IBM-COURSERA DATA SCIENCE CAPSTONE PROJECT

OPENING A VEGAN RESTAURANT IN TORONTO

1. INTRODUCTION

1.1 Background

Toronto is the capital city of the Canadian province of Ontario. It is the most populous city in Canada and the fourth largest city in North America. It is made up of the former cities of Toronto, North York, Scarborough, York, and Etobicoke, and the former borough of East York. The city is home to a large immigrant population and is an international center of business, finance, arts, and culture.

It is recognized as one of the most multicultural and cosmopolitan cities in the world

The restaurant and food service industry are one of the fastest-growing industries in Canada today. The annual sales generated by the restaurant industry in Ontario are $33 Billion and there are 9,1 Million visits to restaurants every day (Source: https://www.restaurantscanada.org/resources/#infographics)

As a cosmopolitan and diverse city the restaurant offer is very wide there are a lot of different cuisines. Italian, Japanese, Indian, American, Mediterranean…and also some which offer vegan cuisine.
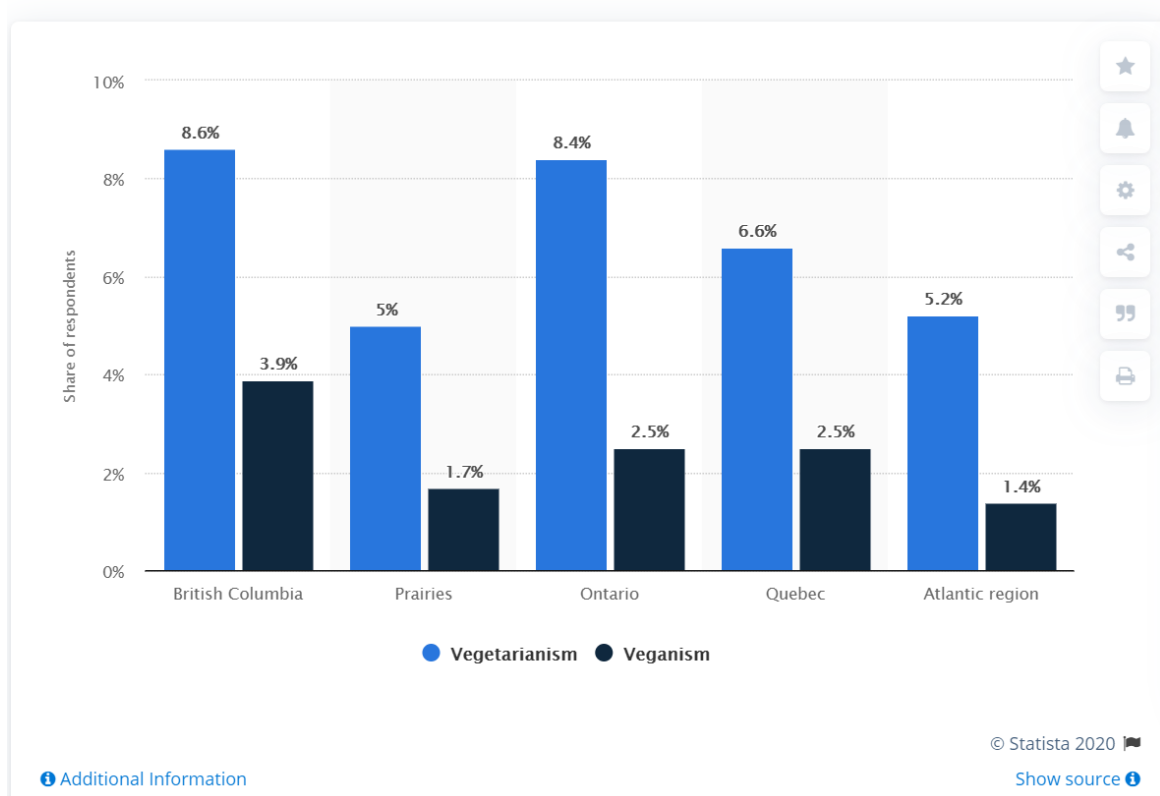
Veganism has emerged as one of the top food and lifestyle trends in the past few years. In 2020, the global market for veggie products (vegan, flexitarian, and vegetarian) will reach 5 billion dollars and exponential growth is expected in successive years.

The main reasons for becoming a veggie are sustainability, health, and respect for animals.

Canada is home to approximately 850,000 vegans and there are nearly 2.3 million vegetarians (source: Statista) That number is increasing, and it's estimated that the number of Canadians going meatless or eating less meat could exceed 10 million by 2025.

The following chart shows the percentage of vegetarians and vegans in Canada as of March 2018, by region (source Statista. https://www.statista.com/topics/3262/vegan-vegetarian-diets-in-canada).

Ontario is the second region with the highest percentage of vegetarians and vegans

Chart showing share of respondents following a vegetarian or vegan diet by Canadian region.

Region	Vegetarianism	Veganism
British Columbia	8.6%	3.9%
Prairies	5%	1.7%
Ontario	8.4%	2.5%
Quebec	6.6%	2.5%
Atlantic region	5.2%	1.4%

● Vegetarianism ● Veganism

❶ Additional Information   Show source ❶

It seems like veganism is catching on, with more vegan restaurants, more vegan options, and more grocery stores carrying vegan food.

Moreover, people, especially in cosmopolitan cities like Toronto, are more open to trying and exploring vegetarian or vegan restaurants even if they do not follow a strictly vegan or vegetarian diet.

1. Business problem

The objective of this project is to determine which might be the 'best' neighborhood in Toronto to open a vegan restaurant.

Location is one of the most important decisions to make when you want to open a new restaurant, and even more when it comes to a niche business like a vegan restaurant.

Several factors would affect the success of a new restaurant that are directly related to the location such as visibility and accessibility (look for spots that have good street visibility and are easily accessible), demographics (you have to consider a location frequented by your potential client target), or competition.

Moreover, veganism is often associated with a healthy and conscious lifestyle, therefore an ideal spot might be near a successful yoga studio or health shop.

1.3 Stakeholders

This study would be extremely useful for any entrepreneur that wants to open a vegan restaurant in the city of Toronto.

The analysis will provide vital information about the more suitable locations for a vegan restaurant that can be used by this target audience.

2. DATA DESCRIPTION

The main data that we use for this project is:

- o List of all boroughs in Toronto and their corresponding neighborhoods
- o Coordinates of all neighborhoods
- o Venues information for each neighborhood

Following the objective of the project, the factors that will help to make a final decision (where is better to open a vegan restaurant) are:

- o Number of existing Vegan restaurants in the neighborhood
- o Number of existing Yoga Studio in the neighborhood
- o Most common venues of each neighborhood

Data sources and acquisition

We make use of a few data sources to get the data required for this project.

1. Wikipedia

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

Wikipedia provides almost all the information about the neighborhoods in Toronto. It includes the postal code, borough, and the name of each neighborhood.

Since the data is not in a format that is suitable for analysis, scraping of the data is done from this site using Beautiful Soup Library. We put the data into a data frame.

	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

2. Geospatial data.

https://cocl.us/Geospatial_data

The second source of data was given by the Coursera course and provides us with the geographical coordinates of all neighborhoods and venues. The file is in CSV format and we can read it using Pandas Library.

	PostalCode	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

3. Foursquare API.

We will get the various location-related data, like the kinds of places in a particular neighborhood, using Foursquare API. This data will include the type of shops, restaurants, cafes, etc. in each neighborhood. We will make calls to Foursquare API using our credentials to acquire these location-related data.

3. METHODOLOGY

3.1 Data Cleaning

After all the data was collected and put into data frames, we need to clean and merge the data to be able to start our analysis.

When we get the data from Wikipedia some cleaning process is required:

1. Some boroughs are shown as "Not assigned". We remove those lines. Only the cells that have an assigned borough will be processed.

2. All neighborhoods with the same postal code will be grouped. For example, M5A is listed twice and has two neighborhoods: Harbourfront and Regent Park. We will combine these two rows into one with the neighborhoods separated by a comma.

3. We will reorder the Columns so the 'PostalCode' appears first for easier readability

3. For a borough with neighborhood = 'Not assigned', the neighborhood will be the same as the borough.

After all these implementations, our data frame remains like shown below:

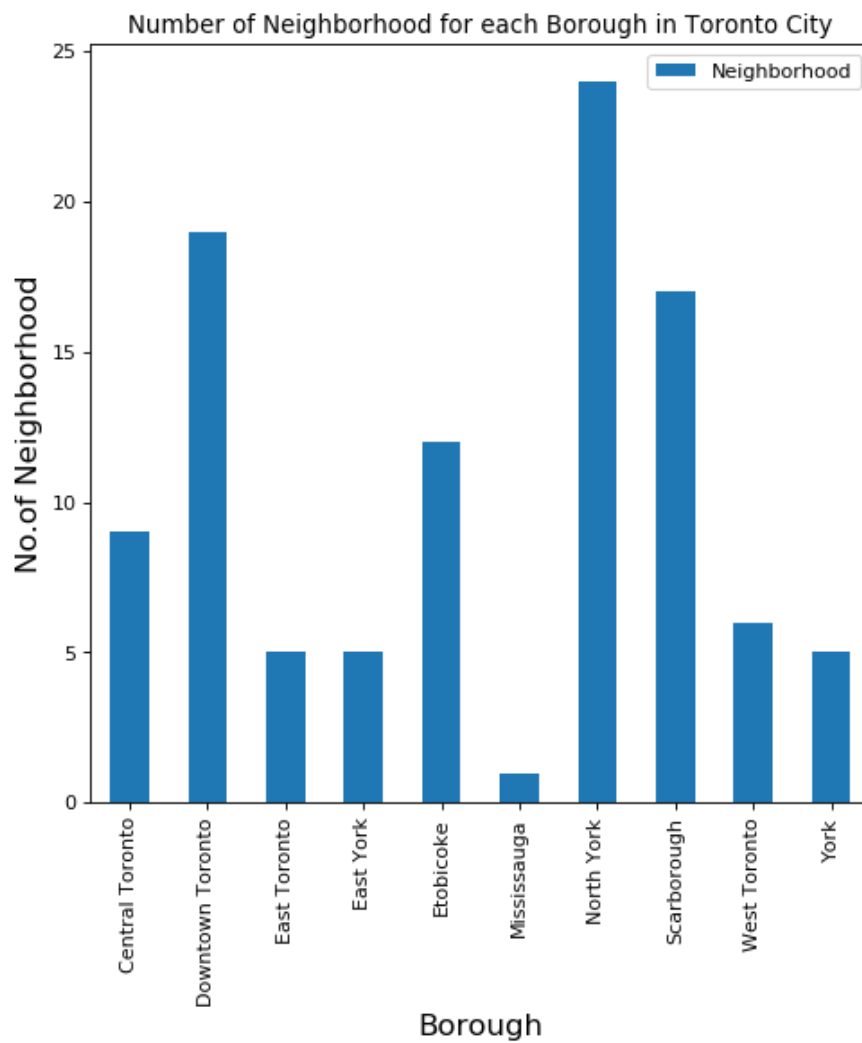|	PostalCode	Borough	Neighborhood
0	M4N	Central Toronto	Lawrence Park
1	M4P	Central Toronto	Davisville North
2	M4R	Central Toronto	North Toronto West, Lawrence Park
3	M4S	Central Toronto	Davisville
4	M4T	Central Toronto	Moore Park, Summerhill East

Now we will merge this data frame with the one we got from the Geocoder package with the geographical coordinates of all neighborhoods. We will merge the two tables based on Postal Code getting the data frame shown below.

|	PostalCode	Borough	Neighborhood	Latitude	Longitude
0	M4N	Central Toronto	Lawrence Park	43.728020	-79.388790
1	M4P	Central Toronto	Davisville North	43.712751	-79.390197
2	M4R	Central Toronto	North Toronto West, Lawrence Park	43.715383	-79.405678
3	M4S	Central Toronto	Davisville	43.704324	-79.388790
4	M4T	Central Toronto	Moore Park, Summerhill East	43.689574	-79.383160

3.1 Data exploration

Now we have our data prepared to start analyzing and exploring it.

We will start comparing the number of neighborhoods in each borough. We will create a bar chart to easily visualize this.

Number of Neighborhood for each Borough in Toronto City

Using Folium we will create a map of Toronto with neighborhoods superimposed on top and color-coded each Neighborhood depending on what borough they are located in
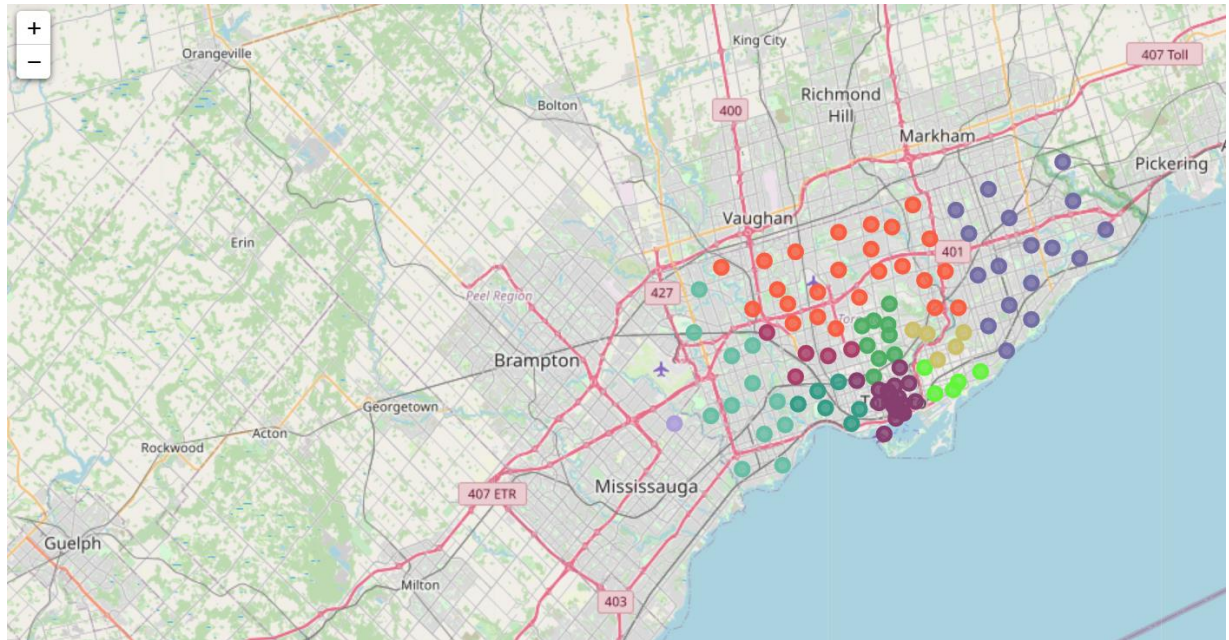
In this map, we can visualize the boroughs and neighborhoods inside.

Next, we will use the Foursquare API to get a list of all the Venues in Toronto and explore each neighborhood attending to this venue information.

First, we will get the venues for all neighborhoods in our dataset and then we will find out how many unique categories can be curated from all the returned venues.

We get a list of all the venues categories and find that there is a total of 273 unique categories of venues.

```
The Venue Categories are ['Park' 'Swim School' 'Bus Line' 'Breakfast Spot' 'Food & Drink Shop'
 'Department Store' 'Hotel' 'Sandwich Place' 'Gym / Fitness Center'
 'Dog Run' 'Dance Studio' 'Yoga Studio' 'Diner' 'Salon / Barbershop'
 'Clothing Store' 'Restaurant' 'Spa' 'Mexican Restaurant' 'Coffee Shop'
 'Chinese Restaurant' 'Fast Food Restaurant' 'Sporting Goods Shop' 'Café'
 'Ice Cream Shop' 'Furniture / Home Store' 'Bagel Shop' 'Dessert Shop'
 'Indian Restaurant' 'Pizza Place' 'Seafood Restaurant'
 'Italian Restaurant' 'Sushi Restaurant' 'Gym' 'Thai Restaurant' 'Brewery'
 'Toy / Game Store' 'Greek Restaurant' 'Gas Station' 'Farmers Market'
 'Gourmet Shop' 'Pharmacy' 'Indoor Play Area' 'Trail' 'Playground'
 'Liquor Store' 'Supermarket' 'American Restaurant' 'Pub'
 'Fried Chicken Joint' 'Bank' 'Vietnamese Restaurant' 'Light Rail Station'
 'Garden' 'Music Venue' 'Jewelry Store' 'BBQ Joint'
 'Middle Eastern Restaurant' 'Burger Joint' 'Donut Shop' 'History Museum'
 'Japanese Restaurant' 'Bakery' 'General Entertainment' 'Butcher'
 'Gastropub' 'Pet Store' 'Deli / Bodega' 'Caribbean Restaurant'
 'Taiwanese Restaurant' 'Gift Shop' 'Market' 'Beer Store' 'Plaza'
 'Grocery Store' 'Snack Place' 'Theme Restaurant' 'Bubble Tea Shop'
 'Beer Bar' 'Ramen Restaurant' 'Juice Bar' 'Bookstore' 'Creperie'
 'Martial Arts School' "Men's Store" 'Gay Bar' 'Escape Room' 'Hobby Shop'
 'Ethiopian Restaurant' 'Steakhouse' 'Smoke Shop' 'Distribution Center'
 'Sake Bar' 'Health & Beauty Service' 'Theater' 'Mediterranean Restaurant'
 'Afghan Restaurant' 'Strip Club' 'Sculpture Garden' 'Historic Site'
 'Chocolate Shop' 'Performing Arts Venue' 'French Restaurant'
 'Event Space' 'Shoe Store' 'Art Gallery' 'Electronics Store'
 'Antique Shop' 'Comic Shop' 'Burrito Place' 'Shopping Mall'
 'New American Restaurant' 'Tanning Salon' 'Movie Theater'
 'Modern European Restaurant' 'Miscellaneous Shop' 'College Rec Center'
 'Cosmetics Shop' 'Lake' 'Tea Room' 'Lounge' 'Video Game Store' 'Wine Bar'
 'Other Great Outdoors' 'Lingerie Store' 'Poutine Place' 'Office'
 'Hookah Bar' 'Food Truck' 'Camera Store' 'Cocktail Bar'
 'Latin American Restaurant' 'Vegetarian / Vegan Restaurant' 'Fountain'
 'Tailor Shop' 'Fish Market' 'German Restaurant' 'Cheese Shop'
 'Comfort Food Restaurant' 'Asian Restaurant' 'Irish Pub'
 'Moroccan Restaurant' 'Belgian Restaurant' 'Kitchen Supply Store'
 'Bistro' 'Concert Hall' 'Museum' 'Basketball Stadium' 'Jazz Club' 'Beach'
 'Eastern European Restaurant' 'Poke Place' 'Art Museum'
 'Portuguese Restaurant' 'Falafel Restaurant' 'Discount Store'
 'Salad Place' 'Korean Restaurant' 'Speakeasy' 'Neighborhood' 'Food Court'
 'Monument / Landmark' 'Colombian Restaurant' 'General Travel'
 'Brazilian Restaurant' 'Gluten-free Restaurant' 'Bar' "Women's Store"
 'Noodle House' 'Building' 'Cupcake Shop' 'Nightclub' 'Skating Rink'
 'IT Services' 'Roof Deck' 'Aquarium' 'Train Station' 'Sports Bar'
 'Scenic Lookout' 'Baseball Stadium' 'Convenience Store'
 'Indie Movie Theater' 'Hotel Bar' 'Molecular Gastronomy Restaurant'
 'College Gym' 'College Arts Building' 'Arts & Crafts Store'
 'Organic Grocery' 'Dumpling Restaurant' 'Gaming Cafe' 'Record Shop'
 'Filipino Restaurant' 'Doner Restaurant' 'Massage Studio'
 'Dim Sum Restaurant' 'Hospital' 'Bed & Breakfast' 'Smoothie Shop'
 'Airport' 'Airport Lounge' 'Harbor / Marina' 'Airport Food Court'
 'Airport Gate' 'Plane' 'Boutique' 'Rental Car Location'
 'Airport Terminal' 'Airport Service' 'Boat or Ferry' 'Church'
 'Optical Shop' 'Hostel' 'Soup Place' 'Opera House' 'Candy Store'
 'Baby Store' 'Athletics & Sports' 'College Auditorium'
 'College Cafeteria' 'Health Food Store' 'Fruit & Vegetable Store'
 'Frozen Yogurt Shop' 'Fish & Chips Shop' 'Stationery Store'
 'Coworking Space' 'Skate Park' 'Garden Center' 'Auto Workshop'
 'Recording Studio' 'Intersection' 'Curling Ice' 'Bike Shop'
 'Warehouse Store' 'Pool' 'River' 'Baseball Field'
 'Construction & Landscaping' 'Wings Joint' 'Supplement Shop'
 'Hardware Store' 'Social Club' 'Kids Store' 'Thrift / Vintage Store'
 'Print Shop' 'Shopping Plaza' 'Drugstore' 'Golf Course'
 'Mobile Phone Shop' 'Luggage Store' 'Bus Station' 'Metro Station'
 'Bridal Shop' 'Home Service' 'Business Service' 'Hockey Arena'
 'Locksmith' 'Accessories Store' 'Medical Center' 'Korean BBQ Restaurant'
 'Hakka Restaurant' 'Soccer Field' 'Motel' 'College Stadium' 'Auto Garage'
 'Cuban Restaurant' 'Malay Restaurant' 'Southern / Soul Food Restaurant'
 'Climbing Gym' 'Stadium' 'Flea Market' 'Cajun / Creole Restaurant'
 'Field' 'Tennis Court' 'Turkish Restaurant']
```

Is there any Vegan/Vegetarian restaurant on this list? The answer is yes and getting this data is crucial for our analysis. We will find out that there is a total of 16 Vegan/Vegetarian restaurants all over Toronto.

EXPLORING THE NEIGHBORHOODS

Now we will start exploring each Neighborhood.

We will merge the Foursquare Venue data with the Neighborhood data. This will give us the nearest Venue for each of the Neighborhood.

	Neighborhood	NeighborhoodLatitude	NeighborhoodLongitude	VenueName	VenueLatitude	VenueLongitude	VenueCategory
0	Lawrence Park	43.728020	-79.388790	Lawrence Park Ravine	43.726963	-79.394382	Park
1	Lawrence Park	43.728020	-79.388790	Zodiac Swim School	43.728532	-79.382860	Swim School
2	Lawrence Park	43.728020	-79.388790	TTC Bus #162 - Lawrence-Donway	43.728026	-79.382805	Bus Line
3	Davisville North	43.712751	-79.390197	Homeway Restaurant & Brunch	43.712641	-79.391557	Breakfast Spot
4	Davisville North	43.712751	-79.390197	Sherwood Park	43.716551	-79.387776	Park

Then we will find out how many venues are in each Neighborhood

Neighborhood	NeighborhoodLatitude	NeighborhoodLongitude	VenueName	VenueLatitude	VenueLongitude	VenueCategory
Agincourt	5	5	5	5	5	5
Alderwood, Long Branch	7	7	7	7	7	7
Bathurst Manor, Wilson Heights, Downsview North	21	21	21	21	21	21
Bayview Village	4	4	4	4	4	4
Bedford Park, Lawrence Manor East	22	22	22	22	22	22
...	...	...	...	...	...	...
Willowdale, Willowdale West	5	5	5	5	5	5
Woburn	4	4	4	4	4	4
Woodbine Heights	8	8	8	8	8	8
York Mills West	2	2	2	2	2	2
York Mills, Silver Hills	1	1	1	1	1	1

We will also find out the most common 10 top venues for each Neighborhood.

	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	Agincourt	Lounge	Skating Rink	Latin American Restaurant	Clothing Store	Breakfast Spot	Falafel Restaurant	Event Space	Ethiopian Restaurant	Escape Room	Discount Store
1	Alderwood, Long Branch	Pizza Place	Pharmacy	Gym	Sandwich Place	Coffee Shop	Pub	Dog Run	Dim Sum Restaurant	Diner	Discount Store
2	Bathurst Manor, Wilson Heights, Downsview North	Bank	Coffee Shop	Fried Chicken Joint	Chinese Restaurant	Bridal Shop	Sandwich Place	Diner	Restaurant	Middle Eastern Restaurant	Supermarket
3	Bayview Village	Japanese Restaurant	Café	Chinese Restaurant	Bank	Distribution Center	Dog Run	Doner Restaurant	Donut Shop	Drugstore	Yoga Studio
4	Bedford Park, Lawrence Manor East	Sandwich Place	Italian Restaurant	Coffee Shop	Greek Restaurant	Thai Restaurant	Locksmith	Liquor Store	Comfort Food Restaurant	Juice Bar	Butcher

Then to continue analyzing the data we will perform a technique in which Categorical Data is transformed into Numerical Data for Machine Learning algorithms. This technique is called One hot encoding. For each of the neighborhoods, individual venues were turned into the frequency at how many of those Venues were in each neighborhood.

	Neighborhoods	Accessories Store	Afghan Restaurant	Airport	Airport Food Court	Airport Gate	Airport Lounge	Airport Service	Airport Terminal	American Restaurant	...	Train Station	Turkish Restaurant	Vegetarian / Vegan Restaurant	Video Game Store	Vietn Rest
0	Lawrence Park	0	0	0	0	0	0	0	0	0	...	0	0	0	0	
1	Lawrence Park	0	0	0	0	0	0	0	0	0	...	0	0	0	0	
2	Lawrence Park	0	0	0	0	0	0	0	0	0	...	0	0	0	0	
3	Davisville North	0	0	0	0	0	0	0	0	0	...	0	0	0	0	
4	Davisville North	0	0	0	0	0	0	0	0	0	...	0	0	0	0	

Now, we will group those rows by Neighborhood and by taking the average of the frequency of occurrence of each Venue Category. We will get the below data frame:

	Neighborhoods	Accessories Store	Afghan Restaurant	Airport	Airport Food Court	Airport Gate	Airport Lounge	Airport Service	Airport Terminal	American Restaurant	...	Train Station	Turkish Restaurant	Vegetarian / Vegan Restaurant	Video Game Store	Vietnamese Restaurant
0	Agincourt	0.000000	0.000000	0.0000	0.0000	0.0000	0.000	0.000	0.0000	0.000000	...	0.00	0.0	0.000000	0.000000	0.000000
1	Alderwood, Long Branch	0.000000	0.000000	0.0000	0.0000	0.0000	0.000	0.000	0.0000	0.000000	...	0.00	0.0	0.000000	0.000000	0.000000
2	Bathurst Manor, Wilson Heights, Downsview North	0.000000	0.000000	0.0000	0.0000	0.0000	0.000	0.000	0.0000	0.000000	...	0.00	0.0	0.000000	0.000000	0.000000
3	Bayview Village	0.000000	0.000000	0.0000	0.0000	0.0000	0.000	0.000	0.0000	0.000000	...	0.00	0.0	0.000000	0.000000	0.000000
4	Bedford Park, Lawrence Manor East	0.000000	0.000000	0.0000	0.0000	0.0000	0.000	0.000	0.0000	0.045455	...	0.00	0.0	0.000000	0.000000	0.000000

As we mentioned before for this project it is crucial to analyze the Vegan/vegetarian restaurants that already exist in Toronto.

We will create a new data frame that shows the Neighborhood names and the mean frequency of Vegan/vegetarian restaurants in that Neighborhood. This made the data much simpler to analyze for our project.

	Neighborhood	Vegetarian / Vegan Restaurant
43	Kensington Market, Chinatown, Grange Park	0.054054
68	Runnymede, Swansea	0.030303
48	Little Portugal, Trinity	0.022222
18	Commerce Court, Victoria Hotel	0.020000
5	Berczy Park	0.018182
...	...	...
32	Glencairn	0.000000
31	Garden District, Ryerson	0.000000
30	Forest Hill North & West, Forest Hill Road Park	0.000000
28	Fairview, Henry Farm, Oriole	0.000000
95	York Mills, Silver Hills	0.000000

Another venue category that is interesting to consider in our analysis is the Yoga Studios. This is an ideal venue to have near our restaurant because the clients of the yoga studio highly likely will be potential customers of a Vegan restaurant.

We will select only the Neighborhood, vegan restaurants, and yoga studio venues so we get a data frame where using the average of the frequency of occurrence, we can see in which neighborhood we can not find either a yoga studio or a vegan restaurant, in which ones we find both and in which ones we find at least one of each.

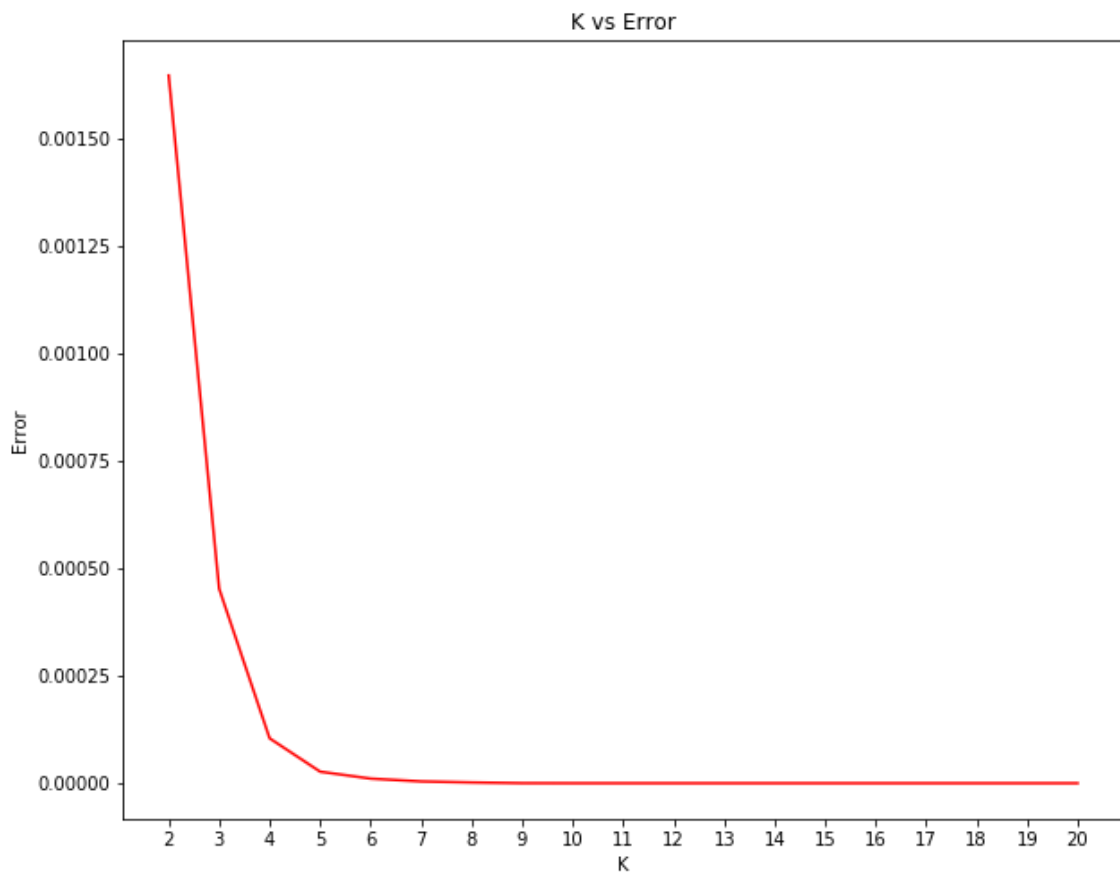|	Neighborhoods	Vegetarian / Vegan Restaurant	Yoga Studio
0	Agincourt	0.000000	0.000000
1	Alderwood, Long Branch	0.000000	0.000000
2	Bathurst Manor, Wilson Heights, Downsview North	0.000000	0.000000
3	Bayview Village	0.000000	0.000000
4	Bedford Park, Lawrence Manor East	0.000000	0.000000
5	Berczy Park	0.018182	0.000000
6	Birch Cliff, Cliffside West	0.000000	0.000000
7	Brockton, Parkdale Village, Exhibition Place	0.000000	0.000000
8	Business reply mail Processing Centre, South C...	0.000000	0.000000
9	CN Tower, King and Spadina, Railway Lands, Har...	0.000000	0.000000
10	Caledonia-Fairbanks	0.000000	0.000000
11	Canada Post Gateway Processing Centre	0.000000	0.000000
12	Cedarbrae	0.000000	0.000000
13	Central Bay Street	0.014706	0.014706
14	Christie	0.000000	0.000000
15	Church and Wellesley	0.000000	0.026667
16	Clarks Corners, Tam O'Shanter, Sullivan	0.000000	0.000000

CLUSTERING THE NEIGHBORHOODS

Now we will cluster the neighborhoods based on the neighborhoods that had similar averages of vegan/vegetarian restaurants in that Neighborhood.

To do this we will use a machine learning technique called K-Means clustering. With the K-Means clustering, objects that are similar based on a certain variable are put into the same cluster.

The first step is to get our optimum K value that was neither overfitting nor underfitting the model. This K value determines how many clusters we will get, and we will use the Elbow Point Technique to get it.

In this technique, we run a test with different numbers of K values and measured the accuracy, and then chose the best K value. The best K value is chosen at the point in which the line has the sharpest turn. In our case, we have the Elbow Point at K = 4. That means we will have a total of 4 clusters.

K vs Error

According to this technique, our neighborhoods will be divided into 4 different clusters according to the ones that have a similar mean frequency of Vegan/Vegetarian restaurants.

With this classification, we will create a new data frame that includes the cluster for each Neighborhood and the occurrence of Vegan/Vegetarian Restaurants.

	Neighborhood	Vegetarian / Vegan Restaurant	ClusterLabels
0	Agincourt	0.0	0
1	Alderwood, Long Branch	0.0	0
2	Bathurst Manor, Wilson Heights, Downsview North	0.0	0
3	Bayview Village	0.0	0
4	Bedford Park, Lawrence Manor East	0.0	0

Now we are going to add info of latitude and longitude and the closest venues to each Neighborhood and sort the results by cluster Labels. This new table would be the basis for analyzing new opportunities for opening a new Vegan restaurant in Toronto.

	Neighborhood	Vegetarian / Vegan Restaurant	ClusterLabels	NeighborhoodLatitude	NeighborhoodLongitude	VenueName	VenueLatitude	VenueLongitude	VenueCategory
0	Agincourt	0.00	0	43.794200	-79.262029	Panagio's Breakfast & Lunch	43.792370	-79.260203	Breakfast Spot
51	Mimico NW, The Queensway West, South of Bloor,...	0.00	0	43.628841	-79.520999	RONA	43.629393	-79.518320	Hardware Store
51	Mimico NW, The Queensway West, South of Bloor,...	0.00	0	43.628841	-79.520999	McDonald's	43.630002	-79.518198	Fast Food Restaurant
51	Mimico NW, The Queensway West, South of Bloor,...	0.00	0	43.628841	-79.520999	Jim & Maria's No Frills	43.631152	-79.518617	Grocery Store
51	Mimico NW, The Queensway West, South of Bloor,...	0.00	0	43.628841	-79.520999	Subway	43.631659	-79.519001	Sandwich Place
...	...	...	...	...	...	...	...	...	...
64	Richmond, Adelaide, King	0.01	3	43.650571	-79.384568	Roots	43.653613	-79.380244	Clothing Store
64	Richmond, Adelaide, King	0.01	3	43.650571	-79.384568	Ted Baker	43.652843	-79.380325	Clothing Store
64	Richmond, Adelaide, King	0.01	3	43.650571	-79.384568	Starbucks	43.649028	-79.381593	Coffee Shop
64	Richmond, Adelaide, King	0.01	3	43.650571	-79.384568	Starbucks	43.646731	-79.383951	Coffee Shop
64	Richmond, Adelaide, King	0.01	3	43.650571	-79.384568	Starbucks	43.650751	-79.388047	Coffee Shop

2139 rows × 9 columns

3.1 Data analysis

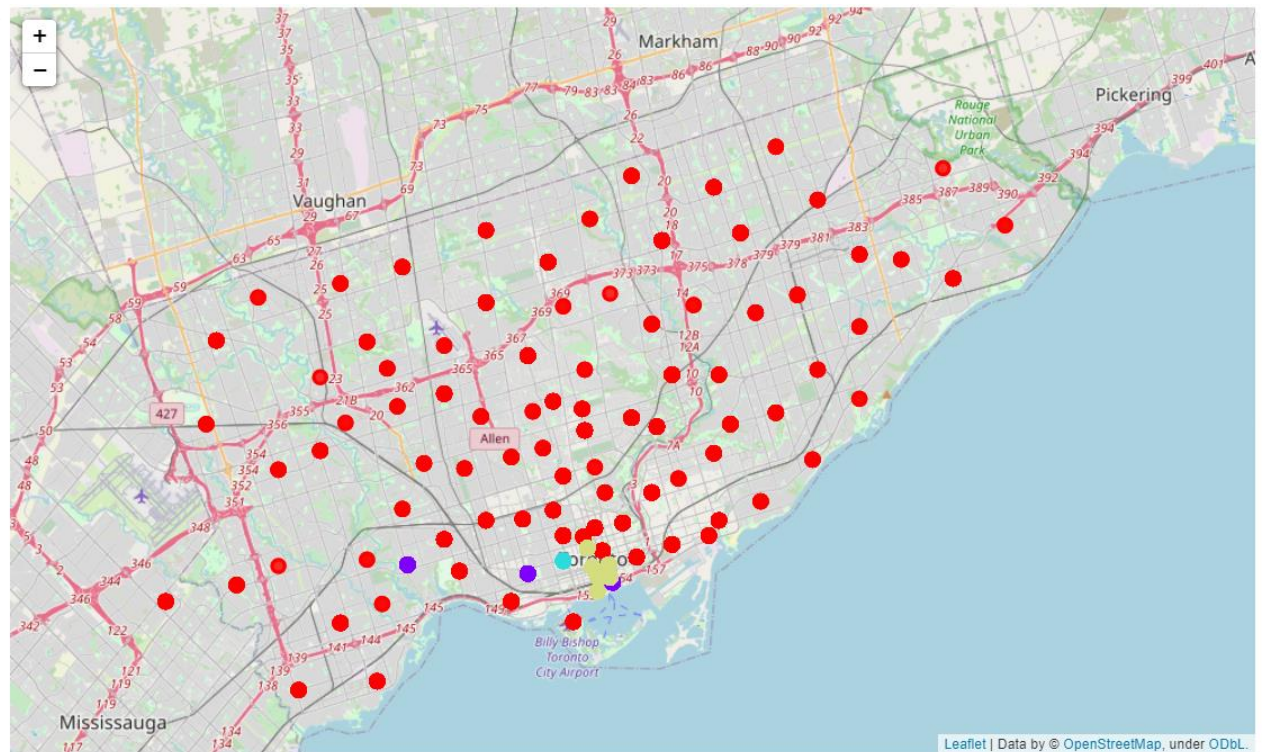Let´s now start to analyze the data and the characteristics of the cluster that we got.

First of all, to better visualize the different clusters and neighborhoods in them we will create a map using Folium Library in where each neighborhood will be colored based on their cluster label.

Cluster 0 — Red

Cluster 1 — Purple

Cluster 2 — Turquoise

Cluster 3— Green Khaki

The map above shows the different clusters that had a similar mean frequency of Vegetarian/Vegan restaurants.

In this map, we can see that the number of neighborhoods in each cluster is quite different. Cluster 0 (red) is the one with more neighborhoods and geographically it is the more dispersed. Cluster 2 (turquoise) just has one neighborhood and cluster 3 (green khaki) have several neighborhoods but remarkably close to the others. Cluster 1 (purple) has 4 neighborhoods and is little dispersed.

Before analyzing each cluster one by one let see the total amount of neighborhoods in each cluster and the average of Vegan/Vegetarian Restaurants in each one.

 Using Matplotlib we create a bar graph where we can compare the number of Neighborhoods per Cluster.

How many Neighborhoods per Cluster
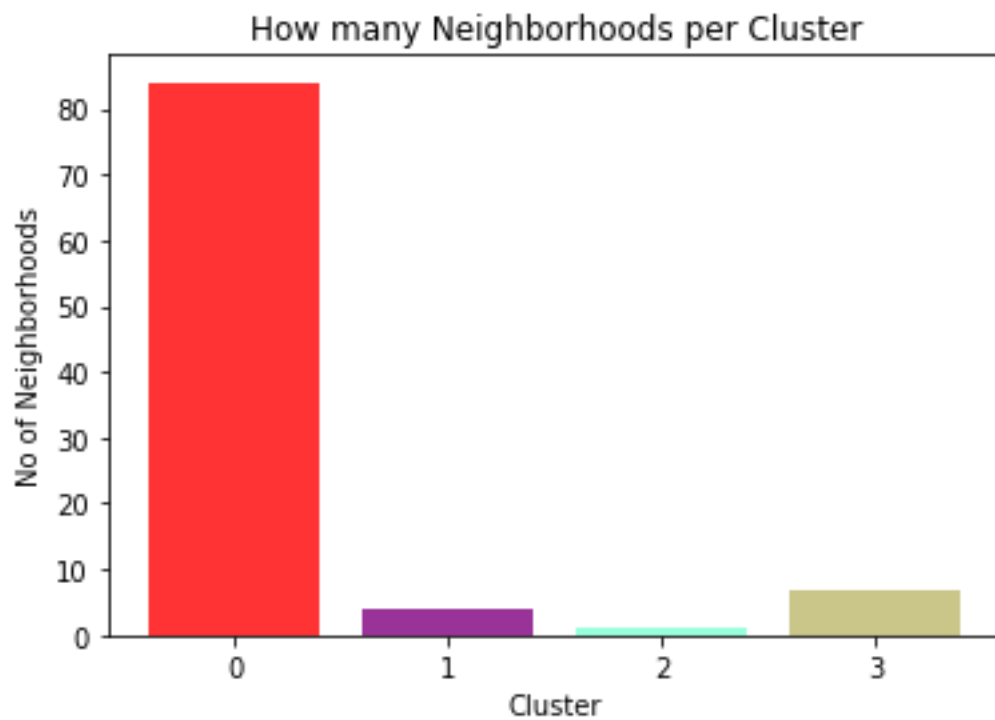
We see that Cluster o has the highest number of neighborhoods (84) while cluster 2 has the lowest number (1).

Now let´s compare the average Vegetarian/Vegan restaurant per cluster.

Average number of Vegetarian / Vegan Restaurant per Cluster
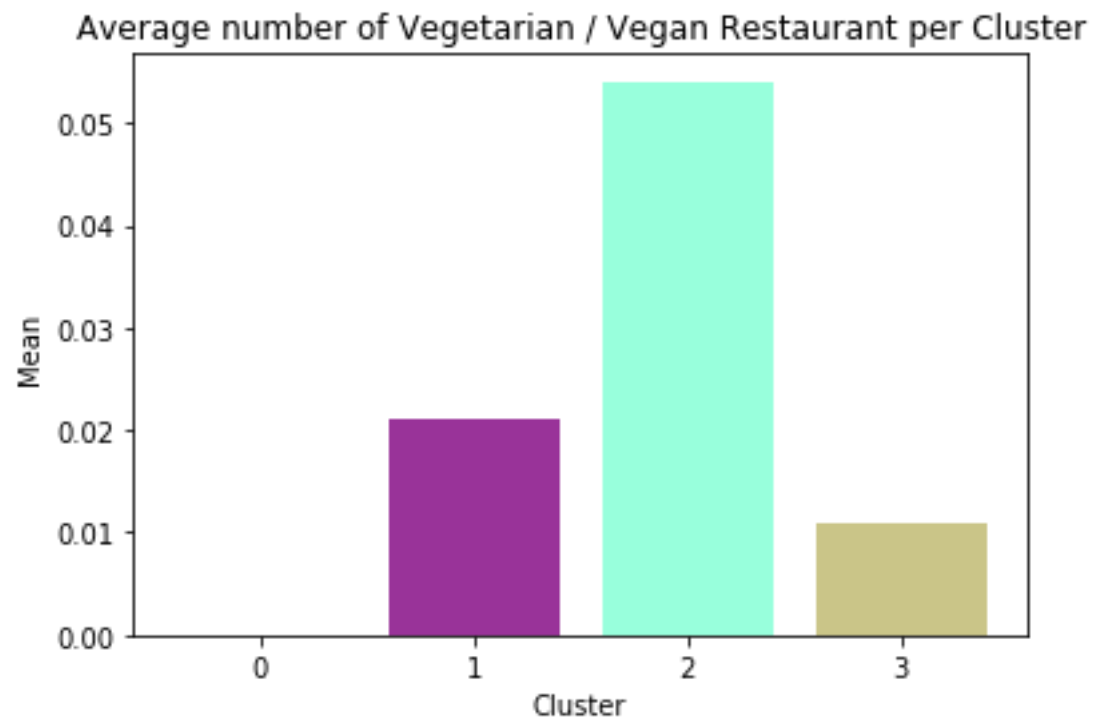
Cluster 0, being the one with the highest number of neighborhoods it has the lowest average of Vegan/Vegetarian Restaurants (0,00) while cluster 2 has the highest average (0,54054) but the lowest number of neighborhoods (1).

Now we will analyze each cluster individually. To do this we will create a data frame with the borough of each neighborhood, and we will merge it with each cluster data frame.

CLUSTER 0

Borough	Neighborhood	Vegetarian / Vegan Restaurant	ClusterLabels	NeighborhoodLatitude	NeighborhoodLongitude	VenueName	VenueLatitude	VenueLongitude	VenueCategory
Central Toronto	Lawrence Park	0.0	0	43.728020	-79.388790	TTC Bus #162 - Lawrence-Donway	43.728026	-79.382805	Bus Line
Central Toronto	Lawrence Park	0.0	0	43.728020	-79.388790	Zodiac Swim School	43.728532	-79.382860	Swim School
Central Toronto	Lawrence Park	0.0	0	43.728020	-79.388790	Lawrence Park Ravine	43.726963	-79.394382	Park
Central Toronto	Davisville North	0.0	0	43.712751	-79.390197	Sherwood Park	43.716551	-79.387776	Park
Central Toronto	Davisville North	0.0	0	43.712751	-79.390197	Summerhill Market North	43.715499	-79.392881	Food & Drink Shop
...	...	...	...	...	...	...	...	...	...
York	Runnymede, The Junction North	0.0	0	43.673185	-79.487262	High Park Brewery	43.669903	-79.483430	Brewery
York	Runnymede, The Junction North	0.0	0	43.673185	-79.487262	195 Jane Rocket	43.672335	-79.492634	Bus Line
York	Runnymede, The Junction North	0.0	0	43.673185	-79.487262	Wonderfood	43.672352	-79.492571	Convenience Store
York	Weston	0.0	0	43.706876	-79.518188	Wallace C. Swanek Park	43.708896	-79.522648	Park
York	Weston	0.0	0	43.706876	-79.518188	Grattan Park	43.706222	-79.521705	Park

84 neighborhoods with 225 unique venue categories and nonvegan restaurants

9 yoga studios. Most of them (5) are situated in boroughs inside Downtown Toronto and the rest in other boroughs (Central Toronto, East Toronto, and East York)

Borough	Neighborhood	Vegetarian / Vegan Restaurant	ClusterLabels	NeighborhoodLatitude	NeighborhoodLongitude	VenueName	VenueLatitude	VenueLongitude	VenueCategory
Central Toronto	North Toronto West, Lawrence Park	0.0	0	43.715383	-79.405678	Barreworks	43.714070	-79.400109	Yoga Studio
Downtown Toronto	Church and Wellesley	0.0	0	43.665860	-79.383160	The Yoga Sanctuary	43.661499	-79.383636	Yoga Studio
Downtown Toronto	Church and Wellesley	0.0	0	43.665860	-79.383160	Bikram Yoga Yonge	43.668205	-79.385780	Yoga Studio
Downtown Toronto	Regent Park, Harbourfront	0.0	0	43.654260	-79.360636	The Yoga Lounge	43.655515	-79.364955	Yoga Studio
Downtown Toronto	University of Toronto, Harbord	0.0	0	43.662696	-79.400049	Sivananda Yoga Centre	43.662754	-79.402951	Yoga Studio
Downtown Toronto	Queen's Park, Ontario Provincial Government	0.0	0	43.662301	-79.389494	The Yoga Sanctuary	43.661499	-79.383636	Yoga Studio
East Toronto	The Danforth West, Riverdale	0.0	0	43.679557	-79.352188	Moksha Yoga Danforth	43.677622	-79.352116	Yoga Studio
East Toronto	Studio District	0.0	0	43.659526	-79.340923	Spirit Loft Yoga	43.663548	-79.341333	Yoga Studio
East York	Thorncliffe Park	0.0	0	43.705369	-79.349372	Bikram Yoga East York	43.705450	-79.351448	Yoga Studio

CLUSTER 1

	Borough	Neighborhood	Vegetarian / Vegan Restaurant	ClusterLabels	NeighborhoodLatitude	NeighborhoodLongitude	VenueName	VenueLatitude	VenueLongitude	VenueCategory
235	West Toronto	Runnymede, Swansea	0.028571	1	43.651571	-79.484450	The Coffee Bouquets	43.648785	-79.485940	Coffee Shop
218	West Toronto	Runnymede, Swansea	0.028571	1	43.651571	-79.484450	Heart	43.650310	-79.480125	Dessert Shop
216	West Toronto	Runnymede, Swansea	0.028571	1	43.651571	-79.484450	Cinelli	43.649916	-79.482146	Salon / Barbershop
215	West Toronto	Runnymede, Swansea	0.028571	1	43.651571	-79.484450	Max's Market	43.650525	-79.479145	Gourmet Shop
214	West Toronto	Runnymede, Swansea	0.028571	1	43.651571	-79.484450	Bloom Restaurant	43.650307	-79.479836	Latin American Restaurant
...	...	...	...	...	...	...	...	...	...	...
34	Downtown Toronto	Berczy Park	0.018182	1	43.644771	-79.373306	Eggspectation	43.646526	-79.375134	Breakfast Spot

There are 4 neighborhoods with a total of 97 unique venue categories in this cluster.

We can find 5 vegan/vegetarian restaurants situated in 4 different neighborhoods.

Borough	Neighborhood	Vegetarian / Vegan Restaurant	ClusterLabels	NeighborhoodLatitude	NeighborhoodLongitude	VenueName	VenueLatitude	VenueLongitude	VenueCategory
Downtown Toronto	Berczy Park	0.018182	1	43.644771	-79.373306	Fresh On Front	43.647815	-79.374453	Vegetarian / Vegan Restaurant
Downtown Toronto	Commerce Court, Victoria Hotel	0.020000	1	43.648198	-79.379817	Fresh On Front	43.647815	-79.374453	Vegetarian / Vegan Restaurant
Downtown Toronto	Commerce Court, Victoria Hotel	0.020000	1	43.648198	-79.379817	Rosalinda	43.650252	-79.385156	Vegetarian / Vegan Restaurant
West Toronto	Little Portugal, Trinity	0.021739	1	43.647927	-79.419750	Veghed	43.649224	-79.422326	Vegetarian / Vegan Restaurant
West Toronto	Runnymede, Swansea	0.028571	1	43.651571	-79.484450	Awai	43.650412	-79.478477	Vegetarian / Vegan Restaurant

We also find 2 yoga studios both situated in neighborhoods where we also find a Vegan/Vegetarian Restaurant.

Borough	Neighborhood	Vegetarian / Vegan Restaurant	ClusterLabels	NeighborhoodLatitude	NeighborhoodLongitude	VenueName	VenueLatitude	VenueLongitude	VenueCategory
West Toronto	Little Portugal, Trinity	0.021739	1	43.647927	-79.41975	YogaSpace	43.647607	-79.420133	Yoga Studio
West Toronto	Runnymede, Swansea	0.028571	1	43.651571	-79.48445	(The New) Moksha Yoga Bloor West	43.648658	-79.485242	Yoga Studio

CLUSTER 2

Borough	Neighborhood	Vegetarian / Vegan Restaurant	ClusterLabels	NeighborhoodLatitude	NeighborhoodLongitude	VenueName	VenueLatitude	VenueLongitude	VenueCategory
Downtown Toronto	Kensington Market, Chinatown, Grange Park	0.054054	2	43.653206	-79.400049	FILM CAFE	43.655109	-79.402342	Comfort Food Restaurant
Downtown Toronto	Kensington Market, Chinatown, Grange Park	0.054054	2	43.653206	-79.400049	Saigon Lotus Restaurant	43.654311	-79.399225	Vietnamese Restaurant
Downtown Toronto	Kensington Market, Chinatown, Grange Park	0.054054	2	43.653206	-79.400049	Thirsty and Miserable	43.654565	-79.401583	Beer Bar
Downtown Toronto	Kensington Market, Chinatown, Grange Park	0.054054	2	43.653206	-79.400049	Meeplemart	43.651628	-79.397410	Gaming Cafe
Downtown Toronto	Kensington Market, Chinatown, Grange Park	0.054054	2	43.653206	-79.400049	Dumpling House	43.653860	-79.398558	Dumpling Restaurant
...	...	...	...	...	...	...	...	...	...
Downtown Toronto	Kensington Market, Chinatown, Grange Park	0.054054	2	43.653206	-79.400049	Pancho y Emiliano	43.654472	-79.401969	Mexican Restaurant
Downtown Toronto	Kensington Market, Chinatown, Grange Park	0.054054	2	43.653206	-79.400049	Trinity Common	43.656590	-79.402761	Bar
Downtown Toronto	Kensington Market, Chinatown, Grange Park	0.054054	2	43.653206	-79.400049	Independent City Market	43.657337	-79.401512	Supermarket
Downtown Toronto	Kensington Market, Chinatown, Grange Park	0.054054	2	43.653206	-79.400049	A & C World	43.657409	-79.399847	Gaming Cafe

In this cluster, we find only one neighborhood with a total of 49 unique venue categories inside.

There are 4 Vegan Restaurant, so it is the cluster with the highest average. We do not find any Yoga Studio.

Borough	Neighborhood	Vegetarian / Vegan Restaurant	ClusterLabels	NeighborhoodLatitude	NeighborhoodLongitude	VenueName	VenueLatitude	VenueLongitude	VenueCategory
Downtown Toronto	Kensington Market, Chinatown, Grange Park	0.054054	2	43.653206	-79.400049	Hibiscus	43.655454	-79.402439	Vegetarian / Vegan Restaurant
Downtown Toronto	Kensington Market, Chinatown, Grange Park	0.054054	2	43.653206	-79.400049	Greens Vegetarian Restaurant	43.652034	-79.402382	Vegetarian / Vegan Restaurant
Downtown Toronto	Kensington Market, Chinatown, Grange Park	0.054054	2	43.653206	-79.400049	Urban Herbivore	43.656193	-79.402673	Vegetarian / Vegan Restaurant
Downtown Toronto	Kensington Market, Chinatown, Grange Park	0.054054	2	43.653206	-79.400049	Buddha's Vegetarian	43.651904	-79.403312	Vegetarian / Vegan Restaurant

CLUSTER 3

Borough	Neighborhood	Vegetarian / Vegan Restaurant	ClusterLabels	NeighborhoodLatitude	NeighborhoodLongitude	VenueName	VenueLatitude	VenueLongitude	VenueCategor
Downtown Toronto	St. James Town	0.011765	3	43.651494	-79.375418	Apple Eaton Centre	43.652818	-79.380617	Electronic Stor
Downtown Toronto	St. James Town	0.011765	3	43.651494	-79.375418	Seafront Fish Market	43.648479	-79.371489	Fish Marke
Downtown Toronto	St. James Town	0.011765	3	43.651494	-79.375418	Schnitzel Queen	43.654239	-79.370533	Germa Restaurar
Downtown Toronto	St. James Town	0.011765	3	43.651494	-79.375418	Buster's Sea Cove	43.648495	-79.371462	Seafoo Restaurar
Downtown Toronto	St. James Town	0.011765	3	43.651494	-79.375418	Scheffler's Deli	43.648643	-79.371537	Cheese Sho
...	...	...	...	...	...	...	...	...	.
Downtown Toronto	First Canadian Place, Underground city	0.010000	3	43.648429	-79.382280	DAVIDsTEA	43.650547	-79.383385	Tea Roor
Downtown Toronto	First Canadian Place, Underground city	0.010000	3	43.648429	-79.382280	Dineen Coffee	43.650497	-79.378765	Caf
Downtown Toronto	First Canadian Place, Underground city	0.010000	3	43.648429	-79.382280	John & Sons Oyster House	43.650602	-79.381555	Seafoo Restaurar
Downtown Toronto	First Canadian Place, Underground city	0.010000	3	43.648429	-79.382280	Shangri-La Toronto	43.649129	-79.386557	Hote
Downtown Toronto	First Canadian Place, Underground city	0.010000	3	43.648429	-79.382280	Kupfert & Kim (First Canadian Place)	43.648547	-79.381624	Gluten-fre Restaurar

There are 7 neighborhoods in Cluster 3, and we find 145 unique venue categories. There are 7 Vegetarian/Vegan Restaurants all of them in neighborhoods inside Down Toronto and where we can also find Yoga Studios (2).

Borough	Neighborhood	Vegetarian / Vegan Restaurant	ClusterLabels	NeighborhoodLatitude	NeighborhoodLongitude	VenueName	VenueLatitude	VenueLongitude	VenueCategory
Downtown Toronto	St. James Town	0.011765	3	43.651494	-79.375418	Fresh On Front	43.647815	-79.374453	Vegetarian / Vegan Restaurant
Downtown Toronto	Central Bay Street	0.014706	3	43.657952	-79.387383	Vegetarian Haven	43.656016	-79.392758	Vegetarian / Vegan Restaurant
Downtown Toronto	Richmond, Adelaide, King	0.010000	3	43.650571	-79.384568	Rosalinda	43.650252	-79.385156	Vegetarian / Vegan Restaurant
Downtown Toronto	Harbourfront East, Union Station, Toronto Islands	0.010000	3	43.640816	-79.381752	Kupfert & Kim	43.641179	-79.378144	Vegetarian / Vegan Restaurant
Downtown Toronto	Toronto Dominion Centre, Design Exchange	0.010000	3	43.647177	-79.381576	Rosalinda	43.650252	-79.385156	Vegetarian / Vegan Restaurant
Downtown Toronto	Stn A PO Boxes	0.010417	3	43.646435	-79.374846	Fresh On Front	43.647815	-79.374453	Vegetarian / Vegan Restaurant
Downtown Toronto	First Canadian Place, Underground city	0.010000	3	43.648429	-79.382280	Rosalinda	43.650252	-79.385156	Vegetarian / Vegan Restaurant

Borough	Neighborhood	Vegetarian / Vegan Restaurant	ClusterLabels	NeighborhoodLatitude	NeighborhoodLongitude	VenueName	VenueLatitude	VenueLongitude	VenueCategory
Downtown Toronto	Central Bay Street	0.014706	3	43.657952	-79.387383	The Yoga Sanctuary	43.661499	-79.383636	Yoga Studio
Downtown Toronto	Stn A PO Boxes	0.010417	3	43.646435	-79.374846	Bikram Yoga Centre	43.649214	-79.375229	Yoga Studio

4. RESULT AND DISCUSSION

Cluster 0 is the one with the highest number of neighborhoods and different category venues but at the same time the only one with any Vegan/Vegetarian Restaurant therefore it has the lowest average of vegan restaurants. On the other hand, there are 9 Yoga Studios in Cluster 0 most of them concentrated in four neighborhoods inside Downtown Toronto.

It is in Cluster 2 where we find the highest average of Vegan/Vegetarian Restaurants (0.054), it has 4 but only 1 Neighborhood, Kensington.

Cluster 1 with 4 Neighborhood and 5 Vegan/Vegetarian Restaurants has the second-highest average of Vegan/Vegetarian Restaurants and it has 2 Yoga Studios, both situated in Neighborhoods where you can also find a Vegan Restaurant.

Cluster 3 is the one with the third-highest average of Vegan/Vegetarian Restaurant. It has a total of 7 each one situated in a different neighborhood but all inside Downtown Toronto. We can also find two yoga studios both situated in neighborhoods that also have a Vegan Restaurant.

According to this analysis, our recommendation is to open a Vegan Restaurant in one of the 4 Neighborhoods inside **Downtown Toronto** and belonging to **Cluster 0**: Church and Wellesley, Regent Park/Harbourfront, University of Toronto/Harbord, and Queen´s Park Ontario/Provincial Government.

In any of them, there is not another Vegan/ Vegetarian Restaurant so we have no direct competition and all of them have a Yoga Studio, even we can find 2 in the Church and Wellesley Neighborhood.

To give a more accurate recommendation among these 4 neighborhoods, we will analyze the results about which are the 10 most common Venues in each of them.

	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
15	Church and Wellesley	Coffee Shop	Gay Bar	Japanese Restaurant	Sushi Restaurant	Restaurant	Yoga Studio	Mediterranean Restaurant	Café	Pub	Hotel

```
1 df_regent=neighborhoods_venues_sorted[neighborhoods_venues_sorted['Neighborhoods']=='Regent Park, Harbourfront']
2 df_regent
```

	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
63	Regent Park, Harbourfront	Coffee Shop	Bakery	Pub	Park	Breakfast Spot	Café	Theater	Distribution Center	Chocolate Shop	Mexican Restaurant

```
1 df_University=neighborhoods_venues_sorted[neighborhoods_venues_sorted['Neighborhoods']=='University of Toronto, Harbord']
2 df_University
```

	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
84	University of Toronto, Harbord	Café	Bookstore	Sandwich Place	Bakery	Bar	Japanese Restaurant	Nightclub	Bank	Italian Restaurant	Beer Bar

```
1 df_Queens=neighborhoods_venues_sorted[neighborhoods_venues_sorted['Neighborhoods']=="Queen's Park, Ontario Provincial Govern
2 df_Queens
```

	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
62	Queen's Park, Ontario Provincial Government	Coffee Shop	Yoga Studio	College Auditorium	Bar	Beer Bar	Smoothie Shop	Sandwich Place	Café	Restaurant	Chinese Restaurant

According to the information in the data frame above we recommend opening the Vegan/Vegetarian Restaurant in **Church and Wellesley** or **Queen´s Park Ontario/Provincial Government**.

In both Neighborhoods, there is a Yoga Studio as one of the most common venues and the type of clients of Yoga studios are people normally open to trying and exploring vegetarian or vegan restaurants even if they do not follow a strictly vegan or vegetarian diet.

The other 10 most common venues are almost all restaurants or café which tells us that people are looking for places to eat to whom we can give a different alternative to the restaurants that already exist.

The analysis does not take into consideration other factors that would affect the success of a new restaurant that are directly related to the location such as visibility or accessibility so the final decision between these two neighborhoods could depend also on the opportunity to find the best suitable local to rent.

5. CONCLUSION

Veganism has emerged as one of the top food and lifestyle trends in the past few years. Moreover, people, especially in cosmopolitan cities like Toronto, are more open to trying and exploring vegetarian or vegan restaurants.

The purpose of this project was to find the best location in Toronto to open a Vegan/Vegetarian Restaurant.

We got a great measure of data from Wikipedia which we scraped with the Beautifulsoup Web scraping Library. We utilized numerous Python libraries to fetch the information and process all the data into clean data frames. We used Foursquare API to investigate the settings in Neighborhoods of Toronto. We also visualized utilizing different plots present in seaborn and Matplotlib libraries.

Applying the K-Means clustering algorithm, we got different clusters of Neighborhoods in Toronto according to the average of Vegan/Vegetarian Restaurants. After having analyzed the clusters one by one we chose the cluster with fewer Vegan/Restaurants on average, therefore with fewer competitors. Moreover, having assumed that the type of clients in a yoga studio can most likely be a potential customer of a vegan restaurant, it should be noted that this cluster is the one with the highest number of Yoga Studios. Taking all that into account we got the most 4 suitable neighborhoods within the cluster. We analyzed each one according to the most common venues in them to be able to give the best recommendation to our stakeholders.

The final decision on optimal Vegan/Vegetarian location will be given to the stakeholders. They will be also encouraged to take into consideration additional factors such as accessibility or locals rent opportunities.