

```

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

airbnb=pd.read_excel("Airbnb.xlsx")

airbnb.head(5)

      id                               NAME    host
id \
0  1001254          Clean & quiet apt home by the park
80014485718
1  1002102          Skylit Midtown Castle
52335172823
2  1002403          THE VILLAGE OF HARLEM....NEW YORK !
78829239556
3  1002755                      NaN
85098326012
4  1003689          Entire Apt: Spacious Studio/Loft by central park
92037596077

      host_identity_verified host name neighbourhood group
neighbourhood \
0            unconfirmed   Madaline           Brooklyn  Kensington
1            verified       Jenna        Manhattan  Midtown
2              NaN         Elise        Manhattan  Harlem
3            unconfirmed     Garry           Brooklyn Clinton Hill
4            verified       Lyndon        Manhattan East Harlem

      lat      long      country ... service fee minimum nights
\
0  40.64749 -73.97237 United States ...      193.0      10.0
1  40.75362 -73.98377 United States ...      28.0      30.0
2  40.80902 -73.94190 United States ...      124.0      3.0
3  40.68514 -73.95976 United States ...      74.0      30.0
4  40.79851 -73.94399 United States ...      41.0      10.0

      number of reviews last review reviews per month review rate number
\
0             9.0  2021-10-19      0.21      4.0

```

```
1          45.0 2022-05-21           0.38        4.0
2          0.0      NaT            NaN        5.0
3         270.0 2019-07-05           4.64        4.0
4          9.0 2018-11-19           0.10        3.0
```

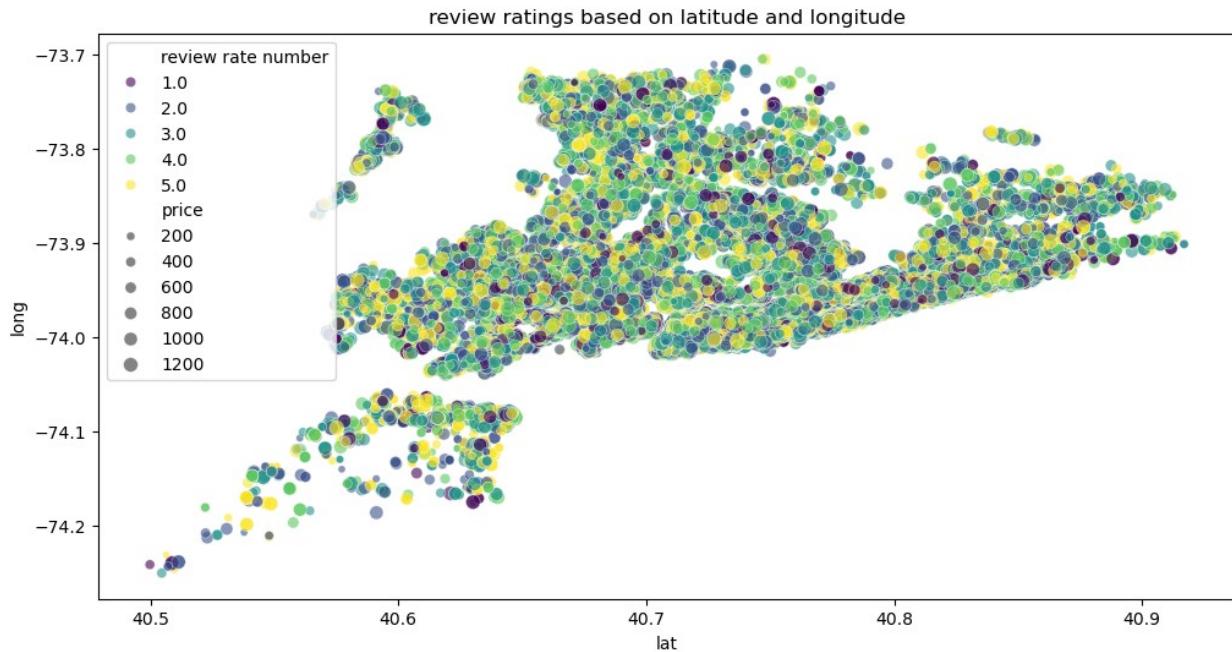
```
calculated host listings count availability 365 \
0                  6.0           286.0
1                  2.0           228.0
2                  1.0           352.0
3                  1.0           322.0
4                  1.0           289.0
```

```
house_rules license
0 Clean up and treat the home the way you'd like...     NaN
1 Pet friendly but please confirm with me if the...     NaN
2 I encourage you to use my kitchen, cooking and...     NaN
3                                         NaN     NaN
4 Please no smoking in the house, porch or on th...     NaN
```

```
[5 rows x 26 columns]
```

```
plt.figure(figsize=(12, 6))
plt.title("review ratings based on latitude and longitude")
sns.scatterplot(x='lat' , y='long',hue='review rate
number',size='price',alpha=0.6, palette="viridis",data=airbnb)

<Axes: title={'center': 'review ratings based on latitude and
longitude'}, xlabel='lat', ylabel='long'>
```



```
plt.figure(figsize=(12, 6))
plt.title("Price Trend")
sns.histplot(airbnb["price"], kde=True, bins=40)

<Axes: title={'center': 'Price Trend'}, xlabel='price',
ylabel='Count'>
```

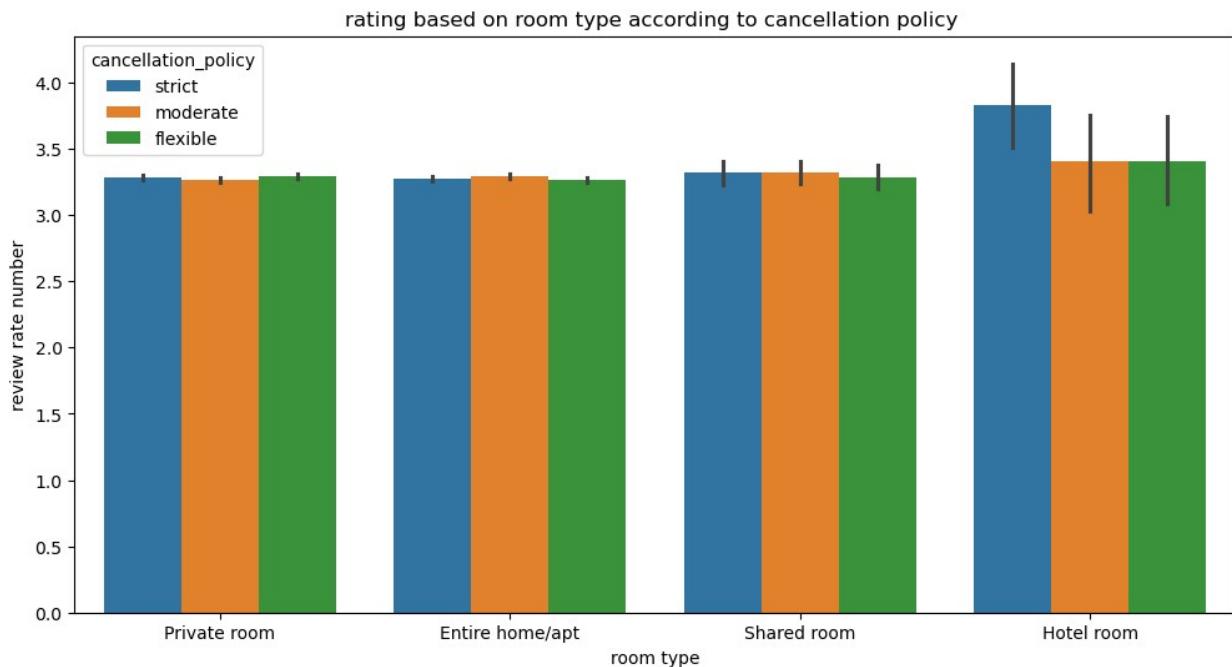


```

plt.figure(figsize=(12, 6))
plt.title("rating based on room type according to cancellation policy")
sns.barplot(x='room type',y='review rate number',hue='cancellation_policy',data=airbnb)

<Axes: title={'center': 'rating based on room type according to cancellation policy'}, xlabel='room type', ylabel='review rate number'>

```



```

tophost=airbnb['host name'].value_counts().head(5)

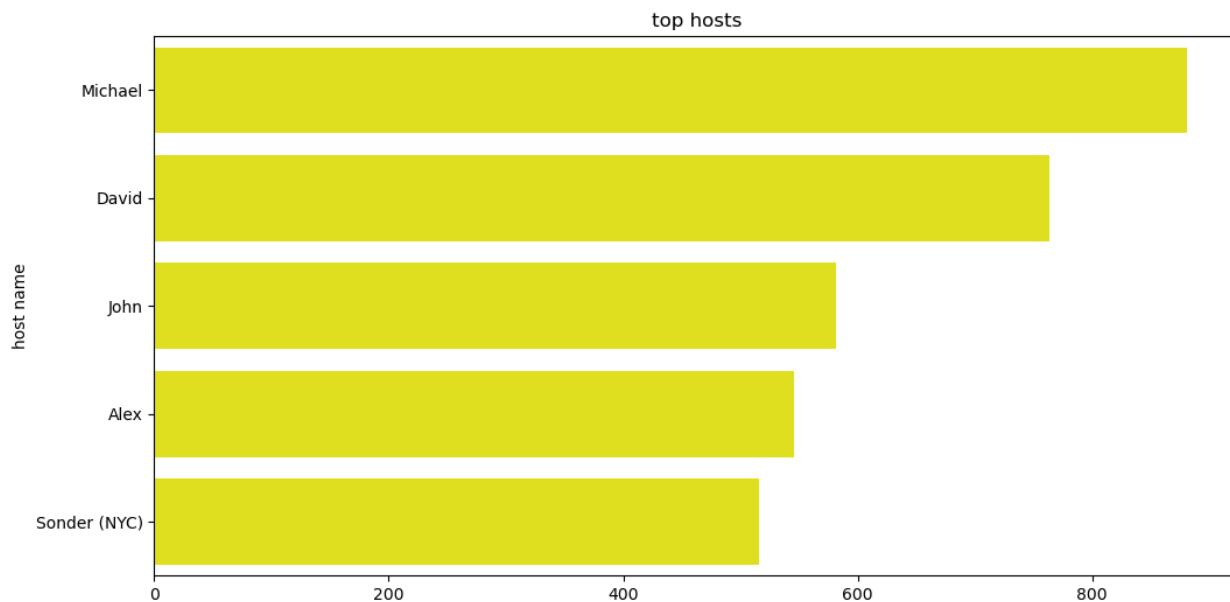
tophost

host name
Michael      881
David        764
John         581
Alex          546
Sonder (NYC)  516
Name: count, dtype: int64

plt.figure(figsize=(12, 6))
plt.title("top hosts")
sns.barplot(x=tophost.values ,y=tophost.index,color='yellow')

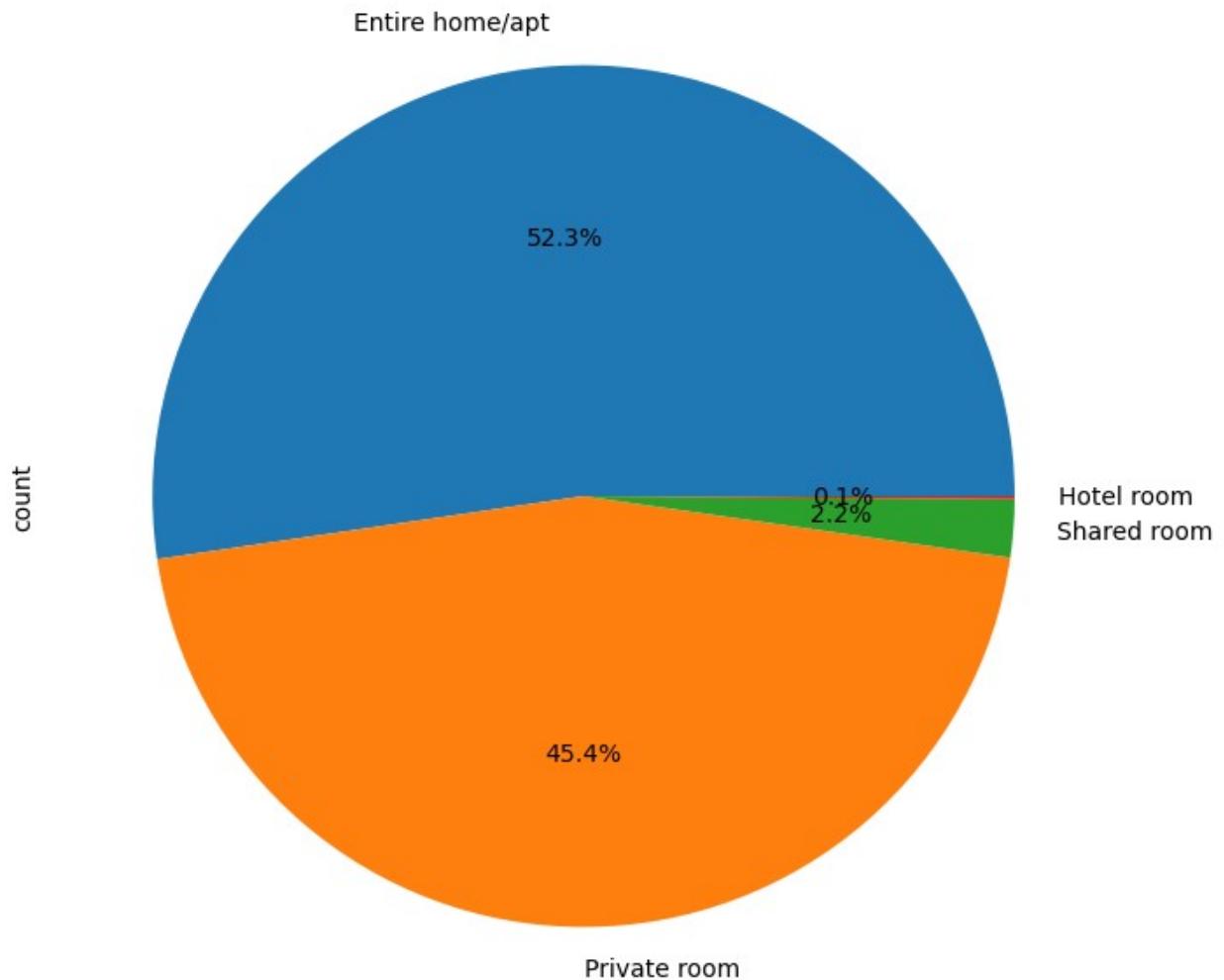
<Axes: title={'center': 'top hosts'}, ylabel='host name'>

```

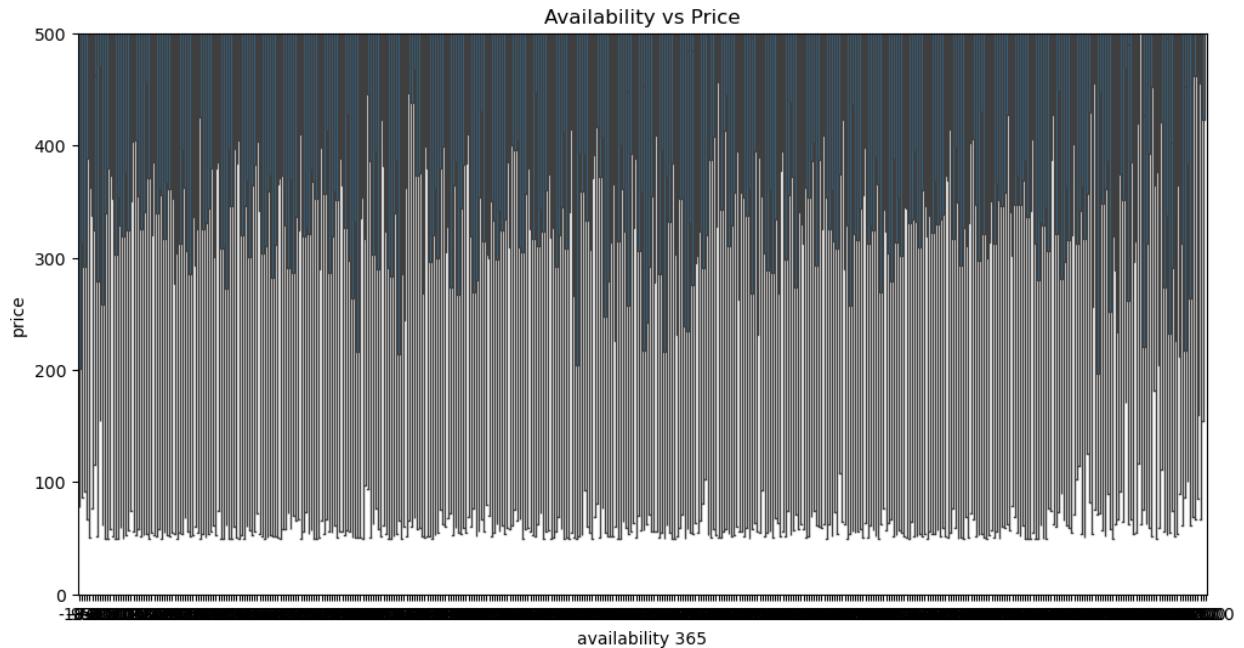


```
airbnb['room type'].value_counts().plot.pie(autopct='%.1f%%',  
figsize=(8,8))  
plt.title("Room Type Distribution")  
plt.show()
```

### Room Type Distribution

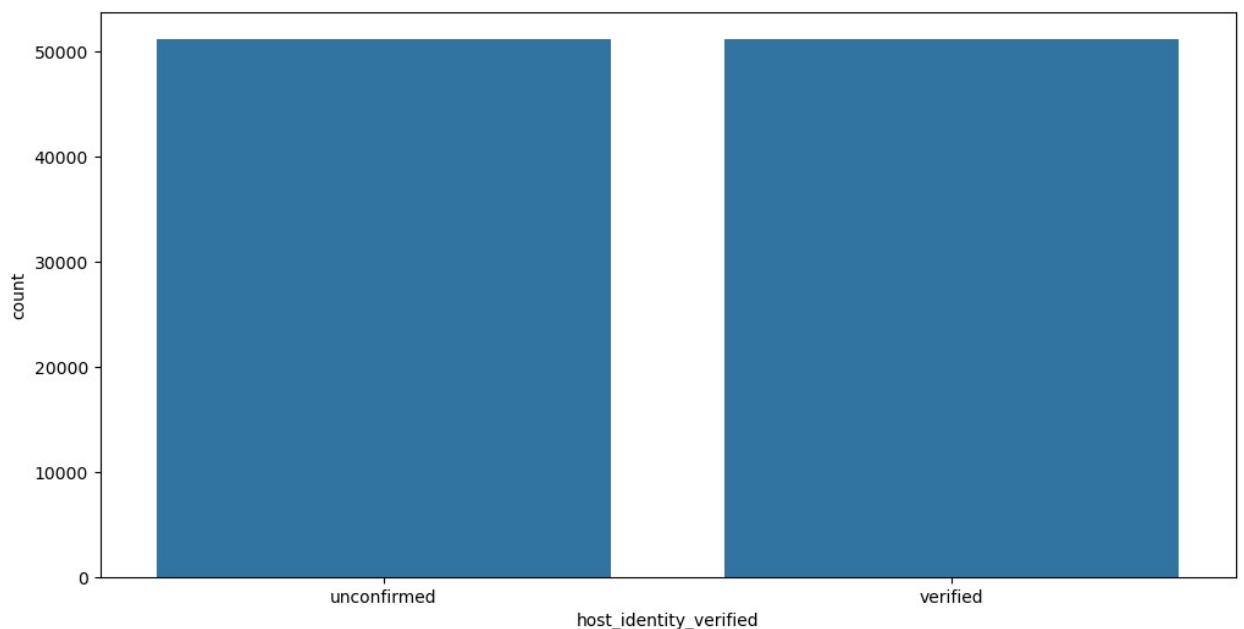


```
plt.figure(figsize=(12, 6))
sns.boxplot(x="availability_365", y="price", data=airbnb)
plt.ylim(0, 500)
plt.title("Availability vs Price")
plt.show()
```



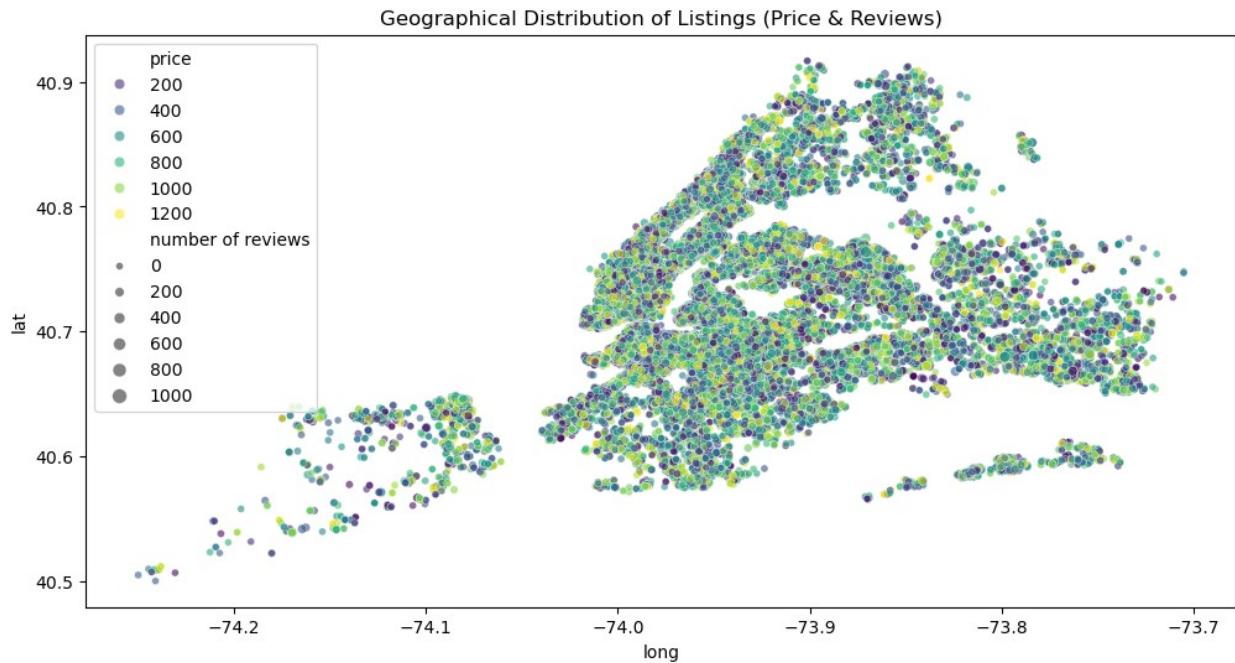
```
plt.figure(figsize=(12, 6))
sns.countplot(x=airbnb['host_identity_verified'])

<Axes: xlabel='host_identity_verified', ylabel='count'>
```

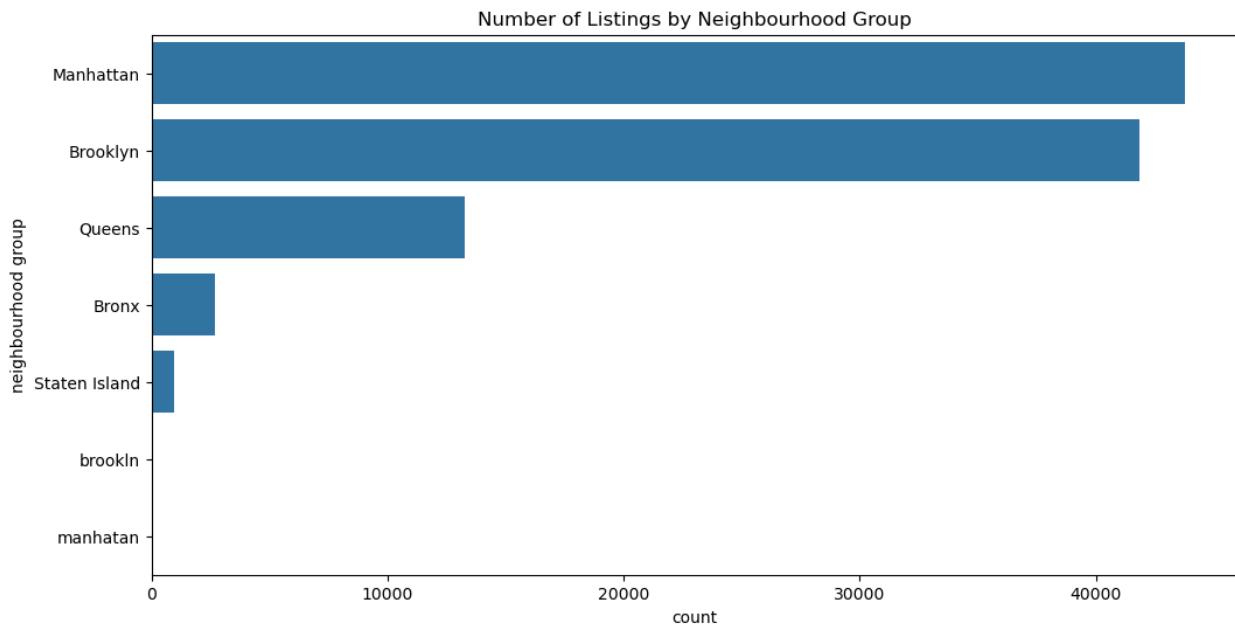


```
plt.figure(figsize=(12, 6))
sns.scatterplot(x="long", y="lat", data=airbnb, hue="price",
size="number of reviews", alpha=0.6, palette="viridis")
```

```
plt.title("Geographical Distribution of Listings (Price & Reviews)")  
plt.show()
```



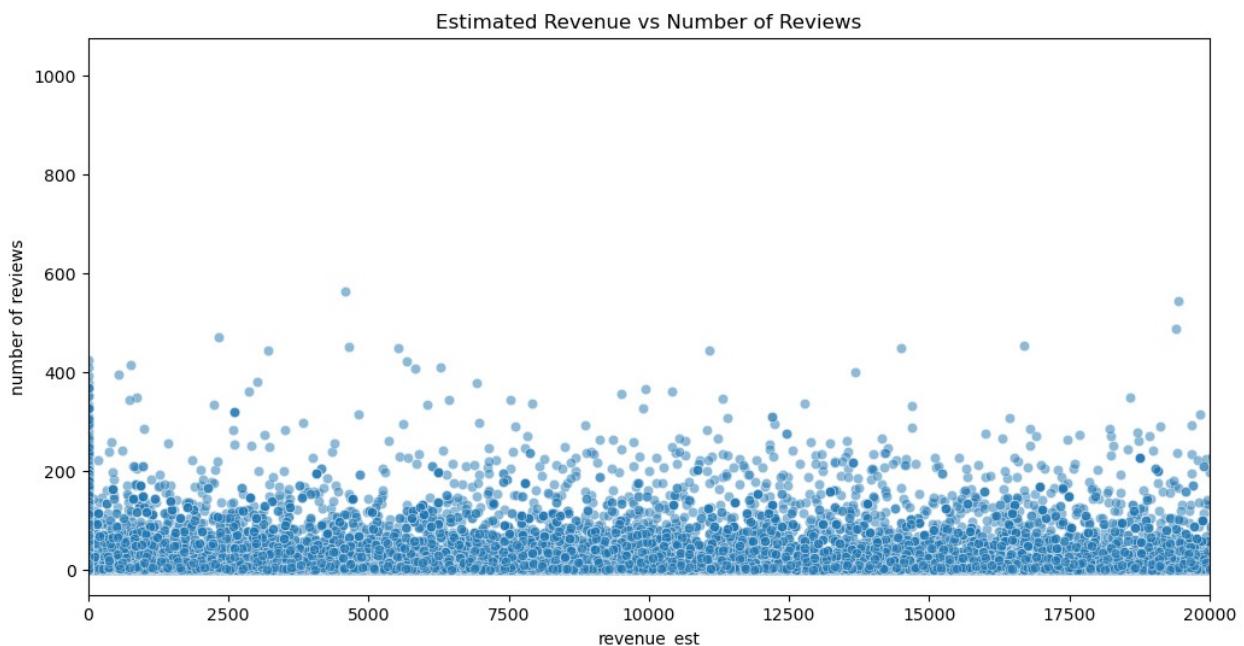
```
plt.figure(figsize=(12, 6))  
sns.countplot(y="neighbourhood group", data=airbnb,  
order=airbnb['neighbourhood group'].value_counts().index)  
plt.title("Number of Listings by Neighbourhood Group")  
plt.show()
```



```

plt.figure(figsize=(12, 6))
airbnb['revenue_est'] = airbnb['price'] * airbnb['availability_365']
sns.scatterplot(x="revenue_est", y="number of reviews", data=airbnb,
alpha=0.5)
plt.xlim(0, 20000)
plt.title("Estimated Revenue vs Number of Reviews")
plt.show()

```



```

pivair=airbnb.pivot_table(values='number of
reviews',index='neighbourhood group',columns='review rate number')
pivair

```

neighbourhood group	review rate number	1.0	2.0	3.0	4.0
Bronx	5.0	45.163636	29.526756	33.365468	31.310185
Brooklyn	27.784380	32.711404	27.954598	27.643935	27.725931
Manhattan	28.663859	29.624131	23.637770	23.489407	23.817243
Queens	23.119491	47.273810	32.159265	31.705226	31.552486
Staten Island	34.895254	38.148148	38.580645	34.470588	32.220588
brookln	NaN	NaN	NaN	167.000000	NaN
	NaN				

```
manhattan          NaN          NaN  260.000000          NaN  
NaN
```

```
plt.figure(figsize=(12, 6))  
plt.title("Heatmap of Ratings vs Number of Reviews Across  
Neighborhoods")  
sns.heatmap(pivair)
```

```
<Axes: title={'center': 'Heatmap of Ratings vs Number of Reviews  
Across Neighborhoods'}, xlabel='review rate number',  
ylabel='neighbourhood group'>
```

