

Compare the Neighborhoods of Manhattan and Brooklyn of New York City & its venues data analysis

1. Introduction

Description & Discussion of the Background

New York City (NYC) or New York (NY), is the most populous city in the United States with an estimated population of 83million distributed over about 302.6 square miles (784 km²). New York City has been described as the cultural, financial, and media capital of the world. New York City is composed of five boroughs: The Bronx, Brooklyn, Manhattan, Queens, and Staten Island.

Manhattan is the geographically smallest and most densely populated borough of New York City. With an estimated 1,628,701 residents, Manhattan's population density is 72,033 people per square mile (27,812/km²) in 2018 which makes it the highest populated city in the United States. Brooklyn is the second-largest among the New York City's five boroughs which is the most populous borough and , the second-most densely populated county in the United States. With an estimated 2,559,903 residents in 2019, Brooklyn's population density is 37,137 people per square mile (14,649/km²) in 2019

From figures, we can see that both Manhattan and Brooklyn are densely populated with high population density. So, when somebody thinks of opening a restaurant or a new business, they must choose a place with high social density or a place with population dense, but no restaurants or shops are available as a trial.

So, the questions are

- Whether both boroughs -Manhattan & Brooklyn are similar or dissimilar with respect to their neighborhoods?
- Where to start a new business or restaurant?
- Which borough is more suitable for starting a new business or restaurant?

In order to obtain the information about places and venues, we can create a map of Manhattan and Brooklyn and see how both the neighborhoods are clustered according to venues. This data could be used to see how similar or dissimilar both boroughs are with respect to venues and which place is best suitable for opening a new business or restaurant.

2. Description of Data

a. Data Source

The data for neighborhoods of New York City was obtained from the Spatial Data Repository of NYU. The .json file obtained has the coordinates for all the neighborhoods of New York City.

b. Data Cleaning

I cleaned and reduced the data to the Boroughs Manhattan and Brooklyn. For that I first transformed the data of NYC to pandas data frame and then sliced it into two new data frames for Manhattan and Brooklyn. Then used geopy library for obtaining the latitude and longitude values of Manhattan and Brooklyn. Then created a map of Manhattan and Brooklyn using Folium.

c. Explore neighborhoods using Foursquare data

After that I used Foursquare API for exploring the venues and Neighborhoods in Manhattan and Brooklyn.

d. Analyses the neighborhoods data

Further analyze each neighborhood by seeing how many venues are there in each neighborhood.

e. Cluster the neighborhoods and examine the clusters for comparing both boroughs and solving the problem of where to start a new business.

Then cluster the neighborhoods and examine the clusters formed to determine discriminating venue categories that distinguish each cluster. Using this cluster information, we can identify how similar or dissimilar are both boroughs and also one can decide where to open a new restaurant or shop.

3. Methodology

The data for neighborhoods of New York City was obtained from the Spatial Data Repository of NYU. It consists of features – Borough, Neighborhood, Latitude and Longitude. After downloading and loading data, I transformed the neighborhood data into Pandas Dataframe. The neighborhood data contain data for 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

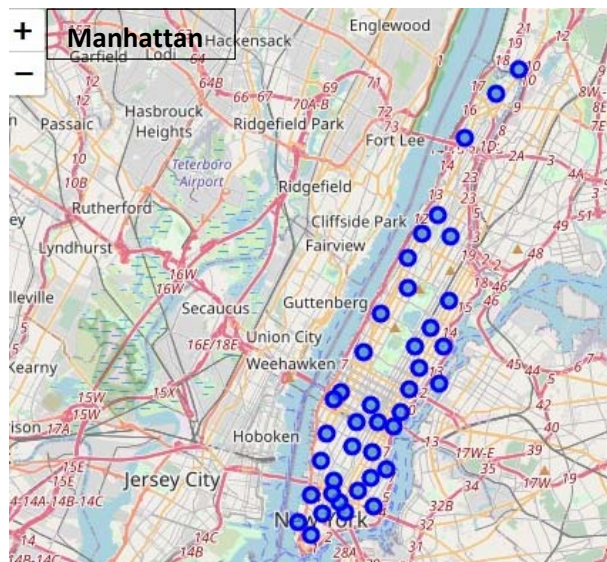
From this I selected the neighborhood data of Manhattan and Brooklyn for comparing their neighborhoods.

	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

	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

Manhattan had 40 neighborhoods whereas Brooklyn had 70 neighborhoods. From data itself Manhattan is smaller when compared to Brooklyn, hence had high population density.

I used the geopy library to obtain the latitude and longitude values and then used Folium to visualize the geographic details of Manhattan and Brooklyn.

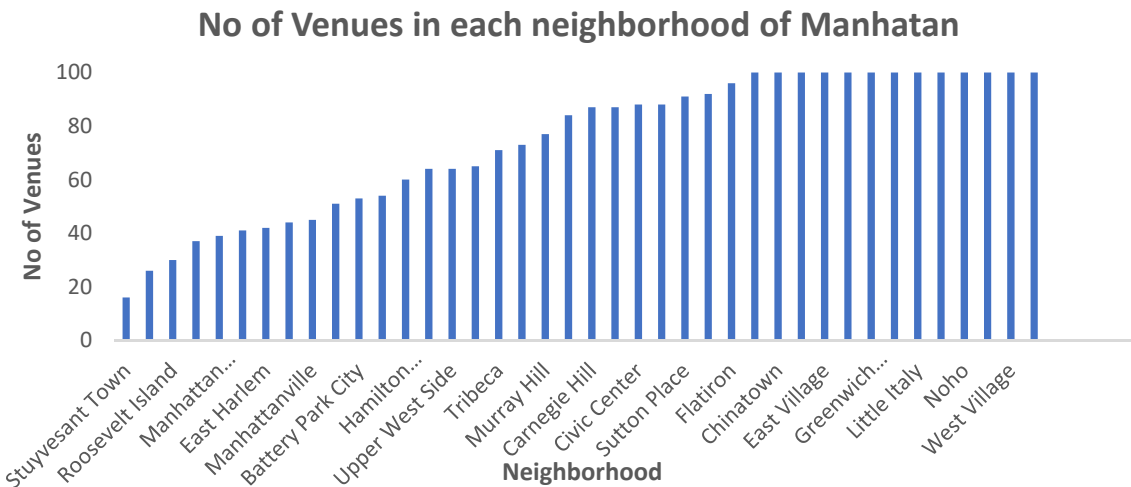


Foursquare API was utilized for exploring the venues of Manhattan and Brooklyn. I used a limit of 100 venues with a radius of 500 meters for both Manhattan and Brooklyn borough. Here is the head of the list of Venues name, category, latitude, and longitude information from the Foursquare API.

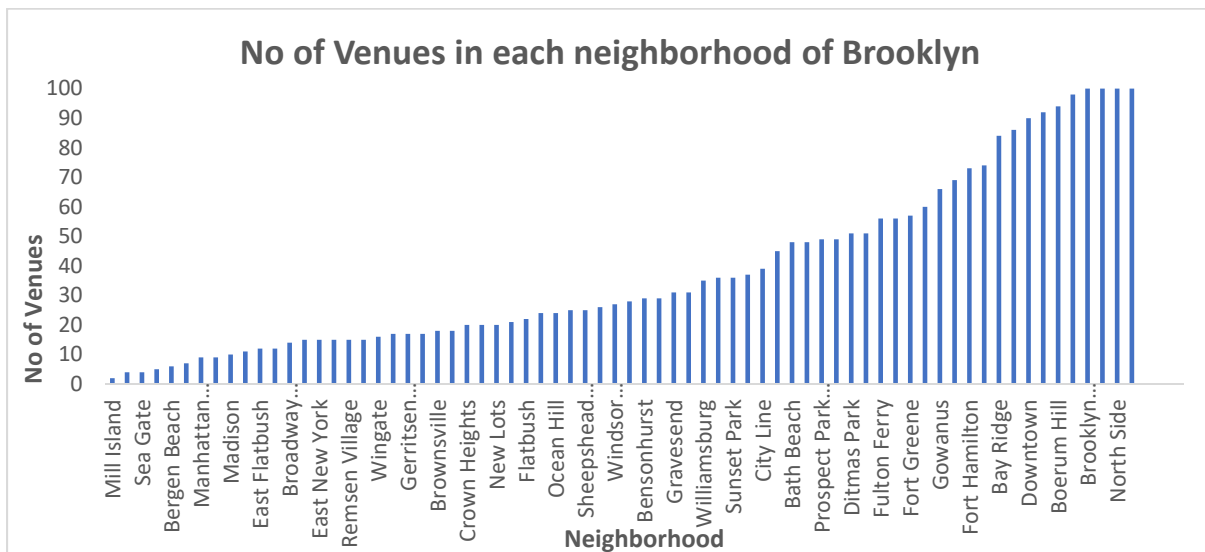
a	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Marble Hill	40.876551	-73.91066	Arturo's	40.874412	-73.910271	Pizza Place
1	Marble Hill	40.876551	-73.91066	Bikram Yoga	40.876844	-73.906204	Yoga Studio
2	Marble Hill	40.876551	-73.91066	Tibbett Diner	40.880404	-73.908937	Diner
3	Marble Hill	40.876551	-73.91066	Starbucks	40.877531	-73.905582	Coffee Shop
4	Marble Hill	40.876551	-73.91066	Dunkin'	40.877136	-73.906666	Donut Shop

	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

When I plotted the venues for each neighborhood, 50% of neighborhoods of Manhattan had more than 80 venues and 13 neighborhoods out of total 40 neighborhoods have reached the 100 limit of venues



But when it comes to Brooklyn, more than 70% of neighborhoods had less than 30 venues and only 4 neighborhoods out of 70 neighborhoods reached the 100 limits for venues.



In summary 323 unique categories were returned for Manhattan and 290 for Brooklyn by Foursquare. Then I created two tables which shows list of top 10 venue category for both boroughs

Manhattan and Brooklyn. The tables are shown below- first one Manhattan and second one 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	Battery Park City	Park	Gym	Hotel	Memorial Site	Wine Shop	Gourmet Shop	Boat or Ferry	Burger Joint	Food Court	Shopping Mall
1	Carnegie Hill	Coffee Shop	Café	Yoga Studio	Pizza Place	Bar	Bookstore	Gym	Gym / Fitness Center	Japanese Restaurant	Italian Restaurant
2	Central Harlem	Bar	Gym / Fitness Center	Chinese Restaurant	African Restaurant	American Restaurant	Seafood Restaurant	French Restaurant	Library	Gym	Beer Bar
3	Chelsea	Art Gallery	Coffee Shop	Italian Restaurant	Seafood Restaurant	Ice Cream Shop	Market	Café	Bakery	Clothing Store	Bookstore
4	Chinatown	Cocktail Bar	Chinese Restaurant	Bakery	Bar	Optical Shop	Coffee Shop	Spa	Salon / Barbershop	American Restaurant	Boutique

	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	Sushi Restaurant	Bubble Tea Shop	Italian Restaurant	Gas Station	Fast Food Restaurant	Pizza Place	Burger Joint	Sandwich Place
1	Bay Ridge	Spa	Italian Restaurant	Pizza Place	American Restaurant	Greek Restaurant	Bar	Thai Restaurant	Sushi Restaurant	Cosmetics Shop	Bagel Shop
2	Bedford Stuyvesant	Pizza Place	Bar	Coffee Shop	Café	Cocktail Bar	Japanese Restaurant	Gift Shop	Gourmet Shop	Boutique	Juice Bar
3	Bensonhurst	Sushi Restaurant	Ice Cream Shop	Grocery Store	Donut Shop	Italian Restaurant	Pet Store	Bagel Shop	Cha Chaan Teng	Liquor Store	Chinese Restaurant
4	Bergen Beach	Harbor / Marina	Donut Shop	Playground	Baseball Field	Athletics & Sports	Women's Store	Falafel Restaurant	Farm	Farmers Market	Fast Food Restaurant

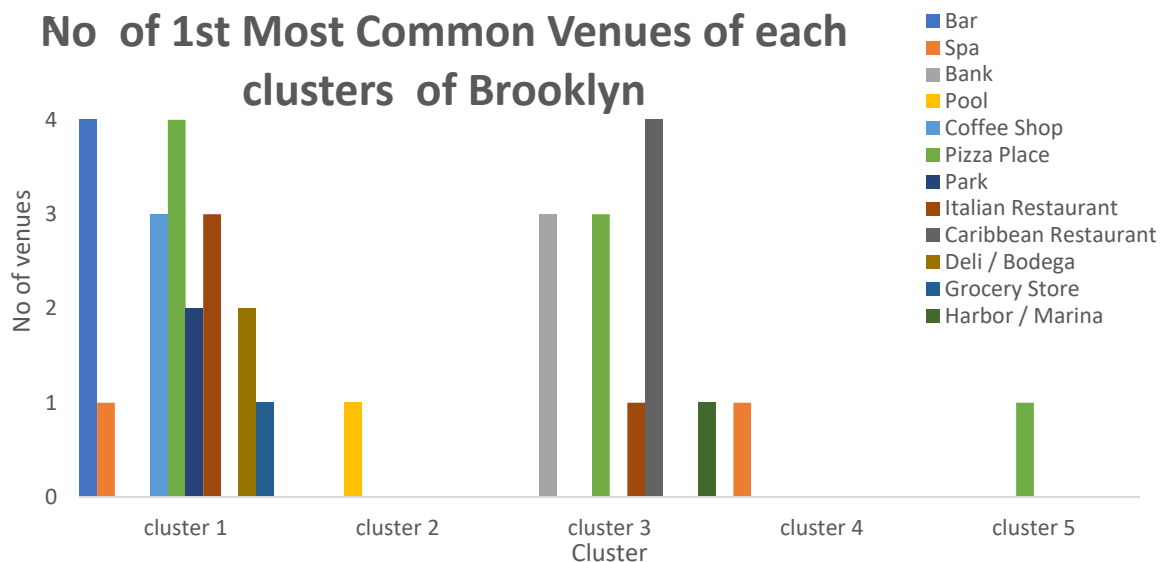
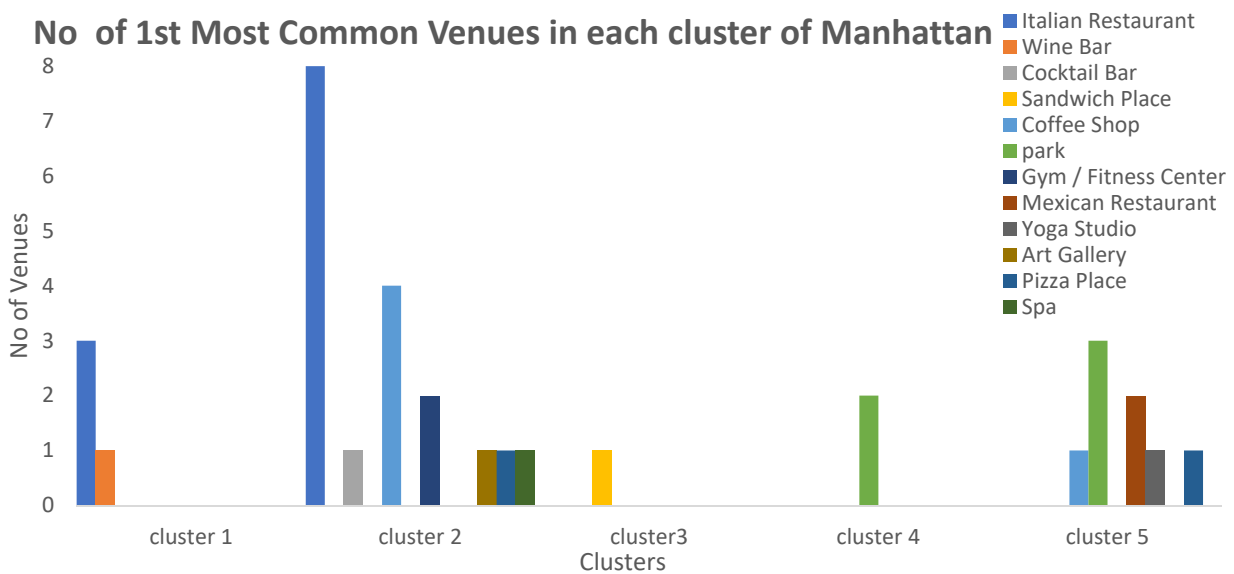
Since my aim was to compare and see whether both boroughs are similar or dissimilar, I used K-Means algorithm which is an unsupervised method to cluster the neighborhoods. For that I run k -means to cluster both the neighborhood into 5 clusters each.

Here is my merged table with cluster labels for both boroughs – Manhattan & Brooklyn.

	Borough	Neighborhood	Latitude	Longitude	Cluster Label	Cluster Labels01	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	2	2	Sandwich Place	American Restaurant	Gym	Coffee Shop	Yoga Studio	Donut Shop	Tennis Stadium	Supplement Shop	Steakhouse	Seafood Restaurant
1	Manhattan	Chinatown	40.715618	-73.994279	1	1	Cocktail Bar	Chinese Restaurant	Bakery	Bar	Optical Shop	Coffee Shop	Spa	Salon / Barbershop	American Restaurant	Boutique
2	Manhattan	Washington Heights	40.851903	-73.936900	4	4	Café	Bakery	Deli / Bodega	Mobile Phone Shop	Chinese Restaurant	Grocery Store	Donut Shop	Tapas Restaurant	Mexican Restaurant	Supplement Shop
3	Manhattan	Inwood	40.867684	-73.921210	4	4	Mexican Restaurant	Bakery	Café	Pizza Place	Lounge Restaurant	Park	Chinese Restaurant	Frozen Yogurt Shop	American Restaurant	
4	Manhattan	Hamilton Heights	40.823604	-73.949688	4	4	Pizza Place	Coffee Shop	Café	Mexican Restaurant	Deli / Bodega	Cocktail Bar	Indian Restaurant	Sushi Restaurant	Park	Yoga Studio

	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	Brooklyn	Bay Ridge	40.625801	-74.030621	0	Spa	Italian Restaurant	Pizza Place	American Restaurant	Greek Restaurant	Bar	Thai Restaurant	Sushi Restaurant	Cosmetics Shop	Bagel Shop
1	Brooklyn	Bensonhurst	40.611009	-73.995180	2	Sushi Restaurant	Ice Cream Shop	Grocery Store	Donut Shop	Italian Restaurant	Pet Store	Bagel Shop	Cha Chaan Teng	Liquor Store	Chinese Restaurant
2	Brooklyn	Sunset Park	40.645103	-74.010316	2	Latin American Restaurant	Pizza Place	Mexican Restaurant	Bank	Bakery	Ice Cream Shop	Mobile Phone Shop	Gym	Fried Chicken Joint	Women's Store
3	Brooklyn	Greenpoint	40.730201	-73.954241	0	Pizza Place	Coffee Shop	Bar	Cocktail Bar	Yoga Studio	Sushi Restaurant	Restaurant	Grocery Store	Tea Room	Salon / Barbershop
4	Brooklyn	Gravesend	40.595260	-73.973471	2	Pizza Place	Lounge	Bus Station	Chinese Restaurant	Bakery	Italian Restaurant	Furniture / Home Store	Hookah Bar	Pharmacy	Metro Station

Now the number of **1st Most Common Venues** in each cluster of both Manhattan and Brooklyn were estimated. The data is plotted as bar graph below.



4. Results & Discussion

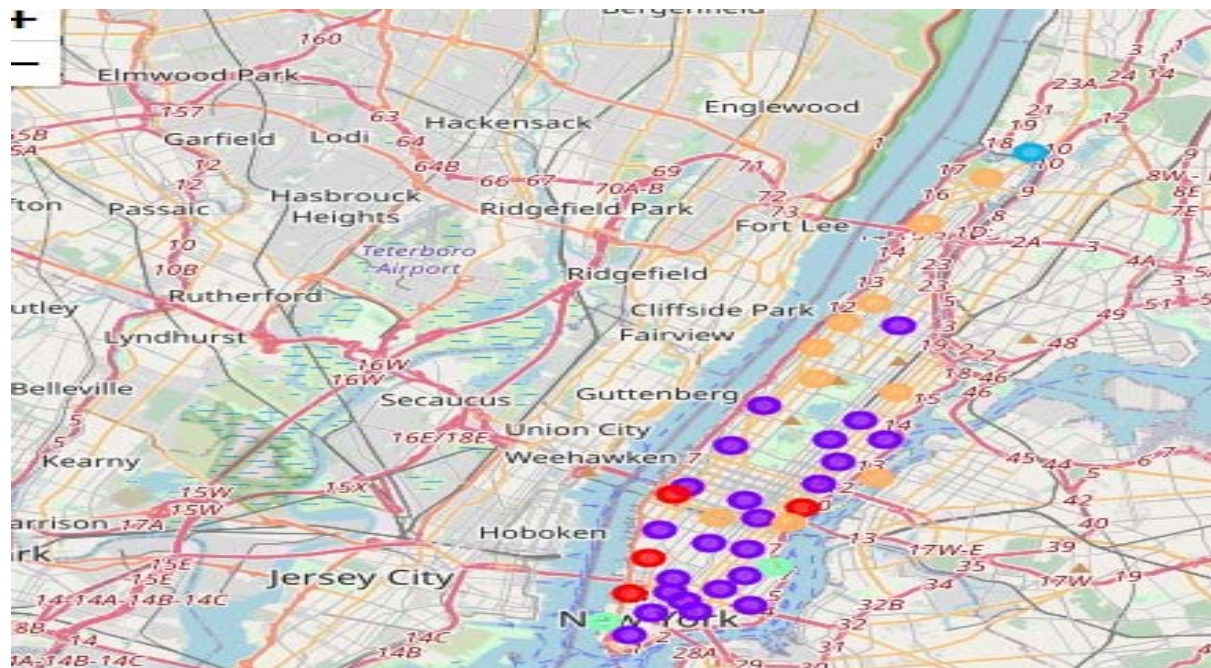
Manhattan and Brooklyn are two populous boroughs of New York city. Geographically Brooklyn is bigger in size than Manhattan. When looking the people per square feet Manhattan's population density is high. Since both cities are densely populated, in order to understand how both boroughs are similar or dissimilar by venues, I used k means clustering for segmenting and examined the neighborhoods. For clustering I used the random k value as 5. So, I clustered both the boroughs into 5clusters each.

When I examine the bar charts of 1st most common venues of each cluster, I labelled clusters with some possible names. The table of possible names are shown below

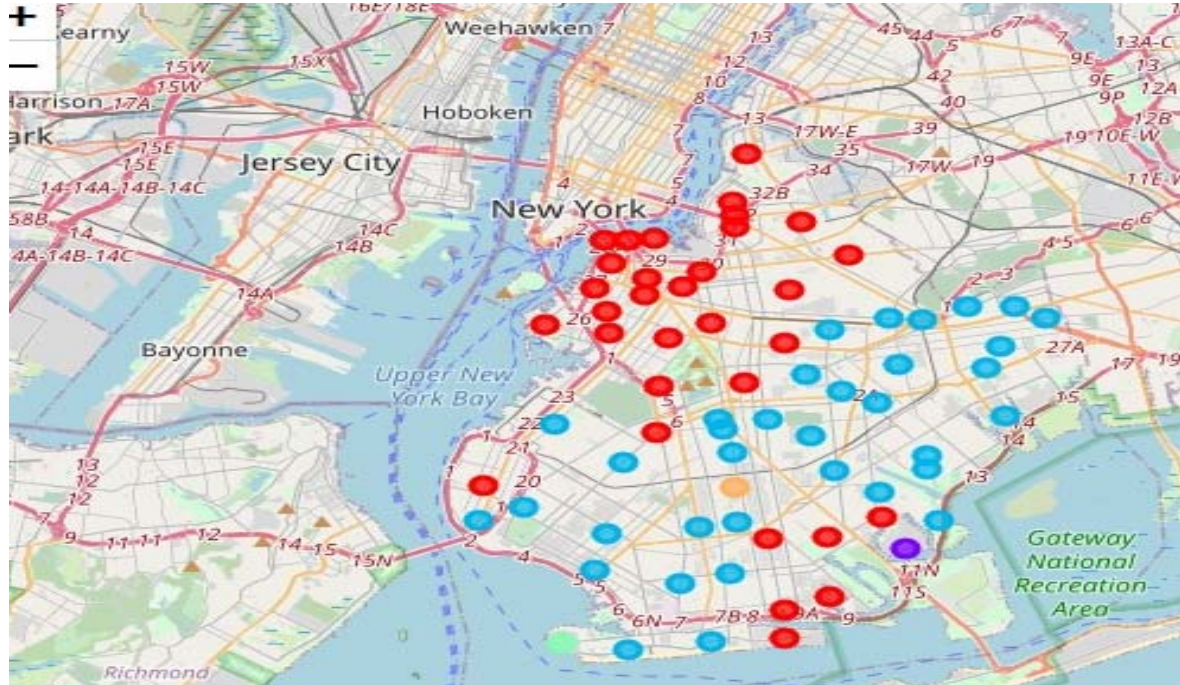
	Manhattan	Brooklyn
Cluster 1	Restaurants & Bar	Food & social venues
Cluster 2	All in one place	Recreation
Cluster 3	Eateries	All in one place
Cluster 4	Recreation	Recreation
Cluster 5	Food & social venues	Pizza place

From this I found that both the boroughs are similar with kind of venues but dissimilar in no of venues per neighborhood. Both boroughs had 2 clusters having almost all types of venues (food, recreation, park, fitness, park etc.) and remaining clusters are with either eateries or recreation venues.

In order to further examine, a clustered map for both boroughs were plotted and visualized using Folium. The clustered map of Manhattan is shown below.



The clustered map of Brooklyn is shown below.



5. Conclusion

Both Manhattan and Brooklyn are densely populated boroughs of New York City. One of my aim of the study was to compare both boroughs for the similarity. I found that both boroughs had similar kind of venues. But coming to number of venues, Manhattan is really congested. One of the reasons for that is its area. Its smaller in size than Brooklyn.

My second aim of the study was where to open a new business or restaurant. In my opinion its best to start in Brooklyn, because many of neighborhoods had less than 40 venues even though it's the second densely populated borough in New York. If some one is planning to open a restaurant in Manhattan, then hey should be really prepared for competition.

As a future study, housing sale price can be incorporated in this data and see for real estate business

6. References

- Wikipedia – New York City, Manhattan, Brooklyn
- Spatial Data Repository of NYU
- Forsquare API