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

In [3]:
```
data=pd.read_csv(r"c:\Users\user\Downloads\8_dataset.csv")
data
```

Out[3]:

	id	diagnosis	radius_mean	texture_mean	perimeter_mean	area_mean	smoothness_
0	842302	M	17.99	10.38	122.80	1001.0	0.
1	842517	M	20.57	17.77	132.90	1326.0	0.(
2	84300903	M	19.69	21.25	130.00	1203.0	0.1
3	84348301	M	11.42	20.38	77.58	386.1	0.1
4	84358402	M	20.29	14.34	135.10	1297.0	0.1
...	...	...	...	...	...	...	
564	926424	M	21.56	22.39	142.00	1479.0	0.
565	926682	M	20.13	28.25	131.20	1261.0	0.(
566	926954	M	16.60	28.08	108.30	858.1	0.(
567	927241	M	20.60	29.33	140.10	1265.0	0.
568	92751	B	7.76	24.54	47.92	181.0	0.(

569 rows × 33 columns

Find mean, median, mode and describe

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

```
id                         3.037183e+07
radius_mean                1.412729e+01
texture_mean               1.928965e+01
perimeter_mean             9.196903e+01
area_mean                  6.548891e+02
smoothness_mean            9.636028e-02
compactness_mean           1.043410e-01
concavity_mean             8.879932e-02
concave points_mean        4.891915e-02
symmetry_mean              1.811619e-01
fractal_dimension_mean     6.279761e-02
radius_se                  4.051721e-01
texture_se                 1.216853e+00
perimeter_se               2.866059e+00
area_se                    4.033708e+01
smoothness_se              7.040979e-03
compactness_se             2.547814e-02
concavity_se               3.189372e-02
concave points_se          1.179614e-02
symmetry_se                2.054230e-02
fractal_dimension_se       3.794904e-03
radius_worst               1.626919e+01
texture_worst              2.567722e+01
perimeter_worst            1.072612e+02
area_worst                 8.805831e+02
smoothness_worst           1.323686e-01
compactness_worst          2.542650e-01
concavity_worst            2.721885e-01
concave points_worst       1.146062e-01
symmetry_worst             2.900756e-01
fractal_dimension_worst    8.394582e-02
Unnamed: 32                         NaN
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 [6]:
```
print(data.mode())
```

```
           id diagnosis  radius_mean  texture_mean  perimeter_mean  \
0        8670         B        12.34         14.93           82.61
1        8913       NaN          NaN         15.70           87.76
2        8915       NaN          NaN         16.84          134.70
3        9047       NaN          NaN         16.85             NaN
4       85715       NaN          NaN         17.46             NaN
..        ...       ...          ...           ...             ...
564  911157302       NaN          NaN           NaN             NaN
565  911296201       NaN          NaN           NaN             NaN
566  911296202       NaN          NaN           NaN             NaN
567  911320501       NaN          NaN           NaN             NaN
568  911320502       NaN          NaN           NaN             NaN

     area_mean  smoothness_mean  compactness_mean  concavity_mean  \
0        512.2           0.1007            0.1147             0.0
1          NaN              NaN            0.1206             NaN
2          NaN              NaN               NaN             NaN
3          NaN              NaN               NaN             NaN
4          NaN              NaN               NaN             NaN
..         ...              ...               ...             ...
564        NaN              NaN               NaN             NaN
565        NaN              NaN               NaN             NaN
566        NaN              NaN               NaN             NaN
567        NaN              NaN               NaN             NaN
568        NaN              NaN               NaN             NaN

     concave points_mean  ...  texture_worst  perimeter_worst  area_worst  \
0                    0.0  ...          17.70            101.7       284.4
1                    NaN  ...          27.26            105.9       402.8
2                    NaN  ...            NaN            117.7       439.6
3                    NaN  ...            NaN              NaN       458.0
4                    NaN  ...            NaN              NaN       472.4
..                   ...  ...            ...              ...         ...
564                  NaN  ...            NaN              NaN         NaN
565                  NaN  ...            NaN              NaN         NaN
566                  NaN  ...            NaN              NaN         NaN
567                  NaN  ...            NaN              NaN         NaN
568                  NaN  ...            NaN              NaN         NaN

     smoothness_worst  compactness_worst  concavity_worst  \
0              0.1216             0.1486              0.0
1              0.1223             0.3416              NaN
2              0.1234                NaN              NaN
3              0.1256                NaN              NaN
4              0.1275                NaN              NaN
..                ...                ...              ...
564               NaN                NaN              NaN
565               NaN                NaN              NaN
566               NaN                NaN              NaN
567               NaN                NaN              NaN
568               NaN                NaN              NaN

     concave points_worst  symmetry_worst  fractal_dimension_worst  \
0                     0.0          0.2226                  0.07427
1                     NaN          0.2369                      NaN
2                     NaN          0.2383                      NaN
3                     NaN          0.2972                      NaN
```

```
4                    NaN          0.3109                    NaN
..                   ...             ...                    ...
564                  NaN             NaN                    NaN
565                  NaN             NaN                    NaN
566                  NaN             NaN                    NaN
567                  NaN             NaN                    NaN
568                  NaN             NaN                    NaN

     Unnamed: 32
0            NaN
1            NaN
2            NaN
3            NaN
4            NaN
..           ...
564          NaN
565          NaN
566          NaN
567          NaN
568          NaN

[569 rows x 33 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 [8]:
```
print(data.sum())
```

```
id                                                         17281572085
diagnosis                MMMMMMMMMMMMMMMMMMBBBMMMMMMMMMMMMMMMBMMMMMMMM...
radius_mean                                                 8038.429
texture_mean                                                10975.81
perimeter_mean                                              52330.38
area_mean                                                   372631.9
smoothness_mean                                               54.829
compactness_mean                                            59.37002
concavity_mean                                              50.526811
concave points_mean                                         27.834994
symmetry_mean                                               103.0811
fractal_dimension_mean                                      35.73184
radius_se                                                   230.5429
texture_se                                                  692.3896
perimeter_se                                                1630.7877
area_se                                                     22951.798
smoothness_se                                               4.006317
compactness_se                                              14.497061
concavity_se                                                18.147525
concave points_se                                           6.712002
symmetry_se                                                 11.688568
fractal_dimension_se                                          2.1593
radius_worst                                                9257.169
texture_worst                                               14610.34
perimeter_worst                                             61031.63
area_worst                                                  501051.8
smoothness_worst                                            75.31773
compactness_worst                                           144.67681
concavity_worst                                             154.875247
concave points_worst                                        65.210941
symmetry_worst                                                165.053
fractal_dimension_worst                                      47.76517
Unnamed: 32                                                       0.0
dtype: object
```

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

```
id                         569
diagnosis                  569
radius_mean                569
texture_mean               569
perimeter_mean             569
area_mean                  569
smoothness_mean            569
compactness_mean           569
concavity_mean             569
concave points_mean        569
symmetry_mean              569
fractal_dimension_mean     569
radius_se                  569
texture_se                 569
perimeter_se               569
area_se                    569
smoothness_se              569
compactness_se             569
concavity_se               569
concave points_se          569
symmetry_se                569
fractal_dimension_se       569
radius_worst               569
texture_worst              569
perimeter_worst            569
area_worst                 569
smoothness_worst           569
compactness_worst          569
concavity_worst            569
concave points_worst       569
symmetry_worst             569
fractal_dimension_worst    569
Unnamed: 32                  0
dtype: int64
```

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

```
id                         911320502
diagnosis                          M
radius_mean                    28.11
texture_mean                   39.28
perimeter_mean                 188.5
area_mean                     2501.0
smoothness_mean               0.1634
compactness_mean              0.3454
concavity_mean                0.4268
concave points_mean           0.2012
symmetry_mean                  0.304
fractal_dimension_mean       0.09744
radius_se                      2.873
texture_se                     4.885
perimeter_se                   21.98
area_se                        542.2
smoothness_se                0.03113
compactness_se                0.1354
concavity_se                   0.396
concave points_se            0.05279
symmetry_se                  0.07895
fractal_dimension_se         0.02984
radius_worst                   36.04
texture_worst                  49.54
perimeter_worst                251.2
area_worst                    4254.0
smoothness_worst              0.2226
compactness_worst              1.058
concavity_worst                1.252
concave points_worst           0.291
symmetry_worst                0.6638
fractal_dimension_worst       0.2075
Unnamed: 32                      NaN
dtype: object
```

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

```
id                            8670
diagnosis                        B
radius_mean                  6.981
texture_mean                  9.71
perimeter_mean               43.79
area_mean                    143.5
smoothness_mean            0.05263
compactness_mean           0.01938
concavity_mean                 0.0
concave points_mean            0.0
symmetry_mean                0.106
fractal_dimension_mean     0.04996
radius_se                   0.1115
texture_se                  0.3602
perimeter_se                 0.757
area_se                      6.802
smoothness_se              0.001713
compactness_se             0.002252
concavity_se                   0.0
concave points_se              0.0
symmetry_se                0.007882
fractal_dimension_se       0.000895
radius_worst                   7.93
texture_worst                 12.02
perimeter_worst               50.41
area_worst                    185.2
smoothness_worst           0.07117
compactness_worst          0.02729
concavity_worst                0.0
concave points_worst           0.0
symmetry_worst              0.1565
fractal_dimension_worst    0.05504
Unnamed: 32                    NaN
dtype: object
```

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

```
st  \
0              0.66560         0.711900              0.265400          0.46
01
1              0.85220         0.953500              0.451400          0.73
51
2              1.27670         1.403900              0.694400          1.09
64
3              2.14300         2.090800              0.951900          1.76
02
4              2.34800         2.490800              1.114400          1.99
66
..                 ...              ...                   ...
...
564          143.24267       153.274747             64.641341        163.87
82
565          143.43487       153.596247             64.804141        164.13
54
566          143.74427       153.936547             64.945941        164.35
72
```

Find covariance and correlation (spearman and pearsons)

In [13]:
```
from numpy import cov
from numpy import mean,std
from numpy.random import randn,seed
from matplotlib import pyplot
```

In [14]:
```
print(mean(data.perimeter_mean),std(data.perimeter_mean))
print(mean(data.area_mean),std(data.area_mean))
```

```
91.96903339191566 24.277619293053174
654.8891036906857 351.6047540632298
```

In [15]:
```
print(cov(data.perimeter_mean,data.area_mean))
```

```
[[   590.44047952   8435.77234508]
 [  8435.77234508 123843.55431768]]
```

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

```
(0.9865068039913903, 0.0)
```

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

```
SpearmanrResult(correlation=0.9970682695182411, pvalue=0.0)
```

In [ ]: