In [4]:
```
import pandas as pd
```

In [5]:
```
df = pd.read_csv("Banglorefiles.csv")
```

In [6]:
```
df.head()
```

Out[6]:

	area_type	availability	location	size	society	total_sqft	bath	balcony	p
0	Super built-up Area	19-Dec	Electronic City Phase II	2 BHK	Coomee	1056	2.0	1.0	3
1	Plot Area	Ready To Move	Chikka Tirupathi	4 Bedroom	Theanmp	2600	5.0	3.0	12
2	Built-up Area	Ready To Move	Uttarahalli	3 BHK	NaN	1440	2.0	3.0	6
3	Super built-up Area	Ready To Move	Lingadheeranahalli	3 BHK	Soiewre	1521	3.0	1.0	9
4	Super built-up Area	Ready To Move	Kothanur	2 BHK	NaN	1200	2.0	1.0	5

In [7]:
```
df.head(n = 10)
```

Out[7]:

	area_type	availability	location	size	society	total_sqft	bath	balcony	p
0	Super built-up Area	19-Dec	Electronic City Phase II	2 BHK	Coomee	1056	2.0	1.0	3
1	Plot Area	Ready To Move	Chikka Tirupathi	4 Bedroom	Theanmp	2600	5.0	3.0	12
2	Built-up Area	Ready To Move	Uttarahalli	3 BHK	NaN	1440	2.0	3.0	6
3	Super built-up Area	Ready To Move	Lingadheeranahalli	3 BHK	Soiewre	1521	3.0	1.0	9
4	Super built-up Area	Ready To Move	Kothanur	2 BHK	NaN	1200	2.0	1.0	5
5	Super built-up Area	Ready To Move	Whitefield	2 BHK	DuenaTa	1170	2.0	1.0	3
6	Super built-up Area	18-May	Old Airport Road	4 BHK	Jaades	2732	4.0	NaN	20
7	Super built-up Area	Ready To Move	Rajaji Nagar	4 BHK	Brway G	3300	4.0	NaN	60
8	Super built-up Area	Ready To Move	Marathahalli	3 BHK	NaN	1310	3.0	1.0	6
9	Plot Area	Ready To Move	Gandhi Bazar	6 Bedroom	NaN	1020	6.0	NaN	37

In [8]:
```
df.tail()
```

Out[8]:

	area_type	availability	location	size	society	total_sqft	bath	balcony
13315	Built-up Area	Ready To Move	Whitefield	5 Bedroom	ArsiaEx	3453	4.0	0.0
13316	Super built-up Area	Ready To Move	Richards Town	4 BHK	NaN	3600	5.0	NaN
13317	Built-up Area	Ready To Move	Raja Rajeshwari Nagar	2 BHK	Mahla T	1141	2.0	1.0
13318	Super built-up Area	18-Jun	Padmanabhanagar	4 BHK	SollyCl	4689	4.0	1.0
13319	Super built-up Area	Ready To Move	Doddathoguru	1 BHK	NaN	550	1.0	1.0

In [9]:
```
df.tail(n = 10)
```

Out[9]:

	area_type	availability	location	size	society	total_sqft	bath	balcony
13310	Super built-up Area	Ready To Move	Rachenahalli	2 BHK	NaN	1050	2.0	2.0
13311	Plot Area	Ready To Move	Ramamurthy Nagar	7 Bedroom	NaN	1500	9.0	2.0
13312	Super built-up Area	Ready To Move	Bellandur	2 BHK	NaN	1262	2.0	2.0
13313	Super built-up Area	Ready To Move	Uttarahalli	3 BHK	Aklia R	1345	2.0	1.0
13314	Super built-up Area	Ready To Move	Green Glen Layout	3 BHK	SoosePr	1715	3.0	3.0
13315	Built-up Area	Ready To Move	Whitefield	5 Bedroom	ArsiaEx	3453	4.0	0.0
13316	Super built-up Area	Ready To Move	Richards Town	4 BHK	NaN	3600	5.0	NaN
13317	Built-up Area	Ready To Move	Raja Rajeshwari Nagar	2 BHK	Mahla T	1141	2.0	1.0
13318	Super built-up Area	18-Jun	Padmanabhanagar	4 BHK	SollyCl	4689	4.0	1.0
13319	Super built-up Area	Ready To Move	Doddathoguru	1 BHK	NaN	550	1.0	1.0

In [10]:
```
df.tail(-3)
```

Out[10]:

	area_type	availability	location	size	society	total_sqft	bath	balcony
3	Super built-up Area	Ready To Move	Lingadheeranahalli	3 BHK	Soiewre	1521	3.0	1.0
4	Super built-up Area	Ready To Move	Kothanur	2 BHK	NaN	1200	2.0	1.0
5	Super built-up Area	Ready To Move	Whitefield	2 BHK	DuenaTa	1170	2.0	1.0
6	Super built-up Area	18-May	Old Airport Road	4 BHK	Jaades	2732	4.0	NaN
7	Super built-up Area	Ready To Move	Rajaji Nagar	4 BHK	Brway G	3300	4.0	NaN
...	...	...	...	...	...	...	...	...
13315	Built-up Area	Ready To Move	Whitefield	5 Bedroom	ArsiaEx	3453	4.0	0.0
13316	Super built-up Area	Ready To Move	Richards Town	4 BHK	NaN	3600	5.0	NaN
13317	Built-up Area	Ready To Move	Raja Rajeshwari Nagar	2 BHK	Mahla T	1141	2.0	1.0
13318	Super built-up Area	18-Jun	Padmanabhanagar	4 BHK	SollyCl	4689	4.0	1.0
13319	Super built-up Area	Ready To Move	Doddathoguru	1 BHK	NaN	550	1.0	1.0

13317 rows × 9 columns

In [11]:
```
df.index
```

Out[11]:
```
RangeIndex(start=0, stop=13320, step=1)
```

In [14]:
```
df.columns
```

Out[14]:
```
Index(['area_type', 'availability', 'location', 'size', 'society',
       'total_sqft', 'bath', 'balcony', 'price'],
      dtype='object')
```

In [15]:
```
df.dtypes
```

Out[15]:
```
area_type        object
availability     object
location         object
size             object
society          object
total_sqft       object
bath            float64
balcony         float64
price           float64
dtype: object
```

In [16]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 13320 entries, 0 to 13319
Data columns (total 9 columns):
 #   Column        Non-Null Count  Dtype  
---  ------        --------------  -----  
 0   area_type     13320 non-null  object 
 1   availability  13320 non-null  object 
 2   location      13319 non-null  object 
 3   size          13304 non-null  object 
 4   society       7818 non-null   object 
 5   total_sqft    13320 non-null  object 
 6   bath          13247 non-null  float64
 7   balcony       12711 non-null  float64
 8   price         13320 non-null  float64
dtypes: float64(3), object(6)
memory usage: 936.7+ KB
```

In [18]:
```
df[df["bath"]>30]
```

Out[18]:

	area_type	availability	location	size	society	total_sqft	bath	balcony	price
4684	Plot Area	Ready To Move	Munnekollal	43 Bedroom	NaN	2400	40.0	0.0	660.0

In [19]:
```
df.sort_values(by = "size")
```

Out[19]:

	area_type	availability	location	size	society	total_sqft	bath	balcony	price
13319	Super built-up Area	Ready To Move	Doddathoguru	1 BHK	NaN	550	1.0	1.0	17.00
12231	Super built-up Area	Ready To Move	Bellandur	1 BHK	Trows M	950	1.0	1.0	40.00
7008	Super built-up Area	Ready To Move	Kudlu Gate	1 BHK	NaN	1400	1.0	1.0	285.00
7747	Super built-up Area	18-Nov	Thanisandra	1 BHK	Bhe 2ko	620 - 934	1.0	0.0	38.53
9246	Super built-up Area	21-Jun	Vijayanagar	1 BHK	Saahaat	492	1.0	1.0	40.00
...	...	...	...	...	...	...	...	...	..
8565	Plot Area	Immediate Possession	Carmelaram	NaN	Ocezes	2000	NaN	NaN	120.00
8703	Plot Area	Immediate Possession	Mysore Road	NaN	Naoden	1200 - 2400	NaN	NaN	42.30
10634	Plot Area	Immediate Possession	Rajankunte	NaN	NaN	1575	NaN	NaN	31.11
11019	Plot Area	Immediate Possession	Yelahanka	NaN	Saaveha	1200 - 1800	NaN	NaN	12.75
11569	Plot Area	Immediate Possession	Hosur Road	NaN	AVeldun	1350	NaN	NaN	8.44

13320 rows × 9 columns

In [20]:
```
df.sort_values(by ="size" , ascending = False)
```

Out[20]:

	area_type	availability	location	size	society	total_sqft	bath	balcony	p
6538	Plot Area	Ready To Move	Hoskote	9 Bedroom	NaN	1800	10.0	3.0	18
648	Built-up Area	Ready To Move	Arekere	9 Bedroom	NaN	4125Perch	9.0	NaN	26
1593	Plot Area	Ready To Move	JP Nagar	9 Bedroom	NaN	2550	9.0	0.0	36
1299	Plot Area	Ready To Move	Chamrajpet	9 Bedroom	NaN	4050	7.0	1.0	120
13221	Plot Area	Ready To Move	T Dasarahalli	9 Bedroom	NaN	1178	9.0	1.0	7
...	...	...	...	...	...	...	...	...	
8565	Plot Area	Immediate Possession	Carmelaram	NaN	Ocezes	2000	NaN	NaN	12
8703	Plot Area	Immediate Possession	Mysore Road	NaN	Naoden	1200 - 2400	NaN	NaN	4
10634	Plot Area	Immediate Possession	Rajankunte	NaN	NaN	1575	NaN	NaN	3
11019	Plot Area	Immediate Possession	Yelahanka	NaN	Saaveha	1200 - 1800	NaN	NaN	1
11569	Plot Area	Immediate Possession	Hosur Road	NaN	AVeldun	1350	NaN	NaN	

13320 rows × 9 columns

In [21]:
```
df.groupby("size").mean()
```

Out[21]:

size	bath	balcony	price
1 BHK	1.054614	0.862264	39.525204
1 Bedroom	1.104762	0.361905	72.851143
1 RK	1.000000	0.461538	28.122308
10 BHK	12.000000	NaN	342.500000
10 Bedroom	9.750000	3.000000	561.250000
11 BHK	10.500000	3.000000	255.000000
11 Bedroom	8.500000	1.500000	160.000000
12 Bedroom	6.000000	2.000000	300.000000
13 BHK	13.000000	0.000000	275.000000
14 BHK	15.000000	0.000000	125.000000
16 BHK	16.000000	NaN	550.000000
18 Bedroom	18.000000	NaN	200.000000
19 BHK	16.000000	NaN	490.000000
2 BHK	1.984609	1.480396	57.428640
2 Bedroom	1.917933	0.658537	93.563465
27 BHK	27.000000	0.000000	230.000000
3 BHK	2.735246	1.895132	104.953519
3 Bedroom	2.981685	1.261860	158.218976
4 BHK	4.303293	1.910020	278.415220
4 Bedroom	3.932763	1.558077	265.067379
43 Bedroom	40.000000	0.000000	660.000000
5 BHK	5.070175	1.777778	425.618644
5 Bedroom	4.638514	1.619772	252.929293
6 BHK	5.700000	1.869565	207.833333
6 Bedroom	5.565445	1.804734	226.455497
7 BHK	6.294118	2.437500	389.294118
7 Bedroom	6.373494	1.869565	214.686747
8 BHK	7.000000	3.000000	172.000000
8 Bedroom	7.321429	1.646154	209.666667
9 BHK	8.875000	2.400000	248.250000
9 Bedroom	8.608696	1.827586	240.847826

In [22]:
```
df["size"]
```

Out[22]:
```
0            2 BHK
1        4 Bedroom
2            3 BHK
3            3 BHK
4            2 BHK
              ...
13315    5 Bedroom
13316        4 BHK
13317        2 BHK
13318        4 BHK
13319        1 BHK
Name: size, Length: 13320, dtype: object
```

In [25]:
```
df['bath'] = df['bath'].astype(str)
```

In [27]:
```
df["bath"].apply(len)
```

Out[27]:
```
0        3
1        3
2        3
3        3
4        3
        ..
13315    3
13316    3
13317    3
13318    3
13319    3
Name: bath, Length: 13320, dtype: int64
```

In [29]:
```
def reverse_name(name):
    return name[: : -1]
df["bath"].apply(reverse_name)
```

Out[29]:
```
0        0.2
1        0.5
2        0.2
3        0.3
4        0.2
        ...
13315    0.4
13316    0.5
13317    0.2
13318    0.4
13319    0.1
Name: bath, Length: 13320, dtype: object
```

In [33]:
```
df.describe()
```

Out[33]:

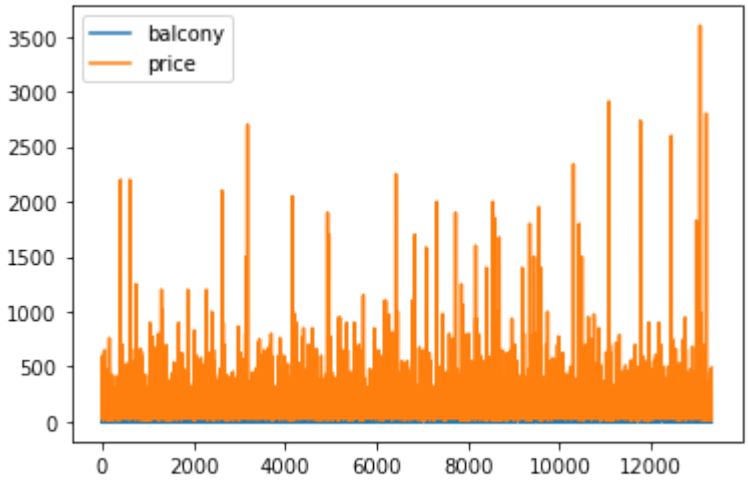
	balcony	price
count	12711.000000	13320.000000
mean	1.584376	112.565627
std	0.817263	148.971674
min	0.000000	8.000000
25%	1.000000	50.000000
50%	2.000000	72.000000
75%	2.000000	120.000000
max	3.000000	3600.000000

In [36]:
```
df['price'].mean()
```

Out[36]:
```
112.56562650150138
```

In [37]:
```
df.plot()
```

Out[37]:
```
<AxesSubplot:>
```



In [*]:

In [ ]: