

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

```
In [2]: data=pd.read_csv(r"c:\Users\user\Downloads\bottle1.csv")  
data
```

Out[2]:

	Cst_Cnt	Btl_Cnt	Sta_ID	Depth_ID	Depthm	T_degC	Salnty	O2ml_L	STheta	O2Sat	...
0	1	1	054.0 056.0	19- 4903CR- HY-060- 0930- 05400560- 0000A-3	0	10.50	33.440	NaN	25.649	NaN	...
1	1	2	054.0 056.0	19- 4903CR- HY-060- 0930- 05400560- 0008A-3	8	10.46	33.440	NaN	25.656	NaN	...
2	1	3	054.0 056.0	19- 4903CR- HY-060- 0930- 05400560- 0010A-7	10	10.46	33.437	NaN	25.654	NaN	...
3	1	4	054.0 056.0	19- 4903CR- HY-060- 0930- 05400560- 0019A-3	19	10.45	33.420	NaN	25.643	NaN	...
4	1	5	054.0 056.0	19- 4903CR- HY-060- 0930- 05400560- 0020A-7	20	10.45	33.421	NaN	25.643	NaN	...
...
5240	173	5241	069.0 144.0	19- 4905CR- HY-125- 1712- 06901440- 0279A-3	279	7.86	33.870	3.81	26.411	57.0	...
5241	173	5242	069.0 144.0	19- 4905CR- HY-125- 1712- 06901440- 0300A-7	300	7.52	33.896	3.55	26.481	52.7	...
5242	173	5243	069.0 144.0	19- 4905CR- HY-125- 1712- 06901440- 0371A-3	371	6.54	33.930	2.67	26.642	38.8	...
5243	173	5244	069.0 144.0	19- 4905CR- HY-125- 1712- 06901440- 0400A-7	400	6.25	33.951	2.32	26.697	33.5	...

	Cst_Cnt	Btl_Cnt	Sta_ID	Depth_ID	Depthm	T_degC	Salnty	O2ml_L	STheta	O2Sat	...
				19-4905CR-HY-125-1712-06901440-0500A-7							
5244	173	5245	069.0144.0		500	5.44	34.027	1.30	26.858	18.4	...

5245 rows × 71 columns

Find mean, median, mode and describe

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

Cst_Cnt 86.191992
Btl_Cnt 2623.000000
Depthm 345.400572
T_degC 8.993144
Salnty 33.827093
...
R_SAMP NaN
DIC1 NaN
DIC2 NaN
TA1 NaN
TA2 NaN
Length: 69, dtype: float64

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

```
id          906024.000000
radius_mean    13.370000
texture_mean   18.840000
perimeter_mean 86.240000
area_mean     551.100000
smoothness_mean 0.095870
compactness_mean 0.092630
concavity_mean 0.061540
concave points_mean 0.033500
symmetry_mean  0.179200
fractal_dimension_mean 0.061540
radius_se      0.324200
texture_se     1.108000
perimeter_se   2.287000
area_se       24.530000
smoothness_se  0.006380
compactness_se 0.020450
concavity_se   0.025890
concave points_se 0.010930
symmetry_se    0.018730
fractal_dimension_se 0.003187
radius_worst   14.970000
texture_worst  25.410000
perimeter_worst 97.660000
area_worst    686.500000
smoothness_worst 0.131300
compactness_worst 0.211900
concavity_worst 0.226700
concave points_worst 0.099930
symmetry_worst  0.282200
fractal_dimension_worst 0.080040
Unnamed: 32      NaN
dtype: float64
```

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

	Cst_Cnt	Btl_Cnt	Sta_ID	Depth_ID
\				
0	118.0	1	081.0 117.0	19-4903CR-HY-060-0930-05400560-0000A-3
1	NaN	2	NaN	19-4903CR-HY-060-0930-05400560-0008A-3
2	NaN	3	NaN	19-4903CR-HY-060-0930-05400560-0010A-7
3	NaN	4	NaN	19-4903CR-HY-060-0930-05400560-0019A-3
4	NaN	5	NaN	19-4903CR-HY-060-0930-05400560-0020A-7
...
5240	NaN	5241	NaN	19-4905CR-HY-125-1712-06901440-0279A-3
5241	NaN	5242	NaN	19-4905CR-HY-125-1712-06901440-0300A-7
5242	NaN	5243	NaN	19-4905CR-HY-125-1712-06901440-0371A-3
5243	NaN	5244	NaN	19-4905CR-HY-125-1712-06901440-0400A-7
5244	NaN	5245	NaN	19-4905CR-HY-125-1712-06901440-0500A-7

	Depthm	T_degC	Salnty	O2ml_L	STheta	O2Sat	...	R_NO2	R_NH4	\
0	0.0	14.59	33.62	0.38	24.990	5.2	...	NaN	NaN	
1	10.0	NaN	NaN	NaN	25.005	NaN	...	NaN	NaN	
2	20.0	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	
3	30.0	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	
4	50.0	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	
...	
5240	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	
5241	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	
5242	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	
5243	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	
5244	NaN	NaN	NaN	NaN	NaN	NaN	...	NaN	NaN	

	R_CHLA	R_PHAEO	R_PREs	R_SAMP	DIC1	DIC2	TA1	TA2
0	NaN	NaN	0.0	NaN	NaN	NaN	NaN	NaN
1	NaN	NaN	10.0	NaN	NaN	NaN	NaN	NaN
2	NaN	NaN	20.0	NaN	NaN	NaN	NaN	NaN
3	NaN	NaN	30.0	NaN	NaN	NaN	NaN	NaN
4	NaN	NaN	50.0	NaN	NaN	NaN	NaN	NaN
...
5240	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
5241	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
5242	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
5243	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
5244	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

[5245 rows x 71 columns]

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

	id	radius_mean	texture_mean	perimeter_mean	area_mean
\					
count	5.690000e+02	569.000000	569.000000	569.000000	569.000000
mean	3.037183e+07	14.127292	19.289649	91.969033	654.889104
std	1.250206e+08	3.524049	4.301036	24.298981	351.914129
min	8.670000e+03	6.981000	9.710000	43.790000	143.500000
25%	8.692180e+05	11.700000	16.170000	75.170000	420.300000
50%	9.060240e+05	13.370000	18.840000	86.240000	551.100000
75%	8.813129e+06	15.780000	21.800000	104.100000	782.700000
max	9.113205e+08	28.110000	39.280000	188.500000	2501.000000

	smoothness_mean	compactness_mean	concavity_mean	concave points_mean
\				
count	569.000000	569.000000	569.000000	569.000000
mean	0.096360	0.104341	0.088799	0.048919
std	0.014064	0.052813	0.079720	0.038803
min	0.052630	0.019380	0.000000	0.000000
25%	0.086370	0.064920	0.029560	0.020310
50%	0.095870	0.092630	0.061540	0.033500
75%	0.105300	0.130400	0.130700	0.074000
max	0.163400	0.345400	0.426800	0.201200

	symmetry_mean	...	texture_worst	perimeter_worst	area_worst	\
count	569.000000	...	569.000000	569.000000	569.000000	
mean	0.181162	...	25.677223	107.261213	880.583128	
std	0.027414	...	6.146258	33.602542	569.356993	
min	0.106000	...	12.020000	50.410000	185.200000	
25%	0.161900	...	21.080000	84.110000	515.300000	
50%	0.179200	...	25.410000	97.660000	686.500000	
75%	0.195700	...	29.720000	125.400000	1084.000000	
max	0.304000	...	49.540000	251.200000	4254.000000	

	smoothness_worst	compactness_worst	concavity_worst	\
count	569.000000	569.000000	569.000000	
mean	0.132369	0.254265	0.272188	
std	0.022832	0.157336	0.208624	
min	0.071170	0.027290	0.000000	
25%	0.116600	0.147200	0.114500	
50%	0.131300	0.211900	0.226700	
75%	0.146000	0.339100	0.382900	
max	0.222600	1.058000	1.252000	

	concave points_worst	symmetry_worst	fractal_dimension_worst	\
count	569.000000	569.000000	569.000000	
mean	0.114606	0.290076	0.083946	
std	0.065732	0.061867	0.018061	
min	0.000000	0.156500	0.055040	
25%	0.064930	0.250400	0.071460	
50%	0.099930	0.282200	0.080040	
75%	0.161400	0.317900	0.092080	
max	0.291000	0.663800	0.207500	

Unnamed: 32

count	0.0
mean	NaN
std	NaN
min	NaN

25%	NaN
50%	NaN
75%	NaN
max	NaN

[8 rows x 32 columns]

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

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

```
Cst_Cnt          452077
Btl_Cnt          13757635
Sta_ID    054.0 056.0054.0 056.0054.0 056.0054.0 056.005...
Depth_ID    19-4903CR-HY-060-0930-05400560-0000A-319-4903C...
Depthm          1811626
...
R_SAMP          0.0
DIC1            0.0
DIC2            0.0
TA1             0.0
TA2             0.0
Length: 71, dtype: object
```

In [6]: `print(data.count())`

```
Cst_Cnt      5245
Btl_Cnt      5245
Sta_ID       5245
Depth_ID     5245
Depthm       5245
...
R_SAMP        0
DIC1          0
DIC2          0
TA1           0
TA2           0
Length: 71, dtype: int64
```

In [14]: `print(data.max())`

```
Cst_Cnt          173
Btl_Cnt          5245
Sta_ID          126.0 132.0
Depth_ID    19-4905CR-HY-125-1712-06901440-0500A-7
Depthm          1547
...
R_SAMP          NaN
DIC1            NaN
DIC2            NaN
TA1             NaN
TA2             NaN
Length: 71, dtype: object
```

```
In [12]: print(data.min())
```

```
Cst_Cnt          1
Btl_Cnt          1
Sta_ID          024.0 120.0
Depth_ID    19-4903CR-HY-060-0930-05400560-0000A-3
Depthm          0

...

R_SAMP          NaN
DIC1            NaN
DIC2            NaN
TA1             NaN
TA2             NaN
Length: 71, dtype: object
```

```
In [15]: print(data.cumsum())
```

```
      Cst_Cnt  Btl_Cnt          Sta_ID
\
0          1          1          054.0 056.0
1          2          3          054.0 056.0054.0 056.0
2          3          6          054.0 056.0054.0 056.0054.0 056.0
3          4         10          054.0 056.0054.0 056.0054.0 056.0054.0 056.0
4          5         15          054.0 056.0054.0 056.0054.0 056.0054.0 056.005...
...
5240    451385    13736661          054.0 056.0054.0 056.0054.0 056.0054.0 056.005...
5241    451558    13741903          054.0 056.0054.0 056.0054.0 056.0054.0 056.005...
5242    451731    13747146          054.0 056.0054.0 056.0054.0 056.0054.0 056.005...
5243    451904    13752390          054.0 056.0054.0 056.0054.0 056.0054.0 056.005...
5244    452077    13757635          054.0 056.0054.0 056.0054.0 056.0054.0 056.005...

      Depth_ID  Depthm  T_degC
\
0          19-4903CR-HY-060-0930-05400560-0000A-3          0          10.50
1    19-4903CR-HY-060-0930-05400560-0000A-319-4903C...          8          20.96
2    19-4903CR-HY-060-0930-05400560-0000A-319-4903C...         18          31.42
3    19-4903CR-HY-060-0930-05400560-0000A-319-4903C...         27          41.07
```

Find covariance and correlation (spearman and pearsons)

```
In [11]: from numpy import cov
from numpy import mean,std
from numpy.random import randn,seed
from matplotlib import pyplot
```

```
In [18]: print(mean(data.Depthm),std(data.Depthm))
print(mean(data.Btl_Cnt),std(data.Btl_Cnt))
```

```
345.40057197330793 356.2641754471208
2623.0 1514.1010534307147
```

```
In [21]: print(cov(data.Depthm,data.Btl_Cnt))
```

```
[[1.26948366e+05 9.51372998e+01]
 [9.51372998e+01 2.29293917e+06]]
```

```
In [20]: from scipy.stats import pearsonr
print(pearsonr(data.Depthm,data.Btl_Cnt))

(0.00017633600454111144, 0.9898131835240318)
```

```
In [22]: from scipy.stats import spearmanr
print(spearmanr(data.Depthm,data.Btl_Cnt))

SpearmanrResult(correlation=0.0020888224774152455, pvalue=0.8797851457218078)
```

```
In [ ]:
```