

```
In [2]: import numpy as np
import pandas as pd
```

```
In [3]: data=pd.read_csv(r"c:\Users\user\Downloads\drug.csv")
data
```

Out[3]:

	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

Find mean, median, mode and describe

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

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

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

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

```
In [27]: print(data.mode())
```

	Country	Region	Happiness	Rank	Happiness	Score \
0	Afghanistan	Sub-Saharan Africa		82.0		5.192
1	Albania	NaN		NaN		NaN
2	Algeria	NaN		NaN		NaN
3	Angola	NaN		NaN		NaN
4	Argentina	NaN		NaN		NaN
..
153	Venezuela	NaN		NaN		NaN
154	Vietnam	NaN		NaN		NaN
155	Yemen	NaN		NaN		NaN
156	Zambia	NaN		NaN		NaN
157	Zimbabwe	NaN		NaN		NaN

	Standard Error	Economy (GDP per Capita)	Family \
0	0.03751	0.00000	0.00000
1	0.03780	0.01530	0.13995
2	0.04394	0.01604	0.30285
3	0.04934	0.06940	0.35386
4	0.05051	0.07120	0.38174
..
153	NaN	1.45900	1.34043
154	NaN	1.52186	1.34951
155	NaN	1.55422	1.36058
156	NaN	1.56391	1.36948
157	NaN	1.69042	1.40223

	Health (Life Expectancy)	Freedom	Trust (Government Corruption) \
0	0.92356	0.00000	0.32524
1	NaN	0.07699	NaN
2	NaN	0.09245	NaN
3	NaN	0.10081	NaN
4	NaN	0.10384	NaN
..
153	NaN	0.65821	NaN
154	NaN	0.65980	NaN
155	NaN	0.66246	NaN
156	NaN	0.66557	NaN
157	NaN	0.66973	NaN

	Generosity	Dystopia	Residual
0	0.00000		0.32858
1	0.00199		0.65429
2	0.02641		0.67042
3	0.05444		0.67108
4	0.05547		0.89991
..
153	0.51535		3.10712
154	0.51752		3.17728
155	0.51912		3.19131
156	0.57630		3.26001
157	0.79588		3.60214

[158 rows x 12 columns]

```
In [6]: 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

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

```
In [7]: print(data.sum())
```

Age	8863
Sex	FMMFFFFMMMFMMFFMMMFMMMFMMFFMMFFMMFMMFFMMFFMMFF...
BP	HIGHLOWLOWNORMALLOWNORMALNORMALLOWNORMALLOWLOW...
Cholesterol	HIGHHIGHHIGHHIGHHIGHHIGHHIGHHIGHHIGHHIGHNORMALHIGH...
Na_to_K	3216.897
Drug	drugYdrugCdrugCdrugXdrugYdrugXdrugYdrugCdrugYd...
dtype:	object

```
In [8]: print(data.count())
```

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

```
In [9]: print(data.max())
```

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

```
In [10]: print(data.min())
```

Age	15
Sex	F
BP	HIGH
Cholesterol	HIGH
Na_to_K	6.269
Drug	drugA
dtype:	object

```
In [11]: print(data.cumsum())
```

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

      BP \
0      HIGH
1     HIGHLOW
2    HIGHLOWLOW
3   HIGHLOWLOWNORMAL
4  HIGHLOWLOWNORMALLOW
```

Find covariance and correlation (spearman and pearsons)

```
In [12]: from numpy import cov
from numpy import mean,std
from numpy.random import randn,seed
from matplotlib import pyplot
```

```
In [14]: print(mean(data.Age),std(data.Age))
print(mean(data.Na_to_K),std(data.Na_to_K))
```

```
44.315 16.502902017524065
16.084484999999999 7.205873008163204
```

```
In [15]: print(cov(data.Age,data.Na_to_K))
```

```
[[273.71434673 -7.54375153]
 [ -7.54375153  52.18553348]]
```

```
In [16]: from scipy.stats import pearsonr
print(pearsonr(data.Age,data.Na_to_K))
```

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

```
In [17]: from scipy.stats import spearmanr
print(spearmanr(data.Age,data.Na_to_K))
```

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

```
In [ ]:
```

