In [1]:
```
## Course Name: DSC530-T301
## Student Name: Harish Kaparwan
## Assignment # : Quiz#2
## Professor Name: Dr. Parajulee
```

In [2]:
```
import pandas as pd
import seaborn as sns
import numpy as np
import matplotlib.pyplot as plt
import matplotlib

data = pd.read_csv('Quiz2_data.csv')
print("\n Rows=",data.shape[0])
print("\n Columns=",data.shape[1])
print("\n Feature=\n\n",data.columns.tolist())
print("\n Missing Value\n\n ",pd.isnull(data).sum().sum())
print("\n Column with Missing Value\n\n ",pd.isnull(data).any())

# Delete unwanted column
data.drop(["id","name", "host_name", "last_review"],axis=1 ,inplace=True )
print(data.head())

data.reviews_per_month.fillna(0, inplace=True)
print("\n After Replacing NAN value with '0' \n\n")
print(data)
```

Rows= 48895

Columns= 16

Feature=

['id', 'name', 'host_id', 'host_name', 'neighbourhood_group', 'neighbourhood', 'latitude', 'longitude', 'room_type', 'price', 'minimum_nights', 'number_of_reviews', 'last_review', 'reviews_per_month', 'calculated_host_listings_count', 'availability_365']

Missing Value

20141

Column with Missing Value

```
id                                False
name                               True
host_id                           False
host_name                          True
neighbourhood_group               False
neighbourhood                     False
latitude                          False
longitude                         False
room_type                         False
price                             False
minimum_nights                    False
number_of_reviews                 False
last_review                        True
reviews_per_month                  True
calculated_host_listings_count    False
availability_365                  False
dtype: bool
```

	host_id	neighbourhood_group	neighbourhood	latitude	longitude	\
0	2787	Brooklyn	Kensington	40.64749	-73.97237	
1	2845	Manhattan	Midtown	40.75362	-73.98377	
2	4632	Manhattan	Harlem	40.80902	-73.94190	
3	4869	Brooklyn	Clinton Hill	40.68514	-73.95976	
4	7192	Manhattan	East Harlem	40.79851	-73.94399	

	room_type	price	minimum_nights	number_of_reviews	\
0	Private room	149	1	9	
1	Entire home/apt	225	1	45	
2	Private room	150	3	0	
3	Entire home/apt	89	1	270	
4	Entire home/apt	80	10	9	

	reviews_per_month	calculated_host_listings_count	availability_365
0	0.21	6	365
1	0.38	2	355
2	NaN	1	365
3	4.64	1	194
4	0.10	1	0

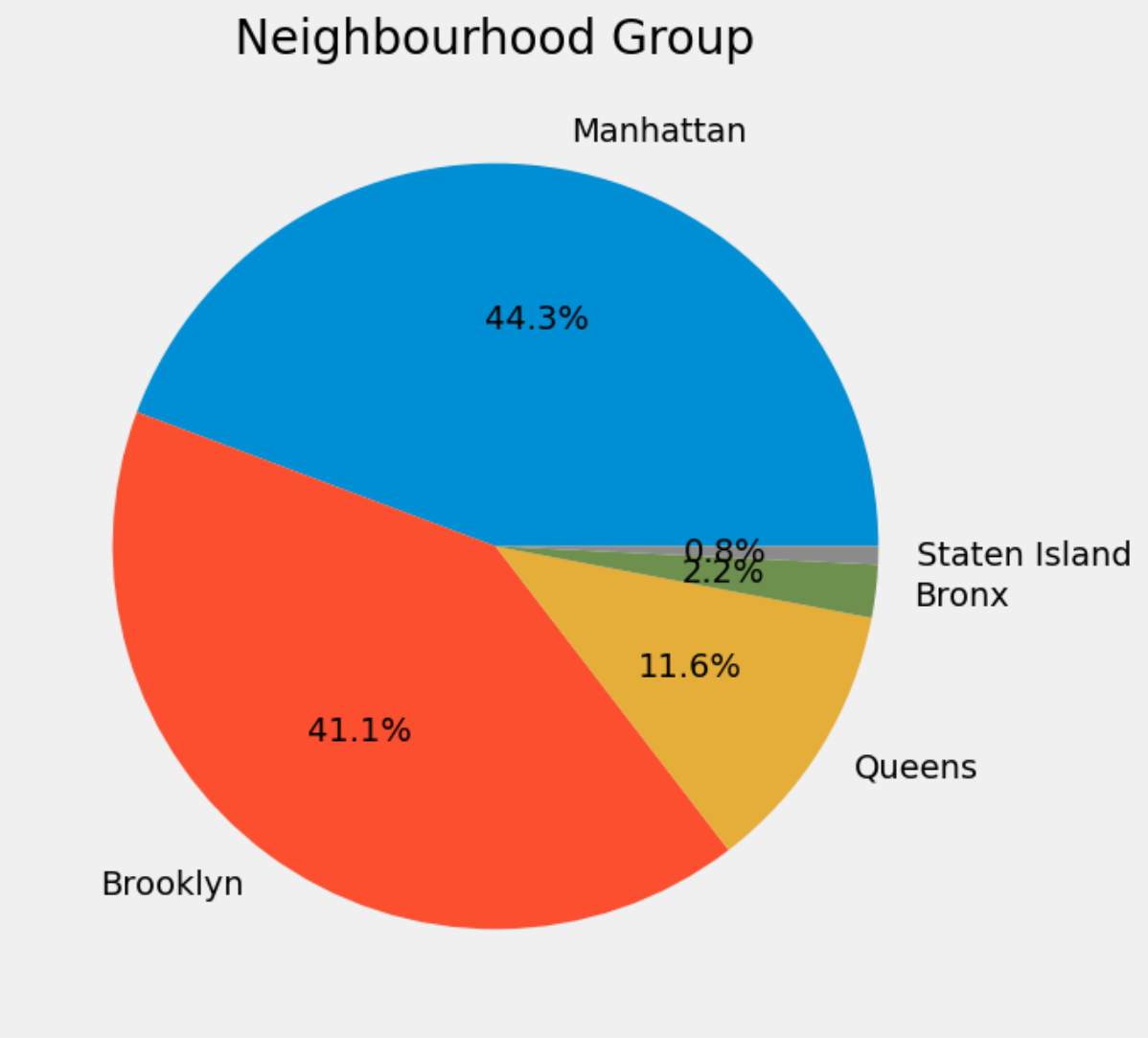
After Replacing NAN value with '0'

	host_id	neighbourhood_group	neighbourhood	latitude	longitude
\					
0	2787	Brooklyn	Kensington	40.64749	-73.97237
1	2845	Manhattan	Midtown	40.75362	-73.98377
2	4632	Manhattan	Harlem	40.80902	-73.94190
3	4869	Brooklyn	Clinton Hill	40.68514	-73.95976

```
4          7192           Manhattan         East Harlem  40.79851  -73.94399
...         ...                 ...                 ...       ...        ...
48890   8232441            Brooklyn  Bedford-Stuyvesant  40.67853  -73.94995
48891   6570630            Brooklyn            Bushwick  40.70184  -73.93317
48892  23492952           Manhattan              Harlem  40.81475  -73.94867
48893  30985759           Manhattan      Hell's Kitchen  40.75751  -73.99112
48894  68119814           Manhattan      Hell's Kitchen  40.76404  -73.98933

             room_type  price  minimum_nights  number_of_reviews  \
0         Private room    149               1                  9
1      Entire home/apt    225               1                 45
2         Private room    150               3                  0
3      Entire home/apt     89               1                270
4      Entire home/apt     80              10                  9
...                ...    ...             ...                ...
48890     Private room     70               2                  0
48891     Private room     40               4                  0
48892  Entire home/apt    115              10                  0
48893      Shared room     55               1                  0
48894     Private room     90               7                  0

       reviews_per_month  calculated_host_listings_count  availability_365
0                   0.21                               6               365
1                   0.38                               2               355
2                   0.00                               1               365
3                   4.64                               1               194
4                   0.10                               1                 0
...                  ...                             ...               ...
48890               0.00                               2                 9
48891               0.00                               2                36
48892               0.00                               1                27
48893               0.00                               6                 2
48894               0.00                               1                23

[48895 rows x 12 columns]
```

In [3]:

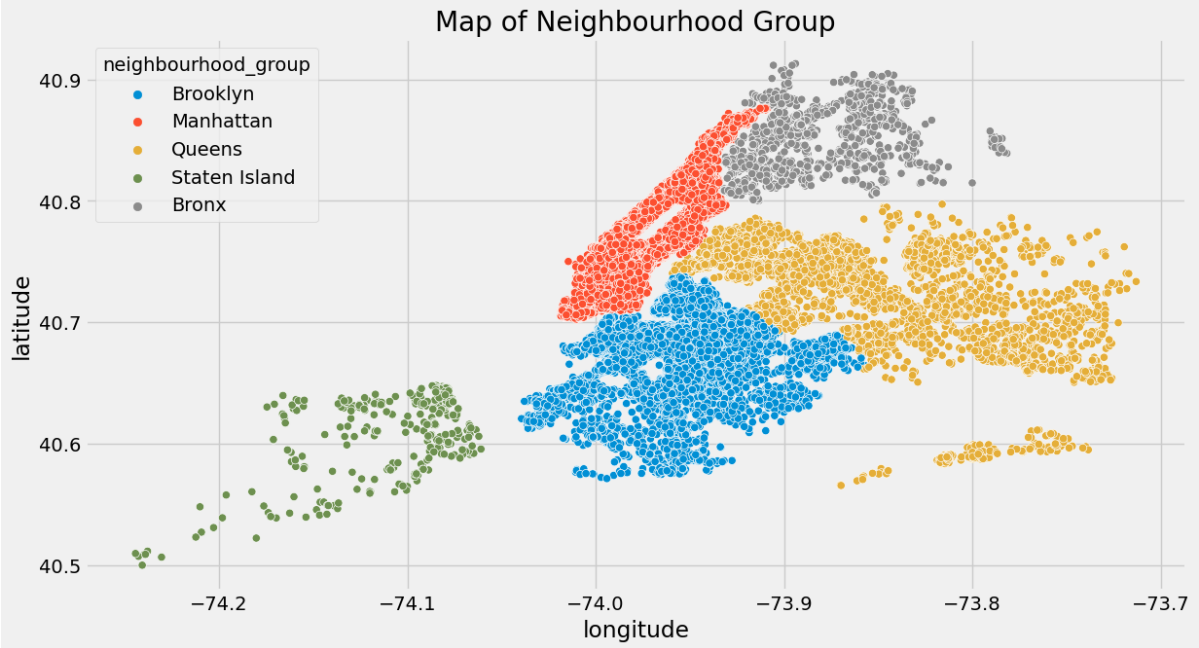
```
# Exploring Data
# Let's explore the data interms of neighbourhood_group
plt.style.use('fivethirtyeight')
plt.figure(figsize=(13,7))
plt.title('Neighbourhood Group')
plt.pie(data.neighbourhood_group.value_counts(), labels = data.neighbourhood
plt.show()
```



The Pie Chart Above shows that Airbnb Listings in newyork are near Mahattan and Brooklyn has highest share of hotels We also know that from this map of neighbourhood Group

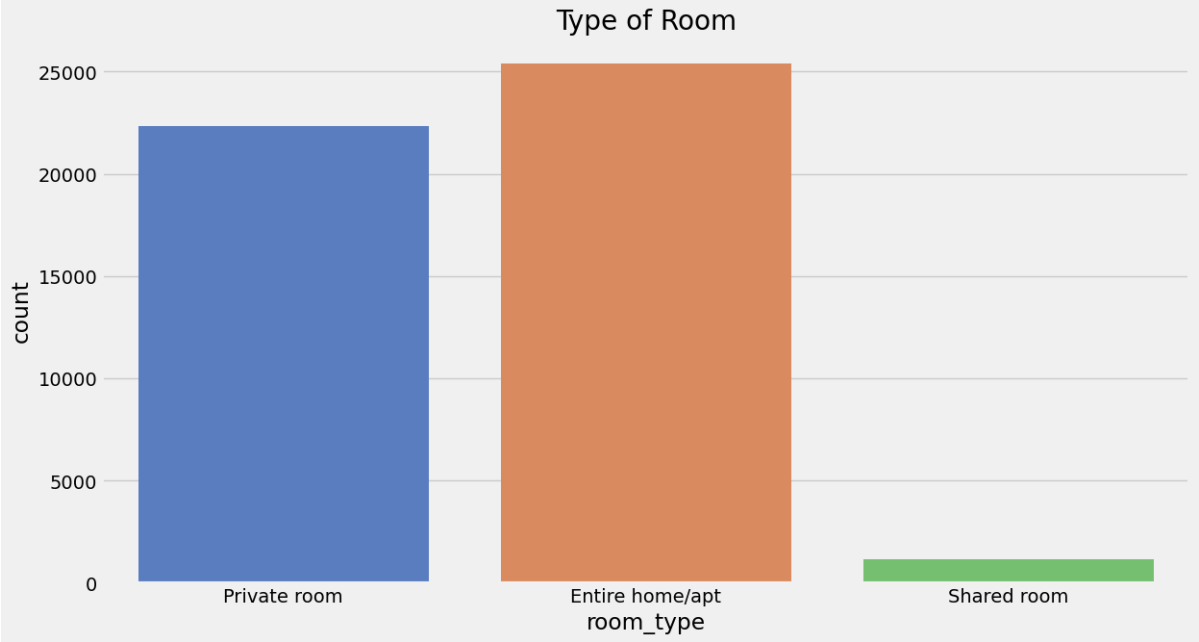
In [4]:
```
#Map of neighbourhood group
plt.figure(figsize=(13,7))
plt.title("Map of Neighbourhood Group")
sns.scatterplot(data.longitude,data.latitude,hue=data.neighbourhood_group)
plt.ioff()
plt.show()
```

```
/opt/anaconda3/lib/python3.8/site-packages/seaborn/_decorators.py:36: Future
Warning: Pass the following variables as keyword args: x, y. From version 0.
12, the only valid positional argument will be `data`, and passing other arg
uments without an explicit keyword will result in an error or misinterpretat
ion.
  warnings.warn(
```



In [5]:
```
# Let's see the types of room
plt.figure(figsize = (13,7))
plt.title("Type of Room")
sns.countplot(data.room_type,palette = "muted")
fig = plt.gcf()
plt.show()
```

```
/opt/anaconda3/lib/python3.8/site-packages/seaborn/_decorators.py:36: FutureWarning: Pass the following variable as a keyword arg: x. From version 0.12, the only valid positional argument will be `data`, and passing other arguments without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
```



In [6]:
```
#Let's see room types occupied by a neighbourhood group
plt.figure(figsize = (13,7))
plt.title("Room Type on Neighbourhood Group")
sns.countplot(data.neighbourhood_group,hue=data.room_type,palette = "muted")
plt.show()
```

```
/opt/anaconda3/lib/python3.8/site-packages/seaborn/_decorators.py:36: Future
Warning: Pass the following variable as a keyword arg: x. From version 0.12,
the only valid positional argument will be `data`, and passing other argumen
ts without an explicit keyword will result in an error or misinterpretation.
  warnings.warn(
```