Desired Venues in Manhattan Neighborhoods

1. Introduction

In this example, there is a client who want's to buy an apartment in Manhattan, but needs specific venues in the vicinity of the apartment, so he hires me to find him a neighborhood that has all the ideal conditions. The task is to find all three required venues

["Supermarket", "Gym", "Pharmacy"]

located in a radius of 500 meters.

2. Data

1. Manhattan neighborhood location dataset

2. Manhattan Venues location dataset

I'm going to need all neighborhood locations in Manhattan. I will use Foursquare API to get all venues' location in each neighborhood. Then explore & modify the data to find out which neighborhoods has all three desired venues.
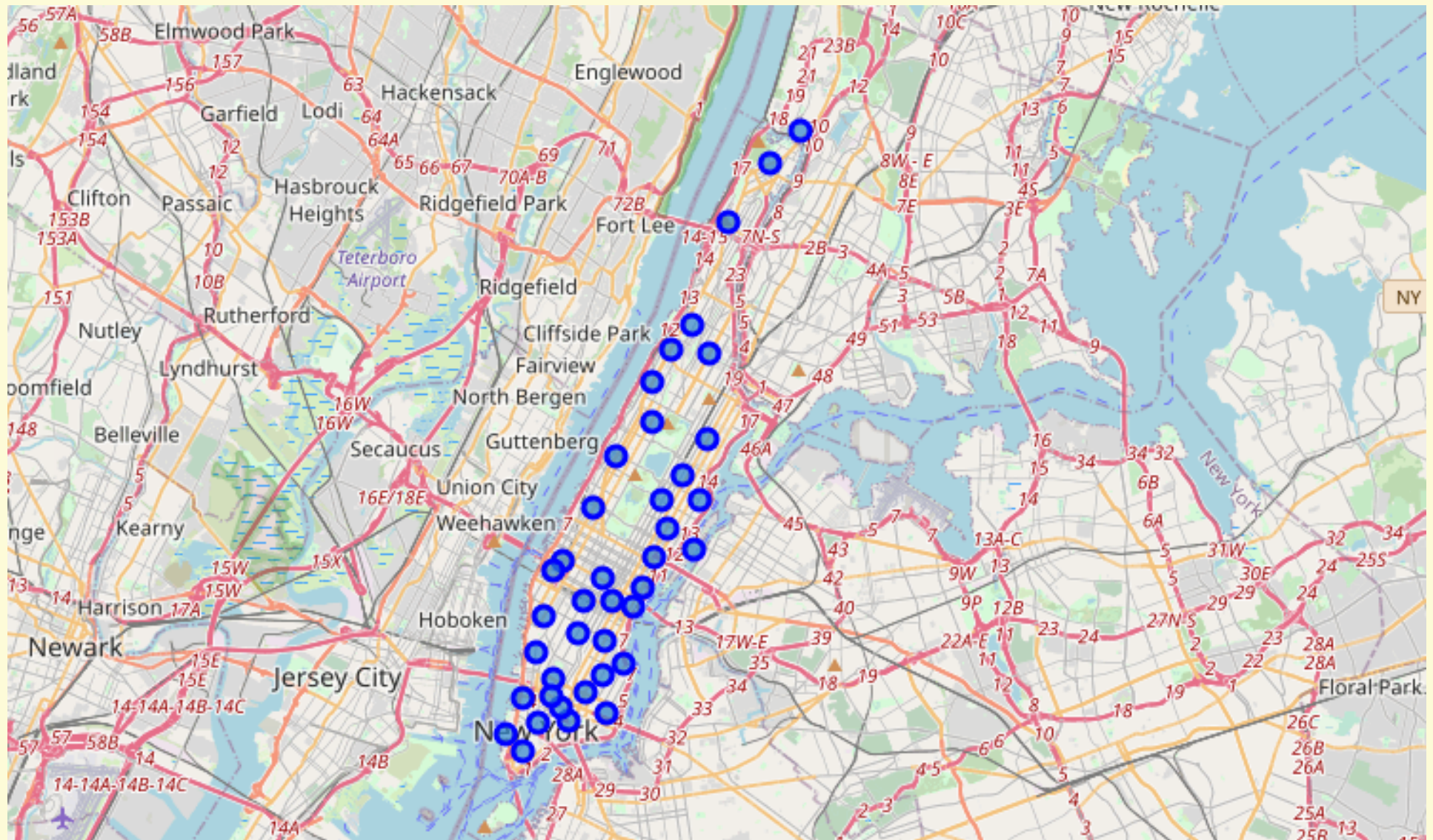
3. Methodology

1. First load the New York Manhattan neighborhood location dataset. And map out where there are located using Python library "folium" (map rendering library).

	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
...	...	...	...	...
35	Manhattan	Turtle Bay	40.752042	-73.967708
36	Manhattan	Tudor City	40.746917	-73.971219
37	Manhattan	Stuyvesant Town	40.731000	-73.974052
38	Manhattan	Flatiron	40.739673	-73.990947
39	Manhattan	Hudson Yards	40.756658	-74.000111

40 rows × 4 columns

3. Methodology

2. Using Python module "Geopy" to get Geographical coordinates of Manhattan to utilized the Foursquare API to segment neighborhoods and get top 100 venues in each neighborhood in Manhattan. each neighborhood in Manhattan. And converted to a DataFrame.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Marble Hill	40.876551	-73.910660	Arturo's	40.874412	-73.910271	Pizza Place
1	Marble Hill	40.876551	-73.910660	Bikram Yoga	40.876844	-73.906204	Yoga Studio
2	Marble Hill	40.876551	-73.910660	Tibbett Diner	40.880404	-73.908937	Diner
3	Marble Hill	40.876551	-73.910660	Starbucks	40.877531	-73.905582	Coffee Shop
4	Marble Hill	40.876551	-73.910660	Dunkin'	40.877136	-73.906666	Donut Shop
...	...	...	...	...	...	...	...
3146	Hudson Yards	40.756658	-74.000111	Cachet Boutique Hotel	40.759773	-73.996460	Hotel
3147	Hudson Yards	40.756658	-74.000111	StarDust	40.759869	-73.996460	Nightclub
3148	Hudson Yards	40.756658	-74.000111	Jake's	40.757954	-74.002296	American Restaurant
3149	Hudson Yards	40.756658	-74.000111	Gray Line New York Sightseeing Cruises - Pier 78	40.759721	-74.003982	Harbor / Marina
3150	Hudson Yards	40.756658	-74.000111	Twilight Cruise By Citysightseeing	40.759744	-74.004096	Boat or Ferry

3151 rows × 7 columns

3. Methodology

3. Methodology

3. The Client only wants three specific categories of venues, so I'll extract rows with the required categories.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Marble Hill	40.876551	-73.910660	Rite Aid	40.875467	-73.908906	Pharmacy
1	Marble Hill	40.876551	-73.910660	Astral Fitness & Wellness Center	40.876705	-73.906372	Gym
2	Marble Hill	40.876551	-73.910660	Blink Fitness	40.877271	-73.905595	Gym
3	Chinatown	40.715618	-73.994279	Hong Kong Supermarket 香港超級市場	40.717596	-73.996173	Supermarket
4	Chinatown	40.715618	-73.994279	Stanley's Pharmacy	40.715782	-73.990544	Pharmacy
...	...	...	...	...	...	...	...
88	Flatiron	40.739673	-73.990947	Equinox Gramercy	40.740749	-73.985771	Gym
89	Flatiron	40.739673	-73.990947	Rowgatta	40.736900	-73.995094	Gym
90	Hudson Yards	40.756658	-74.000111	Brooklyn Fare	40.756130	-73.996614	Supermarket
91	Hudson Yards	40.756658	-74.000111	505W37 Gym	40.757275	-73.997797	Gym
92	Hudson Yards	40.756658	-74.000111	Club 7 By Iowa Sports	40.760827	-73.999120	Gym

93 rows × 7 columns

3. Methodology

4. I've hot encoded the Venue Category and drop all other columns except neighborhood. Grouped the dataset with neighborhoods, it shows the sum of each venue category.

And final, filter out the neighborhoods that contain all 3 required venues within its 500m radius vicinity.

|	Borough	Neighborhood	Latitude	Longitude	Gym	Pharmacy	Supermarket
0	Manhattan	Washington Heights	40.851903	-73.936900	2	1	2
1	Manhattan	Chelsea	40.744035	-74.003116	1	1	1

4. Result

This is the dataset that the client was looking for, and now he knows where to search for his wished-for apartment!

	Borough	Neighborhood	Latitude	Longitude	Gym	Pharmacy	Supermarket
0	Manhattan	Washington Heights	40.851903	-73.936900	2	1	2
1	Manhattan	Chelsea	40.744035	-74.003116	1	1	1

4. Result

- Lastly, I wanted to map out where these two neighborhoods are located, so the client gets an idea where he buys the apartment.
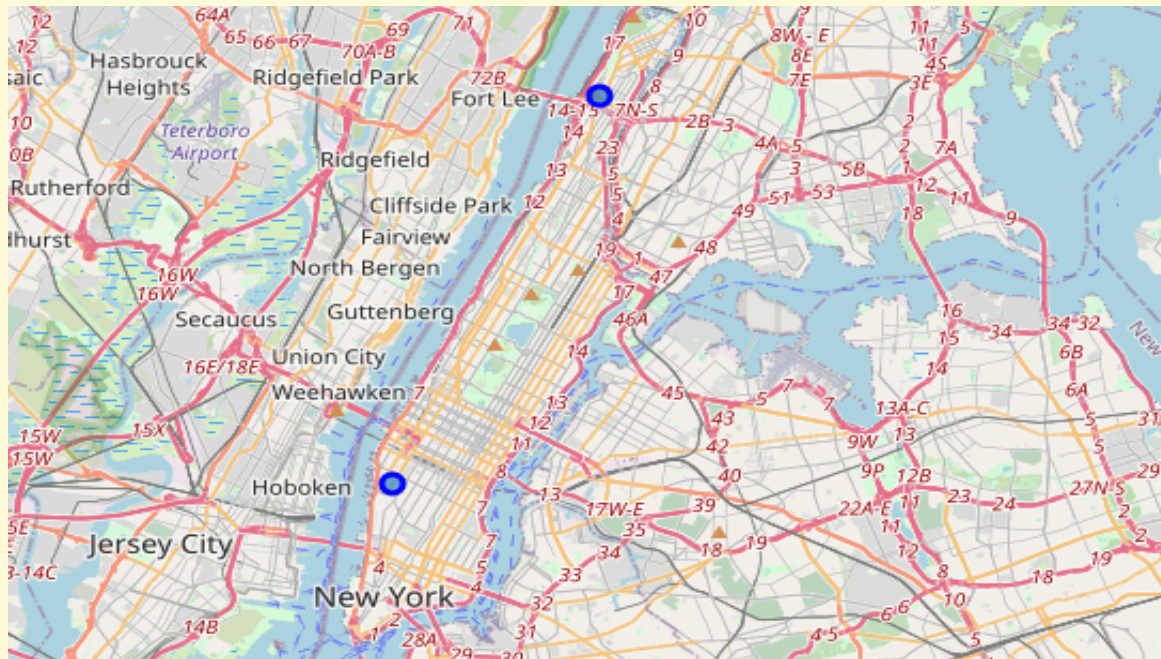
5. Discussion

This is rather an easy example to find a result. Where a client wants to buy an apartment in a vicinity of desired venues.

6. Conclusion

This was one of many ideas about problems that can be solved using location data in addition to other datasets.

There are many other practicable approaches to solving problems that can be solved by using location data in addition to other datasets.

6. Conclusion

For example, if you want to establish a restaurant, and you are looking for the most promising area, getting population-density, age-structure, education level, gender-structure, purchasing-power in addition to location data, you can determine where the best place is for your price level and culinary delicacies you wish to offer.