

Capstone Project

Introduction

What is the proper location to open a new gym in Toronto area?

Purpose of this project is to help a businessman that wants to expand his business around Toronto area, to understand what is the better spot to do so.

Target Audience

Every businessman who are thinking to expand around Toronto.

Data

We are going to handle geospatial data for this exact project, where their source came from Wikipedia.
(wikipedia.org/wiki/List_of_postal_codes_of_Canada)

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

Methodology

Data Preparation

After removing any NA value from the initial dataset, through one hot encoding and filtering of venue categories, we were able to gather the data on several gyms and group them together by Neighborhood.

Out[43]:

	Yoga Studio	Accessories Store	Afghan Restaurant	Airport	Airport Food Court	Airport Gate	Airport Lounge	Airport Service	Airport Terminal	American Restaurant	...	Train Station	Vegetarian / Vegan Restaurant	Video Game Store	Video Store	Vietnamese Restaurant	Warel
0	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	...	0	0	0	0	0	0

5 rows x 269 columns

In [44]:

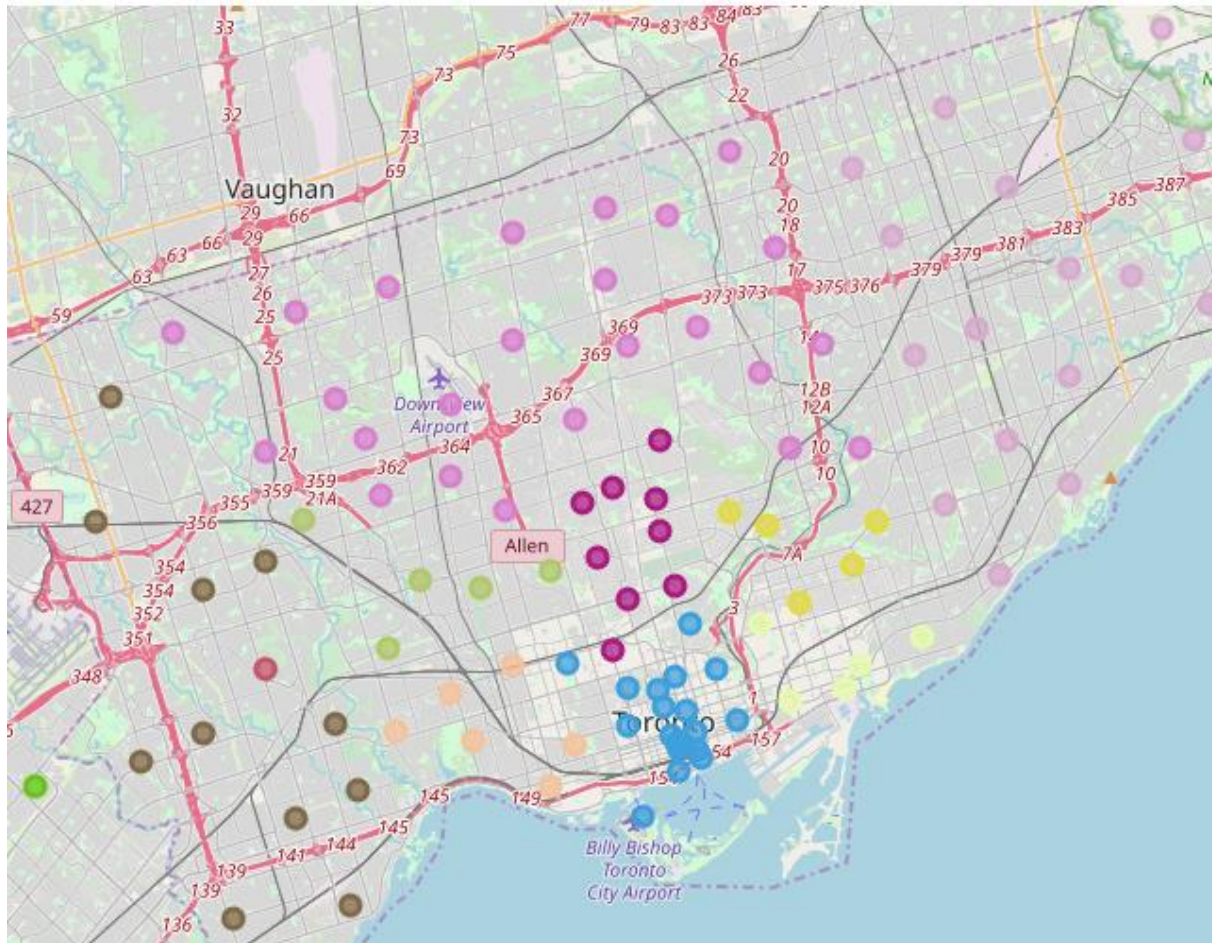
```
toronto_grouped = toronto_onehot.groupby('Neighborhood').mean().reset_index()
toronto_grouped
```

Out[44]:

	Neighborhood	Yoga Studio	Accessories Store	Afghan Restaurant	Airport	Airport Food Court	Airport Gate	Airport Lounge	Airport Service	Airport Terminal	...	Train Station	Vegetarian / Vegan Restaurant	Video Game Store	Video Store
0	Adelaide, King, Richmond	0.000000	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.0	0.020000	0.00	0.000000
1	Agincourt	0.000000	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.0	0.000000	0.00	0.000000
2	Agincourt North, L'Amoreaux East, Milliken, St...	0.000000	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.0	0.000000	0.00	0.000000
3	Albion Gardens, Beaumont Heights, Humbergate, ...	0.000000	0.0	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	...	0.0	0.000000	0.00	0.000000

Clustering

Using K-Means we managed to group several neighborhoods based on their characteristics. As we can see K-Means did a great job.



A detailed map of the Greater Toronto Area (GTA) showing bus routes and stations. The map includes labels for major locations such as Vaughan, Downsview Airport, Allen, Toronto, and Billy Bishop Toronto City Airport. Numerous bus routes are indicated by colored lines (red, blue, green, yellow) and numbers. Stations are marked with colored circles (blue, red, green, yellow). The map also shows major roads, highways, and geographical features like Lake Ontario.

Conclusion

Many gyms are in areas that belong to first two clusters. So, it would be wise to think as an opportunity to search for proper areas that belong to cluster 3,4,5. Specifically, if we search the map, area near Rathburn road and Anglesey boulevard seems a nice spot.

