

D8

```
In [1]: import pandas as pd  
import numpy as np  
import matplotlib.pyplot as plt  
import seaborn as sns
```

```
In [2]: df=pd.read_csv(r"C:\Users\user\Downloads\12_mobile_prices_2023.csv")  
df
```

Out[2]:

	Phone Name	Rating ?/5	Number of Ratings	RAM	ROM/Storage	Back/Rare Camera	Front Camera	Battery	Processor	P
0	POCO C50 (Royal Blue, 32 GB)	4.2	33,561	2 GB RAM	32 GB ROM	8MP Dual Camera	5MP Front Camera	5000 mAh	Mediatek Helio A22 Processor, Upto 2.0 GHz Pro...	:
1	POCO M4 5G (Cool Blue, 64 GB)	4.2	77,128	4 GB RAM	64 GB ROM	50MP + 2MP	8MP Front Camera	5000 mAh	Mediatek Dimensity 700 Processor	₹
2	POCO C51 (Royal Blue, 64 GB)	4.3	15,175	4 GB RAM	64 GB ROM	8MP Dual Rear Camera	5MP Front Camera	5000 mAh	Helio G36 Processor	:
3	POCO C55 (Cool Blue, 64 GB)	4.2	22,621	4 GB RAM	64 GB ROM	50MP Dual Rear Camera	5MP Front Camera	5000 mAh	Mediatek Helio G85 Processor	:
4	POCO C51 (Power Black, 64 GB)	4.3	15,175	4 GB RAM	64 GB ROM	8MP Dual Rear Camera	5MP Front Camera	5000 mAh	Helio G36 Processor	:
...
1831	Infinix Note 7 (Forest Green, 64 GB)	4.3	25,582	4 GB RAM	64 GB ROM	48MP + 2MP + AI Lens Camera	16MP Front Camera	5000 mAh	MediaTek Helio G70 Processor	₹
1832	Infinix Note 7 (Bolivia Blue, 64 GB)	4.3	25,582	4 GB RAM	64 GB ROM	48MP + 2MP + AI Lens Camera	16MP Front Camera	5000 mAh	MediaTek Helio G70 Processor	₹
1833	Infinix Note 7 (Aether Black, 64 GB)	4.3	25,582	4 GB RAM	64 GB ROM	48MP + 2MP + AI Lens Camera	16MP Front Camera	5000 mAh	MediaTek Helio G70 Processor	₹
1834	Infinix Zero 8i (Silver Diamond, 128 GB)	4.2	7,117	8 GB RAM	128 GB ROM	48MP + 8MP + 2MP + AI Lens Camera	16MP + 8MP Dual Front Camera	4500 mAh	MediaTek Helio G90T Processor	₹
1835	Infinix S5 (Quetzal Cyan, 64 GB)	4.3	15,701	4 GB RAM	64 GB ROM	16MP + 5MP + 2MP + Low Light Sensor	32MP Front Camera	4000 mAh	Helio P22 (MTK6762) Processor	₹

1836 rows × 11 columns

In [3]: df.head(10)

Out[3]:

	Phone Name	Rating ?/5	Number of Ratings	RAM	ROM/Storage	Back/Rear Camera	Front Camera	Battery	Processor	Price in INR
0	POCO C50 (Royal Blue, 32 GB)	4.2	33,561	2 GB RAM	32 GB ROM	8MP Dual Camera	5MP Front Camera	5000 mAh	Mediatek Helio A22 Processor, Upto 2.0 GHz Pro...	₹5,649
1	POCO M4 5G (Cool Blue, 64 GB)	4.2	77,128	4 GB RAM	64 GB ROM	50MP + 2MP	8MP Front Camera	5000 mAh	Mediatek Dimensity 700 Processor	₹11,999
2	POCO C51 (Royal Blue, 64 GB)	4.3	15,175	4 GB RAM	64 GB ROM	8MP Dual Rear Camera	5MP Front Camera	5000 mAh	Helio G36 Processor	₹6,999
3	POCO C55 (Cool Blue, 64 GB)	4.2	22,621	4 GB RAM	64 GB ROM	50MP Dual Rear Camera	5MP Front Camera	5000 mAh	Mediatek Helio G85 Processor	₹7,749
4	POCO C51 (Power Black, 64 GB)	4.3	15,175	4 GB RAM	64 GB ROM	8MP Dual Rear Camera	5MP Front Camera	5000 mAh	Helio G36 Processor	₹6,999
5	POCO M4 5G (Power Black, 64 GB)	4.2	77,128	4 GB RAM	64 GB ROM	50MP + 2MP	8MP Front Camera	5000 mAh	Mediatek Dimensity 700 Processor	₹11,999
6	POCO C55 (Power Black, 64 GB)	4.2	22,621	4 GB RAM	64 GB ROM	50MP Dual Rear Camera	5MP Front Camera	5000 mAh	Mediatek Helio G85 Processor	₹7,749
7	POCO C55 (Forest Green, 64 GB)	4.2	22,621	4 GB RAM	64 GB ROM	50MP Dual Rear Camera	5MP Front Camera	5000 mAh	Mediatek Helio G85 Processor	₹7,749
8	POCO C55 (Cool Blue, 128 GB)	4.1	13,647	6 GB RAM	128 GB ROM	50MP Dual Rear Camera	5MP Front Camera	5000 mAh	Mediatek Helio G85 Processor	₹9,249
9	POCO M4 5G (Yellow, 128 GB)	4.2	40,525	6 GB RAM	128 GB ROM	50MP + 2MP	8MP Front Camera	5000 mAh	Mediatek Dimensity 700 Processor	₹13,999

In [4]: `df.info()`

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 1836 entries, 0 to 1835
Data columns (total 11 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Phone Name            1836 non-null   object
1   Rating ?/5           1836 non-null   float64
2   Number of Ratings    1836 non-null   object
3   RAM                   1836 non-null   object
4   ROM/Storage           1662 non-null   object
5   Back/Rare Camera     1827 non-null   object
6   Front Camera          1435 non-null   object
7   Battery               1826 non-null   object
8   Processor             1781 non-null   object
9   Price in INR          1836 non-null   object
10  Date of Scraping      1836 non-null   object
dtypes: float64(1), object(10)
memory usage: 157.9+ KB
```

In [5]: `dff=df.dropna()`

In [6]: `dff.info()`

```
<class 'pandas.core.frame.DataFrame'>
Int64Index: 1291 entries, 0 to 1835
Data columns (total 11 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Phone Name            1291 non-null   object
1   Rating ?/5           1291 non-null   float64
2   Number of Ratings    1291 non-null   object
3   RAM                   1291 non-null   object
4   ROM/Storage           1291 non-null   object
5   Back/Rare Camera     1291 non-null   object
6   Front Camera          1291 non-null   object
7   Battery               1291 non-null   object
8   Processor             1291 non-null   object
9   Price in INR          1291 non-null   object
10  Date of Scraping      1291 non-null   object
dtypes: float64(1), object(10)
memory usage: 121.0+ KB
```

```
In [7]: dff.describe()
```

```
Out[7]:
```

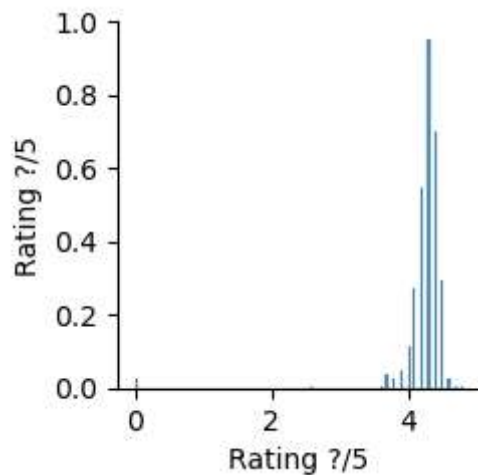
	Rating ?/5
count	1291.000000
mean	4.241208
std	0.427166
min	0.000000
25%	4.200000
50%	4.300000
75%	4.400000
max	4.800000

```
In [8]: dff.columns
```

```
Out[8]: Index(['Phone Name', 'Rating ?/5', 'Number of Ratings', 'RAM', 'ROM/Storage',  
              'Back/Rare Camera', 'Front Camera', 'Battery', 'Processor',  
              'Price in INR', 'Date of Scraping'],  
             dtype='object')
```

```
In [9]: sns.pairplot(dff)
```

```
Out[9]: <seaborn.axisgrid.PairGrid at 0x2890fc1b350>
```



```
In [10]: sns.distplot(df["Rating ?/5"])
```

C:\Users\user\AppData\Local\Temp\ipykernel_7820\1803240722.py:1: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn v0.14.0.

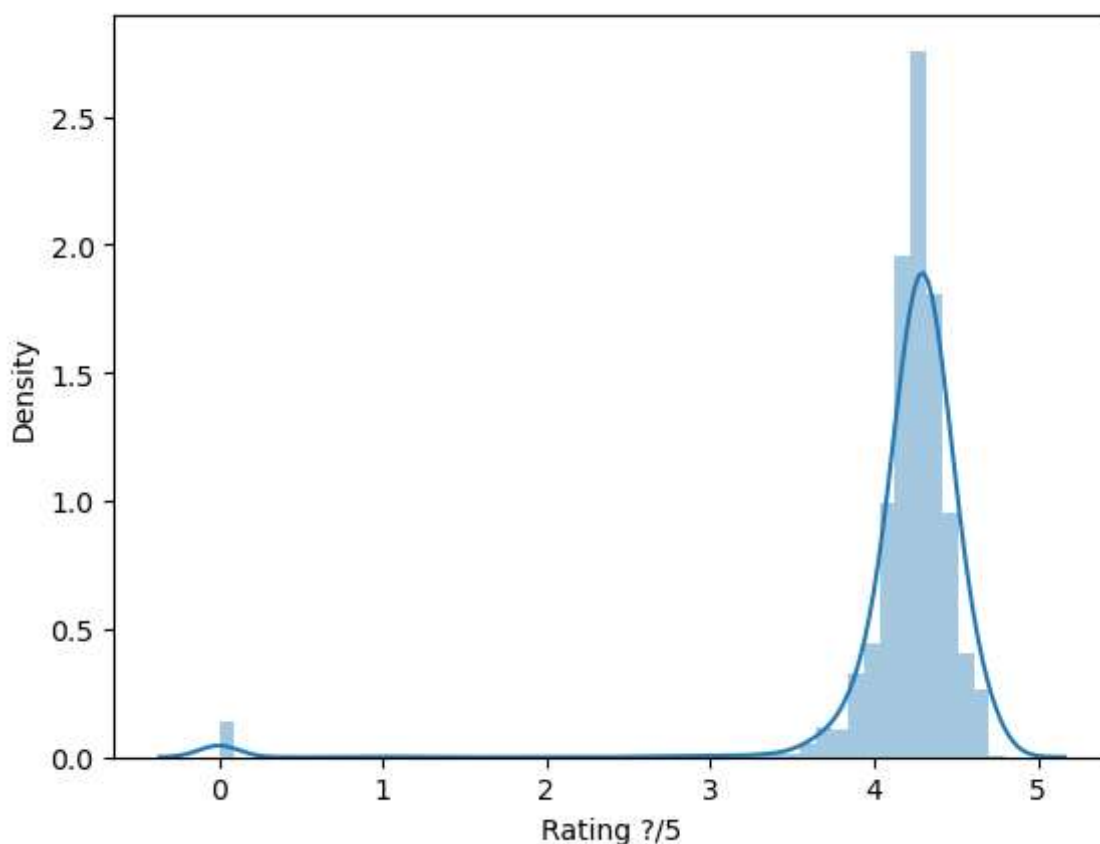
Please adapt your code to use either `displot` (a figure-level function with similar flexibility) or `histplot` (an axes-level function for histograms).

For a guide to updating your code to use the new functions, please see

<https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751> (<https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751>)

```
sns.distplot(df["Rating ?/5"])
```

Out[10]: <Axes: xlabel='Rating ?/5', ylabel='Density'>



```
In [11]: df1=df[['Phone Name', 'Rating ?/5', 'Number of Ratings', 'RAM', 'ROM/Storage',  
                'Back/Rare Camera', 'Front Camera', 'Battery', 'Processor',  
                'Price in INR', 'Date of Scraping']]
```

In [12]: `sns.heatmap(df1.corr())`

C:\Users\user\AppData\Local\Temp\ipykernel_7820\781785195.py:1: FutureWarning: The default value of numeric_only in DataFrame.corr is deprecated. In a future version, it will default to False. Select only valid columns or specify the value of numeric_only to silence this warning.
`sns.heatmap(df1.corr())`

Out[12]: <Axes: >



In [16]: `x=df1[['Rating ?/5']]`
`y=df1['Rating ?/5']`

In [17]: `from sklearn.model_selection import train_test_split`
`x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.3)`

In [18]: `from sklearn.linear_model import LinearRegression`
`lr=LinearRegression()`
`lr.fit(x_train,y_train)`

Out[18]: `LinearRegression`
`LinearRegression()`

In [19]: `print(lr.intercept_)`

1.7763568394002505e-15

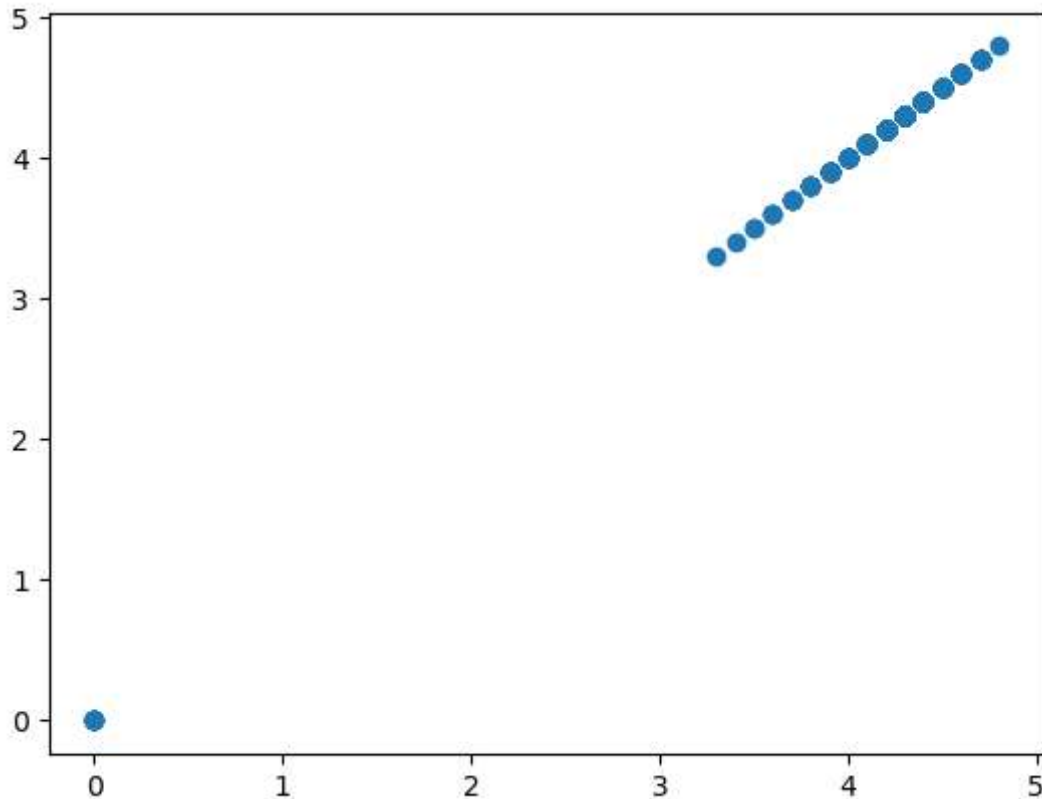

```
In [20]: coeff = pd.DataFrame(lr.coef_,x.columns,columns=['Co-efficient'])  
coeff
```

Out[20]:

	Co-efficient
Rating ?/5	1.0

```
In [21]: prediction=lr.predict(x_test)  
plt.scatter(y_test,prediction)
```

Out[21]: <matplotlib.collections.PathCollection at 0x289154cfc90>



```
In [22]: print(lr.score(x_test,y_test))
```

1.0

In []: