

**Segmenting and clustering New York City and find best location for new Pharmacy using
IBM Watson studio and foursquare website.**

Nada Abouebeid

July 25, 2020

1.1 Introduction

New York City, is the most populous city in the United States. With an estimated 2019 population of 8,336,817 distributed over about 302.6 square miles (784 km²), New York City is also the most densely populated major city in the United States. Located at the southern tip of the U.S. state of New York, the city is the center of the New York metropolitan area, the largest metropolitan area in the world by urban landmass. [1]

1.2 Problem

Due to the increased demand, the interest of entrepreneur to open medical investment project in New York has increased drastically. Which led to the large number of health centers and medical offices in all parts of the city. This fact is why it is so difficult to find the right locations to open new medical office. In this project I am focusing on finding the best locations in the city to open new community pharmacy. Many factors we should keep in our mind before choosing any location.

1.3. Key points to Pick the Best Pharmacy Location

Deciding where to locate is one of the first major choices you'll make while building your pharmacy business. Based on article for Mr. Lester Nathan, MS in Pharmacy times journal-2015-09-14. He stated that "The location you choose for your pharmacy will highly contribute to its success or demise".

The Key points:

1-Adjacent to an upscale residential area

2-Next to a chain store pharmacy

3- Nearby medical arts building

If these characteristics are not simultaneously available, then always go with the first: being in or near an upscale neighborhood. An alternate approach is to be on or close to a highway that connects high-income families to your store. Mr. Lester Nathan stated "positioning the pharmacy inside a medical building is bad today". But being around medical building has positive impact as it increases number of interested patients and customers.

Being in a shopping mall, or free-standing building is good, but being next to a chain store like CVS, Rite Aid, or Walgreens is the best. This has been validated by research from MIT's famous Sloan Management School. Your sales will increase because the chain stores bring in people who are mainly there to find something to aid their

maladies. These people will notice how different you are as long as you truly differentiate yourself.[2]

1.4 Interested

Our Stakeholder are entrepreneur who want to invest in healthcare facilities especially retail or community pharmacy. As demand sharply increased during the current Covid-19 pandemic. Finding the best location can make the difference in tens of thousands of dollars in high-margin sales.

2. Data acquisition and cleaning

2.1 Data sources

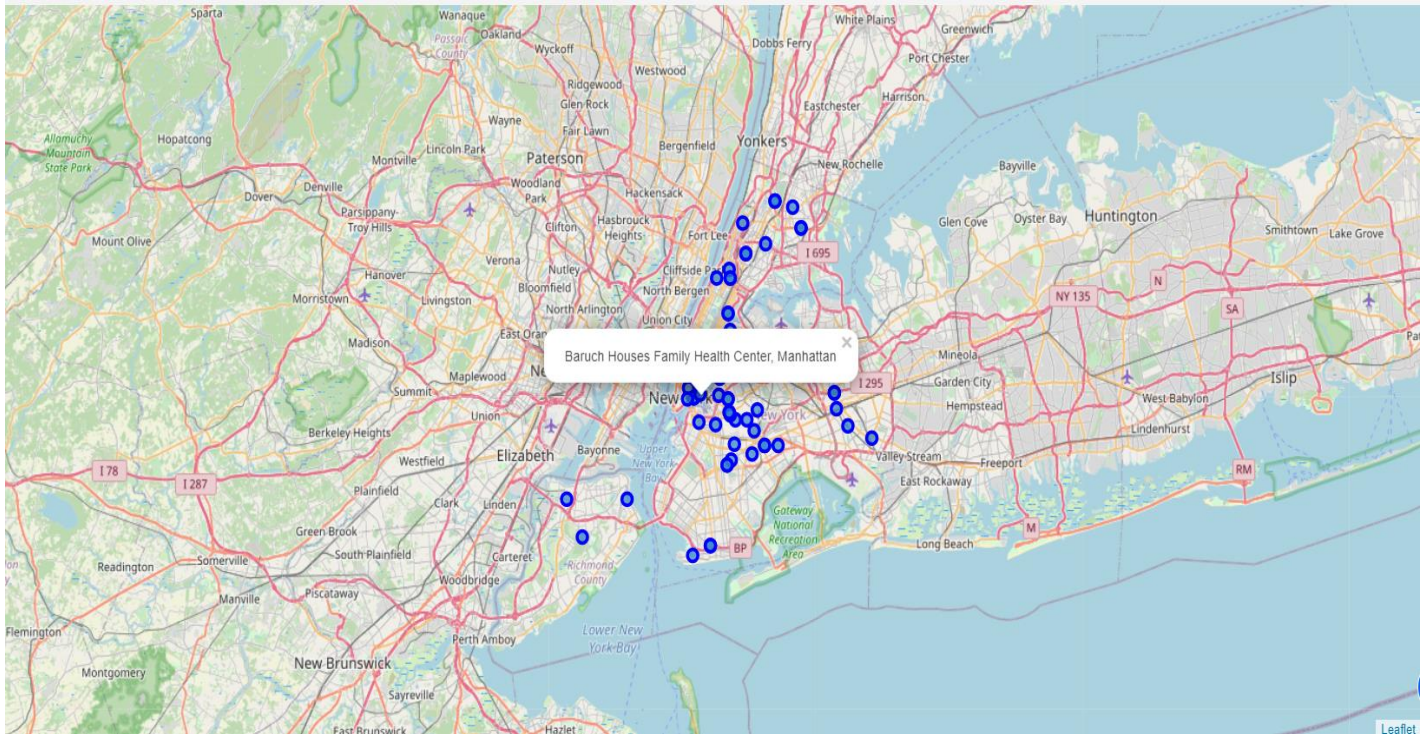
The Data we are using in this project collected from NYC Open Data website. It presents all the healthcare facilities in NYC till July 3, 2019. I decided to fix one of the factors to find the location. So, I chose to fix being nearby medical buildings in this project. It is helpful to limit the options and to grantee the foot flow of interested customers 'patients.

2.2 Data cleaning

The data cleaned and tabled in pandas. Separated the longitude and latitude to prepare the folium map with dot markers of the 60 facilities locations in NYC map.

	borough	facility_name	facility_type	latitude	longitude	postcode
0	Brooklyn	Ida G. Israel Community Health Center	Child Health Center	40.578468	-73.989614	11224.0
1	Queens	South Queens Community Health Center	Child Health Center	40.688615	-73.785593	11434.0
5	Queens	Parsons Communicare Clinic	Child Health Center	40.703826	-73.800483	11432.0
6	Queens	Queens Hospital Center	Child Health Center	40.717250	-73.803433	11432.0
9	Queens	Junction Boulevard Child Health Clinic	Child Health Center	40.754179	-73.872231	11372.0

after finding the geographical coordinate of New York City (40.7127281, -74.0060152) we visualize it in folium map



Out[132]:

	name	categories	lat	lng
0	The Bar Room at Temple Court	Hotel Bar	40.711448	-74.006802
1	The Beekman, A Thompson Hotel	Hotel	40.711173	-74.006702
2	Alba Dry Cleaner & Tailor	Laundry Service	40.711434	-74.006272
3	City Hall Park	Park	40.712415	-74.006724
4	Gibney Dance Center Downtown	Dance Studio	40.713923	-74.005661
5	Aahar Indian Cuisine	Indian Restaurant	40.713307	-74.007994
6	Woolworth Building	Building	40.712559	-74.007964
7	Wafels & Dinges - Pagadder Cart	Dessert Shop	40.712586	-74.004952
8	The Helena Gym	Gym	40.714276	-74.005967
9	Barnes & Noble	Bookstore	40.711914	-74.006145
10	Modell's Sporting Goods	Sporting Goods Shop	40.714198	-74.006142
11	Potbelly Sandwich Shop	Sandwich Place	40.714454	-74.005820
12	Municipal Plaza	Plaza	40.712755	-74.004498
13	CityStore	Gift Shop	40.713266	-74.004004
14	The Beekman Pub	Bar	40.711102	-74.006104
15	Brooklyn Chop House	Steakhouse	40.711528	-74.005857
16	Mary's Coffee Shop	Coffee Shop	40.712786	-74.005944
17	IRT Subway - City Hall (Abandoned)	Museum	40.712781	-74.004727

We used foursquare API to study the venues up to 200 meters away from the healthcare facilities marked on the map. The collected data cleaned and presented in tables with latitude and longitude.

In summary of this table **238** unique categories were returned by Foursquare, then I created a table which shows list of top 10 venue category for each borough in below table.

	facility_name	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	Baruch Houses Family Health Center	Chinese Restaurant	Park	Art Gallery	Yoga Studio	Bakery	Flower Shop	Café	Soccer Field	Pharmacy	Sculpture Garden
1	Bellevue Hospital Center	Bar	American Restaurant	Bagel Shop	Pizza Place	Grocery Store	Ice Cream Shop	Italian Restaurant	Coffee Shop	Yoga Studio	Big Box Store
2	Brownsville Child Health Clinic	Chinese Restaurant	Fried Chicken Joint	Park	Farmers Market	Restaurant	Kids Store	Grocery Store	Trail	Spanish Restaurant	Burger Joint
3	Bushwick Communitcare	Café	Bar	Latin American Restaurant	Pizza Place	Frozen Yogurt Shop	Fast Food Restaurant	Supermarket	Sandwich Place	Cocktail Bar	Coffee Shop
4	Bushwick Community Health Center	Deli / Bodega	Mexican Restaurant	Chinese Restaurant	Pizza Place	Playground	Department Store	Café	Garden	Restaurant	Fried Chicken Joint

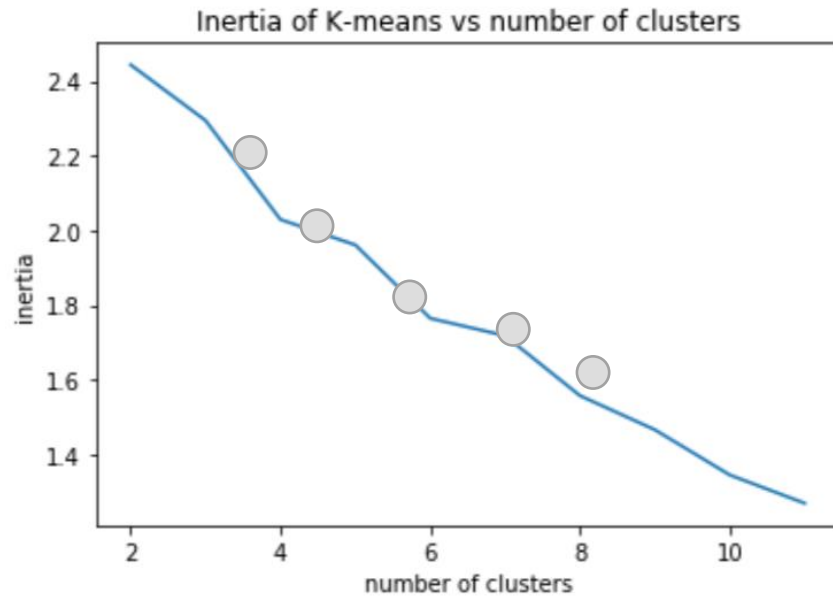
3. Machine learning and clustering

I got some common venue categories. In this reason I used unsupervised learning K-means algorithm to cluster the healthcare facilities locations. K-Means algorithm is one of the most common cluster methods of unsupervised learning.

3.1. K-Means and elbow method

First, will run K-Means to cluster the boroughs into **5 clusters** because when I analyze the K-Means with elbow method it showed me around 6 degrees

for optimum k of the K-Means. 5 clusters quite enough to understand the similarities between the clustered locations.

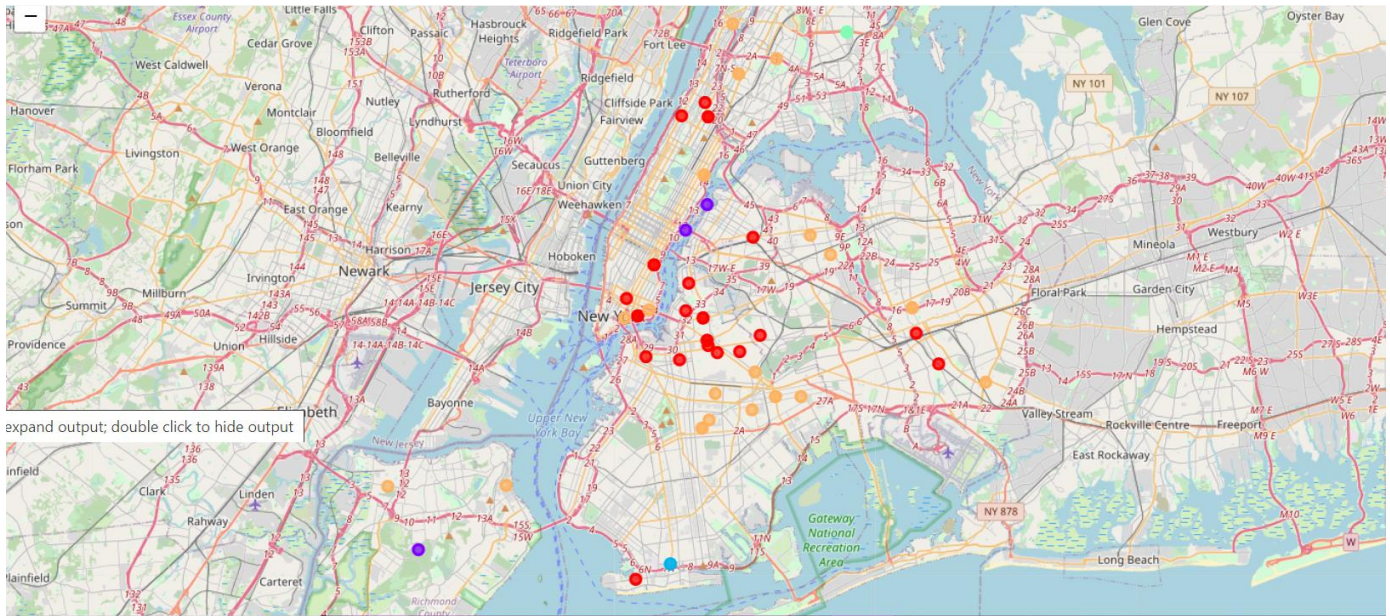


4. Results

Here is my merged table with cluster labels for each borough.

borough	facility_name	facility_type	latitude	longitude	postcode	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	Ida G. Israel Community Health Center	Child Health Center	40.578466	-73.989614	11224.0	0	Liquor Store	Tennis Court	Boutique	Gas Station	Coffee Shop	Fast Food Restaurant	River	Cosmetics Shop	Pharmacy	Intersection
1	Queens	South Queens Community Health Center	Child Health Center	40.686515	-73.785593	11434.0	0	Wine Shop	Electronics Store	Bus Station	Bus Stop	BBQ Joint	Lawyer	Southern / Soul Food Restaurant	Hotel	Fast Food Restaurant	Food Truck
5	Queens	Parsons Communicare Clinic	Child Health Center	40.703826	-73.800483	11432.0	0	Clothing Store	Mexican Restaurant	Mobile Phone Shop	Bank	Coffee Shop	Pizza Place	Southern / Soul Food Restaurant	Supplement Shop	Concert Hall	Sri Lankan Restaurant
6	Queens	Queens Hospital Center	Child Health Center	40.717250	-73.803433	11432.0	4	Pizza Place	Ice Cream Shop	Bakery	Fast Food Restaurant	Convenience Store	Cosmetics Shop	Sushi Restaurant	Gift Shop	Diner	Laundromat
9	Queens	Junction Boulevard Child Health Clinic	Child Health Center	40.754179	-73.872231	11372.0	4	Fast Food Restaurant	South American Restaurant	Bank	Café	Gas Station	Pizza Place	Yoga Studio	Mattress Store	Chinese Restaurant	Bus Station
10	Brooklyn	Woodhull Medical and Mental Health Center	Child Health Center	40.700528	-73.941604	11206.0	0	Pizza Place	Mexican Restaurant	Bakery	Juice Bar	Spanish Restaurant	Farm	Music Venue	Coffee Shop	Bookstore	Bagel Shop
11	Brooklyn	Eleanor Roosevelt Houses Child Health Clinic	Child Health Center	40.694049	-73.934552	11206.0	0	Bar	Pizza Place	Café	Del / Bodega	Chinese Restaurant	Asian Restaurant	Music Venue	Bakery	Dive Bar	Cocktail Bar
12	Bronx	Jacobi Medical Center	Acute Care Hospital	40.857427	-73.847079	10461.0	3	Bakery	Gym	Bar	Farm	Food Truck	Food Court	Food & Drink Shop	Flower Shop	Flea Market	Fish Market
13	Manhattan	Metropolitan Hospital Center	Acute Care Hospital	40.784557	-73.943780	10029.0	4	Pizza Place	Coffee Shop	Pharmacy	Bank	Fast Food Restaurant	Gym	Thai Restaurant	Skating Rink	Sandwich Place	Spa
14	Brooklyn	Bushwick Communicare	Child Health Center	40.694674	-73.919153	11221.0	0	Café	Bar	Latin American Restaurant	Pizza Place	Frozen Yogurt Shop	Fast Food Restaurant	Supermarket	Sandwich Place	Cocktail Bar	Coffee Shop
15	Brooklyn	Lafayette Houses Child Health Clinic	Child Health Center	40.690361	-73.960108	11205.0	0	Pizza Place	Bar	Wine Shop	Coffee Shop	Garden	Yoga Studio	Diner	Cocktail Bar	Café	Arts & Crafts Store

This is a folium map showing the locations in 5 different color mark each for one cluster.



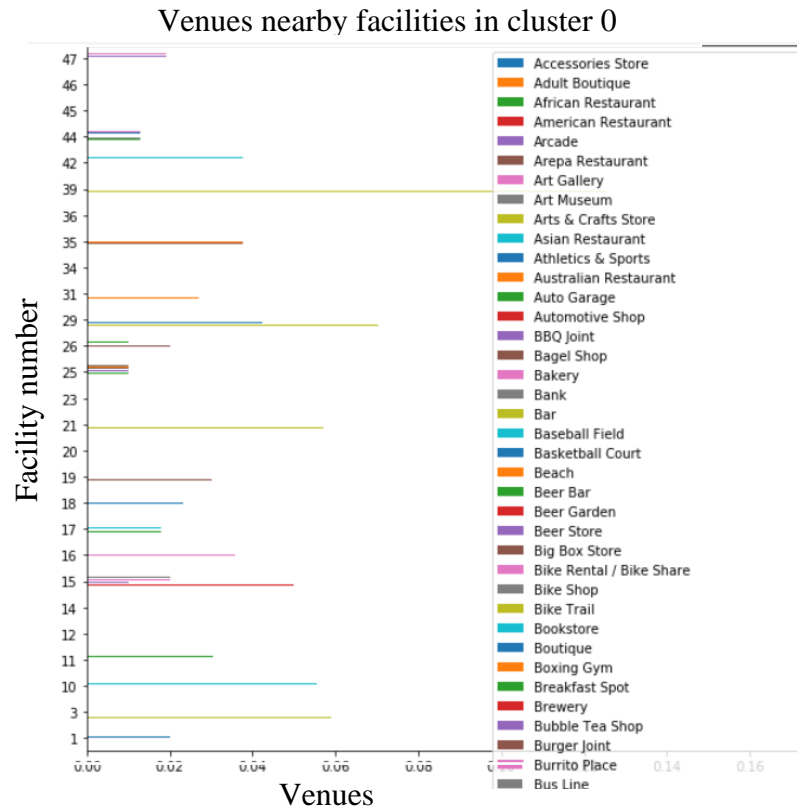
4.1. clusters tables and bar charts

Cluster 0:

Cluster 0 analysis discussion:

no pharmacies but has shopping options and cafes around the areas of the healthcare facilities, but no competitors around. It needs location analysis and feasibility study.

	facility_name	postcode	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	Ida G. Israel Community Health Center	11224.0	0	Liquor Store	Tennis Court	Boutique	Gas Station	Coffee Shop	Fast Food Restaurant	River	Cosmetics Shop	Pharmacy	Intersection
1	South Queens Community Health Center	11434.0	0	Wine Shop	Electronics Store	Bus Station	Bus Stop	BBQ Joint	Lawyer	Southern / Soul Food Restaurant	Hotel	Fast Food Restaurant	Food Truck
5	Parsons Communicare Clinic	11432.0	0	Clothing Store	Mexican Restaurant	Mobile Phone Shop	Bank	Coffee Shop	Pizza Place	Southern / Soul Food Restaurant	Supplement Shop	Concert Hall	Sri Lankan Restaurant
10	Woodhull Medical and Mental Health Center	11206.0	0	Pizza Place	Mexican Restaurant	Bakery	Juice Bar	Spanish Restaurant	Farm	Music Venue	Coffee Shop	Bookstore	Bagel Shop
11	Eleanor Roosevelt Houses Child Health Clinic	11206.0	0	Bar	Pizza Place	Café	Deli / Bodega	Chinese Restaurant	Asian Restaurant	Music Venue	Bakery	Dive Bar	Cocktail Bar



Cluster 1:

Cluster 1 analysis discussion:

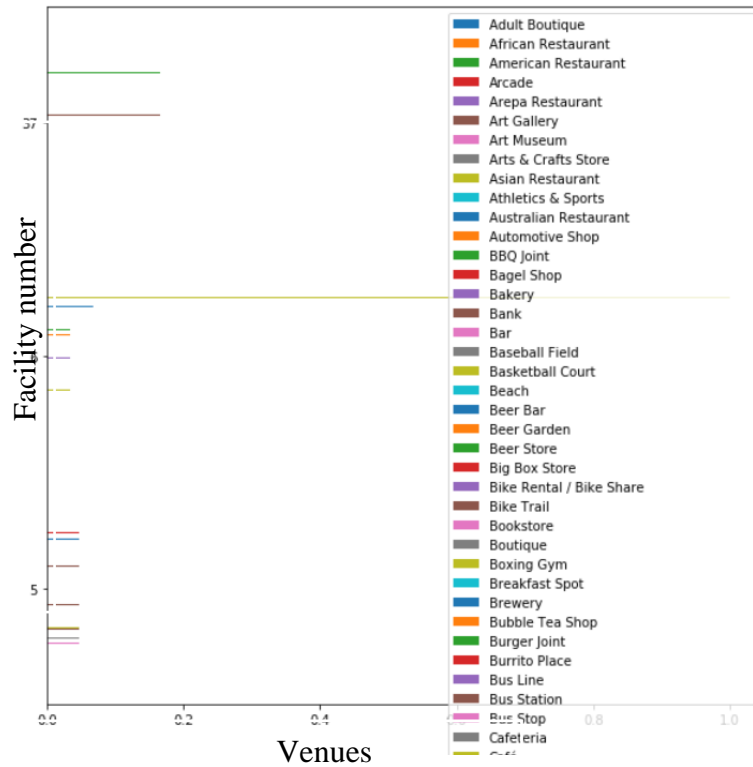
outdoor activities areas - residential areas good option for new community pharmacy location

which basically focus on regular patients. Needs population income study to assure the higher

income patient reaches to the location

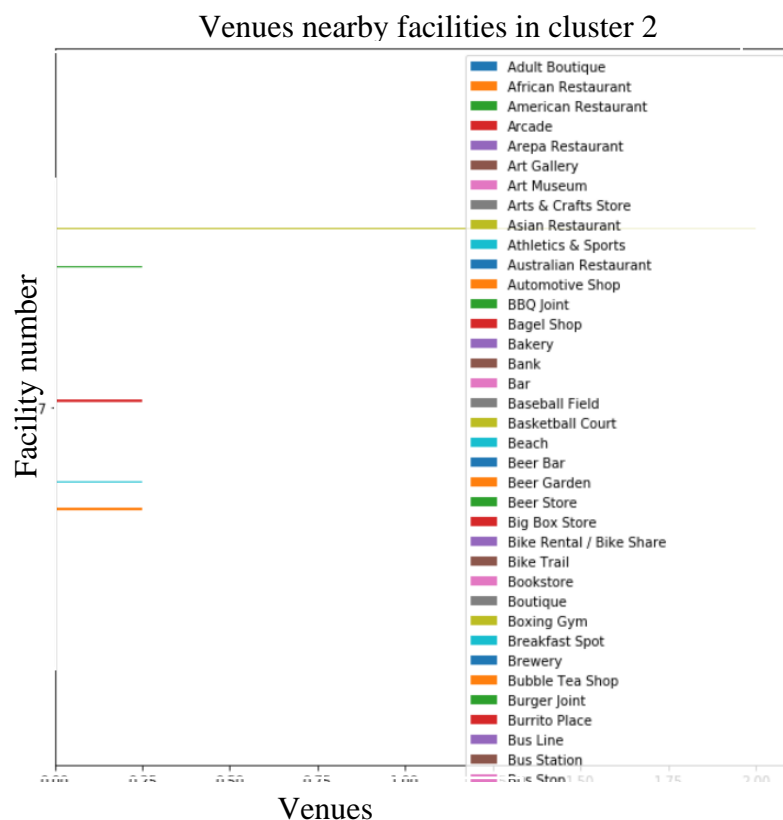
	facility_name	postcode	Cluster Labels	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	4th Most Common Venue	5th Most Common Venue	6th Most Common Venue	7th Most Common Venue	8th Most Common Venue	9th Most Common Venue	10th Most Common Venue
24	Sea View Hospital Rehabilitation Center & Home	10314.0	1	Gym	Historic Site	Liquor Store	Baseball Field	Theater	Moving Target	Discount Store	Dessert Shop	Food Court	Food & Drink Shop
34	Coler-Goldwater Specialty Hospital and Nursing...	10044.0	1	Park	Liquor Store	Food Court	Harbor / Marina	Farmers Market	Baseball Field	Lighthouse	Bus Station	Tennis Court	Garden
58	Coler-Goldwater Specialty Hospital and Nursing...	10044.0	1	Park	Tennis Court	Historic Site	Pharmacy	Scenic Lookout	Bike Trail	Roof Deck	Coffee Shop	Baseball Field	Dog Run

Venues nearby facilities in cluster 1

**Cluster 2:****Cluster 2 analysis discussion:**

expats residential areas - Asian minorities residential areas. not what we are aiming for as we focus on higher income population areas and mixed population more favorable for any business.

	facility_name	postcode	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
43	Coney Island Hospital	11235.0	2	Eastern European Restaurant	Japanese Restaurant	Supermarket	Convenience Store	Yoga Studio	Filipino Restaurant	Farm	Farmers Market	Fast Food Restaurant	Field
65	Coney Island Hospital	11235.0	2	Eastern European Restaurant	Japanese Restaurant	Supermarket	Convenience Store	Yoga Studio	Filipino Restaurant	Farm	Farmers Market	Fast Food Restaurant	Field



Cluster 3:

Cluster 3 analysis discussion:

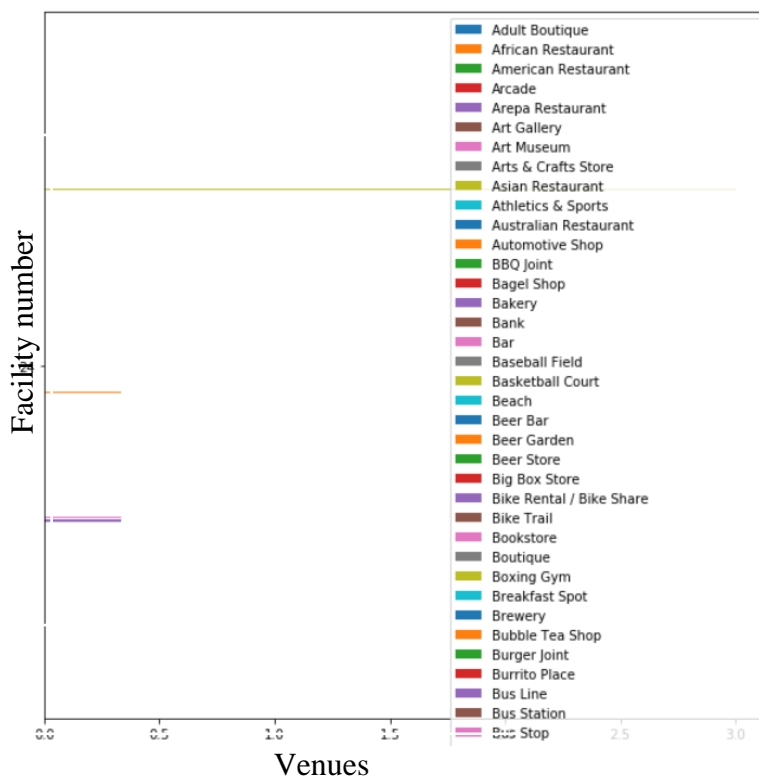
farmer market and supermarket areas - household facilities. classic option for new community

pharmacy as it will assure the foot flow to the store without too much effort to advertise the

location. but no competitors nearby the healthcare facilities

	facility_name	postcode	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
12	Jacobi Medical Center	10461.0	3	Bakery	Bar	Gym	Filipino Restaurant	Farm	Farmers Market	Fast Food Restaurant	Field	Fish & Chips Shop	French Restaurant
30	Jacobi Medical Center	10461.0	3	Bakery	Bar	Gym	Filipino Restaurant	Farm	Farmers Market	Fast Food Restaurant	Field	Fish & Chips Shop	French Restaurant

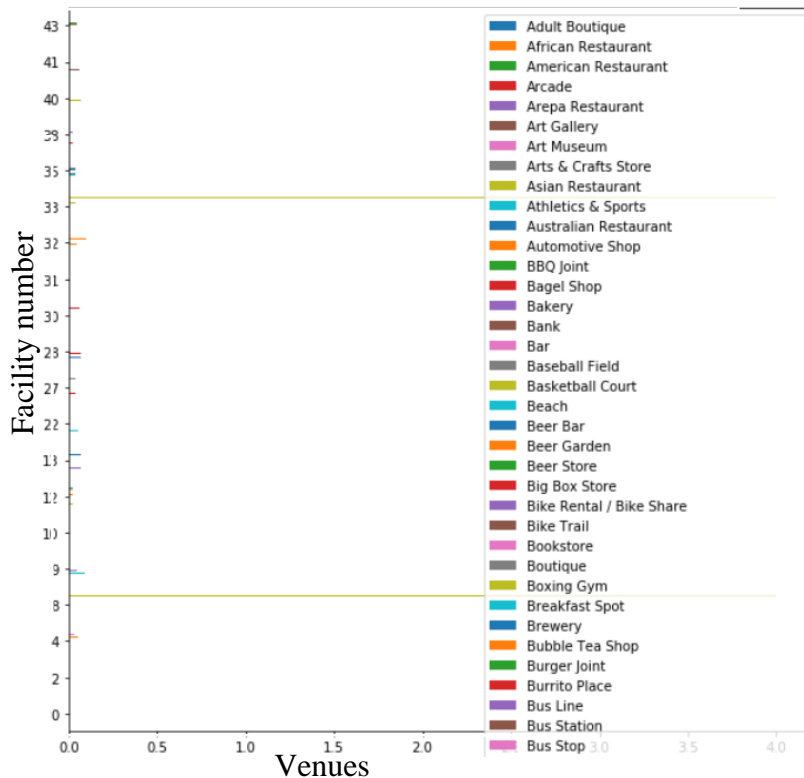
Venues nearby facilities in cluster 3

**Cluster 4:**

Cluster 4 analysis discussion:

it is the ideal areas to open new pharmacy as the are many factors to grantee the foot flow like the household and personal shopping facilities, variety of restaurants and cafes, gym and the most important is the great availability of the competitor pharmacies around the health centers.

Venues nearby facilities in cluster 4



	facility_name	postcode	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
9	Junction Boulevard Child Health Clinic	11372.0	4	Fast Food Restaurant	South American Restaurant	Bank	Pizza Place	Café	Gas Station	Donut Shop	Chinese Restaurant	Salvadoran Restaurant	Sandwich Place
16	East New York Diagnostic & Treatment Center	11207.0	4	Grocery Store	Pizza Place	Mexican Restaurant	Pharmacy	Diner	Convenience Store	Chinese Restaurant	Sandwich Place	Nightclub	Spanish Restaurant
18	North Central Bronx Hospital	10467.0	4	Donut Shop	Pizza Place	Chinese Restaurant	Spanish Restaurant	Pharmacy	Fried Chicken Joint	Latin American Restaurant	Supplement Shop	Supermarket	Caribbean Restaurant
23	Corona Child Health Clinic	11368.0	4	Mexican Restaurant	Supermarket	Pizza Place	Ice Cream Shop	Food Truck	Sandwich Place	Restaurant	Donut Shop	Bakery	Convenience Store
25	Springfield Gardens Medical Center	11413.0	4	Caribbean Restaurant	Pizza Place	Ice Cream Shop	Donut Shop	Fried Chicken Joint	Supermarket	Fast Food Restaurant	New American Restaurant	Sandwich Place	Mobile Phone Shop
27	Morrisania Diagnostic & Treatment Center	10452.0	4	Fried Chicken Joint	Pizza Place	Chinese Restaurant	Fast Food Restaurant	Gym	Pharmacy	Diner	Park	Donut Shop	Sandwich Place
33	Kings County Hospital Center	11203.0	4	Deli / Bodega	Pharmacy	Sandwich Place	Chinese Restaurant	Ice Cream Shop	Fried Chicken Joint	Other Great Outdoors	Donut Shop	Garden Center	Caribbean Restaurant
35	Kings County Hospital Center	11203.0	4	Deli / Bodega	Pharmacy	Sandwich Place	Chinese Restaurant	Ice Cream Shop	Fried Chicken Joint	Other Great Outdoors	Donut Shop	Garden Center	Caribbean Restaurant
41	East New York Diagnostic & Treatment Center	11207.0	4	Grocery Store	Pizza Place	Mexican Restaurant	Pharmacy	Diner	Convenience Store	Chinese Restaurant	Sandwich Place	Nightclub	Spanish Restaurant
42	Brownsville Child Health Clinic	11212.0	4	Chinese Restaurant	Fried Chicken Joint	Pizza Place	Discount Store	Burger Joint	Kids Store	Restaurant	Trail	Grocery Store	Farmers Market
45	Stapleton Child Health Clinic	10304.0	4	Pizza Place	Sandwich Place	Discount Store	Bar	Cosmetics Shop	Bank	Bus Stop	Spanish Restaurant	Restaurant	Residential Building (Apartment / Condo)
48	Mariner's Harbor Houses Child Health Clinic	10303.0	4	Spanish Restaurant	Bar	Pizza Place	Supplement Shop	Latin American Restaurant	Grocery Store	Bagel Shop	Sports Bar	Cosmetics Shop	Discount Store
60	Sutter Avenue Child Health Clinic	11208.0	4	Deli / Bodega	Event Service	Chinese Restaurant	Electronics Store	Bus Station	Spanish Restaurant	Music Venue	Fried Chicken Joint	Metro Station	Fast Food Restaurant
66	Health Center at Tremont	10457.0	4	Donut Shop	Pizza Place	Park	Fast Food Restaurant	Bank	Sandwich Place	Pharmacy	Breakfast Spot	Fried Chicken Joint	Supermarket
67	North Central Bronx Hospital	10467.0	4	Donut Shop	Pizza Place	Chinese Restaurant	Spanish Restaurant	Pharmacy	Fried Chicken Joint	Latin American Restaurant	Supplement Shop	Supermarket	Caribbean Restaurant
68	Crown Heights Child Health Clinic	11213.0	4	Grocery Store	Deli / Bodega	Café	Discount Store	Chinese Restaurant	Southern / Soul Food Restaurant	Fried Chicken Joint	Supermarket	Gas Station	Lounge

5. Discussion:

After clustering the locations based on the nearby venues, we found that cluster 4 has met most of the key points we discussed earlier in the introduction.

Cluster 4 has locations of healthcare facilities which are located in active areas with shopping malls, restaurants, household shopping and gyms. We will diversity of people based on age, culture and most importantly salary income.

Foot flow confirmed.

So, remaining for us to locate the availability of pharmacy especially big chain pharmacies to place our name in between them.

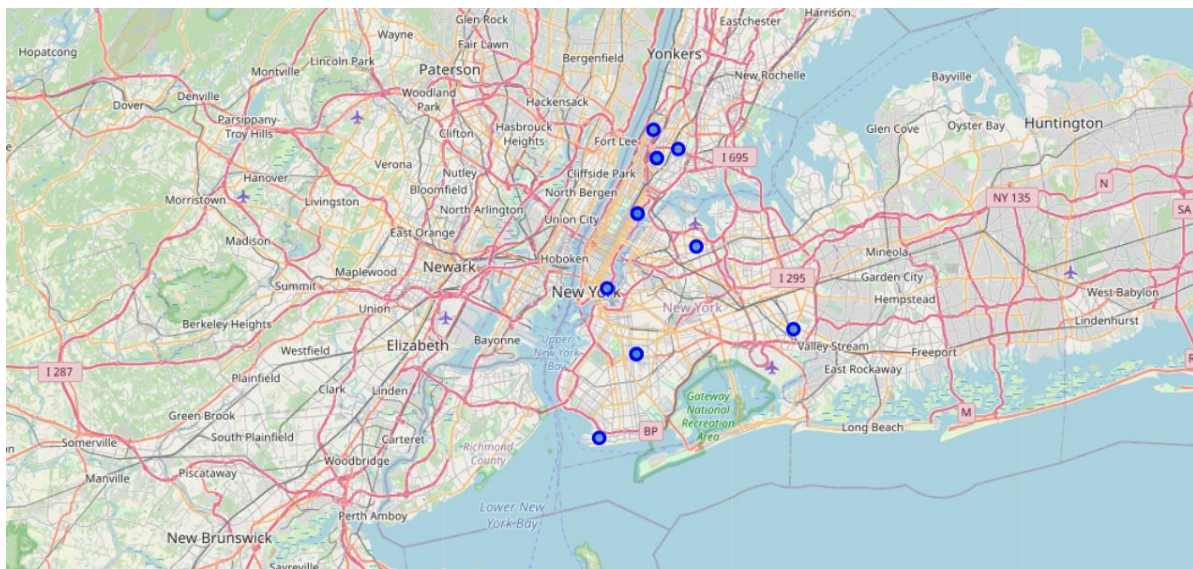
For that I filtered the facilities location in cluster 4 based on the availability of ‘Pharmacy’ venue nearby and what we got 12 locations out of 64 locations of healthcare facilities in NYC.

These locations are located mostly in 4 different boroughs of the city. We got 3 locations in Bronx and Brooklyn, 4 in Manhattan, and 2 in queens.

	borough	facility_name	latitude	longitude		facility_name
0	Brooklyn	Ida G. Israel Community Health Center	40.578468	-73.989614		
9	Queens	Junction Boulevard Child Health Clinic	40.754179	-73.872231		
13	Manhattan	Metropolitan Hospital Center	40.784557	-73.943780		
25	Queens	Springfield Gardens Medical Center	40.678997	-73.754077		
27	Bronx	Morrisania Diagnostic & Treatment Center	40.835957	-73.919986		
33	Brooklyn	Kings County Hospital Center	40.655762	-73.944580		
35	Brooklyn	Kings County Hospital Center	40.655762	-73.944580		
40	Manhattan	Dyckman Clinica De Las Americas	40.861640	-73.924590		
50	Manhattan	Metropolitan Hospital Center	40.784557	-73.943780		
66	Bronx	Health Center at Tremont	40.844083	-73.894482		
69	Manhattan	Baruch Houses Family Health Center	40.716079	-73.980373		
77	Bronx	Morrisania Diagnostic & Treatment Center	40.835957	-73.919986		

borough	facility_name
Bronx	3
Brooklyn	3
Manhattan	4
Queens	2

Finally located the 12 location on my NYC folium map with markers and we are ready to do feasibility study for each location based on many other different factors like tenant location and price.



6. Conclusion:

As a result, investors whom are willing to start a community pharmacy business can achieve better outcomes through their access to platforms where such information are provided. No matter how hard to find the proper data to study and to analyze it is a crucial step to find the best location for your successful future business.

7. References:

1. "World Urban Areas" (PDF). Demographia. 2018. Retrieved March 27,2018.. https://en.wikipedia.org/wiki/New_York_City#cite_note-12
2. Lester Nathan, MS in Pharmacy times journal-2015-09-14 <https://www.pharmacytimes.com/contributor/lester-nathan-ms/2015/09/how-to-pick-the-best-pharmacy-location>
3. NYC Health + Hospitals patient care locations -updated July 3, 2019 <https://data.cityofnewyork.us/Health/NYC-Health-Hospitals-patient-care-locations-2011/f7b6-v6v3>