

Recommendations where to Stay in Manhattan

Introduction

Manhattan is a modern popular city where more and more people traveling here, some of them would be stay here for a short period of days or a few weeks, but some of them decide to stay for a longer time of months or even years. When you are new to a city, you have no information about the neighbourhood and facilities you have around you, here is what we want to be able to help.

Business Problem

It is overwhelming when you move to a new city and you have no one to rely on, we are here to provide the best solutions and recommendations which area to stay. Simply let us know your needs and your daily life routine, we are able to help you shortlist the best locations to stay and save you the trouble of shifting homes

Data Source

The data we have on Manhattan city is where people go all the time from Foursquare. We have the density of every kind of facilities for every neighbourhood. Input from customer is needed to filter the data and plot the neighbourhood and thus to plot all the needed facilities.

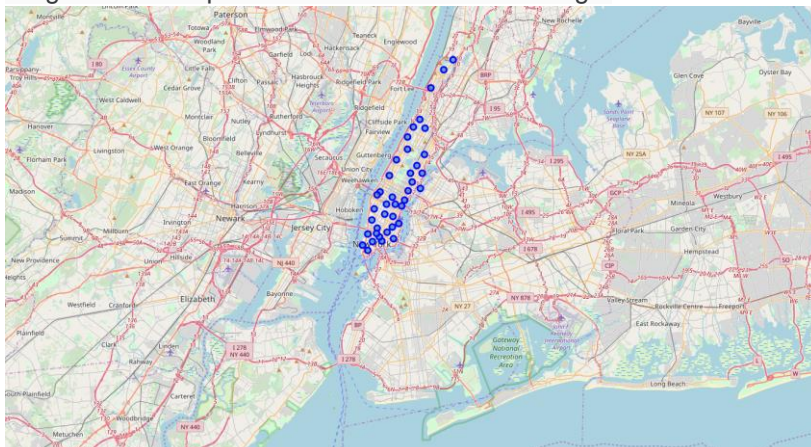
For this customer, the requirement is near to **Gym and nearby there is Chinese restaurant**.

Geolocation data of longitude and latitude already introduced at https://geo.nyu.edu/catalog/nyu_2451_34572 . Load the raw data and convert into data frame.

Details of Manhattan neighbourhood locations coordinates is included in the data frame by filtering "Borough".

	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

Neighborhood is plotted under Manhattan Borough.



Neighborhood venue data retrieved from query to Foursquare, and its geocodes

	name	categories	lat	lng
0	Arturo's	Pizza Place	40.874412	-73.910271
1	Bikram Yoga	Yoga Studio	40.876844	-73.906204
2	Tibbett Diner	Diner	40.880404	-73.908937
3	Dunkin'	Donut Shop	40.877136	-73.906666
4	Starbucks	Coffee Shop	40.877531	-73.905582

All the data is combined for all the neighbourhood under Manhattan and stored in dataframe

Neighborhood	Accessories Store	Adult Boutique	Afghan Restaurant	African Restaurant	American Restaurant	Animal Shelter	Antique Shop	Arcade	Arepa Restaurant	Argentinian Restaurant	Art Gallery	Art Museum	Arts & Crafts Store	Rest
0	Marble Hill	0	0	0	0	0	0	0	0	0	0	0	0	0
1	Marble Hill	0	0	0	0	0	0	0	0	0	0	0	0	0
2	Marble Hill	0	0	0	0	0	0	0	0	0	0	0	0	0
3	Marble Hill	0	0	0	0	0	0	0	0	0	0	0	0	0
4	Marble Hill	0	0	0	0	0	0	0	0	0	0	0	0	0

Based on this special request of near Chinese restaurants and Gym, venue information is filtered based on the requirement

Out[366]:

Chinese Restaurant	Boxing Gym	Climbing Gym	College Gym	Gym	Gym / Fitness Center	Gym Pool	Gymnastics Gym
0	0.010000	0.000000	0.00	0.040000	0.000000	0.00	0.00
1	0.010000	0.000000	0.01	0.030000	0.020000	0.00	0.00
2	0.045455	0.000000	0.00	0.022727	0.045455	0.00	0.00
3	0.010000	0.000000	0.00	0.010000	0.010000	0.00	0.00
4	0.100000	0.000000	0.00	0.010000	0.000000	0.00	0.00
5	0.000000	0.000000	0.00	0.020000	0.050000	0.00	0.00
6	0.010000	0.000000	0.00	0.020000	0.050000	0.00	0.00
7	0.000000	0.000000	0.00	0.023256	0.000000	0.00	0.00
8	0.000000	0.000000	0.00	0.000000	0.000000	0.00	0.00
9	0.040000	0.000000	0.00	0.040000	0.030000	0.00	0.00
10	0.010000	0.000000	0.00	0.050000	0.040000	0.00	0.00
11	0.010000	0.000000	0.00	0.000000	0.010000	0.00	0.00
12	0.000000	0.000000	0.00	0.020000	0.000000	0.00	0.00
13	0.020000	0.000000	0.00	0.000000	0.000000	0.00	0.00
14	0.034483	0.000000	0.00	0.027027	0.040541	0.00	0.00
15	0.000000	0.000000	0.00	0.000000	0.017241	0.00	0.00
16	0.034483	0.000000	0.00	0.030000	0.030000	0.00	0.00
17	0.010000	0.010000	0.00	0.020000	0.060000	0.00	0.00
18	0.000000	0.000000	0.00	0.000000	0.000000	0.00	0.00
19	0.000000	0.000000	0.00	0.000000	0.000000	0.00	0.00
20	0.020000	0.000000	0.00	0.000000	0.016949	0.00	0.00
21	0.052632	0.026316	0.00	0.000000	0.000000	0.00	0.00
22	0.016949	0.000000	0.00	0.041667	0.000000	0.00	0.00
23	0.026316	0.000000	0.00	0.020000	0.000000	0.00	0.00
24	0.000000	0.000000	0.00	0.010000	0.030000	0.00	0.00
25	0.010000	0.000000	0.00	0.000000	0.000000	0.00	0.00
26	0.000000	0.000000	0.00	0.030000	0.020000	0.00	0.00
27	0.020000	0.000000	0.00	0.010000	0.010000	0.00	0.00
28	0.000000	0.000000	0.00	0.037037	0.000000	0.00	0.00

Data then its combined into one data frame to create clustering map

	Borough	Neighborhood	Latitude	Longitude	Cluster Labels	Chinese Restaurant	Gym Total
0	Manhattan	Marble Hill	40.876551	-73.910660	7	0.000000	0.041667
1	Manhattan	Chinatown	40.715618	-73.994279	1	0.100000	0.010000
2	Manhattan	Washington Heights	40.851903	-73.936900	2	0.023810	0.035714
3	Manhattan	Inwood	40.867684	-73.921210	2	0.034483	0.017241
4	Manhattan	Hamilton Heights	40.823604	-73.949688	2	0.034483	0.000000
5	Manhattan	Manhattanville	40.816934	-73.957385	4	0.026316	0.026316
6	Manhattan	Central Harlem	40.815976	-73.943211	1	0.045455	0.068182
7	Manhattan	East Harlem	40.792249	-73.944182	6	0.000000	0.023256
8	Manhattan	Upper East Side	40.775639	-73.960508	1	0.010000	0.040000
9	Manhattan	Yorkville	40.775930	-73.947118	4	0.010000	0.080000
10	Manhattan	Lenox Hill	40.768113	-73.958860	4	0.010000	0.060000
11	Manhattan	Roosevelt Island	40.762160	-73.949168	8	0.000000	0.037037
12	Manhattan	Upper West Side	40.787658	-73.977059	4	0.010000	0.020000
13	Manhattan	Lincoln Square	40.773529	-73.985338	1	0.000000	0.090000
14	Manhattan	Clinton	40.759101	-73.996119	1	0.010000	0.070000
15	Manhattan	Midtown	40.754691	-73.981669	1	0.010000	0.030000
16	Manhattan	Murray Hill	40.748303	-73.978332	1	0.020000	0.060000
17	Manhattan	Chelsea	40.744035	-74.003116	1	0.010000	0.020000
18	Manhattan	Greenwich Village	40.726933	-73.999914	1	0.020000	0.020000
19	Manhattan	East Village	40.727847	-73.982226	4	0.040000	0.000000
20	Manhattan	Lower East Side	40.717807	-73.980890	1	0.052632	0.000000

Clustering is plotted.

Result



Clustering map is plotted and based on the requirement, Roosevelt Island is recommended for this case.

Discussion

Setting of different Cluster K values will produce different results, when the k value is too small , it is hard to set for a single or a few options. If the k values is too large, the output may not be accurate. Testing a few different K values based on the requiremntn is recommended.

Conclusion

This case study illustatest the selection process for given requirement based on the data from FOURSQUARE. More case studies and recommendations can be made customized to different needs.