

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

<http://lib.stat.cmu.edu/datasets/irish.ed>

**Description:**

-----cut here-----

Irish Educational Transitions Data  
-----

Below are shown data on educational transitions for a sample of 500 Irish schoolchildren aged 11 in 1967. The data were collected by Greaney and Kelleghan (1984), and reanalyzed by Raftery and Hout (1985, 1993).

The data were also used, in a simplified form, as an example to illustrate Bayesian model selection methods by Raftery (1988) and Kass and Raftery (1993). In that simplified form, primary terminal leavers and cases with any missing data were removed, leaving 441 cases. The Leaving Certificate variable was used as the dependent variable in a logistic regression analysis.

The variables shown are as follows:

1. Sex: 1=male; 2=female.
2. DVRT (Drumcondra Verbal Reasoning Test Score).
3. Educational level attained:
  - 1 Primary terminal leaver
  - 2 Junior cycle incomplete: vocational school
  - 3 Junior cycle incomplete: secondary school
  - 4 Junior cycle terminal leaver: vocational school
  - 5 Junior cycle terminal leaver: secondary school
  - 6 Senior cycle incomplete: vocational school
  - 7 Senior cycle incomplete: secondary school
  - 8 Senior cycle terminal leaver: vocational school
  - 9 Senior cycle terminal leaver: secondary school
  - 10 3rd level incomplete
  - 11 3rd level complete
4. Leaving Certificate. 1 if Leaving Certificate not taken; 2 if taken.
5. Prestige score for father's occupation  
(calculated by Raftery and Hout, 1985). 0 if missing.
6. Type of school: 1=secondary; 2=vocational; 9=primary terminal leaver.

REFERENCES  
-----

Greaney, V. and Kelleghan, T. (1984). Equality of Opportunity in Irish Schools. Dublin: Educational Company.

Kass, R.E. and Raftery, A.E. (1993). Bayes factors and model uncertainty. Technical Report no. 254, Department of Statistics, University of Washington. Revised version to appear in Journal of the American Statistical Association.

Raftery, A.E. (1988). Approximate Bayes factors for generalized linear models.

Raftery, A.E. and Hout, M. (1985). Does Irish education approach the meritocratic ideal? A logistic analysis. Economic and Social Review, 16, 115-140.

Raftery, A.E. and Hout, M. (1993). Maximally maintained inequality: Expansion, reform and opportunity in Irish schools. Sociology of Education, 66, 41-62.

-----  
OWNERSHIP STATEMENT  
-----

This data belongs to Vincent Greaney and Thomas Kelleghan, Educational Research Centre, St. Patrick's College, Drumcondra, Dublin 9, Ireland, who retain the copyright.

In the form given here, it may be used solely as an example for research on the development of statistical methods. For any other use of the data, permission must be obtained from the owners. Subject to this statement, permission is hereby given to StatLib to distribute this data freely.

Submitted by Adrian Raftery (raftery@stat.washington.edu).

Copyright 1984 Vincent Greaney and Thomas Kelleghan.  
-----

**Descriptive statistics:**

Dataset= irish-ed\_MvsF : n= 500 , d= 5

Class1: n= 250

**Covariance matrix:**

	[,1]	[,2]	[,3]	[,4]	[,5]
[1,]	292.5973	32.5214	4.2575	56.3259	-15.6852
[2,]	32.5214	11.8575	1.5489	15.6488	-3.9244
[3,]	4.2575	1.5489	0.2452	2.4039	-0.4028
[4,]	56.3259	15.6488	2.4039	330.9177	-7.0682
[5,]	-15.6852	-3.9244	-0.4028	-7.0682	4.5138

**Correlation matrix:**

	[,1]	[,2]	[,3]	[,4]	[,5]
[1,]	1.0000	0.5521	0.5026	0.1810	-0.4316
[2,]	0.5521	1.0000	0.9083	0.2498	-0.5364
[3,]	0.5026	0.9083	1.0000	0.2669	-0.3829
[4,]	0.1810	0.2498	0.2669	1.0000	-0.1829
[5,]	-0.4316	-0.5364	-0.3829	-0.1829	1.0000

Median: 101.1616 5.9075 1.3831 37.7007 1.7701

Mean: 101.348 5.956 1.424 38.504 1.984

**MCD-estimated:**

MDC-0.975-Mean: 104.7976 7.0113 1.5674 39.8989 1.4325

MDC-0.750-Mean: 104.7976 7.0113 1.5674 39.8989 1.4325

MDC-0.500-Mean: 104.7976 7.0113 1.5674 39.8989 1.4325

Class2: n= 250

Covariance matrix:

	[,1]	[,2]	[,3]	[,4]	[,5]
[1,]	183.2872	17.6380	2.1570	50.5642	-9.2037
[2,]	17.6380	10.0282	1.4384	18.5028	-3.5348
[3,]	2.1570	1.4384	0.2497	2.3548	-0.3383
[4,]	50.5642	18.5028	2.3548	260.6748	-9.2828
[5,]	-9.2037	-3.5348	-0.3383	-9.2828	4.0105

Correlation matrix:

	[,1]	[,2]	[,3]	[,4]	[,5]
[1,]	1.0000	0.4114	0.3188	0.2313	-0.3395
[2,]	0.4114	1.0000	0.9090	0.3619	-0.5574
[3,]	0.3188	0.9090	1.0000	0.2919	-0.3380
[4,]	0.2313	0.3619	0.2919	1.0000	-0.2871
[5,]	-0.3395	-0.5574	-0.3380	-0.2871	1.0000

Median: 98.8479 6.1999 1.4094 35.9967 1.5333

Mean: 98.956 6.352 1.464 35.316 1.752

MCD-estimated:

MDC-0.975-Mean: 100.0985 7.5757 1.6441 38.1894 1

MDC-0.750-Mean: 100.0985 7.5757 1.6441 38.1894 1

MDC-0.500-Mean: 100.0985 7.5757 1.6441 38.1894 1

Measures:

Mah.Dist: 0.3694

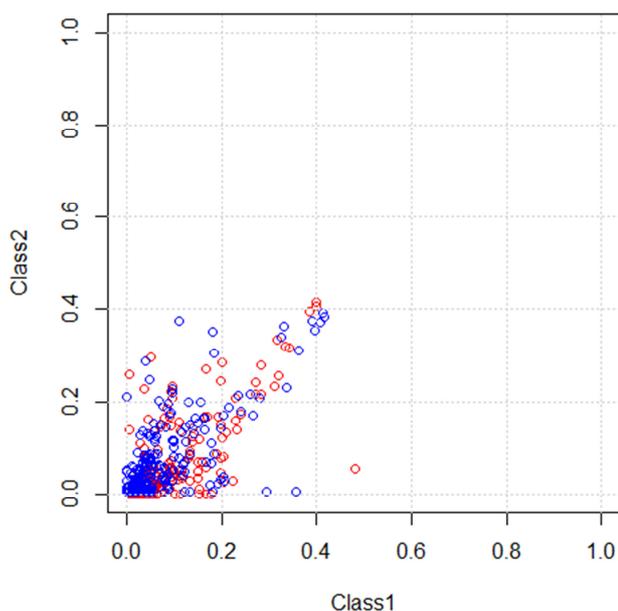
Mah.Dist-MCD-0.975: 1.5484

Mah.Dist-MCD-0.750: 1.5484

Mah.Dist-MCD-0.500: 1.5484

All the MCD estimates have been obtained after a slight perturbation of the data set

DD-Plot (zonoid): irish-ed\_MvsF



DD-Plot (random Tukey): irish-ed\_MvsF

