# Package ‘Flury’

February 14, 2012

**Type** Package

**Version** 0.1-3

**Title** Data Sets from Flury, 1997

**Requires** MASS

**Date** 2007-10-08

**Author** Bernard Flury

**Maintainer** Paul Hewson <paul.hewson@plymouth.ac.uk>

**Description** Contains data sets from Bernard Flury (1997) A First Course in Multivariate Statistics, Springer NY

**License** GPL (>= 2)

**Repository** CRAN

**Date/Publication** 2010-08-02 15:57:03

## R topics documented:

---

`Flury-package` *Data sets from Flury (1997)*

---

## Description

This package contains all the datasets included in Flury (1997), and is intended to support use of the book when teaching with R.

## Details

| | |
|---|---|
| Package: | Flury |
| Type: | Package |
| Version: | 0.1 |
| Date: | 2006-06-15 |
| License: | As per the datasets supplied on the publishers website |

## Author(s)

Bernard Flury

Maintainer: Paul Hewson <paul.hewson@plymouth.ac.uk>

## References

Flury (1997) *A First Course in Multivariate Statistics* New York: Springer ISBN 0-387-98206-X

## See Also

`lookup` cross references the tables in the book with the R object names

---

`angels` *Angels*

---

## Description

Frequency table of the wing length of 130 angels from species Angelus angelus and Angelus diabolicus

## Usage

```
data(angels)
```

## Format

A data frame derived from a frequency table with 10 cells derived from 130 angels

`Wing.Length` Wing length (mm - miracle metres)

`Frequency` Count

## Details

Rather than supply raw data, the original authors only provide information in ten intervals

## Source

Oses, M. and J.Paul (1997) "Morphometric discrimination between two species of angels and estimation of their relative frequency" J.Celest.Morph. 1011:747-767

## References

Flury, B.D. (1997) "A First Course in Multivariate Statistics", Springer NY

## Examples

```
data(angels)
## Not run:
plot(angels, type = "l")
## End(Not run)
```

apples *Apple rootstock data*

## Description

Growth and size measurements on eight apple trees of each of six different rootstocks.

## Usage

`data(apples)`

## Format

A data frame with 48 observations on the following 5 variables.

`Rootstock` a factor with levels `1` `2` `3` `4` `5` `6`

`Girth.4` Trunk girth at 4 years (10cm)

`Growth.4` Extension growth at 4 years (m)

`Girth.15` Trunk girth at 15 years (10cm)

`Weight.15` Weight of tree above ground at 15 years (1000 pounds)

## Source

Andrews D.F. and A.M Herzberg (1985) *Data* New York:Springer

## References

Flury, B.D. (1997) *A First Course in Multivariate Statistics*, New York: Springer

## Examples

```
data(apples)
## Not run:
pairs(apples[,-1],
  lower.panel = function(x, y){ points(x, y,
  pch = unclass(apples[,1]),
  col = as.numeric(apples[,1]))},
  main = "Pairwise scatter plots for apple rootstocks")
## End(Not run)
```

basketball *Basketball Data*

## Description

An experiment conducted on April 18th 1996 at Bryan Park, Bloomington, Indiana. These data record the number of successes within six attempts to hit a basketball from the distance recorded

## Usage

```
data(basketball)
```

## Format

'basketball' is a data frame with 20 observations on the following 5 variables.

`Distance` Distance from the hoop (feet)

`Steve` Steve's score

`Andy` Andy's Score

`Chris` Chris' Score

`Bernard` Bernard's Score (attempts abandoned after 10 feet)

## Details

Further details on the experiment are given in Flury (1997), Table 7.5.4 (page 532)

## Source

Flury, B.D. (1997) "A First Course in Multivariate Statistics", Springer NY

## Examples

```
data(basketball)
## Not run:
bb <- glm(cbind(Andy, 6-Andy) ~ sqrt(Distance),
family = binomial, data = basketball)
with(basketball, (plot(Distance, Andy/6, main = "Andy's Success Rate") ))
lines(basketball$Distance, predict(bb, type = "response"))
## End(Not run)
```

challenger *Challenger O Ring Data*

## Description

These record the number of rocket booster 'O' rings seen to be damaged on each previous shuttle flight prior to the Challenger disaster of 20th January 1986

## Usage

```
data(challenger)
```

## Format

'challenger' is a data frame with 23 observations on the following 2 variables.

`Temp` Ambient temperature (Fahrenheit) at launch time

`Damage` Number of damaged 'O' rings (out of a total of 6)

## Source

Chatterjee, S., M.S. Handcock abd J.S. Simonoff (1995) *A Casebook for a First Course in Statistics and Data Analysis* New York: Wiley

## References

Flury, B.D. (1997) *A First Course in Multivariate Statistics*, New York: Springer

## Examples

```
data(challenger)
## Not run:
ch <- glm(cbind(Damage, 6-Damage) ~ Temp, family = binomial, data = challenger)
with(challenger, plot(Temp, Damage/6))
lines(challenger$Temp, predict(ch, type = "response")  )
abline(v=32, col = "red", lwd = 2)## temp when challenger launched
## End(Not run)
```

## computers — *Computer Repair Data*

### Description

These data give the repair times of N=14 computers.

### Usage

```
data(computers)
```

### Format

A data frame with 14 observations on the following 2 variables.

`Units` The number of parts requiring replacement

`Minutes` Time taken to repair the computer (minutes)

### Source

Chatterjee, S. and B. Price (1977) *Regression Analysis by Example* New York: Wiley

### References

Flury, B.D. (1997) *A First Course in Multivariate Statistics*, New York: Springer

### Examples

```
data(computers)
## Not run:
plot(computers, main = "Computer repair data")
## End(Not run)
```

## cork — *Directional Cork deposit data*

### Description

Cork deposits on North, East, West and South of 28 cork trees.

### Usage

```
data(cork)
```

### Format

A data frame 28 observations on cork deposits from the North, East, South and West.

`North` Deposit on North of tree (mm)

`East` Eastern deposit (mm)

`South` Southern deposit (mm)

`West` Western deposit (mm)

### Source

Rao, C.R. (1948) "Tests of significance in multivariate analysis" *Biometrika* 35:58-79

### References

Flury, B.D. (1997) *A First Course in Multivariate Statistics*, New York: Springer

### Examples

```
data(cork)
## Not run:
pairs(cork)
## End(Not run)
```

---

## dead.beetles *Beetle mortality data*

---

### Description

Beetle mortality data follwoing exposure to insecticide.

### Usage

```
data(dead.beetles)
```

### Format

A data frame with 8 observations on the following 3 variables.

`Dose` log[10] CS[2]mf/l

`tested` Number of beetles tested

`died` Number of beetles killed

### Source

Bliss, C.J. (1975) The calculation of the dosage-mortality curve *Annals of Applied Biology* 22:134-167

### References

Flury, B.D. (1997) *A First Course in Multivariate Statistics*, New York: Springer

### Examples

```
data(dead.beetles)
## Not run:
with(dead.beetles, plot(Dose, died / tested))
dead.glm <- glm(cbind(died, tested-died) ~ Dose, family = binomial, data = dead.beetles)
lines(dead.beetles$Dose, predict(dead.glm, type = "response"))
## End(Not run)
```

---

## deer — *Giant deer premolars*

---

### Description

Breadth of the lower left second premolar of the giant deer (*Megaloceros pachyoseus*)

### Usage

```
data(deer)
```

### Format

'deer' is a vector with 39 observations on the lower left second premolar of the giant deer

### Source

Dong, Z. (1996) Looking into Peking Man's subsistence - a taphonomic analysis of the middle pleistocene *Homo erectus* site in China. Unpublished PhD thesis, Department of Anthropology, Indiana University

### References

Flury, B.D. (1997) *A First Course in Multivariate Statistics*, New York: Springer

### Examples

```
data(deer)
## Not run:
hist(deer, breaks = 10, main = "Deer lower left 2nd premolar")
## End(Not run)
```

## dogwood — *Flowering dogwood trees*

### Description

Data collected on flowering dogwood trees collected in 39 equally sized plots in Hoosier National Forest in Southern Indiana. The total number of trees, the number flowering and the aspect (direction of slope) of the plot were recorded.

### Usage

```
data(dogwood)
```

### Format

'dogwood' is a data frame with observations from 39 plots on the following 3 variables.

`Aspect` Orientation of the plot, i.e. direction of steepest descent, in degrees clockwise from North

`Total.Trees` Total number of trees in the plot

`Flowering.Trees` Number of flowering dogwood trees in the plot

### Source

Data courtesy of K. Dorman and D.Whitehead, Indiana University Department of Biology

### References

Flury, B.D. (1997) *A First Course in Multivariate Statistics*, New York: Springer

### Examples

```
data(dogwood)
```

## electrode — *Electrode Manufacture*

### Description

Comparison of eletrode measurements from two machines

### Usage

```
data(electrode)
```

## Format

A data frame with 100 observations, 50 from each machine, on the following 5 variables.

`Machine` a factor with levels `1` `2`

`X1` Total diameter

`X2` Probe diameter

`X3` Total height

`X4` Basal height

`X5` Internal diameter

## Details

See Figure 5.3.3 in Flury 1997, or the orginal figure 7.20 in Flury and Riedwyl for a better explaination of the five measurements. The actual data presented are linear transformations of the original variables (for reasons of commercial confidentiality).

## Source

Flury, B. and H. Riedwyl (1988) *Multivariate Statistics: A practical approach.* London: Chapman and Hall

## References

Flury, B.D. (1997) *A First Course in Multivariate Statistics* New York: Springer

## Examples

```
data(electrode)
## Not run: pairs(apply(electrode[,-1], 2, jitter, 2),
    lower.panel = function(x, y){ points(x, y,
    pch = unclass(electrode[,1]),
    col = as.numeric(electrode[,1]))},
    main = "Pairwise scatter plots for Electrode data")
## End(Not run)
```

---

`f.swiss.heads` *Female Swiss Heads*

---

## Description

Six measurements on heads of female swiss soldiers.

## Usage

```
data(f.swiss.heads)
```

## Format

A data frame with 59 observations on the following 6 variables.

`MFB` a numeric vector, minimum frontal breadth

`BAM` a numeric vector, breadth of angulus mandibulae

`TFH` a numeric vector, true facial height

`LGAN` a numeric vector, length from glabella to apex nasi

`LTN` a numeric vector, length from tragion to nasion

`LTG` a numeric vector, length from tragion to gnathion

## Details

These data correspond to those reported in siwss.heads, except that here we have data for 59 females.

## Source

Flury, B.D. and H. Riedwyl (1988) *Multivariate Statistics; A Practical Approach* London: Chapman and Hall

## References

Flury, B.D. (1997) *A First Course in Multivariate Statistics* New York: Springer

## Examples

```
data(f.swiss.heads)
## Not run: pairs(f.swiss.heads)
cor(f.swiss.heads) ## how do we compare this with:
cor(swiss.heads)
## End(Not run)
```

---

`f.twins` *Female Twins*

---

## Description

Anthropometric data collected in the 1950's for 36 mono- and 43 dizygotic female twins. See also m.twins

## Usage

```
data(m.twins)
```

### Format

A data frame with 89 observations on the 6 variables and a factor indicating whether the twins are mono- or di-zygotic.

`Type` a factor with levels `monozygotic` `dizygotic`

`STA1` Stature of first twin (cm)

`HIP1` Hip width of first twin (cm)

`CHE1` Chest circumference of first twin (cm)

`STA2` Stature of second twin (cm)

`HIP2` Hip width of second twin (cm)

`CHE2` Chest circumference of second twin (cm)

### Source

Data courtesy of the Institute of Anthropology, University of Hamburg

### References

Flury, B.D. (1997) *A First Course in Multivariate Statistics*, New York: Springer

### Examples

```
data(f.twins)
## Not run: pairs(f.twins[,-1],
  lower.panel = function(x, y){ points(x, y,
  pch = unclass(f.twins[,1]),
  col = as.numeric(f.twins[,1]))},
  main = "Pairwise scatter plots for male twins")
## End(Not run)
```

---

## `f.voles` *Female vole data*

---

### Description

Data on measurements from females of two species of voles, Microtus californicus and M. ochrogaster

### Usage

```
data(f.voles)
```

## Format

A data frame with 86 observations with a factor denoting the species and a further seven variables describing various measurements in units of 0.1mm.

`Species` a factor with levels `californicus` `ochrogaster`

`Age` Age in Days

`L2.Condylo` Condylo incisive length

`L9.Inc.Foramen` Incisive foramen length

`L7.Alveolar` Alveolar length of upper molar tooth row

`B3.Zyg` Zygomatic width

`B4.Interorbital` Interoribital width

`H1.Skull` Skull height

## Details

The letter-number at the start of the variable name refers to the variable names used by the authors in the original publication.

## Source

Airoldi, J.-P. and R.S. Hoffmann (1984) "Age variation in volves (Microtus californicus, M. ochrogaster) and its significance for systematic studies" Occasional papers of the Museum of Natural History, University of Kansas, Lawrence KS 111:1-45

## References

Flury, B.D. (1997) *A First Course in Multivariate Statistics*, New York: Springer

## Examples

```
data(f.voles)
## Not run: pairs(f.voles[,-1],
    lower.panel = function(x, y){ points(x, y,
    pch = unclass(f.voles[,1]),
    col = as.numeric(f.voles[,1]))},
    main = "Pairwise scatter plots for Female vole data")
## End(Not run)
```

## `flea.beetles` *Flea Beetles data*

### Description

Two species of flea beetle (Haltica oleracea, H. carduourum), with 4 body measurements

### Usage

```
data(flea.beetles)
```

### Format

A data frame with 39 observations; 19 from Haltica oleracea and 20 from H. carduourum (denoted by a factor) and four measurements.

`Species` a factor with levels `oleracea` `carduorum`

`TG` Distange of the Transverse Groove to the posterior border of the prothorax (microns)

`Elytra` Length of the Elytra (in units of 0.01mm)

`Second.Antenna` Length of the second antennal joint (microns)

`Third.Antenna` Length of the third antennal joint (microns)

### Source

Lubischew, A.A. (1962) “On the use of discriminant functions in taxonomy”. *Biometrics* 18:455-477

### References

Flury, B.D. (1997) *A First Course in Multivariate Statistics*, New York: Springer

### Examples

```
data(flea.beetles)
## Not run: pairs(flea.beetles[,-1],
  lower.panel = function(x, y){ points(x, y,
  pch = unclass(flea.beetles[,1]),
  col = unclass(flea.beetles[,1]))},
  main = "Pairwise scatter plots for Lubischew's Flea Beetle data")

## End(Not run)
```

ghq *General Health Questionnaire Data*

## Description

Tabulated data on the number of patients requiring psychiatric treatment.

## Usage

```
data(ghq)
```

## Format

'ghq' is a data frame tabulating N= 120 patients according to their gender and GHQ (general health questionnaire) score. There are 17 different combinations of gender and GHQ score, for each combination the number of patients and the number requiring psychiatric treatment are reported.

`Gender` a factor with levels `Male` `Female`

`GHQ` Score on the Generalised Heath Questionnarie

`Number` Number of patients in that combination

`PsychiatricT` Number of patients in that combinations requiring psychiatric treatment

## Source

Silvapulle, M.J. (1981) "On the existence of maximum likelihood estimators for the binomial response model" *Journal of the Royal Statisical Society Series B* 43:310-313

## References

Flury, B.D. (1997) *A First Course in Multivariate Statistics*, New York: Springer

## Examples

```
data(ghq)
```

irisf *Edgar Anderson's Iris Data*

### Description

This famous (Fisher's or Anderson's) iris data set gives the measurements in centimeters of the variables sepal length and width and petal length and width, respectively, for 50 flowers from each of 3 species: "Iris setosa", "I. versicolor", and "I. virginica".

### Usage

```
data(irisf)
```

### Format

'irisf' is a data frame with 150 observations with a factor indicating the species and four quantitative variables:

`Species` a factor with levels `setosa` `versicolor` `virginica`

`Sepal.Length` Sepal length

`Sepal.Width` Sepal width

`Petal.Length` Petal length

`Petal.Width` Petal width

### Source

Fisher, R. A. (1936) "The use of multiple measurements in taxonomic problems". *Annals of Eugenics*, 7:179-188.

Anderson, E (1935) "The irises of the Gaspe Peninsula", *Bulletin of the American Iris Society*, 59:2-5

### References

Flury, B.D. (1997) *A First Course in Multivariate Statistics*, New York: Springer

### See Also

'iris' in datasets, which is presents the same data, but with the species as the fifth column rather than the first

### Examples

```
data(irisf)
## Not run:
pairs(irisf[,-1])
## End(Not run)
```

`lookup` *Lookup table*

## Description

This cross references the data objects and the tables in Flury (1997)

## Usage

```
data(lookup)
```

## Format

A data frame with 33 observations on the following variable.

`Data.Name` R data object name

## References

Flury, B.D. (1997) *A First Course in Multivariate Statistics*, New York: Springer

## Examples

```
data(lookup)
## Not run: lookup
```

`m.twins` *Male Twins*

## Description

Anthropometric data collected in the 1950's for 49 monozygotic and 40 dizygotic male twins. See also female twins (f.twins) .

## Usage

```
data(m.twins)
```

## Format

A data frame with 89 observations on the 6 variables and a factor indicating whether the twins are mono- or di-zygotic.

`Type` a factor with levels `monozygotic` `dizygotic`

`STA1` Stature of first twin (cm)

`HIP1` Hip width of first twin (cm)

CHE1 Chest circumference of first twin (cm)

STA2 Stature of second twin (cm)

HIP2 Hip width of second twin (cm)

CHE2 Chest circumference of second twin (cm)

## Source

Data courtesy of the Institute of Anthropology, University of Hamburg

## References

Flury, B.D. (1997) *A First Course in Multivariate Statistics*, New York: Springer

## Examples

```
data(m.twins)
## Not run:
pairs(m.twins[,-1],
  lower.panel = function(x, y){ points(x, y,
  pch = unclass(m.twins[,1]),
  col = as.numeric(m.twins[,1]))},
  main = "Pairwise scatter plots for male twins")
## End(Not run)
```

---

microtus *Microtus classification (more vole data)*

---

## Description

Microtus multiplex and M. subterraneus are difficult to distinguish morphologically. Here we have 43 known multiplex, 46 known subterraneus and a further 199 unidentified species.

## Usage

```
data(microtus)
```

## Format

A data frame with 288 observations with a factor indicating the species and observations on a further 8 variables.

Group a factor with levels `multiplex` `subterraneus` `unknown`

M1Left Width of upper left molar 1 (0.001mm)

M2Left Width of upper left molar 2 (0.001mm)

M3Left Width of upper left molar 3 (0.001mm)

Foramen Length of incisive foramen (0.001mm)

`Pbone` Length of palatal bone (0.001mm)

`Length` Condylo incisive length or skull length (0.01mm)

`Height` Skull height above bullae (0.01mm)

`Rostrum` Skull width across rostrum (0.01mm)

## Details

89 species have been identified by means of chromosomal analysis.

## Source

Airoldi, J.-P., B. Flury, M. Salvioni (1996) dQuoteDiscrimination between two species of Microtus using both classified and unclassified observations *Journal of Theoretical Biology* 177:247-262

## References

Flury, B.D. (1997) *A First Course in Multivariate Statistics*, New York: Springer

## Examples

```
data(microtus)
## Not run: pairs(microtus[,-1],
    lower.panel = function(x, y){ points(x, y,
    pch = unclass(microtus[,1]),
    col = as.numeric(microtus[,1]))},
    main = "Pairwise scatter plots for Microtus data")
## End(Not run)
```

---

`midge` *Midge Data*

---

## Description

Data on the Antenna length and Wing length of two species of Midge

## Usage

```
data(midge)
```

## Format

A data frame with 15 observations on the following 3 variables.

`Species` a factor with levels `Af` (Amerohelea fasciata) `Apf` (A. pseudofasciata)

`Ant.Length` a numeric vector, Antenna length (mm)

`Wing.Length` a numeric vector, Wing length (mm)

## Details

These data concern two newly discovered species of American predaceous midges. Interest surrounds the possibility of classifying these two species based on external measurements alone. Two possible measures are recorded here.

## Source

Grogan, W.L, and W.W. Wirth (1981) "A new American genus of predaceous midges related to Palpomyia and Bezzia (Diptera: Ceratopogonidae)" *Proceedings of the Biological Society of Washington* 94:1279-1305

## References

Flury, B.D. (1997) *A First Course in Multivariate Statistics*, Springer NY

## Examples

```
data(midge)
## Not run:
with(midge, plot(Ant.Length, Wing.Length,
col = as.numeric(Species), pch = 16, main = "Scatterplot of midge data"))
legend("bottomright", pch = 16, col = c(1,2), legend = c("Af", "Apf"))
## End(Not run)
```

---

# mumps *Mumps antibodies*

---

## Description

Data have been collected on the anti-mumps antibody level in N=385 unvaccinated children aged 14.

## Usage

```
data(mumps)
```

## Format

'mumps' is a vector with 385 observations on children age 14 recording the log concentration of anti-mumps antibodies.

## Source

Data courtesy of Dr. B. Neuenschwander, Swiss Federal Office of Public Health, Division of Epidemiology, 3097 Liebefeld, Switzerland

## References

Flury, B.D. (1997) *A First Course in Multivariate Statistics* New York: Springer

## Examples

```
data(mumps)
## Not run: hist(mumps, breaks = 26,
  main = "Mumps antibody concentration", xlab = "Log antibody concentration" )
## End(Not run)
```

---

pipits *Wing length of Water Pipits*

---

## Description

Measurements of Wing Length for N=381 Water Pipits (*Anthus spinoletta*).

## Usage

```
data(pipits)
```

## Format

A data frame with 17 observations summarising the frequency of wing lengths of 381 Water Pipits.

`Wing.Length` Wing Length (mm)

`Frequency` Frequency observed

## Source

Flury, B.D., J.-P. Airoldi and J.-P. Biber (1992) “Gender identification of water pipits using mixtures of distributions” *Journal of Theoretical Biology* 158:465-480

## References

Flury, B.D. (1997) *A First Course in Multivariate Statistics*, New York: Springer

## Examples

```
data(pipits)
## Not run: plot(pipits, type = "l",
main = "Frequency polygon for wing length of water pipits")
## End(Not run)
```

---

`pipits2` *Wing length of water pipits*

---

### Description

Measurements of Wing Length for 381 Water Pipits caught at Tour du Valat, France.

### Usage

```
data(pipits2)
```

### Format

'pipits2' is a data frame with 15 observations summarising the frequency of wing lengths of 381 Water Pipits.

`Wing.length` Wing Length (mm)

`Number` Frequency observed

### Source

Flury, B.D., J.-P. Airoldi and J.-P. Biber (1992) "Gender identification of water pipits using mixtures of distributions" *Journal of Theoretical Biology* 158:465-480

### References

Flury, B.D. (1997) *A First Course in Multivariate Statistics*, New York: Springer

### Examples

```
data(pipits2)
## Not run:
plot(pipits2, type = "l")
## End(Not run)
```

---

`robustus` *Australopithecus robustus molar size*

---

### Description

*Australopithecus robustus* are early hominids from Africa. The molar size is assumed to be larger in males than female, these data can therefore be used with a mixture model in a study on sexual dimorphism.

## Usage

```
data(robustus)
```

## Format

'robustus' is a vector with 36 observations on the breadth (mm) of the lower first molar of Australopithecus robustus.

## Source

Dong, Z. (1996) Looking into Peking Man's subsistence - a taphonomic analysis of the middle pleistocene *Homo erectus* site in China. Unpublished PhD thesis, Department of Anthropology, Indiana University

## References

Flury, B.D. (1997) *A First Course in Multivariate Statistics*, New York: Springer

## Examples

```
data(robustus)
## Not run:
hist(robustus, breaks =11, main = "Molar breadth")
## End(Not run)
```

---

# sibling.heads *Sibling Heads*

---

## Description

Skull length and breadth measured on siblings in 25 families

## Usage

```
data(sibling.heads)
```

## Format

A data frame with 25 observations on the following 4 variables.

`S1Length` Head length of first son

`S1Breadth` Head breadth of first son

`S2Length` Head length of second son

`S2Breadth` Head breadth of second son

## Source

Frets, G.P. (1921) “Heredity of head form in man” *Genetica* 3:193-384

## References

Flury, B.D. (1997) *A First Course in Multivariate Statistics*, New York: Springer

## Examples

```
data(sibling.heads)
## Not run:
pairs(sibling.heads)
## End(Not run)
```

---

`snailsf` *Aquatic snails hiding behaviour*

---

## Description

Three subjsets of aquatic snail (Potamopyrgus antipodarum) were studied. Data were collected on diseased, juvenile and brooding females to assess the influence of infection on the behaviour of the snails. To avoid predation, snails are usually less likely to be found on the tops of rocks during the day.

## Usage

```
data(snailsf)
```

## Format

’snailsf’ is a data frame with 27 observations on the following 6 variables.

`Infected` a factor with levels `0` (not infected) and `1` (infected)

`Juvenile` a factor with levels `0` and `1` (juvenile)

`Brooding.Female` a factor with levels `0` `1`

`Time` Time of day, in 24 hour clock notation

`Found` Number of snails found on the top of a rock

`Collected` Number of snails collected

## Source

Data courtesy by E.Levri and C.Lively, Indiana University Department of Biology.

## References

Flury, B.D. (1997) *A First Course in Multivariate Statistics*, New York: Springer

### Examples

```
data(snailsf)
## Not run:
S <- sin(snailsf$Time * pi/12)
C <- cos(snailsf$Time * pi/12)
snail.glm <- glm(cbind(Found, Collected-Found) ~ Infected + C + S,
  family = binomial, data = snailsf)
summary(snail.glm)
## End(Not run)
```

---

## steve *Steve's Basketball Data*

---

### Description

Steve shot a basketball six times from each distance (1 to 20 feet), the results of the first throw from each distance are recorded as a hit or a miss

### Usage

```
data(steve)
```

### Format

A data frame with 20 observations on the following 2 variables.

`Distance` Distance between Steve and the basket

`Hit.or.Miss` The result of Steve's first attempt at each distance, a factor with levels `H` `M`

### Source

Flury, B.D. (1997) *A First Course in Multivariate Statistics*, New York: Springer

### Examples

```
data(steve)
## Not run:
steve.glm <- glm(Hit.or.Miss ~ Distance, family = binomial, data =
steve)
plot(steve, main = "Steve's basketball data", pch = 16)
lines(steve$Distance, predict(steve.glm, type = "response"))
## End(Not run)
```

strider *Skeletal dimensions of water striders*

## Description

Water striders live on the surface of lakes and ponds. These insects grow in six distinct stages called instars; at each transition they shed their skin / skeleton. Data has been collected on N=88 female water striders from species Limnoporus canaliculatus. Measurements are reported for the first three instars.

## Usage

```
data(strider)
```

## Format

A data frame with 88 observations on the following 6 variables.

- `Femur1` Average length of femur, instar 1
- `Tibia1` Average length of tibia, instar 1
- `Femur2` Average length of femur, instar 2
- `Tibia2` Average length of tibia, instar 2
- `Femur3` Average length of femur, instar 3
- `Tibia3` Average length of tibia, instar 3

## Details

The actual analysis used the variables $x_j^{(star)} = 100\log(x_j)$, for the j = 1 to 6 variables reported.

## Source

Data courtesy of Dr. C.P. Klingenberg, University of Alberta

## References

Flury, B.D. (1997) *A First Course in Multivariate Statistics*, New York: Springer

## Examples

```
data(strider)
## Not run:
X <- 100 * log(strider)
cov(X) ## etc.
## End(Not run)
```

## swiss.heads *Swiss Heads*

### Description

Six measurements on heads of swiss soldiers.

### Usage

```
data(swiss.heads)
```

### Format

A data frame with 200 observations on the following 6 variables.

- `MFB` a numeric vector, minimum frontal breadth
- `BAM` a numeric vector, breadth of angulus mandibulae
- `TFH` a numeric vector, true facial height
- `LGAN` a numeric vector, length from glabella to apex nasi
- `LTN` a numeric vector, length from tragion to nasion
- `LTG` a numeric vector, length from tragion to gnathion

### Details

Data collected on head measurements of members of the Swiss Army. Six measures are reported here.

swiss.heads contains details for 200 male soldiers, 20 years old f.swiss.heads contains details for 200 female soliders

### Source

Flury, B.D. and H. Riedwyl (1988) *Multivariate Statistics; A Practical Approach* London: Chapman and Hall

### References

Flury, B.D. (1997) *A First Course in Multivariate Statistics* New York: Springer

### Examples

```
data(swiss.heads)
## Not run:
pairs(swiss.heads)
## End(Not run)
```

## treesf *Tree data*

### Description

Measurement on black cherry trees

### Usage

```
data(treesf)
```

### Format

'treesf' is familiar data frame with observations on 31 trees with respect to the following 3 variables.

`Diameter` Diameter measured one foot above ground (inches)

`Height` Height (feet)

`Volume` Volume (cubic feet)

### Details

Note that the diameter needs to be coverted to feet (12 inches in 1 foot) before use. Assuming cone shaped trunks, the colume of wood available should be found from: $V = (\pi/12)D^2H$. By taking logarithms of the variables it is possible to fit $\log V = \log(\pi/12) + 2\log D + \log H$ by using lm().

### Source

Ryan, B.F., B. Joiner and T.A. Ryan (1976) *Minitab Handbook (2nd edition)* Boxton: Duxbury Press (page 278)

### References

Flury, B.D. (1997) *A First Course in Multivariate Statistics*, Springer NY

### Examples

```
data(treesf)
## Not run:
Y <- log(treesf$Volume)
X1 <- log(treesf$Diameter / 12)
X2 <- log(treesf$Height)
trees.lm <- lm(Y ~ X1 + X2)
summary(trees.lm)
log(pi/12)
vcov(trees.lm)
## End(Not run)
```

turtles *Turtle Carapace measurements (Jolicoeur and Mosimann)*

## Description

Measurements on the turtle carapace of 24 male and 24 female painted turtles (Chrysemys picta marginata).

## Usage

```
data(turtles)
```

## Format

'turtle.carapace' is a data frame with 48 observations on the following 4 variables.

`Gender` a factor with levels `Male` `Female`

`Length` carapace length

`Width` carapace width

`Height` carapace height

## Source

Jolicoeur, P. and J.E. Mosimann (1960) "Size and Shape Variation in the Painted Turtle: A Principal Component Analysis", *Growth*, 24:339-354

## References

Flury, B.D. (1997) *A First Course in Multivariate Statistics*, New York: Springer

## Examples

```
data(turtles)
## Not run:
pairs(turtles[,-1],
  lower.panel = function(x, y){ points(x, y,
  pch = unclass(turtles[,1]),
  col = as.numeric(turtles[,1]))},
  main = "Pairwise scatter plots for painted turtles")
## End(Not run)
```

vasoc *Vasoconstriction*

## Description

Effect of the rate and volume of air breathed in by human subjects on vasoconstriction in the skin of their fingers.

## Usage

```
data(vasoc)
```

## Format

'vasoc' is a data frame with 39 observations on the following 3 variables.

- `Volume` Volume of air inspired
- `Rate` Rate of air inspired
- `Y` Binary indicator, 1 = vasoconstriction, 0 = no-vasoconstriction

## Source

Finney, D.J. (1947) The estimation from original records of the relationship between dose and quantal response *Biometrika* 34:320-334

## References

Flury, B.D. (1997) *A First Course in Multivariate Statistics*, New York: Springer

## Examples

```
data(vasoc)
## Not run:
with(vasoc, plot(Volume, Rate, pch = Y+15, col = Y+1))
## End(Not run)
```

wine.sugar *Sugar adulteration in wine production*

## Description

Chemical composition of N=344 commercial samples of concentrated grape must used in wine production. The four variables recorded are considered for discovering adulteration with added sugar from non-grape plants. In unadulterated wines, myoinositol and D/H(I) should follow normal distributions.

## Usage

```
data(wine.sugar)
```

## Format

A data frame with 344 observations on the following 4 variables.

`Myo` Myo-inositol mg/kg sugar

`Scyllo` Scylloinositol mg/kg sugar

`DH.I` D/H(I) (ppm)

`DH.II` D/H(II) (ppm)

## Source

Monetti, A., G. Versini, G. Dalpiaz and F. Reniero (1996) "Sugar adulterations control in concentrated rectified grape musts by finite mixture distribution analysis of the myo- and scyllo-inositol contents and the D/H methyl ratio of fermentative alcohol" *Journal of Agricultural and Food Chemistry* 44:2194-2201

## References

Flury, B.D. (1997) *A First Course in Multivariate Statistics*, New York: Springer

## Examples

```
data(wine.sugar)
## Not run:
with(wine.sugar, plot(log(Myo), DH.I))
require(MASS)
dens <- with(wine.sugar, kde2d(log(Myo), DH.I))
contour(dens)
## End(Not run)
```

wines *Chemical composition of wines*

## Description

These data have been collected on the chemical composition of Weisser Riesling wines from three countries; South Africa,Germany and Italy

## Usage

```
data(wines)
```

## Format

'wines' is a data frame with 26 observations, one factor denoting the country of origin and 15 quantitative variables denoting 15 free monoterpenes and C[13]-norisoprenoids. It is thought these influence the wine's aroma.

`Country` a factor with levels `South Africa` `Germany` `Italy`

`Y1` a numeric vector

`Y2` a numeric vector

`Y3` a numeric vector

`Y4` a numeric vector

`Y5` a numeric vector

`Y6` a numeric vector

`Y7` a numeric vector

`Y8` a numeric vector

`Y9` a numeric vector

`Y10` a numeric vector

`Y11` a numeric vector

`Y12` a numeric vector

`Y13` a numeric vector

`Y14` a numeric vector

`Y15` a numeric vector

## Details

There are a total of nine South African wines, seven German wines (all from Pfalz) and ten from Northern Italy (from both Trentino Alto Adige as Friuli)

### Source

Marais, J., G. Versini, C.J. van Wyj and A. Rapp (1992) "Effect of region on free and bound monoterpene and C[13]-norisoprenoid concentration in Weisser Riesling wines" *South African Journal of Enology and Viniculture* 13:71-77

### References

Flury, B.D. (1997) *A First Course in Multivariate Statistics*, Springer NY

### Examples

```
data(wines)
## Not run: pairs(wines[,-1],
  lower.panel = function(x, y){ points(x, y,
  pch = unclass(wines[,1]),
  col = as.numeric(wines[,1]))},
  main = "Pairwise scatter plots for Marais wine data")
## rather congested scatter plots!
## End(Not run)
```

# Index