Package ‘TRSbook’

May 10, 2018

Type  Package
Title  Functions and Datasets to Accompany the Book "The R Software: Fundamentals of Programming and Statistical Analysis"
Version  1.0.2
Date  2018-05-10
Author  Lafaye de Micheaux Pierre, Drouilhet Remy, Liquet Benoit
Maintainer  Pierre Lafaye de Micheaux <lafaye@unsw.edu.au>
Description  Functions and datasets for the reader of the book "The R Software: Fundamentals of Programming and Statistical Analysis".
License  GPL (>= 2)
LazyLoad  yes
Depends  xtable, RColorBrewer, gdata, IndependenceTests
NeedsCompilation  yes
Repository  CRAN
Date/Publication  2018-05-10 12:17:41 UTC

R topics documented:

  TheRSoftware-package ........................................... 2
  arrowaxis .......................................................... 2
  barchart ........................................................... 3
  bin2dec ............................................................. 4
  BIRTH.WEIGHT ..................................................... 5
  BMI.CHILD .......................................................... 6
  camembert .......................................................... 7
  cor0.test ............................................................ 8
  crosschart ........................................................... 9
  dec2bin ............................................................. 10
  flashy.plot .......................................................... 10
  getaddr ............................................................. 11
  INFARCTION ......................................................... 12
  INTIMA.MEDIA ....................................................... 14

**Description**

This function add an arrow on the extremities of the axes of a plot.

**Details**

- **Package:** TheRSoftware
- **Type:** Package
- **Version:** 1.0
- **Date:** 2014-02-04
- **License:** GPL(>=2.0)
- **LazyLoad:** yes

**Author(s)**

Lafaye de Micheaux Pierre <lafaye@unsw.edu.au>, Remy Drouilhet <Remy.Drouilhet@upmf-grenoble.fr>, Liquet Benoit <b.liquet@uq.edu.au>

**References**

barchart

Usage
arrowaxis(x = TRUE, y = TRUE)

Arguments
x Logical. Default value TRUE indicates an arrow on the x-axis
y Logical. Default value TRUE indicates an arrow on the y-axis

Author(s)
Lafaye de Micheaux Pierre <lafaye@unsw.edu.au>, Remy Drouilhet <Remy.Drouilhet@upmf-grenoble.fr>, Liquet Benoit <b.liquet@uq.edu.au>

References
Chapter 11 (Descriptive Statistics) from the book: The R Software, Fundamentals of Programming and Statistical Analysis

Examples
curve(cos(x), xlim = c(-10, 10))
arrowaxis()

barchart

Bar charts

Description
Pretty bar charts

Usage
barchart(x, col, my.title, pareto = FALSE, freq.cumul = FALSE, family = "Courier")

Arguments
x qualitative variable
col vector of characters for the color of each modality
my.title character. Title of the plot
pareto logical. TRUE for a Pareto diagram. Default os FALSE
freq.cumul logical. TRUE to add a curve of cumulative frequencies. By default freq.cumul is FALSE
family font family for the title. Default is "Courier". Another choice can be, e.g., "HersheyScript"
**bin2dec**

**Description**

To compute the decimal representation of a number written in a binary format

**Usage**

```r
bin2dec(x)
```

**Arguments**

- `x` Numeric. Number in binary format written only with 0s and 1s. See Example below.

**Value**

Decimal representation of the number `x`
Author(s)

Lafaye de Micheaux Pierre <lafaye@unsw.edu.au>, Remy Drouilhet <Remy.Drouilhet@upmf-grenoble.fr>, Liquet Benoit <b.liquet@uq.edu.au>

References

Chapter 5 (Data Manipulation, Functions) from the book: The R Software, Fundamentals of Programming and Statistical Analysis

Examples

bin2dec(10100.101)

Description

This study focused on risks associated with low weight at birth; the data were collected at the Baystate Medical Centre, Massachusetts, in 1986. Physicians have been interested in low weight at birth for several years, because underweight babies have high rates of infant mortality and infant anomalies. The behaviour of the mother-to-be during pregnancy (diet, smoking habits) can have a significant impact on the chances of having a full-term pregnancy, and thus of giving birth to a child of normal weight. The data file includes information on 189 women (identification number: ID) who came to the centre for consultation. Weight at birth is categorized as low if the child weighs less than 2,500 g.

Usage

data(BIRTH.WEIGHT)

Format

A data frame with 189 observations measured on the following 11 variables.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>id</td>
<td>Numeric. Identification.</td>
</tr>
<tr>
<td>age</td>
<td>Numeric. Age of mother.</td>
</tr>
<tr>
<td>lwt</td>
<td>Numeric. Weight of mother at last menstrual period.</td>
</tr>
<tr>
<td>race</td>
<td>1=white, 2=black, 3=other. Race of mother.</td>
</tr>
<tr>
<td>smoke</td>
<td>Yes=1, No=0. Smoking during pregnancy.</td>
</tr>
<tr>
<td>ptl</td>
<td>0=none, 1=one, 2=two, etc. Number of premature births in medical history.</td>
</tr>
<tr>
<td>ht</td>
<td>Yes=1, No=0. Medical history of hypertension.</td>
</tr>
<tr>
<td>ui</td>
<td>Yes=1, No=0. Uterine irritability.</td>
</tr>
<tr>
<td>fvt</td>
<td>0=none, 1=one, etc. Number of medical consultations during first trimester</td>
</tr>
<tr>
<td>bwt</td>
<td>Numeric. Grams.</td>
</tr>
<tr>
<td>low</td>
<td>Yes=1, No=0. Weight at birth less than 2,500 g</td>
</tr>
</tbody>
</table>
**Author(s)**

Lafaye de Micheaux Pierre <lafaye@unsw.edu.au>, Remy Drouilhet <Remy.Drouilhet@upmf-grenoble.fr>, Liquet Benoit <b.liquet@uq.edu.au>

**Source**

http://www.biostatisticien.eu/springeR/

**Examples**

```r
data(BIRTH.WEIGHT)
str(BIRTH.WEIGHT)
```

---

**BMI.CHILD**  
*Body Mass Index of children*

---

**Description**

This data set comes from an epidemiologic study analyzed by a team from the Institut de sante publique d'épidemiologie et de developpement (ISPED) de Bordeaux. A sample of 152 children (3 or 4 years old) in their first year of kindergarten in schools in Bordeaux (Gironde, SouthWest France) underwent a physical check-up in 1996-1997.

**Usage**

```r
data(BMI.CHILD)
```

**Format**

A data frame with 152 observations measured on the 6 following variables:

- **gender**  a factor with levels F and M
- **zep**  a factor with levels Y and N
- **weight** numeric
- **years** numeric
- **months** numeric
- **height** numeric

**Author(s)**

Lafaye de Micheaux Pierre <lafaye@unsw.edu.au>, Remy Drouilhet <Remy.Drouilhet@upmf-grenoble.fr>, Liquet Benoit <b.liquet@uq.edu.au>

**Source**

http://www.biostatisticien.eu/springeR/
Examples

data(BMI.CHILD)
str(BMI.CHILD)

camembert  Pie chart

Description

A variant of the pie function

Usage

camembert(x, col = NULL, family="Courier")

Arguments

x  qualitative variable
col  vector of characters for the color of each modality
family  font family for the title. Default is "Courier". Another choice can be, e.g., "HersheyScript"

Value

A pie chart

Author(s)

Lafaye de Micheaux Pierre <lafaye@unsw.edu.au>, Remy Drouilhet <Remy.Drouilhet@upmf-grenoble.fr>, Liquet Benoit <b.liquet@uq.edu.au>

References

Chapter 11 (Descriptive Statistics) from the book: The R Software, Fundamentals of Programming and Statistical Analysis

See Also

pie

Examples

data(NUTRIELDERLY)
attach(NUTRIELDERLY)
require("RColorBrewer")
col <- brewer.pal(8,"Pastel2")
camembert(fat,col)
detach(NUTRIELDERLY)
cor0.test

Test of the correlation coefficient

Description
Test of the correlation coefficient between two quantitative variables

Usage

cor0.test(x, y, rho0 = 0, alternative = c("two.sided", "less", "greater"))

Arguments

x numeric vector
y numeric vector
rho0 numeric indicating the value of the correlation coefficient under the null. Default is rho0 = 0
alternative Alternative hypothesis for the test. Either two sided ("two.sided"), one sided to the left ("less") or one sided to the right ("greater"). Default is "two.sided".

Value
Returns a list:

statistic Value of the test statistic
p.value p-value of the test

Author(s)
Lafaye de Micheaux Pierre <lafaye@unsw.edu.au>, Remy Drouilhet <Remy.Drouilhet@upmf-grenoble.fr>, Liquet Benoit <b.liquet@uq.edu.au>

References
Chapter 13 (Confidence Intervals and Hypothesis Testing) from the book: The R Software, Fundamentals of Programming and Statistical Analysis

See Also
cor.test

Examples

data(BMI.CHILD)
attach(BMI.CHILD)
cor0.test(weight, height)
detach(BMI.CHILD)
crosschart

A cross chart

Description

A cross chart displays for each observation a small cross above the associated modality.

Usage

crosschart(x, my.title, col, family = "Courier")

Arguments

x
qualitative variable

my.title
character. title of the plot

col
vector of characters for the color of each modality

family
font family for the title. Default is "Courier". Another choice can be, e.g., "HersheyScript"

Value

A cross chart

Author(s)

Lafaye de Micheaux Pierre <lafaye@unsw.edu.au>, Remy Drouilhet <Remy.Drouilhet@upmf-grenoble.fr>, Liquet Benoit <b.liquet@uq.edu.au>

References

Chapter 11 (Descriptive Statistics) from the book: The R Software, Fundamentals of Programming and Statistical Analysis

Examples

data(NUTRIELDERLY)
attach(NUTRIELDERLY)
situation <- as.factor(situation)
levels(situation) <- c("single","couple","family","other")
crosschart(situation, col=c("orange","darkgreen","black","tan"))
detach(NUTRIELDERLY)
**dec2bin**  
*Binary representation of a decimal number*

**Description**
To compute the binary representation of a number written in a decimal format.

**Usage**
```r
dec2bin(x, prec=52)
```

**Arguments**
- `x`: Numeric. Number in a decimal format.
- `prec`: Integer. Precision desired.

**Value**
Binary representation of the number `x`.

**Author(s)**
Lafaye de Micheaux Pierre <lafaye@unsw.edu.au>, Remy Drouilhet <Remy.Drouilhet@upmf-grenoble.fr>, Liquet Benoit <b.liquet@uq.edu.au>

**References**

**Examples**
```r
dec2bin(10.625, 3)
```

---

**flashy.plot**  
*A flashy scatter plot*

**Description**
This function tries to make a nicer plot than the one given by the `plot()` function for two quantitative variables.

**Usage**
```r
flashy.plot(x, y, my.factor, family = "Courier", xlab="", ylab="")
```
getaddr

Arguments

  x        numeric vector
  y        numeric vector
  my.factor factor
  family   font family for the title. Default is "Courier". Another choice can be, e.g., "HersheyScript"
  xlab     character: x label
  ylab     character: y label

Value

A flashy scatter plot

Author(s)

Lafaye de Micheaux Pierre <lafaye@unsw.edu.au>, Remy Drouilhet <Remy.Drouilhet@upmf-grenoble.fr>, Liquet Benoit <b.liquet@uq.edu.au>

References

Chapter 11 (Descriptive Statistics) from the book: The R Software, Fundamentals of Programming and Statistical Analysis

See Also

plot

Examples

```r
  data(NUTRIELDERLY)
  attach(NUTRIELDERLY)
  gender <- as.factor(gender)
  levels(gender) <- c("Male","Female")
  flashy.plot(weight,height,gender,xlab="Height",ylab="Weight")
  detach(NUTRIELDERLY)
```

Description

Retrieve the address in memory of a numeric variable

Usage

`getaddr(x)`
Arguments

- **x**: numeric

Value

Integer value of the address of x

Author(s)

Lafaye de Micheaux Pierre <lafaye@unsw.edu.au>, Remy Drouilhet <Remy.Drouilhet@upmf-grenoble.fr>, Liquet Benoit <b.liquet@uq.edu.au>

References

Chapter 9 (Managing Sessions) from the book: The R Software, Fundamentals of Programming and Statistical Analysis

Examples

```r
## Not run:
x <- c(8L,9L)
addr <- getaddr(x)$addr.int # Gets the address of the first
 # box of the 64-box block where x
 # is stored.
addr
writeaddr(addr,6L) # Write the integer 6 at this address.
x
writeaddr(addr+4L,7L) # An integer is coded over 4 bytes,
 # hence increment the address by 4 to
 # get to x[2].
x
x <- c(12.8,4.5)
x
addr <- getaddr(x)$addr.int # Get the address of the first box
 # of the 128-box block where x is
 # stored.
writeaddr(addr,6.2)
x
writeaddr(addr+8L,7.1) # A double is coded over 8 bytes.
x
## End(Not run)
```
Description

The study for which the following data were collected aimed at examining whether women who use or have used oral contraceptives are at a higher risk of myocardial infarction. The sample includes 149 women who had myocardial infarction (cases) and 300 women who did not (controls). The main exposure factor is usage of oral contraceptives; the data also include age, weight, height, tobacco consumption, hypertension and family history of cardiovascular diseases.

Usage

data(INFARCTION)

Format

A data frame with 449 observations measured on the following 10 variables:

- `infarct`: Identification. 0 = controls; 1 = cases. Myocardial infarction.
- `co`: Usage of oral contraceptives. 0 = never; 1 = yes.
- `tobacco`: Tobacco usage. 0 = no; 1 = smoker; 2 = former smoker.
- `age`: Age in years.
- `weight`: Weight in kg.
- `height`: Height in cm.
- `atcd`: Family history of cardiovascular diseases. 0 = no; 1 = yes.
- `hta`: Hypertension. 0 = no; 1 = yes.

Author(s)

Lafaye de Micheaux Pierre <lafaye@unsw.edu.au>, Remy Drouilhet <Remy.Drouilhet@upmf-grenoble.fr>, Liquet Benoit <b.liquet@uq.edu.au>

Source

http://www.biostatisticien.eu/springeR/

Examples

data(INFARCTION)
str(INFARCTION)
Description

Atherosclerosis is the main cause of death for men above 35 and women above 45 in most developed countries. It is a thickening and hardening of internal artery walls. One of its consequences is myocardial infarction. An artery wall is made of three layers; innermost to outermost, they are called intima, media and adventitia. Intima-media thickness is a marker of atherosclerosis. It was measured by ultrasound on a sample of 110 subjects in 1999 in Bordeaux hospitals. Information on the main risk factors was also collected.

Usage

data(INTIMA.MEDIA)

Format

A data frame with 110 observations measured on the 9 following variables:

- gender 1=male, 2=female. Gender.
- age Age (in years) at date of consultation.
- height Height in cm.
- weight Weight in kg.
- tobacco 0=non smoker, 1=former smoker, 2=smoker. Smoking status.
- packyear Number of packs per year. Estimation of tobacco consumption for smokers and former smokers.
- sport 0=no, 1=yes. Physical activity.
- measure Intima-media thickness in cm
- alcohol 0=non-drinker, 1=occasional drinker, 2=regular drinker. Alcohol consumption.

Author(s)

Lafaye de Micheaux Pierre <lafaye@unsw.edu.au>, Remy Drouilhet <Remy.Drouilhet@upmf-grenoble.fr>, Liquet Benoit <b.liquet@uq.edu.au>

Source

http://www.biostatisticien.eu/springeR/

Examples

data(INTIMA.MEDIA)
str(INTIMA.MEDIA)
Description

Computes the Moore Penrose inverse of a matrix

Usage

```r
mpinv(M, eps=1e-13)
```

Arguments

- `M`: a matrix
- `eps`: real precision

Value

The Moore-Penrose inverse of `M`

Author(s)

Lafaye de Micheaux Pierre <lafaye@unsw.edu.au>, Remy Drouilhet <Remy.Drouilhet@upmf-grenoble.fr>, Liquet Benoit <b.liquet@uq.edu.au>

References

Chapter 10 (Basic Mathematics: Matrix Operations, Integration, and Optimization) from the book: The R Software, Fundamentals of Programming and Statistical Analysis

Examples

```r
A <- matrix(c(2,3,5,1), nrow=2, ncol=2)
solve(A)
mpinv(A)
B <- matrix(c(4,2,8,4), nrow=2, ncol=2)
# solve(B) # gives an error.
mpinv(B)
```
Description

A sample of 226 elderly people living in Bordeaux (Gironde, South-West France) were interviewed in 2000 for a nutritional study.

Usage

```r
data(NUTRIELDERLY)
```

Format

A data frame with 226 observations measured on the 13 following variables:

- `gender`: 2 = female; 1 = male
- `situation`: 1 = single; 2 = living with spouse; 3 = living with family; 4 = living with someone else; Family status.
- `tea`: Number of cups. Daily consumption of tea.
- `coffee`: Number of cups. Daily consumption of coffee
- `height`: Height in cm.
- `weight`: Weight in cm.
- `age`: Age in years at date of interview.
- `meat`: 0 = never; 1 = less than once a week; 2 = Once a week; 3 = 2/3 times a week; 4 = 4/6 times a week; 5 = every day. Consumption of meat.
- `fat`: 1 = butter; 2 = margarine; 3 = peanut oil; 4 = sunflower oil; 5 = olive oil; 6 = mix of vegetable oils (e.g., Isio4); 7 = colza oil; 8 = duck or goose fat. Type of fat used for cooking.

Author(s)

Lafaye de Micheaux Pierre <lafaye@unsw.edu.au>, Remy Drouilhet <Remy.Drouilhet@upmf-grenoble.fr>, Liquet Benoit <b.liquet@uq.edu.au>

Source

`http://www.biostatisticien.eu/springeR/`

Examples

```r
data(NUTRIELDERLY)
str(NUTRIELDERLY)
```
**Description**

Comparing the theoretical variance with a reference value.

**Usage**

```r
sigma2.test(x, alternative = "two.sided", var0 = 1, conf.level = 0.95)
```

**Arguments**

- `x` numeric vector
- `alternative` Alternative hypothesis for the test. Either two sided ("two.sided"), one sided to the left ("less") or one sided to the right ("greater"). Default is "two.sided".
- `var0` value of reference for the variance
- `conf.level` confidence level

**Value**

Returns a list:

- `statistic` Value of the test statistic
- `parameter` degrees of freedom
- `p.value` `p`-value of the test
- `conf.int` confidence interval
- `estimate` sample variance
- `null.value` value of reference for the variance
- `alternative` Alternative hypothesis for the test
- `method` "One-sample Chi-squared test for given variance"
- `data.name` name of the data set

**Author(s)**

Lafaye de Micheaux Pierre <lafaye@unsw.edu.au>, Remy Drouilhet <Remy.Drouilhet@upmf-grenoble.fr>,
Liquet Benoit <b.liquet@uq.edu.au>

**References**


**Examples**

```r
data(NUTRIELDERLY)
sigma2.test(NUTRIELDERLY$weight, conf.level=0.9)$conf
```
twosample.cor.test

Comparing statistically two correlation coefficients

Description

Test of the equality of two correlation coefficients

Usage

twosample.cor.test(x1, y1, x2, y2, alpha = 0.05, alternative =
c("two.sided", "less", "greater"))

Arguments

- x1: x1 is a numeric vector associated to y1
- y1: y1 is a numeric vector associated to x1
- x2: x2 is a numeric vector associated to y2
- y2: y2 is a numeric vector associated to x2
- alpha: significance level of the test
- alternative: Alternative hypothesis for the test. Either two sided ("two.sided"), one sided to the left ("less") or one sided to the right ("greater"). Default is "two.sided".

Value

Returns a list:

- statistic: Value of the test statistic
- p.value: p-value of the test

Author(s)

Lafaye de Micheaux Pierre <lafaye@unsw.edu.au>, Remy Drouilhet <Remy.Drouilhet@upmf-grenoble.fr>, Liquet Benoit <b.liquet@uq.edu.au>

References

Chapter 13 (Confidence Intervals and Hypothesis Testing) from the book: The R Software, Fundamentals of Programming and Statistical Analysis

See Also

cor0.test
VectorAddr

Examples

data(BMI.CHILD)
attach(BMI.CHILD)
indf <- which(GENDER=="F")  # To retrieve indices of the females.
indm <- which(GENDER=="M")  # To retrieve indices of the males.
twosample.cor.test(height[indf],weight[indf],
                    height[indm],weight[indm])
detach(BMI.CHILD)

vectoraddr

Address of vector

Description

Object representing an address of numeric vector

Usage

VectorAddr(x)

Arguments

x

Vector.

Value

An object of class VectorAddr.

Author(s)

Lafaye de Micheaux Pierre <lafaye@unsw.edu.au>, Remy Drouilhet <Remy.Drouilhet@upmf-grenoble.fr>,
Liquet Benoit <b.liquet@uq.edu.au>

References

Chapter 9 (Managing Sessions) from the book: The R Software, Fundamentals of Programming
and Statistical Analysis

Examples

## Not run:
x <- c(8L,9L)
addr <- VectorAddr(x) # Gets the address of the first
# box of the 64-box block where x
# is stored.
addr
update(addr,6L) # Write the integer 6 at this address.
x
update(addr+4L,7L) # An integer is coded over 4 bytes,
# hence increment the address by 4 to
# get to x[2].

x
x <- c(12.8, 4.5)

addr <- VectorAddr(x)  # Get the address of the first box of the 128-box block where x is stored.

update(addr, 6.2)

update(addr + 8L, 7.1)  # A double is coded over 8 bytes.

## End (Not run)

---

**writeaddr**

*Writing a value at some memory address*

**Description**

Writing a value at some memory address

**Usage**

`writeaddr(addr, newval)`

**Arguments**

- `addr`  Integer value. Address in memory.
- `newval`  New value to write at this address.

**Value**

Nothing is returned.

**Author(s)**

Lafaye de Micheaux Pierre <lafaye@unsw.edu.au>, Remy Drouilhet <Remy.Drouilhet@upmf-grenoble.fr>, Liquet Benoit <b.liquet@uq.edu.au>

**References**

Chapter 9 (Managing Sessions) from the book: *The R Software, Fundamentals of Programming and Statistical Analysis*
Examples

```r
## Not run:
x <- c(8L, 9L)
addr <- getaddr(x)$addr.int # Gets the address of the first
   # box of the 64-box block where x
   # is stored.

addr
writeaddr(addr, 6L) # Write the integer 6 at this address.
x
writeaddr(addr+4L, 7L) # An integer is coded over 4 bytes,
   # hence increment the address by 4 to
   # get to x[2].

x
x <- c(12.8, 4.5)
x
addr <- getaddr(x)$addr.int # Get the address of the first box
   # of the 128-box block where x is
   # stored.

writeaddr(addr, 6.2)
x
writeaddr(addr+8L, 7.1) # A double is coded over 8 bytes.
x

## End(Not run)
```
Index

*Topic **datasets**
  BIRTH. WEIGHT, 5
  BMI. CHILD, 6
  INFACTION, 12
  INTIMA. MEDIA, 14
  NUTRIELDERLY, 16

*Topic **package**
  TheRSoftware-package, 2

arrowaxis, 2
barchart, 3
barplot, 4
bin2dec, 4
BIRTH. WEIGHT, 5
BMI. CHILD, 6
camembert, 7
cor.test, 8
cor0.test, 8, 18
crosschart, 9
dec2bin, 10
flashy.plot, 10
getaddr, 11
getAddrDb1 (VectorAddr), 19
getAddrInt (VectorAddr), 19
INFACTION, 12
INTIMA. MEDIA, 14

mpinv, 15
NUTRIELDERLY, 16
Ops.VectorAddr (VectorAddr), 19

pie, 7
plot, 11
print.VectorAddr (VectorAddr), 19
printAddr (VectorAddr), 19
sigma2.test, 17
TheRSoftware (TheRSoftware-package), 2
TheRSoftware-package, 2
twosample.cor.test, 18
update.VectorAddr (VectorAddr), 19
VectorAddr, 19
writeaddr, 20
writeAtAddrDb1 (VectorAddr), 19
writeAtAddrInt (VectorAddr), 19