In [1]:
```
pip install pandas
```
```
Requirement already satisfied: pandas in c:\users\divya gupta\anaconda3\lib\site-packages (2.2.2)
Requirement already satisfied: numpy>=1.26.0 in c:\users\divya gupta\anaconda3\lib\site-packages (from pandas) (1.26.4)
Requirement already satisfied: python-dateutil>=2.8.2 in c:\users\divya gupta\anaconda3\lib\site-packages (from pandas) (2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in c:\users\divya gupta\anaconda3\lib\site-packages (from pandas) (2024.1)
Requirement already satisfied: tzdata>=2022.7 in c:\users\divya gupta\anaconda3\lib\site-packages (from pandas) (2023.3)
Requirement already satisfied: six>=1.5 in c:\users\divya gupta\anaconda3\lib\site-packages (from python-dateutil>=2.8.2->pandas) (1.16.0)
Note: you may need to restart the kernel to use updated packages.
```

In [37]:
```
pip install numpy
```
```
Requirement already satisfied: numpy in c:\users\divya gupta\anaconda3\lib\site-packages (1.26.4)
Note: you may need to restart the kernel to use updated packages.
```

In [39]:
```
pip install matplotlib
```
```
Requirement already satisfied: matplotlib in c:\users\divya gupta\anaconda3\lib\site-packages (3.9.2)
Requirement already satisfied: contourpy>=1.0.1 in c:\users\divya gupta\anaconda3\lib\site-packages (from matplotlib) (1.2.0)
Requirement already satisfied: cycler>=0.10 in c:\users\divya gupta\anaconda3\lib\site-packages (from matplotlib) (0.11.0)
Requirement already satisfied: fonttools>=4.22.0 in c:\users\divya gupta\anaconda3\lib\site-packages (from matplotlib) (4.51.0)
Requirement already satisfied: kiwisolver>=1.3.1 in c:\users\divya gupta\anaconda3\lib\site-packages (from matplotlib) (1.4.4)
Requirement already satisfied: numpy>=1.23 in c:\users\divya gupta\anaconda3\lib\site-packages (from matplotlib) (1.26.4)
Requirement already satisfied: packaging>=20.0 in c:\users\divya gupta\anaconda3\lib\site-packages (from matplotlib) (24.1)
Requirement already satisfied: pillow>=8 in c:\users\divya gupta\anaconda3\lib\site-packages (from matplotlib) (10.4.0)
Requirement already satisfied: pyparsing>=2.3.1 in c:\users\divya gupta\anaconda3\lib\site-packages (from matplotlib) (3.1.2)
Requirement already satisfied: python-dateutil>=2.7 in c:\users\divya gupta\anaconda3\lib\site-packages (from matplotlib) (2.9.0.post0)
Requirement already satisfied: six>=1.5 in c:\users\divya gupta\anaconda3\lib\site-packages (from python-dateutil>=2.7->matplotlib) (1.16.0)
Note: you may need to restart the kernel to use updated packages.
```

In [45]:
```
pip install seaborn
```
```
Requirement already satisfied: seaborn in c:\users\divya gupta\anaconda3\lib\site-packages (0.13.2)Note: you may need to restart the kernel to use updated packages.

Requirement already satisfied: numpy!=1.24.0,>=1.20 in c:\users\divya gupta\anaconda3\lib\site-packages (from seaborn) (1.26.4)
Requirement already satisfied: pandas>=1.2 in c:\users\divya gupta\anaconda3\lib\site-packages (from seaborn) (2.2.2)
Requirement already satisfied: matplotlib!=3.6.1,>=3.4 in c:\users\divya gupta\anaconda3\lib\site-packages (from seaborn) (3.9.2)
Requirement already satisfied: contourpy>=1.0.1 in c:\users\divya gupta\anaconda3\lib\site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (1.2.0)
Requirement already satisfied: cycler>=0.10 in c:\users\divya gupta\anaconda3\lib\site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (0.11.0)
Requirement already satisfied: fonttools>=4.22.0 in c:\users\divya gupta\anaconda3\lib\site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (4.51.0)
Requirement already satisfied: kiwisolver>=1.3.1 in c:\users\divya gupta\anaconda3\lib\site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (1.4.4)
Requirement already satisfied: packaging>=20.0 in c:\users\divya gupta\anaconda3\lib\site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (24.1)
Requirement already satisfied: pillow>=8 in c:\users\divya gupta\anaconda3\lib\site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (10.4.0)
Requirement already satisfied: pyparsing>=2.3.1 in c:\users\divya gupta\anaconda3\lib\site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (3.1.2)
Requirement already satisfied: python-dateutil>=2.7 in c:\users\divya gupta\anaconda3\lib\site-packages (from matplotlib!=3.6.1,>=3.4->seaborn) (2.9.0.post0)
Requirement already satisfied: pytz>=2020.1 in c:\users\divya gupta\anaconda3\lib\site-packages (from pandas>=1.2->seaborn) (2024.1)
Requirement already satisfied: tzdata>=2022.7 in c:\users\divya gupta\anaconda3\lib\site-packages (from pandas>=1.2->seaborn) (2023.3)
Requirement already satisfied: six>=1.5 in c:\users\divya gupta\anaconda3\lib\site-packages (from python-dateutil>=2.7->matplotlib!=3.6.1,>=3.4->seaborn) (1.16.0)
```

In [47]:
```
import pandas as pd
import numpy as np
import matplotlib as plt
import seaborn as sns
```

In [19]:
```
df=pd.read_csv("airbnb.csv")
df
```

Out[19]:

	id	name	host_id	host_name	neighbourhood_group	neighbourhood	latitude	longitude	room_type	price	...	last_review	reviews_per_month	calculated_host_listings_count	availability_365	number_of_reviews_ltm	license
0	1.312228e+06	Rental unit in Brooklyn · ★5.0 · 1 bedroom	7130382	Walter	Brooklyn	Clinton Hill	40.683710	-73.964610	Private room	55.0	...	20/12/15	0.03	1.0	0.0	0.0	No License
1	4.527754e+07	Rental unit in New York · ★4.67 · 2 bedrooms ·...	51501835	Jeniffer	Manhattan	Hell's Kitchen	40.766610	-73.988100	Entire home/apt	144.0	...	01/05/23	0.24	139.0	364.0	2.0	No License
2	9.710000e+17	Rental unit in New York · ★4.17 · 1 bedroom · ...	528871354	Joshua	Manhattan	Chelsea	40.750764	-73.994605	Entire home/apt	187.0	...	18/12/23	1.67	1.0	343.0	6.0	Exempt
3	3.857863e+06	Rental unit in New York · ★4.64 · 1 bedroom · ...	19902271	John And Catherine	Manhattan	Washington Heights	40.835600	-73.942500	Private room	120.0	...	17/09/23	1.38	2.0	363.0	12.0	No License
4	4.089661e+07	Condo in New York · ★4.91 · Studio · 1 bed · 1...	61391963	Stay With Vibe	Manhattan	Murray Hill	40.751120	-73.978600	Entire home/apt	85.0	...	03/12/23	0.24	133.0	335.0	3.0	No License
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
20765	2.473690e+07	Rental unit in New York · ★4.75 · 1 bedroom · ...	186680487	Henry D	Manhattan	Lower East Side	40.711380	-73.991560	Private room	45.0	...	29/09/23	1.81	1.0	157.0	12.0	No License
20766	2.835711e+06	Rental unit in New York · ★4.46 · 1 bedroom · ...	3237504	Aspen	Manhattan	Greenwich Village	40.730580	-74.000700	Entire home/apt	105.0	...	01/07/23	0.48	1.0	0.0	1.0	No License
20767	5.182527e+07	Rental unit in New York · ★4.93 · 1 bedroom · ...	304317395	Jeff	Manhattan	Hell's Kitchen	40.757350	-73.993430	Entire home/apt	299.0	...	08/12/23	2.09	1.0	0.0	27.0	No License
20768	7.830000e+17	Rental unit in New York · ★5.0 · 1 bedroom · 1...	163083101	Marissa	Manhattan	Chinatown	40.713750	-73.991470	Entire home/apt	115.0	...	17/09/23	0.91	1.0	363.0	7.0	No License
20769	5.660000e+17	Rental unit in Queens · ★4.89 · 1 bedroom · 1 ...	93827372	Glenroy	Queens	Rosedale	40.658874	-73.728651	Private room	102.0	...	10/12/23	4.50	1.0	0.0	62.0	OS STRRE 00005

20770 rows × 22 columns

In [23]:
```
#row and column number
df.shape
```

Out[23]: (20770, 22)

In [25]:
```
#data type
df.dtypes
```

Out[25]:
```
id                                float64
name                               object
host_id                             int64
host_name                          object
neighbourhood_group                object
neighbourhood                      object
latitude                          float64
longitude                         float64
room_type                          object
price                             float64
minimum_nights                    float64
number_of_reviews                 float64
last_review                        object
reviews_per_month                 float64
calculated_host_listings_count    float64
availability_365                  float64
number_of_reviews_ltm             float64
license                            object
rating                             object
bedrooms                           object
beds                                int64
baths                              object
dtype: object
```

In [33]:
```
df.info()
```
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 20770 entries, 0 to 20769
Data columns (total 22 columns):
 #   Column                          Non-Null Count  Dtype  
---  ------                          --------------  -----  
 0   id                              20770 non-null  float64
 1   name                            20770 non-null  object 
 2   host_id                         20770 non-null  int64  
 3   host_name                       20770 non-null  object 
 4   neighbourhood_group             20770 non-null  object 
 5   neighbourhood                   20763 non-null  object 
 6   latitude                        20763 non-null  float64
 7   longitude                       20763 non-null  float64
 8   room_type                       20763 non-null  object 
 9   price                           20736 non-null  float64
 10  minimum_nights                  20763 non-null  float64
 11  number_of_reviews               20763 non-null  float64
 12  last_review                     20763 non-null  object 
 13  reviews_per_month               20763 non-null  float64
 14  calculated_host_listings_count  20763 non-null  float64
 15  availability_365                20763 non-null  float64
 16  number_of_reviews_ltm           20763 non-null  float64
 17  license                         20770 non-null  object 
 18  rating                          20770 non-null  object 
 19  bedrooms                        20770 non-null  object 
 20  beds                            20770 non-null  int64  
 21  baths                           20770 non-null  object 
dtypes: float64(10), int64(2), object(10)
memory usage: 3.5+ MB
```

In [31]:
```
df.describe()
```

Out[31]:

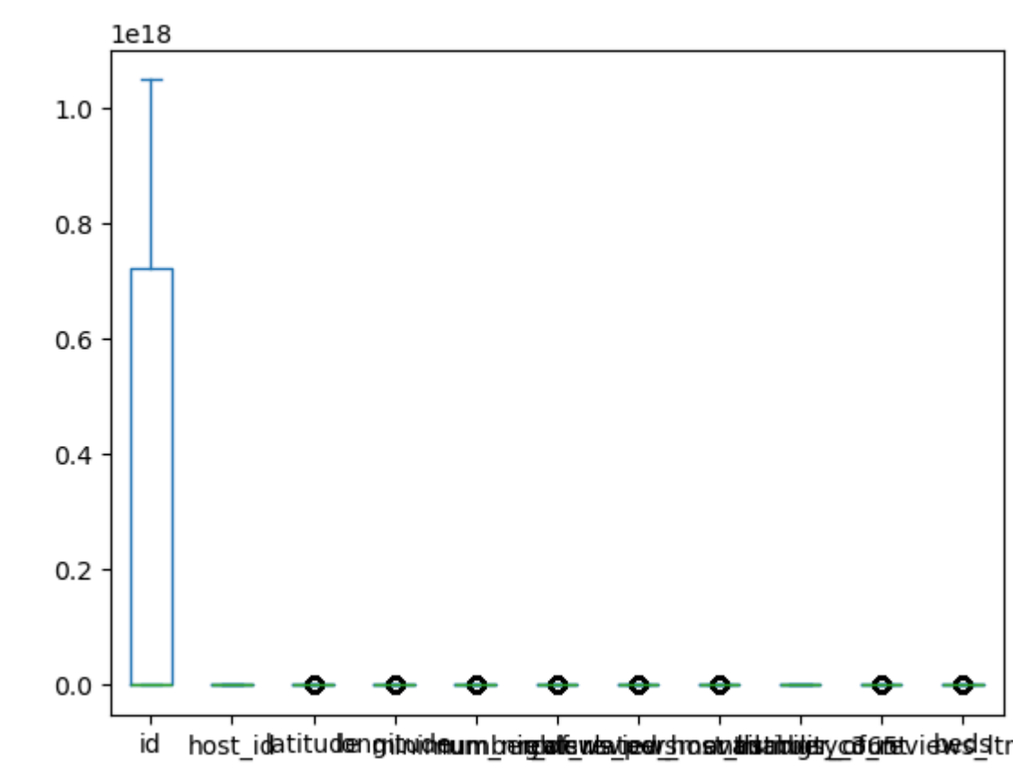
	id	host_id	latitude	longitude	price	minimum_nights	number_of_reviews	reviews_per_month	calculated_host_listings_count	availability_365	number_of_reviews_ltm	beds
count	2.077000e+04	2.077000e+04	20763.000000	20763.000000	20736.000000	20763.000000	20763.000000	20763.000000	20763.000000	20763.000000	20763.000000	20770.000000
mean	3.033858e+17	1.749049e+08	40.726821	-73.939179	187.714940	28.558493	42.610605	1.257589	18.866686	206.067957	10.848962	1.723592
std	3.901221e+17	1.725657e+08	0.060293	0.061403	1023.245124	33.532697	73.523401	1.904472	70.921443	135.077259	21.354876	1.211993
min	2.595000e+03	1.678000e+03	40.500314	-74.249840	10.000000	1.000000	1.000000	0.010000	1.000000	0.000000	0.000000	1.000000
25%	2.707260e+07	2.041184e+07	40.684159	-73.980755	80.000000	30.000000	4.000000	0.210000	1.000000	87.000000	1.000000	1.000000
50%	4.992852e+07	1.086990e+08	40.722890	-73.949597	125.000000	30.000000	14.000000	0.650000	2.000000	215.000000	3.000000	1.000000
75%	7.220000e+17	3.143997e+08	40.763106	-73.917475	199.000000	30.000000	49.000000	1.800000	5.000000	353.000000	15.000000	2.000000
max	1.050000e+18	5.504035e+08	40.911147	-73.713650	100000.000000	1250.000000	1865.000000	75.490000	713.000000	365.000000	1075.000000	42.000000

In [35]:
```
df["availability_365"].describe()
```

Out[35]:
```
count    20763.000000
mean       206.067957
std        135.077259
min          0.000000
25%         87.000000
50%        215.000000
75%        353.000000
max        365.000000
Name: availability_365, dtype: float64
```

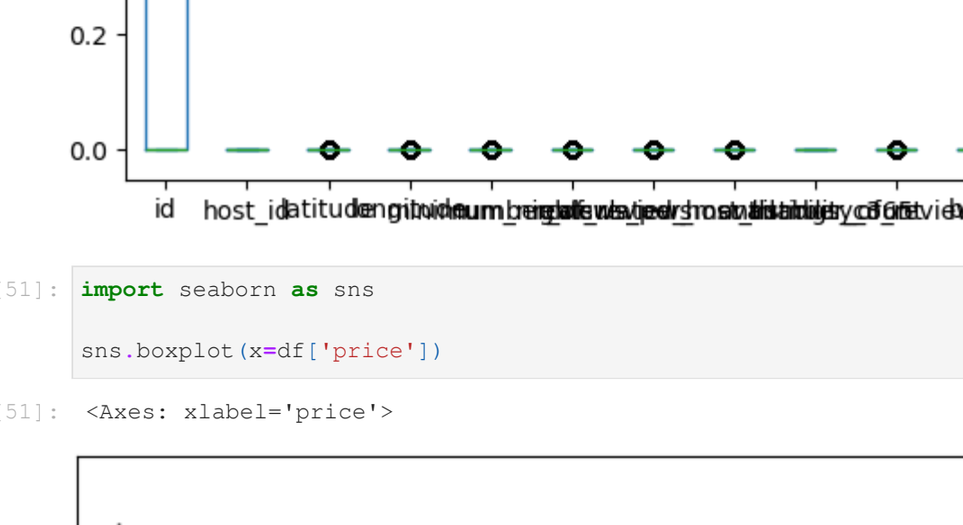
In [49]:
```
df.plot(x='price', kind='box')
```

Out[49]: <Axes: >



In [51]:
```
import seaborn as sns

sns.boxplot(x=df['price'])
```

Out[51]: <Axes: xlabel='price'>



In [57]:
```
#remove outlier part i.e100000
sum(df['price'] >15000)
```

Out[57]: 2

In [59]:
```
#subsetting, where clause in sql,
df[df['price'] >15000]
```

Out[59]:

	id	name	host_id	host_name	neighbourhood_group	neighbourhood	latitude	longitude	room_type	price	...	last_review	reviews_per_month	calculated_host_listings_count	availability_365	number_of_reviews_ltm	license
3995	1.716029e+07	Rental unit in Brooklyn · ★4.48 · 1 bedroom · ...	110361431	Bobbi	Brooklyn	Bedford-Stuyvesant	40.69085	-73.93806	Private room	100000.0	...	20/10/23	0.96	2.0	346.0	10.0	No License
5497	6.050000e+17	Rental unit in Brooklyn · ★4.33 · 1 bedroom · ...	110361431	Bobbi	Brooklyn	Bedford-Stuyvesant	40.69254	-73.93636	Private room	100000.0	...	31/10/23	0.45	2.0	365.0	5.0	No License

2 rows × 22 columns

In [61]:
```
sum(df['availability_365'] >180)
```

Out[61]: 10982

In [63]:
```
df[(df['availability_365'] >180)]
```

Out[63]:

	id	name	host_id	host_name	neighbourhood_group	neighbourhood	latitude	longitude	room_type	price	...	last_review	reviews_per_month	calculated_host_listings_count	availability_365	number_of_reviews_ltm	license
1	4.527754e+07	Rental unit in New York · ★4.67 · 2 bedrooms ·...	51501835	Jeniffer	Manhattan	Hell's Kitchen	40.766610	-73.988100	Entire home/apt	144.0	...	01/05/23	0.24	139.0	364.0	2.0	No License
2	9.710000e+17	Rental unit in New York · ★4.17 · 1 bedroom · ...	528871354	Joshua	Manhattan	Chelsea	40.750764	-73.994605	Entire home/apt	187.0	...	18/12/23	1.67	1.0	343.0	6.0	Exempt
3	3.857863e+06	Rental unit in New York · ★4.64 · 1 bedroom · ...	19902271	John And Catherine	Manhattan	Washington Heights	40.835600	-73.942500	Private room	120.0	...	17/09/23	1.38	2.0	363.0	12.0	No License
4	4.089661e+07	Condo in New York · ★4.91 · Studio · 1 bed · 1...	61391963	Stay With Vibe	Manhattan	Murray Hill	40.751120	-73.978600	Entire home/apt	85.0	...	03/12/23	0.24	133.0	335.0	3.0	No License
5	4.958498e+07	Rental unit in New York · ★5.0 · 1 bedroom · 1...	51501835	Jeniffer	Manhattan	Hell's Kitchen	40.759950	-73.992960	Entire home/apt	115.0	...	29/07/23	0.16	139.0	276.0	2.0	No License
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
20759	7.480000e+17	Rental unit in New York · 1 bedroom · 1 bed · ...	334376746	Leo	Manhattan	Harlem	40.802970	-73.958090	Private room	53.0	...	03/12/23	0.86	9.0	361.0	1.0	No License
20761	4.862769e+07	Rental unit in New York · ★4.0 · 1 bedroom · 1...	51501835	Jeniffer	Manhattan	Lower East Side	40.718290	-73.993720	Entire home/apt	145.0	...	29/08/22	0.09	139.0	275.0	0.0	No License
20763	1.576774e+07	Rental unit in New York · 1 bedroom · 1 bed · ...	22541573	Furnished Quarters	Manhattan	West Village	40.731810	-74.008190	Entire home/apt	215.0	...	28/06/17	0.01	164.0	358.0	0.0	No License
20764	8.970000e+17	Rental unit in New York · 1 bedroom · 1 bed · ...	15828752	Gideon	Manhattan	Midtown	40.764520	-73.977920	Entire home/apt	140.0	...	31/07/23	0.19	1.0	267.0	1.0	No License
20768	7.830000e+17	Rental unit in New York · ★5.0 · 1 bedroom · 1...	163083101	Marissa	Manhattan	Chinatown	40.713750	-73.991470	Entire home/apt	115.0	...	17/09/23	0.91	1.0	363.0	7.0	No License

10982 rows × 22 columns

In [67]:
```
df[(df['availability_365'] >355)]
```

Out[67]:

	id	name	host_id	host_name	neighbourhood_group	neighbourhood	latitude	longitude	room_type	price	...	last_review	reviews_per_month	calculated_host_listings_count	availability_365	number_of_reviews_ltm	license
1	4.527754e+07	Rental unit in New York · ★4.67 · 2 bedrooms ·...	51501835	Jeniffer	Manhattan	Hell's Kitchen	40.76661	-73.98810	Entire home/apt	144.0	...	01/05/23	0.24	139.0	364.0	2.0	No License
3	3.857863e+06	Rental unit in New York · ★4.64 · 1 bedroom · ...	19902271	John And Catherine	Manhattan	Washington Heights	40.83560	-73.94250	Private room	120.0	...	17/09/23	1.38	2.0	363.0	12.0	No License
6	4.527754e+07	Rental unit in New York · ★4.67 · 2 bedrooms ·...	51501835	Jeniffer	Manhattan	Hell's Kitchen	40.76661	-73.98810	Entire home/apt	144.0	...	01/05/23	0.24	139.0	364.0	2.0	No License
8	3.857863e+06	Rental unit in New York · ★4.64 · 1 bedroom · ...	19902271	John And Catherine	Manhattan	Washington Heights	40.83560	-73.94250	Private room	120.0	...	17/09/23	1.38	2.0	363.0	12.0	No License
11	4.545705e+07	Rental unit in New York · ★4.33 · Studio · 1 b...	51501835	Jeniffer	Manhattan	Hell's Kitchen	40.76737	-73.98787	Entire home/apt	105.0	...	31/08/22	0.10	139.0	364.0	0.0	No License
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
20752	1.622917e+07	Loft in New York · ★4.89 · 1 bedroom · 2 beds ·...	185745	Alessandra	Manhattan	Lower East Side	40.71415	-73.98999	Entire home/apt	160.0	...	03/12/23	1.77	3.0	365.0	36.0	No License
20757	7.610000e+17	Rental unit in New York · ★4.75 · 1 bedroom · ...	103729	Ann	Manhattan	Lower East Side	40.71863	-73.98795	Entire home/apt	112.0	...	30/11/23	0.30	1.0	365.0	2.0	No License
20759	7.480000e+17	Rental unit in New York · 1 bedroom · 1 bed · ...	334376746	Leo	Manhattan	Harlem	40.80297	-73.95809	Private room	53.0	...	03/12/23	0.86	9.0	361.0	1.0	No License
20763	1.576774e+07	Rental unit in New York · 1 bedroom · 1 bed · ...	22541573	Furnished Quarters	Manhattan	West Village	40.73181	-74.00819	Entire home/apt	215.0	...	28/06/17	0.01	164.0	358.0	0.0	No License
20768	7.830000e+17	Rental unit in New York · ★5.0 · 1 bedroom · 1...	163083101	Marissa	Manhattan	Chinatown	40.71375	-73.99147	Entire home/apt	115.0	...	17/09/23	0.91	1.0	363.0	7.0	No License

5080 rows × 22 columns

In [73]:
```
df.groupby('neighbourhood_group')['rating'].max()
```

Out[73]:
```
neighbourhood_group
Bronx            No rating
Brooklyn         No rating
Manhattan        No rating
Queens           No rating
Staten Island    No rating
Name: rating, dtype: object
```

In [75]:
```
df['rating']
```

Out[75]:
```
0           5
1        4.67
2        4.17
3        4.64
4        4.91
         ... 
20765    4.75
20766    4.46
20767    4.93
20768       5
20769    4.89
Name: rating, Length: 20770, dtype: object
```

In [83]:
```
df[df['rating'] == 'No rating']['rating']
```

Out[83]:
```
15       No rating
17       No rating
18       No rating
20       No rating
21       No rating
           ...    
20756    No rating
20758    No rating
20759    No rating
20763    No rating
20764    No rating
Name: rating, Length: 3595, dtype: object
```

In [85]:
```
df['rating'] = df['rating'].astype('float', errors='ignore')
df.info()
```
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 20770 entries, 0 to 20769
Data columns (total 22 columns):
 #   Column                          Non-Null Count  Dtype  
---  ------                          --------------  -----  
 0   id                              20770 non-null  float64
 1   name                            20770 non-null  object 
 2   host_id                         20770 non-null  int64  
 3   host_name                       20770 non-null  object 
 4   neighbourhood_group             20770 non-null  object 
 5   neighbourhood                   20763 non-null  object 
 6   latitude                        20763 non-null  float64
 7   longitude                       20763 non-null  float64
 8   room_type                       20763 non-null  object 
 9   price                           20736 non-null  float64
 10  minimum_nights                  20763 non-null  float64
 11  number_of_reviews               20763 non-null  float64
 12  last_review                     20763 non-null  object 
 13  reviews_per_month               20763 non-null  float64
 14  calculated_host_listings_count  20763 non-null  float64
 15  availability_365                20763 non-null  float64
 16  number_of_reviews_ltm           20763 non-null  float64
 17  license                         20770 non-null  object 
 18  rating                          20770 non-null  object 
 19  bedrooms                        20770 non-null  object 
 20  beds                            20770 non-null  int64  
 21  baths                           20770 non-null  object 
dtypes: float64(10), int64(2), object(10)
memory usage: 3.5+ MB
```

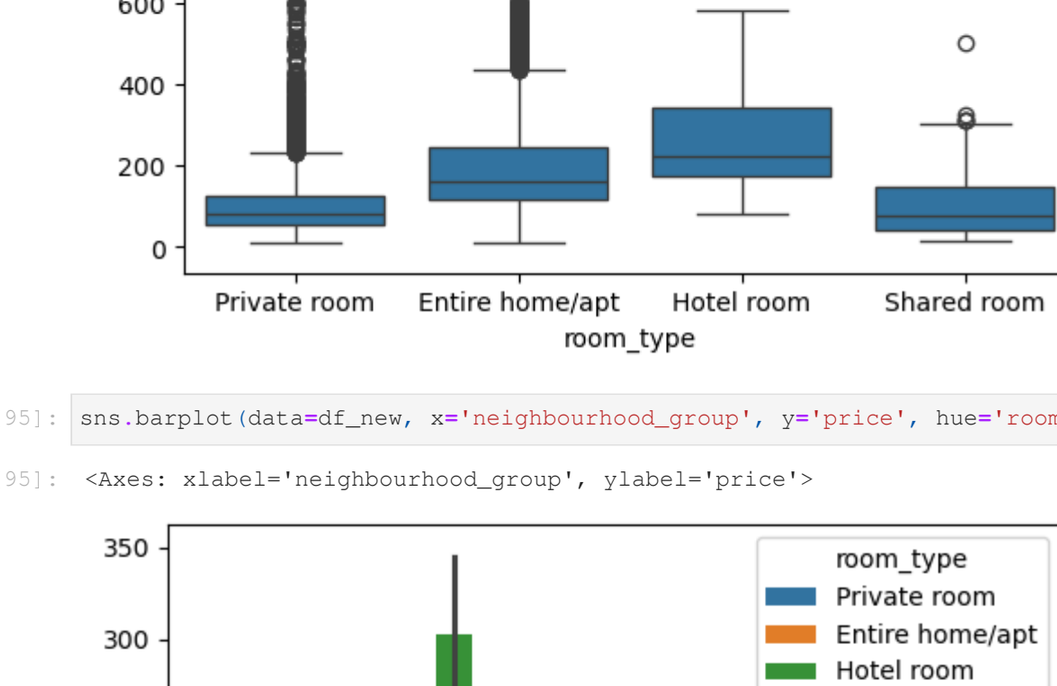
In [87]:
```
df['rating'] = df['rating'].replace('No rating', None)

df.info()
```
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 20770 entries, 0 to 20769
Data columns (total 22 columns):
 #   Column                          Non-Null Count  Dtype  
---  ------                          --------------  -----  
 0   id                              20770 non-null  float64
 1   name                            20770 non-null  object 
 2   host_id                         20770 non-null  int64  
 3   host_name                       20770 non-null  object 
 4   neighbourhood_group             20770 non-null  object 
 5   neighbourhood                   20763 non-null  object 
 6   latitude                        20763 non-null  float64
 7   longitude                       20763 non-null  float64
 8   room_type                       20763 non-null  object 
 9   price                           20736 non-null  float64
 10  minimum_nights                  20763 non-null  float64
 11  number_of_reviews               20763 non-null  float64
 12  last_review                     20763 non-null  object 
 13  reviews_per_month               20763 non-null  float64
 14  calculated_host_listings_count  20763 non-null  float64
 15  availability_365                20763 non-null  float64
 16  number_of_reviews_ltm           20763 non-null  float64
 17  license                         20770 non-null  object 
 18  rating                          17175 non-null  object 
 19  bedrooms                        20770 non-null  object 
 20  beds                            20770 non-null  int64  
 21  baths                           20770 non-null  object 
dtypes: float64(10), int64(2), object(10)
memory usage: 3.5+ MB
```

In [89]:
```
df['rating'] = df['rating'].astype('float')
df.info()
```
```
---------------------------------------------------------------------------
ValueError                                Traceback (most recent call last)
Cell In[89], line 1
----> 1 df['rating'] = df['rating'].astype('float')
      2 df.info()

File ~\anaconda3\Lib\site-packages\pandas\core\generic.py:6643, in NDFrame.astype(self, dtype, copy, errors)
   6637     results = [
   6638         ser.astype(dtype, copy=copy, errors=errors) for _, ser in self.items()
   6639     ]
   6641 else:
   6642     # else, only a single dtype is given
-> 6643     new_data = self._mgr.astype(dtype=dtype, copy=copy, errors=errors)
   6644     res = self._constructor_from_mgr(new_data, axes=new_data.axes)
   6645     return res.__finalize__(self, method="astype")

File ~\anaconda3\Lib\site-packages\pandas\core\internals\managers.py:430, in BaseBlockManager.astype(self, dtype, copy, errors)
    427 elif using_copy_on_write():
    428     copy = False
--> 430 return self.apply(
    431     "astype",
    432     dtype=dtype,
    433     copy=copy,
    434     errors=errors,
    435     using_cow=using_copy_on_write(),
    436 )

File ~\anaconda3\Lib\site-packages\pandas\core\internals\managers.py:363, in BaseBlockManager.apply(self, f, align_keys, **kwargs)
    361         applied = b.apply(f, **kwargs)
    362     else:
--> 363         applied = getattr(b, f)(**kwargs)
    364     result_blocks = extend_blocks(applied, result_blocks)
    366 out = type(self).from_blocks(result_blocks, self.axes)

File ~\anaconda3\Lib\site-packages\pandas\core\internals\blocks.py:758, in Block.astype(self, dtype, copy, errors, using_cow, squeeze)
    755         raise ValueError("Can not squeeze with more than one column.")
    756     values = values[0, :]  # type: ignore[call-overload]
--> 758 new_values = astype_array_safe(values, dtype, copy=copy, errors=errors)
    760 new_values = maybe_coerce_values(new_values)
    762 refs = None

File ~\anaconda3\Lib\site-packages\pandas\core\dtypes\astype.py:237, in astype_array_safe(values, dtype, copy, errors)
    234     dtype = dtype.numpy_dtype
    236 try:
--> 237     new_values = astype_array(values, dtype, copy=copy)
    238 except (ValueError, TypeError):
    239     # e.g. _astype_nansafe can fail on object-dtype of strings
    240     #  trying to convert to float
    241     if errors == "ignore":

File ~\anaconda3\Lib\site-packages\pandas\core\dtypes\astype.py:182, in astype_array(values, dtype, copy)
    179     values = values.astype(dtype, copy=copy)
    181 else:
--> 182     values = _astype_nansafe(values, dtype, copy=copy)
    184 # in pandas we don't store numpy str dtypes, so convert to object
    185 if isinstance(dtype, np.dtype) and issubclass(values.dtype.type, str):

File ~\anaconda3\Lib\site-packages\pandas\core\dtypes\astype.py:133, in _astype_nansafe(arr, dtype, copy, skipna)
    129     raise ValueError(msg)
    131 if copy or arr.dtype == object or dtype == object:
    132     # Explicit copy, or required since NumPy can't view from / to object.
--> 133     return arr.astype(dtype, copy=True)
    135 return arr.astype(dtype, copy=copy)

ValueError: could not convert string to float: 'New '
```

In [91]:
```
df_new = df[df['price'] < 1500]
```

In [93]:
```
sns.boxplot(data=df_new, x='room_type', y='price')
```

Out[93]: <Axes: xlabel='room_type', ylabel='price'>



In [95]:
```
sns.barplot(data=df_new, x='neighbourhood_group', y='price', hue='room_type')
```

Out[95]: <Axes: xlabel='neighbourhood_group', ylabel='price'>