



COURSERA CAPSTONE PROJECT-THE BATTLE FOR NEIGHBORHOODS

Baichuan Chu

Introduction and Problems

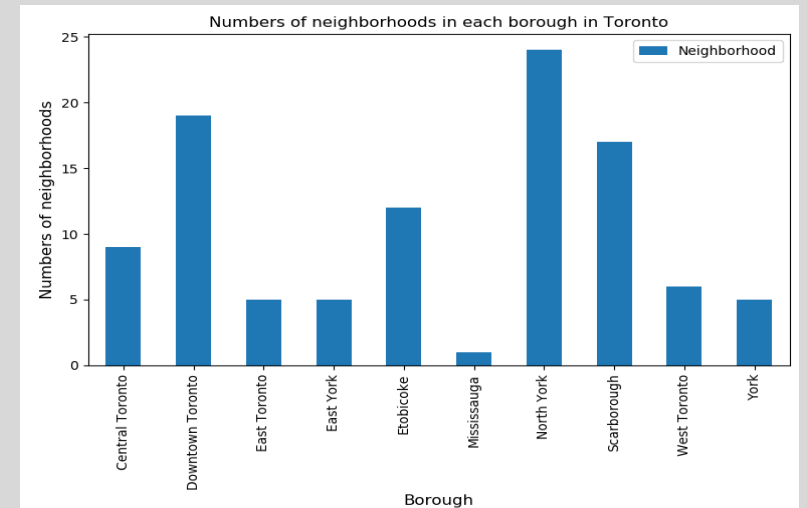
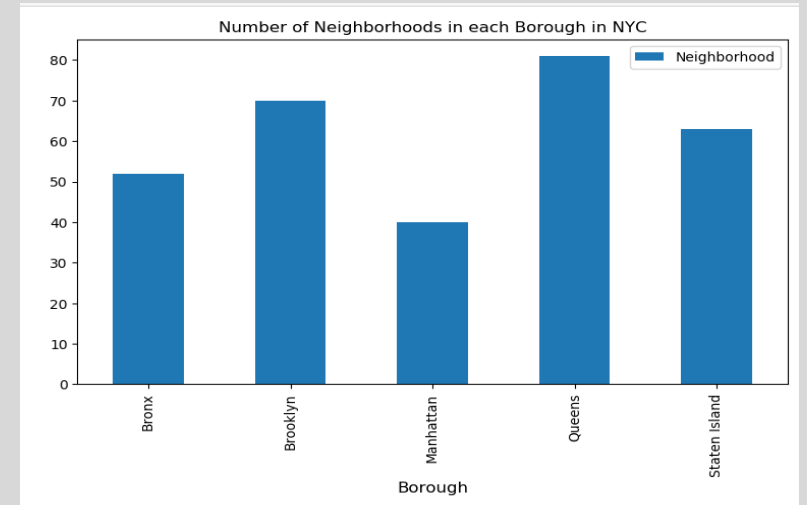
- Today, people are traveling around the world. For example, a lot of students travel to America in order to study here. Meanwhile, there are people who like to have businesses in other countries.
- Under the trend, United States and Canada are the popular choices to those people. Among all those popular cities in American and Canada, I will mainly focus on **Toronto** and **New York City**
- I am trying to solve some of the problems:
 1. Within Toronto and New York City, which boroughs of those two cities contain the most neighborhood?
 2. What are the differences between the two boroughs? (Mainly discuss the differences in their venues.)
 3. Which one of the cities is good for people to travel and study?
 4. Which one of the cities is good for people to open up a Chinese restaurant?

Data

- I used several datasets in order to do this project:
 1. Geospace data that allows me to get all coordinates of the boroughs.
 2. The Toronto Neighborhoods data, a Wikipedia data, which allows me to explore all neighborhoods in the Toronto. https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M
 3. New York data which give me all information about New York City's boroughs.
 4. FourSqaure is also an important data set to me. It's a tool, but this tools can let me obtain all the information I need for my analysis. It could tell me the venues around each neighborhoods.

Boroughs: North York, Toronto V.S Queen New York City

- Because I mainly focus on Toronto and New York City, I need to pull out the data of neighborhoods in those two cities and create data frames for them.
- After having all data in the table, I could use matplotlib function to make bar chart in order to pick the boroughs which have the most neighborhoods.
- From the graph I created, we can see that Queen has the most neighborhoods in New York City.
- For Toronto, North York is also the most populated one.
- Thus, I decide to analyze those two areas.



Method Part I:

- **Toronto:**

- In previous slides, I choose North York to continue my analysis.
- I clean and sort the data from Toronto and pull out all data under neighborhood name "North York".
- After I have the North York data, I use FourSquare to find out the venues around it. Within North York, there are **236** venues in **24** neighborhoods.
- Then, I used Kmean to make them into clusters for further analysis.(Cluster picture on the right)

Out[32]:

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Hillcrest Village	43.803762	-79.363452	Eagle's Nest Golf Club	43.805455	-79.364186	Golf Course
1	Hillcrest Village	43.803762	-79.363452	AY Jackson Pool	43.804515	-79.366138	Pool
2	Hillcrest Village	43.803762	-79.363452	Villa Madina	43.801685	-79.363938	Mediterranean Restaurant
3	Hillcrest Village	43.803762	-79.363452	Duncan Creek Park	43.805539	-79.360695	Dog Run
4	Fairview, Henry Farm, Oriole	43.778517	-79.346556	The LEGO Store	43.778207	-79.343483	Toy / Game Store
5	Fairview, Henry Farm, Oriole	43.778517	-79.346556	SilverCity	43.778681	-79.344085	Movie Theater
6	Fairview, Henry Farm, Oriole	43.778517	-79.346556	CF Fairview Mall	43.777994	-79.343665	Shopping Mall
7	Fairview, Henry Farm, Oriole	43.778517	-79.346556	New York Fries - Fairview Mall	43.778605	-79.343577	Restaurant
8	Fairview, Henry Farm, Oriole	43.778517	-79.346556	Apple Fairview	43.777883	-79.343789	Electronics Store
9	Fairview, Henry Farm, Oriole	43.778517	-79.346556	Michel's Baguette	43.777082	-79.344557	Bakery

Out[143]:

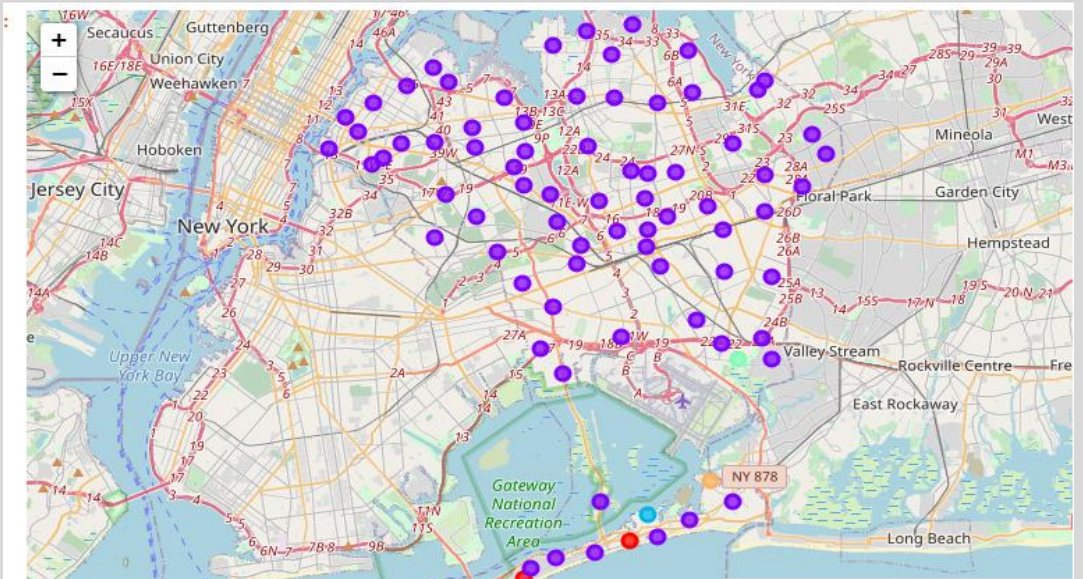


Method Part II:

New York:

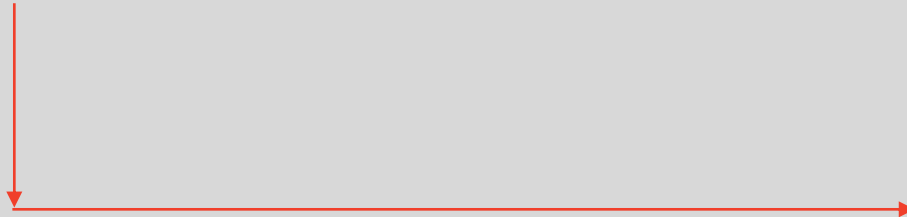
- For New York, I choose Queen.
- Almost did the same thing in Toronto.
 - Pull out and sort the Queen data.
 - Use FourSquare to find the venues around that area.
 - Use Kmean cluster to make the data into different clusters.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Astoria	40.768509	-73.915654	Favela Grill	40.767348	-73.917897	Brazilian Restaurant
1	Astoria	40.768509	-73.915654	Orange Blossom	40.769856	-73.917012	Gourmet Shop
2	Astoria	40.768509	-73.915654	Titan Foods Inc.	40.769198	-73.919253	Gourmet Shop
3	Astoria	40.768509	-73.915654	Simply Fit Astoria	40.769114	-73.912403	Gym
4	Astoria	40.768509	-73.915654	CrossFit Queens	40.769404	-73.918977	Gym
5	Astoria	40.768509	-73.915654	Off The Hook	40.767200	-73.918104	Seafood Restaurant
6	Astoria	40.768509	-73.915654	Al-sham Sweets and Pastries	40.768077	-73.911561	Middle Eastern Restaurant
7	Astoria	40.768509	-73.915654	Sitan Muay Thai	40.766108	-73.913224	Martial Arts School
8	Astoria	40.768509	-73.915654	Irish Rover	40.765835	-73.914647	Pub
9	Astoria	40.768509	-73.915654	Noisette	40.768544	-73.911248	Dessert Shop



Results: Toronto V.S New York

- I made 5 cluster for both Queen and North York.
- Let's compare the clusters from both which contains the most neighborhood.



- Comparison:
 - First of all, Queens has much more venues than North York does.
 - Queen has 2088 distinct venues. On the other hand, North York has only 236 venues.
 - In North York, cluster 1 has the most neighborhood. There is only 1 Chinese restaurant appear in the first most common column.
 - On the other hand, cluster 1 also has the most neighborhood. There are more than 5 Chinese restaurants appear in the first most common column.

Conclusion: Answering the Question

- Answer 1: Throughout my project, I found that North York and Queens are the boroughs contains the most neighborhoods in Toronto and New York City.
- Answer 2: There are many differences between each neighborhoods:
 1. Queen is way bigger than North York in terms of neighborhoods.
 2. Queen contains way more venue categories than North York, especially Chinese restaurants.
 3. Compare to Queen, North York could be a place for people to open up more business because there is a lack of a lot of venues in North York.
- Answer 3: Within my analysis, I deeply believe New York should be a place for people to study and travel. Only with the Queens area, it has contained so many different venues like pizza place, bank, yoga studio, coffee shops, and donut shop etc. People always find themselves need those venues when traveling and studying in a foreign land.
- Answer 4: With all the benefits of traveling to New York, North York should be the best place for people to open up businesses. Along with my research, North York is a big place but doesn't develop much. A lot of the venues there are living necessities. North York needs more people to invest and develop; thus, it would be a great place for people to make investments and do business.