```
In [1]: import pandas as pd
        import numpy as np
```

```
In [2]: data=pd.read_csv(r"D:\Download2\HousePrice.csv")
```

```
In [3]: data
```

Out[3]:

	POSTED_BY	UNDER_CONSTRUCTION	RERA	BHK_NO.	BHK_OR_RK	SQUARE_FT	READY
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

```
In [4]: data["POSTED_BY"].unique()
```

Out[4]: array(['Owner', 'Dealer', 'Builder'], dtype=object)

```
In [5]: data["POSTED_BY"].replace("Owner",1,inplace=True)
```

```
In [6]: data["POSTED_BY"].replace("Dealer",0,inplace=True)
```

```
In [7]: data["POSTED_BY"].replace("Builder",2,inplace=True)
```

In [8]: `data`

Out[8]:

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

29451 rows × 12 columns

In [9]: `data["POSTED_BY"].unique()`

Out[9]: `array([1, 0, 2], dtype=int64)`

In [10]:
```python
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 29451 entries, 0 to 29450
Data columns (total 12 columns):
 #   Column                 Non-Null Count  Dtype
---  ------                 --------------  -----
 0   POSTED_BY              29451 non-null  int64
 1   UNDER_CONSTRUCTION     29451 non-null  int64
 2   RERA                   29451 non-null  int64
 3   BHK_NO.                29451 non-null  int64
 4   BHK_OR_RK              29451 non-null  object
 5   SQUARE_FT              29451 non-null  float64
 6   READY_TO_MOVE          29451 non-null  int64
 7   RESALE                 29451 non-null  int64
 8   ADDRESS                29451 non-null  object
 9   LONGITUDE              29451 non-null  float64
 10  LATITUDE               29451 non-null  float64
 11  TARGET(PRICE_IN_LACS)  29451 non-null  float64
dtypes: float64(4), int64(6), object(2)
memory usage: 2.7+ MB
```

In [11]:
```python
data=pd.read_csv(r"D:\Download2\HousePrice.csv")
data
```

Out[11]:

	POSTED_BY	UNDER_CONSTRUCTION	RERA	BHK_NO.	BHK_OR_RK	SQUARE_FT	READY
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

In [12]:
```python
a1=data["POSTED_BY"].str.replace("Own","self")
a1
```

Out[12]:
```
0        selfer
1        Dealer
2        selfer
3        selfer
4        Dealer
          ...
29446    selfer
29447    selfer
29448    Dealer
29449    selfer
29450    Dealer
Name: POSTED_BY, Length: 29451, dtype: object
```

In [13]:
```python
data=pd.read_csv(r"D:\Download2\HousePrice.csv")
data
```

Out[13]:

	POSTED_BY	UNDER_CONSTRUCTION	RERA	BHK_NO.	BHK_OR_RK	SQUARE_FT	READ\
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

In [14]:
```python
data["POSTED_BY"]=data["POSTED_BY"].str.replace("Own","self")
data
```

Out[14]:

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

29451 rows × 12 columns

In [15]:
```python
old_str=input("Enter the value..... ")
New_str=input("Enter the new values.......")
data["POSTED_BY"]=data["POSTED_BY"].str.replace(old_str,New_str)
```

```
Enter the value..... Dealer
Enter the new values.......broker
```

In [16]: `data`

Out[16]:

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

29451 rows × 12 columns

◄ | | ►

How to Make Round of Value

In [17]: `np.round(data["SQUARE_FT"],2)`

Out[17]:
```
0        1300.24
1        1275.00
2         933.16
3         929.92
4         999.01
          ...
29446    2500.00
29447     769.23
29448    1022.64
29449     927.08
29450     896.77
Name: SQUARE_FT, Length: 29451, dtype: float64
```

In [18]: `data["SQUARE_FT"]=np.round(data["SQUARE_FT"])`

In [19]: `data`

Out[19]:

	POSTED_BY	UNDER_CONSTRUCTION	RERA	BHK_NO.	BHK_OR_RK	SQUARE_FT	READY
0	selfer	0	0	2	BHK	1300.0	
1	broker	0	0	2	BHK	1275.0	
2	selfer	0	0	2	BHK	933.0	
3	selfer	0	1	2	BHK	930.0	
4	broker	1	0	2	BHK	999.0	
...	...	...	...	...	...	...	
29446	selfer	0	0	3	BHK	2500.0	
29447	selfer	0	0	2	BHK	769.0	
29448	broker	0	0	2	BHK	1023.0	
29449	selfer	0	0	2	BHK	927.0	
29450	broker	0	1	2	BHK	897.0	

29451 rows × 12 columns

In [20]: `data=np.round(data,2)`

In [21]: `data`

Out[21]:

	POSTED_BY	UNDER_CONSTRUCTION	RERA	BHK_NO.	BHK_OR_RK	SQUARE_FT	READY
0	selfer	0	0	2	BHK	1300.0	
1	broker	0	0	2	BHK	1275.0	
2	selfer	0	0	2	BHK	933.0	
3	selfer	0	1	2	BHK	930.0	
4	broker	1	0	2	BHK	999.0	
...	...	...	...	...	...	...	
29446	selfer	0	0	3	BHK	2500.0	
29447	selfer	0	0	2	BHK	769.0	
29448	broker	0	0	2	BHK	1023.0	
29449	selfer	0	0	2	BHK	927.0	
29450	broker	0	1	2	BHK	897.0	

29451 rows × 12 columns

In [22]:
```python
a=12.34556
ab=np.round(a,2)
ab
```

Out[22]: 12.35

How to work with For Loop when dataset we have...

In [23]:
```python
for x in range(3):
    print(x)
```

```
0
1
2
```

In [24]:
```python
for x in data["POSTED_BY"]:
    print(x)
```

```
selrer
broker
broker
broker
selfer
broker
selfer
broker
selfer
selfer
selfer
broker
selfer
broker
selfer
broker
selfer
broker
selfer
selfer
```

In [25]:
```python
data["POSTED_BY"].unique()
```

Out[25]: array(['selfer', 'broker', 'Builder'], dtype=object)

In [26]:
```python
ab=[]
for x in data["POSTED_BY"]:
    if x=="broker":
        ab.append(1)
    elif x=="selfer":
        ab.append(2)
    else:
        ab.append(3)
```

In [29]: `ab`

```
1,
1,
1,
2,
1,
1,
2,
1,
1,
1,
1,
2,
2,
1,
1,
1,
2,
1,
2,
...]
```

In [30]: `data`

Out[30]:

	POSTED_BY	UNDER_CONSTRUCTION	RERA	BHK_NO.	BHK_OR_RK	SQUARE_FT	READY
0	selfer	0	0	2	BHK	1300.0	
1	broker	0	0	2	BHK	1275.0	
2	selfer	0	0	2	BHK	933.0	
3	selfer	0	1	2	BHK	930.0	
4	broker	1	0	2	BHK	999.0	
...	...	...	...	...	...	...	
29446	selfer	0	0	3	BHK	2500.0	
29447	selfer	0	0	2	BHK	769.0	
29448	broker	0	0	2	BHK	1023.0	
29449	selfer	0	0	2	BHK	927.0	
29450	broker	0	1	2	BHK	897.0	

29451 rows × 12 columns

Create a new column in current dataset

In [31]: `data["Frenchise"]=ab`

In [32]: `data`

Out[32]:

	POSTED_BY	UNDER_CONSTRUCTION	RERA	BHK_NO.	BHK_OR_RK	SQUARE_FT	READY
0	selfer	0	0	2	BHK	1300.0	
1	broker	0	0	2	BHK	1275.0	
2	selfer	0	0	2	BHK	933.0	
3	selfer	0	1	2	BHK	930.0	
4	broker	1	0	2	BHK	999.0	
...	...	...	...	...	...	...	
29446	selfer	0	0	3	BHK	2500.0	
29447	selfer	0	0	2	BHK	769.0	
29448	broker	0	0	2	BHK	1023.0	
29449	selfer	0	0	2	BHK	927.0	
29450	broker	0	1	2	BHK	897.0	

29451 rows × 13 columns

In [33]: `data.insert(1,"French_Owner",ab)`

In [34]: data

Out[34]:

	POSTED_BY	French_Owner	UNDER_CONSTRUCTION	RERA	BHK_NO.	BHK_OR_RK	SQU
0	selfer	2	0	0	2	BHK	
1	broker	1	0	0	2	BHK	
2	selfer	2	0	0	2	BHK	
3	selfer	2	0	1	2	BHK	
4	broker	1	1	0	2	BHK	
...	...	...	...	...	...	...	
29446	selfer	2	0	0	3	BHK	
29447	selfer	2	0	0	2	BHK	
29448	broker	1	0	0	2	BHK	
29449	selfer	2	0	0	2	BHK	
29450	broker	1	0	1	2	BHK	

29451 rows × 14 columns
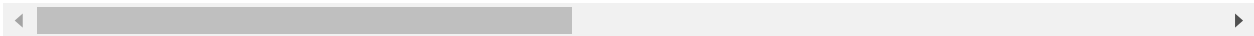
How to remove col from dataset

In [35]: **del** data['POSTED_BY']

In [36]: `data`

Out[36]:

	French_Owner	UNDER_CONSTRUCTION	RERA	BHK_NO.	BHK_OR_RK	SQUARE_FT	REA
0	2	0	0	2	BHK	1300.0	
1	1	0	0	2	BHK	1275.0	
2	2	0	0	2	BHK	933.0	
3	2	0	1	2	BHK	930.0	
4	1	1	0	2	BHK	999.0	
...	...	...	...	...	...	...	
29446	2	0	0	3	BHK	2500.0	
29447	2	0	0	2	BHK	769.0	
29448	1	0	0	2	BHK	1023.0	
29449	2	0	0	2	BHK	927.0	
29450	1	0	1	2	BHK	897.0	

29451 rows × 13 columns
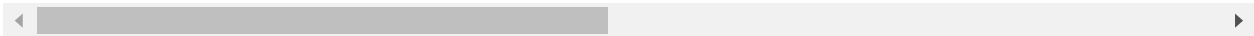
In [37]:
```python
del data['French_Owner']
```

In [38]: `data`

Out[38]:

	UNDER_CONSTRUCTION	RERA	BHK_NO.	BHK_OR_RK	SQUARE_FT	READY_TO_MOVE	
0	0	0	2	BHK	1300.0	1	
1	0	0	2	BHK	1275.0	1	
2	0	0	2	BHK	933.0	1	
3	0	1	2	BHK	930.0	1	
4	1	0	2	BHK	999.0	0	
...	...	...	...	...	...	...	
29446	0	0	3	BHK	2500.0	1	
29447	0	0	2	BHK	769.0	1	
29448	0	0	2	BHK	1023.0	1	
29449	0	0	2	BHK	927.0	1	
29450	0	1	2	BHK	897.0	1	

29451 rows × 12 columns

In [39]: `data["Frenchise"].value_counts()`

Out[39]:
```
1    18291
2    10538
3      622
Name: Frenchise, dtype: int64
```

In [40]: `dataset=data.drop('LATITUDE',1)`

```
C:\Users\Sai\AppData\Local\Temp\ipykernel_11268\3908617382.py:1: FutureWarning:
In a future version of pandas all arguments of DataFrame.drop except for the ar
gument 'labels' will be keyword-only.
  dataset=data.drop('LATITUDE',1)
```

In [41]: `dataset`

Out[41]:

	UNDER_CONSTRUCTION	RERA	BHK_NO.	BHK_OR_RK	SQUARE_FT	READY_TO_MOVE	
0	0	0	2	BHK	1300.0	1	
1	0	0	2	BHK	1275.0	1	
2	0	0	2	BHK	933.0	1	
3	0	1	2	BHK	930.0	1	
4	1	0	2	BHK	999.0	0	
...	...	...	...	...	...	...	
29446	0	0	3	BHK	2500.0	1	
29447	0	0	2	BHK	769.0	1	
29448	0	0	2	BHK	1023.0	1	
29449	0	0	2	BHK	927.0	1	
29450	0	1	2	BHK	897.0	1	

29451 rows × 11 columns

In [ ]: