

# Introduction

## Description and Discussion of the Background.

Toronto is the most populous city in Canada with 2,73 million people in 2016 and the fourth most populous city in North America.

Toronto is an international center for business and finance. It has a high concentration of banks and brokerage firms. The city is the headquarter of many large Canadian and multinational corporations. Different companies in technology, design, life sciences, education, arts, fashion, aerospace, environmental innovation, food services and tourism are represented here. The city also is an important center for media, publishing, telecommunication, information technology and film production industries.

We are looking for to open a chain of yoga studios in Toronto. Yoga is getting more and more popular among business people and employees as after intense, nervous and sedentary work, yoga helps to restore physical and mental strength. Regular users may find that yoga prevents pain, increases body awareness and reduces stress that accompanies a person in such a metropolis as Toronto.

## Business Problem

The objective of the capstone project to find the most suitable locations for yoga studios. For this it is necessary to investigate the situation with yoga studios in different neighborhoods of the city, to determine in which neighborhood the need for more yoga studios is possible.

## Target Audience

This report will help entrepreneurs looking for to open various centers or studios of yoga or other wellness practices.

## Data Description.

To solve the problem, we will use following data:

- List of neighborhoods in Toronto;
- Latitude and Longitude data of Toronto neighborhoods;
- Venue data related to yoga, gym or other wellness/fitness centers. This will help to find the neighborhoods more suitable to open yoga studios.

## Data sources.

- Toronto neighborhoods from Wikipedia;
- Latitude and Longitude data of Toronto neighborhoods from Geocoder package;
- Venue data in Toronto neighborhoods from Foursquare API.

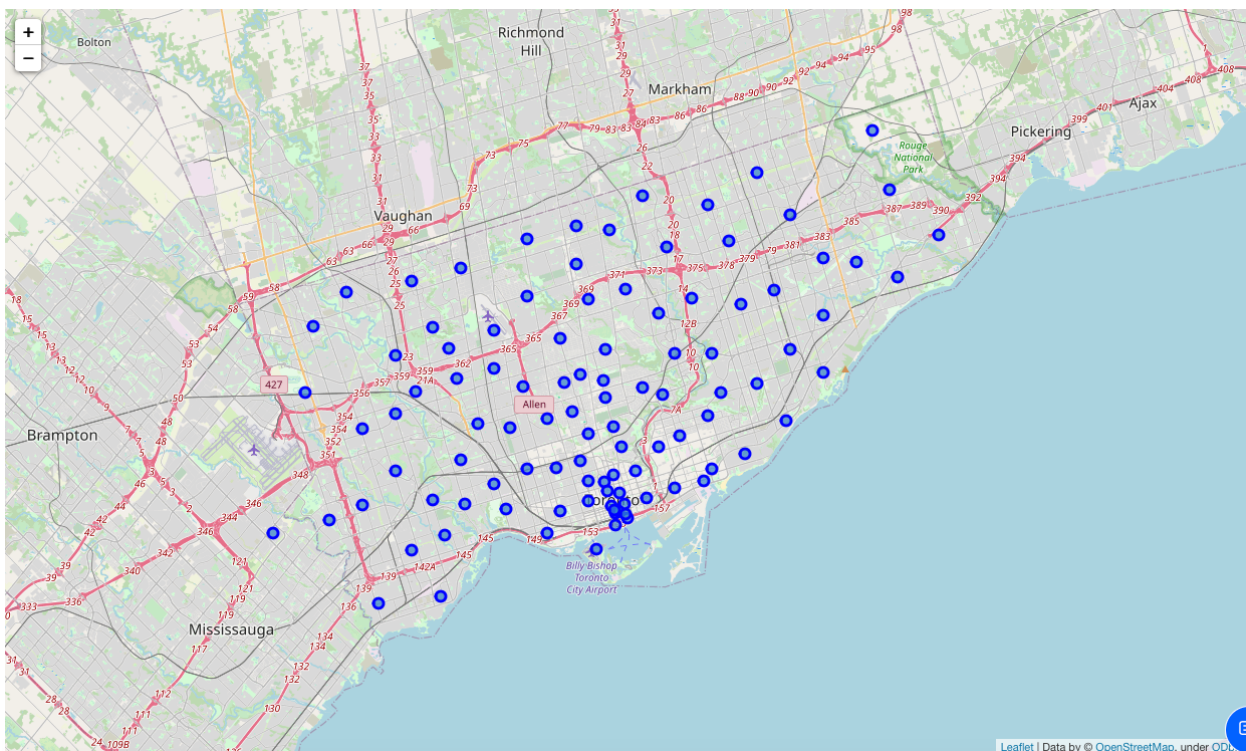
## Methodology.

As a database, I used GitHub repository in my study.

I took information about Postal Codes, Boroughs and Neighborhoods of Toronto city from Wikipedia, then added Latitude and Longitude data for each postal code from Geocoder Package. One borough can have several postal codes, one postal code can have several neighborhoods which are combined in the column Neighborhood.

|   | PostalCode | Borough          | Neighborhood                                | Latitude  | Longitude  |
|---|------------|------------------|---|-----------|------------|
| 0 | M3A        | North York       | Parkwoods                                   | 43.753259 | -79.329656 |
| 1 | M4A        | North York       | Victoria Village                            | 43.725882 | -79.315572 |
| 2 | M5A        | Downtown Toronto | Regent Park, Harbourfront                   | 43.654260 | -79.360636 |
| 3 | M6A        | North York       | Lawrence Manor, Lawrence Heights            | 43.718518 | -79.464763 |
| 4 | M7A        | Downtown Toronto | Queen's Park, Ontario Provincial Government | 43.662301 | -79.389494 |

I used Python Folium library to visualize geographic details of Toronto, its boroughs and neighborhoods. Latitude and longitude values were used to get the visual as below:



I utilized the Foursquare API to explore the boroughs and neighborhoods. I designed the limit as 100 venues and the radius 500 meter for each postal code from their given latitude and longitude information. Here is the part of the list venues name, category, latitude and longitude from Foursquare API.

|   | name                                      | categories            | lat       | lng        |
|---|---|-----------------------|-----------|------------|
| 0 | Victoria Village Arena                    | Hockey Arena          | 43.723481 | -79.315635 |
| 1 | Portugril                                 | Portuguese Restaurant | 43.725819 | -79.312785 |
| 2 | Tim Hortons                               | Coffee Shop           | 43.725517 | -79.313103 |
| 3 | The Frig                                  | French Restaurant     | 43.727051 | -79.317418 |
| 4 | Eglinton Ave E & Sloane Ave/Bermondsey Rd | Intersection          | 43.726086 | -79.313620 |

2141 venues were returned by Foursquare. Here is a merged table of boroughs and venues:

|   | Borough          | Borough Latitude | Borough Longitude | Venue                             | Venue Latitude | Venue Longitude | Venue Category             |
|---|------------------|------------------|-------------------|-----------------------------------|----------------|-----------------|----------------------------|
| 0 | Parkwoods        | 43.753259        | -79.329656        | Brookbanks Park                   | 43.751976      | -79.332140      | Park                       |
| 1 | Parkwoods        | 43.753259        | -79.329656        | Variety Store                     | 43.751974      | -79.333114      | Food & Drink Shop          |
| 2 | Parkwoods        | 43.753259        | -79.329656        | Corrosion Service Company Limited | 43.752432      | -79.334661      | Construction & Landscaping |
| 3 | Victoria Village | 43.725882        | -79.315572        | Victoria Village Arena            | 43.723481      | -79.315635      | Hockey Arena               |
| 4 | Victoria Village | 43.725882        | -79.315572        | Portugril                         | 43.725819      | -79.312785      | Portuguese Restaurant      |

Here is the number of venues in neighborhoods:

|  | Neighborhood   | Neighborhood Latitude | Neighborhood Longitude | Venue | Venue Latitude | Venue Longitude | Venue Category |
|--|--|-----------------------|------------------------|-------|----------------|-----------------|----------------|
|  | Agincourt  | 4                     | 4                      | 4     | 4              | 4               | 4              |
|  | Alderwood, Long Branch   | 7                     | 7                      | 7     | 7              | 7               | 7              |
|  | Bathurst Manor, Wilson Heights, Downsview North                                      | 21                    | 21                     | 21    | 21             | 21              | 21             |
|  | Bayview Village  | 4                     | 4                      | 4     | 4              | 4               | 4              |
|  | Bedford Park, Lawrence Manor East  | 23                    | 23                     | 23    | 23             | 23              | 23             |
|  | Berczy Park  | 58                    | 58                     | 58    | 58             | 58              | 58             |
|  | Birch Cliff, Cliffside West  | 4                     | 4                      | 4     | 4              | 4               | 4              |
|  | Brockton, Parkdale Village, Exhibition Place   | 22                    | 22                     | 22    | 22             | 22              | 22             |
|  | Business reply mail Processing Centre, South Central Letter Processing Plant Toronto | 16                    | 16                     | 16    | 16             | 16              | 16             |

As limit for venues was at 100 and limit for radius was at 500 meters, so not all venues were counted here. If we increase limits, we can get more venues.

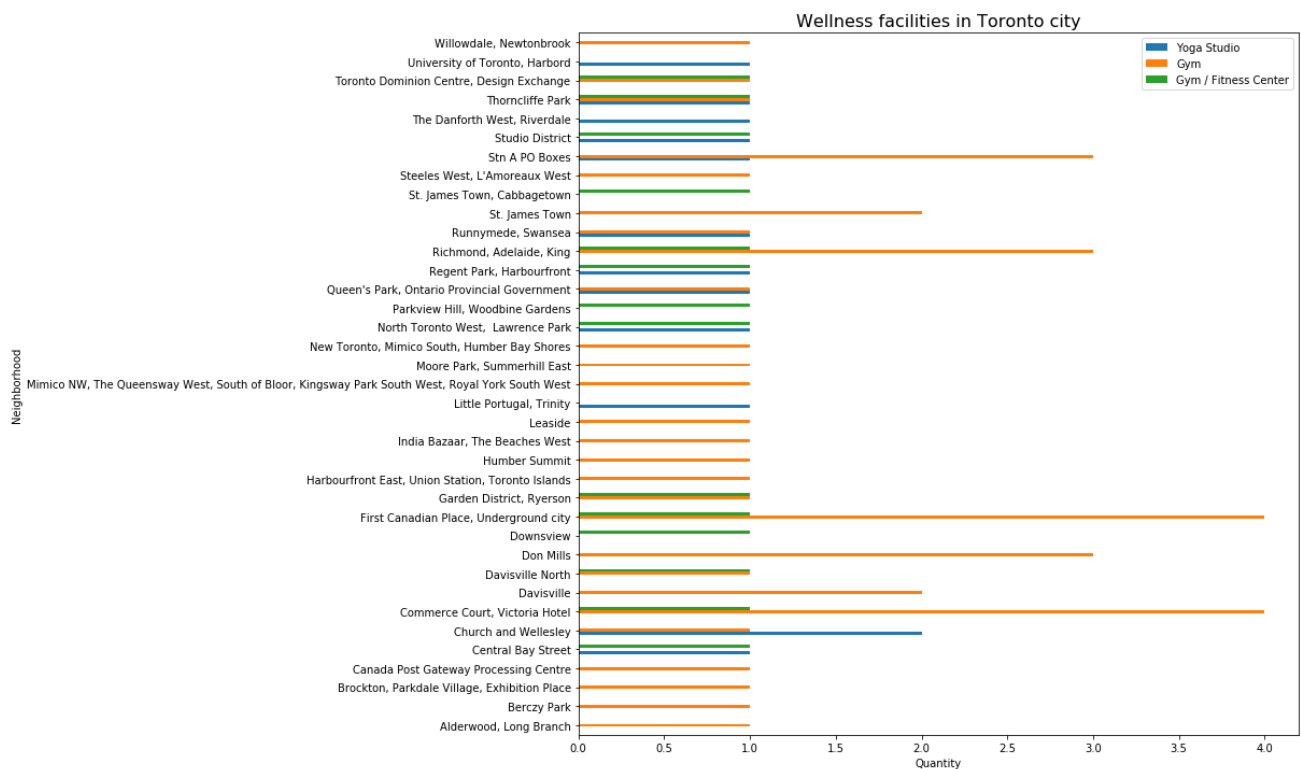
Venues of 269 unique categories were returned by Foursquare. To find which of them could be related to wellness practices I made a list of their names. Here is the part of it:

```
[ 'Yoga Studio',
  'Accessories Store',
  'Afghan Restaurant',
  'Airport',
  'Airport Food Court',
  'Airport Gate',
  'Airport Lounge',
  'Airport Service',
  'Airport Terminal',
  'American Restaurant',
  'Antique Shop',
  'Aquarium',
  'Art Gallery',
  'Art Museum',
  'Arts & Crafts Store',
  'Asian Restaurant',
  'Athletics & Sports',
  'Auto Garage',
  'Auto Workshop',
```

I found category of 'Yoga Studio'. Then I assumed that categories of "Gym" and "Gym/Fitness Center" also can provide some wellness practices as yoga or similar and decided to review them together with 'Yoga Studios'. Then I defined which neighborhoods have venues of 'Yoga Studio', "Gym" and "Gym/Fitness Center" categories.

|  | Yoga Studio | Gym | Gym / Fitness Center |
|--|-------------|-----|----------------------|
| Neighborhood                                 |             |     |                      |
| Alderwood, Long Branch                       | 0           | 1   | 0                    |
| Berczy Park                                  | 0           | 1   | 0                    |
| Brockton, Parkdale Village, Exhibition Place | 0           | 1   | 0                    |
| Canada Post Gateway Processing Centre        | 0           | 1   | 0                    |
| Central Bay Street                           | 1           | 0   | 1                    |

In total 37 neighborhoods have 'Yoga Studio' or "Gym" or "Gym/Fitness Center" venues. So here is no need to open a yoga studio and compete with existing venues.



Then I found neighborhoods which do not have venues of 'Yoga Studio', "Gym" and "Gym/Fitness Center" categories. Here is the part of that list with neighborhood names.

|  | Yoga Studio | Gym | Gym / Fitness Center |
|--|-------------|-----|----------------------|
| Neighborhood   |             |     |                      |
| Agincourt  | 0           | 0   | 0                    |
| Bathurst Manor, Wilson Heights, Downsview North  | 0           | 0   | 0                    |
| Bayview Village  | 0           | 0   | 0                    |
| Bedford Park, Lawrence Manor East  | 0           | 0   | 0                    |
| Birch Cliff, Cliffside West  | 0           | 0   | 0                    |
| Business reply mail Processing Centre, South Central Letter Processing Plant Toronto                       | 0           | 0   | 0                    |
| CN Tower, King and Spadina, Railway Lands, Harbourfront West, Bathurst Quay, South Niagara, Island airport | 0           | 0   | 0                    |
| Caledonia-Fairbanks  | 0           | 0   | 0                    |
| Cedarbrae  | 0           | 0   | 0                    |
| Christie   | 0           | 0   | 0                    |
| Clarks Corners, Tam O'Shanter, Sullivan  | 0           | 0   | 0                    |
| Cliffside, Cliffcrest, Scarborough Village West  | 0           | 0   | 0                    |
| Del Ray, Mount Dennis, Keelsdale and Silverthorn   | 0           | 0   | 0                    |
| Dorset Park, Wexford Heights, Scarborough Town Centre  | 0           | 0   | 0                    |
| Dufferin, Dovercourt Village   | 0           | 0   | 0                    |
| East Toronto, Broadview North (Old East York)  | 0           | 0   | 0                    |
| Eringate, Bloordale Gardens, Old Burnhamthorpe, Markland Wood  | 0           | 0   | 0                    |
| Fairview, Henry Farm, Oriole   | 0           | 0   | 0                    |
| Forest Hill North & West, Forest Hill Road Park  | 0           | 0   | 0                    |
| Glencairn  | 0           | 0   | 0                    |
| Golden Mile, Clairlea, Oakridge  | 0           | 0   | 0                    |
| Guildwood, Morningside, West Hill  | 0           | 0   | 0                    |
| High Park, The Junction South  | 0           | 0   | 0                    |
| Hillcrest Village  | 0           | 0   | 0                    |
| Humberlea, Emery   | 0           | 0   | 0                    |
| Humewood-Cedarvale   | 0           | 0   | 0                    |

So 59 neighborhoods do not have venues that provide wellness practices and they can be considered as right areas to open yoga studios.

Further, I decided to see if wellness practices are popular in any neighborhoods and found that in some neighborhoods gym and yoga studios are in 5 most common venues:

----First Canadian Place, Underground city----

|   | venue       | freq |
|---|-------------|------|
| 0 | Coffee Shop | 0.10 |
| 1 | Café        | 0.08 |
| 2 | Hotel       | 0.05 |
| 3 | Gym         | 0.04 |
| 4 | Restaurant  | 0.04 |

----Forest Hill North & West, Forest Hill Road Park----

|   | venue              | freq |
|---|--------------------|------|
| 0 | Trail              | 0.25 |
| 1 | Sushi Restaurant   | 0.25 |
| 2 | Mexican Restaurant | 0.25 |
| 3 | Jewelry Store      | 0.25 |
| 4 | Yoga Studio        | 0.00 |

So potentially new yoga studios have good chances to attract members as people have interest in them.

Then I decided to apply K-means algorithm to cluster neighborhoods with some common venue categories. K-means can group data only unsupervised based on the similarity of neighborhoods to

each other. I ran K-means to cluster the neighborhoods into 5 clusters. Here is my merged table with cluster labels for each neighborhood.

|   | PostalCode | 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 Most Common Venue | 8th Most Common Venue | 9th Most Common Venue | 10th Most Common Venue |
|---|------------|------------------|---|-----------|------------|----------------|----------------------------|------------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|------------------------|
| 0 | M3A        | North York       | Parkwoods                                   | 43.753259 | -79.329656 | 2.0            | Construction & Landscaping | Park                   | Food & Drink Shop     | Women's Store         | Deli / Bodega         | Department Store      | Dessert Shop          | Dim Sum Restaurant    | Diner                 | Discount Store         |
| 1 | M4A        | North York       | Victoria Village                            | 43.725882 | -79.315572 | 0.0            | French Restaurant          | Pizza Place            | Intersection          | Hockey Arena          | Portuguese Restaurant | Coffee Shop           | Donut Shop            | Doner Restaurant      | Dog Run               | Distribution Center    |
| 2 | M5A        | Downtown Toronto | Regent Park, Harbourfront                   | 43.654260 | -79.360636 | 0.0            | Coffee Shop                | Park                   | Pub                   | Bakery                | Theater               | Breakfast Spot        | Café                  | Health Food Store     | Beer Store            | Shoe Store             |
| 3 | M6A        | North York       | Lawrence Manor, Lawrence Heights            | 43.718518 | -79.464763 | 0.0            | Clothing Store             | Furniture / Home Store | Accessories Store     | Miscellaneous Shop    | Vietnamese Restaurant | Event Space           | Coffee Shop           | Boutique              | Women's Store         | Diner                  |
| 4 | M7A        | Downtown Toronto | Queen's Park, Ontario Provincial Government | 43.662301 | -79.389494 | 0.0            | Coffee Shop                | Diner                  | College Cafeteria     | Bar                   | Beer Bar              | Smoothie Shop         | Sandwich Place        | Burrito Place         | Café                  | Park                   |

## Results.

Here are 59 neighborhoods that do not have venues that provide wellness practices and they can be considered as right areas to open yoga studios. Here is the part of the list:

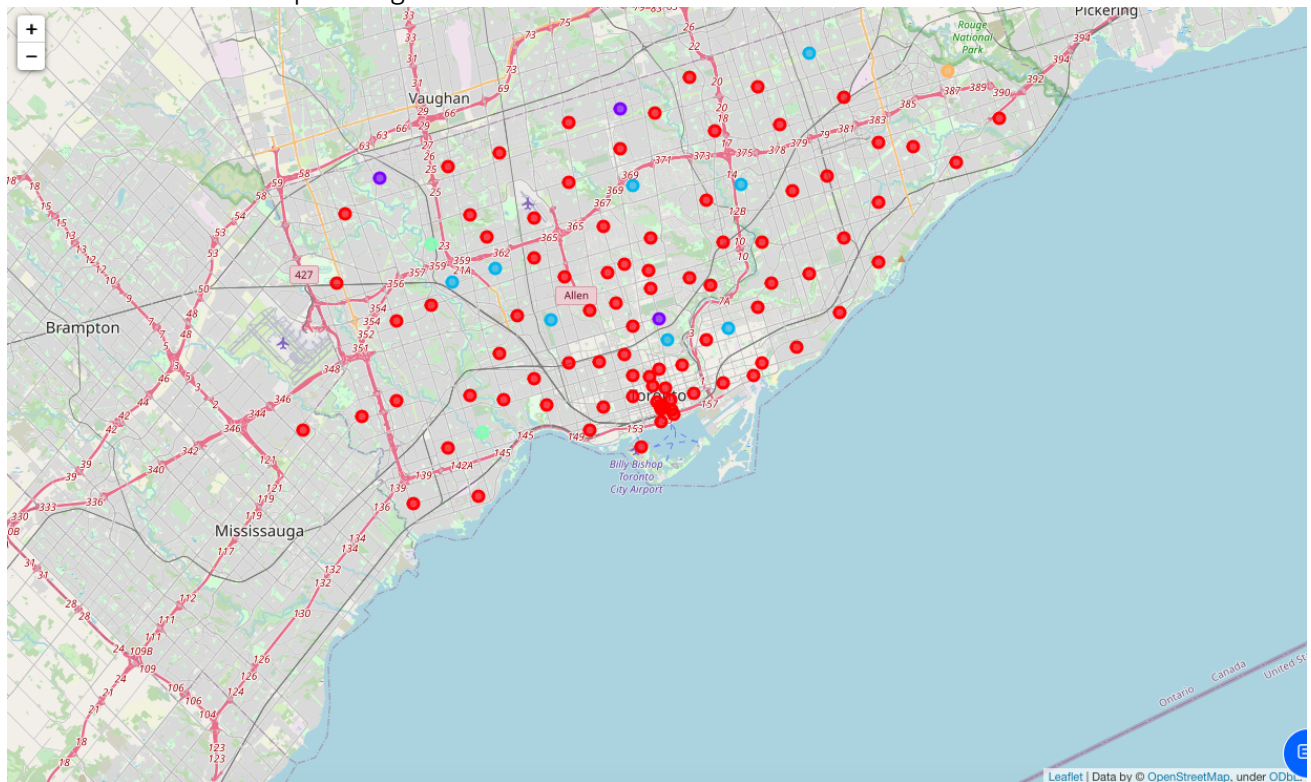
|  | Neighborhood   | Yoga Studio | Gym | Gym / Fitness Center |
|--|--|-------------|-----|----------------------|
|  | Agincourt  | 0           | 0   | 0                    |
|  | Bathurst Manor, Wilson Heights, Downsview North  | 0           | 0   | 0                    |
|  | Bayview Village  | 0           | 0   | 0                    |
|  | Bedford Park, Lawrence Manor East  | 0           | 0   | 0                    |
|  | Birch Cliff, Cliffside West  | 0           | 0   | 0                    |
|  | Business reply mail Processing Centre, South Central Letter Processing Plant Toronto                       | 0           | 0   | 0                    |
|  | CN Tower, King and Spadina, Railway Lands, Harbourfront West, Bathurst Quay, South Niagara, Island airport | 0           | 0   | 0                    |
|  | Caledonia-Fairbanks  | 0           | 0   | 0                    |
|  | Cedarbrae  | 0           | 0   | 0                    |
|  | Christie   | 0           | 0   | 0                    |
|  | Clarks Corners, Tam O'Shanter, Sullivan  | 0           | 0   | 0                    |
|  | Cliffside, Cliffcrest, Scarborough Village West  | 0           | 0   | 0                    |
|  | Del Ray, Mount Dennis, Keelsdale and Silverthorn   | 0           | 0   | 0                    |
|  | Dorset Park, Wexford Heights, Scarborough Town Centre  | 0           | 0   | 0                    |
|  | Dufferin, Dovercourt Village   | 0           | 0   | 0                    |
|  | East Toronto, Broadview North (Old East York)  | 0           | 0   | 0                    |
|  | Eringate, Bloordale Gardens, Old Burnhamthorpe, Markland Wood  | 0           | 0   | 0                    |
|  | Fairview, Henry Farm, Oriole   | 0           | 0   | 0                    |
|  | Forest Hill North & West, Forest Hill Road Park  | 0           | 0   | 0                    |
|  | Glencairn  | 0           | 0   | 0                    |
|  | Golden Mile, Clairlea, Oakridge  | 0           | 0   | 0                    |
|  | Guildwood, Morningside, West Hill  | 0           | 0   | 0                    |
|  | High Park, The Junction South  | 0           | 0   | 0                    |
|  | Hillcrest Village  | 0           | 0   | 0                    |
|  | Humberlea, Emery   | 0           | 0   | 0                    |
|  | Humewood-Cedarvale   | 0           | 0   | 0                    |

Yoga studios and gym are around 5 common venues in some neighborhoods.

Clusters that I got using K-mean method confirmed that population in Toronto city has strong interest in different sport activities, including yoga and gym:

- Cluster 1 (All-rounders) is the largest. Most common venues are Coffee Shops, restaurants and stores. However, gym, fitness and yoga are among 10 common venues also.
- Cluster 2 (Sportsmen) has gym as the 1<sup>st</sup> common venue in all neighborhoods.
- Cluster 3 (Joggers and Hikers) has park as the 1<sup>st</sup> common venue in most neighborhoods and as the 2<sup>nd</sup> common venue in the left.
- Cluster 4 (Baseball Fans) consist of 2 neighborhoods and has Baseball Field as the 1<sup>st</sup> common venue.
- Cluster 5 (Eaters and Buyers) consist of 1 neighborhood and prefers just stores and restaurants.

Here is a clustered map of neighborhoods of Toronto:





## Discussion.

Toronto is an international center for business and finance with population of 2,73 million people in 2016. So, the demand in different sport and wellness facilities is quite high.

I considered venues relative to 103 postal codes and their corresponding neighborhoods. Adding limits at 100 and using radius just 500 meters to simplify my calculations for this project I reviewed not all possible venues that could give us more detailed picture of existing sport and wellness facilities in Toronto city.

Then some postal codes include several neighborhoods. Examining each neighborhood separately also can give more details.

I used the K-means algorithm as part of this clustering study. Clusters 2-5 consist of a small number of neighborhoods, but Cluster 1 is quite big and does not exhibit any special property for combining it under some single obvious rule. So, it would make sense to also break cluster 1 into its own clusters and consider the result. However, if initially the breakdown was strictly by neighborhoods, rather than by postal code, and with larger limits and radius, the clustering result could be clearer. The total metro area population of Toronto (including Toronto city) in 2016 was about 6 million people. Part of the population that works in city but lives not in the city also could become active users of the wellness facilities such yoga studios. This fact could also be included in consideration.

## Conclusion.

As a result, the project to open yoga studios in Toronto city has high potentials.

High population of Toronto city, strong interest in sport and wellness practices and lack of such facilities confirm the high potential of the project. Here are 59 neighborhoods/postal codes that do not have venues that provide wellness practices and can be considered as right areas to open yoga studios.

This project can be used also by city managers to examine if there are enough sport and wellness facilities in Toronto city.

## References.

- Toronto – Wikipedia.
- Foursquare API
- <https://www.toronto.ca/>
- <https://www.macrotrends.net/cities/20402/toronto/population>