Package ‘sherlock’

June 12, 2023

Title  Graphical Displays for Structured Problem Solving and Diagnosis
Version  0.7.0
Description  Powerful graphical displays and statistical tools for structured problem solving and diagnosis. The functions of the 'sherlock' package are especially useful for applying the process of elimination as a problem diagnosis technique.

The 'sherlock' package was designed to seamlessly work with the 'tidyverse' set of packages and provides a collection of graphical displays built on top of the 'ggplot' and 'plotly' packages, such as different kinds of small multiple plots as well as helper functions such as adding reference lines, normalizing observations, reading in data or saving analysis results in an Excel file.

References:

License  MIT + file LICENSE
Encoding  UTF-8
RoxygenNote  7.2.0
URL  https://github.com/gaboraszabo/sherlock,
     https://gaboraszabo.github.io/sherlock/
BugReports  https://github.com/gaboraszabo/sherlock/issues
Imports  magrittr, rlang (>= 0.4.11), forcats, ggplot2, dplyr, tidyr, cowplot, scales, ggb4x, stringr, plotly, readr, openxlsx, purrr, fs, rstudioapi, tidytext
Suggests  roxygen2
Depends  R (>= 2.10)
LazyData  true
NeedsCompilation  no
Author  Gabor Szabo [aut, cre]
Maintainer  Gabor Szabo <gabor.attila.szabo@gmail.com>
Repository  CRAN
Date/Publication  2023-06-11 23:50:02 UTC
R topics documented:

create_project_folder ................................. 2
draw_cartesian_small_multiples ....................... 3
draw_categorical_scatterplot ......................... 4
draw_horizontal_reference_line ....................... 6
draw_interaction_plot ................................ 7
draw_multivari_plot ................................... 8
draw_multivari_plot_count ............................ 9
draw_pareto_chart ..................................... 10
draw_pareto_chart_grouped ............................ 11
draw_polar_small_multiples ............................ 13
draw_process_behavior_chart ......................... 15
draw_small_multiples_line_plot ....................... 15
draw_timeseries_scatterplot ......................... 17
draw_vertical_reference_line ......................... 19
draw_youden_plot .................................... 19
load_file ............................................. 21
load_files ............................................. 21
multi_vari_data ...................................... 22
multi_vari_data_2 .................................... 23
normalize_observations ............................... 23
plot_tukey_duckworth_paired_test ................... 24
plot_tukey_duckworth_test ............................ 24
polar_small_multiples_data ......................... 25
save_analysis ........................................ 26
scale_color_sherlock ................................ 26
scale_fill_sherlock .................................. 27
select_low_high_units ............................... 27
select_low_high_units_manual ....................... 28
small_multiples_data ................................ 28
theme_sherlock ....................................... 29
timeseries_scatterplot_data ......................... 29
youden_plot_data .................................... 30
youden_plot_data_2 .................................. 30

Index 31

create_project_folder  Create Project Folder

Description

Creates a project folder on your computer

Usage

create_project_folder(folder_name, path, subfolders = "generic")
**Arguments**

- **folder_name**: Set name of the folder. Examples: "Analysis_20221212", "01_application" (required)
- **path**: Set path for folder. Example: "R/Projects/" (required)
- **subfolders**: Set type of subfolder structure. Options are "generic" or "shiny". The "generic" option has the following subfolders: "01_data", "02_reports", "03_images", "04_scripts" and "05_misc". The "shiny" option has the following subfolders: "data", "css", "www", "images" and "scripts". By default, it is set to "generic". (optional)

**Value**

A project folder and sub-folder structure and corresponding .Rproj file on your computer

---

**draw_cartesian_small_multiples**

*Draw Cartesian Small Multiples Plot*

---

**Description**

Draws a cartesian small multiples plot

**Usage**

```r
draw_cartesian_small_multiples(
  data,
  x_coord,
  y_coord,
  grouping_var,
  faceting_var_1,
  faceting_var_2,
  interactive = FALSE,
  size = 2,
  alpha = 0.4,
  analysis_desc_label = NULL,
  x_axis_label = NULL,
  y_axis_label = NULL,
  n_breaks_x_axis = 10,
  n_breaks_y_axis = 10,
  accuracy = 0.001,
  show_axis_values = TRUE
)
```
**Arguments**

- **data**: Input dataset to be plotted (required)
- **x_coord**: Column for X coordinate values (required)
- **y_coord**: Column for Y coordinate values (required)
- **grouping_var**: Grouping variable. Each group is displayed in a different color. (optional)
- **faceting_var_1**: Set first faceting variable (optional)
- **faceting_var_2**: Set second faceting variable (optional)
- **interactive**: Set plot interactivity. By default, it is set to FALSE (optional)
- **size**: Set point size. By default, it is set to 2 (optional)
- **alpha**: Set transparency. By default, it is set to 0.4 (optional)
- **analysis_desc_label**: Label (subtitle) for analysis description. By default, it is set to NULL (optional)
- **x_axis_label**: Label for x axis. By default, it is set to display x axis column name (optional)
- **y_axis_label**: Label for y axis. By default, it is set to display y axis column name (optional)
- **n_breaks_x_axis**: Set number of breaks on X axis. By default, it is set to 10 (optional)
- **n_breaks_y_axis**: Set number of breaks on Y axis. By default, it is set to 10 (optional)
- **accuracy**: Set number of decimal places to be displayed on X and Y axes. Examples: 0.1 - one decimal place, 0.01 - two decimal places, 0.001 - three decimal places etc. By default, it is set to 0.001 (optional)
- **show_axis_values**: Logical. If FALSE, default, axis values are not shown (optional)

**Value**

A 'ggplot' or 'plotly' object

---

**draw_categorical_scatterplot**

*Draw Categorical Scatter Plot*

**Description**

Draws a Categorical Scatter Plot
**draw_categorical_scatterplot**

**Usage**

```r
draw_categorical_scatterplot(
  data,
  y_var,
  grouping_var_1,
  grouping_var_2,
  grouping_var_3,
  grouping_var_4,
  plot_means = FALSE,
  connect_means = FALSE,
  group_color = FALSE,
  point_size = 2,
  alpha = 0.5,
  jitter = FALSE,
  x_axis_text_size = 11,
  panel_text_size = 11
)
```

**Arguments**

- `data`: Input dataset to be plotted (required)
- `y_var`: Y variable to be plotted on Y axis (required)
- `grouping_var_1`: Select column for lowest level grouping variable (optional)
- `grouping_var_2`: Select column for second level grouping variable (optional)
- `grouping_var_3`: Select column for third level grouping variable (optional)
- `grouping_var_4`: Select column for fourth level grouping variable (optional)
- `plot_means`: Logical. If TRUE, means for lowest-level grouping variable are plotted. By default, it is set to FALSE. (optional)
- `connect_means`: Logical. If TRUE, means for lowest-level grouping variable are connected with a line. By default, it is set to FALSE. (optional)
- `group_color`: Set whether to color by grouping_var_1. By default, it is set to FALSE (optional)
- `point_size`: Set point size. By default, it is set to 2 (optional)
- `alpha`: Set transparency. By default, it is set to 0.5 (optional)
- `jitter`: Set whether to add jitter. By default, it is set to FALSE (optional)
- `x_axis_text_size`: Set X axis text size. By default, it is set to 11 (optional)
- `panel_text_size`: Set panel text size. By default, it is set to 11 (optional)

**Value**

A 'ggplot' object
Examples

```r
multi_vari_data_2 %>%
  draw_categorical_scatterplot(y_var = Length,
    grouping_var_1 = Part,
    grouping_var_2 = Operator,
    jitter = FALSE)
```

---

draw_horizontal_reference_line

*Draw horizontal reference line*

Description

Draws a horizontal reference line or multiple reference lines to plots

Usage

```r
draw_horizontal_reference_line(
  reference_line,
  color = "grey",
  linetype = "dashed",
  size = 0.7
)
```

Arguments

- **reference_line**: input y coordinate of reference line(s). for multiple reference lines, concatenate individual values into a vector (required)
- **color**: change reference line color. options are "grey", "blue" and "red". by default, it is set to "grey" (optional)
- **linetype**: change line type. identical to linetype ggplot2 aesthetic. by default, it is set to "dashed" (optional)
- **size**: change line thickness. identical to size ggplot2 aesthetic. by default, it is set to 0.7 (optional)

Value

A horizontal reference line plotted on top of a 'ggplot' object
**draw_interaction_plot**  
*Draw Interaction Plot*

---

**Description**

Draws an Interaction Plot

**Usage**

```r
draw_interaction_plot(
  data, 
  y_var, 
  x_var_1_levels, 
  x_var_2_levels, 
  point_size = 4, 
  line_size = 1, 
  alpha = 0.5, 
  analysis_desc_label = NULL
)
```

**Arguments**

- **data**: input dataset to be plotted (required)
- **y_var**: Y variable to be plotted on Y axis (required)
- **x_var_1_levels**: First grouping variable levels, e.g. -1/1 or "low"/"high" (required)
- **x_var_2_levels**: Second grouping variable levels, e.g. -1/1 or "low"/"high" (required)
- **point_size**: Set point size. By default, it is set to 4 (optional)
- **line_size**: Set line size. By default, it is set to 1 (optional)
- **alpha**: Set transparency. By default, it is set to 0.5 (optional)
- **analysis_desc_label**: analysis_desc_label Label (subtitle) for analysis description. By default, it is set to NULL (optional)

**Value**

A 'ggplot' object
draw_multivari_plot  Draw Multivari Plot

Description

Draws a multivari small multiples plot

Usage

```
draw_multivari_plot(
  data,  
  y_var,  
  grouping_var_1,  
  grouping_var_2,  
  grouping_var_3,  
  grouping_var_4,  
  data_point_label = NULL,  
  plot_means = FALSE,  
  x_axis_text_size = 11,  
  panel_text_size = 14,  
  point_size = 2.5,  
  line_size = 0.7,  
  alpha = 0.6
)
```

Arguments

- `data`  
  Input dataset to be plotted (required)
- `y_var`  
  Response variable, Y (required)
- `grouping_var_1`  
  Select column for lowest level grouping variable (required)
- `grouping_var_2`  
  Select column for second level grouping variable (required)
- `grouping_var_3`  
  Select column for third level grouping variable (optional)
- `grouping_var_4`  
  Select column for fourth level grouping variable (optional)
- `data_point_label`  
  Select column to label data points (optional)
- `plot_means`  
  Logical. if FALSE, default, means for mid-level factor are not plotted (optional)
- `x_axis_text_size`  
  Set X axis text size. By default, it is set to 11 (optional)
- `panel_text_size`  
  Set panel text size. By default, it is set to 14 (optional)
- `point_size`  
  Set point size. By default, it is set to 2.5 (optional)
- `line_size`  
  Set line size. By default, it is set to 0.7 (optional)
- `alpha`  
  Set transparency. By default, it is set to 0.6 (optional)
### draw_multivari_plot_count

**Value**

A `ggplot` object

**Examples**

```r
library(dplyr)
library(ggh4x)

polar_small_multiples_data %>%
  filter(ID_Measurement_Angle %in% c(0, 45, 90, 135)) %>%
  normalize_observations(y_var = ID,
    grouping_var = Tip_Bottom,
    ref_values = c(0.2075, 0.2225)) %>%
  draw_multivari_plot(y_var = ID_normalized,
    grouping_var_1 = ID_Measurement_Angle,
    grouping_var_2 = Mold_Cavity_Number,
    grouping_var_3 = Tip_Bottom,
    x_axis_text = 6) +
  draw_horizontal_reference_line(reference_line = 0)
```

---

**draw_multivari_plot_count**

*Draw Multivari Plot for Counts*

**Description**

Draws a multivari small multiples plot for count data

**Usage**

```r
draw_multivari_plot_count(
  data, 
  y_var, 
  grouping_var_1, 
  grouping_var_2, 
  grouping_var_3, 
  grouping_var_4, 
  x_axis_text_size = 11, 
  panel_text_size = 14, 
  alpha = 0.6
)
```

**Arguments**

- `data` : Input dataset to be plotted (required)
- `y_var` : Response variable, Y (required)
draw_pareto_chart

Description

Draws a Pareto Chart

Usage

draw_pareto_chart(
  data,
  cat_var,
  summarize = FALSE,
  continuous_var,
  drop_na = TRUE,
  highlight_first_n_items = 0,
  lump_last_n_items = 0,
  lumped_cat_name = "Other",
  column_fill = scale_fill_sherlock(3),
  scale = "numeric",
  accuracy = 1,
  title_label = "Pareto Chart",
  analysis_desc_label = NULL,
  axis_text_size = 10
)

Arguments

data input dataset to be plotted (required)
cat_var Categorical variable (required)
**draw_pareto_chart_grouped**

**Description**

Draws a small multiples type of Pareto Chart grouped by a categorical variable.

**Value**

A `ggplot` object

**Arguments**

- **summarize**  
  Logical. If FALSE, default, the function expects total counts of each category of the categorical variable. If TRUE, individual values within each category are automatically summed up and ranked. (required)

- **continuous_var**  
  Continuous variable to rank by (e.g. sum, frequency etc.). Not required if summarize argument is set to TRUE. (required)

- **drop_na**  
  Logical. If TRUE, default, NA values of the categorical variable are dropped. (required)

- **highlight_first_n_items**  
  Specify the top n items to be highlighted. By default, it is set to 0. (optional)

- **lump_last_n_items**  
  Specify the last n items to be lumped into one category. By default, it is set to 0. (optional)

- **lumped_cat_name**  
  Name lumped category. By default, it is set to "Other". (optional)

- **column_fill**  
  Column fill color. By default, it is set to scale_fill_sherlock(3) (optional)

- **scale**  
  Specify an acceptable argument for scale. Acceptable arguments are "numeric", "percent", "dollar", "dollar-k" or "dollar-M". By default, it is set to "numeric" (optional)

- **accuracy**  
  Number to round to. Default value is set to 1. If NULL, values will be rounded to the nearest integer. (optional)

- **title_label**  
  Specify plot title. By default, it is set to display "Pareto Chart" (optional)

- **analysis_desc_label**  
  Specify plot analysis desc label (subtitle). By default, it is set to display CONTINUOUS VARIABLE COLUMN NAME "by" CATEGORICAL VARIABLE COLUMN NAME (optional)

- **axis_text_size**  
  Set axis text size. By default, it is set at 10. (optional)
draw_pareto_chart_grouped

Usage

```r
draw_pareto_chart_grouped(
  data,
  cat_var,
  grouping_var,
  summarize = FALSE,
  continuous_var,
  drop_na = TRUE,
  highlight_first_n_items = 0,
  lump_last_n_items = 0,
  lumped_cat_name = "Other",
  color = "one",
  scale = "numeric",
  accuracy = 1,
  title_label = "Pareto Chart",
  analysis_desc_label = NULL,
  axis_text_size = 10,
  x_axis_span = "free"
)
```

Arguments

data input dataset to be plotted (required)
cat_var Categorical variable (required)
grouping_var Grouping variable (required)
summarize Logical. If FALSE, default, the function expects total counts of each category of the categorical variable. If TRUE, individual values within each category are automatically summed up and ranked. (required)
continuous_var Continuous variable to rank by (e.g. sum, frequency etc.). Not required if summarize argument is set to TRUE. (required)
drop_na Logical. If TRUE, default, NA values of the categorical variable are dropped. (required)
highlight_first_n_items Specify the top n items to be highlighted. By default, it is set to 0. (optional)
lump_last_n_items Specify the last n items to be lumped into one category. By default, it is set to 0. (optional)
lumped_cat_name Name lumped category. By default, it is set to "Other". (optional)
color Set panel fill color for facets. Options are "one" (one color) or "multi" (each panel is a different color). By default, it is set to "one". (optional)
scale Specify an acceptable argument for scale. Acceptable arguments are "numeric", "percent", "dollar", "dollar-k" or "dollar-M". By default, it is set to "numeric" (optional)
accuracy Number to round to. Default value is set to 1. If NULL, values will be rounded to the nearest integer. (optional)
**draw_polar_small_multiples**

**Description**

Draws a Polar Small Multiples Plot

**Usage**

```r
draw_polar_small_multiples(
  data,
  angular_axis,
  x_y_coord_axis,
  grouping_var,
  faceting_var_1,
  faceting_var_2,
  connect_with_lines = FALSE,
  connect_start_and_end_points = TRUE,
  x_y_coord_axis_limits = c(0, NA),
  point_size = 2,
  line_size = 0.6,
  point_alpha = 0.6,
  line_alpha = 0.5,
  label_text_size = 11,
  analysis_desc_label = ""
)
```
**Arguments**

- **data**: input dataset to be plotted (required)
- **angular_axis**: angular coordinate values (required)
- **x_y_coord_axis**: x-y coordinate values (required)
- **grouping_var**: grouping variable (required)
- **faceting_var_1**: Set first faceting variable (optional)
- **faceting_var_2**: Set second faceting variable (optional)
- **connect_with_lines**: Logical. If set to TRUE, values within each group are connected with a line. By default, it is set to FALSE (optional)
- **connect_start_and_end_points**: Logical. If set to TRUE, the start and end points of the lines get connected. It is useful when trying to draw a complete circle but may not be useful when only trying to draw a shape different than that (e.g. a semicircle). By default, it is set to TRUE (optional)
- **x_y_coord_axis_limits**: Set x-y coordinate axis limits. By default, it is set to start at 0. (optional)
- **point_size**: Set point size. By default, it is set to 2 (optional)
- **line_size**: Set line size. By default, it is set to 0.6 (optional)
- **point_alpha**: Set point transparency. By default, it is set to 0.6 (optional)
- **line_alpha**: Set line transparency. By default, it is set to 0.5 (optional)
- **label_text_size**: Size of text for labels. By default, it is set to 11. (optional)
- **analysis_desc_label**: Label (subtitle) for analysis description. By default, it is set to NULL. (optional)

**Value**

A `'ggplot'` object

**Examples**

```r
draw_polar_small_multiples
```

```r
library(dplyr)
polar_small_multiples_data %>%
  filter(Mold_Cavity_Number %in% c(4, 6)) %>%
  draw_polar_small_multiples(angular_axis = ID_Measurement_Angle,
                           x_y_coord_axis = ID_2,
                           grouping_var = Tip_Bottom,
                           faceting_var_1 = Mold_Cavity_Number,
                           point_size = 0.5,
                           connect_with_lines = TRUE,
                           label_text_size = 7)
```
**draw_process_behavior_chart**

*Draw Process Behavior Chart*

**Description**

Draws a Process Behavior Chart

**Usage**

```r
draw_process_behavior_chart(
  data,
  y_var,
  grouping_var,
  limits = TRUE,
  interactive = TRUE
)
```

**Arguments**

- `data`: input dataset to be plotted (required)
- `y_var`: Y variable to be plotted on Y axis (required)
- `grouping_var`: Variable to group by (optional)
- `limits`: Logical. If TRUE, natural process limits (control limits) are plotted. By default, it is set to FALSE (optional)
- `interactive`: Set plot interactivity. By default, it is set to TRUE (optional)

**Value**

A `ggplot` or `plotly` object

**draw_small_multiples_line_plot**

*Draw Small Multiples Line Plot*

**Description**

Draws a Small Multiples Line Plot
Usage

draw_small_multiples_line_plot(
  data,
  x_axis_var,
  y_axis_var,
  grouping_var,
  faceting_var_1,
  faceting_var_2,
  plot_max_values = FALSE,
  lowest_highest_units,
  unique_color_by_group = FALSE,
  size = 0.7,
  alpha = 0.4,
  interactive = TRUE,
  analysis_desc_label = NULL,
  x_axis_label = NULL,
  y_axis_label = NULL,
  n_breaks_x_axis = 10,
  n_breaks_y_axis = 10,
  accuracy = 0.01
)

Arguments

data                  input dataset to be plotted (required)
 x_axis_var             variable to be plotted on x axis (required)
 y_axis_var             variable to be plotted on x axis (required)
 grouping_var           set grouping variable (required)
 faceting_var_1        Set first faceting variable (optional)
 faceting_var_2        Set second faceting variable (optional)
 plot_max_values       Highlights maximum values per group. By default, it is set to FALSE (optional)
 lowest_highest_units  takes a vector of strings corresponding to the lowest/highest units to be highlighted (optional)
 unique_color_by_group set whether to display each group in a unique color. By default, it is set to FALSE (optional)
 size                   Set line size. By default, it is set to 0.7 (optional)
 alpha                  Set transparency. By default, it is set to 0.4 (optional)
 interactive            set plot interactivity. By default, it is set to TRUE (optional)
 analysis_desc_label    Label (subtitle) for analysis description. By default, it is set to NULL (optional)
 x_axis_label           Label for x axis. By default, it is set to display x axis column name (optional)
**draw_timeseries_scatterplot**

**y_axis_label**  
Label for y axis. By default, it is set to display y axis column name (optional)

**n_breaks_x_axis**  
Set number of breaks on X axis. By default, it is set to 10 (optional)

**n_breaks_y_axis**  
Set number of breaks on Y axis. By default, it is set to 10 (optional)

**accuracy**  
Set number of decimal places to be displayed on X and Y axes. Examples: 0.1 - one decimal place, 0.01 - two decimal places, 0.001 - three decimal places etc. By default, it is set to 0.01 (optional)

**Value**

A 'ggplot' or 'plotly' object

**draw_timeseries_scatterplot**  
*Draw Timeseries Scatterplot*

**Description**

Draws a Timeseries Scatterplot

**Usage**

```r
draw_timeseries_scatterplot(
  data,
  y_var,
  grouping_var_1,
  grouping_var_1_type = "date-time",
  grouping_var_2,
  faceting = FALSE,
  limits = FALSE,
  date_breaks = "1 month",
  date_labels = "%b %Y",
  analysis_desc_label = NULL,
  x_axis_text_size = 11,
  point_size = 1,
  alpha = 0.3,
  line_size = 1,
  interactive = TRUE
)
```

**Arguments**

- **data**: input dataset to be plotted (required)
- **y_var**: Y variable to be plotted on Y axis (required)
- **grouping_var_1**: Time variable to be plotted on x axis (required)
draw_timeseries_scatterplot

- **grouping_var_1_type**: Time variable type. Options are "date-time" or "factor"
- **grouping_var_2**: Additional variable for faceting (optional)
- **faceting**: Set whether to display each group in a separate plot. By default, it is set to FALSE (optional)
- **limits**: Logical. If TRUE, process behavior chart control limits for the individual group means are plotted. By default, it is set to FALSE (optional)
- **date_breaks**: Set date breaks. Takes a string, for example "1 week" or "2 days". By default, it is set to "1 month" (optional)
- **date_labels**: Set date labels. Identical to the date labels argument of the scale_x_date() ggplot function (optional)
- **analysis_desc_label**: Label (subtitle) for analysis description. By default, it is set to NULL (optional)
- **x_axis_text_size**: X axis text size. By default, it is set to 11. (optional)
- **point_size**: Set point size. By default, it is set to 1 (optional)
- **alpha**: Set transparency for individual observations. Identical to the alpha ggplot argument. By default, it is set to 0.3 (optional)
- **line_size**: Set line size. By default, it is set to 1 (optional)
- **interactive**: Set plot interactivity. By default, it is set to TRUE (optional)

**Value**

A 'ggplot' or 'plotly' object

**Examples**

timeseries_scatterplot_data %>%
  draw_timeseries_scatterplot(y_var = y,
                grouping_var_1 = date,
                grouping_var_2 = cavity,
                faceting = TRUE,
                limits = TRUE,
                alpha = 0.15,
                line_size = 0.5,
                x_axis_text = 7,
                interactive = FALSE)
**draw_vertical_reference_line**

*Draw vertical reference line*

**Description**

Draws a vertical reference line or multiple reference lines to plots.

**Usage**

```r
draw_vertical_reference_line(
  reference_line,
  color = "grey",
  linetype = "dashed",
  size = 0.7
)
```

**Arguments**

- `reference_line`: input x coordinate of reference line(s). for multiple reference lines, concatenate individual values into a vector (required)
- `color`: change reference line color. options are "grey", "blue" and "red". by default, it is set to "grey" (optional)
- `linetype`: change line type. identical to linetype ggplot2 aesthetic. by default, it is set to "dashed" (optional)
- `size`: change line thickness. identical to size ggplot2 aesthetic. by default, it is set to 0.7 (optional)

**Value**

A vertical reference line plotted on top of 'ggplot' object

**draw_youden_plot**

*Draw Youden Plot*

**Description**

Draws a Youden Plot.
Usage

draw_youden_plot(
    data,
    x_axis_var,
    y_axis_var,
    grouping_var,
    lsl,
    usl,
    median_line = FALSE,
    size = 2,
    alpha = 0.4,
    analysis_desc_label = NULL,
    x_axis_label = NULL,
    y_axis_label = NULL
)

Arguments

data input dataset to be plotted (required)
x_axis_var variable to be plotted on x axis (required)
y_axis_var variable to be plotted on x axis (required)

Arguments

grouping_var grouping variable (optional)
lsl lower specification limit (optional)
usl upper specification limit (optional)
median_line logical. If TRUE, a median bias line is plotted. By default, it is set to FALSE (optional)

Arguments

size Set point size. By default, it is set to 2 (optional)
alpha Set transparency. By default, it is set to 0.4 (optional)

Arguments

analysis_desc_label Label (subtitle) for analysis description. By default, it is set to NULL (optional)

Arguments

x_axis_label Label for x axis. By default, it is set to display x axis column name (optional)
y_axis_label Label for y axis. By default, it is set to display y axis column name (optional)

Value

A 'ggplot' object

Examples

youden_plot_data %>%
  draw_youden_plot(x_axis_var = measurement_1,
                   y_axis_var = measurement_2,
                   grouping_var = location)

youden_plot_data_2 %>%
  draw_youden_plot(x_axis_var = gage_1,
load_file

\[
y\_axis\_var = \text{gage}._2, \\
\text{median\_line} = \text{TRUE}
\]

---

### load_file

**Load File**

#### Description
Reads either an .xlsx or a .csv file into a table

#### Usage

```r
load_file(path, filetype = ".xlsx", col_names = TRUE)
```

#### Arguments

- **path**
  - path for the file (required)

- **filetype**
  - set whether to read an .xlsx file or a .csv file. It takes either ".xlsx" or "csv". By default, it is set to ".xlsx" (optional)

- **col_names**
  - Either TRUE, FALSE or a character vector of column names. If TRUE, the first row of the input will be used as the column names, and will not be included in the data frame. If FALSE, column names will be generated automatically: X1, X2, X3 etc. If col_names is a character vector, the values will be used as the names of the columns, and the first row of the input will be read into the first row of the output data frame.

#### Value
Returns data in the form of a `tibble` object.

---

### load_files

**Load Files**

#### Description
Reads a series of either .xlsx or .csv files into a table. Particularly useful when reading in multiple files having the same variables, for example reading in data from an experiment where data was logged and saved separately for each individual unit. Integration of a custom data cleaning function.
Usage

```r
load_files(
  folder,
  filetype = ".csv",
  data_cleaning_function = NULL,
  id_by_filename = FALSE,
  id_col_name = "index"
)
```

Arguments

- **folder**
  Folder where the files to be read in are located (required)
- **filetype**
  Set whether to read in .xlsx or .csv files. It takes either ".xlsx" or ".csv". By
default, it is set to ".csv" (required)
- **data_cleaning_function**
  Add a custom data cleaning function built for individual files. Use no brackets
  when referencing the function, for example clean_data_from_data_logger. The
  function being added must be saved in the environment (optional)
- **id_by_filename**
  Logical. If set to TRUE, the output will contain a column, storing the name
  of each file being read in. Ideally, the names of the individual files should be
  pertinent to their content, e.g. if 20 files are being read in with experimental data
  from parts 1 through 20, the files should be named 1-20. (optional)
- **id_col_name**
  Specify a name for the .id column. By default, it is set to "index" (optional)

Value

Returns data in the form of a tibble object.

---

### multi_vari_data

**Multivari Plot Sample Dataset 1**

Description

Contains a sample Multivari Plot dataset

Usage

```r
multi_vari_data
```

Format

An object of class tbl_df (inherits from tbl.data.frame) with 18 rows and 4 columns.

Examples

```r
multi_vari_data
```
**multi_vari_data_2**

**Multi-Vari Plot Sample Dataset 2**

**Description**
Contains a sample Multi-Vari Plot dataset

**Usage**
multi_vari_data_2

**Format**
An object of class `tbl_df` (inherits from `tbl, data.frame`) with 54 rows and 4 columns.

**Examples**
multi_vari_data_2

**normalize_observations**

*Normalize observations*

**Description**
This function takes an input dataset and normalizes observations

**Usage**
normalize_observations(data, y_var, grouping_var, ref_values)

**Arguments**
- **data**: input dataset to be plotted (required)
- **y_var**: response variable, Y (required)
- **grouping_var**: select grouping variable to normalize by (required)
- **ref_values**: add reference (nominal) values. Takes a string of values with values appearing in the same order as in grouping variable. String length must be equal to unique values in grouping variable (required)

**Value**
A tibble object with observations normalized and saved in a new column.
Examples

```r
library(dplyr)

polar_small_multiples_data %>%
  filter(ID_Measurement_Angle %in% c(0, 45, 90, 135)) %>%
  normalize_observations(y_var = ID,
                         grouping_var = Tip_Bottom,
                         ref_values = c(0.2075, 0.2225))
```

Description

Plots Tukey-Duckworth Paired Test

Usage

```r
plot_tukey_duckworth_paired_test(data, y_var, x_vars, arrows = FALSE)
```

Arguments

- `data` input dataset (required)
- `y_var` Y variable of interest (required)
- `x_vars` X variables of interest (required)
- `arrows` Set whether to display arrows in the plot. By default, it is set to FALSE (optional)

Value

A `ggplot` object
Usage

plot_tukey_duckworth_test(
  data,
  y_var,
  x_var_levels,
  point_size = 3,
  point_type = "solid",
  split_levels = FALSE,
  analysis_desc_label = NULL
)

Arguments

data input dataset (required)
y_var Y variable of interest (required)
x_var_levels Levels of X variable of interest (required)
point_size Set point size. By default, it is set to 3. (optional)
point_type Set point size. Options are "solid" (default) and "no fill". (optional)
split_levels Set whether to plot the two levels in separately on the X axis. By default, it is set to FALSE (optional)
analysis_desc_label Label (subtitle) for analysis description. By default, it is set to NULL (optional)

Value

A `ggplot` object

polar_small_multiples_data

Polar Small Multiples Sample Dataset

Description

Contains a sample dataset to demonstrate the use of Polar Small Multiples plot

Usage

polar_small_multiples_data

Format

An object of class tbl_df (inherits from tbl.data.frame) with 144 rows and 5 columns.

Examples

polar_small_multiples_data
**save_analysis**  
*Save Analysis*

**Description**

Saves analysis results, both data and plot, into an .xlsx file

**Usage**

```r
save_analysis(data, plot, filename, filepath)
```

**Arguments**

- `data` Data to be saved (required)
- `plot` Plot to be saved (optional)
- `filename` Name of the Excel file in a string format without the .xlsx extension. Example: "analysis_results" (required)
- `filepath` Path for the file. Example: "Documents/" (required)

**Value**

An Excel file

---

**scale_color_sherlock**  
*Sherlock Color Palettes*

**Description**

Set color scheme to one of the Sherlock color palettes

**Usage**

```r
scale_color_sherlock(palette = 1)
```

**Arguments**

- `palette` color palette to be used (required). options are 1, 2 and 3 (2 and 3 are only one color for no grouping). by default it is set to 1.

**Value**

A `ggplot` color scheme that uses one of the Sherlock color palettes
scale_fill_sherlock  

**Sherlock Fill Color Palettes**

**Description**
Set fill color scheme to one of the Sherlock color palettes

**Usage**
scale_fill_sherlock(palette = 1)

**Arguments**

- **palette** fill color palette to be used (required). options are 1, 2 and 3 (2 and 3 are only one color for no grouping). by default it is set to 1.

**Value**
A 'ggplot' color scheme that uses one of the Sherlock color fill palettes

select_low_high_units  

**Select Low-High Units**

**Description**
Automatically selects low-high units in a tibble as well as assigns them into groups

**Usage**
select_low_high_units(data, var, number_of_pairs)

**Arguments**

- **data** input dataset (required)
- **var** variable of interest (required)
- **number_of_pairs** Number of low-high pairs to be created. Takes a numeric value (required)

**Value**
A tibble object filtered down to the low-high units selected
select_low_high_units_manual

Select Low-High Units

Description
Select low-high units manually in a tibble and assign them into groups

Usage
select_low_high_units_manual(
data,
select_units_by = "row_number",
lowest_units,
highest_units,
part_id_col
)

Arguments

<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>data</td>
<td>input dataset (required)</td>
</tr>
<tr>
<td>select_units_by</td>
<td>Set to select units either based on row number or part ID. Options are &quot;row_number&quot; and &quot;part_id&quot;. By default, it is set to &quot;row_number&quot;. (required)</td>
</tr>
<tr>
<td>lowest_units</td>
<td>A numerical or character vector of the lowest units selected. Examples: c(1, 6, 8, 12). c(&quot;part5&quot;, &quot;part45&quot;, &quot;part9&quot;, &quot;part23&quot;). (required)</td>
</tr>
<tr>
<td>highest_units</td>
<td>A numerical or character vector of the lowest units selected. Examples: c(1, 6, 8, 12). c(&quot;part5&quot;, &quot;part45&quot;, &quot;part9&quot;, &quot;part23&quot;). (required)</td>
</tr>
<tr>
<td>part_id_col</td>
<td>Set column for part id. Only to be used when select_units_by is set to &quot;part_id&quot;.</td>
</tr>
</tbody>
</table>

Value
A tibble object filtered down to the low-high units selected

small_multiples_data
Small Multiples Sample Dataset

Description
Contains a sample dataset for small multiples

Usage
small_multiples_data
theme_sherlock

Format
An object of class tbl_df (inherits from tbl.data.frame) with 2900 rows and 4 columns.

Examples
small_multiples_data

description

A 'theme' object with Sherlock plot theme

Usage
theme_sherlock(axis_text_size = "normal")

Arguments
axis_text_size set axis text and axis title size. options are "normal" or "small". by default, it is set to "normal"

Value

A 'theme' object with Sherlock plot theme

timeseries_scatterplot_data

Timeseries Scatterplot Sample Dataset

description

Contains a sample Timeseries Scatterplot dataset

Usage
timeseries_scatterplot_data

Format
An object of class tbl_df (inherits from tbl.data.frame) with 1170 rows and 5 columns.

Examples
timeseries_scatterplot_data
### Youden Plot Sample Dataset

**Description**

Contains a sample Youden Plot dataset

**Usage**

`youden_plot_data`

**Format**

An object of class `data.frame` with 40 rows and 3 columns.

**Examples**

`youden_plot_data`

---

### Youden Plot Sample Dataset 2

**Description**

Contains a sample Youden Plot dataset

**Usage**

`youden_plot_data_2`

**Format**

An object of class `data.frame` with 30 rows and 2 columns.

**Examples**

`youden_plot_data_2`
Index

* datasets
  multi_vari_data, 22
  multi_vari_data_2, 23
  polar_small_multiples_data, 25
  small_multiples_data, 28
  timeseries_scatterplot_data, 29
  youden_plot_data, 30
  youden_plot_data_2, 30

create_project_folder, 2

draw_cartesian_small_multiples, 3
draw_categorical_scatterplot, 4
draw_categorical_scatterplot, 4

draw_horizontal_reference_line, 6
draw_interaction_plot, 7
draw_multivari_plot, 8
draw_multivari_plot_count, 9
draw_pareto_chart, 10
draw_pareto_chart_grouped, 11
draw_polar_small_multiples, 13
draw_process_behavior_chart, 15
draw_small_multiples_line_plot, 15
draw_timeseries_scatterplot, 17
draw_vertical_reference_line, 19
draw_youden_plot, 19

load_file, 21
load_files, 21

multi_vari_data, 22
multi_vari_data_2, 23

normalize_observations, 23

plot_tukey_duckworth_paired_test, 24
plot_tukey_duckworth_test, 24
polar_small_multiples_data, 25

save_analysis, 26
scale_color_sherlock, 26
scale_fill_sherlock, 27

select_low_high_units, 27
select_low_high_units_manual, 28
small_multiples_data, 28

timeseries_scatterplot_data, 29

youden_plot_data, 30
youden_plot_data_2, 30