

Opening a New Fitness Centre in Toronto, Canada

Vijay Saddi

Data Science, IBM (Coursera), Toronto, Canada
E-mail: vijayks102@gmail.com

Completed: 2020-03-31
Published? No

Abstract

With the popularity of fitness centres and gyms on the rise in recent years, a lucrative business opportunity exists for those seeking to open one of their own in Toronto, Canada. This research project examined the ideal location to open a fitness centre in Toronto, based on characteristics within local neighbourhoods. Neighbourhood data was extracted from Wikipedia, neighbourhood data was pulled using the Foursquare Develop API, and transformed in such a way to create 5 clusters for the k -Means Clustering machine learning algorithm. Results showed that Downtown Toronto has the most local venues for which consider when opening the business. The findings suggested that the financial district has little to no competition, which might yield the most success for the business venture.

Keywords: Business, gym, fitness, profitable, opportunity, investment

1. Introduction

For some individuals, a fitness centre/gym is essential to maintain fitness and conditioning for a specific sport or career. For others, vanity may be a main driver in the goal for a more appealing physique – whether it be for a vacation or other personal reasons. One thing for certain, is that gyms have become increasingly popular in recent years and present a lucrative business venture. Therefore, it may be a prime opportunity for a new entrepreneur or existing business owner to establish or expand a current business. In modern society, fitness has exploded and become more relevant to a wide range of individuals with different personality types. With these main premises in mind, this project is aimed to aid an individual or business to find the most suitable location in Toronto.

1.1 Business Problem

The main objective of this project is to determine the most suitable location to open a new fitness centre in Toronto, Canada. By using various data science methods/tools along with machine learning algorithms, the end goal is to provide a clear and concise solution to the following business question: *What would be the most ideal location to open a fitness centre in Toronto, Canada?*

This study will explore the city of Toronto, Canada in order to determine the most ideal neighbourhood to open a fitness centre.

1.2 Target Audience

The intended audience for this study is the aspiring or existing business owner who seeks to open a fitness centre in Toronto, Canada. Research indicates that as of 2018, the timing is optimal due to significant urban growth in the city. A report published by the city of Toronto suggests that urban plans aim to direct growth away from residential areas into avenues, centres and downtown.¹ The report suggests that the current project aimed at strong development prospects to bring more people and jobs into the city.

Strong growth in the downtown core is ideal, as it will bring more jobs, residents and general business into the city over the next decade. Coupled with the increasing trend in fitness and health, the argument for a successful gym seems to strengthen.

2. Data

The results of this study will be heavily influenced by data sources and data transformations. This means that it is critical that the input data must be credible and accurate.

Similarly, any modification to the original data set must be warranted as to not provide inaccurate results.

2.1 Data as the Main Driver

The main driver behind the results produced from this study is the data. In order to solve the business problem, data will be gathered based on:

1. Toronto neighbourhood data. This defines the scope of the project which is limited to the city.
2. Latitude and longitudes pertaining to neighbourhoods within Toronto. This is also required to generate maps, and pull venue information.
3. Venue data from each individual neighbourhood. We will use the data to create clusters based on features within the neighbourhoods.

2.2 Transforming the Data

In order to render the data into a usable format, the following methods will be used to gain a better understanding:

1. Web scraping Toronto neighbourhood data, taken from Wikipedia (as of February 2020)
2. Retrieving latitude and longitude data for each neighbourhood using a python Geocoder package
3. Utilizing a developer API provided by Foursquare, to retrieve venue data for each neighbourhood for further investigations

The combination of the source data along with required data transformations will allow for much more meaningful insights for which to provide the target audience.

3. Methodology

This section will outline in more detail the proposed steps taken from 2.2, as we perform transformations to the source data.

3.1.1 Web Scraping

The first step is the gather the neighbourhood data for the city of Toronto. There are many ways to obtain this data, however we will use the data available on the Wikipedia² page, version February of 2020. This is performed by using a python library called *BeautifulSoup*, which specializes in this task.

3.1.2 Retrieving Coordinates using Geocoder

The newly obtained data is helpful, however it is essentially a table with postal codes, boroughs, and neighbourhood

names. In order to search for venues using the *Foursquare API*, we need to provide it with coordinates. To do so, we use the python library called *Geocoder* which can convert our location names into geographical coordinates. The new coordinates were merged to the master table which includes the Toronto data.

In order to verify that the coordinates are in fact correct, we use python's *Folium* package which can create a map from the data provided. In our case, the map correctly displayed the city of Toronto.

3.1.3 Pulling Venues using Foursquare API

The next step involves using the Foursquare Developer API to get the top 100 venues within a 500 meter radius. This is done through performing various API calls which pass in the exact coordinates (latitude and longitude). In turn, the API will use the information in our API call to query its database and return a JSON file with the venue details. From this file, the relevant data is extracted and reshaped into a python Pandas data frame for further exploration and processing. We perform some statistics and aggregate functions to explore what kind of venues are nearby. This step will aid in finding nearby venues which may complement a suitable location for the fitness centre.

3.1.4 Performing k-Means Clustering

The final step is the perform a clustering machine learning algorithm on the transformed data. In this case, the best algorithm to use is the k-means clustering method. This algorithm identifies a user-defined, k number of centroids, and attempts to group similar data points towards a specific cluster based on its properties. A successful cluster is clearly visible and keeps the distance between the centroid and data point at a minimum.

In this case, we will specify 5 k clusters from which to determine which neighbourhoods have nearby attractions, already have gyms and overall appear to be suitable for a gym.

4. Results

As mentioned, we specified 5 k clusters for the algorithm to calculate, based on gyms being the 1st most common venue. The clusters can be viewed by superimposing them on a map of Toronto. The following clusters can be viewed in figure 1.0:

- Cluster 1
- Cluster 2
- Cluster 3
- Cluster 4
- Cluster 5

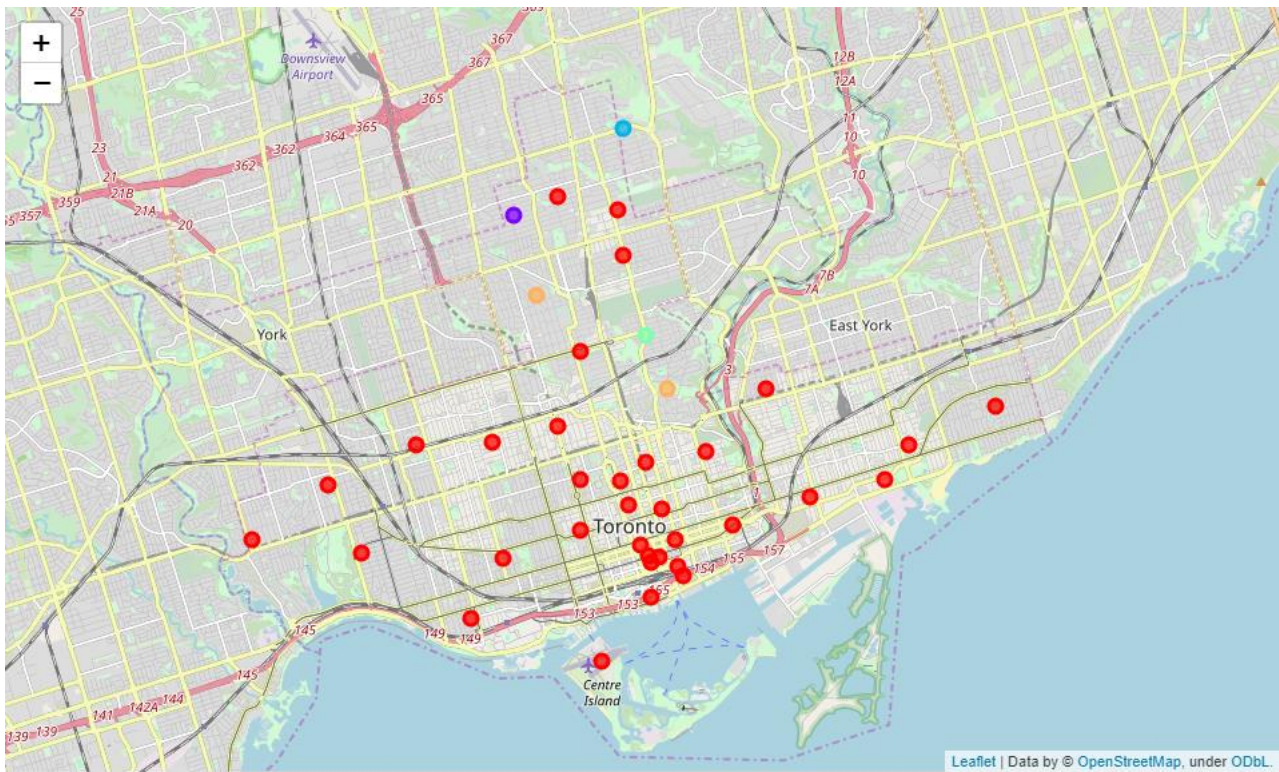


Figure 1.0: k-Means Clusters in Toronto

Cluster 1 has 39 local venues, Clusters 2-4 all have 1 local venue and Cluster 5 has 2 local venues.

When considering the venues within each cluster, Clusters 2-5 do not have any gyms or fitness centers listed in their 1st-

common venues in Cluster 1. Coffee Shops are the most common, with a count of 16.

5. Discussion

The clusters displayed in Figure 1.0 show that the overwhelming majority of nearby venues belong to Cluster 1. These Cluster points seem to reside mainly Downtown, with fewer locations dispersed throughout Central Toronto, East Toronto and West Toronto. This makes sense, as Downtown Toronto has a much higher population density (8,210/Km²)³ than other neighbourhoods such as North York (4,915.5/km²).⁴ With this information in mind, it can be recommended that Cluster 1 contains several ideal locations for which to open a fitness centre.

As we delve deeper into Cluster 1, we would need to determine if there is specific district which can yield the largest return in terms of success of the business. This is because Cluster 1 appears to be quite large and the data points disperse all over Central and Downtown Toronto, as seen in Figure 1.0. We can explore this further by considering the frequency of Boroughs within the cluster. As shown in Figure 3.0, Downtown Toronto has overwhelmingly more districts than the others combined, at 18.

If we want to consider 1st-5th most common venues in Downtown Toronto, we do not see fitness centres listed at all. That is not to say that they do not exist at all – but as the label suggests, they are not a ‘Most common Venue’. In fact,

1st Most Common Venue	
Coffee Shop	16
Café	3
Sandwich Place	2
Bar	2
Park	1
Greek Restaurant	1
Pharmacy	1
Gift Shop	1
Light Rail Station	1
Japanese Restaurant	1
Airport Service	1
Neighborhood	1
Gym	1
Grocery Store	1
Thai Restaurant	1

Figure 2.0: Cluster 1- 1st Most Common Venues

10th most common venues. Cluster 1 has a gym located in Central Toronto as its 1st most common venue. Also within Cluster 1, Downtown Toronto does not have a Gym listed as its 1st most common venue. Figure 2.0 shows the 1st most

Central Toronto is the only district which has a Gym listed at all, within the data. Looking back at Cluster 1 in Figure 1.0, we may also notice that the frequency of data points tends towards the Downtown core – circling the intersection of Bay

Borough	
Downtown Toronto	18
West Toronto	6
East Toronto	5
Central Toronto	5

Figure 3.0: Boroughs Within Cluster 1

Street and Richmond Street. Figure 4.0 best depicts this observation. This area also contains the main Financial District, as well as company head offices and local businesses. As Figure 2.0 indicates, there also tend to be a high number of coffee shops and restaurants as common venues. Therefore, it

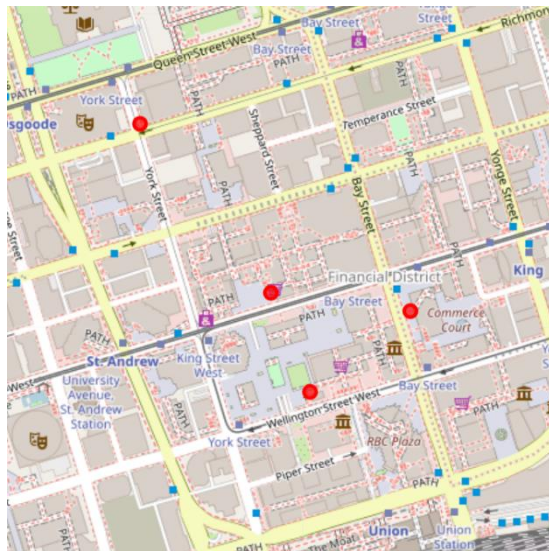


Figure 4.0: Common Venues Within Cluster 1

can be suggested that a fitness centre located nearby would become an attraction to the local community.

As such, this project can recommend current or future entrepreneurs to capitalize on these findings to open a new fitness centre within Cluster 1, tending towards the Downtown Core area of Bay & Richmond Street. The timing seems to be in favour as of March 2020, because there appears to be little to no competition.

6. Limitations and Suggestions

This project is aimed at providing an aspiring fitness centre owner with some valuable points to consider when entering the business venture. It is by no means a fully comprehensive study for which to provide a guaranteed result. For example,

the data obtained only provides common venues and frequency of other local gyms in the area. There are many other financial barriers to consider such as: property construction, cost of rent, utilities, marketing and more. This project also does not evaluate ideal customer interests and behaviour, or whether the recommended location contains the ‘correct’ customer base who would purchase the service.

Further research could be conducted to determine whether operating costs are ideal given the location, as well as local customer interests and purchasing habits.

References

- [1] City of Toronto, 2018, *How Does the City Grow?* url: <https://www.toronto.ca/wp-content/uploads/2018/07/9108-City-Planning-How-Does-The-City-Grow-July-2018.pdf>
- [2] Wikipedia, 2020, *List of postal codes of Canada: M* url: https://en.wikipedia.org/w/index.php?title=List_of_postal_codes_of_Canada:_M&diff=942851379&oldid=942655599
- [3] Wikipedia, 2020, *Old Toronto* url: https://en.wikipedia.org/wiki/Old_Toronto
- [4] Wikipedia, 2020, *North York* url: https://en.wikipedia.org/wiki/North_York