Santhosh Gopi B

import libraries

In [1]:
```
import pandas as pd
import numpy as np
```

Read file

In [2]:
```
df=pd.read_csv(r"E:\154\4_drug200.csv")
```

Display File

In [3]:
```
display(df)
```

	Age	Sex	BP	Cholesterol	Na_to_K	Drug
0	23	F	HIGH	HIGH	25.355	drugY
1	47	M	LOW	HIGH	13.093	drugC
2	47	M	LOW	HIGH	10.114	drugC
3	28	F	NORMAL	HIGH	7.798	drugX
4	61	F	LOW	HIGH	18.043	drugY
...	...	...	...	...	...	...
195	56	F	LOW	HIGH	11.567	drugC
196	16	M	LOW	HIGH	12.006	drugC
197	52	M	NORMAL	HIGH	9.894	drugX
198	23	M	NORMAL	NORMAL	14.020	drugX
199	40	F	LOW	NORMAL	11.349	drugX

200 rows × 6 columns

MEAN

In [4]:
```
df.mean()
```

Out[4]:
```
Age         44.315000
Na_to_K     16.084485
dtype: float64
```

Median

In [5]:
```
df.median()
```

Out[5]:
```
Age        45.0000
Na_to_K    13.9365
dtype: float64
```

Mode

In [6]:
```
df.mode()
```

Out[6]:

	Age	Sex	BP	Cholesterol	Na_to_K	Drug
0	47.0	M	HIGH	HIGH	12.006	drugY
1	NaN	NaN	NaN	NaN	18.295	NaN

Descibe

In [7]:
```
df.describe()
```

Out[7]:

	Age	Na_to_K
count	200.000000	200.000000
mean	44.315000	16.084485
std	16.544315	7.223956
min	15.000000	6.269000
25%	31.000000	10.445500
50%	45.000000	13.936500
75%	58.000000	19.380000
max	74.000000	38.247000

Sum

In [8]:
```
df.sum()
```

Out[8]:
```
Age                                                          8863
Sex            FMMFFFFMMMFFMFFFMMMFMMMFFFMFFMFMMFMMMMFMFFMMFF...
BP             HIGHLOWLOWNORMALLOWNORMALNORMALLOWNORMALLOWLOW...
Cholesterol    HIGHHIGHHIGHHIGHHIGHHIGHHIGHHIGHHIGHNORMALHIGH...
Na_to_K                                                  3216.897
Drug           drugYdrugCdrugCdrugXdrugYdrugXdrugYdrugCdrugYd...
dtype: object
```

Cumulative Sum

In [9]:
```
df.cumsum()
```

Out[9]:

	Age	Sex	
0	23	F	
1	70	FM	
2	117	FMM	
3	145	FMMF	
4	206	FMMFF	
...	...	...	
195	8732	FMMFFFFMMMFFMFFFMMMFMMMFFFMFFMFMMFMMMMFMFFMMFF...	HIGHLOWLOWN
196	8748	FMMFFFFMMMFFMFFFMMMFMMMFFFMFFMFMMFMMMMFMFFMMFF...	HIGHLOWLOWN
197	8800	FMMFFFFMMMFFMFFFMMMFMMMFFFMFFMFMMFMMMMFMFFMMFF...	HIGHLOWLOWN
198	8823	FMMFFFFMMMFFMFFFMMMFMMMFFFMFFMFMMFMMMMFMFFMMFF...	HIGHLOWLOWN
199	8863	FMMFFFFMMMFFMFFFMMMFMMMFFFMFFMFMMFMMMMFMFFMMFF...	HIGHLOWLOWN

Minimum Values

In [10]:
```
df.min()
```

Out[10]:
```
Age              15
Sex               F
BP             HIGH
Cholesterol    HIGH
Na_to_K       6.269
Drug          drugA
dtype: object
```

Maximum Values

In [11]:
```
df.max()
```

Out[11]:
```
Age              74
Sex               M
BP           NORMAL
Cholesterol  NORMAL
Na_to_K      38.247
Drug          drugY
dtype: object
```

Correlation

In [13]:
```
from scipy.stats import spearmanr
print(spearmanr(df['Age'],df['Na_to_K']))
```

```
SpearmanrResult(correlation=-0.047273882688479915, pvalue=0.5062200581387418)
```

Co variance

In [14]:
```
from scipy.stats import pearsonr
print(pearsonr(df['Na_to_K'],df['Age']))
```

```
(-0.06311949726772592, 0.3745756399034559)
```

In [ ]:
```

```