Package 'PBImisc'

January 20, 2025

Version 1.0

Type Package

Title A Set of Datasets Used in My Classes or in the Book 'Modele Liniowe i Mieszane w R, Wraz z Przykladami w Analizie Danych'

Author Przemyslaw Biecek <przemyslaw.biecek@gmail.com>

Maintainer Przemyslaw Biecek <przemyslaw.biecek@gmail.com>

Description A set of datasets and functions used in the book 'Modele liniowe i mieszane w R, wraz z przykladami w analizie danych'. Datasets either come from real studies or are created to be as similar as possible to real studies.

Repository CRAN

License GPL (>= 2)

LazyLoad yes

LazyData yes

URL http://www.biecek.pl/R/

Depends R (>= 2.8.0)

Imports lme4, Matrix

Suggests ggplot2, ca, lattice

NeedsCompilation no

Date/Publication 2016-02-25 22:46:59

Contents

`PBImisc-package` *Set of supplementary datasets and functions*

Description

A set of datasets and functions used in the book „Modele liniowe i mieszane w R, wraz z przykładami w analizie danych”

Details

Package:	PBImisc
Type:	Package
Version:	1.0
Date:	2016-02-15
License:	GPL-2

General Description

A set of datasets some of them are my original ones, some are taken from other packages of literature.

Author(s)

Przemyslaw Biecek

Maintainer: You should complain to Przemyslaw Biecek <przemyslaw.biecek@gmail.com>

References

Przemyslaw Biecek „Modele liniowe i mieszane w R, wraz z przykladami w analizie danych" 2013, Wydawnictwo PWN

Examples

```
# here you will find some examples
#
```

AML *Acute myeloid leukemia AML study*

Description

This dataset bases on blood samples for patients with Acute myeloid leukemia.

Usage

```
data(AML)
```

Format

data.frame with 66 obs. and 5 variables

`Mutation` Factor w/ 4 levels CBFbeta, FLT3, None, Other

`CD14.control` CD14 level in the control group

`CD14.D3` CD14 level after D3 treatment

`CD14.1906` CD14 level after D3 homolog 1906 treatment

`CD14.2191` CD14 level after D3 homolog 2191 treatment

Details

Mutation - mutated gene that causes leucemia, one of following CBFbeta, FLT3, None, Other CD14.control, CD14.D3, CD14.1906, CD14.2191 - effects in vitamin D3 or its homologues

Source

Artificial dataset generated to be consistent with Ewa M. study

Examples

```
library(lattice)
data(AML)
AML2 = reshape(AML, direction="long", varying=colnames(AML)[2:5])
bwplot(CD14~time|Mutation, AML2)
interaction.plot(AML2$time,AML2$Mutation, AML2$CD14)
```

---

apartments *Apartment prices in Warsaw in years 2007-2009*

Description

Dataset downloaded from website http://www.oferty.net/. Dataset contains offer and transictional prices for apartments sold in in Warsaw in years 2007-2009.

Usage

```
data(apartments)
```

Format

data.frame with 973 obs. and 16 variables

`year` data year of the transaction

`month` data month of the transaction

`surface` apartment area in m2

`city` city (all transactions are from Warsaw)

`district` district in which the apartment is located, factor with 28 levels

`street` steet in which the apartment is located

`n.rooms` number of rooms

`floor` floor

`construction.date` the construction year

`type` ownership rights

`offer.price` price in the offer

`transaction.price` declared price in the transaction

`m2.price` price per m2

`condition` apartment condition, factor with 5 levels

`lat`, `lon` latitude and longitude coordinates for district center

Details

This and other related dataset you may find here http://www.oferty.net/.

Source

website http://www.oferty.net/

Examples

```
data(apartments)
library(lattice)
xyplot(m2.price~construction.date|district, apartments, type=c("g","p"))

#
# apartments2 = na.omit(apartments[,c(13,1,3,5,7,8,9,10,14,15,16)])
# wsp = (bincombinations(10)==1)[-1,]
# params = matrix(0, nrow(wsp), 3)
# for (i in 1:nrow(wsp)) {
#    model = lm(m2.price~., data=apartments2[,c(TRUE,wsp[i,])])
#    params[i,1] = AIC(model, k=log(nrow(apartments2)))
#     params[i,2] = model$rank
#     params[i,3] = summary(model)$adj.r.squared
# }
# plot(params[,2], params[,3], xlab="no. of regressors", ylab="adj R^2")
#
```

---

boxplotpp *boxplot plus plus*

Description

boxplotpp

Usage

```
boxplotpp(x, xname=seq(1:ncol(x)), utitle="", addLines=TRUE,
           color = ifelse(addLines, "white","lightgrey"), ...)

boxplotInTime(x, xname, additional=T, color = ifelse(additional,
     "white","lightgrey"), main="", ylim=range(unlist(x),na.rm=T), ...,
     points = dim(x)[2], at = 1:points)
```

Arguments

x	TODO
xname	TODO
utitle	TODO
addLines	TODO
color	TODO
additional	TODO

main	TODO
points	TODO
at	TODO
ylim	TODO
...	TODO

Details

TODO

Value

TODO

Author(s)

Przemyslaw Biecek

Examples

```
#TODO
```

---

corn *A datasets relatead to gene expression in corn*

Description

Dataset from the book „Modele liniowe i mieszane w R, wraz z przykladami w analizie danych".

Usage

```
data(corn)
```

Format

data.frame with 5339 obs. and 36 variables

A dataset with expression of 5339 genes. Each column corresponds to a single experiment. Column name codes the setup of experiment. For example DH.C.1 is related to line DH in the condition C and it is a first technical replicate of this set of conditions.

Note that a noise injection was added to this data, in order to obtain the original dataset please contact with the package maintainer.

Details

Dataset from the book „Modele liniowe i mieszane w R, wraz z przykladami w analizie danych".

Used as an example of modeling of data from expression microarrays with the use of models with mixed effects.

Examples

```
## Not run:

require(lme4)

names <- colnames(corn)
X <- t(matrix(unlist(strsplit(names, ".", fixed=T)), 3, 36))
X <- data.frame(X)
colnames(X) <- c("spec", "temp", "plant")

summary(X)

y <- corn[4662,]
lmer(y~spec*temp + (1|plant:spec:temp), data=X)


## End(Not run)
```

`dementia` *A set of datasets relatead to dementia*

Description

Dataset from the book „Modele liniowe i mieszane w R, wraz z przykladami w analizie danych".

Usage

```
data(dementia)
```

Format

data.frame with 1000 obs. and 4 variables

`demscore` score of dementia

`age` age, a factor with two levels

`sex` sex, a factor with two levels

`study` a source of data, a factor with 10 levels

Details

Dataset from the book „Modele liniowe i mieszane w R, wraz z przykladami w analizie danych".

Used as an example of mixed modeling in meta analysis.

Examples

```
## Not run:
  require(lme4)
  modelFullI <- lmer(demscore~age*sex+(age*sex|study), data=dementia,
                  REML=FALSE)
  summary(modelFullI)

## End(Not run)
```

`Drosophila` *Drosophila datasets and QTL mapping study*

Description

Two datasets with genotypes and phenotypes for backcrossed Drosophilas.

Usage

```
data(Drosophila)
```

Format

Two datasets with genotypes and phenotypes for backcrossed Drosophilas.

The set of 41 markers describes genotypes while 5 variables describe phenotypes. See references for more details.

`bm` A data.frame with 370 obs. and 46 variables, first 41 are genotypes of gene markers, last five describes genotypes

`bs` A data.frame with 402 obs. and 46 variables, first 41 are genotypes of gene markers, last five describes genotypes

`chr` Factor w/ 4 levels CBFbeta, FLT3, None, Other

`pos` Markers position on chromosom in centimorgnas

Details

The phonotype pc1 is nicely described by genotype in both backcrossed datasets.

Source

Genetic Architecture of a Morphological Shape Difference Between Two Drosophila Species Zhao-Bang Zenga, Jianjun Liu, Lynn F. Stamb, Chen-Hung Kao, John M. Mercer, Cathy C. Laurie Genetics, Vol. 154, 299-310, January 2000

Examples

```
data(Drosophila)
library(lattice)
# calculate log likelihoods
pval1 = numeric(41)
for (i in 1:41) {
  y = Drosophila$bm$pc1
  x = factor(Drosophila$bm[,i])
  pval1[i] = logLik(lm(y~x))
}
# loglikelihood plot
xyplot(pval1~pos|chr, data=Drosophila, type=c("p","l"),
     pch=19, ylab="log likelihood")
```

---

ecap *Epidemiology of Allergic Disease in Poland*

Description

This dataset touch one particular aspect from ECAP dataset. The original dataset is much more richer.

Usage

```
data(ecap)
```

Format

data.frame with 2102 obs. and 9 variables

`city`, `district` City and district, city is a factor with nine levels, the district effect is nested in the city effect

`sex` Sex

`weight`, `height` Weight and height

`house.surface` Surface of house in which the pearson live

`PNIF` Peak Nasal Inspiratory Flow

`age` Age of the pearson

`allergenes` Number of allergens

Details

PNIF stands for Peak Nasal Inspiratory Flow

Source

Artificial dataset generated to be consistent with ECAP (Epidemiologia Chorob Alergicznych w Polsce) study http://www.ecap.pl/

Examples

```
data(ecap)
library(lattice)
xyplot(PNIF~age|city, data=ecap, type=c("p","g","smooth"))
```

---

eden *European day hospital evaluation*

Description

This dataset bases on origical study of European day hospital evaluation

Artificial dataset (subset from real dataset with some random modifications). Do not use it for derivation of real conclusions.

Usage

```
data(eden)
```

Format

data.frame with 642 obs. and 12 variables

`mdid` Medical doctor id, there are 24 different MDs which examine patients

`center` City in which the examination takes place

`BPRS.Maniac`, `BPRS.Negative`, `BPRS.Positive`, `BPRS.Depression` BPRS stands for Brief Psychiatric Rating Scale, scores are averaged in four subscales

`BPRS.Average` Average from 24 questions

`MANSA` Scale which measures Quality of Life (Manchester Short Assessment of Quality of Life)

`sex` Sex

`children` Number of childs

`years.of.education` Number of years of education

`day` Hospitalization mode, day or stationary

Details

This dataset touch one particular aspect from EDEN dataset. The original dataset is much more richer.

Source

Artificial dataset generated to be consistent with Joanna R. study.

Bases on European day hospital evaluation, http://www.edenstudy.com/

Examples

```
data(eden)
library(lattice)
xyplot(BPRS.Average~MANSA|center, data=eden, type=c("p","g","smooth"))
```

---

elastase *Relation between graft function and elastase*

Description

Relation between graft function and elastase from nephrology study.

Usage

```
data(elastase)
```

Format

data.frame with 54 obs. and 5 variables

`sex`, `age`, `weight` Patient's sex, age and weight

`elastase` Elastase concentration

`GFR` Patient's GFR (glomerular filtration rate)

Details

Artificial dataset (real one with some random modifications). Do not use it for medical reasoning.

Source

Artificial dataset generated to be consistent with Malgorzata L. study

Examples

```
data(elastase)
library(lattice)
xyplot(GFR~elastase, data=elastase, type=c("p","r","g"))
```

endometriosis *Endometriosis study*

Description

How the endometriosis affects concetration of alpha and beta factors in the blood.

Usage

```
data(endometriosis)
```

Format

data.frame with 165 obs. and 4 variables

`disease` disease, blood samples were taken from women with endometriosis of from healthy ones

`phase` phase in the menstrual cycle as the examination day (proliferative or secretory)

`alpha.factor`, `beta.factor` concentration of alpha and beta factors in blood

Details

Dataset used as example of ANCOVA

Source

Artificial dataset generated to be consistent with Ula S. study

Examples

```
data(endometriosis)
library(lattice)
xyplot(log(alpha.factor)~log(beta.factor)|disease*phase,
            data=endometriosis, type=c("p", "r"))
summary(aov(alpha.factor~beta.factor*disease*phase, data=endometriosis))
```

eunomia *European Evaluation of Coercion in Psychiatry and Harmonisation of Best Clinical Practise*

Description

This dataset touch one particular aspect from EUNOMIA dataset. The original dataset is much more richer.

Usage

```
data(eunomia)
```

Format

data.frame with 2008 obs. and 15 variables

`CENTRE13` Center in which the patient is hospitalized, factor with 13 levels

`SUBJECT` Patients ID

`GENDER`, `AGE`, `NUM.HOSP` Gender, age and number of hospitalizations of given patient

`CAT.T1`, `CAT.T2`, `CAT.T3` Clients Scale for Assessment of Treatment, short assessment, which measures the impact of COPD on a patients life, measured in times: T1, T2 and T3

`BPRS.T1`, `BPRS.T2`, `BPRS.T3` Average score for Brief Psychiatric Rating Scale, measured in times: T1, T2 and T3

`MANSA.T1`, `MANSA.T2`, `MANSA.T3` Scale which measures Quality of Life (Manchester Short Assessment of Quality of Life), measured in times: T1, T2 and T3

`ICD10` International Statistical Classification of Diseases and Related Health Problems 10th Revision (ICD-10)

Details

Artificial dataset generated to be consistent with Eunomia study (European Evaluation of Coercion in Psychiatry and Harmonisation of Best Clinical Practise)

Source

Artificial dataset generated to be consistent with Joanna R. study.

Eunomia dataset, http://www.eunomia-study.net/

Examples

```
data(eunomia)
library(lattice)
bwplot(CENTRE13~BPRS.T1, data=eunomia)
xyplot(BPRS.T1~MANSA.T1|CENTRE13, data=eunomia, type=c("p","g","smooth"))
```

---

flu *Numbers of flu occurences in the 10 years period in the Poland.*

Description

Data from National Institute of Hygiene reports. Each row correspond to one record from NIH institute.

Usage

```
data(flu)
```

Format

data.frame with 6384 obs. and 11 variables

`region` Region for which given report was taken. A factor with 16 levels

`inception.no` Number of flu occurences in given region for given report period (one or two weeks)

`inception.no` Number of flu occurences in given region for given report period (one or two weeks)

`inception.rate` Number of flu occurences normalized to 100k people

`inception.no.0-14`, `inception.no.15+`, `inception.rate.0-14`, `inception.rate.15+` Absolute and normalized numbers of flu occurences calculated for age group 0-14 or 15+

`date` Date of given report

`date.id` Report id, there is 38 reports per year

`latitude`, `longitude` Geographical coordinates for region

Details

Dataset used during the third edition of WZUR conference, see http://www.biecek.pl/WZUR3/wzurDane.html for more information.

Source

Reports from National Institute of Public Health - National Institute of Hygiene, see: http://www.pzh.gov.pl

More information: http://www.biecek.pl/WZUR3/wzurDane.html

Examples

```
data(flu)
library(ggplot2)
subflu = flu[flu$region=="Mazowieckie", ]
# linear scale
qplot(date, inception.rate,data=subflu, geom="line")+scale_y_sqrt() +theme_bw()

# polar coordinates
qplot(1 + date.id*12/38, inception.rate,data=subflu, geom="path", xlab="month")+
        scale_y_sqrt()+geom_smooth(span=0.1,se=FALSE, size=2,col="red") +
        coord_polar() +theme_bw()
```

---

genomes *724 bacterial genomes data*

Description

Few parameters gathered for 724 bacterial species.

Usage

```
data(genomes)
```

Format

data.frame with 724 obs. and 7 variables

`organism` Organism name, unique value for every row

`group` Group, a factor with 22 levels

`size` Genome size in Mbp

`CG` GC content for genome sequence

`habitat`, `temp.group`, `temperature` Where does this bacteria live?

Details

This dataset is prepared by Pawel M., data are taken from NCBI repository.

See http://www.ncbi.nlm.nih.gov/ for more details

Source

Pawel M. study

Examples

```
data(genomes)
library(ggplot2)
# is this relation linear ?
qplot(size,GC, data=genomes) + theme_bw()
# or linear in log scales?
qplot(size,GC, data=genomes, log="xy") + theme_bw()
```

---

heights *Husband and Wife heights*

Description

A dataset from „A modern approach to regression with R". Simon J. Sheather 2009 . Paired heights for husbands and wifes.

Usage

```
data(heights)
```

Format

data.frame with 96 obs. and 2 variables

`Husband`, `Wife` Height of husband and wife.

Details

The dataset from „A modern approach to regression with R". Simon J. Sheather 2009

Source

A modern approach to regression with R. Simon J. Sheather 2009

Examples

```
data(heights)
plot(Husband~Wife, data=heights, pch=19)
abline(lm(Husband~Wife, data=heights), col="red")
abline(lm(Husband~Wife-1, data=heights), col="blue")
```

---

histpp *hist plus plus*

Description

histpp

Usage

```
histpp(x, xname="", utitle="")
```

Arguments

x	TODO
xname	TODO
utitle	TODO

Details

TODO

Value

TODO

Author(s)

Przemyslaw Biecek

References

TODO

Examples

```
# TODO
```

kidney *Graft function after kidney transplantation*

Description

Artificial dataset (subset from real dataset with some random modifications)

Usage

```
data(kidney)
```

Format

data.frame with 334 obs. and 16 variables

`recipient.age`, `donor.age` Age od donor and recipient

`CIT` Cold ischemia time

`discrepancy.AB`, `discrepancy.DR` discrepancies in AB and DR antibodies

`therapy` scheme of immunosuppression

`diabetes` diabetes

`bpl.drugs` number of drugs for blood pressure lowering

`MDRD7`, `MDRD30`, `MDRD3`, `MDRD6`, `MDRD12`, `MDRD24`, `MDRD36`, `MDRD60` MDRD (Modification of Diet in Renal Disease) as a estiamtor of glomerular filtration rate (GFR) from serum creatinine, measured 7, 30 days and 3, 6, 12, 24, 36 and 60 months after kidney transplantation

Details

Example of longitudinal study, note that graft for all patients survives 5 years after kidney transplantation.

Source

Artificial dataset generated to be consistent with Maria M. study

Examples

```
data(kidney)
boxplotInTime(kidney[,9:16], colnames(kidney[,9:16]), additional=TRUE)
```

`Likelihood displacements` *Log-likelihood displacements for single observation and single grouping variable*

Description

Functions for log-likelihood displacements for each observation or each level of given factor

Usage

```
recalculateLogLik(model, fixef = fixef(model), vcor = VarCorr(model))

groupDisp(formula, data, var)

obsDisp(formula, data, inds=1:nrow(data))
```

Arguments

`model`	a mixed model of the class mer,
`fixef`, `vcor`	model parameters log-likelihood evaluation, if not provided then the estimates extracted from the 'model' parameter will be used
`formula`	a model formula that will be passes to the nlme function
`data`	a data frame
`var`	a name of grouping variable (factor) for which the group log-likelihood displacement will be performed
`inds`	indexes of observations for which observation log-likelihood displacement will be performed

Details

Likelihood displacement is defined as a difference of likelihoods calculated on full dataset for two models with different sets of parameters. The first model is a model with ML estimates obtained for full dataset, while the second model is a model with ML estimates obtained on dataset without a selected observation or group of observations.

Likelihood displacements are used in model diagnostic.

Note that these functions reestimate coefficients in a set of model may be a time consuming.

The function recalculateLogLik() calculated a log-likelihood for model defined by the object model and model parameters defined in following function arguments.

The functions groupDisp() and obsDisp() calculates how the log-likelihood will decrees if selected groups or selected observations will not be used for parameter estimates. Note that log-likelihood is calculated on full dataset.

Author(s)

Przemyslaw Biecek

Examples

```
data(eunomia)
require(lme4)
set.seed(1313)
eunomias <- eunomia[sample(1:2000,100),]
groupDisp(formula = BPRS.T2~ (1|CENTRE13), data=eunomias, var="CENTRE13")

obsDisp(formula = BPRS.T2~ (1|CENTRE13), data=eunomias, inds = 1:10)

obsDisp(formula = BPRS.T2~ (1|CENTRE13), data=eunomias)
```

---

milk *Milk yield data*

Description

Milk yield data for 10 unrelated cows

Usage

```
data(milk)
```

Format

data.frame with 40 obs. and 2 variables

`cow` cow id, a factor with 10 levels

`milk.amount` milk amount in kgs per week

Details

Weekly milk yield amount for 10 cows. For every cow 5 measurements are taken.

Examples

```
data(milk)
library(lattice)
# change the order of levels
milk$cow = reorder(milk$cow, milk$milk.amount, mean)
#plot it
dotplot(cow~milk.amount, data=milk)
```

`milkgene` *Mutation in BTN3A1 gene and milk yield*

Description

It is known that BTN3A1 (Butyrophilin subfamily 3 member A1) has a crucial function in the secretion of lipids into milk. Doeas the SNP mutation in it change the average milk yield?

Usage

```
data(milkgene)
```

Format

data.frame with 1000 obs. and 5 variables

`cow.id` cow id, there is 465 cows in this study

`btn3a1` btn3a1 genotype, a factor with two levels

`lactation` for some cows there are milk yileds for four lactations for other only for the first one

`milk`, `fat` milk and fat amount in kgs per lactation

Details

Milk and fat yields for 465 cows. For every cow also the genotype of btn3a1 is measured.

Source

Artificial dataset generated to be consistent with Joanna Sz. study

Examples

```
data(milkgene)
library(lattice)
xyplot(milk~fat, data=milkgene)
bwplot(milk~lactation, data=milkgene)
```

musculus *A dataset relatead to mice musculus growth which depends on diet and genetic structure*

Description

Dataset from the book „Modele liniowe i mieszane w R, wraz z przykladami w analizie danych".

Usage

```
data(musculus)
```

Format

data.frame with 30 obs. and 10 variables

`id` an individual id

`dadid` id of father, 0 for founders

`momid` id of mother, 0 for founders

`sex` sex

`sigma` maximal stress

`diet` diet, D1 or D2

`k1` resilience coefficient in point 1

`k2` resilience coefficient in point 2

`E1` Younga module in point 1

`E2` Younga module in point 2

Details

Dataset from the book „Modele liniowe i mieszane w R, wraz z przykladami w analizie danych".

Used as an example of model with mixed effects where random effects have know dependency structure, here related to the kinship coefficient.

Examples

```
## Not run:
   require(kinship2)
   pedmus <- pedigree(musculus$id, musculus$dadid, musculus$momid, musculus$sex)
   plot(pedmus, affected=musculus$diet)

   fam  <- makefamid(musculus$id, musculus$dadid, musculus$momid)
   kmatrix <- makekinship(fam, musculus$id, musculus$dadid, musculus$momid)
   kmatrix[1:5,1:15]

## End(Not run)
```

`Plot results from post hoc testing`

A function for visual representation of pairwise testing (both for pairwise.t.test and pairwise.wilcox.test)

Description

Plot sets of groups in which means of medians are not significantly different.

On the veritical axis the means are marked. Then in a greedy fashion means that are not significantly different are linked by a line.

Usage

```
plotPairwiseTests(p.vals, means, alpha=0.05, digits=3, mar=c(2,10,3,1), ...)
```

Arguments

`p.vals`	A slot `$p.value` from the object returned by pairwise.*.test function
`means`	A vector of means or medians corresponding to p.vals object (the order of groups should be the same in both objects)
`alpha`	A threshold for p.value
`digits`	Number of significant digits to be ploted with means.
`mar`	Figure margins, left margin should be large enought to handle names of groups
`...`	These arguments are passed to the plot function.

Author(s)

Przemyslaw Biecek

Examples

```
data(iris)
tmp1 <- pairwise.wilcox.test(iris$Sepal.Width, iris$Species)
tmp2 <- tapply(iris$Sepal.Width, iris$Species, median, na.rm=TRUE)
plotPairwiseTests(tmp1$p.value, tmp2, alpha=0.001)
```

schizophrenia *Genetic backgroud of schizophrenia*

Description

Dataset with genotypes and phenotypes for 98 patients with schizophrenia disorder.

Usage

```
data(schizophrenia)
```

Format

data.frame with 98 obs. and 9 variables

`NfkB`, `CD28`, `IFN` Genotypes for SNP mutations in selected three genes

`Dikeos.manic`, `Dikeos.reality.distortion`, `Dikeos.depression`, `Dikeos.disorganization`, `Dikeos.negative` Dikeos scores for schizophrenia measured in five domains

`Dikeos.sum` Sum of Dikeos scores

Details

Alleles for two SNPs in genes: Nuclear Factor-Kappa Beta (NfkB) and Cluster of Differentiation 28 (CD28) were examined as well as mental health described by five scales (see Dikeos 2008 for more details).

Source

Artificial dataset generated to be consistent with Dorota F. study

Examples

```
data(schizophrenia)
attach(schizophrenia)
interaction.plot(CD28, NfkB, Dikeos.sum)
interaction.plot(NfkB, CD28, Dikeos.sum)
model.tables(aov(Dikeos.sum~NfkB*CD28))
```

score.cardiovascular *SCORE for Cardiovascular Risk*

Description

Calculation of risk SCORE for use in the clinical management of cardiovascular risk in European.

Usage

```
calculateScoreEur(age, cholesterol, SBP, currentSmoker,
  gender = "Men", risk = "Low risk")
```

Arguments

age	age in years
cholesterol	in mmol/L
SBP	Systolic blood pressure in mmHg
currentSmoker	the current smoker status, 1 for current smokers, 0 for non smokers
gender	"Men" or "Women"
risk	is it "Low risk" or "High risk" group

Details

Calculation of SCORE based on the paper

„Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project" R.M. Conroy et. al. Eur Heart J (2003) 24 (11): 987-1003. doi: 10.1016/S0195-668X(03)00114-3

Author(s)

Przemyslaw Biecek

SejmSenat *SejmSenat*

Description

Changes in word usage in consecutive Sejm and Senate cadencies

Usage

```
data(SejmSenat)
```

Format

contingency matrix with 973 27 rows and 8 columns

`Sejm.I`, `Sejm.II`, `Sejm.III`, `Sejm.IV`, summary of records from four Sejm cadencies

`Senat.II`, `Senat.III`, `Senat.IV`, `Senat.V`, summary of records from four Senate cadencies

`adj`, `adja`, `adjp`, `adv`, `aglt`, `bedzie`,`conj`, `depr`, `fin`, `ger`, `ign`, `imps`, `impt`, `inf`, `interp`,`num`, `pact`, `pant`, `pcon`, `ppas`, `pra`
word modes

Details

Word usage statistics generated from Sejm and Senat records

Source

The IPI PAN Corpus webpage http://korpus.pl/

Examples

```
data(SejmSenat)
library(ca)
# can you see some patterns?
plot(ca(SejmSenat[-15,]), mass =c(TRUE,TRUE), arrows =c(FALSE,TRUE))
```

`vaccination` *Effective dose study*

Description

What is the minimal dose that is effective?

Usage

```
data(vaccination)
```

Format

data.frame with 100 obs. and 2 variables

`response` a reaction effect

`dose` a dose that was applied

Details

Responses for different doses of treatment.

Source

Artificial dataset generated to be consistent with Karolina P. study

Examples

```
data(vaccination)
library(lattice)
bwplot(response~dose, data=vaccination)
```

YXZ *Artificial dataset which shows the differences between tests type I and III (sequential vs. marginal)*

Description

Artificial dataset, shows inconsistency for test type I and III

Usage

```
data(YXZ)
```

Format

data.frame with 100 obs. and 3 variables

`X, Z` explanatory variables

`Y` response variable

Details

See the example, results for staistical tests are inconsistet due to correlation between X and Z variables

Source

Artificial dataset, generated by PBI

Examples

```
attach(YXZ)
summary(lm(Y~X+Z))
anova(lm(Y~Z+X))
anova(lm(Y~X))
anova(lm(Y~Z))
```

Index