import libraries

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

import dataset

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

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

mean

In [3]:
```
print(data.mean())
```

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

median

In [4]:
```
print(data.median())
```

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

mode

In [5]:
```
print(data.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
```

In [8]:
```
df = pd.DataFrame(data[["Age","Na_to_K"]])
df
```

Out[8]:

	Age	Na_to_K
0	23	25.355
1	47	13.093
2	47	10.114
3	28	7.798
4	61	18.043
...	...	...
195	56	11.567
196	16	12.006
197	52	9.894
198	23	14.020
199	40	11.349

200 rows × 2 columns

In [9]:
```
print(df.mode())
```

```
    Age  Na_to_K
0  47.0   12.006
1   NaN   18.295
```

In [10]:
```
print(df.mean())
```

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

In [11]:
```
print(df.median())
```

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

In [12]:
```
print(df.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
```

In [13]:
```
print(df.sum())
```

```
Age        8863.000
Na_to_K    3216.897
dtype: float64
```

In [14]:
```
print(df.cumsum())
```

```
      Age   Na_to_K
0      23    25.355
1      70    38.448
2     117    48.562
3     145    56.360
4     206    74.403
..    ...       ...
195  8732  3169.628
196  8748  3181.634
197  8800  3191.528
198  8823  3205.548
199  8863  3216.897

[200 rows x 2 columns]
```

In [15]:
```
print(df.min())
```

```
Age        15.000
Na_to_K     6.269
dtype: float64
```

In [16]:
```
print(df.max())
```

```
Age        74.000
Na_to_K    38.247
dtype: float64
```

In [17]:
```
print(df.count())
```

```
Age        200
Na_to_K    200
dtype: int64
```

In [18]:
```
from numpy import cov
```

In [19]:
```
print(cov(df))
```

```
[[  2.7730125 -39.9254925 -43.433265  ... -49.579815  -10.57395
  -33.7365525]
 [-39.9254925 574.8423245 625.346801  ... 713.844071  152.24243
  485.7347285]
 [-43.433265  625.346801  680.288498  ... 776.560958  165.61814
  528.410393 ]
 ...
 [-49.579815  713.844071  776.560958  ... 886.457618  189.05594
  603.189503 ]
 [-10.57395   152.24243   165.61814   ... 189.05594    40.3202
  128.64299  ]
 [-33.7365525 485.7347285 528.410393  ... 603.189503  128.64299
  410.4399005]]
```

In [20]:
```
from scipy.stats import pearsonr
```

In [22]:
```
df1 = df["Age"]
df2 = df["Na_to_K"]
df1
df2
```

Out[22]:
```
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 [23]:
```
print(pearsonr(df1,df2))
```

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

In [24]:
```
from scipy.stats import spearmanr
```

In [25]:
```
print(spearmanr(df1,df2))
```

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

In [ ]:
```

```