Package 'isdals'

July 2, 2014

Version 2.0-1

Title Provides datasets for Introduction to Statistical Data Analysis for the Life Sciences

Author Claus Ekstrom <ekstrom@sund.ku.dk> and Helle Sorensen <helle@math.ku.dk>

Maintainer Claus Ekstrom <ekstrom@sund.ku.dk>

Description Provides datasets for Introduction to Statistical Data Analysis for the Life Sciences

License GPL-2

Suggests VGAM

NeedsCompilation no

Repository CRAN

Date/Publication 2013-10-01 11:18:49

R topics documented:

isdals-package *Data sets for Introduction to Statistical Data Analysis for the Life Sciences*

Description

The package provides the data sets from the book "Introduction to Statistical Data Analysis for the Life Sciences" and is also used for the course in Statistical Data Analysis 1 held at the University of Copenhagen.

Details

Package:	isdals
Type:	Package
Version:	1.0
Date:	2009-08-03
License:	GPL2
LazyLoad:	yes

Author(s)

Maintainer: Claus Ekstrom <ekstrom@life.ku.dk>

References

Claus Thorn Ekstrom and Helle Sorensen (2010). Introduction to Statistical Data Analysis for the Life Sciences. CRC Press

Examples

```
data(chicken)
data(tartar)
```

agefat *Age and body fat percentage*

Description

In order to relate the body fat percentage to age, researchers selected nine healthy adults and determined their body fat percentage.

Usage

```
data(agefat)
```

Format

A data frame with 9 observations on the following 2 variables.

`age` age of the subject

`fatpct` body fat percentage

Source

Ib Skovgaard (2004).Basal Biostatistik 2, Samfundslitteratur.

Examples

```
data(agefat)
```

---

aids *Aids prevalence data*

Description

Number of aids cases and deaths for a 19-year period.

Usage

```
data(aids)
```

Format

A data frame with 19 observations on the following 3 variables.

`year` a numeric vector

`cases` a numeric vector

`deaths` a numeric vector

Examples

```
data(aids)
```

alligator *Alligator food preference*

Description

Data on food preference for 59 alligators. It is of interest to examine if different sized alligators have different food preferences.

Usage

```
data(alligator)
```

Format

A data frame with 59 observations on the following 2 variables.

`length` length of the alligator (in meters)

`food` a factor with levels `Fish` `Invertebrates` `Other` representing the food preference

Source

Agresti, A. (2007). An Introduction to Categorical Data Analysis. Wiley

Examples

```
data(alligator)
library(VGAM)
model <- vglm(food ~ length, family=multinomial, data=alligator)
summary(model)
```

antibio *Decomposition of organic material*

Description

The amount of organic material in heifer dung was measured after eight weeks of decomposition. The data come from 36 heifers from six treatment groups. The treatments are different types of antibiotics. Only 34 observations are available.

Usage

```
data(antibio)
```

Format

A data frame with 34 observations on the following 2 variables.

`type` a factor with the antibiotic treatments. Level: `Alfacyp` `Control` `Enroflox` `Fenbenda` `Ivermect` `Spiramyc`

`org` a numeric vector with the amount of organic matrial

Source

C. Sommer and B. M. Bibby (2002). The influence of veterinary medicines on the decomposition of dung organic matter in soil. European Journal of Soil Biology", 38, 115-159.

Examples

```
data(antibio)
```

`binding` *Binding of antibiotics*

Description

When an antibiotic is injected into the bloodstream, a certain part of it will bind to serum protein. This binding reduces the medical effect. As part of a larger study, the binding rate was measured for 12 cows which were given one of three types of antibiotics: chloramphenicol, erythromycin, and tetracycline

Usage

```
data(binding)
```

Format

A data frame with 12 observations on the following 2 variables.

`antibiotic` antibiotic type. Factor with levels `Chlor` `Eryth` `Tetra`

`binding` binding rate

Source

G. Ziv and F. G. Sulman (1972). Binding of antibiotics to bovine and ovine serum. Antimicrobial Agents and Chemotherapy, 2, 206-213.

Examples

```
data(binding)
```

birthweight *Birth weight of boys and girls*

Description

Data from a study that was undertaken to investigate how the sex of the baby and the age of the fetus influence birth weight during the last weeks of the pregnancy.

Usage

```
data(birthweight)
```

Format

A data frame with 361 observations on the following 5 variables.

`sex` a factor with levels `male` `female`

`age` a numeric vector

`weight` a numeric vector

Source

Anette Dobson (2001). An Introduction to Generalized Linear Models (2nd ed.) Chapman and Hall.

Examples

```
data(birthweight)
## maybe str(birthweight) ; plot(birthweight) ...
```

bodyfat *Body fat in women*

Description

It is expensive and cumbersome to determine the body fat in humans as it involves immersion of the person in water. This dataset provides information on body fat, triceps skinfold thickness, thigh circumference, and mid-arm circumference for twenty healthy females aged 20 to 34. It is desirable if a model could provide reliable predictions of the amount of body fat, since the measurements needed for the predictor variables are easy to obtain.

Usage

```
data(bodyfat)
```

Format

A data frame with 20 observations on the following 4 variables.

`Fat` body fat

`Triceps` triceps skinfold measurement

`Thigh` thigh circumference

`Midarm` mid-arm circumference

Source

J. Neter and M.H. Kutner and C.J. Nachtsheim and W. Wasserman (1996). Applied Linear Statistical Models. McGraw-Hill

Examples

```
data(bodyfat)
```

---

butterfat *Butterfat and dairy cattle*

Description

Average butterfat content (percentages) for random samples of 20 cows (10 two year olds and 10 mature (greater than four years old)) from each of five breeds.

Usage

```
data(butterfat)
```

Format

A data frame with 100 observations on the following 3 variables.

`Butterfat` a numeric vector

`Breed` a factor with levels `Ayrshire` `Canadian` `Guernsey` `Holstein-Fresian` `Jersey`

`Age` a factor with levels `2year` `Mature`

Source

Hand et al. (1993). A Handbook of Small Data Sets. Chapman and Hall

Examples

```
data(butterfat)
```

cabbage *Cabbage yield*

Description

Cabbage yield for different treatment methods and different fields

Usage

```
data(cabbage)
```

Format

A data frame with 16 observations on the following 3 variables.

`method` a factor with levels `A C K N`

`yield` a numeric vector

`field` a numeric vector

Examples

```
data(cabbage)
```

cancer2 *Tumor size and emission of radioactivity*

Description

An experiment involved 21 cancer tumors. For each tumor the weight was registered as well as the emitted radioactivity obtained with a special medical technique (scintigraphic images). Three data points from large tumors were removed.

Usage

```
data(cancer2)
```

Format

A data frame with 18 observations on the following 3 variables.

`id` tumor id (numeric)

`tumorwgt` tumor weight

`radioact` emitted radioactivity (numeric)

Source

Shin et al. (2005). Noninvasive imaging for monitoring of viable cencer cells using a dual-imaging reporter gene. The Journal of Nuclear Medicine, 45, 2109-2115.

Examples

```
data(cancer2)
```

---

cattle *Hormone concentration in cattle*

Description

As part of a larger cattle study, the effect of a particular type of feed on the concentration of a certain hormone was investigated. Nine cows were given the feed for a period, and the hormone concentration was measured initially and at the end of the period.

Usage

```
data(cattle)
```

Format

A data frame with 9 observations on the following 3 variables.

`cow` cow id

`initial` initial homorne concentration (before treatment)

`final` final hormone concentration (after treatment)

Examples

```
data(cattle)
```

---

chicken *Weight gain for chickens*

Description

Twenty chickens were fed with four different feed types - five chickens for each type - and the weight gain was registered for each chicken after a period.

Usage

```
data(chicken)
```

Format

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

`feed` id of feed type. Numeric but it should be used a factor

`gain` Weight gain (numeric)

Source

Anonymous (1949). Query 70. Biometrics, 250–251.

Examples

```
data(chicken)
```

---

chloro *Chlorophyll concentration in winter wheat*

Description

An experiment with winter wheat was carried in order to investigate if the concentration of nitrogen in the soil can be predicted from the concentration of chlorophyll in the plants. The chlorophyll concentration in the leaves as well as the nitrogen concentration in the soil were measured for 18 plants.

Usage

```
data(chloro)
```

Format

A data frame with 18 observations on the following 2 variables.

`chloro` chlorophyll concentration in leaves

`nit` nitrogen concentration in soil

Source

Experiment was carried out at the Royal Veterinary and Agricultural University in Denmark.

Examples

```
data(chloro)
```

cooling *Tenderness of pork*

Description

Two different cooling methods for pork meat were compared in an experiment with 18 pigs from two different groups: low or high pH content. After slaughter, each pig was split in two and one side was exposed to rapid cooling while the other was put through a cooling tunnel. After the experiment, the tenderness of the meat was measured.

Usage

```
data(cooling)
```

Format

A data frame with 18 observations on the following 4 variables.

`pig` a numeric vector with the id of the pig

`ph` pH concentration level. A factor with levels `high` `low`

`tunnel` Tenderness observed from tunnel cooling

`rapid` Tenderness observed from rapid cooling

References

A. J. Moller and E. Kirkegaard and T. Vestergaard (1987). Tenderness of Pork Muscles as Influenced by Chilling Rate and Altered Carcass Suspension. Meat Science, 27, p. 275–286.

Examples

```
data(cooling)
hist(cooling$tunnel[cooling$ph=="low"], main="",
     xlab="Tenderness (low pH)", col="lightgray", ylim=c(0,5), xlim=c(3,9))
hist(cooling$tunnel[cooling$ph=="high"], main="",
     xlab="Tenderness (high pH)", col="lightgray", ylim=c(0,5), xlim=c(3,9))

hist(cooling$tunnel[cooling$ph=="low"], freq=FALSE, main="",
     xlab="Tenderness (low pH)", col="lightgray", ylim=c(0,.5), xlim=c(3,9))
hist(cooling$tunnel[cooling$ph=="high"], freq=FALSE, main="",
     xlab="Tenderness (high pH)", col="lightgray", ylim=c(0,.5), xlim=c(3,9))

plot(cooling$tunnel, cooling$rapid,
     xlim=c(3,9), ylim=c(3,9),
     xlab="Tenderness (tunnel)", ylab="Tenderness (rapid)")

boxplot(cooling$tunnel, cooling$rapid, names=c("Tunnel", "Rapid"),
        ylab="Tenderness score")
```

cornyield *Yield of corn after fertilizer treatment*

Description

Two varieties of corn were randomly assigned to the 8 plots in a completely randomized design so that each variety was planted on 4 plots. Four amounts of fertilizer (5, 10, 15, and 20 units) were randomly assigned to the 4 plots in which variety A was planted. Likewise, the same four amounts of fertilizer were randomly assigned to the 4 plots in which variety B was planted. Yield in bushels per acre was recorded for each plot at the end of the experiment.

Usage

```
data(cornyield)
```

Format

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

`yield` a numeric vector

`variety` a factor with levels `A` `B`

`fertilizer` a numeric vector

Examples

```
data(cornyield)
```

crabs *Weight of crabs*

Description

The length and weight of 361 crabs. The crabs were measured at three different days and they were raised in three different vat types.

Usage

```
data(crabs)
```

Format

A data frame with 361 observations on the following 5 variables.

`day` id for day of measurement (a numeric vector)

`date` date of measurement (a numeric vector)

`kar` id of the vat type (a numeric vector

`lgth` length of the crab in cm

`wgt` weight of the crab in grams

Details

Only crabs from day 1 (190692) are used in the isdals book.

Source

Experiment carried out at the Royal Veterinary and Agricultural University of Copenhagen.

Examples

```
data(crabs)
```

---

cuckoo *Hatching of cuckoo eggs*

Description

Cuckoos place their eggs in other birds' nests for hatching and rearing. Researchers investigated 154 cuckoo eggs and measured their size. The adoptive species is also registered (three types). It is believed that cuckoos choose the "adoptive parents" such that the cuckoo eggs are similar in size to the eggs of the adoptive species.

Usage

```
data(cuckoo)
```

Format

A data frame with 154 observations on the following 2 variables.

`spec` adoptive species. Factor with levels `redstart` `whitethroat` `wren`

`width` width of egg (unit: half millimeters)

Source

O.H. Latter (1905). The egg of Cuculus Canorus: An attempt to ascertain from the dimensions of the cuckoo's egg if the species is tending to break up into sub-species, each exhibiting a preference for some one foster-parent. Biometrika, 4, 363-373.

Examples

```
data(cuckoo)
```

cucumber *Disease spread in cucumber*

Description

Spread of a disease in cucumbers depends on climate and amount of fertilizer. The amount of infection on standardized plants was recorded after a number of days, and two plants were examined for each combination of climate and dose.

Usage

```
data(cucumber)
```

Format

A data frame with 12 observations on the following 3 variables.

`disease` a numeric vector

`climate` a factor with levels `A` (change to day temperature 3 hours before sunrise) and `B` (normal change to day temperature)

`dose` a numeric vector with dose of applied fertilizer

Source

de Neergaard, E. et al (1993). Studies of Didymella bryoniae: the influence of nutrition and cultural practices on the occurrence of stem lesions and internal and external fruit rot on different cultivars of cucumber. Netherlands Journal of Plant Pathology. 99:335-343

Examples

```
data(cucumber)
```

dhl *Running times from relay race*

Description

Running times from 5 times 5 km relay race in Copenhagen 2006, held over four days. The sex distribution in the team classifies the teams into six groups. Total running time for a team (not each participant) is registered.

Usage

```
data(dhl)
```

Format

A data frame with 24 observations on the following 6 variables.

`day` race day. A factor with levels `Monday` `Thursday` `Tuesday` `Wednesday`

`men` number of men on the team (numeric)

`women` number of men on the team (numeric)

`hours` hours of running (should be combined with minutes and seconds)

`minutes` minutes of running (should be combined with hours and seconds)

`seconds` seconds of running (should be combined with hours and minutes)

Details

The total running time for the team (not for each participant) is registered. On average, there are 800 teams per combination of race day and sex group. The dataset contains median running times.

Source

http://www.sparta.dk

Examples

```
data(dhl)
attach(dhl)
totaltime <- 60*60*hours + 60*minutes + seconds ## Total time in seconds
```

---

digestcoefs *Effect of NaOH treatment of straw on digesitibility*

Description

In an experiment with six horses the digestibility coefficient was measured twice for each horse: once after the horse had been fed straw treated with NaOH and once after the horse had been treated ordinary straw.

Usage

```
data(digestcoefs)
```

Format

A data frame with 6 observations on the following 3 variables.

`horse` horse id

`ordinary` digestibility coefficient corresponding to ordinary straw

`naoh` digestibility coefficient corresponding to NaOH treated straw

Source

Ib Skovgaard (2004). Basal Biostatistik 2. Samfundslitteratur.

Examples

```
data(digestcoefs)
```

`duckweed` *Growth of duckweed*

Description

Growth of duckweed (Lemna) by counting the number of leaves every day over a two-week period

Usage

```
data(duckweed)
```

Format

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

`days` a numeric vector

`leaves` a numeric vector

Source

E. Ashby and T. A. Oxley (1935). The interactions of factors in the growth of Lemna. Annals of Botany. 49:309-336

Examples

```
data(duckweed)
```

elisa — *Optical density for dilutions of a standard dissolution with ubiquitin antibody*

Description

As part of a so-called ELISA experiment, the optical density was measured for various dilutions of two different dissolutions with ubiquitin antibody. One dissolution was standard, whereas the other was serum from mice. For each dilution, the mixture proportion describes how many times the original ubiquitin dissolution has been thinned.

Usage

```
data(elisa)
```

Format

A data frame with 16 observations on the following 3 variables.

`type` type of dissolution. Factor with levels `mouse` `std`

`mix` a numeric vector describing how many times the original ubiquitin dissolution was thinned

`od` optical density

Source

The data was generated by Marianne Freisleben in her work for the master's thesis at the University of Copenhagen.

Examples

```
data(elisa)
```

farmprice — *Relation between soil area and price for farms*

Description

In February 2010, 12 production farms were for sale in a municipality on Fuen island in Denmark. The dataset contains the soil area in thousands of square meters and the price in thousands of DKK.

Usage

```
data(farmprice)
```

Format

A data frame with 12 observations on the following 2 variables.

`area` area of soil in thousands of square meters

`price` price in thousands of DKK

Examples

```
data(farmprice)
```

---

fev *Forced expiratory volume in children*

Description

Dataset to examine if respiratory function in children was influenced by exposure to smoking at home.

Usage

```
data(fev)
```

Format

A data frame with 654 observations on the following 5 variables.

`Age` age in years

`FEV` forced expiratory volume in liters

`Ht` height measured in inches

`Gender` gender (0=female, 1=male)

`Smoke` exposure to smoking (0=no, 1=yes)

Source

I. Tager and S. Weiss and B. Rosner and F. Speizer (1979). Effect of Paternal Cigarette Smoking on the Pulmonary Function of Children. American Journal of Epidemiology. 110:15-26

Examples

```
data(fev)
```

geneexp *Gene expression*

Description

Two groups were compared in an experiment with six microarrays. Two conditions (the test group and the reference group) were examined on each array and the amount of protein synthesized by the gene (also called the gene expression) was registered.

Usage

```
data(geneexp)
```

Format

A data frame with 6 observations on the following 3 variables.

`array` array id

`test` gene expression level for test group

`reference` gene expression level for reference group

Source

Fictious data.

Examples

```
data(geneexp)
```

gestation *Gestation period for 13 horses*

Description

The length of the gestation period (the period from conception to birth) was registered for 13 horses.

Usage

```
data(gestation)
```

Format

A data frame with 13 observations on the following variable.

`gest` length of gestation period

Source

Fictious (but realistic) data.

Examples

```
data(gestation)
```

`hazard` *Sorption of hazardous organic solvents*

Description

The sorption was measured for a variety of hazardous organic solvents. The solvents were classified into three types (esters, aromatics, and chloroalkanes), and the purpose was to examine differences between the three types.

Usage

```
data(hazard)
```

Format

A data frame with 32 observations on the following 2 variables.

`type` type of solvent. Factor with levels `aromatic` `chlor` `estere`

`sorption` sorption measurements

Source

J.D. Ortego, T.M Aminabhavi, S.F. Harlapur, R.H. Balundgi (1995). A review of polymeric geosynthetics used in hazardous waste facilities. Journal of Hazardous Materials, 42, 115-156.

Examples

```
data(hazard)
```

herring *Nematodes in herring fillets*

Description

An experiment was carried out in order to investigate the migration of nematodes in Danish herrings. The fish were allocated to eight different treatment groups corresponding to different combinations of storage time and storage conditions until filleting. After filleting, it was determined whether nematodes were present in the fillet or not.

Usage

```
data(herring)
```

Format

A data frame with 884 observations on the following 4 variables.

`group` a numeric vector that is the combination of storage and time

`time` a numeric vector that contains the duration of storage in hours before the fish is filleted

`condi` a numeric vector representing the storage condition

`fillet` a numeric vector to indicate the presence of nematodes (1) or absence of nematodes (0)

Details

The variable group is the combination of storage condition and storage time. Notice that a storage time 0 is equivalent to storage condition 0 and that no fish were stored 132 hours under condition 4. Hence, there are only 8 combinations; i.e., 8 levels of the group variable.

Source

A. Roepstorff and H. Karl and B. Bloemsma and H. H. Huss (1993). Catch handling and the possible migration of Anisakis larvae in herring, Clupea harengus. Journal of Food Protection. 56:783-787.

Examples

```
data(herring)
## maybe str(herring) ; plot(herring) ...
```

hormone *Hormone concentration in cattle*

Description

As part of a larger cattle study, the effect of a particular ype of feed on the concentration of a certain hormone was investigated. Twenty cows were given the feed for a period, and the hormone concentration was measured initially and at the end of the period.

Usage

```
data(hormone)
```

Format

A data frame with 20 observations on the following 3 variables.

`feed` a numeric vector

`initial` a numeric vector

`final` a numeric vector

Examples

```
data(hormone)
```

interspike *Interspike intervals for neureon from guinea pigs*

Description

A study of the membrane potential for neurons from guinea pigs was carried out. The data consists of 312 measurements of interspike intervals; that is, the length of the time period between spontaneous firings from a neuron.

Usage

```
data(interspike)
```

Format

A data frame with 312 observations on the following variable.

`interval` length of the interspike intervals

Source

Petr Lansky, Pavel Sanda and Jufang He (2006). The parameters of the stochastic leaky integrate-and-fire neuronal model. Journal of Computational Neuroscience, 21, 211-223.

Examples

```
data(interspike)
```

`lameness` *Lameness scores for horses*

Description

A score measuring the symmetry of the gait for eight trotting horses. Each horse was tested twice, namely while it was clinically healthy and after mechanical induction of lameness in a fore limb.

Usage

```
data(lameness)
```

Format

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

`horse` a numeric vector with an id of the horse

`lame` the symmetry score when the horse is lame

`healthy` the symmetry score when the horse is healthy

Source

A.T. Jensen, H. Sorensen, M.H. Thomsen and P.H. Andersen (2010). Quantification of symmetry for functional data with application to equine lameness classification. Submitted manuscript.

Examples

```
data(lameness)
```

lifespan *Length of gestation period and lifespan for horses*

Description

Length of the gestation period (period from conception to birth) and the lifespan (duration of life) for seven horses.

Usage

```
data(lifespan)
```

Format

A data frame with 7 observations on the following 2 variables.

`lifespan` duration of life (years)

`gestation` length of gestation period (days)

Source

Probably fictitous data.

Examples

```
data(lifespan)
```

logit *Calculate the logit transform*

Description

Calculate the logit transform

Usage

```
logit(p)
```

Arguments

p	numeric vector

Details

Calculates the logit transform of p, ie., log(p/(1-p))

Value

Calculates the logit transform of p

Author(s)

Claus Ekstrom <ekstrom@life.ku.dk>

Examples

```
p <- 0.3
logit(p)
```

`lucerne` *Fertility of lucerne*

Description

Ten plants were used in an experiment of fertility of lucerne Two clusters of flowers were selected from each plant and pollinated. One cluster was bent down, whereas the other was exposed to wind and sun. At the end of the experiment, the average number of seeds per pod was counted for each cluster and the weight of 1000 seeds was registered for each cluster.

Usage

```
data(lucerne)
```

Format

A data frame with 10 observations on the following 5 variables.

`plant` plant id

`seeds.exp` average number of seeds per pod from cluster exposed to sun and wind

`wgt.exp` weight of 1000 seeds from cluster exposed to sun and wind

`seeds.bent` average number of seeds per pod from cluster that was bent down

`wgt.bent` weight of 1000 seeds from cluster that was bent down

Source

H.L. Petersen (1954). Pollination and seed setting in lucerne. Kgl. Veterinaer og Landbohojskole, Aarsskrift 1954, 138-169.

Examples

```
data(lucerne)
```

mackerel *Nematodes in mackerel*

Description

Data to examine if cooling right after catching prevents nematodes (roundworms) from moving from the belly of mackerel to the fillet. A total of 150 mackerels were investigated and their length, number of nematodes in the belly, and time before counting the nematodes in the fillet were registered. The response variable is binary: presence or absence of nematodes in the fillet.

Usage

```
data(mackerel)
```

Format

A data frame with 150 observations on the following 7 variables.

`length` a numeric vector

`visc` a numeric vector

`left` a numeric vector

`right` a numeric vector

`filet` a numeric vector

`portion` a numeric vector

`time` a numeric vector

Source

A. Roepstorff and H. Karl and B. Bloemsma and H. H. Huss (1993). Catch handling and the possible migration of Anisakis larvae in herring, Clupea harengus. Journal of Food Protection. 56:783-787.

Examples

```
data(mackerel)
## maybe str(mackerel) ; plot(mackerel) ...
```

malaria *Parasite counts for children with malaria*

Description

A medical researcher took blood samples from 31 children who were infected with malaria and determined for each child the number of malaria parasites in 1 ml of blood.

Usage

```
data(malaria)
```

Format

A data frame with 31 observations on the following variable.

`parasites` the number of malaria parasites

Source

M.L. Samuels and J.A. Witmer (2003). Statistics for the Life Sciences (3rd ed.). Pearson Education, Inc., New Jersey.

References

C. B. Williams (1964) Patterns in the Balance of Nature. Academic Press, London.

Examples

```
data(malaria)
```

massspec *Comparison of mass spectrometry methods*

Description

Two common methods are GC-MS (gas chromatography-mass spectrometry) and HPLC (high performance liquid chromatography). The biggest difference between the two methods is that one uses gas while the other uses liquid. We wish to determine if the two methods measure the same amount of muconic acid in human urine.

Usage

```
data(massspec)
```

Format

A data frame with 16 observations on the following 3 variables.

`sample` a numeric vector

`hplc` a numeric vector

`gcms` a numeric vector

Examples

```
data(massspec)
```

`mincedmeat` *Weight of packs with minced meat*

Description

In meat production, packs of minced meat are specified to contain 500 grams of minced meat. A sample of ten packs was drawn at random and the weights (in grams) of the content was recorded.

Usage

```
data(mincedmeat)
```

Format

A data frame with 10 observations on the following variable.

`wgt` weight of minced meat in grams

Source

Fictitious data.

Examples

```
data(mincedmeat)
```

oilvit *Utilization of vitamin A*

Description

In an experiment on the utilization of vitamin A, 20 rats were given vitamin A over a period of three days. Ten rats were fed vitamin A in corn oil and ten rats were fed vitamin A in castor oil (American oil). On the fourth day, the liver of each rat was examined and the vitamin A concentration in the liver was determined.

Usage

```
data(oilvit)
```

Format

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

`type` type of oil. A factor with levels `am` `corn`

`avit` vitamin A concentration in liver

Source

C.I.Bliss (1967). Statistics in Biology. McGraw-Hill, New York

Examples

```
data(oilvit)
```

paperstr *Tensile strength of Kraft paper*

Description

Tensile strength in pound-force per square inch of Kraft paper (used in brown paper bags) for various amounts of hardwood contents in the paper pulp.

Usage

```
data(paperstr)
```

Format

A data frame with 19 observations on the following 2 variables.

`hardwood` hardwood content

`strength` tensile strength in pound-force per square inch

Source

G. Joglekar and J. H. Schuenemeyer and V. LaRiccia (1989). Lack-of-Fit Testing When Replicates Are Not Available. The American Statistician. 43:135-143

Examples

```
data(paperstr)
```

---

phosphor *Phosphor concentration in plants during growth*

Description

In a plant physiological experiment the amount of water-soluble phosphorous (among others) was measured in the plants, as a percentage of dry matter. The phosphorous concentration was measured nine weeks during the growth season, and the averages over the plants in the experiments was reported.

Usage

```
data(phosphor)
```

Format

A data frame with 9 observations on the following 2 variables.

`week` week number

`phos` phosphor concentration (average over the plants)

Source

Ib Skovgaard (2004). Basal Biostatistik 2, Samfundslitteratur.

Examples

```
data(phosphor)
```

pillbug *Effect of stimuli on pillbugs*

Description

An experiment on the effect of different stimuli was carried out with 60 pillbugs. The bugs were split into three groups: 20 bugs were exposed to strong light, 20 bugs were exposed to moisture, and 20 bugs were used as controls. For each bug it was registered how many seconds it used to move six inches.

Usage

```
data(pillbug)
```

Format

A data frame with 60 observations on the following 2 variables.

`time` number of seconds it took the pillbug to move six inches

`group` treatment. A factor with levels `Control` `Light` `Moisture`

Source

Samuels and Witmer (2003). Statistics for the Life Sciences (3rd ed.). Pearson Education, Inc., New Jersey.

Examples

```
data(pillbug)
```

pine *Height and diameter of pines*

Description

The data consist of height and diameter (in breast height) measurements from 18 pine trees.

Usage

```
data(pine)
```

Format

A data frame with 18 observations on the following 2 variables.

`diam` diameter of the pine tree

`height` height of the pine tree

Source

J.N.R. Jeffers (1959). Experimental Design and Analysis in Forest Research. Almqvist \& Wiksell, Stockholm.

Examples

```
data(pine)
```

`poison` *Effects of insecticides on mortality*

Description

The data concerns three insecticides (rotenone, deguelin, and a mixture of those). A total of 818 insects were exposed to different doses of one of the three insecticides. After exposure, it was recorded if the insect died or not.

Usage

```
data(poison)
```

Format

A data frame with 818 observations on the following 3 variables.

`status` status of insect: dead=1, alive=0 (numeric vector)

`poison` type of insecticide. A factor with levels `D` (deguelin) `M` (mixture)) `R` (rotenone)

`logdose` natural logarithm of dose of insecticide

Source

D.J. Finney (1952). Probit analysis. Cambridge University Press, England.

Examples

```
data(poison)
```

pork *Pork colour over time*

Description

Investigation of meat quality of pork through color stability of pork chops. The color was measured from a pork chop from each of ten pigs at days 1, 4, and 6 after storage.

Usage

```
data(pork)
```

Format

A data frame with 30 observations on the following 3 variables.

`brightness` a numeric vector

`day` a numeric vector

`pig` a numeric vector

Examples

```
data(pork)
```

ratliver *Drugs in rat's livers*

Description

An experiment was undertaken to investigate the amount of drug present in the liver of a rat. Nineteen rats were randomly selected, weighed, placed under a light anesthetic, and given an oral dose of the drug. It was believed that large livers would absorb more of a given dose than a small liver, so the actual dose given was approximately determined as 40 mg of the drug per kilogram of body weight. After a fixed length of time, each rat was sacrificed, the liver weighed, and the percent dose in the liver was determined.

Usage

```
data(ratliver)
```

Format

A data frame with 19 observations on the following 4 variables.

`BodyWt` body weight of each rat in grams

`LiverWt` weight of liver in grams

`Dose` relative dose of the drug given to each rat as a fraction of the largest dose

`DoseInLiver` proportion of the dose in the liver

Source

S. Weisberg (1985). Applied Linear Regression (2nd ed.). John Wiley and Sons

Examples

```
data(ratliver)
```

`ratweight` *Weight gain of rats*

Description

Data contains the weight gain for rats fed on four different diets: combinations of protein source (beef or cereal) and protein amount (low and high)

Usage

```
data(ratweight)
```

Format

A data frame with 40 observations on the following 3 variables.

`Gain` a numeric vector

`Protein` a factor with levels `Beef` `Cereal`

`Amount` a factor with levels `High` `Low`

Source

Hand et al. (1993). A Handbook of Small Data Sets. Chapman and Hall

Examples

```
data(ratweight)
```

`residualplot` *Plots a standardaized residual*

Description

Plots a standardized residual plot from an lm object and provides additional graphics to help evaluate the variance homogeneity and mean.

Usage

```
residualplot(object, bandwidth = 0.3, ...)
```

Arguments

`object`	lm object
`bandwidth`	The width of the window used to calculate the local smoothed version of the mean and the variance. Value should be between 0 and 1 and determines the percentage of the windowwidth used
`...`	Other arguments passed to the plot function

Details

Plots a standardized residual plot from an lm object and provides additional graphics to help evaluate the variance homogeneity and mean.

The brown area is a smoothed estimate of 1.96*SD of the standardized residuals in a window around the predicted value. The brown area should largely be rectangular if the standardized residuals have more or less the same variance.

The dashed line shows the smoothed mean of the standardized residuals and should generally follow the horizontal line through (0,0).

Value

Produces a standardized residual plot

Author(s)

Claus Ekstrom <ekstrom@life.ku.dk>

See Also

`rstandard`, `predict`

Examples

```
# Linear regression example
x <- rnorm(100)
y <- rnorm(100, mean=.5*x)
model <- lm(y ~ x)
residualplot(model)
```

`ricestraw` *Weight increase for cattle fed with rice straw*

Description

Weight gain of cattle fed with rice straw to see if rice straw can replace wheat straw as potential feed for slaughter cattle in Tanzania

Usage

```
data(ricestraw)
```

Format

A data frame with 35 observations on the following 2 variables.

`time` number of days that the calf has been fed rice straw

`weight` weight gain (in kg) since the calf was first fed rice straw

Source

Ph.D. project at the Faculty of LIFE Sciences, University of Copenhagen

Examples

```
data(ricestraw)
plot(ricestraw$time, ricestraw$weight)
lm(weight ~ time, data=ricestraw)
```

salmon *Parasite counts for salmons*

Description

An experiment with two difference salmon stocks, from River Conon in Scotland and from River Atran in Sweden, was carried out. Thirteen fish from each stock were infected and after four weeks the number of a certain type of parasites was counted for each of the 26 fish.

Usage

```
data(salmon)
```

Format

A data frame with 26 observations on the following 2 variables.

`stock` origin of the fish. A factor with levels `atran` `conon`

`parasites` a numeric vector with the parasite counts

Source

Heinecke, R. D, Martinussen, T. and Buchmann, K. (2007). Microhabitat selection of Gyrodactylus salaris Malmberg on different salmonids. Journal of Fish Diseases, 30, 733-743.

Examples

```
data(salmon)
```

sarcomere *Sarcomere length and meat tenderness*

Description

The average sarcomere length in the meat and the corresponding tenderness as scored by a panel of sensory judges was examined. A high score corresponds to tender meat.

Usage

```
data(sarcomere)
```

Format

A data frame with 24 observations on the following 3 variables.

`pig` factor with levels 1–24. Pid id

`sarc.length` numeric Sarcomere length

`tenderness` numeric Meat tenderness score

References

A. J. Moller and E. Kirkegaard and T. Vestergaard (1987). Tenderness of Pork Muscles as Influenced by Chilling Rate and Altered Carcass Suspension. Meat Science, 27, p. 275–286.

Examples

```
data(sarcomere)
cor(sarcomere$sarc.length, sarcomere$tenderness)
```

`seal` *Size of seal population from 1952 to 1962*

Description

The number of seals in a population were counted each year during a period of 11 years, freom 1952 to 1962.

Usage

```
data(seal)
```

Format

A data frame with 11 observations on the following 2 variables.

`year` year of seal count

`size` number of seals in population

Source

J. Verzani (1005). Using R for Introductory Statistics. Chapman \& Hall/CRC, London

Examples

```
data(seal)
```

soybean *Stress and growth for soybeans*

Description

An experiment was carried out with 26 soybean plants. The plants were pairwise genetically identical, so there were 13 pairs in total. For each pair, one of the plants was 'stressed' by being shaken daily, whereas the other plant was not shaken. After a period the plants were harvested and the total leaf area was measured for each plant.

Usage

```
data(soybean)
```

Format

A data frame with 13 observations on the following 3 variables.

`pair` id of the pair of plants

`stress` Total leaf area of stressed plant

`nostress` total leaf area of control plant

Examples

```
data(soybean)
```

stearicacid *Digestibility percentage of fat for various levels of stearic acid*

Description

The average digestibility percent was measured for nine different levels of stearic acid proportion

Usage

```
data(stearicacid)
```

Format

A data frame with 9 observations on the following 2 variables.

`stearic.acid` Percentage of stearic acid

`digest` Average digestibility percentage

Source

Jorgensen, G. and Hansen, N.G. (1973). Fedtsyresammensaetningens indflydelse paa fedstoffers fordojelighed. Landokonomisk Forsogslaboratorium.

Examples

```
data(stearicacid)
lm(digest ~ stearic.acid, data=stearicacid)
```

---

stomach *Stomach experiment*

Description

Fifteen subjects participated in an experiment related to overweight and got a standardized meal.The interest was, among others, to find relationships between the time it takes from a meal until the stomach is empty again and the concentration of a certain hormone.

Usage

```
data(stomach)
```

Format

A data frame with 15 observations on the following 2 variables.

`conc` hormone concentration

`empty` time from meal until the stomach is empty

Source

Ib Skovgaard (2004). Basal Biostatistik 2. Samfundslitteratur.

Examples

```
data(stomach)
```

tartar *Tartar for dogs*

Description

A dog experiment was carried out in order to examine the effect of two treatments on the development of tartar. Apart from the two treatment groups there was also a control group. Twenty-six dogs were used and allocated to one of the three groups. After four weeks each dog was examined, and the development of tartar was summarized by an index.

Usage

```
data(tartar)
```

Format

A data frame with 26 observations on the following 2 variables.

`treat` treatment. A factor with levels `Control` `HMP` `P2O7`

`index` a numeric vector with the tartar index

Examples

```
data(tartar)
```

thumbtack *Throwing thumbtacks*

Description

A brass thumbtack was thrown 100 times and it was registered whether the pin was pointing up or down towards the table upon landing.

Usage

```
data(thumbtack)
```

Format

The format is: int [1:100] 1 1 0 0 1 1 0 1 0 0 ...

Details

1 corresponds to "tip pointing down" and 0 corresponds to "tip pointing up"

References

Mats Rudemo (1979). Statistik og sandsynlighedslaere med biologiske anvendelser. Del 1: Grundbegreber.

Examples

```
data(thumbtack)
mean(thumbtack)
```

`turtles` *Clutch size of turtles*

Description

Data to examine the effect of turtle carapace length on the clutch size of turtles.

Usage

```
data(turtles)
```

Format

A data frame with 18 observations on the following 2 variables.

`length` a numeric vector

`clutch` a numeric vector

Source

K. G. Ashton and R. L. Burke and J. N. Layne (2007). Geographic variation in body and clutch size of gopher tortoises. Copeia. 49:355-363.

Examples

```
data(turtles)
## maybe str(turtles) ; plot(turtles) ...
```

vitamina *Food intake for Danish people 1985*

Description

The daily food intake was studied for 2224 subjects, and the content of many different vitamins and substances were meaured,

Usage

```
data(vitamina)
```

Format

A data frame with 2224 observations on the following 20 variables.

`person` subject id (a numeric vector)

`wt` weight (kg)

`ht` height (cm)

`sex` sex: 1 for male, 2= for female

`age` age

`bmr` basal metabolic rate

`E_bmr` energy divided by bmr

`energi` energy content (kJ)

`Avit` vitamin A (RE)

`retinol` retinol (microgram)

`betacar` beta-caroten (microgram)

`Dvit` vitamin D (microgram

`Evit` vitamin E (alphaTE)

`B1vit` vitamin B1 (milligram)

`B2vit` vitamin B2 (milligram)

`niacin` niacin (NE)

`B6vit` vitamin B6 (milligram)

`folacin` folacin (microgram)

`B12vit` vitamin B12 (microgram)

`Cvit` vitamin C (milliggram)

Details

Only variables Avit and bmr are used in the isdals book.

Source

J. Haraldsdottir, J.H. Jensen, A. Moller (1985). Danskernes kostvaner 1985, Hovedresultater. Levnedsmiddelstyrelsen, publikation nr.\ 138.

Examples

```
data(vitamina)
```

Index