Package 'sherlock'

November 17, 2022

Title Graphical Displays to Aid Structured Problem Solving and Diagnosis

Version 0.5.1

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:
David Hartshorne (2019, ISBN: 978-1-5272-5139-7).
Stefan H. Steiner, R. Jock MacKay (2005, ISBN: 0873896467).

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, cowplot, scales, ggh4x, stringr, plotly, readr, openxlsx

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 2022-11-17 14:10:02 UTC

R topics documented:

`draw_cartesian_small_multiples` *Draw Cartesian Small Multiple Plot*

Description

Draws a cartesian small multiples plot

Usage

```
draw_cartesian_small_multiples(
  data,
  x_coord,
  y_coord,
  grouping_var_1,
  grouping_var_2,
  four_quadrants = FALSE,
  show_axis_values = FALSE,
```

```
  faceted = TRUE,
  size = 2,
  alpha = 0.4,
  analysis_desc_label = NULL,
  interactive = FALSE
)
```

Arguments

data	Input dataset to be plotted (required)
x_coord	x coordinate values (required)
y_coord	y coordinate values (required)
grouping_var_1	Grouping variable 1 (required)
grouping_var_2	Grouping variable 2 (optional)
four_quadrants	Logical. Set whether to display four quadrant with both axes starting at zero. By default, it is set to FALSE (optional)
show_axis_values	Logical. if FALSE, default, axis values are not shown (optional)
faceted	Logical. if TRUE, default, plot will be faceted. Note: Cartesian plot is always faceted when there are two grouping variables. Drop grouping variable 2 for no faceting. (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)
interactive	Set plot interactivity. By default, it is set to FALSE (optional)

Value

A 'ggplot' or 'plotly' object

---

draw_categorical_scatterplot

Draw Categorical Scatterplot

Description

Draws a Categorical Scatterplot

Usage

```
draw_categorical_scatterplot(
  data,
  y_var,
  grouping_var_1,
  grouping_var_2,
  grouping_var_3,
  group_color = FALSE,
  size = 2,
  alpha = 0.5,
  jitter = TRUE,
  interactive = FALSE
)
```

Arguments

data	input dataset to be plotted (required)
y_var	Y variable to be plotted on Y axis (required)
grouping_var_1	First grouping variable (optional)
grouping_var_2	Second, higher-level grouping variable (optional)
grouping_var_3	Third, highest-level grouping variable (optional)
group_color	Set whether to color by grouping_var_1. 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 0.5 (optional)
jitter	Set whether to add jitter. By default, it is set to TRUE (optional)
interactive	Set plot interactivity. By default, it is set to FALSE (optional)

Value

A 'ggplot' or 'plotly' object

Examples

```
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

```
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

```
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,
  response,
  factor_1,
  factor_2,
  factor_3,
  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)
response	response variable, Y (required)
factor_1	lowest level factor (required)
factor_2	mid-level factor (required)
factor_3	top level factor (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. options are "normal" (default), "small", "xs" and "none" (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)

Value

A 'ggplot' object

Examples

```
library(dplyr)
library(ggh4x)

polar_small_multiples_data %>%
  filter(ID_Measurement_Angle %in% c(0, 45, 90, 135)) %>%
  normalize_observations(response = ID,
                         grouping_var = Tip_Bottom,
                         ref_values = c(0.2075, 0.2225)) %>%
  draw_multivari_plot(response    = ID_normalized,
                      factor_1    = ID_Measurement_Angle,
                      factor_2    = Mold_Cavity_Number,
                      factor_3    = Tip_Bottom,
                      x_axis_text = 6) +
  draw_horizontal_reference_line(reference_line = 0)
```

`draw_pareto_chart` *Draw Pareto Chart*

Description

Draws a Pareto Chart

Usage

```
draw_pareto_chart(
  data,
  cat_var,
  continuous_var,
  highlight_first_n_items = 0,
  lump_last_n_items = 0,
  lumped_cat_name = "Other",
  column_fill = scale_fill_sherlock(3),
  scale = "numeric",
  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)
`continuous_var`	Continuous variable to rank by (e.g. sum, frequency etc.) (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)
`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)

Value

A 'ggplot' object

`draw_polar_small_multiples` *Draw Polar Small Multiples*

Description

Draws a Polar Small Multiple Plot

Usage

```
draw_polar_small_multiples(
  data,
  angular_axis,
  x_y_coord_axis,
  grouping_var,
  faceting_var_1,
  faceting_var_2,
  connect_with_lines = FALSE,
  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 FALSE, default, values within each group are not connected with a line (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

```
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

```
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,
  lowest_highest_units,
  faceting = FALSE,
  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
)
```

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)
`lowest_highest_units`	takes a vector of strings corresponding to the lowest/highest units to be highlighted (optional)
`faceting`	set whether to display each group in a separate plot. By default, it is set to FALSE (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)
y_axis_label	Label for y axis. By default, it is set to display y axis column name (optional)

Value

A 'ggplot' or 'plotly' object

`draw_timeseries_scatterplot` *Draw Timeseries Scatterplot*

Description

Draws a Timeseries Scatterplot

Usage

```
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)
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

```
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)
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)
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)

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,
                     y_axis_var  = gage_2,
                     median_line = TRUE)
```

---

load_file *Load File*

Description

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

Usage

```
load_file(path, filetype = ".xlsx")
```

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)

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

```
multi_vari_data
```

Format

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

Examples

```
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, response, grouping_var, ref_values)
```

Arguments

`data`	input dataset to be plotted (required)
`response`	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

```
library(dplyr)

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

`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

```
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

```
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

small_multiples_data *Small Multiples Sample Dataset*

Description

Contains a sample dataset for small multiples

Usage

```
small_multiples_data
```

Format

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

Examples

```
small_multiples_data
```

theme_sherlock *Theme Sherlock*

Description

Set 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 Timerseries 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_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_data_2 *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