

```
import pandas as pd
import matplotlib.pyplot as plt
from matplotlib import rcParams
import seaborn as sns

df=pd.read_csv('/content/House Price India.csv')
df.head()
```

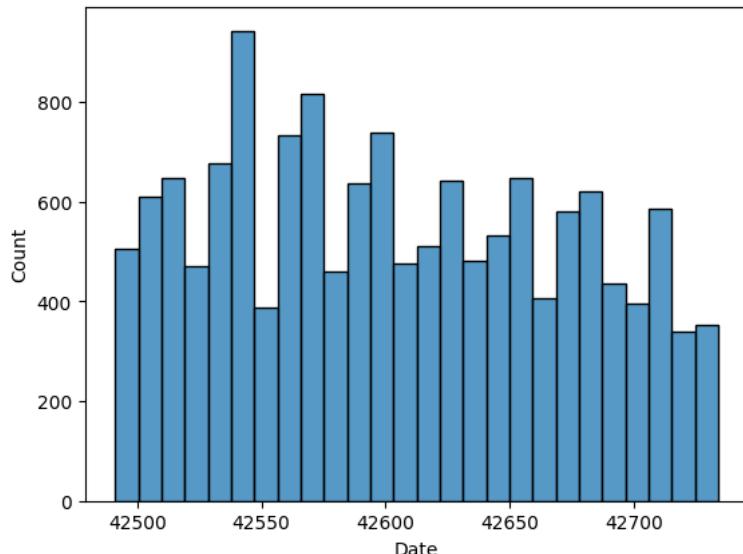
	id	Date	number of bedrooms	number of bathrooms	living area	lot area	number of floors	waterfront present	number of views	condition of the house	...	Built Year	Rer
0	6762810145	42491	5	2.50	3650	9050	2.0	0	4	5	...	1921	
1	6762810635	42491	4	2.50	2920	4000	1.5	0	0	5	...	1909	
2	6762810998	42491	5	2.75	2910	9480	1.5	0	0	3	...	1939	
3	6762812605	42491	4	2.50	3310	42998	2.0	0	0	3	...	2001	
4	6762812919	42491	3	2.00	2710	4500	1.5	0	0	4	...	1929	

5 rows x 23 columns

UNIVARIATE

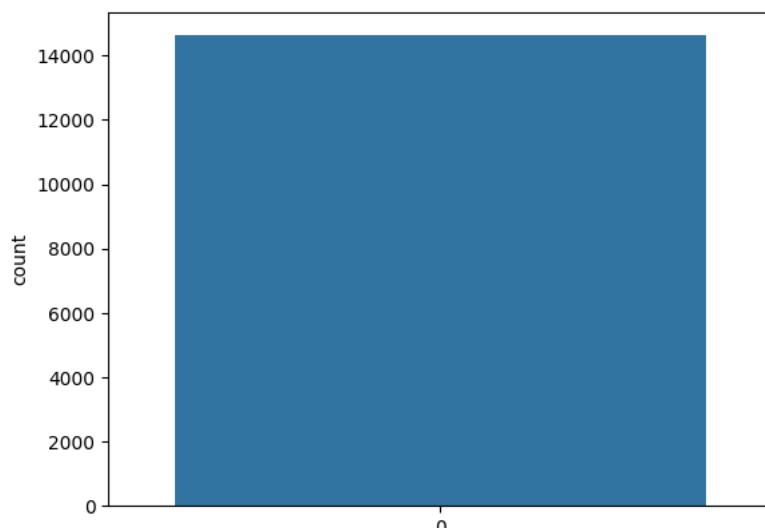
```
sns.histplot(df.Date)
```

<Axes: xlabel='Date', ylabel='Count'>



```
sns.countplot(df.id)
```

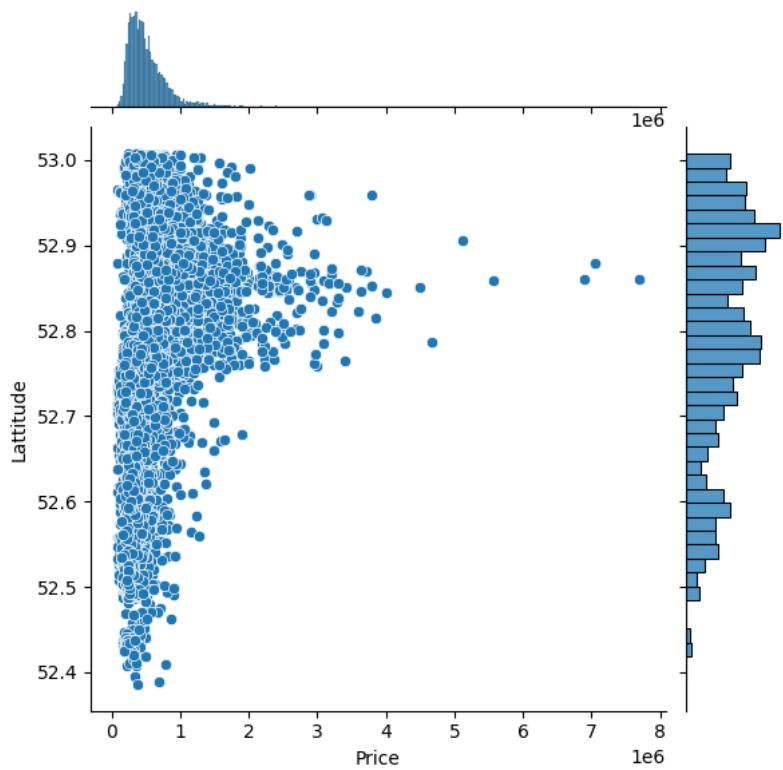
<Axes: ylabel='count'>



BIVARIATE

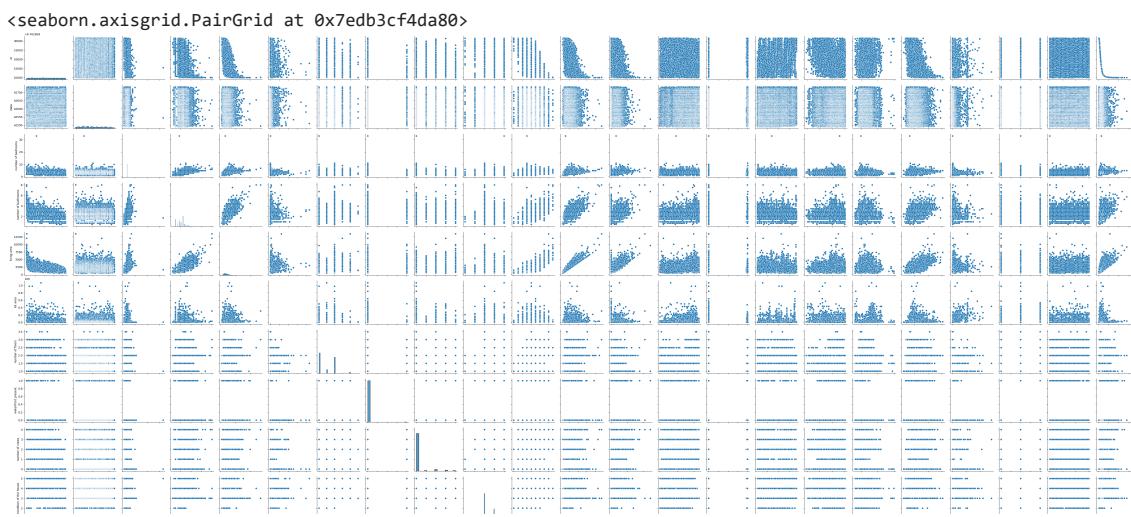
```
sns.jointplot(x='Price',y='Latitude',data=df)
```

```
<seaborn.axisgrid.JointGrid at 0x7edb3d5669e0>
```



MULTIVARIATE

```
sns.pairplot(df)
```



```
df.describe()
```

number of floors	waterfront present	number of views	condition of the house	...	Built Year	Renovation Year
14620.000000	14620.000000	14620.000000	14620.000000	...	14620.000000	14620.000000
1.502360	0.007661	0.233105	3.430506	...	1970.926402	90.924001
0.540239	0.087193	0.766259	0.664151	...	29.493625	416.216666
1.000000	0.000000	0.000000	1.000000	...	1900.000000	0.000000
1.000000	0.000000	0.000000	3.000000	...	1951.000000	0.000000
1.500000	0.000000	0.000000	3.000000	...	1975.000000	0.000000
2.000000	0.000000	0.000000	4.000000	...	1997.000000	0.000000
3.500000	1.000000	4.000000	5.000000	...	2015.000000	2015.000000

```
df.isnull().any() # no null values
```

```

id                         False
Date                        False
number of bedrooms          False
number of bathrooms          False
living area                  False
lot area                     False
number of floors             False
waterfront present          False
number of views              False
condition of the house      False
grade of the house          False
Area of the house(excluding basement) False
Area of the basement        False
Built Year                   False
Renovation Year              False
Postal Code                  False
Latitude                     False
Longitude                    False
living_area_renov            False
lot_area_renov               False
Number of schools nearby    False
Distance from the airport    False
Price                        False
dtype: bool

```

```
df.dropna() #drops null values
```

	id	Date	number of bedrooms	number of bathrooms	living area	lot area	number of floors	waterfront present	number of views	condition of the house	...	Built Year
0	6762810145	42491	5	2.50	3650	9050	2.0	0	4	5	...	1921
1	6762810635	42491	4	2.50	2920	4000	1.5	0	0	5	...	1909
2	6762810998	42491	5	2.75	2910	9480	1.5	0	0	3	...	1939
3	6762812605	42491	4	2.50	3310	42998	2.0	0	0	3	...	2001
4	6762812919	42491	3	2.00	2710	4500	1.5	0	0	4	...	1929
...
14610	6762830618	42734	2	1.00	1070	6120	1.0	0	0	3	...	1962
14618	6762830709	42734	4	1.00	1030	6621	1.0	0	0	4	...	1955
14619	6762831463	42734	3	1.00	900	4770	1.0	0	0	3	...	1969

14620 rows × 23 columns

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

✓ 0s completed at 6:29 PM

