

The data set (and description) can be downloaded here:

<http://archive.ics.uci.edu/ml/machine-learning-databases/liver-disorders/bupa.data>

Description:

1. Title: BUPA liver disorders

2. Source information:

-- Creators: BUPA Medical Research Ltd.

-- Donor: Richard S. Forsyth
8 Grosvenor Avenue
Mapperley Park
Nottingham NG3 5DX
0602-621676

-- Date: 5/15/1990

3. Past usage:

--

None known other than what is shown in the PC/BEAGLE User's Guide
(written by Richard S. Forsyth).

4. Relevant information:

-- The first 5 variables are all blood tests which are thought
to be sensitive to liver disorders that might arise from
excessive alcohol consumption. Each line in the bupa.data file
constitutes the record of a single male individual.

--

It appears that drinks>5 is some sort of a selector on this database.
See the PC/BEAGLE User's Guide for more information.

5. Number of instances: 345

6. Number of attributes: 7 overall

7. Attribute information:

1. mcv mean corpuscular volume

2. alkphos alkaline phosphatase

3. sgpt alamine aminotransferase

4. sgot aspartate aminotransferase

5. gammagt gamma-glutamyl transpeptidase

6. drinks number of half-pint equivalents of alcoholic beverages
drunk per day

7. selector field used to split data into two sets

8. Missing values: none

Citation Request:

Please refer to the repository <http://archive.ics.uci.edu/ml> (see
citation policy).

See also Frank, A. & Asuncion, A. (2010). UCI Machine Learning
Repository

[<http://archive.ics.uci.edu/ml>].

Irvine, CA: University of California, School of Information and
Computer Science.

Descriptive statistics:

Dataset= bupa : n= 345 , d= 6

Class1: n= 145

Covariance matrix:

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]
[1,]	14.9696	5.4507	13.1248	8.0324	39.0269	6.4979
[2,]	5.4507	345.6176	4.4418	23.4608	107.0530	4.2127
[3,]	13.1248	4.4418	248.9430	83.5237	315.4976	26.0226
[4,]	8.0324	23.4608	83.5237	59.8776	143.1381	13.1582
[5,]	39.0269	107.0530	315.4976	143.1381	1103.9025	63.9217
[6,]	6.4979	4.2127	26.0226	13.1582	63.9217	15.4427

Correlation matrix:

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]
[1,]	1.0000	0.0758	0.2150	0.2683	0.3036	0.4274
[2,]	0.0758	1.0000	0.0151	0.1631	0.1733	0.0577
[3,]	0.2150	0.0151	1.0000	0.6841	0.6018	0.4197
[4,]	0.2683	0.1631	0.6841	1.0000	0.5567	0.4327
[5,]	0.3036	0.1733	0.6018	0.5567	1.0000	0.4896
[6,]	0.4274	0.0577	0.4197	0.4327	0.4896	1.0000

Median: 90.1771 69.5565 27.5063 20.9639 22.0795 2.795

Mean: 90.6345 71.9793 31.2069 22.7862 31.5448 3.5414

MCD-estimated:

MDC-0.975-Mean: 89.9813 68.757 26.0467 20.4206 18.8318 2.3598
MDC-0.750-Mean: 89.8889 68.6481 26.0556 20.4907 19.0185 2.2593
MDC-0.500-Mean: 90.0571 68.9333 25.5238 20.1143 18.4381 2.2571

Class2: n= 200

Covariance matrix:

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]
[1,]	23.0862	1.0180	12.1835	9.8273	42.8859	3.2715
[2,]	1.0180	326.2356	41.8236	34.5713	106.1680	7.3935
[3,]	12.1835	41.8236	477.2004	192.5309	445.0601	4.3705
[4,]	9.8273	34.5713	192.5309	127.4371	241.0319	6.9160
[5,]	42.8859	106.1680	445.0601	241.0319	1807.8202	31.7621
[6,]	3.2715	7.3935	4.3705	6.9160	31.7621	8.0751

Correlation matrix:

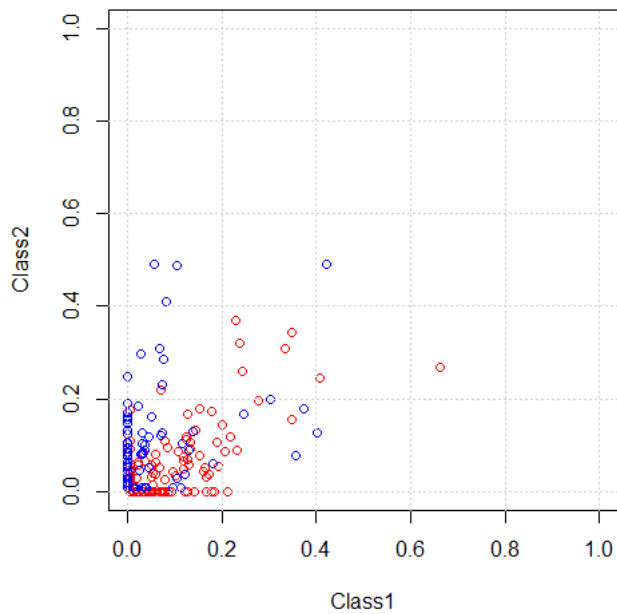
	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]
[1,]	1.0000	0.0117	0.1161	0.1812	0.2099	0.2396
[2,]	0.0117	1.0000	0.1060	0.1696	0.1382	0.1440
[3,]	0.1161	0.1060	1.0000	0.7807	0.4792	0.0704
[4,]	0.1812	0.1696	0.7807	1.0000	0.5022	0.2156
[5,]	0.2099	0.1382	0.4792	0.5022	1.0000	0.2629
[6,]	0.2396	0.1440	0.0704	0.2156	0.2629	1.0000

Median:	89.8908	65.6591	24.9448	23.2974	30.8386	3.0388
Mean:	89.815	68.34	29.825	25.99	43.17	3.3925
MCD-estimated:						
MDC-0.975-Mean:	89.554	66.0072	22.7338	21.9137	25.6115	2.8273
MDC-0.750-Mean:	89.554	66.0072	22.7338	21.9137	25.6115	2.8273
MDC-0.500-Mean:	89.554	66.0072	22.7338	21.9137	25.6115	2.8273

Measures:

Mah. Dist:	0.7934
Mah. Dist-MCD-0.975:	1.0677
Mah. Dist-MCD-0.750:	1.0604
Mah. Dist-MCD-0.500:	1.1049

DD-Plot (zonoid): bupa



DD-Plot (random Tukey): bupa

