

"Vegetarian / Vegan Restaurants" Neighbourhoods



Photo by Humankind Ventures Ltd, 2020

Abstract

Internationally, the rates of obesity-related chronic diseases have become a public health concern. Most people are transforming into distinct lifestyles and dietary changes to reduce the high risk of chronic infection caused by modern nutritional patterns. Vegetarianism is the best diet to enhance health. There is a limited number of vegetarian restaurants; therefore, there is an opportunity to create employment through vegetarian businesses. The targeted audience for this study is entrepreneurs coveting to opening Vegetarian / Vegan Restaurants. The initial stage of implementing this assignment project was to use the Notebook to build the code to scrape the following Wikipedia page,

https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M BeautifulSoup package applied for web scraping. Foursquare API determined various Vegetarian restaurant locations in every neighborhood. A Geocoder package verified the neighborhood's geographical location and venue data of Vegetarian restaurants. The first neighborhood explored in our data set is 'Malvern, Rouge'. The algorithm formed 4 clusters. When comparing all the clusters (cluster 1, cluster 2, cluster 3, and cluster 4), cluster 4 has the maximum average number of Vegetarian / Vegan Restaurants. There are no Vegetarian / Vegan Restaurants in cluster 1.

TABLE OF CONTENTS

1	BUSINESS PROBLEM	3
1.1	Problem statement	3
1.2	Targeted audience	4
2	DATA	4
2.1	Data from Wikipedia	4
2.2	Data cleaning	4
2.3	Geographical Location data using Geocoder Package	5
2.4	Determining data through Foursquare API.....	5
3	METHODOLOGY	5
3.1	Libraries.....	5
3.2	Data loading and cleaning.....	6
3.3	FOURSQUARE API.....	7
3.4	K-means clustering.....	7
4	RESULTS.....	8
4.1	Data from Wikipedia	8
4.2	Data cleaning	8
4.3	Latitudes and longitudes from Geocoder package	9
4.4	FOURSQUARE API.....	10
4.5	K-means clustering.....	12
5	DATA ANALYSIS	12
6	DISCUSSION.....	14
7	CONCLUSION.....	16
8	REFERENCES.....	17

LIST OF FIGURES

Figure 1: Python libraries used	6
Figure 2: Toronto map with markers	10
Figure 3: Average Neighbourhoods in each cluster	14
Figure 4: Average Vegetarian/Vegan Restaurants per cluster	15

LIST OF TABLES

Table 1: Data not cleaned	8
Table 2: Cleaned data.....	8
Table 3: Toronto latitudes and longitudes	9

1 BUSINESS PROBLEM

Globally, the rates of obesity-related chronic diseases have become a public health concern. A modern diet is mainly processed food high in total fat, saturated fat, and simple sugars, which increases the risk of developing obesity [1]. In 2010, over 27 million (approximately 11%) adults were diagnosed with heart disease. Nearly 20 million Americans had cancer as of 2010, while 21 million Americans had diabetes mellitus. Most people are transforming into distinct lifestyles and dietary changes to reduce the high risk of chronic disease caused by modern nutritional patterns. Lately, doctors recommend that patients should live a healthy life and not depend on reactionary medicine. Vegetarianism is the best diet to enhance health. The study of [2] stipulates that vegetarians warranted a reputation of being overall healthier people. They are often associated with living more healthily and mindfully by participating in more regular exercise, avoiding fattening or sugary foods, and avoiding alcohol and smoking. Obesity is linked to lower rates of a vegetarian diet, heart disease, and diabetes mellitus (DiMarino, 2013).

1.1 Problem statement

There is a limited number of vegetarian restaurants; therefore, there is an opportunity to create employment through vegetarian businesses. Globally, unemployment is a considerable challenge. According to [4] internationally, young people are three times unemployed than adults, and over 350 million young people are not engaged in education, employment, or training. Therefore, there is a need to create work. This study is significant since segmentation and clustering neighborhoods' information provides entrepreneurs with knowledge; for example, (entrepreneurs will know their competitor's locations and areas where there is a need for their products).

1.2 Targeted audience

The entrepreneurs can use segmentation and clustering neighborhoods' information as a prototype when exploring new opportunities to open a thriving Vegetarian restaurant business in Toronto. The skills gained on the segmentation and clustering neighborhoods assignment can contribute immensely to the body of knowledge since the study findings can be used as a reference when conducting research. The target audience of the project is academics and non-academic researchers. Academics can use the information to educate learners and enhance quality education. Tourists can study this kind of assignment to understand the area of interest before visiting Toronto.

2 DATA

In this assignment, I explored neighborhoods in Toronto to determine where the entrepreneurs can open a new Vegetarian restaurant. The initial stage of implementing this assignment project was to use the Notebook to build the code to scrape the following Wikipedia page, https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M. I applied the head method to display the columns and rows of the data. I used the BeautifulSoup package for web scraping.

2.1 Data from Wikipedia

The code below was used to read from Wikipedia into Pandas Dataframe.

Read the data into a Pandas Dataframe

```
soup = BeautifulSoup(str(data_html))
neigh = soup.find('table')
table_str = str(neigh.extract())
```

```
df = pd.read_html(table_str)[0]
df.head(12)
```

2.2 Data cleaning

Columns that were not relevant dropped using the drop () method. I used the codes illustrated below to clean the data. After that, I processed the cells that have an assigned borough. The dataframe consists of three columns: PostalCode, Borough, and Neighborhood. Then I ignored the cells with a borough that is not assigned using the code below.

Remove Not assigned Boroughs

```
df_dropna = df[df.Borough != 'Not assigned'].reset_index(drop=True)
```

Rename Postal Code to PostalCode

```
df_dropna.rename(columns={'Postal Code' : 'PostalCode'}, inplace=True)  
df = df_dropna  
df.head(12)
```

2.3 Geographical Location data using Geocoder Package

I used A Geocoder package to determine the neighborhoods' geographical location and venue data of Vegetarian restaurants. Below is the Geographical location code using the Geocoder.

Latitudes and Longitudes aligned to the different PostalCodes

```
1 cor_url = "https://cocl.us/Geospatial_data"  
2  
3 df_geo = pd.read_csv(cor_url)  
4 df_geo.rename(columns={'Postal Code': 'PostalCode'}, inplace=True)  
5 df_geo.head()
```

2.4 Determining data through Foursquare API

I used the Foursquare API to determine various Vegetarian restaurant locations in every neighborhood. The Venue data assists in deciding the area, which is best suitable to open Vegetarian restaurants.

3 METHODOLOGY

3.1 Libraries

The folium was pip installed for visualizing geospatial data as follows

pip install folium

The geocoder was imported and initiated as follows

from geopy.geocoders import Nominatim

Figure 1 illustrates the libraries imported during data manipulation and data visualization.

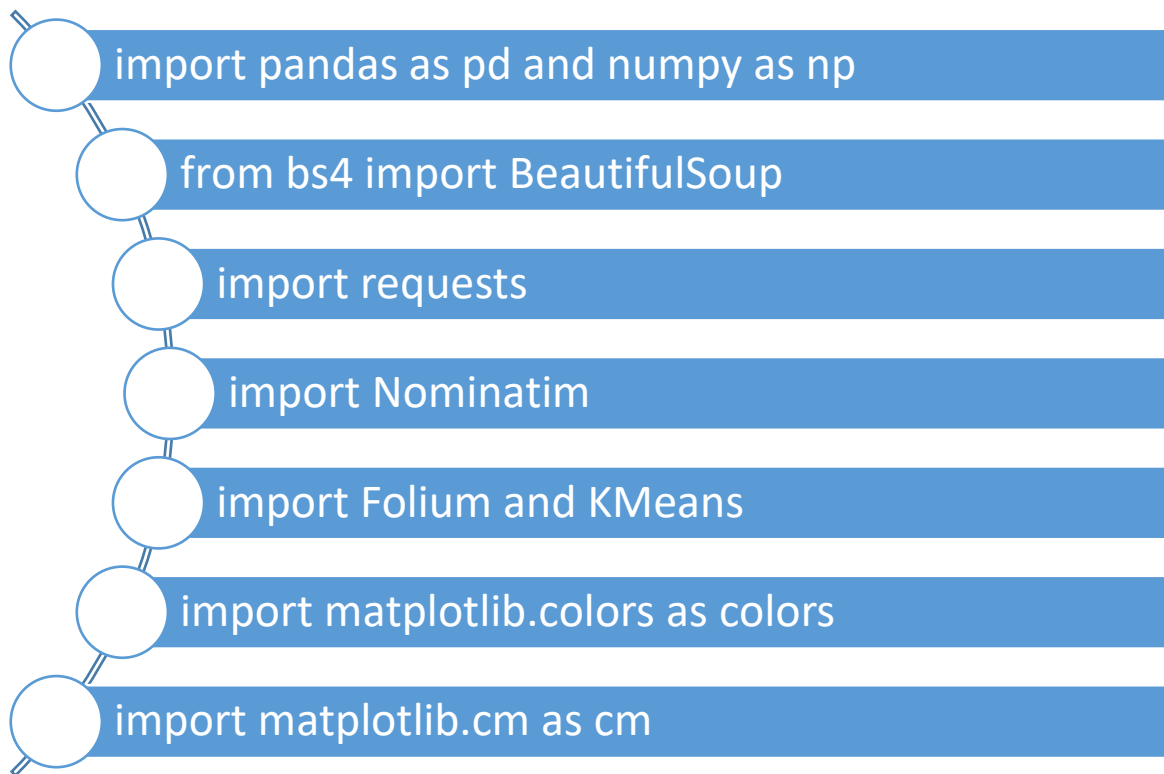


Figure 1: Python libraries used

3.2 Data loading and cleaning

I followed step by step instructions from Coursera course to assess and clean data:

- The Notebook to build the code to scrape the following Wikipedia page, https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
- The dataframe consists of three columns: PostalCode, Borough, and Neighborhood
- Only process the cells that have an assigned borough. Ignore cells with a borough that is **Not Assigned**.

After removing the cells with Not Assigned, I renamed Postal Code to PostalCode and grouped the data based on Borough using the codes below.

```
1 df_dropna.rename(columns={'Postal Code' : 'PostalCode'}, inplace=True)
2 df = df_dropna
3 df.head()
```

Group Borough Data

```
1 df_grouped = df.groupby(['Borough', 'PostalCode'], as_index=False).agg(lambda x:','.join(x))
2 df_grouped.head()
```


The number of columns and rows were determined by the .shape method. The columns were rearranged so that the Postal code appears first.

```
1 df = df.groupby(["PostalCode", "Borough"])["Neighbourhood"].apply(", ".join).reset_index()
2 df.head()
```

A table was created with the latitudes and longitudes collected from the Geocoder package, corresponding to the different Postal Codes as illustrated below.

Latitudes and Longitudes aligned to the different PostalCodes

```
1 cor_url = "https://cocl.us/Geospatial_data"
2
3 df_geo = pd.read_csv(cor_url)
4 df_geo.rename(columns={'Postal Code': 'PostalCode'}, inplace=True)
5 df_geo.head()
```

I visualized a map using Folium and each Neighborhood color-coded depending on the location of each Borough.

3.3 FOURSQUARE API

I accessed the Foursquare API by adding the credentials (CLIENT_ID, CLIENT_SECRET, and Foursquare version). I used Foursquare to determine various Vegetarian restaurant neighborhood. The Venue data assists in deciding the area, which is best suitable to open a Vegetarian restaurant. Below are the codes for determining the first neighborhood in df_toronto, averaging latitude and longitude values, and the top hundred number of venues returned by Foursquare API.

```
1 LIMIT = 100 # limit of number of venues returned by Foursquare API
2 radius = 500 # define radius
3 url = 'https://api.foursquare.com/v2/venues/explore?&client_id={}&client_secret={}&v={}&ll={},{&radius={}&limit={}'.format(
4     CLIENT_ID,
5     CLIENT_SECRET,
6     VERSION,
7     neighborhood_latitude,
8     neighborhood_longitude,
9     radius,
10    LIMIT)
11
12 # get the result to a json file
13 results = requests.get(url).json()
```

3.4 K-means clustering

K-means clustering is applied to the dataset to explore different venues. The neighbourhoods with analogous averages were clustered. The Elbow Point Technique was used to obtain optimum K value that was either overfitting or underfitting the model.

4 RESULTS

4.1 Data from Wikipedia

The data below was read from Wikipedia into Pandas Dataframe.

Table 1: Data not cleaned

	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
5	M6A	North York	Lawrence Manor, Lawrence Heights
6	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government
7	M8A	Not assigned	Not assigned
8	M9A	Etobicoke	Islington Avenue, Humber Valley Village
9	M1B	Scarborough	Malvern, Rouge
10	M2B	Not assigned	Not assigned
11	M3B	North York	Don Mills

4.2 Data cleaning

The table below demonstrates the cleaned data.

Table 2: Cleaned data

	PostalCode	Borough	Neighbourhood
0	M3A	North York	Parkwoods
1	M4A	North York	Victoria Village
2	M5A	Downtown Toronto	Regent Park, Harbourfront
3	M6A	North York	Lawrence Manor, Lawrence Heights
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government
5	M9A	Etobicoke	Islington Avenue, Humber Valley Village
6	M1B	Scarborough	Malvern, Rouge
7	M3B	North York	Don Mills
8	M4B	East York	Parkview Hill, Woodbine Gardens
9	M5B	Downtown Toronto	Garden District, Ryerson
10	M6B	North York	Glencairn
11	M9B	Etobicoke	West Deane Park, Princess Gardens, Martin Grov...

4.3 Latitudes and longitudes from Geocoder package

I created a table with the latitudes and longitudes collected from the Geocoder package, corresponding to the different Postal Codes as illustrated below.

Table 3: Toronto latitudes and longitudes

	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

I visualized a map using Folium and each Neighborhood color-coded depending on each Borough's location as illustrated below.

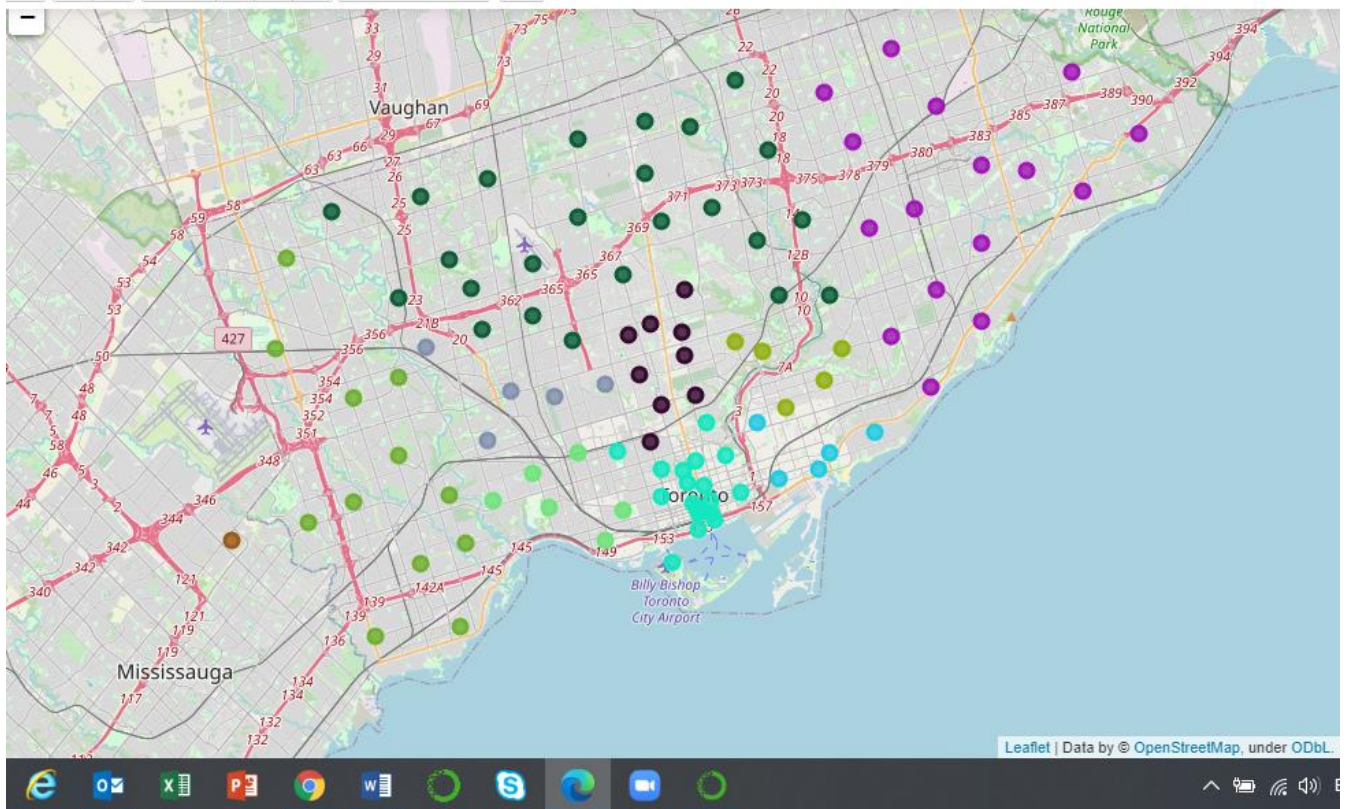


Figure 2: Toronto map with markers

4.4 FOURSQUARE API

The first neighborhood explored in our data set is 'Malvern, Rouge.' There are 273 unique categories, and I illustrated some of the venue categories below. The "Vegetarian / Vegan Restaurant" is available in the list of venues.

Count unique venues

```
1 print('There are {} unique categories.'.format(len(toronto_venues['Venue Category'].unique())))
```

There are 273 unique categories.

```
1 print("The Venue Categories are", toronto_venues['Venue Category'].unique())
```

The Venue Categories are ['Fast Food Restaurant' 'Bar' 'Construction & Landscaping' 'Bank' 'Electronics Store' 'Restaurant' 'Mexican Restaurant' 'Rental Car Location' 'Medical Center' 'Intersection' 'Breakfast Spot' 'Coffee Shop' 'Korean BBQ Restaurant' 'Caribbean Restaurant' 'Hakka Restaurant' 'Thai Restaurant' 'Athletics & Sports' 'Gas Station' 'Bakery' 'Fried Chicken Joint' 'Playground' 'Jewelry Store' 'Smoke Shop' 'Department Store' 'Hobby Shop' 'Train Station' 'Ice Cream Shop' 'Bus Line' 'Metro Station' 'Bus Station' 'Park' 'Soccer Field' 'Motel' 'American Restaurant' 'Café' 'General Entertainment' 'Skating Rink' 'College Stadium' 'Chinese Restaurant' 'Indian Restaurant' 'Pet Store' 'Vietnamese Restaurant' 'Sandwich Place' 'Middle Eastern Restaurant' 'Auto Garage' 'Accessories Store' 'Latin American Restaurant' 'Lounge' 'Clothing Store' 'Italian Restaurant' 'Noodle House' 'Pizza Place' 'Pharmacy' 'Grocery Store' 'Gym' 'Golf Course' 'Pool' 'Mediterranean Restaurant' 'Dog Run' 'Shopping Mall' 'Toy / Game Store' 'Movie Theater' 'Chocolate Shop' 'Salon / Barbershop' 'Juice Bar' 'Burger Joint' 'Theater' 'Liquor Store' 'Video Game Store' 'Food Court' 'Japanese Restaurant' 'Asian Restaurant' 'Cosmetics Shop' 'Burrito Place' 'Sporting Goods Shop' 'Women's Store' 'Mobile Phone Shop' 'Boutique' 'Miscellaneous Shop' 'Tea Room' 'Supplement Shop' 'Shoe Store' 'Luggage Store' 'Greek Restaurant' 'Dessert Shop' 'Spa' 'Convenience Store' 'Baseball Field' 'Martial Arts School' 'Ramen Restaurant' 'Steakhouse' 'Arts & Crafts Store' 'Plaza']

I grouped the "Neighbourhoods" of "Vegetarian / Vegan Restaurant," as illustrated below. "Neighbourhoods" was renamed to "Neighbourhood."

```
1 fd = to_grouped[["Neighbourhoods", "Vegetarian / Vegan Restaurant"]]
2 fd.head()
```

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

```
1 fd = fd.rename(columns={'Neighbourhoods': 'Neighbourhood'})
```

There are a total of 16 locations with Vegetarian / Vegan Restaurants in Toronto.

4.5 K-means clustering

I used the Elbow Point Technique to determine the optimum K value neither overfitting nor underfitting the model; I find out that Elbow Point at K = 4. That means there are a total of 4 clusters.

```
1 kclusters = 4
2
3 toronto_grouped_clustering = fd.drop('Neighbourhood', 1)
4
5 # run k-means clustering
6 kmeans = KMeans(n_clusters=kclusters, random_state=0).fit(toronto_grouped_clustering)
7
8 # check cluster labels generated for each row in the dataframe
9 kmeans.labels_[0:10]
```

```
array([0, 0, 0, 0, 0, 1, 0, 0, 0, 0])
```

```
1 # create a new dataframe that includes the cluster as well as the top 10 venues for each neighborhood.
2 to_merged = fd.copy()
3
4 # add clustering labels
5 to_merged["Cluster Labels"] = kmeans.labels_
6 to_merged.head()
```

	Neighbourhood	Vegetarian / Vegan Restaurant	Cluster Labels
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

5 DATA ANALYSIS

There are no Vegetarian / Vegan Restaurants in cluster 1, as demonstrated below.

```
1 clus1 = to_merged.loc[to_merged['Cluster Labels'] == 0]
2 df_clus1 = pd.merge(df_new, clus1, on='Neighbourhood')
3 df_clus1
```

	Borough	Neighbourhood	Vegetarian / Vegan Restaurant	Cluster Labels	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Scarborough	Malvern, Rouge	0.0	0	43.806686	-79.194353	Wendy's	43.807448	-79.199056	Fast Food Restaurant
1	Scarborough	Rouge Hill, Port Union, Highland Creek	0.0	0	43.784535	-79.160497	SEBS Engineering Inc. (Sustainable Energy and ...)	43.782371	-79.156820	Construction & Landscaping
2	Scarborough	Rouge Hill, Port Union, Highland Creek	0.0	0	43.784535	-79.160497	Royal Canadian Legion	43.782533	-79.163085	Bar
3	Scarborough	Guildwood, Morningside, West Hill	0.0	0	43.763573	-79.188711	Eggsmart	43.767800	-79.190466	Breakfast Spot
4	Scarborough	Guildwood, Morningside, West Hill	0.0	0	43.763573	-79.188711	Lawrence Ave E & Kingston Rd	43.767704	-79.189490	Intersection
...
1251	Etobicoke	South Steeles, Silverstone, Humbergate, Jamest...	0.0	0	43.739416	-79.588437	Shoppers Drug Mart	43.741685	-79.584487	Pharmacy
1252	Etobicoke	Northwest, West Humber - Clairville	0.0	0	43.706748	-79.594054	Saand Rexdale	43.705072	-79.598725	Drugstore
1253	Etobicoke	Northwest, West Humber - Clairville	0.0	0	43.706748	-79.594054	Logistics Distribution	43.707554	-79.589252	Bar

There are 0.030303 Vegetarian / Vegan Restaurants in cluster 2, as demonstrated below.

1	df_clus2.sort_values(["Vegetarian / Vegan Restaurant"], ascending=False)									
	Borough	Neighbourhood	Vegetarian / Vegan Restaurant	Cluster Labels	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
232	West Toronto	Runnymede, Swansea	0.030303	1	43.651571	-79.484450	Fat Bastard Burrito	43.649779	-79.482894	Burrito Place
208	West Toronto	Runnymede, Swansea	0.030303	1	43.651571	-79.484450	Max's Market	43.650525	-79.479145	Gourmet Shop
214	West Toronto	Runnymede, Swansea	0.030303	1	43.651571	-79.484450	Java Joe's	43.650424	-79.479755	Café
213	West Toronto	Runnymede, Swansea	0.030303	1	43.651571	-79.484450	Bloor-Jane Restaurant	43.649513	-79.484070	Diner
212	West Toronto	Runnymede, Swansea	0.030303	1	43.651571	-79.484450	RBC Royal Bank	43.650142	-79.480274	Bank
...
34	Downtown Toronto	Berczy Park	0.018182	1	43.644771	-79.373306	Oyshi Sushi	43.642340	-79.375853	Sushi Restaurant
33	Downtown Toronto	Berczy Park	0.018182	1	43.644771	-79.373306	Brookfield Place	43.646791	-79.378769	Shopping Mall
32	Downtown Toronto	Berczy Park	0.018182	1	43.644771	-79.373306	Scotland Yard Pub	43.647053	-79.374060	Pub
31	Downtown Toronto	Berczy Park	0.018182	1	43.644771	-79.373306	Spirit of Hockey	43.647047	-79.377537	Sporting Goods Shop
0	Downtown Toronto	Berczy Park	0.018182	1	43.644771	-79.373306	The Reservoir Lounge	43.648517	-79.374556	Jazz Club

There are 0.11765 Vegetarian / Vegan Restaurants in cluster 3, as demonstrated below.

1	clus3 = to_merged.loc[to_merged['Cluster Labels'] == 2]									
2	df_clus3 = pd.merge(df_new, clus3, on='Neighbourhood')									
3	df_clus3									
	Borough	Neighbourhood	Vegetarian / Vegan Restaurant	Cluster Labels	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Downtown Toronto	St. James Town	0.011765	2	43.651494	-79.375418	Richmond Station	43.651569	-79.379266	American Restaurant
1	Downtown Toronto	St. James Town	0.011765	2	43.651494	-79.375418	Beerbistro	43.649419	-79.377237	Gastropub
2	Downtown Toronto	St. James Town	0.011765	2	43.651494	-79.375418	The Poet Cafe	43.650637	-79.371276	Café
3	Downtown Toronto	St. James Town	0.011765	2	43.651494	-79.375418	Sukhothai	43.648487	-79.374547	Thai Restaurant
4	Downtown Toronto	St. James Town	0.011765	2	43.651494	-79.375418	Craft Beer Market	43.649872	-79.378398	Beer Bar
...
644	Downtown Toronto	First Canadian Place, Underground city	0.010000	2	43.648429	-79.382280	Bay Adelaide Centre	43.650879	-79.380003	Office
645	Downtown Toronto	First Canadian Place, Underground city	0.010000	2	43.648429	-79.382280	Copacabana Grilled Brazilian	43.648333	-79.388151	Brazilian Restaurant
646	Downtown Toronto	First Canadian Place, Underground city	0.010000	2	43.648429	-79.382280	McEwan Foods	43.647160	-79.381044	Deli / Bodega
647	Downtown Toronto	First Canadian Place, Underground city	0.010000	2	43.648429	-79.382280	Piper's Gastropub	43.645468	-79.381779	Cocktail Bar

When we compare all the clusters (cluster 1, cluster 2, cluster 3, and cluster 4), cluster 4 illustrated below has the maximum average number of Vegetarian / Vegan Restaurants equating to (0.054054). Vegetarian / Vegan Restaurants in cluster 4 is located in Kensington Market and Chinatown, Grange Park.


```

1 clus4 = to_merged.loc[to_merged['Cluster Labels'] == 3]
2 df_clus4 = pd.merge(df_new, clus4, on='Neighbourhood')
3 df_clus4

```

	Borough	Neighbourhood	Vegetarian / Vegan Restaurant	Cluster Labels	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Downtown Toronto	Kensington Market, Chinatown, Grange Park	0.054054	3	43.653206	-79.400049	El Trompo	43.655832	-79.402561	Mexican Restaurant
1	Downtown Toronto	Kensington Market, Chinatown, Grange Park	0.054054	3	43.653206	-79.400049	Essence of Life Organics	43.654111	-79.400431	Organic Grocery
2	Downtown Toronto	Kensington Market, Chinatown, Grange Park	0.054054	3	43.653206	-79.400049	Wanda's Pie in the Sky	43.656163	-79.402630	Bakery
3	Downtown Toronto	Kensington Market, Chinatown, Grange Park	0.054054	3	43.653206	-79.400049	Krispy Kreme Doughnuts	43.655834	-79.399417	Donut Shop
4	Downtown Toronto	Kensington Market, Chinatown, Grange Park	0.054054	3	43.653206	-79.400049	A & C World	43.657409	-79.399847	Gaming Cafe
...
69	Downtown Toronto	Kensington Market, Chinatown, Grange Park	0.054054	3	43.653206	-79.400049	Fudao Noodle House	43.654645	-79.398874	Noodle House
70	Downtown Toronto	Kensington Market, Chinatown, Grange Park	0.054054	3	43.653206	-79.400049	Torteria San Cosme	43.654702	-79.400646	Mexican Restaurant
71	Downtown Toronto	Kensington Market, Chinatown, Grange Park	0.054054	3	43.653206	-79.400049	Saigon Lotus Restaurant	43.654311	-79.399225	Vietnamese Restaurant
72	Downtown Toronto	Kensington Market, Chinatown, Grange Park	0.054054	3	43.653206	-79.400049	Thai Country Kitchen	43.656159	-79.399423	Thai Restaurant
73	Downtown Toronto	Kensington Market, Chinatown, Grange Park	0.054054	3	43.653206	-79.400049	Sonic Boom	43.650859	-79.396985	Record Shop

6 DISCUSSION

In Figure 3, I compared the number of Neighborhoods per Cluster. We see that cluster 1 has the highest neighborhoods (84), while cluster 2 has (7), cluster 3 has (4) neighborhoods, and cluster 4 has only (1).

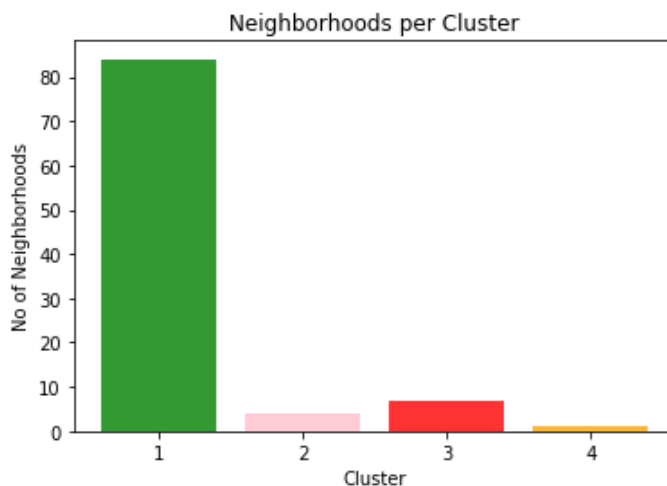


Figure 3: Average Neighbourhoods in each cluster

Figure 4 below illustrates that most Vegetarian / Vegan Restaurants are situated in cluster 4, followed by cluster 3 and cluster 2. Cluster 1 does not have any bar, which demonstrates no

Vegetarian / Vegan Restaurants.

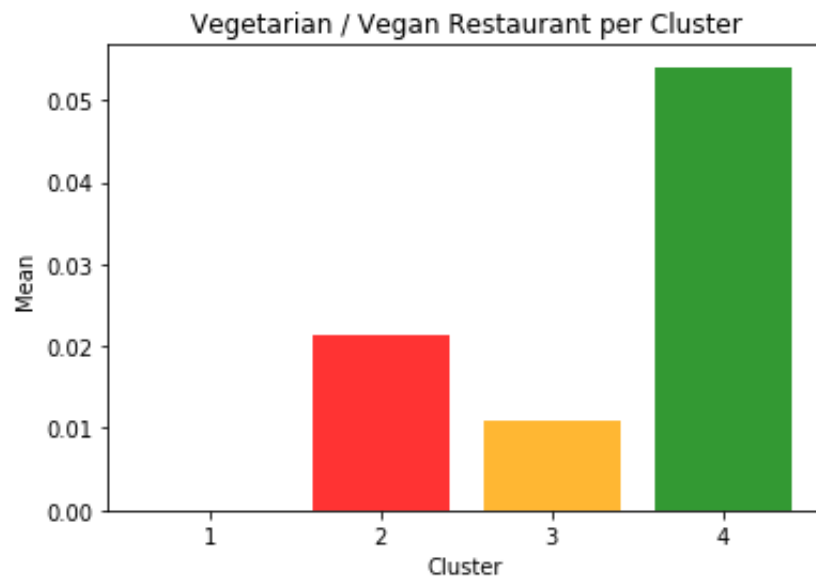


Figure 4: Average Vegetarian/Vegan Restaurants per cluster

7 CONCLUSION

The entrepreneurs can use segmentation and clustering neighborhoods' information as a prototype when exploring new opportunities to open a thriving Vegetarian restaurant business in Toronto. The skills gained on the segmentation and clustering neighborhoods assignment can contribute immensely to the body of knowledge since the study findings can be used as a reference when conducting research. The target audience of the project is academics and non-academic researchers. Academics can use the information to educate learners and enhance quality education. Tourists can study this kind of assignment to understand the area of interest before visiting Toronto. The following step by step instructions from Coursera course work were followed to assess and clean data:

- The Notebook to build the code to scrape the following Wikipedia page,
https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
- The dataframe consists of three columns: PostalCode, Borough, and Neighborhood
- Only process the cells that have an assigned borough. Ignore cells with a borough that is **Not Assigned**.

After removing the cells with Not Assigned, I renamed Postal Code to PostalCode and grouped the data based on Borough using the codes below

I accessed the Foursquare API by adding the credentials (CLIENT_ID, CLIENT_SECRET, and Foursquare version). I used Foursquare to determine various Vegetarian restaurant neighborhood. The Venue data assists in deciding the area, which is best suitable to open a Vegetarian restaurant. The first neighborhood explored in our data set is 'Malvern, Rouge.' There are a total of 16 locations with Vegetarian / Vegan Restaurants in Toronto. Cluster 1 has the highest neighbourhoods (84), while cluster 2 has (7), cluster 3 has (4) neighbourhoods, and cluster 4 has only (1). Most of the Vegetarian / Vegan Restaurants are situated in cluster 4, followed by cluster 3 and cluster 2. Cluster 1 does not have any Vegetarian / Vegan Restaurants; therefore, this an opportunity to establish a new Vegetarian / Vegan Restaurants business in this area. The study was successful because the entrepreneurs will have a clear picture of the location of Vegetarian / Vegan Restaurants in Toronto.

8 REFERENCES

1. DiMarino, A., 2013. *A Comparison of Vegetarian Diets and the Standardized Western Diet in Nutrient Adequacy and Weight Status* (Doctoral dissertation, The Ohio State University).
2. Berich, H.M., 2015. Knowledge and Perceptions of Vegetarian Diets Among College-Aged Students.
3. Humankind Ventures Ltd, 2020. How to Avoid Deficiencies on a Vegan Diet? How to Avoid Deficiencies on a Vegan Diet - Forth Life (forthwithlife.co.uk)
4. Hanna, A.L., 2014. The Global Youth Unemployment Crisis: Exploring Successful Initiatives and Partnering with Youths. *Durham, NC: Sanford School of Public Policy, Duke University. Retrieved August, 14, p.2018.*
5. Khan, 2020, Battle of the neighborhoods. From: <https://towardsdatascience.com/battle-of-the-neighborhoods-b72ec0dc76b5> (Accessed 06 December 2020).