

Package ‘ggstatsplot’

March 17, 2019

Type Package

Title 'ggplot2' Based Plots with Statistical Details

Version 0.0.10

Maintainer Indrajeet Patil <patilindrajeet.science@gmail.com>

Description Extension of 'ggplot2', 'ggstatsplot' creates graphics with details from statistical tests included in the plots themselves. It is targeted primarily at behavioral sciences community to provide a one-line code to generate information-rich plots for statistical analysis of continuous (violin plots, scatterplots, histograms, dot plots, dot-and-whisker plots) or categorical (pie and bar charts) data. Currently, it supports only the most common types of statistical tests: parametric, nonparametric, robust, and bayesian versions of t-test/anova, correlation analyses, contingency table analysis, and regression analyses.

License GPL-3 | file LICENSE

URL <https://indrajeetpatil.github.io/ggstatsplot/>,
<https://github.com/IndrajeetPatil/ggstatsplot>

BugReports <https://github.com/IndrajeetPatil/ggstatsplot/issues>

Depends R (>= 3.5.0)

Imports BayesFactor (>= 0.9.12-4.2), boot (>= 1.3-20), broomExtra, cowplot (>= 0.9.4), crayon (>= 1.3.4), dplyr (>= 0.8.0.1), ellipsis (>= 0.1.0), ez (>= 4.4-0), ggcorrplot (>= 0.1.2), ggExtra (>= 0.8), ggplot2 (>= 3.1.0), ggrepel (>= 0.8.0), ggsignif (>= 0.5.0), glue (>= 1.3.0), grid, groupedstats (>= 0.0.5), jmv (>= 0.9.6), magrittr (>= 1.5), metafor (>= 2.0-0), methods, paletteer (>= 0.2.1), psych (>= 1.8.12), purrr (>= 0.3.2), purrrlyr (>= 0.0.5), rcompanion (>= 2.1.1), rlang (>= 0.3.1), scales (>= 1.0.0), sjstats (>= 0.17.3), stats, stringr (>= 1.3.1), tibble (>= 2.1.1), tidyr (>= 0.8.3), utils, WRS2 (>= 0.10-0)

Suggests broom, broom.mixed, glmmTMB, gmm, knitr, lme4, MCMCglmm, ordinal, readr, rmarkdown, robust, spelling, survival, testthat

VignetteBuilder knitr

Encoding UTF-8

Language en-US

LazyData true

RoxygenNote 6.1.1

NeedsCompilation no

Author Indrajeet Patil [cre, aut, ctb]

(<<https://orcid.org/0000-0003-1995-6531>>),

Chuck Powell [aut, ctb] (<<https://orcid.org/0000-0002-3606-2188>>)

Repository CRAN

Date/Publication 2019-03-17 17:50:02 UTC

R topics documented:

ggstatsplot-package	3
bartlett_message	4
bf_caption_maker	6
bf_contingency_tab	7
bf_corr_test	8
bf_extractor	9
bf_oneway_anova	10
bf_one_sample_ttest	11
bf_two_sample_ttest	12
combine_plots	14
ggbarstats	17
ggbetweenstats	21
ggcoefstats	26
ggcorrmatrix	33
ggdotplotstats	36
gghistostats	39
ggpiestats	43
ggplot_converter	46
ggscatterstats	47
grouped_ggbarstats	51
grouped_ggbetweenstats	56
grouped_ggcorrmatrix	62
grouped_ggdotplotstats	67
grouped_gghistostats	71
grouped_ggpiestats	76
grouped_ggscatterstats	81
intent_morality	86
iris_long	88
movies_long	89
movies_wide	90
normality_message	91

outlier_df	92
pairwise_p	93
subtitle_anova_bayes	96
subtitle_anova_parametric	97
subtitle_anova_robust	99
subtitle_contingency_tab	100
subtitle_friedman_nonparametric	102
subtitle_ggscatterstats	103
subtitle_kw_nonparametric	104
subtitle_mann_nonparametric	106
subtitle_meta_ggcoefstats	108
subtitle_onesample_proptest	110
subtitle_template	111
subtitle_t_bayes	112
subtitle_t_onesample	114
subtitle_t_parametric	115
subtitle_t_robust	117
theme_corrmat	118
theme_ggstatsplot	119
theme_pie	120
Titanic_full	120
VR_dilemma	121

Index	123
--------------	------------

ggstatsplot-package	<i>ggstatsplot: 'ggplot2' Based Plots with Statistical Details</i>
---------------------	--

Description

ggstatsplot is an extension of ggplot2 package for creating graphics with details from statistical tests included in the plots themselves and targeted primarily at behavioral sciences community to provide a one-line code to produce information-rich plots. In a typical exploratory data analysis workflow, data visualization and statistical modelling are two different phases: visualization informs modelling, and modelling in its turn can suggest a different visualization method, and so on and so forth. The central idea of ggstatsplot is simple: combine these two phases into one in the form of graphics with statistical details, which makes data exploration simpler and faster. Currently, it supports only the most common types of statistical tests (parametric, nonparametric, bayesian, and robust versions of t-test/anova, correlation, regression, and contingency tables analyses).

Details

ggstatsplot

The main functions are-

- [ggbetweenstats](#) function to produce information-rich comparison plot between different groups or conditions with ggplot2 and details from the statistical tests in the subtitle.

- `ggscatterstats` function to produce ggplot2 scatterplots along with a marginal histograms/boxplots/density plots from ggExtra and details from the statistical tests in the subtitle.
- `ggpiestats` function to produce pie chart with details from the statistical tests in the subtitle.
- `ggbarstats` function to produce stacked bar chart with details from the statistical tests in the subtitle.
- `gghistostats` function to produce histogram for a single variable with results from one sample test displayed in the subtitle.
- `ggdotplotstats` function to produce Cleveland-style dot plots/charts for a single variable with labels and results from one sample test displayed in the subtitle.
- `ggcorrmat` function to visualize the correlation matrix.
- `ggcoefstats` function to visualize results from regression analyses.
- `combine_plots` helper function to combine multiple ggstatsplot plots using `cowplot::plot_grid()` with a combination of title, caption, and annotation label.
- `theme_ggstatsplot` default theme used for this package.

For more documentation, see the dedicated [Website](#).

Author(s)

Maintainer: Indrajeet Patil <patilindrajeet.science@gmail.com> (0000-0003-1995-6531) [contributor]

Authors:

- Chuck Powell <ibecav@gmail.com> (0000-0002-3606-2188) [contributor]

See Also

Useful links:

- <https://indrajeetpatil.github.io/ggstatsplot/>
- <https://github.com/IndrajeetPatil/ggstatsplot>
- Report bugs at <https://github.com/IndrajeetPatil/ggstatsplot/issues>

bartlett_message

Display homogeneity of variance test as a message

Description

A note to the user about the validity of assumptions for the default linear model.

Usage

```
bartlett_message(data, x, y, lab = NULL, k = 2, output = "message",
...)
```

Arguments

data	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
x	The grouping variable from the dataframe data.
y	The response (a.k.a. outcome or dependent) variable from the dataframe data.
lab	A character describing label for the variable. If NULL, variable name will be used.
k	Number of digits after decimal point (should be an integer) (Default: k = 2).
output	What output is desired: "message" (default) or "stats" (or "tidy") objects.
...	Additional arguments (ignored).

Value

A list of class "htest" containing the following components:

statistic	Bartlett's K-squared test statistic.
parameter	the degrees of freedom of the approximate chi-squared distribution of the test statistic.
p.value	the p-value of the test.
method	the character string "Bartlett test of homogeneity of variances".
data.name	a character string giving the names of the data.

Author(s)

Indrajeet Patil

See Also

[ggbetweenstats](#)

Other helper_messages: [effsize_ci_message](#), [ggcorrmat_matrix_message](#), [grouped_message](#), [normality_message](#), [pairwise_p](#), [palette_message](#)

Examples

```
# getting message
ggstatsplot::bartlett_message(
  data = iris,
  x = Species,
  y = Sepal.Length,
  lab = "Iris Species"
)

# getting results from the test
ggstatsplot::bartlett_message(
  data = mtcars,
  x = am,
```

```

    y = wt,
    output = "tidy"
  )

```

bf_caption_maker	<i>Prepare caption with bayes factor in favor of null</i>
------------------	---

Description

Convenience function to write a caption message with bayes factors in favor of the null hypothesis.

Usage

```
bf_caption_maker(bf.df, k = 2, caption = NULL, ...)
```

Arguments

bf.df	A dataframe containing two columns <code>log_e_bf01</code> and <code>bf.prior</code> . If dataframe contains more than two rows, only the first row will be used.
k	Number of digits after decimal point (should be an integer) (Default: <code>k = 2</code>).
caption	Text to display as caption (will be displayed on top of the bayes factor caption/message).
...	Additional arguments (ignored).

Examples

```

set.seed(123)

# dataframe containing results
bf_results <-
  bf_extractor(BayesFactor::correlationBF(
    x = iris$Sepal.Length,
    y = iris$Petal.Length
  )) %>%
  dplyr::mutate(.data = ., bf.prior = 0.707)

# creating caption
ggstatsplot::bf_caption_maker(
  bf.df = bf_results,
  k = 3,
  caption = "Note: Iris dataset"
)

```

bf_contingency_tab *Bayesian contingency table analysis.*

Description

Bayesian contingency table analysis.

Usage

```
bf_contingency_tab(data, main, condition, sampling.plan = "indepMulti",
  fixed.margin = "rows", prior.concentration = 1, caption = NULL,
  output = "caption", k = 2, ...)
```

Arguments

data	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
main	The variable to use as the rows in the contingency table.
condition	The variable to use as the columns in the contingency table.
sampling.plan	Character describing the sampling plan. Possible options are "indepMulti" (independent multinomial; default), "poisson", "jointMulti" (joint multinomial), "hypergeom" (hypergeometric). For more, see ?BayesFactor::contingencyTableBF().
fixed.margin	For the independent multinomial sampling plan, which margin is fixed ("rows" or "cols"). Defaults to "rows".
prior.concentration	Specifies the prior concentration parameter, set to 1 by default. It indexes the expected deviation from the null hypothesis under the alternative, and corresponds to Gunel and Dickey's (1974) "a" parameter.
caption	The text for the plot caption.
output	Can either be "caption" (which will contain text for evidence in favor of null) or "results" (which will return the dataframe with results).
k	Number of digits after decimal point (should be an integer) (Default: k = 2).
...	further arguments to be passed to or from methods.

Author(s)

Indrajeet Patil

See Also

[bf_corr_test](#), [bf_oneway_anova](#), [bf_two_sample_ttest](#)

Examples

```
# for reproducibility
set.seed(123)

# to get caption (default)
bf_contingency_tab(
  data = mtcars,
  main = am,
  condition = cyl,
  fixed.margin = "cols"
)

# to see results
bf_contingency_tab(
  data = mtcars,
  main = am,
  condition = cyl,
  sampling.plan = "jointMulti",
  fixed.margin = "rows",
  prior.concentration = 1
)
```

bf_corr_test	<i>Bayesian correlation test.</i>
--------------	-----------------------------------

Description

Bayesian correlation test.

Usage

```
bf_corr_test(data, x, y, bf.prior = 0.707, caption = NULL,
  output = "caption", k = 2, ...)
```

Arguments

data	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
x	second continuous variable
y	first continuous variable
bf.prior	A number between 0.5 and 2 (default 0.707), the prior width to use in calculating Bayes factors.
caption	The text for the plot caption.
output	Can either be "caption" (which will contain text for evidence in favor of null) or "results" (which will return the dataframe with results).
k	Number of digits after decimal point (should be an integer) (Default: k = 2).
...	further arguments to be passed to or from methods.

Author(s)

Indrajeet Patil

See Also[bf_contingency_tab](#), [bf_oneway_anova](#), [bf_two_sample_ttest](#)**Examples**

```
# for reproducibility
set.seed(123)

# to get caption (default)
bf_corr_test(
  data = anscombe,
  x = x1,
  y = y4,
  bf.prior = 1
)

# to see results
bf_corr_test(
  data = anscombe,
  x = x1,
  y = y4,
  bf.prior = 0.8,
  output = "results"
)
```

bf_extractor	<i>Convenience function to extract bayes factors from BayesFactor model object.</i>
--------------	---

Description

Convenience function to extract bayes factors from BayesFactor model object.

Usage

```
bf_extractor(bf.object, ...)
```

Arguments

bf.object	An object from BayesFactor package test results.
...	Currently ignored.

Examples

```
# getting only bayes factors
ggstatsplot::bf_extractor(
  BayesFactor::anovaBF(Sepal.Length ~ Species,
    data = iris,
    progress = FALSE
  )
)
```

bf_oneway_anova	<i>Bayesian one-way analysis of variance.</i>
-----------------	---

Description

Bayesian one-way analysis of variance.

Usage

```
bf_oneway_anova(data, x, y, bf.prior = 0.707, caption = NULL,
  output = "caption", k = 2, ...)
```

Arguments

data	a data frame containing data for all factors in the formula
x	The grouping variable from the dataframe data.
y	The response (a.k.a. outcome or dependent) variable from the dataframe data.
bf.prior	A number between 0.5 and 2 (default 0.707), the prior width to use in calculating Bayes factors.
caption	The text for the plot caption.
output	Can either be "caption" (which will contain text for evidence in favor of null) or "results" (which will return the dataframe with results).
k	Number of digits after decimal point (should be an integer) (Default: k = 2).
...	Additional arguments.

Author(s)

Indrajeet Patil

See Also

[bf_contingency_tab](#), [bf_corr_test](#), [bf_two_sample_ttest](#)

Examples

```
# to get caption (default)
bf_oneway_anova(
  data = iris,
  x = Species,
  y = Sepal.Length,
  bf.prior = 0.8
)

# to get results dataframe
bf_oneway_anova(
  data = iris,
  x = Species,
  y = Sepal.Length,
  bf.prior = 0.8,
  output = "results"
)
```

bf_one_sample_ttest	<i>Bayesian one-sample t-test.</i>
---------------------	------------------------------------

Description

Bayesian one-sample t -test.

Usage

```
bf_one_sample_ttest(data = NULL, x, test.value = 0, bf.prior = 0.707,
  caption = NULL, output = "caption", k = 2, ...)
```

Arguments

data	for use with formula, a data frame containing all the data
x	a vector of observations for the first (or only) group
test.value	A number specifying the value of the null hypothesis (Default: 0).
bf.prior	A number between 0.5 and 2 (default 0.707), the prior width to use in calculating Bayes factors.
caption	The text for the plot caption.
output	Can either be "caption" (which will contain text for evidence in favor of null) or "results" (which will return the dataframe with results).
k	Number of digits after decimal point (should be an integer) (Default: k = 2).
...	further arguments to be passed to or from methods.

Author(s)

Indrajeet Patil

See Also

[bf_contingency_tab](#), [bf_oneway_anova](#), [bf_two_sample_ttest](#)

Examples

```
# to get caption (default)
bf_one_sample_ttest(
  data = iris,
  x = Sepal.Length,
  test.value = 5.85,
  bf.prior = 0.8,
  output = "caption", k = 2
)

# to get results dataframe
bf_one_sample_ttest(
  data = iris,
  x = Sepal.Length,
  test.value = 5.85,
  bf.prior = 0.8,
  output = "results"
)
```

bf_two_sample_ttest *Bayesian two-samples t-test.*

Description

Bayesian two-samples *t*-test.

Usage

```
bf_two_sample_ttest(data, x, y, paired = FALSE, bf.prior = 0.707,
  caption = NULL, output = "caption", k = 2, ...)
```

Arguments

data	for use with formula, a data frame containing all the data
x	a vector of observations for the first (or only) group
y	a vector of observations for the second group (or condition, for paired)
paired	if TRUE, observations are paired
bf.prior	A number between 0.5 and 2 (default 0.707), the prior width to use in calculating Bayes factors.
caption	The text for the plot caption.
output	Can either be "caption" (which will contain text for evidence in favor of null) or "results" (which will return the dataframe with results).

k Number of digits after decimal point (should be an integer) (Default: k = 2).
... further arguments to be passed to or from methods.

Author(s)

Indrajeet Patil

See Also

[bf_contingency_tab](#), [bf_corr_test](#), [bf_oneway_anova](#)

Examples

```
# for reproducibility
set.seed(123)

# to get caption (default)
bf_two_sample_ttest(
  data = mtcars,
  x = am,
  y = wt,
  paired = FALSE,
  bf.prior = 0.880
)

# to see results
bf_two_sample_ttest(
  data = mtcars,
  x = am,
  y = wt,
  paired = FALSE,
  output = "results"
)

# for paired sample test
bf_two_sample_ttest(
  data = dplyr::filter(
    ggstatsplot::intent_morality,
    condition %in% c("accidental", "attempted"),
    harm == "Poisoning"
  ),
  x = condition,
  y = rating,
  paired = TRUE,
  output = "results"
)
```

Description

Wrapper around [plot_grid](#) that will return a plotgrid along with a combination of title, caption, and annotation label

Usage

```
combine_plots(..., title.text = NULL, title.color = "black",
  title.size = 16, title.vjust = 0.5, title.hjust = 0.5,
  title.fontface = "bold", caption.text = NULL,
  caption.color = "black", caption.size = 10, caption.vjust = 0.5,
  caption.hjust = 0.5, caption.fontface = "plain", sub.text = NULL,
  sub.color = "black", sub.size = 12, sub.vjust = 0.5,
  sub.hjust = 0.5, sub.fontface = "plain", sub.x = 0.5,
  sub.y = 0.5, sub.vpadding = grid::unit(1, "lines"), sub.angle = 0,
  sub.lineheight = 0.9, title.rel.heights = c(0.1, 1.2),
  caption.rel.heights = c(1.2, 0.1), title.caption.rel.heights = c(0.1,
  1.2, 0.1))
```

Arguments

- ... Arguments passed on to `cowplot::plot_grid`
- plotlist** (optional) List of plots to display. Alternatively, the plots can be provided individually as the first `n` arguments of the function `plot_grid` (see examples).
- align** (optional) Specifies whether graphs in the grid should be horizontally ("h") or vertically ("v") aligned. Options are "none" (default), "hv" (align in both directions), "h", and "v".
- axis** (optional) Specifies whether graphs should be aligned by the left ("l"), right ("r"), top ("t"), or bottom ("b") margins. Options are "none" (default), or a string of any combination of l, r, t, and b in any order (e.g. "tblr" or "rlbt" for aligning all margins). Must be specified if any of the graphs are complex (e.g. faceted) and alignment is specified and desired. See [align_plots\(\)](#) for details.
- nrow** (optional) Number of rows in the plot grid.
- ncol** (optional) Number of columns in the plot grid.
- rel_widths** (optional) Numerical vector of relative columns widths. For example, in a two-column grid, `rel_widths = c(2, 1)` would make the first column twice as wide as the second column.
- rel_heights** (optional) Numerical vector of relative columns heights. Works just as `rel_widths` does, but for rows rather than columns.

labels (optional) List of labels to be added to the plots. You can also set `labels="AUTO"` to auto-generate upper-case labels or `labels="auto"` to auto-generate lower-case labels.

label_size (optional) Numerical value indicating the label size. Default is 14.

label_fontfamily (optional) Font family of the plot labels. If not provided, is taken from the current theme.

label_fontface (optional) Font face of the plot labels. Default is "bold".

label_colour (optional) Color of the plot labels. If not provided, is taken from the current theme.

label_x (optional) Single value or vector of x positions for plot labels, relative to each subplot. Defaults to 0 for all labels. (Each label is placed all the way to the left of each plot.)

label_y (optional) Single value or vector of y positions for plot labels, relative to each subplot. Defaults to 1 for all labels. (Each label is placed all the way to the top of each plot.)

hjust Adjusts the horizontal position of each label. More negative values move the label further to the right on the plot canvas. Can be a single value (applied to all labels) or a vector of values (one for each label). Default is -0.5.

vjust Adjusts the vertical position of each label. More positive values move the label further down on the plot canvas. Can be a single value (applied to all labels) or a vector of values (one for each label). Default is 1.5.

scale Individual number or vector of numbers greater than 0. Enables you to scale the size of all or select plots. Usually it's preferable to set margins instead of using scale, but scale can sometimes be more powerful.

cols Deprecated. Use `ncol`.

rows Deprecated. Use `nrow`.

<code>title.text</code>	String or plotmath expression to be drawn as title for the <i>combined plot</i> .
<code>title.color</code>	Text color for title.
<code>title.size</code>	Point size of title text.
<code>title.vjust</code>	Vertical justification for title. Default = 0.5 (centered on y). 0 = baseline at y, 1 = ascender at y.
<code>title.hjust</code>	Horizontal justification for title. Default = 0.5 (centered on x). 0 = flush-left at x, 1 = flush-right.
<code>title.fontface</code>	The font face ("plain", "bold" (default), "italic", "bold.italic") for title.
<code>caption.text</code>	String or plotmath expression to be drawn as the caption for the <i>combined plot</i> .
<code>caption.color</code>	Text color for caption.
<code>caption.size</code>	Point size of title text.
<code>caption.vjust</code>	Vertical justification for caption. Default = 0.5 (centered on y). 0 = baseline at y, 1 = ascender at y.
<code>caption.hjust</code>	Horizontal justification for caption. Default = 0.5 (centered on x). 0 = flush-left at x, 1 = flush-right.

caption.fontface	The font face ("plain" (default), "bold", "italic", "bold.italic") for caption.
sub.text	The label with which the <i>combined plot</i> should be annotated. Can be a plotmath expression.
sub.color	Text color for annotation label (Default: "black").
sub.size	Point size of annotation text (Default: 12).
sub.vjust	Vertical justification for annotation label (Default: 0.5).
sub.hjust	Horizontal justification for annotation label (Default: 0.5).
sub.fontface	The font face ("plain" (default), "bold", "italic", "bold.italic") for the annotation label.
sub.x	The x position of annotation label (Default: 0.5).
sub.y	The y position of annotation label (Default: 0.5).
sub.vpadding	Vertical padding. The total vertical space added to the label, given in grid units. By default, this is added equally above and below the label. However, by changing the y and vjust parameters, this can be changed (Default: <code>grid::unit(1, "lines")</code>).
sub.angle	Angle at which annotation label is to be drawn (Default: 0).
sub.lineheight	Line height of annotation label.
title.rel.heights	Numerical vector of relative columns heights while combining (title, plot).
caption.rel.heights	Numerical vector of relative columns heights while combining (plot, caption).
title.caption.rel.heights	Numerical vector of relative columns heights while combining (title, plot, caption).

Value

Combined plot with title and/or caption and/or annotation label

Author(s)

Indrajeet Patil

References

https://indrajeetpatil.github.io/ggstatsplot/articles/web_only/combine_plots.html

Examples

```
# loading the necessary libraries
library(ggplot2)

# preparing the first plot
p1 <-
```



```

ggplot2::ggplot(
  data = subset(iris, iris$Species == "setosa"),
  aes(x = Sepal.Length, y = Sepal.Width)
) +
geom_point() +
labs(title = "setosa")

# preparing the second plot
p2 <-
ggplot2::ggplot(
  data = subset(iris, iris$Species == "versicolor"),
  aes(x = Sepal.Length, y = Sepal.Width)
) +
geom_point() +
labs(title = "versicolor")

# combining the plot with a title and a caption
combine_plots(
  p1,
  p2,
  labels = c("(a)", "(b)"),
  title.text = "Dataset: Iris Flower dataset",
  caption.text = "Note: Only two species of flower are displayed",
  title.color = "red",
  caption.color = "blue"
)

```

Description

Bar charts for categorical data with statistical details included in the plot as a subtitle.

Usage

```

ggbarstats(data, main, condition = NULL, counts = NULL, ratio = NULL,
  paired = FALSE, labels.legend = NULL, results.subtitle = TRUE,
  stat.title = NULL, sample.size.label = TRUE, label.separator = " ",
  label.text.size = 4, label.fill.color = "white",
  label.fill.alpha = 1, bar.outline.color = "black",
  bf.message = FALSE, sampling.plan = "jointMulti",
  fixed.margin = "rows", prior.concentration = 1, title = NULL,
  subtitle = NULL, caption = NULL, legend.position = "right",
  x.axis.orientation = NULL, conf.level = 0.95, nboot = 100,
  simulate.p.value = FALSE, B = 2000, legend.title = NULL,
  xlab = NULL, ylab = "Percent", k = 2, perc.k = 0,
  data.label = "percentage", bar.proptest = TRUE,
  ggtheme = ggplot2::theme_bw(), ggstatsplot.layer = TRUE,

```

```
package = "RColorBrewer", palette = "Dark2", direction = 1,
ggplot.component = NULL, messages = TRUE)
```

Arguments

<code>data</code>	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
<code>main</code>	The variable to use as the rows in the contingency table.
<code>condition</code>	The variable to use as the columns in the contingency table.
<code>counts</code>	A string naming a variable in data containing counts, or NULL if each row represents a single observation (Default).
<code>ratio</code>	A vector of numbers: the expected proportions for the proportion test. Default is NULL, which means if there are two levels <code>ratio = c(1,1)</code> , etc.
<code>paired</code>	Logical indicating whether data came from a within-subjects design study (Default: FALSE). If TRUE, McNemar test subtitle will be returned. If FALSE, Pearson's chi-square test will be returned.
<code>labels.legend</code>	A character vector with custom labels for levels of the main variable displayed in the legend.
<code>results.subtitle</code>	Decides whether the results of statistical tests are to be displayed as a subtitle (Default: TRUE). If set to FALSE, only the plot will be returned.
<code>stat.title</code>	Title for the effect being investigated with the chi-square test. The default is NULL, i.e. no title will be added to describe the effect being shown. An example of a <code>stat.title</code> argument will be something like "main x condition" or "interaction".
<code>sample.size.label</code>	Logical that decides whether sample size information should be displayed for each level of the grouping variable <code>condition</code> (Default: TRUE).
<code>label.separator</code>	If "both" counts and proportion information is to be displayed in a label, this argument decides whether these two pieces of information are going to be on the same line (" ") or on separate lines ("\n").
<code>label.text.size</code>	Numeric that decides text size for slice/bar labels (Default: 4).
<code>label.fill.color</code>	Character that specifies fill color for slice/bar labels (Default: white).
<code>label.fill.alpha</code>	Numeric that specifies fill color transparency or "alpha" for slice/bar labels (Default: 1 range 0 to 1).
<code>bar.outline.color</code>	Character specifying color for bars (default: "black").
<code>bf.message</code>	Logical that decides whether to display a caption with results from bayes factor test in favor of the null hypothesis (default: FALSE).
<code>sampling.plan</code>	Character describing the sampling plan. Possible options are "indepMulti" (independent multinomial; default), "poisson", "jointMulti" (joint multinomial), "hypergeom" (hypergeometric). For more, see <code>?BayesFactor::contingencyTableBF()</code> .

<code>fixed.margin</code>	For the independent multinomial sampling plan, which margin is fixed ("rows" or "cols"). Defaults to "rows".
<code>prior.concentration</code>	Specifies the prior concentration parameter, set to 1 by default. It indexes the expected deviation from the null hypothesis under the alternative, and corresponds to Gunel and Dickey's (1974) "a" parameter.
<code>title</code>	The text for the plot title.
<code>subtitle</code>	The text for the plot subtitle. Will work only if <code>results.subtitle = FALSE</code> .
<code>caption</code>	The text for the plot caption.
<code>legend.position</code>	The position of the legend "none", "left", "right", "bottom", "top" (Default: "right").
<code>x.axis.orientation</code>	The orientation of the x axis labels one of "slant" or "vertical" to change from the default horizontal orientation (Default: NULL which is horizontal).
<code>conf.level</code>	Scalar between 0 and 1. If unspecified, the defaults return 95% lower and upper confidence intervals (0.95).
<code>nboot</code>	Number of bootstrap samples for computing confidence interval for the effect size (Default: 100).
<code>simulate.p.value</code>	a logical indicating whether to compute p-values by Monte Carlo simulation.
<code>B</code>	an integer specifying the number of replicates used in the Monte Carlo test.
<code>legend.title</code>	Title text for the legend.
<code>xlab</code>	Custom text for the x axis label (Default: NULL, which will cause the x axis label to be the main variable).
<code>ylab</code>	Custom text for the y axis label (Default: "percent").
<code>k</code>	Number of digits after decimal point (should be an integer) (Default: <code>k = 2</code>).
<code>perc.k</code>	Numeric that decides number of decimal places for percentage labels (Default: 0).
<code>data.label</code>	Character decides what information needs to be displayed on the label in each pie slice. Possible options are "percentage" (default), "counts", "both".
<code>bar.proptest</code>	Decides whether proportion test for main variable is to be carried out for each level of condition (Default: TRUE).
<code>ggtheme</code>	A function, ggplot2 theme name. Default value is <code>ggplot2::theme_bw()</code> . Any of the ggplot2 themes, or themes from extension packages are allowed (e.g., <code>ggthemes::theme_fivethirtyeight()</code> , <code>hrbrthemes::theme_ipsum_ps()</code> , etc.).
<code>ggstatsplot.layer</code>	Logical that decides whether <code>theme_ggstatsplot</code> theme elements are to be displayed along with the selected <code>ggtheme</code> (Default: TRUE).
<code>package</code>	Name of package from which the palette is desired as string or symbol.
<code>palette</code>	If a character string (e.g., "Set1"), will use that named palette. If a number, will index into the list of palettes of appropriate type. Default palette is "Dark2".

direction	Either 1 or -1. If -1 the palette will be reversed.
ggplot.component	A ggplot component to be added to the plot prepared by ggstatsplot. This argument is primarily helpful for grouped_ variant of the current function. Default is NULL. The argument should be entered as a function. If the given function has an argument <code>axes.range.restrict</code> and if it has been set to TRUE, the added ggplot component <i>might</i> not work as expected.
messages	Decides whether messages references, notes, and warnings are to be displayed (Default: TRUE).

Value

Unlike a number of statistical softwares, ggstatsplot doesn't provide the option for Yates' correction for the Pearson's chi-squared statistic. This is due to compelling amount of Monte-Carlo simulation research which suggests that the Yates' correction is overly conservative, even in small sample sizes. As such it is recommended that it should not ever be applied in practice (Camilli & Hopkins, 1978, 1979; Feinberg, 1980; Larntz, 1978; Thompson, 1988).

Author(s)

Chuck Powell, Indrajeet Patil

Examples

```
# for reproducibility
set.seed(123)

# simple function call with the defaults (with condition)
ggstatsplot::ggbarstats(
  data = datasets::mtcars,
  main = vs,
  condition = cyl,
  bf.message = TRUE,
  nboot = 10,
  labels.legend = c("0 = V-shaped", "1 = straight"),
  legend.title = "Engine"
)
## Not run:
# simple function call with the defaults (with count data)
library(jmv)

ggstatsplot::ggbarstats(
  data = as.data.frame(HairEyeColor),
  main = Eye,
  condition = Hair,
  counts = Freq
)
## End(Not run)
```

ggbetweenstats	<i>Box/Violin plots for group or condition comparisons in between-subjects designs.</i>
----------------	---

Description

A combination of box and violin plots along with jittered data points for between-subjects designs with statistical details included in the plot as a subtitle.

Usage

```
ggbetweenstats(data, x, y, plot.type = "boxviolin",
  type = "parametric", pairwise.comparisons = FALSE,
  pairwise.annotation = "asterisk", pairwise.display = "significant",
  p.adjust.method = "holm", effsize.type = "unbiased",
  partial = TRUE, effsize.noncentral = TRUE, bf.prior = 0.707,
  bf.message = FALSE, results.subtitle = TRUE, xlab = NULL,
  ylab = NULL, caption = NULL, title = NULL, subtitle = NULL,
  sample.size.label = TRUE, k = 2, var.equal = FALSE,
  conf.level = 0.95, nboot = 100, tr = 0.1,
  axes.range.restrict = FALSE, mean.label.size = 3,
  mean.label.fontface = "bold", mean.label.color = "black",
  notch = FALSE, notchwidth = 0.5, linetype = "solid",
  outlier.tagging = FALSE, outlier.shape = 19, outlier.label = NULL,
  outlier.label.color = "black", outlier.color = "black",
  outlier.coef = 1.5, mean.plotting = TRUE, mean.ci = FALSE,
  mean.size = 5, mean.color = "darkred", point.jitter.width = NULL,
  point.jitter.height = 0, point.dodge.width = 0.6,
  ggtheme = ggplot2::theme_bw(), ggstatsplot.layer = TRUE,
  package = "RColorBrewer", palette = "Dark2", direction = 1,
  ggplot.component = NULL, messages = TRUE)
```

Arguments

data	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
x	The grouping variable from the dataframe data.
y	The response (a.k.a. outcome or dependent) variable from the dataframe data.
plot.type	Character describing the <i>type</i> of plot. Currently supported plots are "box" (for pure boxplots), "violin" (for pure violin plots), and "boxviolin" (for a combination of box and violin plots; default).
type	Type of statistic expected ("parametric" or "nonparametric" or "robust" or "bayes"). Corresponding abbreviations are also accepted: "p" (for parametric), "np" (nonparametric), "r" (robust), or "bf" resp.

<code>pairwise.comparisons</code>	Logical that decides whether pairwise comparisons are to be displayed. Only significant comparisons will be shown by default. (default: FALSE). To change this behavior, select appropriate option with <code>pairwise.display</code> argument.
<code>pairwise.annotation</code>	Character that decides the annotations to use for pairwise comparisons. Either "p.value" or "asterisk" (default).
<code>pairwise.display</code>	Decides which pairwise comparisons to display. Available options are "significant" (abbreviation accepted: "s") or "non-significant" (abbreviation accepted: "ns") or "everything"/"all". The default is "significant". You can use this argument to make sure that your plot is not uber-cluttered when you have multiple groups being compared and scores of pairwise comparisons being displayed.
<code>p.adjust.method</code>	Adjustment method for <i>p</i> -values for multiple comparisons. Possible methods are: "holm" (default), "hochberg", "hommel", "bonferroni", "BH", "BY", "fdr", "none".
<code>effsize.type</code>	Type of effect size needed for <i>parametric</i> tests. The argument can be "biased" ("d" for Cohen's <i>d</i> for t-test ; "partial_eta" for partial eta-squared for anova) or "unbiased" ("g" Hedge's <i>g</i> for t-test ; "partial_omega" for partial omega-squared for anova).
<code>partial</code>	Logical that decides if partial eta-squared or omega-squared are returned (Default: TRUE). If FALSE, eta-squared or omega-squared will be returned. Valid only for objects of class <code>lm</code> , <code>aov</code> , <code>anova</code> , or <code>aovlist</code> .
<code>effsize.noncentral</code>	Logical indicating whether to use non-central <i>t</i> -distributions for computing the confidence interval for Cohen's <i>d</i> or Hedge's <i>g</i> (Default: TRUE).
<code>bf.prior</code>	A number between 0.5 and 2 (default 0.707), the prior width to use in calculating Bayes factors.
<code>bf.message</code>	Logical that decides whether to display Bayes Factor in favor of the <i>null</i> hypothesis for parametric test (Default: FALSE).
<code>results.subtitle</code>	Decides whether the results of statistical tests are to be displayed as a subtitle (Default: TRUE). If set to FALSE, only the plot will be returned.
<code>xlab, ylab</code>	Labels for x and y axis variables. If NULL (default), variable names for x and y will be used.
<code>caption</code>	The text for the plot caption.
<code>title</code>	The text for the plot title.
<code>subtitle</code>	The text for the plot subtitle. Will work only if <code>results.subtitle</code> = FALSE.
<code>sample.size.label</code>	Logical that decides whether sample size information should be displayed for each level of the grouping variable <i>x</i> (Default: TRUE).
<code>k</code>	Number of digits after decimal point (should be an integer) (Default: <code>k = 2</code>).

<code>var.equal</code>	a logical variable indicating whether to treat the variances in the samples as equal. If TRUE, then a simple F test for the equality of means in a one-way analysis of variance is performed. If FALSE, an approximate method of Welch (1951) is used, which generalizes the commonly known 2-sample Welch test to the case of arbitrarily many samples.
<code>conf.level</code>	Scalar between 0 and 1. If unspecified, the defaults return 95% lower and upper confidence intervals (0.95).
<code>nboot</code>	Number of bootstrap samples for computing confidence interval for the effect size (Default: 100).
<code>tr</code>	Trim level for the mean when carrying out robust tests. If you get error stating "Standard error cannot be computed because of Winsorized variance of 0 (e.g., due to ties). Try to decrease the trimming level.", try to play around with the value of <code>tr</code> , which is by default set to 0.1. Lowering the value might help.
<code>axes.range.restrict</code>	Logical that decides whether to restrict the axes values ranges to min and max values of the axes variables (Default: FALSE), only relevant for functions where axes variables are of numeric type.
<code>mean.label.size</code> , <code>mean.label.fontface</code> , <code>mean.label.color</code>	Aesthetics for the label displaying mean. Defaults: 3, "bold", "black", respectively.
<code>notch</code>	A logical. If FALSE (default), a standard box plot will be displayed. If TRUE, a notched box plot will be used. Notches are used to compare groups; if the notches of two boxes do not overlap, this suggests that the medians are significantly different. In a notched box plot, the notches extend $1.58 * \text{IQR} / \sqrt{n}$. This gives a roughly 95% confidence interval for comparing medians. IQR: Inter-Quartile Range.
<code>notchwidth</code>	For a notched box plot, width of the notch relative to the body (default 0.5).
<code>linetype</code>	Character strings ("blank", "solid", "dashed", "dotted", "dotdash", "longdash", and "twodash") specifying the type of line to draw box plots (Default: "solid"). Alternatively, the numbers 0 to 6 can be used (0 for "blank", 1 for "solid", etc.).
<code>outlier.tagging</code>	Decides whether outliers should be tagged (Default: FALSE).
<code>outlier.shape</code>	Hiding the outliers can be achieved by setting <code>outlier.shape = NA</code> . Importantly, this does not remove the outliers, it only hides them, so the range calculated for the y-axis will be the same with outliers shown and outliers hidden.
<code>outlier.label</code>	Label to put on the outliers that have been tagged.
<code>outlier.label.color</code>	Color for the label to put on the outliers that have been tagged (Default: "black").
<code>outlier.color</code>	Default aesthetics for outliers (Default: "black").
<code>outlier.coef</code>	Coefficient for outlier detection using Tukey's method. With Tukey's method, outliers are below (1st Quartile) or above (3rd Quartile) <code>outlier.coef</code> times the Inter-Quartile Range (IQR) (Default: 1.5).
<code>mean.plotting</code>	Logical that decides whether mean is to be highlighted and its value to be displayed (Default: TRUE).

<code>mean.ci</code>	Logical that decides whether 95 is to be displayed (Default: FALSE).
<code>mean.size</code>	Point size for the data point corresponding to mean (Default: 5).
<code>mean.color</code>	Color for the data point corresponding to mean (Default: "darkred").
<code>point.jitter.width</code>	Numeric specifying the degree of jitter in x direction. Defaults to 40% of the resolution of the data.
<code>point.jitter.height</code>	Numeric specifying the degree of jitter in y direction. Defaults to 0.1.
<code>point.dodge.width</code>	Numeric specifying the amount to dodge in the x direction. Defaults to 0.60.
<code>ggtheme</code>	A function, ggplot2 theme name. Default value is <code>ggplot2::theme_bw()</code> . Any of the ggplot2 themes, or themes from extension packages are allowed (e.g., <code>ggthemes::theme_fivethirtyeight()</code> , <code>hrbrthemes::theme_ipsum_ps()</code> , etc.).
<code>ggstatsplot.layer</code>	Logical that decides whether <code>theme_ggstatsplot</code> theme elements are to be displayed along with the selected <code>ggtheme</code> (Default: TRUE).
<code>package</code>	Name of package from which the palette is desired as string or symbol.
<code>palette</code>	If a character string (e.g., "Set1"), will use that named palette. If a number, will index into the list of palettes of appropriate type. Default palette is "Dark2".
<code>direction</code>	Either 1 or -1. If -1 the palette will be reversed.
<code>ggplot.component</code>	A ggplot component to be added to the plot prepared by <code>ggstatsplot</code> . This argument is primarily helpful for <code>grouped_</code> variant of the current function. Default is NULL. The argument should be entered as a function. If the given function has an argument <code>axes.range.restrict</code> and if it has been set to TRUE, the added ggplot component <i>might</i> not work as expected.
<code>messages</code>	Decides whether messages references, notes, and warnings are to be displayed (Default: TRUE).

Details

For parametric tests, Welch's ANOVA/*t*-test are used as a default (i.e., `var.equal = FALSE`).
References:

- ANOVA: Delacre, Leys, Mora, & Lakens, *PsyArXiv*, 2018
- *t*-test: Delacre, Lakens, & Leys, *International Review of Social Psychology*, 2017

If robust tests are selected, following tests are used is .

- ANOVA: one-way ANOVA on trimmed means (see `?WRS2::t1way`)
- *t*-test: Yuen's test for trimmed means (see `?WRS2::yuen`)

Variant of this function `ggwithinstats` is currently in progress. You *can* still use this function just to prepare the **plot** for exploratory data analysis, but the statistical details displayed in the subtitle will be incorrect. You can remove them by adding + `ggplot2::labs(subtitle = NULL)`.

Author(s)

Indrajeet Patil

References

https://indrajeetpatil.github.io/ggstatsplot/articles/web_only/ggbetweenstats.html

See Also

[grouped_ggbetweenstats](#), [pairwise_p](#)

Examples

```
# to get reproducible results from bootstrapping
set.seed(123)

# simple function call with the defaults
ggstatsplot::ggbetweenstats(
  data = mtcars,
  x = am,
  y = mpg,
  title = "Fuel efficiency by type of car transmission",
  caption = "Transmission (0 = automatic, 1 = manual)",
  bf.message = TRUE
)

## Not run:
# more detailed function call
ggstatsplot::ggbetweenstats(
  data = datasets::morley,
  x = Expt,
  y = Speed,
  plot.type = "box",
  conf.level = 0.99,
  xlab = "The experiment number",
  ylab = "Speed-of-light measurement",
  pairwise.comparisons = TRUE,
  pairwise.annotation = "p.value",
  p.adjust.method = "fdr",
  outlier.tagging = TRUE,
  outlier.label = Run,
  nboot = 10,
  ggtheme = ggplot2::theme_grey(),
  ggstatsplot.layer = FALSE,
  bf.message = TRUE
)

## End(Not run)
```

ggcoefstats

Model coefficients for fitted models with the model summary as a caption.

Description

Model coefficients for fitted models with the model summary as a caption.

Usage

```
ggcoefstats(x, output = "plot", statistic = NULL, scales = NULL,
  conf.method = "Wald", conf.type = "Wald", component = "survival",
  p.kr = TRUE, p.adjust.method = "none", coefficient.type = c("beta",
    "location", "coefficient"), by.class = FALSE, effsize = "eta",
  partial = TRUE, nboot = 500, meta.analytic.effect = FALSE,
  point.color = "blue", point.size = 3, point.shape = 16,
  conf.int = TRUE, conf.level = 0.95, se.type = "nid", k = 2,
  k.caption.summary = 0, exclude.intercept = TRUE,
  exponentiate = FALSE, errorbar.color = "black",
  errorbar.height = 0, errorbar.linetype = "solid",
  errorbar.size = 0.5, vline = TRUE, vline.color = "black",
  vline.linetype = "dashed", vline.size = 1, sort = "none",
  xlab = "regression coefficient", ylab = "term", title = NULL,
  subtitle = NULL, stats.labels = TRUE, caption = NULL,
  caption.summary = TRUE, stats.label.size = 3,
  stats.label.fontface = "bold", stats.label.color = NULL,
  label.r = 0.15, label.size = 0.25, label.box.padding = 1,
  label.label.padding = 0.25, label.point.padding = 0.5,
  label.segment.color = "grey50", label.segment.size = 0.5,
  label.segment.alpha = NULL, label.min.segment.length = 0.5,
  label.force = 1, label.max.iter = 2000, label.nudge.x = 0,
  label.nudge.y = 0, label.xlim = c(NA, NA), label.ylim = c(NA, NA),
  label.direction = "y", package = "RColorBrewer", palette = "Dark2",
  direction = 1, ggtheme = ggplot2::theme_bw(),
  ggstatsplot.layer = TRUE, messages = FALSE, ...)
```

Arguments

- | | |
|--------|---|
| x | A model object to be tidied with <code>broom::tidy</code> , or a tidy data frame containing results. If a data frame is to be plotted, it <i>must</i> contain columns named <code>term</code> (names of predictors), or <code>estimate</code> (corresponding estimates of coefficients or other quantities of interest). Other optional columns are <code>conf.low</code> and <code>conf.high</code> (for confidence intervals); <code>p.value</code> . It is important that all term names should be unique. |
| output | Character describing the expected output from this function: <code>"plot"</code> (visualization of regression coefficients) or <code>"tidy"</code> (tidy dataframe of results from |

	broom::tidy) or "glance" (object from broom::glance) or "augment" (object from broom::augment).
statistic	Which statistic is to be displayed (either "t" or "f" or "z") in the label. This is especially important if the x argument in ggcoefstats is a dataframe in which case the function wouldn't know what kind of model it is dealing with.
scales	scales on which to report the variables: for random effects, the choices are "sd-cor" (standard deviations and correlations: the default if scales is NULL) or "vcov" (variances and covariances). NA means no transformation, appropriate e.g. for fixed effects.
conf.method	Character describing method for computing confidence intervals (for more, see ?lme4::confint.merMod and ?broom.mixed::tidy.brmsfit). This argument has different defaults depending on the model object. For the merMod class model objects (lmer, glmer, nlmer, etc.), the default is "Wald" (other options are: "profile", "boot"). For MCMC or brms fit model objects (Stan, JAGS, etc.), the default is "quantile", while the only other options is "HPDinterval".
conf.type	Whether to use "profile" or "Wald" confidence intervals, passed to the type argument of <code>ordinal::confint.clm()</code> . Defaults to "profile".
component	Character specifying whether to tidy the survival or the longitudinal component of the model. Must be either "survival" or "longitudinal". Defaults to "survival".
p.kr	Logical, if TRUE, the computation of <i>p</i> -values for lmer is based on conditional F-tests with Kenward-Roger approximation for the df. For details, see ?sjstats::p_value.
p.adjust.method	Adjustment method for <i>p</i> -values for multiple comparisons. Possible methods are: "holm", "hochberg", "hommel", "bonferroni", "BH", "BY", "fdr", "none". Default is no correction ("none"). This argument is relevant for multiplicity correction for multiway ANOVA designs (see, Cramer et al., 2015).
coefficient.type	Relevant only for ordinal regression models (clm, clmm, "svyolr", and polr), this argument decides which parameters are display in the plot. Available parameters are: parameter that measures the intercept , i.e. the log-odds distance between response values ("alpha"); effects on the location ("beta"); or effects on the scale ("zeta"). For clm and clmm models, by default, only "beta" (a vector of regression parameters) parameters will be show. Other options are "alpha" (a vector of threshold parameters) or "both". For polr models, by default, only "coefficient" will be shown. Other option is to show "zeta" parameters. Note that, from broom 0.7.0 onward, coefficients will be renamed and "intercept" type coefficients will correspond to "alpha" parameters, "location" type coefficients will correspond to "beta" parameters, and "scale" type coefficients will correspond to "zeta" parameters.
by.class	A logical indicating whether or not to show performance measures broken down by class. Defaults to FALSE. When by.class = FALSE only returns a tibble with accuracy and kappa statistics. Mostly relevant for an object of class "confusionMatrix".
effsize	Character describing the effect size to be displayed: "eta" (default) or "omega". This argument is relevant only for models objects of class aov, anova, and aovlist.

<code>partial</code>	Logical that decides if partial eta-squared or omega-squared are returned (Default: TRUE). If FALSE, eta-squared or omega-squared will be returned. Valid only for objects of class <code>aov</code> , <code>anova</code> , or <code>aovlist</code> .
<code>nboot</code>	Number of bootstrap samples for confidence intervals for partial eta-squared and omega-squared (Default: 500). This argument is relevant only for models objects of class <code>aov</code> , <code>anova</code> , and <code>aovlist</code> .
<code>meta.analytic.effect</code>	Logical that decides whether subtitle for meta-analysis via linear (mixed-effects) models - as implemented in the <code>metafor</code> package - is to be displayed (default: FALSE). If TRUE, input to argument <code>subtitle</code> will be ignored. This will be mostly relevant if a data frame with estimates and their standard errors is entered as input to <code>x</code> argument.
<code>point.color</code>	Character describing color for the point (Default: "blue").
<code>point.size</code>	Numeric specifying size for the point (Default: 3).
<code>point.shape</code>	Numeric specifying shape to draw the points (Default: 16 (a dot)).
<code>conf.int</code>	Logical. Decides whether to display confidence intervals as error bars (Default: TRUE).
<code>conf.level</code>	Numeric deciding level of confidence intervals (Default: 0.95). For MCMC model objects (Stan, JAGS, etc.), this will be probability level for CI.
<code>se.type</code>	Character specifying the method used to compute standard standard errors for quantile regression (Default: "nid"). To see all available methods, see <code>quantreg::summary.rq()</code> .
<code>k</code>	Number of decimal places expected for results displayed in labels (Default : <code>k = 2</code>).
<code>k.caption.summary</code>	Number of decimal places expected for results displayed in captions (Default : <code>k.caption.summary = 0</code>).
<code>exclude.intercept</code>	Logical that decides whether the intercept should be excluded from the plot (Default: TRUE).
<code>exponentiate</code>	If TRUE, the x-axis will be logarithmic (Default: FALSE).
<code>errorbar.color</code>	Character deciding color of the error bars (Default: "black").
<code>errorbar.height</code>	Numeric specifying the height of the error bars (Default: 0).
<code>errorbar.linetype</code>	Line type of the error bars (Default: "solid").
<code>errorbar.size</code>	Numeric specifying the size of the error bars (Default: 0.5).
<code>vline</code>	Decides whether to display a vertical line (Default: "TRUE").
<code>vline.color</code>	Character specifying color of the vertical line (Default: "black").
<code>vline.linetype</code>	Character specifying line type of the vertical line (Default: "dashed").
<code>vline.size</code>	Numeric specifying the size of the vertical line (Default: 1).
<code>sort</code>	If "none" (default) do not sort, "ascending" sort by increasing coefficient value, or "descending" sort by decreasing coefficient value.

<code>xlab</code>	Label for x axis variable (Default: "estimate").
<code>ylab</code>	Label for y axis variable (Default: "term").
<code>title</code>	The text for the plot title.
<code>subtitle</code>	The text for the plot subtitle. The input to this argument will be ignored if <code>meta.analytic.effect</code> is set to TRUE.
<code>stats.labels</code>	Logical. Decides whether the statistic and p-values for each coefficient are to be attached to each dot as a text label using <code>ggrepel</code> (Default: TRUE).
<code>caption</code>	The text for the plot caption.
<code>caption.summary</code>	Logical. Decides whether the model summary should be displayed as a cation to the plot (Default: TRUE). Color of the line segment. Defaults to the same color as the text.
<code>stats.label.size</code> , <code>stats.label.fontface</code> , <code>stats.label.color</code>	Aesthetics for the labels. Defaults: 3, "bold", NULL, resp. If <code>stats.label.color</code> is NULL, colors will be chosen from the specified package (Default: "RColorBrewer") and palette (Default: "Dark2").
<code>label.r</code> ,	Radius of rounded corners, as unit or number. Defaults to 0.15. (Default unit is lines).
<code>label.size</code>	Size of label border, in mm. Defaults to 0.25.
<code>label.box.padding</code>	Amount of padding around bounding box, as number. Defaults to 1. (Default unit is lines).
<code>label.label.padding</code>	Amount of padding around label, as number. Defaults to 0.25. (Default unit is lines).
<code>label.point.padding</code>	Amount of padding around labeled point, as number. Defaults to 0. (Default unit is lines).
<code>label.segment.color</code>	Color of the line segment (Default: "grey50").
<code>label.segment.size</code>	Width of line segment connecting the data point to the text label, in mm. Defaults to 0.5.
<code>label.segment.alpha</code>	Transparency of the line segment. Defaults to the same transparency as the text.
<code>label.min.segment.length</code>	Skip drawing segments shorter than this. Defaults to 0.5. (Default unit is lines).
<code>label.force</code>	Force of repulsion between overlapping text labels. Defaults to 1.
<code>label.max.iter</code>	Maximum number of iterations to try to resolve overlaps. Defaults to 2000.
<code>label.nudge.x</code> , <code>label.nudge.y</code>	Horizontal and vertical adjustments to nudge the starting position of each text label. Defaults to 0.
<code>label.xlim</code> , <code>label.ylim</code>	Limits for the x and y axes. Text labels will be constrained to these limits. By default, text labels are constrained to the entire plot area. Defaults to <code>c(NA, NA)</code> .

<code>label.direction</code>	Character ("both", "x", or "y") – direction in which to adjust position of labels (Default: "y").
<code>package</code>	Name of package from which the palette is desired as string or symbol.
<code>palette</code>	Name of palette as string or symbol.
<code>direction</code>	Either 1 or -1. If -1 the palette will be reversed.
<code>ggtheme</code>	A function, ggplot2 theme name. Default value is <code>ggplot2::theme_bw()</code> . Any of the ggplot2 themes, or themes from extension packages are allowed (e.g., <code>ggthemes::theme_fivethirtyeight()</code> , <code>hrbrthemes::theme_ipsum_ps()</code> , etc.).
<code>ggstatsplot.layer</code>	Logical that decides whether <code>theme_ggstatsplot</code> theme elements are to be displayed along with the selected <code>ggtheme</code> (Default: TRUE).
<code>messages</code>	Decides whether messages references, notes, and warnings are to be displayed (Default: TRUE).
<code>...</code>	Additional arguments to tidying method.

Value

Plot with the regression coefficients' point estimates as dots with confidence interval whiskers.

Author(s)

Indrajeet Patil

References

https://indrajeetpatil.github.io/ggstatsplot/articles/web_only/ggcoefstats.html

Examples

```
# for reproducibility
set.seed(123)

# ----- with model object -----

# model object
mod <- lm(formula = mpg ~ cyl * am, data = mtcars)

# to get a plot
ggstatsplot::ggcoefstats(x = mod, output = "plot")

# to get a tidy dataframe
ggstatsplot::ggcoefstats(x = mod, output = "tidy")

# to get a glance summary
ggstatsplot::ggcoefstats(x = mod, output = "glance")

# to get augmented dataframe
```

```

ggstatsplot::ggcoefstats(x = mod, output = "augment")

# ----- with custom dataframe -----

# creating a dataframe
df <-
  structure(
    list(
      term = structure(
        c(3L, 4L, 1L, 2L, 5L),
        .Label = c(
          "Africa",
          "Americas", "Asia", "Europe", "Oceania"
        ),
        class = "factor"
      ),
      estimate = c(
        0.382047603321706,
        0.780783111514665,
        0.425607573765058,
        0.558365541235078,
        0.956473848429961
      ),
      std.error = c(
        0.0465576338644502,
        0.0330218199731529,
        0.0362834986178494,
        0.0480571500648261,
        0.062215818388157
      ),
      statistic = c(
        8.20590677855356,
        23.6444603038067,
        11.7300588415607,
        11.6187818146078,
        15.3734833553524
      ),
      conf.low = c(
        0.290515146096969,
        0.715841986960399,
        0.354354575031406,
        0.46379116008131,
        0.827446138277154
      ),
      conf.high = c(
        0.473580060546444,
        0.845724236068931,
        0.496860572498711,
        0.652939922388847,
        1.08550155858277
      ),
      p.value = c(
        3.28679518728519e-15,

```

```

      4.04778497135963e-75,
      7.59757330804449e-29,
      5.45155840151592e-26,
      2.99171217913312e-13
    ),
    df.residual = c(
      394L, 358L, 622L,
      298L, 22L
    )
  ),
  row.names = c(NA, -5L),
  class = c(
    "tbl_df",
    "tbl", "data.frame"
  )
)

# plotting the dataframe
ggstatsplot::ggcoefstats(
  x = df,
  statistic = "t",
  meta.analytic.effect = TRUE
)

# ----- getting model summary -----

# model
library(lme4)
lmm1 <- lme4::lmer(
  formula = Reaction ~ Days + (Days | Subject),
  data = sleepstudy
)

# dataframe with model summary
ggstatsplot::ggcoefstats(x = lmm1, output = "glance")

# ----- getting augmented dataframe -----

# setup
set.seed(123)
library(survival)

# fit
cfits <-
  survival::coxph(formula = Surv(time, status) ~ age + sex, data = lung)

# augmented dataframe
ggstatsplot::ggcoefstats(
  x = cfits,
  data = lung,
  output = "augment",
  type.predict = "risk"
)

```


ggcorrmat

*Visualization of a correlalogram (or correlation matrix)***Description**

Visualization of a correlalogram (or correlation matrix)

Usage

```
ggcorrmat(data, cor.vars = NULL, cor.vars.names = NULL,
  output = "plot", matrix.type = "full", method = "square",
  corr.method = "pearson", type = NULL, exact = FALSE,
  continuity = TRUE, beta = 0.1, digits = 2, k = NULL,
  sig.level = 0.05, p.adjust.method = "none", hc.order = FALSE,
  hc.method = "complete", lab = TRUE, package = "RColorBrewer",
  palette = "Dark2", direction = 1, colors = c("#E69F00", "white",
  "#009E73"), outline.color = "black", ggtheme = ggplot2::theme_bw(),
  ggstatsplot.layer = TRUE, title = NULL, subtitle = NULL,
  caption = NULL, caption.default = TRUE, lab.col = "black",
  lab.size = 5, insig = "pch", pch = 4, pch.col = "black",
  pch.cex = 11, tl.cex = 12, tl.col = "black", tl.srt = 45,
  axis.text.x.margin.l = 0, axis.text.x.margin.t = 0,
  axis.text.x.margin.r = 0, axis.text.x.margin.b = 0,
  messages = TRUE)
```

Arguments

<code>data</code>	Dataframe from which variables specified are preferentially to be taken.
<code>cor.vars</code>	List of variables for which the correlation matrix is to be computed and visualized. If <code>NULL</code> (default), all numeric variables from data will be used.
<code>cor.vars.names</code>	Optional list of names to be used for <code>cor.vars</code> . The names should be entered in the same order.
<code>output</code>	Character that decides expected output from this function: "plot" (for visualization matrix) or "correlations" (or "corr" or "r"; for correlation matrix) or "p-values" (or "p.values" or "p"; for a matrix of <i>p</i> -values) or "ci" (for a tibble with confidence intervals for unique correlation pairs; not available for robust correlation) or "n" (or "sample.size" for a tibble with sample sizes for each correlation pair).
<code>matrix.type</code>	Character, "full" (default), "upper" or "lower", display full matrix, lower triangular or upper triangular matrix.
<code>method</code>	Character argument that decides the visualization method of correlation matrix to be used. Allowed values are "square" (default), "circle"
<code>corr.method, type</code>	A character string indicating which correlation coefficient is to be computed ("pearson" (default) or "kendall" or "spearman"). "robust" can also be entered but only if output argument is set to either "correlations" or "p-values".

	The robust correlation used is percentage bend correlation (see ?WRS2::pball). Abbreviations will also work: "p" (for parametric/Pearson's r), "np" (nonparametric/Spearman's ρ), "r" (robust).
exact	A logical indicating whether an exact p -value should be computed. Used for Kendall's τ and Spearman's ρ . For more details, see ?stats::cor.test.
continuity	A logical. If TRUE, a continuity correction is used for Kendall's τ and Spearman's ρ when not computed exactly (Default: TRUE).
beta	A numeric bending constant for robust correlation coefficient (Default: 0.1).
digits, k	Decides the number of decimal digits to be displayed (Default: 2).
sig.level	Significance level (Default: 0.05). If the p -value in p -value matrix is bigger than sig.level, then the corresponding correlation coefficient is regarded as insignificant and flagged as such in the plot. This argument is relevant only when output = "plot".
p.adjust.method	What adjustment for multiple tests should be used? ("holm", "hochberg", "hommel", "bonferroni", "BH", "BY", "fdr", "none"). See stats::p.adjust for details about why to use "holm" rather than "bonferroni". Default is "none". If adjusted p -values are displayed in the visualization of correlation matrix, the adjusted p -values will be used for the upper triangle, while unadjusted p -values will be used for the lower triangle of the matrix.
hc.order	Logical value. If TRUE, correlation matrix will be hc.ordered using hclust function (Default is FALSE).
hc.method	The agglomeration method to be used in hclust (see ?hclust).
lab	Logical value. If TRUE, correlation coefficient values will be displayed in the plot.
package	Name of package from which the palette is desired as string or symbol.
palette	Name of palette as string or symbol.
direction	Either 1 or -1. If -1 the palette will be reversed.
colors	A vector of 3 colors for low, mid, and high correlation values. If set to NULL, manual specification of colors will be turned off and 3 colors from the specified palette from package will be selected.
outline.color	The outline color of square or circle. Default value is "gray".
ggtheme	A function, ggplot2 theme name. Default value is ggplot2::theme_bw(). Any of the ggplot2 themes, or themes from extension packages are allowed (e.g., ggthemes::theme_fivethirtyeight(), hrbrthemes::theme_ipsum_ps(), etc.).
ggstatsplot.layer	Logical that decides whether theme_ggstatsplot theme elements are to be displayed along with the selected ggtheme (Default: TRUE).
title	The text for the plot title.
subtitle	The text for the plot subtitle.
caption	The text for the plot caption. If not specified (if it is NULL, i.e.), a default caption will be shown.

<code>caption.default</code>	Logical decides whether the default caption should be shown.
<code>lab.col</code>	Color to be used for the correlation coefficient labels (applicable only when <code>lab = TRUE</code>).
<code>lab.size</code>	Size to be used for the correlation coefficient labels (applicable only when <code>lab = TRUE</code>).
<code>insig</code>	Character used to show specialized insignificant correlation coefficients ("pch" (default) or "blank"). If "blank", the corresponding glyphs will be removed; if "pch" is used, characters (see <code>?pch</code> for details) will be added on the corresponding glyphs.
<code>pch</code>	Decides the glyphs (read point shapes) to be used for insignificant correlation coefficients (only valid when <code>insig = "pch"</code>). Default value is <code>pch = 4</code> .
<code>pch.col</code> , <code>pch.cex</code>	The color and the cex (size) of pch (only valid when <code>insig = "pch"</code>). Defaults are <code>pch.col = "#F0E442"</code> and <code>pch.cex = 10</code> .
<code>tl.cex</code> , <code>tl.col</code> , <code>tl.srt</code>	The size, the color, and the string rotation of text label (variable names, i.e.).
<code>axis.text.x.margin.t</code> , <code>axis.text.x.margin.r</code> , <code>axis.text.x.margin.b</code> , <code>axis.text.x.margin.l</code>	Margins between x-axis and the variable name texts (t: top, r: right, b: bottom, l:left), especially useful in case the names are slanted, i.e. when the <code>tl.srt</code> is between 45 and 75 (Defaults: 0, 0, 0, 0, resp.).
<code>messages</code>	Decides whether messages references, notes, and warnings are to be displayed (Default: TRUE).

Value

Correlation matrix plot or correlation coefficient matrix or matrix of p-values.

Author(s)

Indrajeet Patil

References

https://indrajeetpatil.github.io/ggstatsplot/articles/web_only/ggcorrmat.html

See Also

[grouped_ggcorrmat](#) [ggscatterstats](#) [grouped_ggscatterstats](#)

Examples

```
# for reproducibility
set.seed(123)

# if `cor.vars` not specified, all numeric variables used
ggstatsplot::ggcorrmat(data = iris)
```

```

# to get the correlalogram
# note that the function will run even if the vector with variable names is
# not of same length as the number of variables
ggstatsplot::ggcorrmat(
  data = ggplot2::msleep,
  cor.vars = sleep_total:bodywt,
  cor.vars.names = c("total sleep", "REM sleep")
) + # further modification using `ggplot2`
  ggplot2::scale_y_discrete(position = "right")

# to get the correlation matrix
ggstatsplot::ggcorrmat(
  data = ggplot2::msleep,
  cor.vars = sleep_total:bodywt,
  output = "r"
)

# setting output = "p-values" (or "p") will return the p-value matrix
ggstatsplot::ggcorrmat(
  data = ggplot2::msleep,
  cor.vars = sleep_total:bodywt,
  corr.method = "r",
  p.adjust.method = "bonferroni",
  output = "p"
)

# setting `output = "ci"` will return the confidence intervals for unique
# correlation pairs
ggstatsplot::ggcorrmat(
  data = ggplot2::msleep,
  cor.vars = sleep_total:bodywt,
  p.adjust.method = "BH",
  output = "ci"
)

# modifying elements of the correlation matrix by changing function defaults
ggstatsplot::ggcorrmat(
  data = datasets::iris,
  cor.vars = c(Sepal.Length, Sepal.Width, Petal.Length, Petal.Width),
  sig.level = 0.01,
  ggtheme = ggplot2::theme_bw(),
  hc.order = TRUE,
  matrix.type = "lower",
  outline.col = "white",
  title = "Dataset: Iris"
)

```

Description

A dot chart with statistical details from one-sample test included in the plot as a subtitle.

Usage

```
ggdotplotstats(data, x, y, xlab = NULL, ylab = NULL, title = NULL,
  subtitle = NULL, caption = NULL, type = "parametric",
  test.value = 0, bf.prior = 0.707, bf.message = FALSE,
  robust.estimator = "onestep", conf.level = 0.95, nboot = 100,
  k = 2, results.subtitle = TRUE, ggtheme = ggplot2::theme_bw(),
  ggstatsplot.layer = TRUE, point.color = "black", point.size = 3,
  point.shape = 16, centrality.para = "mean",
  centrality.color = "blue", centrality.size = 1,
  centrality.linetype = "dashed", centrality.line.labeller = TRUE,
  centrality.k = 2, test.value.line = FALSE,
  test.value.color = "black", test.value.size = 1,
  test.value.linetype = "dashed", test.line.labeller = TRUE,
  test.k = 0, ggplot.component = NULL, messages = TRUE)
```

Arguments

data	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
x	A numeric variable.
y	Label or grouping variable.
xlab	Labels for x and y axis variables. If NULL (default), variable names for x and y will be used.
ylab	Label for y axis variable.
title	The text for the plot title.
subtitle	The text for the plot subtitle. Will work only if results.subtitle = FALSE.
caption	The text for the plot caption.
type	Type of statistic expected ("parametric" or "nonparametric" or "robust" or "bayes").Corresponding abbreviations are also accepted: "p" (for parametric), "np" (nonparametric), "r" (robust), or "bf"resp.
test.value	A number specifying the value of the null hypothesis (Default: 0).
bf.prior	A number between 0.5 and 2 (default 0.707), the prior width to use in calculating Bayes factors.
bf.message	Logical that decides whether to display Bayes Factor in favor of the <i>null</i> hypothesis for parametric test (Default: FALSE).
robust.estimator	If test = "robust" robust estimator to be used ("onestep" (Default), "mom", or "median"). For more, see ?WRS2::onesampb.
conf.level	Scalar between 0 and 1. If unspecified, the defaults return 95% lower and upper confidence intervals (0.95).

nboot	Number of bootstrap samples for computing confidence interval for the effect size (Default: 100).
k	Number of digits after decimal point (should be an integer) (Default: k = 2).
results.subtitle	Decides whether the results of statistical tests are to be displayed as a subtitle (Default: TRUE). If set to FALSE, only the plot will be returned.
ggtheme	A function, ggplot2 theme name. Default value is ggplot2::theme_bw(). Any of the ggplot2 themes, or themes from extension packages are allowed (e.g., ggthemes::theme_fivethirtyeight(), hrbrthemes::theme_ipsum_ps(), etc.).
ggstatsplot.layer	Logical that decides whether theme_ggstatsplot theme elements are to be displayed along with the selected ggtheme (Default: TRUE).
point.color	Character describing color for the point (Default: "black").
point.size	Numeric specifying size for the point (Default: 3).
point.shape	Numeric specifying shape to draw the points (Default: 16 (a dot)).
centrality.param	Decides <i>which</i> measure of central tendency ("mean" or "median") is to be displayed as a vertical line.
centrality.color	Decides color for the vertical line for centrality parameter (Default: "blue").
centrality.size	Decides size for the vertical line for centrality parameter (Default: 1.2).
centrality.linetype	Decides linetype for the vertical line for centrality parameter (Default: "dashed").
centrality.line.labeller	A logical that decides whether line labels should be displayed for the centrality.param line (Default: TRUE).
centrality.k	Integer denoting the number of decimal places expected for centrality parameter label. (Default: 2).
test.value.line	Decides whether test value is to be displayed as a vertical line (Default: FALSE).
test.value.color	Decides color for the vertical line denoting test value (Default: "black").
test.value.size	Decides size for the vertical line for test value (Default: 1.2).
test.value.linetype	Decides linetype for the vertical line for test value (Default: "dashed").
test.line.labeller	A logical that decides whether line labels should be displayed for test.value line (Default: TRUE).
test.k	Integer denoting the number of decimal places expected for test.value label. (Default: 0).

ggplot.component	A ggplot component to be added to the plot prepared by ggstatsplot. This argument is primarily helpful for grouped_ variant of the current function. Default is NULL. The argument should be entered as a function. If the given function has an argument axes.range.restrict and if it has been set to TRUE, the added ggplot component <i>might</i> not work as expected.
messages	Decides whether messages references, notes, and warnings are to be displayed (Default: TRUE).

Author(s)

Indrajeet Patil

Examples

```
# for reproducibility
set.seed(123)

# plot
ggdotplotstats(
  data = ggplot2::mpg,
  x = cty,
  y = manufacturer,
  conf.level = 0.99,
  test.value = 15,
  test.value.line = TRUE,
  test.line.labeller = TRUE,
  test.value.color = "red",
  centrality.para = "median",
  centrality.k = 0,
  title = "Fuel economy data",
  xlab = "city miles per gallon",
  bf.message = TRUE,
  caption = substitute(
    paste(italic("Source"), ": EPA dataset on http://fuelconomy.gov")
  )
)
```

gghistostats

Histogram for distribution of a numeric variable

Description

Histogram with statistical details from one-sample test included in the plot as a subtitle.

Usage

```
gghistostats(data = NULL, x, binwidth = NULL, bar.measure = "count",
  xlab = NULL, title = NULL, subtitle = NULL, caption = NULL,
  type = "parametric", test.value = 0, bf.prior = 0.707,
  bf.message = FALSE, robust.estimator = "onestep",
  effsize.type = "g", effsize.noncentral = TRUE, conf.level = 0.95,
  nboot = 100, k = 2, ggtheme = ggplot2::theme_bw(),
  ggstatsplot.layer = TRUE, fill.gradient = FALSE,
  low.color = "#0072B2", high.color = "#D55E00", bar.fill = "grey50",
  results.subtitle = TRUE, centrality.para = "mean",
  centrality.color = "blue", centrality.size = 1,
  centrality.linetype = "dashed", centrality.line.labeller = TRUE,
  centrality.k = 2, test.value.line = FALSE,
  test.value.color = "black", test.value.size = 1,
  test.value.linetype = "dashed", test.line.labeller = TRUE,
  test.k = 0, normal.curve = FALSE, normal.curve.color = "black",
  normal.curve.linetype = "solid", normal.curve.size = 1,
  ggplot.component = NULL, messages = TRUE)
```

Arguments

data	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
x	A numeric variable.
binwidth	The width of the histogram bins. Can be specified as a numeric value, or a function that calculates width from x. The default is to use the $\max(x) - \min(x) / \sqrt{N}$. You should always check this value and explore multiple widths to find the best to illustrate the stories in your data.
bar.measure	Character describing what value needs to be represented as height in the bar chart. This can either be "count", which shows number of points in bin, or "density", which density of points in bin, scaled to integrate to 1, or "proportion", which shows relative frequencies of observations in each bin, or "mix", which shows <i>both</i> count and proportion in the same plot.
xlab	Labels for x and y axis variables. If NULL (default), variable names for x and y will be used.
title	The text for the plot title.
subtitle	The text for the plot subtitle. Will work only if results.subtitle = FALSE.
caption	The text for the plot caption.
type	Type of statistic expected ("parametric" or "nonparametric" or "robust" or "bayes"). Corresponding abbreviations are also accepted: "p" (for parametric), "np" (nonparametric), "r" (robust), or "bf" resp.
test.value	A number specifying the value of the null hypothesis (Default: 0).
bf.prior	A number between 0.5 and 2 (default 0.707), the prior width to use in calculating Bayes factors.

bf.message	Logical that decides whether to display Bayes Factor in favor of the <i>null</i> hypothesis for parametric test (Default: FALSE).
robust.estimator	If test = "robust" robust estimator to be used ("onestep" (Default), "mom", or "median"). For more, see ?WRS2::onesampb.
effsize.type	Type of effect size needed for <i>parametric</i> tests. The argument can be "biased" ("d" for Cohen's <i>d</i>) or "unbiased" ("g" Hedge's <i>g</i> for t-test). The default is
effsize.noncentral	Logical indicating whether to use non-central <i>t</i> -distributions for computing the confidence interval for Cohen's <i>d</i> or Hedge's <i>g</i> (Default: TRUE).
conf.level	Scalar between 0 and 1. If unspecified, the defaults return 95% lower and upper confidence intervals (0.95).
nboot	Number of bootstrap samples for computing confidence interval for the effect size (Default: 100).
k	Number of digits after decimal point (should be an integer) (Default: k = 2).
ggtheme	A function, ggplot2 theme name. Default value is ggplot2::theme_bw(). Any of the ggplot2 themes, or themes from extension packages are allowed (e.g., ggthemes::theme_fivethirtyeight(), hrbrthemes::theme_ipsum_ps(), etc.).
ggstatsplot.layer	Logical that decides whether theme_ggstatsplot theme elements are to be displayed along with the selected ggtheme (Default: TRUE).
fill.gradient	Logical decides whether color fill gradient is to be displayed (Default: FALSE). If FALSE, the legend and the color gradient will also be removed. The default is set to FALSE because the gradient provides redundant information in light of y-axis labels.
low.color, high.color	Colors for low and high ends of the gradient. Defaults are colorblind-friendly.
bar.fill	If fill.gradient = FALSE, then bar.fill decides which color will uniformly fill all the bars in the histogram (Default: "grey50").
results.subtitle	Decides whether the results of statistical tests are to be displayed as a subtitle (Default: TRUE). If set to FALSE, only the plot will be returned.
centrality.param	Decides <i>which</i> measure of central tendency ("mean" or "median") is to be displayed as a vertical line.
centrality.color	Decides color for the vertical line for centrality parameter (Default: "blue").
centrality.size	Decides size for the vertical line for centrality parameter (Default: 1.2).
centrality.linetype	Decides linetype for the vertical line for centrality parameter (Default: "dashed").
centrality.line.labeller	A logical that decides whether line labels should be displayed for the centrality.param line (Default: TRUE).

<code>centrality.k</code>	Integer denoting the number of decimal places expected for centrality parameter label. (Default: 2).
<code>test.value.line</code>	Decides whether test value is to be displayed as a vertical line (Default: FALSE).
<code>test.value.color</code>	Decides color for the vertical line denoting test value (Default: "black").
<code>test.value.size</code>	Decides size for the vertical line for test value (Default: 1.2).
<code>test.value.linetype</code>	Decides linetype for the vertical line for test value (Default: "dashed").
<code>test.line.labeller</code>	A logical that decides whether line labels should be displayed for test.value line (Default: TRUE).
<code>test.k</code>	Integer denoting the number of decimal places expected for test.value label. (Default: 0).
<code>normal.curve</code>	Logical decides whether to super-impose a normal curve using <code>stats::dnorm(mean(x), sd(x))</code> . Default is FALSE.
<code>normal.curve.color</code> , <code>normal.curve.linetype</code> , <code>normal.curve.size</code>	If <code>normal.curve = TRUE</code> , then these arguments can be used to modify color (Default: "black"), size (default: 1.0), linetype (default: "solid").
<code>ggplot.component</code>	A ggplot component to be added to the plot prepared by <code>ggstatsplot</code> . This argument is primarily helpful for <code>grouped_</code> variant of the current function. Default is NULL. The argument should be entered as a function. If the given function has an argument <code>axes.range.restrict</code> and if it has been set to TRUE, the added ggplot component <i>might</i> not work as expected.
<code>messages</code>	Decides whether messages references, notes, and warnings are to be displayed (Default: TRUE).

Author(s)

Indrajeet Patil

References

https://indrajeetpatil.github.io/ggstatsplot/articles/web_only/gghistostats.html

See Also

[grouped_gghistostats](#)

Examples

```
# most basic function call with the defaults
# this is the only function where data argument can be `NULL`
ggstatsplot::gghistostats(
  x = ToothGrowth$len,
```

```

    xlab = "Tooth length",
    centrality.param = "median"
  )

  # a detailed function call
  ggstatsplot::gghistostats(
    data = iris,
    x = Sepal.Length,
    bar.measure = "mix",
    type = "p",
    bf.message = TRUE,
    caption = substitute(paste(italic("Note"), ": Iris dataset by Fisher.")),
    bf.prior = 0.8,
    test.value = 3,
    test.value.line = TRUE,
    binwidth = 0.10,
    bar.fill = "grey50"
  )

```

ggpiestats

Pie charts with statistical tests

Description

Pie charts for categorical data with statistical details included in the plot as a subtitle.

Usage

```

ggpiestats(data, main, condition = NULL, counts = NULL, ratio = NULL,
  paired = FALSE, results.subtitle = TRUE, factor.levels = NULL,
  stat.title = NULL, sample.size.label = TRUE,
  label.separator = "\n", label.text.size = 4,
  label.fill.color = "white", label.fill.alpha = 1,
  bf.message = FALSE, sampling.plan = "indepMulti",
  fixed.margin = "rows", prior.concentration = 1, title = NULL,
  subtitle = NULL, caption = NULL, conf.level = 0.95, nboot = 100,
  simulate.p.value = FALSE, B = 2000, legend.title = NULL,
  facet.wrap.name = NULL, k = 2, perc.k = 0,
  slice.label = "percentage", facet.proptest = TRUE,
  ggtheme = ggplot2::theme_bw(), ggstatsplot.layer = TRUE,
  package = "RColorBrewer", palette = "Dark2", direction = 1,
  ggplot.component = NULL, messages = TRUE)

```

Arguments

data	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
main	The variable to use as the rows in the contingency table.

<code>condition</code>	The variable to use as the columns in the contingency table.
<code>counts</code>	A string naming a variable in data containing counts, or NULL if each row represents a single observation (Default).
<code>ratio</code>	A vector of numbers: the expected proportions for the proportion test. Default is NULL, which means if there are two levels <code>ratio = c(1,1)</code> , etc.
<code>paired</code>	Logical indicating whether data came from a within-subjects design study (Default: FALSE). If TRUE, McNemar test subtitle will be returned. If FALSE, Pearson's chi-square test will be returned.
<code>results.subtitle</code>	Decides whether the results of statistical tests are to be displayed as a subtitle (Default: TRUE). If set to FALSE, only the plot will be returned.
<code>factor.levels</code>	A character vector with labels for factor levels of main variable.
<code>stat.title</code>	Title for the effect being investigated with the chi-square test. The default is NULL, i.e. no title will be added to describe the effect being shown. An example of a <code>stat.title</code> argument will be something like "main x condition" or "interaction".
<code>sample.size.label</code>	Logical that decides whether sample size information should be displayed for each level of the grouping variable <code>condition</code> (Default: TRUE).
<code>label.separator</code>	If "both" counts and proportion information is to be displayed in a label, this argument decides whether these two pieces of information are going to be on the same line (" ") or on separate lines ("\n").
<code>label.text.size</code>	Numeric that decides text size for slice/bar labels (Default: 4).
<code>label.fill.color</code>	Character that specifies fill color for slice/bar labels (Default: white).
<code>label.fill.alpha</code>	Numeric that specifies fill color transparency or "alpha" for slice/bar labels (Default: 1 range 0 to 1).
<code>bf.message</code>	Logical that decides whether to display a caption with results from bayes factor test in favor of the null hypothesis (default: FALSE).
<code>sampling.plan</code>	Character describing the sampling plan. Possible options are "indepMulti" (independent multinomial; default), "poisson", "jointMulti" (joint multinomial), "hypergeom" (hypergeometric). For more, see <code>?BayesFactor::contingencyTableBF()</code> .
<code>fixed.margin</code>	For the independent multinomial sampling plan, which margin is fixed ("rows" or "cols"). Defaults to "rows".
<code>prior.concentration</code>	Specifies the prior concentration parameter, set to 1 by default. It indexes the expected deviation from the null hypothesis under the alternative, and corresponds to Gunel and Dickey's (1974) "a" parameter.
<code>title</code>	The text for the plot title.
<code>subtitle</code>	The text for the plot subtitle. Will work only if <code>results.subtitle = FALSE</code> .
<code>caption</code>	The text for the plot caption.

<code>conf.level</code>	Scalar between 0 and 1. If unspecified, the defaults return 95% lower and upper confidence intervals (0.95).
<code>nboot</code>	Number of bootstrap samples for computing confidence interval for the effect size (Default: 100).
<code>simulate.p.value</code>	a logical indicating whether to compute p-values by Monte Carlo simulation.
<code>B</code>	an integer specifying the number of replicates used in the Monte Carlo test.
<code>legend.title</code>	Title text for the legend.
<code>facet.wrap.name</code>	The text for the facet_wrap variable label.
<code>k</code>	Number of digits after decimal point (should be an integer) (Default: <code>k = 2</code>).
<code>perc.k</code>	Numeric that decides number of decimal places for percentage labels (Default: 0).
<code>slice.label</code>	Character decides what information needs to be displayed on the label in each pie slice. Possible options are "percentage" (default), "counts", "both".
<code>facet.proptest</code>	Decides whether proportion test for main variable is to be carried out for each level of condition (Default: TRUE).
<code>ggtheme</code>	A function, ggplot2 theme name. Default value is <code>ggplot2::theme_bw()</code> . Any of the ggplot2 themes, or themes from extension packages are allowed (e.g., <code>ggthemes::theme_fivethirtyeight()</code> , <code>hrbrthemes::theme_ipsum_ps()</code> , etc.).
<code>ggstatsplot.layer</code>	Logical that decides whether theme_ggstatsplot theme elements are to be displayed along with the selected ggtheme (Default: TRUE).
<code>package</code>	Name of package from which the palette is desired as string or symbol.
<code>palette</code>	If a character string (e.g., "Set1"), will use that named palette. If a number, will index into the list of palettes of appropriate type. Default palette is "Dark2".
<code>direction</code>	Either 1 or -1. If -1 the palette will be reversed.
<code>ggplot.component</code>	A ggplot component to be added to the plot prepared by ggstatsplot. This argument is primarily helpful for grouped_ variant of the current function. Default is NULL. The argument should be entered as a function. If the given function has an argument <code>axes.range.restrict</code> and if it has been set to TRUE, the added ggplot component <i>might</i> not work as expected.
<code>messages</code>	Decides whether messages references, notes, and warnings are to be displayed (Default: TRUE).

Value

Unlike a number of statistical softwares, ggstatsplot doesn't provide the option for Yates' correction for the Pearson's chi-squared statistic. This is due to compelling amount of Monte-Carlo simulation research which suggests that the Yates' correction is overly conservative, even in small sample sizes. As such it is recommended that it should not ever be applied in practice (Camilli & Hopkins, 1978, 1979; Feinberg, 1980; Larntz, 1978; Thompson, 1988).

Author(s)

Indrajeet Patil

References

https://indrajeetpatil.github.io/ggstatsplot/articles/web_only/ggpiestats.html

Examples

```
# for reproducibility
set.seed(123)

# simple function call with the defaults (without condition)
ggstatsplot::ggpiestats(
  data = ggplot2::msleep,
  main = vore,
  perc.k = 1,
  k = 2
)

# simple function call with the defaults (with condition)
ggstatsplot::ggpiestats(
  data = datasets::mtcars,
  main = vs,
  condition = cyl,
  bf.message = TRUE,
  nboot = 10,
  factor.levels = c("0 = V-shaped", "1 = straight"),
  legend.title = "Engine"
)

# simple function call with the defaults (without condition; with count data)
library(jmv)

ggstatsplot::ggpiestats(
  data = as.data.frame(HairEyeColor),
  main = Eye,
  counts = Freq
)
```

ggplot_converter

Transform object of any other class to an object of class ggplot.

Description

Transform object of any other class to an object of class ggplot.

Usage

```
ggplot_converter(plot)
```

Arguments

`plot` A plot that needs to be converted to object of class `ggplot`.

Examples

```
library(ggplot2)

# creating a plot that is not of class `ggplot`
p <- ggExtra::ggMarginal(ggplot(mtcars, aes(wt, mpg)) + geom_point())

# checking class of object
class(p)

# checking class of converted plot
p_converted <- ggstatsplot::ggplot_converter(p)
class(p_converted)
```

ggscatterstats

Scatterplot with marginal distributions

Description

Scatterplots from `ggplot2` combined with marginal histograms/boxplots/density plots with statistical details added as a subtitle.

Usage

```
ggscatterstats(data, x, y, type = "pearson", conf.level = 0.95,
  bf.prior = 0.707, bf.message = FALSE, label.var = NULL,
  label.expression = NULL, xlab = NULL, ylab = NULL, method = "lm",
  method.args = list(), formula = y ~ x, point.color = "black",
  point.size = 3, point.alpha = 0.4, point.width.jitter = 0,
  point.height.jitter = 0, line.size = 1.5, line.color = "blue",
  marginal = TRUE, marginal.type = "histogram", marginal.size = 5,
  margins = c("both", "x", "y"), package = "wesanderson",
  palette = "Royal1", direction = 1, xfill = "#009E73",
  yfill = "#D55E00", xalpha = 1, yalpha = 1, xsize = 0.7,
  ysize = 0.7, centrality.param = NULL, results.subtitle = TRUE,
  title = NULL, subtitle = NULL, caption = NULL, nboot = 100,
  beta = 0.1, k = 2, axes.range.restrict = FALSE,
  ggtheme = ggplot2::theme_bw(), ggstatsplot.layer = TRUE,
  ggplot.component = NULL, messages = TRUE)
```

Arguments

data	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
x	The column in data containing the explanatory variable to be plotted on the x axis. Can be entered either as a character string (e.g., "x") or as a bare expression (e.g, x).
y	The column in data containing the response (outcome) variable to be plotted on the y axis. Can be entered either as a character string (e.g., "y") or as a bare expression (e.g, y).
type	Type of association between paired samples required ("parametric": Pearson's product moment correlation coefficient" or "nonparametric": Spearman's rho" or "robust": percentage bend correlation coefficient" or "bayes": Bayes Factor for Pearson's r "). Corresponding abbreviations are also accepted: "p" (for parametric/pearson's), "np" (nonparametric/spearman), "r" (robust), "bf" (for bayes factor), resp.
conf.level	Scalar between 0 and 1. If unspecified, the defaults return 95% lower and upper confidence intervals (0.95).
bf.prior	A number between 0.5 and 2 (default 0.707), the prior width to use in calculating Bayes factors.
bf.message	Logical that decides whether to display Bayes Factor in favor of the <i>null</i> hypothesis for parametric test (Default: FALSE).
label.var	Variable to use for points labels. Can be entered either as a character string (e.g., "var1") or as a bare expression (e.g, var1).
label.expression	An expression evaluating to a logical vector that determines the subset of data points to label. This argument can be entered either as a character string (e.g., "y < 4 & z < 20") or as a bare expression (e.g., y < 4 & z < 20).
xlab	Labels for x and y axis variables. If NULL (default), variable names for x and y will be used.
ylab	Labels for x and y axis variables. If NULL (default), variable names for x and y will be used.
method	<p>Smoothing method (function) to use, accepts either a character vector, e.g. "auto", "lm", "glm", "gam", "loess" or a function, e.g. MASS::rlm or mgcv::gam, base::lm, or base::loess.</p> <p>For method = "auto" the smoothing method is chosen based on the size of the largest group (across all panels). <code>loess()</code> is used for less than 1,000 observations; otherwise <code>mgcv::gam()</code> is used with formula = $y \sim s(x, bs = "cs")$. Somewhat anecdotally, loess gives a better appearance, but is $O(N^2)$ in memory, so does not work for larger datasets.</p> <p>If you have fewer than 1,000 observations but want to use the same <code>gam()</code> model that method = "auto" would use, then set method = "gam", formula = $y \sim s(x, bs = "cs")$.</p>
method.args	List of additional arguments passed on to the modelling function defined by method.
formula	Formula to use in smoothing function, eg. $y \sim x$, $y \sim \text{poly}(x, 2)$, $y \sim \log(x)$

<code>point.color</code> , <code>point.size</code> , <code>point.alpha</code>	Aesthetics specifying geom point (defaults: <code>point.color = "black"</code> , <code>point.size = 3</code> , <code>point.alpha =</code>
<code>point.width.jitter</code> , <code>point.height.jitter</code>	Degree of jitter in x and y direction, respectively. Defaults to 0 (0 data).
<code>line.size</code>	Size for the regression line.
<code>line.color</code>	color for the regression line.
<code>marginal</code>	Decides whether <code>ggExtra::ggMarginal()</code> plots will be displayed; the default is TRUE.
<code>marginal.type</code>	Type of marginal distribution to be plotted on the axes ("histogram", "boxplot", "density", "violin", "densigram").
<code>marginal.size</code>	Integer describing the relative size of the marginal plots compared to the main plot. A size of 5 means that the main plot is 5x wider and 5x taller than the marginal plots.
<code>margins</code>	Character describing along which margins to show the plots. Any of the following arguments are accepted: "both", "x", "y".
<code>package</code>	Name of package from which the palette is desired as string or symbol.
<code>palette</code>	Name of palette as string or symbol.
<code>direction</code>	Either 1 or -1. If -1 the palette will be reversed.
<code>xfill</code> , <code>yfill</code>	Character describing color fill for x and y axes marginal distributions (default: "#009E73" (for x) and "#D55E00" (for y)). If set to NULL, manual specification of colors will be turned off and 2 colors from the specified palette from package will be selected.
<code>xalpha</code> , <code>yalpha</code>	Numeric deciding transparency levels for the marginal distributions. Any numbers from 0 (transparent) to 1 (opaque). The default is 1 for both axes.
<code>xsize</code> , <code>ysize</code>	Size for the marginal distribution boundaries (Default: 0.7).
<code>centrality.param</code>	Decides <i>which</i> measure of central tendency ("mean" or "median") is to be displayed as vertical (for x) and horizontal (for y) lines.
<code>results.subtitle</code>	Decides whether the results of statistical tests are to be displayed as a subtitle (Default: TRUE). If set to FALSE, only the plot will be returned.
<code>title</code>	The text for the plot title.
<code>subtitle</code>	The text for the plot subtitle. Will work only if <code>results.subtitle = FALSE</code> .
<code>caption</code>	The text for the plot caption.
<code>nboot</code>	Number of bootstrap samples for computing confidence interval for the effect size (Default: 100).
<code>beta</code>	bending constant (Default: 0.1). For more, see <code>?WRS2::pbcor</code> .
<code>k</code>	Number of digits after decimal point (should be an integer) (Default: <code>k = 2</code>).
<code>axes.range.restrict</code>	Logical that decides whether to restrict the axes values ranges to min and max values of the axes variables (Default: FALSE), only relevant for functions where axes variables are of numeric type.

ggtheme	A function, ggplot2 theme name. Default value is <code>ggplot2::theme_bw()</code> . Any of the ggplot2 themes, or themes from extension packages are allowed (e.g., <code>ggthemes::theme_fivethirtyeight()</code> , <code>hrbrthemes::theme_ipsum_ps()</code> , etc.).
ggstatsplot.layer	Logical that decides whether <code>theme_ggstatsplot</code> theme elements are to be displayed along with the selected <code>ggtheme</code> (Default: <code>TRUE</code>).
ggplot.component	A ggplot component to be added to the plot prepared by <code>ggstatsplot</code> . This argument is primarily helpful for <code>grouped_</code> variant of the current function. Default is <code>NULL</code> . The argument should be entered as a function. If the given function has an argument <code>axes.range.restrict</code> and if it has been set to <code>TRUE</code> , the added ggplot component <i>might</i> not work as expected.
messages	Decides whether messages references, notes, and warnings are to be displayed (Default: <code>TRUE</code>).

Note

- `marginal.type = "densigram"` will work only with the development version of `ggExtra` that you can download from GitHub: `devtools::install_github("daattali/ggExtra")`
- The plot uses `ggrepel::geom_label_repel` to attempt to keep labels from over-lapping to the largest degree possible. As a consequence plot times will slow down massively (and the plot file will grow in size) if you have a lot of labels that overlap.

Author(s)

Indrajeet Patil, Chuck Powell

References

https://indrajeetpatil.github.io/ggstatsplot/articles/web_only/ggscatterstats.html

See Also

[grouped_ggscatterstats](#), [ggcorrmat](#), [grouped_ggcorrmat](#)

Examples

```
# to get reproducible results from bootstrapping
set.seed(123)

# creating dataframe
mtcars_new <- mtcars %>%
  tibble::rownames_to_column(., var = "car") %>%
  tibble::as_tibble(x = .)

# simple function call with the defaults
ggstatsplot::ggscatterstats(
  data = mtcars_new,
```

```

x = wt,
y = mpg,
type = "np",
label.var = car,
label.expression = wt < 4 & mpg < 20,
axes.range.restrict = TRUE,
centrality.para = "median",
xfill = NULL
)

```

grouped_ggbarstats *Grouped bar (column) charts with statistical tests*

Description

Helper function for `ggstatsplot::ggbarstats` to apply this function across multiple levels of a given factor and combining the resulting plots using `ggstatsplot::combine_plots`.

Usage

```

grouped_ggbarstats(data, main, condition, counts = NULL, grouping.var,
  title.prefix = NULL, ratio = NULL, paired = FALSE,
  results.subtitle = TRUE, labels.legend = NULL, stat.title = NULL,
  sample.size.label = TRUE, label.separator = " ",
  label.text.size = 4, label.fill.color = "white",
  label.fill.alpha = 1, bar.outline.color = "black",
  bf.message = FALSE, sampling.plan = "jointMulti",
  fixed.margin = "rows", prior.concentration = 1, subtitle = NULL,
  caption = NULL, legend.position = "right",
  x.axis.orientation = NULL, conf.level = 0.95, nboot = 100,
  simulate.p.value = FALSE, B = 2000, legend.title = NULL,
  xlab = NULL, ylab = "Percent", k = 2, perc.k = 0,
  data.label = "percentage", bar.proptest = TRUE,
  ggtheme = ggplot2::theme_bw(), ggstatsplot.layer = TRUE,
  package = "RColorBrewer", palette = "Dark2", direction = 1,
  ggplot.component = NULL, messages = TRUE, ...)

```

Arguments

<code>data</code>	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
<code>main</code>	The variable to use as the rows in the contingency table.
<code>condition</code>	The variable to use as the columns in the contingency table.
<code>counts</code>	A string naming a variable in data containing counts, or NULL if each row represents a single observation (Default).
<code>grouping.var</code>	A single grouping variable (can be entered either as a bare name <code>x</code> or as a string <code>"x"</code>).

<code>title.prefix</code>	Character string specifying the prefix text for the fixed plot title (name of each factor level) (Default: NULL). If NULL, the variable name entered for <code>grouping.var</code> will be used.
<code>ratio</code>	A vector of numbers: the expected proportions for the proportion test. Default is NULL, which means if there are two levels <code>ratio = c(1,1)</code> , etc.
<code>paired</code>	Logical indicating whether data came from a within-subjects design study (Default: FALSE). If TRUE, McNemar test subtitle will be returned. If FALSE, Pearson's chi-square test will be returned.
<code>results.subtitle</code>	Decides whether the results of statistical tests are to be displayed as a subtitle (Default: TRUE). If set to FALSE, only the plot will be returned.
<code>labels.legend</code>	A character vector with custom labels for levels of the main variable displayed in the legend.
<code>stat.title</code>	Title for the effect being investigated with the chi-square test. The default is NULL, i.e. no title will be added to describe the effect being shown. An example of a <code>stat.title</code> argument will be something like "main x condition" or "interaction".
<code>sample.size.label</code>	Logical that decides whether sample size information should be displayed for each level of the grouping variable condition (Default: TRUE).
<code>label.separator</code>	If "both" counts and proportion information is to be displayed in a label, this argument decides whether these two pieces of information are going to be on the same line (" ") or on separate lines ("\n").
<code>label.text.size</code>	Numeric that decides text size for slice/bar labels (Default: 4).
<code>label.fill.color</code>	Character that specifies fill color for slice/bar labels (Default: white).
<code>label.fill.alpha</code>	Numeric that specifies fill color transparency or "alpha" for slice/bar labels (Default: 1 range 0 to 1).
<code>bar.outline.color</code>	Character specifying color for bars (default: "black").
<code>bf.message</code>	Logical that decides whether to display a caption with results from bayes factor test in favor of the null hypothesis (default: FALSE).
<code>sampling.plan</code>	Character describing the sampling plan. Possible options are "indepMulti" (independent multinomial; default), "poisson", "jointMulti" (joint multinomial), "hypergeom" (hypergeometric). For more, see <code>?BayesFactor::contingencyTableBF()</code> .
<code>fixed.margin</code>	For the independent multinomial sampling plan, which margin is fixed ("rows" or "cols"). Defaults to "rows".
<code>prior.concentration</code>	Specifies the prior concentration parameter, set to 1 by default. It indexes the expected deviation from the null hypothesis under the alternative, and corresponds to Gunel and Dickey's (1974) "a" parameter.
<code>subtitle</code>	The text for the plot subtitle. Will work only if <code>results.subtitle = FALSE</code> .

caption	The text for the plot caption.
legend.position	The position of the legend "none", "left", "right", "bottom", "top" (Default: "right").
x.axis.orientation	The orientation of the x axis labels one of "slant" or "vertical" to change from the default horizontal orientation (Default: NULL which is horizontal).
conf.level	Scalar between 0 and 1. If unspecified, the defaults return 95% lower and upper confidence intervals (0.95).
nboot	Number of bootstrap samples for computing confidence interval for the effect size (Default: 100).
simulate.p.value	a logical indicating whether to compute p-values by Monte Carlo simulation.
B	an integer specifying the number of replicates used in the Monte Carlo test.
legend.title	Title text for the legend.
xlab	Custom text for the x axis label (Default: NULL, which will cause the x axis label to be the main variable).
ylab	Custom text for the y axis label (Default: "percent").
k	Number of digits after decimal point (should be an integer) (Default: k = 2).
perc.k	Numeric that decides number of decimal places for percentage labels (Default: 0).
data.label	Character decides what information needs to be displayed on the label in each pie slice. Possible options are "percentage" (default), "counts", "both".
bar.proptest	Decides whether proportion test for main variable is to be carried out for each level of condition (Default: TRUE).
ggtheme	A function, ggplot2 theme name. Default value is ggplot2::theme_bw(). Any of the ggplot2 themes, or themes from extension packages are allowed (e.g., ggthemes::theme_fivethirtyeight(), hrbrthemes::theme_ipsum_ps(), etc.).
ggstatsplot.layer	Logical that decides whether theme_ggstatsplot theme elements are to be displayed along with the selected ggtheme (Default: TRUE).
package	Name of package from which the palette is desired as string or symbol.
palette	If a character string (e.g., "Set1"), will use that named palette. If a number, will index into the list of palettes of appropriate type. Default palette is "Dark2".
direction	Either 1 or -1. If -1 the palette will be reversed.
ggplot.component	A ggplot component to be added to the plot prepared by ggstatsplot. This argument is primarily helpful for grouped_ variant of the current function. Default is NULL. The argument should be entered as a function. If the given function has an argument axes.range.restrict and if it has been set to TRUE, the added ggplot component <i>might</i> not work as expected.

messages	Decides whether messages references, notes, and warnings are to be displayed (Default: TRUE).
...	Arguments passed on to <code>combine_plots</code>
	title.text String or plotmath expression to be drawn as title for the <i>combined plot</i> .
	title.color Text color for title.
	title.size Point size of title text.
	title.vjust Vertical justification for title. Default = 0.5 (centered on y). 0 = baseline at y, 1 = ascender at y.
	title.hjust Horizontal justification for title. Default = 0.5 (centered on x). 0 = flush-left at x, 1 = flush-right.
	title.fontface The font face ("plain", "bold" (default), "italic", "bold.italic") for title.
	caption.text String or plotmath expression to be drawn as the caption for the <i>combined plot</i> .
	caption.color Text color for caption.
	caption.size Point size of title text.
	caption.vjust Vertical justification for caption. Default = 0.5 (centered on y). 0 = baseline at y, 1 = ascender at y.
	caption.hjust Horizontal justification for caption. Default = 0.5 (centered on x). 0 = flush-left at x, 1 = flush-right.
	caption.fontface The font face ("plain" (default), "bold", "italic", "bold.italic") for caption.
	sub.text The label with which the <i>combined plot</i> should be annotated. Can be a plotmath expression.
	sub.color Text color for annotation label (Default: "black").
	sub.size Point size of annotation text (Default: 12).
	sub.x The x position of annotation label (Default: 0.5).
	sub.y The y position of annotation label (Default: 0.5).
	sub.hjust Horizontal justification for annotation label (Default: 0.5).
	sub.vjust Vertical justification for annotation label (Default: 0.5).
	sub.vpadding Vertical padding. The total vertical space added to the label, given in grid units. By default, this is added equally above and below the label. However, by changing the y and vjust parameters, this can be changed (Default: <code>grid::unit(1, "lines")</code>).
	sub.fontface The font face ("plain" (default), "bold", "italic", "bold.italic") for the annotation label.
	sub.angle Angle at which annotation label is to be drawn (Default: 0).
	sub.lineheight Line height of annotation label.
	title.caption.rel.heights Numerical vector of relative columns heights while combining (title, plot, caption).
	title.rel.heights Numerical vector of relative columns heights while combining (title, plot).
	caption.rel.heights Numerical vector of relative columns heights while combining (plot, caption).

Value

Unlike a number of statistical softwares, ggstatsplot doesn't provide the option for Yates' correction for the Pearson's chi-squared statistic. This is due to compelling amount of Monte-Carlo simulation research which suggests that the Yates' correction is overly conservative, even in small sample sizes. As such it is recommended that it should not ever be applied in practice (Camilli & Hopkins, 1978, 1979; Feinberg, 1980; Larntz, 1978; Thompson, 1988).

Author(s)

Indrajeet Patil, Chuck Powell

See Also

[ggbarstats](#)

Examples

```
## Not run:
# with condition and with count data
library(jmv)

ggstatsplot::grouped_ggbarstats(
  data = as.data.frame(HairEyeColor),
  main = Hair,
  condition = Eye,
  counts = Freq,
  grouping.var = Sex
)

# the following will take slightly more amount of time
# for reproducibility
set.seed(123)

# let's create a smaller dataframe
diamonds_short <- ggplot2::diamonds %>%
  dplyr::filter(.data = ., cut %in% c("Very Good", "Ideal")) %>%
  dplyr::filter(.data = ., clarity %in% c("SI1", "SI2", "VS1", "VS2")) %>%
  dplyr::sample_frac(tbl = ., size = 0.05)

# plot
ggstatsplot::grouped_ggbarstats(
  data = diamonds_short,
  main = color,
  condition = clarity,
  grouping.var = cut,
  bf.message = TRUE,
  sampling.plan = "poisson",
  title.prefix = "Quality",
  data.label = "both",
  messages = FALSE,
```

```

    perc.k = 1,
    nrow = 2
  )

## End(Not run)

```

grouped_ggbetweenstats

Violin plots for group or condition comparisons in between-subjects designs repeated across all levels of a grouping variable.

Description

A combined plot of comparison plot created for levels of a grouping variable.

Usage

```

grouped_ggbetweenstats(data, x, y, grouping.var, title.prefix = NULL,
  plot.type = "boxviolin", type = "parametric",
  pairwise.comparisons = FALSE, pairwise.annotation = "asterisk",
  pairwise.display = "significant", p.adjust.method = "holm",
  effsize.type = "unbiased", partial = TRUE,
  effsize.noncentral = TRUE, bf.prior = 0.707, bf.message = FALSE,
  results.subtitle = TRUE, xlab = NULL, ylab = NULL,
  subtitle = NULL, caption = NULL, sample.size.label = TRUE, k = 2,
  var.equal = FALSE, conf.level = 0.95, nboot = 100, tr = 0.1,
  axes.range.restrict = FALSE, mean.label.size = 3,
  mean.label.fontface = "bold", mean.label.color = "black",
  notch = FALSE, notchwidth = 0.5, linetype = "solid",
  outlier.tagging = FALSE, outlier.label = NULL,
  outlier.label.color = "black", outlier.color = "black",
  outlier.shape = 19, outlier.coef = 1.5, mean.plotting = TRUE,
  mean.ci = FALSE, mean.size = 5, mean.color = "darkred",
  point.jitter.width = NULL, point.jitter.height = 0,
  point.dodge.width = 0.6, ggtheme = ggplot2::theme_bw(),
  ggstatsplot.layer = TRUE, package = "RColorBrewer",
  palette = "Dark2", direction = 1, ggplot.component = NULL,
  messages = TRUE, ...)

```

Arguments

data	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
x	The grouping variable from the dataframe data.
y	The response (a.k.a. outcome or dependent) variable from the dataframe data.
grouping.var	A single grouping variable (can be entered either as a bare name x or as a string "x").

title.prefix	Character string specifying the prefix text for the fixed plot title (name of each factor level) (Default: NULL). If NULL, the variable name entered for grouping.var will be used.
plot.type	Character describing the <i>type</i> of plot. Currently supported plots are "box" (for pure boxplots), "violin" (for pure violin plots), and "boxviolin" (for a combination of box and violin plots; default).
type	Type of statistic expected ("parametric" or "nonparametric" or "robust" or "bayes").Corresponding abbreviations are also accepted: "p" (for parametric), "np" (nonparametric), "r" (robust), or "bf"resp.
pairwise.comparisons	Logical that decides whether pairwise comparisons are to be displayed. Only significant comparisons will be shown by default. (default: FALSE). To change this behavior, select appropriate option with pairwise.display argument.
pairwise.annotation	Character that decides the annotations to use for pairwise comparisons. Either "p.value" or "asterisk" (default).
pairwise.display	Decides which pairwise comparisons to display. Available options are "significant" (abbreviation accepted: "s") or "non-significant" (abbreviation accepted: "ns") or "everything"/"all". The default is "significant". You can use this argument to make sure that your plot is not uber-cluttered when you have multiple groups being compared and scores of pairwise comparisons being displayed.
p.adjust.method	Adjustment method for <i>p</i> -values for multiple comparisons. Possible methods are: "holm" (default), "hochberg", "hommel", "bonferroni", "BH", "BY", "fdr", "none".
effsize.type	Type of effect size needed for <i>parametric</i> tests. The argument can be "biased" ("d" for Cohen's <i>d</i> for t-test ; "partial_eta" for partial eta-squared for anova) or "unbiased" ("g" Hedge's <i>g</i> for t-test ; "partial_omega" for partial omega-squared for anova)).
partial	Logical that decides if partial eta-squared or omega-squared are returned (Default: TRUE). If FALSE, eta-squared or omega-squared will be returned. Valid only for objects of class lm, aov, anova, or aovlist.
effsize.noncentral	Logical indicating whether to use non-central <i>t</i> -distributions for computing the confidence interval for Cohen's <i>d</i> or Hedge's <i>g</i> (Default: TRUE).
bf.prior	A number between 0.5 and 2 (default 0.707), the prior width to use in calculating Bayes factors.
bf.message	Logical that decides whether to display Bayes Factor in favor of the <i>null</i> hypothesis for parametric test (Default: FALSE).
results.subtitle	Decides whether the results of statistical tests are to be displayed as a subtitle (Default: TRUE). If set to FALSE, only the plot will be returned.
xlab	Labels for x and y axis variables. If NULL (default), variable names for x and y will be used.

<code>ylab</code>	Labels for x and y axis variables. If NULL (default), variable names for x and y will be used.
<code>subtitle</code>	The text for the plot subtitle. Will work only if <code>results.subtitle = FALSE</code> .
<code>caption</code>	The text for the plot caption.
<code>sample.size.label</code>	Logical that decides whether sample size information should be displayed for each level of the grouping variable x (Default: TRUE).
<code>k</code>	Number of digits after decimal point (should be an integer) (Default: <code>k = 2</code>).
<code>var.equal</code>	a logical variable indicating whether to treat the variances in the samples as equal. If TRUE, then a simple F test for the equality of means in a one-way analysis of variance is performed. If FALSE, an approximate method of Welch (1951) is used, which generalizes the commonly known 2-sample Welch test to the case of arbitrarily many samples.
<code>conf.level</code>	Scalar between 0 and 1. If unspecified, the defaults return 95% lower and upper confidence intervals (0.95).
<code>nboot</code>	Number of bootstrap samples for computing confidence interval for the effect size (Default: 100).
<code>tr</code>	Trim level for the mean when carrying out robust tests. If you get error stating "Standard error cannot be computed because of Winsorized variance of 0 (e.g., due to ties). Try to decrease the trimming level.", try to play around with the value of <code>tr</code> , which is by default set to 0.1. Lowering the value might help.
<code>axes.range.restrict</code>	Logical that decides whether to restrict the axes values ranges to min and max values of the axes variables (Default: FALSE), only relevant for functions where axes variables are of numeric type.
<code>mean.label.size</code>	Aesthetics for the label displaying mean. Defaults: 3, "bold", "black", respectively.
<code>mean.label.fontface</code>	Aesthetics for the label displaying mean. Defaults: 3, "bold", "black", respectively.
<code>mean.label.color</code>	Aesthetics for the label displaying mean. Defaults: 3, "bold", "black", respectively.
<code>notch</code>	A logical. If FALSE (default), a standard box plot will be displayed. If TRUE, a notched box plot will be used. Notches are used to compare groups; if the notches of two boxes do not overlap, this suggests that the medians are significantly different. In a notched box plot, the notches extend $1.58 * IQR / \sqrt{n}$. This gives a roughly 95% confidence interval for comparing medians. IQR: Inter-Quartile Range.
<code>notchwidth</code>	For a notched box plot, width of the notch relative to the body (default 0.5).
<code>linetype</code>	Character strings ("blank", "solid", "dashed", "dotted", "dotdash", "longdash", and "twodash") specifying the type of line to draw box plots (Default: "solid"). Alternatively, the numbers 0 to 6 can be used (0 for "blank", 1 for "solid", etc.).

<code>outlier.tagging</code>	Decides whether outliers should be tagged (Default: FALSE).
<code>outlier.label</code>	Label to put on the outliers that have been tagged.
<code>outlier.label.color</code>	Color for the label to to put on the outliers that have been tagged (Default: "black").
<code>outlier.color</code>	Default aesthetics for outliers (Default: "black").
<code>outlier.shape</code>	Hiding the outliers can be achieved by setting <code>outlier.shape = NA</code> . Importantly, this does not remove the outliers, it only hides them, so the range calculated for the y-axis will be the same with outliers shown and outliers hidden.
<code>outlier.coef</code>	Coefficient for outlier detection using Tukey's method. With Tukey's method, outliers are below (1st Quartile) or above (3rd Quartile) <code>outlier.coef</code> times the Inter-Quartile Range (IQR) (Default: 1.5).
<code>mean.plotting</code>	Logical that decides whether mean is to be highlighted and its value to be displayed (Default: TRUE).
<code>mean.ci</code>	Logical that decides whether 95 is to be displayed (Default: FALSE).
<code>mean.size</code>	Point size for the data point corresponding to mean (Default: 5).
<code>mean.color</code>	Color for the data point corresponding to mean (Default: "darkred").
<code>point.jitter.width</code>	Numeric specifying the degree of jitter in x direction. Defaults to 40% of the resolution of the data.
<code>point.jitter.height</code>	Numeric specifying the degree of jitter in y direction. Defaults to 0.1.
<code>point.dodge.width</code>	Numeric specifying the amount to dodge in the x direction. Defaults to 0.60.
<code>ggtheme</code>	A function, ggplot2 theme name. Default value is <code>ggplot2::theme_bw()</code> . Any of the ggplot2 themes, or themes from extension packages are allowed (e.g., <code>ggthemes::theme_fivethirtyeight()</code> , <code>hrbrthemes::theme_ipsum_ps()</code> , etc.).
<code>ggstatsplot.layer</code>	Logical that decides whether <code>theme_ggstatsplot</code> theme elements are to be displayed along with the selected <code>ggtheme</code> (Default: TRUE).
<code>package</code>	Name of package from which the palette is desired as string or symbol.
<code>palette</code>	If a character string (e.g., "Set1"), will use that named palette. If a number, will index into the list of palettes of appropriate type. Default palette is "Dark2".
<code>direction</code>	Either 1 or -1. If -1 the palette will be reversed.
<code>ggplot.component</code>	A ggplot component to be added to the plot prepared by <code>ggstatsplot</code> . This argument is primarily helpful for <code>grouped_</code> variant of the current function. Default is NULL. The argument should be entered as a function. If the given function has an argument <code>axes.range.restrict</code> and if it has been set to TRUE, the added ggplot component <i>might</i> not work as expected.
<code>messages</code>	Decides whether messages references, notes, and warnings are to be displayed (Default: TRUE).

...

Arguments passed on to `combine_plots`**title.text** String or plotmath expression to be drawn as title for the *combined plot*.**title.color** Text color for title.**title.size** Point size of title text.**title.vjust** Vertical justification for title. Default = 0.5 (centered on y). 0 = baseline at y, 1 = ascender at y.**title.hjust** Horizontal justification for title. Default = 0.5 (centered on x). 0 = flush-left at x, 1 = flush-right.**title.fontface** The font face ("plain", "bold" (default), "italic", "bold.italic") for title.**caption.text** String or plotmath expression to be drawn as the caption for the *combined plot*.**caption.color** Text color for caption.**caption.size** Point size of title text.**caption.vjust** Vertical justification for caption. Default = 0.5 (centered on y). 0 = baseline at y, 1 = ascender at y.**caption.hjust** Horizontal justification for caption. Default = 0.5 (centered on x). 0 = flush-left at x, 1 = flush-right.**caption.fontface** The font face ("plain" (default), "bold", "italic", "bold.italic") for caption.**sub.text** The label with which the *combined plot* should be annotated. Can be a plotmath expression.**sub.color** Text color for annotation label (Default: "black").**sub.size** Point size of annotation text (Default: 12).**sub.x** The x position of annotation label (Default: 0.5).**sub.y** The y position of annotation label (Default: 0.5).**sub.hjust** Horizontal justification for annotation label (Default: 0.5).**sub.vjust** Vertical justification for annotation label (Default: 0.5).**sub.vpadding** Vertical padding. The total vertical space added to the label, given in grid units. By default, this is added equally above and below the label. However, by changing the y and vjust parameters, this can be changed (Default: `grid::unit(1, "lines")`).**sub.fontface** The font face ("plain" (default), "bold", "italic", "bold.italic") for the annotation label.**sub.angle** Angle at which annotation label is to be drawn (Default: 0).**sub.lineheight** Line height of annotation label.**title.caption.rel.heights** Numerical vector of relative columns heights while combining (title, plot, caption).**title.rel.heights** Numerical vector of relative columns heights while combining (title, plot).**caption.rel.heights** Numerical vector of relative columns heights while combining (plot, caption).

Details

For parametric tests, Welch's ANOVA/*t*-test are used as a default (i.e., `var.equal = FALSE`).
References:

- ANOVA: Delacre, Leys, Mora, & Lakens, *PsyArXiv*, 2018
- *t*-test: Delacre, Lakens, & Leys, *International Review of Social Psychology*, 2017

If robust tests are selected, following tests are used is .

- ANOVA: one-way ANOVA on trimmed means (see `?WRS2::t1way`)
- *t*-test: Yuen's test for trimmed means (see `?WRS2::yuen`)

Variant of this function `ggwithinstats` is currently in progress. You *can* still use this function just to prepare the **plot** for exploratory data analysis, but the statistical details displayed in the subtitle will be incorrect. You can remove them by adding `+ ggplot2::labs(subtitle = NULL)`.

Author(s)

Indrajeet Patil, Chuck Powell

References

https://indrajeetpatil.github.io/ggstatsplot/articles/web_only/ggbetweenstats.html

See Also

[ggbetweenstats](#)

Examples

```
# to get reproducible results from bootstrapping
set.seed(123)

# the most basic function call
ggstatsplot::grouped_ggbetweenstats(
  data = dplyr::filter(ggplot2::mpg, drv != "4"),
  x = year,
  y = hwy,
  grouping.var = drv,
  conf.level = 0.99,
  bf.message = TRUE
)
## Not run:
# modifying individual plots using `ggplot.component` argument
ggstatsplot::grouped_ggbetweenstats(
  data = dplyr::filter(
    ggstatsplot::movies_long,
    genre %in% c("Action", "Comedy"),
    mpaa %in% c("R", "PG")
  ),
```

```

x = genre,
y = rating,
grouping.var = mpaa,
results.subtitle = FALSE,
ggplot.component = ggplot2::scale_y_continuous(breaks = seq(1, 9, 1)),
messages = FALSE
)

## End(Not run)

```

grouped_ggcorrmat	<i>Visualization of a correlalogram (or correlation matrix) for all levels of a grouping variable</i>
-------------------	---

Description

Helper function for `ggstatsplot::ggcorrmat` to apply this function across multiple levels of a given factor and combining the resulting plots using `ggstatsplot::combine_plots`.

Usage

```

grouped_ggcorrmat(data, cor.vars = NULL, cor.vars.names = NULL,
  grouping.var, title.prefix = NULL, output = "plot",
  matrix.type = "full", method = "square", corr.method = "pearson",
  type = NULL, exact = FALSE, continuity = TRUE, beta = 0.1,
  digits = 2, k = NULL, sig.level = 0.05, p.adjust.method = "none",
  hc.order = FALSE, hc.method = "complete", lab = TRUE,
  package = "RColorBrewer", palette = "Dark2", direction = 1,
  colors = c("#E69F00", "white", "#009E73"), outline.color = "black",
  ggtheme = ggplot2::theme_bw(), ggstatsplot.layer = TRUE,
  subtitle = NULL, caption = NULL, caption.default = TRUE,
  lab.col = "black", lab.size = 5, insig = "pch", pch = 4,
  pch.col = "black", pch.cex = 11, tl.cex = 12, tl.col = "black",
  tl.srt = 45, axis.text.x.margin.t = 0, axis.text.x.margin.r = 0,
  axis.text.x.margin.b = 0, axis.text.x.margin.l = 0,
  messages = TRUE, ...)

```

Arguments

<code>data</code>	Dataframe from which variables specified are preferentially to be taken.
<code>cor.vars</code>	List of variables for which the correlation matrix is to be computed and visualized. If <code>NULL</code> (default), all numeric variables from data will be used.
<code>cor.vars.names</code>	Optional list of names to be used for <code>cor.vars</code> . The names should be entered in the same order.
<code>grouping.var</code>	A single grouping variable (can be entered either as a bare name <code>x</code> or as a string <code>"x"</code>).

title.prefix	Character string specifying the prefix text for the fixed plot title (name of each factor level) (Default: NULL). If NULL, the variable name entered for grouping.var will be used.
output	Character that decides expected output from this function: "plot" (for visualization matrix) or "correlations" (or "corr" or "r"; for correlation matrix) or "p-values" (or "p.values" or "p"; for a matrix of p -values) or "ci" (for a tibble with confidence intervals for unique correlation pairs; not available for robust correlation) or "n" (or "sample.size" for a tibble with sample sizes for each correlation pair).
matrix.type	Character, "full" (default), "upper" or "lower", display full matrix, lower triangular or upper triangular matrix.
method	Character argument that decides the visualization method of correlation matrix to be used. Allowed values are "square" (default), "circle"
corr.method	A character string indicating which correlation coefficient is to be computed ("pearson" (default) or "kendall" or "spearman"). "robust" can also be entered but only if output argument is set to either "correlations" or "p-values". The robust correlation used is percentage bend correlation (see ?WRS2::pball). Abbreviations will also work: "p" (for parametric/Pearson's r), "np" (nonparametric/Spearman's ρ), "r" (robust).
type	A character string indicating which correlation coefficient is to be computed ("pearson" (default) or "kendall" or "spearman"). "robust" can also be entered but only if output argument is set to either "correlations" or "p-values". The robust correlation used is percentage bend correlation (see ?WRS2::pball). Abbreviations will also work: "p" (for parametric/Pearson's r), "np" (nonparametric/Spearman's ρ), "r" (robust).
exact	A logical indicating whether an exact p -value should be computed. Used for Kendall's τ and Spearman's ρ . For more details, see ?stats::cor.test.
continuity	A logical. If TRUE, a continuity correction is used for Kendall's τ and Spearman's ρ when not computed exactly (Default: TRUE).
beta	A numeric bending constant for robust correlation coefficient (Default: 0.1).
digits	Decides the number of decimal digits to be displayed (Default: 2).
k	Decides the number of decimal digits to be displayed (Default: 2).
sig.level	Significance level (Default: 0.05). If the p -value in p -value matrix is bigger than sig.level, then the corresponding correlation coefficient is regarded as insignificant and flagged as such in the plot. This argument is relevant only when output = "plot".
p.adjust.method	What adjustment for multiple tests should be used? ("holm", "hochberg", "hommel", "bonferroni", "BH", "BY", "fdr", "none"). See stats::p.adjust for details about why to use "holm" rather than "bonferroni". Default is "none". If adjusted p -values are displayed in the visualization of correlation matrix, the adjusted p -values will be used for the upper triangle, while unadjusted p -values will be used for the lower triangle of the matrix.
hc.order	Logical value. If TRUE, correlation matrix will be hc.ordered using hclust function (Default is FALSE).

hc.method	The agglomeration method to be used in hclust (see ?hclust).
lab	Logical value. If TRUE, correlation coefficient values will be displayed in the plot.
package	Name of package from which the palette is desired as string or symbol.
palette	Name of palette as string or symbol.
direction	Either 1 or -1. If -1 the palette will be reversed.
colors	A vector of 3 colors for low, mid, and high correlation values. If set to NULL, manual specification of colors will be turned off and 3 colors from the specified palette from package will be selected.
outline.color	The outline color of square or circle. Default value is "gray".
ggtheme	A function, ggplot2 theme name. Default value is ggplot2::theme_bw(). Any of the ggplot2 themes, or themes from extension packages are allowed (e.g., ggthemes::theme_fivethirtyeight(), hrbrthemes::theme_ipsum_ps(), etc.).
ggstatsplot.layer	Logical that decides whether theme_ggstatsplot theme elements are to be displayed along with the selected ggtheme (Default: TRUE).
subtitle	The text for the plot subtitle.
caption	The text for the plot caption. If not specified (if it is NULL, i.e.), a default caption will be shown.
caption.default	Logical decides whether the default caption should be shown.
lab.col	Color to be used for the correlation coefficient labels (applicable only when lab = TRUE).
lab.size	Size to be used for the correlation coefficient labels (applicable only when lab = TRUE).
insig	Character used to show specialized insignificant correlation coefficients ("pch" (default) or "blank"). If "blank", the corresponding glyphs will be removed; if "pch" is used, characters (see ?pch for details) will be added on the corresponding glyphs.
pch	Decides the glyphs (read point shapes) to be used for insignificant correlation coefficients (only valid when insig = "pch"). Default value is pch = 4.
pch.col	The color and the cex (size) of pch (only valid when insig = "pch"). Defaults are pch.col = "#F0E442" and pch.cex = 10.
pch.cex	The color and the cex (size) of pch (only valid when insig = "pch"). Defaults are pch.col = "#F0E442" and pch.cex = 10.
tl.cex	The size, the color, and the string rotation of text label (variable names, i.e.).
tl.col	The size, the color, and the string rotation of text label (variable names, i.e.).
tl.srt	The size, the color, and the string rotation of text label (variable names, i.e.).
axis.text.x.margin.t	Margins between x-axis and the variable name texts (t: top, r: right, b: bottom, l:left), especially useful in case the names are slanted, i.e. when the tl.srt is between 45 and 75 (Defaults: 0, 0, 0, 0, resp.).

`axis.text.x.margin.r`

Margins between x-axis and the variable name texts (t: top, r: right, b: bottom, l:left), especially useful in case the names are slanted, i.e. when the `tl.srt` is between 45 and 75 (Defaults: 0, 0, 0, 0, resp.).

`axis.text.x.margin.b`

Margins between x-axis and the variable name texts (t: top, r: right, b: bottom, l:left), especially useful in case the names are slanted, i.e. when the `tl.srt` is between 45 and 75 (Defaults: 0, 0, 0, 0, resp.).

`axis.text.x.margin.l`

Margins between x-axis and the variable name texts (t: top, r: right, b: bottom, l:left), especially useful in case the names are slanted, i.e. when the `tl.srt` is between 45 and 75 (Defaults: 0, 0, 0, 0, resp.).

`messages`

Decides whether messages references, notes, and warnings are to be displayed (Default: TRUE).

...

Arguments passed on to `combine_plots`

title.text String or plotmath expression to be drawn as title for the *combined plot*.

title.color Text color for title.

title.size Point size of title text.

title.vjust Vertical justification for title. Default = 0.5 (centered on y). 0 = baseline at y, 1 = ascender at y.

title.hjust Horizontal justification for title. Default = 0.5 (centered on x). 0 = flush-left at x, 1 = flush-right.

title.fontface The font face ("plain", "bold" (default), "italic", "bold.italic") for title.

caption.text String or plotmath expression to be drawn as the caption for the *combined plot*.

caption.color Text color for caption.

caption.size Point size of title text.

caption.vjust Vertical justification for caption. Default = 0.5 (centered on y). 0 = baseline at y, 1 = ascender at y.

caption.hjust Horizontal justification for caption. Default = 0.5 (centered on x). 0 = flush-left at x, 1 = flush-right.

caption.fontface The font face ("plain" (default), "bold", "italic", "bold.italic") for caption.

sub.text The label with which the *combined plot* should be annotated. Can be a plotmath expression.

sub.color Text color for annotation label (Default: "black").

sub.size Point size of annotation text (Default: 12).

sub.x The x position of annotation label (Default: 0.5).

sub.y The y position of annotation label (Default: 0.5).

sub.hjust Horizontal justification for annotation label (Default: 0.5).

sub.vjust Vertical justification for annotation label (Default: 0.5).

sub.vpadding Vertical padding. The total vertical space added to the label, given in grid units. By default, this is added equally above and below the label. However, by changing the `y` and `vjust` parameters, this can be changed (Default: `grid::unit(1, "lines")`).

sub.fontface The font face ("plain" (default), "bold", "italic", "bold.italic") for the annotation label.

sub.angle Angle at which annotation label is to be drawn (Default: 0).

sub.lineheight Line height of annotation label.

title.caption.rel.heights Numerical vector of relative columns heights while combining (title, plot, caption).

title.rel.heights Numerical vector of relative columns heights while combining (title, plot).

caption.rel.heights Numerical vector of relative columns heights while combining (plot, caption).

Value

Correlation matrix plot or correlation coefficient matrix or matrix of p-values.

Author(s)

Indrajeet Patil, Chuck Powell

References

https://indrajeetpatil.github.io/ggstatsplot/articles/web_only/ggcorrmat.html

See Also

[ggcorrmat](#), [ggscatterstats](#), [grouped_ggscatterstats](#)

Examples

```
# for reproducibility
set.seed(123)

# for plot
# (without specifying needed variables; all numeric variables will be used)
ggstatsplot::grouped_ggcorrmat(
  data = ggplot2::msleep,
  grouping.var = vore
)

# for getting plot
ggstatsplot::grouped_ggcorrmat(
  data = ggplot2::msleep,
  grouping.var = vore,
  cor.vars = sleep_total:bodywt,
  corr.method = "r",
```

```

    p.adjust.method = "holm",
    colors = NULL,
    package = "wesanderson",
    palette = "BottleRocket2",
    nrow = 2
  )

  # for getting correlations
  ggstatsplot::grouped_ggcorrmat(
    data = ggplot2::msleep,
    grouping.var = vore,
    cor.vars = sleep_total:bodywt,
    output = "correlations"
  )

  # for getting confidence intervals
  # confidence intervals are not available for **robust** correlation
  ggstatsplot::grouped_ggcorrmat(
    data = datasets::iris,
    grouping.var = Species,
    corr.method = "r",
    p.adjust.method = "holm",
    cor.vars = Sepal.Length:Petal.Width,
    output = "ci"
  )

```

grouped_ggdotplotstats

Grouped histograms for distribution of a labelled numeric variable

Description

Helper function for `ggstatsplot::ggdotplotstats` to apply this function across multiple levels of a given factor and combining the resulting plots using `ggstatsplot::combine_plots`.

Usage

```

grouped_ggdotplotstats(data, x, y, grouping.var, title.prefix = NULL,
  xlab = NULL, ylab = NULL, subtitle = NULL, caption = NULL,
  type = "parametric", test.value = 0, bf.prior = 0.707,
  bf.message = FALSE, robust.estimator = "onestep",
  conf.level = 0.95, nboot = 100, k = 2,
  ggtheme = ggplot2::theme_bw(), ggstatsplot.layer = TRUE,
  point.color = "black", point.size = 3, point.shape = 16,
  results.subtitle = TRUE, centrality.para = "mean",
  centrality.color = "blue", centrality.size = 1,
  centrality.linetype = "dashed", centrality.line.labeller = TRUE,
  centrality.k = 2, test.value.line = FALSE,
  test.value.color = "black", test.value.size = 1,

```

```
test.value.linetype = "dashed", test.line.labeller = TRUE,
test.k = 0, ggplot.component = NULL, messages = TRUE, ...)
```

Arguments

data	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
x	A numeric variable.
y	Label or grouping variable.
grouping.var	A single grouping variable (can be entered either as a bare name x or as a string "x").
title.prefix	Character string specifying the prefix text for the fixed plot title (name of each factor level) (Default: NULL). If NULL, the variable name entered for grouping.var will be used.
xlab	Labels for x and y axis variables. If NULL (default), variable names for x and y will be used.
ylab	Label for y axis variable.
subtitle	The text for the plot subtitle. Will work only if results.subtitle = FALSE.
caption	The text for the plot caption.
type	Type of statistic expected ("parametric" or "nonparametric" or "robust" or "bayes").Corresponding abbreviations are also accepted: "p" (for parametric), "np" (nonparametric), "r" (robust), or "bf" resp.
test.value	A number specifying the value of the null hypothesis (Default: 0).
bf.prior	A number between 0.5 and 2 (default 0.707), the prior width to use in calculating Bayes factors.
bf.message	Logical that decides whether to display Bayes Factor in favor of the <i>null</i> hypothesis for parametric test (Default: FALSE).
robust.estimator	If test = "robust" robust estimator to be used ("onestep" (Default), "mom", or "median"). For more, see ?WRS2::onesampb.
conf.level	Scalar between 0 and 1. If unspecified, the defaults return 95% lower and upper confidence intervals (0.95).
nboot	Number of bootstrap samples for computing confidence interval for the effect size (Default: 100).
k	Number of digits after decimal point (should be an integer) (Default: k = 2).
ggtheme	A function, ggplot2 theme name. Default value is ggplot2::theme_bw(). Any of the ggplot2 themes, or themes from extension packages are allowed (e.g., ggthemes::theme_fivethirtyeight(), hrbrthemes::theme_ipsum_ps(), etc.).
ggstatsplot.layer	Logical that decides whether theme_ggstatsplot theme elements are to be displayed along with the selected ggtheme (Default: TRUE).
point.color	Character describing color for the point (Default: "black").

<code>point.size</code>	Numeric specifying size for the point (Default: 3).
<code>point.shape</code>	Numeric specifying shape to draw the points (Default: 16 (a dot)).
<code>results.subtitle</code>	Decides whether the results of statistical tests are to be displayed as a subtitle (Default: TRUE). If set to FALSE, only the plot will be returned.
<code>centrality.param</code>	Decides <i>which</i> measure of central tendency ("mean" or "median") is to be displayed as a vertical line.
<code>centrality.color</code>	Decides color for the vertical line for centrality parameter (Default: "blue").
<code>centrality.size</code>	Decides size for the vertical line for centrality parameter (Default: 1.2).
<code>centrality.linetype</code>	Decides linetype for the vertical line for centrality parameter (Default: "dashed").
<code>centrality.line.labeller</code>	A logical that decides whether line labels should be displayed for the centrality.param line (Default: TRUE).
<code>centrality.k</code>	Integer denoting the number of decimal places expected for centrality parameter label. (Default: 2).
<code>test.value.line</code>	Decides whether test value is to be displayed as a vertical line (Default: FALSE).
<code>test.value.color</code>	Decides color for the vertical line denoting test value (Default: "black").
<code>test.value.size</code>	Decides size for the vertical line for test value (Default: 1.2).
<code>test.value.linetype</code>	Decides linetype for the vertical line for test value (Default: "dashed").
<code>test.line.labeller</code>	A logical that decides whether line labels should be displayed for test.value line (Default: TRUE).
<code>test.k</code>	Integer denoting the number of decimal places expected for test.value label. (Default: 0).
<code>ggplot.component</code>	A ggplot component to be added to the plot prepared by ggstatsplot. This argument is primarily helpful for grouped_ variant of the current function. Default is NULL. The argument should be entered as a function. If the given function has an argument <code>axes.range.restrict</code> and if it has been set to TRUE, the added ggplot component <i>might</i> not work as expected.
<code>messages</code>	Decides whether messages references, notes, and warnings are to be displayed (Default: TRUE).
<code>...</code>	Arguments passed on to <code>combine_plots</code>
	title.text String or plotmath expression to be drawn as title for the <i>combined plot</i> .
	title.color Text color for title.

title.size Point size of title text.

title.vjust Vertical justification for title. Default = 0.5 (centered on y). 0 = baseline at y, 1 = ascender at y.

title.hjust Horizontal justification for title. Default = 0.5 (centered on x). 0 = flush-left at x, 1 = flush-right.

title.fontface The font face ("plain", "bold" (default), "italic", "bold.italic") for title.

caption.text String or plotmath expression to be drawn as the caption for the *combined plot*.

caption.color Text color for caption.

caption.size Point size of title text.

caption.vjust Vertical justification for caption. Default = 0.5 (centered on y). 0 = baseline at y, 1 = ascender at y.

caption.hjust Horizontal justification for caption. Default = 0.5 (centered on x). 0 = flush-left at x, 1 = flush-right.

caption.fontface The font face ("plain" (default), "bold", "italic", "bold.italic") for caption.

sub.text The label with which the *combined plot* should be annotated. Can be a plotmath expression.

sub.color Text color for annotation label (Default: "black").

sub.size Point size of annotation text (Default: 12).

sub.x The x position of annotation label (Default: 0.5).

sub.y The y position of annotation label (Default: 0.5).

sub.hjust Horizontal justification for annotation label (Default: 0.5).

sub.vjust Vertical justification for annotation label (Default: 0.5).

sub.vpadding Vertical padding. The total vertical space added to the label, given in grid units. By default, this is added equally above and below the label. However, by changing the y and vjust parameters, this can be changed (Default: `grid::unit(1, "lines")`).

sub.fontface The font face ("plain" (default), "bold", "italic", "bold.italic") for the annotation label.

sub.angle Angle at which annotation label is to be drawn (Default: 0).

sub.lineheight Line height of annotation label.

title.caption.rel.heights Numerical vector of relative columns heights while combining (title, plot, caption).

title.rel.heights Numerical vector of relative columns heights while combining (title, plot).

caption.rel.heights Numerical vector of relative columns heights while combining (plot, caption).

Author(s)

Indrajeet Patil

See Also[ggdotplotstats](#)**Examples**

```
# for reproducibility
set.seed(123)

# removing factor level with very few no. of observations
df <- dplyr::filter(.data = ggplot2::mpg, cyl %in% c("4", "6", "8"))

# plot
ggstatsplot::grouped_ggdotplotstats(
  data = df,
  x = "cty",
  y = "manufacturer",
  grouping.var = "cyl",
  test.value = 15.5,
  title.prefix = "cylinder count",
  ggplot.component = ggplot2::scale_x_continuous(
    sec.axis = ggplot2::dup_axis(),
    limits = c(12, 24),
    breaks = seq(12, 24, 2)
  ),
  messages = FALSE
)
```

grouped_gghistostats *Grouped histograms for distribution of a numeric variable*

Description

Helper function for `ggstatsplot::gghistostats` to apply this function across multiple levels of a given factor and combining the resulting plots using `ggstatsplot::combine_plots`.

Usage

```
grouped_gghistostats(data, x, grouping.var, title.prefix = NULL,
  binwidth = NULL, bar.measure = "count", xlab = NULL,
  subtitle = NULL, caption = NULL, type = "parametric",
  test.value = 0, bf.prior = 0.707, bf.message = FALSE,
  robust.estimator = "onestep", effsize.type = "g",
  effsize.noncentral = TRUE, conf.level = 0.95, nboot = 100, k = 2,
  ggtheme = ggplot2::theme_bw(), ggstatsplot.layer = TRUE,
  fill.gradient = FALSE, low.color = "#0072B2",
  high.color = "#D55E00", bar.fill = "grey50",
  results.subtitle = TRUE, centrality.param = "mean",
```

```

centrality.color = "blue", centrality.size = 1,
centrality.linetype = "dashed", centrality.line.labeller = TRUE,
centrality.k = 2, test.value.line = FALSE,
test.value.color = "black", test.value.size = 1,
test.value.linetype = "dashed", test.line.labeller = TRUE,
test.k = 0, normal.curve = FALSE, normal.curve.color = "black",
normal.curve.linetype = "solid", normal.curve.size = 1,
ggplot.component = NULL, messages = TRUE, ...)

```

Arguments

data	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
x	A numeric variable.
grouping.var	A single grouping variable (can be entered either as a bare name x or as a string "x").
title.prefix	Character string specifying the prefix text for the fixed plot title (name of each factor level) (Default: NULL). If NULL, the variable name entered for grouping.var will be used.
binwidth	The width of the histogram bins. Can be specified as a numeric value, or a function that calculates width from x. The default is to use the $\max(x) - \min(x) / \sqrt{N}$. You should always check this value and explore multiple widths to find the best to illustrate the stories in your data.
bar.measure	Character describing what value needs to be represented as height in the bar chart. This can either be "count", which shows number of points in bin, or "density", which density of points in bin, scaled to integrate to 1, or "proportion", which shows relative frequencies of observations in each bin, or "mix", which shows <i>both</i> count and proportion in the same plot.
xlab	Labels for x and y axis variables. If NULL (default), variable names for x and y will be used.
subtitle	The text for the plot subtitle. Will work only if results.subtitle = FALSE.
caption	The text for the plot caption.
type	Type of statistic expected ("parametric" or "nonparametric" or "robust" or "bayes").Corresponding abbreviations are also accepted: "p" (for parametric), "np" (nonparametric), "r" (robust), or "bf"resp.
test.value	A number specifying the value of the null hypothesis (Default: 0).
bf.prior	A number between 0.5 and 2 (default 0.707), the prior width to use in calculating Bayes factors.
bf.message	Logical that decides whether to display Bayes Factor in favor of the <i>null</i> hypothesis for parametric test (Default: FALSE).
robust.estimator	If test = "robust" robust estimator to be used ("onestep" (Default), "mom", or "median"). For more, see ?WRS2::onesampb.
effsize.type	Type of effect size needed for <i>parametric</i> tests. The argument can be "biased" ("d" for Cohen's <i>d</i>) or "unbiased" ("g" Hedge's <i>g</i> for t-test). The default is

<code>effsize.noncentral</code>	Logical indicating whether to use non-central t -distributions for computing the confidence interval for Cohen's d or Hedge's g (Default: TRUE).
<code>conf.level</code>	Scalar between 0 and 1. If unspecified, the defaults return 95% lower and upper confidence intervals (0.95).
<code>nboot</code>	Number of bootstrap samples for computing confidence interval for the effect size (Default: 100).
<code>k</code>	Number of digits after decimal point (should be an integer) (Default: $k = 2$).
<code>ggtheme</code>	A function, ggplot2 theme name. Default value is <code>ggplot2::theme_bw()</code> . Any of the ggplot2 themes, or themes from extension packages are allowed (e.g., <code>ggthemes::theme_fivethirtyeight()</code> , <code>hrbrthemes::theme_ipsum_ps()</code> , etc.).
<code>ggstatsplot.layer</code>	Logical that decides whether <code>theme_ggstatsplot</code> theme elements are to be displayed along with the selected <code>ggtheme</code> (Default: TRUE).
<code>fill.gradient</code>	Logical decides whether color fill gradient is to be displayed (Default: FALSE). If FALSE, the legend and the color gradient will also be removed. The default is set to FALSE because the gradient provides redundant information in light of y-axis labels.
<code>low.color</code>	Colors for low and high ends of the gradient. Defaults are colorblind-friendly.
<code>high.color</code>	Colors for low and high ends of the gradient. Defaults are colorblind-friendly.
<code>bar.fill</code>	If <code>fill.gradient = FALSE</code> , then <code>bar.fill</code> decides which color will uniformly fill all the bars in the histogram (Default: "grey50").
<code>results.subtitle</code>	Decides whether the results of statistical tests are to be displayed as a subtitle (Default: TRUE). If set to FALSE, only the plot will be returned.
<code>centrality.para</code>	Decides <i>which</i> measure of central tendency ("mean" or "median") is to be displayed as a vertical line.
<code>centrality.color</code>	Decides color for the vertical line for centrality parameter (Default: "blue").
<code>centrality.size</code>	Decides size for the vertical line for centrality parameter (Default: 1.2).
<code>centrality.linetype</code>	Decides linetype for the vertical line for centrality parameter (Default: "dashed").
<code>centrality.line.labeller</code>	A logical that decides whether line labels should be displayed for the centrality.para line (Default: TRUE).
<code>centrality.k</code>	Integer denoting the number of decimal places expected for centrality parameter label. (Default: 2).
<code>test.value.line</code>	Decides whether test value is to be displayed as a vertical line (Default: FALSE).
<code>test.value.color</code>	Decides color for the vertical line denoting test value (Default: "black").

<code>test.value.size</code>	Decides size for the vertical line for test value (Default: 1.2).
<code>test.value.linetype</code>	Decides linetype for the vertical line for test value (Default: "dashed").
<code>test.line.labeller</code>	A logical that decides whether line labels should be displayed for test.value line (Default: TRUE).
<code>test.k</code>	Integer denoting the number of decimal places expected for test.value label. (Default: 0).
<code>normal.curve</code>	Logical decides whether to super-impose a normal curve using <code>stats::dnorm(mean(x), sd(x))</code> . Default is FALSE.
<code>normal.curve.color</code>	If <code>normal.curve = TRUE</code> , then these arguments can be used to modify color (Default: "black"), size (default: 1.0), linetype (default: "solid").
<code>normal.curve.linetype</code>	If <code>normal.curve = TRUE</code> , then these arguments can be used to modify color (Default: "black"), size (default: 1.0), linetype (default: "solid").
<code>normal.curve.size</code>	If <code>normal.curve = TRUE</code> , then these arguments can be used to modify color (Default: "black"), size (default: 1.0), linetype (default: "solid").
<code>ggplot.component</code>	A ggplot component to be added to the plot prepared by <code>ggstatsplot</code> . This argument is primarily helpful for <code>grouped_</code> variant of the current function. Default is NULL. The argument should be entered as a function. If the given function has an argument <code>axes.range.restrict</code> and if it has been set to TRUE, the added ggplot component <i>might</i> not work as expected.
<code>messages</code>	Decides whether messages references, notes, and warnings are to be displayed (Default: TRUE).
<code>...</code>	Arguments passed on to <code>combine_plots</code>
	title.text String or plotmath expression to be drawn as title for the <i>combined plot</i> .
	title.color Text color for title.
	title.size Point size of title text.
	title.vjust Vertical justification for title. Default = 0.5 (centered on y). 0 = baseline at y, 1 = ascender at y.
	title.hjust Horizontal justification for title. Default = 0.5 (centered on x). 0 = flush-left at x, 1 = flush-right.
	title.fontface The font face ("plain", "bold" (default), "italic", "bold.italic") for title.
	caption.text String or plotmath expression to be drawn as the caption for the <i>combined plot</i> .
	caption.color Text color for caption.
	caption.size Point size of title text.
	caption.vjust Vertical justification for caption. Default = 0.5 (centered on y). 0 = baseline at y, 1 = ascender at y.

caption.hjust Horizontal justification for caption. Default = 0.5 (centered on x). 0 = flush-left at x, 1 = flush-right.

caption.fontface The font face ("plain" (default), "bold", "italic", "bold.italic") for caption.

sub.text The label with which the *combined plot* should be annotated. Can be a plotmath expression.

sub.color Text color for annotation label (Default: "black").

sub.size Point size of annotation text (Default: 12).

sub.x The x position of annotation label (Default: 0.5).

sub.y The y position of annotation label (Default: 0.5).

sub.hjust Horizontal justification for annotation label (Default: 0.5).

sub.vjust Vertical justification for annotation label (Default: 0.5).

sub.vpadding Vertical padding. The total vertical space added to the label, given in grid units. By default, this is added equally above and below the label. However, by changing the y and vjust parameters, this can be changed (Default: `grid::unit(1, "lines")`).

sub.fontface The font face ("plain" (default), "bold", "italic", "bold.italic") for the annotation label.

sub.angle Angle at which annotation label is to be drawn (Default: 0).

sub.lineheight Line height of annotation label.

title.caption.rel.heights Numerical vector of relative columns heights while combining (title, plot, caption).

title.rel.heights Numerical vector of relative columns heights while combining (title, plot).

caption.rel.heights Numerical vector of relative columns heights while combining (plot, caption).

Author(s)

Indrajeet Patil, Chuck Powell

References

https://indrajeetpatil.github.io/ggstatsplot/articles/web_only/gghistostats.html

See Also

[gghistostats](#)

Examples

```
ggstatsplot::grouped_gghistostats(
  data = iris,
  x = Sepal.Length,
  test.value = 5,
  bf.message = TRUE,
  grouping.var = Species,
```

```

bar.fill = "orange",
nrow = 1,
ggplot.component = list(
  ggplot2::scale_x_continuous(breaks = seq(3, 9, 1), limits = (c(3, 9))),
  ggplot2::scale_y_continuous(breaks = seq(0, 25, 5), limits = (c(0, 25)))
),
messages = FALSE
)

```

grouped_ggpiestats *Grouped pie charts with statistical tests*

Description

Helper function for `ggstatsplot::ggpiestats` to apply this function across multiple levels of a given factor and combining the resulting plots using `ggstatsplot::combine_plots`.

Usage

```

grouped_ggpiestats(data, main, condition = NULL, counts = NULL,
  grouping.var, title.prefix = NULL, ratio = NULL, paired = FALSE,
  results.subtitle = TRUE, factor.levels = NULL, stat.title = NULL,
  sample.size.label = TRUE, label.separator = "\n",
  label.text.size = 4, label.fill.color = "white",
  label.fill.alpha = 1, bf.message = FALSE,
  sampling.plan = "indepMulti", fixed.margin = "rows",
  prior.concentration = 1, subtitle = NULL, caption = NULL,
  conf.level = 0.95, nboot = 100, simulate.p.value = FALSE,
  B = 2000, legend.title = NULL, facet.wrap.name = NULL, k = 2,
  perc.k = 0, slice.label = "percentage", facet.proptest = TRUE,
  ggtheme = ggplot2::theme_bw(), ggstatsplot.layer = TRUE,
  package = "RColorBrewer", palette = "Dark2", direction = 1,
  ggplot.component = NULL, messages = TRUE, ...)

```

Arguments

<code>data</code>	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
<code>main</code>	The variable to use as the rows in the contingency table.
<code>condition</code>	The variable to use as the columns in the contingency table.
<code>counts</code>	A string naming a variable in data containing counts, or NULL if each row represents a single observation (Default).
<code>grouping.var</code>	A single grouping variable (can be entered either as a bare name <code>x</code> or as a string <code>"x"</code>).
<code>title.prefix</code>	Character string specifying the prefix text for the fixed plot title (name of each factor level) (Default: NULL). If NULL, the variable name entered for <code>grouping.var</code> will be used.

<code>ratio</code>	A vector of numbers: the expected proportions for the proportion test. Default is NULL, which means if there are two levels <code>ratio = c(1,1)</code> , etc.
<code>paired</code>	Logical indicating whether data came from a within-subjects design study (Default: FALSE). If TRUE, McNemar test subtitle will be returned. If FALSE, Pearson's chi-square test will be returned.
<code>results.subtitle</code>	Decides whether the results of statistical tests are to be displayed as a subtitle (Default: TRUE). If set to FALSE, only the plot will be returned.
<code>factor.levels</code>	A character vector with labels for factor levels of main variable.
<code>stat.title</code>	Title for the effect being investigated with the chi-square test. The default is NULL, i.e. no title will be added to describe the effect being shown. An example of a <code>stat.title</code> argument will be something like "main x condition" or "interaction".
<code>sample.size.label</code>	Logical that decides whether sample size information should be displayed for each level of the grouping variable condition (Default: TRUE).
<code>label.separator</code>	If "both" counts and proportion information is to be displayed in a label, this argument decides whether these two pieces of information are going to be on the same line (" ") or on separate lines ("\n").
<code>label.text.size</code>	Numeric that decides text size for slice/bar labels (Default: 4).
<code>label.fill.color</code>	Character that specifies fill color for slice/bar labels (Default: white).
<code>label.fill.alpha</code>	Numeric that specifies fill color transparency or "alpha" for slice/bar labels (Default: 1 range 0 to 1).
<code>bf.message</code>	Logical that decides whether to display a caption with results from bayes factor test in favor of the null hypothesis (default: FALSE).
<code>sampling.plan</code>	Character describing the sampling plan. Possible options are "indepMulti" (independent multinomial; default), "poisson", "jointMulti" (joint multinomial), "hypergeom" (hypergeometric). For more, see <code>?BayesFactor::contingencyTableBF()</code> .
<code>fixed.margin</code>	For the independent multinomial sampling plan, which margin is fixed ("rows" or "cols"). Defaults to "rows".
<code>prior.concentration</code>	Specifies the prior concentration parameter, set to 1 by default. It indexes the expected deviation from the null hypothesis under the alternative, and corresponds to Gunel and Dickey's (1974) "a" parameter.
<code>subtitle</code>	The text for the plot subtitle. Will work only if <code>results.subtitle = FALSE</code> .
<code>caption</code>	The text for the plot caption.
<code>conf.level</code>	Scalar between 0 and 1. If unspecified, the defaults return 95% lower and upper confidence intervals (0.95).
<code>nboot</code>	Number of bootstrap samples for computing confidence interval for the effect size (Default: 100).

<code>simulate.p.value</code>	a logical indicating whether to compute p-values by Monte Carlo simulation.
<code>B</code>	an integer specifying the number of replicates used in the Monte Carlo test.
<code>legend.title</code>	Title text for the legend.
<code>facet.wrap.name</code>	The text for the <code>facet_wrap</code> variable label.
<code>k</code>	Number of digits after decimal point (should be an integer) (Default: <code>k = 2</code>).
<code>perc.k</code>	Numeric that decides number of decimal places for percentage labels (Default: <code>0</code>).
<code>slice.label</code>	Character decides what information needs to be displayed on the label in each pie slice. Possible options are "percentage" (default), "counts", "both".
<code>facet.proptest</code>	Decides whether proportion test for main variable is to be carried out for each level of condition (Default: <code>TRUE</code>).
<code>ggtheme</code>	A function, ggplot2 theme name. Default value is <code>ggplot2::theme_bw()</code> . Any of the ggplot2 themes, or themes from extension packages are allowed (e.g., <code>ggthemes::theme_fivethirtyeight()</code> , <code>hrbrthemes::theme_ipsum_ps()</code> , etc.).
<code>ggstatsplot.layer</code>	Logical that decides whether <code>theme_ggstatsplot</code> theme elements are to be displayed along with the selected <code>ggtheme</code> (Default: <code>TRUE</code>).
<code>package</code>	Name of package from which the palette is desired as string or symbol.
<code>palette</code>	If a character string (e.g., "Set1"), will use that named palette. If a number, will index into the list of palettes of appropriate type. Default palette is "Dark2".
<code>direction</code>	Either 1 or -1. If -1 the palette will be reversed.
<code>ggplot.component</code>	A ggplot component to be added to the plot prepared by <code>ggstatsplot</code> . This argument is primarily helpful for <code>grouped_</code> variant of the current function. Default is <code>NULL</code> . The argument should be entered as a function. If the given function has an argument <code>axes.range.restrict</code> and if it has been set to <code>TRUE</code> , the added ggplot component <i>might</i> not work as expected.
<code>messages</code>	Decides whether messages references, notes, and warnings are to be displayed (Default: <code>TRUE</code>).
<code>...</code>	Arguments passed on to <code>combine_plots</code>
	title.text String or plotmath expression to be drawn as title for the <i>combined plot</i> .
	title.color Text color for title.
	title.size Point size of title text.
	title.vjust Vertical justification for title. Default = <code>0.5</code> (centered on y). <code>0</code> = baseline at y, <code>1</code> = ascender at y.
	title.hjust Horizontal justification for title. Default = <code>0.5</code> (centered on x). <code>0</code> = flush-left at x, <code>1</code> = flush-right.
	title.fontface The font face ("plain", "bold" (default), "italic", "bold.italic") for title.

caption.text String or plotmath expression to be drawn as the caption for the *combined plot*.

caption.color Text color for caption.

caption.size Point size of title text.

caption.vjust Vertical justification for caption. Default = 0.5 (centered on y). 0 = baseline at y, 1 = ascender at y.

caption.hjust Horizontal justification for caption. Default = 0.5 (centered on x). 0 = flush-left at x, 1 = flush-right.

caption.fontface The font face ("plain" (default), "bold", "italic", "bold.italic") for caption.

sub.text The label with which the *combined plot* should be annotated. Can be a plotmath expression.

sub.color Text color for annotation label (Default: "black").

sub.size Point size of annotation text (Default: 12).

sub.x The x position of annotation label (Default: 0.5).

sub.y The y position of annotation label (Default: 0.5).

sub.hjust Horizontal justification for annotation label (Default: 0.5).

sub.vjust Vertical justification for annotation label (Default: 0.5).

sub.vpadding Vertical padding. The total vertical space added to the label, given in grid units. By default, this is added equally above and below the label. However, by changing the y and vjust parameters, this can be changed (Default: grid::unit(1, "lines")).

sub.fontface The font face ("plain" (default), "bold", "italic", "bold.italic") for the annotation label.

sub.angle Angle at which annotation label is to be drawn (Default: 0).

sub.lineheight Line height of annotation label.

title.caption.rel.heights Numerical vector of relative columns heights while combining (title, plot, caption).

title.rel.heights Numerical vector of relative columns heights while combining (title, plot).

caption.rel.heights Numerical vector of relative columns heights while combining (plot, caption).

Value

Unlike a number of statistical softwares, ggstatsplot doesn't provide the option for Yates' correction for the Pearson's chi-squared statistic. This is due to compelling amount of Monte-Carlo simulation research which suggests that the Yates' correction is overly conservative, even in small sample sizes. As such it is recommended that it should not ever be applied in practice (Camilli & Hopkins, 1978, 1979; Feinberg, 1980; Larntz, 1978; Thompson, 1988).

Author(s)

Indrajeet Patil, Chuck Powell

References

https://indrajeetpatil.github.io/ggstatsplot/articles/web_only/ggpiestats.html

See Also

[ggpiestats](#)

Examples

```
# grouped one-sample proportion tests
ggstatsplot::grouped_ggpiestats(
  data = mtcars,
  grouping.var = am,
  main = cyl
)

# without condition and with count data
library(jmv)

ggstatsplot::grouped_ggpiestats(
  data = as.data.frame(HairEyeColor),
  main = Hair,
  counts = Freq,
  grouping.var = Sex
)

# the following will take slightly more amount of time
## Not run:
# for reproducibility
set.seed(123)

# let's create a smaller dataframe
diamonds_short <- ggplot2::diamonds %>%
  dplyr::filter(.data = ., cut %in% c("Fair", "Very Good", "Ideal")) %>%
  dplyr::sample_frac(tbl = ., size = 0.10)

# plot
ggstatsplot::grouped_ggpiestats(
  data = diamonds_short,
  main = color,
  condition = clarity,
  grouping.var = cut,
  bf.message = TRUE,
  sampling.plan = "poisson",
  title.prefix = "Quality",
  slice.label = "both",
  messages = FALSE,
  perc.k = 1,
  nrow = 3
)
```



```
## End(Not run)
```

```
grouped_ggscatterstats
```

Scatterplot with marginal distributions for all levels of a grouping variable

Description

Grouped scatterplots from ggplot2 combined with marginal histograms/boxplots/density plots with statistical details added as a subtitle.

Usage

```
grouped_ggscatterstats(data, x, y, type = "pearson", conf.level = 0.95,
  bf.prior = 0.707, bf.message = FALSE, label.var = NULL,
  label.expression = NULL, grouping.var, title.prefix = NULL,
  xlab = NULL, ylab = NULL, method = "lm", method.args = list(),
  formula = y ~ x, point.color = "black", point.size = 3,
  point.alpha = 0.4, line.size = 1.5, point.width.jitter = 0,
  point.height.jitter = 0, line.color = "blue", marginal = TRUE,
  marginal.type = "histogram", marginal.size = 5, margins = c("both",
  "x", "y"), package = "wesanderson", palette = "Royal1",
  direction = 1, xfill = "#009E73", yfill = "#D55E00", xalpha = 1,
  yalpha = 1, xsize = 0.7, ysize = 0.7, centrality.param = NULL,
  results.subtitle = TRUE, caption = NULL, subtitle = NULL,
  nboot = 100, beta = 0.1, k = 2, axes.range.restrict = FALSE,
  ggtheme = ggplot2::theme_bw(), ggstatsplot.layer = TRUE,
  ggplot.component = NULL, messages = TRUE, ...)
```

Arguments

data	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
x	The column in data containing the explanatory variable to be plotted on the x axis. Can be entered either as a character string (e.g., "x") or as a bare expression (e.g, x).
y	The column in data containing the response (outcome) variable to be plotted on the y axis. Can be entered either as a character string (e.g., "y") or as a bare expression (e.g, y).
type	Type of association between paired samples required ("parametric": Pearson's product moment correlation coefficient" or "nonparametric": Spearman's rho" or "robust": percentage bend correlation coefficient" or "bayes": Bayes Factor for Pearson's r "). Corresponding abbreviations are also accepted: "p" (for parametric/pearson's), "np" (nonparametric/spearman), "r" (robust), "bf" (for bayes factor), resp.

<code>conf.level</code>	Scalar between 0 and 1. If unspecified, the defaults return 95% lower and upper confidence intervals (0.95).
<code>bf.prior</code>	A number between 0.5 and 2 (default 0.707), the prior width to use in calculating Bayes factors.
<code>bf.message</code>	Logical that decides whether to display Bayes Factor in favor of the <i>null</i> hypothesis for parametric test (Default: FALSE).
<code>label.var</code>	Variable to use for points labels. Can be entered either as a character string (e.g., "var1") or as a bare expression (e.g, var1).
<code>label.expression</code>	An expression evaluating to a logical vector that determines the subset of data points to label. This argument can be entered either as a character string (e.g., "y < 4 & z < 20") or as a bare expression (e.g, y < 4 & z < 20).
<code>grouping.var</code>	A single grouping variable (can be entered either as a bare name x or as a string "x").
<code>title.prefix</code>	Character string specifying the prefix text for the fixed plot title (name of each factor level) (Default: NULL). If NULL, the variable name entered for <code>grouping.var</code> will be used.
<code>xlab</code>	Labels for x and y axis variables. If NULL (default), variable names for x and y will be used.
<code>ylab</code>	Labels for x and y axis variables. If NULL (default), variable names for x and y will be used.
<code>method</code>	<p>Smoothing method (function) to use, accepts either a character vector, e.g. "auto", "lm", "glm", "gam", "loess" or a function, e.g. MASS::rlm or mgcv::gam, base::lm, or base::loess.</p> <p>For method = "auto" the smoothing method is chosen based on the size of the largest group (across all panels). <code>loess()</code> is used for less than 1,000 observations; otherwise <code>mgcv::gam()</code> is used with formula = y ~ s(x, bs = "cs"). Somewhat anecdotally, loess gives a better appearance, but is $O(N^2)$ in memory, so does not work for larger datasets.</p> <p>If you have fewer than 1,000 observations but want to use the same <code>gam()</code> model that method = "auto" would use, then set method = "gam", formula = y ~ s(x, bs = "cs").</p>
<code>method.args</code>	List of additional arguments passed on to the modelling function defined by method.
<code>formula</code>	Formula to use in smoothing function, eg. y ~ x, y ~ poly(x, 2), y ~ log(x)
<code>point.color</code>	Aesthetics specifying geom point (defaults: point.color = "black", point.size = 3, point.alpha =
<code>point.size</code>	Aesthetics specifying geom point (defaults: point.color = "black", point.size = 3, point.alpha =
<code>point.alpha</code>	Aesthetics specifying geom point (defaults: point.color = "black", point.size = 3, point.alpha =
<code>line.size</code>	Size for the regression line.
<code>point.width.jitter</code>	Degree of jitter in x and y direction, respectively. Defaults to 0 (0 data).
<code>point.height.jitter</code>	Degree of jitter in x and y direction, respectively. Defaults to 0 (0 data).
<code>line.color</code>	color for the regression line.

marginal	Decides whether <code>ggExtra::ggMarginal()</code> plots will be displayed; the default is TRUE.
marginal.type	Type of marginal distribution to be plotted on the axes ("histogram", "boxplot", "density", "violin", "densigram").
marginal.size	Integer describing the relative size of the marginal plots compared to the main plot. A size of 5 means that the main plot is 5x wider and 5x taller than the marginal plots.
margins	Character describing along which margins to show the plots. Any of the following arguments are accepted: "both", "x", "y".
package	Name of package from which the palette is desired as string or symbol.
palette	Name of palette as string or symbol.
direction	Either 1 or -1. If -1 the palette will be reversed.
xfill	Character describing color fill for x and y axes marginal distributions (default: "#009E73" (for x) and "#D55E00" (for y)). If set to NULL, manual specification of colors will be turned off and 2 colors from the specified palette from package will be selected.
yfill	Character describing color fill for x and y axes marginal distributions (default: "#009E73" (for x) and "#D55E00" (for y)). If set to NULL, manual specification of colors will be turned off and 2 colors from the specified palette from package will be selected.
xalpha	Numeric deciding transparency levels for the marginal distributions. Any numbers from 0 (transparent) to 1 (opaque). The default is 1 for both axes.
yalpha	Numeric deciding transparency levels for the marginal distributions. Any numbers from 0 (transparent) to 1 (opaque). The default is 1 for both axes.
xsize	Size for the marginal distribution boundaries (Default: 0.7).
ysize	Size for the marginal distribution boundaries (Default: 0.7).
centrality.param	Decides <i>which</i> measure of central tendency ("mean" or "median") is to be displayed as vertical (for x) and horizontal (for y) lines.
results.subtitle	Decides whether the results of statistical tests are to be displayed as a subtitle (Default: TRUE). If set to FALSE, only the plot will be returned.
caption	The text for the plot caption.
subtitle	The text for the plot subtitle. Will work only if <code>results.subtitle = FALSE</code> .
nboot	Number of bootstrap samples for computing confidence interval for the effect size (Default: 100).
beta	bending constant (Default: 0.1). For more, see <code>?WRS2::pbcor</code> .
k	Number of digits after decimal point (should be an integer) (Default: k = 2).
axes.range.restrict	Logical that decides whether to restrict the axes values ranges to min and max values of the axes variables (Default: FALSE), only relevant for functions where axes variables are of numeric type.

ggtheme	A function, ggplot2 theme name. Default value is <code>ggplot2::theme_bw()</code> . Any of the ggplot2 themes, or themes from extension packages are allowed (e.g., <code>ggthemes::theme_fivethirtyeight()</code> , <code>hrbrthemes::theme_ipsum_ps()</code> , etc.).
ggstatsplot.layer	Logical that decides whether <code>theme_ggstatsplot</code> theme elements are to be displayed along with the selected <code>ggtheme</code> (Default: <code>TRUE</code>).
ggplot.component	A ggplot component to be added to the plot prepared by <code>ggstatsplot</code> . This argument is primarily helpful for <code>grouped_</code> variant of the current function. Default is <code>NULL</code> . The argument should be entered as a function. If the given function has an argument <code>axes.range.restrict</code> and if it has been set to <code>TRUE</code> , the added ggplot component <i>might</i> not work as expected.
messages	Decides whether messages references, notes, and warnings are to be displayed (Default: <code>TRUE</code>).
...	Arguments passed on to <code>combine_plots</code>
	title.text String or plotmath expression to be drawn as title for the <i>combined plot</i> .
	title.color Text color for title.
	title.size Point size of title text.
	title.vjust Vertical justification for title. Default = 0.5 (centered on y). 0 = baseline at y, 1 = ascender at y.
	title.hjust Horizontal justification for title. Default = 0.5 (centered on x). 0 = flush-left at x, 1 = flush-right.
	title.fontface The font face ("plain", "bold" (default), "italic", "bold.italic") for title.
	caption.text String or plotmath expression to be drawn as the caption for the <i>combined plot</i> .
	caption.color Text color for caption.
	caption.size Point size of title text.
	caption.vjust Vertical justification for caption. Default = 0.5 (centered on y). 0 = baseline at y, 1 = ascender at y.
	caption.hjust Horizontal justification for caption. Default = 0.5 (centered on x). 0 = flush-left at x, 1 = flush-right.
	caption.fontface The font face ("plain" (default), "bold", "italic", "bold.italic") for caption.
	sub.text The label with which the <i>combined plot</i> should be annotated. Can be a plotmath expression.
	sub.color Text color for annotation label (Default: "black").
	sub.size Point size of annotation text (Default: 12).
	sub.x The x position of annotation label (Default: 0.5).
	sub.y The y position of annotation label (Default: 0.5).
	sub.hjust Horizontal justification for annotation label (Default: 0.5).
	sub.vjust Vertical justification for annotation label (Default: 0.5).

sub.vpadding Vertical padding. The total vertical space added to the label, given in grid units. By default, this is added equally above and below the label. However, by changing the `y` and `vjust` parameters, this can be changed (Default: `grid::unit(1, "lines")`).

sub.fontface The font face ("plain" (default), "bold", "italic", "bold.italic") for the annotation label.

sub.angle Angle at which annotation label is to be drawn (Default: 0).

sub.lineheight Line height of annotation label.

title.caption.rel.heights Numerical vector of relative columns heights while combining (title, plot, caption).

title.rel.heights Numerical vector of relative columns heights while combining (title, plot).

caption.rel.heights Numerical vector of relative columns heights while combining (plot, caption).

Author(s)

Indrajeet Patil, Chuck Powell

References

https://indrajeetpatil.github.io/ggstatsplot/articles/web_only/ggscatterstats.html

See Also

[ggscatterstats](#), [ggcorrmatrix](#), [grouped_ggcorrmatrix](#)

Examples

```
## Not run:
# to ensure reproducibility
set.seed(123)

# basic function call
ggstatsplot::grouped_ggscatterstats(
  data = dplyr::filter(
    ggstatsplot::movies_long,
    genre == "Comedy" |
    genre == "Drama"
  ),
  x = length,
  y = rating,
  method = "lm",
  formula = y ~ x + I(x^3),
  grouping.var = genre
)

# using labeling
# (also show how to modify basic plot from within function call)
```

```

ggstatsplot::grouped_ggscatterstats(
  data = dplyr::filter(ggplot2::mpg, cyl != 5),
  x = displ,
  y = hwy,
  grouping.var = cyl,
  title.prefix = "Cylinder count",
  type = "robust",
  label.var = manufacturer,
  label.expression = hwy > 25 & displ > 2.5,
  xfill = NULL,
  ggplot.component = ggplot2::scale_y_continuous(sec.axis = ggplot2::dup_axis()),
  package = "yarr",
  palette = "appletv",
  messages = FALSE
)

# labeling without expression

ggstatsplot::grouped_ggscatterstats(
  data = dplyr::filter(
    .data = ggstatsplot::movies_long,
    rating == 7,
    genre %in% c("Drama", "Comedy")
  ),
  x = budget,
  y = length,
  grouping.var = genre,
  bf.message = TRUE,
  label.var = "title",
  marginal = FALSE,
  title.prefix = "Genre",
  caption.text = "All movies have IMDB rating equal to 7."
)

## End(Not run)

```

intent_morality

Moral judgments about third-party moral behavior.

Description

Moral judgments about third-party moral behavior.

Usage

```
intent_morality
```

Format

A data frame with 4016 rows and 8 variables

- id. Participant id.
- gender. Participant's gender.
- item. Which story/vignette participants read for a given condition.
- harm. What kind of harm was involved in the item.
- belief. What kind of belief the actor had (neutral or negative/harmful).
- outcome. What kind of outcome the actor caused (neutral or negative/harmful).
- condition. Type of harm, composed of belief and outcome.
- question. Type of moral judgment asked (wrongness or punishment).
- rating. Moral judgment rating on a scale of 1 to 7.

Details

This dataset contains data from a recent study about how people judge behavior of others when they unintentionally or intentionally cause harm to others.

Participants responded to four different vignettes that contains four different types of conditions-

- accidental harm. neutral belief, harmful/negative outcome
- intentional harm. harmful/negative belief, harmful/negative outcome
- attempted harm. harmful/negative belief, neutral outcome
- neutral harm. neutral belief, neutral outcome

Additionally, participants saw one of the four variants for each of the four items. Each of the item had a different type of harm.

Source

<https://www.nature.com/articles/s41598-017-05299-9>

Examples

```
dim(intent_morality)
head(intent_morality)
dplyr::glimpse(intent_morality)
```

`iris_long`*Edgar Anderson's Iris Data in long format.*

Description

Edgar Anderson's Iris Data in long format.

Usage

```
iris_long
```

Format

A data frame with 600 rows and 5 variables

- `id`. Dummy identity number for each flower (150 flowers in total).
- `Species`. The species are *Iris setosa*, *versicolor*, and *virginica*.
- `condition`. Factor giving a detailed description of the attribute (Four levels: "Petal.Length", "Petal.Width", "Sepal.Length", "Sepal.Width").
- `attribute`. What attribute is being measured ("Sepal" or "Petal").
- `measure`. What aspect of the attribute is being measured ("Length" or "Width").
- `value`. Value of the measurement.

Details

This famous (Fisher's or Anderson's) iris data set gives the measurements in centimeters of the variables sepal length and width and petal length and width, respectively, for 50 flowers from each of 3 species of iris. The species are *Iris setosa*, *versicolor*, and *virginica*.

This is a modified dataset from `datasets` package.

Source

<https://stat.ethz.ch/R-manual/R-devel/library/datasets/html/iris.html>

Examples

```
dim(iris_long)
head(iris_long)
dplyr::glimpse(iris_long)
```

movies_long	<i>Movie information and user ratings from IMDB.com (long format).</i>
-------------	--

Description

Movie information and user ratings from IMDB.com (long format).

Usage

```
movies_long
```

Format

A data frame with 1,579 rows and 8 variables

- title. Title of the movie.
- year. Year of release.
- budget. Total budget (if known) in US dollars
- length. Length in minutes.
- rating. Average IMDB user rating.
- votes. Number of IMDB users who rated this movie.
- mpaa. MPAA rating.
- genre. Different genres of movies (action, animation, comedy, drama, documentary, romance, short).

Details

Modified dataset from ggplot2movies package.

The internet movie database, <http://imdb.com/>, is a website devoted to collecting movie data supplied by studios and fans. It claims to be the biggest movie database on the web and is run by amazon. More about information imdb.com can be found online, http://imdb.com/help/show_leaf?about, including information about the data collection process, http://imdb.com/help/show_leaf?infosource.

Movies were are identical to those selected for inclusion in movies_wide but this dataset has been constructed such that every movie appears in one and only one genre category.

Source

<https://CRAN.R-project.org/package=ggplot2movies>

Examples

```
dim(movies_long)
head(movies_long)
dplyr::glimpse(movies_long)
```

movies_wide	<i>Movie information and user ratings from IMDB.com (wide format).</i>
-------------	--

Description

Movie information and user ratings from IMDB.com (wide format).

Usage

```
movies_wide
```

Format

A data frame with 1,579 rows and 13 variables

- title. Title of the movie.
- year. Year of release.
- budget. Total budget in millions of US dollars
- length. Length in minutes.
- rating. Average IMDB user rating.
- votes. Number of IMDB users who rated this movie.
- mpaa. MPAA rating.
- action, animation, comedy, drama, documentary, romance, short. Binary variables representing if movie was classified as belonging to that genre.
- NumGenre. The number of different genres a film was classified in an integer between one and four

Details

Modified dataset from ggplot2movies package.

The internet movie database, <http://imdb.com/>, is a website devoted to collecting movie data supplied by studios and fans. It claims to be the biggest movie database on the web and is run by amazon. More information about imdb.com can be found online, http://imdb.com/help/show_leaf?about, including information about the data collection process, http://imdb.com/help/show_leaf?infosource.

Movies were selected for inclusion if they had a known length and had been rated by at least one imdb user. Small categories such as documentaries and NC-17 movies were removed.

Source

<https://CRAN.R-project.org/package=ggplot2movies>

Examples

```
dim(movies_wide)
head(movies_wide)
dplyr::glimpse(movies_wide)
```

normality_message	<i>Display normality test result as a message.</i>
-------------------	--

Description

A note to the user about the validity of assumptions for the default linear model.

Usage

```
normality_message(x, lab = NULL, k = 2, output = "message", ...)
```

Arguments

x	A numeric vector.
lab	A character describing label for the variable. If NULL, a generic "x" label will be used.
k	Number of digits after decimal point (should be an integer) (Default: k = 2).
output	What output is desired: "message" (default) or "stats" (or "tidy") objects.
...	Additional arguments (ignored).

Value

A list with class "htest" containing the following components:

statistic	the value of the Shapiro-Wilk statistic.
p.value	an approximate p-value for the test. This is said in Royston (1995) to be adequate for $p.value < 0.1$.
method	the character string "Shapiro-Wilk normality test".
data.name	a character string giving the name(s) of the data.

Author(s)

Indrajeet Patil

See Also

[ggbetweenstats](#)

Other helper_messages: [bartlett_message](#), [effsize_ci_message](#), [ggcorrmatrix_message](#), [grouped_message](#), [pairwise_p](#), [palette_message](#)

Examples

```
# message
normality_message(
  x = anscombe$x1,
  lab = "x1",
  k = 3
)

# statistical test object
ggstatsplot::normality_message(
  x = anscombe$x2,
  output = "tidy"
)
```

outlier_df	<i>Adding a column to dataframe describing outlier status.</i>
------------	--

Description

This function is mostly helpful for internal operations of some of the functions in this package.

Usage

```
outlier_df(data, x, y, outlier.label, outlier.coef = 1.5, ...)
```

Arguments

data	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
x	The grouping variable from the dataframe data.
y	The response (a.k.a. outcome or dependent) variable from the dataframe data.
outlier.label	Label to put on the outliers that have been tagged.
outlier.coef	Coefficient for outlier detection using Tukey's method. With Tukey's method, outliers are below (1st Quartile) or above (3rd Quartile) outlier.coef times the Inter-Quartile Range (IQR) (Default: 1.5).
...	Additional arguments.

Author(s)

Indrajeet Patil

Examples

```
# adding column for outlier and a label for that outlier
ggstatsplot::outlier_df(
  data = morley,
  x = Expt,
  y = Speed,
  outlier.label = Run,
  outlier.coef = 2
) %>%
  dplyr::arrange(outlier)
```

pairwise_p	<i>Pairwise comparison tests</i>
------------	----------------------------------

Description

Calculate pairwise comparisons between group levels with corrections for multiple testing.

Usage

```
pairwise_p(data, x, y, type = "parametric", tr = 0.1, paired = FALSE,
  var.equal = FALSE, p.adjust.method = "holm", k = 2,
  messages = TRUE, ...)
```

Arguments

data	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
x	The grouping variable from the dataframe data.
y	The response (a.k.a. outcome or dependent) variable from the dataframe data.
type	Type of statistic expected ("parametric" or "nonparametric" or "robust" or "bayes").Corresponding abbreviations are also accepted: "p" (for parametric), "np" (nonparametric), "r" (robust), or "bf"resp.
tr	Trim level for the mean when carrying out robust tests. If you get error stating "Standard error cannot be computed because of Winsorized variance of 0 (e.g., due to ties). Try to decrease the trimming level.", try to play around with the value of tr, which is by default set to 0.1. Lowering the value might help.
paired	a logical indicating whether you want a paired t-test.
var.equal	a logical variable indicating whether to treat the variances in the samples as equal. If TRUE, then a simple F test for the equality of means in a one-way analysis of variance is performed. If FALSE, an approximate method of Welch (1951) is used, which generalizes the commonly known 2-sample Welch test to the case of arbitrarily many samples.

<code>p.adjust.method</code>	Adjustment method for p -values for multiple comparisons. Possible methods are: "holm" (default), "hochberg", "hommel", "bonferroni", "BH", "BY", "fdr", "none".
<code>k</code>	Number of digits after decimal point (should be an integer) (Default: <code>k = 2</code>).
<code>messages</code>	Decides whether messages references, notes, and warnings are to be displayed (Default: <code>TRUE</code>).
<code>...</code>	Additional arguments.

Author(s)

Indrajeet Patil

See Also[ggbetweenstats](#), [grouped_ggbetweenstats](#)Other helper_messages: [bartlett_message](#), [effsize_ci_message](#), [ggcorrmatrix_message](#), [grouped_message](#), [normality_message](#), [palette_message](#)**Examples**

```
# time consuming, so not run on `CRAN` machines
## Not run:
# show all columns in a tibble
options(tibble.width = Inf)

# for reproducibility
set.seed(123)

#----- between-subjects design -----

# parametric
# if `var.equal = TRUE`, then Student's *t*-test will be run
ggstatsplot::pairwise_p(
  data = ggplot2::msleep,
  x = vore,
  y = brainwt,
  type = "p",
  var.equal = TRUE,
  paired = FALSE,
  p.adjust.method = "bonferroni"
)

# if `var.equal = FALSE`, then Games-Howell test will be run
ggstatsplot::pairwise_p(
  data = ggplot2::msleep,
  x = vore,
  y = brainwt,
  type = "p",
  var.equal = FALSE,
```

```

    paired = FALSE,
    p.adjust.method = "bonferroni"
  )

# non-parametric
ggstatsplot::pairwise_p(
  data = ggplot2::msleep,
  x = vore,
  y = brainwt,
  type = "np",
  paired = FALSE,
  p.adjust.method = "none"
)

# robust
ggstatsplot::pairwise_p(
  data = ggplot2::msleep,
  x = vore,
  y = brainwt,
  type = "r",
  paired = FALSE,
  p.adjust.method = "fdr"
)

## End(Not run)
#----- within-subjects design -----

set.seed(123)
library(jmv)
data("bugs", package = "jmv")

# converting to long format
bugs_long <- bugs %>%
  tibble::as_tibble(.) %>%
  tidyr::gather(., key, value, LDLF:HDHF)

# parametric
ggstatsplot::pairwise_p(
  data = bugs_long,
  x = key,
  y = value,
  type = "p",
  paired = TRUE,
  p.adjust.method = "BH"
)

# non-parametric
ggstatsplot::pairwise_p(
  data = bugs_long,
  x = key,
  y = value,
  type = "np",
  paired = TRUE,

```

```

    p.adjust.method = "BY"
  )

  # robust
  ggstatsplot::pairwise_p(
    data = bugs_long,
    x = key,
    y = value,
    type = "r",
    paired = TRUE,
    p.adjust.method = "hommel"
  )

```

subtitle_anova_bayes *Making text subtitle for the between-subject one-way anova designs.*

Description

Making text subtitle for the between-subject one-way anova designs.

Usage

```

subtitle_anova_bayes(data, x, y, effsize.type = "unbiased",
  partial = TRUE, var.equal = FALSE, bf.prior = 0.707,
  paired = FALSE, k = 2, ...)

```

Arguments

data	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
x	The grouping variable from the dataframe data.
y	The response (a.k.a. outcome or dependent) variable from the dataframe data.
effsize.type	Type of effect size needed for <i>parametric</i> tests. The argument can be "biased" ("d" for Cohen's <i>d</i> for t-test ; "partial_eta" for partial eta-squared for anova) or "unbiased" ("g" Hedge's <i>g</i> for t-test ; "partial_omega" for partial omega-squared for anova)).
partial	Logical that decides if partial eta-squared or omega-squared are returned (Default: TRUE). If FALSE, eta-squared or omega-squared will be returned. Valid only for objects of class lm, aov, anova, or aovlist.
var.equal	a logical variable indicating whether to treat the variances in the samples as equal. If TRUE, then a simple F test for the equality of means in a one-way analysis of variance is performed. If FALSE, an approximate method of Welch (1951) is used, which generalizes the commonly known 2-sample Welch test to the case of arbitrarily many samples.
bf.prior	A number between 0.5 and 2 (default 0.707), the prior width to use in calculating Bayes factors.

paired	a logical indicating whether you want a paired t-test.
k	Number of digits after decimal point (should be an integer) (Default: k = 2).
...	Additional arguments.

Author(s)

Indrajeet Patil

Examples

```
## Not run:
# with defaults
subtitle_anova_bayes(
  data = ggplot2::msleep,
  x = vore,
  y = sleep_rem,
  k = 2,
  bf.prior = 0.8
)

# modifying the defaults
subtitle_anova_bayes(
  data = ggplot2::msleep,
  x = vore,
  y = sleep_rem,
  effsize.type = "partial_eta",
  var.equal = TRUE
)

## End(Not run)
```

subtitle_anova_parametric

Making text subtitle for the between-subject anova designs.

Description

Making text subtitle for the between-subject anova designs.

Usage

```
subtitle_anova_parametric(data, x, y, effsize.type = "unbiased",
  partial = TRUE, conf.level = 0.95, nboot = 100,
  var.equal = FALSE, k = 2, messages = TRUE, ...)
```

Arguments

<code>data</code>	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
<code>x</code>	The grouping variable from the dataframe <code>data</code> .
<code>y</code>	The response (a.k.a. outcome or dependent) variable from the dataframe <code>data</code> .
<code>effsize.type</code>	Type of effect size needed for <i>parametric</i> tests. The argument can be "biased" ("d" for Cohen's <i>d</i> for t-test ; "partial_eta" for partial eta-squared for anova) or "unbiased" ("g" Hedge's <i>g</i> for t-test ; "partial_omega" for partial omega-squared for anova)).
<code>partial</code>	Logical that decides if partial eta-squared or omega-squared are returned (Default: TRUE). If FALSE, eta-squared or omega-squared will be returned. Valid only for objects of class <code>lm</code> , <code>aov</code> , <code>anova</code> , or <code>aovlist</code> .
<code>conf.level</code>	Scalar between 0 and 1. If unspecified, the defaults return 95% lower and upper confidence intervals (0.95).
<code>nboot</code>	Number of bootstrap samples for computing confidence interval for the effect size (Default: 100).
<code>var.equal</code>	a logical variable indicating whether to treat the variances in the samples as equal. If TRUE, then a simple F test for the equality of means in a one-way analysis of variance is performed. If FALSE, an approximate method of Welch (1951) is used, which generalizes the commonly known 2-sample Welch test to the case of arbitrarily many samples.
<code>k</code>	Number of digits after decimal point (should be an integer) (Default: <code>k = 2</code>).
<code>messages</code>	Decides whether messages references, notes, and warnings are to be displayed (Default: TRUE).
<code>...</code>	Additional arguments.

Author(s)

Indrajeet Patil

Examples

```
# with defaults
subtitle_anova_parametric(
  data = ggplot2::msleep,
  x = vore,
  y = sleep_rem,
  k = 3
)

# modifying the defaults
subtitle_anova_parametric(
  data = ggplot2::msleep,
  x = vore,
  y = sleep_rem,
  effsize.type = "biased",
  partial = FALSE,
```

```
var.equal = TRUE,
nboot = 10
)
```

subtitle_anova_robust *Making text subtitle for the robust ANOVA (between-subjects designs).*

Description

Making text subtitle for the robust ANOVA (between-subjects designs).

Usage

```
subtitle_anova_robust(data, x, y, tr = 0.1, nboot = 100,
  conf.level = 0.95, conf.type = "norm", messages = TRUE, k = 2,
  ...)
```

Arguments

data	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
x	The grouping variable from the dataframe data.
y	The response (a.k.a. outcome or dependent) variable from the dataframe data.
tr	Trim level for the mean when carrying out robust tests. If you get error stating "Standard error cannot be computed because of Winsorized variance of 0 (e.g., due to ties). Try to decrease the trimming level.", try to play around with the value of tr, which is by default set to 0.1. Lowering the value might help.
nboot	Number of bootstrap samples for computing confidence interval for the effect size (Default: 100).
conf.level	Scalar between 0 and 1. If unspecified, the defaults return 95% lower and upper confidence intervals (0.95).
conf.type	A vector of character strings representing the type of intervals required. The value should be any subset of the values "norm", "basic", "perc", "bca". For more, see ?boot::boot.ci.
messages	Decides whether messages references, notes, and warnings are to be displayed (Default: TRUE).
k	Number of digits after decimal point (should be an integer) (Default: k = 2).
...	Additional arguments.

Author(s)

Indrajeet Patil

Examples

```
# examples not executed due to time constraints
## Not run:
# for reproducibility
set.seed(123)

# going with the defaults
subtitle_anova_robust(
  data = ggplot2::midwest,
  x = state,
  y = percbelowpoverty,
  nboot = 10
)

# changing defaults
subtitle_anova_robust(
  data = ggplot2::midwest,
  x = state,
  y = percollege,
  tr = 0.2,
  nboot = 10
)

## End(Not run)
```

subtitle_contingency_tab

Making text subtitle for contingency analysis (Pearson's chi-square test for independence for between-subjects design or McNemar's test for within-subjects design)

Description

Making text subtitle for contingency analysis (Pearson's chi-square test for independence for between-subjects design or McNemar's test for within-subjects design)

Usage

```
subtitle_contingency_tab(data, main, condition, counts = NULL,
  nboot = 100, paired = FALSE, stat.title = NULL,
  conf.level = 0.95, conf.type = "norm", simulate.p.value = FALSE,
  B = 2000, bias.correct = FALSE, k = 2, messages = TRUE, ...)
```

Arguments

data	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
------	---

main	The variable to use as the rows in the contingency table.
condition	The variable to use as the columns in the contingency table.
counts	A string naming a variable in data containing counts, or NULL if each row represents a single observation (Default).
nboot	Number of bootstrap samples for computing confidence interval for the effect size (Default: 100).
paired	Logical indicating whether data came from a within-subjects design study (Default: FALSE). If TRUE, McNemar test subtitle will be returned. If FALSE, Pearson's chi-square test will be returned.
stat.title	Title for the effect being investigated with the chi-square test. The default is NULL, i.e. no title will be added to describe the effect being shown. An example of a stat.title argument will be something like "main x condition" or "interaction".
conf.level	Scalar between 0 and 1. If unspecified, the defaults return 95% lower and upper confidence intervals (0.95).
conf.type	A vector of character strings representing the type of intervals required. The value should be any subset of the values "norm", "basic", "perc", "bca". For more, see ?boot::boot.ci.
simulate.p.value	a logical indicating whether to compute p-values by Monte Carlo simulation.
B	an integer specifying the number of replicates used in the Monte Carlo test.
bias.correct	If TRUE, a bias correction will be applied to Cramer's V.
k	Number of digits after decimal point (should be an integer) (Default: k = 2).
messages	Decides whether messages references, notes, and warnings are to be displayed (Default: TRUE).
...	Additional arguments (currently ignored).

Author(s)

Indrajeet Patil

See Also

[ggpiestats](#)

Examples

```
# without counts data
subtitle_contingency_tab(
  data = mtcars,
  main = am,
  condition = cyl,
  nboot = 15
)
```

```
# with counts data
# in case of no variation, a `NULL` will be returned.
library(jmv)

as.data.frame(HairEyeColor) %>%
  dplyr::filter(.data = ., Sex == "Male") %>%
  subtitle_contingency_tab(
    data = .,
    main = Hair,
    condition = Sex,
    counts = Freq
  )
```

```
subtitle_friedman_nonparametric
```

Making text subtitle for the Friedman Rank Sum Test (nonparametric ANOVA) (within-subjects designs).

Description

Making text subtitle for the Friedman Rank Sum Test (nonparametric ANOVA) (within-subjects designs).

Usage

```
subtitle_friedman_nonparametric(data, x, y, messages = TRUE, k = 2,
  ...)
```

Arguments

data	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
x	The grouping variable from the dataframe data.
y	The response (a.k.a. outcome or dependent) variable from the dataframe data.
messages	Decides whether messages references, notes, and warnings are to be displayed (Default: TRUE).
k	Number of digits after decimal point (should be an integer) (Default: k = 2).
...	Additional arguments (ignored).

Author(s)

Indrajeet Patil

Examples

```
# setup
set.seed(123)
library(ggstatsplot)
library(jmv)
data("bugs", package = "jmv")

# converting to long format
data_bugs <- bugs %>%
  tibble::as_tibble(.) %>%
  tidyr::gather(., key, value, LDLF:HDHF)

# creating the subtitle
ggstatsplot::subtitle_friedman_nonparametric(
  data = data_bugs,
  x = key,
  y = value,
  k = 2
)
```

subtitle_ggscatterstats

Making text subtitle for the correlation test.

Description

Making text subtitle for the correlation test.

Usage

```
subtitle_ggscatterstats(data, x, y, nboot = 100, beta = 0.1,
  type = "pearson", bf.prior = 0.707, conf.level = 0.95,
  conf.type = "norm", messages = TRUE, k = 2, ...)
```

Arguments

data	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
x	A vector containing the explanatory variable.
y	The response - a vector of length the number of rows of x.
nboot	Number of bootstrap samples for computing confidence interval for the effect size (Default: 100).
beta	bending constant (Default: 0.1). For more, see ?WRS2::pbcor.
type	Type of association between paired samples required ("parametric": Pearson's product moment correlation coefficient" or "nonparametric": Spearman's rho" or "robust": percentage bend correlation coefficient" or "bayes":

	Bayes Factor for Pearson's r "). Corresponding abbreviations are also accepted: "p" (for parametric/pearson's), "np" (nonparametric/spearman), "r" (robust), "bf" (for bayes factor), resp.
bf.prior	A number between 0.5 and 2 (default 0.707), the prior width to use in calculating Bayes factors.
conf.level	Scalar between 0 and 1. If unspecified, the defaults return 95% lower and upper confidence intervals (0.95).
conf.type	A vector of character strings representing the type of intervals required. The value should be any subset of the values "norm", "basic", "perc", "bca". For more, see ?boot::boot.ci.
messages	Decides whether messages references, notes, and warnings are to be displayed (Default: TRUE).
k	Number of digits after decimal point (should be an integer) (Default: k = 2).
...	further arguments to be passed to or from methods.

Author(s)

Indrajeet Patil

Examples

```
# without changing defaults
subtitle_ggscatterstats(
  data = ggplot2::midwest,
  x = area,
  y = percblack
)

# changing defaults
subtitle_ggscatterstats(
  data = ggplot2::midwest,
  x = area,
  y = percblack,
  nboot = 25,
  beta = 0.2,
  type = "r",
  k = 1
)
```

 subtitle_kw_nonparametric

Making text subtitle for the Kruskal-Wallis test (nonparametric ANOVA) (between-subjects designs).

Description

Making text subtitle for the Kruskal-Wallis test (nonparametric ANOVA) (between-subjects designs).

Usage

```
subtitle_kw_nonparametric(data, x, y, messages = TRUE, k = 2,
  nboot = 100, conf.level = 0.95, conf.type = "norm", ...)
```

Arguments

data	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
x	The grouping variable from the dataframe data.
y	The response (a.k.a. outcome or dependent) variable from the dataframe data.
messages	Decides whether messages references, notes, and warnings are to be displayed (Default: TRUE).
k	Number of digits after decimal point (should be an integer) (Default: k = 2).
nboot	Number of bootstrap samples for computing confidence interval for the effect size (Default: 100).
conf.level	Scalar between 0 and 1. If unspecified, the defaults return 95% lower and upper confidence intervals (0.95).
conf.type	A vector of character strings representing the type of intervals required. The value should be any subset of the values "norm", "basic", "perc", "bca". For more, see ?boot::boot.ci.
...	Additional arguments (ignored).

Author(s)

Indrajeet Patil

Examples

```
subtitle_kw_nonparametric(
  data = ggplot2::msleep,
  x = vore,
  y = sleep_rem
)
```

```
subtitle_mann_nonparametric
```

Making text subtitle for the Mann-Whitney U-test (between-subjects designs).

Description

Making text subtitle for the Mann-Whitney *U*-test (between-subjects designs).

Usage

```
subtitle_mann_nonparametric(data, x, y, paired = FALSE, k = 2,
  conf.level = 0.95, conf.type = "norm", nboot = 100,
  messages = TRUE, ...)
```

```
subtitle_t_nonparametric(data, x, y, paired = FALSE, k = 2,
  conf.level = 0.95, conf.type = "norm", nboot = 100,
  messages = TRUE, ...)
```

Arguments

<code>data</code>	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
<code>x</code>	The grouping variable from the dataframe data.
<code>y</code>	The response (a.k.a. outcome or dependent) variable from the dataframe data.
<code>paired</code>	a logical indicating whether you want a paired t-test.
<code>k</code>	Number of digits after decimal point (should be an integer) (Default: <code>k = 2</code>).
<code>conf.level</code>	Scalar between 0 and 1. If unspecified, the defaults return 95% lower and upper confidence intervals (0.95).
<code>conf.type</code>	A vector of character strings representing the type of intervals required. The value should be any subset of the values "norm", "basic", "perc", "bca". For more, see <code>?boot::boot.ci</code> .
<code>nboot</code>	Number of bootstrap samples for computing confidence interval for the effect size (Default: 100).
<code>messages</code>	Decides whether messages references, notes, and warnings are to be displayed (Default: TRUE).
<code>...</code>	Additional arguments.

Details

Two-sample Wilcoxon test, also known as Mann-Whitney test, is carried out.

For the two independent samples case, the Mann-Whitney *U*-test is calculated and *W* is reported from `stats::wilcox.test`. For the paired samples case the Wilcoxon signed rank test is run and *V* is reported.

Since there is no single commonly accepted method for reporting effect size for these tests we are computing and reporting r (computed as Z/\sqrt{N}) along with the confidence intervals associated with the estimate.

Note: The `stats::wilcox.test` function does not follow the same convention as `stats::t.test`. The sign of the V test statistic will always be positive since it is **the sum of the positive signed ranks**. Therefore V will vary in magnitude but not significance based solely on the order of the grouping variable. Consider manually reordering your factor levels if appropriate as shown in the second example below.

Author(s)

Indrajeet Patil, Chuck Powell

Examples

```
## Not run:
set.seed(123)

# ----- between-subjects design -----
# simple function call
ggstatsplot::subtitle_mann_nonparametric(
  data = sleep,
  x = group,
  y = extra
)

# creating a smaller dataset
msleep_short <- dplyr::filter(
  .data = ggplot2::msleep,
  vore %in% c("carni", "herbi")
)

# modifying few things
ggstatsplot::subtitle_mann_nonparametric(
  data = msleep_short,
  x = vore,
  y = sleep_rem,
  nboot = 200,
  conf.level = 0.99,
  conf.type = "bca"
)

# The order of the grouping factor matters when computing *V*
# Changing default alphabeical order manually
msleep_short$vore <- factor(msleep_short$vore,
  levels = c("herbi", "carni")
)

# note the change in the reported *V* value but the identical
# value for *p* and the reversed effect size
ggstatsplot::subtitle_mann_nonparametric(
  data = msleep_short,
```

```

    x = vore,
    y = sleep_rem
  )

# ----- within-subjects design -----
# using dataset included in the package
ggstatsplot::subtitle_mann_nonparametric(
  data = VR_dilemma,
  x = modality,
  y = score,
  paired = TRUE,
  conf.level = 0.90,
  conf.type = "perc",
  nboot = 200,
  k = 5
)

## End(Not run)

```

```
subtitle_meta_ggcoefstats
```

Prepare subtitle with meta-analysis results

Description

Making text subtitle for meta-analysis via linear (mixed-effects) models as implemented in the metafor package.

Usage

```

subtitle_meta_ggcoefstats(data, k = 2, messages = TRUE,
  output = "subtitle", caption = NULL, ...)

```

Arguments

data	A dataframe. It must contain columns named estimate (corresponding estimates of coefficients or other quantities of interest) and std.error (the standard error of the regression term).
k	Number of digits after decimal point (should be an integer) (Default: k = 2).
messages	Decides whether messages references, notes, and warnings are to be displayed (Default: TRUE).
output	Character describing the desired output. If "subtitle", a formatted subtitle with summary effect and statistical details will be returned, and if "caption", expression containing details from model summary will be returned. The other option is to return "tidy" data frame with coefficients or "glance" dataframe with model summaries.
caption	The text for the plot caption.
...	Additional arguments (ignored).

Author(s)

Indrajeet Patil

Examples

```

# let's create a dataframe
df_results <-
  structure(
    .Data = list(estimate = c(
      0.382047603321706, 0.780783111514665,
      0.425607573765058, 0.558365541235078, 0.956473848429961
    ), std.error = c(
      0.0465576338644502,
      0.0330218199731529, 0.0362834986178494, 0.0480571500648261, 0.062215818388157
    ), t.value = c(
      8.20590677855356, 23.6444603038067, 11.7300588415607,
      11.6187818146078, 15.3734833553524
    ), conf.low = c(
      0.290515146096969,
      0.715841986960399, 0.354354575031406, 0.46379116008131, 0.827446138277154
    ), conf.high = c(
      0.473580060546444, 0.845724236068931, 0.496860572498711,
      0.652939922388847, 1.08550155858277
    ), p.value.x = c(
      3.28679518728519e-15,
      4.04778497135963e-75, 7.59757330804449e-29, 5.45155840151592e-26,
      2.99171217913312e-13
    ), df.residual = c(
      394L, 358L, 622L, 298L,
      22L
    )),
    row.names = c(NA, -5L),
    class = c("tbl_df", "tbl", "data.frame")
  )

# making subtitle
ggstatsplot::subtitle_meta_ggcoefstats(
  data = df_results,
  k = 3,
  messages = FALSE
)

# getting tidy data frame with coefficients
ggstatsplot::subtitle_meta_ggcoefstats(
  data = df_results,
  messages = FALSE,
  output = "tidy"
)

# making caption
ggstatsplot::subtitle_meta_ggcoefstats(
  data = df_results,

```

```

    k = 2,
    messages = FALSE,
    output = "caption"
  )

  # getting dataframe with model summary
  ggstatsplot::subtitle_meta_ggcoefstats(
    data = df_results,
    messages = FALSE,
    output = "glance"
  )

```

subtitle_onesample_proptest

Making text subtitle for Proportion Test (N Outcomes)

Description

This is going to be a chi-squared Goodness of fit test.

Usage

```

subtitle_onesample_proptest(data, main, counts = NULL, ratio = NULL,
  legend.title = NULL, k = 2, ...)

```

Arguments

data	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
main	The variable to use as the rows in the contingency table.
counts	A string naming a variable in data containing counts, or NULL if each row represents a single observation (Default).
ratio	A vector of numbers: the expected proportions for the proportion test. Default is NULL, which means if there are two levels ratio = c(1,1), etc.
legend.title	Title text for the legend.
k	Number of digits after decimal point (should be an integer) (Default: k = 2).
...	Additional arguments (currently ignored).

Author(s)

Indrajeet Patil

Examples

```
# with counts
library(jmv)

subtitle_onesample_proptest(
  data = as.data.frame(HairEyeColor),
  main = Eye,
  counts = Freq
)

# in case no variation, only sample size will be shown
subtitle_onesample_proptest(
  data = cbind.data.frame(x = rep("a", 10)),
  main = x
)
```

subtitle_template	<i>Template for subtitles with statistical details for tests with a single parameter (e.g., t, chi-squared, etc.)</i>
-------------------	---

Description

Template for subtitles with statistical details for tests with a single parameter (e.g., t, chi-squared, etc.)

Usage

```
subtitle_template(no.parameters, stat.title = NULL, statistic.text,
  statistic, parameter = NULL, parameter2 = NULL, p.value,
  effsize.text, effsize.estimate, effsize.LL, effsize.UL, n,
  conf.level = 0.95, k = 3L, k.parameter = 0L)
```

Arguments

no.parameters	An integer that specifies that the number of parameters for the statistical test. Can be 0 for non-parametric tests, 1 for tests based on <i>t</i> -statistic or chi-squared statistic, 2 for tests based on <i>F</i> -statistic.
stat.title	A character describing the test being run, which will be added as a prefix in the subtitle. The default is NULL. An example of a stat.title argument will be something like "Student's t-test: ".
statistic.text	A character that specifies the relevant test statistic. For example, for tests with <i>t</i> -statistic, statistic.text = "t". If you want to use plotmath, you will have to quote the argument (e.g., quote(italic("t"))).
statistic	The numeric value of a statistic.
parameter	The numeric value of a parameter being modeled (often degrees of freedom for the test). Default is NULL to accommodate non-parametric tests.

parameter2	Relevant only if the statistic in question has two degrees of freedom (default: NULL).
p.value	The two-sided p-value associated with the observed statistic.
effsize.text	A character that specifies the relevant effect size. For example, for Cohen's <i>d</i> statistic, <code>effsize.text = "d"</code> . If you want to use <code>plotmath</code> , you will have to quote the argument (e.g., <code>quote(italic("d"))</code>).
effsize.estimate, effsize.LL, effsize.UL	The estimated value of the effect size, its lower bound, and its upper.
n	An integer specifying the sample size used for the test.
conf.level	Scalar between 0 and 1. If unspecified, the defaults return 95% lower and upper confidence intervals (0.95).
k	Number of decimal places to display (default: 3).
k.parameter	Number of decimal places to display for the parameter (default: 0).

Author(s)

Indrajeet Patil

Examples

```
set.seed(123)

# subtitle for *t*-statistic with Cohen's *d* as effect size
ggstatsplot::subtitle_template(
  no.parameters = 1L,
  statistic.text = quote(italic("t")),
  statistic = 5.494,
  parameter = 29.234,
  p.value = 0.00001,
  effsize.text = quote(italic("d")),
  effsize.estimate = -1.980,
  effsize.LL = -2.873,
  effsize.UL = -1.088,
  n = 32L,
  conf.level = 0.95,
  k = 3L,
  k.parameter = 3L
)
```

subtitle_t_bayes

Making text subtitle for the bayesian t-test.

Description

Making text subtitle for the bayesian t-test.

Usage

```
subtitle_t_bayes(data, x, y, bf.prior = 0.707, paired = FALSE, k = 2,
  ...)
```

Arguments

data	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
x	The grouping variable from the dataframe data.
y	The response (a.k.a. outcome or dependent) variable from the dataframe data.
bf.prior	A number between 0.5 and 2 (default 0.707), the prior width to use in calculating Bayes factors.
paired	a logical indicating whether you want a paired t-test.
k	Number of digits after decimal point (should be an integer) (Default: k = 2).
...	Additional arguments.

Author(s)

Indrajeet Patil

Examples

```
# for reproducibility
set.seed(123)

# between-subjects design

subtitle_t_bayes(
  data = mtcars,
  x = am,
  y = wt,
  paired = FALSE
)

# within-subjects design

subtitle_t_bayes(
  data = dplyr::filter(
    ggstatsplot::intent_morality,
    condition %in% c("accidental", "attempted"),
    harm == "Poisoning"
  ),
  x = condition,
  y = rating,
  paired = TRUE
)
```

subtitle_t_onesample *Making text subtitle for one sample t-test and its nonparametric and robust equivalents.*

Description

Making text subtitle for one sample t-test and its nonparametric and robust equivalents.

Usage

```
subtitle_t_onesample(data, x, type = "parametric", test.value = 0,
  bf.prior = 0.707, robust.estimator = "onestep", effsize.type = "g",
  effsize.noncentral = TRUE, conf.level = 0.95, conf.type = "norm",
  nboot = 100, k = 2, messages = TRUE, ...)
```

Arguments

data	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
x	A numeric variable.
type	Type of statistic expected ("parametric" or "nonparametric" or "robust" or "bayes").Corresponding abbreviations are also accepted: "p" (for parametric), "np" (nonparametric), "r" (robust), or "bf"resp.
test.value	A number specifying the value of the null hypothesis (Default: 0).
bf.prior	A number between 0.5 and 2 (default 0.707), the prior width to use in calculating Bayes factors.
robust.estimator	If test = "robust" robust estimator to be used ("onestep" (Default), "mom", or "median"). For more, see ?WRS2::onesampb.
effsize.type	Type of effect size needed for <i>parametric</i> tests. The argument can be "biased" ("d" for Cohen's <i>d</i>) or "unbiased" ("g" Hedge's <i>g</i> for t-test). The default is
effsize.noncentral	Logical indicating whether to use non-central <i>t</i> -distributions for computing the confidence interval for Cohen's <i>d</i> or Hedge's <i>g</i> (Default: TRUE).
conf.level	Scalar between 0 and 1. If unspecified, the defaults return 95% lower and upper confidence intervals (0.95).
conf.type	A vector of character strings representing the type of intervals required. The value should be any subset of the values "norm", "basic", "perc", "bca". For more, see ?boot::boot.ci.
nboot	Number of bootstrap samples for computing confidence interval for the effect size (Default: 100).
k	Number of digits after decimal point (should be an integer) (Default: k = 2).
messages	Decides whether messages references, notes, and warnings are to be displayed (Default: TRUE).
...	Additional arguments.

Author(s)

Indrajeet Patil

See Also[gghistostats](#)**Examples**

```
# for reproducibility
set.seed(123)

ggstatsplot::subtitle_t_onesample(
  data = iris,
  x = Sepal.Length,
  test.value = 5,
  type = "r"
)
```

subtitle_t_parametric *Making text subtitle for the t-test (between-/within-subjects designs).*

Description

Making text subtitle for the t-test (between-/within-subjects designs).

Usage

```
subtitle_t_parametric(data, x, y, paired = FALSE, effsize.type = "g",
  effsize.noncentral = TRUE, conf.level = 0.95, var.equal = FALSE,
  k = 2, ...)
```

Arguments

data	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
x	The grouping variable from the dataframe data.
y	The response (a.k.a. outcome or dependent) variable from the dataframe data.
paired	a logical indicating whether you want a paired t-test.
effsize.type	Type of effect size needed for <i>parametric</i> tests. The argument can be "biased" ("d" for Cohen's <i>d</i> for t-test ; "partial_eta" for partial eta-squared for anova) or "unbiased" ("g" Hedge's <i>g</i> for t-test ; "partial_omega" for partial omega-squared for anova)).
effsize.noncentral	Logical indicating whether to use non-central <i>t</i> -distributions for computing the confidence interval for Cohen's <i>d</i> or Hedge's <i>g</i> (Default: TRUE).

<code>conf.level</code>	Scalar between 0 and 1. If unspecified, the defaults return 95% lower and upper confidence intervals (0.95).
<code>var.equal</code>	a logical variable indicating whether to treat the variances in the samples as equal. If TRUE, then a simple F test for the equality of means in a one-way analysis of variance is performed. If FALSE, an approximate method of Welch (1951) is used, which generalizes the commonly known 2-sample Welch test to the case of arbitrarily many samples.
<code>k</code>	Number of digits after decimal point (should be an integer) (Default: <code>k = 2</code>).
<code>...</code>	Additional arguments.

Details

Cohen's d is calculated in the traditional fashion as the difference between means or mean minus μ divided by the estimated standardized deviation. By default Hedge's correction is applied $(N-3)/(N-2.25)$ to produce g . For independent samples t -test, there are two possibilities implemented. If the t -test did not make a homogeneity of variance assumption, (the Welch test), the variance term will mirror the Welch test, otherwise a pooled and weighted estimate is used. If a paired samples t -test was requested, then effect size desired is based on the standard deviation of the differences.

The computation of the confidence intervals defaults to a use of non-central Student- t distributions (`effsize.noncentral = TRUE`); otherwise a central distribution is used.

When computing confidence intervals the variance of the effect size d or g is computed using the conversion formula reported in Cooper et al. (2009)

- $((n1+n2)/(n1*n2) + .5*d^2/df) * ((n1+n2)/df)$ (independent samples)
- $\text{sqrt}(((1 / n) + (d^2 / n)) * 2 * (1 - r))$ (paired case)

Author(s)

Indrajeet Patil, Chuck Powell

See Also

`subtitle_t_parametric`

Examples

```
# creating a smaller dataset
msleep_short <- dplyr::filter(
  .data = ggplot2::msleep,
  vore %in% c("carni", "herbi")
)

# with defaults
subtitle_t_parametric(
  data = msleep_short,
  x = vore,
  y = sleep_rem
)
```

```
# changing defaults
subtitle_t_parametric(
  data = msleep_short,
  x = vore,
  y = sleep_rem,
  var.equal = TRUE,
  k = 2,
  effsize.type = "d"
)
```

subtitle_t_robust	<i>Making text subtitle for the robust t-test (between- and within-subjects designs).</i>
-------------------	---

Description

Making text subtitle for the robust t-test (between- and within-subjects designs).

Usage

```
subtitle_t_robust(data, x, y, tr = 0.1, paired = FALSE, nboot = 100,
  conf.level = 0.95, conf.type = "norm", k = 2, messages = TRUE,
  ...)
```

Arguments

data	A dataframe (or a tibble) from which variables specified are to be taken. A matrix or tables will not be accepted.
x	The grouping variable from the dataframe data.
y	The response (a.k.a. outcome or dependent) variable from the dataframe data.
tr	Trim level for the mean when carrying out robust tests. If you get error stating "Standard error cannot be computed because of Winsorized variance of 0 (e.g., due to ties). Try to decrease the trimming level.", try to play around with the value of tr, which is by default set to 0.1. Lowering the value might help.
paired	a logical indicating whether you want a paired t-test.
nboot	Number of bootstrap samples for computing confidence interval for the effect size (Default: 100).
conf.level	Scalar between 0 and 1. If unspecified, the defaults return 95% lower and upper confidence intervals (0.95).
conf.type	A vector of character strings representing the type of intervals required. The value should be any subset of the values "norm", "basic", "perc", "bca". For more, see ?boot::boot.ci.
k	Number of digits after decimal point (should be an integer) (Default: k = 2).
messages	Decides whether messages references, notes, and warnings are to be displayed (Default: TRUE).
...	Additional arguments.

Author(s)

Indrajeet Patil

Examples

```
# with defaults
subtitle_t_robust(
  data = sleep,
  x = group,
  y = extra
)

# changing defaults
subtitle_t_robust(
  data = ToothGrowth,
  x = supp,
  y = len,
  nboot = 10,
  k = 1,
  tr = 0.2
)

# within-subjects design
ggstatsplot::subtitle_t_robust(
  data = dplyr::filter(
    ggstatsplot::intent_morality,
    condition %in% c("accidental", "attempted"),
    harm == "Poisoning"
  ),
  x = condition,
  y = rating,
  paired = TRUE,
  nboot = 25
)
```

theme_corrmat

*Default theme used for correlation matrix***Description**

Default theme used for correlation matrix

Usage

theme_corrmat()

Value

A ggplot2 object with the theme_ggstatsplot overlaid.

Author(s)

Indrajeet Patil

theme_ggstatsplot	<i>Default theme used in all ggstatsplot package plots</i>
-------------------	--

Description

Common theme used across all plots generated in ggstatsplot and *assumed* by the author to be aesthetically pleasing to the user/reader.

Usage

```
theme_ggstatsplot(ggtheme = ggplot2::theme_bw(),  
  ggstatsplot.layer = TRUE)  
  
theme_mprl(ggtheme = ggplot2::theme_bw(), ggstatsplot.layer = TRUE)
```

Arguments

ggtheme	A function, ggplot2 theme name. Default value is ggplot2::theme_bw(). Any of the ggplot2 themes, or themes from extension packages are allowed (e.g., ggthemes::theme_fivethirtyeight(), hrbrthemes::theme_ipsum_ps(), etc.).
ggstatsplot.layer	Logical that decides whether theme_ggstatsplot theme elements are to be displayed along with the selected ggtheme (Default: TRUE).

Value

A ggplot2 object with the theme_ggstatsplot theme.

Author(s)

Indrajeet Patil

References

https://indrajeetpatil.github.io/ggstatsplot/articles/web_only/theme_ggstatsplot.html

theme_pie	<i>Default theme used for pie chart</i>
-----------	---

Description

Default theme used for pie chart

Usage

```
theme_pie(ggtheme = ggplot2::theme_bw(), ggstatsplot.layer = TRUE)
```

Arguments

- ggtheme A function, ggplot2 theme name. Default value is ggplot2::theme_bw(). Any of the ggplot2 themes, or themes from extension packages are allowed (e.g., ggthemes::theme_fivethirtyeight(), hrbrthemes::theme_ipsum_ps(), etc.).
- ggstatsplot.layer Logical that decides whether theme_ggstatsplot theme elements are to be displayed along with the selected ggtheme (Default: TRUE).

Value

A ggplot2 object with the theme_ggstatsplot theme.

Author(s)

Indrajeet Patil

Titanic_full	<i>Titanic dataset.</i>
--------------	-------------------------

Description

Titanic dataset.

Usage

```
Titanic_full
```


Format

A data frame with 2201 rows and 5 variables

- id. Dummy identity number for each person.
- Class. 1st, 2nd, 3rd, Crew.
- Sex. Male, Female.
- Age. Child, Adult.
- Survived. No, Yes.

Details

This data set provides information on the fate of passengers on the fatal maiden voyage of the ocean liner 'Titanic', summarized according to economic status (class), sex, age and survival.

This is a modified dataset from datasets package.

Source

<https://stat.ethz.ch/R-manual/R-devel/library/datasets/html/Titanic.html>

Examples

```
dim(Titanic_full)
head(Titanic_full)
dplyr::glimpse(Titanic_full)
```

VR_dilemma

Virtual reality moral dilemmas.

Description

Virtual reality moral dilemmas.

Usage

```
VR_dilemma
```

Format

A data frame with 68 rows and 4 variables

- id. Dummy identity number for each participant.
- order. The order in which the participants completed the two sessions: "text_first" (0) or "text_second" (1).
- modality. Describes how the moral dilemmas were presented to the participants: either in text format ("text") or in Virtual Reality ("vr").
- score. Proportion of "utilitarian" decisions. In other words, of the 4 decisions, how many affirmative were responses. Range: 0 (all utilitarian) - 1 (none utilitarian).

Details

Dataset from a study where participants completed identical moral dilemmas in two different sessions held on separate days: in one session, they read text description of the scenario, while in another session they completed the same scenarios in Virtual Reality (videos: <https://www.youtube.com/watch?v=ebdU3HhhYs8>). The study investigated if there was a discrepancy between how people judged the same scenarios while reading them in text versus experiencing them in virtual reality.

Source

<https://psyarxiv.com/ry3ap/>

Examples

```
dim(VR_dilemma)
head(VR_dilemma)
dplyr::glimpse(VR_dilemma)
```

Index

*Topic **datasets**

- intent_morality, 86
- iris_long, 88
- movies_long, 89
- movies_wide, 90
- Titanic_full, 120
- VR_dilemma, 121
- _PACKAGE (ggstatsplot-package), 3
- align_plots(), 14
- bartlett_message, 4, 91, 94
- bf_caption_maker, 6
- bf_contingency_tab, 7, 9, 10, 12, 13
- bf_corr_test, 7, 8, 10, 13
- bf_extractor, 9
- bf_one_sample_ttest, 11
- bf_oneway_anova, 7, 9, 10, 12, 13
- bf_two_sample_ttest, 7, 9, 10, 12, 12
- combine_plots, 4, 14
- effsize_ci_message, 5, 91, 94
- ggbarstats, 4, 17, 55
- ggbetweenstats, 3, 5, 21, 61, 91, 94
- ggcoefstats, 4, 26
- ggcorrmatrix, 4, 33, 50, 66, 85
- ggcorrmatrix_message, 5, 91, 94
- ggdotchartstats (ggdotplotstats), 36
- ggdotplotstats, 4, 36, 71
- gghistostats, 4, 39, 75, 115
- ggpiestats, 4, 43, 80, 101
- ggplot_converter, 46
- ggscatterstats, 4, 35, 47, 66, 85
- ggstatsplot (ggstatsplot-package), 3
- ggstatsplot-package, 3
- grouped_ggbarstats, 51
- grouped_ggbetweenstats, 25, 56, 94
- grouped_ggcorrmatrix, 35, 50, 62, 85
- grouped_ggdotplotstats, 67
- grouped_gghistostats, 42, 71
- grouped_ggpiestats, 76
- grouped_ggscatterstats, 35, 50, 66, 81
- grouped_message, 5, 91, 94
- intent_morality, 86
- iris_long, 88
- loess(), 48, 82
- mgcv::gam(), 48, 82
- movies_long, 89
- movies_wide, 90
- normality_message, 5, 91, 94
- ordinal::confint.clm(), 27
- outlier_df, 92
- pairwise_p, 5, 25, 91, 93
- palette_message, 5, 91, 94
- plot_grid, 14
- subtitle_anova_bayes, 96
- subtitle_anova_parametric, 97
- subtitle_anova_robust, 99
- subtitle_contingency_tab, 100
- subtitle_friedman_nonparametric, 102
- subtitle_ggscatterstats, 103
- subtitle_kw_nonparametric, 104
- subtitle_mann_nonparametric, 106
- subtitle_meta_ggcoefstats, 108
- subtitle_onesample_proptest, 110
- subtitle_t_bayes, 112
- subtitle_t_nonparametric
(subtitle_mann_nonparametric),
106
- subtitle_t_onesample, 114
- subtitle_t_parametric, 115
- subtitle_t_robust, 117
- subtitle_template, 111

theme_corrmat, [118](#)
theme_ggstatsplot, [4](#), [119](#)
theme_mprl (theme_ggstatsplot), [119](#)
theme_pie, [120](#)
Titanic_full, [120](#)

VR_dilemma, [121](#)