

# Final Capstone Project

## Analysis of neighbourhoods in North York

Georgios Papanikolaou

### 1.1 Introduction:

North York is an eclectic, multicultural district home to the hands-on Ontario Science Centre and the Aga Khan Museum, with exhibits on Islamic culture in a striking modern building. In the area's north, Black Creek Pioneer Village is an 1800s living museum. Sprawling Downsview Park includes a lake, event spaces, and a flea and farmers' market, while Edwards Gardens has a greenhouse, fountains, and botanic gardens. In this district there is also the Earl Bales Park which is a very large park. The area is well developed and therefore many new businesses are interested in investing in this specific region therefore the market is highly competitive. *[Wikipedia]*

### 1.2 Problem Description

A Mediterranean restaurant named 'Mediterrano' is interested to open its new business unit in North York but first they have requested from us an analysis in the different neighborhoods of North York and the most common venues in order to find the optimal location to open their new restaurant. 'Mediterrano' it's a very successful restaurant in Ontario and now they want to be sure that they will continue their success also in North York.

### 1.3 Interest

The management team of 'Mediterrano' are the main people of interest of this research. Other businesses which might be interested in investing in this region and belonging in the food industry belong could also be benefit from this analysis.

## 2. Data acquisition and cleaning

### 2.1 Data Sources

Our goal is to segment and cluster the neighborhoods of North York in the city of Toronto as those will be used to analyze the existing venues in the region of North York.

- i. [Postal codes of Canada](#). This Wikipedia page was used in order to scrap the postal codes of the whole city of Toronto.
- ii. Geospatial\_Coordinates.csv was provided from Applied Data Science Capstone course from IBM in coursera which contains the geospatial coordinates of the different regions of Toronto.
- iii. Venue's data was extracted using an API call from Foursquare

## 2.2 Data cleaning

Two tables were created using the data from i & ii sources. The first table included the postal codes, the boroughs and the neighborhoods of the city of Toronto and the second one the postal codes and the geospatial coordinates. The purpose was to combine those two tables but there were some differences in the naming of the columns so there was some data processing that was required in order to make the tables consistent and combine them.

After this step the source iii was used in order to extract the Venues data and combine it with the other datasets in order to create a homogeneous dataset which will later on be used for the k-means clustering technique.

## 2.3 Feature selection

The features that will be used have been collected from the location provider Foursquare.com, which we used to extract data for the different types of venues that exist in every neighborhood. The data for every neighborhood consisted of the features below:

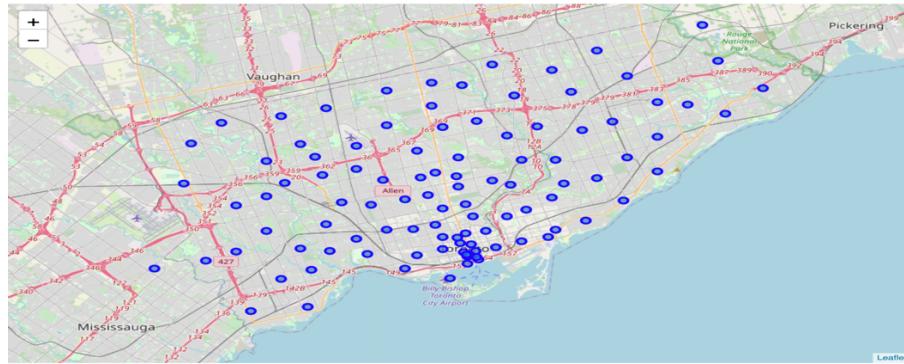
- Neighborhood
- Neighborhood Latitude
- Neighborhood Longitude
  - Venue
  - Venue Latitude
  - Venue Category

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Hillcrest Village	43.803762	-79.363452	New York Fries	43.803664	-79.363905	Fast Food Restaurant
1	Hillcrest Village	43.803762	-79.363452	Eagle's Nest Golf Club	43.805455	-79.364186	Golf Course
2	Hillcrest Village	43.803762	-79.363452	AY Jackson Pool	43.804515	-79.366138	Pool
3	Hillcrest Village	43.803762	-79.363452	Villa Madina	43.801685	-79.363938	Mediterranean Restaurant
4	Hillcrest Village	43.803762	-79.363452	Duncan Creek Park	43.805539	-79.360695	Dog Run

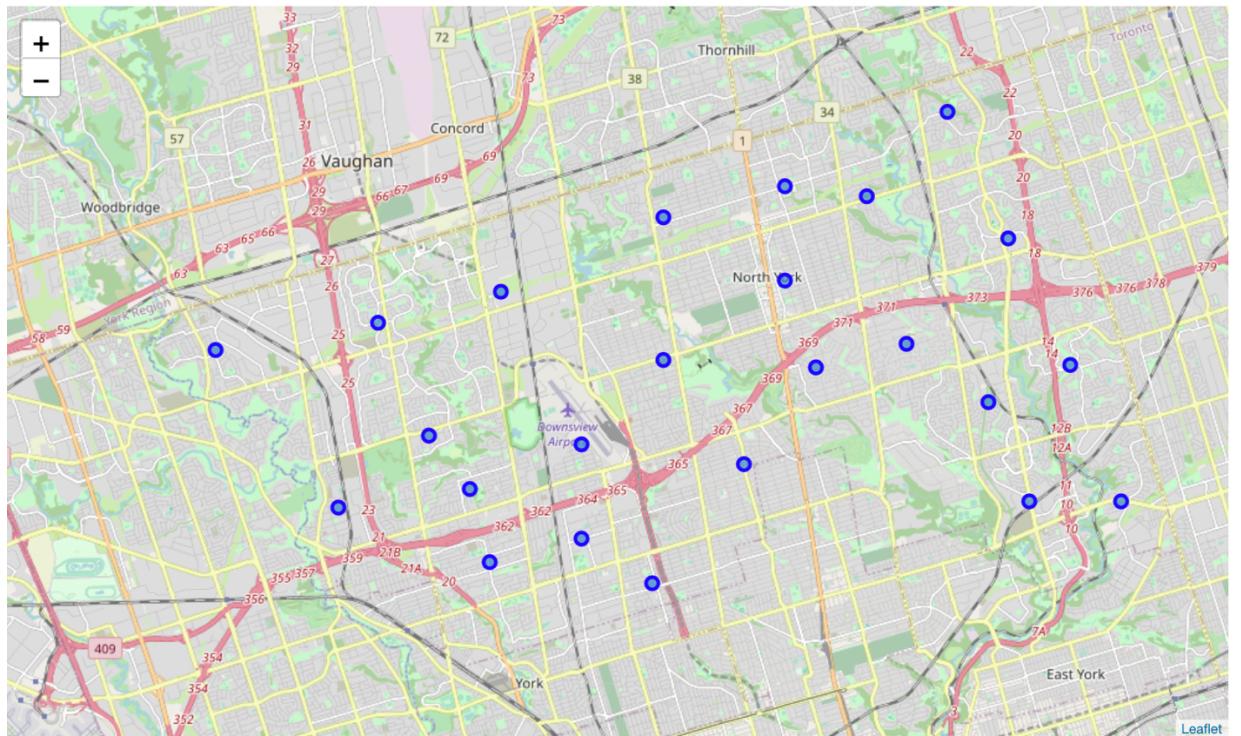
### 3.0 Methodology

First step was the data scrapping from the Wikipedia that contains postal codes of the whole city of Toronto and their coordinates.

Additionally we need to find the coordinates and locations of North York, and therefore the geocoder API has been used for retrieving this data. This is an important step as this data will be used as an input in the Foursquare API call to gain access in the dataset with the venues across the different neighborhoods of North York.



In order to start our analysis in North York we needed to merge the data from both sources (Wikipedia+Foursquare) and then map them using folium.



A total of 24 unique neighborhoods in the district of North York were found and those are listed below:

```
[ 'Hillcrest Village' 'Fairview, Henry Farm, Oriole' 'Bayview Village'  
  'York Mills, Silver Hills' 'Willowdale, Newtonbrook' 'Willowdale South'  
  'York Mills West' 'Willowdale West' 'Parkwoods' 'Don Mills North'  
  'Don Mills South' 'Bathurst Manor, Wilson Heights, Downsview North'  
  'Northwood Park, York University' 'Downsview East' 'Downsview West'  
  'Downsview Central' 'Downsview Northwest' 'Victoria Village'  
  'Bedford Park, Lawrence Manor East' 'Lawrence Manor, Lawrence Heights'  
  'Glencairn' 'North Park, Maple Leaf Park, Upwood Park' 'Humber Summit'  
  'Humberlea, Emery' ]
```

The information for the different venues of North York were found using Foursquare. We retrieved data which showed 99 unique venue categories in the different neighborhoods of North York.

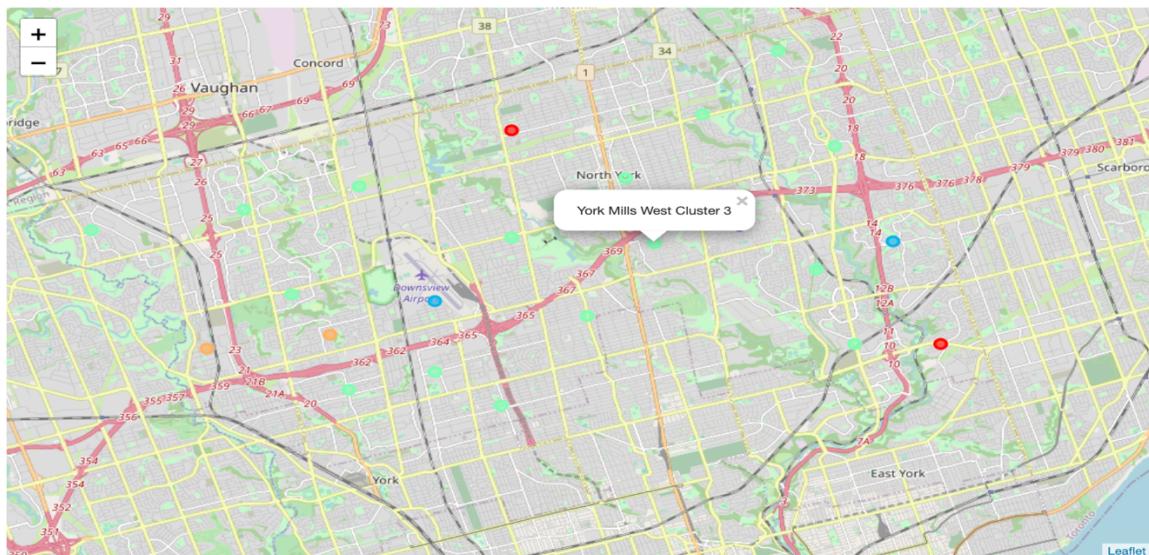
Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	
0	Hillcrest Village	43.803762	-79.363452	New York Fries	43.803664	-79.363905	Fast Food Restaurant
1	Hillcrest Village	43.803762	-79.363452	Eagle's Nest Golf Club	43.805455	-79.364186	Golf Course
2	Hillcrest Village	43.803762	-79.363452	AY Jackson Pool	43.804515	-79.366138	Pool
3	Hillcrest Village	43.803762	-79.363452	Villa Madina	43.801685	-79.363938	Mediterranean Restaurant
4	Hillcrest Village	43.803762	-79.363452	Duncan Creek Park	43.805539	-79.360695	Dog Run

There are 99 uniques categories.

K-Means which is a machine learning technique, was used to segment and cluster these neighborhoods so that we can group them together to understand their similarities and to be able to find which is the optimal location for the new business unit of ‘Mediterrano’. This is an important step as the use of this algorithm will give us a better look in the demand of the region.

## 4.0 Results

Using the K-means clustering we were able to find the top 5 clusters of similar neighborhoods in North York. Those neighborhoods are grouped using similar venues in every neighborhood. This data will help us better estimate the best location for the ‘Mediterrano’ to open its new business unit.



With Foursquare API we managed to collect the data of the most common venues in every neighborhood which are listed in the table below.

Postcode	Latitude	Longitude	Borough	Neighborhood	Cluster Label	1st Most Common	2nd Most Common	3rd Most Common	4th Most Common	5th Most Common	6th Most Common	7th Most Common	8th Most Common	9th Most Common	10th Most Common	
0	M2H 43.8038	-79.36	North York	Hillcrest Village	3	Golf Course	Mediterranean Restaurant	Fast Food Restaurant	Pool	Dog Run	Deli / Bodega	Clothing Store	Coffee Shop	Comfort Food Restaurant	Construction & Landscaping	
1	M2J 43.7865	-79.35	North York	North York Fair	3	Clothing Store	Japanese Restaurant	Fast Food Restaurant	Pool	Japanese Restaurant	Japanese Restaurant	Bank	Bar	Vintage	Sporting Goods Shop	
2	M2K 43.7865	-79.35	North York	Bayview Village	2	Chinese Rest	Japanese Restaurant	Cafe	Japanese Restaurant	Japanese Restaurant	Coffee Shop	Comfort Food & Construction & Lands Convenience Store	Cosmetics Shop	Cupcake Shop		
3	M2L 43.7575	-79.37	North York	York Mills, Silver Hil	1	Cafeteria	Women's Store	Dim Sum Restaurant	Clothing Store	Coffee Shop	Comfort Food & Construction & Lands Convenience Store	Cosmetics Shop	Cupcake Shop			
5	M2N 43.7701	-79.41	North York	Willowdale South	3	Ramen Restau	Sushi Restaurant	Pizza Place	Sandwich Place	Coffee Shop	Japanese Restaurant	Cafe	Lounge	Ice Cream Shop	Plaza	
6	M2P 43.7528	-79.4	North York	York Mills West	3	Park	Construction & Landscaping	Convenience Store	Dim Sum Restaurant	Chocolate Shop	Clothing Store	Coffee Shop	Comfort Food Restaurant	Cosmetics Shop	Cupcake Shop	
7	M2R 43.7827	-79.44	North York	Willowdale West	0	Park	Pharmacy	Pizza Place	Coffee Shop	Women's Store	Deli / Bodega	Chocolate Shop	Clothing Store	Comfort Food Restaurant	Construction & Landscaping	
8	M3A 43.7533	-79.33	North York	Parkwood	2	Park	Food & Drink Shop	Dim Sum Restaurant	Chocolate Shop	Clothing Store	Coffee Shop	Comfort Food Restaurant	Construction & Land Convenience Store	Cosmetics Shop		
9	M3B 43.7616	-79.33	North York	Don Mills North	3	Japanese Restau	Caribbean Restaurant	Cafe	Gym	Dim Sum Restaurant	Coffee Shop	Comfort Food Restaurant	Construction & Land Convenience Store	Cosmetics Shop		
10	M3C 43.7259	-79.34	North York	Don Mills South	0	Restaurant	Gym	Coffee Shop	Bike Shop	Sandwich Place	Dim Sum Restaurant	Discount Store	Art Gallery	Beer Store	Clothing Store	
11	M3D 43.7543	-79.44	North York	Bathurst Manor, V	3	Coffee Shop	Bank	Sandwich Place	Ice Cream Shop	Health Food Store	Park	Pet Store	Pharmacy	Pizza Place	Bridal Shop	
12	M3E 43.7543	-79.44	North York	Wilson Heights, V	0	Intersection	Home	Caribbean Restaurant	Coffee Shop	Macaroni & Cheese	Miscellaneous Store	Electronics Store	French Onion Joint	Clothing Store		
13	M3F 43.7375	-79.46	North York	Downsview East	2	Airport	Park	Business Service	Women's Store	Dim Sum Restaurant	Coffee Shop	Comfort Food Restaurant	Construction & Land Convenience Store	Cosmetics Shop		
14	M3G 43.7339	-79.46	North York	Downsview West	3	Grocery Store	Park	Bank	Hotel	Shopping Mall	Women's Store	Department Store	Clothing Store	Coffee Shop	Comfort Food Restaurant	
15	M3M 43.7285	-79.5	North York	Downsview Central	4	Food Truck	Baseball Field	Women's Store	Dim Sum Restaurant	Clothing Store	Coffee Shop	Comfort Food Restaurant	Construction & Land Convenience Store	Cosmetics Shop		
16	M3N 43.7616	-79.5	North York	Downsview North	3	Grocery Store	Gym / Fitness Center	Athletics & Sports	Liquor Store	Women's Store	Dim Sum Restaurant	Coffee Shop	Comfort Food Restaurant	Construction & Land Convenience Store	Cosmetics Shop	
17	M4A 43.7259	-79.52	North York	Victoria Village	0	Coffee Shop	Pizza Place	Hockey Arena	Portuguese Restau	Women's Store	Department Store	Clothing Store	Comfort Food Restaurant	Construction & Land Convenience Store	Cosmetics Shop	
18	M4A 43.7259	-79.52	North York	Bedford Park, Law	3	Italian Restauran	Coffee Shop	Pizza Place	Sandwich Place	Women's Store	Restaurant	Café	Butcher	Fast Food Restaurant	Indian Restaurant	
19	M4A 43.7185	-79.46	North York	Lawrence Manor, L	3	Clothing Store	Accessories Store	Show Store	Vietnamese Restaurant	Furniture / Home Store	Coffee Shop	Boutique	Gift Shop	Bakery	Distribution Center	
20	M6B 43.7096	-79.45	North York	Glencain	3	Italian Restauran	Pizza Place	Bakery	Japanese Restaurant	Women's Store	Coffee Shop	Comfort Food Restaurant	Construction & Land Convenience Store	Cosmetics Shop		
21	M6L 43.7138	-79.49	North York	North Park, Maple	3	Park	Massage Studio	Construction & Land Bakery	Women's Store	Department Store	Clothing Store	Coffee Shop	Comfort Food Restaurant	Convenience Store		
22	M9L 43.7563	-79.57	North York	Humber Summit	3	Furniture / Hom	Pizza Place	Intersection	Dim Sum Restaurant	Clothing Store	Coffee Shop	Comfort Food Restaurant	Construction & Land Convenience Store	Cosmetics Shop		
23	M9M 43.7248	-79.53	North York	Humberley, Emery	4	Construction &	Baseball Field	Women's Store	Diner	Clothing Store	Coffee Shop	Comfort Food Restaurant	Convenience Store	Cosmetics Shop	Cupcake Shop	

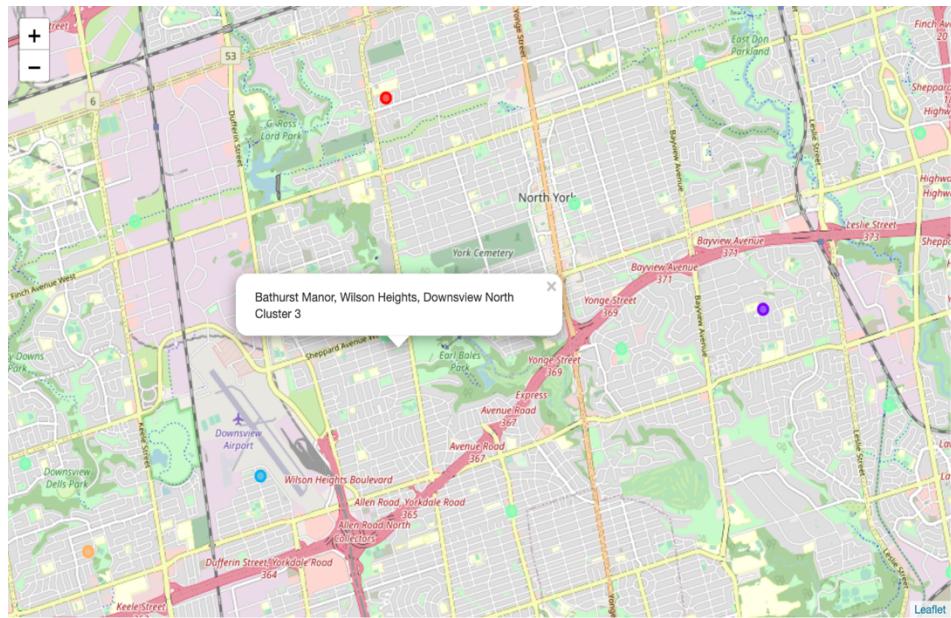
If we compare the results of the different clusters, we understand that **Cluster 3** is the most dominant with many different choices and especially in the Food industry. Therefore, we should find a district in this cluster where the competition is lower, and the new restaurant will have more advantages and a faster growth rate which is important in the very first steps.

## 5.0 Discussion

Comparing the number of restaurants in each neighborhood of cluster we recommend a neighborhood where the supply is lower and 'Mediterrano' will be able to succeed even faster. The neighborhood which is selected as the optimal is Bathurst Manor, Wilson Heights, Downsview North.

Postcode	Latitude	Longitude	Borough	Neighborhood	Cluster Label	1st Most Common	2nd Most Common	3rd Most Common	4th Most Common	5th Most Common	6th Most Common	7th Most Common	8th Most Common	9th Most Common	10th Most Common
0	M2H 43.80376	-79.36345	North York	Hillcrest Village	3	Golf Course	Mediterranean Restaurant	Fast Food Restaurant	Pool	Dog Run	Deli / Bodega	Clothing Store	Coffee Shop	Comfort Food Construction & Landscaping	
1	M2J 43.77852	-79.34556	North York	Bathurst Manor, Oriole	3	Clothing Store	Japanese Restaurant	Fast Food Restaurant	Pool	Japanese Restaurant	Japanese Restaurant	Bank	Bakery	Women's Store	Sporting Goods Shop
2	M2K 43.78695	-79.3598	North York	Bayview Village	3	Chinese Resta	Japanese Rest	Cafe	Bank	Coffee Shop	Comfort Food Construction	Convenience	Cosmetics Shop	Cupcake Shop	
5	M2N 43.77012	-79.40849	North York	Willowdale South	3	Ramen Restau	Sushi Restaurant	Pizza Place	Sandwich Place	Coffee Shop	Japanese Restaurant	Cafe	Lounge	Ice Cream Shop	Plaza
9	M3B 43.74591	-79.35219	North York	Don Mills North	3	Japanese Rest	Caribbean Restaurant	Gym	Coffee Shop	Dim Sum Restaurant	Comfort Food Restaurant	Construction	Convenience	Cosmetics Shop	
10	M3C 43.7259	-79.3529	North York	Don Mills South	3	Restaurant	Gym	Coffee Shop	Bike Shop	Sandwich Place	Plai Dim Sum Rest	Discount Stor	Art Gallery	Beer Store	Clothing Store
11	M3H 43.7543	-79.44226	North York	Bathurst Manor, Wilson Heights, Downsview North	3	Coffee Shop	Bank	Sandwich Place	Ice Cream Shop	Health Food S	5 Park	Pet Store	Pharmacy	Pizza Place	Bridal Shop
12	M3I 43.76198	-79.48726	North York	Northwest Park, York University	3	Furniture / Hc	Ball	Caribbean Rest	Coffee Shop	Message Shop	Miscellaneous	Fast Food Rest	Electronics Store	Stuffed Chicken	Clothing Store
14	M3L 43.7622	-79.50594	North York	Downsview West	3	Grocery Store	Park	Bank	Women's Store	Dim Sum Rest	Comfort Food Restaurant	Construction	Convenience	Convenience Store	
16	M3N 43.75163	-79.5	North York	Downsview Northwest	3	Grocery Store	Gym / Fitness	Athletics & Sports	Women's Store	Dim Sum Rest	Coffee Shop	Comfort Food	Construction	Convenience	Convenience Store
18	M3M 43.73228	-79.41975	North York	Bedford Park, Lawrence Manor East	3	Italian Restau	Coffee Shop	Pizza Place	Sandwich Place	Plai Women's Store	Restaurant	Cafe	Butcher	Fast Food Rest	Indian Restaurant
19	M6A 43.71852	-79.46476	North York	Lawrence Manor, Lawrence Heights	3	Clothing Store	Accessories St	Shoe Store	Vietnamese Ri	Furniture / Hc	Coffee Shop	Butique	Gift Shop	Bakery	Distribution Center
20	M6B 43.70958	-79.44507	North York	Glencain	3	Italian Restau	Pizza Place	Bakery	Japanese Rest	Women's Store	Coffee Shop	Comfort Food	Construction	Convenience	Cosmetics Shop
21	M6L 43.7137	-79.49007	North York	North Park, Maple Leaf Park, Upwood Park	3	Park	Massage Studio	Construction	Bakery	Women's Store	Department S	Clothing Store	Coffee Shop	Comfort Food	Convenience Store
22	M9L 43.7563	-79.56396	North York	Humber Summit	3	Furniture / Hc	Pizza Place	Intersection	Dim Sum Rest	Clothing Store	Coffee Shop	Comfort Food Construction	Convenience	Cosmetics Shop	

In the map below you can find where exactly those neighborhoods are located.



The existing venues in this location are listed below.

[Bathurst Manor, Wilson Heights, Downsview North](#)

3 Coffee Shop | Bank | Sandwich Place | Ice Cream Shop | Health Food Store | Park | Pet Store | Pharmacy | Pizza Place | Bridal Shop

## 6.0 Conclusion

With the information which was discussed previously we conclude that our best recommendation for 'Mediterrano' is to start their new business unit in the Bathurst Manor, Wilson Heights, Downsview North.

- Region: 3<sup>rd</sup> Cluster
- Neighborhood names: Bathurst Manor, Wilson Heights, Downsview North.
- Other offers in the district: Sandwich Place, Ice Cream Shop, Health Food Store, Park, Pet Store, Pharmacy, Pizza Place

Supplier in the same district offer primarily fast food, therefore this specific region is missing of a restaurant especially that, that can offer Mediterranean food. It is also important that in the region there is also Earl Bales Park which is a large city Park which we strongly believe that it will help 'Mediterrano' to attract new customers.