

```
##Import the libraries
import pandas as pd
import numpy as np
import seaborn as sns
```

```
df = pd.read_csv("/content/airbnblistings.csv")
```

```
df.head(5)
```

	id	name	host_id	host_name	host_since_Year	neighbourhood	latitude	longitude	room_type	price	...	last_r
0	2818	Quiet Garden View Room & Super Fast WiFi	3159	Daniel	2008	Oostelijk Havengebied - Indische Buurt	52.36435	4.94358	Private room	49	...	2022-
1	20168	Studio with private bathroom in the centre 1	59484	Alexander	2009	Centrum-Oost	52.36407	4.89393	Private room	106	...	2020-
2	27886	Romantic, stylish B&B houseboat in canal district	97647	Flip	2010	Centrum-West	52.38761	4.89188	Private room	126	...	2022-
3	28871	Comfortable double room	124245	Edwin	2010	Centrum-West	52.36775	4.89092	Private room	75	...	2022-
4	29051	Comfortable single room	124245	Edwin	2010	Centrum-Oost	52.36584	4.89111	Private room	55	...	2022-

5 rows × 22 columns

```
def make_upper(x):
    return x.upper()
```

```
df['name']=df['name'].apply(make_upper)
```

```
df['name'].str.contains("AMSTERDAM").count()
```

6173

```
df[df['name'].str.contains("AMSTERDAM")]["price"].mean()
```

200.10332326283987

```
df[df['name'].str.contains("AMSTERDAM")]["review_scores_rating"].mean()
```

4.794828767123287

```
import matplotlib.pyplot as plt
```

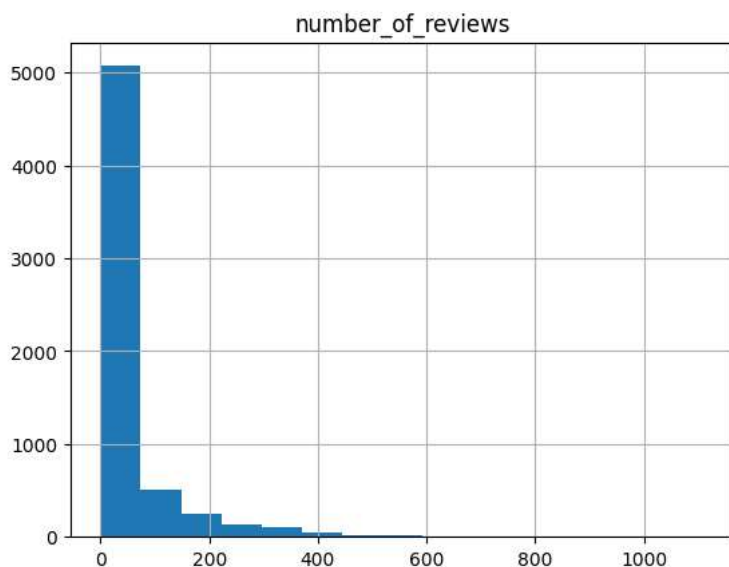
```
df.hist(column='price',bins=20,rwidth=2,color = "yellow")
```

```
array([[<Axes: title={'center': 'price'}>]], dtype=object)
```



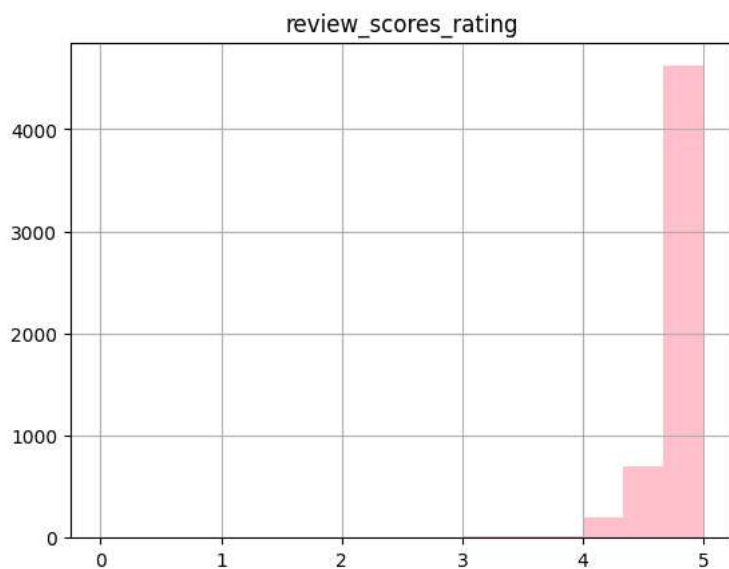
```
df.hist(column='number_of_reviews',bins=15,rwidth=2)
```

```
array([[<Axes: title={'center': 'number_of_reviews'}>]], dtype=object)
```



```
df.hist(column='review_scores_rating',bins=15,rwidth=2,color="pink")
```

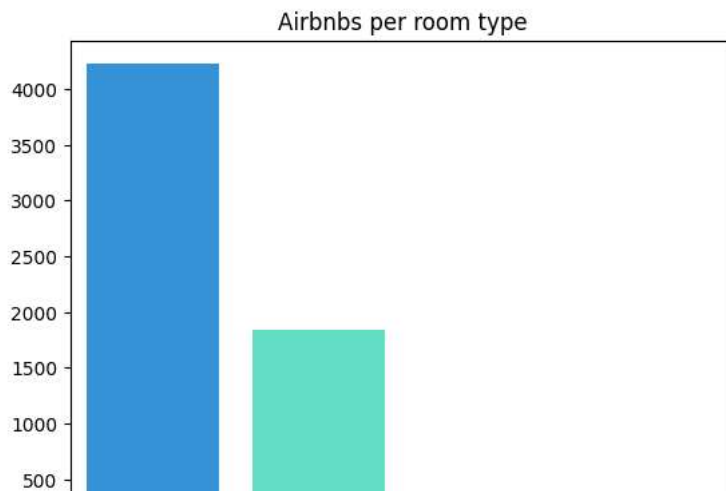
```
array([[<Axes: title={'center': 'review_scores_rating'}>]], dtype=object)
```



```
%matplotlib inline
airbnb_by_rooms = df["room_type"].value_counts()
x=airbnb_by_rooms.index
y=airbnb_by_rooms.values
sns.barplot(x=x, y=y, palette="rainbow")
plt.title("Airbnbs per room type")
```



```
Text(0.5, 1.0, 'Airbnbs per room type')
```

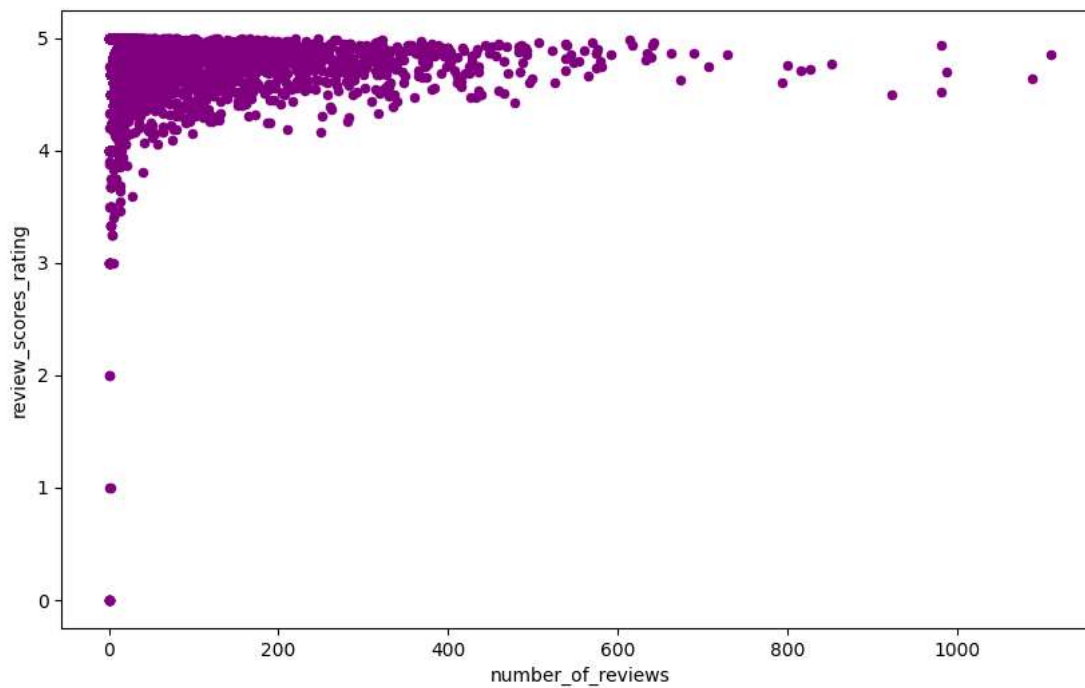


```
df["room_type"].value_counts().nlargest(1)
```

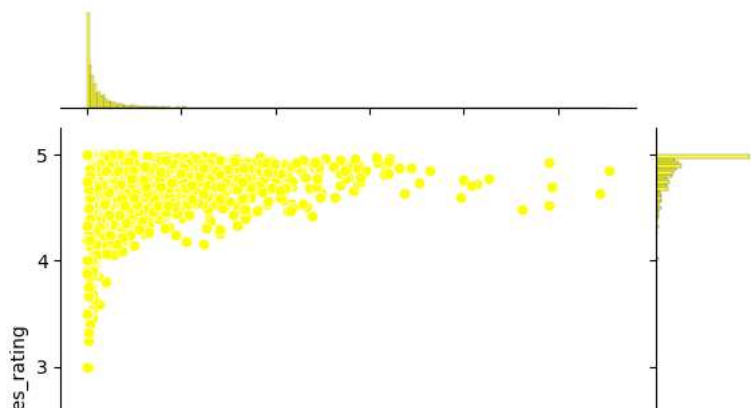
```
Entire home/apt    4223
Name: room_type, dtype: int64
```

```
df.plot(kind='scatter',x='number_of_reviews', y='review_scores_rating', figsize=(10,6),color="purple")
```

```
<Axes: xlabel='number_of_reviews', ylabel='review_scores_rating'>
```



```
sns.jointplot(data=df,kind='scatter',x='number_of_reviews', y='review_scores_rating',color="yellow")
plt.show()
```

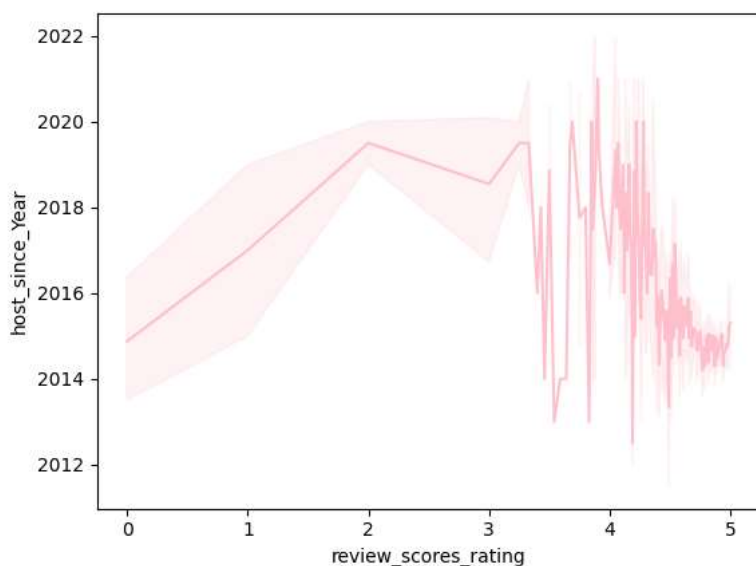


```
import seaborn as sns
```

```
df = df[df['es_rating'] > 4]
```

```
sns.lineplot(x = "review_scores_rating", y = "host_since_Year", data=df, color="pink")
```

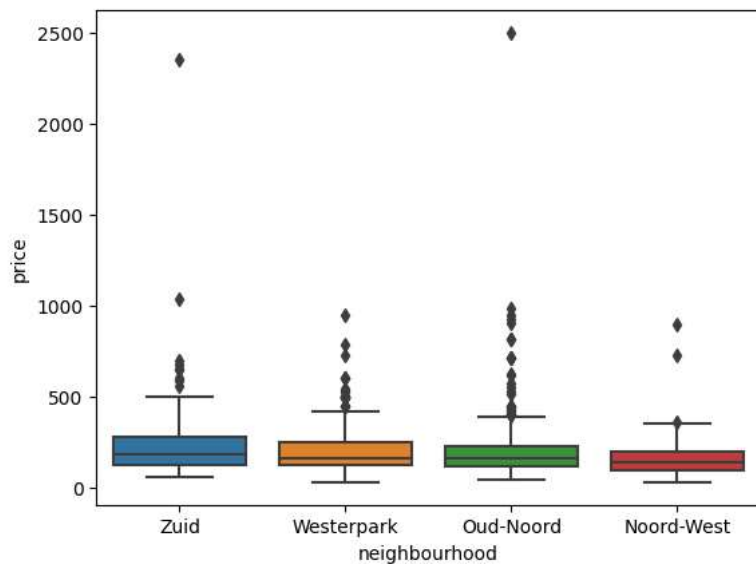
```
<Axes: xlabel='review_scores_rating', ylabel='host_since_Year'>
```



```
df2=df[(df.neighbourhood=='Westerpark') | (df.neighbourhood=='Oud-Noord') | (df.neighbourhood=='Noord-West') | (df.neighbourhood=='Zuid')]
```

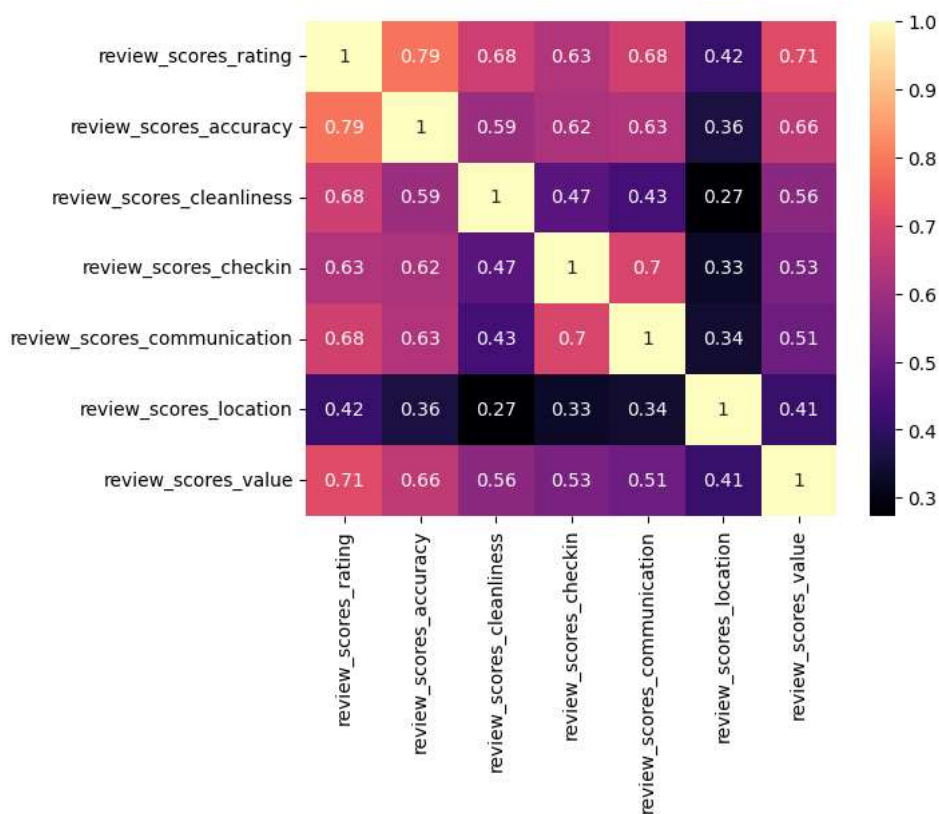
```
sns.boxplot(data=df2, y='price', x=df2['neighbourhood'])
```

```
<Axes: xlabel='neighbourhood', ylabel='price'>
```



```
review_columns = ['review_scores_rating',
                  'review_scores_accuracy',
                  'review_scores_cleanliness',
                  'review_scores_checkin',
                  'review_scores_communication',
                  'review_scores_location',
                  'review_scores_value']
```

```
dataplot=sns.heatmap(df.loc[:,review_columns].corr(),cmap="magma",annot=True)
# displaying heatmap
plt.show()
```



[Colab paid products](#) - [Cancel contracts here](#)

