```
pip install pandas
```

```
Requirement already satisfied: pandas in /usr/local/lib/python3.12/dist-p
Requirement already satisfied: numpy>=1.26.0 in /usr/local/lib/python3.12
Requirement already satisfied: python-dateutil>=2.8.2 in /usr/local/lib/p
Requirement already satisfied: pytz>=2020.1 in /usr/local/lib/python3.12/
Requirement already satisfied: tzdata>=2022.7 in /usr/local/lib/python3.1
Requirement already satisfied: six>=1.5 in /usr/local/lib/python3.12/dist
```

```
data = {
    "Product_ID": [101, 102, 103, 104, 105, 106, 107, 108, 109, 110],
    "Product_Name": ["Echo Dot", "Fire TV Stick", "Kindle Paperwhite", "AmazonI
    "Category": ["Electronics", "Electronics", "Books", "Home & Kitchen", "Ele
    "Price": [4499, 3999, 12999, 499, 8999, 7999, 299, 12999, 11999, 24999],
    "Rating": [4.5, 4.3, 4.7, 4.2, 4.4, 4.1, 4.0, 4.6, 4.5, 4.8]
}
```

```
import pandas as pd
```

```
df = pd.DataFrame(data)
```

```
df
```

	Product_ID	Product_Name	Category	Price	Rating
0	101	Echo Dot	Electronics	4499	4.5
1	102	Fire TV Stick	Electronics	3999	4.3
2	103	Kindle Paperwhite	Books	12999	4.7
3	104	AmazonBasics Battery	Home & Kitchen	499	4.2
4	105	Echo Show	Electronics	8999	4.4
5	106	Fire HD 8 Tablet	Electronics	7999	4.1
6	107	Amazon Prime Mug	Home & Kitchen	299	4.0
7	108	Echo Buds	Electronics	12999	4.6
8	109	Fire TV Cube	Electronics	11999	4.5
9	110	Kindle Oasis	Books	24999	4.8

```
df.head()
```

	Product_ID	Product_Name	Category	Price	Rating
0	101	Echo Dot	Electronics	4499	4.5
1	102	Fire TV Stick	Electronics	3999	4.3
2	103	Kindle Paperwhite	Books	12999	4.7
3	104	AmazonBasics Battery	Home & Kitchen	499	4.2
4	105	Echo Show	Electronics	8999	4.4

```
df.tail()
```

	Product_ID	Product_Name	Category	Price	Rating
5	106	Fire HD 8 Tablet	Electronics	7999	4.1
6	107	Amazon Prime Mug	Home & Kitchen	299	4.0
7	108	Echo Buds	Electronics	12999	4.6
8	109	Fire TV Cube	Electronics	11999	4.5
9	110	Kindle Oasis	Books	24999	4.8

```
df.head(2)
```

	Product_ID	Product_Name	Category	Price	Rating
0	101	Echo Dot	Electronics	4499	4.5
1	102	Fire TV Stick	Electronics	3999	4.3

```
#Show data type of rows
df.dtypes
```

	0
Product_ID	int64
Product_Name	object
Category	object
Price	int64
Rating	float64

dtype: object

```
#No. of rows and columns (show)

df.size
```

```
50
```

```
#Show stats for data.
df.describe()
```

	Product_ID	Price	Rating
count	10.00000	10.000000	10.000000
mean	105.50000	8929.000000	4.410000
std	3.02765	7387.233733	0.260128
min	101.00000	299.000000	4.000000
25%	103.25000	4124.000000	4.225000
50%	105.50000	8499.000000	4.450000
75%	107.75000	12749.000000	4.575000
max	110.00000	24999.000000	4.800000

```
#Find mean of mean of rating or average rating.

Avg= df['Rating'].mean()
Avg
```

```
np.float64(4.41)
```

```
#Find median:
Median=df['Rating'].median()
Median
```

```
4.45
```

```
#Find mode:
Mode=df['Rating'].mode()
Mode
```

	Rating
0	4.5

dtype: float64

```
#Show all column names.

df.columns
```

```
Index(['Product_ID', 'Product_Name', 'Category', 'Price', 'Rating'],
dtype='object')
```

```
#All values:

df.values
```

```
array([[101, 'Echo Dot', 'Electronics', 4499, 4.5],
       [102, 'Fire TV Stick', 'Electronics', 3999, 4.3],
       [103, 'Kindle Paperwhite', 'Books', 12999, 4.7],
       [104, 'AmazonBasics Battery', 'Home & Kitchen', 499, 4.2],
       [105, 'Echo Show', 'Electronics', 8999, 4.4],
       [106, 'Fire HD 8 Tablet', 'Electronics', 7999, 4.1],
       [107, 'Amazon Prime Mug', 'Home & Kitchen', 299, 4.0],
       [108, 'Echo Buds', 'Electronics', 12999, 4.6],
       [109, 'Fire TV Cube', 'Electronics', 11999, 4.5],
       [110, 'Kindle Oasis', 'Books', 24999, 4.8]], dtype=object)
```

```
#Show dataset summary:

df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 10 entries, 0 to 9
Data columns (total 5 columns):
 #   Column        Non-Null Count  Dtype
---  ------        --------------  -----
 0   Product_ID    10 non-null     int64
 1   Product_Name  10 non-null     object
 2   Category      10 non-null     object
 3   Price         10 non-null     int64
 4   Rating        10 non-null     float64
dtypes: float64(1), int64(2), object(2)
memory usage: 532.0+ bytes
```

```
#Access info with integer indexing:

df.iloc[5,1]
```

```
'Fire HD 8 Tablet'
```

```
#Access info with column label.

df.loc[0,'Product_Name']
```

```
'Echo Dot'
```

```
df.loc[1,'Product_Name']
```

```
'Fire TV Stick'
```

```
df.iloc[5,1]
```

```
'Fire HD 8 Tablet'
```

```
#Filter for Price>5000.

Filter_salary=df.query('Price > 5000')
Filter_salary
```

	Product_ID	Product_Name	Category	Price	Rating
2	103	Kindle Paperwhite	Books	12999	4.7
4	105	Echo Show	Electronics	8999	4.4
5	106	Fire HD 8 Tablet	Electronics	7999	4.1
7	108	Echo Buds	Electronics	12999	4.6
8	109	Fire TV Cube	Electronics	11999	4.5
9	110	Kindle Oasis	Books	24999	4.8

```
#Remove any any column:

drop_2 =df.drop(columns=['Price'])
drop_2
```

	Product_ID	Product_Name	Category	Rating
0	101	Echo Dot	Electronics	4.5
1	102	Fire TV Stick	Electronics	4.3
2	103	Kindle Paperwhite	Books	4.7
3	104	AmazonBasics Battery	Home & Kitchen	4.2
4	105	Echo Show	Electronics	4.4
5	106	Fire HD 8 Tablet	Electronics	4.1
6	107	Amazon Prime Mug	Home & Kitchen	4.0
7	108	Echo Buds	Electronics	4.6
8	109	Fire TV Cube	Electronics	4.5
9	110	Kindle Oasis	Books	4.8

```python
#Change Product_Name to Name:

Rename_1=df.rename(columns={'Product_Name' : 'Name'})
Rename_1
```

	Product_ID	Name	Category	Price	Rating
0	101	Echo Dot	Electronics	4499	4.5
1	102	Fire TV Stick	Electronics	3999	4.3
2	103	Kindle Paperwhite	Books	12999	4.7
3	104	AmazonBasics Battery	Home & Kitchen	499	4.2
4	105	Echo Show	Electronics	8999	4.4
5	106	Fire HD 8 Tablet	Electronics	7999	4.1
6	107	Amazon Prime Mug	Home & Kitchen	299	4.0
7	108	Echo Buds	Electronics	12999	4.6
8	109	Fire TV Cube	Electronics	11999	4.5
9	110	Kindle Oasis	Books	24999	4.8

```
#Sort(makes data in ascending order)

Sort=df.sort_values(by='Rating')
Sort
```

	Product_ID	Name	Category	Price	Rating
6	107	Amazon Prime Mug	Home & Kitchen	299	4.0
5	106	Fire HD 8 Tablet	Electronics	7999	4.1
3	104	AmazonBasics Battery	Home & Kitchen	499	4.2
1	102	Fire TV Stick	Electronics	3999	4.3
4	105	Echo Show	Electronics	8999	4.4
0	101	Echo Dot	Electronics	4499	4.5
8	109	Fire TV Cube	Electronics	11999	4.5
7	108	Echo Buds	Electronics	12999	4.6
2	103	Kindle Paperwhite	Books	12999	4.7
9	110	Kindle Oasis	Books	24999	4.8

```
#Fill missing values with 0

Fillna_0=df.fillna(0)
Fillna_0
```

	Product_ID	Name	Category	Price	Rating
0	101	Echo Dot	Electronics	4499	4.5
1	102	Fire TV Stick	Electronics	3999	4.3
2	103	Kindle Paperwhite	Books	12999	4.7
3	104	AmazonBasics Battery	Home & Kitchen	499	4.2
4	105	Echo Show	Electronics	8999	4.4
5	106	Fire HD 8 Tablet	Electronics	7999	4.1
6	107	Amazon Prime Mug	Home & Kitchen	299	4.0
7	108	Echo Buds	Electronics	12999	4.6
8	109	Fire TV Cube	Electronics	11999	4.5
9	110	Kindle Oasis	Books	24999	4.8

```python
#Remove duplicate objects if any.

Drop_dupl=df.drop_duplicates()
Drop_dupl
```

	Product_ID	Name	Category	Price	Rating
0	101	Echo Dot	Electronics	4499	4.5
1	102	Fire TV Stick	Electronics	3999	4.3
2	103	Kindle Paperwhite	Books	12999	4.7
3	104	AmazonBasics Battery	Home & Kitchen	499	4.2
4	105	Echo Show	Electronics	8999	4.4
5	106	Fire HD 8 Tablet	Electronics	7999	4.1
6	107	Amazon Prime Mug	Home & Kitchen	299	4.0
7	108	Echo Buds	Electronics	12999	4.6
8	109	Fire TV Cube	Electronics	11999	4.5
9	110	Kindle Oasis	Books	24999	4.8

```
#Replace and rename have a difference:

#Rename: to change label, ex: City to location.
#Replace: to change value, ex: kindle oasis to bridge store
```

```
#Replace Echo Dot with Echoes in the dataset.

Replace_a=df.replace({'Echo Dot' : 'Echoes'})
Replace_a
```

	Product_ID	Name	Category	Price	Rating
0	101	Echoes	Electronics	4499	4.5
1	102	Fire TV Stick	Electronics	3999	4.3
2	103	Kindle Paperwhite	Books	12999	4.7
3	104	AmazonBasics Battery	Home & Kitchen	499	4.2
4	105	Echo Show	Electronics	8999	4.4
5	106	Fire HD 8 Tablet	Electronics	7999	4.1
6	107	Amazon Prime Mug	Home & Kitchen	299	4.0
7	108	Echo Buds	Electronics	12999	4.6
8	109	Fire TV Cube	Electronics	11999	4.5
9	110	Kindle Oasis	Books	24999	4.8

```
#Use group by for Salary and Name:

Group_df=df.groupby('Name')['Price'].sum()
Group_df
```

Name	Price
Amazon Prime Mug	299
AmazonBasics Battery	499
Echo Buds	12999
Echo Dot	4499
Echo Show	8999
Fire HD 8 Tablet	7999
Fire TV Cube	11999
Fire TV Stick	3999
Kindle Oasis	24999
Kindle Paperwhite	12999

dtype: int64

```
#Now write upto 5 names in grouped data.
Group_df.head(5)
```

Price

Name

Amazon Prime Mug	299
AmazonBasics Battery	499
Echo Buds	12999
Echo Dot	4499
Echo Show	8999

dtype: int64

```
#Find out aggregate for price with names.
Agg_df=df.groupby('Name').agg({'Price':'mean'})
Agg_df
```

Price

Name

Amazon Prime Mug	299.0
AmazonBasics Battery	499.0
Echo Buds	12999.0
Echo Dot	4499.0
Echo Show	8999.0
Fire HD 8 Tablet	7999.0
Fire TV Cube	11999.0
Fire TV Stick	3999.0
Kindle Oasis	24999.0
Kindle Paperwhite	12999.0

```
#Counting non null values.
Count=df.groupby('Name')['Price'].count()
Count
```

| | Price |
Name	
Amazon Prime Mug	1
AmazonBasics Battery	1
Echo Buds	1
Echo Dot	1
Echo Show	1
Fire HD 8 Tablet	1
Fire TV Cube	1
Fire TV Stick	1
Kindle Oasis	1
Kindle Paperwhite	1

```
#remove rows or columns with missing value.

Df_clean=df.dropna()
Df_clean
```

Product_ID	Name	Category	Price	Rating