# D8

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

In [23]:
```
df=pd.read_csv(r"C:\Users\user\Downloads\12_mobile_prices_2023.csv")
df
```

Out[23]:

| | 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 + 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 + 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 + 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 [24]:
```
df.head(10)
```

Out[24]:

| | Phone Name | Rating ?/5 | Number of Ratings | RAM | ROM/Storage | Back/Rare Camera | Front Camera | Battery | Processor | Price ir INF |
|---|---|---|---|---|---|---|---|---|---|---|
| 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 [25]:
```
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 [26]:
```
dff=df.dropna()
```

In [27]:
```
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 [28]:
```
dff.describe()
```

Out[28]:

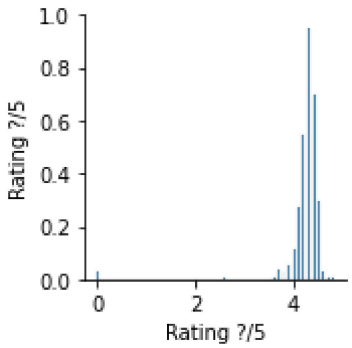
| | Rating ?/5 |
|---|---|
| **count** | 1291.000000 |
| **mean** | 4.241208 |
| **std** | 0.427166 |
| **min** | 0.000000 |
| **25%** | 4.200000 |
| **50%** | 4.300000 |
| **75%** | 4.400000 |
| **max** | 4.800000 |

In [29]:
```
dff.columns
```

Out[29]:
```
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 [30]:
```
sns.pairplot(dff)
```

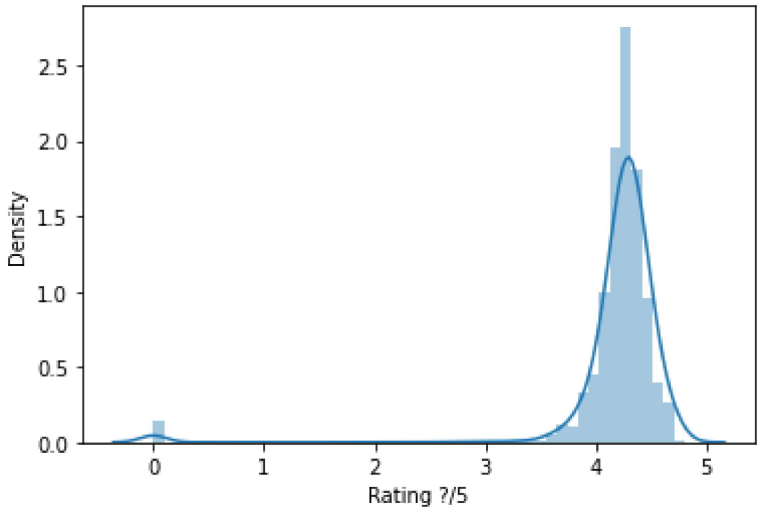
Out[30]:
```
<seaborn.axisgrid.PairGrid at 0x1dda121afd0>
```



In [31]:
```
sns.distplot(df["Rating ?/5"])
```

```
C:\ProgramData\Anaconda3\lib\site-packages\seaborn\distributions.py:2557: Fut
ureWarning: `distplot` is a deprecated function and will be removed in a futu
re version. Please adapt your code to use either `displot` (a figure-level fu
nction with similar flexibility) or `histplot` (an axes-level function for hi
stograms).
  warnings.warn(msg, FutureWarning)
```

Out[31]: <AxesSubplot:xlabel='Rating ?/5', ylabel='Density'>




In [32]:
```
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 [33]:
```
sns.heatmap(df1.corr())
```

Out[33]: <AxesSubplot:>


In [34]:
```
x=df1['Rating ?/5']
y=df1['Rating ?/5']
```

In [35]:
```
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 [36]:
```
from sklearn.linear_model import LinearRegression
lr=LinearRegression()
lr.fit(x_train,y_train)
```

```
---------------------------------------------------------------------------
ValueError                                Traceback (most recent call last)
<ipython-input-36-c1d813126ab8> in <module>
      1 from sklearn.linear_model import LinearRegression
      2 lr=LinearRegression()
----> 3 lr.fit(x_train,y_train)

C:\ProgramData\Anaconda3\lib\site-packages\sklearn\linear_model\_base.py in f
it(self, X, y, sample_weight)
    516         accept_sparse = False if self.positive else ['csr', 'csc', 'c
oo']
    517
--> 518         X, y = self._validate_data(X, y, accept_sparse=accept_sparse,
    519                                    y_numeric=True, multi_output=True)
    520

C:\ProgramData\Anaconda3\lib\site-packages\sklearn\base.py in _validate_data
(self, X, y, reset, validate_separately, **check_params)
    431                 y = check_array(y, **check_y_params)
    432             else:
--> 433                 X, y = check_X_y(X, y, **check_params)
    434             out = X, y
    435

C:\ProgramData\Anaconda3\lib\site-packages\sklearn\utils\validation.py in inn
er_f(*args, **kwargs)
     61             extra_args = len(args) - len(all_args)
     62             if extra_args <= 0:
---> 63                 return f(*args, **kwargs)
     64
     65             # extra_args > 0

C:\ProgramData\Anaconda3\lib\site-packages\sklearn\utils\validation.py in che
ck_X_y(X, y, accept_sparse, accept_large_sparse, dtype, order, copy, force_al
l_finite, ensure_2d, allow_nd, multi_output, ensure_min_samples, ensure_min_f
eatures, y_numeric, estimator)
    812         raise ValueError("y cannot be None")
    813
--> 814     X = check_array(X, accept_sparse=accept_sparse,
    815                     accept_large_sparse=accept_large_sparse,
    816                     dtype=dtype, order=order, copy=copy,

C:\ProgramData\Anaconda3\lib\site-packages\sklearn\utils\validation.py in inn
er_f(*args, **kwargs)
     61             extra_args = len(args) - len(all_args)
     62             if extra_args <= 0:
---> 63                 return f(*args, **kwargs)
     64
     65             # extra_args > 0

C:\ProgramData\Anaconda3\lib\site-packages\sklearn\utils\validation.py in che
ck_array(array, accept_sparse, accept_large_sparse, dtype, order, copy, force
_all_finite, ensure_2d, allow_nd, ensure_min_samples, ensure_min_features, es
timator)
    635             # If input is 1D raise error
    636             if array.ndim == 1:
--> 637                 raise ValueError(
```

```
    638                     "Expected 2D array, got 1D array instead:\narray=
{}.\n"
    639                     "Reshape your data either using array.reshape(-1,
1) if "

ValueError: Expected 2D array, got 1D array instead:
array=[4.3 4.2 4.2 ... 3.7 4.2 4.5].
Reshape your data either using array.reshape(-1, 1) if your data has a single
feature or array.reshape(1, -1) if it contains a single sample.
```

In [22]:
```
print(lr.intercept_)
```

```
---------------------------------------------------------------------------
AttributeError                            Traceback (most recent call last)
<ipython-input-22-182bb45ab960> in <module>
----> 1 print(lr.intercept_)

AttributeError: 'LinearRegression' object has no attribute 'intercept_'
```

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

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

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

In [ ]:
```
from sklearn.linear_model import Ridge,Lasso
```

In [ ]:
```
rr=Ridge(alpha=10)
rr.fit(x_train,y_train)
```

In [ ]:
```
rr.score(x_test,y_test)
```

In [ ]:
```
la=Lasso(alpha=10)
la.fit(x_train,y_train)
```

In [ ]:
```
la.score(x_test,y_test)
```