

# Task 1: Visualizing geospatial data

```
In [35]: import pandas as pd
```



```
In [2]: pip install folium
```



```
Collecting folium
  Downloading folium-0.14.0-py2.py3-none-any.whl (102 kB)
    0.0/102.3 kB ? eta ------
----- 61.4/102.3 kB 1.1 MB/s eta 0:00:01
----- 102.3/102.3 kB 1.2 MB/s eta 0:00:00

Collecting branca>=0.6.0 (from folium)
  Downloading branca-0.6.0-py3-none-any.whl (24 kB)
Requirement already satisfied: jinja2>=2.9 in d:\apps\anaconda\files\lib\site-packages (from folium) (2.11.3)
Requirement already satisfied: numpy in d:\apps\anaconda\files\lib\site-packages (from folium) (1.23.5)
Requirement already satisfied: requests in d:\apps\anaconda\files\lib\site-packages (from folium) (2.22.0)
Requirement already satisfied: MarkupSafe>=0.23 in d:\apps\anaconda\files\lib\site-packages (from jinja2>=2.9->folium) (2.0.1)
Requirement already satisfied: chardet<3.1.0,>=3.0.2 in d:\apps\anaconda\files\lib\site-packages (from requests->folium) (3.0.4)
Requirement already satisfied: idna<2.9,>=2.5 in d:\apps\anaconda\files\lib\site-packages (from requests->folium) (2.8)
Requirement already satisfied: urllib3!=1.25.0,!=1.25.1,<1.26,>=1.21.1 in d:\apps\anaconda\files\lib\site-packages (from requests->folium) (1.25.11)
Requirement already satisfied: certifi>=2017.4.17 in d:\apps\anaconda\files\lib\site-packages (from requests->folium) (2022.12.7)
Installing collected packages: branca, folium
Successfully installed branca-0.6.0 folium-0.14.0
Note: you may need to restart the kernel to use updated packages.
```

```
WARNING: Ignoring invalid distribution - (d:\apps\anaconda\files\lib\site-packages)
WARNING: Ignoring invalid distribution -ensorflow-intel (d:\apps\anaconda\files\lib\site-packages)
WARNING: Ignoring invalid distribution -rotobuf (d:\apps\anaconda\files\lib\site-packages)
WARNING: Ignoring invalid distribution - (d:\apps\anaconda\files\lib\site-packages)
WARNING: Ignoring invalid distribution -ensorflow-intel (d:\apps\anaconda\files\lib\site-packages)
WARNING: Ignoring invalid distribution -rotobuf (d:\apps\anaconda\files\lib\site-packages)

[notice] A new release of pip is available: 23.1.2 -> 23.2.1
[notice] To update, run: D:\apps\anaconda\files\python.exe -m pip install --upgrade pip
```

```
In [36]: import folium
```



## Load New York Dataset

```
In [37]: df = pd.read_csv("AB_NYC_2019.csv")
df
```



Out[37]:

	<b>id</b>	<b>name</b>	<b>host_id</b>	<b>host_name</b>	<b>neighbourhood_group</b>	<b>neighbourhood</b>	<b>latitude</b>	<b>longitude</b>	<b>room_type</b>	<b>price</b>	<b>minimum_nights</b>
<b>0</b>	2539	Clean & quiet apt home by the park	2787	John	Brooklyn	Kensington	40.64749	-73.97237	Private room	149	1
<b>1</b>	2595	Skylit Midtown Castle	2845	Jennifer	Manhattan	Midtown	40.75362	-73.98377	Entire home/apt	225	1
<b>2</b>	3647	THE VILLAGE OF HARLEM....NEW YORK !	4632	Elisabeth	Manhattan	Harlem	40.80902	-73.94190	Private room	150	3
<b>3</b>	3831	Cozy Entire Floor of Brownstone	4869	LisaRoxanne	Brooklyn	Clinton Hill	40.68514	-73.95976	Entire home/apt	89	1
<b>4</b>	5022	Entire Apt: Spacious Studio/Loft by central park	7192	Laura	Manhattan	East Harlem	40.79851	-73.94399	Entire home/apt	80	10
...	...	...	...	...	...	...	...	...	...	...	...
<b>48890</b>	36484665	Charming one bedroom - newly renovated rowhouse	8232441	Sabrina	Brooklyn	Bedford-Stuyvesant	40.67853	-73.94995	Private room	70	2
<b>48891</b>	36485057	Affordable room in Bushwick/East Williamsburg	6570630	Marisol	Brooklyn	Bushwick	40.70184	-73.93317	Private room	40	4
<b>48892</b>	36485431	Sunny Studio at Historical Neighborhood	23492952	Ilgar & Aysel	Manhattan	Harlem	40.81475	-73.94867	Entire home/apt	115	10
<b>48893</b>	36485609	43rd St. Time Square-cozy single bed	30985759	Taz	Manhattan	Hell's Kitchen	40.75751	-73.99112	Shared room	55	1
<b>48894</b>	36487245	Trendy duplex in the very	68119814	Christophe	Manhattan	Hell's Kitchen	40.76404	-73.98933	Private room	90	7

id	name	host_id	host_name	neighbourhood_group	neighbourhood	latitude	longitude	room_type	price	minimum_nights
	heart of Hell's Kitchen									

18805 rows × 11 columns

## Data Pre-Processing

```
In [38]: df.isna().sum()
```

```
Out[38]: id 0  
name 16  
host_id 0  
host_name 21  
neighbourhood_group 0  
neighbourhood 0  
latitude 0  
longitude 0  
room_type 0  
price 0  
minimum_nights 0  
number_of_reviews 0  
last_review 10052  
reviews_per_month 10052  
calculated_host_listings_count 0  
availability_365 0  
dtype: int64
```

```
In [39]: df.name.fillna("None", inplace=True)  
df.host_name.fillna("None", inplace=True)  
df.last_review.fillna(df.last_review.mode()[0], inplace=True)  
df.reviews_per_month.fillna(df.reviews_per_month.mean(), inplace=True)
```

```
In [40]: df.isna().sum()
```

```
Out[40]: id          0  
name        0  
host_id      0  
host_name    0  
neighbourhood_group 0  
neighbourhood 0  
latitude      0  
longitude     0  
room_type     0  
price         0  
minimum_nights 0  
number_of_reviews 0  
last_review    0  
reviews_per_month 0  
calculated_host_listings_count 0  
availability_365 0  
dtype: int64
```



## Visualize New York Data On Map

```
In [41]: # New York Center  
ny_lat = df.latitude.mean()  
ny_lon = df.longitude.mean()
```

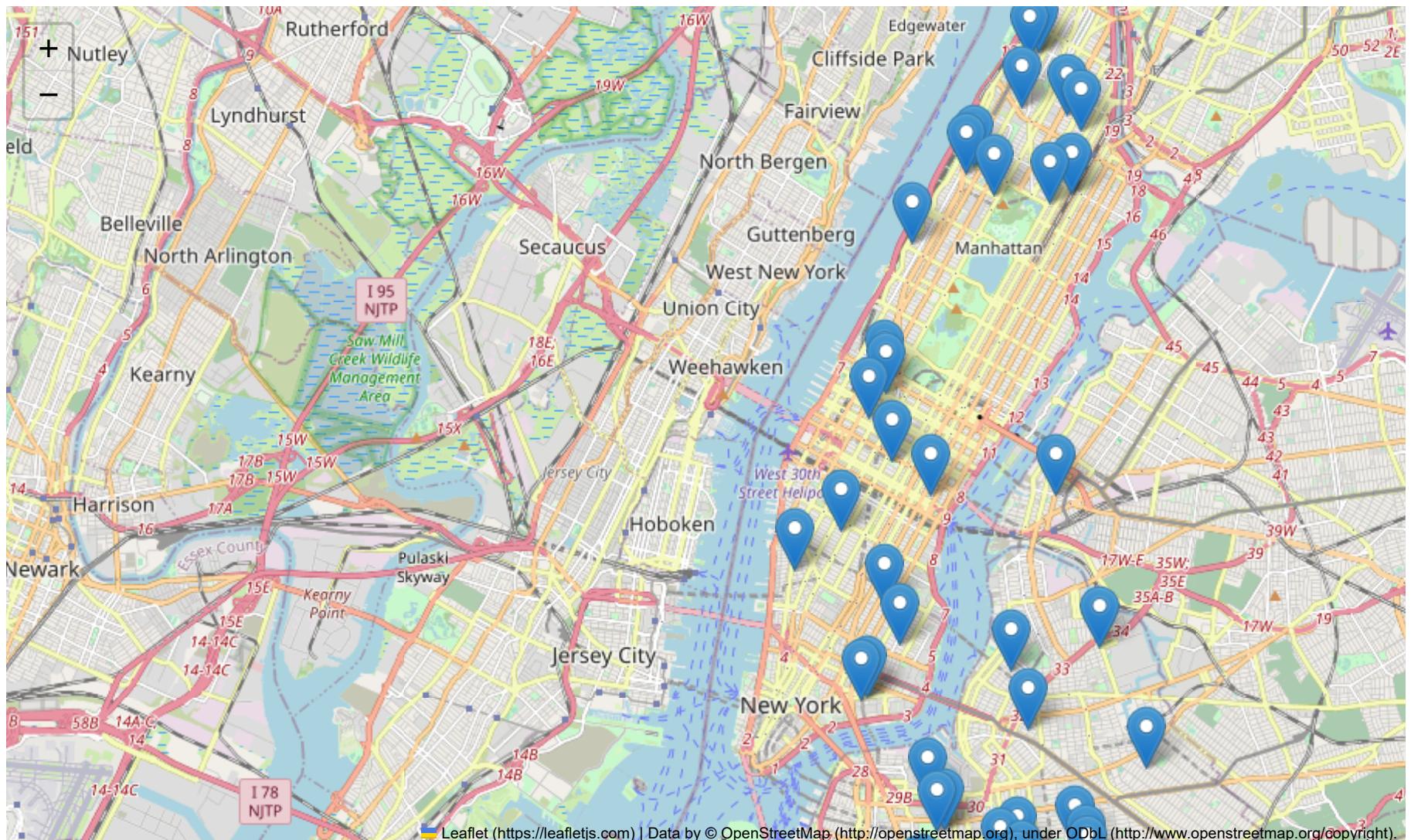


```
In [42]: m = folium.Map(location=[ny_lat, ny_lon])  
  
for index, row in df.head(50).iterrows():  
    folium.Marker(location=[row.latitude, row.longitude],  
                  popup=f'host Name: {row.host_name}<br><br><b>Neighbor:</b> {row.neighbourhood}')
```

```
m
```



Out[42]:



## Air BNB Data in Clustered Form

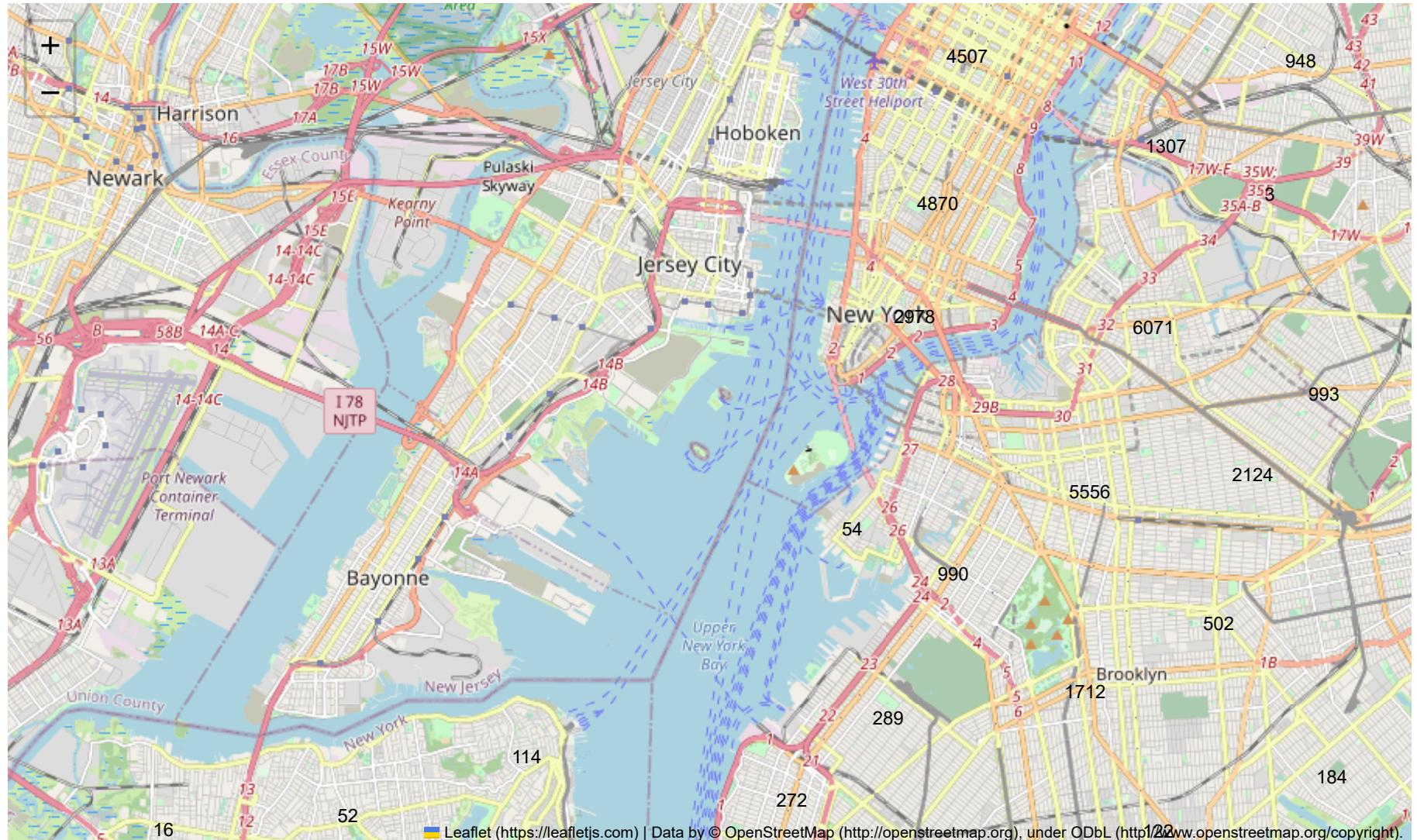
In [43]:

```
from folium.plugins import FastMarkerCluster

map1 = folium.Map(location=[ny_lat, ny_lon])
locations = list(zip(df.latitude, df.longitude))
```

```
FastMarkerCluster(data=locations).add_to(map1)
map1
```

Out[43]:

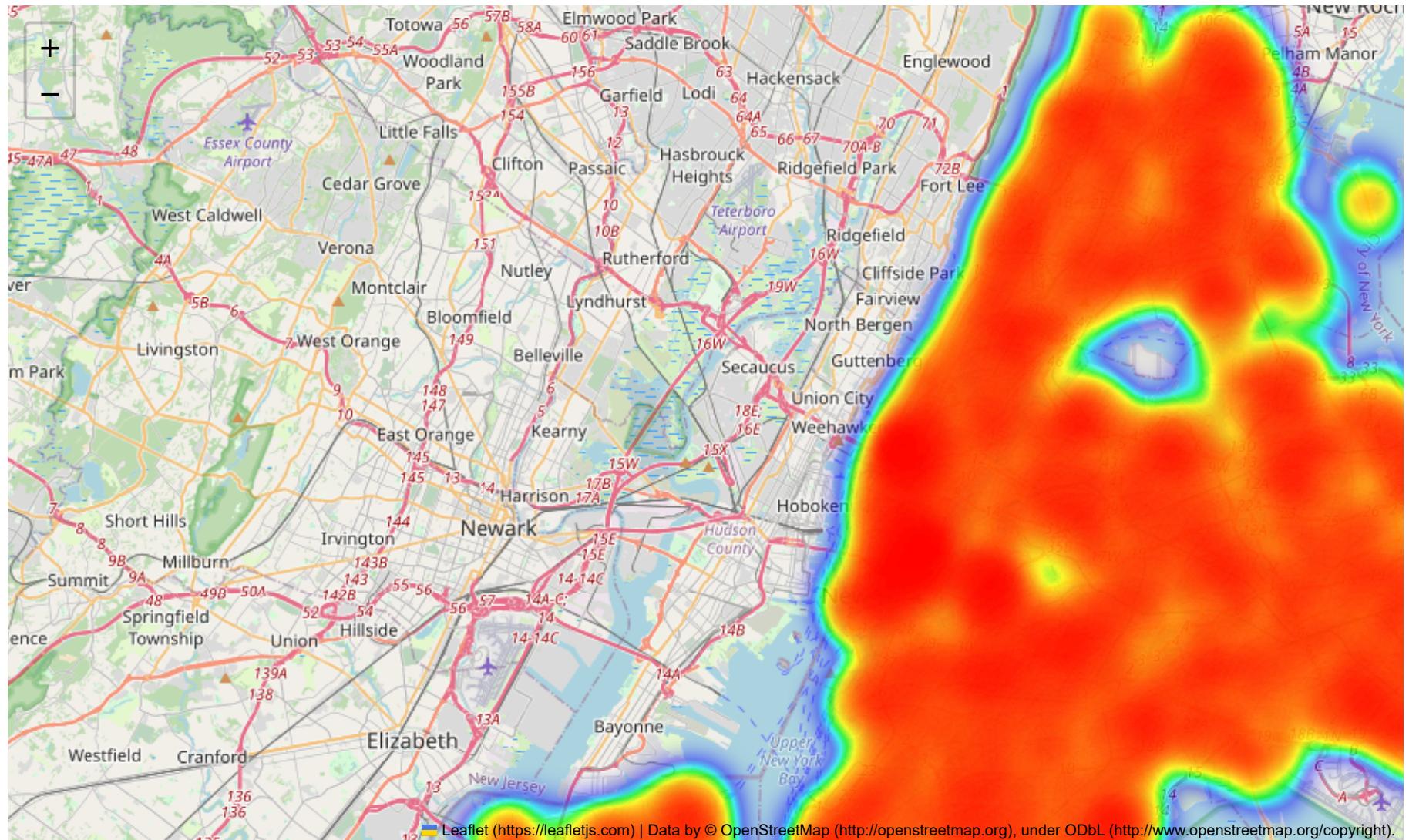


## AirBnB HeatMap in NYC

```
In [44]: from folium.plugins import HeatMap
```

```
map2 = folium.Map(location=[ny_lat, ny_lon])  
  
HeatMap(data=locations).add_to(map2)  
map2
```

Out[44]:



```
In [45]: # df.groupby("host_name").value_counts().reset_index()
```

```
# df['number_of_reviews'].value_counts(ascending=False).reset_index()  
df.groupby(['neighbourhood', 'host_name']).value_counts(ascending=False).reset_index()
```

Out[45]:

	neighbourhood	host_name	id	name	host_id	neighbourhood_group	latitude	longitude	room_type	price	minimum_nights
0	Allerton	Alex	33363084	E community that is commercially (Website hidd...	242175033	Bronx	40.86003	-73.86584	Entire home/apt	60	1
1	Allerton	Amoyiem	34941479	Place to be Private room 2	263266237	Bronx	40.86997	-73.84867	Private room	36	2
2	Allerton	Amoyiem	34941499	Private room 1	263266237	Bronx	40.86958	-73.84872	Private room	38	2
3	Allerton	Amoyiem	34941506	Place to be( 3bedroom apartment)	263266237	Bronx	40.86880	-73.84726	Private room	120	2
4	Allerton	Amoyiem	35237543	The Place to be private room 3	263266237	Bronx	40.86871	-73.84740	Private room	40	2
...	...	...	...	...	...	...	...	...	...	...	...
48890	Woodside	Zain	21283357	Brand New Fully Remodeled Modern House	153988027	Queens	40.74267	-73.89960	Entire home/apt	255	4
48891	Woodside	Zeyna	34976241	Cozy Apartment perfect for NYC vacation	254119216	Queens	40.74479	-73.91193	Entire home/apt	108	2
48892	Woodside	Zia	15286645	Newly Renovated Studio Near Manhattan	97300191	Queens	40.74767	-73.90721	Entire home/apt	110	2
48893	Woodside	Zoe	31445795	Cozy clean private room in Woodside. 20mins to...	86322604	Queens	40.75396	-73.90111	Private room	56	2

	neighbourhood	host_name	id	name	host_id	neighbourhood_group	latitude	longitude	room_type	price	minimum_nights
48894	Woodside	Zumunta	29096185	Welcome to Woodside!	56177749	Queens	40.74218	-73.90403	Private room	31	1

18895 rows x 17 columns

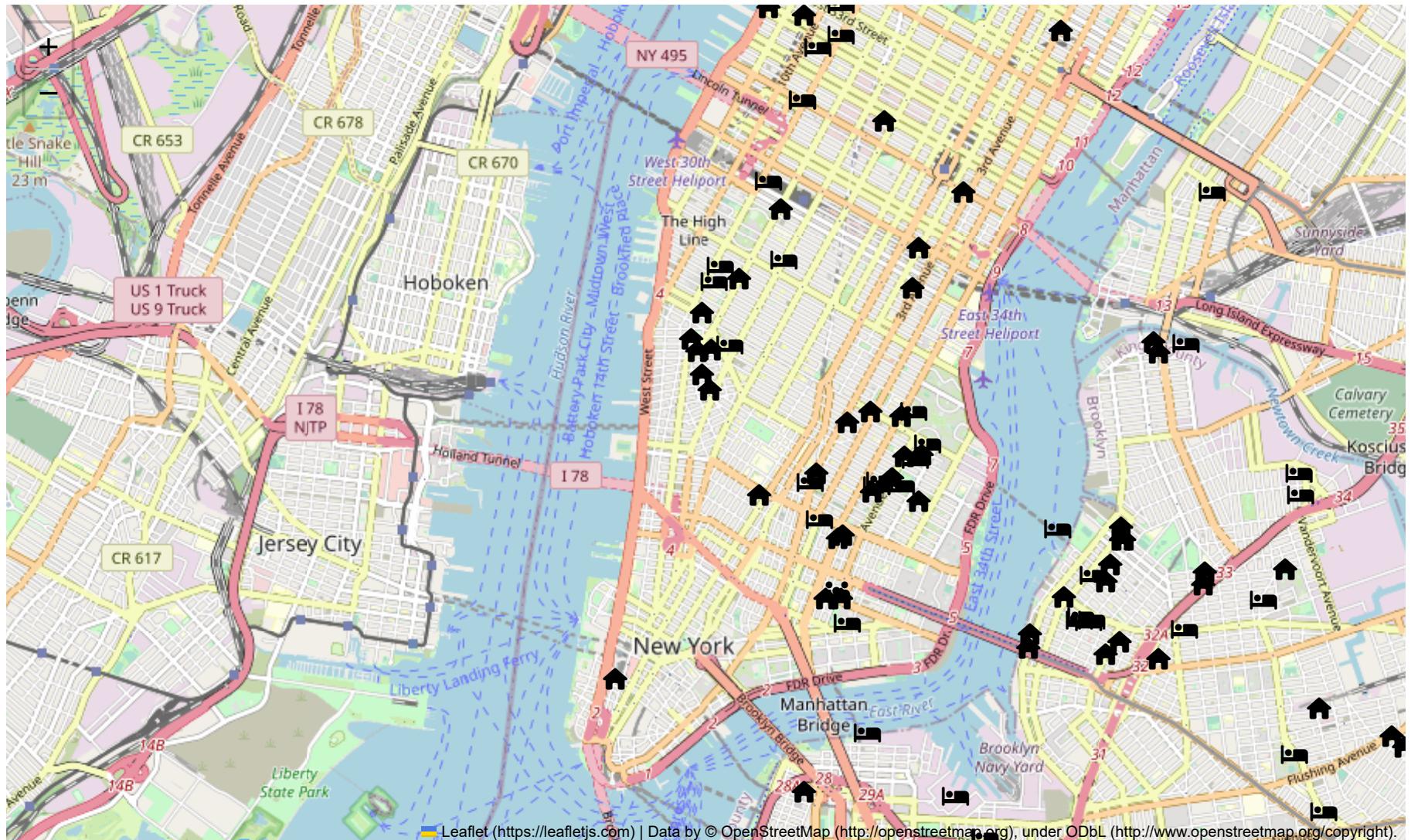
## Visualizing Different Types of Room Types

```
In [46]: m = folium.Map(location=[ny_lat, ny_lon])

for index, row in df.head(200).iterrows():
    if row.room_type == 'Entire home/apt':
        folium.Marker(location=[row.latitude, row.longitude], icon=folium.Icon(color="blue",icon="home", prefix='fa')).add_to(m)
    elif row.room_type == 'Private room':
        folium.Marker(location=[row.latitude, row.longitude], icon=folium.Icon(color="red",icon="bed", prefix='fa')).add_to(m)
    elif row.room_type == 'Shared room':
        folium.Marker(location=[row.latitude, row.longitude], icon=folium.Icon(color="green",icon="users", prefix='fa')).add_to(m)

m
```

Out[46]:

In [47]: `df.room_type.unique()`Out[47]: `array(['Private room', 'Entire home/apt', 'Shared room'], dtype=object)`In [48]: `df.neighbourhood = df.neighbourhood.apply(lambda x: x.split(",")[0])`

## Top 20 Most Expensive Places

In [49]:

```
from folium.plugins import HeatMap

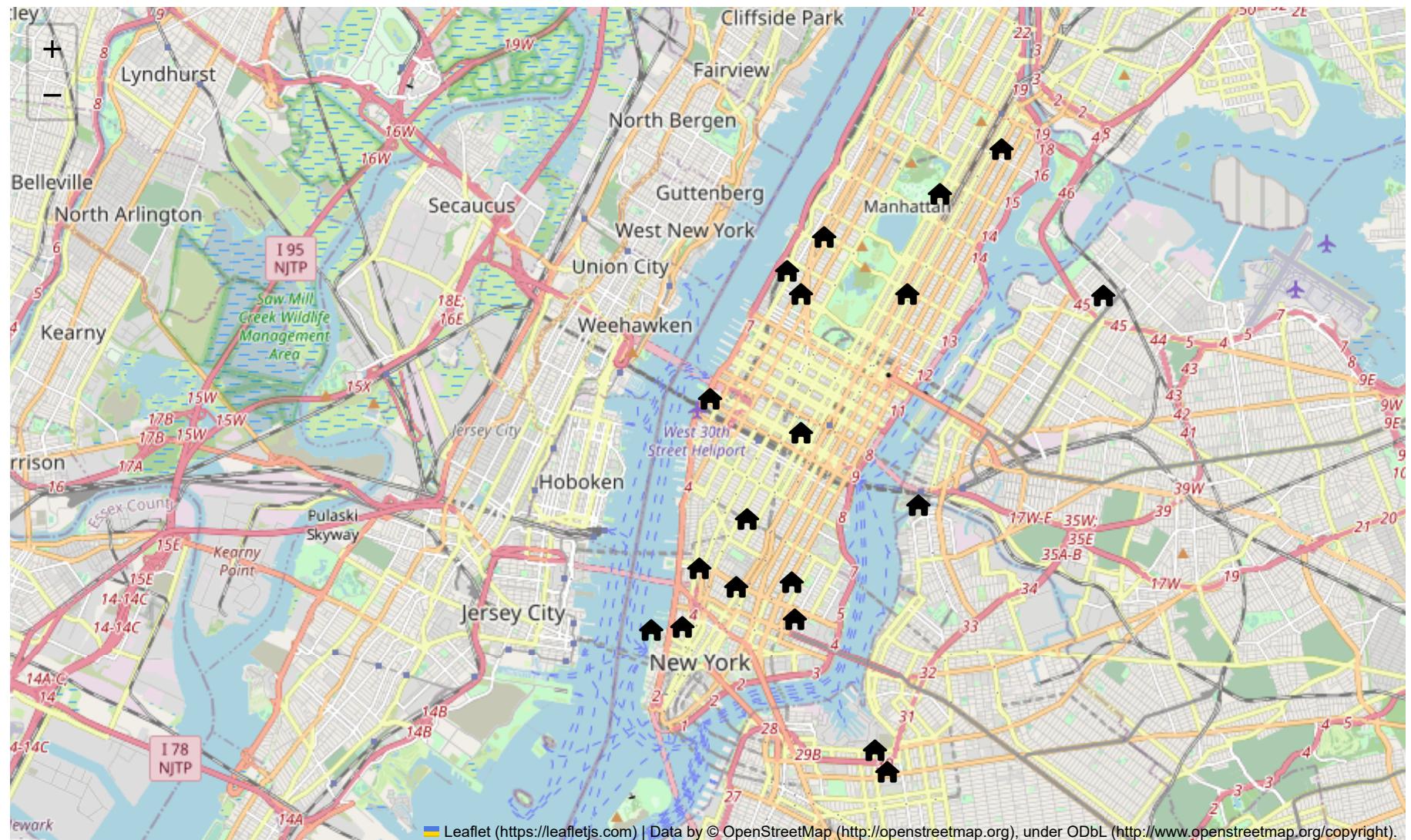
expensive_20_airbnbs = df.sort_values(by='price', ascending=False).head(20)
m = folium.Map(location=[ny_lat, ny_lon], zoom_start=12)

for index, row in expensive_20_airbnbs.iterrows():
    folium.Marker(location=[row.latitude, row.longitude], popup=row.host_name, icon=folium.Icon(color="purple", icon="home", prefix="fa"))

m
```



Out[49]:



## Task 2: Industry project

Looking for someone to write a script I can use over and over to scrape a restaurant menu and all the modifiers and prices. Prepare a csv for the data.

```
In [1]: pip install webdriver_manager
```

```
In [33]: # menuItems = doc.findall("div", {"class": ["flex", "sm12"]})  
menuItems = doc.select("div.flex.sm12.v-card__title")
```

```
final = []
for item in menuItems:
    menu = {'title':'', 'price':'', 'description': ''}
    if item.select(".itemLabel"):
        menu['title'] = item.select(".itemLabel")[0].text
        if item.select(".priceStyle"):
            menu['price'] = item.select(".priceStyle")[0].text
        if item.select(".itemDescription"):
            menu['description'] = item.select(".itemDescription")[0].text
    final.append(menu)
```

```
# Len(menuItems)
# final

MenusDataframe = pd.DataFrame(final, columns=['title', 'price', 'description'])
MenusDataframe.to_csv("MenuDetail.csv", index=None)
```

In [34]: `menudf = pd.read_csv("MenuDetail.csv")  
menudf`

Out[34]:

		title	price	description
0		Classic Nachos	9.99	Choice of one topping: refried beans, black b...
1		Nachos Fajitas	13.99	Grilled steak or chicken cooked with bell pep...
2		Azul Special Nachos	12.99	Choice of one topping: refried beans, black b...
3		Seafood Nachos	15.50	Scallops, shrimp, bell peppers, onions, and t...
4		Nachos Mexicanos	12.75	Chorizo, shredded chicken, topped with cilant...
...		...	...	...
173		Separate Pickup and Delivery Menus	0.00	Allows you to remove items for delivery.
174		Advance Ordering	0.00	Choose before they start an order so they get...
175		Auto POS Price Sync with \$Amount or %Percenta...	0.00	You don't have to do the math!\nOnly availabl...
176		Catering Menus	0.00	Create separate catering menus integrated wit...
177		Kiosk Mode	0.00	

178 rows × 3 columns

In [ ]:

In [ ]:

In [ ]: