Package ‘summarytools’

March 2, 2020

Type Package

Title Tools to Quickly and Neatly Summarize Data

Version 0.9.6

Description Data frame summaries, cross-tabulations, weight-enabled frequency tables and common descriptive (univariate) statistics in concise tables available in a variety of formats (plain ASCII, Markdown and HTML). A good point-of-entry for exploring data, both for experienced and new R users.

Imports base64enc, checkmate, dplyr, grDevices, htmltools, lubridate, magick, matrixStats, methods, pander, pryr, rapportools, stats, tcltk, tibble, tidyr, utils

Suggestsforcats, kableExtra, knitr, magrittr, rmarkdown, rstudioapi

Depends R (>= 2.10)

VignetteBuilder knitr

LazyData true

License GPL-2

URL https://github.com/dcomtois/summarytools

BugReports https://github.com/dcomtois/summarytools/issues

Encoding UTF-8

RoxygenNote 7.0.2

NeedsCompilation no

Author Dominic Comtois [aut, cre]

Maintainer Dominic Comtois <dominic.comtois@gmail.com>

Repository CRAN

Date/Publication 2020-03-02 07:20:02 UTC
**R topics documented:**

summarytools-package .................................................. 2
bleartmp ................................................................. 3
cetable ................................................................. 4
define_keywords .......................................................... 6
descr ................................................................. 9
dfSummary .............................................................. 11
examens ............................................................ 15
exams .............................................................. 16
freq .............................................................. 16
label ............................................................. 19
print.list ............................................................. 20
print.stby ............................................................. 21
print.summarytools ...................................................... 22
stby ............................................................ 25
st_css ............................................................ 26
st_options ............................................................. 27
tabagisme ............................................................. 30
tb .............................................................. 31
tobacco ............................................................ 32
unlabel ............................................................ 32
use_custom_lang .......................................................... 33
view ............................................................ 33
what.is ............................................................ 35

**Index**

37

**summarytools-package  Tools to Quickly and Neatly Summarize Data**

**Description**

**summarytools** provides users with functions to neatly and quickly summarize numerical and categorical data. Data frame summaries, frequency tables and cross-tabulations, as well as common descriptive (univariate) statistics can be produced in a straightforward manner. Users with little to no prior R programming experience but who are familiar with popular commercial statistical software such as SAS, SPSS and Stata should feel right at home.

**Details**

These are the four core functions:

**dfSummary** Extensive yet legible data frame summaries.

**freq** Frequency tables supporting weights and displaying proportions of valid and of total data, including cumulative proportions.

**descr** All common univariate descriptive stats for single vectors or for all numerical vectors in a data frame.
**etable** Cross-tabulations for two categorical vectors or factors. Choose between *Total*, *Columns* or *Rows* proportions.

**Output formats** are:

- **plain ascii** Ideal when looking at results in the console.
- **rmarkdown** Ideal when writing short papers or presentations.
- **html** This format is well integrated in RStudio (but will work with any browser). Use the `view()` function to see results appear directly in RStudio’s Viewer or in your default Web Browser.

**Author(s)**

**Maintainer**: Dominic Comtois <dominic.comtois@gmail.com>

**See Also**

Useful links:

- [https://github.com/dcomtois/summarytools](https://github.com/dcomtois/summarytools)
- Report bugs at [https://github.com/dcomtois/summarytools/issues](https://github.com/dcomtois/summarytools/issues)

---

### cleartmp

#### Delete Temporary Html Files

**Description**

Delete temporary files created when using generic print method with `method='browser'` or `method='viewer'`, or when calling `view()` function.

**Usage**

```
cleartmp(all = TRUE, silent = FALSE, verbose = FALSE)
```

**Arguments**

- **all** Logical. When `TRUE` (default), all temporary summarytools files are deleted. When `FALSE`, only the latest file is.
- **silent** Logical. Hide confirmation messages (`FALSE` by default).
- **verbose** Logical. Display a message for every file that is deleted. `FALSE` by default.

**Details**

All temporary files are deleted automatically when R session is ended. This function is thus an overkill in most circumstances.

**Author(s)**

Dominic Comtois, <dominic.comtois@gmail.com>
ctable

Cross-Tabulation

Description

Cross-tabulation for a pair of categorical variables (or factors) with either row, column, or total proportions, as well as marginal sums.

Usage

cetable(
  x,
  y,
  prop = st_options("ctable.prop"),
  useNA = "ifany",
  totals = st_options("ctable.totals"),
  style = st_options("style"),
  round.digits = 1,
  justify = "right",
  plain.ascii = st_options("plain.ascii"),
  headings = st_options("headings"),
  display.labels = st_options("display.labels"),
  split.tables = Inf,
  dnn = c(substitute(x), substitute(y)),
  chisq = FALSE,
  OR = FALSE,
  RR = FALSE,
  weights = NA,
  rescale.weights = FALSE,
  ...
)

Arguments

x
  First categorical variable - values will appear as row names.

y
  Second categorical variable - values will appear in as column names.

prop
  Proportions to display; “r” for rows (default), “c” for columns, “t” for total, or “n” for none. This option can be set globally; see st_options.

useNA
  Argument passed on to table; One of “ifany” (default), “no”, or “always”.

totals
  Logical. Should row and column totals be displayed? Defaults to TRUE. To change this default value globally, see st_options.

style
  Style to be used by pander when rendering output table; One of “simple” (default), “grid”, or “markdown”. This option can be set globally; see st_options.

round.digits
  Number of significant digits to display. Defaults to 1. To change this default value globally, see st_options.
Table 5

Justify

String indicating alignment of columns; one of “l” (left) “c” (center), or “r” (right). Defaults to “r”.

Plain.ascii

Logical. `pander` argument; when `TRUE`, no markup characters will be used (useful when printing to console). Defaults to `TRUE` unless `style = 'markdown'`, in which case it will be set to `FALSE` automatically. To change the default value globally, use `st_options`.

Headings

Logical. Set to `FALSE` to omit heading section. Can be set globally via `st_options`.

Display.labels

Logical. Should variable / data frame label be displayed in the title section? Default is `TRUE`. To change this default value globally, use `st_options`.

Split.tables

Pander argument that specifies how many characters wide a table can be. Inf by default.

DNN

Names to be used in output table. Vector of two strings; By default, the character values for arguments x and y are used.

Chisq

Logical. Display chisq statistic along with p-value.

OR

Logical or numeric. Display odds ratio with the specified confidence level (typically .95). Can be set to `TRUE`, in which case 95 interval is given. Confidence intervals are calculated using Wald’s method (normal approximation).

RR

Logical or numeric. Display risk ratio (also called relative risk) with the specified confidence level (typically .95). Can be set to `TRUE`, in which case 95 calculated using Wald’s method (normal approximation).

Weights

Vector of weights; must be of the same length as x.

Rescale.weights

Logical parameter. When set to `TRUE`, the total count will be the same as the unweighted x. FALSE by default.

... Additional arguments passed to `pander`.

Details

Rmarkdown does not, to this day, support multi-header tables. Therefore, until such support is available, the recommended way to display cross-tables in .Rmd documents is to use `method=render` with the `print()` or `view()` functions. See package vignettes for examples.

Value

A frequency table of classes `matrix` and `summarytools` with added attributes used by `print` method.

Author(s)

Dominic Comtois, <dominic.comtois@gmail.com>

See Also

table, xtabs
Examples

```r
data("tobacco")
ctable(tobacco$gender, tobacco$smoker)

# Use with() to simplify syntax
with(tobacco, ctable(smoker, diseased))

# Show column proportions, without totals
with(tobacco, ctable(smoker, diseased, prop = "c", totals = FALSE))

# Simple 2 x 2 table with odds ratio and risk ratio
with(tobacco, ctable(gender, smoker, totals = FALSE, headings = FALSE, prop = "n",
                      OR = TRUE, RR = TRUE))

# Grouped cross-tabulations
with(tobacco, stby(list(x = smoker, y = diseased), gender, ctable))

## Not run:
ct <- ctable(tobacco$gender, tobacco$smoker)

# Show html results in browser
print(ct, method = "browser")

# Save results to html file
print(ct, file = "ct_gender_smoker.html")

# Save results to text file
print(ct, file = "ct_gender_smoker.txt")

## End(Not run)
```

---

**define_keywords**

*Modify Keywords Used In Outputs*

**Description**

As an alternative to `use_custom_lang`, this allows temporarily modifying the keywords used in the outputs.

**Usage**

```r
define_keywords(...)```

**Arguments**

```r
... one or more pairs of keywords and their new values see Details.```
Details

On systems with GUI capabilities, a window will pop-up, allowing the modification of the custom column. The changes will be active as long as the package is loaded. A dialog will show up prompting the user to save the modified set of keywords in a custom csv language file.

Here is the full list of modifiable keywords. Check out the language_template.csv file in the package’s includes directory.

- **title.freq**: main heading for `freq()`
- **title.freq.weighted**: main heading for `freq()` (weighted)
- **title.ctable**: main heading for `ctable()`
- **title.ctable.weighted**: main heading `ctable()` (weighted)
- **title.ctable.row**: indicates what proportions are displayed
- **title.ctable.col**: indicates what proportions are displayed
- **title.ctable.tot**: indicates what proportions are displayed
- **title descr**: main heading for `descr()`
- **title.descr.weighted**: main heading for `descr()` (weighted)
- **title.dfSummary**: main heading for `dfSummary()`
- n heading item used in `descr()`
- **dimensions**: heading item used in `dfSummary()`
- **duplicates**: heading item used in `dfSummary()`
- **data.frame**: heading item (all functions)
- **label**: heading item (all functions) & column name in `dfSummary()`
- **variable**: heading item (all functions) & column name in `dfSummary()`
- **group**: heading item (all functions when used with `stby()`)
- **by**: heading item for `descr()` when used with `stby()`
- **weights**: heading item - `descr()` & `freq()`
- **type**: heading item for `freq()`
- **logical**: heading item - type in `freq()`
- **character**: heading item - type in `freq()`
- **numeric**: heading item - type in `freq()`
- **factor**: heading item - type in `freq()`
- **factor.ordered**: heading item - type in `freq()`
- **date**: heading item - type in `freq()`
- **datetime**: heading item - type in `freq()`
- **freq**: column name in `freq()`
- **pct**: column name in `freq()` when `report.nas=FALSE`
- **pct.valid.f**: column name in `freq()`
- **pct.valid.cum**: column name in `freq()`
\texttt{pct.total} column name in \texttt{freq()}
\texttt{pct.total.cum} column name in \texttt{freq()}
\texttt{pct.cum} column name in \texttt{freq()}
\texttt{valid} column name in \texttt{freq()} and \texttt{dfSummary()} & column content in \texttt{dfSummary()}
\texttt{invalid} column content in \texttt{dfSummary()} (emails)
\texttt{total} column grouping in \texttt{freq()}, html version
\texttt{mean} row name in \texttt{descr()}
\texttt{sd.long} row name in \texttt{descr()}
\texttt{sd} cell content (\texttt{dfSummary()})
\texttt{min} row name in \texttt{descr()}
\texttt{q1} row name in \texttt{descr()} - 1st quartile
\texttt{med} row name in \texttt{descr()}
\texttt{q3} row name in \texttt{descr()} - 3rd quartile
\texttt{max} row name in \texttt{descr()}
\texttt{mad} row name in \texttt{descr()} - Median Absolute Deviation
\texttt{iqr} row name in \texttt{descr()} - Inter-Quartile Range
\texttt{cv} row name in \texttt{descr()} - Coefficient of Variation
\texttt{skewness} row name in \texttt{descr()}
\texttt{se.skewness} row name in \texttt{descr()} - Std. Error for Skewness
\texttt{kurtosis} row name in \texttt{descr()}
\texttt{n.valid} row name in \texttt{descr()} - Count of non-missing values
\texttt{pct.valid} row name in \texttt{descr()} - pct. of non-missing values
\texttt{no} column name in \texttt{dfSummary()} - position of column in the data frame
\texttt{stats.values} column name in \texttt{dfSummary()}
\texttt{freqs.pct.valid} column name in \texttt{dfSummary()}
\texttt{graph} column name in \texttt{dfSummary()}
\texttt{missing} column name in \texttt{dfSummary()}
\texttt{distinct.value} cell content in \texttt{dfSummary()} - singular form
\texttt{distinct.values} cell content in \texttt{dfSummary()} - plural form
\texttt{all.nas} cell content in \texttt{dfSummary()} - column has only NAs
\texttt{all.empty.str} cell content in \texttt{dfSummary()} - column has only empty strings
\texttt{all.empty.str.nas} cell content in \texttt{dfSummary()} - col. has only NAs and empty strings
\texttt{no.levelsDefined} cell content in \texttt{dfSummary()} - factor has no levels defined
\texttt{int.sequence} cell content in \texttt{dfSummary()}
\texttt{rounded} cell content in \texttt{dfSummary()} - note appearing in Stats/Values
\texttt{others} cell content in \texttt{dfSummary()} - nbr of values not displayed
\texttt{codes} cell content in \texttt{dfSummary()} - When UPC codes are detected
**mode**  cell content in `dfSummary()` - mode = most frequent value

**med.short**  cell content in `dfSummary()` - median (shortened term)

**start**  cell content in `dfSummary()` - earliest date for date-type cols

**end**  cell content in `dfSummary()` - latest date for data-type cols

**emails**  cell content in `dfSummary()`

**generated.by**  footnote content

**version**  footnote content

**date.fmt**  footnote - date format (see `strptime`)

### Examples

```r
## Not run:
define_keywords(n = "Nb. Obs.")

## End(Not run)
```

---

**descr**  

*Univariate Statistics for Numerical Data*

### Description

Calculates mean, sd, min, Q1*, median, Q3*, max, MAD, IQR*, CV, skewness*, SE.skewness*, and kurtosis* on numerical vectors. (*) Not available when using sampling weights.

### Usage

```r
descr(
  x,
  var = NULL,
  stats = st_options("descr.stats"),
  na.rm = TRUE,
  round.digits = st_options("round.digits"),
  transpose = st_options("descr.transpose"),
  style = st_options("style"),
  plain.ascii = st_options("plain.ascii"),
  justify = "r",
  headings = st_options("headings"),
  display.labels = st_options("display.labels"),
  split.tables = 100,
  weights = NA,
  rescale.weights = FALSE,
  ...)
```

...
Arguments

- **x**: A numerical vector or a data frame.
- **var**: Unquoted expression referring to a specific column in `x`. Provides support for piped function calls (e.g., `df %>% descr(some_var)`).
- **stats**: Which stats to produce. Either “all” (default), “fivenum”, “common” (see Details), or a selection of: “mean”, “sd”, “min”, “q1”, “med”, “q3”, “max”, “mad”, “iqr”, “cv”, “skewness”, “se.skewness”, “kurtosis”, “n.valid”, and “pct.valid”. This can be set globally via `st_options` (“descr.stats”).
- **na.rm**: Argument to be passed to statistical functions. Defaults to TRUE. Can be set globally; see `st_options`.
- **round.digits**: Number of significant digits to display. Defaults to 2, and can be set globally (see `st_options`).
- **transpose**: Logical. Makes variables appear as columns, and stats as rows. Defaults to FALSE. To change this default value, see `st_options` (option “descr.transpose”).
- **style**: Style to be used by `pander` when rendering output table; One of “simple” (default), “grid”, or “rmarkdown” This option can be set globally; see `st_options`.
- **plain.ascii**: Logical. `pander` argument; when TRUE, no markup characters will be used (useful when printing to console). Defaults to TRUE unless style = 'rmarkdown', in which case it will be set to FALSE automatically. To change the default value globally, see `st_options`.
- **justify**: Alignment of numbers in cells; “l” for left, “c” for center, or “r” for right (default). Has no effect on html tables.
- **headings**: Logical. Set to FALSE to omit heading section. Can be set globally via `st_options`. TRUE by default.
- **display.labels**: Logical. Should variable / data frame labels be displayed in the title section? Default is TRUE. To change this default value globally, see `st_options`.
- **split.tables**: Pander argument that specifies how many characters wide a table can be. 100 by default.
- **weights**: Vector of weights having same length as `x`. NA (default) indicates that no weights are used.
- **rescale.weights**: Logical. When set to TRUE, the total count will be the same as the unweighted `x`. FALSE by default.
- **...**: Additional arguments passed to `pander`.

Value

An object having classes `matrix` and `summarytools` containing the statistics, with extra attributes used by `print` method.

Author(s)

Dominic Comtois, <dominic.comtois@gmail.com>
dfSummary

Examples

```r
data("exams")

# All stats for all numerical variables
descr(exams)

# Only common statistics
descr(exams, stats = "common")

# Arbitrary selection of statistics, transposed
descr(exams, stats = c("mean", "sd", "min", "max"), transpose = TRUE)

# Rmarkdown-ready
descr(exams, plain.ascii = FALSE, style = "rmarkdown")

# Grouped statistics
data("tobacco")
with(tobacco, stby(BMI, gender, descr))

# Grouped statistics, transposed
with(tobacco, stby(BMI, age.gr, descr, stats = "common", transpose = TRUE))

## Not run:
# Show in Viewer (or browser if not in RStudio)
view(descr(exams))

# Save to html file with title
print(descr(exams),
      file = "descr_exams.html",
      report.title = "BMI by Age Group",
      footnote = "<b>Schoolyear:</b> 2018-2019<br/><b>Semester:</b> Fall")

## End(Not run)
```

---

**dfSummary**

*Data frame Summary*

**Description**

Summary of a data frame consisting of: variable names, labels if any, factor levels, frequencies and/or numerical summary statistics, and valid/missing observation counts.

**Usage**

```r
dfSummary(
  x,
  round.digits = st_options("round.digits"),
```
Arguments

- **x**  
  A data frame.

- **round.digits**  
  Number of significant digits to display. Defaults to 2 and can be set globally; see **st_options**.

- **varnumbers**  
  Logical. Should the first column contain variable number? Defaults to TRUE. Can be set globally; see **st_options**, option “dfSummary.varnumbers”.

- **labels.col**  
  Logical. If TRUE, variable labels (as defined with **rapportools, Hmisc or summarytools** label functions) will be displayed. TRUE by default, but the **labels** column is only shown if at least one column has a defined label. This option can also be set globally; see **st_options**, option “dfSummary.labels.col”.

- **valid.col**  
  Logical. Include column indicating count and proportion of valid (non-missing) values. TRUE by default, but can be set globally; see **st_options**, option “dfSummary.valid.col”.

- **na.col**  
  Logical. Include column indicating count and proportion of missing (NA) values. TRUE by default, but can be set globally; see **st_options**, option “dfSummary.na.col”.

- **graph.col**  
  Logical. Display barplots / histograms column in html reports. TRUE by default, but can be set globally; see **st_options**, option “dfSummary.graph.col”.

- **graph.magnif**  
  Numeric. Magnification factor, useful if the graphs show up too large (then use a value < 1) or too small (use a value > 1). Must be positive. Default to 1. Can be set globally; see **st_options**, option “dfSummary.graph.magnif”.

- **style**  
  Style to be used by **pander** when rendering output table. Defaults to “multiline”. The only other valid option is “grid”. Style “simple” is not supported for this particular function, and “rmarkdown” will fallback to “multiline”.

- **...**
**dfSummary**

plain.ascii Logical. `pander` argument; when TRUE, no markup characters will be used (useful when printing to console). Defaults to TRUE. Set to FALSE when in context of markdown rendering. To change the default value globally, see `st_options`.

justify String indicating alignment of columns; one of “l” (left), “c” (center), or “r” (right). Defaults to “l”.

col.widths Numeric or character. Vector of column widths. If numeric, values are assumed to be numbers of pixels. Otherwise, any CSS-supported units can be used. NA by default, meaning widths are calculated automatically.

headings Logical. Set to FALSE to omit headings. To change this default value globally, see `st_options`.

display.labels Logical. Should data frame label be displayed in the title section? Default is TRUE. To change this default value globally, see `st_options`.

max.distinct.values The maximum number of values to display frequencies for. If variable has more distinct values than this number, the remaining frequencies will be reported as a whole, along with the number of additional distinct values. Defaults to 10.

trim.strings Logical; for character variables, should leading and trailing white space be removed? Defaults to FALSE. See details section.

max.string.width Limits the number of characters to display in the frequency tables. Defaults to 25.

split.cells A numeric argument passed to `pander`. It is the number of characters allowed on a line before splitting the cell. Defaults to 40.

split.tables `pander` argument which determines the maximum width of a table. Keeping the default value (`Inf`) is recommended.

tmp.img.dir Character. Directory used to store temporary images when rendering `dfSummary()` with `method = "pander"`, `plain.ascii = TRUE` and `style = "grid"`. See Details.

silent Logical. Hide console messages. FALSE by default. To change this value globally, see `st_options`.

... Additional arguments passed to `pander`.

**Details**

The default `plain.ascii = TRUE` option is there to make results appear cleaner in the console. When used in a context of `markdown` rendering, set this option to FALSE.

When the `trim.strings` is set to TRUE, trimming is done before calculating frequencies, so those will be impacted accordingly.

Specifying `tmp.img.dir` allows producing results consistent with pandoc styling while also showing png graphs. Due to the fact that in Pandoc, column widths are determined by the length of cell contents even if said content is merely a link to an image, we cannot use the standard R temporary directory to store the images. We need a shorter path; on Mac OS and Linux, using “/tmp” is a sensible choice, since this directory is cleaned up automatically on a regular basis. On Windows however, there is no such convenient directory and the user will have to choose a directory and cleanup the temporary images manually after the document has been rendered. Providing a relative path such as “img” is recommended. The maximum length for this parameter is set to 5 characters. It can be set globally using `st_options`; for example: `st_options(tmp.img.dir = ".\")`. 


**Value**

A data frame with additional class `summarytools` containing as many rows as there are columns in `x`, with attributes to inform `print` method. Columns in the output data frame are:

- **No**: Number indicating the order in which column appears in the data frame.
- **Variable**: Name of the variable, along with its class(es).
- **Label**: Label of the variable (if applicable).
- **Stats / Values**: For factors, a list of their values, limited by the `max.distinct.values` parameter. For character variables, the most common values (in descending frequency order), also limited by `max.distinct.values`. For numerical variables, common univariate statistics (mean, std. deviation, min, med, max, IQR and CV).
- **Freqs (% of Valid)**: For factors and character variables, the frequencies and proportions of the values listed in the previous column. For numerical vectors, number of distinct values, or frequency of distinct values if their number is not greater than `max.distinct.values`.
- **Text Graph**: An ascii histogram for numerical variables, and ascii barplot for factors and character variables.
- **Valid**: Number and proportion of valid values.
- **Missing**: Number and proportion of missing (NA and NAN) values.

**Author(s)**

Dominic Comtois, <dominic.comtois@gmail.com>

**Examples**

```r
data("tobacco")
saved_x11_option <- st_options("use.x11")
st_options(use.x11 = FALSE)
dfSummary(tobacco)

# Exclude some columns
dfSummary(tobacco, varnumbers = FALSE, valid.col = FALSE)

# Limit number of categories to be displayed for factors / categorical data
dfSummary(tobacco, max.distinct.values = 5, style = "grid")

# Using stby()
stby(tobacco, tobacco$gender, dfSummary)

st_options(use.x11 = saved_x11_option)

## Not run:
# Show in Viewer or browser (view: no capital V!)
view(dfSummary(iris))

# Rmarkdown-ready
dfSummary(tobacco, style = "rmarkdown", plain.ascii = TRUE,
```
examens

varnumbers = FALSE, valid.col = FALSE, tmp.img.dir = "/img")

# Using group_by()
tobacco %>% group_by(gender) %>% dfSummary()

## End(Not run)

---

### Description

Jeu de données simulées contenant les notes de 30 étudiants, avec les colonnes suivantes:

- **étudiant** Nom de l’étudiant.
- **sexe** Variable categorielle (facteur). Deux niveaux: “Fille”, “Garçon”.
- **francais** Note en français (numérique).
- **math** Note en maths (numérique).
- **geographie** Note en géographie (numérique).
- **histoire** Note en histoire (numérique).
- **economie** Note en économie (numérique).
- **anglais** Note en anglais (numérique).

### Usage

data(examens)

### Format

Un data frame de 30 rangees et 8 colonnes

### Details

Donnees simulées. Les notes de chaque étudiant sont centeres autour d’une moyenne personnelle et ecart-type randomises. 

A copy of this dataset is available in English under the name “exams”.

---
Description
A simulated dataset with grades for hypothetical 30 students, with the following variables:

- student Student’s name.
- gender Factor with 2 levels: “Girl”, “Boy”.
- french French Grade (numerical).
- math Math Grade (numerical).
- geography Geography Grade (numerical).
- history History Grade (numerical).
- economics Economics Grade (numerical).
- english English Grade (numerical).

Usage

data(exams)

Format
A data frame with 30 rows and 8 variables

Details
All names and grades are simulated. Grades for each student are centered around a personal randomized average and standard deviation.

A copy of this dataset is also available in French under the name “examens”.

Description
Displays weighted or unweighted frequencies, including <NA> counts and proportions.
freq

Usage

freq(
  x,
  var = NULL,
  round.digits = st_options("round.digits"),
  order = "default",
  style = st_options("style"),
  plain.ascii = st_options("plain.ascii"),
  justify = "default",
  cumul = st_options("freq.cumul"),
  totals = st_options("freq.totals"),
  report.nas = st_options("freq.report.nas"),
  rows = numeric(),
  missing = "",
  display.type = TRUE,
  display.labels = st_options("display.labels"),
  headings = st_options("headings"),
  weights = NA,
  rescale.weights = FALSE,
  ...
)

Arguments

x Factor or vector, or data frame when y is also provided, usually in a piped call.
var Unquoted expression referring to a specific column in x. Provides support for
piped function calls (e.g. df %>% freq(some_var).
round.digits Number of significant digits to display. Defaults to 2 and can be set globally; see st_options.
order Ordering of rows in frequency table; “names” (default for non-factors), “levels”
debug for factors), or “freq” (from most frequent to less frequent). To invert
the order, place a minus sign before or after the word. “-freq” will thus display
the items starting from the lowest in frequency to the highest, and so forth.
style Style to be used by pander when rendering output table; One of “simple” (de-
default), “grid”, or “rmarkdown” This option can be set globally; see st_options.
plain.ascii Logical. pander argument; when TRUE, no markup characters will be used (use-
ful when printing to console). Defaults to TRUE unless style = 'rmarkdown',
in which case it will be set to FALSE automatically. To change the default value
globally, see st_options.
justify String indicating alignment of columns. By default (“default”), “right” is used
for text tables and “center” is used for html tables. You can force it to one of
“left”, “center”, or “right”.
cumul Logical. Set to FALSE to hide cumulative proportions from results. TRUE by
default. To change this value globally, see st_options.
totals Logical. Set to FALSE to hide totals from results. TRUE by default. To change
this value globally, see st_options.
freq

report.nas Logical. Set to FALSE to turn off reporting of missing values. To change this default value globally, see st_options.

rows Character or numeric vector allowing subsetting of the results. The order given here will be reflected in the resulting table. If a single string is used, it will be used as a regular expression to filter row names.

missing Characters to display in NA cells. Defaults to "".

display.type Logical. Should variable type be displayed? Default is TRUE.

display.labels Logical. Should variable / data frame labels be displayed? Default is TRUE. To change this default value globally, see st_options.

headings Logical. Set to FALSE to omit heading section. Can be set globally via st_options.

weights Vector of weights; must be of the same length as x.

rescale.weights Logical parameter. When set to TRUE, the total count will be the same as the unweighted x. FALSE by default.

... Additional arguments passed to pander.

Details

The default plain.ascii = TRUE option is there to make results appear cleaner in the console. To avoid rmarkdown rendering problems, this option is automatically set to FALSE whenever style = "rmarkdown" (unless plain.ascii = TRUE is made explicit in the function call).

Value

A frequency table of class matrix and summarytools with added attributes used by print method.

Author(s)

Dominic Comtois, <dominic.comtois@gmail.com>

See Also

table

Examples

data(tobacco)
freq(tobacco$gender)
freq(tobacco$gender, totals = FALSE)

# Ignore NA's, don't show totals, omit headings
freq(tobacco$gender, report.nas = FALSE, totals = FALSE, headings = FALSE)

# In .Rmd documents, use the two following arguments, minimally
freq(tobacco$gender, style="rmarkdown", plain.ascii = FALSE)

# Grouped Frequencies
with(tobacco, stby(diseased, smoker, freq))
(fr_smoker_by_gender <- with(tobacco, stby(smoker, gender, freq)))

# Print html Source
print(fr_smoker_by_gender, method = "render", footnote = NA)

# Order by frequency (+ to -)
freq(tobacco$age.gr, order = "freq")

# Order by frequency (- to +)
freq(tobacco$age.gr, order = "-freq")

# Use the 'rows' argument to display only the 10 most common items
freq(tobacco$age.gr, order = "freq", rows = 1:10)

## Not run:
# Display rendered html results in RStudio's Viewer
# notice 'view()' is NOT written with capital V
# If working outside RStudio, Web browser is used instead
# A temporary file is stored in temp dir
view(fr_smoker_by_gender)

# Display rendered html results in default Web browser
# A temporary file is stored in temp dir here too
print(fr_smoker_by_gender, method = "browser")

# Write results to text file (.txt, .md, .Rmd) or html file (.html)
print(fr_smoker_by_gender, method = "render", file = "fr_smoker_by_gender.md")
print(fr_smoker_by_gender, method = "render", file = "fr_smoker_by_gender.html")

## End(Not run)

---

**label**

*Get or Set Variable or Data Frame Labels*

**Description**

Assigns a label to a vector or data frame, or returns value stored in the object’s *label* attribute (or NA if none exists).

**Usage**

```r
label(x, all = FALSE, fallback = FALSE, simplify = FALSE)
label(x) <- value
```

**Arguments**

- `x` An R object to extract labels from.
print.list

all Logical. When x is a data frame, setting this argument to TRUE will make the function return all variable labels. By default, its value is FALSE, so that if x is a data frame, it is the data frame’s label that will be returned.

fallback a logical value indicating if labels should fallback to object name(s). Defaults to FALSE.

simplify When x is a data frame and all = TRUE, coerce results to a vector and remove NA’s. Default is FALSE.

value String to be used as label. To clear existing labels, use NA or NULL.

Note

Loosely based on Gergely Daróczi’s label function.

Author(s)

Dominic Comtois, <dominic.comtois@gmail.com>,

print.list Print Method for Objects of Class “list”

Description

Displays a list comprised of summarytools objects created with lapply.

Usage

## S3 method for class ’list’
print(x, method = "pander", file = "", append = FALSE, report.title = NA, table.classes = NA, bootstrap.css = st_options(’bootstrap.css’), custom.css = st_options(’custom.css’), silent = FALSE, footnote = st_options(’footnote’), collapse = 0, escape.pipe = st_options(’escape.pipe’), ...)

Arguments

x A summarytools object that was generated with freq, descr, ctable or dfSummary.

method One of “pander”, “viewer”, “browser”, or “render”. For print(), default is “pander”; for view(), default is “viewer”. If “viewer” is used outside RStudio, “browser” will be used instead. Use “render” if function is called from an Rmd document.

file File name to write output to. Defaults to “”.

append Logical. When file argument is supplied, this indicates whether to append output to existing file. FALSE by default.

report.title For html reports, this goes into the <title> tag. Defaults to NA, in which case <title> will be generic.
print.stby

Description

Displays a list comprised of summarytools objects created with stby.

Usage

```r
## S3 method for class 'stby'
print(x, method = "pander", file = "",
      append = FALSE, report.title = NA, table.classes = NA,
      bootstrap.css = st_options('bootstrap.css'),
      custom.css = st_options('custom.css'), silent = FALSE,
      footnote = st_options('footnote'),
      escape.pipe = st_options('escape.pipe'), ...)
```

Arguments

- `x` A summarytools object that was generated with `freq, descr, ctable` or `dfSummary`.
- `method` One of “pander”, “viewer”, “browser”, or “render”. For `print()`, default is “pander”; for `view()`, default is “viewer”. If “viewer” is used outside RStudio, “browser” will be used instead. Use “render” if function is called from an Rmd document.
file
File name to write output to. Defaults to "".

append
Logical. When file argument is supplied, this indicates whether to append output to existing file. FALSE by default.

report.title
For html reports, this goes into the <title> tag. Defaults to NA, in which case <title> will be generic.

table.classes
Character. Additional classes to assign to output tables. All Bootstrap css classes can be used. It also allows user-defined classes (see custom.css parameter). See details section. NA by default.

bootstrap.css
Logical. Set to FALSE to omit Bootstrap css. TRUE by default. To change this default value globally, see st_options.

custom.css
Path to a user-defined .css file. Classes defined in this file can be used in the table.classes parameter. NA by default. To change this default value globally, see st_options.

silent
Hide console messages (such as ignored variables or NaN to NA transformations).

footnote
footnote in html output. When set to “default”, this is the package name and version, R version, and current date). Has no effect when method is “pander”. Set to “default”, provide your own text, or set to NA to omit. To change this default value globally, see st_options.

escape.pipe
Logical. Set to TRUE when using style='grid' and file argument is supplied if the intent is to generate a text file that can be converted to other formats using Pandoc. To change this default value globally, see st_options.

... Additional arguments can be used to override parameters stored as attributes in the object being printed. See Details section.

Description
Display summarytools objects in the console, in Web Browser or in RStudio’s Viewer, or write content to file.

Usage

## S3 method for class 'summarytools'
print(x, method = "pander", file = "", append = FALSE, report.title = NA, table.classes = NA, bootstrap.css = st_options('bootstrap.css'), custom.css = st_options('custom.css'), silent = FALSE, footnote = st_options('footnote'), max.tbl.height = Inf, collapse = 0, escape.pipe = st_options("escape.pipe"), ...)
Arguments

- **x**: A `summarytools` object that was generated with `freq`, `descr`, `ctable` or `dfSummary`.
- **method**: One of “pander”, “viewer”, “browser”, or “render”. For `print()`, default is “pander”; for `view()`, default is “viewer”. If “viewer” is used outside `RStudio`, “browser” will be used instead. Use “render” if function is called from an Rmd document.
- **file**: File name to write output to. Defaults to “”.
- **append**: Logical. When `file` argument is supplied, this indicates whether to append output to existing file. FALSE by default.
- **report.title**: For html reports, this goes into the `<title>` tag. Defaults to NA, in which case `<title>` will be generic.
- **table.classes**: Character. Additional classes to assign to output tables. All Bootstrap css classes can be used. It also allows user-defined classes (see `custom.css` parameter). See details section. NA by default.
- **bootstrap.css**: Logical. Set to FALSE to omit Bootstrap css. TRUE by default. To change this default value globally, see `st_options`.
- **custom.css**: Path to a user-defined `.css` file. Classes defined in this file can be used in the `table.classes` parameter. NA by default. To change this default value globally, see `st_options`.
- **silent**: Hide console messages (such as ignored variables or NaN to NA transformations).
- **footnote**: footnote in html output. When set to “default”, this is the package name and version, R version, and current date). Has no effect when `method` is “pander”. Set to “default”, provide your own text, or set to NA to omit. To change this default value globally, see `st_options`.
- **max.tbl.height**: Maximum table height (in pixels) allowed in rendered `dfSummary()` tables. When this argument is used, results will show up in a `<div>` with the specified height and a scroll bar. Intended to be used in Rmd documents. Has no effect when `method` is “pander”. Inf by default.
- **collapse**: Numeric. 0 by default. Set to 1 to make `freq()` sections collapsible (when clicking on the variable name). Future versions might provide alternate collapsing options.
- **escape.pipe**: Logical. Set to TRUE when using style='grid' and file argument is supplied if the intent is to generate a text file that can be converted to other formats using Pandoc. To change this default value globally, see `st_options`.
- **...**: Additional arguments can be used to override parameters stored as attributes in the object being printed. See Details section.

Details

Plain ascii and markdown tables are generated via `pander`. See References section for a list of all available `pander` options.

The following additional arguments can be used to override formatting attributes stored in the object to be printed. Refer to the function’s documentation for details on these arguments.
• style
• round.digits (except for dfSummary objects)
• plain.ascii
• justify
• headings
• display.labels
• varnumbers (dfSummary objects)
• labels.col (dfSummary objects)
• graph.col (dfSummary objects)
• valid.col (dfSummary objects)
• na.col (dfSummary objects)
• col.widths (dfSummary objects)
• split.tables
• report.nas (freq objects)
• display.type (freq objects)
• missing (freq objects)
• totals (freq and ctable objects)
• caption (freq and ctable objects)

The following additional arguments can be used to override heading elements to be printed:

• Data.frame
• Data.frame.label
• Variable
• Variable.label
• Group
• date
• Weights (freq & descr objects)
• Data.type (freq objects)
• Row.variable (ctable objects)
• Col.variable (ctable objects)

Value

NULL when method="pander"; a file path (returned invisibly) when method="viewer" or method="browser". In the latter case, the file path is also passed to shell.exec so the document is opened in default Web Browser.

Author(s)

Dominic Comtois, <dominic.comtois@gmail.com>
stby

Obtain Grouped Statistics With summarytools

Description

This is essentially the base by function, except for the class of the returned object.

Usage

stby(data, INDICES, FUN, ..., simplify = TRUE)

Arguments

data an R object, normally a data frame, possibly a matrix.
INDICES a grouping variable or a list of grouping variables, each of length nrow(data).
FUN a function to be applied to (usually data-frame) subsets of data.
... Further arguments to FUN.
simplify Logical. Essentially a placeholder to maintain full compatibility with base by. For more details, see tapply.

Value

An object having classes “list” and “summarytools”.

See Also

by, tapply

References

RStudio Summarytools on GitHub List of pander options on GitHub Bootstrap Cascading Stylesheets

See Also

pander

Examples

## Not run:
data(tobacco)
view(dfSummary(tobacco), footnote = NA)
## End(Not run)
data(exams)
print(freq(exams$gender), style = 'rmarkdown')
print(descr(exams), headings = FALSE)
Examples

```r
data("tobacco")
with(tobacco, stby(BMI, gender, descr))
```

---

**st_css**

*Include summarytools' css Into Active Document*

---

**Description**

Generates the `css` needed by `summarytools` in *Rmarkdown* documents.

**Usage**

```r
st_css(main = TRUE, global = FALSE, bootstrap = FALSE, style.tag = TRUE, ...)
```

**Arguments**

- **main** Logical. Include `summarytools.css` file. TRUE by default. Affects only `summarytools` objects.
- **global** Logical. Include the additional `summarytools-global.css` file, which affects all content in the document. Provides control over objects that were not `html-rendered`; in particular, table widths and vertical alignment are modified to improve layout. FALSE by default.
- **bootstrap** Logical. Include `bootstrap.min.css`. FALSE by default.
- **style.tag** Logical. Includes the opening and closing `<style>` tags. TRUE by default.
- **...** Character. Path to additional `css` file(s) to include.

**Details**

Typically the function is called right after the initial setup chunk of an *Rmarkdown* document, in a chunk having options `echo=FALSE` and `results="asis"`.

**Author(s)**

Dominic Comtois, <dominic.comtois@gmail.com>
**st_options**  
Query and set summarytools global options

**Description**

To list all summarytools global options, call without arguments. To display the value of one or several options, enter the name(s) of the option(s) in a character vector as sole argument. To **reset** all options, use single unnamed argument 'reset' or 0.

**Usage**

```r
st_options(
  option = NULL,
  value = NULL,
  style = "simple",
  plain.ascii = TRUE,
  round.digits = 2,
  headings = TRUE,
  footnote = "default",
  display.labels = TRUE,
  bootstrap.css = TRUE,
  custom.css = NA,
  escape.pipe = FALSE,
  freq.cumul = TRUE,
  freq.totals = TRUE,
  freq.report.nas = TRUE,
  freq.ignore.threshold = 25,
  freq.silent = FALSE,
  ctable.prop = "r",
  ctable.totals = TRUE,
  descr.stats = "all",
  descr.transpose = FALSE,
  descr.silent = FALSE,
  dfSummary.style = "multiline",
  dfSummary.varnumbers = TRUE,
  dfSummary.labels.col = TRUE,
  dfSummary.valid.col = TRUE,
  dfSummary.na.col = TRUE,
  dfSummary.graph.col = TRUE,
  dfSummary.graph.magnif = 1,
  dfSummary.silent = FALSE,
  tmp.img.dir = NA,
  subtitle.emphasis = TRUE,
  lang = "en",
  use.x11 = TRUE
)
```
## Arguments

<table>
<thead>
<tr>
<th>Option</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>option</td>
<td>option(s) name(s) to query (optional). Can be a single string or a vector of strings to query multiple values.</td>
</tr>
<tr>
<td>value</td>
<td>The value you wish to assign to the option specified in the first argument. This is for backward-compatibility, as all options can now be set via their own parameter. That is, instead of <code>st_options('plain.ascii',FALSE))</code>, use <code>st_options(plain.ascii = FALSE)</code>.</td>
</tr>
<tr>
<td>style</td>
<td>Character. One of “simple” (default), “rmarkdown”, or “grid”. Does not apply to <code>dfSummary</code>.</td>
</tr>
<tr>
<td>plain.ascii</td>
<td>Logical. TRUE by default. Set to FALSE when using summarytools with a rendering tool such as <code>knitr</code> or when creating rmarkdown output files to be converted with Pandoc. Note however that its value will automatically be set to FALSE whenever <code>style</code> is set to “rmarkdown”).</td>
</tr>
<tr>
<td>round.digits</td>
<td>Numeric. Defaults to 2.</td>
</tr>
<tr>
<td>headings</td>
<td>Logical. Set to FALSE to remove all headings from outputs. Only the tables will be printed out, except when <code>by</code> or <code>lapply</code> are used. In that case, the variable or the group will still appear before the tables. FALSE by default.</td>
</tr>
<tr>
<td>footnote</td>
<td>Character. When the default value “default” is used, the package name &amp; version, as well as the R version number are displayed below <code>html</code> outputs. Set no NA to omit the footnote, or provide a custom string. Applies only to <code>html</code> outputs.</td>
</tr>
<tr>
<td>display.labels</td>
<td>Logical. TRUE by default. Set to FALSE to omit data frame and variable labels in the headings section.</td>
</tr>
<tr>
<td>bootstrap.css</td>
<td>Logical. Specifies whether to Include Bootstrap css in <code>html</code> reports <code>head</code> section outputs. Defaults to TRUE. Set to FALSE when using the “render” method inside a shiny app to avoid interacting with the app’s layout.</td>
</tr>
<tr>
<td>custom.css</td>
<td>Character. Path to an additional, user-provided, CSS file. NA by default.</td>
</tr>
<tr>
<td>escape.pipe</td>
<td>Logical. Set to TRUE if Pandoc conversion is your goal and you have unsatisfying results with grid or multiline tables. FALSE by default.</td>
</tr>
<tr>
<td>freq.cumul</td>
<td>Logical. Corresponds to the <code>cumul</code> parameter of <code>freq</code>. TRUE by default.</td>
</tr>
<tr>
<td>freq.totals</td>
<td>Logical. Corresponds to the <code>totals</code> parameter of <code>freq</code>. TRUE by default.</td>
</tr>
<tr>
<td>freq.report.nas</td>
<td>Logical. Corresponds to the <code>display.nas</code> parameter of <code>freq</code>. TRUE by default.</td>
</tr>
<tr>
<td>freq.ignore.threshold</td>
<td>Numeric. Number of distinct values above which numerical variables are ignored when calling <code>freq</code> with a whole data frame as main argument. Defaults to 25.</td>
</tr>
<tr>
<td>freq.silent</td>
<td>Logical. Hide console messages. FALSE by default.</td>
</tr>
<tr>
<td>ctable.prop</td>
<td>Character. Corresponds to the <code>prop</code> parameter of <code>ctable</code>. Defaults to “r” (row).</td>
</tr>
<tr>
<td>ctable.totals</td>
<td>Logical. Corresponds to the <code>totals</code> parameter of <code>ctable</code>. TRUE by default.</td>
</tr>
<tr>
<td>descr.stats</td>
<td>Character. Corresponds to the <code>stats</code> parameter of <code>descr</code>. Defaults to “all”.</td>
</tr>
<tr>
<td>descr.transpose</td>
<td>Logical. Corresponds to the <code>transpose</code> parameter of <code>descr</code>. FALSE by default.</td>
</tr>
</tbody>
</table>
**st_options**

```

descr.silent Logical. Hide console messages. FALSE by default.
dfSummary.varnumbers Logical. In dfSummary, display variable numbers in the first column. Defaults to TRUE.
dfSummary.labels.col Logical. In dfSummary, display variable labels Defaults to TRUE.
dfSummary.valid.col Logical. In dfSummary, include column indicating count and proportion of valid (non-missing). TRUE by default.
dfSummary.na.col Logical. In dfSummary, include column indicating count and proportion of missing (NA) values. TRUE by default.
dfSummary.graph.col Logical. Display barplots / histograms column in dfSummary html reports. TRUE by default.
dfSummary.graph.magnif Numeric. Magnification factor, useful if dfSummary graphs show up too large (then use a value between 0 and 1) or too small (use a value > 1). Must be positive. Default to 1.
dfSummary.silent Logical. Hide console messages. FALSE by default.
tmp.img.dir Character. Directory used to store temporary images. See Details section of dfSummary. NA by default.
subtitle.emphasis Logical. Controls the formatting of the “subtitle” (the data frame or variable name, depending on context. When TRUE (default), “h4” is used, while with FALSE, “bold” / “strong” is used. Hence the default value gives it stronger emphasis.
use.x11 Logical. TRUE by default. In console-only environments, setting this to FALSE will prevent errors occurring when dfSummary tries to generate html “Base64-encoded” graphs.
```

**Details**

To learn more about summarytools options, see the project’s GitHub page.

**Examples**

```
## Not run:
# show all summarytools global options
st_options()

# show a specific option
```
tabagisme

Usage du Tabac et etat de Sante (Donnees simulees)

Description

Jeu de donnees simulees de 1000 sujets, avec les colonnes suivantes:

- age Numerique.
- age.gr Groupe d’age - variable categorielle, 4 niveaux.
- IMC Indice de masse corporelle (numerique).
- fumeur Variable categorielle, 2 niveaux (“Oui” / “Non”).
- cigs.par.jour Nombre de cigarettes fumees par jour (numerique).
- malade Variable categorielle, 2 niveaux (“Oui” / “Non”).
- maladie Champs texte.
- ponderation Poids echantillonal (numerique).

Usage

data(tabagisme)

Format

Un data frame de 1000 rangees et 9 colonnes
Details

Note sur la simulation des données: la probabilité pour un sujet de tomber dans la catégorie “malade” est basée sur une fonction arbitraire faisant intervenir l’âge, l’IMC et le nombre de cigarettes fumées par jour.

A copy of this dataset is available in English under the name “tobacco”.

---

**tb**

*Convert Summarytools Objects into Tibbles*

**Description**

Make a tidy dataset out of freq() or descr() outputs

**Usage**

```r
tb(x, order = 1, na.rm = FALSE, drop.var.col = FALSE)
```

**Arguments**

- `x` a freq() or descr() output object.
- `order` Integer. Useful for grouped results produced with stby or dplyr::group_by. When set to 1 (default), the ordering is done using the grouping variables first. When set to 2, the ordering is first determined by the variable column for descr or the column displaying the variable values for freq. When set to 3, the same ordering as with 2 is used, but columns are rearranged to reflect this sort order.
- `na.rm` Logical. For freq objects, remove <NA> rows (or Missing) rows if NA values were made explicit with forcats::fct_explicit_na(). Has no effect on descr objects.
- `drop.var.col` Logical. For descr objects, drop the variable column. This is possible only when statistics are produced for a single variable; for multiple variables, this argument is ignored. FALSE by default.

**Value**

A tibble which is constructed following the tidy principles.

**Examples**

```r
tb(freq(iris$Species))
tb(descr(iris))
data("tobacco")
tb(stby(tobacco, tobacco$gender, descr))
```
**Tobacco Use and Health - Simulated Dataset**

**Description**

A simulated datasets of 1,000 subjects, with the following variables:

**Usage**

```r
data(tobacco)
```

**Format**

A data frame with 1000 rows and 9 variables

**Details**

- gender Factor with 2 levels: “F” and “M”, having roughly 500 of each.
- age Numerical.
- age.gr Factor with 4 age categories.
- BMI Body Mass Index (numerical).
- smoker Factor (“Yes” / “No”).
- cigs.per.day Number of cigarettes smoked per day (numerical).
- diseased Factor (“Yes” / “No”).
- disease Character.
- samp.wgts Sampling weights (numerical).

A note on simulation: probability for an individual to fall into category “diseased” is based on an arbitrary function involving age, BMI and number of cigarettes per day.

A copy of this dataset is also **available in French** under the name “tabagisme”.

---

**Clear Variable and Data Frame Label(s)**

**Description**

Returns the object with all labels removed. Both the “label” attribute and Hmisc’s “labelled” class are removed.

**Usage**

```r
unlabel(x)
```
use_custom_lang

Arguments

x  An R object to remove labels from.

Author(s)

Dominic Comtois, <dominic.comtois@gmail.com>,

See Also

label

Description

If your language is not available or if you wish to customize the outputs’ language to suit your preference, you can set up a translations file (see details) and import it with this function.

Usage

use_custom_lang(file)

Arguments

file  Character. The path to the translations file.

Details

To build the translations file, copy the language_template.csv file located in the installed package’s includes directory and fill out the ‘custom’ column using a text editor, leaving column titles unchanged. The file must also retain its UTF-8 encoding.

Description

Visualize results in RStudio’s Viewer or in Web Browser
Usage

view(x, method = "viewer", file = ",", append = FALSE,
    report.title = NA, table.classes = NA,
    bootstrap.css = st_options("bootstrap.css"),
    custom.css = st_options("custom.css"), silent = FALSE,
    footnote = st_options("footnote"),
    max.tbl.height = Inf,
    collapse = 0,
    escape.pipe = st_options("escape.pipe"), ...)

Arguments

x A summarytools object that was generated with freq, descr, ctable or dfSummary.
method One of "pander", "viewer", "browser", or "render". For print(), default is "pander"; for view(), default is "viewer". If "viewer" is used outside RStudio, "browser" will be used instead. Use "render" if function is called from an Rmd document.
file File name to write output to. Defaults to ".".
append Logical. When file argument is supplied, this indicates whether to append output to existing file. FALSE by default.
report.title For html reports, this goes into the <title> tag. Defaults to NA, in which case <title> will be generic.
table.classes Character. Additional classes to assign to output tables. All Bootstrap css classes can be used. It also allows user-defined classes (see custom.css parameter). See details section. NA by default.
bootstrap.css Logical. Set to FALSE to omit Bootstrap css. TRUE by default. To change this default value globally, see st_options.
custom.css Path to a user-defined .css file. Classes defined in this file can be used in the table.classes parameter. NA by default. To change this default value globally, see st_options.
silent Hide console messages (such as ignored variables or NaN to NA transformations).
footnote footnote in html output. When set to "default", this is the package name and version, R version, and current date). Has no effect when method is "pander". Set to "default", provide your own text, or set to NA to omit. To change this default value globally, see st_options.
max.tbl.height Maximum table height (in pixels) allowed in rendered dfSummary() tables. When this argument is used, results will show up in a <div> with the specified height and a scroll bar. Intended to be used in Rmd documents. Has no effect when method is "pander". Inf by default.
collapse Numeric. 0 by default. Set to 1 to make freq() sections collapsible (when clicking on the variable name). Future versions might provide alternate collapsing options.
escape.pipe Logical. Set to TRUE when using style='grid' and file argument is supplied if the intent is to generate a text file that can be converted to other formats using Pandoc. To change this default value globally, see st_options.
Additional arguments can be used to override parameters stored as attributes in the object being printed. See Details section.

Details

Creates html outputs and opens them in the Viewer, in a browser or renders the html code appropriate for Rmarkdown documents.

For objects of class "summarytools", this function is simply a wrapper around print.summarytools with method set to "viewer".

Objects of class “by” or “list” are dispatched to the present function, as it can manage multiple objects, whereas print.summarytools can only manage one object at a time.

what.is

Obtain Extended Properties of Objects

Description

Combination of most common “macro-level” functions that describe an object.

Usage

what.is(x, show.all = FALSE, ignore.size.warn = FALSE)

Arguments

- **x** Any object.
- **show.all** Logical. When TRUE, all logical results from the “is.” identifier functions will be displayed, with a warning message when the result applies only to the first element in the structure. FALSE by default.
- **ignore.size.warn** Set to TRUE to force execution of the function for large (> 20 K-bytes) objects. Defaults to FALSE.

Details

An alternative to calling in turn class, typeof, dim, and so on. A call to this function will readily give all this information at once.

Value

A list with following elements:

- **properties** A data frame with the class(es), type, mode and storage mode of the object as well as the dim, length and object.size.
- **attributes.lengths** A named character vector giving all attributes (c.f. “names”, “row.names”, “class”, “dim”, and so forth) along with their length.
- **extensive.is** A character vector of all the identifier functions. (starting with “is.”) that yield TRUE when used with x as argument.
- **function.type** When x is a function, results of ftype are added.
Author(s)

Dominic Comtois, <dominic.comtois@gmail.com>

See Also

class, typeof, mode, storage.mode, dim, length, is.object, otype, object.size, ftype

Examples

what.is(1)
what.is(NaN)
what.is(iris3)
what.is(print)
what.is(what.is)
Index

*Topic **IO**
cleartmp, 3

*Topic **attribute**
dfSummary, 11
what.is, 35

*Topic **category**
ctable, 4
dfSummary, 11
freq, 16

*Topic **classes**
ctable, 4
dfSummary, 11
freq, 16
what.is, 35

*Topic **datasets**
examens, 15
exams, 16
tabagisme, 30
tobacco, 32

*Topic **methods**
print.summarytools, 22

*Topic **print**
print.summarytools, 22

*Topic **univar**
descr, 9
dfSummary, 11
freq, 16

*Topic **utilities**
define_keywords, 6
descr, 9, 20, 21, 23, 24, 28, 31, 34
dfSummary, 11, 20, 21, 23, 24, 28, 29, 34
dim, 35, 36

exams, 15
freq, 16, 20, 21, 23, 24, 28, 31, 34
ftype, 35, 36

is.object, 36

label, 19, 20, 33
label<-(label), 19
lapply, 28
length, 36

mode, 36

object.size, 36
otype, 36

pander, 4, 5, 10, 12, 13, 17, 18, 23, 25
print, 5, 10
print.list, 20
print.stby, 21
print.summarytools, 22, 35

st_css, 26
st_options, 4, 5, 10, 12, 13, 17, 18, 21–23, 27, 34
stby, 25, 31
storage.mode, 36
strptime, 9
summarytools (summarytools-package), 2
summarytools-package, 2

tabagisme, 30
table, 4, 5, 18
tapply, 25
tb, 31

tibble, 31

tobacco, 32

typeof, 35, 36

unlabel, 32

use_custom_lang, 6, 33

view, 33

what.is, 35

xtabs, 5