```
import numpy as np
import pandas as pd
from numpy import linalg as lg
from numpy import cov
```

5 Data Sets

```
In [2]: a=pd.read_csv("drug.csv")
a
```

Out[2]:		Age	Sex	ВР	Cholesterol	Na_to_K	Drug
	0	23	F	HIGH	HIGH	25.355	drugY
	1	47	М	LOW	HIGH	13.093	drugC
	2	47	М	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	М	LOW	HIGH	12.006	drugC
	197	52	М	NORMAL	HIGH	9.894	drugX
	198	23	М	NORMAL	NORMAL	14.020	drugX
	199	40	F	LOW	NORMAL	11.349	drugX

200 rows × 6 columns

a) Find mean, median, mode and describe

```
In [3]:
         print(a.mean)
         print("Median:")
         print(a.median)
         print("Mode:")
         print(a.mode)
         print("Describe")
         print(a.describe())
        <bound method NDFrame._add_numeric_operations.<locals>.mean of
                                                                              Age Sex
                                                                                           BP
        Cholesterol Na_to_K
        0
              23
                  F
                         HIGH
                                     HIGH
                                            25.355 drugY
        1
              47
                   Μ
                          LOW
                                            13.093 drugC
                                     HIGH
        2
              47
                          LOW
                                            10.114 drugC
                   Μ
                                     HIGH
                  F
        3
              28
                      NORMAL
                                     HIGH
                                             7.798 drugX
        4
                   F
                                            18.043
              61
                          LOW
                                     HIGH
                                                    drugY
                          . . .
                                      . . .
        195
                          LOW
                                            11.567
                                                    drugC
              56
                   F
                                     HIGH
                                            12.006
        196
              16
                   Μ
                          LOW
                                     HIGH
                                                    drugC
        197
                      NORMAL
                                            9.894 drugX
              52
                   Μ
                                     HIGH
                                            14.020 drugX
        198
              23
                   Μ
                      NORMAL
                                   NORMAL
        199
                          LOW
                                   NORMAL
                                            11.349 drugX
```

```
[200 rows x 6 columns]>
Median:
                                                                  Age Sex
<bound method NDFrame._add_numeric_operations.<locals>.median of
P Cholesterol Na_to_K Drug
              HIGH
                                 25.355 drugY
     23
                          HIGH
                          HIGH 13.093 drugC
1
     47
          Μ
               LOW
2
     47
               LOW
                          HIGH 10.114 drugC
         М
3
        F NORMAL
                                7.798 drugX
     28
                          HIGH
4
          F
                          HIGH
                                18.043 drugY
     61
               LOW
                . . .
                          . . .
195
        F
    56
                LOW
                          HIGH
                                 11.567
                                        drugC
196
                          HIGH
                                12.006 drugC
     16 M
                LOW
197
                          HIGH
                                 9.894 drugX
     52
        M NORMAL
198
                                 14.020 drugX
     23
        M NORMAL
                        NORMAL
199
        F
                        NORMAL
                                 11.349 drugX
     40
               LOW
[200 rows x 6 columns]>
Mode:
                                               BP Cholesterol Na_to_K
<bound method DataFrame.mode of</pre>
                                  Age Sex
     23
             HIGH HIGH
                                 25.355 drugY
         F
                          HIGH
1
     47
               LOW
                                13.093 drugC
          Μ
2
     47
         М
                LOW
                         HIGH
                                10.114 drugC
3
        F NORMAL
                                7.798 drugX
     28
                          HIGH
                                 18.043 drugY
4
     61 F
               LOW
                          HIGH
                . . .
                          . . .
195
    56 F
                                 11.567 drugC
                LOW
                          HIGH
                                 12.006 drugC
196
               LOW
                          HIGH
    16 M
                                 9.894 drugX
197
        M NORMAL
                          HIGH
     52
                                 14.020 drugX
198
        M NORMAL
                        NORMAL
     23
                                 11.349 drugX
199
        F
     40
               LOW
                        NORMAL
[200 \text{ rows x 6 columns}]
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
       74.000000
                  38.247000
```

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

```
In [4]:
         print(a.sum())
        Age
                                                                    8863
                       FMMFFFFMMMFFMFFFMMMFMMFFFMFFMMFMMMMFMFFMMFF...
        Sex
                       HIGHLOWLOWNORMALLOWNORMALLOWNORMALLOWLOW...
        Cholesterol
                       HIGHHIGHHIGHHIGHHIGHHIGHHIGHNORMALHIGH...
        Na_to_K
                                                                3216.897
                       drugYdrugCdrugXdrugYdrugXdrugYdrugCdrugYd...
        Drug
        dtype: object
In [5]:
         print(a.cumsum())
                                                                 Sex
              Age
               23
                                                                   F
               70
                                                                  FΜ
        1
        2
                                                                 FMM
              117
        3
              145
                                                                FMMF
              206
                                                               FMMFF
```

```
195
                  FMMFFFFMMMFMFFFMMMFMMMFFFMFMMMMMMMFMFFMMFF...
            8732
        196
            8748
                  EMMEEEEMMMEMMEEMEEMMMEMMMEMEMMEMMMMEMEEMMEE...
        197
            8800
                  EMMEEEEMMMEMMEEMEEMMMEMMMEMEMMEMMMMEMEEMMEE...
        198
                  FMMFFFFMMMFMFFFMMMFMMMFFFMFMMMMMMMFMFFMMFF...
            8823
                  FMMFFFFMMMFFMFFFMMMFMMMFFFMFMMMMMMFMFFMMFF...
        199
            8863
                                                          ΒP
        0
                                                        HIGH
        1
                                                     HTGHI OW
        2
                                                  HIGHLOWLOW
        3
                                             HIGHLOWLOWNORMAL
        Δ
                                          HIGHLOWLOWNORMALLOW
            HIGHLOWLOWNORMALLOWNORMALLOWNORMALLOWLOW...
        195
            HIGHLOWLOWNORMALLOWNORMALLOWNORMALLOWLOW...
        196
            HIGHLOWLOWNORMALLOWNORMALLOWNORMALLOWLOW...
        197
            HIGHLOWLOWNORMALLOWNORMALLOWNORMALLOWLOW...
        198
            HIGHLOWLOWNORMALLOWNORMALLOWNORMALLOWLOW...
        199
                                                  Cholesterol
                                                               Na to K
                                                        HIGH
                                                                25.355
        0
                                                                38.448
        1
                                                    HIGHHIGH
                                                                48.562
        2
                                                 HIGHHIGHHIGH
        3
                                                                56.360
                                             HIGHHIGHHIGH
                                                                74.403
        4
                                         HIGHHIGHHIGHHIGH
        195
                                                              3169.628
            HIGHHIGHHIGHHIGHHIGHHIGHHIGHNORMALHIGH...
            HIGHHIGHHIGHHIGHHIGHHIGHHIGHNORMALHIGH...
        196
                                                              3181.634
            HIGHHIGHHIGHHIGHHIGHHIGHHIGHNORMALHIGH...
        197
                                                              3191.528
            HIGHHIGHHIGHHIGHHIGHHIGHHIGHHIGHNORMALHIGH...
                                                             3205.548
        198
            HIGHHIGHHIGHHIGHHIGHHIGHHIGHNORMALHIGH...
        199
                                                              3216.897
                                                        Drug
        0
                                                       drugY
        1
                                                  drugYdrugC
        2
                                              drugYdrugCdrugC
        3
                                         drugYdrugCdrugCdrugX
        4
                                    drugYdrugCdrugXdrugY
        195
            drugYdrugCdrugXdrugYdrugXdrugYdrugCdrugYd...
        196
            drugYdrugCdrugXdrugYdrugXdrugYdrugCdrugYd...
        197
            drugYdrugCdrugXdrugYdrugXdrugYdrugCdrugYd...
            drugYdrugCdrugXdrugYdrugXdrugYdrugCdrugYd...
        199
            drugYdrugCdrugXdrugYdrugXdrugYdrugCdrugYd...
        [200 rows x 6 columns]
In [6]:
        print(a.count())
        print(a.min())
        print(a.max())
        Age
                      200
        Sex
                      200
        RP
                      200
        Cholesterol
                      200
        Na_to_K
                      200
                      200
        Drug
        dtype: int64
                         15
        Age
        Sex
                          F
        BP
                       HIGH
        Cholesterol
                       HIGH
                      6.269
        Na_to_K
                      drugA
        Drug
        dtype: object
                          74
        Age
        Sex
```

BP NORMAL
Cholesterol NORMAL
Na_to_K 38.247
Drug drugY
dtype: object

c) Find covariance and correlation (spearman and pearsons)

```
In [7]:
          print(a.cov())
                                  Na_to_K
                          Age
                   273.714347
                                -7.543752
          Age
                               52.185533
          Na_to_K
                    -7.543752
 In [8]:
          b=a.dropna(axis=1,how='any')
 Out[8]:
              Age Sex
                             BP Cholesterol Na_to_K
            0
                23
                      F
                           HIGH
                                       HIGH
                                              25.355 drugY
            1
                47
                     Μ
                            LOW
                                       HIGH
                                              13.093 drugC
            2
                            LOW
                                              10.114 drugC
                47
                                      HIGH
                     M
            3
                28
                       NORMAL
                                      HIGH
                                               7.798 drugX
                                              18.043 drugY
                61
                      F
                            LOW
            4
                                      HIGH
                                              11.567 drugC
          195
                     F
                            LOW
                                      HIGH
                56
                            LOW
          196
                16
                                       HIGH
                                              12.006 drugC
          197
                52
                     M NORMAL
                                       HIGH
                                               9.894 drugX
          198
                23
                     Μ
                       NORMAL
                                   NORMAL
                                              14.020 drugX
                     F
          199
                40
                            LOW
                                   NORMAL
                                              11.349 drugX
         200 rows × 6 columns
In [11]:
          from scipy.stats import pearsonr
          print(pearsonr(a['Age'],a['Na_to_K']))
          (-0.06311949726772592, 0.3745756399034559)
In [13]:
          from scipy.stats import spearmanr
          print(spearmanr(a['Age'],a['Na_to_K']))
          SpearmanrResult(correlation=-0.047273882688479915, pvalue=0.5062200581387418)
 In [ ]:
```