In [32]:
```
import pandas as pd
import numpy as np
```

In [33]:
```
df=pd.read_csv("data.csv")
df
```

Out[33]:

	Product	Prices	Quantities
0	samsang	10000	5
1	apple	15000	10
2	metrola	9000	4
3	lava	80000	5
4	oppo	10000	6
5	vivo	11000	5

In [34]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 6 entries, 0 to 5
Data columns (total 3 columns):
 #   Column      Non-Null Count  Dtype 
---  ------      --------------  ----- 
 0   Product     6 non-null      object
 1   Prices      6 non-null      int64 
 2   Quantities  6 non-null      int64 
dtypes: int64(2), object(1)
memory usage: 276.0+ bytes
```

In [35]:
```
df["Prices"]=df["Prices"].astype(int)
df["Quantities"]=df["Quantities"].astype(int)
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 6 entries, 0 to 5
Data columns (total 3 columns):
 #   Column      Non-Null Count  Dtype 
---  ------      --------------  ----- 
 0   Product     6 non-null      object
 1   Prices      6 non-null      int32 
 2   Quantities  6 non-null      int32 
dtypes: int32(2), object(1)
memory usage: 228.0+ bytes
```

In [36]:
```
df.columns
```

Out[36]:
```
Index(['Product', 'Prices', 'Quantities'], dtype='object')
```

In [37]:
```
df["Revenue"]=df["Prices"] *  df["Quantities"]
```

In [38]:
```
df
```

Out[38]:

	Product	Prices	Quantities	Revenue
0	samsang	10000	5	50000
1	apple	15000	10	150000
2	metrola	9000	4	36000
3	lava	80000	5	400000
4	oppo	10000	6	60000
5	vivo	11000	5	55000

In [39]:
```
df["email"]=["sunny@gmail.com", "sunil@gmail.com", "kartik@gmail.com",
             "shankar@gmail.com","amjad@gmail.com","saif@gmail.com"]
df
```

Out[39]:

	Product	Prices	Quantities	Revenue	email
0	samsang	10000	5	50000	sunny@gmail.com
1	apple	15000	10	150000	sunil@gmail.com
2	metrola	9000	4	36000	kartik@gmail.com
3	lava	80000	5	400000	shankar@gmail.com
4	oppo	10000	6	60000	amjad@gmail.com
5	vivo	11000	5	55000	saif@gmail.com

In [40]:
```
# Extract usernames
df['Username'] = df['email'].str.split('@').str[0]

print(df)
```

```
   Product  Prices  Quantities  Revenue             email Username
0  samsang   10000           5    50000   sunny@gmail.com    sunny
1    apple   15000          10   150000   sunil@gmail.com    sunil
2  metrola    9000           4    36000  kartik@gmail.com   kartik
3     lava   80000           5   400000  shankar@gmail.com  shankar
4     oppo   10000           6    60000   amjad@gmail.com    amjad
5     vivo   11000           5    55000    saif@gmail.com     saif
```

In [41]:
```
df["discount_percentage%"]=[20,10,15,20,30,40]
```

In [42]:
```
df
```

Out[42]:

	Product	Prices	Quantities	Revenue	email	Username	discount_percentage%
0	samsang	10000	5	50000	sunny@gmail.com	sunny	2
1	apple	15000	10	150000	sunil@gmail.com	sunil	1
2	metrola	9000	4	36000	kartik@gmail.com	kartik	1
3	lava	80000	5	400000	shankar@gmail.com	shankar	2
4	oppo	10000	6	60000	amjad@gmail.com	amjad	3
5	vivo	11000	5	55000	saif@gmail.com	saif	4

In [43]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 6 entries, 0 to 5
Data columns (total 7 columns):
 #   Column               Non-Null Count  Dtype 
---  ------               --------------  ----- 
 0   Product              6 non-null      object
 1   Prices               6 non-null      int32 
 2   Quantities           6 non-null      int32 
 3   Revenue              6 non-null      int32 
 4   email                6 non-null      object
 5   Username             6 non-null      object
 6   discount_percentage%  6 non-null      int64 
dtypes: int32(3), int64(1), object(3)
memory usage: 396.0+ bytes
```

In [44]:
```
df["discount_percentage%"]=df["discount_percentage%"].astype(int)
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 6 entries, 0 to 5
Data columns (total 7 columns):
 #   Column               Non-Null Count  Dtype 
---  ------               --------------  ----- 
 0   Product              6 non-null      object
 1   Prices               6 non-null      int32 
 2   Quantities           6 non-null      int32 
 3   Revenue              6 non-null      int32 
 4   email                6 non-null      object
 5   Username             6 non-null      object
 6   discount_percentage%  6 non-null      int32 
dtypes: int32(4), object(3)
memory usage: 372.0+ bytes
```

In [45]:
```
# Calculate the final price after applying the discount
df['Final_Price'] = df['Prices'] * (1 - df['discount_percentage%']/100)

# Display the result
df['Final_Price']
```

Out[45]:
```
0     8000.0
1    13500.0
2     7650.0
3    64000.0
4     7000.0
5     6600.0
Name: Final_Price, dtype: float64
```

In [46]:
```
df['Ratings']= [4.5, 3.8, 4.2, 4.0,5.0,3.5]
df
```

Out[46]:

	Product	Prices	Quantities	Revenue	email	Username	discount_percentage%
0	samsang	10000	5	50000	sunny@gmail.com	sunny	2
1	apple	15000	10	150000	sunil@gmail.com	sunil	1
2	metrola	9000	4	36000	kartik@gmail.com	kartik	1
3	lava	80000	5	400000	shankar@gmail.com	shankar	2
4	oppo	10000	6	60000	amjad@gmail.com	amjad	3
5	vivo	11000	5	55000	saif@gmail.com	saif	4

In [47]:
```
# Corrected code: Sort by Price (ascending)
sorted_df = df.sort_values(by=['Prices'], ascending=[True])

print(sorted_df)
```

```
   Product  Prices  Quantities  Revenue             email Username  \
2  metrola    9000           4    36000   kartik@gmail.com   kartik
0  samsang   10000           5    50000    sunny@gmail.com    sunny
4     oppo   10000           6    60000    amjad@gmail.com    amjad
5     vivo   11000           5    55000     saif@gmail.com     saif
1    apple   15000          10   150000    sunil@gmail.com    sunil
3     lava   80000           5   400000  shankar@gmail.com  shankar

   discount_percentage%  Final_Price  Ratings
2                    15       7650.0      4.2
0                    20       8000.0      4.5
4                    30       7000.0      5.0
5                    40       6600.0      3.5
1                    10      13500.0      3.8
3                    20      64000.0      4.0
```

In [48]:
```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 6 entries, 0 to 5
Data columns (total 9 columns):
 #   Column               Non-Null Count  Dtype  
---  ------               --------------  -----  
 0   Product              6 non-null      object 
 1   Prices               6 non-null      int32  
 2   Quantities           6 non-null      int32  
 3   Revenue              6 non-null      int32  
 4   email                6 non-null      object 
 5   Username             6 non-null      object 
 6   discount_percentage%  6 non-null      int32  
 7   Final_Price          6 non-null      float64
 8   Ratings              6 non-null      float64
dtypes: float64(2), int32(4), object(3)
memory usage: 468.0+ bytes
```

In [49]:
```
df["Ratings"]=df["Ratings"].astype(float)
df["Ratings"]
```

Out[49]:
```
0    4.5
1    3.8
2    4.2
3    4.0
4    5.0
5    3.5
Name: Ratings, dtype: float64
```

In [50]:
```
df1 = df.sort_values(by=['Ratings', 'Final_Price'], ascending=[False, True])
df1
```

Out[50]:

	Product	Prices	Quantities	Revenue	email	Username	discount_percentage%
4	oppo	10000	6	60000	amjad@gmail.com	amjad	3
0	samsang	10000	5	50000	sunny@gmail.com	sunny	2
2	metrola	9000	4	36000	kartik@gmail.com	kartik	1
3	lava	80000	5	400000	shankar@gmail.com	shankar	2
1	apple	15000	10	150000	sunil@gmail.com	sunil	1
5	vivo	11000	5	55000	saif@gmail.com	saif	4

In [51]:
```
df.sort_values(by='Ratings',ascending=[False], inplace=True)
df
```

Out[51]:

	Product	Prices	Quantities	Revenue	email	Username	discount_percentage%
4	oppo	10000	6	60000	amjad@gmail.com	amjad	3
0	samsang	10000	5	50000	sunny@gmail.com	sunny	2
2	metrola	9000	4	36000	kartik@gmail.com	kartik	1
3	lava	80000	5	400000	shankar@gmail.com	shankar	2
1	apple	15000	10	150000	sunil@gmail.com	sunil	1
5	vivo	11000	5	55000	saif@gmail.com	saif	4

In [52]:
```
df.sort_values(by='Final_Price',ascending=[True], inplace=True)
df
```

Out[52]:

	Product	Prices	Quantities	Revenue	email	Username	discount_percentage%
5	vivo	11000	5	55000	saif@gmail.com	saif	4
4	oppo	10000	6	60000	amjad@gmail.com	amjad	3
2	metrola	9000	4	36000	kartik@gmail.com	kartik	1
0	samsang	10000	5	50000	sunny@gmail.com	sunny	2
1	apple	15000	10	150000	sunil@gmail.com	sunil	1
3	lava	80000	5	400000	shankar@gmail.com	shankar	2

In [53]:
```
df1=pd.read_csv("data1.csv")
df1
```

Out[53]:

	Product_name	Prices
0	samsang	10000
1	apple	15000
2	Ericson	NaN
3	metrola	9000
4	lava	80000
5	Google	NaN
6	oppo	10000
7	vivo	11000
8	Carbon	$12,000
9	lenova	$15,000

In [54]:
```
df1.isnull().sum()
```

Out[54]:
```
Product_name    0
Prices          2
dtype: int64
```

In [55]:
```
df1.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10 entries, 0 to 9
Data columns (total 2 columns):
 #   Column        Non-Null Count  Dtype 
---  ------        --------------  ----- 
 0   Product_name  10 non-null     object
 1   Prices        8 non-null      object
dtypes: object(2)
memory usage: 292.0+ bytes
```

In [56]:
```
 # Remove currency symbols
df1['Prices'] = df1['Prices'].replace('[\$,]', '', regex=True)
df1['Prices']
```

Out[56]:
```
0    10000
1    15000
2      NaN
3     9000
4    80000
5      NaN
6    10000
7    11000
8    12000
9    15000
Name: Prices, dtype: object
```

In [57]:
```
df1["Prices"] = df1["Prices"] .fillna(df1["Prices"] .median())
df1
```

Out[57]:

	Product_name	Prices
0	samsang	10000
1	apple	15000
2	Ericson	11500.0
3	metrola	9000
4	lava	80000
5	Google	11500.0
6	oppo	10000
7	vivo	11000
8	Carbon	12000
9	lenova	15000

In [58]:
```
df1["Product_name"]=df1["Product_name"].str.lower().str.strip()
```

In [59]:
```
df1
```

Out[59]:

	Product_name	Prices
0	samsang	10000
1	apple	15000
2	ericson	11500.0
3	metrola	9000
4	lava	80000
5	google	11500.0
6	oppo	10000
7	vivo	11000
8	carbon	12000
9	lenova	15000

In [60]:
```
df1['Product_name'] = df1['Product_name'].str.title()
```

In [61]:
```
df1
```

Out[61]:

	Product_name	Prices
0	Samsang	10000
1	Apple	15000
2	Ericson	11500.0
3	Metrola	9000
4	Lava	80000
5	Google	11500.0
6	Oppo	10000
7	Vivo	11000
8	Carbon	12000
9	Lenova	15000