

Where the Yoga Studios are in Toronto

Coursera - Data Science Capstone Project



Why this project matters?

- Yoga has become more mainstream over the past few years and more studios are popping up in different cities.
- This project attempts to answer where a studio should be placed in Toronto:
 - Does the neighborhood matter?
 - Do the venues nearby matter?



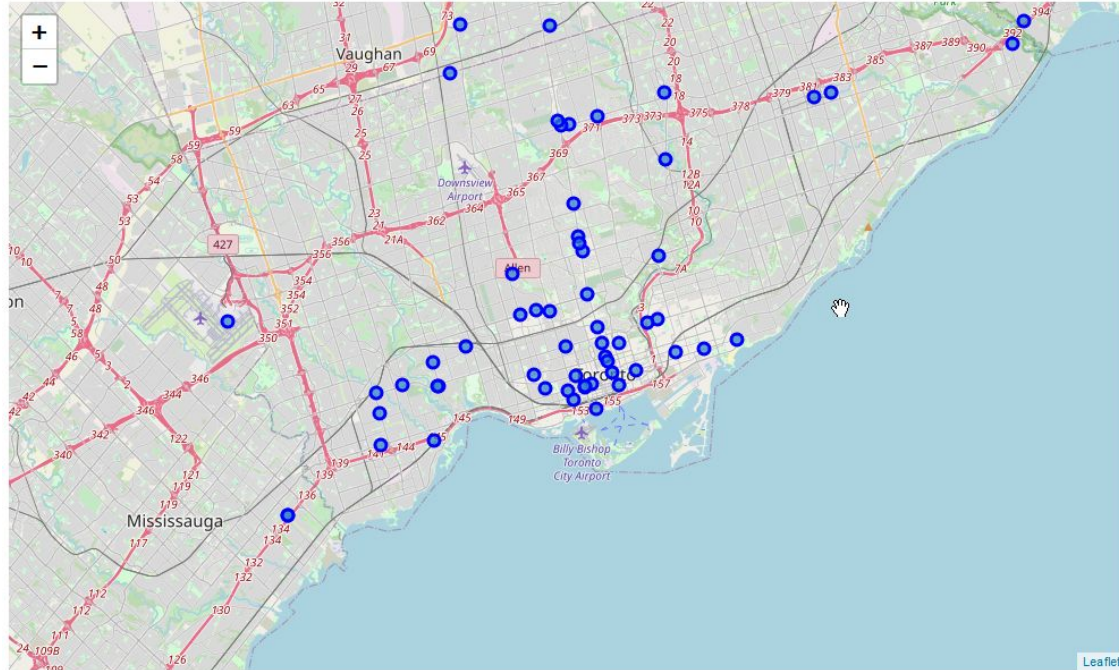
Data Acquisition and Cleaning

For this project, we would need 3 main types of data:

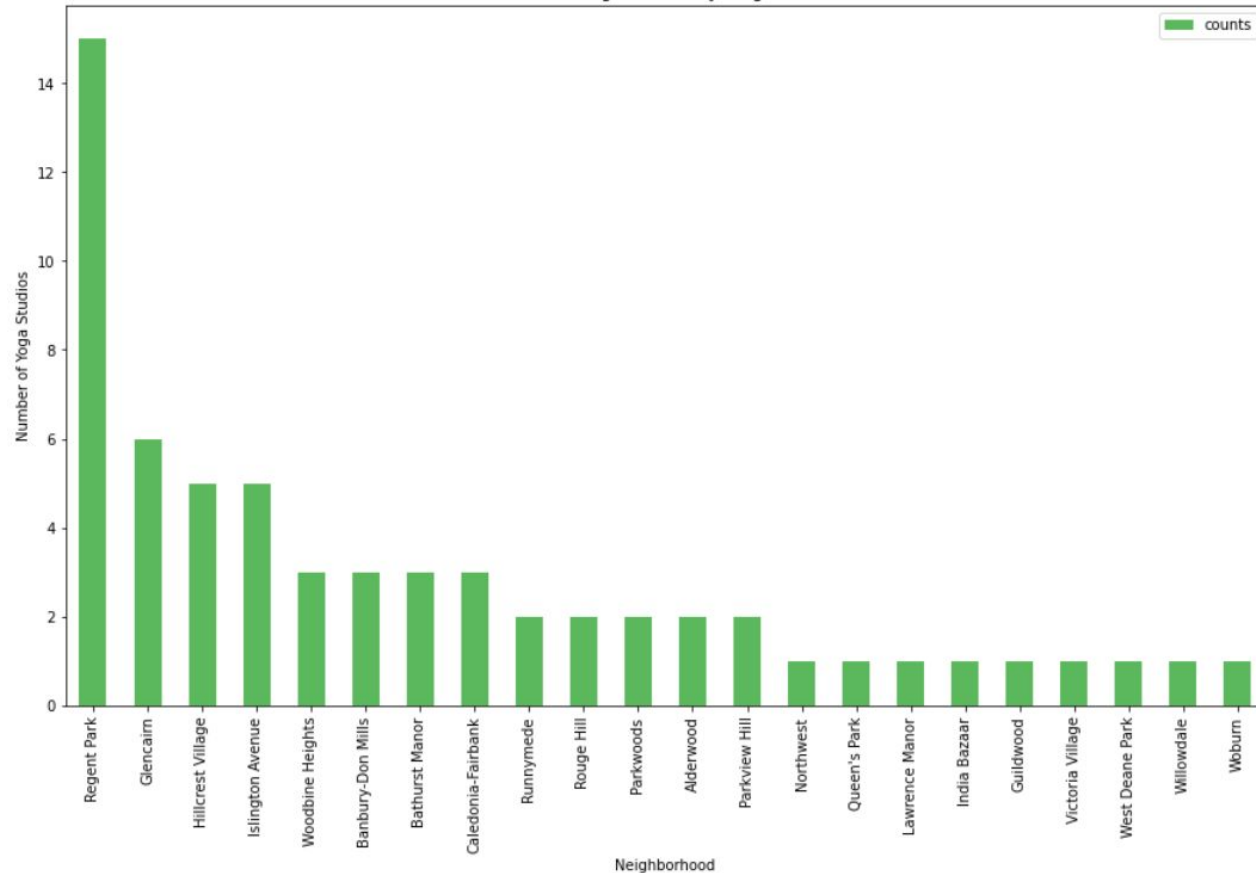
1. Neighborhood data containing the neighborhood names, postal code, latitude and longitude
 - a. Acquisition: fortunately, we already have this data from the previous assignment.
 - b. Cleaning: merging the dataframe containing a list of the neighborhoods in Toronto and the dataframe containing the geographical coordinates of each neighborhood
2. Yoga Studios in each neighborhood:
 - a. Acquisition: obtained by passing each neighborhood latitude and longitude data in the Foursquare API
 - b. Cleaning: removing empty results and only keeping the columns that are needed
3. Venues within a 200m radius of each yoga studio
 - a. Acquisition: Foursquare API
 - b. Cleaning: removing the venues that are tagged as Yoga studio

Data Exploration

A visual of the yoga venues returned from the Foursquare API shows that the majority of the venues are located along **subway lines** and around **downtown Toronto**.

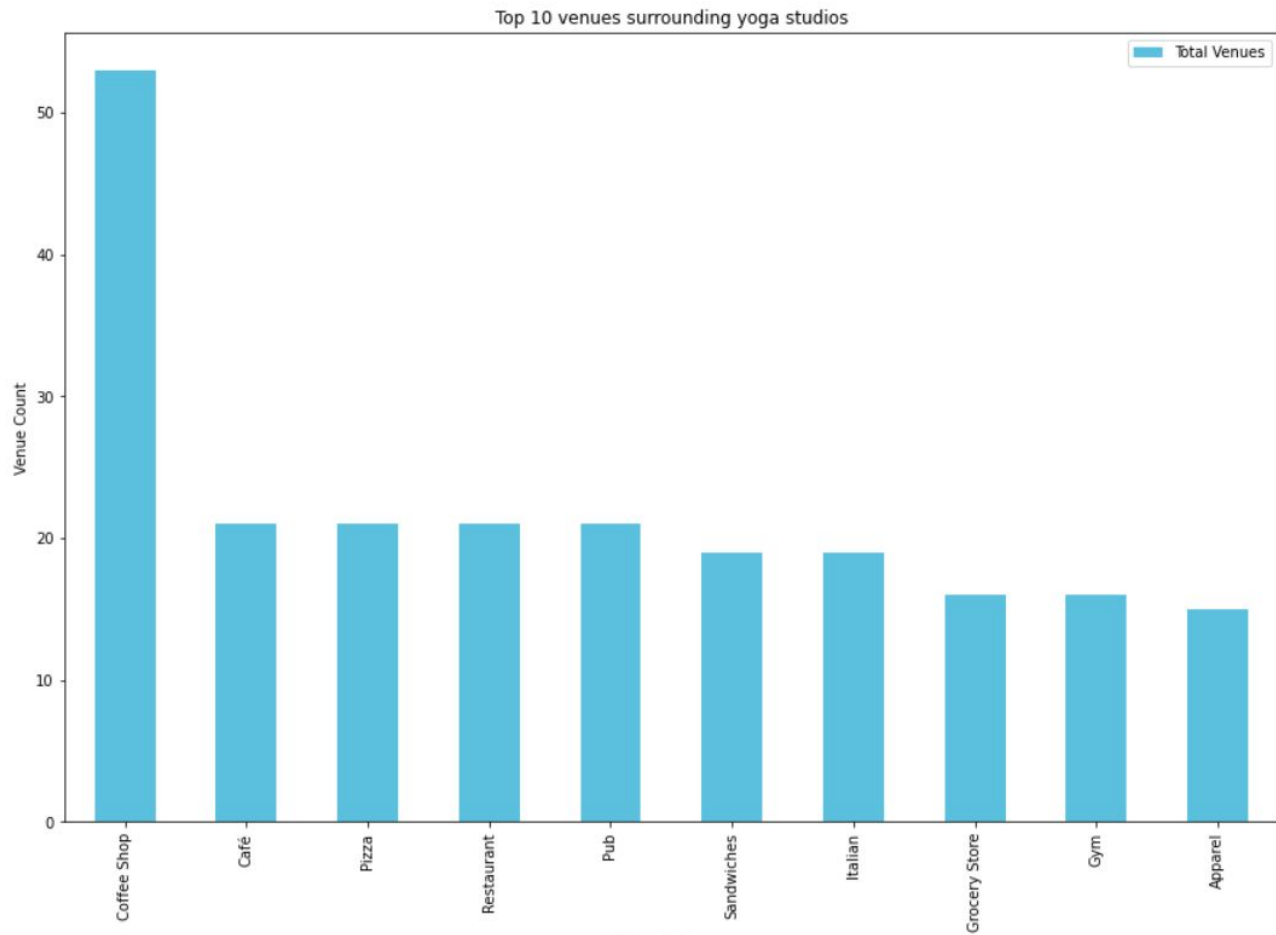


Number of Yoga Studios by Neighborhood



Regent Park has the most yoga studios of all the neighborhoods

Neighborhood	counts
Regent Park	15
Glencairn	6
Hillcrest Village	5
Islington Avenue	5
Woodbine Heights	3



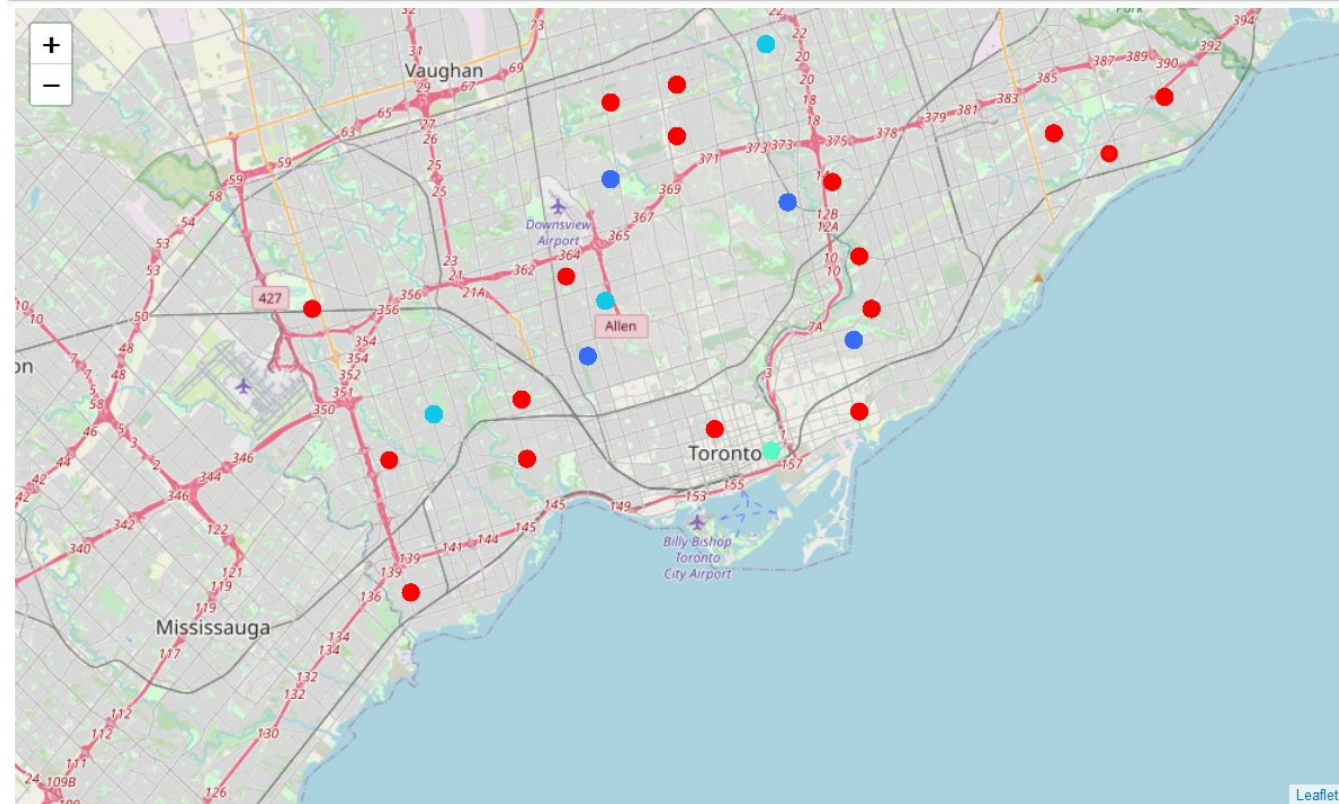
Interestingly, the majority of yoga studios are surrounded by coffee shops (200m radius)


Total Venues	
Venue Category	
Coffee Shop	53
Café	21
Pizza	21
Restaurant	21
Pub	21
Sandwiches	19
Italian	19

Clustering

Using a k-mean clustering to analyze the top venues for each neighborhood cluster.

K = 8 was used and it shows 6 distinct clusters in the Greater Toronto Area





Neighborhood Cluster Label	Top Venue in that Category
0	Restaurant
1	Vegetarian/Vegan
2	Spa
3	Fast Food
4	Movie Theater
5	Coffee Shop
6	Coffee Shop
7	Pub



Conclusion

- Many Yoga studios are set around coffee shops. It's worth pointing that there may or may not be a correlation or association between coffee shops and yoga studios. This warrants further study and is out of scope due to the lack of data.
- Regent Parks has the highest number of yoga studio.
- In the cluster analysis, it has been determined that these venues are also present near yoga studios: Coffee Shops, Vegan/Vegetarian and Pubs yielded the highest count of yoga studios