

Capstone Project - The Battle of Neighborhoods

Introduction/Business Problem:

Opening a tutoring point is all about location, location, location. However, not every tutoring point is suitable for every location, and vice versa. It comes down to a combination of tutoring point style, target audience, your competitors. If you can define your tutoring point type and identify your target demographic and its most populated areas, you'll be well on your way to choosing a tutoring point location that sets your business up for success. There's a lot of work, planning and preparation that goes into opening a tutoring point. Here are a few tips on how to open a tutoring point in Toronto and some steps you'll need to take in order to achieve your goal.

The first thing you need to decide is what tutoring point you want to open and the students it will hold. Having this set in stone will make the other steps a lot

easier. To make things a little easier, here are three categories you need to consider:

Location and Premises:

There are many factors that will contribute to where you decide your premises to be. It can be challenging to find a venue that will factor in all of your conditions, so it is likely you will have to compromise on a few things. However, here are the main factors you should consider when finding premises for your new tutoring point:

- Location: How accessible is the location to students? Are there many competitors nearby?
- Building type: What type of building would you like your tutoring point to be in?
- Target market: Make sure it is in an area where your target market is also.
- Cost: Your ideal location may prove to be a little costly. Make sure that there is a realistic price point.

Your Target Students:

In order to open a tutoring point in Toronto, you need to be familiar with exactly who you are aiming to bring into your venue. Are you targeting audience from Indian,

Pakistani, Bangladeshi and Chinese origin? You need to also be aware of the age group, the location, the amount of money they are willing to spend. Subjects they are interested in.

Who are your competitors?

As well as familiarising yourself with your audience, you must do so with your competitors. Look at similar businesses in your area and well-renowned tutoring points of

the same category. Consider the ways in which they market themselves and the service they provide. What works and what doesn't? You'll need to be able to compete with these establishments, so take on board what they're already doing and better it.

Data Section:

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

Will use York, North York, East York dataset which we scrapped from wikipedia on Week 3. Dataset consisting of latitude and longitude, postal codes.

Foursquare API Data:

We will need data about different venues in different neighborhoods of that specific borough.

In order to gain that information we will use "Foursquare" locational information. Foursquare is a location data provider with information about all manner of venues and events within an area of interest. Such information includes venue names, locations, menus and even photos. As such, the foursquare location platform will be used as the sole data source since all the stated required information can be obtained through the API.

After finding the list of neighborhoods, we then connect to the Foursquare API to gather information about venues inside each and every neighborhood. For each neighborhood, we have chosen the radius to be 100 meter.

The data retrieved from Foursquare contained information of venues within a specified distance of the longitude and latitude of the postcodes. The information obtained per venue as follows:

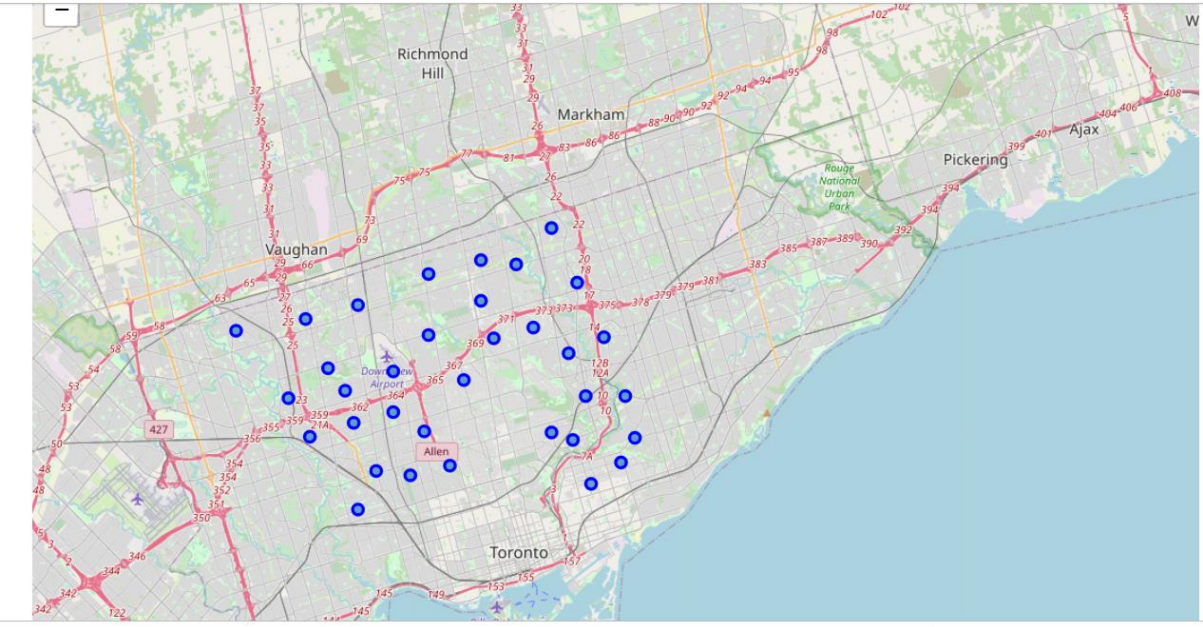
1. Neighborhood
2. Neighborhood Latitude
3. Neighborhood Longitude
4. Venue
5. Name of the venue e.g. the name of a store or restaurant
6. Venue Latitude
7. Venue Longitude
8. Venue Category

Data of School Zones is taken from the following website.

<https://open.toronto.ca/dataset/school-locations-all-types/>

Count of Schools in each postal zone is determined in York, North York and East York to place tutorial point in most school zone.

Map of York, NY, EY



Methodology Section

Clustering Approach:

To compare the similarities of two cities, we decided to explore neighborhoods, segment them, and group them into clusters to find similar neighborhoods in a big city like Toronto. To be able to do that, we need to cluster data which is a form of unsupervised machine learning: k-means clustering algorithm.

Using K-Means Clustering Approach | Most Common Venue

[350]:

	Neighborhoods	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	Bathurst Manor, Wilson Heights, Downsview North	Coffee Shop	Bank	Pharmacy	Supermarket	Fried Chicken Joint	Pizza Place	Deli / Bodega	Restaurant	Sandwich Place	Bridal Shop
1	Bayview Village	Chinese Restaurant	Café	Bank	Japanese Restaurant	Yoga Studio	Electronics Store	Cosmetics Shop	Curling Ice	Deli / Bodega	Department Store
2	Bedford Park, Lawrence Manor East	Sandwich Place	Italian Restaurant	Coffee Shop	Restaurant	Greek Restaurant	Grocery Store	Indian Restaurant	Juice Bar	Liquor Store	Comfort Food Restaurant
3	Caledonia-Fairbanks	Park	Women's Store	Pool	Yoga Studio	Dog Run	Construction & Landscaping	Convenience Store	Cosmetics Shop	Curling Ice	Deli / Bodega
4	Del Ray, Mount Dennis, Keelsdale and Silverthorn	Turkish Restaurant	Sandwich Place	Bar	Discount Store	Yoga Studio	Dog Run	Construction & Landscaping	Convenience Store	Cosmetics Shop	Curling Ice
5	Don Mills	Restaurant	Coffee Shop	Japanese Restaurant	Asian Restaurant	Gym	Beer Store	Caribbean Restaurant	Dim Sum Restaurant	Italian Restaurant	Discount Store
6	Downsview	Grocery Store	Park	Gym / Fitness Center	Shopping Mall	Snack Place	Construction & Landscaping	Hotel	Baseball Field	Bank	Liquor Store
7	East Toronto, Broadview North (Old East York)	Park	Convenience Store	Metro Station	Yoga Studio	Gift Shop	Construction & Landscaping	Cosmetics Shop	Curling Ice	Deli / Bodega	Department Store
8	Fairview, Henry Farm, Oriole	Clothing Store	Coffee Shop	Fast Food Restaurant	Restaurant	Japanese Restaurant	Cosmetics Shop	Convenience Store	Tea Room	Electronics Store	Bank
9	Glencairn	Park	Pub	Sushi Restaurant	Japanese Restaurant	Yoga Studio	Distribution Center	Construction & Landscaping	Convenience Store	Cosmetics Shop	Curling Ice

Most Common Venues near Neighborhood | Using Clustering

[248]:

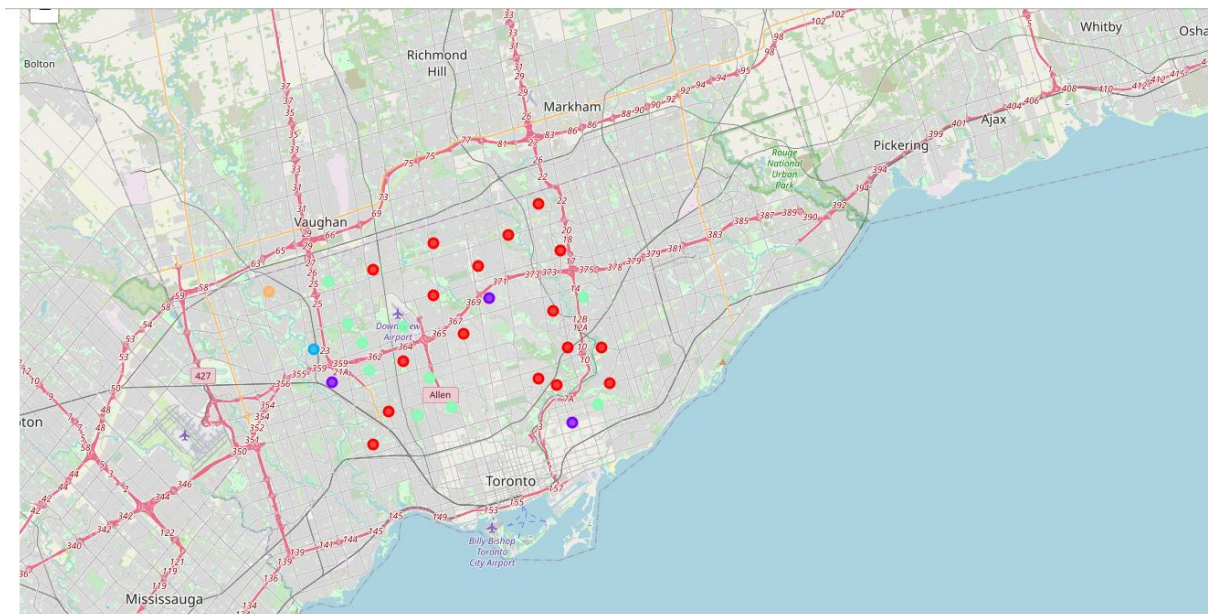
	PostalCode	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	M6C	York	Humewood-Cedarvale	43.693781	-79.428191	3.0	Park	Hockey Arena	Trail	Field	Yoga Studio	Department Store	Discount Store	Diner	Dim Sum Restaurant	Dessert Shop
1	M6E	York	Caledonia-Fairbanks	43.689026	-79.453512	3.0	Park	Women's Store	Pool	Yoga Studio	Dog Run	Construction & Landscaping	Convenience Store	Cosmetics Shop	Curling Ice	Deli / Bodega
2	M6M	York	Del Ray, Mount Dennis, Keelsdale and Silverthorn	43.691116	-79.476013	0.0	Turkish Restaurant	Sandwich Place	Bar	Discount Store	Yoga Studio	Dog Run	Construction & Landscaping	Convenience Store	Cosmetics Shop	Curling Ice
3	M6N	York	Runnymede, The Junction North	43.673185	-79.487262	0.0	Grocery Store	Convenience Store	Breakfast Spot	Pizza Place	Yoga Studio	Dog Run	Construction & Landscaping	Cosmetics Shop	Curling Ice	Deli / Bodega
4	M9N	York	Weston	43.706876	-79.518188	1.0	Convenience Store	Park	Yoga Studio	Golf Course	Construction & Landscaping	Cosmetics Shop	Curling Ice	Deli / Bodega	Department Store	Dessert Shop
5	M2H	North York	Hillcrest Village	43.803762	-79.363452	0.0	Golf Course	Dog Run	Pool	Mediterranean Restaurant	Construction & Landscaping	Convenience Store	Cosmetics Shop	Curling Ice	Deli / Bodega	Department Store
6	M2J	North York	Fairview, Henry Farm, Oriole	43.778517	-79.346556	0.0	Clothing Store	Coffee Shop	Fast Food Restaurant	Restaurant	Japanese Restaurant	Cosmetics Shop	Convenience Store	Tea Room	Electronics Store	Bank
7	M2K	North York	Bayview Village	43.786947	-79.385975	0.0	Chinese Restaurant	Café	Bank	Japanese Restaurant	Yoga Studio	Electronics Store	Cosmetics Shop	Curling Ice	Deli / Bodega	Department Store
10	M2N	North York	Willowdale, Willowdale East	43.770120	-79.408493	0.0	Ramen Restaurant	Restaurant	Sandwich Place	Pizza Place	Coffee Shop	Café	Sushi Restaurant	Shopping Mall	Steakhouse	Bubble Tea Shop

Work Flow:

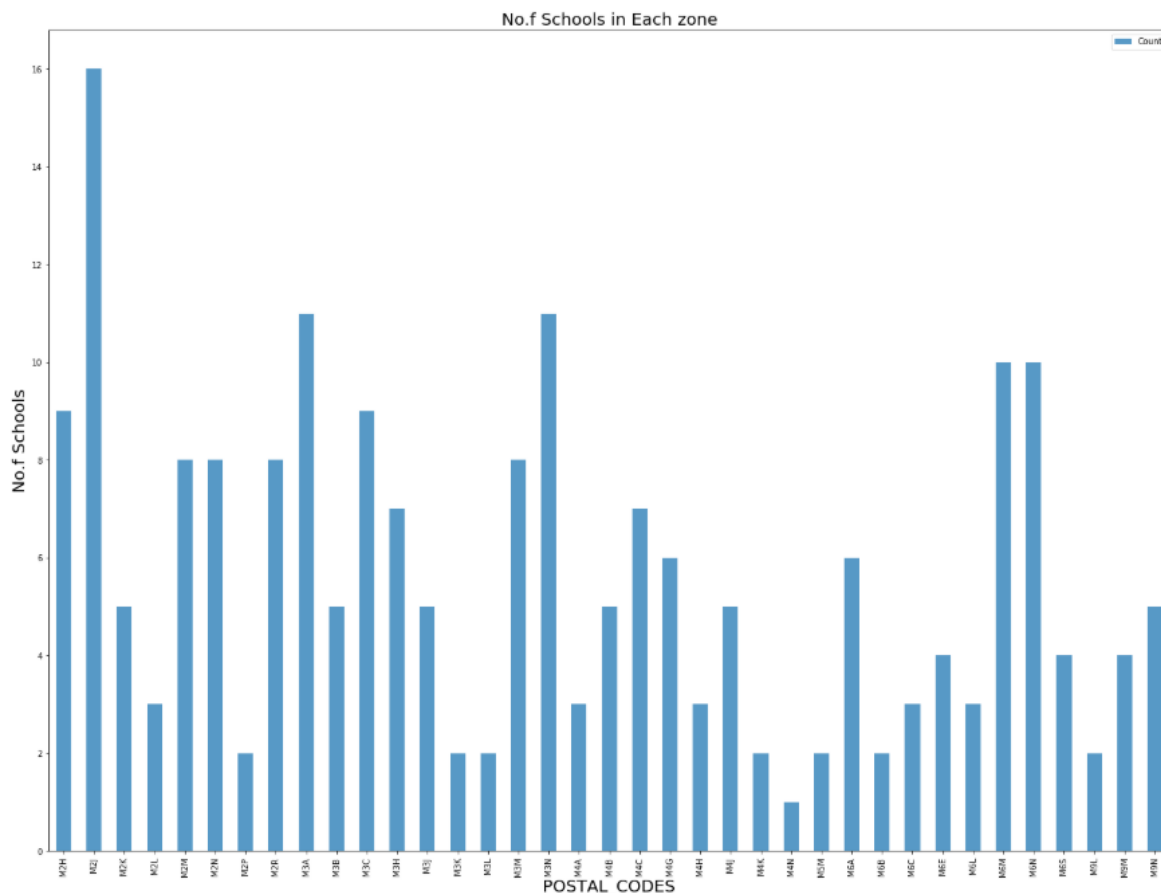
Using credentials of Foursquare API features of near-by places of the neighborhoods would be mined. Due to http request limitations the number of places per neighborhood parameter would reasonably be set to 100 and the radius parameter would be set to 500.

Results Section

Map of Clusters in York, NY, EY



No.of Schools by Clusters in each postal zone in York, NY, EY



Discussion Section

The major purpose of this project, is to suggest a better neighborhood in a York, North York and East York of the city for the person who is planning to start a Tutorial point. Places that have good feasibility to coffee shops, small pizza, breakfast points, and feasibility to schools are much preferred. But it should not be noisy or disturbing at the sametime.

Conclusion Section

In this Capstone project, using k-means cluster algorithm I separated the neighborhood into 10(Ten) different clusters and for 103 different latitude and logitude from dataset, which have very-similar neighborhoods around them. Of them Zones with Parks, Coffee Shops, and Schools are considered the apt places to start a tutorial point. These will be more accessible to Students. These include Fairview, Henry Farm, Oriole, Downsview, Runnymede, The Junction North are considered as good places.

This project has shown me a practical application to resolve a real situation that has impacting personal and financial impact using Data Science tools. The mapping with Folium is a very powerful technique to consolidate information and make the analysis and decision better with confidence.

Future Works:

This Capstone project can be continued for making it more precise in terms to find best place to start tutorial point in York. Best means on the basis of all required things around and also in terms of school areas, education levels of people around.

Libraries Which are Used to Develop the Project:

Pandas: For creating and manipulating dataframes.

Folium: Python visualization library would be used to visualize the neighborhoods cluster distribution of using interactive leaflet map.

Scikit Learn: For importing k-means clustering.

JSON: Library to handle JSON files.

XML: To separate data from presentation and XML stores data in plain text format.

Geocoder: To retrieve Location Data.

Beautiful Soup and Requests: To scrap and library to handle http requests.

Matplotlib: Python Plotting Module.