

Package ‘colleyRstats’

January 9, 2026

Type Package

Title Functions to Streamline Statistical Analysis and Reporting

Version 0.0.2

Description Built upon popular R packages such as 'ggstatsplot' and 'ARTool', this collection of-
fers a wide array of tools for simplifying reproducible analyses, generating high-quality visual-
izations, and producing 'APA'-compliant outputs. The primary goal of this package is to signifi-
cantly reduce repetitive coding efforts, allowing you to focus on interpreting re-
sults. Whether you're dealing with ANOVA assumptions, reporting effect sizes, or creating publi-
cation-ready visualizations, this package makes these tasks easier.

URL <https://github.com/M-Colley/colleyRstats>

BugReports <https://github.com/M-Colley/colleyRstats/issues>

License MIT + file LICENSE

Encoding UTF-8

RoxygenNote 7.3.3

Language en-US

Depends R (>= 4.5.0)

Imports ARTool (>= 0.11.2), car, clipr, conflicted, dplyr (>= 1.1.4),
effectsize (>= 1.0.1), FSA, ggplot2 (>= 4.0.1), ggpmisc (>=
0.6.3), ggsignif, ggstatsplot (>= 0.13.4), ggtext, purrr,
readxl, report (>= 0.6.3), rlang, rstatix, see, stats, stringr,
tidyr, utils, writexl, xtable

Suggests afex, BayesFactor, bayestestR, Cairo, dunn.test, DT, emmeans,
emoa, flexdashboard, Hmisc, moocore (>= 0.1.10), nparLD,
pkgload, psych, reporttools, rstantools, scales, sjPlot,
stargazer, styler, tibble, testthat (>= 3.3.0)

Config/testthat/edition 3

Config/Needs/check anthonymnorth/roxyglobals

Config/roxyglobals/unique TRUE

Config/testthat/parallel false

NeedsCompilation no

Author Mark Colley [aut, cre, cph] (ORCID:
<https://orcid.org/0000-0001-5207-5029>)

Maintainer Mark Colley <mark.colley@yahoo.de>

Repository CRAN

Date/Publication 2026-01-08 23:40:02 UTC

Contents

add_pareto_emoa_column	3
checkAssumptionsForAnova	4
check_homogeneity_by_group	5
check_normality_by_group	5
colleyRstats_setup	6
data the data frame	7
debug_contr_error	8
generateEffectPlot	9
generateMoboPlot	10
generateMoboPlot2	12
ggbetweenstatsWithPriorNormalityCheck	13
ggbetweenstatsWithPriorNormalityCheckAsterisk	14
ggwithinstatsWithPriorNormalityCheck	16
ggwithinstatsWithPriorNormalityCheckAsterisk	17
latexify_report	18
na.zero	19
normalize	20
not_empty	21
not_in	21
n_fun	22
pathPrep	22
remove_outliers_REI	23
reportART	23
reportDunnTest	24
reportDunnTestTable	25
reportggstatsplot	27
reportggstatsplotPostHoc	28
reportMeanAndSD	29
reportNparLD	30
reportNPAV	31
reshape_data	32
rFromNPAV	33
rFromWilcox	34
rFromWilcoxAdjusted	35
stat_sum_df	35

Index	37
--------------	-----------

`add_pareto_emoa_column`*Add Pareto EMOA Column to a Data Frame*

Description

This function calculates the Pareto front for a given set of objectives in a data frame and adds a new column, PARETO_EMOA, which indicates whether each row in the data frame belongs to the Pareto front.

Usage

```
add_pareto_emoa_column(data, objectives)
```

Arguments

<code>data</code>	A data frame containing the data, including the objective columns.
<code>objectives</code>	A character vector specifying the names of the objective columns in data. These columns should be numeric and will be used to calculate the Pareto front.

Value

A data frame with the same columns as `data`, along with an additional column, PARETO_EMOA, which is TRUE for rows that are on the Pareto front and FALSE otherwise.

Examples

```
# Define objective columns
objectives <- c("trust", "predictability", "perceivedSafety", "Comfort")

# Example data frame
main_df <- data.frame(
  trust = runif(10),
  predictability = runif(10),
  perceivedSafety = runif(10),
  Comfort = runif(10)
)

# Add the Pareto front column
main_df <- add_pareto_emoa_column(data = main_df, objectives)
head(main_df)
```

`checkAssumptionsForAnova`

Check the assumptions for an ANOVA with a variable number of factors: Normality and Homogeneity of variance assumption.

Description

Check the assumptions for an ANOVA with a variable number of factors: Normality and Homogeneity of variance assumption.

Usage

```
checkAssumptionsForAnova(data, y, factors)
```

Arguments

<code>data</code>	the data frame
<code>y</code>	The dependent variable for which assumptions should be checked
<code>factors</code>	A character vector of factor names

Value

A message indicating whether to use parametric or non-parametric ANOVA

Examples

```
set.seed(123)

main_df <- data.frame(
  tlx_mental = rnorm(40),
  Video      = factor(rep(c("A", "B"), each = 20)),
  DriverPosition = factor(rep(c("Left", "Right"), times = 20))
)

checkAssumptionsForAnova(
  data = main_df,
  y     = "tlx_mental",
  factors = c("Video", "DriverPosition")
)
```

`check_homogeneity_by_group`*Check homogeneity of variances across groups*

Description

Check homogeneity of variances across groups

Usage

```
check_homogeneity_by_group(data, x, y)
```

Arguments

<code>data</code>	the data frame
<code>x</code>	the grouping variable (column name as string)
<code>y</code>	the dependent variable (column name as string)

Value

TRUE if Levene's test is non-significant ($p \geq .05$), FALSE otherwise

`check_normality_by_group`*Check normality for groups*

Description

Check normality for groups

Usage

```
check_normality_by_group(data, x, y)
```

Arguments

<code>data</code>	the data frame
<code>x</code>	the x column
<code>y</code>	the y column

Value

TRUE if all groups are normal, FALSE otherwise. For groups with $n > 5000$, Shapiro-Wilk is skipped and the function returns FALSE with a warning.

colleyRstats_setup	<i>Configure Global R Environment for colleyRstats</i>
--------------------	--

Description

Sets ggplot2 themes and conflict preferences to match the standards used in the colleyRstats workflow.

Usage

```
colleyRstats_setup(  
  set_options = TRUE,  
  set_theme = TRUE,  
  set_conflicts = TRUE,  
  print_citation = TRUE,  
  verbose = TRUE  
)
```

Arguments

set_options	Logical. If TRUE, prints a notice that global options are no longer changed automatically. Default is TRUE.
set_theme	Logical. If TRUE, sets the default ggplot2 theme to <code>see::theme_lucid</code> with custom modifications. Default is TRUE.
set_conflicts	Logical. If TRUE, sets conflicted preferences to favor dplyr and other tidyverse packages. Default is TRUE.
print_citation	Logical. If TRUE, prints the citation information for this package. Default is TRUE.
verbose	Logical. If TRUE, emit informational messages. Default is TRUE.

Value

Invisibly returns NULL.

Examples

```
# Runs everywhere, no extra packages, no session side effects  
colleyRstats::colleyRstats_setup(  
  set_options = FALSE,  
  set_theme = FALSE,  
  set_conflicts = FALSE,  
  print_citation = FALSE,  
  verbose = FALSE  
)  
  
# Full setup (requires suggested packages; changes session defaults)
```

```

if (requireNamespace("ggplot2", quietly = TRUE) &&
    requireNamespace("see", quietly = TRUE)) {
  local({
    old_theme <- ggplot2::theme_get()
    on.exit(ggplot2::theme_set(old_theme), add = TRUE)

    colleyRstats::colleyRstats_setup(
      set_options = FALSE,
      set_conflicts = FALSE, # avoid persisting conflict prefs in checks
      print_citation = FALSE,
      verbose = TRUE
    )

    ggplot2::ggplot(mtcars, ggplot2::aes(mpg, wt)) +
      ggplot2::geom_point()
  })
}

```

data the data frame *Replace values across a data frame*

Description

Replace all occurrences of given values in all columns of a data frame.

The data data frame contains a collection of records, with attributes organized in columns. It may include various types of values, such as numerical, categorical, or textual data.

Usage

```
replace_values(data, to_replace, replace_with)
```

Arguments

<code>data</code>	The input data frame to be modified.
<code>to_replace</code>	A vector of values to be replaced within the data frame. This must be the same length as <code>replace_with</code> .
<code>replace_with</code>	A vector of corresponding replacement values. This must be the same length as <code>to_replace</code> .

Value

Modified data frame with specified values replaced.

Examples

```
data <- data.frame(
  q1 = c("neg2", "neg1", "0"),
  q2 = c("1", "neg2", "neg1")
)

replace_values(
  data,
  to_replace = c("neg2", "neg1"),
  replace_with = c("-2", "-1")
)
```

debug_contr_error	<i>Debug contrast errors in ANOVA-like models</i>
-------------------	---

Description

Debug contrast errors in ANOVA-like models

Usage

```
debug_contr_error(dat, subset_vec = NULL)
```

Arguments

dat	A data frame of predictors.
subset_vec	Optional logical or numeric index vector used to subset rows before checks.

Value

A list with two elements:

nlevels Integer vector giving the number of levels for each factor variable in dat.

levels List of factor level labels for each factor variable in dat.

Examples

```
dat <- data.frame(
  group = factor(rep(letters[1:3], each = 3)),
  score = rnorm(9)
)

debug_contr_error(dat = dat)
```

generateEffectPlot	<i>Function to define a plot, either showing the main or interaction effect in bold.</i>
--------------------	--

Description

Function to define a plot, either showing the main or interaction effect in bold.

Usage

```
generateEffectPlot(
  data,
  x,
  y,
  fillColourGroup,
  ytext = "testylab",
  xtext = "testxlab",
  legendPos = c(0.1, 0.23),
  legendHeading = NULL,
  shownEffect = "main",
  effectLegend = FALSE,
  effectDescription = NULL,
  xLabelsOverwrite = NULL,
  useLatexMarkup = FALSE,
  numberColors = 6
)
```

Arguments

data	the data frame
x	factor shown on the x-axis
y	dependent variable
fillColourGroup	group to color
ytext	label for y-axis
xtext	label for x-axis
legendPos	position for legend
legendHeading	custom heading for legend
shownEffect	either "main" or "interaction"
effectLegend	TRUE: show legend for effect (Default: FALSE)
effectDescription	custom label for effect
xLabelsOverwrite	custom labels for x-axis
useLatexMarkup	use latex font and markup
numberColors	number of colors

Value

a plot

Examples

```
set.seed(123)
main_df <- data.frame(
  strategy = factor(rep(c("A", "B"), each = 20)),
  Emotion   = factor(rep(c("Happy", "Sad"), times = 20)),
  trust_mean = rnorm(40, mean = 5, sd = 1)
)

generateEffectPlot(
  data = main_df,
  x = "strategy",
  y = "trust_mean",
  fillColourGroup = "Emotion",
  ytext = "Trust",
  xtext = "Strategy",
  legendPos = c(0.1, 0.23)
)
```

generateMoboPlot

Generate a Multi-objective Optimization Plot

Description

This function generates a multi-objective optimization plot using ggplot2. The plot visualizes the relationship between the x and y variables, grouping and coloring by a fill variable, with the option to customize legend position, labels, and annotation of sampling and optimization phases.

Usage

```
generateMoboPlot(
  data,
  x,
  y,
  fillColourGroup = "ConditionID",
  ytext,
  legendPos = c(0.65, 0.85),
  numberSamplingSteps = 5,
  labelPosFormulaY = "top",
  verticalLinePosY = 0.75
)
```

Arguments

<code>data</code>	A data frame containing the data to be plotted.
<code>x</code>	A string representing the column name in data to be used for the x-axis. Can be either numeric or factor.
<code>y</code>	A string representing the column name in data to be used for the y-axis. This should be a numeric variable.
<code>fillColourGroup</code>	A string representing the column name in data that defines the fill color grouping for the plot. Default is "ConditionID".
<code>ytext</code>	A custom label for the y-axis. If not provided, the y-axis label will be the title-cased version of y.
<code>legendPos</code>	A numeric vector of length 2 specifying the position of the legend inside the plot. Default is <code>c(0.65, 0.85)</code> .
<code>numberSamplingSteps</code>	An integer specifying the number of initial sampling steps before the optimization phase begins. Default is 5.
<code>labelPosFormulaY</code>	A string specifying the vertical position of the polynomial equation label in the plot. Acceptable values are "top", "center", or "bottom". Default is "top".
<code>verticalLinePosY</code>	A numeric value of the y-coordinate where the "sampling" and "optimization" line should be drawn.

Value

A ggplot object representing the multi-objective optimization plot, ready to be rendered.

Examples

```
library(ggplot2)
library(ggpmisc)

# Example with numeric x-axis
df <- data.frame(
  x = 1:20,
  y = rnorm(20),
  ConditionID = rep(c("A", "B"), 10)
)
generateMoboPlot(df, x = "x", y = "y")

# Example with factor x-axis
df <- data.frame(
  x = factor(rep(1:5, each = 4)),
  y = rnorm(20),
  ConditionID = rep(c("A", "B"), 10)
)
generateMoboPlot(df, x = "x", y = "y", numberSamplingSteps = 3)
```

Description

This function generates a multi-objective optimization plot using ggplot2. The plot visualizes the relationship between the x and y variables, grouping and coloring by a fill variable, with the option to customize legend position, labels, and annotation of sampling and optimization phases. Appropriate if you use <https://github.com/Pascal-Jansen/Bayesian-Optimization-for-Unity> in version 1.1.0 or higher.

Usage

```
generateMoboPlot2(
  data,
  x = "Iteration",
  y,
  phaseCol = "Phase",
  fillColourGroup = "",
  ytext,
  legendPos = c(0.65, 0.85),
  labelPosFormulaY = "top",
  verticalLinePosY = 0.75
)
```

Arguments

data	A data frame containing the data to be plotted.
x	A string representing the column name in data to be used for the x-axis. Can be either numeric or factor. Default is "Iteration".
y	A string representing the column name in data to be used for the y-axis. This should be a numeric variable.
phaseCol	the name of the column for the color of the phase (sampling or optimization)
fillColourGroup	A string representing the column name in data that defines the fill color grouping for the plot. Default is "ConditionID".
ytext	A custom label for the y-axis. If not provided, the y-axis label will be the title-cased version of y.
legendPos	A numeric vector of length 2 specifying the position of the legend inside the plot. Default is c(0.65, 0.85).
labelPosFormulaY	A string specifying the vertical position of the polynomial equation label in the plot. Acceptable values are "top", "center", or "bottom". Default is "top".
verticalLinePosY	A numeric value of the y-coordinate where the "sampling" and "optimization" line should be drawn.

Value

A ggplot object representing the multi-objective optimization plot, ready to be rendered.

Examples

```
library(ggplot2)
library(ggpmisc)

# Example with numeric x-axis
df <- data.frame(
  x = 1:20,
  y = rnorm(20),
  ConditionID = rep(c("A", "B"), 10),
  Phase = rep(c("Sampling", "Optimization"), 10)
)
generateMoboPlot2(data = df, x = "x", y = "y")
```

ggbetweenstatsWithPriorNormalityCheck

Check the data's distribution. If non-normal, take the non-parametric variant of ggbetweenstats. x and y have to be in parentheses, e.g., "ConditionID".

Description

Check the data's distribution. If non-normal, take the non-parametric variant of *ggbetweenstats*. x and y have to be in parentheses, e.g., "ConditionID".

Usage

```
ggbetweenstatsWithPriorNormalityCheck(
  data,
  x,
  y,
  ylab,
  xlabels = NULL,
  showPairwiseComp = TRUE,
  plotType = "boxviolin"
)
```

Arguments

data	the data frame
x	the independent variable, most likely "ConditionID"
y	the dependent variable under investigation
ylab	label to be shown for the dependent variable
xlabels	labels to be used for the x-axis

showPairwiseComp whether to show pairwise comparisons, TRUE as default

plotType either "box", "violin", or "boxviolin" (default)

Value

A ggplot object produced by ggstatsplot::ggbetweenstats, which can be printed or further modified with +.

Examples

```
set.seed(123)

# Toy within-subject style data
main_df <- data.frame(
  Participant = factor(rep(1:20, each = 3)),
  CondID      = factor(rep(c("A", "B", "C"), times = 20)),
  tlx_mental  = rnorm(60, mean = 50, sd = 10)
)

# Custom x-axis labels
labels_xlab <- c("Condition A", "Condition B", "Condition C")

ggbetweenstatsWithPriorNormalityCheck(
  data = main_df,
  x = "CondID",
  y = "tlx_mental", ylab = "Mental Demand",
  xlabels = labels_xlab,
  showPairwiseComp = TRUE
)
```

ggbetweenstatsWithPriorNormalityCheckAsterisk

Check the data's distribution. If non-normal, take the non-parametric variant of ggbetweenstats. x and y have to be in parentheses, e.g., "ConditionID".

Description

Check the data's distribution. If non-normal, take the non-parametric variant of ggbetweenstats. x and y have to be in parentheses, e.g., "ConditionID".

Usage

```
ggbetweenstatsWithPriorNormalityCheckAsterisk(  
  data,  
  x,  
  y,  
  ylab,  
  xlabels,  
  plotType = "boxviolin"  
)
```

Arguments

data	the data frame
x	the independent variable, most likely "ConditionID"
y	the dependent variable under investigation
ylab	label to be shown for the dependent variable
xlabels	labels to be used for the x-axis
plotType	either "box", "violin", or "boxviolin" (default)

Value

A ggplot object produced by `ggstatsplot::ggbetweenstats` with additional significance annotations, which can be printed or modified.

Examples

```
set.seed(123)  
  
# Toy within-subject style data  
main_df <- data.frame(  
  Participant = factor(rep(1:20, each = 3)),  
  CondID      = factor(rep(c("A", "B", "C"), times = 20)),  
  tlx_mental  = rnorm(60, mean = 50, sd = 10)  
)  
  
# Custom x-axis labels  
labels_xlab <- c("Condition A", "Condition B", "Condition C")  
  
ggbetweenstatsWithPriorNormalityCheckAsterisk(  
  data = main_df,  
  x = "CondID", y = "tlx_mental", ylab = "Mental Demand", xlabels = labels_xlab  
)
```

```
ggwithinstatsWithPriorNormalityCheck
```

Check the data's distribution. If non-normal, take the non-parametric variant of ggwithinstats. x and y have to be in parentheses, e.g., "ConditionID".

Description

Check the data's distribution. If non-normal, take the non-parametric variant of *ggwithinstats*. *x* and *y* have to be in parentheses, e.g., "ConditionID".

Usage

```
ggwithinstatsWithPriorNormalityCheck(
  data,
  x,
  y,
  ylab,
  xlabels = NULL,
  showPairwiseComp = TRUE,
  plotType = "boxviolin"
)
```

Arguments

<code>data</code>	the data frame
<code>x</code>	the independent variable, most likely "ConditionID"
<code>y</code>	the dependent variable under investigation
<code>ylab</code>	label to be shown for the dependent variable
<code>xlabels</code>	labels to be used for the x-axis
<code>showPairwiseComp</code>	whether to show pairwise comparisons, TRUE as default
<code>plotType</code>	either "box", "violin", or "boxviolin" (default)

Value

A ggplot object produced by `ggstatsplot::ggwithinstats` with additional significance annotations, which can be printed or modified.

Examples

```
#' set.seed(123)

# Toy within-subject style data
main_df <- data.frame(
```

```

Participant = factor(rep(1:20, each = 3)),
CondID      = factor(rep(c("A", "B", "C"), times = 20)),
tlx_mental  = rnorm(60, mean = 50, sd = 10)
)

# Custom x-axis labels
labels_xlab <- c("Condition A", "Condition B", "Condition C")

ggwithinstatsWithPriorNormalityCheck(
  data = main_df,
  x = "CondID", y = "tlx_mental",
  ylab = "Mental Demand",
  xlabels = labels_xlab,
  showPairwiseComp = TRUE
)

```

```
ggwithinstatsWithPriorNormalityCheckAsterisk
```

Check the data's distribution. If non-normal, take the non-parametric variant of ggwithinstats. x and y have to be in parentheses, e.g., "ConditionID". Add Asterisks instead of p-values.

Description

Check the data's distribution. If non-normal, take the non-parametric variant of *ggwithinstats*. x and y have to be in parentheses, e.g., "ConditionID". Add Asterisks instead of p-values.

Usage

```

ggwithinstatsWithPriorNormalityCheckAsterisk(
  data,
  x,
  y,
  ylab,
  xlabels,
  plotType = "boxviolin"
)

```

Arguments

data	the data frame
x	the independent variable, most likely "ConditionID"
y	the dependent variable under investigation
ylab	label to be shown for the dependent variable
xlabels	labels to be used for the x-axis
plotType	either "box", "violin", or "boxviolin" (default)

Value

A ggplot object produced by `ggstatsplot::ggwithinstats` with additional significance annotations, which can be printed or modified.

Examples

```
set.seed(123)

# Toy within-subject style data
main_df <- data.frame(
  Participant = factor(rep(1:20, each = 3)),
  CondID      = factor(rep(c("A", "B", "C"), times = 20)),
  tlx_mental  = rnorm(60, mean = 50, sd = 10)
)

# Custom x-axis labels
labels_xlab <- c("Condition A", "Condition B", "Condition C")

ggwithinstatsWithPriorNormalityCheckAsterisk(
  data = main_df,
  x = "CondID", y = "tlx_mental",
  ylab = "Mental Demand", xlabels = labels_xlab
)
```

code latexify_report

Transform text from report::report() into LaTeX-friendly output.

Description

This function transforms the text output from `report::report()` by performing several substitutions to prepare the text for LaTeX typesetting. In particular, it replaces instances of R2, %, and ~ with the corresponding LaTeX code. Additionally, it provides options to:

- Omit bullet items marked as "non-significant" (when `only_sig = TRUE`).
- Remove a concluding note about standardized parameters (when `remove_std = TRUE`).
- Wrap bullet items in a LaTeX `itemize` environment or leave them as plain text (controlled by `itemize`).

Usage

```
latexify_report(
  x,
  print_result = TRUE,
  only_sig = FALSE,
  remove_std = FALSE,
  itemize = TRUE
)
```

Arguments

<code>x</code>	Character vector or a single string containing the report text.
<code>print_result</code>	Logical. If TRUE (default), the formatted text is printed to the console.
<code>only_sig</code>	Logical. If TRUE, bullet items containing "non-significant" are omitted. Default is FALSE.
<code>remove_std</code>	Logical. If TRUE, the final standardized parameters note is removed. Default is FALSE.
<code>itemize</code>	Logical. If TRUE (default), bullet items are wrapped in a LaTeX <code>itemize</code> environment; otherwise the bullet markers are simply removed.

Value

A single string with the LaTeX-friendly formatted report text.

Examples

```
if (requireNamespace("report", quietly = TRUE)) {
  # Simple linear model on the iris dataset
  model <- stats::lm(
    Sepal.Length ~ Sepal.Width + Petal.Length,
    data = datasets::iris
  )

  # Format the report output, showing only significant items, removing the
  # standard note, and wrapping bullet items in an itemize environment.
  report_text <- try(report::report(model), silent = TRUE)
  if (!inherits(report_text, "try-error")) {
    latexify_report(
      report_text,
      only_sig = TRUE,
      remove_std = TRUE,
      itemize = TRUE
    )
  }
}
```

na.zero

Replace NA values with zero

Description

Replace NA values with zero

Usage

```
na.zero(x)
```

Arguments

x A vector.

Value

A vector with NAs replaced by zeros.

Examples

```
na.zero(c(NA, 1, NA, 2))
```

normalize	<i>This function normalizes the values in a vector to the range [new_min, new_max] based on their original range [old_min, old_max].</i>
-----------	--

Description

This function normalizes the values in a vector to the range [new_min, new_max] based on their original range [old_min, old_max].

Usage

```
normalize(x_vector, old_min, old_max, new_min, new_max)
```

Arguments

x_vector	A numeric vector that you want to normalize.
old_min	The minimum value in the original scale of the data.
old_max	The maximum value in the original scale of the data.
new_min	The minimum value in the new scale to which you want to normalize the data.
new_max	The maximum value in the new scale to which you want to normalize the data.

Value

A numeric vector with the normalized values.

Examples

```
normalize(c(1, 2, 3, 4, 5), 1, 5, 0, 1)
```

not_empty	<i>Ensure input is not empty</i>
-----------	----------------------------------

Description

Stops execution if x is NULL, empty, or contains only NAs.

Usage

```
not_empty(x, msg = "Input must not be empty.")
```

Arguments

x	The object to check
msg	The error message to display

Value

Invisible TRUE if valid.

not_in	<i>Negate %in% membership</i>
--------	-------------------------------

Description

Negate %in% membership

Usage

```
not_in(x, y)
```

```
x %!in% y
```

Arguments

x	Vector of values to test.
y	Vector of values to match against.

Value

Logical vector indicating non-membership.

n_fun	<i>Build a median/size label for plot annotations</i>
-------	---

Description

Build a median/size label for plot annotations

Usage

```
n_fun(x)
```

Arguments

x	A numeric vector.
---	-------------------

Value

A data frame with the median and label.

pathPrep	<i>Convert Windows paths to R-friendly format</i>
----------	---

Description

Convert Windows paths to R-friendly format

Usage

```
pathPrep(path = "clipboard", read_fn = NULL, write_fn = NULL)
```

Arguments

path	Path to convert or the string "clipboard" to read from the clipboard.
read_fn	Optional custom function to read from the clipboard.
write_fn	Optional custom function to write to the clipboard.

Value

A normalized path string.

remove_outliers_REI	<i>Remove outliers and calculate REI</i>
---------------------	--

Description

This function takes a data frame, optional header information, variables to consider, and a range for a Likert scale. It then calculates the Response Entropy Index (REI) and flags suspicious entries based on percentiles.

Usage

```
remove_outliers_REI(df, header = FALSE, variables = "", range = c(1, 5))
```

Arguments

df	Data frame containing the data.
header	Logical indicating if the data frame has a header. Defaults to FALSE.
variables	Character string specifying which variables to consider, separated by commas.
range	Numeric vector specifying the range of the Likert scale. Defaults to c(1, 5).

Details

For more information on the REI method, refer to: [Response Entropy Index Method](#)

Value

A data frame with calculated REI, percentile, and a 'Suspicious' flag.

Examples

```
df <- data.frame(var1 = c(1, 2, 3), var2 = c(2, 3, 4))
result <- remove_outliers_REI(df, TRUE, "var1,var2", c(1, 5))
```

reportART	<i>Generate the Latex-text based on the ARTool (see https://github.com/mjskay/ARTool). The ART result must be piped into an anova(). Only significant main and interaction effects are reported. P-values are rounded for the third digit. Attention: Effect sizes are not calculated! Attention: the independent variables of the formula and the term specifying the participant must be factors (i.e., use as.factor()).</i>
-----------	--

Description

To easily copy and paste the results to your manuscript, the following commands must be defined in

Latex: `\newcommand{\F}[3]{\mathcal{F}(\{ \#1 \}, \{ \#2 \}) = \{ \#3 \}}` `\newcommand{\p}{\textit{p=}}` `\newcommand{\pminor}{\textit{p=}}`

Usage

```
reportART(model, dv = "Testdependentvariable", write_to_clipboard = FALSE)
```

Arguments

`model` the model of the art

`dv` the name of the dependent variable that should be reported

`write_to_clipboard` whether to write to the clipboard

Value

A message describing the statistical results.

Examples

```
if (requireNamespace("ARTool", quietly = TRUE)) {
  set.seed(123)

  main_df <- data.frame(
    tlx_mental = stats::rnorm(80),
    Video      = factor(rep(c("A", "B"), each = 40)),
    gesture    = factor(rep(c("G1", "G2"), times = 40)),
    eHMI       = factor(rep(c("On", "Off"), times = 40)),
    UserID     = factor(rep(1:20, each = 4))
  )

  art_model <- ARTool::art(
    tlx_mental ~ Video * gesture * eHMI +
      Error(UserID / (gesture * eHMI)),
    data = main_df
  )

  model_anova <- stats::anova(art_model)
  reportART(model_anova, dv = "mental demand")
}
```

reportDunnTest	<i>Report dunnTest as text. Required commands in LaTeX:</i>
	<code>\newcommand{\padjminor}{\textit{p\$_{adj}\$<}}</code>
	<code>\newcommand{\padj}{\textit{p\$_{adj}\$=}}</code>
	<code>\newcommand{\rankbiserial}[1]{\$_{rb}\$ = #1\$}</code>

Description

Report dunnTest as text. Required commands in LaTeX: `\newcommand{\padjminor}{\textit{p$_{adj}$<}}`
`\newcommand{\padj}{\textit{p$_{adj}$=}}` `\newcommand{\rankbiserial}[1]{$_{rb}$ = #1$}`

Usage

```
reportDunnTest(d, data, iv = "testiv", dv = "testdv")
```

Arguments

d	the dunn test object
data	the data frame
iv	independent variable
dv	dependent variable

Value

A message describing the statistical results.

Examples

```
if (requireNamespace("FSA", quietly = TRUE)) {
  # Use built-in iris data
  data(iris)

  # Dunn test on Sepal.Length by Species
  d <- FSA::dunnTest(Sepal.Length ~ Species,
    data = iris,
    method = "holm"
  )

  # Report the Dunn test
  reportDunnTest(d,
    data = iris,
    iv = "Species",
    dv = "Sepal.Length"
  )
}
```

reportDunnTestTable	<i>report Dunn test as a table. Customizable with sensible defaults. Required commands in LaTeX:</i> $\backslash\text{newcommand}\{\backslash\text{p}_{\text{adj}}\text{<}\}\{\backslash\text{textit}\{\text{p}_{\text{adj}}\text{<}\}\}$ $\backslash\text{newcommand}\{\backslash\text{p}_{\text{adj}}\}\{\backslash\text{textit}\{\text{p}_{\text{adj}}\text{=}\}\}$ $\backslash\text{newcommand}\{\backslash\text{rankbiserial}[1]\{\text{\$r}_{\text{rb}}\} = \#1\}$
---------------------	---

Description

report Dunn test as a table. Customizable with sensible defaults. Required commands in LaTeX:
 $\backslash\text{newcommand}\{\backslash\text{p}_{\text{adj}}\text{<}\}\{\backslash\text{textit}\{\text{p}_{\text{adj}}\text{<}\}\}$
 $\backslash\text{newcommand}\{\backslash\text{p}_{\text{adj}}\}\{\backslash\text{textit}\{\text{p}_{\text{adj}}\text{=}\}\}$
 $\backslash\text{newcommand}\{\backslash\text{rankbiserial}[1]\{\text{\$r}_{\text{rb}}\} = \#1\}$

Usage

```
reportDunnTestTable(  
  d = NULL,  
  data,  
  iv = "testiv",  
  dv = "testdv",  
  orderByP = FALSE,  
  numberDigitsForPValue = 4,  
  latexSize = "small",  
  orderText = TRUE  
)
```

Arguments

d	the dunn test object
data	the data frame
iv	independent variable
dv	dependent variable
orderByP	whether to order by the p value
numberDigitsForPValue	the number of digits to show
latexSize	which size for the text
orderText	whether to order the text

Value

A message describing the statistical results in a table.

Examples

```
if (requireNamespace("FSA", quietly = TRUE)) {  
  # Use built-in iris data  
  data(iris)  
  
  # Dunn test on Sepal.Length by Species  
  d <- FSA::dunnTest(Sepal.Length ~ Species,  
    data = iris,  
    method = "holm"  
  )  
  
  # Report the Dunn test  
  reportDunnTestTable(d,  
    data = iris,  
    iv = "Species",  
    dv = "Sepal.Length"  
  )  
}
```

reportggstatsplot	<i>Report statistical details for ggstatsplot.</i>
-------------------	--

Description

Report statistical details for ggstatsplot.

Usage

```
reportggstatsplot(  
  p,  
  iv = "independent",  
  dv = "Testdependentvariable",  
  write_to_clipboard = FALSE  
)
```

Arguments

p	the object returned by ggwithinstats or ggbetweenstats
iv	the independent variable
dv	the dependent variable
write_to_clipboard	whether to write to the clipboard

Value

A message describing the statistical results.

Examples

```
library(ggstatsplot)  
library(dplyr)  
  
# Generate a plot  
plt <- ggbetweenstats(mtcars, am, mpg)  
  
reportggstatsplot(plt, iv = "am", dv = "mpg")
```

reportggstatsplotPostHoc

Report significant post-hoc pairwise comparisons

Description

This function extracts significant pairwise comparisons from a ggstatsplot object, calculates the mean and standard deviation for the groups involved using the raw data, and prints LaTeX-formatted sentences reporting the results.

Usage

```
reportggstatsplotPostHoc(
  data,
  p,
  iv = "testiv",
  dv = "testdv",
  label_mappings = NULL
)
```

Arguments

data	A data frame containing the raw data used to generate the plot.
p	A ggstatsplot object (e.g., returned by ggbetweenstats) containing the pairwise comparison statistics.
iv	Character string. The column name of the independent variable (grouping variable).
dv	Character string. The column name of the dependent variable.
label_mappings	Optional named list or vector. Used to rename factor levels in the output text (e.g., list("old_name" = "New Label")).

Value

No return value. The function prints LaTeX-formatted text to the console.

LaTeX Requirements

To easily copy and paste the results to your manuscript, the following commands (or similar) must be defined in your LaTeX preamble, as the function outputs commands taking arguments (e.g., `\m{value}`):

```
\newcommand{\m}[1]{\textit{M}=#1}
\newcommand{\sd}[1]{\textit{SD}=#1}
\newcommand{\padj}[1]{\textit{p}_{adj}=#1}
\newcommand{\padjminor}[1]{\textit{p}_{adj}<#1}
```

Examples

```
library(ggstatsplot)
library(dplyr)

# Generate a plot
plt <- ggbetweenstats(mtcars, am, mpg)

# Report stats
reportggstatsplotPostHoc(
  data = mtcars,
  p = plt,
  iv = "am",
  dv = "mpg",
  label_mappings = list("0" = "Automatic", "1" = "Manual")
)
```

reportMeanAndSD	<i>Report the mean and standard deviation of a dependent variable for all levels of an independent variable rounded to the 2nd digit.</i>
-----------------	---

Description

#' To easily copy and paste the results to your manuscript, the following commands must be defined in Latex: `\newcommand{\m}{\textit{M=}}` `\newcommand{\sd}{\textit{SD=}}`

Usage

```
reportMeanAndSD(data, iv = "testiv", dv = "testdv")
```

Arguments

data	the data frame
iv	the independent variable
dv	the dependent variable

Value

Mean and SD values

Examples

```
example_data <- data.frame(Condition = rep(c("A", "B", "C"),
each = 10), TLX1 = stats::rnorm(30))

reportMeanAndSD(example_data, iv = "Condition", dv = "TLX1")
```

reportNparLD	<i>Report the model produced by nparLD. The model provided must be the model generated by the command 'nparLD' nparLD (see https://CRAN.R-project.org/package=nparLD).</i>
--------------	--

Description

#' Only significant main and interaction effects are reported. P-values are rounded for the third digit and relative treatment effects (RTE) are included when available. Attention: the independent variables of the formula and the term specifying the participant must be factors (i.e., use as.factor()).

Usage

```
reportNparLD(model, dv = "Testdependentvariable", write_to_clipboard = FALSE)
```

Arguments

model	the model
dv	the dependent variable
write_to_clipboard	whether to write to the clipboard

Details

#' To easily copy and paste the results to your manuscript, the following commands must be defined in LaTeX: `\newcommand{\F}{\textit{F=}}` `\newcommand{\p}{\textit{p=}}` `\newcommand{\pminor}{\textit{p$<$}}`

Value

A message describing the statistical results.

Examples

```
if (requireNamespace("nparLD", quietly = TRUE)) {
  # Small toy data set for nparLD
  set.seed(123)
  example_data <- data.frame(
    Subject = factor(rep(1:10, each = 3)),
    Time     = factor(rep(c("T1", "T2", "T3"), times = 10)),
    TLX1     = stats::rnorm(30, mean = 50, sd = 10)
  )

  # Fit nparLD model
  model <- nparLD::nparLD(
    TLX1 ~ Time,
    data      = example_data,
    subject   = "Subject",
    description = FALSE
  )
}
```

```

)

# Report the nparLD result
reportNparLD(model, dv = "TLX1")
}

```

reportNPAV

Generate the Latex-text based on the NPAV by Lüpsen (see <https://www.uni-koeln.de/~{}luepsen/R/>). Only significant main and interaction effects are reported. P-values are rounded for the third digit and partial eta squared values are provided when possible. Attention: the independent variables of the formula and the term specifying the participant must be factors (i.e., use `as.factor()`).

Description

Deprecated: `reportNPAV()` will be removed in `colleyRstats` 0.1.0. Use `reportART()` with `ARTool` instead.

Usage

```
reportNPAV(model, dv = "Testdependentvariable", write_to_clipboard = FALSE)
```

Arguments

<code>model</code>	the model of the <code>np.anova</code>
<code>dv</code>	the name of the dependent variable that should be reported
<code>write_to_clipboard</code>	whether to write to the clipboard

Details

To easily copy and paste the results to your manuscript, the following commands must be defined in

Latex: `\newcommand{\F}[3]{\mathcal{F}(\{#1\},\{#2\})=\{#3\}}` `\newcommand{\p}{\textit{p=}}` `\newcommand{\pminor}{\textit{p=}}`

Value

A message describing the statistical results.

Examples

```

model <- data.frame(
  Df = c(1, 1, 10),
  `F value` = c(6.12, 5.01, NA),
  `Pr(>F)` = c(0.033, 0.045, NA),
  check.names = FALSE
)
rownames(model) <- c("Video", "gesture:eHMI", "Residuals")
reportNPAV(model, dv = "mental workload")

```

reshape_data	<i>Reshape Excel Data Based on Custom Markers and Include Custom ID Column</i>
--------------	--

Description

This function takes an Excel file with data in a wide format and transforms it to a long format. It includes a customizable "ID" column in the first position and repeats it for each slice. The function identifies sections of columns between markers that start with a user-defined string (default is "videoinfo") and appends those sections under the first section, aligning by column index.

Usage

```
reshape_data(
  input_filepath,
  sheetName = "Results",
  marker = "videoinfo",
  id_col = "ID",
  output_filepath
)
```

Arguments

input_filepath	String, the file path of the input Excel file.
sheetName	String, the name of the sheet to read from the Excel file. Default is "Results".
marker	String, the string that identifies the start of a new section of columns. Default is "videoinfo".
id_col	String, the name of the column to use as the ID column. Default is "ID".
output_filepath	String, the file path for the output Excel file.

Details

Relevant if you receive data in wide-format but cannot use built-in functionality due to naming (e.g., in LimeSurvey)

Value

None, writes the reshaped data to an Excel file specified by output_filepath.

Examples

```
if (requireNamespace(c("write_xlsx", "readxl"), quietly = TRUE)) {
  tmp_in <- tempfile(fileext = ".xlsx")
  tmp_out <- tempfile(fileext = ".xlsx")

  # Minimal toy input that includes your required pieces:
```

```

# an ID column and something that contains the marker value.
toy <- data.frame(
  ID = c(1, 1, 2, 2),
  section = c("videoinfo", "videoinfo", "videoinfo", "videoinfo"),
  key = c("fps", "duration_s", "fps", "duration_s"),
  value = c(30, 12.3, 25, 9.8),
  stringsAsFactors = FALSE
)

writexl::write_xlsx(toy, tmp_in)

reshape_data(
  input_filepath = tmp_in,
  marker = "videoinfo",
  id_col = "ID",
  output_filepath = tmp_out
)

out <- readxl::read_excel(tmp_out)
print(out)
}

```

rFromNPAV

Calculation based on Rosenthal's formula (1994). N stands for the number of measurements. Necessary command:

Description

Calculation based on Rosenthal's formula (1994). N stands for the *number of measurements*. Necessary command:

Usage

```
rFromNPAV(pvalue, N)
```

Arguments

pvalue	p value
N	number of measurements in the experiment

Value

Invisibly returns a list with components:

- r: effect size as a numeric scalar.
- z: corresponding z-statistic.
- text: LaTeX-formatted character string that is also sent to the console.

Examples

```
rFromNPAV(0.02, N = 180)
```

rFromWilcox	<i>Calculation based on Rosenthal's formula (1994). N stands for the number of measurements.</i>
-------------	--

Description

Calculation based on Rosenthal's formula (1994). N stands for the *number of measurements*.

Usage

```
rFromWilcox(wilcoxModel, N)
```

Arguments

wilcoxModel	the Wilcox model
N	number of measurements in the experiment

Value

Invisibly returns a list with components:

- r: effect size as a numeric scalar.
- z: corresponding z-statistic.
- text: character string that is also sent to the console.

Examples

```
set.seed(1)
d <- data.frame(
  group = rep(c("A", "B"), each = 10),
  value = rnorm(20)
)
w <- stats::wilcox.test(value ~ group, data = d, exact = FALSE)
rFromWilcox(w, N = nrow(d))
```

rFromWilcoxAdjusted	<i>rFromWilcoxAdjusted</i>
---------------------	----------------------------

Description

rFromWilcoxAdjusted

Usage

```
rFromWilcoxAdjusted(wilcoxModel, N, adjustFactor)
```

Arguments

wilcoxModel	the Wilcox model
N	number of measurements in the experiment
adjustFactor	ad adjustment factor

Value

Invisibly returns a list with components:

- r: adjusted effect size as a numeric scalar.
- z: adjusted z-statistic.
- text: character string that is also sent to the console.

Examples

```
set.seed(1)
d <- data.frame(
  group = rep(c("A", "B"), each = 10),
  value = rnorm(20)
)
w <- stats::wilcox.test(value ~ group, data = d, exact = FALSE)
rFromWilcoxAdjusted(w, N = nrow(d), adjustFactor = 2)
```

stat_sum_df	<i>Generating the sum and adding a crossbar.</i>
-------------	--

Description

Generating the sum and adding a crossbar.

Usage

```
stat_sum_df(fun, geom = "crossbar", ...)
```

Arguments

<code>fun</code>	function
<code>geom</code>	geom to be shown
<code>...</code>	Additional arguments passed to <code>stat_summary</code>

Value

A `ggplot2` layer that can be added to a `ggplot` object.

Examples

```
# Simple summary function: use the mean as y, ymin, and ymax
mean_fun <- function(x) {
  m <- mean(x, na.rm = TRUE)
  data.frame(y = m, ymin = m, ymax = m)
}

ggplot2::ggplot(mtcars, ggplot2::aes(x = factor(cyl), y = mpg)) +
  stat_sum_df(mean_fun)
```

Index

`%!in%(not_in)`, [21](#)

`add_pareto_emoa_column`, [3](#)

`check_homogeneity_by_group`, [5](#)
`check_normality_by_group`, [5](#)
`checkAssumptionsForAnova`, [4](#)
`colleyRstats_setup`, [6](#)

`data the data frame`, [7](#)
`debug_contr_error`, [8](#)

`generateEffectPlot`, [9](#)
`generateMoboPlot`, [10](#)
`generateMoboPlot2`, [12](#)
`ggbetweenstatsWithPriorNormalityCheck`,
[13](#)
`ggbetweenstatsWithPriorNormalityCheckAsterisk`,
[14](#)
`ggwithinstatsWithPriorNormalityCheck`,
[16](#)
`ggwithinstatsWithPriorNormalityCheckAsterisk`,
[17](#)

`latexify_report`, [18](#)

`n_fun`, [22](#)
`na.zero`, [19](#)
`normalize`, [20](#)
`not_empty`, [21](#)
`not_in`, [21](#)
`nparLD`, [30](#)

`pathPrep`, [22](#)

`remove_outliers_REI`, [23](#)
`replace_values(data the data frame)`, [7](#)
`reportART`, [23](#)
`reportDunnTest`, [24](#)
`reportDunnTestTable`, [25](#)
`reportggstatsplot`, [27](#)
`reportggstatsplotPostHoc`, [28](#)
`reportMeanAndSD`, [29](#)
`reportNparLD`, [30](#)
`reportNPAV`, [31](#)
`reshape_data`, [32](#)
`rFromNPAV`, [33](#)
`rFromWilcox`, [34](#)
`rFromWilcoxAdjusted`, [35](#)

`stat_sum_df`, [35](#)