Package: summarySCI (via r-universe)

May 19, 2026

Title Produces Publication-Ready Summary Tables

Version 0.1.1

Description Produces tables with descriptive statistics for continuous, categorical and dichotomous variables. It is largely based on the package 'gtsummary'; Sjoberg DD et al. (2021) <doi:10.32614/RJ-2021-053>.

License LGPL-3

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.2

Imports Hmisc, cardx, dplyr, gtsummary (>= 2.3.0), forcats, labelled, purrr, rlang, tidyr, flextable, officer

Suggests knitr, tidyverse, survival, testthat (>= 3.0.0)

VignetteBuilder knitr

URL https://github.com/SAKK-Statistics/summarySCI/

BugReports https://github.com/SAKK-Statistics/summarySCI/issues

Config/testthat/edition 3

Config/pak/sysreqs libcairo2-dev cmake libfontconfig1-dev libfreetype6-dev libfribidi-dev make libharfbuzz-dev libicu-dev libjpeg-dev libpng-dev libtiff-dev libuv1-dev libwebp-dev libxml2-dev libssl-dev libnode-dev libx11-dev zlib1g-dev

Repository https://sakk-statistics.r-universe.dev

Date/Publication 2026-02-18 09:53:30 UTC

RemoteUrl https://github.com/sakk-statistics/summarysci

RemoteRef HEAD

RemoteSha 05591f02ba87b3e6f6e66db9725544add7a93306

Contents

summaryByVisit *Creates publication-ready summary tables for continuous data grouped, by visit*

Description

Creates publication-ready summary tables for continuous data grouped, by visit

Usage

```
summaryByVisit(
  data,
  vars = NULL,
  group = NULL,
  labels = NULL,
  stat_cont = "median_range",
  visit = "visit",
  order = NULL,
  visitgroup = NULL,
  digits_cont = 1,
  add_n = FALSE,
  overall = FALSE,
  as_flex_table = TRUE,
  border = TRUE,
  word_output = FALSE,
  file_name = paste0("SummaryByVisit_", format(Sys.Date(), "%Y%m%d"), ".docx")
)
```

Arguments

data	A data frame or tibble containing the data to be summarized.
vars	Continuous variables to include in the summary table. Need to be specified with quotes, e.g. "age" or c("age", "response"). Default to all variables present in the data except group.
group	A single column from data. Need to be specified with quotes, e.g. "treatment". Summary statistics will be stratified according to this variable. Default to NULL. A maximum of 3 groups are currently supported.
labels	A list containing the labels that should be used for the variables in the table. If NULL, labels are automatically taken from the dataset. If no label present, the variable name is taken.
stat_cont	Summary statistic to display for continuous variables. Options include "median_IQR", "median_range" (default), "mean_sd", "mean_se" and "geomMean_sd".
visit	Name of the stratum for which summary statistics are displayed by line. Typically, this would be "visit".

order	A numerical variable defining the visit order.
visitgroup	A grouping variable for the stratum for which summary statistics are displayed by line. Must be an ordered factor. Typically, this would be a visit group such as e.g., baseline, follow-up etc.
digits_cont	Digits for summary statistics and CI of continuous variables. Default to 1.
add_n	Logical. If TRUE, an additional column with the total number of non-missing observations for each variable is added.
overall	Logical. If TRUE, an additional column with the total is added to the table. Ignored, if no groups are defined. Default to FALSE.
as_flex_table	Logical. If TRUE (default) the gtsummary object is converted to a flextable object. Useful when rendering to Word.
border	Logical. If TRUE, a border will be drawn around the table. Only available if flex_table = TRUE. Default is TRUE.
word_output	Logical. If TRUE, the table is also saved in a word document.
file_name	Character string. Specify the name of the Word document containing the table. Only used when word_output is TRUE. Needs to end with ".docx".

Value

A table of class "flextable" or c("tbl_strata_nested_stack", "tbl_stack", "gtsummary"). Optionally returns a .docx file in the specified folder.

---

summaryLevels *summaryLevels*

Description

Collapses factor levels from multiple columns into one and creates summary table.

Usage

```
summaryLevels(
  data,
  vars = NULL,
  group = NULL,
  label = NULL,
  levels = NULL,
  stat_cat = "n_percent",
  test = FALSE,
  test_cat = "fisher.test",
  ci = FALSE,
  ci_cat = "wilson",
  conf_level = 0.95,
  digits_cat = 0,
  overall = FALSE,
```

```
  as_flex_table = TRUE,
  border = TRUE,
  word_output = FALSE,
  file_name = paste0("SummaryLevels_", format(Sys.Date(), "%Y%m%d"), ".docx")
)
```

Arguments

data	A data frame or tibble containing the data to be summarized.
vars	Variables to include in the summary table. Need to be specified with quotes, e.g. "score" or c("score", "age_cat"). Default to all variables present in the data except group.
group	A single column from data. Need to be specified with quotes, e.g. "treatment". Summary statistics will be stratified according to this variable. Default to NULL.
label	A label for the new variable to be created. If no label present, the variable name is taken.
levels	= A vector containing the values indicating presence of the factor level. Included by default are "1", "yes", "Yes".
stat_cat	Summary statistic to display for categorical variables. Options include "n_percent" (default) and "n", and "n_N".
test	Logical. Indicates whether p-values are displayed (TRUE) or not (FALSE). Default to FALSE
test_cat	Test type used to calculated the p-value for categorical variables. Only used if test = TRUE. Options include "fisher.test" (default), "chisq.test", "chisq.test.no.correct". If NULL, the function decides itself: "chisq.test.no.correct" for categorical variables with all expected cell counts >=5, and "fisher.test" for categorical variables with any expected cell count <5.
ci	Logical. Indicates whether CI are displayed (TRUE) or not (FALSE). Default to FALSE.
ci_cat	Confidence interval method for categorical variables. Options include "wilson" (default), "wilson.no.correct", "clopper.pearson", "wald", "wald.no.correct", "agresti.coull" and "jeffreys". If NULL, no CI will be displayed.
conf_level	Numeric. Confidence level. Default to 0.95.
digits_cat	Numeric. Digits for summary statistics and CI of categorical variables. Default to 0.
overall	Logical. If TRUE, an additional column with the total is added to the table. Default to FALSE.
as_flex_table	Logical. If TRUE (default) the gtsummary object is converted to a flextable object. Useful when rendering to Word.
border	Logical. If TRUE, a border will be drawn around the table. Only available if flex_table = TRUE. Default is TRUE.
word_output	Logical. If TRUE, the table is also saved in a word document.
file_name	Character string. Specify the name of the Word document containing the table. Only used when word_output is TRUE. Needs to end with ".docx".

Value

A table of class "flextable" or c("tbl_stack", "gtsummary"). Optionally returns a .docx file in the specified folder.

---

summaryTable *Creates publication-ready summary tables*

Description

Creates publication-ready summary tables based on the gtsummary package.

Usage

```
summaryTable(
  data,
  vars = NULL,
  group = NULL,
  labels = NULL,
  stat_cont = "median_range",
  stat_cat = "n_percent",
  continuous_as = "continuous",
  dichotomous_as = "dichotomous",
  value = NULL,
  test = FALSE,
  test_cont = "wilcox.test",
  test_cat = "fisher.test",
  ci = FALSE,
  ci_cont = "wilcox.test",
  ci_cat = "wilson",
  conf_level = 0.95,
  digits_cont = 1,
  digits_cat = 0,
  missing = TRUE,
  missing_percent = TRUE,
  missing_text = "Missing",
  overall = FALSE,
  add_n = TRUE,
  as_flex_table = TRUE,
  border = TRUE,
  word_output = FALSE,
  file_name = paste0("SummaryTable_", format(Sys.Date(), "%Y%m%d"), ".docx")
)
```

Arguments

data	A data frame or tibble containing the data to be summarized.

vars	Variables to include in the summary table. Need to be specified with quotes, e.g. "age" or c("age", "response"). Default to all variables present in the data except group.
group	A single column from data. Need to be specified with quotes, e.g. "treatment". Summary statistics will be stratified according to this variable. Default to NULL.
labels	A list containing the labels that should be used for the variables in the table. If NULL, labels are automatically taken from the dataset. If no label present, the variable name is taken.
stat_cont	Summary statistic to display for continuous variables. Options include "median_IQR", "median_range" (default), "mean_sd", "mean_se" and "geomMean_sd".
stat_cat	Summary statistic to display for categorical variables. Options include "n_percent" (default) and "n", and "n_N".
continuous_as	Type for the continuous variables. Can either be "continous" (default) or "categorical".
dichotomous_as	Type for the dichotomous variables. Can either be "categorical" (default, one row per level) or "dichotomous" (only one row with reference level (see argument value), only works if missing = "FALSE" or missing_percent = FALSE.
value	Specifies the reference level of a variable to display on a single row. Default is NULL. The syntax is as follows: value = list(varname ~ "level to show").
test	Logical. Indicates whether p-values are displayed (TRUE) or not (FALSE). Default to FALSE
test_cont	Test type used to calculate the p-value for continuous variables. Only used if test = TRUE. Options include "t.test", "oneway.test", "kruskal.test", "wilcox.test" (default), "paired.t.test", "paired.wilcox.test"
test_cat	Test type used to calculated the p-value for categorical variables. Only used if test = TRUE. Options include "fisher.test" (default), "chisq.test", "chisq.test.no.correct". If NULL, the function decides itself: "chisq.test.no.correct" for categorical variables with all expected cell counts >=5, and "fisher.test" for categorical variables with any expected cell count <5.
ci	Logical. Indicates whether CI are displayed (TRUE) or not (FALSE). Default to FALSE.
ci_cont	Confidence interval method for continuous variables. Only used if ci = TRUE. Options include "t.test" and "wilcox.test" (default).
ci_cat	Confidence interval method for categorical variables. Options include "wilson" (default), "wilson.no.correct", "clopper.pearson", "wald", "wald.no.correct", "agresti.coull" and "jeffreys". If NULL, no CI will be displayed.
conf_level	Numeric. Confidence level. Default to 0.95.
digits_cont	Numeric. Digits for summary statistics and CI of continuous variables. Default to 1.
digits_cat	Numeric. Digits for summary statistics and CI of categorical variables. Default to 0.
missing	Logical. If TRUE (default), the missing values are shown.

`missing_percent`
: Indicates whether percentages for missings are shown (TRUE, default) or not (FALSE) for categorical variables. If "both", then both options are displayed next to each other.

`missing_text`
: String indicating text shown on missing row. Default to "Missing".

`overall`
: Logical. If TRUE, an additional column with the total is added to the table. Default to FALSE.

`add_n`
: Logical. If TRUE (default), an additional column with the total number of non-missing observations for each variable is added.

`as_flex_table`
: Logical. If TRUE (default) the gtsummary object is converted to a flextable object. Useful when rendering to Word.

`border`
: Logical. If TRUE, a border will be drawn around the table. Only available if flex_table = TRUE. Default is TRUE.

`word_output`
: Logical. If TRUE, the table is also saved in a word document.

`file_name`
: Character string. Specify the name of the Word document containing the table. Only used when `word_output` is TRUE. Needs to end with ".docx".

Value

A table of class "`flextable`" or `c("tbl_summary", "gtsummary")`. Optionally returns a .docx file in the specified folder.

Examples

```
library(survival)
data("cancer")
summaryTable(data = cancer,vars = c("inst", "time","age", "ph.ecog"),
             labels = list(inst = "Institution code",
                           time = "Time",
                           age = "Age",
                           ph.ecog = "ECOG score"))
```

Index