```python
import numpy as np
import pandas as pd
from numpy import mean,std
import matplotlib.pyplot as pp
from numpy import cov
from scipy.stats import pearsonr
from scipy.stats import spearmanr
```

```python
a=pd.read_csv(r"C:\Users\user\Downloads\4_drug200.csv")
a
```

	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

a) Find mean, median, mode and describe

```python
a.mean()
```

```
Age        44.315000
Na_to_K    16.084485
dtype: float64
```

```
a.median()
```

```
Age        45.0000
Na_to_K    13.9365
dtype: float64
```

```
a.mode()
```

	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

```
a.describe()
```

	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

b) Find sum(), cumsum(), count, min and max values

In [34]:

```
a.sum()
```

Out[34]:

```
Age                                                      8863
Sex             FMMFFFFMMMFFMFFFMMMFMMMFFFFMFFMFMMFMMMMFMFFMMFF...
BP              HIGHLOWLOWNORMALLOWNORMALNORMALLOWNORMALLOWLOW...
Cholesterol     HIGHHIGHHIGHHIGHHIGHHIGHHIGHHIGHHIGHNORMALHIGH...
Na_to_K                                                3216.897
Drug            drugYdrugCdrugCdrugXdrugYdrugXdrugYdrugCdrugYd...
dtype: object
```

In [35]:

```
a.cumsum()
```

Out[35]:

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

200 rows × 6 columns

In [36]:

```
a.count()
```

Out[36]:

```
Age            200
Sex            200
BP             200
Cholesterol    200
Na_to_K        200
Drug           200
dtype: int64
```

```
a.count()
```

Out[37]:

```
Age            200
Sex            200
BP             200
Cholesterol    200
Na_to_K        200
Drug           200
dtype: int64
```

In [38]:

```
a.max()
```

Out[38]:

```
Age                 74
Sex                  M
BP              NORMAL
Cholesterol     NORMAL
Na_to_K         38.247
Drug             drugY
dtype: object
```

c) Find covariance and correlation (spearman and pearsons)

In [39]:

```
d1=a["Age"]
d2=a["Na_to_K"]
cov(d1,d2)
```

Out[39]:

```
array([[273.71434673,  -7.54375153],
       [ -7.54375153,  52.18553348]])
```

In [40]:

```
pearsonr(d1,d2)
```

Out[40]:

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

In [41]:

```
spearmanr(d1,d2)
```

Out[41]:

```
SpearmanrResult(correlation=-0.047273882688479915, pvalue=0.50622005813874
18)
```

In [ ]: