How similar is New York City and Toronto?

Applied Data Science Capstone Project

Introduction

- New York City is the largest city in the United States with a population of 8,836,817
- Toronto is the largest city in Canada with a population of 6,471,850
- Are they have similar venues in their neighborhoods?
- If we want to visit or invest some places in north America, which neighborhoods are the good choices than can meet our criteria?

Data acquisition and cleaning

- Borough, latitude and longitude of each neighborhood for New York City were downloaded from new_york_dataset and then transformed into a pandas dataframe.
- The first five rows of the dataframe is displayed as follows.

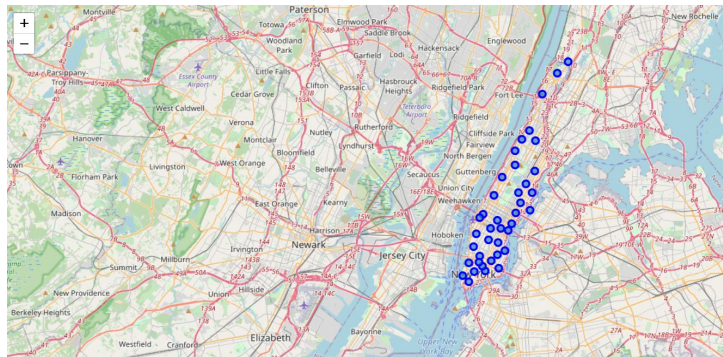|	Borough	Neighborhood	Latitude	Longitude
0	Manhattan	Marble Hill	40.876551	-73.910660
1	Manhattan	Chinatown	40.715618	-73.994279
2	Manhattan	Washington Heights	40.851903	-73.936900
3	Manhattan	Inwood	40.867684	-73.921210
4	Manhattan	Hamilton Heights	40.823604	-73.949688

Data acquisition and cleaning

- Boroughs that contain Toronto was scraped from a [Wikipedia page](#) and transformed it into a pandas dataframe using the BeautifulSoup package.
- The first five rows of the dataframe is displayed as follows.

	Borough	Neighborhood	Latitude	Longitude
0	Downtown Toronto	Regent Park , Harbourfront	43.654260	-79.360636
1	Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937
2	Downtown Toronto	St. James Town	43.651494	-79.375418
3	East Toronto	The Beaches	43.676357	-79.293031
4	Downtown Toronto	Berczy Park	43.644771	-79.373306

Visualize geographic details and boroughs of Manhattan

The map of Manhattan with neighborhoods superimposed on top

Visualize geographic details and boroughs of Toronto

The map of Toronto with neighborhoods superimposed on top

Explore the venues of each neighborhood of Manhantan using Foursquare API

Dataframe that includes the cluster as well as top 10 venues for neighborhoods in Manhattan

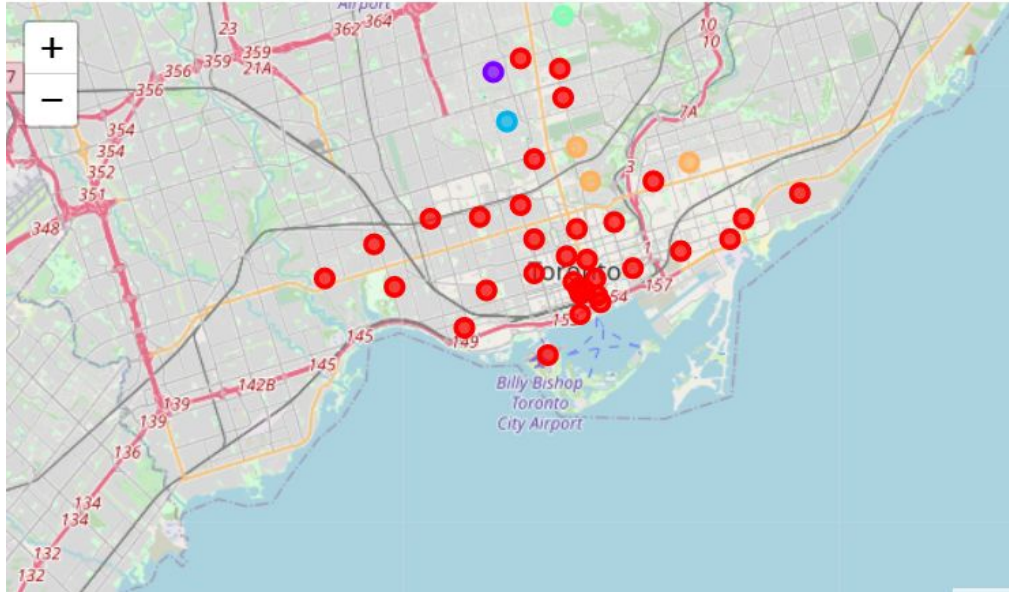|	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Manhattan	Marble Hill	40.876551	-73.910660	4	Gym	Sandwich Place	Yoga Studio	Pizza Place	Steakhouse	Shopping Mall	Seafood Restaurant	Clothing Store	Coffee Shop	Deli / Bodega
1	Manhattan	Chinatown	40.715618	-73.994279	1	Bakery	Chinese Restaurant	Cocktail Bar	American Restaurant	Hotpot Restaurant	Ice Cream Shop	Dessert Shop	Spa	Salon / Barbershop	Asian Restaurant
2	Manhattan	Washington Heights	40.851903	-73.936900	4	Café	Bakery	Deli / Bodega	Bank	Mobile Phone Shop	Grocery Store	Spanish Restaurant	Pizza Place	Tapas Restaurant	Sandwich Place
3	Manhattan	Inwood	40.867684	-73.921210	4	Mexican Restaurant	Café	Restaurant	Chinese Restaurant	Lounge	Spanish Restaurant	Bakery	Caribbean Restaurant	Pizza Place	Park
4	Manhattan	Hamilton Heights	40.823604	-73.949688	4	Pizza Place	Coffee Shop	Café	Deli / Bodega	Mexican Restaurant	Bakery	Park	Cocktail Bar	Sandwich Place	Chinese Restaurant

Explore the venues of each neighborhood of Toronto using Foursquare API

Dataframe that includes the cluster as well as top 10 venues for neighborhoods in Toronto

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	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
Downtown Toronto	Regent Park , Harbourfront	43.654260	-79.360636	0	Coffee Shop	Bakery	Café	Pub	Park	Breakfast Spot	Theater	Yoga Studio	Dessert Shop	Shoe Store
Downtown Toronto	Garden District, Ryerson	43.657162	-79.378937	0	Coffee Shop	Clothing Store	Café	Bubble Tea Shop	Middle Eastern Restaurant	Japanese Restaurant	Cosmetics Shop	Hotel	Pizza Place	Burger Joint
Downtown Toronto	St. James Town	43.651494	-79.375418	0	Coffee Shop	Café	Beer Bar	Gastropub	Cosmetics Shop	Cocktail Bar	Gym	Farmers Market	Hotel	Italian Restaurant
East Toronto	The Beaches	43.676357	-79.293031	0	Asian Restaurant	Health Food Store	Trail	Pub	Wings Joint	Donut Shop	Distribution Center	Dog Run	Doner Restaurant	Eastern European Restaurant
Downtown Toronto	Berczy Park	43.644771	-79.373306	0	Coffee Shop	Cocktail Bar	Farmers Market	Seafood Restaurant	Bakery	Restaurant	Pharmacy	Cheese Shop	Beer Bar	Japanese Restaurant

Run k-means to cluster the neighborhoods of Manhattan into 5 clusters

Run k-means to cluster the neighborhoods of Toronto into 5 clusters

Select all the neighborhoods in Manhattan with Chinese restaurants, hotels and parks

	Neighborhood	Chinese Restaurant	Hotel	Park
0	Battery Park City	0.012658	0.050633	0.101266
3	Chelsea	0.010000	0.020000	0.020000
6	Clinton	0.010000	0.030000	0.020000
10	Flatiron	0.010000	0.010000	0.020000
12	Greenwich Village	0.020000	0.010000	0.010000
17	Lincoln Square	0.010526	0.021053	0.010526
23	Midtown	0.010000	0.100000	0.010000
31	Sutton Place	0.020000	0.020000	0.030000

Select all the neighborhoods in Toronto with Chinese restaurants, hotels and parks

	Neighborhood	Chinese Restaurant	Hotel	Park
14	Garden District, Ryerson	0.01	0.02	0.01
15	Harbourfront East , Union Station , Toronto Is...	0.01	0.04	0.02

Analysis Results

-

	Number of Neighborhood	Unique venue categories	Number of neighborhood meets the criteria
Manhatan	40	329	8
Toronto	39	234	2

Discussion and future work

- Based on the neighborhood and venue analysis we would say that although New York City and Toronto have a similar population, New York City is more diverse, convenient and has more choices for tourism and investments.
- However, we did not compare other facts such as price, security, environment and traffic.
- In the future work, we can take more facts into consideration and create an user interface for the users to view the comparison by selecting the facts they are more interested in.