

Coursera Final Capstone Project

Introduction/Business Problem

We will compare two neighbourhoods of Toronto and New York by selecting 100 venues in the neighbourhoods of these cities with a radius of 500 m. We will compare and show that how these neighbourhoods are similar or dissimilar on various parameters and reflect by using data science techniques that how clustering of two neighbourhoods can be done based on their latitude and longitude.

Required Data

dataset of venues with latitude and longitude by using Foursquare API. Webscraping to obtain data of toronto and new york.

Methodology

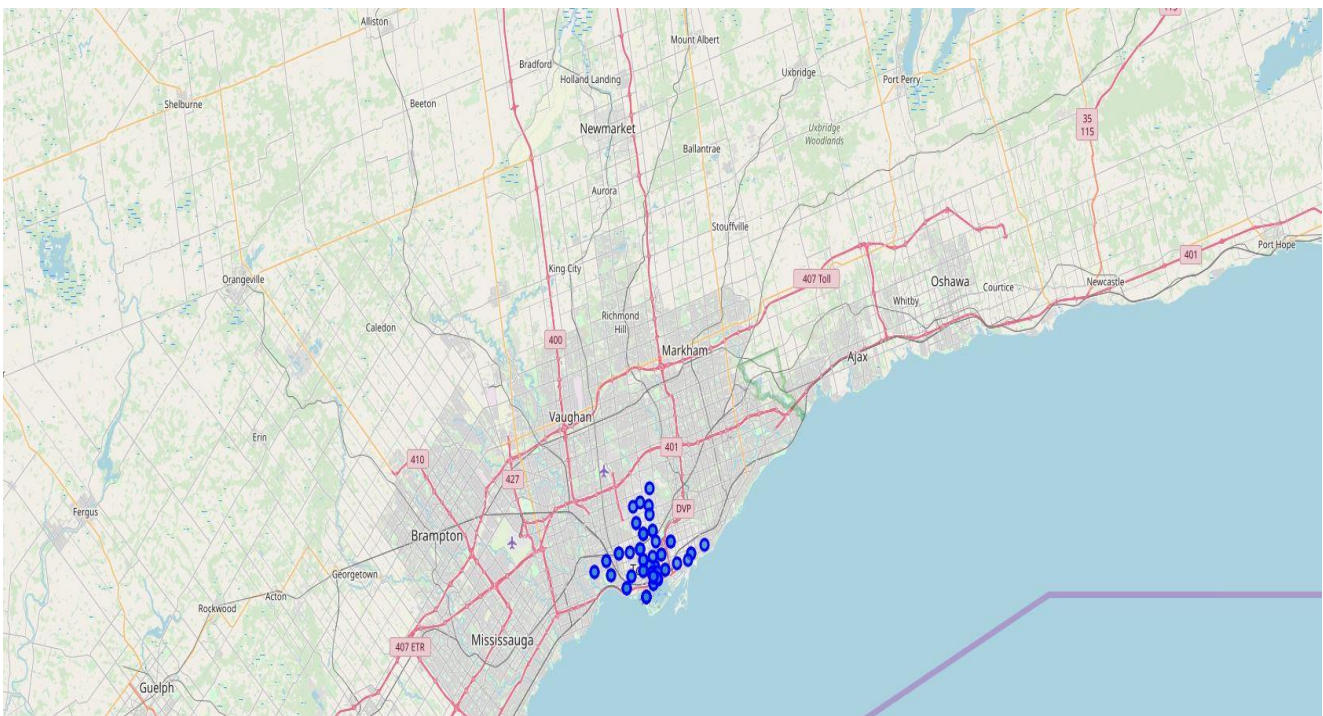
We will collect data of 100 venues by using Foursquare API. then our motive is to collect the best 10 common venues in neighbourhoods.

Firstly we have created the dataset for cities having Bouroughs and Neighbourhoods. After this we have added this dataset with latitudes and longitudes received from geo dataset kit.

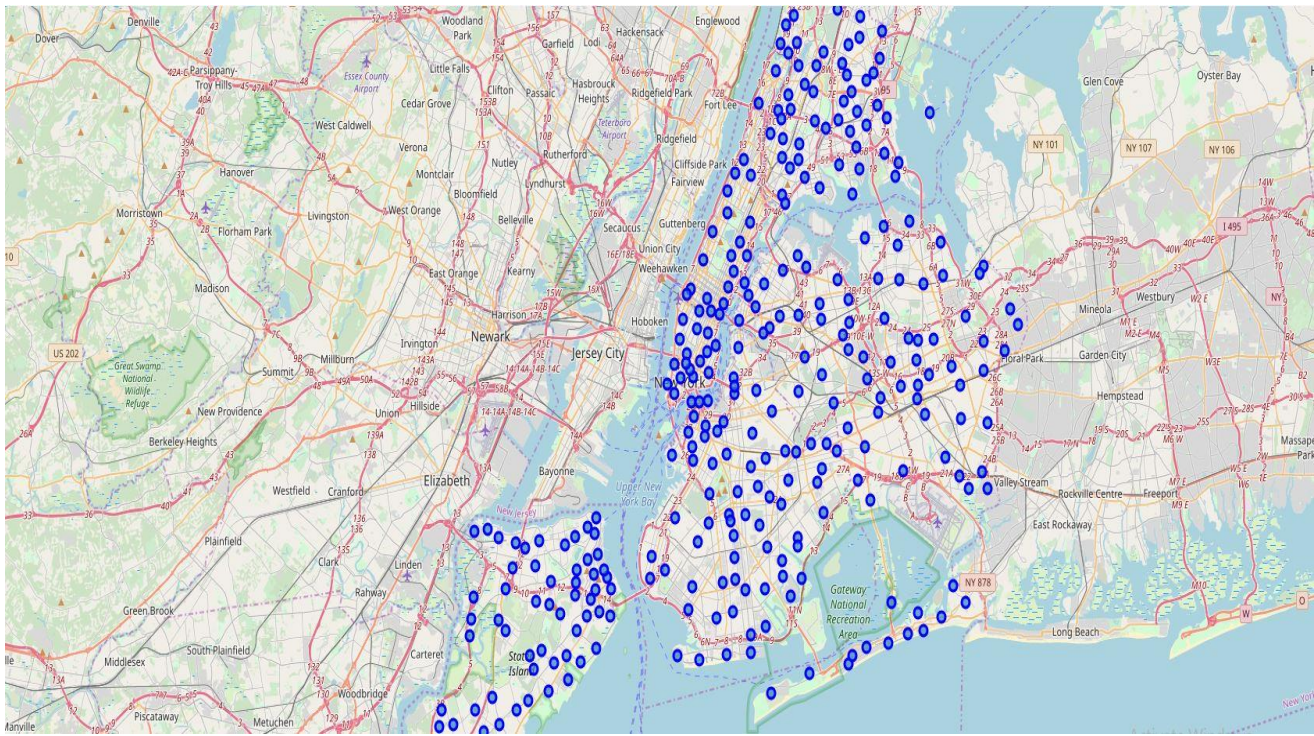
Secondly we have received Foursquare data of the two neighbourhoods which are the top ten places in the two neighbourhoods. Finally we have merged the two data of both cities by using K-means clustering algorithm.

Analysis

Toronto Map with 100 venues



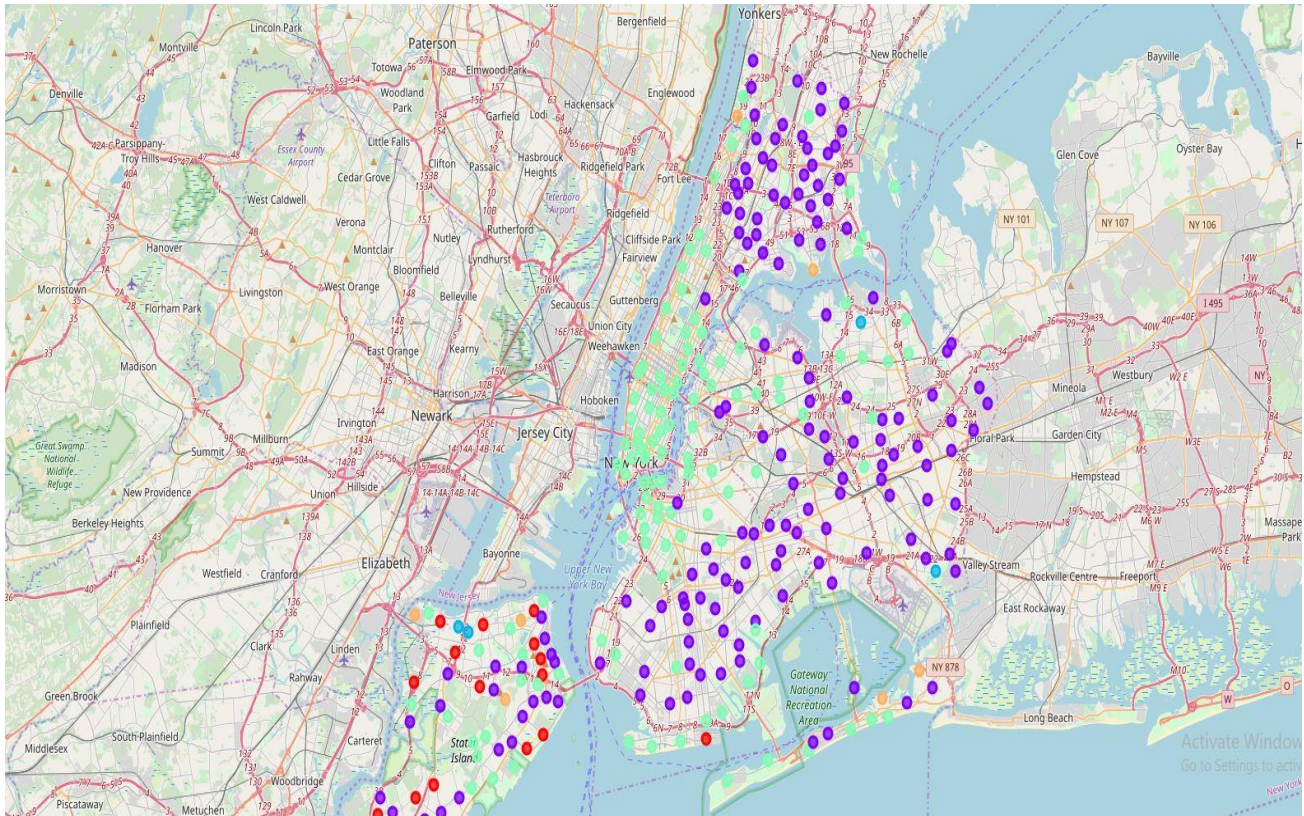
New York Map with 100 venues



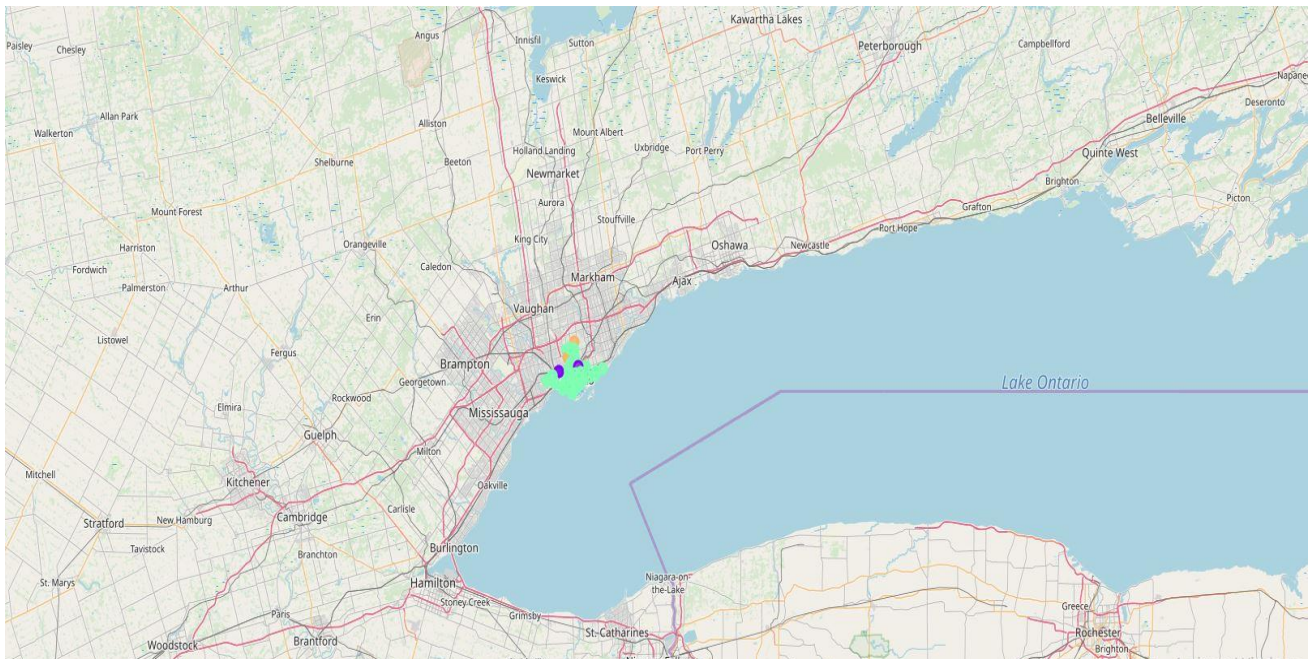
Top 5 rows data of Merged data of two cities.

	Borough	Latitude	Longitude	Neighborhood	Postcode
2	Downtown Toronto	43.654260	-79.360636	Harbourfront	M5A
3	Downtown Toronto	43.654260	-79.360636	Regent Park	M5A
13	Downtown Toronto	43.657162	-79.378937	Ryerson	M5B
14	Downtown Toronto	43.657162	-79.378937	Garden District	M5B
27	Downtown Toronto	43.651494	-79.375418	St. James Town	M5C

Common clusters on New York map



Common clusters on Toronto map



Finally top10 venues list

	Latitude	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
112	43.728020	4	Park	Swim School	Bus Line	Wings Joint	Diner	Falafel Restaurant	Event Space	Ethiopian Restaurant	Electronics Store	Eastern European Restaurant
122	43.696948	4	Park	Trail	Jewelry Store	Sushi Restaurant	Wings Joint	Diner	Falafel Restaurant	Event Space	Ethiopian Restaurant	Electronics Store
123	43.696948	4	Park	Trail	Jewelry Store	Sushi Restaurant	Wings Joint	Diner	Falafel Restaurant	Event Space	Ethiopian Restaurant	Electronics Store
183	43.679563	4	Park	Playground	Trail	Wings Joint	Dim Sum Restaurant	Falafel Restaurant	Event Space	Ethiopian Restaurant	Electronics Store	Eastern European Restaurant
27	40.806551	4	Park	South American Restaurant	Home Service	Grocery Store	Pool	Boat or Ferry	Bus Stop	Event Service	Event Space	Exhibit
35	40.881395	4	Park	Bank	Thai Restaurant	Tennis Stadium	Tennis Court	Pharmacy	Pizza Place	Falafel Restaurant	Empanada Restaurant	English Restaurant
169	40.659816	4	Park	Playground	Trail	Wings Joint	Dim Sum Restaurant	Falafel Restaurant	Event Space	Ethiopian Restaurant	Electronics Store	Eastern European Restaurant
192	40.597711	4	Park	Women's Store	Field	Ethiopian Restaurant	Event Service	Event Space	Exhibit	Eye Doctor	Factory	Falafel Restaurant
203	40.597069	4	Park	Women's Store	Field	Ethiopian Restaurant	Event Service	Event Space	Exhibit	Eye Doctor	Factory	Falafel Restaurant
256	40.635630	4	Park	Bagel Shop	Deli / Bodega	Bus Stop	Filipino Restaurant	Event Service	Event Space	Exhibit	Eye Doctor	Factory
303	40.611322	4	Playground	Park	Women's Store	Ethiopian Restaurant	Event Service	Event Space	Exhibit	Eye Doctor	Factory	Falafel Restaurant

Discussion of Results

the above analysis reflects that both toronto and new york are same in many characteristics.

Similarities are that both of them have such neighbourhoods which are closer to parks, children's playing field and are ethnically diverse. restaurants in both cities serve wide variety of cuisines catering to diversity in both cities. there are many tourist spots in both cities which lie on waterfronts.

Dissimilarities are that new york is a much larger city than toronto and has more bus halts.

Conclusion

with the help of Foursquare data and K means algorithm we clustered the two neighbourhoods of toronto and new york on various similarities. this approach can be used to compare neighbourhoods on various parameters to judge the living or business conditions prevalent.