The Battle of Neighbourhoods

Contents:

1. Introduction (Business Problem)

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 -
The data set-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% as seen by the data set given in the below link https://assets.simpleviewinc.com/simpleview/image/upload/v1/clients/newyorkcity/FYI_Hotel_reports_February_2019_8607015b-b32a-4c7f-9fbd-84cd2a93cbe6.pdf

Visitors often prefer short stays that are often over weekends - averaging 2.4 nights. This is shown in the link below https://aka.nyc/content/uploads/2017/12/new_york_city_travel_and_tourism_trend_report_2017.pdf

Problem Description

New York city has around 300 hotels with over 75,000 rooms. Airbnb has more than 50,000 apartment listings in New York City as of the year 2018. Due to this it can be hard to find the right fit or to know what one can afford. In this project we will try and find out the most optimal borough in Manhattan where tourists can rent an accommodation through Airbnb service. This will ensure that the tourist has a pleasant stay in New York and a possibility to be near the most visited attractions like Central Park, Times Square etc.

Target Audience

This analysis will interest visitors or tourists of New York city who prefers short stays and want to select the best neighbourhoods in Manhattan, New York.

Success Criteria

The success criteria of this project is a recommendation with the set of apartments clusters that have the best score calculated based on
- Cost of accommodation
- Location of accommodation
- Attractions in 100m radius of the accommodation
- Crime rate in a radius of 100m from the accommodation

2. Data

For this analysis we are using free and publicly availabel datasets.

We will find available Airbnb accommodations in Manhattan, New York and define the most affordable apartment clusters for tourists.

The factors that will contribute to this analysis are:
- Average price of accommodation in the neighbourhood
- Number of tourist attractions near the accommodation
- Crime rate near the accommodation

The datasets for the above listed factors are available on the links given below
- Apartment listing in New York City from *Inside Airbnb* :
  http://data.insideairbnb.com/united-states/ny/new-york-city/2019-12-04/data/listings.csv.gz
- Tabulation of New York neighbourhoods :
  https://data.cityofnewyork.us/api/geospatial/cpf4-rkhq?method=export&format=GeoJSON
- New York Crime data-
  https://data.cityofnewyork.us/api/views/5uac-w243/rows.csv?accessType=DOWNLOAD
- Foursquare API to extract data about venues -museums, restaurants, art galleries, shopping complexes, sightseeing attractions, concert stadiums etc.

Airbnb data

The records were filtered by below given criteria for our project:
- Brough- Manhattan, New York
- Number of reviews >= 10
- Last review- later than 10/01/2019
- Minimum nights >=1
- Availability >=10 days/year
- Exclude shared rooms, hostels and RV

This led to below data set

	id	listing_url	scrape_id	last_scraped	name	summary	space	description	experiences_offered	neighborhood_overview	nc
4	5178	https://www.airbnb.com/rooms/5178	20191204162729	2019-12-05	Large Furnished Room Near B'way	Please don't expect the luxury here just a basic room in the center of Manhattan.	You will use one large, furnished, private room of a two-bedroom apartment and share a bathroom with the host. The apartment is located a few blocks away from Central Park between 8th and 9th Avenue. The closest subway station is Columbus Circle 59th Street. Great restaurants, Broadway and all transportation are easily accessible. The cost of the room is $79 per night. Weekly rate is available. There is a $12.00 fee per guest. The apartment also features hardwood floors and a second-floor walk-up. There is a full-sized bed,TV, microwave, and a	Please don't expect the luxury here just a basic room in the center of Manhattan. You will use one large, furnished, private room of a two-bedroom apartment and share a bathroom with the host. The apartment is located a few blocks away from Central Park between 8th and 9th Avenue. The closest subway station is Columbus Circle 59th Street. Great restaurants, Broadway and all transportation are easily accessible. The cost of the room is $79 per night. Weekly rate is available. There is a $12.00 fee per guest. The apartment also features hardwood floors	none	Theater district, many restaurants around here.	Reserva shoul made at l a few c before da arrival reserva for the will be ta so please not availability

New York Crime Records

We need to read the crime records of New York city
Only felony and misdemeanour were considered for this analysis

	CMPLNT_NUM	ADDR_PCT_CD	BORO_NM	CMPLNT_FR_DT	CMPLNT_FR_TM	CMPLNT_TO_DT	CMPLNT_TO_TM	CRM_ATPT_CPTD_CD	HADEVELOF
0	314773184	48	BRONX	31.12.19	06:00:00 PM	NaN	NaN	COMPLETED	NaN
1	289837961	25	MANHATTAN	30.12.19	08:30:00 PM	31.12.19	10:00:00 AM	COMPLETED	NaN
2	535744284	77	BROOKLYN	24.12.19	04:55:00 PM	24.12.19	05:00:00 PM	COMPLETED	NaN
3	895678119	52	BRONX	30.12.19	07:32:00 PM	NaN	NaN	COMPLETED	NaN
4	299841674	18	MANHATTAN	30.12.19	03:30:00 PM	30.12.19	04:50:00 PM	COMPLETED	NaN

Data Cleaning

Needed Column Description:

- id: Listing identifier
- name: Accommodation's name
- last_review: Date accommodation was last reviewed
- listing_url: Accommodation's url
- picture_url: Accommodation's picture url
- neighbourhood_group_clensed: New York's brough names e.g Manhattan, Bronx etc. (We will be using accommodations only from Manhattan)
- neighbourhood_cleansed: Airbnb neighbourhood name (They are not the same as NYC neighbourhood tabulation areas
- review_score_rating: Accommodation's weighted sum of other scores like location
- latitude: Accommodation location's latitude
- longitude: Accommodation location's longitude
- property_type: Accommodation's type eg. Private room, entire apartment
- room_type: Accommodation's room type
- bathrooms: Number of bathrooms
- bedrooms: Number of bedrooms
- square_feet: Accommodation's area
- price: price per night
- security_deposite: security deposite amount
- cleaning_fee: Additional fee
- minimum_nights: Minimum nights for rent
- number_of_reviews_ltm: Number of review for the last month
- reviews_per_month: Average number of reviews per month
- number_of_reviews: Overall number of reviews
- availability_365: Available days/year

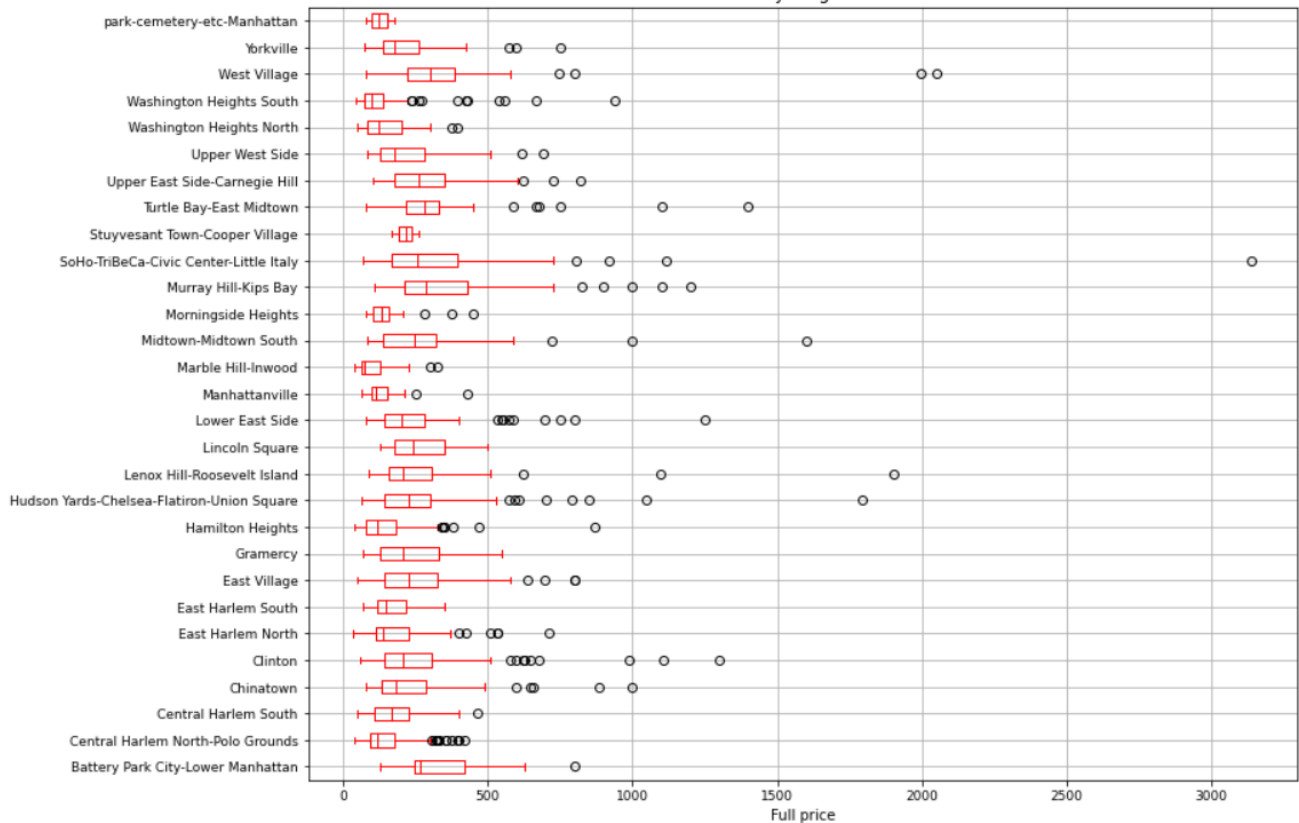We should clean different Prices columns :

fill in empty values; convert String to Float, e.g. $2,100.00 => 2100.00.

We clean different price columns:

- Fill in empty values
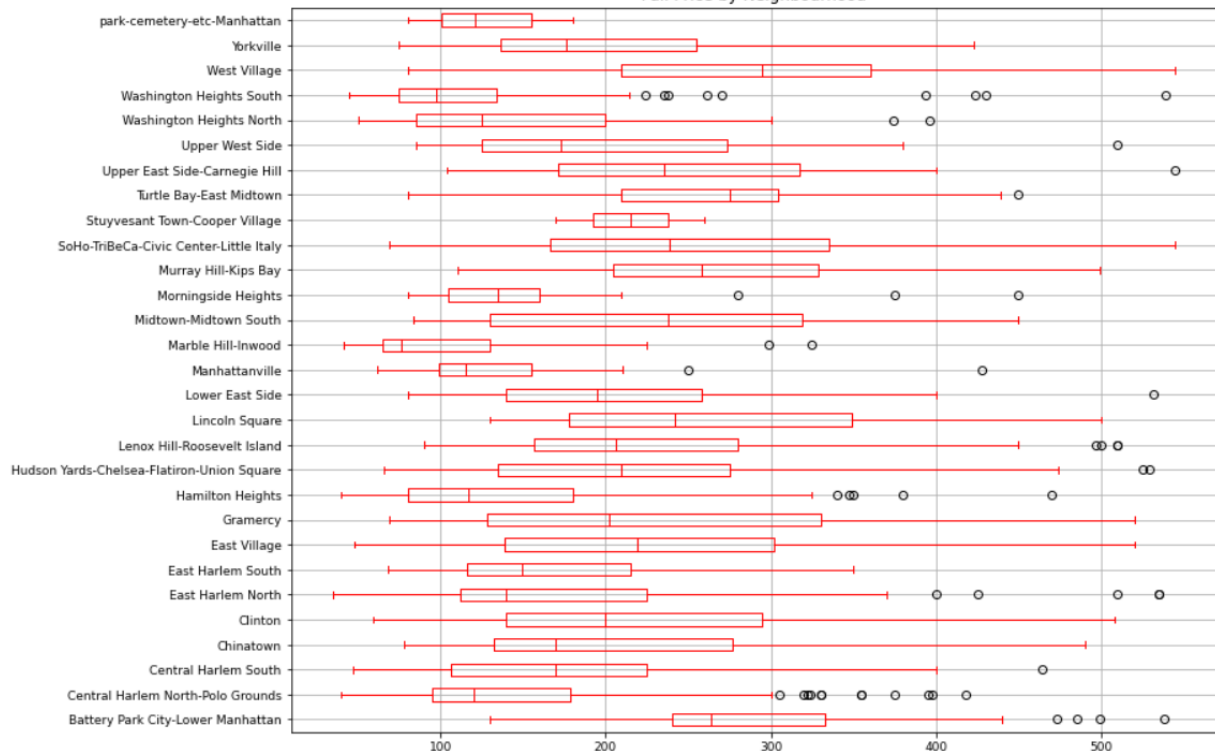- Convert string to float eg. $2,100.00- 2,100.00

Outliners

Full Price by Neighbourhood

In the above chart we can see some outliers - observation points that are distant from other observations. In our case they are the indication of variance in the data. So we remove them.

Full Price by Neighbourhood

In this project we tried to detect Manhattan's Neighborhoods that have accommodations for rent with positive reviews, reasonable prices, low number of crimes and tourists' attractions nearby.

3. Procedure

In the first step we collected the following data:
- Airbnb Accommodations with their NYC Tabulation Area (official neighbourhood names)
- Airbnb Accommodation's number of crimes nearby
- Defined NYC Tabulation Area (official neighbourhood name) for each Manhattan's crime case.

The second step in our analysis will be a calculation and exploration different neighbourhoods of Manhattan. We will explore the following characteristics:
- number of crimes in the area
- average price per person
- number of accommodations available.

In third step we select Top-100 Airbnb accommodations based on:
- Summary rating, number of crimes and price per person
- Invoke Foursquare API to find Top accommodations' nearby venues create
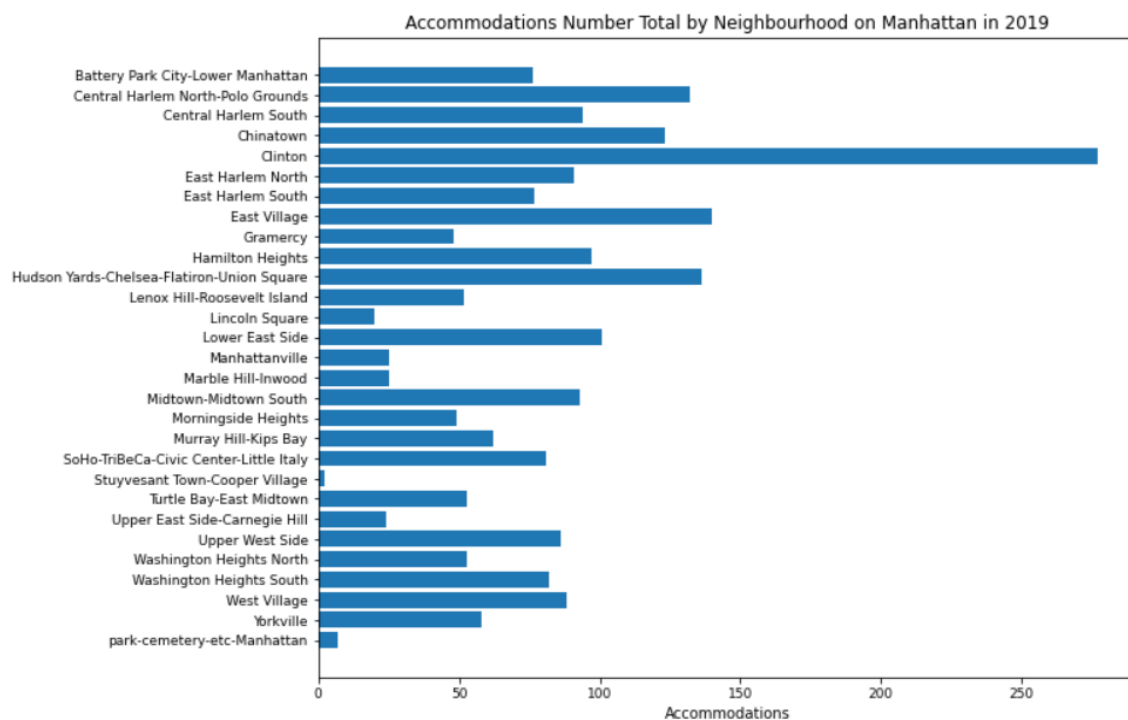
4. Analysis

Average price per neighbourhood

Out[ ]:

	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
22	Upper East Side-Carnegie Hill	96.988889	24	62.833333

Apartments Total by Neighbourhood Chart

Accommodations Number Total by Neighbourhood on Manhattan in 2019

Horizontal bar graph of neighbourhood crime records

Neighborhoods Total Crimes Records on Manhattan in 2019
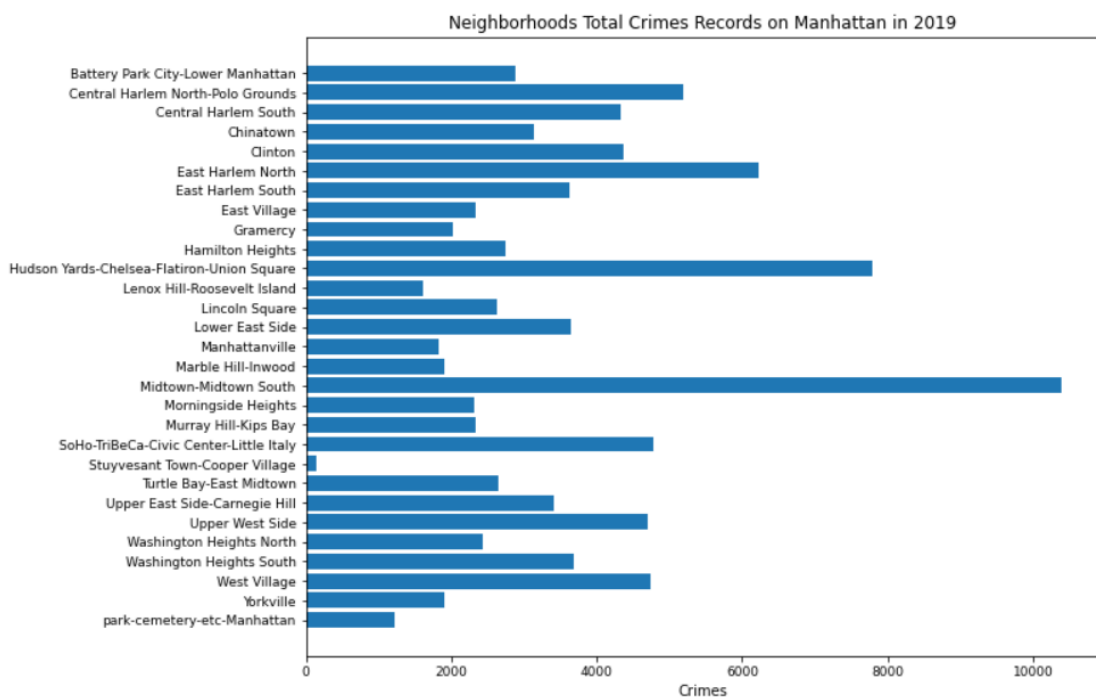
Analysis Map
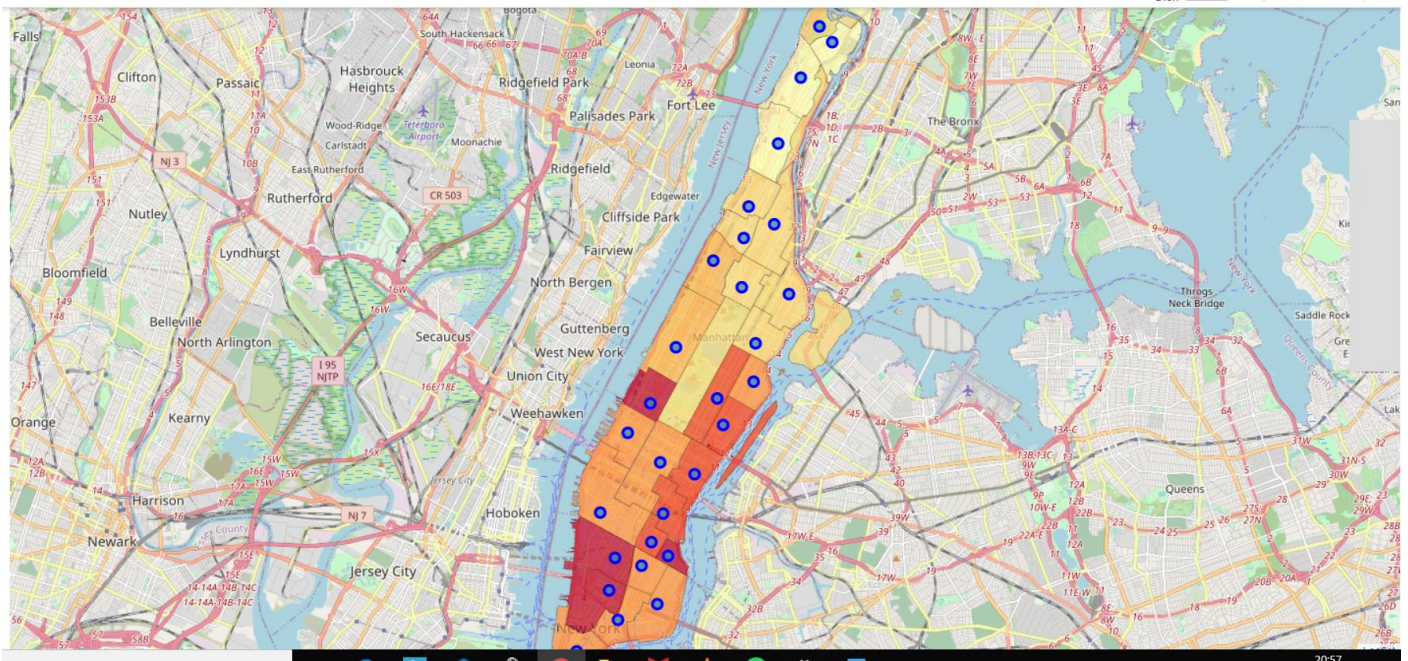
Crimes cluster map
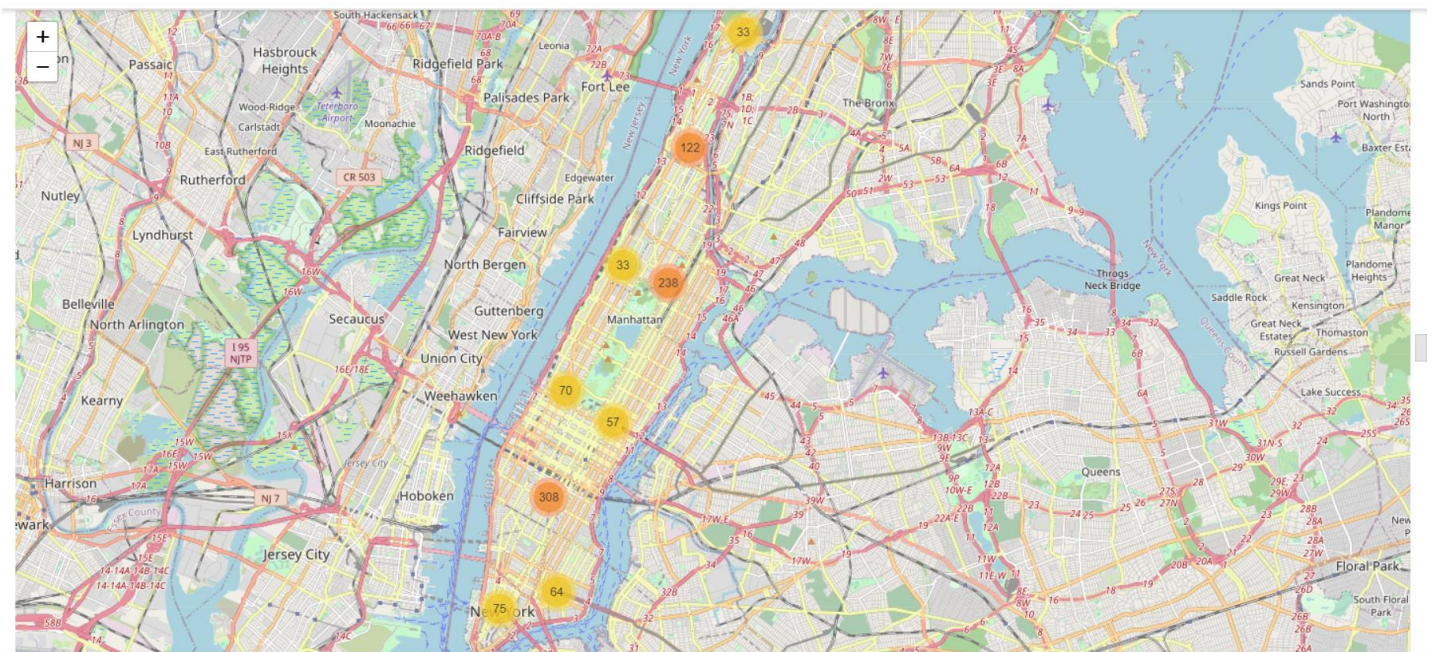
Top-5 Neighbourhoods with the Highest Crime level:

- Midtown-Midtown South - 10,397
- Hudson Yards-Chelsea-Flatiron-Union Square - 7,788
- East Harlem North - 6,221
- Central Harlem North-Polo Grounds - 5,186
- SoHo-TriBeCa-Civic Center-Little Italy - 4,789

Top-5 Neighbourhoods with the Lowest Crime level in 2019 year :

- Stuyvesant Town-Cooper Village - 145
- park-cemetery-etc-Manhattan - 1,213
- Lenox Hill-Roosevelt Island - 1,604
- Manhattanville - 1,832

- Yorkville - 1,898

Foursquare API Neighbourhoods Analysis

We choose Top accommodations by
- review_scores_rating - overall accommodations rating - from maximum 100 to lower values
- full_price - from lower price to higher
- price_per_person - from lower price to higher
- crimes - from lower number to higher

In [ ]: `df_airnb_top_venues.head()`

Out[ ]:

|	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	Harlem Public	40.828803	-73.948685	Bar
2	Private Bedroom in Cozy Hamilton Heights Apartment	40.82749	-73.94461	Buddha Taco Bar	40.829007	-73.948089	Bar
3	Private Bedroom in Cozy Hamilton Heights Apartment	40.82749	-73.94461	Monkey Cup	40.825694	-73.947234	Coffee Shop
4	Private Bedroom in Cozy Hamilton Heights Apartment	40.82749	-73.94461	The Honeywell	40.828549	-73.948739	Speakeasy

Custom top level categories

	name	latitude	longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
5	Private Bedroom in Cozy Hamilton Heights Apartment	40.82749	-73.94461	Wat's On Your Plate	40.825632	-73.948494	Caribbean Restaurant
6	Private Bedroom in Cozy Hamilton Heights Apartment	40.82749	-73.94461	Barepas	40.827717	-73.946173	Arepa Restaurant
7	Private Bedroom in Cozy Hamilton Heights Apartment	40.82749	-73.94461	Bono Trattoria	40.830180	-73.947420	Italian Restaurant
8	Private Bedroom in Cozy Hamilton Heights Apartment	40.82749	-73.94461	Convent Garden	40.828560	-73.942531	Garden
12	Private Bedroom in Cozy Hamilton Heights Apartment	40.82749	-73.94461	Riverside Drive Running Path	40.827738	-73.942124	Scenic Lookout

K-means Clustering

	Cluster Labels	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	0	**Stylish, Quiet, Centrally Located (9th & 52nd)	Food Place	Entertainment	Sightseeing	0.64	0.36	0
1	2	157-C	Food Place	Sightseeing	Fine Art	0.58	0.25	0.12
2	2	A neat bedroom in a cozy 3-bedroom apartment	Sightseeing	Food Place	Shopping	0.45	0.45	0.03
3	1	Art filled peaceful paradise EV Union Square	Food Place	Entertainment	Sightseeing	0.67	0.17	0.08
4	1	Artsy Parisian Apt in Greenwich Village	Food Place	Entertainment	Sightseeing	0.65	0.25	0.1
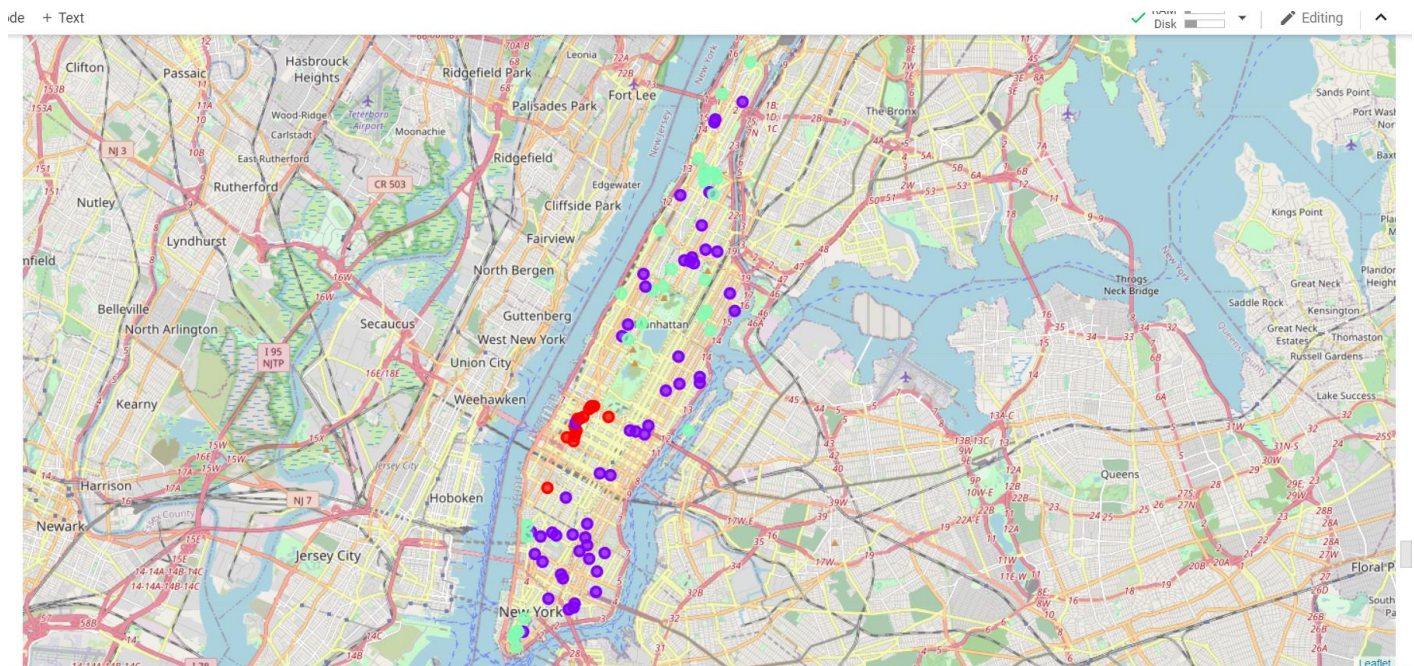
5. Result and discussion

The analysis defined 3 clusters All clusters have a 'Food Place' category as the first Common Venues. This is common among the clusters defines 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 - This 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
- 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 1 : It contains

- average price_per_person - $110
- average crimes rate - 67
- mix of all Venue Categories
- contains 58% from all accommodations selected from analysis (Top-100 Airbnb accommodations)
- spreads almost on all Manhattan's areas.

Cluster 2 Sightseeing is the cheapest one with many Sightseeing attractions nearby

- 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

Note

- Our analysis was limited to Manhattan Brugh's only
- Foursquare developer account has a limitation of 950 calls/day so maybe it's worth to upgrade our free account to analyze Top-1000 Airbnb accommodations instead of Top-100.

6. Conclusion

The basic data analysis was performed to identify Manhattan's Neighborhoods clusters for a short stay visit. During the analysis, we cleaned 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 0 - * 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 1 - Mix. It has average prices and spreads over almost all Manhattan's neighborhoods.