

Univariate Analysis

Dataset:

mob_dataset														
	product id	price	sale	weight	resolution	ppi	cpu core	cpu freq	internal mem	ram	rearcam	front cam	battery	thickness
0	203	2357	10.0	135.0	5.20	424.0	8	1.350	16.0	3.000	13.00	8.0	2610	7.4
1	880	1749	10.0	125.0	4.00	233.0	2	1.300	4.0	1.000	3.15	0.0	1700	9.9
2	40	1916	10.0	110.0	4.70	312.0	4	1.200	8.0	1.500	13.00	5.0	2000	7.6
3	99	1315	11.0	118.5	4.00	233.0	2	1.300	4.0	0.512	3.15	0.0	1400	11.0
4	880	1749	11.0	125.0	4.00	233.0	2	1.300	4.0	1.000	3.15	0.0	1700	9.9
...
156	1206	3551	899.5	178.0	5.46	538.0	4	1.875	68.0	6.000	12.00	16.0	4080	8.4
157	1296	3211	899.5	170.0	5.50	534.0	4	1.975	68.0	6.000	20.00	8.0	3400	7.9
158	856	3260	899.5	150.0	5.50	401.0	8	2.200	64.0	4.000	20.00	20.0	3000	6.8
159	1296	3211	899.5	170.0	5.50	534.0	4	1.975	68.0	6.000	20.00	8.0	3400	7.9
160	1131	2536	899.5	202.0	6.00	367.0	8	1.500	16.0	3.000	21.50	16.0	2700	8.4

161 rows x 14 columns

Quantitative Dataset:

	product id	price	sale	weight	resolution	ppi	cpu core	cpu freq	internal mem	ram	rearcam
mean	675.559006	2214.329193	259.571429	154.198758	5.159627	333.993789	4.857143	1.514478	21.520348	2.204994	10.378261
median	774.0	2258.0	106.0	153.0	5.15	294.0	4.0	1.4	16.0	2.0	12.0
mode	14	1734	899.5	223.85	5.5	294.0	4	1.2	16.0	1.0	13.0
Q1:25%	237.0	1734.0	37.0	134.1	4.8	233.0	4.0	1.2	8.0	1.0	5.0
Q2:50%	774.0	2258.0	106.0	153.0	5.15	294.0	4.0	1.4	16.0	2.0	12.0
Q3:75%	1026.0	2744.0	382.0	170.0	5.5	428.0	8.0	1.875	32.0	3.0	16.0
99%	1331.8	4005.8	899.5	223.85	6.55	634.4	8.0	2.7	68.0	6.0	22.1
Q4:100%	1339.0	4259.0	899.5	223.85	6.55	720.5	8.0	2.7	68.0	6.0	23.0
IQR	789.0	1010.0	345.0	35.9	0.7	195.0	4.0	0.675	24.0	2.0	11.0
1.5rule	1183.5	1515.0	517.5	53.85	1.05	292.5	6.0	1.0125	36.0	3.0	16.5
lesser_outlier	-946.5	219.0	-480.5	80.25	3.75	-59.5	-2.0	0.1875	-28.0	-2.0	-11.5
greater_outlier	2209.5	4259.0	899.5	223.85	6.55	720.5	14.0	2.8875	68.0	6.0	32.5
min	10.0	614.0	10.0	80.25	3.75	121.0	0.0	0.1875	0.0	0.0	0.0
max	1339.0	4259.0	899.5	223.85	6.55	720.5	8.0	2.7	68.0	6.0	23.0
kurtosis	-1.318761	-0.14346	0.095916	0.152568	-0.159611	-0.521551	-0.863429	0.453594	0.528894	0.044367	-0.960575
skew	-0.216869	0.020099	1.250054	0.115564	-0.17094	0.41233	-0.009077	-0.27533	1.164121	0.792698	0.106929
var	168799.023059	584769.184705	94414.183929	1219.109998	0.535324	17261.828086	5.973214	0.326577	379.78112	2.591557	38.211993
std_deviation	410.851583	764.702023	307.268911	34.915756	0.731658	131.384276	2.444016	0.57147	19.487974	1.609831	6.181585

Descriptive Analysis:

We can see the exact value of mean, median and percentile details in the above screenshots so I will explain the observation of remaining analysis

Skewness and Kurtosis:

product id:

As per the skew , it has -ve value so it has negative Skewness

As per the Kurtosis, it has -ve which is <3 so it Platykurtic (data points are present in high proximity to mean)

price:

As per the skew , it has +ve value so it has positive Skewness

As per the Kurtosis, it has -ve value which is <3 so it Platykurtic (data points are present in high proximity to mean)

sales:

As per the skew , it has +ve value so it has positive Skewness

As per the Kurtosis, it has +ve which is <3 so it Platykurtic (data points are present in high proximity to mean)

weight:

As per the skew , it has +ve value so it has positive Skewness

As per the Kurtosis, it has +ve which is <3 so it Platykurtic (data points are present in high proximity to mean)

resolution:

As per the skew , it has -ve value so it has negative Skewness

As per the Kurtosis, it has -ve which is <3 so it Platykurtic (data points are present in high proximity to mean)

ppi:

As per the skew , it has 0 value so it has Normal Skewness

As per the Kurtosis, it has 0 which is <3 so it Platykurtic (data points are present in high proximity to mean)

cpu core:

As per the skew , it has -ve value so it has negative Skewness

As per the Kurtosis, it has -ve which is <3 so it Platykurtic (data points are present in high proximity to mean)

cpu freq:

As per the skew , it has +ve value so it has positive Skewness

As per the Kurtosis, it has -ve value which is <3 so it Platykurtic (data points are present in high proximity to mean)

internal mem and ram:

As per the skew , it has +ve value so it has positive Skewness

As per the Kurtosis, it has +ve which is <3 so it Platykurtic (data points are present in high proximity to mean)

rear mem:

As per the skew , it has -ve value so it has negative Skewness

As per the Kurtosis, it has +ve which is <3 so it Platykurtic (data points are present in high proximity to mean)

front cam:

As per the skew , it has +ve value so it has positive Skewness

As per the Kurtosis, it has +ve which is <3 so it Platykurtic (data points are present in high proximity to mean)

battery:

As per the skew , it has -ve value so it has negative Skewness

As per the Kurtosis, it has +ve which is <3 so it Platykurtic (data points are present in high proximity to mean)

thickness:

As per the skew , it has -ve value so it has negative Skewness

As per the Kurtosis, it has +ve which is <3 so it Platykurtic (data points are present in high proximity to mean)

We could see some outliers in those columns

lesser

['weight', 'resolution', 'cpu freq']

greater

['price',
'sale',
'weight',
'resolution',
'ppi',
'internal mem',
'battery',
'thickness']

As per the above analysis , most of the data are present in high proximity to mean and

All the features have normal, positive skewness

We just replace the outliers in all the column by comparing each data with min and max values