Package 'datavis'

May 10, 2022

Type Package

Title Functions for data visualization

Version 0.0.1

Author Dimitris Kokoretsis

Maintainer Dimitris Kokoretsis <dkokoret@gmail.com>

Description These functions were designed to simplify and streamline the creation, manipulation and export of bar plots and box plots of publication-grade quality.

License MIT

Encoding UTF-8

RoxygenNote 7.1.2

Roxygen list(markdown = TRUE)

Imports ggplot2, ggthemes, data.table, rlang, ggiraphExtra, stringr, measurements, Cairo

RemoteType github

RemoteHost api.github.com

RemoteRepo datavis

RemoteUsername dimitriskokoretsis

RemoteRef HEAD

RemoteSha 0eec19b1dd693e6072ae29ea87005a212ac3f0ae

GithubRepo datavis

GithubUsername dimitriskokoretsis

GithubRef HEAD

GithubSHA1 0eec19b1dd693e6072ae29ea87005a212ac3f0ae

NeedsCompilation no

R topics documented:

bar_point_plot *Bar plots of grouped data with summary statistics*

Description

Create bar plots of grouped data means plus/minus standard deviations (or custom, manually calculated error measure).

Usage

```
bar_point_plot(
  d,
  x = NULL,
  y,
  color.group = NULL,
  x.axis = NULL,
  y.axis = NULL,
  legend.title = NULL,
  x.order = NULL,
  group.order = NULL,
  x.first = NULL,
  group.first = NULL,
  points = TRUE,
  barwidth = 0.7,
  jitterwidth = 1,
  pointsize = 1,
  whisker.width = 1,
  mean.type = "arithmetic",
  error.lower = NULL,
  error.upper = NULL
)
```

Arguments

d	data.frame with data to be plotted.
x	Character. The name of the column to be used for the x axis (categorical data). Defaults to NULL, but if NULL, color.group needs to be supplied.
y	Character. The name of the column to be used for the y axis (numeric data).
color.group	Character. The name of the column to be used for color grouping (categorical data). Defaults to NULL, but if NULL, x needs to be supplied.
x.axis	Character. The title of the x axis. If NULL, x axis title is given the value of x. Defaults to NULL.
y.axis	Character. The title of the y axis. If NULL, y axis title is given the value of y. Defaults to NULL.
legend.title	Character. The title of the legend. If NULL, legend title is given the value of color.group. Defaults to NULL.
x.order	Character vector of length equal to the number of x categories. Determines the order of categories in the x axis, respectively. Defaults to alphabetical order.

group.order	Character vector of length equal to the number of color grouping categories. Determines the order of color groups. Defaults to alphabetical order.
x.first	Character. Places a specific x axis category first. Ignored if x.order is supplied. Defaults to NULL.
group.first	Character. Places a specific color group category first. Ignored if group.order is supplied. Defaults to NULL.
points	Logical. Sets whether or not to plot individual data points. Defaults to TRUE.
barwidth	Numeric. Sets the width of the bars. Defaults to 0.7.
jitterwidth	Numeric. The horizontal dispersion of individual data points. Defaults to 1.
pointsize	Numeric. Sets the size of the individual data points. Defaults to 1.
whisker.width	Numeric. Sets the width of the error bar whiskers. Defaults to 1.
mean.type	Character, either "arithmetic" or "geometric". Sets which type of mean and standard deviation to plot. Defaults to "arithmetic".
error.lower	Character, name of a column in the d data.frame. Plot custom lower error bar, calculated by the user and included in the d data frame as a column. Defaults to NULL, which plots mean - SD.
error.upper	Character, name of a column in the d data.frame. Plot custom upper error bar, calculated by the user and included in the d data frame as a column. Defaults to NULL, which plots mean + SD.

Details

The calculations for mean and standard deviation are performed automatically. If other error measure is be shown, it needs to be calculated manually and included as two fields in the supplied data.frame (field names are supplied as error.lower and error.upper arguments). To change between arithmetic and geometric mean plus/minus standard deviation, set mean.type argument to "arithmetic" or "geometric".

The order of groups shown in the x axis and color groups defaults to alphabetical. To change it, supply x.order or group.order arguments. These should be character vectors with the desired order of each factor. Do not include categories that don't exist in the supplied data.frame. To bring just one category first, supply x.first or group.first arguments. If x.order is supplied, x.first will be ignored. The same holds for group.order and group.first.

Adjust other supplied arguments to customize the plot aesthetically.

Value

A plot based on ggplot2.

---

box_mean_plot *Box plots of grouped data with optional mean value*

Description

Create standard box plots of grouped data (line: median, box: 25 to 75 percentiles, whiskers: 1.5*IQR (interquartile range)).

Usage

```
box_mean_plot(
  d,
  x = NULL,
  y,
  color.group = NULL,
  x.axis = NULL,
  y.axis = NULL,
  legend.title = NULL,
  x.order = NULL,
  group.order = NULL,
  x.first = NULL,
  group.first = NULL,
  means = FALSE,
  boxwidth = 0.7,
  whisker.width = 1,
  mean.size = 1,
  points = TRUE,
  jitterwidth = 1,
  pointsize = 1,
  mean.type = "arithmetic"
)
```

Arguments

d	data.frame with data to be plotted.
x	Character. The name of the column to be used for the x axis (categorical data). Defaults to NULL, but then color.group needs to be supplied.
y	Character. The name of the column to be used for the y axis (numeric data).
color.group	Character. The name of the column to be used for color grouping (categorical data). Defaults to NULL, but then x needs to be supplied.
x.axis	Character. The title of the x axis. If NULL, x axis title is given the value of x. Defaults to NULL.
y.axis	Character. The title of the y axis. If NULL, y axis title is given the value of y. Defaults to NULL.
legend.title	Character. The title of the legend. If NULL, legend title is given the value of color.group. Defaults to NULL.
x.order	Character vector of length equal to the number of x categories. Sets the order of categories in the x axis. Defaults to alphabetical order.
group.order	Character vector of length equal to the number of color grouping categories. Sets the order of color groups. Defaults to alphabetical order.
x.first	Character. Places a specific x axis category first. Ignored if x.order is supplied. Defaults to NULL.
group.first	Character. Places a specific color group category first. Ignored if group.order is supplied. Defaults to NULL.
means	Logical. Sets whether or not to plot group means. Defaults to FALSE.
boxwidth	Numeric. Sets the width of the bars. Defaults to 0.7.
whisker.width	Numeric. Sets the width of the error bar whiskers. Defaults to 1.

mean.size	Numeric. Sets the size of group mean points. Defaults to 1.
points	Logical. Sets whether or not to plot individual data points. Defaults to TRUE. Note that, even if set to FALSE, outliers will still be shown as individual data points, as is the norm with box plots.
jitterwidth	Numeric. The horizontal dispersion of individual data points. Defaults to 1.
pointsize	Numeric. Sets the size of the individual data points. Defaults to 1.
mean.type	Character. Sets from arithmetic to geometric mean and SD. Defaults to "arithmetic".

Details

Can include the mean value for each group. To include it, set means argument to TRUE. To change between arithmetic or geometric mean, set the mean.type argument to "arithmetic" or "geometric".

The order of groups shown in the x axis and color groups defaults to alphabetical. To change it, supply x.order or group.order arguments. These should be character vectors with the desired order of each factor. Do not include categories that don't exist in the supplied data.frame. To bring just one category first, supply x.first or group.first arguments. If x.order is supplied, x.first will be ignored. The same holds for group.order and group.first.

Adjust other supplied arguments to customize the plot aesthetically.

Value

A plot based on ggplot2.

---

plot_save *Plot export in multiple formats*

Description

Exports plots based on ggplot2 in vector (PDF, SVG), raster (PNG) and R data format (Rds).

Usage

```
plot_save(plot, filepath, height, width, unit = "in", dpi = 600)
```

Arguments

plot	Plot to be exported. Needs to have been produced using ggplot2 (functions described in this guide comply with this).
filepath	Character. The directory and file name for the exported plot. Do not add file extension, the appropriate extension will be added to each exported file.
height	Numeric. The height of the exported plot in the desired length unit.
width	Numeric. The width of the exported plot in the desired length unit.
unit	Character. The unit to be used as height and width. "in" for inches, "cm" for centimeters and "mm" for millimeters. Defaults to "in".
dpi	Integer. The resolution of the exported PNG image in dots per inch. Defaults to 600.

Details

PNG and PDF files are exported using the ggplot2-based ggsave function. Exported PNG images are anti-aliased using the Cairo graphics library. Exported SVG graphics are drawn using the R-native SVG graphics device. This is because SVG graphics exported by ggsave are not rendered properly by the graphic design software Affinity Designer. The downside of exporting SVG with the native graphics device is that the text is drawn as paths instead of text, so it cannot be easily modified in any graphic design software.

Value

Exports plot without returning a value.

---

plot_stats *Add labels on top of bars or boxes*

Description

Adds text labels on top of bars or boxes of a ggplot2-based plot. Useful for annotating statistical analysis information.

Usage

```
plot_stats(plot, d, labels, position = "identity", size = 1, y.adj = 0)
```

Arguments

plot	Plot onto which to add labels. Requires a discrete x scale, optionally dodged by "fill" aesthetic.
d	data.frame containing the labels to be plotted. Must have categorical fields corresponding to the data used for the original plot.
labels	Character. The name of the column with the labels to be plotted.
position	Character, either "identity" or "dodge". Sets where (horizontally) the labels should be plotted. Defaults to "identity". If plot is dodged, "dodge" plots labels above each bar or box. "identity" plots them above the middle of each group of bars or boxes. If plot is not dodged, position is automatically set to "identity".
size	Numeric. Sets the size of labels. Defaults to 1.
y.adj	Numeric. Adjusts the height of labels above each bar, box or group. Defaults to 0. Positive sets it higher, negative sets it lower.

Details

plot_stats detects the maximum plotted value of each data group (be it bar, box, error bar or individual data point), and plots labels above it.

The supplied plot needs to be based on ggplot2 and to have a discrete x scale. The supplied data.frame needs to have a character field with the labels to be plotted (whose name is specified in the labels argument) and factor field(s) for the groups in which each label belongs.

Value

A plot based on ggplot2.

Index