In [2]:
```
import pandas as pd
```

In [38]:
```
data = pd.read_excel("data.xlsx")
```

In [5]:
```
data.shape
```

Out[5]:
```
(30, 4)
```

In [6]:
```
data.head(5)
```

Out[6]:

	Roll No.	Name	Marks	Gender
0	1	Ava Johnson	76	Female
1	2	Liam Smith	89	Male
2	3	Olivia Brown	45	Female
3	4	Ethan Davis	64	Male
4	5	Emma Wilson	92	Female

In [7]:
```
data.tail(5)
```

Out[7]:

	Roll No.	Name	Marks	Gender
25	26	Daniel Harris	77	Male
26	27	Ella Robinson	64	Female
27	28	Matthew Collins	92	Male
28	29	Avery Lee	85	Female
29	30	Sahil Jadhav	900	Male

In [8]:
```
data.describe()
```

Out[8]:

	Roll No.	Marks
count	30.000000	30.000000
mean	15.500000	98.466667
std	8.803408	152.465651
min	1.000000	18.000000
25%	8.250000	60.250000
50%	15.500000	74.500000
75%	22.750000	87.250000
max	30.000000	900.000000

In [9]:
```
data.columns
```

Out[9]:
```
Index(['Roll No.', 'Name', 'Marks', 'Gender'], dtype='object')
```

In [10]:
```
data.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 30 entries, 0 to 29
Data columns (total 4 columns):
 #   Column    Non-Null Count  Dtype 
---  ------    --------------  ----- 
 0   Roll No.  30 non-null     int64 
 1   Name      30 non-null     object
 2   Marks     30 non-null     int64 
 3   Gender    30 non-null     object
dtypes: int64(2), object(2)
memory usage: 1.1+ KB
```

In [46]:
```
data.isna().sum()
```

Out[46]:
```
Roll No.    0
Name        0
Marks       1
Gender      0
dtype: int64
```

In [12]:
```
import numpy as np
```

In [13]:
```
mean_value = np.mean(data['Marks'])
```

In [14]:
```
print("Mean of the column:", mean_value)
```

```
Mean of the column: 98.46666666666667
```

In [17]:
```
median_value = np.median(data["Marks"])
```

In [18]:
```
print("Median:", median_value)
```

```
Median: 74.5
```

In [22]:
```
column_data = data["Marks"].values

unique_values, counts = np.unique(column_data, return_counts=True)

mode_value = unique_values[np.argmax(counts)]
```

In [23]:
```
print("Mode:", mode_value)
```

```
Mode: 64
```

In [25]:
```
data.describe()
```

Out[25]:

	Roll No.	Marks
count	30.000000	30.000000
mean	15.500000	98.466667
std	8.803408	152.465651
min	1.000000	18.000000
25%	8.250000	60.250000
50%	15.500000	74.500000
75%	22.750000	87.250000
max	30.000000	900.000000

In [31]:
```
data.std(numeric_only=True)
```

Out[31]:
```
Roll No.      8.803408
Marks       152.465651
dtype: float64
```

In [32]:
```
data.mean(numeric_only=True)
```

Out[32]:
```
Roll No.    15.500000
Marks       98.466667
dtype: float64
```

In [34]:
```
data.median(numeric_only=True)
```

Out[34]:
```
Roll No.    15.5
Marks       74.5
dtype: float64
```

In [35]:
```
data.var(numeric_only=True)
```

Out[35]:
```
Roll No.       77.500000
Marks       23245.774713
dtype: float64
```

In [36]:
```
data.skew(numeric_only=True)
```

Out[36]:
```
Roll No.    0.0000
Marks       5.3521
dtype: float64
```

In [40]:
```
data.columns
```

Out[40]:
```
Index(['Roll No.', 'Name', 'Marks', 'Gender'], dtype='object')
```

In [41]:
```
#Using mean fill null values
```

In [42]:
```
df2 = data.copy()
```

In [48]:
```
for cols in data.columns:
    if df2[cols].isnull().any():
        df2[cols] = df2[cols].fillna(value = df2[cols].mean())
```

In [49]:
```
df2.isnull().sum()
```

Out[49]:
```
Roll No.    0
Name        0
Marks       0
Gender      0
dtype: int64
```

In [50]:
```
df3 = data.copy()
```

In [59]:
```
for cols in data.columns:
    if df3[cols].isnull().any():
        df3[cols] = df3[cols].ffill()
```

In [55]:
```
df3.isnull().sum()
```

Out[55]:
```
Roll No.    0
Name        0
Marks       0
Gender      0
dtype: int64
```

In [53]:
```
df3.shape
```

Out[53]:
```
(30, 4)
```

In [56]:
```
df3
```

Out[56]:

	Roll No.	Name	Marks	Gender
0	1	Ava Johnson	76.0	Female
1	2	Liam Smith	89.0	Male
2	3	Olivia Brown	45.0	Female
3	4	Ethan Davis	64.0	Male
4	5	Emma Wilson	92.0	Female
5	6	Noah Taylor	57.0	Male
6	7	Isabella Lee	57.0	Female
7	8	James Harris	83.0	Male
8	9	Sophia Clark	52.0	Female
9	10	Jackson Lewis	71.0	Male
10	11	Mia Walker	98.0	Female
11	12	Alexander Hall	65.0	Male
12	13	Charlotte Allen	47.0	Female
13	14	Lucas Young	81.0	Male
14	15	Amelia Scott	69.0	Female
15	16	Mason King	60.0	Male
16	17	Harper Perez	88.0	Female
17	18	William White	73.0	Male
18	19	Abigail Green	95.0	Female
19	20	Benjamin Adams	50.0	Male
20	21	Grace Nelson	76.0	Female
21	22	Samuel Carter	82.0	Male
22	23	Chloe Mitchell	61.0	Female
23	24	Oliver Moore	90.0	Male
24	25	Lily Turner	54.0	Female
25	26	Daniel Harris	77.0	Male
26	27	Ella Robinson	64.0	Female
27	28	Matthew Collins	92.0	Male
28	29	Avery Lee	85.0	Female
29	30	Sahil Jadhav	900.0	Male

In [57]:
```
df4 = data.copy()
```

In [58]:
```
for cols in data.columns:
    if df4[cols].isnull().any():
        df4[cols] = df4[cols].bfill()
```

In [61]:
```
df4
```

Out[61]:

	Roll No.	Name	Marks	Gender
0	1	Ava Johnson	76.0	Female
1	2	Liam Smith	89.0	Male
2	3	Olivia Brown	45.0	Female
3	4	Ethan Davis	64.0	Male
4	5	Emma Wilson	92.0	Female
5	6	Noah Taylor	57.0	Male
6	7	Isabella Lee	83.0	Female
7	8	James Harris	83.0	Male
8	9	Sophia Clark	52.0	Female
9	10	Jackson Lewis	71.0	Male
10	11	Mia Walker	98.0	Female
11	12	Alexander Hall	65.0	Male
12	13	Charlotte Allen	47.0	Female
13	14	Lucas Young	81.0	Male
14	15	Amelia Scott	69.0	Female
15	16	Mason King	60.0	Male
16	17	Harper Perez	88.0	Female
17	18	William White	73.0	Male
18	19	Abigail Green	95.0	Female
19	20	Benjamin Adams	50.0	Male
20	21	Grace Nelson	76.0	Female
21	22	Samuel Carter	82.0	Male
22	23	Chloe Mitchell	61.0	Female
23	24	Oliver Moore	90.0	Male
24	25	Lily Turner	54.0	Female
25	26	Daniel Harris	77.0	Male
26	27	Ella Robinson	64.0	Female
27	28	Matthew Collins	92.0	Male
28	29	Avery Lee	85.0	Female
29	30	Sahil Jadhav	900.0	Male

In [ ]: