

Title: Understanding the Business of “Sports & Fitness Service” in Toronto

Assignment:

Name: CHAN Chak Yu

Date: 12-Dec-2020

1. Introduction/ Business Problem

1.1 Background

In recent years, with the improvement of healthcare services and living standards, consumers are becoming more health-conscious. This trend is limited to any country or ethnic but remain true all around the globe. Consumers are open to all means to improve their health condition, from enhanced diet to regular health care check-up. Among all means, regular sports and exercising are emerging more popular and important. To fit and improve health, consumers are eager and willing to attempt all kinds of sports activities, including going to gym, fitness centres, yoga studio, etc. This implies unlimited business potential to the “Sports & Fitness” industry.

Being the capital city of Ontario, Toronto is the most populous city in Canada. It accommodates a large population exceeding 2.7M in 2016 and population is still on the rise. It leads to a huge demand for “Sports & Fitness Service”.

However, given its history and all past establishment, Toronto is also a most developed and thus complicated city in Canada. Over the past decades, a huge amount of geographical data (e.g. coordinates, venue address, category, etc.) have accumulated for the city. It is vital to deploy that latest and most comprehensive data for our investigation. Therefore, a Data Science project on Toronto is called for in order to address the business objectives.

1.2 Business Issues and Objectives

It is essential for brands and service providers in “Sports & Fitness” industries to grasp a better and more profound understanding of Toronto market. It includes:

- To gauge the quantities of sports service centres in Toronto and their distribution by neighborhood.
- To identify areas with the most and the least sports service centres.
- To discover potentials for provision of “Sports & Fitness Services” for the city.

2. Data Description

2.1 Data Sources

In this study, we will deploy latest geographical data on Toronto, Canada. Data are all sourced from online public domain in order to be cost-effective and most updated at time. In order to make a comprehensive analysis, data are to be extracted and scrapped from multiple sources, including:

List of postal codes on Wikipedia page

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

There is a table of postal codes for Toronto on the page. Yet, the data in the table are not readily applicable. Libraries (e.g. pandas, or BeautifulSoup package) are deployed so as to transform the data in the table on the Wikipedia page into the pandas DataFrame. We will read pandas data and map distribution based on: Postal Codes, Borough and Neighborhood.

Remarks: We shall only process the cells that have an assigned borough; ignore cells with a borough that is “Not Assigned”.

	Postal Code	Borough	Neighbourhood
0	M1A	Not assigned	Not assigned
1	M2A	Not assigned	Not assigned
2	M3A	North York	Parkwoods
3	M4A	North York	Victoria Village
4	M5A	Downtown Toronto	Regent Park, Harbourfront

Geographical data given by IBM at “Applied Data Science Capstone” course

http://cocl.us/Geospatial_data

This is a link to a CSV file that has the longitude and latitude coordinates of each postal code. The data will help to map.

	Postal Code	Latitude	Longitude
0	M1B	43.806686	-79.194353
1	M1C	43.784535	-79.160497
2	M1E	43.763573	-79.188711
3	M1G	43.770992	-79.216917
4	M1H	43.773136	-79.239476

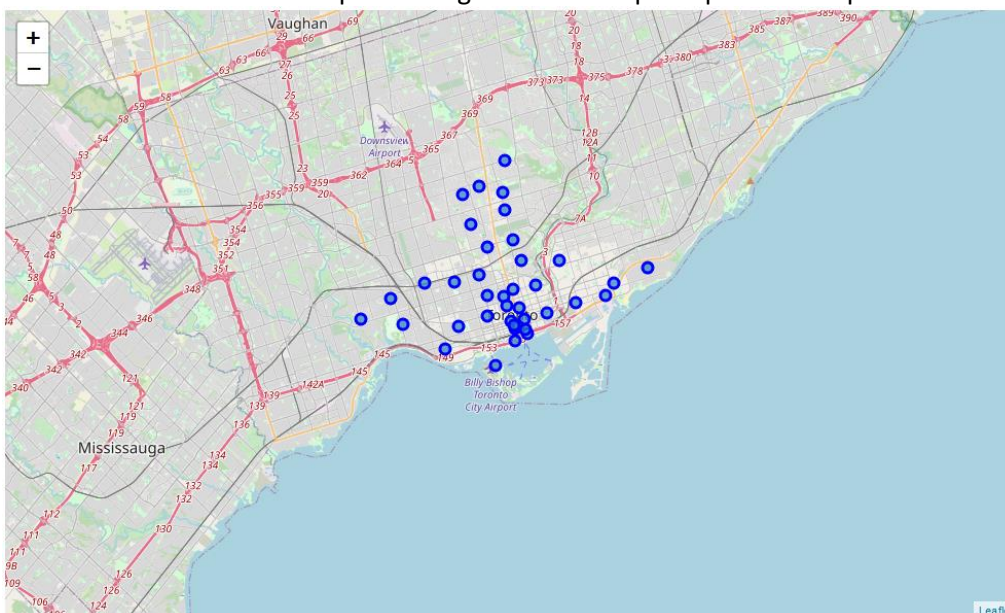
Latest geo-data from various libraries, including:

- Geopy: to get the latitude and longitude values of Toronto
- Folium: to Create a map of Toronto with neighborhoods superimposed on top for visualization analysis
- Foursquare API: to get location data, i.e. name of venues, respective categories, neighborhood, location coordinates for later analysis

To simplify the analysis, we will emphasize the key and downtown areas in Toronto. Therefore, we will screen boroughs with Toronto. Namely, these key boroughs are:

- *Downtown Toronto*
- *East Toronto*
- *West Toronto*
- *Central Toronto*

Here comes the Toronto Map with neighborhoods superimposed on top.



2.2 Data Preparation, Cleaning & Wrangling

Since data come from multiple sources, they may not be readily compatible. We should bear in mind to clean them properly. Furthermore, given a prolonged process of data wrangling, it is a good practice to clean the working notebook and add Markdown cells to explain steps and assumptions that we are making.

3. Methodology and Exploratory Data Analysis

3.1 First Observations

At first glances, we can see the location cover a very wide spectrum of venues and categories, include restaurants and café. In order to focus on “Sports & Fitness Service”, we will focus to look at related categories. They include:

- Athletics & Sports
- Climbing Gym
- College Gym
- Dance Studio
- Gym
- Gym / Fitness Center
- Martial Arts School
- Sports Bar
- Swim School
- Yoga Studio

Remarks: For data analysis in the later stage, let's insert back the column “Borough” into the dataset.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	Borough
0	Regent Park, Harbourfront	43.654260	-79.360636	The Extension Room	43.653313	-79.359725	Gym / Fitness Center	Downtown Toronto
1	Regent Park, Harbourfront	43.654260	-79.360636	The Yoga Lounge	43.655515	-79.364955	Yoga Studio	Downtown Toronto
2	Queen's Park, Ontario Provincial Government	43.662301	-79.389494	The Yoga Sanctuary	43.661499	-79.383636	Yoga Studio	Downtown Toronto
3	Queen's Park, Ontario Provincial Government	43.662301	-79.389494	Hart House Gym	43.664172	-79.394888	Gym	Downtown Toronto
4	Garden District, Ryerson	43.657162	-79.378937	Hard Candy Fitness	43.659556	-79.382440	Gym / Fitness Center	Downtown Toronto

Now, we are about to map all “Sports & Fitness” venues on Toronto.



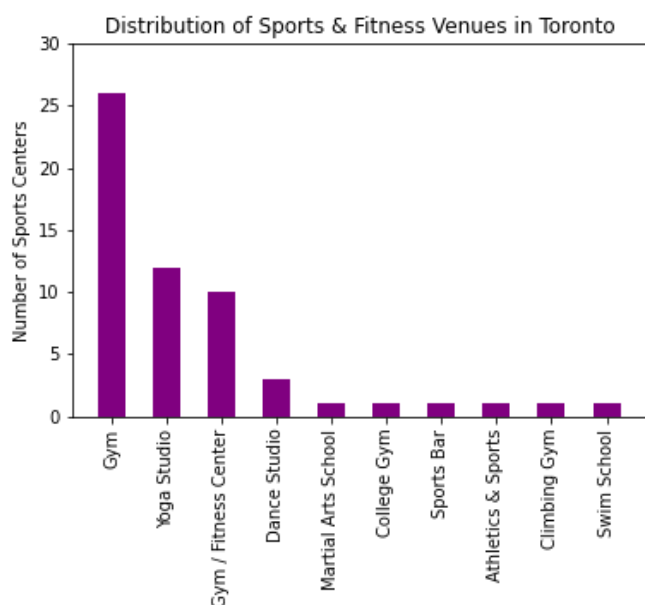
3.2 Intermediate Observations

Here is the number of venues by types of Sports & Fitness provision (arranged in descending order):

Venue Category	
Gym	26
Yoga Studio	12
Gym / Fitness Center	10
Dance Studio	3
Martial Arts School	1
College Gym	1
Sports Bar	1
Athletics & Sports	1
Climbing Gym	1
Swim School	1

We can observe these patterns:

- There are 26 “Gym” altogether – the most common types of venue in Toronto.
- After “Gym”, “Yoga Studio” and “Gym/ Fitness Center” follow closely, 12 and 10 respectively.
- There are very few “Dance Studio” and all remaining types.



4. Results

Now, we have a clearer picture of distribution of "Sports & Fitness" centres in Toronto. To make the data more readable, we will group all below categories (with minor occurrence) into a broader groups: 'Other Sports Center':

- Dance Studio
- Martial Arts School
- College Gym
- Sports Bar
- Athletics & Sports
- Climbing Gym
- Swim School

Venue Re-Cat	
Gym	26
Yoga Studio	12
Gym / Fitness Center	10
Other Sports Center	9

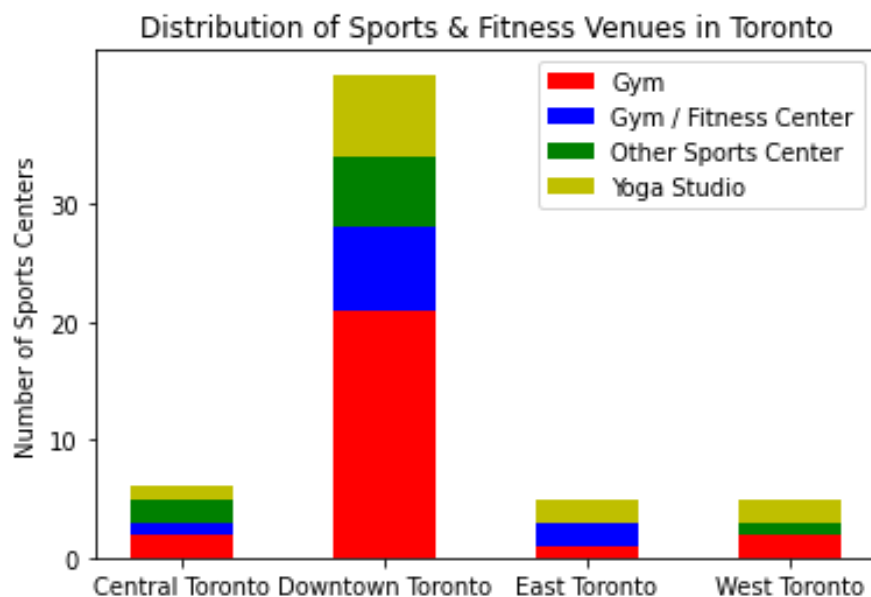
Next, we use pivot table to read data by borough.

Borough	Central Toronto	Downtown Toronto	East Toronto	West Toronto
Venue Re-Cat				
Gym	2.0	21.0	1.0	2.0
Gym / Fitness Center	1.0	7.0	2.0	0.0
Other Sports Center	2.0	6.0	0.0	1.0
Yoga Studio	1.0	7.0	2.0	2.0

Altogether, we can make a few observations already:

- Downtown Toronto accommodates the majority of "Sports & Fitness", 41. It is more than total number of "Sports & Fitness" centres in 3 other key areas combined.
- "Gym" is the most common type of "Sports & Fitness" whereas they are most often in Downtown.
- Compared to "Gym", all other types of "Sports & Fitness" is less common in other areas.
- There are no "Other Sports Center" in East Toronto. Also, "Gym/ Fitness" is absent in West Toronto.

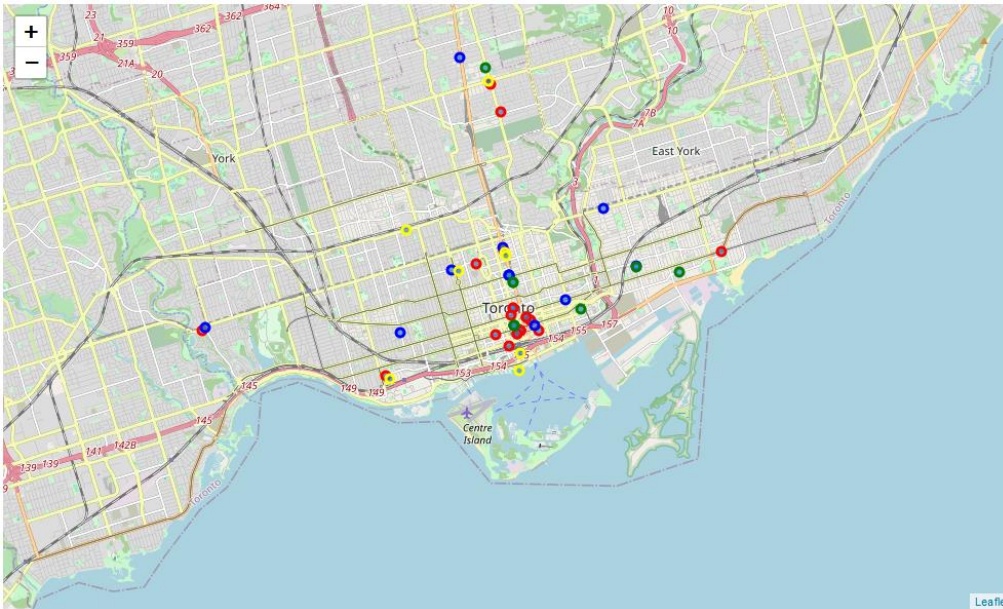
Then, we shall dissent number and distribution of "Sports & Fitness" centres by key boroughs in Toronto on a stacked bar plot



Let's map all Sports & Fitness venues by different colors on Toronto. It will give us a clearer picture of their distribution and concentration by areas.

- Gym (red)
- Gym / Fitness Center (blue)
- Other Sports Center (green)
- Yoga Studio (yellow)

Here comes the Toronto Map again with Sports & Fitness venues marked by different colors according to types.



5. Discussion

By now, we have gathered sufficient information to answer the business objectives.

Areas with the most and least sports service centres:

Downtown Toronto is, obviously, the areas most concentrated with “Sports & Fitness” venues. Other areas (namely, East, West and Central) contain just a few venues.

Also, these sport venues are heavily skewed towards “Gym”, distantly followed by “Yoga Studio” and “Fitness”. Other specialty (e.g. Martial Arts, Climbing) are lacking.

Discover potentials for provision of “Sports & Fitness Services”:

- From the map, we see a lot of “Sports & Fitness” venues are concentrated in city center, “Downtown”. The peripheral of city areas are less concentrated. As competition in Downtown areas is getting more and more fierce, it is time to consider put new “Sports and Fitness” venues in other areas where competition is still low.
- We have an impression that specialty centers (e.g. “Yoga Studio” and “Fitness”) have potential, given there are just very few in the city currently. As consumers are more aware of these sports centers, they are likely to have higher demand of such centers. Service providers are advised to plan setting up these centres in the future and gather capital for the future.

6. Conclusion

Like we what have set up in background, “Sports & Fitness” industry become very important to consumers nowadays. It is not just limited to a niche portion of elite citizens but all residents in general. In this analysis, we identify a few potentials for consideration for service providers, for example:

- What types of “Sports & Fitness” venues are missing / lagging behind in quantity?
- How are “Sports & Fitness” venues distributed in Toronto?

These are important information for service providers. However, given the ever-changing situation in Toronto, it is vital to keep monitoring the latest data. Therefore, data analysis become the best tool in this exercise, as we re-run

the analysis with the same codes and execution in regular time interval without much hassles. We would recommend service providers to continuously keep track of changing distribution of “Sports & Fitness” (e.g. every quarter) in the coming future.