

Capstone project-The Battle of Neighborhoods

Introduction

New York is one of the largest and well-known cities in US and across the world. It has become top choices for local and foreign communities for housing, job employment, education, shopping and tourism.

Business Problem

The objective of this report is to study and analyze Queens and Brooklyn neighborhoods of separately and group them into similar clusters and, to analyze those clusters to gather meaningful information. That information can be used to find out similarities and dissimilarities between two boroughs Queens and Brooklyn. Therefore, people can make better decision between the two boroughs on all the great amenities and other types of venues that exist in the neighborhood, such as gourmet fast food joints, pharmacies, parks, graduate schools and so on.

Data Description

To explore neighborhoods of New York, we will download the data from server: https://cocl.us/new_york_dataset. Then transfer the data into pandas frame. Use Geopy to get the latitude and the longitude coordinates of all the neighborhoods of New York. Map the neighborhoods using Folium. Use Foursquare API to get some information about venues around the neighborhoods. At last, focusing on to find similarities and dissimilarities between two boroughs use machine learning technique, "Clustering" to segment the neighborhoods.

Methodology

1. Download the data from server and transfer the data into a panda frame. This dataframe has 5 boroughs and 306 neighborhoods.

	Borough	Neighborhood	Latitude	Longitude
0	Bronx	Wakefield	40.894705	-73.847201
1	Bronx	Co-op City	40.874294	-73.829939
2	Bronx	Eastchester	40.887556	-73.827806
3	Bronx	Fieldston	40.895437	-73.905643
4	Bronx	Riverdale	40.890834	-73.912585

Create a new dataframe for borough-Queens

	Borough	Neighborhood	Latitude	Longitude
0	Queens	Astoria	40.768509	-73.915654
1	Queens	Woodside	40.746349	-73.901842
2	Queens	Jackson Heights	40.751981	-73.882821
3	Queens	Elmhurst	40.744049	-73.881656
4	Queens	Howard Beach	40.654225	-73.838138

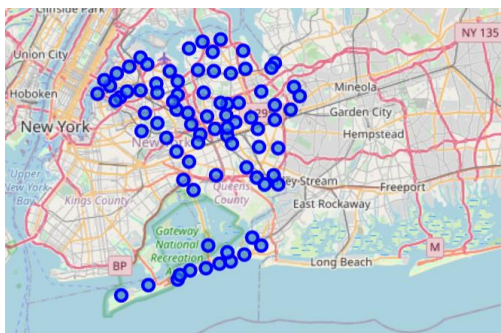
Create a new dataframe for borough-Queens

	Borough	Neighborhood	Latitude	Longitude
0	Brooklyn	Bay Ridge	40.625801	-74.030621
1	Brooklyn	Bensonhurst	40.611009	-73.995180
2	Brooklyn	Sunset Park	40.645103	-74.010316
3	Brooklyn	Greenpoint	40.730201	-73.954241
4	Brooklyn	Gravesend	40.595260	-73.973471

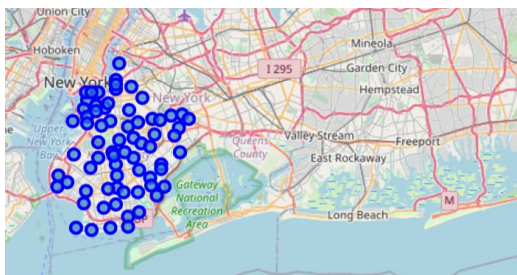
2. Generate a map of New York and plot the neighborhoods data on it. We then use the python **folium** library to visualize geographic details of New York and its boroughs. I created a map of New York with boroughs superimposed on top using the latitude and longitude values to get the visual as below:



We do the same to visualize Queens with the neighborhoods in it:



Do it one more time to visualize Brooklyn with the neighborhoods in it:



3. Utilize FourSquare API to explore the neighborhoods of Queens and Brooklyn.

We set the LIMIT parameter to 100, which would limit the number of venues returned by the Foursquare API and the radius of 500 meter. Here is a head of the list of Nearby Venues for the first neighborhood, Astoria in Queens. There are 268 unique venue categories in Queens.

	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	CrossFit Queens	40.769404	-73.918977	Gym
4	Astoria	40.768509	-73.915654	Off The Hook	40.767200	-73.918104	Seafood Restaurant

We will repeat the process above and here is a head of the list of Nearby Venues for the first neighborhood, Bay Bridge in Brooklyn. There are 290 unique venue categories in Brooklyn.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Bay Ridge	40.625801	-74.030621	Pilo Arts Day Spa and Salon	40.624748	-74.030591	Spa
1	Bay Ridge	40.625801	-74.030621	Bagel Boy	40.627896	-74.029335	Bagel Shop
2	Bay Ridge	40.625801	-74.030621	Leo's Casa Calamari	40.624200	-74.030931	Pizza Place
3	Bay Ridge	40.625801	-74.030621	Pegasus Cafe	40.623168	-74.031186	Breakfast Spot
4	Bay Ridge	40.625801	-74.030621	The Bookmark Shoppe	40.624577	-74.030562	Bookstore

4. Analyze each neighborhoods.

We use One Hot Encoding, use the neighborhood to group data, and find out the top ten venues present in each neighborhood in Queens and Brooklyn.

Queens:

	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	Arverne	Surf Spot	Sandwich Place	Metro Station	Thai Restaurant	Bus Stop	Beach	Coffee Shop	Donut Shop	Board Shop	Playground
1	Astoria	Bar	Middle Eastern Restaurant	Hookah Bar	Greek Restaurant	Deli / Bodega	Indian Restaurant	Mediterranean Restaurant	Seafood Restaurant	Bakery	Food Truck
2	Astoria Heights	Playground	Italian Restaurant	Plaza	Bus Station	Bowling Alley	Supermarket	Liquor Store	Burger Joint	Bakery	Food
3	Auburndale	Miscellaneous Shop	Discount Store	Supermarket	Korean Restaurant	Furniture / Home Store	Fast Food Restaurant	Gymnastics Gym	Toy / Game Store	Athletics & Sports	Train
4	Bay Terrace	Clothing Store	Women's Store	Kids Store	Lingerie Store	Mobile Phone Shop	Shoe Store	American Restaurant	Donut Shop	Cosmetics Shop	Coffee Shop

Brooklyn:

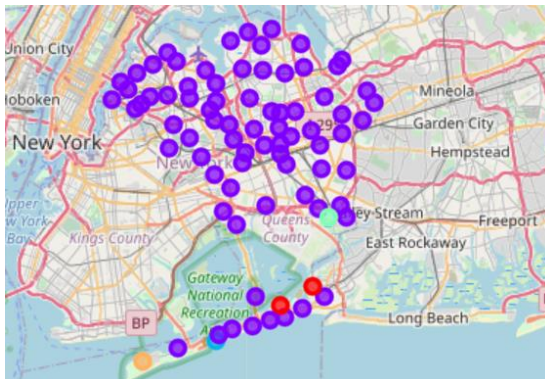
	Neighborhood	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
0	Bath Beach	Chinese Restaurant	Pharmacy	Gas Station	Bubble Tea Shop	Fast Food Restaurant	Italian Restaurant	Sushi Restaurant	Check Cashing Service	Surf Spot	Bike Rental / Bike Share
1	Bay Ridge	Spa	Italian Restaurant	Pizza Place	Grocery Store	American Restaurant	Greek Restaurant	Chinese Restaurant	Bar	Ice Cream Shop	Bagel Shop
2	Bedford Stuyvesant	Deli / Bodega	Coffee Shop	Bar	Café	Pizza Place	Tiki Bar	New American Restaurant	Gift Shop	Gourmet Shop	Juice Bar
3	Bensonhurst	Chinese Restaurant	Ice Cream Shop	Italian Restaurant	Donut Shop	Sushi Restaurant	Pizza Place	Bridal Shop	Coffee Shop	Shabu-Shabu Restaurant	Flower Shop
4	Bergen Beach	Harbor / Marina	Athletics & Sports	Other Repair Shop	Playground	Donut Shop	Baseball Field	Women's Store	Factory	Falafel Restaurant	Farm

We have some common venue categories in the neighborhoods. We use the unsupervised learning K-means algorithm to cluster the neighborhoods. We use a K-cluster value of 5 to split the neighborhoods into 5 different clusters based on the similarity the venues they contain.

Results

We use the matplotlib and folium packages to visualize the clusters on a map of Queens and Brooklyn.

Queens:



Cluster 1

	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
63	Somerville	Park	Women's Store	Fast Food Restaurant	Empanada Restaurant	Event Space	Falafel Restaurant	Farm	Farmers Market	Filipino Restaurant	Egyptian Restaurant
79	Bayswater	Park	Playground	Women's Store	Fast Food Restaurant	Empanada Restaurant	Event Space	Falafel Restaurant	Farm	Farmers Market	Filipino Restaurant

Cluster 2

	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	Astoria	Bar	Middle Eastern Restaurant	Hookah Bar	Greek Restaurant	Deli / Bodega	Indian Restaurant	Mediterranean Restaurant	Seafood Restaurant	Bakery	Food
1	Woodside	Thai Restaurant	Grocery Store	Filipino Restaurant	Latin American Restaurant	Bakery	American Restaurant	Pub	Donut Shop	Bar	Ice Cream Shop
2	Jackson Heights	Latin American Restaurant	Peruvian Restaurant	South American Restaurant	Bakery	Mexican Restaurant	Thai Restaurant	Mobile Phone Shop	Pizza Place	Spanish Restaurant	Empire Restaurant

Cluster 3

	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
50	Neponsit	Beach	Women's Store	Filipino Restaurant	Event Space	Falafel Restaurant	Farm	Farmers Market	Fast Food Restaurant	Fish & Chips Shop	French Restaurant

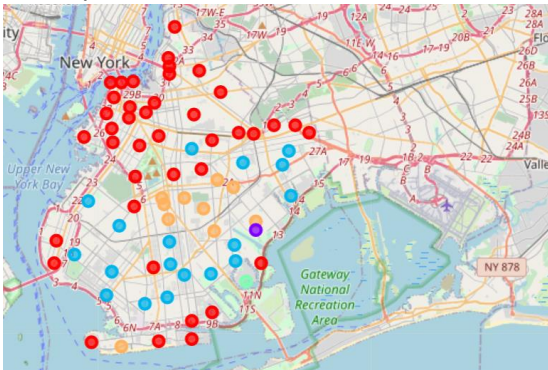
Cluster 4

	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
64	Brookville	Deli / Bodega	Women's Store	Fish & Chips Shop	Event Space	Falafel Restaurant	Farm	Farmers Market	Fast Food Restaurant	Filipino Restaurant	Fish Market

Cluster 5

	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
43	Breezy Point	Monument / Landmark	Trail	Bus Stop	Beach	Women's Store	Filipino Restaurant	Falafel Restaurant	Farm	Farmers Market	Fast Food Restaurant

Brooklyn:



Cluster 1

	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	Bay Ridge	Spa	Italian Restaurant	Pizza Place	Grocery Store	American Restaurant	Greek Restaurant	Chinese Restaurant	Bar	Ice Cream Shop	Bagel Shop
3	Greenpoint	Bar	Pizza Place	Coffee Shop	Cocktail Bar	Yoga Studio	Grocery Store	Thrift / Vintage Store	Sushi Restaurant	Bakery	Café
5	Brighton Beach	Beach	Restaurant	Russian Restaurant	Eastern European Restaurant	Sushi Restaurant	Gourmet Shop	Mobile Phone Shop	Bank	Bakery	Taco Place

Cluster 2

	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
59	Paerdegat Basin	Asian Restaurant	Moving Target	Food	Auto Garage	Harbor / Marina	Women's Store	Falafel Restaurant	Farm	Farmers Market	Fast Food Restaurant

Cluster 3

	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
1	Bensonhurst	Chinese Restaurant	Ice Cream Shop	Italian Restaurant	Donut Shop	Sushi Restaurant	Pizza Place	Bridal Shop	Coffee Shop	Shabu-Shabu Restaurant	Flower Shop
2	Sunset Park	Latin American Restaurant	Bakery	Bank	Pizza Place	Mexican Restaurant	Pharmacy	Deli / Bodega	Fried Chicken Joint	Mobile Phone Shop	Gym
4	Gravesend	Pizza Place	Bakery	Chinese Restaurant	Bus Station	Martial Arts Dojo	Lounge	Hookah Bar	Diner	Pharmacy	Donut Shop

Cluster 4

	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
30	Mill Island	Other Repair Shop	Pool	Women's Store	Field	Factory	Falafel Restaurant	Farm	Farmers Market	Fast Food Restaurant	Filipino Restaurant

Cluster 5

	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
8	Flatbush	Coffee Shop	Deli / Bodega	Mexican Restaurant	Caribbean Restaurant	Bank	Plaza	Middle Eastern Restaurant	Lounge	Liquor Store	Sandwich Place
10	East Flatbush	Pharmacy	Wine Shop	Print Shop	Caribbean Restaurant	Department Store	Food & Drink Shop	Hardware Store	Supermarket	Chinese Restaurant	Fast Food Restaurant
28	Canarsie	Gym	Asian Restaurant	Chinese Restaurant	Caribbean Restaurant	Food	Flower Shop	Fish Market	Fish & Chips Shop	Filipino Restaurant	Event Service

Discussion

After clustering the data of the respective neighborhoods, both Boroughs have venues which can be explored and attract people. The neighborhoods are much similar in features like food places, clubs, gyms, parks, shops etc. We can see that both boroughs has varieties of stores and places for convenience and leisure in life. Another observation is both boroughs have a variety of cuisines, Asian, Middle Eastern, Mexico, Italian, etc., which indicating a lot diversities in the neighborhoods. As far as concern to dissimilarity, it differs in terms of some unique places like landmark and monuments which we did not see any of historical places in Brooklyn neighborhoods. Maybe Brooklyn is newly developed than Queens and has a shorter history.

Conclusion

Queens and Brooklyn neighborhoods have more like similar venues. As we know that every place is unique in its own way, so that's argument is present in both neighborhoods. The dissimilarity exists in terms of some different venues and facilities but not on a larger extent.

