

AI-ASSIGNMENT -01

| | |
|---------------------|---------------|
| Assignment Date | 24 APRIL 2023 |
| Student Name | A.SATHIYAN |
| Student Roll Number | 923020106303 |
| Maximum Marks | 2 Marks |

01 Importing the required libraries

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

02 LOAD THE DATASET

```
#READ THE DATASET
```

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

```
df.head()
```

```
      id  Date  number of bedrooms  number of bathrooms  living area \
0  6762810145  42491             5             2.50         3650
1  6762810635  42491             4             2.50         2920
2  6762810998  42491             5             2.75         2910
3  6762812605  42491             4             2.50         3310
4  6762812919  42491             3             2.00         2710
```

```
      lot area  number of floors  waterfront present  number of views \
0         9050             2.0             0             4
1         4000             1.5             0             0
2         9480             1.5             0             0
3        42998             2.0             0             0
4         4500             1.5             0             0
```

```
      condition of the house  ...  Built Year  Renovation Year  Postal Code \
0              5  ...      1921             0      122003
1              5  ...      1909             0      122004
2              3  ...      1939             0      122004
3              3  ...      2001             0      122005
4              4  ...      1929             0      122006
```

```
      Latitude  Longitude  living_area_renov  lot_area_renov \
0    52.8645    -114.557         2880         5400
1    52.8878    -114.470         2470         4000
```

| | | | | |
|---|---------|----------|------|-------|
| 2 | 52.8852 | -114.468 | 2940 | 6600 |
| 3 | 52.9532 | -114.321 | 3350 | 42847 |
| 4 | 52.9047 | -114.485 | 2060 | 4500 |

| | Number of schools nearby | Distance from the airport | Price |
|---|--------------------------|---------------------------|---------|
| 0 | 2 | 58 | 2380000 |
| 1 | 2 | 51 | 1400000 |
| 2 | 1 | 53 | 1200000 |
| 3 | 3 | 76 | 838000 |
| 4 | 1 | 51 | 805000 |

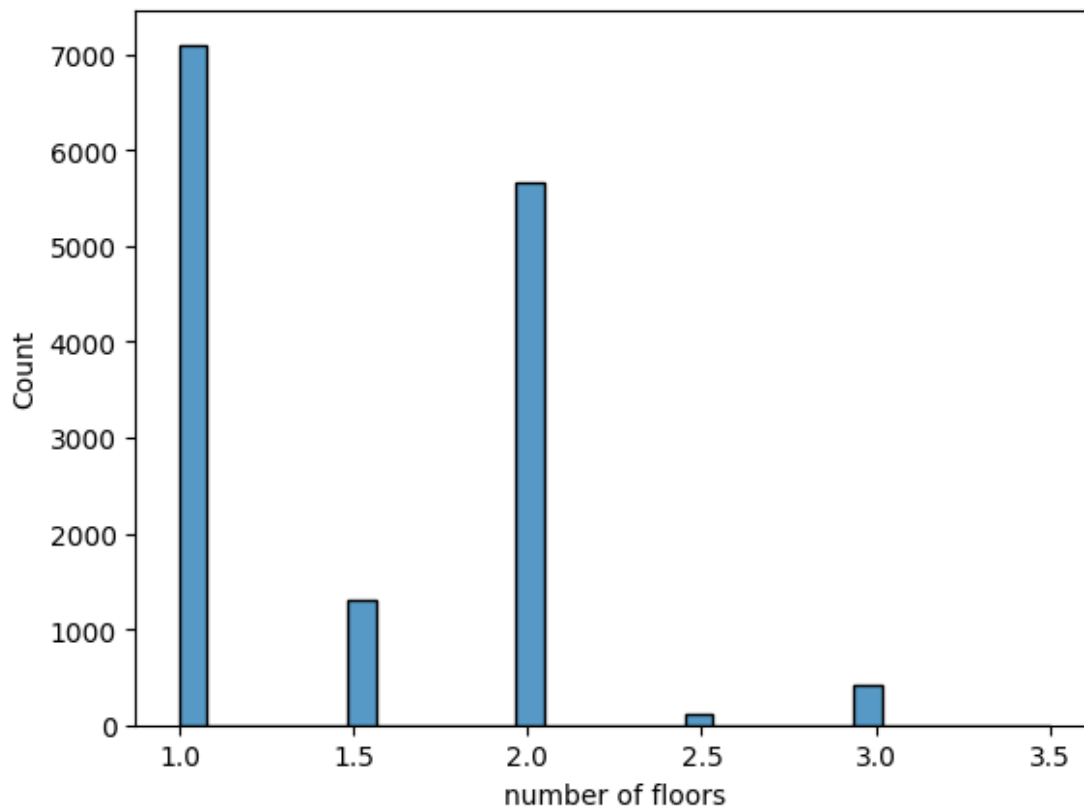
[5 rows x 23 columns]

03 Perform the Below Visualizations

(3a) *Univariate Analysis*

```
sns.histplot(df['number of floors'])
```

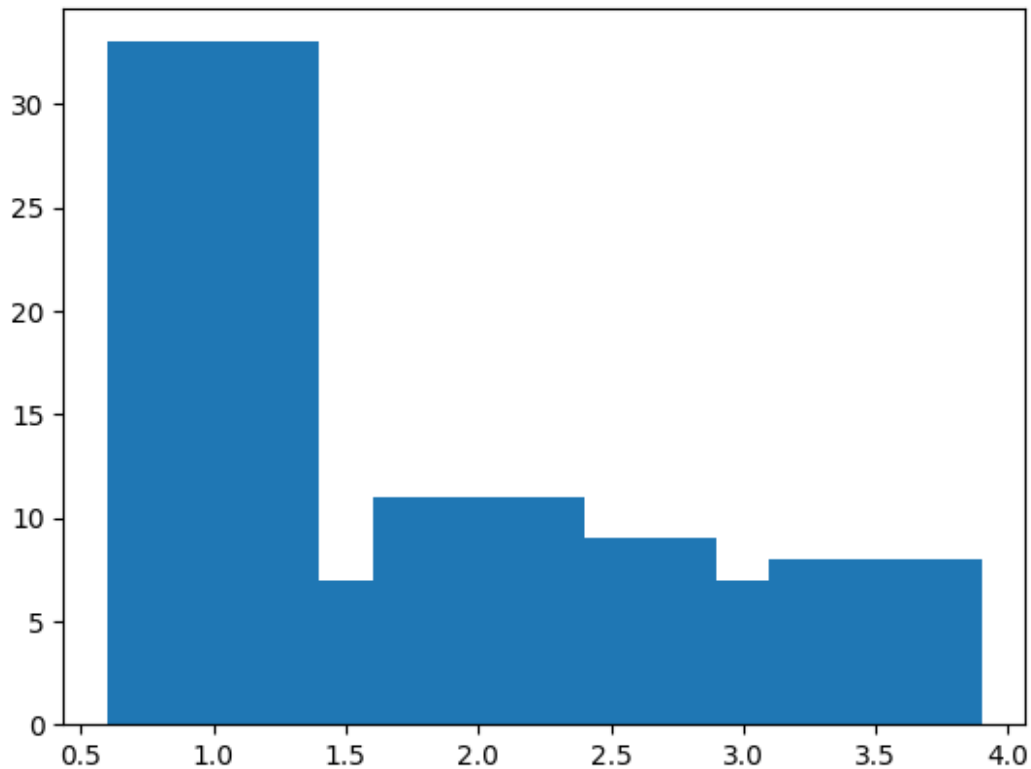
<Axes: xlabel='number of floors', ylabel='Count'>



```
#(3b)Bi - Variate Analysis
```

```
plt.bar(df['number of floors'],df['number of bedrooms'])
```

```
<BarContainer object of 14620 artists>
```



```
# (3c) Multi-Variate Analysis
```

```
import seaborn as sns
```

```
/usr/local/lib/python3.9/dist-packages/seaborn/axisgrid.py:2095: UserWarning:  
The `size` parameter has been renamed to `height`; please update your code.  
warnings.warn(msg, UserWarning)
```

```
<seaborn.axisgrid.PairGrid at 0x7fac541132e0>
```

```
sns.pairplot(df,hue='Built Year',size= 3)
```

```
/usr/local/lib/python3.9/dist-packages/seaborn/axisgrid.py:2095: UserWarning:  
The `size` parameter has been renamed to `height`; please update your code.  
warnings.warn(msg, UserWarning)
```

```
<seaborn.axisgrid.PairGrid at 0x7fac32bff670>
```

04 Perform descriptive statistics on the dataset

```
df['living area'].describe()
```

```
count    14620.000000
mean      2098.262996
std       928.275721
min       370.000000
25%      1440.000000
50%      1930.000000
75%      2570.000000
max      13540.000000
Name: living area, dtype: float64
```

```
df['Number of schools nearby'].describe()
```

```
count    14620.000000
mean       2.012244
std       0.817284
min       1.000000
25%       1.000000
50%       2.000000
75%       3.000000
max       3.000000
Name: Number of schools nearby, dtype: float64
```

05 Handle the Missing values

```
df.isnull()
```

| | id | Date | number of bedrooms | number of bathrooms | living area | \ |
|-------|-------|-------|--------------------|---------------------|-------------|---|
| 0 | False | False | False | False | False | |
| 1 | False | False | False | False | False | |
| 2 | False | False | False | False | False | |
| 3 | False | False | False | False | False | |
| 4 | False | False | False | False | False | |
| ... | ... | ... | ... | ... | ... | |
| 14615 | False | False | False | False | False | |
| 14616 | False | False | False | False | False | |
| 14617 | False | False | False | False | False | |
| 14618 | False | False | False | False | False | |
| 14619 | False | False | False | False | False | |

| | lot area | number of floors | waterfront present | number of views | \ |
|-------|----------|------------------|--------------------|-----------------|---|
| 0 | False | False | False | False | |
| 1 | False | False | False | False | |
| 2 | False | False | False | False | |
| 3 | False | False | False | False | |
| 4 | False | False | False | False | |
| ... | ... | ... | ... | ... | |
| 14615 | False | False | False | False | |

| | | | | |
|-------|-------|-------|-------|-------|
| 14616 | False | False | False | False |
| 14617 | False | False | False | False |
| 14618 | False | False | False | False |
| 14619 | False | False | False | False |

| | condition of the house | ... | Built Year | Renovation Year | Postal Code |
|-------|------------------------|-----|------------|-----------------|-------------|
| \ | | | | | |
| 0 | False | ... | False | False | False |
| 1 | False | ... | False | False | False |
| 2 | False | ... | False | False | False |
| 3 | False | ... | False | False | False |
| 4 | False | ... | False | False | False |
| ... | ... | ... | ... | ... | ... |
| 14615 | False | ... | False | False | False |
| 14616 | False | ... | False | False | False |
| 14617 | False | ... | False | False | False |
| 14618 | False | ... | False | False | False |
| 14619 | False | ... | False | False | False |

| | Lattitude | Longitude | living_area_renov | lot_area_renov | \ |
|-------|-----------|-----------|-------------------|----------------|---|
| 0 | False | False | False | False | |
| 1 | False | False | False | False | |
| 2 | False | False | False | False | |
| 3 | False | False | False | False | |
| 4 | False | False | False | False | |
| ... | ... | ... | ... | ... | |
| 14615 | False | False | False | False | |
| 14616 | False | False | False | False | |
| 14617 | False | False | False | False | |
| 14618 | False | False | False | False | |
| 14619 | False | False | False | False | |

| | Number of schools nearby | Distance from the airport | Price |
|-------|--------------------------|---------------------------|-------|
| 0 | False | False | False |
| 1 | False | False | False |
| 2 | False | False | False |
| 3 | False | False | False |
| 4 | False | False | False |
| ... | ... | ... | ... |
| 14615 | False | False | False |
| 14616 | False | False | False |
| 14617 | False | False | False |
| 14618 | False | False | False |
| 14619 | False | False | False |

[14620 rows x 23 columns]