roozara_hw1

January 16, 2019

- **ECE 657A: Data and Knowledge Modelling and Analysis**
- **Winter 2019**
- **WATIAM:roozara ID: 20801583**
- **Homework 2:Data Summarization.....**

References used:pandas: powerful Python data analysis toolkit

Pandas library was imported to convert the csv dataset to data frame format for manipulation. Also similarly, numpy was imported so that the one dimensional arrays for features can be used to represnt mean,mode,standard deviation,skew and variance

Finally to plot the histogram , matplotlib.pyplot was imported

```
In [5]: #Import pandas package as pd
        import pandas as pd
        import numpy as np
        import matplotlib.pyplot as plt


        file_url = 'https://archive.ics.uci.edu/ml/machine-learning-databases/breast-cancer-wis
        data = pd.read_csv(file_url,sep= ',',header = None)
        #data = data.iloc[:,:32]
        #create the features columns
        features = ['ID number', 'Diagnosis', 'mean radius','mean texture','mean perimeter','me
                    'mean concavity','mean concave points','mean symmetry','mean fractal dimens
                    'SE area','SE smoothness','SE compactness','SE concavity', 'SE concave poi
                    'Worst radius', 'Worst texture', 'Worst perimeter', 'Worst area','Worst sm
                    'Worst concave points','Worst symmetry','Worst fractal dimension']

        data.columns = features
        data = data.set_index('ID number')
        data.head()
```

Out[5]:

ID number	Diagnosis	mean radius	mean texture	mean perimeter	mean area	\
842302	M	17.99	10.38	122.80	1001.0	
842517	M	20.57	17.77	132.90	1326.0	
84300903	M	19.69	21.25	130.00	1203.0	
84348301	M	11.42	20.38	77.58	386.1	
84358402	M	20.29	14.34	135.10	1297.0	

```
           mean smoothness  mean compactness  mean concavity  \
ID number
842302             0.11840           0.27760          0.3001
842517             0.08474           0.07864          0.0869
84300903           0.10960           0.15990          0.1974
84348301           0.14250           0.28390          0.2414
84358402           0.10030           0.13280          0.1980

           mean concave points  mean symmetry           ...             \
ID number                                                 ...
842302                 0.14710         0.2419           ...
842517                 0.07017         0.1812           ...
84300903               0.12790         0.2069           ...
84348301               0.10520         0.2597           ...
84358402               0.10430         0.1809           ...

           Worst radius  Worst texture  Worst perimeter  Worst area  \
ID number
842302            25.38          17.33           184.60      2019.0
842517            24.99          23.41           158.80      1956.0
84300903          23.57          25.53           152.50      1709.0
84348301          14.91          26.50            98.87       567.7
84358402          22.54          16.67           152.20      1575.0

           Worst smoothness  Worst compactness  Worst concavity  \
ID number
842302               0.1622             0.6656           0.7119
842517               0.1238             0.1866           0.2416
84300903             0.1444             0.4245           0.4504
84348301             0.2098             0.8663           0.6869
84358402             0.1374             0.2050           0.4000

           Worst concave points  Worst symmetry  Worst fractal dimension
ID number
842302                   0.2654          0.4601                  0.11890
842517                   0.1860          0.2750                  0.08902
84300903                 0.2430          0.3613                  0.08758
84348301                 0.2575          0.6638                  0.17300
84358402                 0.1625          0.2364                  0.07678

[5 rows x 31 columns]
```

1 MEAN

```
In [30]: mean_results = data.loc[:,features[2:32]].mean()
         mean_results
```

```
Out[30]: mean radius                14.127292
         mean texture               19.289649
         mean perimeter             91.969033
         mean area                 654.889104
         mean smoothness             0.096360
         mean compactness            0.104341
         mean concavity              0.088799
         mean concave points         0.048919
         mean symmetry               0.181162
         mean fractal dimension      0.062798
         SE radius                   0.405172
         SE texture                  1.216853
         SE perimeter                2.866059
         SE area                    40.337079
         SE smoothness               0.007041
         SE compactness              0.025478
         SE concavity                0.031894
         SE concave points           0.011796
         SE symmetry                 0.020542
         SE fractal dimension        0.003795
         Worst radius               16.269190
         Worst texture              25.677223
         Worst perimeter           107.261213
         Worst area                880.583128
         Worst smoothness            0.132369
         Worst compactness           0.254265
         Worst concavity             0.272188
         Worst concave points        0.114606
         Worst symmetry              0.290076
         Worst fractal dimension     0.083946
         dtype: float64
```

2 Mode

```
In [31]: mode_results=data.loc[:,features[2:32]].mode().iloc[0,:]
         mode_results
```

```
Out[31]: mean radius                12.340000
         mean texture               14.930000
         mean perimeter             82.610000
         mean area                 512.200000
         mean smoothness             0.100700
         mean compactness            0.114700
         mean concavity              0.000000
         mean concave points         0.000000
         mean symmetry               0.160100
         mean fractal dimension      0.056670
```

```
SE radius                    0.220400
SE texture                   0.856100
SE perimeter                 1.778000
SE area                     16.640000
SE smoothness                0.005080
SE compactness               0.011040
SE concavity                 0.000000
SE concave points            0.000000
SE symmetry                  0.013440
SE fractal dimension         0.001784
Worst radius                12.360000
Worst texture               17.700000
Worst perimeter            101.700000
Worst area                 284.400000
Worst smoothness             0.121600
Worst compactness            0.148600
Worst concavity              0.000000
Worst concave points         0.000000
Worst symmetry               0.222600
Worst fractal dimension      0.074270
Name: 0, dtype: float64
```

3 skew

```
In [32]: skew_results = data.loc[:,features[2:32]].skew()
         skew_results
```

```
Out[32]: mean radius                0.942380
         mean texture               0.650450
         mean perimeter             0.990650
         mean area                  1.645732
         mean smoothness            0.456324
         mean compactness           1.190123
         mean concavity             1.401180
         mean concave points        1.171180
         mean symmetry              0.725609
         mean fractal dimension     1.304489
         SE radius                  3.088612
         SE texture                 1.646444
         SE perimeter               3.443615
         SE area                    5.447186
         SE smoothness              2.314450
         SE compactness             1.902221
         SE concavity               5.110463
         SE concave points          1.444678
         SE symmetry                2.195133
         SE fractal dimension       3.923969
```

```
Worst radius               1.103115
Worst texture              0.498321
Worst perimeter            1.128164
Worst area                 1.859373
Worst smoothness           0.415426
Worst compactness          1.473555
Worst concavity            1.150237
Worst concave points       0.492616
Worst symmetry             1.433928
Worst fractal dimension    1.662579
dtype: float64
```

4 Standard Deviation

```
In [33]: SD_results= data.loc[:,features[2:32]].std()
         SD_results
```

```
Out[33]: mean radius                3.524049
         mean texture               4.301036
         mean perimeter            24.298981
         mean area                351.914129
         mean smoothness            0.014064
         mean compactness           0.052813
         mean concavity             0.079720
         mean concave points        0.038803
         mean symmetry              0.027414
         mean fractal dimension     0.007060
         SE radius                  0.277313
         SE texture                 0.551648
         SE perimeter               2.021855
         SE area                   45.491006
         SE smoothness              0.003003
         SE compactness             0.017908
         SE concavity               0.030186
         SE concave points          0.006170
         SE symmetry                0.008266
         SE fractal dimension       0.002646
         Worst radius               4.833242
         Worst texture              6.146258
         Worst perimeter           33.602542
         Worst area               569.356993
         Worst smoothness           0.022832
         Worst compactness          0.157336
         Worst concavity            0.208624
         Worst concave points       0.065732
         Worst symmetry             0.061867
         Worst fractal dimension    0.018061
         dtype: float64
```

5 Variance

```
In [34]: var_results=data.loc[:,features[2:32]].var()
         var_results
```

```
Out[34]: mean radius                  12.418920
         mean texture                 18.498909
         mean perimeter              590.440480
         mean area                123843.554318
         mean smoothness              0.000198
         mean compactness             0.002789
         mean concavity               0.006355
         mean concave points          0.001506
         mean symmetry                0.000752
         mean fractal dimension       0.000050
         SE radius                    0.076902
         SE texture                   0.304316
         SE perimeter                 4.087896
         SE area                   2069.431583
         SE smoothness                0.000009
         SE compactness               0.000321
         SE concavity                 0.000911
         SE concave points            0.000038
         SE symmetry                  0.000068
         SE fractal dimension         0.000007
         Worst radius                23.360224
         Worst texture               37.776483
         Worst perimeter           1129.130847
         Worst area              324167.385102
         Worst smoothness             0.000521
         Worst compactness            0.024755
         Worst concavity              0.043524
         Worst concave points         0.004321
         Worst symmetry               0.003828
         Worst fractal dimension      0.000326
         dtype: float64
```

6 Correlations of features using PCC

calculating the correlations of only the continuous value features by dropping the diagnosis . computation done below

```
In [16]: correlations = data.drop('Diagnosis', 1).corr(method='pearson')
         display(correlations)
```

```
                        mean radius  mean texture  mean perimeter  mean area  \
mean radius                1.000000      0.323782        0.997855   0.987357
mean texture               0.323782      1.000000        0.329533   0.321086
```

mean perimeter	0.997855	0.329533	1.000000	0.986507
mean area	0.987357	0.321086	0.986507	1.000000
mean smoothness	0.170581	-0.023389	0.207278	0.177028
mean compactness	0.506124	0.236702	0.556936	0.498502
mean concavity	0.676764	0.302418	0.716136	0.685983
mean concave points	0.822529	0.293464	0.850977	0.823269
mean symmetry	0.147741	0.071401	0.183027	0.151293
mean fractal dimension	-0.311631	-0.076437	-0.261477	-0.283110
SE radius	0.679090	0.275869	0.691765	0.732562
SE texture	-0.097317	0.386358	-0.086761	-0.066280
SE perimeter	0.674172	0.281673	0.693135	0.726628
SE area	0.735864	0.259845	0.744983	0.800086
SE smoothness	-0.222600	0.006614	-0.202694	-0.166777
SE compactness	0.206000	0.191975	0.250744	0.212583
SE concavity	0.194204	0.143293	0.228082	0.207660
SE concave points	0.376169	0.163851	0.407217	0.372320
SE symmetry	-0.104321	0.009127	-0.081629	-0.072497
SE fractal dimension	-0.042641	0.054458	-0.005523	-0.019887
Worst radius	0.969539	0.352573	0.969476	0.962746
Worst texture	0.297008	0.912045	0.303038	0.287489
Worst perimeter	0.965137	0.358040	0.970387	0.959120
Worst area	0.941082	0.343546	0.941550	0.959213
Worst smoothness	0.119616	0.077503	0.150549	0.123523
Worst compactness	0.413463	0.277830	0.455774	0.390410
Worst concavity	0.526911	0.301025	0.563879	0.512606
Worst concave points	0.744214	0.295316	0.771241	0.722017
Worst symmetry	0.163953	0.105008	0.189115	0.143570
Worst fractal dimension	0.007066	0.119205	0.051019	0.003738

	mean smoothness	mean compactness	mean concavity	\
mean radius	0.170581	0.506124	0.676764	
mean texture	-0.023389	0.236702	0.302418	
mean perimeter	0.207278	0.556936	0.716136	
mean area	0.177028	0.498502	0.685983	
mean smoothness	1.000000	0.659123	0.521984	
mean compactness	0.659123	1.000000	0.883121	
mean concavity	0.521984	0.883121	1.000000	
mean concave points	0.553695	0.831135	0.921391	
mean symmetry	0.557775	0.602641	0.500667	
mean fractal dimension	0.584792	0.565369	0.336783	
SE radius	0.301467	0.497473	0.631925	
SE texture	0.068406	0.046205	0.076218	
SE perimeter	0.296092	0.548905	0.660391	
SE area	0.246552	0.455653	0.617427	
SE smoothness	0.332375	0.135299	0.098564	
SE compactness	0.318943	0.738722	0.670279	
SE concavity	0.248396	0.570517	0.691270	
SE concave points	0.380676	0.642262	0.683260	

SE symmetry	0.200774	0.229977	0.178009
SE fractal dimension	0.283607	0.507318	0.449301
Worst radius	0.213120	0.535315	0.688236
Worst texture	0.036072	0.248133	0.299879
Worst perimeter	0.238853	0.590210	0.729565
Worst area	0.206718	0.509604	0.675987
Worst smoothness	0.805324	0.565541	0.448822
Worst compactness	0.472468	0.865809	0.754968
Worst concavity	0.434926	0.816275	0.884103
Worst concave points	0.503053	0.815573	0.861323
Worst symmetry	0.394309	0.510223	0.409464
Worst fractal dimension	0.499316	0.687382	0.514930

	mean concave points	mean symmetry	\
mean radius	0.822529	0.147741	
mean texture	0.293464	0.071401	
mean perimeter	0.850977	0.183027	
mean area	0.823269	0.151293	
mean smoothness	0.553695	0.557775	
mean compactness	0.831135	0.602641	
mean concavity	0.921391	0.500667	
mean concave points	1.000000	0.462497	
mean symmetry	0.462497	1.000000	
mean fractal dimension	0.166917	0.479921	
SE radius	0.698050	0.303379	
SE texture	0.021480	0.128053	
SE perimeter	0.710650	0.313893	
SE area	0.690299	0.223970	
SE smoothness	0.027653	0.187321	
SE compactness	0.490424	0.421659	
SE concavity	0.439167	0.342627	
SE concave points	0.615634	0.393298	
SE symmetry	0.095351	0.449137	
SE fractal dimension	0.257584	0.331786	
Worst radius	0.830318	0.185728	
Worst texture	0.292752	0.090651	
Worst perimeter	0.855923	0.219169	
Worst area	0.809630	0.177193	
Worst smoothness	0.452753	0.426675	
Worst compactness	0.667454	0.473200	
Worst concavity	0.752399	0.433721	
Worst concave points	0.910155	0.430297	
Worst symmetry	0.375744	0.699826	
Worst fractal dimension	0.368661	0.438413	

	mean fractal dimension	...	\
mean radius	-0.311631	...	
mean texture	-0.076437	...	

```
mean perimeter                -0.261477          ...
mean area                     -0.283110          ...
mean smoothness                0.584792          ...
mean compactness               0.565369          ...
mean concavity                 0.336783          ...
mean concave points            0.166917          ...
mean symmetry                  0.479921          ...
mean fractal dimension         1.000000          ...
SE radius                      0.000111          ...
SE texture                     0.164174          ...
SE perimeter                   0.039830          ...
SE area                       -0.090170          ...
SE smoothness                  0.401964          ...
SE compactness                 0.559837          ...
SE concavity                   0.446630          ...
SE concave points              0.341198          ...
SE symmetry                    0.345007          ...
SE fractal dimension           0.688132          ...
Worst radius                  -0.253691          ...
Worst texture                 -0.051269          ...
Worst perimeter               -0.205151          ...
Worst area                    -0.231854          ...
Worst smoothness               0.504942          ...
Worst compactness              0.458798          ...
Worst concavity                0.346234          ...
Worst concave points           0.175325          ...
Worst symmetry                 0.334019          ...
Worst fractal dimension        0.767297          ...

                        Worst radius  Worst texture  Worst perimeter  \
mean radius                 0.969539       0.297008         0.965137
mean texture                0.352573       0.912045         0.358040
mean perimeter              0.969476       0.303038         0.970387
mean area                   0.962746       0.287489         0.959120
mean smoothness             0.213120       0.036072         0.238853
mean compactness            0.535315       0.248133         0.590210
mean concavity              0.688236       0.299879         0.729565
mean concave points         0.830318       0.292752         0.855923
mean symmetry               0.185728       0.090651         0.219169
mean fractal dimension     -0.253691      -0.051269        -0.205151
SE radius                   0.715065       0.194799         0.719684
SE texture                 -0.111690       0.409003        -0.102242
SE perimeter                0.697201       0.200371         0.721031
SE area                     0.757373       0.196497         0.761213
SE smoothness              -0.230691      -0.074743        -0.217304
SE compactness              0.204607       0.143003         0.260516
SE concavity                0.186904       0.100241         0.226680
SE concave points           0.358127       0.086741         0.394999
```

SE symmetry	-0.128121	-0.077473	-0.103753
SE fractal dimension	-0.037488	-0.003195	-0.001000
Worst radius	1.000000	0.359921	0.993708
Worst texture	0.359921	1.000000	0.365098
Worst perimeter	0.993708	0.365098	1.000000
Worst area	0.984015	0.345842	0.977578
Worst smoothness	0.216574	0.225429	0.236775
Worst compactness	0.475820	0.360832	0.529408
Worst concavity	0.573975	0.368366	0.618344
Worst concave points	0.787424	0.359755	0.816322
Worst symmetry	0.243529	0.233027	0.269493
Worst fractal dimension	0.093492	0.219122	0.138957

	Worst area	Worst smoothness	Worst compactness	\
mean radius	0.941082	0.119616	0.413463	
mean texture	0.343546	0.077503	0.277830	
mean perimeter	0.941550	0.150549	0.455774	
mean area	0.959213	0.123523	0.390410	
mean smoothness	0.206718	0.805324	0.472468	
mean compactness	0.509604	0.565541	0.865809	
mean concavity	0.675987	0.448822	0.754968	
mean concave points	0.809630	0.452753	0.667454	
mean symmetry	0.177193	0.426675	0.473200	
mean fractal dimension	-0.231854	0.504942	0.458798	
SE radius	0.751548	0.141919	0.287103	
SE texture	-0.083195	-0.073658	-0.092439	
SE perimeter	0.730713	0.130054	0.341919	
SE area	0.811408	0.125389	0.283257	
SE smoothness	-0.182195	0.314457	-0.055558	
SE compactness	0.199371	0.227394	0.678780	
SE concavity	0.188353	0.168481	0.484858	
SE concave points	0.342271	0.215351	0.452888	
SE symmetry	-0.110343	-0.012662	0.060255	
SE fractal dimension	-0.022736	0.170568	0.390159	
Worst radius	0.984015	0.216574	0.475820	
Worst texture	0.345842	0.225429	0.360832	
Worst perimeter	0.977578	0.236775	0.529408	
Worst area	1.000000	0.209145	0.438296	
Worst smoothness	0.209145	1.000000	0.568187	
Worst compactness	0.438296	0.568187	1.000000	
Worst concavity	0.543331	0.518523	0.892261	
Worst concave points	0.747419	0.547691	0.801080	
Worst symmetry	0.209146	0.493838	0.614441	
Worst fractal dimension	0.079647	0.617624	0.810455	

	Worst concavity	Worst concave points	\
mean radius	0.526911	0.744214	
mean texture	0.301025	0.295316	

mean perimeter	0.563879	0.771241
mean area	0.512606	0.722017
mean smoothness	0.434926	0.503053
mean compactness	0.816275	0.815573
mean concavity	0.884103	0.861323
mean concave points	0.752399	0.910155
mean symmetry	0.433721	0.430297
mean fractal dimension	0.346234	0.175325
SE radius	0.380585	0.531062
SE texture	-0.068956	-0.119638
SE perimeter	0.418899	0.554897
SE area	0.385100	0.538166
SE smoothness	-0.058298	-0.102007
SE compactness	0.639147	0.483208
SE concavity	0.662564	0.440472
SE concave points	0.549592	0.602450
SE symmetry	0.037119	-0.030413
SE fractal dimension	0.379975	0.215204
Worst radius	0.573975	0.787424
Worst texture	0.368366	0.359755
Worst perimeter	0.618344	0.816322
Worst area	0.543331	0.747419
Worst smoothness	0.518523	0.547691
Worst compactness	0.892261	0.801080
Worst concavity	1.000000	0.855434
Worst concave points	0.855434	1.000000
Worst symmetry	0.532520	0.502528
Worst fractal dimension	0.686511	0.511114

	Worst symmetry	Worst fractal dimension
mean radius	0.163953	0.007066
mean texture	0.105008	0.119205
mean perimeter	0.189115	0.051019
mean area	0.143570	0.003738
mean smoothness	0.394309	0.499316
mean compactness	0.510223	0.687382
mean concavity	0.409464	0.514930
mean concave points	0.375744	0.368661
mean symmetry	0.699826	0.438413
mean fractal dimension	0.334019	0.767297
SE radius	0.094543	0.049559
SE texture	-0.128215	-0.045655
SE perimeter	0.109930	0.085433
SE area	0.074126	0.017539
SE smoothness	-0.107342	0.101480
SE compactness	0.277878	0.590973
SE concavity	0.197788	0.439329
SE concave points	0.143116	0.310655

```
SE symmetry                  0.389402                0.078079
SE fractal dimension         0.111094                0.591328
Worst radius                 0.243529                0.093492
Worst texture                0.233027                0.219122
Worst perimeter              0.269493                0.138957
Worst area                   0.209146                0.079647
Worst smoothness             0.493838                0.617624
Worst compactness            0.614441                0.810455
Worst concavity              0.532520                0.686511
Worst concave points         0.502528                0.511114
Worst symmetry               1.000000                0.537848
Worst fractal dimension      0.537848                1.000000

[30 rows x 30 columns]
```

```
In [37]: correlations_ = correlations.where(np.triu(np.ones(correlations.shape)).astype(np.bool
         correlations_ = correlations_.stack().reset_index()
         correlations_.columns = ['By ROW', 'By COLUMN', 'VALUE']
         correlations_.loc[correlations_['By ROW'] == correlations_['By COLUMN'], 'VALUE'] = np
         correlations_ = correlations_.sort_values(by=['VALUE'], ascending=False).dropna()
         display(correlations_)
```

	By ROW	By COLUMN	VALUE
2	mean radius	mean perimeter	0.997855
412	Worst radius	Worst perimeter	0.993708
3	mean radius	mean area	0.987357
60	mean perimeter	mean area	0.986507
413	Worst radius	Worst area	0.984015
430	Worst perimeter	Worst area	0.977578
257	SE radius	SE perimeter	0.972794
79	mean perimeter	Worst perimeter	0.970387
20	mean radius	Worst radius	0.969539
77	mean perimeter	Worst radius	0.969476
22	mean radius	Worst perimeter	0.965137
104	mean area	Worst radius	0.962746
107	mean area	Worst area	0.959213
106	mean area	Worst perimeter	0.959120
258	SE radius	SE area	0.951830
80	mean perimeter	Worst area	0.941550
23	mean radius	Worst area	0.941082
295	SE perimeter	SE area	0.937655
166	mean concavity	mean concave points	0.921391
50	mean texture	Worst texture	0.912045
209	mean concave points	Worst concave points	0.910155
451	Worst compactness	Worst concavity	0.892261
185	mean concavity	Worst concavity	0.884103
141	mean compactness	mean concavity	0.883121

```
160        mean compactness       Worst compactness  0.865809
186          mean concavity    Worst concave points  0.861323
204     mean concave points         Worst perimeter  0.855923
456         Worst concavity    Worst concave points  0.855434
64            mean perimeter     mean concave points  0.850977
142         mean compactness     mean concave points  0.831135
..                       ...                     ...       ...
38             mean texture  mean fractal dimension -0.076437
390             SE symmetry           Worst texture -0.077473
75            mean perimeter             SE symmetry -0.081629
287              SE texture              Worst area -0.083195
68            mean perimeter              SE texture -0.086761
238  mean fractal dimension                 SE area -0.090170
289              SE texture       Worst compactness -0.092439
11              mean radius              SE texture -0.097317
342           SE smoothness    Worst concave points -0.102007
286              SE texture         Worst perimeter -0.102242
391             SE symmetry         Worst perimeter -0.103753
18              mean radius             SE symmetry -0.104321
343           SE smoothness          Worst symmetry -0.107342
392             SE symmetry              Worst area -0.110343
284              SE texture            Worst radius -0.111690
291              SE texture    Worst concave points -0.119638
389             SE symmetry            Worst radius -0.128121
292              SE texture          Worst symmetry -0.128215
98                mean area           SE smoothness -0.166777
338           SE smoothness              Worst area -0.182195
71            mean perimeter           SE smoothness -0.202694
247  mean fractal dimension         Worst perimeter -0.205151
337           SE smoothness         Worst perimeter -0.217304
14              mean radius           SE smoothness -0.222600
335           SE smoothness            Worst radius -0.230691
248  mean fractal dimension              Worst area -0.231854
245  mean fractal dimension            Worst radius -0.253691
66            mean perimeter  mean fractal dimension -0.261477
93                 mean area  mean fractal dimension -0.283110
9               mean radius  mean fractal dimension -0.311631

[435 rows x 3 columns]
```

from above: We infer that the cell nucleus radius and cell nucleus perimeter is having a positive correlation between them. so increase in the cell nucleus radius also yields increase in the perimeter and vice versa. Thus the perimeter feature could replace the radius feature in dataset. The perimeter and the fractal dimension is having a negative correlation between them ,so as perimeter increases ,the fractal dimension decreases and vice versa.

7 Histograms

The histogram below is a plot of Perimeter feature for Beningn patients and Malignant patients

```
In [15]: plt.hist= data.hist(column = 'mean perimeter',by = ['Diagnosis'])

         plt.xlabel('perimeter')
         plt.ylabel('count')
         plt.show()
```