In [5]:
```
!jupyter nbconvert --to pdf --execute --template classic airbnb-geovis-activities.html
```

Exploratory data analysis (EDA)

In [ ]:
```
# Install nbconvert and pyppeteer for the current user
!pip install --user nbconvert pyppeteer

# Export the notebook to HTML
!jupyter nbconvert --to html --template classic airbnb-geovis-activities.ipynb

# Export HTML to PDF using pyppeteer
!jupyter nbconvert --to pdf --execute --template classic airbnb-geovis-activities.html
```

You will learn how to systematically approach investigating an unknown dataset while maintaining a creative and open mind to search for insights.

Context

Airbnb is an online marketplace for people to rent places to stay.

Airbnb has rolled out a new service to help listers set prices. Airbnb makes a percentage commission off of the listings, so they are incentivized to help listers price optimally; that is, at the maximum possible point where they will still close a deal. You are an Airbnb consultant helping with this new pricing service.

Goal

We are going to focus on a question: which features are helpful for finding out the appropriate listing price?

Load Data

In [2]:
```
import pandas as pd
#from matplotlib import pyplot as plt
import seaborn as sns
import folium
```

In [3]:
```
listings = pd.read_csv('data/airbnb_nyc.csv')
```

In [4]:
```
listings
```

Out[4]:

	id	name	summary	description	experiences_offered	neighborhood_overview	transit	house_rules	ho
0	2539	Clean & quiet apt home by the park	Renovated apt home in elevator building.	Renovated apt home in elevator building. Spaci...	none	Close to Prospect Park and Historic Ditmas Park	Very close to F and G trains and Express bus i...	-The security and comfort of all our guests is...	
1	3647	THE VILLAGE OF HARLEM....NEW YORK !	NaN	WELCOME TO OUR INTERNATIONAL URBAN COMMUNITY T...	none	NaN	NaN	Upon arrival please have a legibile copy of yo...	
2	7750	Huge 2 BR Upper East Cental Park	NaN	Large Furnished 2BR one block to Central Park...	none	NaN	NaN	NaN	1
3	8505	Sunny Bedroom Across Prospect Park	Just renovated sun drenched bedroom in a quiet...	Just renovated sun drenched bedroom in a quiet...	none	Quiet and beautiful Windsor Terrace. The apart...	Ten minutes walk to the 15th sheet F&G train s...	- No shoes in the house - Quiet hours after 11...	2
4	8700	Magnifique Suite au N de Manhattan - vue Cloitres	Suite de 20 m2 a 5 min des 2 lignes de metro a...	Suite de 20 m2 a 5 min des 2 lignes de metro a...	none	NaN	Metro 1 et A	NaN	2
...	...	...	...	...	...	...	...	...	
30174	36484363	QUIT PRIVATE HOUSE	THE PUBLIC TRANSPORTATION: THE TRAIN STATION I...	THE PUBLIC TRANSPORTATION: THE TRAIN STATION I...	none	QUIT QUIT QUIT !!!!!!	TRAIN STATION 5 MINUTE UBER OR 15 MINUTE WALK ...	Guest should not wear shoes, no smoking mariju...	1077
30175	36484665	Charming one bedroom - newly renovated rowhouse	This one bedroom in a large, newly renovated r...	This one bedroom in a large, newly renovated r...	none	There's an endless number of new restaurants, ...	We are three blocks from the G subway and abou...	NaN	823
30176	36485057	Affordable room in Bushwick/East Williamsburg	NaN	NaN	none	NaN	NaN	NaN	657
30177	36485609	43rd St. Time Square-cozy single bed	NaN	NaN	none	NaN	NaN	NaN	3098
30178	36487245	Trendy duplex in the very heart of Hell's Kitchen	Private room in a nice duplex in the very hear...	Private room in a nice duplex in the very hear...	none	NaN	NaN	NaN	681

30179 rows × 81 columns

In [5]:
```
listings.columns
```

Out[5]:
```
Index(['id', 'name', 'summary', 'description', 'experiences_offered',
       'neighborhood_overview', 'transit', 'house_rules', 'host_id',
       'host_since', 'host_response_time', 'host_response_rate',
       'host_is_superhost', 'host_listings_count', 'host_identity_verified',
       'street', 'neighbourhood', 'latitude', 'longitude', 'property_type',
       'room_type', 'accommodates', 'bathrooms', 'bedrooms', 'beds',
       'bed_type', 'amenities', 'price', 'guests_included', 'extra_people',
       'minimum_nights', 'calendar_updated', 'has_availability',
       'availability_30', 'availability_60', 'availability_90',
       'availability_365', 'number_of_reviews', 'number_of_reviews_ltm',
       'review_scores_rating', 'review_scores_accuracy',
       'review_scores_cleanliness', 'review_scores_checkin',
       'review_scores_communication', 'review_scores_location',
       'review_scores_value', 'instant_bookable', 'cancellation_policy',
       'calculated_host_listings_count',
       'calculated_host_listings_count_entire_homes',
       'calculated_host_listings_count_private_rooms',
       'calculated_host_listings_count_shared_rooms', 'reviews_per_month',
       'check_in_24h', 'air_conditioning', 'high_end_electronics', 'bbq',
       'balcony', 'nature_and_views', 'bed_linen', 'breakfast', 'tv',
       'coffee_machine', 'cooking_basics', 'white_goods', 'elevator', 'gym',
       'child_friendly', 'parking', 'outdoor_space', 'host_greeting',
       'hot_tub_sauna_or_pool', 'internet', 'long_term_stays', 'pets_allowed',
       'private_entrance', 'secure', 'self_check_in', 'smoking_allowed',
       'accessible', 'event_suitable'],
      dtype='object')
```

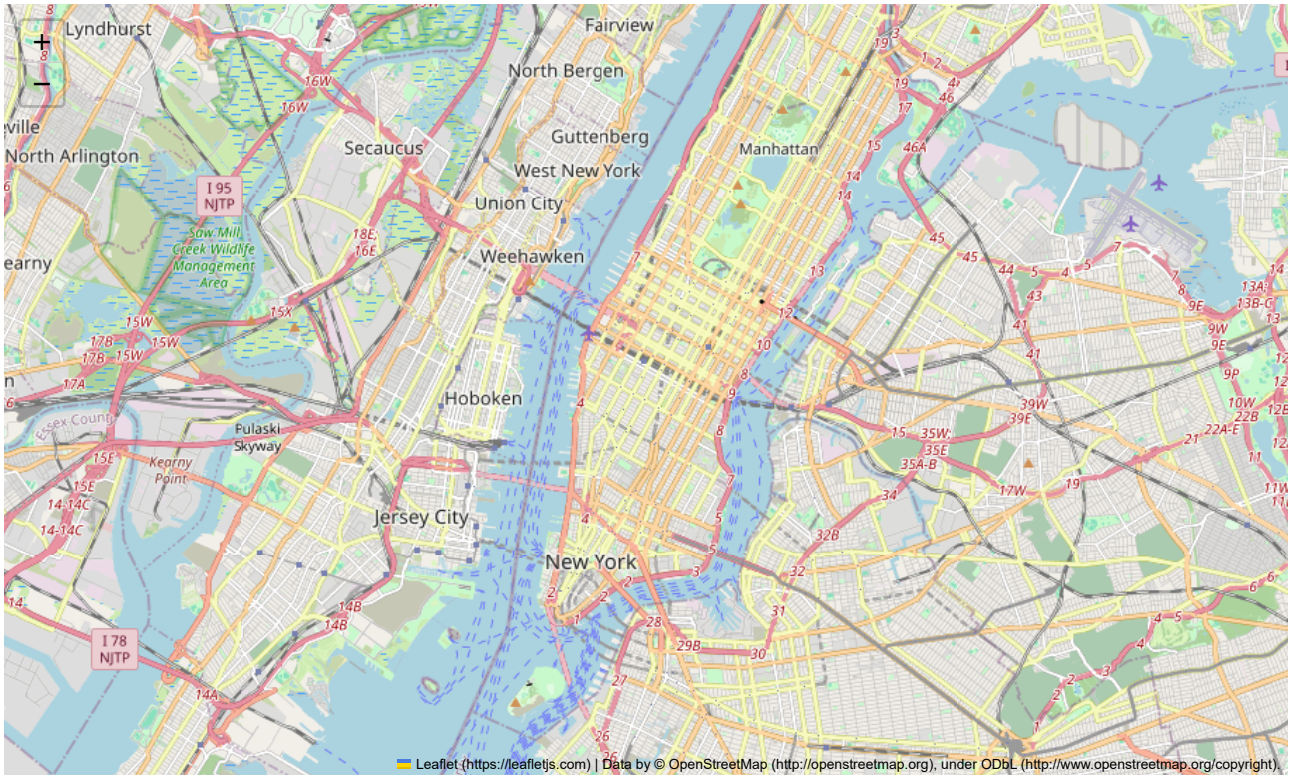
Please check out data dictionary here (https://docs.google.com/spreadsheets/d/1iWCNJcSutYqpULSQHlNyGInUvHg2BoUGoNRIGa6Szc4/edit#gid=982310896)

Activities

Q: Pick an `latitude` and `longitude` from the dataset. Using the two coordinates, create a map. See that you can zoom in and out based on the Leaflet maps.

In [6]:
```
latitude = listings.loc[9, 'latitude']
longitude = listings.loc[9, 'longitude']

my_nyc_map = folium.Map(location=[latitude, longitude], zoom_start=12)
my_nyc_map
```

Out[6]:



Q: Folium map has different styles available. Using the following `tiles` argument:

- `Stamen Toner`
- `Stamen Terrain`
- `Stamen Watercolor`
- `CartoDB positron`
- `CartoDB dark_matter`

argument, play around with the maps

In [10]:
```
my_nyc_map = folium.Map(location=[latitude, longitude], zoom_start=12, tiles = "CartoDB positron")
my_nyc_map
```

Out[10]:

In [11]:
```
my_nyc_map = folium.Map(location=[latitude, longitude], zoom_start=12,  tiles = "CartoDB dark_matter")
my_nyc_map
```

Out[11]:

Q: Using the heat maps, let's understand the price distribution with the location.

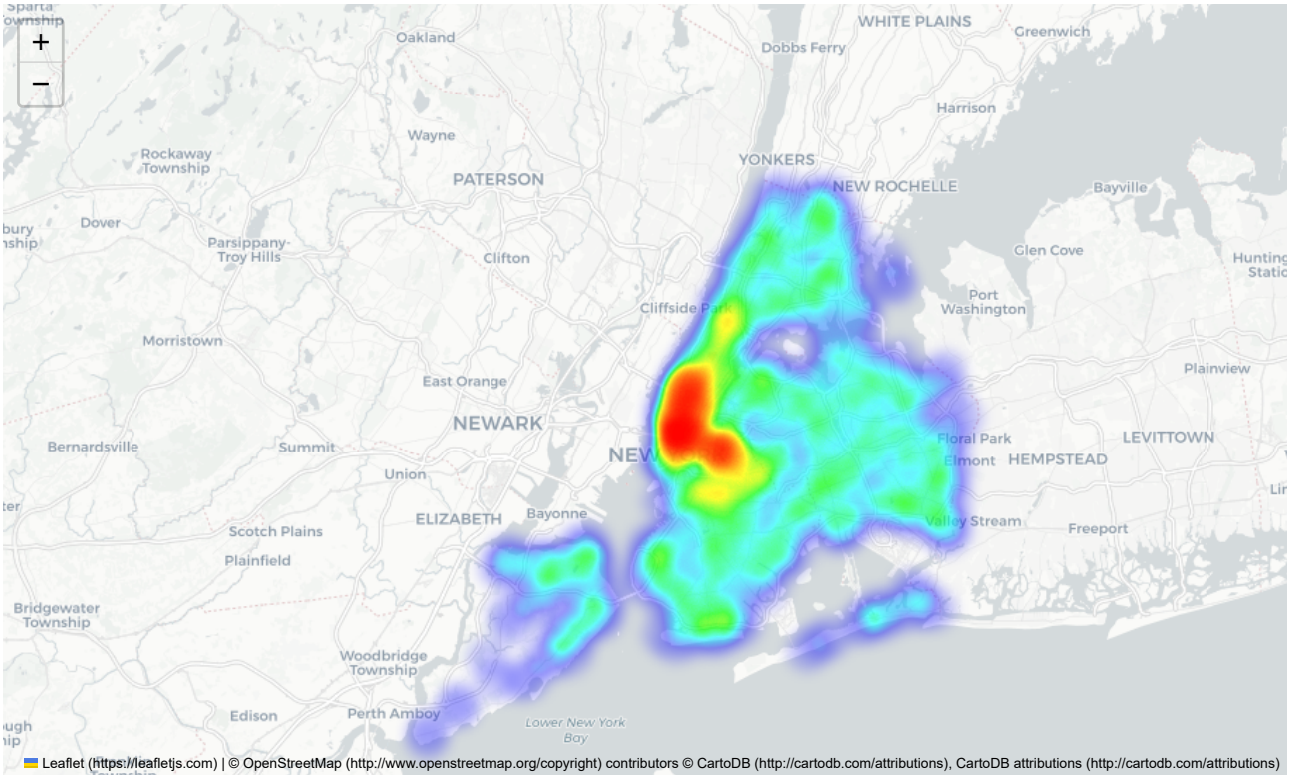
Notice that Folium has different parameters available such as:

- `min_opacity`
- `radius`
- `blur`
- and etc.

Play around with these parameters

In [12]:
```
# added new plugin
from folium import plugins
from folium.plugins import HeatMap
my_nyc_map = folium.Map(location=[latitude, longitude], zoom_start=10, tiles = "CartoDB positron")
latitude=latitude.astype(float)
longitude=longitude.astype(float)
price = listings['price'].astype(float)
listings = listings.dropna(axis=0, subset=['latitude', 'longitude', 'price'])
heat_data = listings[['latitude', 'longitude', 'price']].values.tolist()

heatmap = HeatMap(heat_data,
                  min_opacity=0.2,
                  radius=15,
                  blur=10
                 )

heatmap.add_to(my_nyc_map)
my_nyc_map
```

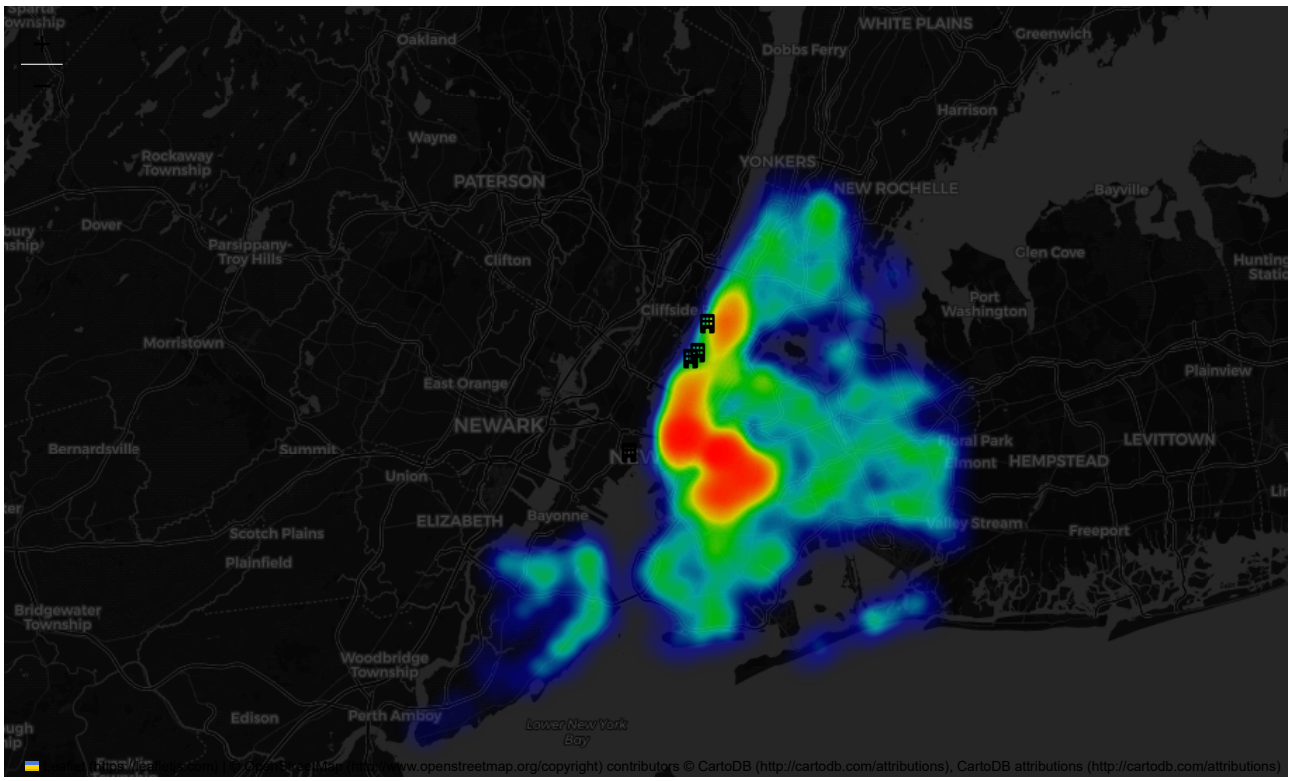
Out[12]:



Q: Using the heat maps, let's understand the `review_scores_rating` distribution with the location. What problem did you encounter? How are you solving this problem?

In [13]:
```
from folium import plugins
from folium.plugins import HeatMap
my_nyc_map = folium.Map(location=[latitude, longitude], zoom_start=10, tiles = "CartoDB dark_matter")
latitude=latitude.astype(float)
longitude=longitude.astype(float)
review_scores_rating = listings['review_scores_rating'].astype(float)
listings = listings.dropna(axis=0, subset=['latitude','longitude', 'review_scores_rating'])
heat_data = listings[['latitude', 'longitude', 'review_scores_rating']].values.tolist()

heatmap = HeatMap(heat_data,
                  min_opacity=0.2,
                  radius=15,
                  blur=10
                 )

heatmap.add_to(my_nyc_map)

list_tourist_attractions = [
    (40.785091, -73.968285, 'Central Park'),
    (40.758896, -73.985130, 'Times Square'),
    (40.689247, -74.044502, 'Statue of Liberty'),
     (40.763342, -73.978301, 'Empire State Building')
]

for latitude, longitude, popup in list_tourist_attractions:
    folium.Marker(
        location=[latitude, longitude],
        popup=popup,
        icon=folium.Icon(color='blue', icon='building', prefix='fa')
    ).add_to(my_nyc_map)

my_nyc_map
```

Out[13]:



References

"New York", Inside Airbnb, http://insideairbnb.com/get-the-data.html (http://insideairbnb.com/get-the-data.html)