3-praveenk-chintatejdeepreddy-code

May 21, 2023

Name : Praveen K Roll No: 211AI028 Name : Chinta Tejdeep Reddy Roll No: 211AI013 ASSIGN-MENT - 3

```
[93]: import pandas as pd
      import numpy as np;
      import matplotlib.pyplot as plt
      import seaborn as sns
      from sklearn.decomposition import PCA
      from sklearn.preprocessing import StandardScaler
      from sklearn.preprocessing import LabelEncoder
      from sklearn.preprocessing import MinMaxScaler
[61]: #Loading the Dataframe
      df = pd.read_csv(r"C:
       →\Users\Praveen\Desktop\ASSIGNMENTSEM-III\DS\3\Mobile_Recommendation_System-main\mainDataset

csv",encoding='ISO-8859-1')
[62]: #Printing the dataframe
      print(df)
                                                             Rating RAM Gb
                                                                              ROM Gb
                                                       Name
     0
          Samsung Galaxy Note 10 Plus (Aura Glow, 256 GB)
                                                                 4.6
                                                                          12
                                                                                 256
                                   Lenovo A7 (Blue, 32 GB)
     1
                                                                 4.0
                                                                           2
                                                                                  32
     2
             Redmi Note 9 Pro Max (Champagne Gold, 64 GB)
                                                                 4.3
                                                                           6
                                                                                  64
     3
                       Realme X50 Pro (Moss Green, 256 GB)
                                                                 4.1
                                                                          12
                                                                                 256
     4
                                    Lava Z61 (Gold, 16 GB)
                                                                 4.2
                                                                           2
                                                                                  16
     . .
                                                                 4.0
                                                                           2
     956
                       Lava Z61 Pro (Lavender Blue, 16 GB)
                                                                                  16
     957
                                  Lenovo A7 (Black, 32 GB)
                                                                 4.0
                                                                           2
                                                                                  32
                     Itel VISION 1 (GRADATION BLUE, 32 GB)
                                                                           2
                                                                                  32
     958
                                                                 2.5
     959
                     Itel Vision1 (Gradation Green, 32 GB)
                                                                 4.1
                                                                           3
                                                                                  32
     960
                      Itel Vision1 (Gradation Blue, 32 GB)
                                                                 4.1
                                                                           3
                                                                                  32
          Expandable GB
                          Size Cm Size Inch R1 Cam MP
                                                          R2 Cam MP
                                                                    R3 Cam MP
     0
                     1.0
                            17.27
                                        6.800
                                                      12
                                                                 0.0
     1
                   128.0
                            15.47
                                        6.090
                                                      13
                                                                 0.0
                                                                              0
     2
                            16.94
                                                      64
                                                                 0.0
                     NaN
                                        6.670
                                                                              0
     3
                     NaN
                            16.36
                                        6.440
                                                      64
                                                                12.0
                                                                              8
```

```
4
                      64.0
                                13.84
                                             5.450
                                                              8
                                                                        0.0
                       •••
                     128.0
                                13.84
                                            5.450
      956
                                                              8
                                                                        0.0
                                                                                        0
      957
                     128.0
                                15.47
                                            6.090
                                                             13
                                                                        0.0
                                                                                        0
                     128.0
                                15.46
                                            6.088
                                                              8
                                                                                        0
      958
                                                                        0.0
      959
                     128.0
                                15.46
                                             6.088
                                                              8
                                                                        0.3
                                                                                        0
      960
                     128.0
                                15.46
                                             6.088
                                                              8
                                                                        0.3
            Brand name
                         Unnamed: 17
                                         Unnamed: 18 Unnamed: 19 Unnamed: 20 Unnamed: 21
                                   NaN
                                                  NaN
                                                                NaN
                                                                              NaN
                                                                                            NaN
      0
               Samsung
                                   NaN
                                                  NaN
                                                                NaN
                                                                              {\tt NaN}
      1
                Lenovo
                                                                                            NaN
      2
                 Redmi
                                   NaN
                                                  NaN
                                                                NaN
                                                                              NaN
                                                                                            NaN
      3
                                   NaN
                                                                              {\tt NaN}
                Realme
                                                  NaN
                                                                NaN
                                                                                             NaN
      4
                                   NaN
                                                  NaN
                                                                NaN
                                                                              NaN
                   Lava
                                                                                             NaN
      . .
                    •••
      956
                                   {\tt NaN}
                                                  NaN
                                                                {\tt NaN}
                                                                              NaN
                                                                                            NaN
                   Lava
      957
                Lenovo
                                   {\tt NaN}
                                                  NaN
                                                                {\tt NaN}
                                                                              NaN
                                                                                            NaN
      958
                   Itel
                                   {\tt NaN}
                                                  NaN
                                                                NaN
                                                                              NaN
                                                                                            NaN
      959
                   Itel
                                   NaN
                                                  NaN
                                                                {\tt NaN}
                                                                              NaN
                                                                                            NaN
      960
                   Itel
                                   NaN
                                                  NaN
                                                                NaN
                                                                              NaN
                                                                                            NaN
          Unnamed: 22
                         Unnamed: 23
                                         Unnamed: 24
                                                        NaN
      0
                    NaN
                                   NaN
                                                  NaN
                    NaN
                                   NaN
                                                  NaN
      1
                                                        NaN
      2
                    NaN
                                   NaN
                                                  NaN
                                                        NaN
      3
                    NaN
                                   {\tt NaN}
                                                  NaN
                                                        NaN
      4
                    NaN
                                   {\tt NaN}
                                                  NaN
                                                        NaN
      . .
      956
                    NaN
                                                  NaN
                                                        NaN
                                   NaN
      957
                    NaN
                                   NaN
                                                  NaN
                                                        NaN
      958
                    NaN
                                   NaN
                                                  NaN
                                                        NaN
      959
                    NaN
                                   {\tt NaN}
                                                  NaN
                                                        NaN
      960
                    NaN
                                   {\tt NaN}
                                                  NaN
                                                        NaN
      [961 rows x 26 columns]
[63]: #Printing the first 5 elements of the dataframe
       df.head()
[63]:
                                                                   Rating
                                                                            RAM Gb
                                                                                      ROM Gb \
                                                            Name
          Samsung Galaxy Note 10 Plus (Aura Glow, 256 GB)
                                                                       4.6
                                                                                 12
                                                                                          256
       1
                                      Lenovo A7 (Blue, 32 GB)
                                                                       4.0
                                                                                   2
                                                                                           32
       2
             Redmi Note 9 Pro Max (Champagne Gold, 64 GB)
                                                                       4.3
                                                                                   6
                                                                                           64
       3
                        Realme X50 Pro (Moss Green, 256 GB)
                                                                                 12
                                                                                          256
                                                                       4.1
       4
                                       Lava Z61 (Gold, 16 GB)
                                                                       4.2
                                                                                   2
                                                                                           16
```

Expandable GB Size Cm Size Inch R1 Cam MP R2 Cam MP R3 Cam MP ... \

```
0
                     1.0
                            17.27
                                          6.80
                                                        12
                                                                   0.0
                                                                                 0
      1
                  128.0
                            15.47
                                          6.09
                                                        13
                                                                   0.0
                                                                                 0
      2
                    NaN
                            16.94
                                          6.67
                                                        64
                                                                   0.0
      3
                    NaN
                            16.36
                                          6.44
                                                        64
                                                                  12.0
                                                                                 8
      4
                   64.0
                                          5.45
                                                         8
                                                                   0.0
                                                                                 0
                            13.84
                                     Unnamed: 18 Unnamed: 19 Unnamed: 20 Unnamed: 21
         Brand name
                      Unnamed: 17
      0
             Samsung
                               NaN
                                              NaN
                                                           NaN
                                                                        NaN
                                                                                      NaN
      1
              Lenovo
                               NaN
                                              NaN
                                                           NaN
                                                                        NaN
                                                                                      NaN
      2
               Redmi
                               NaN
                                                           NaN
                                                                        NaN
                                              NaN
                                                                                      NaN
      3
              Realme
                               NaN
                                              NaN
                                                           NaN
                                                                        NaN
                                                                                      NaN
      4
                Lava
                               NaN
                                              NaN
                                                                        NaN
                                                                                      NaN
                                                           NaN
        Unnamed: 22
                       Unnamed: 23
                                     Unnamed: 24
                                                   Pro
      0
                 NaN
                               NaN
                                              NaN
                                                   NaN
      1
                 NaN
                               NaN
                                              NaN
                                                   NaN
      2
                 NaN
                               NaN
                                              NaN
                                                   NaN
      3
                 NaN
                               NaN
                                              NaN
                                                   NaN
      4
                 NaN
                               NaN
                                                   NaN
                                              NaN
      [5 rows x 26 columns]
[64]: #Describe the dataset and all the measures
      df.describe()
[64]:
                  Rating
                               RAM Gb
                                             ROM Gb
                                                      Expandable GB
                                                                          Size Cm
      count
              936.000000
                                                         715.000000
                           961.000000
                                        961.000000
                                                                      961.000000
                4.227137
                             4.500520
                                          78.817898
                                                         298.641958
                                                                       15.682352
      mean
      std
                0.368420
                             2.094201
                                          65.408979
                                                         151.091947
                                                                        1.162916
                2.300000
                             2.000000
                                          16.000000
                                                           1.000000
                                                                       12.700000
      min
      25%
                4.200000
                             3.000000
                                          32.000000
                                                         256.000000
                                                                       15.210000
      50%
                4.300000
                             4.000000
                                          64.000000
                                                         256.000000
                                                                       16.000000
      75%
                             6.000000
                                        128.000000
                                                         512.000000
                4.400000
                                                                        16.510000
                                        512.000000
      max
                5.000000
                            12.000000
                                                         512.000000
                                                                        17.780000
               Size Inch
                            R1 Cam MP
                                         R2 Cam MP
                                                       R3 Cam MP
                                                                    R4 Cam MP
              961.000000
                           961.000000
                                        961.000000
                                                      961.000000
                                                                   961.000000
      count
      mean
                6.173996
                            26.070760
                                           4.323205
                                                        1.637877
                                                                     0.514048
      std
                0.457630
                            20.893111
                                           5.162666
                                                        2.862469
                                                                     1.113394
                5.000000
                             5.000000
                                          0.00000
                                                        0.000000
                                                                     0.000000
      min
      25%
                            13.000000
                                          0.00000
                                                                     0.000000
                5.990000
                                                        0.00000
      50%
                6.300000
                            13.000000
                                           2.000000
                                                        0.000000
                                                                     0.000000
      75%
                6.500000
                            48.000000
                                           8.000000
                                                        2.000000
                                                                     0.000000
      max
                7.000000
                           108.000000
                                          48.000000
                                                       16.000000
                                                                     8.000000
                                             Unnamed: 18
                                                           Unnamed: 19
                                                                                        \
                   Price Rs
                              Unnamed: 17
                                                                         Unnamed: 20
                 959.000000
                                       0.0
                                                      0.0
                                                                    0.0
                                                                                  0.0
      count
```

	mean	17537.970803	Nal	N Na	N Na	aN	NaN	
	std	14675.059110	Nal	N Na	N Na	aN	NaN	
	min	3899.000000	Nal	N Na	N Na	aN	NaN	
	25%	9560.000000) Nal	N Na	N Na	aN	NaN	
	50%	13999.000000	Nal	N Na	N Na	aN	NaN	
	75%	18734.000000) Nal	N Na	N Na	aN	NaN	
	max	124999.000000) Nal	Na Na	N Na	aN	NaN	
		Unnamed: 21	Unnamed: 22	Unnamed: 23	Unnamed: 24	Pro		
	count	0.0	0.0	0.0	0.0	0.0		
	mean	NaN	NaN	NaN	NaN	NaN		
	std	NaN	NaN	NaN	NaN	NaN		
	min	NaN	NaN	NaN	NaN	NaN		
	25%	NaN	NaN	NaN	NaN	NaN		
	50%	NaN	NaN	NaN	NaN	NaN		
	75%	NaN	NaN	NaN	NaN	NaN		
	max	NaN	NaN	NaN	NaN	NaN		
	[8 rows	s x 21 columns	.1					
	[O TOWE	o x 21 columne						
[117]:	#Shape	of the datase	et					
	df.shap	oe .						
[117]:	(961, 2	29)						
[66]:		ing the number	of null val	ues in each f	eature			
	df.isnu	ill().sum()						
[66]:	Name		0					
	Rating	2	25					
	RAM Gb		0					

[66]: RAM Gb ROM Gb 0 Expandable GB 246 ${\tt Size}\ {\tt Cm}$ 0 Size Inch 0 R1 Cam MP 0 R2 Cam MP 0 R3 Cam MP 0 R4 Cam MP 0 Battery Mah 0 Price Rs 2 Processor 0 Image 0 0 Processor name Brand name 0 Unnamed: 17 961 Unnamed: 18 961

```
Unnamed: 19
                        961
      Unnamed: 20
                        961
      Unnamed: 21
                        961
      Unnamed: 22
                        961
      Unnamed: 23
                        961
      Unnamed: 24
                        961
     Pro
                        961
      dtype: int64
     Finding the Correlation between the various features to see how they vary with each other.
[67]: correlation = df['Price Rs'].corr(df['Rating'])
      print("Correlation:")
      print(correlation)
     Correlation:
     0.30696556781567247
[68]: correlation = df['Battery Mah'].corr(df['Rating'])
      print("Correlation:")
      print(correlation)
     Correlation:
     0.451432302331625
[69]: correlation = df['RAM Gb'].corr(df['Rating'])
      print("Correlation:")
      print(correlation)
     Correlation:
     0.38343858940106756
[70]: correlation = df['RAM Gb'].corr(df['Price Rs'])
      print("Correlation:")
      print(correlation)
     Correlation:
     0.6668573282484104
[71]: correlation = df['ROM Gb'].corr(df['Rating'])
      print("Correlation:")
      print(correlation)
     Correlation:
     0.33083043521783173
[72]: correlation = df['ROM Gb'].corr(df['Price Rs'])
```

print("Correlation:")

```
print(correlation)
```

Correlation:

0.6427890816574197

```
[73]: correlation = df['Battery Mah'].corr(df['Price Rs'])
print("Correlation:")
print(correlation)
```

Correlation:

0.07362529314706924

```
[74]: pd = df
    label_encoder = LabelEncoder()
    pd['Processor_enc'] = label_encoder.fit_transform(df['Processor'])
    correlation = pd['Processor_enc'].corr(df['Price Rs'])
    print("Correlation:")
    print(correlation)
```

Correlation:

-0.016568678246270813

Plotting the correlation heat map between the various features.

```
[75]: plt.figure(figsize=(14,14))
cols = ['Rating','RAM Gb','ROM Gb','Expandable GB','Size Cm','Size Inch','R1

→Cam MP','R2 Cam MP','R3 Cam MP','R4 Cam MP','Battery Mah','Price Rs']
pd = df[cols]
corr = pd.corr()
sns.heatmap(corr, cbar=False, square= True, fmt='.2%', annot=True,

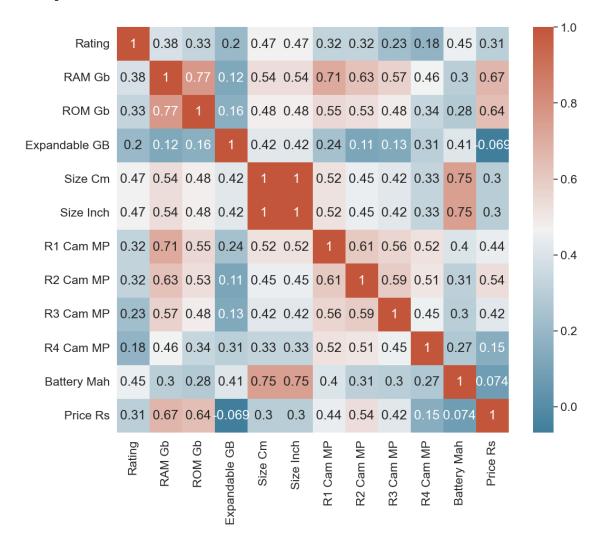
→cmap='Greens')
```

[75]: <AxesSubplot: >

Rating	100.00%	38.34%	33.08%	19.93%	46.67%	46.68%	32.29%	32.42%	22.71%	18.00%	45.14%	30.70%
RAM Gb	38.34%	100.00%	76.93%	12.39%	53.64%	53.62%	70.98%	63.07%	57.22%	46.49%	30.08%	66.69%
ROM Gb	33.08%	76.93%	100.00%	15.96%	47.71%	47.69%	55.17%	53.01%	48.50%	34.02%	28.19%	64.28%
Expandable GB	19.93%	12.39%	15.96%	100.00%	41.60%	41.57%	23.69%	11.30%	12.79%	30.58%	40.97%	-6.87%
Size Cm	46.67%	53.64%	47.71%	41.60%	100.00%	100.00%	51.98%	44.85%	41.92%	32.88%	74.75%	29.71%
Size Inch	46.68%	53.62%	47.69%	41.57%	100.00%	100.00%	51.97%	44.84%	41.90%	32.87%	74.73%	29.69%
R1 Cam MP	32.29%	70.98%	55.17%	23.69%	51.98%	51.97%	100.00%	60.65%	56.43%	52.07%	40.09%	44.31%
R2 Cam MP	32.42%	63.07%	53.01%	11.30%	44.85%	44.84%	60.65%	100.00%	59.20%	51.00%	31.04%	53.66%
R3 Cam MP	22.71%	57.22%	48.50%	12.79%	41.92%	41.90%	56.43%	59.20%	100.00%	45.33%	30.01%	41.65%
R4 Cam MP	18.00%	46.49%	34.02%	30.58%	32.88%	32.87%	52.07%	51.00%	45.33%	100.00%	27.20%	15.24%
Battery Mah	45.14%	30.08%	28.19%	40.97%	74.75%	74.73%	40.09%	31.04%	30.01%	27.20%	100.00%	7.36%
Price Rs	30.70%	66.69%	64.28%	-6.87%	29.71%	29.69%	44.31%	53.66%	41.65%	15.24%	7.36%	100.00%
	Rating	RAM Gb	ROM Gb	Expandable GB	Size Cm	Size Inch	R1 Cam MP	R2 Cam MP	R3 Cam MP	R4 Cam MP	Battery Mah	Price Rs

```
# Configure a custom diverging colormap
#
cmap = sns.diverging_palette(230, 20, as_cmap=True)
#
# Draw the heatmap
#
sns.heatmap(corr, annot=True, cmap=cmap)
```

[76]: <AxesSubplot: >



Finding the Covariance between the different features.

```
[77]: covariance = df['Price Rs'].cov(df['Rating'])

# Print the covariance
print("Covariance:")
```

```
Covariance:
     1673.9839854857914
[78]: covariance = df['Size Cm'].cov(df['Battery Mah'])
      # Print the covariance
      print("Covariance:")
      print(covariance)
     Covariance:
     751.345585598769
[79]: covariance = df['RAM Gb'].cov(df['Battery Mah'])
      # Print the covariance
      print("Covariance:")
      print(covariance)
     Covariance:
     544.4242976066597
[80]: covariance = df['R1 Cam MP'].cov(df['Rating'])
      # Print the covariance
      print("Covariance:")
      print(covariance)
     Covariance:
     2.492478632478632
[81]: covariance = df['Expandable GB'].cov(df['Battery Mah'])
      # Print the covariance
      print("Covariance:")
      print(covariance)
     Covariance:
     51086.74128812366
     Covariance Heat Map
[82]: cols = ['Rating', 'RAM Gb', 'ROM Gb', 'Expandable GB', 'Size Cm', 'Size Inch', 'R1
       ⇔Cam MP','R2 Cam MP','R3 Cam MP','R4 Cam MP','Battery Mah','Price Rs']
      pd = df[cols]
      plt.figure(figsize=(24,24))
```

print(covariance)

```
corr = pd.cov()
sns.heatmap(corr, cbar=False, square= True, fmt='.2%', annot=True,

cmap='Greens')
```

[82]: <AxesSubplot: >

Rating	13.57%	29.54%	802.12%	756.77%	20.03%	7.88%	249.25%	61.93%	24.20%	7.43%	14307.82%	167398.40%
RAM Gb	29.54%	438.57%	10537.56%	3608.19%	130.62%	51.39%	3105.83%	681.87%	343.04%	108.41%	54442.43%	2051436.43%
ROM Gb	802.12%	10537.56%	427833.45%	153922.10%	3628.71%	1427.51%	75400.04%	17901.91%	9080.48%	2477.50%	1593832.92%	61744599.36%
Expandable GB	756.77%	3608.19%	153922.10%	2282877.64%	7032.09%	2765.30%	71560.12%	8321.06%	5011.21%	5158.92%	5108674.13%	13879643.04%
Size Cm	20.03%	130.62%	3628.71%	7032.09%	135.24%	53.22%	1263.03%	269.24%	139.53%	42.57%	75134.56%	507077.04%
Size Inch	7.88%	51.39%	1427.51%	2765.30%	53.22%	20.94%	496.94%	105.94%	54.89%	16.75%	29558.74%	199428.03%
R1 Cam MP	249.25%	3105.83%	75400.04%	71560.12%	1263.03%	496.94%	43652.21%	6542.21%	3374.86%	1211.15%	723872.53%	13585449.86%
R2 Cam MP	61.93%	681.87%	17901.91%	8321.06%	269.24%	105.94%	6542.21%	2665.31%	874.82%	293.16%	138503.62%	3913075.00%
R3 Cam MP	24.20%	343.04%	9080.48%	5011.21%	139.53%	54.89%	3374.86%	874.82%	819.37%	144.47%	74253.26%	1746546.74%
R4 Cam MP	7.43%	108.41%	2477.50%	5158.92%	42.57%	16.75%	1211.15%	293.16%	144.47%	123.96%	26174.77%	247061.62%
Battery Mah	14307.82%	54442.43%	1593832.92%	5108674.13%	75134.56%	29558.74%	723872.53%	138503.62%	74253.26%	26174.77%	74698230.02%	93355011.78%
Price Rs	167398.40%	2051436.43%	61744599.369	#13879643.04%	507077.04%	199428.03%	13585449.86%	ଥ913075.00%	1746546.74%	247061.62%	93355011.78%	535735986.97
	Rating	RAM Gb	ROM Gb	Expandable GB	Size Cm	Size Inch	R1 Cam MP	R2 Cam MP	R3 Cam MP	R4 Cam MP	Battery Mah	Price Rs

Plotting the covariance matrix showing correlations coefficients

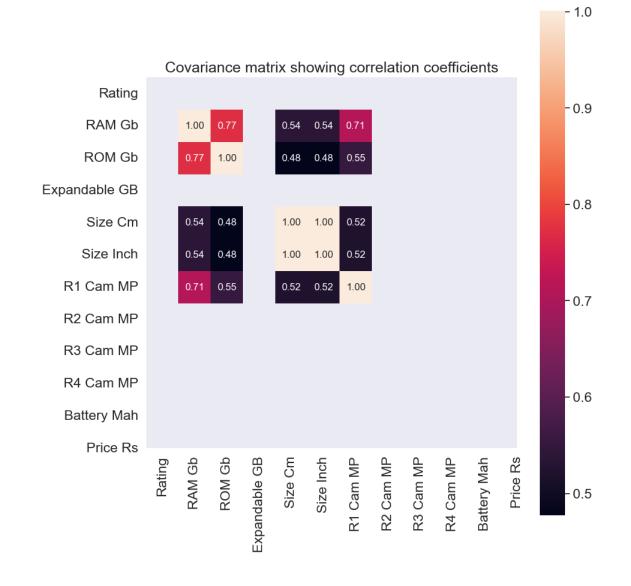
```
[83]: cols = ['Rating','RAM Gb','ROM Gb','Expandable GB','Size Cm','Size Inch','R1_

Cam MP','R2 Cam MP','R3 Cam MP','R4 Cam MP','Battery Mah','Price Rs']

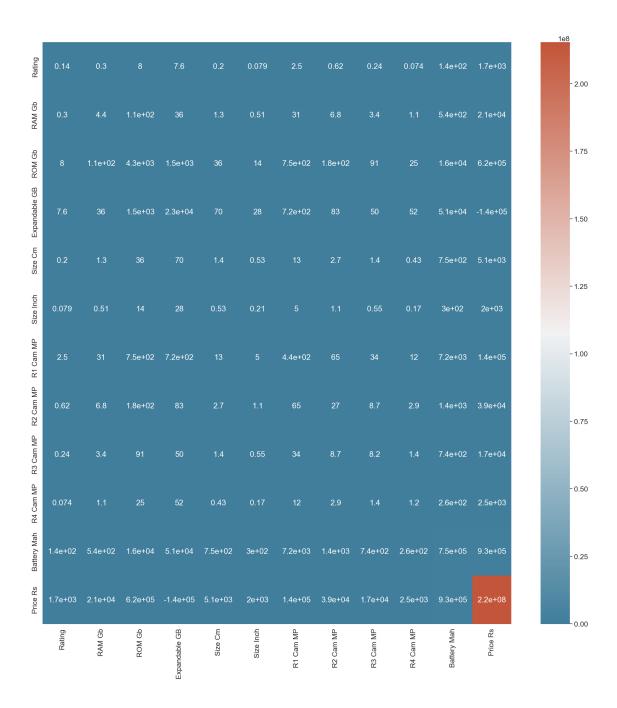
from sklearn.preprocessing import StandardScaler

stdsc = StandardScaler()

X_std = stdsc.fit_transform(df[cols].iloc[:,range(0,7)].values)
```



[84]: <AxesSubplot: >



Applying PCA on the camera features and reducing it to a single feature.

```
[86]: cam_features = df[['R1 Cam MP', 'R2 Cam MP', 'R3 Cam MP', 'R4 Cam MP']]
    scaler = StandardScaler()
    cam_features_std = scaler.fit_transform(cam_features)
    pca = PCA(n_components=1)
    cam_features_pca = pca.fit_transform(cam_features_std)
    df['Cam Reduced'] = cam_features_pca
```

```
print("Explained Variance Ratio:", pca.explained_variance_ratio_)
      Explained Variance Ratio: [0.65677317]
      Applying PCA on the Size features and reducing it to a single feature.
[87]: siz_features = df[['Size Cm', 'Size Inch']]
       scaler = StandardScaler()
       siz_features_std = scaler.fit_transform(siz_features)
       pca = PCA(n_components=1)
       siz_features_pca = pca.fit_transform(siz_features_std)
       df['Size Reduced'] = siz_features_pca
       print("Explained Variance Ratio:", pca.explained_variance_ratio_)
      Explained Variance Ratio: [0.99999894]
[88]: print(df.columns)
      Index(['Name', 'Rating', 'RAM Gb', 'ROM Gb', 'Expandable GB', 'Size Cm',
             'Size Inch', 'R1 Cam MP', 'R2 Cam MP', 'R3 Cam MP', 'R4 Cam MP',
              'Battery Mah', 'Price Rs', 'Processor', 'Image', 'Processor name',
             'Brand name', 'Unnamed: 17', 'Unnamed: 18', 'Unnamed: 19',
             'Unnamed: 20', 'Unnamed: 21', 'Unnamed: 22', 'Unnamed: 23',
              'Unnamed: 24', 'Pro', 'Processor_enc', 'Cam Reduced', 'Size Reduced'],
            dtype='object')
      Creating the new dataframe with reduced features.
[109]: new_pdf = df[['Name', 'Rating', 'RAM Gb', 'ROM Gb', 'Expandable GB', 'Size_
        →Reduced', 'Cam Reduced',
              'Battery Mah', 'Price Rs', 'Processor', 'Processor name',
              'Brand name']]
[110]: new_pdf.head()
[110]:
                                                      Name
                                                            Rating RAM Gb
                                                                             ROM Gb \
       O Samsung Galaxy Note 10 Plus (Aura Glow, 256 GB)
                                                               4.6
                                                                         12
                                                                                256
       1
                                  Lenovo A7 (Blue, 32 GB)
                                                               4.0
                                                                          2
                                                                                 32
       2
             Redmi Note 9 Pro Max (Champagne Gold, 64 GB)
                                                               4.3
                                                                          6
                                                                                 64
       3
                      Realme X50 Pro (Moss Green, 256 GB)
                                                               4.1
                                                                         12
                                                                                256
       4
                                   Lava Z61 (Gold, 16 GB)
                                                               4.2
                                                                                 16
```

```
0
                    1.0
                            -1.933641
                                           0.292194
                                                             4300
                                                                    85000.0
                  128.0
       1
                             0.259041
                                           0.333936
                                                             4000
                                                                     8490.0
       2
                    0.0
                            -1.531907
                                           2.462778
                                                             5020
                                                                    17767.0
       3
                    0.0
                            -0.823487
                                           8.276223
                                                             4200
                                                                    47999.0
                                                                     5999.0
       4
                   64.0
                             2.240084
                                           0.125226
                                                             3000
                                               Processor Processor name Brand name
       0
                                   Exynos 9820 Processor
                                                                  Exynos
                                                                            Samsung
                      Octa-core Unisoc SC9863 Processor
       1
                                                               Octa-core
                                                                             Lenovo
       2
           Qualcomm Snapdragon 720G octa-core processor
                                                                Qualcomm
                                                                              Redmi
       3
                      Qualcomm Snapdragon 865 Processor
                                                                Qualcomm
                                                                             Realme
       4
                  1.5 GHz Quad Core Processor Processor
                                                                     1.5
                                                                               Lava
      Standardizing the Cam Reduced features to a range of values between 1 to 16
[111]: scaler = MinMaxScaler(feature_range=(0, 16))
       # Select the column you want to scale
       column_name = 'Cam Reduced'
       # Extract the values of the selected column
       column_values = new_pdf[column_name].values
       # Reshape the column values to a 2D array
       reshaped_values = column_values.reshape(-1, 1)
       # Apply Min-Max scaling to the column values
       scaled_values = scaler.fit_transform(reshaped_values)
       # Assign the scaled values back to the dataframe
       new_pdf[column_name] = scaled_values.flatten()
      C:\Users\Praveen\AppData\Local\Temp\ipykernel_21376\1600270228.py:15:
      SettingWithCopyWarning:
      A value is trying to be set on a copy of a slice from a DataFrame.
      Try using .loc[row_indexer,col_indexer] = value instead
      See the caveats in the documentation: https://pandas.pydata.org/pandas-
      docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
        new_pdf[column_name] = scaled_values.flatten()
[112]: new_pdf.head()
[112]:
                                                             Rating
                                                                     RAM Gb
                                                                             ROM Gb \
                                                                                256
          Samsung Galaxy Note 10 Plus (Aura Glow, 256 GB)
                                                                4.6
                                                                         12
       1
                                   Lenovo A7 (Blue, 32 GB)
                                                                4.0
                                                                                 32
```

Size Reduced Cam Reduced

Battery Mah

Price Rs \

Expandable GB

```
2
      Redmi Note 9 Pro Max (Champagne Gold, 64 GB)
                                                          4.3
                                                                            64
                                                                     6
3
                Realme X50 Pro (Moss Green, 256 GB)
                                                          4.1
                                                                    12
                                                                           256
4
                              Lava Z61 (Gold, 16 GB)
                                                          4.2
                                                                     2
                                                                             16
                   Size Reduced Cam Reduced Battery Mah Price Rs
   Expandable GB
0
              1.0
                      -1.933641
                                     0.292194
                                                       4300
                                                               85000.0
            128.0
                       0.259041
                                                       4000
1
                                     0.333936
                                                                8490.0
2
              0.0
                      -1.531907
                                     2.462778
                                                       5020
                                                               17767.0
              0.0
                      -0.823487
3
                                     8.276223
                                                       4200
                                                               47999.0
             64.0
                        2.240084
                                     0.125226
4
                                                       3000
                                                                5999.0
                                         Processor Processor name Brand name
                             Exynos 9820 Processor
0
                                                            Exynos
                                                                       Samsung
1
                Octa-core Unisoc SC9863 Processor
                                                         Octa-core
                                                                        Lenovo
2
    Qualcomm Snapdragon 720G octa-core processor
                                                          Qualcomm
                                                                         Redmi
3
                Qualcomm Snapdragon 865 Processor
                                                          Qualcomm
                                                                        Realme
            1.5 GHz Quad Core Processor Processor
4
                                                                1.5
                                                                          Lava
Standardizing the Size Reduced features to a range of values between 1 to 16
```

```
[113]: scaler = MinMaxScaler(feature_range=(0,6))
# Select the column you want to scale
column_name = 'Size Reduced'

# Extract the values of the selected column
column_values = new_pdf[column_name].values

# Reshape the column values to a 2D array
reshaped_values = column_values.reshape(-1, 1)

# Apply Min-Max scaling to the column values
scaled_values = scaler.fit_transform(reshaped_values)

# Assign the scaled values back to the dataframe
new_pdf[column_name] = scaled_values.flatten()
```

```
C:\Users\Praveen\AppData\Local\Temp\ipykernel_21376\2745112538.py:15:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy
   new_pdf[column_name] = scaled_values.flatten()
```

```
[114]: new_pdf.head()
```

[114]:				Name	Ratin	g RAM Gb	ROM Gb	\
0	Samsung Galaxy	Note 10 Plus	(Aura Glow, 2	256 GB)	4.	6 12	256	
1		Leno	vo A7 (Blue,	32 GB)	4.	0 2	32	
2	Redmi Note	9 Pro Max (Cha	mpagne Gold,	64 GB)	4.	3 6	64	
3	Re	alme X50 Pro (Moss Green, 2	256 GB)	4.	1 12	256	
4		Lav	a Z61 (Gold,	16 GB)	4.	2 2	16	
	Expandable GB	Size Reduced	Cam Reduced	Batter	ry Mah	Price Rs	\	
0	1.0	0.601181	0.292194		4300	85000.0		
1	128.0	2.729173	0.333936		4000	8490.0		
2	0.0	0.991063	2.462778		5020	17767.0		
3	0.0	1.678583	8.276223		4200	47999.0		
4	64.0	4.651771	0.125226		3000	5999.0		
						.	,	
		_				r name Bra		
0		•	os 9820 Proce			Exynos	Samsung	
1	Oc	ta-core Unisoc	: SC9863 Proce	essor	Oct	a-core	Lenovo	
2	Qualcomm Snap	dragon 720G oc	ta-core proce	essor	Qu	${ t alcomm}$	Redmi	
3	Qu	alcomm Snapdra	gon 865 Proce	ssor	Qu	${ t alcomm}$	Realme	
4	1.5 GH	z Quad Core Pr	ocessor Proce	essor		1.5	Lava	

Imputing. Fill the null values woth 0 in the Expandable GB feature.

[115]: new_pdf['Expandable GB'].fillna(0, inplace=True)

 $\begin{tabular}{l} $C:\Users\Praveen\AppData\Local\Temp\ipykernel_21376\2747860075.py:1: SettingWithCopyWarning: \end{tabular}$

A value is trying to be set on a copy of a slice from a DataFrame

See the caveats in the documentation: https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy new_pdf['Expandable GB'].fillna(0, inplace=True)

[116]: new_pdf.head()

[116]:					Name	Rating	RAM Gb	ROM Gb	\	
	0	Samsung Galaxy	Note 10 Plus	(Aura Glow, 2	56 GB)	4.6	12	256		
	1		Leno	vo A7 (Blue,	32 GB)	4.0	2	32		
	2	Redmi Note	9 Pro Max (Cha	mpagne Gold,	64 GB)	4.3	6	64		
	3	Re	alme X50 Pro (Moss Green, 2	56 GB)	4.1	12	256		
	4		Lav	a Z61 (Gold,	16 GB)	4.2	2	16		
		Expandable GB	Size Reduced	Cam Reduced	Batter	y Mah F	Price Rs	\		
	0	1.0	0.601181	0.292194		4300	85000.0			
	1	128.0	2.729173	0.333936		4000	8490.0			
	2	0.0	0.991063	2.462778		5020	17767.0			
	3	0.0	1.678583	8.276223		4200	47999.0			
	4	64.0	4.651771	0.125226		3000	5999.0			

		Processor	Processor name	Brand name	
0	Exynos 9820	Processor	Exynos	Samsung	
1	Octa-core Unisoc SC9863	Processor	Octa-core	Lenovo	
2	Qualcomm Snapdragon 720G octa-core	processor	Qualcomm	Redmi	
3	Qualcomm Snapdragon 865	Processor	Qualcomm	Realme	
4	1.5 GHz Quad Core Processor	Processor	1.5	Lava	
[]:					
[]:					