Introduction

In this report we are interested in learning what are the most common venues in each neighborhood of Toronto. With this knowledge, we can then make suggestions to people to decide in which neighborhood would be best to open a new restaurant or set up the office for their own business.

Data

To solve this problem, we will be using the following data sources:

- List of postal codes for Toronto, Ontario, Canada:
  https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
  Each row contains Postal code and their corresponding borough and neighborhood. Below is an example of the dataset:

	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

- Geocodes for each neighborhood in Toronto: http://cocl.us/Geospatial_data
  Each row contains postal code and their corresponding latitude and longitude.
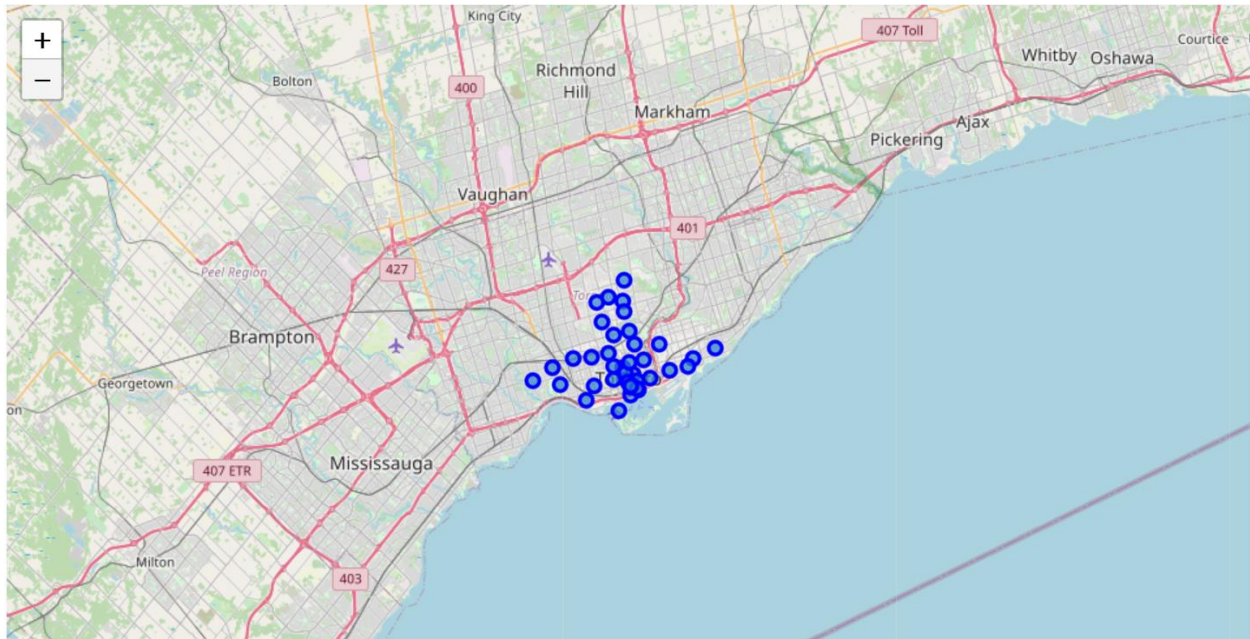- Location data from Foursqure API

Methodology

First, we use the first postal code dataset to build a data frame 'df_borough' of the postal code of each neighborhood along with the borough name and neighborhood name.

To use the Foursquare location, we combine the data frame 'df_borough' above with a csv file with the latitude and longitude coordinates of each neighborhood. To work with boroughs with 'Toronto' in them only, we further clean up the df_borough by dropping rows which do not contain 'Toronto' in 'borough' column. The result looks like below:

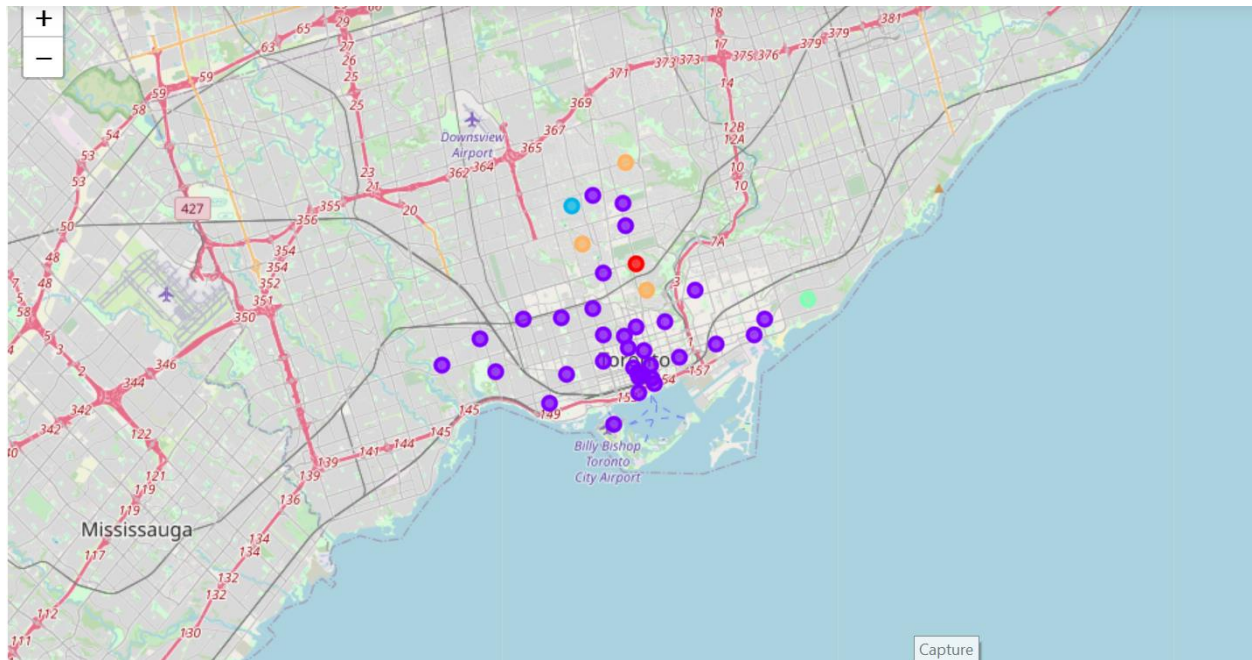|	Postal Code	Borough	Neighbourhood	Latitude	Longitude
0	M5A	Downtown Toronto	Regent Park, Harbourfront	43.654260	-79.360636
1	M7A	Downtown Toronto	Queen's Park, Ontario Provincial Government	43.662301	-79.389494
2	M5B	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937
3	M5C	Downtown Toronto	St. James Town	43.651494	-79.375418
4	M4E	East Toronto	The Beaches	43.676357	-79.293031

We proceed to retrieve the latitude and longitude of Toronto and use folium to visualize it on the map 'map_Toronto'.

After initializing the Foursquare credentials, we start by exploring the first neighborhood 'Regent park, Harbourfront' by inspecting the latitude, longitude and the top 100 venues that are within 500 meters radius of Regent park. We use helper function 'get_category_type' to the category for each neighboring venue. We expand the exploration by getting all the venues and their categories in data frame 'df_toronto' and get the result 'toronto_venues'. We count the frequencies of each category for each neighborhood. Then we created the new data frame and display the top 10 venues for each neighborhood.

	Neighborhood	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
0	Berczy Park	Coffee Shop	Café	Farmers Market	Seafood Restaurant	Bakery	Cocktail Bar	Beer Bar	Cheese Shop	Restaurant	Italian Restaurant
1	Brockton, Parkdale Village, Exhibition Place	Café	Bakery	Breakfast Spot	Performing Arts Venue	Coffee Shop	Gym	Gym / Fitness Center	Pet Store	Nightclub	Italian Restaurant
2	Business reply mail Processing Centre, South C...	Light Rail Station	Brewery	Recording Studio	Restaurant	Fast Food Restaurant	Farmers Market	Auto Workshop	Butcher	Spa	Burrito Place
3	CN Tower, King and Spadina, Railway Lands, Har...	Airport Service	Airport Lounge	Coffee Shop	Harbor / Marina	Boutique	Boat or Ferry	Rental Car Location	Bar	Plane	Sculpture Garden
4	Central Bay Street	Coffee Shop	Sandwich Place	Café	Italian Restaurant	Japanese Restaurant	Department Store	Burger Joint	Bubble Tea Shop	Salad Place	Portuguese Restaurant

Lastly we use k-means to cluster the neighborhoods into 5 clusters and visualize the result.

Results

We obtained the 10 most common venue categories for each Toronto cluster based on the analysis above:

- Cluster 1

	Borough	Cluster Labels	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
4	East Toronto	0	Trail	Health Food Store	Pub	Women's Store	Deli / Bodega	Electronics Store	Eastern European Restaurant	Dumpling Restaurant	Donut Shop	Doner Restaurant

- Cluster 2

	Borough	Cluster Labels	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
0	Downtown Toronto	1	Coffee Shop	Pub	Park	Bakery	Breakfast Spot	Theater	Café	Spa	Ice Cream Shop	Beer Store
1	Downtown Toronto	1	Coffee Shop	Pub	Park	Bakery	Breakfast Spot	Theater	Café	Spa	Ice Cream Shop	Beer Store
2	Downtown Toronto	1	Clothing Store	Coffee Shop	Café	Japanese Restaurant	Italian Restaurant	Cosmetics Shop	Bubble Tea Shop	Electronics Store	Lingerie Store	Middle Eastern Restaurant
3	Downtown Toronto	1	Café	Coffee Shop	Clothing Store	Restaurant	Cocktail Bar	Cosmetics Shop	American Restaurant	Diner	Seafood Restaurant	Beer Bar
5	Downtown Toronto	1	Coffee Shop	Bakery	Café	Farmers Market	Cocktail Bar	Restaurant	Beer Bar	Seafood Restaurant	Cheese Shop	Concert Hall
6	Downtown Toronto	1	Coffee Shop	Sandwich Place	Café	Japanese Restaurant	Italian Restaurant	Salad Place	Bubble Tea Shop	Burger Joint	Department Store	New American Restaurant

- Cluster 3

	Borough	Cluster Labels	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
29	Central Toronto	2	Restaurant	Women's Store	Ethiopian Restaurant	Electronics Store	Eastern European Restaurant	Dumpling Restaurant	Donut Shop	Doner Restaurant	Dog Run	Distribution Center

- Cluster 4

	Borough	Cluster Labels	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
19	Central Toronto	3	Home Service	Garden	Women's Store	Deli / Bodega	Ethiopian Restaurant	Electronics Store	Eastern European Restaurant	Dumpling Restaurant	Donut Shop	Doner Restaurant

- Cluster 5

	Borough	Cluster Labels	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
18	Central Toronto	4	Park	Bus Line	Swim School	Event Space	Electronics Store	Eastern European Restaurant	Dumpling Restaurant	Donut Shop	Doner Restaurant	Dog Run
33	Downtown Toronto	4	Park	Trail	Playground	Dance Studio	Electronics Store	Eastern European Restaurant	Dumpling Restaurant	Donut Shop	Doner Restaurant	Dog Run

Discussion

Based on the results above, if someone is looking to open a restaurant in Toronto, we should recommend them to open it in Cluster 3 Central Toronto where most of the venues are restaurants. Similarly, if a contractor is trying to start their own business, we should recommend them to start in Cluster 4 where majority of venues are home service.

Conclusion

In this report, we used Toronto neighborhood postal code data scraped from Wikipedia and Foursquare API data, and applied k-means to cluster the neighborhoods to find the most common venues in each neighborhood cluster. With this information, we were able to advise the best location to open a new restaurant or a new home office. The best location to open a new restaurant is in Cluster 3 where most of the venues are restaurants. Similarly, the best location to start their own business is in Cluster 4 where most of them are home services.