Package 'etable'

February 19, 2015

Version 1.2.0

Date 2013-12-11

Title Easy Table

Author Andreas Schulz

Maintainer Andreas Schulz <ades-s@web.de>

Depends R (>= 3.0.0), xtable, Hmisc

Description A table function for descriptive statistics in tabular format, using variables in a data.frame. You can create simple or highly customized tables.

License GPL (>= 2)

NeedsCompilation no

Repository CRAN

Date/Publication 2013-12-11 17:19:33

R topics documented:

`etable-package` *Easy Table*

Description

The package comes without any warranty.

Details

Package:	etable
Title:	Easy Table
Type:	Package
Version:	1.2.0
Date:	2013-12-11
Depends:	epade, Hmisc, R (>= 2.13)
License:	GPL version 2 or newer
LazyLoad:	yes

Author(s)

Andreas Schulz
Maintainer: <ades-s@web.de>

`combi_cell` *Dichotomous and continuous variables combination Cell FUN*

Description

To calculate different statistics depending on the type of variable.

Usage

```
combi_cell(x, y, z, w, cell_ids, row_ids, col_ids, vnames, vars, n_min,
             digits=3, style=1)
```

Arguments

`x`	The x variable for calculations, if not using y
`y`	The y variable for calculations, if not using x
`z`	NOT USED
`w`	Weights for x or y variable.

cell_ids	Index vector for selecting values in cell.
row_ids	NOT USED
col_ids	NOT USED
vnames	NOT USED
vars	NOT USED
n_min	Minimum n in the cell for useful calculation. Cells with n<n_min deliver no output.
digits	Integer indicating the number of significant digits.
style	Type of representation.

- 1 N, Proportion, Median, Q1, Q3
- 2 N, Proportion, Mean, SD

Author(s)

ADES <ades-s@web.de>

Examples

```
sex     <- factor(rbinom(1000, 1, 0.4),  labels=c('Men', 'Women'))
height  <- rnorm(1000, mean=1.7, sd=0.1)
weight  <- rnorm(1000, mean=70, sd=5)
bmi     <- weight/height^2
event   <- factor(rbinom(1000, 1, 0.1), labels=c('no',  'yes'))
d<-data.frame(sex, height, weight, bmi, event)
tabular.ade(x_vars=names(d), cols=c('sex','ALL'), rnames=c('Gender'),
            data=d, FUN=combi_cell)
```

---

corr_p_cell *Correlation Cell FUN*

Description

To calculating Pearson product-moment correlation coefficient.

Usage

```
corr_p_cell(x, y, z, w, cell_ids, row_ids, col_ids, vnames, vars, n_min,
digits = 3)
```

Arguments

x	The x variable
y	The y variable
z	NOT USED
w	Weights for x and y variable.
cell_ids	Index vector for selecting values in cell.
row_ids	NOT USED
col_ids	NOT USED
vnames	NOT USED
vars	NOT USED
n_min	Minimum n in the cell for useful calculation. Cells with n<n_min deliver no output.
digits	Integer indicating the number of decimal places.

Author(s)

ADES <ades-s@web.de>

Examples

```
sex     <- factor(rbinom(1000, 1, 0.4),  labels=c('Men', 'Women'))
height  <- rnorm(1000, mean=1.70, sd=0.1)
weight  <- rnorm(1000, mean=70, sd=5)
bmi     <- weight/height^2
d<-data.frame(sex, bmi, height, weight)
tabular.ade(x_vars=c('bmi','height','weight'), xname=c('BMI','Height','Weight'),
            y_vars=c('bmi','height','weight'), yname=c('BMI','Height','Weight'),
            rows=c('sex','ALL'), rnames=c('Gender'), data=d, FUN=corr_p_cell)
```

---

eventpct_cell *Factor level frequencies Cell FUN.*

Description

For calculating frequencies or proportions of a certain level from factor x.

Usage

```
eventpct_cell(x, y, z, w, cell_ids, row_ids, col_ids, vnames, vars, n_min,
digits=1, digits2=0, event=2, type=1)
```

Arguments

x	The factor x for calculations
y	NOT USED
z	NOT USED
w	Weights for x factor. Only if calculating weigted frequences.
cell_ids	Index vector for selecting values in cell.
row_ids	NOT USED
col_ids	NOT USED
vnames	NOT USED
vars	NOT USED
n_min	Minimum n in the cell for useful calculation. Cells with n<n_min deliver no output.
digits	Integer indicating the number of decimal places (percentages)
digits2	Integer indicating the number of decimal places (N, needed if N is not integer because of weighting)
event	The Number of factor level to calculate frequencies. from 1 to nlevels(x)
type	Type of representation, one of following. • 1, pct (n) • 2, n (pct) • 3, pct • 4, n • 5, pct (n/N)

Author(s)

ADES <ades-s@web.de>

Examples

```
sex     <- factor(rbinom(1000, 1, 0.4), labels=c('Men', 'Women'))
event   <- factor(rbinom(1000, 1, 0.1), labels=c('no',  'yes'))
decades <- rbinom(1000, 3, 0.5)
decades <- factor(decades, labels=c('[35,45)','[45,55)','[55,65)','[65,75)'))
d<-data.frame(sex, decades, event)
tabular.ade(x_vars=c('event'), xname=c('Event'),
   rows=c('sex','ALL'), rnames=c('Gender'),
   cols=c('decades', 'ALL'),   cnames=c('Age decades'),
   data=d, FUN=eventpct_cell)
```

iqr_cell *Median IQR Cell FUN.*

Description

For calculate median and interquartile range. (weighting is possible)

Usage

```
iqr_cell(x, y, z, w, cell_ids, row_ids, col_ids, vnames, vars, n_min,
digits = 3, add_n=FALSE)
```

Arguments

x	The x variable for calculations
y	NOT USED
z	NOT USED
w	Weights for x variable.
cell_ids	Index vector for selecting values in cell.
row_ids	NOT USED
col_ids	NOT USED
vnames	NOT USED
vars	NOT USED
n_min	Minimum n in the cell for useful calculation. Cells with n<n_min deliver no output.
digits	Integer indicating the number of significant digits.
add_n	Logical asking whether to draw N for each cell.

Author(s)

ADES <ades-s@web.de>

Examples

```
sex     <- factor(rbinom(1000, 1, 0.4),  labels=c('Men', 'Women'))
height  <- rnorm(1000, mean=1.66, sd=0.1)
height[which(sex=='Men')]<-height[which(sex=='Men')]+0.1
weight  <- rnorm(1000, mean=70, sd=5)
decades <- rbinom(1000, 3, 0.5)
decades <- factor(decades, labels=c('[35,45)','[45,55)','[55,65)','[65,75)'))
d<-data.frame(sex, decades, height, weight)
tabular.ade(x_vars=c('height', 'weight'), xname=c('Height [m]','Weight [kg]'),
   rows=c('sex','ALL'), rnames=c('Gender'),
   cols=c('decades'),   cnames=c('Age decades'),
   data=d, FUN=iqr_cell, add_n=TRUE)
```

mean_sd_cell *Mean and SD Cell FUN*

Description

To calculate mean and SD or weighted mead and SD.

Usage

```
mean_sd_cell(x, y, z, w, cell_ids, row_ids, col_ids, vnames, vars, n_min,
             digits = 3, style=1, nsd=1)
```

Arguments

x	The x variable for calculations
y	NOT USED
z	NOT USED
w	Weights for x variable.
cell_ids	Index vector for selecting values in cell.
row_ids	NOT USED
col_ids	NOT USED
vnames	NOT USED
vars	NOT USED
n_min	Minimum n in the cell for useful calculation. Cells with n<n_min deliver no output.
digits	Integer indicating the number of significant digits.
style	Type of representation. • 1. mean (sd) • 2. mean (mean-sd*nsd, mean+sd*nsd) • 3. mean plus-minus sd
nsd	Multiplier for sd in stlyle 2. (for normal distribution) • nsd=1 –> 68.27 % values • nsd=1.645 –> 90 % values • nsd=1.96 –> 95 % values • nsd=2 –> 95.45 % values • nsd=2.575 –> 99 % values • nsd=3 –> 99.73 % values

Author(s)

ADES <ades-s@web.de>

Examples

```
sex     <- factor(rbinom(1000, 1, 0.4),  labels=c('Men', 'Women'))
height  <- rnorm(1000, mean=1.66, sd=0.1)
height[which(sex=='Men')]<-height[which(sex=='Men')]+0.1
weight  <- rnorm(1000, mean=70, sd=5)
decades <- rbinom(1000, 3, 0.5)
decades <- factor(decades, labels=c('[35,45)','[45,55)','[55,65)','[65,75)'))
d<-data.frame(sex, decades, height, weight)
tabular.ade(x_vars=c('height', 'weight'), xname=c('Height [m]','Weight [kg]'),
   rows=c('sex','ALL'), rnames=c('Gender'),
   cols=c('decades'),   cnames=c('Age decades'),
   data=d, FUN=mean_sd_cell, style=2, nsd=1.96)
```

`miss_cell` *missing values Cell FUN*

Description

For counting the number of missing values in each cell.

Usage

```
miss_cell(x, y, z, w, cell_ids, row_ids, col_ids, vnames, vars, n_min,
          pct = FALSE, digits = 0, prefix='', suffix='')
```

Arguments

`x`	The x variable
`y`	NOT USED
`z`	NOT USED
`w`	NOT USED (The number of missing will not be weighted!).
`cell_ids`	Index vector for selecting values in cell.
`row_ids`	NOT USED
`col_ids`	NOT USED
`vnames`	NOT USED
`vars`	NOT USED
`n_min`	NOT USED
`pct`	Logical asking whatever to draw absolute or relative frequency of missing values.
`digits`	Integer indicating the number of decimal places.
`prefix`	Free text added in each cell bevor results.
`suffix`	Free text added in each cell after results.

Author(s)

ADES <ades-s@web.de>

Examples

```
sex     <- factor(rbinom(1000, 1, 0.4),  labels=c('Men', 'Women'))
height  <- rnorm(1000, mean=1.66, sd=0.1)
height[which(sex=='Men')]<-height[which(sex=='Men')]+0.1
weight  <- rnorm(1000, mean=70, sd=5)
decades <- rbinom(1000, 3, 0.5)
decades <- factor(decades, labels=c('[35,45)','[45,55)','[55,65)','[65,75)'))
d<-data.frame(sex, decades, height, weight)
d$height[round(runif(250,1,1000))]<- NA
d$weight[round(runif(25 ,1,1000))]<- NA
tabular.ade(x_vars=c('height', 'weight'), xname=c('Height [m]','Weight [kg]'),
        cols=c('sex','decades','ALL'), cnames=c('Gender', 'Age decades'),
        data=d, FUN=miss_cell, prefix='Miss:')
```

---

mode_cell *Mode Cell FUN*

Description

Find most frequently value

Usage

```
mode_cell(x, y, z, w, cell_ids, row_ids, col_ids, vnames, vars, n_min,
       digits=3)
```

Arguments

x	The x variable
y	NOT USED
z	NOT USED
w	Weights for x variable. Only if calculating weigted mode.
cell_ids	Index vector for selecting values in cell.
row_ids	Index vector for selecting values in row.
col_ids	Index vector for selecting values in col.
vnames	NOT USED
vars	NOT USED
n_min	NOT USED
digits	Integer indicating the number of significant digits.

Author(s)

ADES <ades-s@web.de>

Examples

```
sex     <- factor(rbinom(1000, 1, 0.4),  labels=c('Men', 'Women'))
note    <- as.factor(rbinom(1000, 4, 0.5)+1)
decades <- rbinom(1000, 3, 0.5)
decades <- factor(decades, labels=c('[35,45)','[45,55)','[55,65)','[65,75)'))
d<-data.frame(sex, decades, note)
tabular.ade(x_vars=c('note'), xname=c('Noten'),
       rows=c('sex','ALL','decades'), rnames=c('Gender', 'Age decades'),
       data=d, FUN=mode_cell)
```

---

n_cell *Frequency Cell FUN*

Description

For calculating relative or absolute frequencies.

Usage

```
n_cell(x, y, z, w, cell_ids, row_ids, col_ids, vnames, vars, n_min,
       digits=0, digits2=1, type="n")
```

Arguments

x	The x variable (can be easy 1:N if no missings)
y	NOT USED
z	NOT USED
w	Weights for x variable. Only if calculating weigted frequences.
cell_ids	Index vector for selecting values in cell.
row_ids	Index vector for selecting values in row.
col_ids	Index vector for selecting values in col.
vnames	NOT USED
vars	NOT USED
n_min	NOT USED
digits	Integer indicating the number of decimal places (N)
digits2	Integer indicating the number of decimal places (percent)
type	Type of frequencies, one of following. • n, Number in cell. • pct, Overall percentages.

- pctn, Overall percentages and n.
- rowpct, Percentages of rows.
- colpct, Percentages of cols.
- rowpctn, Percentages of rows and n.
- colpctn, Percentages of cols and n.
- all, Overall, row, col percentages.

Details

The function calculate frequencies for cell. If x has no missings the frequencies are independent from x.

Author(s)

ADES <ades-s@web.de>

Examples

```
sex     <- factor(rbinom(1000, 1, 0.4),  labels=c('Men', 'Women'))
decades <- rbinom(1000, 3, 0.5)
decades <- factor(decades, labels=c('[35,45)','[45,55)','[55,65)','[65,75)'))
d<-data.frame(sex, decades)
tabular.ade(x_var='sex',  rows=c('sex',     'ALL'), rnames=c('Gender'),
                          cols=c('decades', 'ALL'), cnames=c('Age decades'),
            data=d, FUN=n_cell, , type="all")
```

---

quantile_cell *quantile Cell FUN*

Description

For calculating sample or weighted quantiles

Usage

```
quantile_cell(x, y, z, w, cell_ids, row_ids, col_ids, vnames, vars, n_min,
              digits = 3, probs = 0.5, plabels=FALSE)
```

Arguments

x	The x variable for calculations
y	NOT USED
z	NOT USED
w	Weights for x variable.
cell_ids	Index vector for selecting values in cell.
row_ids	NOT USED

col_ids	NOT USED
vnames	NOT USED
vars	NOT USED
n_min	Minimum n in the cell for useful calculation. Cells with n<n_min deliver no output.
digits	Integer indicating the number of significant digits.
probs	A single or a vector of numeric probabilities for sample quantile with values in [0,1].
plabels	Logical asking whether to label the quantile in the cell or only draw the value.

Author(s)

ADES <ades-s@web.de>

Examples

```
sex     <- factor(rbinom(1000, 1, 0.4),  labels=c('Men', 'Women'))
height  <- rnorm(1000, mean=1.66, sd=0.1)
height[which(sex=='Men')]<-height[which(sex=='Men')]+0.1
weight  <- rnorm(1000, mean=70, sd=5)
decades <- rbinom(1000, 3, 0.5)
decades <- factor(decades, labels=c('[35,45)','[45,55)','[55,65)','[65,75)'))
d<-data.frame(sex, decades, height, weight)
tabular.ade(x_vars=c('height', 'weight'), xname=c('Height [m]','Weight [kg]'),
   rows=c('sex',     'ALL'), rnames=c('Gender'),
   cols=c('decades', 'ALL'), cnames=c('Age decades'),
   data=d, FUN=quantile_cell, probs = 0.99)
```

---

stat_cell *Diverse statistics Cell FUN*

Description

To calculate values of several statistics.

Usage

```
stat_cell(x, y, z, w, cell_ids, row_ids, col_ids, vnames, vars, n_min,
digits = 3, digits2=1)
```

Arguments

x	The x variable
y	NOT USED
z	NOT USED
w	Weights for x variable.

cell_ids	Index vector for selecting values in cell.
row_ids	NOT USED
col_ids	NOT USED
vnames	NOT USED
vars	a vector of character strings with names of variables in data.frame for x,y and z. Uses names of x or y as keywords, to choose the statistik.
n_min	Minimum n in the cell for useful calculation. Cells with n<n_min deliver no output.
digits	Integer indicating the number of significant digits.
digits2	Integer indicating the number of decimal places for percentages.

Details

Keywords are:

- N: number in this cell
- MIN: minimum
- MAX: maximum
- SUM: sum
- MEAN: mean
- SD: standard deviation
- MSD: mean, standard deviation
- M2SD: mean-2SD, mean+2SD
- VAR: variance
- MEDIAN: median
- MD: mean deviation from the mean (*1.253)
- MAD: median absolute deviation (*1.4826)
- IQR: interquartile range
- MQQ: median (Q1/Q3)
- PROP: proportion
- POP: proportion of level 2 (only binar)
- RANGE: range
- CV: coefficient of variation
- MODE: mode
- MISS: number of missing values
- PNM: proportion of non missing values
- COMB: POP for binar and MQQ for continues
- SKEW: skewness
- KURT: excess kurtosis

- P1: 1th Quantile
- P2.5: 2.5th Quantile
- P5: 5th Quantile
- P10: 10th Quantile
- P25: 25th Quantile
- P50: 50th Quantile
- P75: 75th Quantile
- P90: 90th Quantile
- P95: 95th Quantile
- P97.5: 97.5th Quantile
- P99: 99th Quantile

Author(s)

ADES <ades-s@web.de>

Examples

```
sex     <- factor(rbinom(1000, 1, 0.4),  labels=c('Men', 'Women'))
height  <- rnorm(1000, mean=1.66, sd=0.1)
height[which(sex=='Men')]<-height[which(sex=='Men')]+0.1
weight  <- rnorm(1000, mean=70, sd=5)
decades <- rbinom(1000, 3, 0.5)
decades <- factor(decades, labels=c('[35,45)','[45,55)','[55,65)','[65,75)'))
d<-data.frame(sex, decades, height, weight)
tabular.ade(x_vars=c('height', 'weight'), xname=c('Height [m]','Weight [kg]'),
   y_vars=c('N', 'MEAN', 'SD', 'SKEW', 'KURT'),
   rows=c('sex', 'ALL', 'decades', 'ALL'), rnames=c('Gender', 'Age decades'),
   data=d, FUN=stat_cell)
```

`tabular.ade` *Tabular representation of any statistics.*

Description

Create tables for any descriptive statistic, using variables in a data.frame. You can create simple to highly customized tables, with or without weighting.

Usage

```
tabular.ade(x_vars, xname=NULL, y_vars=NULL, yname=NULL,
            z_vars=NULL, zname=NULL,
            rows=NULL, rnames=NULL, cols=NULL, cnames=NULL, w=NULL,
            data=NULL, FUN, allnames=FALSE, nonames=TRUE, alllabel='Total',
            inset='?', remove='', n_min=0, ...)
```

Arguments

x_vars	This variable will be used to calculate the statistics for it. • a character string with the name of the variable in the data.frame • a vector of character strings with names of variables in data.frame
xname	Labels for x. • a character string with the label for x • a vector of character strings with labels for x, with same length as x.
y_vars	This variable can be used to calculate bivariable statistics. • a character string with the name of the variable in the data.frame • a vector of character strings with names of variables in data.frame
yname	Labels for y. • a character string with the label for y • a vector of character strings with labels for y, with same length as x.
z_vars	This variable can be used for additional calculations. • a character string with the name of the variable in the data.frame
zname	Labels for z. • a character string with the label for y
rows	This factors will be used to separate the rows of the table in subgroups. • a character string with the name of the factor variable in the data.frame • a vector of character strings with names of factor variables in data.frame (max 6) • a vector with names of factors and/or Keyword 'ALL', add extra overall group for leading factor.
rnames	Labels for rows. • a character string with the label for rows • a vector of character strings with labels for rows, with same length as rows. • a vector with names of factors and/or Keyword 'ALL', add extra overall group for leading factor.
cols	This factors will be used to separate the columns of the table in subgroups. • a character string with the name of the factor variable in the data.frame • a vector of character strings with names of factor variables in data.frame (max 6)
cnames	Labels for cols. • a character string with the label for cols • a vector of character strings with labels for rows, with same length as cols.
w	This numeric variable will be used to weighting the table. • a character string with the name of the factor variable in the data.frame
data	A data frame with all used variables.
FUN	A abstract cell function to calculate statistic in every cell of the table. See details.

allnames	Logical asking whether to fill every cell with labels or only the first one.
nonames	Logical asking whether to use dimnames for variable labels or make all labeling in the table self.
alllabel	Label for overall Group without splitting in this Factor.
inset	Inset text in each cell, '?' will be replaced with the value of the cell.
remove	Remove a character string from each cell.
n_min	min N in each cell, it will be only passed in the cell function. But it is necessary to not calculate statistics from 1 or 2 values.
...	additional parameters passed to the FUN

Details

FUN can be a cell function from this package or your own function. If you wanna writing you own cell function. It must take following parameters. But it must not use them.

- x, The whole x variable.
- y, The whole y variable.
- z, The whole z variable.
- w, The whole w variable.
- cell_ids, Index vector to select values that belong in this cell.
- row_ids, Index vector to select values that belong in this row.
- col_ids, Index vector to select values that belong in this col.
- vnames, A vector of length 3, with labels of variables (x,y,z)
- vars, A vector of length 3, with names of variables (x,y,z)
- n_min , Min needed N for calculation.
- ... , additional own parameters.

An example with simple mean see below.

Value

A character Matrix.(Table)

Author(s)

ADES <ades-s@web.de>

Examples

```
# 1) simple own FUN cell function.
s_mean<- function(x, y, z, w, cell_ids, row_ids, col_ids, vnames, vars, n_min, ds=3){
out<- ''
if(length(cell_ids)>= n_min){
out<-  format(mean(x[cell_ids], na.rm=TRUE), digits=ds)
}
return(out)
```

```
}
########################################
# 2) simple 2 x 2 table of means
sex   <- factor(rbinom(5000, 1, 0.5), labels=c('Men', 'Women'))
age   <- round(runif(5000, 18, 89))
treat <- factor(rbinom(5000, 1, 0.3), labels=c('control', 'treated'))
d<-data.frame(sex, age, treat)
tabular.ade(x_vars='age', xname='Age [y]', rows='sex', rnames='Sex', cols='treat',
cnames='Treatment', data=d, nonames=FALSE, FUN=s_mean)
########################################
# 3) Relative frequency table
d$dosis <- round(runif(5000, 0.5, 6.49))
tabular.ade(x_vars='age', xname='Age [y]', rows=c('sex', 'treat'),
rnames=c('Sex', 'Treatment'), cols='dosis', cnames='Dosis', data=d, FUN=n_cell,
type='pct')
########################################
# 4) Weighted median table
d$w <- runif(5000, 0.1, 5)
d$bmi <- rnorm(5000, 30, 3)
tabular.ade(x_vars=c('age', 'bmi'), xname=c('Age', 'BMI'),
cols=c('sex', 'ALL', 'treat'),
cnames=c('Sex', 'Treatment'), w='w', data=d, FUN=quantile_cell)
########################################
# 5) Correlation table between age and bmi
tabular.ade(x_vars='age', xname='Age', y_vars='bmi', yname='BMI',
rows=c('dosis'), rnames=c('Dosis'), cols=c('sex', 'treat'),
cnames=c('Sex', 'Treatment'), data=d, FUN=corr_p_cell)
########################################
# 6) Multiple statistics
tabular.ade(x_vars=c('N', 'MEAN', 'SD', 'SKEW', 'KURT', 'RANGE'),
y_vars=c('age', 'bmi'), yname=c('Age', 'BMI'),
cols=c('sex', 'ALL', 'treat'), cnames=c('Sex', 'Treatment'),
w='w', data=d, FUN=stat_cell)
```

Index