

Capstone Project - The Battle of Neighborhoods

1. Introduction

This project is a kind of summary of the learned material covered in IBM course of data science. In this work, we used the algorithm – K-means clustering - applied to the neighborhoods in Toronto.

The objective of this Capstone is to learn about segmenting and k-means clustering, to understand how to use the Github platform, to learn the Foursquare API and to try to show the results of studying by publishing the final project on Github. The main goal is to understand the practical side of the application of techniques and algorithms of data science in real life tasks.

2. Data

The data has been collected from Wikipedia page, https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M.

This html table was converted to Pandas DataFrame for cleaning and preprocessing. We have dropped the rows where Borough is 'Not assigned' and combined the neighbourhoods with same Postal code.

Then we have imported the csv file containing the latitudes and longitudes for various neighbourhoods in Canada and merged two tables.

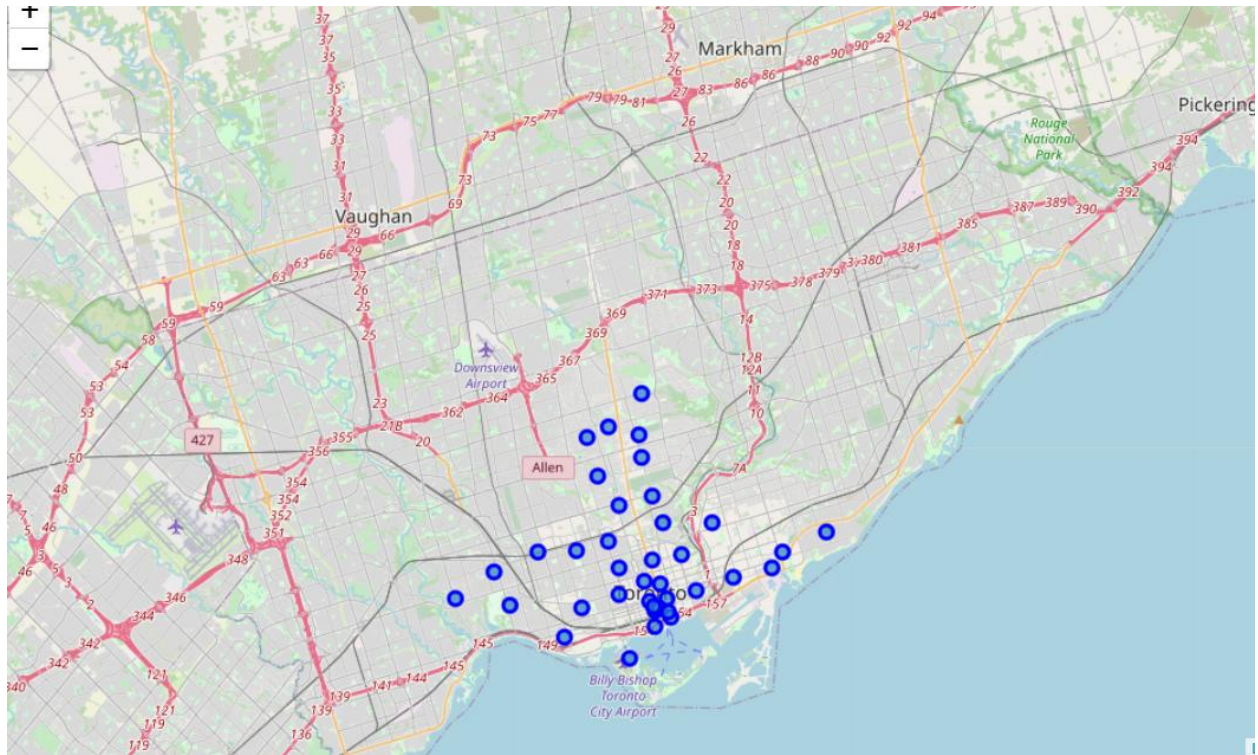
We will analyze all the rows from the data frame which contains Toronto in their Borough.

3. Segmenting and Clustering Neighbourhoods in Toronto

Wiki page contains 103 rows × 3 columns.

According to the merged table (Wiki + latitudes and longitudes for various neighbourhoods in Canada) we have found several neighbourhoods which contains Toronto.

Now we can visualize all the Neighbourhoods of the above data frame “Toronto” using Folium and the result we can see on the map:



The first neighborhood's name in our data frame is 'The Beaches'.

Name	Latitude	Longitude
The Beaches	43.67635739999999	-79.2930312

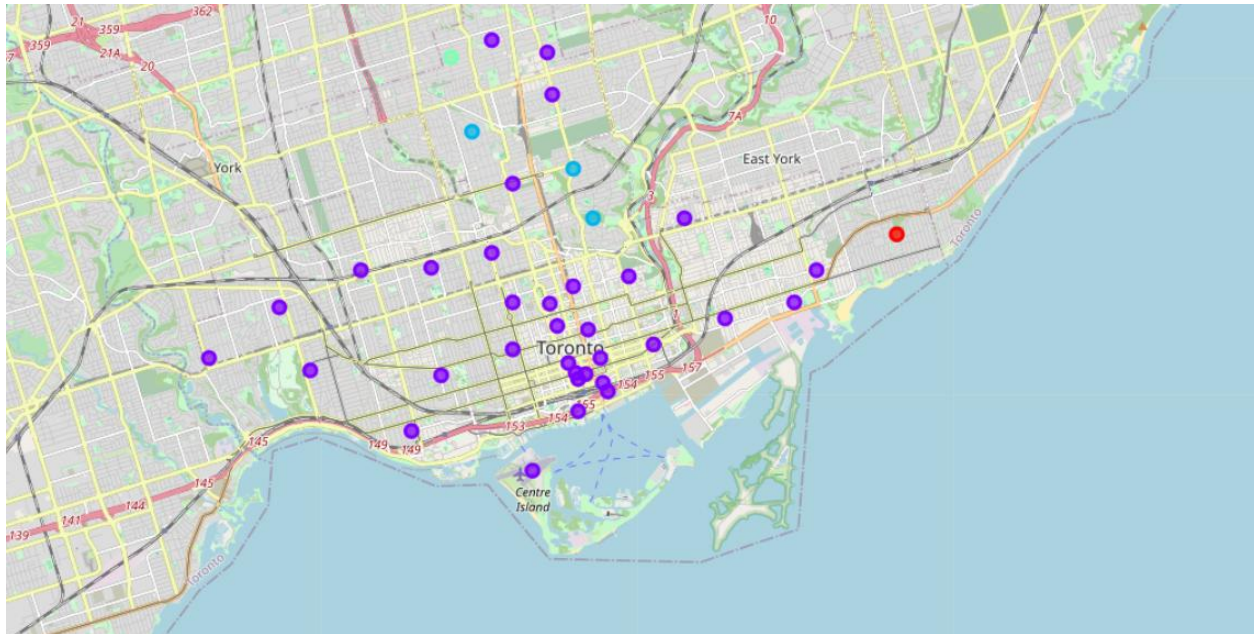
Using Foursquare API we could get the top 100 venues that are in The Beaches within a radius of 500 meters.

	name	categories	lat	lng
0	Glen Manor Ravine	Trail	43.676821	-79.293942
1	The Big Carrot Natural Food Market	Health Food Store	43.678879	-79.297734
2	Grover Pub and Grub	Pub	43.679181	-79.297215
3	Upper Beaches	Neighborhood	43.680563	-79.292869

3.1. Cluster neighborhoods

For clustering we will use K—Means clustering method. The set number of clusters = 5.

By the next step we will create a new dataframe that includes the cluster as well as the top 10 venues for each neighborhood. Visualization of the resulting clusters is on the map below:



By the last step we will examine each cluster.

Cluster 1

Borough	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
East Toronto	Pub	Trail	Health Food Store	Wine Bar	Dog Run	Dessert Shop	Diner	Discount Store	Distribution Center	Donut Shop

Cluster 2

Borough	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
East Toronto	Greek Restaurant	Coffee Shop	Italian Restaurant	Bookstore	Furniture / Home Store	Ice Cream Shop	Spa	Japanese Restaurant	Juice Bar	Brewery
East Toronto	Park	Fast Food Restaurant	Sushi Restaurant	Ice Cream Shop	Fish & Chips Shop	Italian Restaurant	Liquor Store	Movie Theater	Pet Store	Pizza Place
East Toronto	Coffee Shop	Brewery	Gastropub	Café	Bakery	American Restaurant	Yoga Studio	Convenience Store	Cheese Shop	Clothing Store
Central Toronto	Gym	Hotel	Park	Sandwich Place	Food & Drink Shop	Department Store	Dance Studio	Breakfast Spot	Gym / Fitness Center	Pizza Place
Central Toronto	Coffee Shop	Clothing Store	Pet Store	Seafood Restaurant	Salon / Barber Shop	Restaurant	Café	Chinese Restaurant	Yoga Studio	Sporting Goods Shop
Central Toronto	Pizza Place	Dessert Shop	Sandwich Place	Sushi Restaurant	Café	Coffee Shop	Gym	Italian Restaurant	Farmers Market	Diner
Central Toronto	Coffee Shop	Pizza Place	Liquor Store	Restaurant	Pub	Bank	Supermarket	Bagel Shop	Fried Chicken Joint	Sushi Restaurant
Downtown Toronto	Restaurant	Café	Coffee Shop	Pizza Place	Bakery	Chinese Restaurant	Italian Restaurant	Pub	Park	Snack Place
Downtown Toronto	Coffee Shop	Sushi Restaurant	Japanese Restaurant	Restaurant	Gay Bar	Fast Food Restaurant	Yoga Studio	Café	Pub	Mediterranean Restaurant
Downtown Toronto	Coffee Shop	Park	Bakery	Café	Pub	Breakfast Spot	Theater	Cosmetics Shop	Shoe Store	Brewery
Downtown Toronto	Coffee Shop	Clothing Store	Hotel	Japanese Restaurant	Cosmetics Shop	Middle Eastern Restaurant	Café	Bubble Tea Shop	Theater	Bookstore
Downtown Toronto	Coffee Shop	Café	American Restaurant	Gastropub	Cocktail Bar	Gym	Clothing Store	Moroccan Restaurant	Lingerie Store	Seafood Restaurant
Downtown	Coffee	Cocktail Bar	Bakery	Beer Bar	Farmers	Restaurant	Cheese Shop	Seafood	Hotel	Basketball

Toronto	Shop				Market			Restaurant		Stadium
Downtown Toronto	Coffee Shop	Italian Restaurant	Café	Sandwich Place	Salad Place	Thai Restaurant	Burger Joint	Bubble Tea Shop	Ramen Restaurant	Portuguese Restaurant
Downtown Toronto	Coffee Shop	Café	Restaurant	Clothing Store	Deli / Bodega	Hotel	Thai Restaurant	Gym	Pizza Place	Burrito Place
Downtown Toronto	Coffee Shop	Aquarium	Café	Hotel	Scenic Lookout	Restaurant	Brewery	Italian Restaurant	Fried Chicken Joint	History Museum
Downtown Toronto	Coffee Shop	Hotel	Café	Restaurant	Italian Restaurant	Japanese Restaurant	Seafood Restaurant	Salad Place	American Restaurant	Deli / Bodega
Downtown Toronto	Coffee Shop	Restaurant	Café	Hotel	Gym	American Restaurant	Italian Restaurant	Japanese Restaurant	Deli / Bodega	Seafood Restaurant
Central Toronto	Sandwich Place	Café	Coffee Shop	Donut Shop	Middle Eastern Restaurant	Pub	Liquor Store	BBQ Joint	Indian Restaurant	History Museum
Downtown Toronto	Café	Bookstore	Bar	Japanese Restaurant	Bakery	Sandwich Place	Beer Bar	Beer Store	Italian Restaurant	Comfort Food Restaurant
Downtown Toronto	Café	Coffee Shop	Vegetarian / Vegan Restaurant	Vietnamese Restaurant	Mexican Restaurant	Bar	Gaming Cafe	Grocery Store	Bakery	Caribbean Restaurant
Downtown Toronto	Airport Lounge	Airport Terminal	Boutique	Harbor / Marina	Boat or Ferry	Rental Car Location	Plane	Coffee Shop	Sculpture Garden	Airport Service
Downtown Toronto	Coffee Shop	Seafood Restaurant	Cocktail Bar	Restaurant	Japanese Restaurant	Italian Restaurant	Beer Bar	Hotel	Café	Gym
Downtown Toronto	Coffee Shop	Café	Hotel	Gym	Japanese Restaurant	Restaurant	Salad Place	Seafood Restaurant	Deli / Bodega	Steakhouse
Downtown Toronto	Grocery Store	Café	Park	Candy Store	Restaurant	Baby Store	Italian Restaurant	Coffee Shop	Nightclub	Diner
West Toronto	Pharmacy	Bakery	Bus Stop	Bar	Grocery Store	Music Venue	Bank	Park	Supermarket	Coffee Shop

West Toronto	Bar	Coffee Shop	Vegetarian/Vegan Restaurant	Asian Restaurant	Restaurant	Vietnamese Restaurant	Café	Men's Store	Wine Bar	Korean Restaurant
West Toronto	Café	Coffee Shop	Breakfast Spot	Gym	Stadium	Burrito Place	Restaurant	Climbing Gym	Pet Store	Bakery
West Toronto	Café	Mexican Restaurant	Thai Restaurant	Speakeasy	Diner	Bar	Flea Market	Bakery	Fried Chicken Joint	Italian Restaurant
West Toronto	Breakfast Spot	Gift Shop	Dessert Shop	Cuban Restaurant	Eastern European Restaurant	Dog Run	Italian Restaurant	Bar	Restaurant	Movie Theater
West Toronto	Café	Coffee Shop	Pub	Pizza Place	Sushi Restaurant	Italian Restaurant	Gym	Gourmet Shop	Restaurant	Latin American Restaurant
Downtown Toronto	Coffee Shop	College Cafeteria	Sushi Restaurant	Gym	Fast Food Restaurant	Restaurant	Portuguese Restaurant	Park	Nightclub	Mexican Restaurant
East Toronto	Light Rail Station	Yoga Studio	Garden Center	Farmers Market	Fast Food Restaurant	Burrito Place	Restaurant	Brewery	Auto Workshop	Garden

Cluster 3

Borough	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
Central Toronto	Park	Trail	Playground	Tennis Court	Wine Bar	Dog Run	Dessert Shop	Diner	Discount Store	Distribution Center
Downtown Toronto	Park	Trail	Playground	Dance Studio	Escape Room	Electronics Store	Eastern European Restaurant	Donut Shop	Doner Restaurant	Dog Run
Central Toronto	Park	Trail	Jewelry Store	Sushi Restaurant	Wine Bar	Dessert Shop	Ethiopian Restaurant	Escape Room	Electronics Store	Eastern European Restaurant

Cluster 4

Borough	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
Central Toronto	Home Service	Garden	Wine Bar	Department Store	Event Space	Ethiopian Restaurant	Escape Room	Electronics Store	Eastern European Restaurant	Donut Shop

Cluster 5

Borough	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
Central Toronto	Park	Bus Line	Swim School	Falafel Restaurant	Ethiopian Restaurant	Escape Room	Electronics Store	Eastern European Restaurant	Donut Shop	Doner Restaurant

4. Conclusion

During these 5 weeks we have understand the basics of K-Means Clustering, using of Foursquare API and repository of Github.

We have got 5 different clusters of Toronto. The most common venues are:

Cluster	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue
1	Pub	Trail	Health Food Store
2	Coffee Shop	Restaurant	Café
3	Park	Trail	Playground
4	Home Service	Garden	Wine Bar
5	Park	Bus Line	Swim School