

AIRBNB VISUALIZATION AND ANALYSIS/ Retno Prabaningrum

(retnoprabaningrum81@gmail.com)

1. Import Required Libraries

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
import seaborn as sns
```

2. Load Dataset

```
data= pd.read_csv('/content/AB_NYC_2019.csv')
data.head(5)
```

	id	name	host_id	host_name	neighbourhood_group	neighbourhood	1.
0	2539	Clean & quiet apt home by the park	2787	John	Brooklyn	Kensington	4
1	2595	Skylit Midtown Castle	2845	Jennifer	Manhattan	Midtown	4
2	3647	THE VILLAGE OF HARLEM....NEW YORK !	4632	Elisabeth	Manhattan	Harlem	4
3	3831	Cozy Entire Floor of Brownstone	4869	LisaRoxanne	Brooklyn	Clinton Hill	4
4	5022	Entire Apt: Spacious Studio/Loft by central park	7192	Laura	Manhattan	East Harlem	4



```
data.shape
```

```
(22059, 16)
```

```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
```

RangeIndex: 22059 entries, 0 to 22058

Data columns (total 16 columns):

#	Column	Non-Null Count	Dtype
0	id	22059 non-null	int64
1	name	22044 non-null	object
2	host_id	22059 non-null	int64
3	host_name	22045 non-null	object
4	neighbourhood_group	22059 non-null	object
5	neighbourhood	22059 non-null	object
6	latitude	22059 non-null	float64
7	longitude	22059 non-null	float64
8	room_type	22058 non-null	object
9	price	22058 non-null	float64
10	minimum_nights	22058 non-null	float64
11	number_of_reviews	22058 non-null	float64
12	last_review	18358 non-null	object
13	reviews_per_month	18358 non-null	float64
14	calculated_host_listings_count	22058 non-null	float64
15	availability_365	22058 non-null	float64

dtypes: float64(8), int64(2), object(6)

memory usage: 2.7+ MB

3. Remove the Duplicates Data (if any) and Check the Null Values

```
data.duplicated().sum()
```

```
data.drop_duplicates(inplace=True)
```

```
data.isnull().sum()
```

id	0
name	15
host_id	0
host_name	14
neighbourhood_group	0
neighbourhood	0
latitude	0
longitude	0
room_type	1
price	1
minimum_nights	1
number_of_reviews	1
last_review	3701
reviews_per_month	3701
calculated_host_listings_count	1
availability_365	1
dtype: int64	

4. Check the Data Changes

```
data.head(5)
```

	id	name	host_id	host_name	neighbourhood_group	neighbourhood	1.
0	2539	Clean & quiet apt home by the park	2787	John	Brooklyn	Kensington	4
1	2595	Skylit Midtown Castle	2845	Jennifer	Manhattan	Midtown	4
2	3647	THE VILLAGE OF HARLEM....NEW YORK !	4632	Elisabeth	Manhattan	Harlem	4
3	3831	Cozy Entire Floor of Brownstone	4869	LisaRoxanne	Brooklyn	Clinton Hill	4
4	5022	Entire Apt: Spacious Studio/Loft by central park	7192	Laura	Manhattan	East Harlem	4

5. Check the Continous Variables

```
data.describe()
```

	id	host_id	latitude	longitude	price	minimum_r
count	2.205900e+04	2.205900e+04	22059.000000	22059.000000	22058.000000	22058.0
mean	8.511062e+06	2.447159e+07	40.729520	-73.956852	151.245852	7.4
std	5.258526e+06	2.707545e+07	0.053118	0.039701	235.534537	23.8
min	2.539000e+03	2.571000e+03	40.499790	-74.242850	10.000000	1.0
25%	3.996390e+06	3.977023e+06	40.689765	-73.983265	73.000000	2.0
50%	8.520836e+06	1.428111e+07	40.723370	-73.958010	110.000000	3.0
75%	1.324998e+07	3.657868e+07	40.763950	-73.941490	175.000000	5.0
max	1.775024e+07	1.210969e+08	40.908040	-73.000000	10000.000000	1250.0



6. Print Columns Names

```
data.columns
```

```
Index(['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'],
dtype='object')

```

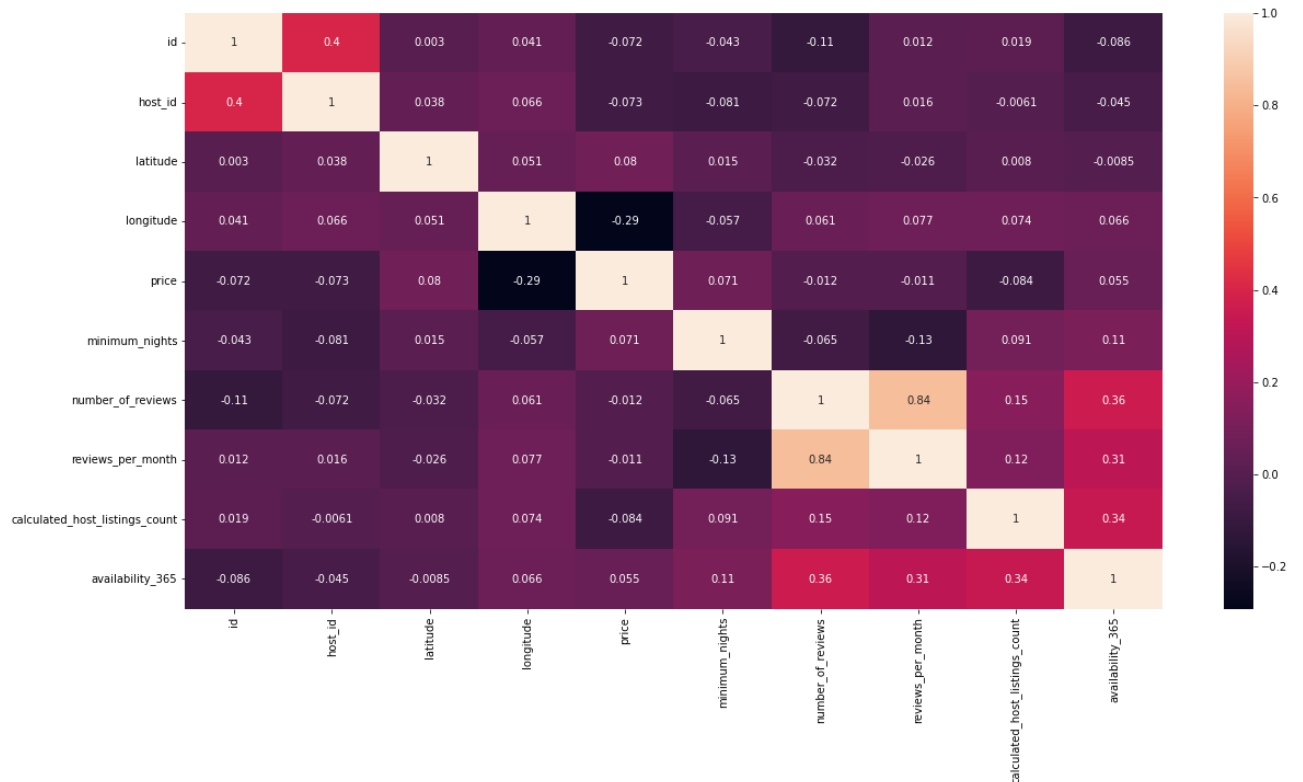
Get Correlation

```

corr = data.corr('kendall')
plt.figure(figsize=(20,10))
sns.heatmap(corr, annot=True)

```

<matplotlib.axes._subplots.AxesSubplot at 0x7f5e31207d50>



7. Data Visualization

```
data['neighbourhood_group'].unique()
```

```

array(['Brooklyn', 'Manhattan', 'Queens', 'Staten Island', 'Bronx'],
      dtype=object)

```

```
print(data.corr)
```

```
<bound method DataFrame.corr of                                     id
0          2539          Clean & quiet apt home by the park          2787
1          2595          Skylit Midtown Castle          2845
2          3647          THE VILLAGE OF HARLEM....NEW YORK !          4632
3          3831          Cozy Entire Floor of Brownstone          4869
4          5022  Entire Apt: Spacious Studio/Loft by central park          7192
...          ...          ...          ...
22054  17744351          1 Bd furnished Bayridge Brooklyn NY  119029523
22055  17746222  Perfect bedroom. Near Subways Columbia CityCol...  16721721
22056  17748284  Guest BedRm in quaint NYC neighborhood. NEAR T...  25196982
22057  17749273          1 br Woodside - close to LGA!  121096858
22058  17750238          Great room in a spacious Inwood Apartment  13860679
```

```
          host_name  neighbourhood_group  neighbourhood  latitude  longitude \
0          John          Brooklyn          Kensington  40.64749  -73.97237
1        Jennifer          Manhattan          Midtown  40.75362  -73.98377
2        Elisabeth          Manhattan          Harlem  40.80902  -73.94190
3      LisaRoxanne          Brooklyn          Clinton Hill  40.68514  -73.95976
4          Laura          Manhattan          East Harlem  40.79851  -73.94399
...          ...          ...          ...          ...
22054          Ebada          Brooklyn          Fort Hamilton  40.62241  -74.02863
22055        Federico          Manhattan          Harlem  40.81530  -73.95080
22056        Nikolas          Queens          Astoria  40.76924  -73.91702
22057        Caressa          Queens          Woodside  40.74176  -73.90631
22058          Max          Manhattan          Inwood  40.86348  -73.00000
```

```
          room_type  price  minimum_nights  number_of_reviews  last_review \
0      Private room  149.0          1.0          9.0  2018-10-19
1  Entire home/apt  225.0          1.0         45.0  2019-05-21
2      Private room  150.0          3.0          0.0          NaN
3  Entire home/apt   89.0          1.0        270.0  2019-07-05
4  Entire home/apt   80.0         10.0          9.0  2018-11-19
...          ...          ...          ...          ...
22054  Entire home/apt  115.0          4.0          3.0  2019-05-19
22055      Private room   65.0          2.0         18.0  2018-11-04
22056      Private room   65.0          1.0         11.0  2019-06-17
22057      Private room   35.0          1.0          0.0          NaN
22058          NaN          NaN          NaN          NaN          NaN
```

```
          reviews_per_month  calculated_host_listings_count  availability_365
0              0.21              6.0              365.0
1              0.38              2.0              355.0
2              NaN              1.0              365.0
3              4.64              1.0              194.0
4              0.10              1.0               0.0
...          ...          ...          ...
22054              0.48              3.0              295.0
22055              0.64              1.0               0.0
22056              3.24              1.0               9.0
22057              NaN              1.0               0.0
22058              NaN              NaN              NaN
```

```
[22059 rows x 16 columns]>
```



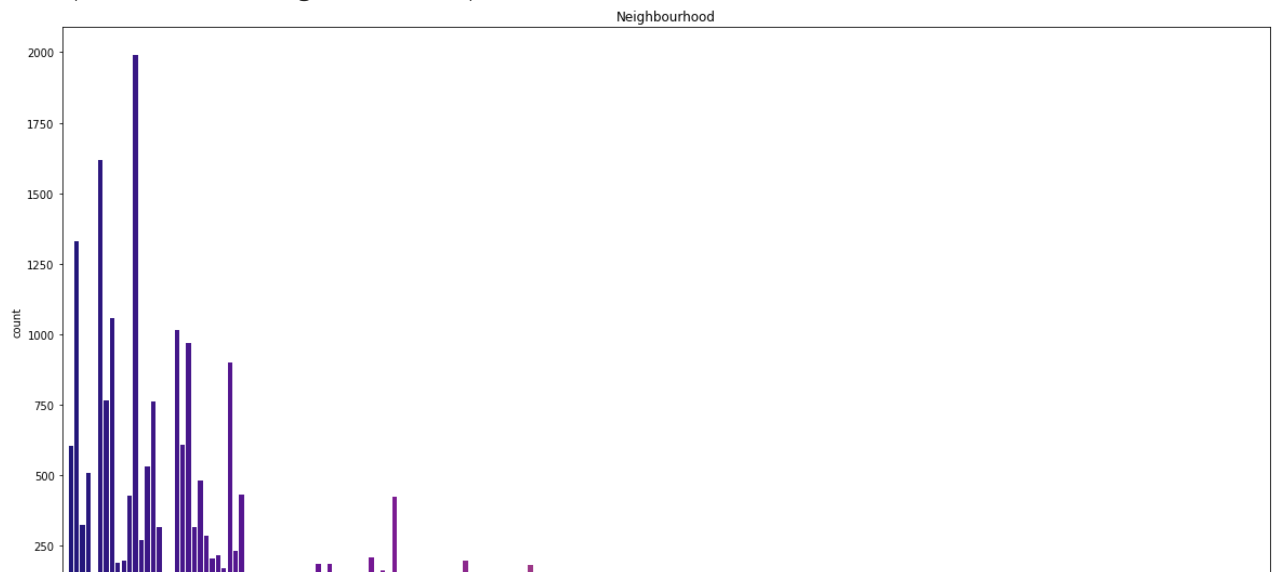
```
data.head(10)
```

	id	name	host_id	host_name	neighbourhood_group	neighbourhood	1.
0	2539	Clean & quiet apt home by the park	2787	John	Brooklyn	Kensington	4
1	2595	Skylit Midtown Castle	2845	Jennifer	Manhattan	Midtown	4
2	3647	THE VILLAGE OF HARLEM....NEW YORK !	4632	Elisabeth	Manhattan	Harlem	4
3	3831	Cozy Entire Floor of Brownstone	4869	LisaRoxanne	Brooklyn	Clinton Hill	4
4	5022	Entire Apt: Spacious Studio/Loft by central park	7192	Laura	Manhattan	East Harlem	4
5	5099	Large Cozy 1 BR Apartment In Midtown East	7322	Chris	Manhattan	Murray Hill	4
6	5121	BlissArtsSpace!	7356	Garon	Brooklyn	Bedford- Stuyvesant	4
7	5178	Large Furnished Room Near B'way	8967	Shunichi	Manhattan	Hell's Kitchen	4
8	5203	Cozy Clean Guest Room - Family Apt	7490	MaryEllen	Manhattan	Upper West Side	4
9	5238	Cute & Cozy Lower East Side 1 bdrm	7549	Ben	Manhattan	Chinatown	4



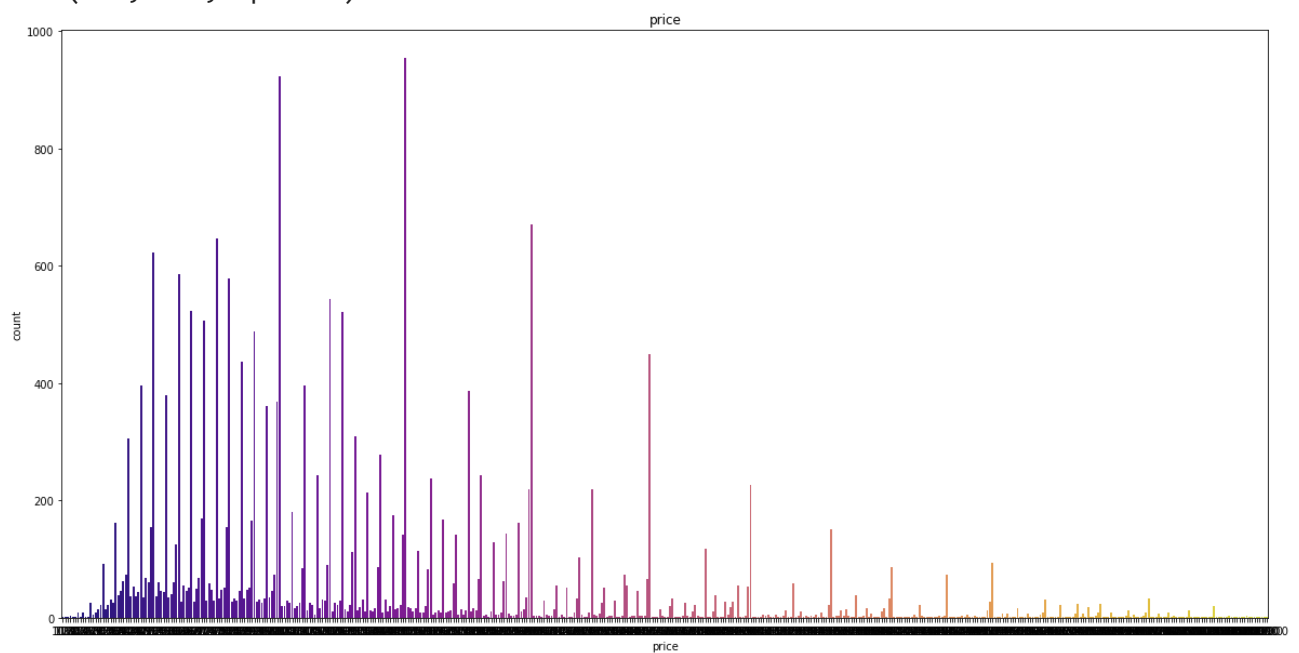
```
sns.countplot(data['neighbourhood'], palette="plasma")
fig = plt.gcf()
fig.set_size_inches(20, 10)
plt.title('Neighbourhood')
```

```
/usr/local/lib/python3.7/dist-packages/seaborn/_decorators.py:43: FutureWarning: Pas
FutureWarning
Text(0.5, 1.0, 'Neighbourhood')
```



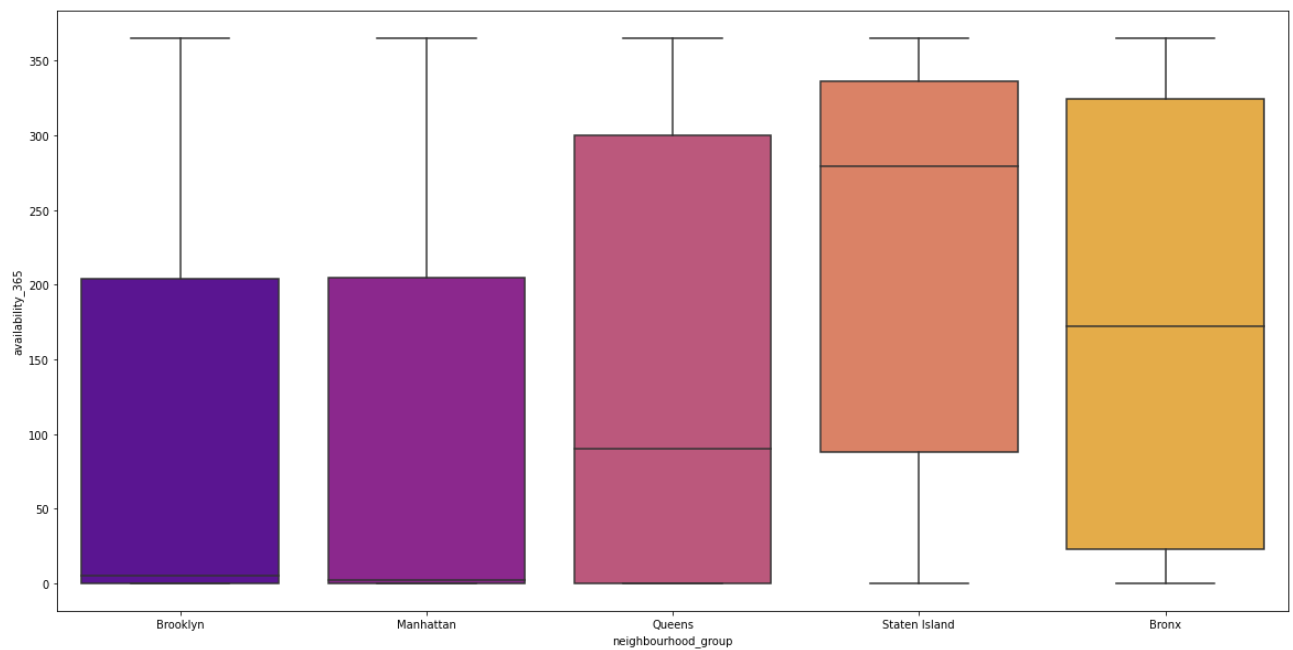
```
sns.countplot(data['price'], palette="plasma")
fig = plt.gcf()
fig.set_size_inches(20, 10)
plt.title('price')
```

```
/usr/local/lib/python3.7/dist-packages/seaborn/_decorators.py:43: FutureWarning: Pas
FutureWarning
Text(0.5, 1.0, 'price')
```



8. Check the Relation Between Neighborhood and Availability Room

```
plt.figure(figsize=(20, 10))
ax = sns.boxplot(data=data, x='neighbourhood_group', y='availability_365', palette='plasma')
```



9. Map of Availability Room

```
plt.figure(figsize=(10,5))
sns.scatterplot(data.longitude,data.latitude, hue=data.availability_365)
plt.ioff()
```


/usr/local/lib/python3.7/dist-packages/seaborn/_decorators.py:43: FutureWarning: Pas
FutureWarning

data.values

```
array([[2539, 'Clean & quiet apt home by the park', 2787, ..., 0.21, 6.0,  
       365.0],  
       [2595, 'Skylit Midtown Castle', 2845, ..., 0.38, 2.0, 355.0],  
       [3647, 'THE VILLAGE OF HARLEM....NEW YORK !', 4632, ..., nan, 1.0,  
       365.0],  
       ...,  
       [17748284, 'Guest BedRm in quaint NYC neighborhood. NEAR TRAIN',  
       25196982, ..., 3.24, 1.0, 9.0],  
       [17749273, '1 br Woodside - close to LGA!', 121096858, ..., nan,  
       1.0, 0.0],  
       [17750238, 'Great room in a spacious Inwood Apartment', 13860679,  
       ..., nan, nan, nan]], dtype=object)
```

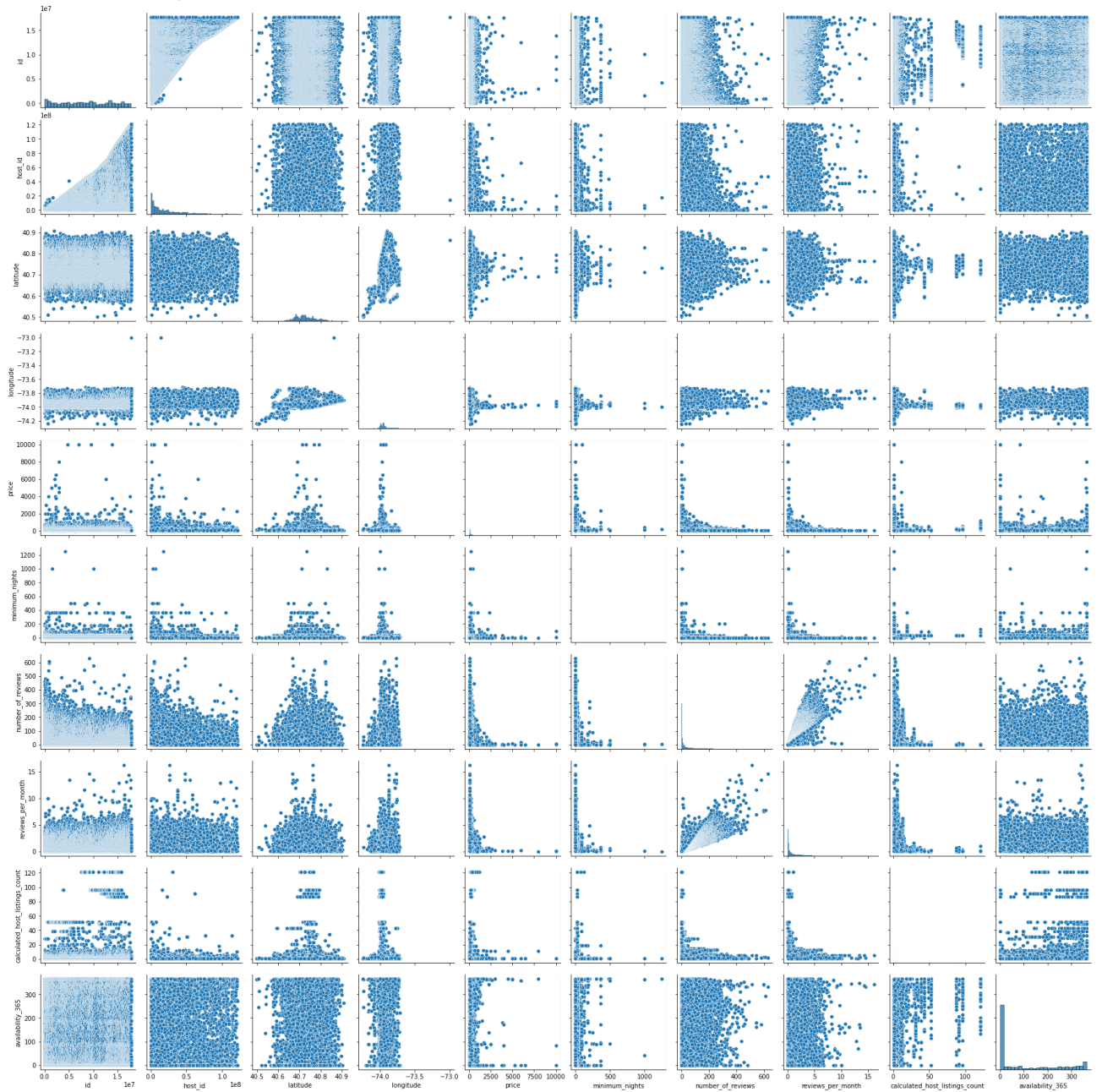
T **B** **I** <>       ψ  

10. Regression Analyst

10. Regression Analyst

sns.pairplot(data)

<seaborn.axisgrid.PairGrid at 0x7f5e29fb2d10>



! 0s completed at 5:19 PM

● ✕