

The battle of cities

By Hydilla Cai

2020.4.2

- Introduction
- Data acquisition and cleaning
- Methodology
- Analysis
- Results and conclusion

Introduction

- This project will mainly focus on the comparison of three districts: New York City, Toronto and Los Angeles. They are all economic and cultural centers in their area. They are all districts of my interests so in this project I will compare the segmentations of these three cities and find out the similarity of these three cities.

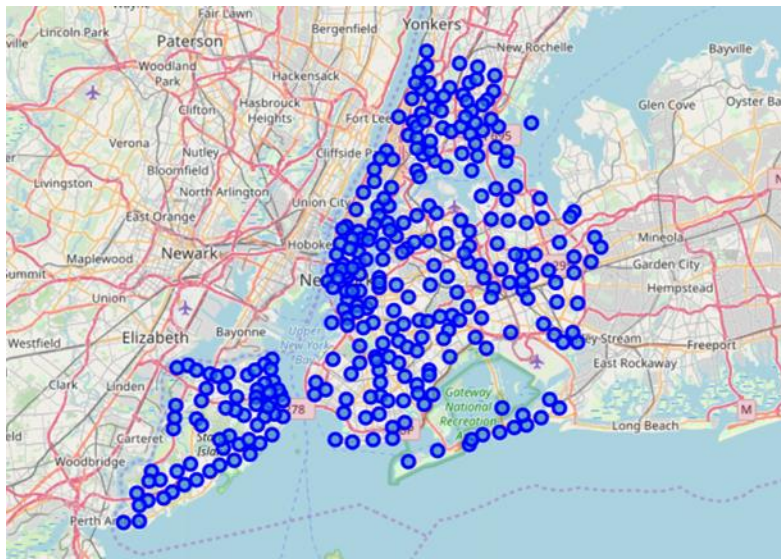
Data acquisition and cleaning

- We get the data of three districts from different sources. Since the data of New York city can be obtain from the Internet, we can use it directly, but we should turn other sources to get the data of other districts. We combine the postal code in Wikipedia and coordinates from official website to form the data. After that we get that the Los Angeles county has 88 cities so we combine the data to form a csv.

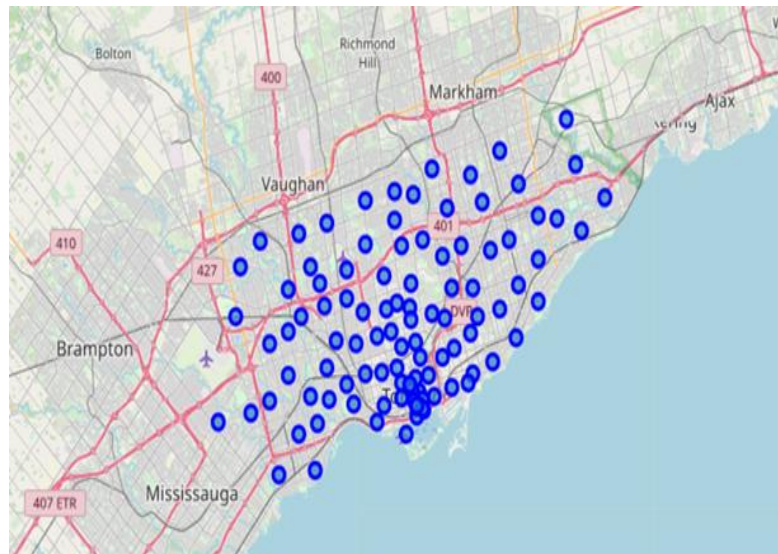
Methodology

- Clustering
- K-means

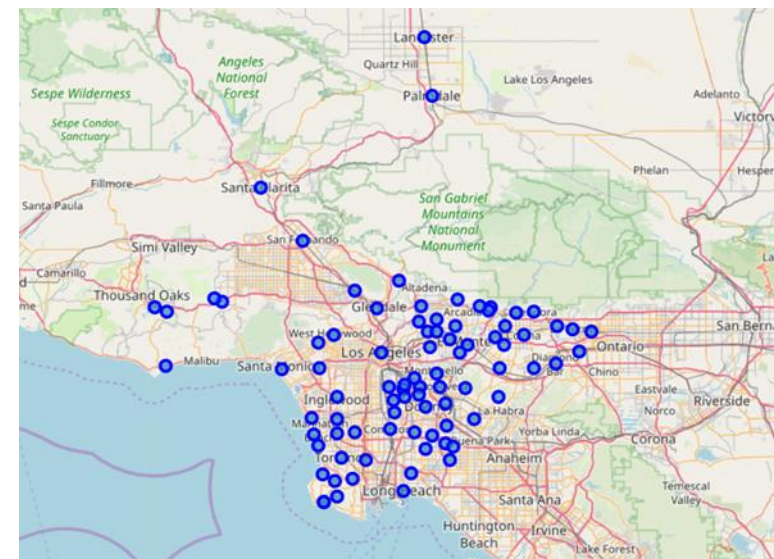
Analysis



New York city

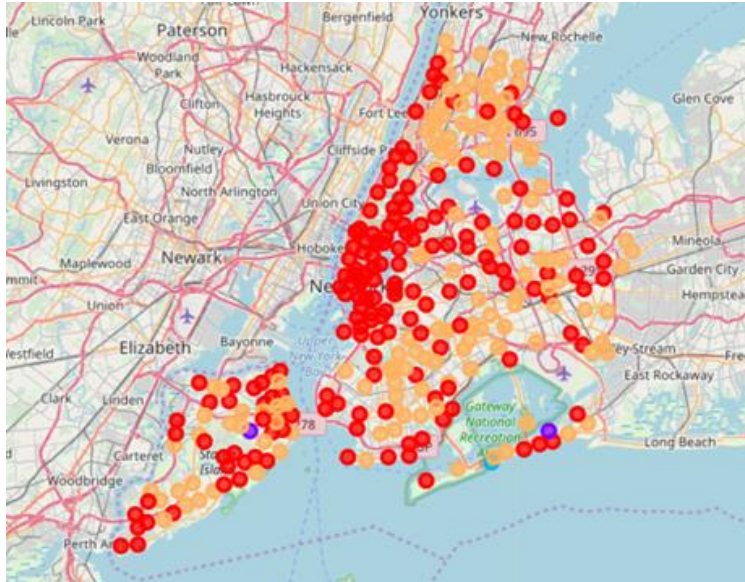


Toronto

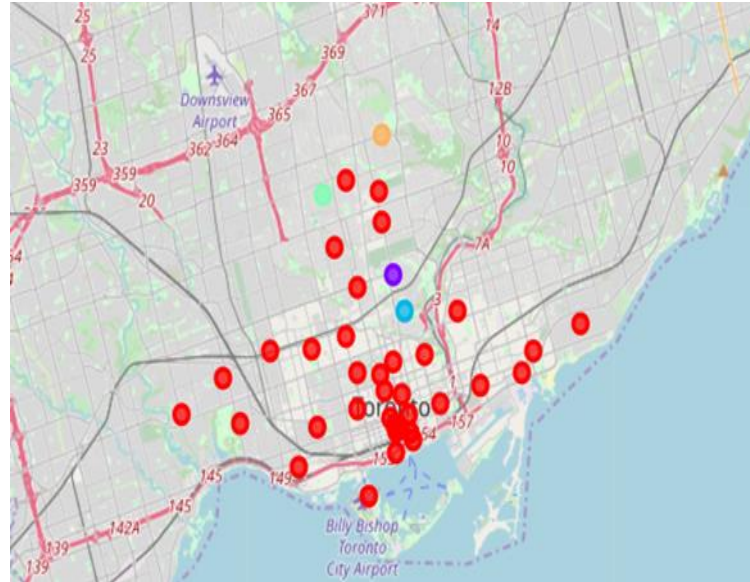


Los Angeles

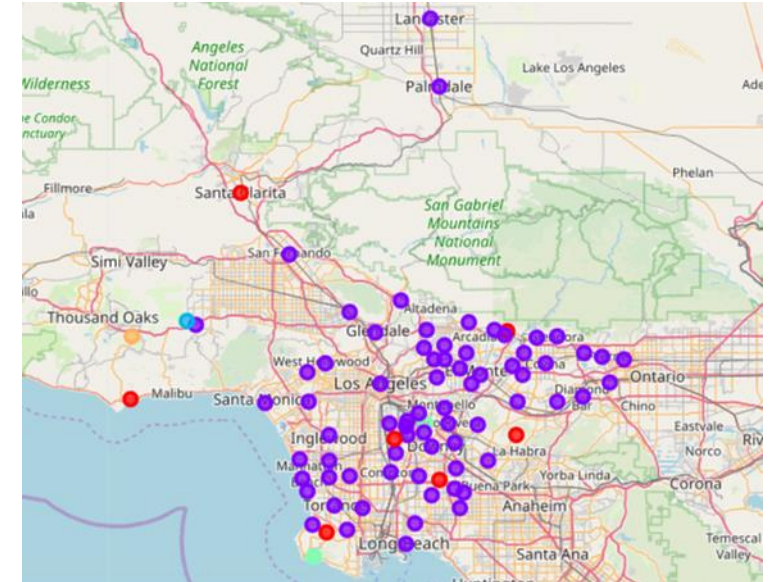
K-means



New York city



Toronto



Los Angeles

Clustering-NY

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
4	Riverdale	Park	Baseball Field	Plaza	Medical Supply Store	Bank
6	Marble Hill	Sandwich Place	Gym	Coffee Shop	Yoga Studio	Diner
10	Baychester	Donut Shop	Supermarket	Gym / Fitness Center	Music Venue	Men's Store
12	City Island	Harbor / Marina	Thrift / Vintage Store	Boat or Ferry	Seafood Restaurant	Pharmacy
16	Fordham	Mobile Phone Shop	Bank	Shoe Store	Supplement Shop	Donut Shop
22	Port Morris	Latin American Restaurant	Furniture / Home Store	Storage Facility	Restaurant	Distillery
24	Hunts Point	Juice Bar	Restaurant	Home Service	Gourmet Shop	Bank

Clustering-NY

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
192	Somerville	Park	Women's Store	Field	Entertainment Service	Ethiopian Restaurant
203	Todt Hill	Park	Women's Store	Field	Entertainment Service	Ethiopian Restaurant

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
179	Neponsit	Beach	Women's Store	Filipino Restaurant	Ethiopian Restaurant	Event Space

Clustering-NY

	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
0	Wakefield	Pharmacy	Laundromat	Donut Shop	Pizza Place	Sandwich Place
1	Co-op City	Baseball Field	Bus Station	Restaurant	Park	Chinese Restaurant
2	Eastchester	Caribbean Restaurant	Diner	Intersection	Fast Food Restaurant	Bakery
3	Fieldston	Bus Station	River	Plaza	Women's Store	Field
5	Kingsbridge	Pizza Place	Sandwich Place	Mexican Restaurant	Supermarket	Spanish Restaurant
7	Woodlawn	Pub	Pizza Place	Deli / Bodega	Playground	Convenience Store
8	Norwood	Pizza Place	Park	Bank	Pharmacy	Deli / Bodega
9	Williamsbridge	Caribbean Restaurant	Bar	Soup Place	Nightclub	Entertainment Service

Clustering-Toronto

	Borough	Clustering Label	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
37	East Toronto	0	Trail	Pub	Health Food Store	Wings Joint	Cupcake Shop
41	East Toronto	0	Greek Restaurant	Coffee Shop	Italian Restaurant	Ice Cream Shop	Bookstore
42	East Toronto	0	Sandwich Place	Pet Store	Intersection	Steakhouse	Restaurant
43	East Toronto	0	Café	Coffee Shop	Bakery	Gastropub	Brewery
45	Central Toronto	0	Food & Drink Shop	Hotel	Park	Breakfast Spot	Gym
46	Central Toronto	0	Clothing Store	Sporting Goods Shop	Coffee Shop	Restaurant	Mexican Restaurant
47	Central Toronto	0	Sandwich Place	Dessert Shop	Pizza Place	Gym	Café
49	Central Toronto	0	Pub	Coffee Shop	Sushi Restaurant	Bank	Sports Bar
51	Downtown Toronto	0	Coffee Shop	Park	Bakery	Restaurant	Italian Restaurant
52	Downtown Toronto	0	Gastropub	Gay Bar	Coffee Shop	Men's Store	Restaurant
53	Downtown Toronto	0	Coffee Shop	Pub	Park	Mexican Restaurant	Breakfast Spot

Clustering-Toronto

	Latitude	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
9	33.88200	Concert Hall	Farmers Market	Dumpling Restaurant	Eastern European Restaurant
11	34.14695	Home Service	Fast Food Restaurant	Flower Shop	Flea Market
38	33.96100	Home Service	Yoshoku Restaurant	Dry Cleaner	Flea Market
49	34.02600	Beach	Harbor / Marina	Yoshoku Restaurant	Dry Cleaner
65	33.78792	Park	Farm	Donut Shop	Flea Market
71	34.39200	Home Service	Construction & Landscaping	Carpet Store	Farm
77	33.95500	Park	Liquor Store	Farm	Eastern European Restaurant

Clustering-LA

	Latitude	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
1	34.09500	Cocktail Bar	Chinese Restaurant	Seafood Restaurant	Café
3	33.86607	Chinese Restaurant	Grocery Store	Bubble Tea Shop	Vietnamese Restaurant
4	33.34300	Seafood Restaurant	Mexican Restaurant	Pizza Place	Ice Cream Shop
5	34.13400	Pizza Place	Pharmacy	Big Box Store	Bank
6	34.08501	Ice Cream Shop	Discount Store	Pharmacy	Fast Food Restaurant
7	33.97800	Discount Store	South American Restaurant	Latin American Restaurant	Shoe Store
8	33.96507	Seafood Restaurant	Mexican Restaurant	Fried Chicken Joint	Latin American Restaurant
10	34.07400	Boutique	Park	Salon / Barbershop	American Restaurant
12	34.18100	Japanese Restaurant	Pizza Place	Sandwich Place	Ice Cream Shop

Clustering-LA

	Latitude	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue
2	33.98095	Wine Bar	Dry Cleaner	Flower Shop	Flea Market
62	33.74400	Dry Cleaner	Flower Shop	Flea Market	Fish & Chips Shop
63	33.74400	Dry Cleaner	Flower Shop	Flea Market	Fish & Chips Shop

Results and conclusion

- We can see from our analysis that the three districts have the similarities and the differences. If we divide the area in the districts of the following parts-food, recreation&relaxing,chores and so on. We find that both New York city and Los Angeles have the clustering that contains the living area. The sites are highly concentrated and New York city and Los Angeles both have an area of the living cluster.