

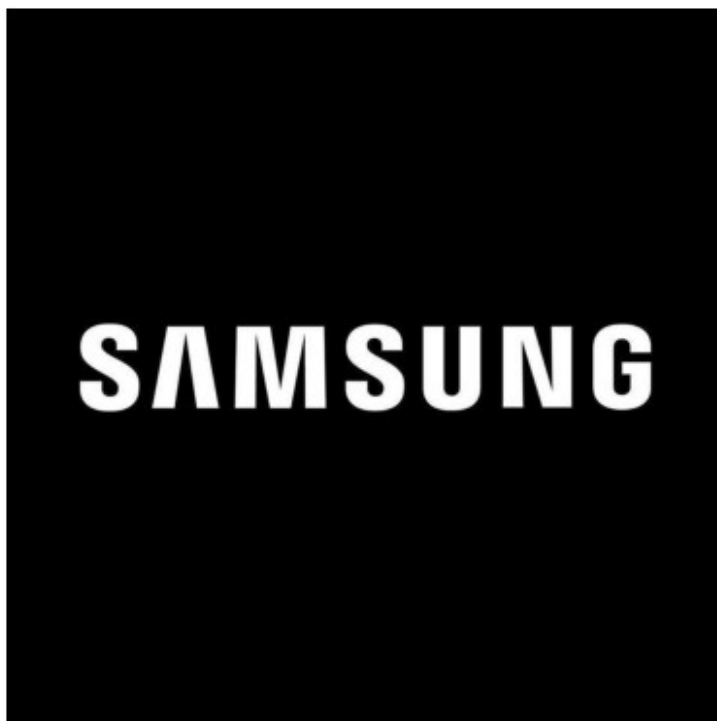
## EDA-OF-SAMSUNG-GALAXY

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import seaborn as sns

import html5lib
```

## LOGO OF SAMSUNG

```
import matplotlib.pyplot as plt
import matplotlib.image as mpimg
img = mpimg.imread('fw44YPe89yxA.jpg')
plt.imshow(img)
plt.axis('off')
plt.show()
```



```
samsung_df = pd.read_csv('Samsung_mobile.csv')
samsung_df.head()
```

	Mobile_Phone	Model	RAM_GB	ROM_GB	Review
Latest_Price_Taka \					
0	Samsung Galaxy	Note 20 ultra	8	256	good
					134999.0
1	Samsung Galaxy	S20 ultra	12	128	good

```

129900.0
2   Samsung Galaxy          S20+      8      128    good
84999.0
3   Samsung Galaxy          Lite       6      128    good
55999.0
4   Samsung Galaxy          A51        6      128    good
45990.0

```

```

    Payment_Method  Discount  Selling_Price
0             cash      3000         131999
1             cash      3000         126900
2             cash      3000           81999
3             cash      2500          53499
4             cash      2500          43490

```

```

samsung_df = pd.read_csv('Samsung_mobile.csv')
samsung_df.tail()

```

```

    Mobile_Phone  Model  RAM_GB  ROM_GB  Review
Latest_Price_Taka \
16 Samsung Galaxy  A10s      2      32  normal
12999.0
17 Samsung Galaxy  A10s      2      32  normal
12999.0
18 Samsung Galaxy  A01       2      16  usual
9999.0
19 Samsung Metro  metro 350      0      0  usual
3550.0
20 Samsung Metro  metro 313      0      0  usual
2750.0

```

```

    Payment_Method  Discount  Selling_Price
16          rocket      1500         11499
17          rocket      1500         11499
18          rocket      1000           8999
19          rocket       500          3050
20          rocket       250          2500

```

```

samsung_df.info()

```

```

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 21 entries, 0 to 20
Data columns (total 9 columns):
#   Column              Non-Null Count  Dtype
---  -
0   Mobile_Phone        21 non-null    object
1   Model               21 non-null    object
2   RAM_GB              21 non-null    int64
3   ROM_GB              21 non-null    int64
4   Review              21 non-null    object

```

```

5 Latest_Price_Taka 21 non-null float64
6 Payment_Method    21 non-null object
7 Discount          21 non-null int64
8 Selling_Price     21 non-null int64

```

```
dtypes: float64(1), int64(4), object(4)
```

```
memory usage: 1.6+ KB
```

```
samsung_df.shape
```

```
(21, 9)
```

```
samsung_df.columns
```

```
Index(['Mobile_Phone', 'Model', 'RAM_GB', 'ROM_GB', 'Review',
      'Latest_Price_Taka', 'Payment_Method', 'Discount',
      'Selling_Price'],
      dtype='object')
```

```
samsung_df.describe(include = 'object')
```

	Mobile_Phone	Model	Review	Payment_Method
count	21	21	21	21
unique	2	15	3	3
top	Samsung Galaxy	A51	good	bkash
freq	19	2	12	8

```
samsung_df.describe()
```

	RAM_GB	ROM_GB	Latest_Price_Taka	Discount
Selling_Price				
count	21.000000	21.000000	21.000000	21.000000
mean	5.142857	93.714286	37610.952381	1940.476190
std	2.920372	62.026722	36891.663754	766.097097
min	0.000000	0.000000	2750.000000	250.000000
25%	3.000000	32.000000	15999.000000	1500.000000
50%	6.000000	128.000000	27490.000000	2000.000000
75%	6.000000	128.000000	42999.000000	2500.000000
max	12.000000	256.000000	134999.000000	3000.000000

```
samsung_df.isnull().sum()
```

Mobile_Phone	0
Model	0

```
RAM_GB      0
ROM_GB      0
Review      0
Latest_Price_Taka  0
Payment_Method  0
Discount    0
Selling_Price  0
dtype: int64
```

>> Data visualization is the representation of data through use of common graphics, such as charts, plots, infographics and even animations.

### *COUNTPLOT*

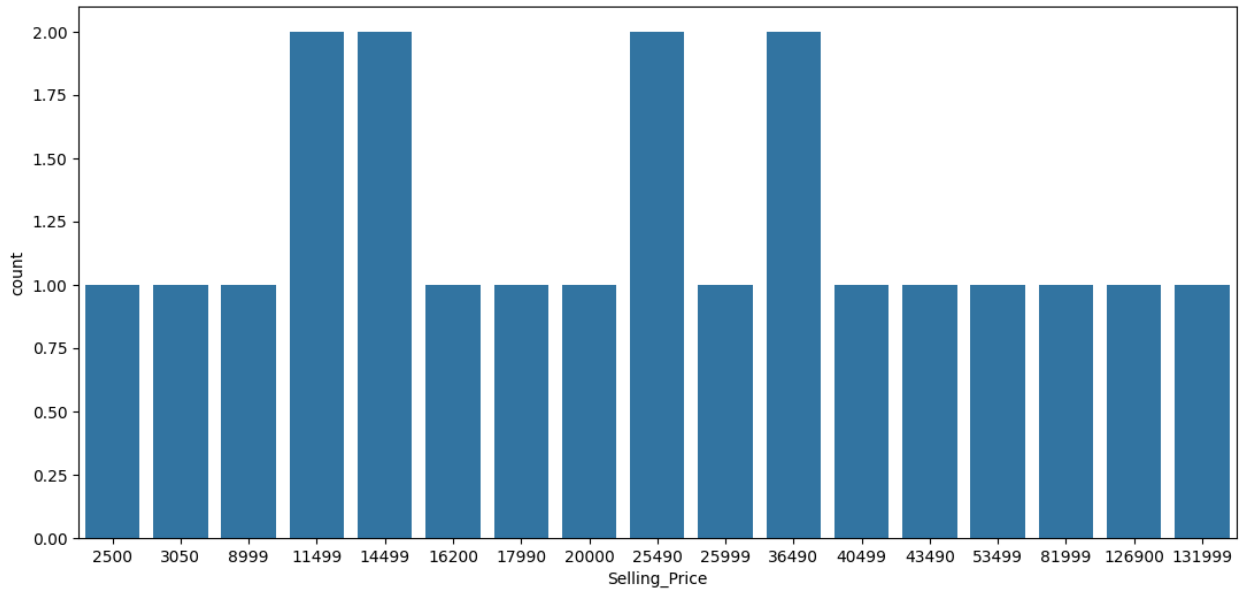
>>A countplot is a statistical graph that uses bars to show the number of observations in each category of a categorical variable

```
samsung_df.Selling_Price
0      131999
1      126900
2       81999
3       53499
4       43490
5       40499
6       36490
7       36490
8       25999
9       25490
10      25490
11      20000
12      17990
13      16200
14      14499
15      14499
16      11499
17      11499
18       8999
19       3050
20       2500
Name: Selling_Price, dtype: int64
```

```
plt.figure(figsize=(13,6))

sns.countplot(x='Selling_Price', data = samsung_df)

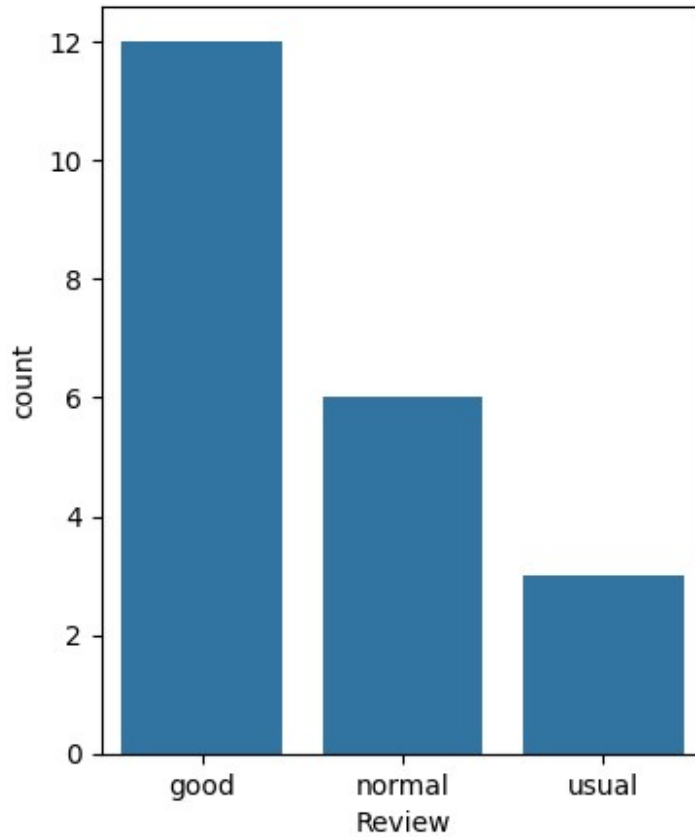
plt.show()
```



```
samsung_df.Review
```

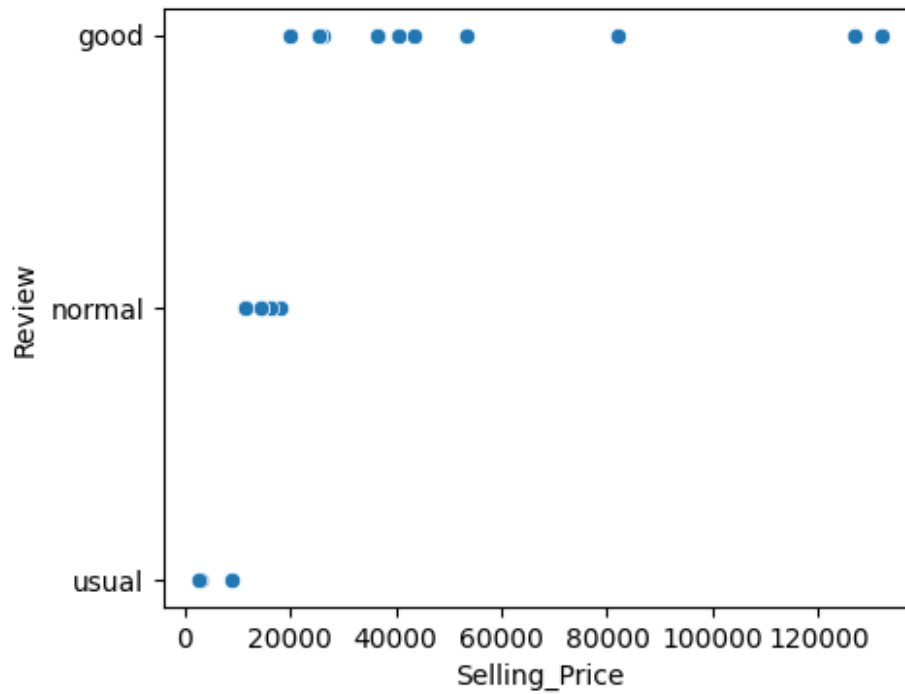
```
0      good
1      good
2      good
3      good
4      good
5      good
6      good
7      good
8      good
9      good
10     good
11     good
12     normal
13     normal
14     normal
15     normal
16     normal
17     normal
18     usual
19     usual
20     usual
Name: Review, dtype: object
```

```
plt.figure(figsize=(4,5))  
  
sns.countplot(x='Review', data = samsung_df)  
plt.show()
```



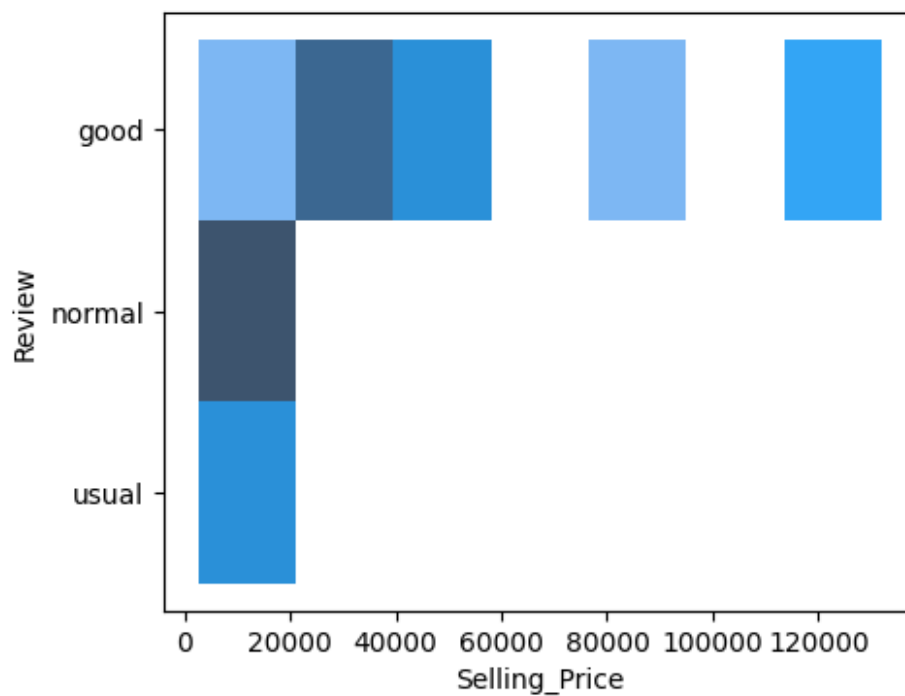
## >>SCATTERPLOT

```
plt.figure(figsize=(5,4))  
  
sns.scatterplot(x='Selling_Price',y='Review',data = samsung_df)  
plt.show()
```



## >>HISTPLOT

```
plt.figure(figsize=(5,4))  
sns.histplot(x='Selling_Price',y='Review',data = samsung_df)  
plt.show()
```



samsung\_df.RAM\_GB

0	8
1	12
2	8
3	6
4	6
5	8
6	6
7	6
8	6
9	6
10	6
11	6
12	6
13	6
14	3
15	3
16	2
17	2
18	2
19	0
20	0

Name: RAM\_GB, dtype: int64

samsung\_df.ROM\_GB

0	256
1	128
2	128
3	128
4	128
5	128
6	128
7	128
8	128
9	128
10	128
11	64
12	64
13	128
14	64
15	32
16	32
17	32
18	16
19	0

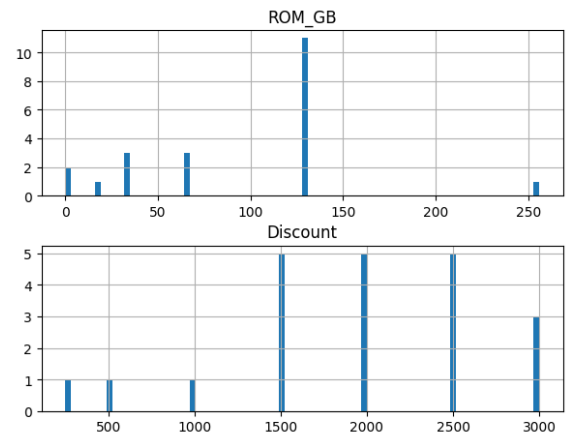
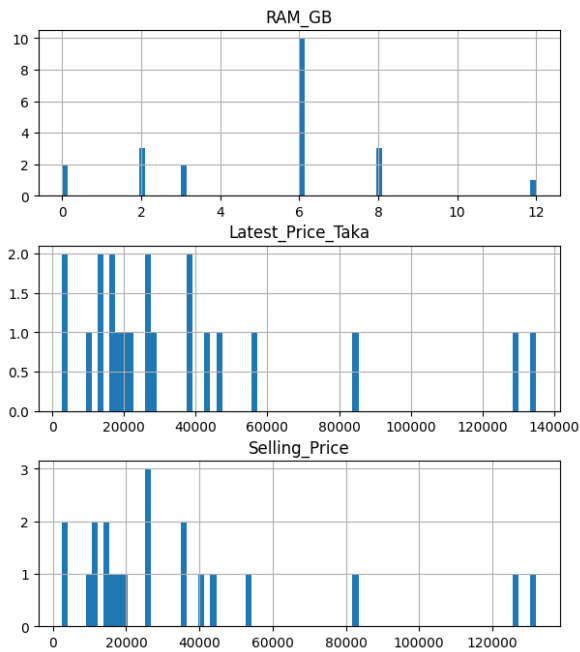


```

20      0
Name: ROM_GB, dtype: int64

samsung_df.hist(bins=80,figsize=(16,8))
plt.show()

```



>>PIE

```

samsung_df.Model

0      Note 20 ultra
1          S20 ultra
2              S20+
3              Lite
4              A51
5              A71
6              A70
7              A70
8              A51
9              M30s
10             M30s
11             F41
12             M31
13             M31
14             A20s
15             A20s
16             A10s
17             A10s
18             A01

```

```
19         metro 350
20         metro 313
Name: Model, dtype: object
```

```
samsung_df.Review
```

```
0      good
1      good
2      good
3      good
4      good
5      good
6      good
7      good
8      good
9      good
10     good
11     good
12   normal
13   normal
14   normal
15   normal
16   normal
17   normal
18   usual
19   usual
20   usual
```

```
Name: Review, dtype: object
```

```
plt.figure(figsize=(4,4))
```

```
# declaring data
```

```
data = samsung_df["Review"].value_counts()
```

```
keys = ['good', 'normal', 'usual']
```

```
# define Seaborn color palette to use
```

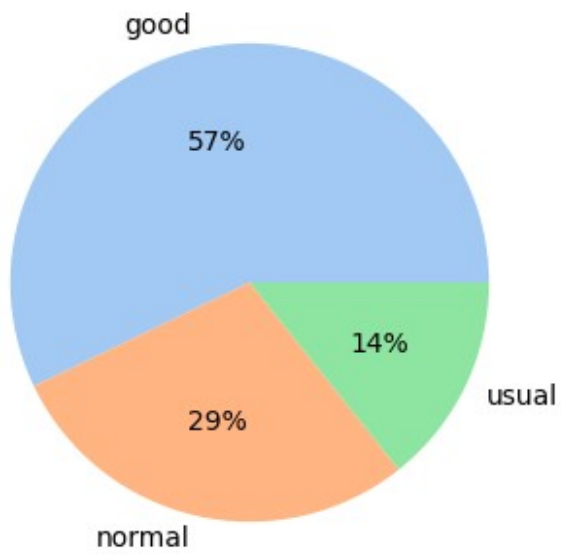
```
palette_color = sns.color_palette('pastel')
```

```
# plotting data on chart
```

```
plt.pie(data, labels=keys, colors=palette_color, autopct='%.0f%%')
```

```
# displaying chart
```

```
plt.show()
```



THE END