

```
In [1]: from PIL import Image
from statistics import *
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

```
In [2]: import cse351_hw1_zhang_junhui_112895310
```

```
In [3]: airbnb_data = pd.read_csv("AB_NYC_2019.csv")
```

```
In [4]: # Question 1
removedAnomaly = cse351_hw1_zhang_junhui_112895310.removeAnomaly(airbnb_data);
```

```
In [5]: removedAnomaly
```

Out[5]:

	id	name	host_id	host_name	neighbourhood_group	neighbourhood	latitude	longitude	room_type	price	minimum
0	2539	Clean & quiet apt home by the park	2787	John	Brooklyn	Kensington	40.64749	-73.97237	Private room	149	
1	2595	Skylit Midtown Castle	2845	Jennifer	Manhattan	Midtown	40.75362	-73.98377	Entire home/apt	225	
2	3647	THE VILLAGE OF HARLEM....NEW YORK !	4632	Elisabeth	Manhattan	Harlem	40.80902	-73.94190	Private room	150	
3	3831	Cozy Entire Floor of Brownstone	4869	LisaRoxanne	Brooklyn	Clinton Hill	40.68514	-73.95976	Entire home/apt	89	
5	5099	Large Cozy 1 BR Apartment In Midtown East	7322	Chris	Manhattan	Murray Hill	40.74767	-73.97500	Entire home/apt	200	
...
48890	36484665	Charming one bedroom - newly renovated rowhouse	8232441	Sabrina	Brooklyn	Bedford-Stuyvesant	40.67853	-73.94995	Private room	70	
48891	36485057	Affordable room in Bushwick/East Williamsburg	6570630	Marisol	Brooklyn	Bushwick	40.70184	-73.93317	Private room	40	
48892	36485431	Sunny Studio at Historical Neighborhood	23492952	Ilgar & Aysel	Manhattan	Harlem	40.81475	-73.94867	Entire home/apt	115	
48893	36485609	43rd St. Time Square-cozy single bed	30985759	Taz	Manhattan	Hell's Kitchen	40.75751	-73.99112	Shared room	55	
48894	36487245	Trendy duplex in the very heart of Hell's Kitchen	68119814	Christophe	Manhattan	Hell's Kitchen	40.76404	-73.98933	Private room	90	

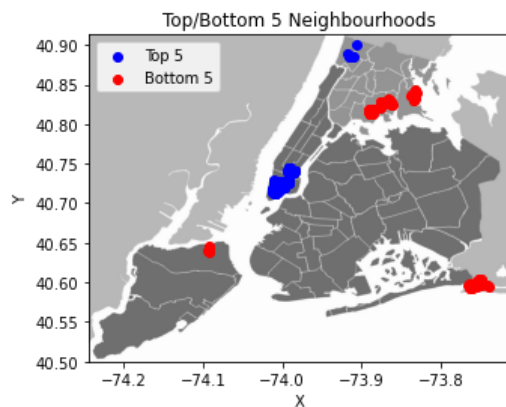
31103 rows × 16 columns

```
In [6]: # Question 2A
top5bot5List = cse351_hw1_zhang_junhui_112895310.getTop5Bot5Price(removedAnomaly)
```

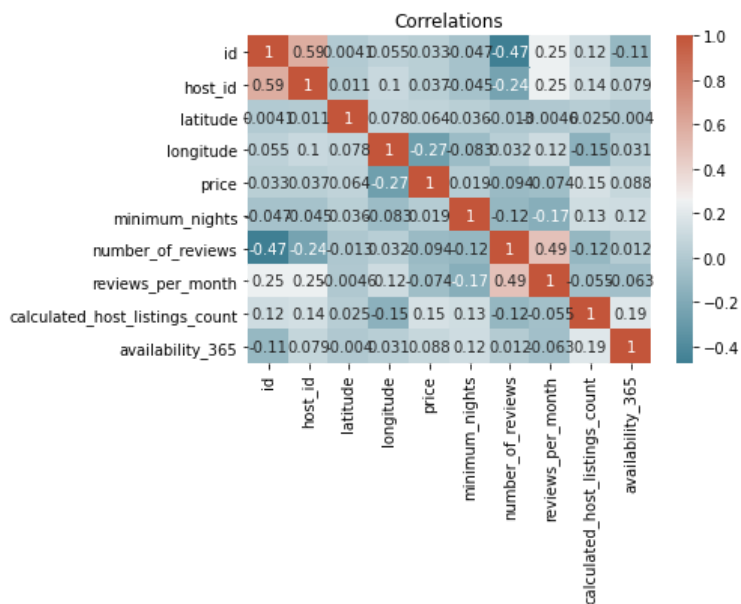
```
In [7]: top5bot5List
```

```
Out[7]: {'Top5': ['Riverdale', 'Tribeca', 'Flatiron District', 'NoHo', 'SoHo'],
        'Bot5': ['Schuylerville',
                  'Hunts Point',
                  'New Brighton',
                  'Soundview',
                  'Far Rockaway']}
```

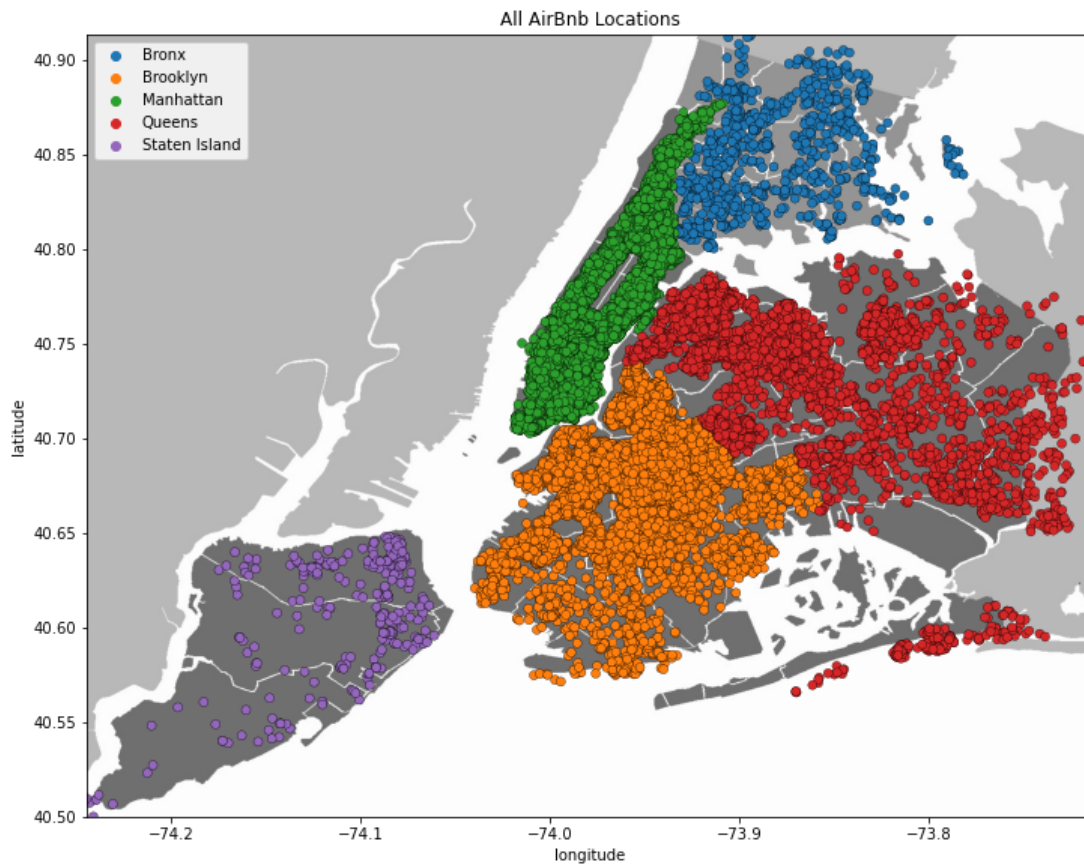
```
In [8]: # Question 2B
cse351_hw1_zhang_junhui_112895310.plotTop5Bot5(top5bot5List, removedAnomaly)
```



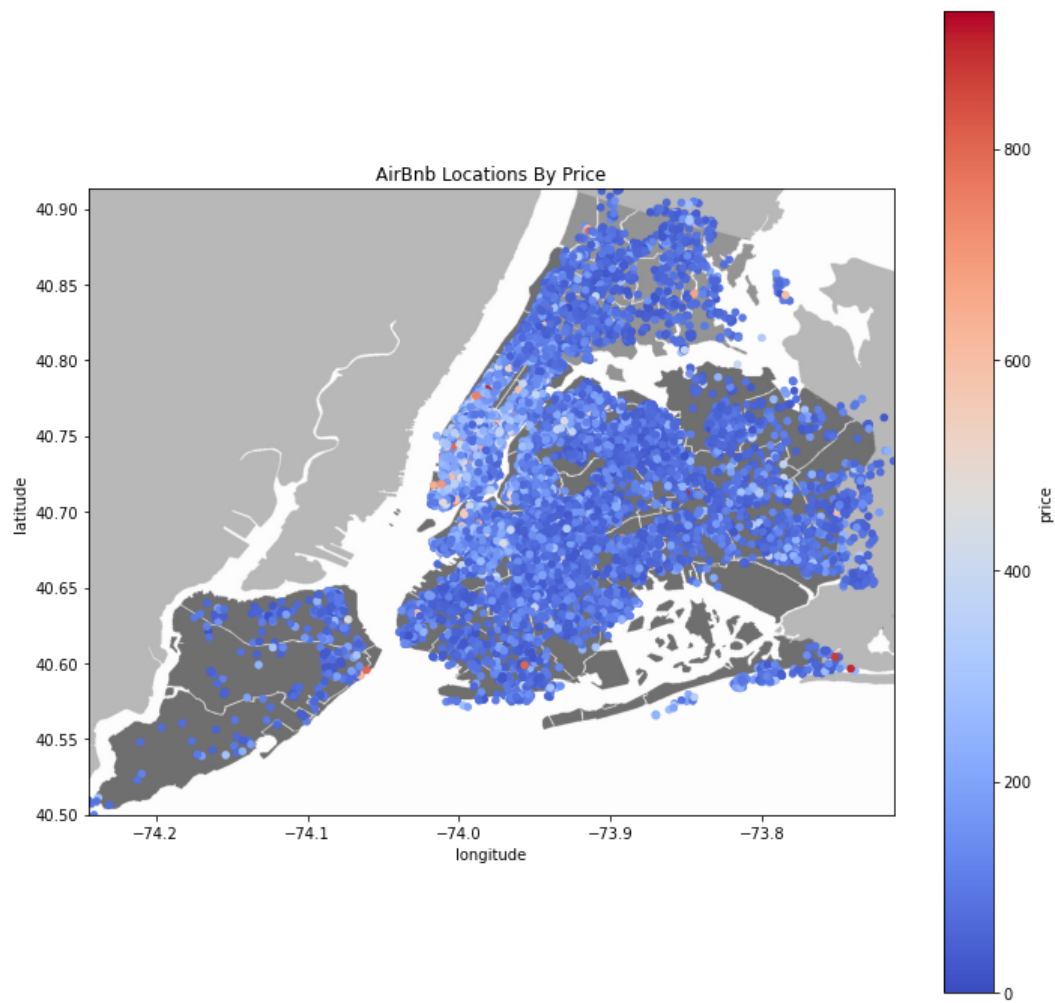
```
In [9]: # Question 3
cse351_hw1_zhang_junhui_112895310.createCorrelationHeatMap(removedAnomaly)
```



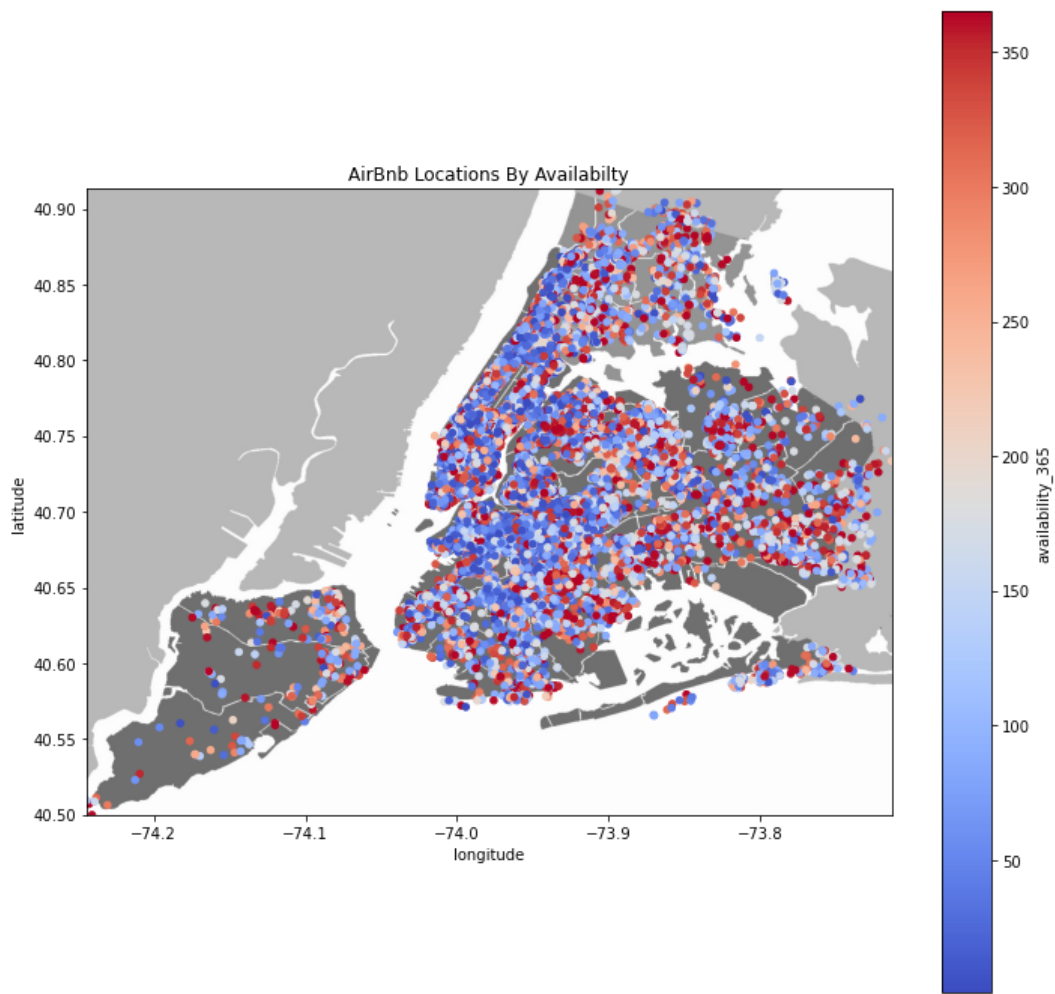
```
In [10]: # Question 4A
cse351_hw1_zhang_junhui_112895310.plotByBorough(removedAnomaly)
```



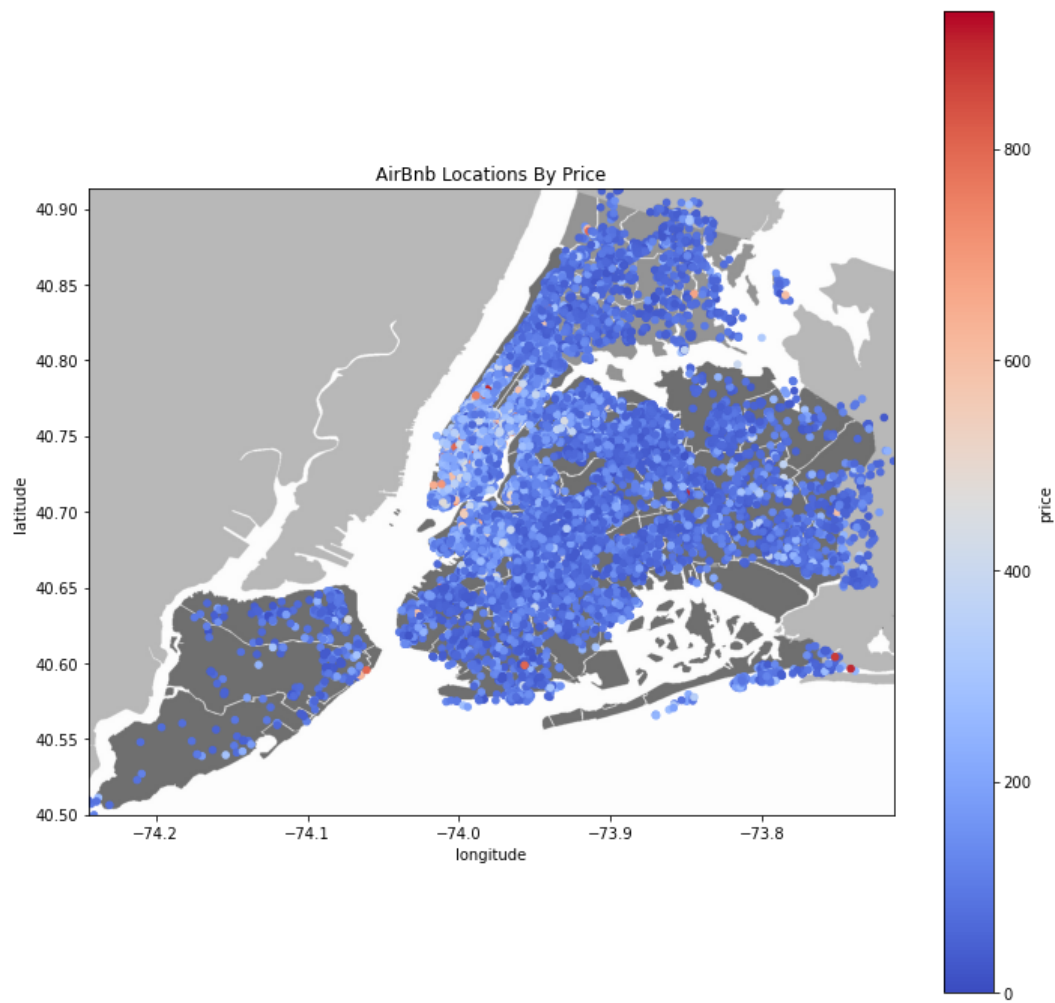
```
In [11]: # Question 4B  
cse351_hw1_zhang_junhui_112895310.generatePlotByPrice(removedAnomaly)
```



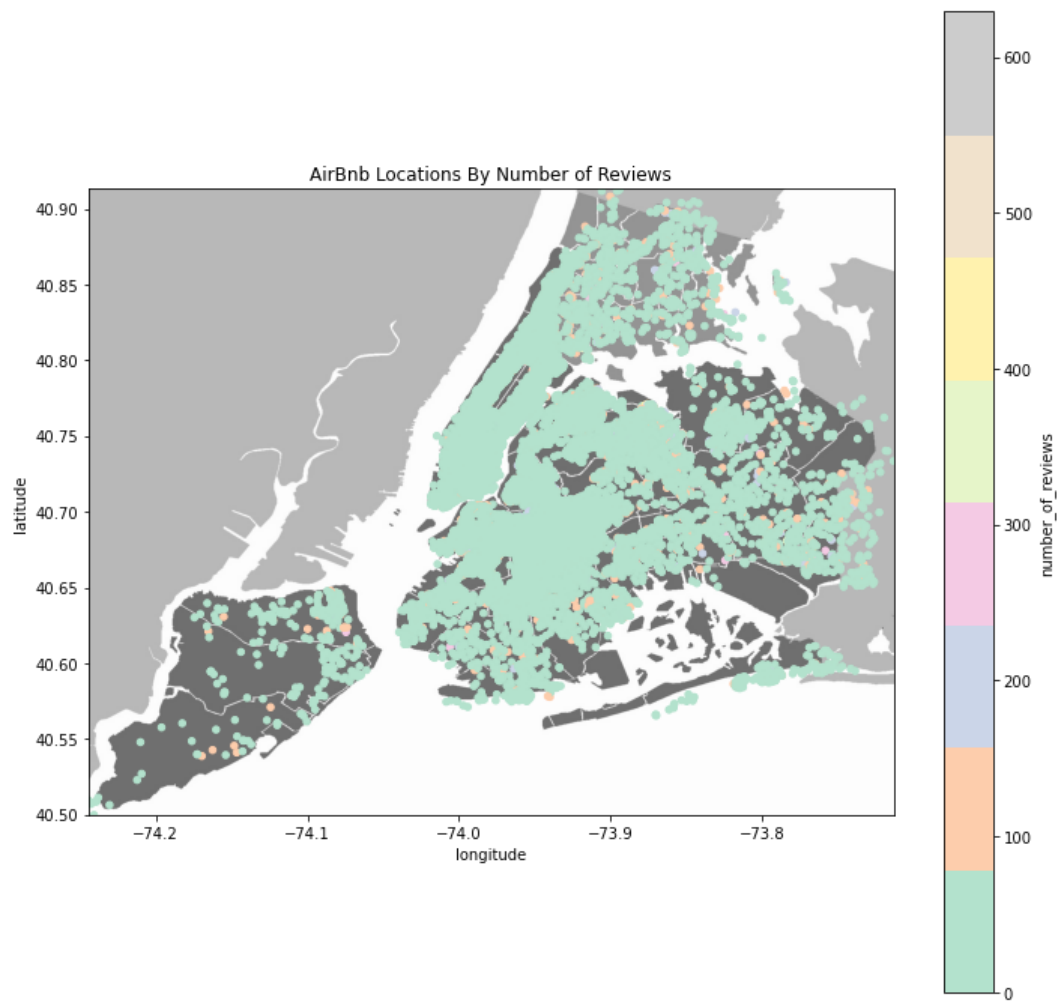
```
In [12]: # Question 5
cse351_hw1_zhang_junhui_112895310.generateWordCloud(removedAnomaly)
```

```
In [15]: cse351_hw1_zhang_junhui_112895310.generatePlotByPrice(removedAnomaly)
```



```
In [16]: cse351_hw1_zhang_junhui_112895310.generatePlotByReviews(removedAnomaly)
```



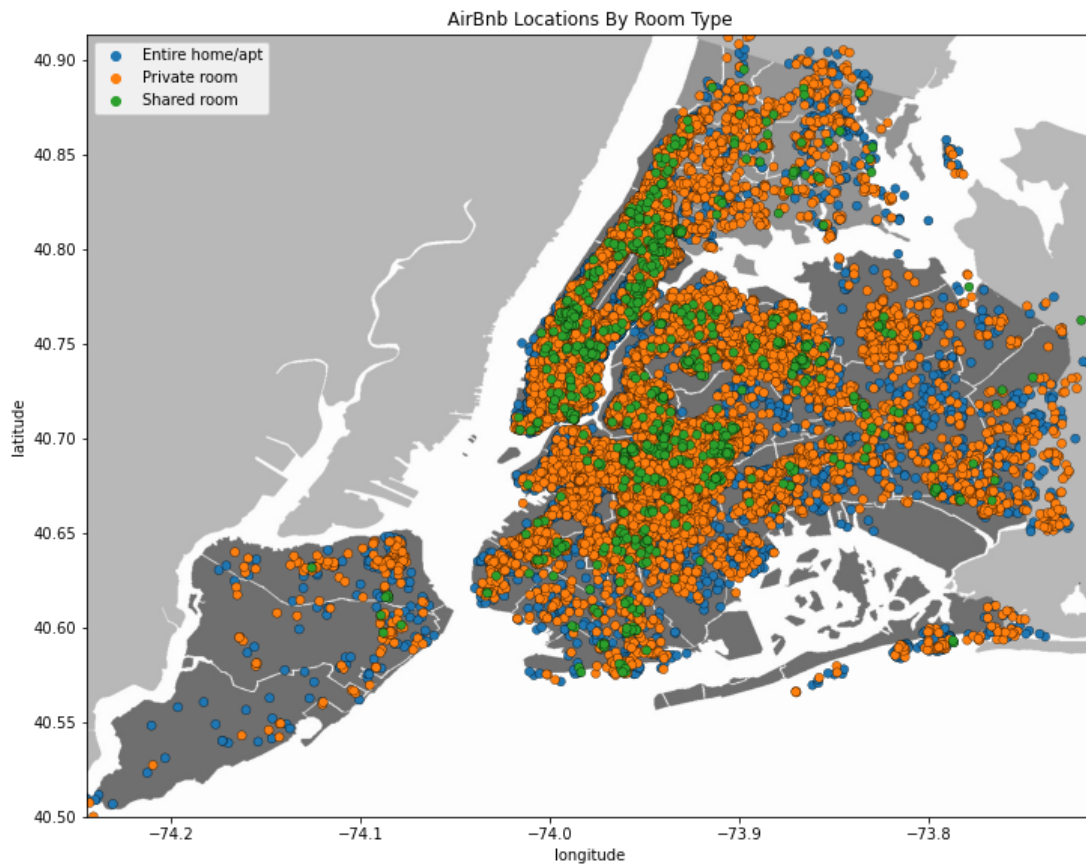
In []:

In []:

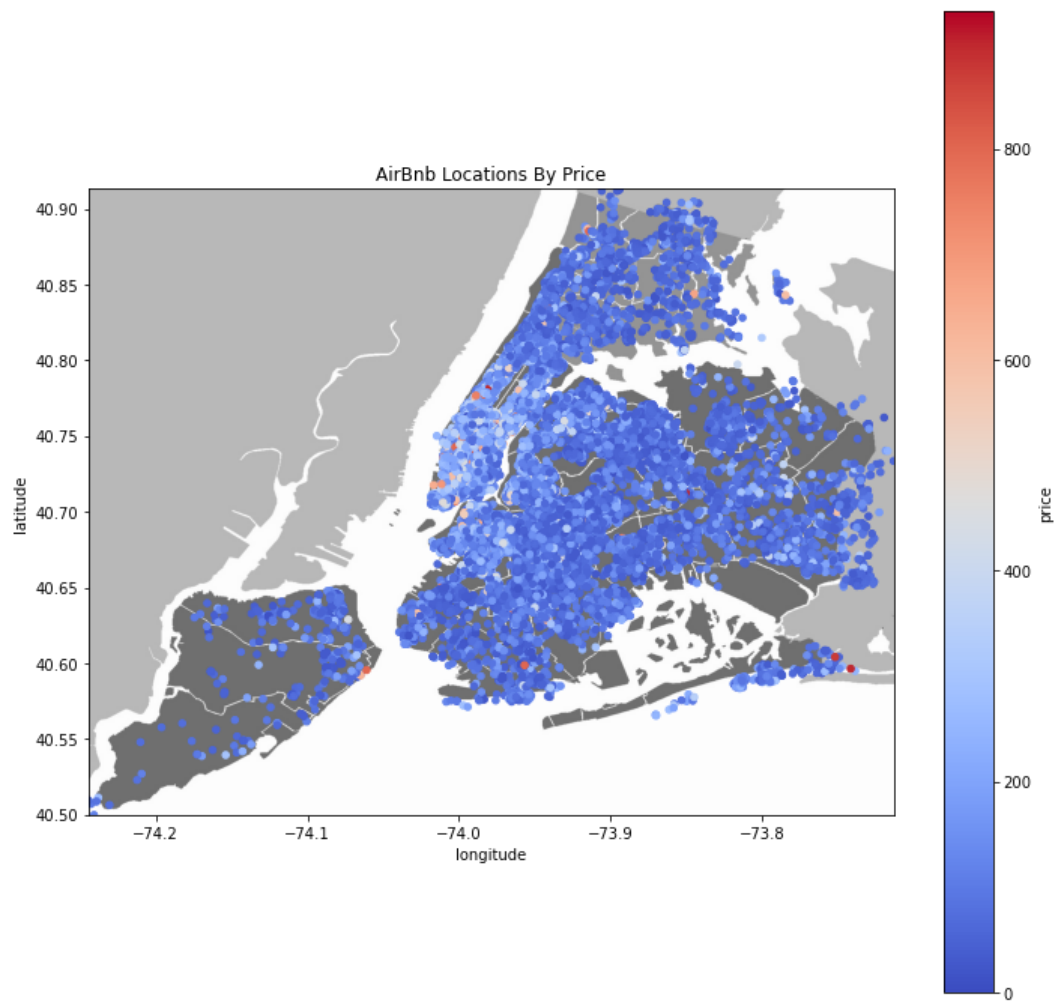
In []:

In [17]:

```
# Question 7
cse351_hw1_zhang_junhui_112895310.generatePlotByRoomType(airbnb_data)
```

In [18]: `cse351_hw1_zhang_junhui_112895310.generatePlotByPrice(removedAnomaly)`



In [19]: *# Explanations in py file*

In []: