The Battle of Neighborhoods

1. Introduction: Business Problem

Background: According the latest NYC&Company release New York City welcomed about 65.2 million tourists in 2018 year — 51.6 million domestic and 13.5 million international visitors. And these numbers are continuously increasing from year to year — https://en.wikipedia.org/wiki/Tourism_in_New_York_City. New York City has the largest selection of lodging choices in the country — from the hostels to the luxury hotels. The prices vary from 100$ till several thousand dollars with average price 292 USD per night. The Hotel Occupancy rate is also high — in 2018 year it was 88% — https://assets.simpleviewinc.com/simpleview/image/upload/v1/clients/newyorkcity/FYI_Hotel_reports_February_2019_8607015b-b32a-4c7f-9fbd-84cd2a93cbe6.pdf Visitors prefer short stays that are often over weekends — averaging 2.4 nights — https://aka.nyc/content/uploads/2017/12/new_york_city_travel_and_tourism_trend_report_2017.pdf

Problem description In New York City there are almost 300 hotels with over 75,000 hotel rooms and Airbnb has more than 50,000 apartment listings in New York City in 2018 year — it can be hard to find the right fit or know how much you will get with your money. In this project we will try to find the most optimal borough in Manhattan where a tourist can

rent an accommodation via Airbnb service and have a pleasant stay in NYC and a possibility to attend the most visited attractions like Central Park, Times Square and so on.

Target Audience This investigation would interest New York City's visitors who prefers short stays (from 1 night) and wants to select the best neighborhoods in Manhattan, New York.

Success Criteria The success criteria of this project will be a recommendation with the set of apartments clusters have the best score calculated based on

Accommodation price with fees; Location of the accommodation; Venues in radius of 1000 meters from the accommodation; Crime rate in radius of 100 meters from the accommodation.
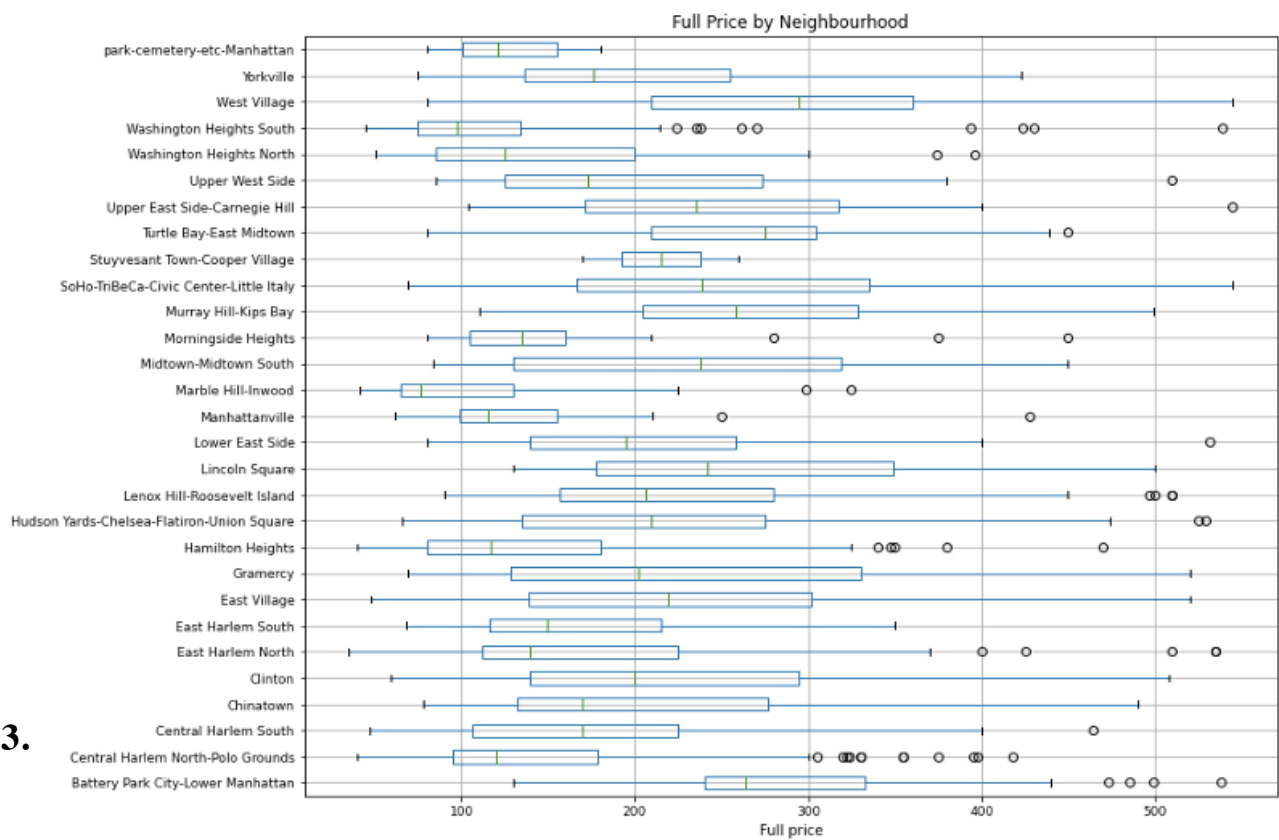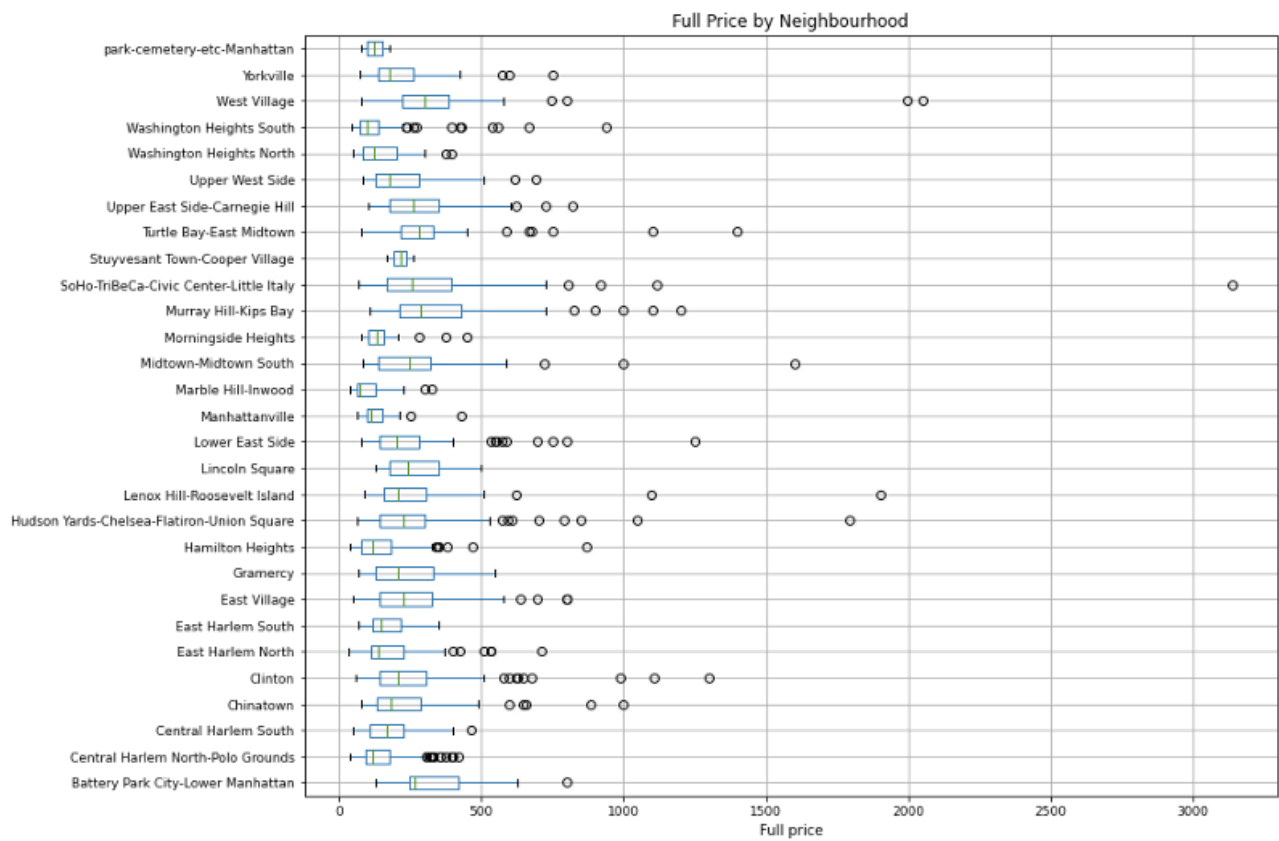
2. Data

In our investigation we will use the free and public available datasets. We will try to evaluate available Airbnb 2019-year accommodations on Manhattan, New York and define the most reasonable apartments sets(clusters) for the visitors.

Based on definition of our problem, we suppose that factors that will help us are: accommodation prices average by neighborhood; number of tourist attractions near the accommodation; number of crimes nearby the accommodation. Following data sources is needed for our project:

New York City apartment listing from the Inside Airbnb site — http://data.insideairbnb.com/united-states/ny/new-york-city/2019-12-04/data/listings.csv.gz New York Neighborhoods Tabulation Areas — https://data.cityofnewyork.us/api/geospatial/cpf4-rkhq?method=export&format=GeoJSON Foursquare API to extract data about venues — food places, museums, galleries, shopping centers, sightseeing attractions, concert halls and so

on New York Crime data for 2019 year — https://data.cityofnewyork.us/api/views/5uac-w243/rows.csv?accessType=DOWNLOAD

Full Price by Neighbourhood

Full Price by Neighbourhood

3.

Methodology

In this project we are trying to detect Manhattan's Neighborhoods that have accommodations for rent with positive reviews, reasonable prices, low number of crimes and tourists' attractions nearby.
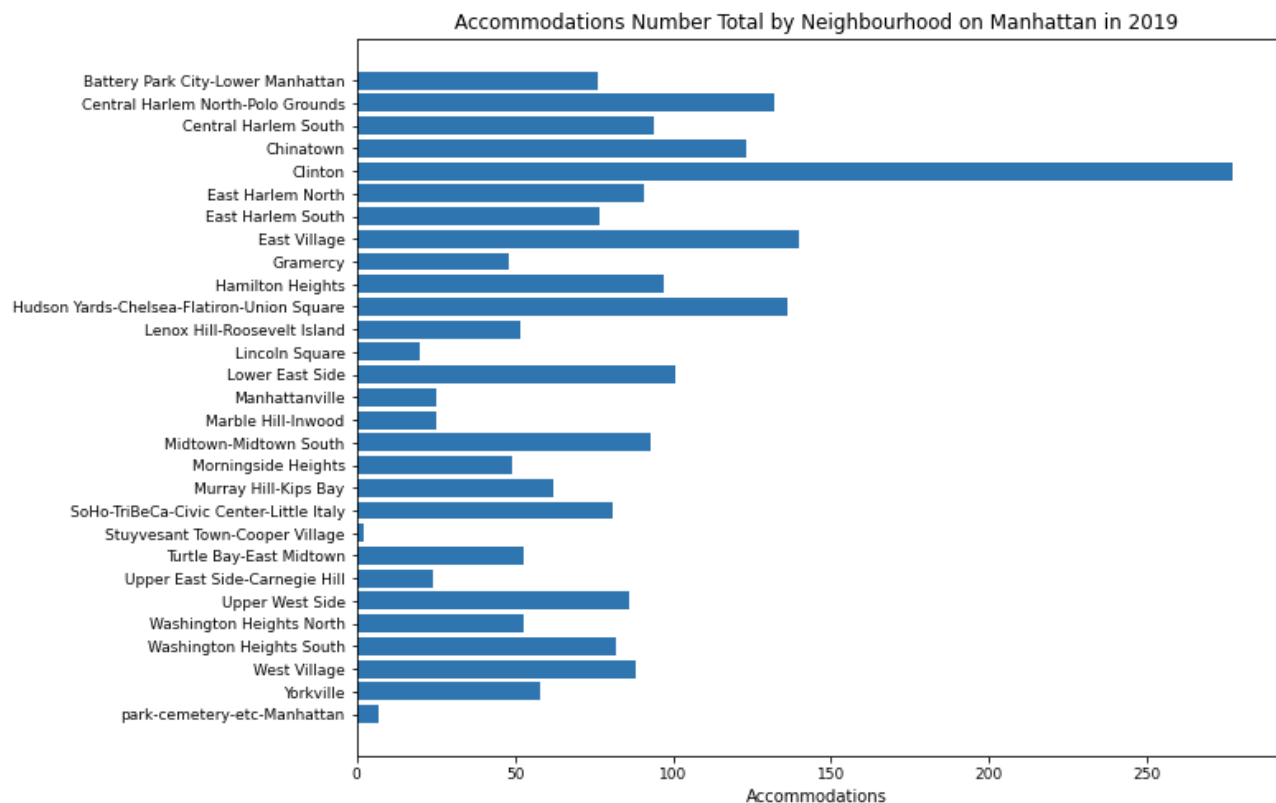
In the first step we have collected the following data:

Airbnb Accommodations with their NYC Tabulation Area (official neighborhood names); Airbnb Accommodation's number of crimes nearby; Defined NYC Tabulation Area (official neighborhood name) for each Manhattan's crime case. The second step in our analysis will be a calculation and exploration different neighborhoods of Manhattan. We will explore the following characteristics:

number of crimes in the area; average price per person; number of accommodations available. In third and final step we will
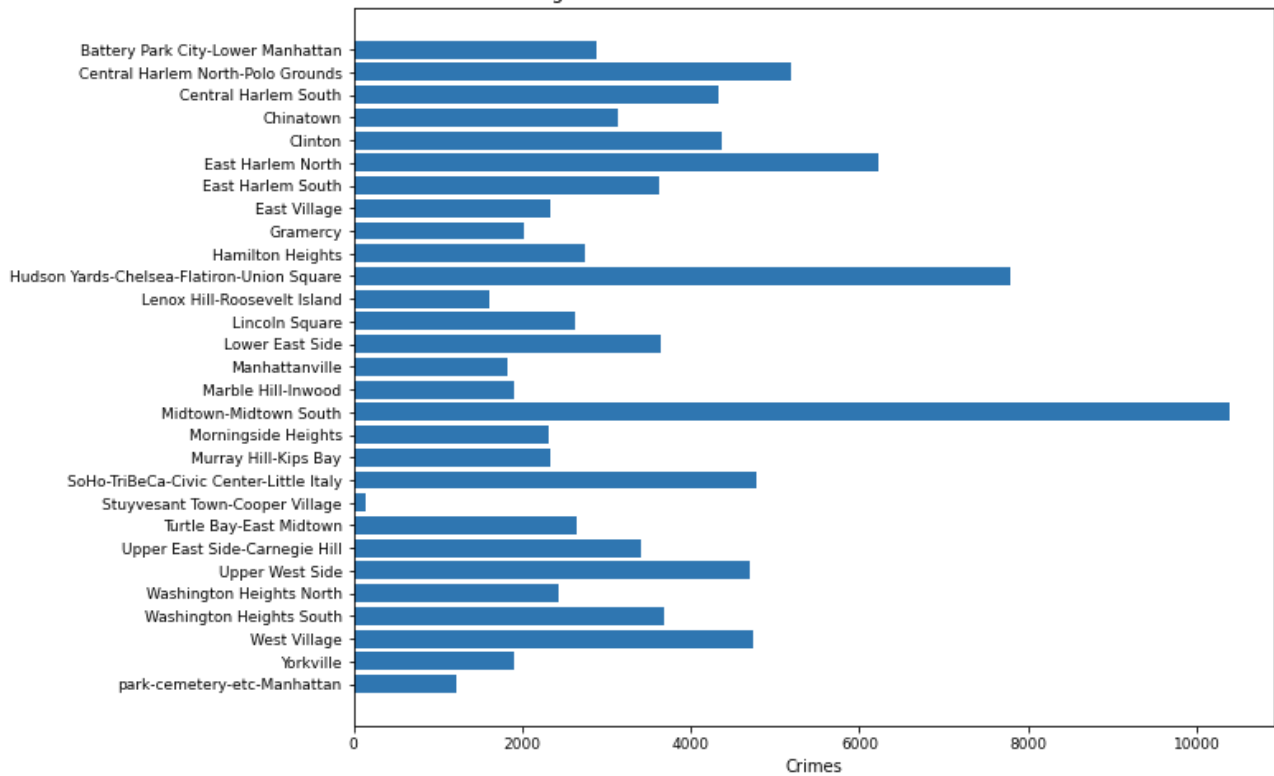
select Top-100 Airbnb accommodations based on summary rating, number of crimes and price per person, and invoke Foursquare API to find Top accommodations' nearby venues create and investigate clusters (using k-means clustering) for our accommodations to make some recommendations to our tourists.

	tab_area	mean_price_per_person	accommodates	mean_crimes
15	Marble Hill-Inwood	45.475238	25	54.440000
25	Washington Heights South	46.793172	82	57.682927
24	Washington Heights North	54.738050	53	50.886792
1	Central Harlem North-Polo Grounds	57.038497	132	73.265152
14	Manhattanville	59.746667	25	70.520000
9	Hamilton Heights	60.069404	97	66.206186
28	park-cemetery-etc-Manhattan	61.404762	7	13.428571
5	East Harlem North	62.275497	91	104.824176
6	East Harlem South	64.950000	77	90.051948
2	Central Harlem South	67.000823	94	87.148936
17	Morningside Heights	71.386054	49	56.204082
23	Upper West Side	78.127907	86	57.813953
16	Midtown-Midtown South	81.491705	93	139.000000
0	Battery Park City-Lower Manhattan	82.133130	76	90.394737
3	Chinatown	82.362703	123	93.081301
4	Clinton	84.308819	277	84.314079
7	East Village	84.842758	140	74.385714
13	Lower East Side	88.452617	101	72.415842
10	Hudson Yards-Chelsea-Flatiron-Union Square	88.591179	136	77.816176
27	Yorkville	89.692529	58	65.000000
18	Murray Hill-Kips Bay	90.410167	62	48.016129
8	Gramercy	94.364931	48	84.791667
21	Turtle Bay-East Midtown	94.817296	53	60.509434
11	Lenox Hill-Roosevelt Island	96.636218	52	45.019231

Accommodations Number Total by Neighbourhood on Manhattan in 2019

	tab_area	crimes_summary
20	Stuyvesant Town-Cooper Village	145
28	park-cemetery-etc-Manhattan	1213
11	Lenox Hill-Roosevelt Island	1604
14	Manhattanville	1832
27	Yorkville	1898
15	Marble Hill-Inwood	1900
8	Gramercy	2021
17	Morningside Heights	2321
18	Murray Hill-Kips Bay	2333
7	East Village	2340
24	Washington Heights North	2440
12	Lincoln Square	2636
21	Turtle Bay-East Midtown	2643
9	Hamilton Heights	2738
0	Battery Park City-Lower Manhattan	2883
3	Chinatown	3137
22	Upper East Side-Carnegie Hill	3403
6	East Harlem South	3629
13	Lower East Side	3650
25	Washington Heights South	3682
2	Central Harlem South	4320
4	Clinton	4364
23	Upper West Side	4700
26	West Village	4748
19	SoHo-TriBeCa-Civic Center-Little Italy	4789
1	Central Harlem North-Polo Grounds	5186

Neighborhoods Total Crimes Records on Manhattan in 2019

	tab_area	crimes_summary	Latitude	Longitude
0	Battery Park City-Lower Manhattan	2883	40.706784	-74.010147
1	Central Harlem North-Polo Grounds	5186	40.819400	-73.941200
2	Central Harlem South	4320	40.802900	-73.952300
3	Chinatown	3137	40.715100	-73.995500
4	Clinton	4364	40.764423	-73.992392

	name	latitude	longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Private Bedroom in Cozy Hamilton Heights Apartment	40.82749	-73.94461	Branson Got Juice!	40.825525	-73.943681	Juice Bar
1	Private Bedroom in Cozy Hamilton Heights Apartment	40.82749	-73.94461	Camila's Tasting Room	40.827680	-73.946073	Bar
2	Private Bedroom in Cozy Hamilton Heights Apartment	40.82749	-73.94461	Harlem Public	40.828803	-73.948685	Bar
3	Private Bedroom in Cozy Hamilton Heights Apartment	40.82749	-73.94461	The Honeywell	40.828549	-73.948739	Speakeasy
4	Private Bedroom in Cozy Hamilton Heights Apartment	40.82749	-73.94461	Bono Trattoria	40.830180	-73.947420	Italian Restaurant

	name	latitude	longitude	Venue	Venue Latitude	Venue Longitude	Venue Type	Venue Category
4	Private Bedroom in Cozy Hamilton Heights Apartment	40.82749	-73.94461	Bono Trattoria	40.830180	-73.947420	Italian Restaurant	Food Place
8	Private Bedroom in Cozy Hamilton Heights Apartment	40.82749	-73.94461	Benyam Cuisine	40.825075	-73.940358	Ethiopian Restaurant	Food Place
9	Private Bedroom in Cozy Hamilton Heights Apartment	40.82749	-73.94461	Convent Garden	40.828560	-73.942531	Garden	Sightseeing
12	Private Bedroom in Cozy Hamilton Heights Apartment	40.82749	-73.94461	Wat's On Your Plate	40.825632	-73.948494	Caribbean Restaurant	Food Place
15	Private Bedroom in Cozy Hamilton Heights Apartment	40.82749	-73.94461	Riverside Drive Running Path	40.827738	-73.942124	Scenic Lookout	Sightseeing

	name	1st Most Common Venue	2nd Most Common Venue	3rd Most Common Venue	1st Most Common Venue Share	2nd Most Common Venue Share	3rd Most Common Venue Share
0	**Stylish, Quiet, Centrally Located (9th & 52nd)	Food Place	Entertainment	Sightseeing	0.69	0.31	0
1	157-C	Food Place	Sightseeing	Fine Art	0.57	0.24	0.14
2	A neat bedroom in a cozy 3-bedroom apartment	Food Place	Sightseeing	Shopping	0.48	0.41	0.04
3	Art filled peaceful paradise EV Union Square	Food Place	Entertainment	Sightseeing	0.71	0.14	0.07
4	Artsy Parisian Apt in Greenwich Village	Food Place	Entertainment	Sightseeing	0.73	0.18	0.09

	review_scores_rating	accommodates	full_price	price_per_person	crimes	bathrooms	bedrooms	Cluster Labels
count	51.000000	51.000000	51.000000	51.000000	51.000000	51.000000	51.000000	51.0
mean	99.960784	2.647059	208.352941	85.655229	73.058824	1.117647	1.078431	2.0
std	0.196039	1.634913	128.457903	49.044886	63.690945	0.431141	0.483452	0.0
min	99.000000	1.000000	54.000000	27.000000	4.000000	0.000000	0.000000	2.0
25%	100.000000	2.000000	90.000000	54.500000	35.000000	1.000000	1.000000	2.0
50%	100.000000	2.000000	175.000000	67.000000	62.000000	1.000000	1.000000	2.0
75%	100.000000	3.500000	282.500000	105.500000	83.500000	1.000000	1.000000	2.0
max	100.000000	10.000000	519.000000	259.500000	385.000000	3.000000	3.000000	2.0

	name	tab_area	neighbourhood_cleansed	review_scores_rating	property_type	room_type	accommodates	full_price	price_per_person
2029	Private Bedroom in Cozy Hamilton Heights Apartment	Hamilton Heights	Harlem	100.0	Apartment	Private room	2	54.0	27.000000
1801	Mr. B - Room Apartment in NYC	Washington Heights South	Washington Heights	100.0	Apartment	Private room	1	67.0	67.000000
1354	Cozy sanctuary/Big master bedroom/NYC/The Heights	Washington Heights North	Washington Heights	100.0	Apartment	Private room	2	74.0	37.000000
1529	Little Safe Haven	Hamilton Heights	Harlem	100.0	Apartment	Private room	1	80.0	80.000000
654	One cozy private BR close to the mecca of shopping	Turtle Bay-East Midtown	Midtown	100.0	Apartment	Private room	1	80.0	80.000000
2130	Spacious and Modern Hideout | Close to Everything	Hamilton Heights	Harlem	100.0	Apartment	Private room	2	85.0	42.500000
1694	Master Suite near NYP/CUMC and 2 blks from Subway	Washington Heights South	Washington Heights	100.0	Apartment	Private room	2	85.0	42.500000
2129	Gorgeous Manhattan Hideout | Close to Everything	Hamilton Heights	Harlem	100.0	Apartment	Private room	2	85.0	42.500000
1219	Old School New York	Upper West Side	Upper West Side	100.0	Apartment	Private room	2	87.0	43.500000

5. Results and Discussion

During the analysis, three clusters were defined. All clusters have a 'Food Place' category as the first Common Venues. This is what we have in common among our clusters. But they are distinguished by the other characteristics as

average Price per person average Crimes Rate the second Common Venues number of available Airbnb accommodations neighborhoods location Cluster 0 — Mix is the most generic cluster — it has a

average price_per_person — $110; average crimes rate — 67 (but very varying — depends on the neighborhood, from 3 to 385 crime cases in radius of 100 meters from the accommodation); mix of all Venue Categories (Fine Arts, Shopping, Entertainment); contains 58% from all accommodations selected from analysis (Top-100 Airbnb accommodations); spreads almost on all Manhattan's areas. Cluster 1 — Entertainment is the smallest cluster with the following particular qualities (Nightclub, Stadium, Pub, Theater, Concert and so on):

highest average price_per_person among all clusters — $111; highest average crimes rate among all clusters — 102; Entertainment is 1st and the 2nd Top Common Venue Categories; contains 15% from all top accommodations (Top-100 Airbnb accommodations); spreads on Chelsea, Hell's Kitchen, and Midtown Airbnb's Neighborhoods. Cluster 2 — Sightseeing is the cheapest one with many Sightseeing attractions nearby (Monument/Landmark, Memorial Site, Historic Site, Lake, Park, Pier, and so on)

lowest average price_per_person — $59; lowest crimes rate among all clusters — 65; Sightseeing is the second top Common Venue Category; contains 27% from all top accommodations (Top-100 Airbnb accommodations); spreads on East Harlem, Financial District, Harlem, Inwood, Morningside Heights, Roosevelt Island, Upper West Side, Washington Heights, West Village. We identified three clusters from which a visitor could choose an appropriate accommodation based on his/her preferences or needs.

Top Neighborhoods Statistics Top-5 Manhattan Tab Area (Airbnb Neighborhoods) with Lowest Average Price per Person in 2019 year:

Marble Hill-Inwood (Marble Hill, Inwood) — 45.48 USD — 25 accommodations Washington Heights South (Washington Heights) — 46.79 USD — 82 accommodations Washington Heights North (Inwood, Washington Heights) — 54.74 USD — 53 accommodations Central Harlem North-Polo Grounds (Harlem, East Harlem) — 57.00 USD — 132 accommodations Manhattanville (Harlem) — 59.75 USD — 25 accommodations Top-5 Manhattan Tab Area (Airbnb Neighborhoods) with the Lowest Crime level in 2019 year :Stuyvesant Town-Cooper Village (Stuyvesant Town) -145 park-cemetery-etc-Manhattan (Inwood, Washington Heights) — 1,213 Lenox Hill-Roosevelt Island (Roosevelt Island, Upper East Side) — 1,604 Manhattanville (Harlem) — 1,832 Yorkville (Upper East Side) — 1,898

6. Conclusion

To conclude, the basic data analysis was performed to identify Manhattan's Neighborhoods clusters for a short stay visit. During the analysis, we cleansed and investigated Manhattan Neighborhoods' datasets, found some statistical characteristics and visualize them.

The aim of this project is to help Manhattan visitors select the Airbnb neighborhoods where to stay based on the most common venues, price policy, and safety characteristics:

if a person is interested in Entertainment (Nightlife, Pubs, Concerts, Movies) we recommend paying attention for accommodations from the Cluster 1 — Entertainment: Chelsea, Hell's Kitchen, and Midtown Airbnb's Neighborhoods. But the person should take into the consideration the high prices and crime rate for this location; if a person is looking for a neighborhood with lower prices and nice views nearby we recommend looking at Cluster 2 — Sightseeing: Chelsea, Hell's Kitchen, and Midtown Airbnb's Neighborhoods; if a person does not have any preferences — investigate proposals from Cluster 0 — Mix. It has average prices and spreads over almost all Manhattan's neighborhoods.