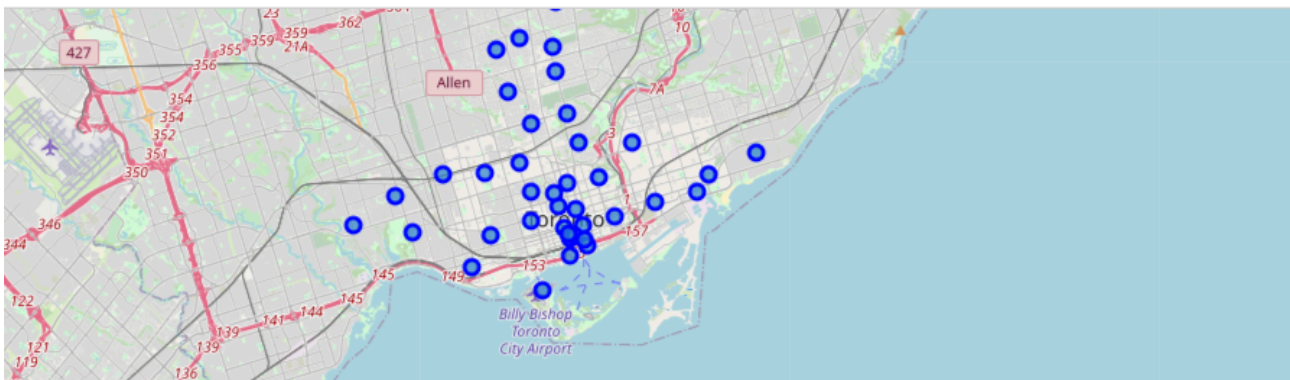
Geographical co-ordinates of Toronto were extracted, and a map was created using Folium library, centred around Toronto. All the prominent neighbourhoods were marked on the map, with the blue dots indicating centre of the neighbourhoods.



Boroughs containing the name Toronto were filtered to create a subset to act upon during the course of our project.

	Postal Code	Borough	Neighborhood	Latitude	Longitude
0	M4E	East Toronto	The Beaches	43.676357	-79.293031
1	M4K	East Toronto	The Danforth West, Riverdale	43.679557	-79.352188
2	M4L	East Toronto	India Bazaar, The Beaches West	43.668999	-79.315572
3	M4M	East Toronto	Studio District	43.659526	-79.340923
4	M4N	Central Toronto	Lawrence Park	43.728020	-79.388790
5	M4P	Central Toronto	Davisville North	43.712751	-79.390197
6	M4R	Central Toronto	North Toronto West, Lawrence Park	43.715383	-79.405678
7	M4S	Central Toronto	Davisville	43.704324	-79.388790
8	M4T	Central Toronto	Moore Park, Summerhill East	43.689574	-79.383160
9	M4V	Central Toronto	Summerhill West, Rathnelly, South Hill, Forest...	43.686412	-79.400049

The same data frame was represented on a map display.



Foursquare credentials including CLIENT_ID, CLIENT_SECRET and VERSION were saved to variables, and then a function named getNearbyGyms() was written to explore venues within Toronto area, with Venue Category as Gyms/Fitness Centres and various subsets of the same.

Foursquare category code for Gyms/ Fitness Centres was obtained from Foursquare website, and used to make the call specific to the type of venues.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	The Beaches	43.676357	-79.293031	Beaches BJJ	43.672703	-79.289013	Martial Arts Dojo
1	The Beaches	43.676357	-79.293031	Link Integrated Health	43.683174	-79.300078	Gym
2	The Beaches	43.676357	-79.293031	Bold Steps Dance Studio	43.673555	-79.283204	Gym / Fitness Center
3	The Beaches	43.676357	-79.293031	Beaches Rec Centre	43.673659	-79.298634	College Gym
4	The Beaches	43.676357	-79.293031	Beach Pilates	43.671808	-79.292115	Pilates Studio
5	The Beaches	43.676357	-79.293031	Prana Fitness	43.671070	-79.295043	Yoga Studio
6	The Beaches	43.676357	-79.293031	Fitness Works For Women	43.670846	-79.294619	Gym
7	The Beaches	43.676357	-79.293031	Horizon Martial Arts	43.672480	-79.289242	Martial Arts Dojo
8	The Beaches	43.676357	-79.293031	Glanville Mediation Services	43.678463	-79.299467	Yoga Studio
9	The Beaches	43.676357	-79.293031	Classical Martial Arts Centre CMAC Beaches Dojo	43.669539	-79.301278	Martial Arts Dojo

It resulted in a data frame as shown above.

Based on one-hot encoding, neighbourhoods were grouped according to the categories of fitness centres available in the location. A data frame was created to indicate the top 5 most common fitness centres available against each neighbourhood.

Machine Learning: K-means Clustering

K-Means Clustering, an unsupervised Machine Learning algorithm, was used to cluster neighbourhoods based on their most commonly used fitness venues. The clusters of neighbourhoods were further used to determine the end result, which is a subset of the same.

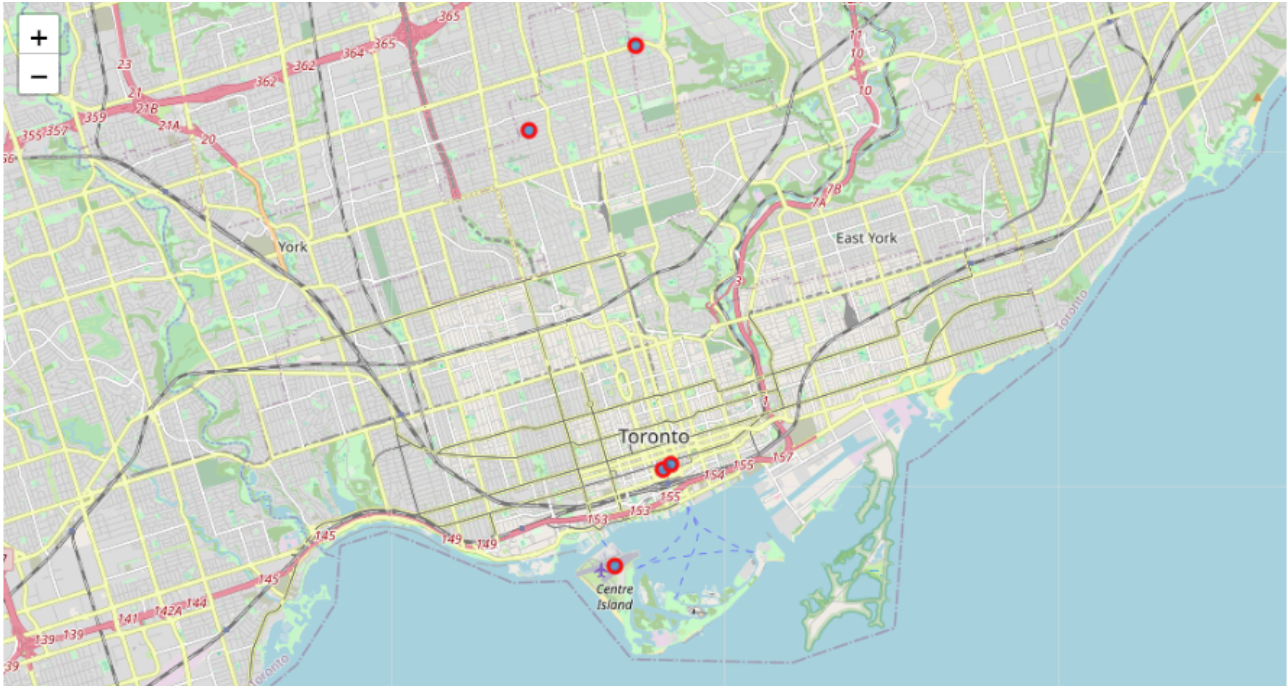
	Neighborhood	1st Most Common Fitness Centre	2nd Most Common Fitness Centre	3rd Most Common Fitness Centre	4th Most Common Fitness Centre	5th Most Common Fitness Centre
0	Berczy Park	Gym	Gym / Fitness Center	Gym Pool	Yoga Studio	Pilates Studio
1	Brockton, Parkdale Village, Exhibition Place	Gym	Gym / Fitness Center	Yoga Studio	Martial Arts Dojo	Boxing Gym
2	Business reply mail Processing Centre, South C...	Gym / Fitness Center	Gym	Yoga Studio	Sports Club	Farmers Market
3	CN Tower, King and Spadina, Railway Lands, Har...	Gym	Gym / Fitness Center	Track	Sports Club	Pilates Studio
4	Central Bay Street	Gym	Gym / Fitness Center	Yoga Studio	Gym Pool	Bike Shop
5	Christie	Gym / Fitness Center	Yoga Studio	Martial Arts Dojo	Gym	Pilates Studio
6	Church and Wellesley	Gym	Gym / Fitness Center	Gym Pool	Yoga Studio	Pilates Studio
7	Commerce Court, Victoria Hotel	Gym	Gym / Fitness Center	Pilates Studio	Track Stadium	Gym Pool
8	Davisville	Gym	Yoga Studio	Gym / Fitness Center	Martial Arts Dojo	Track
9	Davisville North	Gym	Yoga Studio	Gym / Fitness Center	Pilates Studio	Track
10	Dufferin, Dovercourt Village	Gym	Gym / Fitness Center	Martial Arts Dojo	Boxing Gym	Yoga Studio

Now, we have a list of neighbourhoods with the 5 most common fitness centres available in the locality. The data frame was analysed to identify the neighbourhoods wherein an existing Yoga Studio did not form a part of the top 5 fitness centres. It returned a data frame as indicated below.

	Postal Code	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Fitness Centre	2nd Most Common Fitness Centre	3rd Most Common Fitness Centre	4th Most Common Fitness Centre	5th Most Common Fitness Centre
0	M4N	Central Toronto	Lawrence Park	43.728020	-79.388790	4	Gym / Fitness Center	Gym	College Gym	Gym Pool	Bike Shop
1	M5K	Downtown Toronto	Toronto Dominion Centre, Design Exchange	43.647177	-79.381576	0	Gym	Gym / Fitness Center	Pilates Studio	Track Stadium	Gym Pool
2	M5L	Downtown Toronto	Commerce Court, Victoria Hotel	43.648198	-79.379817	0	Gym	Gym / Fitness Center	Pilates Studio	Track Stadium	Gym Pool
3	M5N	Central Toronto	Roselawn	43.711695	-79.416936	3	Gym	Gym / Fitness Center	Martial Arts Dojo	Pilates Studio	Gym Pool
4	M5V	Downtown Toronto	CN Tower, King and Spadina, Railway Lands, Har...	43.628947	-79.394420	0	Gym	Gym / Fitness Center	Track	Sports Club	Pilates Studio

The list consisted of a total of five neighbourhoods, with three clusters from Central Toronto and two clusters from Downtown Toronto. A total of 10 neighbourhoods formed a part of all the clusters combined together.

The final list was visualised, with the locations marked on a map generated using Folium. The final representation is as shown below:



4. Results

As already indicated above, the analysis resulted with 5 locations most suitable for launching a new chain of Yoga Studios. The general fitness culture, as indicated by the most common fitness centres, indicate an inclination towards Gyms, Pilates, Martial Arts, Pools and Track, and presents a viable opportunity to generate a hype and create an impact by launching an excellent alternative to the conventional fitness routines.

	Postal Code	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Fitness Centre	2nd Most Common Fitness Centre	3rd Most Common Fitness Centre	4th Most Common Fitness Centre	5th Most Common Fitness Centre
0	M4N	Central Toronto	Lawrence Park	43.728020	-79.388790	4	Gym / Fitness Center	Gym	College Gym	Gym Pool	Bike Shop
1	M5K	Downtown Toronto	Toronto Dominion Centre, Design Exchange	43.647177	-79.381576	0	Gym	Gym / Fitness Center	Pilates Studio	Track Stadium	Gym Pool
2	M5L	Downtown Toronto	Commerce Court, Victoria Hotel	43.648198	-79.379817	0	Gym	Gym / Fitness Center	Pilates Studio	Track Stadium	Gym Pool
3	M5N	Central Toronto	Roselawn	43.711695	-79.416936	3	Gym	Gym / Fitness Center	Martial Arts Dojo	Pilates Studio	Gym Pool
4	M5V	Downtown Toronto	CN Tower, King and Spadina, Railway Lands, Har...	43.628947	-79.394420	0	Gym	Gym / Fitness Center	Track	Sports Club	Pilates Studio

5. Discussions

While clustering various neighbourhoods using the K-Means Clustering method, similarities in the most common types of fitness centres were utilised. From this data frame, neighbourhoods that did not have a Yoga Studio at all amongst the top 5 most common fitness centres were filtered out as a subset. Here, the following assumptions have been made:

- a. It was assumed that normal gyms or fitness centres tagged as such in the Foursquare data does not include Yoga sessions, or if it did, the sessions were not very popular.
- b. It is also being assumed that there were no Yoga Studios in the top 5 most common venues, because:
 - i. There were no such studios available in the locality, or
 - ii. The ones that were available were not very popular due to questionable standards.

and not because the people were not inclined towards Yoga as a form of fitness training.

The above mentioned results were drawn on the basis of these assumptions, in combination with the analysed data frames containing location information.

6. Conclusion

Based on the results, and the ensuing discussions, where in various assumptions were considered at a qualitative level, the location clusters identified from the project has been found feasible for the launch of A1's Yoga Studios. The exact locations would further depend upon a market feasibility study which would include cost analysis, infrastructure and various local rules and regulations.

It needs to be noted that Clustering methods, such as K-Means Clustering, are extremely powerful machine learning algorithms which can work on unlabelled data, and generate insights based on the generated clusters - identifying characteristics that indicate high degrees of intra cluster similarities as well as inter-cluster dissimilarities. Though we have utilised only K-Means Clustering to work upon the location information from Foursquare, various different methodologies such as Prediction and Classification also exist, which work on labelled data (supervised learning) to derive actionable insights.