The perfect location for the perfect restaurant

1. Introduction

An important restauration group is thinking about to open a restaurant in the city of Toronto. This group has developed different types of restaurants around the world and have the knowledge to set up any restaurant type: Italian, Belgian, Thai, Chinese, etc. However, their knowledge about Toronto area and their citizens likes is none.

The idea is to define a data science methodology, that by using data from Foursquare API, makes possible to find:

- Which type of restaurants better fits the customer likes in Toronto?
- Where is the best location to set this new restaurant?

This methodology could be implemented in other cities and by others restauration groups.

2. Data

The data from Toronto will be picked from different sources:

- Wikipedia: List of postal codes of Canada.
  'https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M'
- Geospatial data of these neighbourhoods.
  'https://cocl.us/Geospatial_data'
- Using Foursquare API:
  - The recommended restaurants in each neighbourhood in an area of radius 750 meters and its name, latitude, longitude and venue category.
  - The likes that people give to each of these restaurants.

Then, the dataset used is a list of the restaurants of the different neighbourhoods. The dataset has the next attributes:

- Neighbourhood name.
- Neighbourhood latitude.
- Neighbourhood longitude.
- Venue name.
- Venue Latitude.
- Venue Longitude.
- Venue Category: Pizza Place, Thai Restaurant, etc.
- Venue number of likes.

Example:

|	Neighbourhood	Neighbourhood Latitude	Neighbourhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category	Venue likes
6	Regent Park, Harbourfront	43.65426	-79.360636	Roselle Desserts	43.653447	-79.362017	Bakery	48
7	Regent Park, Harbourfront	43.65426	-79.360636	Impact Kitchen	43.656369	-79.356980	Restaurant	51
8	Regent Park, Harbourfront	43.65426	-79.360636	Morning Glory Cafe	43.653947	-79.361149	Breakfast Spot	36
9	Regent Park, Harbourfront	43.65426	-79.360636	Souk Tabule	43.653756	-79.354390	Mediterranean Restaurant	53
10	Regent Park, Harbourfront	43.65426	-79.360636	Figs Breakfast & Lunch	43.655675	-79.364503	Breakfast Spot	55

Jorge Quintero Bermejo

3. Data exploratory analysis

At first, the restaurants with less than 20 likes has been neglected. As they have low likes values, it means that they are not remarkably interesting for the customers.

In addition, neighbourhoods with less than 5 restaurants are neglected. These areas are not considered potential to attract customers.

Then, it has been detected that the neighbourhoods with more restaurants are:

- 'Kensington Market, Chinatown, Grange Park' with 69 restaurants.
- 'Richmond, Adelaide, King' with 63 restaurants.
- 'Commerce Court, Victoria Hotel' and 'Toronto Dominion Centre, Design Exchange' with 61 restaurants.

Number of restaurants in studied Toronto neighbourhoods
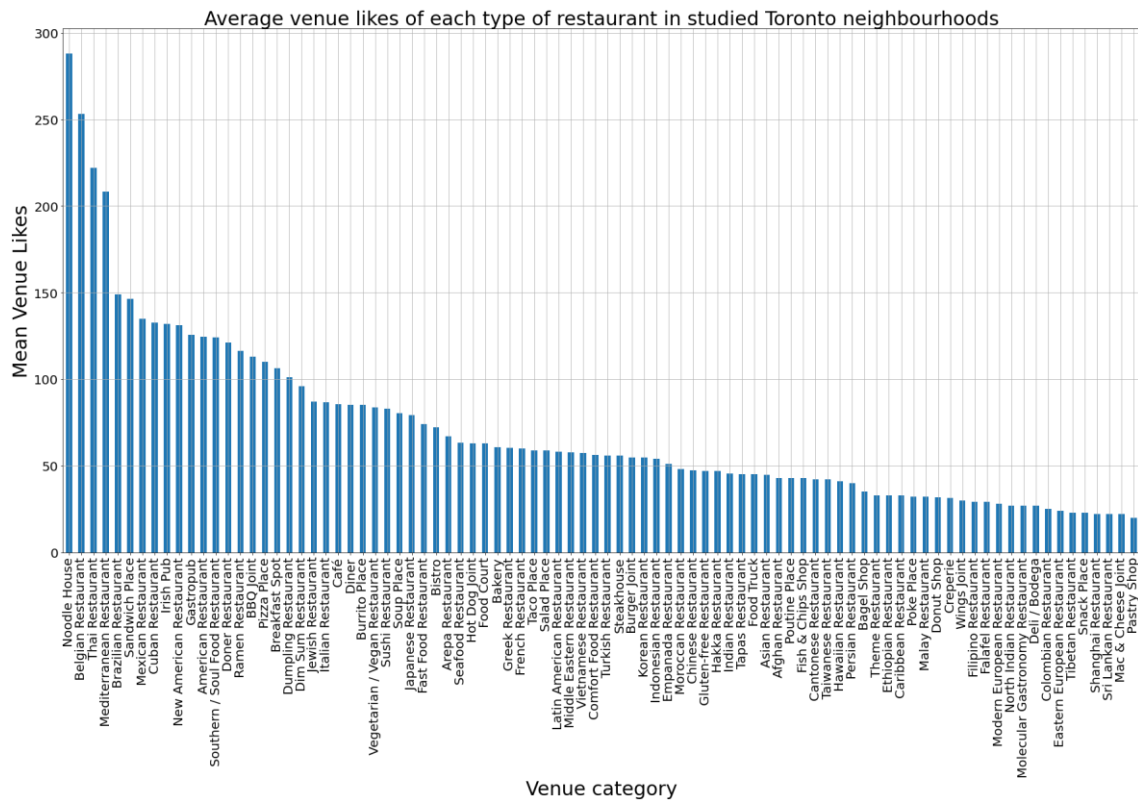
Jorge Quintero Bermejo
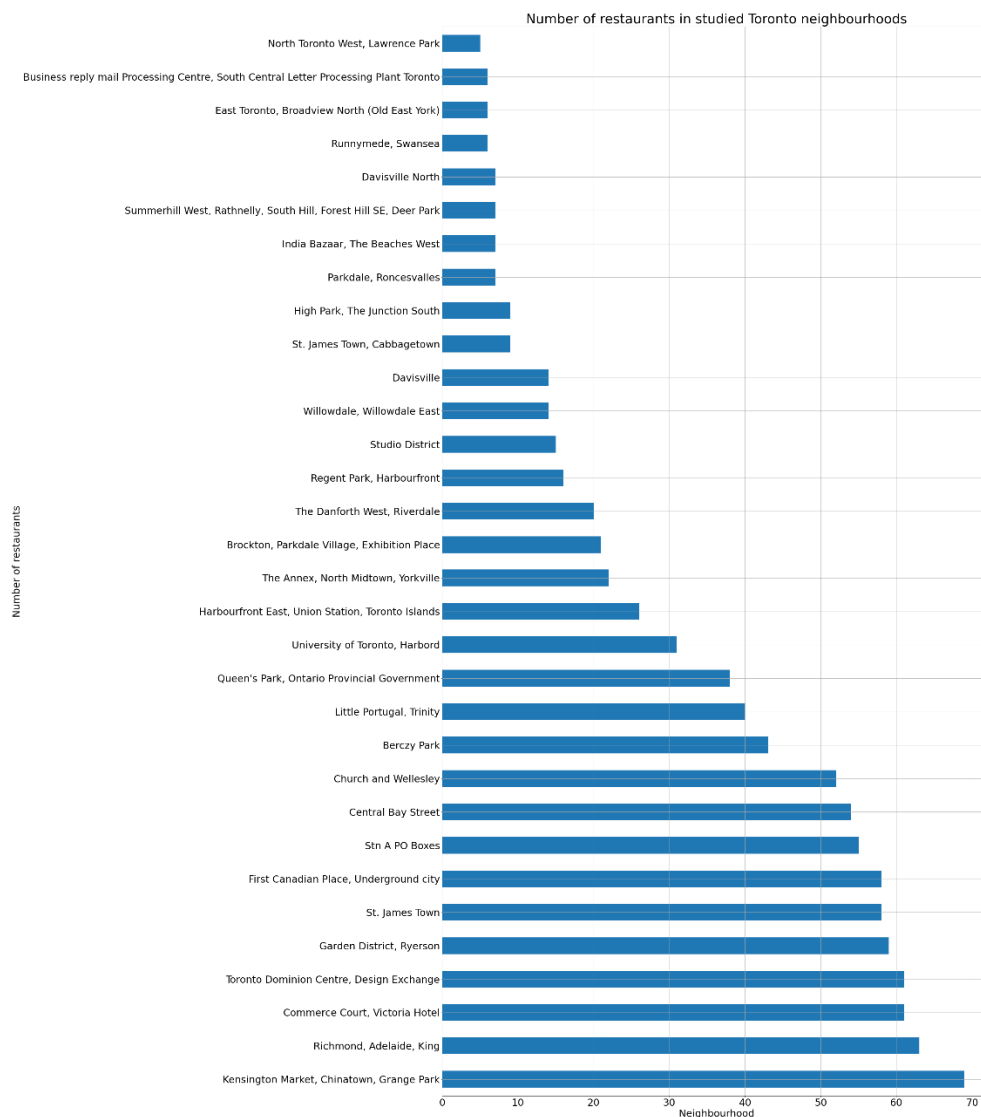
The restaurant categories that more can be found in Toronto are:

- 'Café' with 114 restaurants.
- 'Italian restaurant' with 66 restaurants.
- 'Japanese Restaurant' with 58 restaurants.
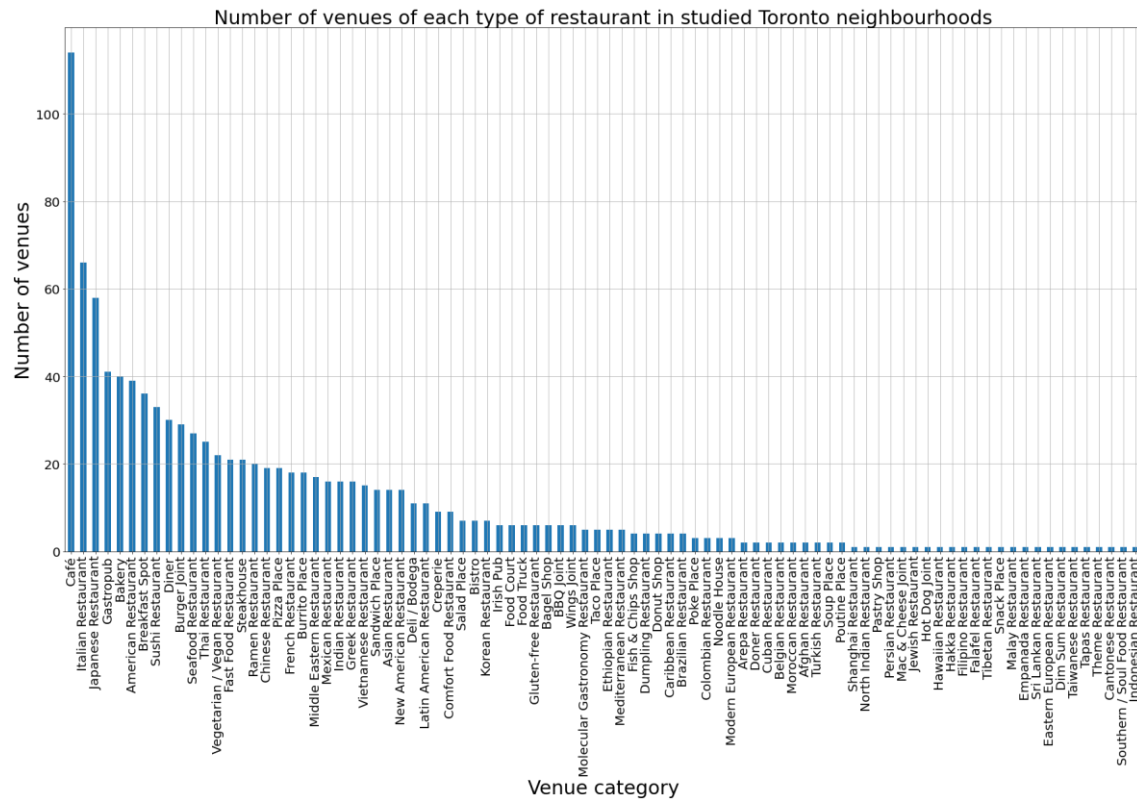
However, the restaurants that have more average likes per restaurant category are:

- 'Noddle house' with 288 likes.
- 'Belgian restaurant' with 253 likes.
- 'Thai restaurant' with 222 restaurants.

Jorge Quintero Bermejo

Average venue likes of each type of restaurant in studied Toronto neighbourhoods

Going through a scatter plot, the relationship between number of venues and mean venue likes of each restaurant category is studied.

Mean Venue likes vs Number of venues
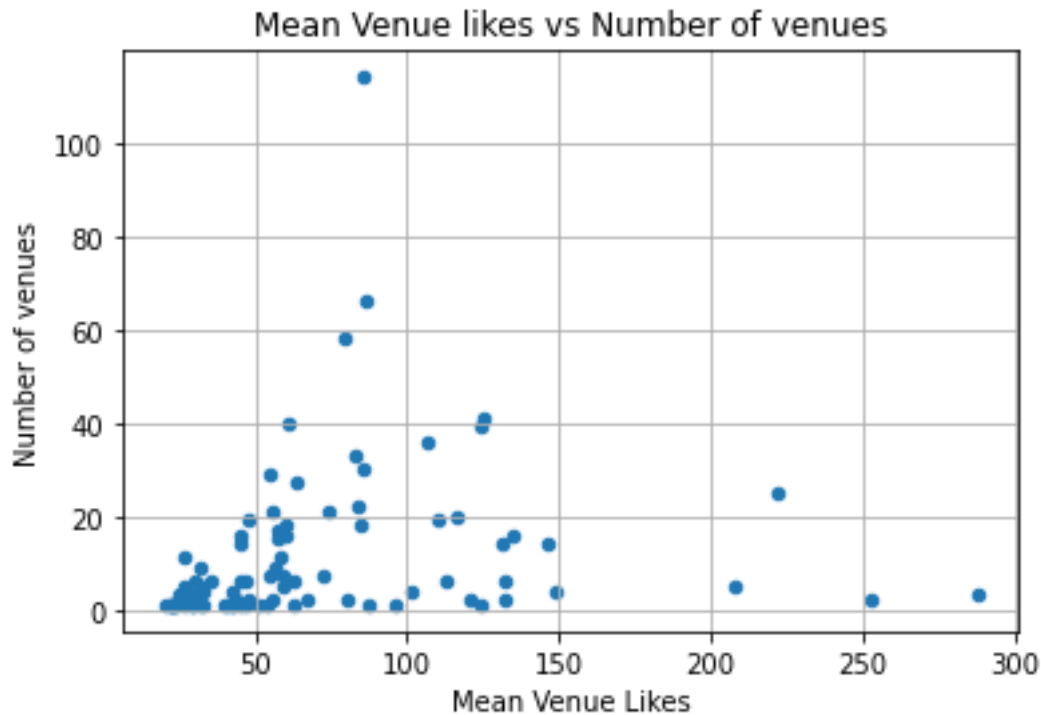
- Café is the most typical venue category (114 restaurants) with an average number of likes of 85.

Jorge Quintero Bermejo

- Noddle house, Belgian Restaurant and Mediterranean restaurant have an average number of likes of more than 200, whereas there are very few of them. (3, 2 and 5 respectively)
- Thai restaurants have an average number of likes higher than 200 while there are 25 of this type in Toronto.
- Gastro bar and American restaurants are 2 solid categories as they have 120 mean venue likes and 40 restaurants.

4. K-means clustering

K-means clustering with k=5, considering mean likes and number of restaurants per neighbourhood (features to apply the k-means algorithm) is done to analyse the relationship between likes, number of restaurants and which categories are most popular in each neighbourhood. This unsupervised technique is considered as attractive method to see if there is relationship between geospatial data and number of venues and likes.

- Cluster 0: (Red) Medium mean venue likes and medium number of venues.

	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	Venue likes	Number of venues
1	Queen's Park, Ontario Provincial Government	Café	Japanese Restaurant	Diner	65.368421	38
4	Berczy Park	Café	Italian Restaurant	Japanese Restaurant	80.465116	43
9	Little Portugal, Trinity	Café	Bakery	Italian Restaurant	80.900000	40
23	University of Toronto, Harbord	Thai Restaurant	Bakery	Sushi Restaurant	60.516129	31
30	Church and Wellesley	Japanese Restaurant	Café	Sandwich Place	70.653846	52

- Cluster 1: (Purple) Low mean venue likes and low number of venues.

	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	Venue likes	Number of venues
7	East Toronto, Broadview North (Old East York)	Café	Pizza Place	Turkish Restaurant	52.000000	6
10	The Danforth West, Riverdale	Greek Restaurant	Café	Burger Joint	58.300000	20
13	India Bazaar, The Beaches West	Indian Restaurant	Fast Food Restaurant	Asian Restaurant	59.142857	7
15	Studio District	Bakery	Diner	Latin American Restaurant	48.800000	15
16	Willowdale, Willowdale East	Ramen Restaurant	Sushi Restaurant	Korean Restaurant	46.357143	14
17	Davisville North	Café	Indian Restaurant	Vegetarian / Vegan Restaurant	48.142857	7
18	High Park, The Junction South	Café	Indian Restaurant	Vietnamese Restaurant	59.888889	9
19	North Toronto West, Lawrence Park	Café	Diner	Ramen Restaurant	45.200000	5
20	The Annex, North Midtown, Yorkville	Italian Restaurant	Thai Restaurant	Wings Joint	58.318182	22
22	Davisville	Café	Indian Restaurant	Taco Place	41.785714	14
24	Runnymede, Swansea	Café	American Restaurant	Italian Restaurant	34.333333	6
26	Summerhill West, Rathnelly, South Hill, Forest...	Cantonese Restaurant	American Restaurant	Italian Restaurant	42.857143	7

- Cluster 2: (Blue) High mean venue likes and high number of venues.

	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	Venue likes	Number of venues
2	Garden District, Ryerson	Japanese Restaurant	Diner	Italian Restaurant	119.745763	59
6	Richmond, Adelaide, King	Café	Japanese Restaurant	Italian Restaurant	117.079365	63
29	First Canadian Place, Underground city	Café	American Restaurant	Japanese Restaurant	112.982759	58

- Cluster 3: (Green) Medium mean venue likes and low number of venues.

Jorge Quintero Bermejo

	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	Venue likes	Number of venues
0	Regent Park, Harbourfront	Italian Restaurant	Café	Bakery	90.125000	16
8	Harbourfront East, Union Station, Toronto Islands	Italian Restaurant	Café	American Restaurant	77.230769	26
12	Brockton, Parkdale Village, Exhibition Place	Café	Bakery	Burrito Place	70.904762	21
21	Parkdale, Roncesvalles	Pizza Place	American Restaurant	Cuban Restaurant	90.857143	7
28	St. James Town, Cabbagetown	Diner	Japanese Restaurant	Gastropub	69.666667	9
31	Business reply mail Processing Centre, South C…	Burrito Place	Fast Food Restaurant	Italian Restaurant	71.500000	6

- Cluster 4: (Orange) Medium mean venue likes and high number of venues.

|	Neighbourhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	Venue likes	Number of venues
3	St. James Town	Café	Italian Restaurant	Gastropub	95.275862	58
5	Central Bay Street	Café	Japanese Restaurant	Chinese Restaurant	101.592593	54
11	Toronto Dominion Centre, Design Exchange	Café	Japanese Restaurant	Italian Restaurant	103.327869	61
14	Commerce Court, Victoria Hotel	Café	Japanese Restaurant	American Restaurant	104.377049	61
25	Kensington Market, Chinatown, Grange Park	Café	Vegetarian / Vegan Restaurant	Mexican Restaurant	84.347826	69
27	Stn A PO Boxes	Café	American Restaurant	Gastropub	91.927273	55
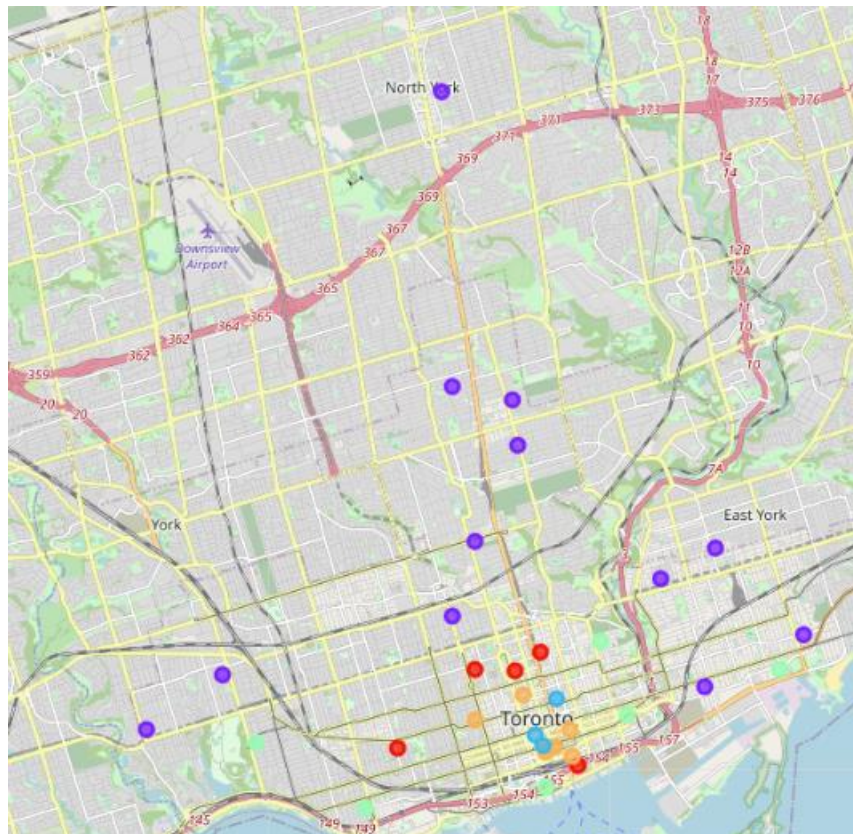
5. Results and discussion

From the exploratory analysis, it can be concluded that:

- Regarding the type of restaurant that is better to set up to achieve success, different categories can be considered:
  - It seems that noodle house, Belgian Restaurant and Mediterranean restaurant are attractive as they have high number mean venue likes (>200 likes). On the other hand, there are few businesses that are dedicated to this (<5 restaurants).

Jorge Quintero Bermejo

- A solid business could be a gastrobar and/or American restaurant, as they have high mean venue likes (125 likes) and different businesses dedicated to this (40 restaurants). It is a type of restaurant that is going to work, but maybe it is hard to make the difference.
- It seems that Thai restaurant are extraordinarily successful in Toronto. They have more than 200 likes per average and they are consolidated as there are more than 20 restaurants of these types.

Regarding the location of the restaurant, the most attractive neighbourhoods to set up a restaurant are:

- 'Kensington Market, Chinatown, Grange Park' with 69 restaurants, 'Richmond, Adelaide, King' with 63 restaurants, 'Commerce Court, Victoria Hotel' and 'Toronto Dominion Centre, Design Exchange' with 61 restaurants, considering only the places that most venues have.
- Going through k-means and the relationship between venue likes and number of restaurants, that restaurants in cluster 2(blue) are potentially more successful than others. Looking at the next picture, when closer to the centre (black x/ blue dots), the restaurants have more average number of likes and the number of restaurants are higher in the neighbourhoods. When further from the centre in all directions, first the number of likes decreases (orange dots) and after, the number of restaurants in the area decreases (red dots).

Jorge Quintero Bermejo
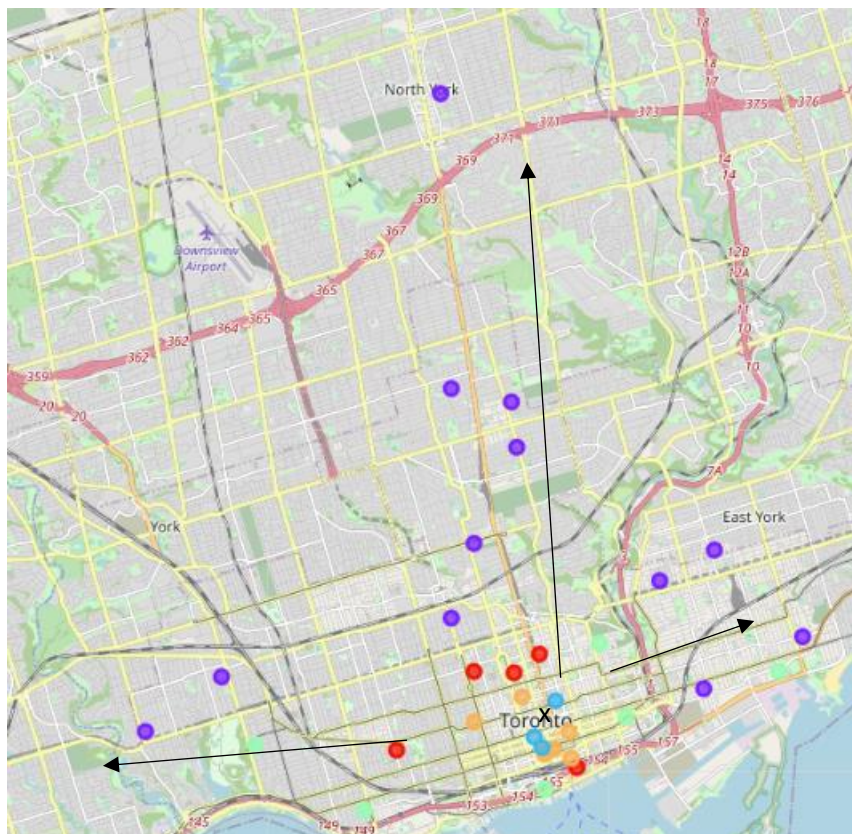
6.Conclusion

Based on the analysis done, if it is concluded that the restaurant category that has more potential is Thai restaurants, as Toronto citizens like these restaurants and there are several of them around the city, that means that different Thai restaurants succeed previously.

The best location possible is when closer to the centre of the city. Based on k-means, the areas of the blue cluster are the best, followed orange and red cluster areas.

Jorge Quintero Bermejo