Assignment 3

In [40]:

```
from matplotlib import pyplot as plt
import pandas as pd

df = pd.read_csv("houses.csv")
df
```

Out[40]:

	POSTED_BY	UNDER_CONSTRUCTION	RERA	BHK_NO.	BHK_OR_RK	SQUARE_FT	REA
0	Owner	0	0	2	BHK	1300.236407	
1	Dealer	0	0	2	BHK	1275.000000	
2	Owner	0	0	2	BHK	933.159722	
3	Owner	0	1	2	BHK	929.921143	
4	Dealer	1	0	2	BHK	999.009247	
...	...	...	...	...	...	...	
29446	Owner	0	0	3	BHK	2500.000000	
29447	Owner	0	0	2	BHK	769.230769	
29448	Dealer	0	0	2	BHK	1022.641509	
29449	Owner	0	0	2	BHK	927.079009	
29450	Dealer	0	1	2	BHK	896.774194	

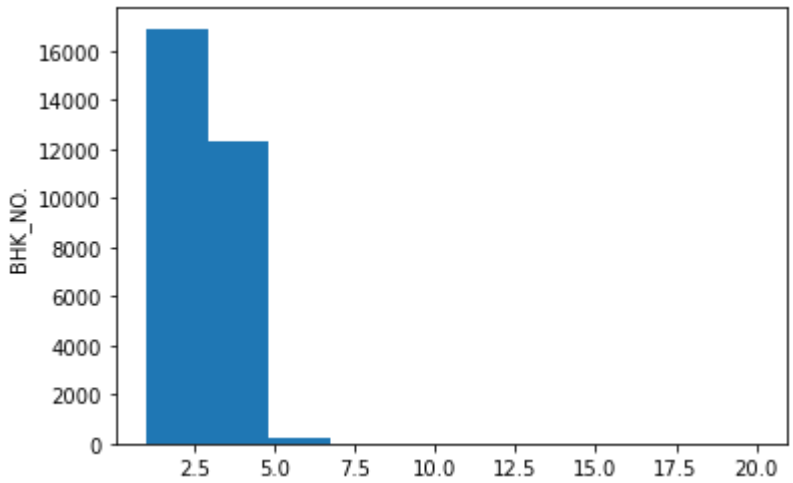
29451 rows × 12 columns

Histogram

- A histogram shows the frequency on the vertical axis and the horizontal axis is another dimension. Usually it has bins where every bin has as minimum and maximum value.
- Each bin has a frequency between x and infinite

In [35]:

```
plt.hist(df["BHK_NO."])
plt.ylabel("BHK_NO.")
plt.show()
```

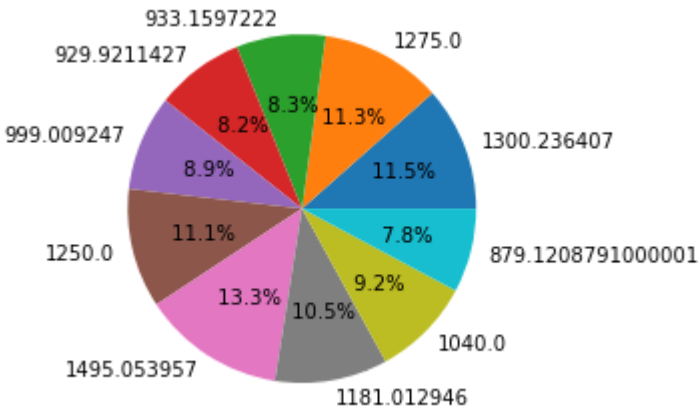


Pie Chart

- A pie chart can only display one series of data. Pie chart shows the size of items in one data series, proportional to the sum of the items.
- Matplot API has pie() function that generates a pie diagram representing data in a array. The fractional area of each wedge is given by x/sum(x)

In [36]:

```
plt.pie(df["SQUARE_FT"][:10],labels = df["SQUARE_FT"][:10], autopct = "% 1.1f%%")
plt.show()
```

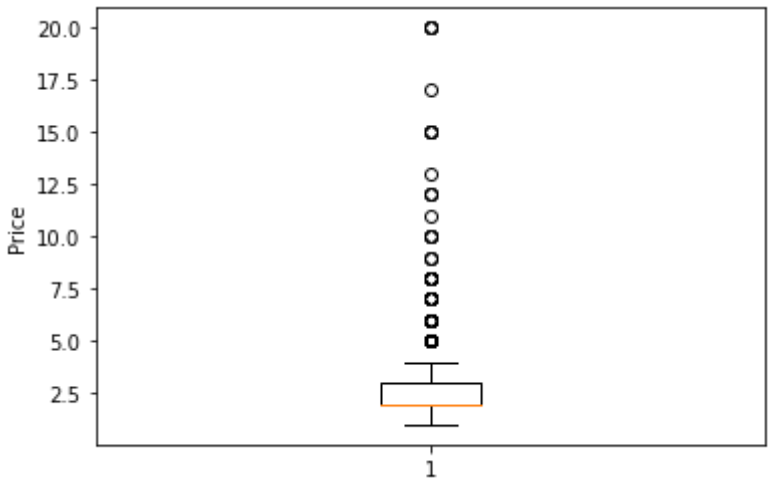


Box Plot Chart

- A box plot represents the summary of a set of data containing the minimum, first quartile, median, third quartile and maximum.

In [37]:

```
plt.boxplot(df["BHK_NO."])
plt.ylabel("Price")
plt.show()
```

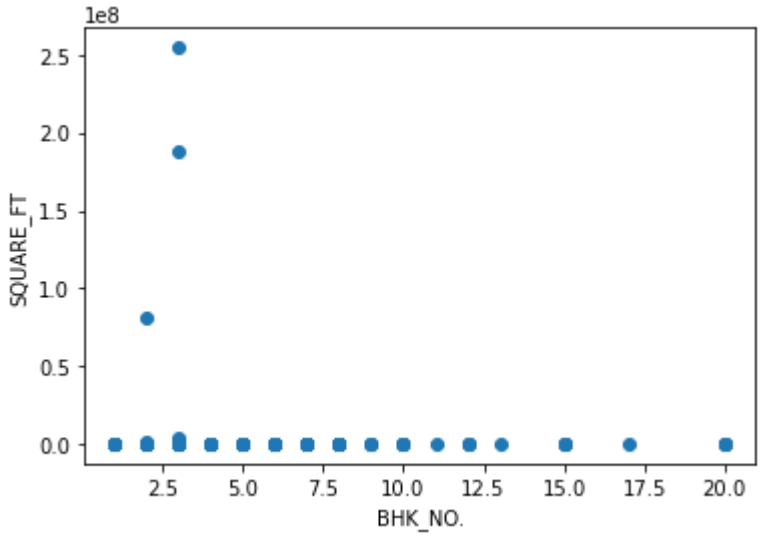


Scatter Plot

- A scatter plot is a type of plot that shows the data as a collection of points.
- The position of a point depends on its two dimensional value, where each vaue is a position on either the horizontal or vertical dimensions

In [38]:

```
plt.scatter(df["BHK_NO."],df["SQUARE_FT"])
plt.xlabel("BHK_NO.")
plt.ylabel("SQUARE_FT")
plt.show()
```



Column Chart or Bar graph

- Column chart or commonly called as bar chart is a chart that presents categorical data with rectangular bars to the values that they represent. - A bar chart shows comparisions among discrete categories. One axis of the chart shows the specific categories being compared and the other axis represents a measured value.

In [39]:

```
plt.bar(df["POSTED_BY"],df["BHK_NO."])
plt.ylabel("BHK_NO.")
plt.show()
```