

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

```
In [2]: a=pd.read_csv(r"C:\Users\user\Downloads\4_drug200 - 4_drug200.csv")
a
```

Out[2]:

	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

```
In [3]: a.mean()
```

Out[3]: Age 44.315000  
Na\_to\_K 16.084485  
dtype: float64

```
In [4]: a.median()
```

Out[4]: Age 45.0000  
Na\_to\_K 13.9365  
dtype: float64

```
In [5]: a.mode()
```

Out[5]:

	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

```
In [6]: a.describe()
```

Out[6]:

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

```
In [7]: a.sum()
```

```
Out[7]: Age                                     8863
Sex      FMMFFFFMMMFMMFFMMMFMMMFMMFFMFMMFMMFMMMFMMFF...
BP      HIGHLOWLOWNORMALLOWNORMALNORMALLOWNORMALLOWLOW...
Cholesterol  HIGHHHIGHHHIGHHHIGHHHIGHHHIGHHHIGHHHIGHH...
Na_to_K                                     3216.897
Drug      drugYdrugCdrugCdrugXdrugYdrugXdrugYdrugCdrugYd...
dtype: object
```

```
In [8]: a.cumsum()
```

Out[8]:

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

200 rows × 6 columns

```
In [9]: a.count()
```

```
Out[9]: Age          200  
Sex          200  
BP           200  
Cholesterol  200  
Na_to_K      200  
Drug         200  
dtype: int64
```

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

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

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

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

```
In [16]: from numpy import cov
```

```
In [19]: d1=a['Age']  
d2=a['Na_to_K']  
d1  
d2
```

```
Out[19]: 0          25.355  
1          13.093  
2          10.114  
3           7.798  
4          18.043  
...  
195        11.567  
196        12.006  
197         9.894  
198        14.020  
199        11.349  
Name: Na_to_K, Length: 200, dtype: float64
```

```
In [20]: cov(d1,d2)
```

```
Out[20]: array([[273.71434673, -7.54375153],  
                [-7.54375153, 52.18553348]])
```

```
In [21]: from scipy.stats import pearsonr  
print(pearsonr(d1,d2))  
  
(-0.06311949726772592, 0.3745756399034559)
```

```
In [22]: from scipy.stats import spearmanr  
print(spearmanr(d1,d2))  
  
SpearmanrResult(correlation=-0.047273882688479915, pvalue=0.5062200581387418)
```

```
In [ ]:
```