



# **Study of business viability in Toronto**

**Determination of the optimal location**

# Which is the best location in Toronto to start a business?

- Before choosing a location, It is vital for entrepreneurs to know their own future businesses, understanding their needs, and identifying who your customers are and how you can best meet their needs.
- If your new business does not manufacture products, wholesale, or sell strictly over the Internet, then finding the right location to set up your business, from your customers' point of view, can be critical to your success.
- The present study will analyze commercial patterns in each particular borough in Toronto, focusing on to scope out the competition, accessibility, and cost.

# Data acquisition and cleaning

**The data used here will be a combination of:**

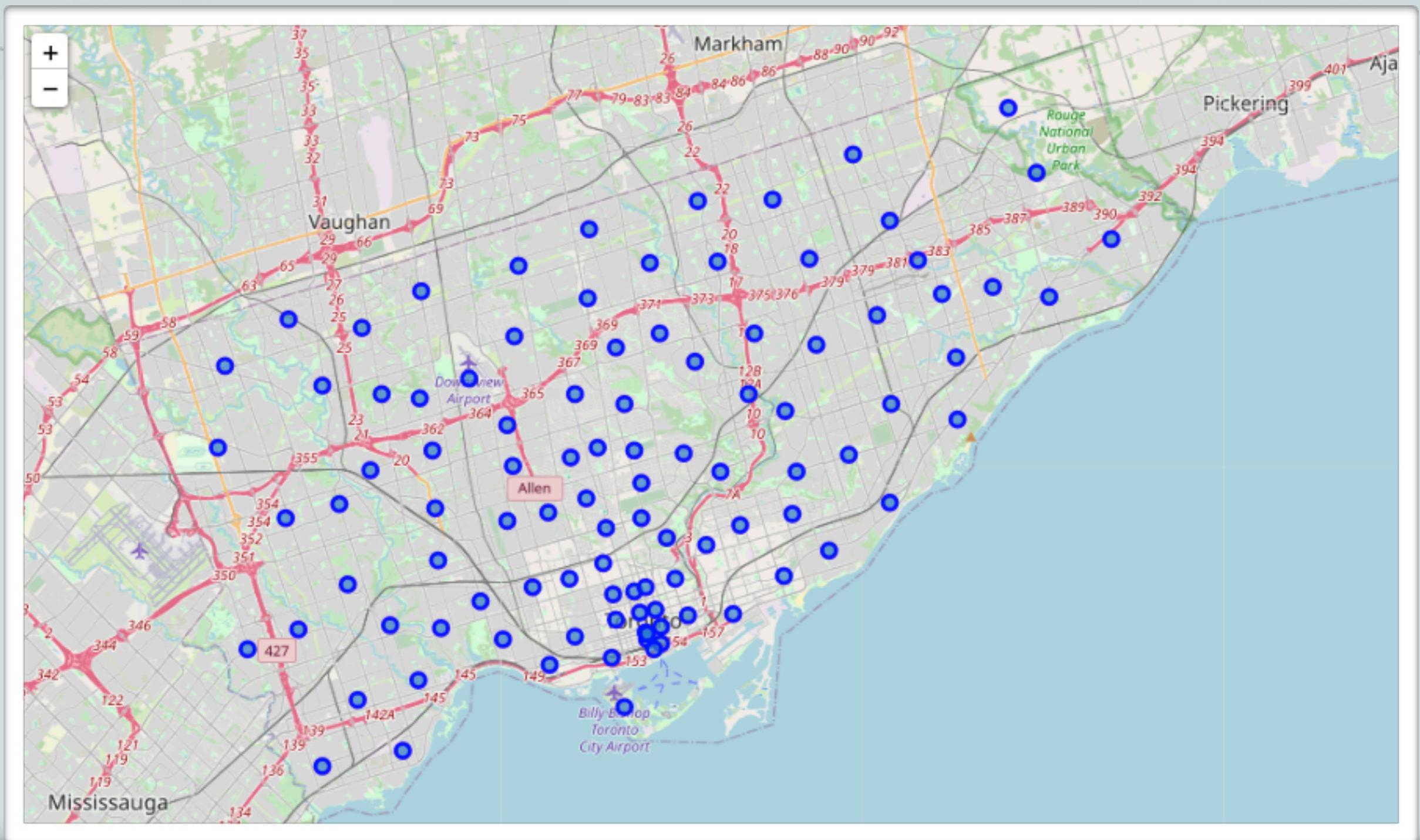
- Postal codes, borough, and neighborhood names obtained by web scrapping ([Link](#)).
- Geographical coordinates of each postal code obtained using the Pgeocode Python library.
- Venues data of each borough obtained using the Foursquare API.
- Duplicate, highly similar or highly correlated features were dropped.
- Cleaned data contains 2,165 rows and 262 features in the raw dataset.

# Methodology

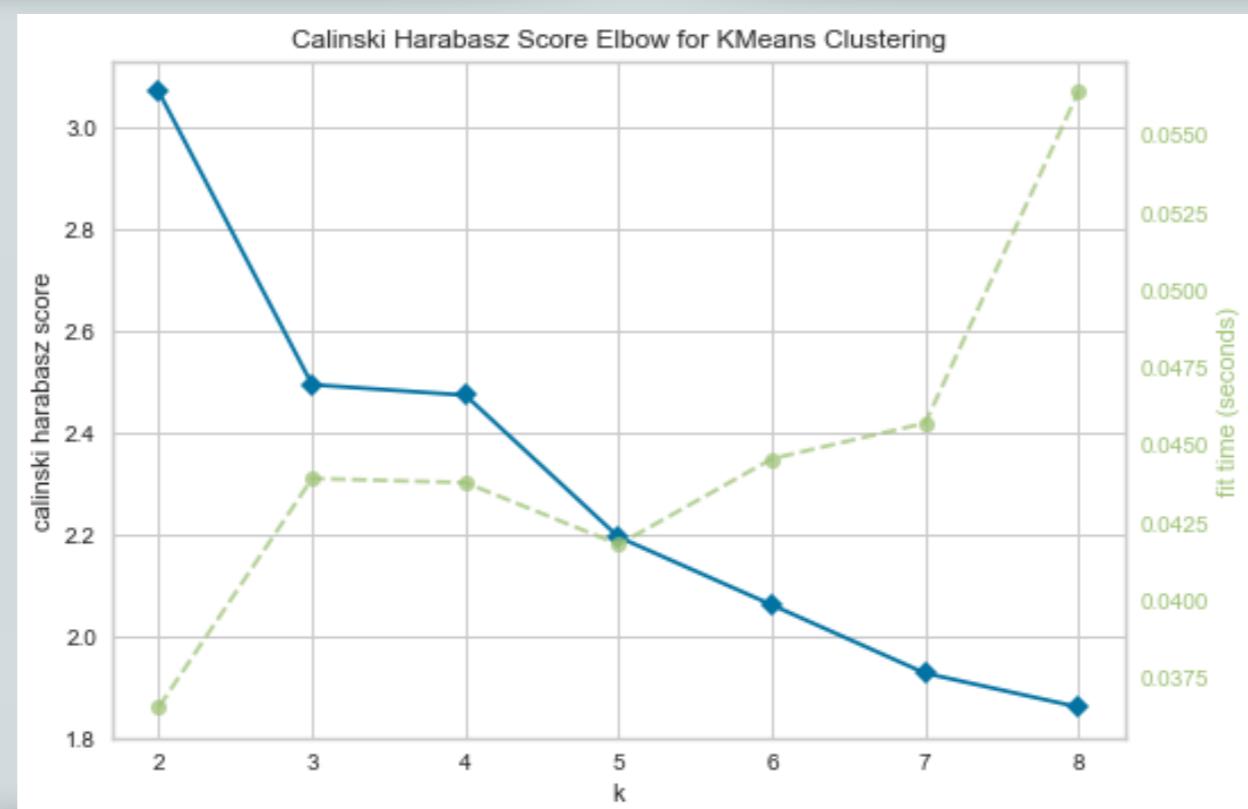
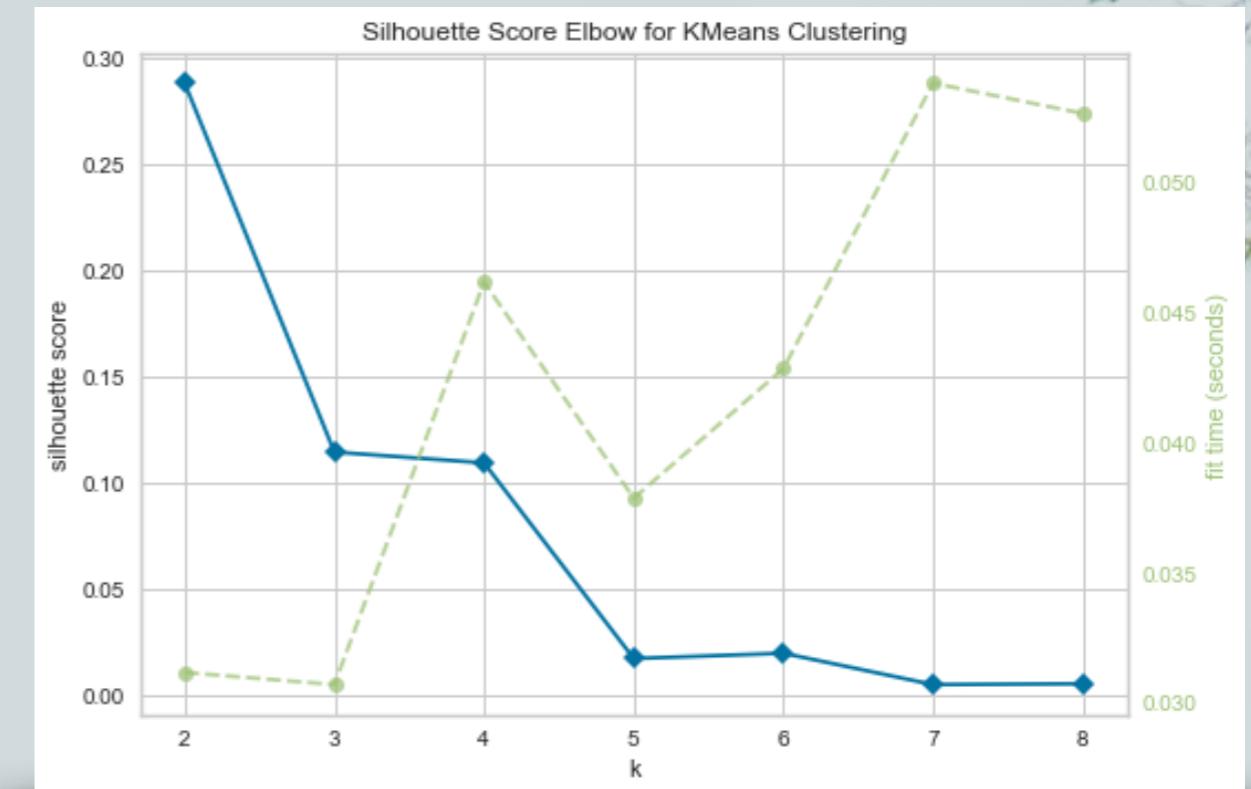
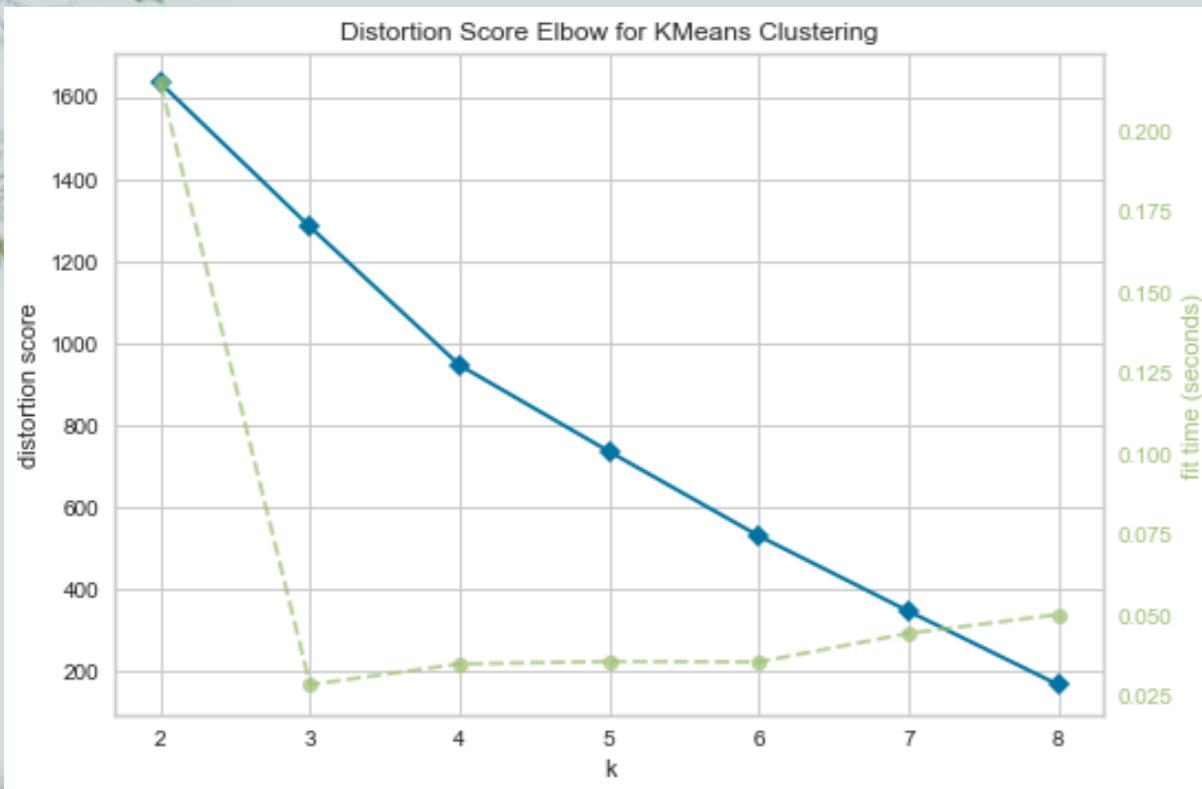
**The steps adopted in this project are summarized below:**

1. Web scrape to obtain the postal codes, borough, and neighborhood names.
2. Get the geographical coordinates of each postal code using the Pgeocode Python library.
3. Join all the previous information obtained so far in a pandas dataframe.
4. Use the Foursquare API to explore the 100 most common venues for each particular borough.
5. Analyze the most common venues by borough considering the mean of the frequency as our metric and, assing the results to a dataframe.
6. Apply k-means clustering to group together boroughs with similar venues' data and discover underlying patterns.
7. Interactively visualize spatial data that has been manipulated so far, using the Folium library.
8. Discuss the insights observed in each cluster.

# Visualizing boroughs in Toronto



# The optimal number of clusters

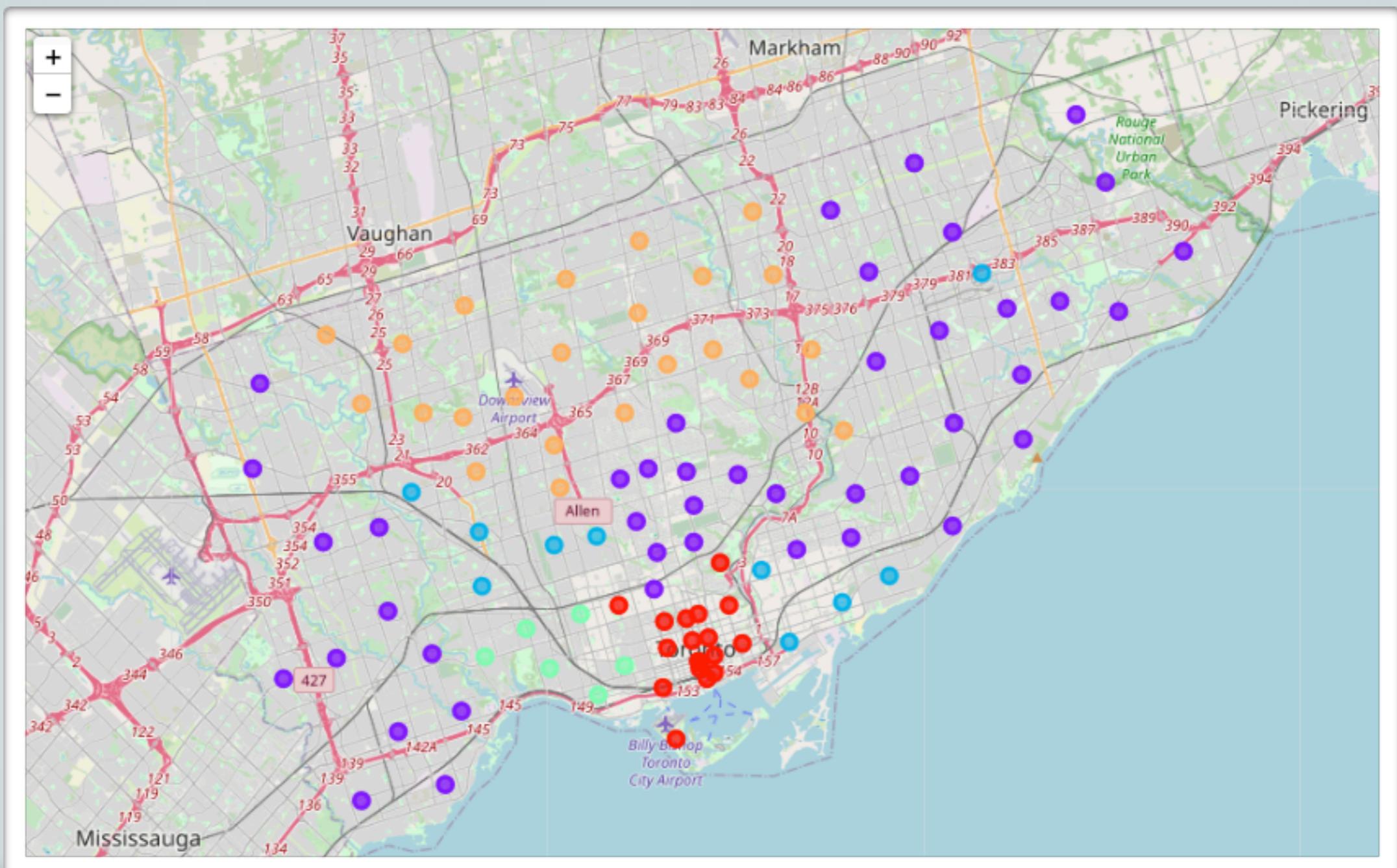


# The top five most popular venues per borough

	Postal Code	Borough	Neighborhood	latitude	longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	M3A	North York	Parkwoods	43.7545	-79.3300	4	Clothing Store	Coffee Shop	Pizza Place	Park	Sandwich Place
1	M4A	North York	Victoria Village	43.7276	-79.3148	4	Clothing Store	Coffee Shop	Pizza Place	Park	Sandwich Place
2	M5A	Downtown Toronto	Regent Park, Harbourfront	43.6555	-79.3626	0	Coffee Shop	Café	Restaurant	Hotel	Japanese Restaurant
3	M6A	North York	Lawrence Manor, Lawrence Heights	43.7223	-79.4504	4	Clothing Store	Coffee Shop	Pizza Place	Park	Sandwich Place
4	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.6641	-79.3889	0	Coffee Shop	Café	Restaurant	Hotel	Japanese Restaurant
5	M9A	Etobicoke	Islington Avenue, Humber Valley Village	43.6662	-79.5282	1	Pizza Place	Coffee Shop	Pharmacy	Sandwich Place	Bank
6	M1B	Scarborough	Malvern, Rouge	43.8113	-79.1930	1	Coffee Shop	Pizza Place	Pharmacy	Bank	Convenience Store
95	M4X	Downtown Toronto	St. James Town, Cabbagetown	43.6684	-79.3689	0	Coffee Shop	Café	Restaurant	Hotel	Japanese Restaurant
96	M5X	Downtown Toronto	First Canadian Place, Underground city	43.6492	-79.3823	0	Coffee Shop	Café	Restaurant	Hotel	Japanese Restaurant
97	M8X	Etobicoke	The Kingsway, Montgomery Road, Old Mill North	43.6518	-79.5076	1	Pizza Place	Coffee Shop	Pharmacy	Sandwich Place	Bank
98	M4Y	Downtown Toronto	Church and Wellesley	43.6656	-79.3830	0	Coffee Shop	Café	Restaurant	Hotel	Japanese Restaurant
99	M7Y	East Toronto	Business reply mail Processing Centre, South C...	43.7804	-79.2505	2	Greek Restaurant	Restaurant	Italian Restaurant	Coffee Shop	Pub
100	M8Y	Etobicoke	Old Mill South, King's Mill Park, Sunnylea, Hu...	43.6325	-79.4939	1	Pizza Place	Coffee Shop	Pharmacy	Sandwich Place	Bank
101	M8Z	Etobicoke	Mimico NW, The Queensway West, South of Bloor,...	43.6256	-79.5231	1	Pizza Place	Coffee Shop	Pharmacy	Sandwich Place	Bank

# Clustering boroughs in Toronto

We are interested in grouping together boroughs with similar venues' data, discovering underlying patterns.



# Clusters

**The insights observed, considering the five most common venues categories per cluster, are summarized below:**

- **Cluster 0:** The boroughs contained in this cluster have shown balanced venues categories, in which you can find restaurants, fast foods, coffee shops, pharmacies, banks, and home services pretty much almost everything that any person could need. Besides, entertainment (Hockey arena) can also be found.
- **Cluster 1:** It seems to be a business center, with a focus on coffee shops, cafés, having hotels and restaurants available for a possible need.
- **Cluster 2:** There is a focus on gastronomical variety here, you can find several types of restaurants. Besides, entertainment for adults (pubs) is also available.
- **Cluster 3:** It seems to be a wealthy residential location with a lot of places to hang out with friends (bar), to chill (coffee shops), and good gastronomical options (restaurants, cafés, and bakeries).
- **Cluster 4:** This cluster appears to be a typical middle-class residential location, with parks for recreation, clothing stores for a possible need, coffee shops to chill, and some gastronomical options (pizza place and restaurant).

# Conclusion

**The present study aimed at the analysis of commercial patterns in each particular borough in Toronto, focusing on to scope out the competition, to find the optimal location to open a new restaurant. Based on the patterns encountered, we can conclude the following:**

- Cluster 2 boroughs would represent the hardest competition due to the gastronomical variety.
- The higher investment would be needed at boroughs in Clusters 1 and 3, due to the high quality provided by competing restaurants.
- Even though Cluster 3 boroughs represent a balanced environment, it can be seen a reasonable variety of restaurants and fast food. High creativity and quality advertising would be necessary to attract customers.
- The most reasonable choice, based in our clustering analysis, would be boroughs at Cluster 4. They have not so many gastronomical options, summed with the middle class-like environment, would not require a very sophisticated restaurant and excessive investment in advertisements.