# Package 'GGally'

March 25, 2020

```
Version 1.5.0
License GPL (>= 2.0)
Title Extension to 'ggplot2'
Type Package
LazyLoad yes
LazyData true
URL https://ggobi.github.io/ggally, https://github.com/ggobi/ggally
BugReports https://github.com/ggobi/ggally/issues
Description The R package 'ggplot2' is a plotting system based on the grammar of graphics.
      'GGally' extends 'ggplot2' by adding several functions
      to reduce the complexity of combining geometric objects with transformed data.
      Some of these functions include a pairwise plot matrix, a two group pairwise plot
      matrix, a parallel coordinates plot, a survival plot, and several functions to
      plot networks.
Depends R (>= 3.1), ggplot2 (> 2.2.0)
Imports grDevices, grid, gtable (>= 0.2.0), plyr (>= 1.8.3), progress,
      RColorBrewer, reshape (>= 0.8.5), utils, rlang
Suggests broom (>= 0.4.0), chemometrics, geosphere (>= 1.5-1), igraph
      (>= 1.0.1), intergraph (>= 2.0-2), maps (>= 3.1.0), mapproj,
      network (>= 1.12.0), scagnostics, scales (>= 0.4.0), sna (>=
      2.3-2), survival, packagedocs (>= 0.4.0), rmarkdown, roxygen2,
      testthat, crosstalk
RoxygenNote 7.1.0
VignetteBuilder packagedocs
SystemRequirements openssl
Encoding UTF-8
NeedsCompilation no
Author Barret Schloerke [aut, cre] (author for ggpairs, ggduo, ggnostic, ggts,
      ggfacet, and ggally_*. Contributor for all functions.),
      Jason Crowley [aut] (ggparcoord),
```

Di Cook [aut, ths] (ggscatmat, gglyph),
Heike Hofmann [ths],
Hadley Wickham [ths],
Francois Briatte [aut] (ggcorr, ggnet, ggnet2),
Moritz Marbach [aut] (ggnet, ggnet2),
Edwin Thoen [aut] (ggsurv),
Amos Elberg [aut] (ggnetworkmap),
Joseph Larmarange [aut] (ggcoef),
Ott Toomet [ctb]

Maintainer Barret Schloerke <schloerke@gmail.com>

Repository CRAN

**Date/Publication** 2020-03-25 18:20:15 UTC

## **R** topics documented:

.gg	. 4
dd_ref_boxes	. 5
dd_ref_lines	. 5
ustralia_PISA2012	. 6
rew_colors	. 7
roomify	. 8
val_data_col	. 8
nd_plot_type	. 9
ea	. 9
n_switch	. 10
etPlot	. 11
gally_barDiag	. 12
gally_blank	. 12
gally_box	. 13
gally_cor	. 14
gally_density	. 15
gally_densityDiag	. 16
gally_denstrip	. 17
gally_diagAxis	. 17
gally_dot	. 18
gally_dot_and_box	. 19
gally_facetbar	. 20
gally_facetdensity	. 21
gally_facetdensitystrip	. 21
gally_facethist	. 22
gally_na	. 23
gally_nostic_cooksd	. 23
gally_nostic_hat	. 24
gally_nostic_line	. 25
gally_nostic_resid	. 26
gally_nostic_se_fit	. 28
gally_nostic_sigma	. 29

Index

98

gally_nostic_std_resid	 30
gally_points	
gally_ratio	
gally_smooth	
gally_text	33
gcoef	34
gcorr	35
c gduo	38
gfacet	43
glegend	44
gmatrix	 45
gmatrix_gtable	 48
gmatrix_progress	 49
gnet	 49
gnet2	
gnetworkmap	 59
gnostic	 63
gpairs	 65
gparcoord	70
gscatmat	74
gsurv	75
gts	77
lyphplot	78
lyphs	79
rab_legend	
appy	
s_horizontal	82
owertriangle	83
napping_string	83
napping_swap_x_y	
asa	
igs	
rint.ggmatrix	
rint_if_interactive	
sychademic	
utPlot	
escale01	90
cag_order	90
catmat	91
ngleClassOrder	92
kewness	92
r.ggmatrix	93
witter_spambots	93
ppertriangle	94
1_ggmatrix_theme	 95
rap_fn_with_param_arg	95

4.gg

+.gg

Modify a ggmatrix object by adding an ggplot2 object to all plots

## **Description**

This operator allows you to add ggplot2 objects to a ggmatrix object.

## Usage

```
## S3 method for class 'gg'
e1 + e2
```

## **Arguments**

- e1 An object of class ggplot or theme
- e2 A component to add to e1

## **Details**

If the first object is an object of class ggmatrix, you can add the following types of objects, and it will return a modified ggplot object.

• theme: update plot theme

The + operator completely replaces elements with elements from e2.

## See Also

+.gg and theme

```
data(tips, package = "reshape")
pm <- ggpairs(tips[, 2:3])
## change to black and white theme
pm + ggplot2::theme_bw()
## change to linedraw theme
# pm + ggplot2::theme_linedraw()
## change to custom theme
# pm + ggplot2::theme(panel.background = ggplot2::element_rect(fill = "lightblue"))
## add a list of information
extra <- list(ggplot2::theme_bw(), ggplot2::labs(caption = "My caption!"))
pm + extra</pre>
```

add\_ref\_boxes 5

244	rof	boxes
auu	rer	boxes

Add reference boxes around each cell of the glyphmap.

## **Description**

Add reference boxes around each cell of the glyphmap.

## Usage

```
add_ref_boxes(
  data,
  var_fill = NULL,
  color = "white",
  size = 0.5,
  fill = NA,
  ...
)
```

## Arguments

data	A glyphmap structure.	
var_fill	Variable name to use to set the fill color	
color	Set the color to draw in, default is "white'	
size	Set the line size, default is 0.5	
fill	fill value used if var_fill is NULL	
	other arguments passed onto geom_rect	

add\_ref\_lines

Add reference lines for each cell of the glyphmap.

## Description

Add reference lines for each cell of the glyphmap.

## Usage

```
add_ref_lines(data, color = "white", size = 1.5, ...)
```

## **Arguments**

data	A glyphmap structure.
color	Set the color to draw in, default is "white"
size	Set the line size, default is 1.5
	other arguments passed onto geom_line

6 australia\_PISA2012

australia\_PISA2012

Programme for International Student Assessment (PISA) 2012 Data for Australia

## **Description**

About PISA

## Usage

data(australia\_PISA2012)

#### **Format**

A data frame with 8247 rows and 32 variables

## **Details**

The Programme for International Student Assessment (PISA) is a triennial international survey which aims to evaluate education systems worldwide by testing the skills and knowledge of 15-year-old students. To date, students representing more than 70 economies have participated in the assessment.

While 65 economies took part in the 2012 study, this data set only contains information from the country of Australia.

```
• gender : Factor w/ 2 levels "female", "male": 1 1 2 2 2 1 1 1 2 1 ...
```

- age: Factor w/ 4 levels "4", "5", "6", "7": 2 2 2 4 3 1 2 2 2 2 ...
- homework: num 5 5 9 3 2 3 4 3 5 1 ...
- desk: num 1 0 1 1 1 1 1 1 1 1 ...
- room: num 1 1 1 1 1 1 1 1 1 1 ...
- study: num 1 1 1 1 1 1 1 1 1 1 ...
- computer: num 1 1 1 1 1 1 1 1 1 1 ...
- software: num 1 1 1 1 1 1 1 1 1 1 ...
- internet: num 1 1 1 1 1 1 1 1 1 1 ...
- literature : num 0 0 1 0 1 1 1 1 1 0 ...
- poetry: num 0 0 1 0 1 1 0 1 1 1 ...
- art: num 1 0 1 0 1 1 0 1 1 1 ...
- textbook: num 1 1 1 1 1 0 1 1 1 1 ...
- dictionary: num 1 1 1 1 1 1 1 1 1 1 ...
- dishwasher: num 1 1 1 1 0 1 1 1 1 1 ...
- PV1MATH: num 562 565 602 520 613 ...
- PV2MATH: num 569 557 594 507 567 ...

brew\_colors 7

- PV3MATH: num 555 553 552 501 585 ...
- PV4MATH: num 579 538 526 521 596 ...
- PV5MATH: num 548 573 619 547 603 ...
- PV1READ: num 582 617 650 554 605 ...
- PV2READ: num 571 572 608 560 557...
- PV3READ: num 602 560 594 517 627...
- PV4READ: num 572 564 575 564 597 ...
- PV5READ: num 585 565 620 572 598 ...
- PV1SCIE: num 583 627 668 574 639 ...
- PV2SCIE: num 579 600 665 612 635 ...
- PV3SCIE : num 593 574 620 571 666 ...
- PV4SCIE: num 567 582 592 598 700 ...
- PV5SCIE: num 587 625 656 662 670 ...
- SENWGT\_STU: num 0.133 0.133 0.141 0.141 0.141 ...
- possessions: num 10 8 12 9 11 11 10 12 12 11 ...

#### Source

http://www.oecd.org/pisa/pisaproducts/database-cbapisa2012.htm

brew\_colors

RColorBrewer Set1 colors

## **Description**

RColorBrewer Set1 colors

## Usage

brew\_colors(col)

## **Arguments**

col

standard color name used to retrieve hex color value

8 eval\_data\_col

broomify

Broomify a model

## **Description**

broom::augment a model and add broom::glance and broom::tidy output as attributes. X and Y variables are also added.

## Usage

```
broomify(model, lmStars = TRUE)
```

## **Arguments**

model model to be sent to broom::augment, broom::glance, and broom::tidy

1mStars boolean that determines if stars are added to labels

## Value

broom::augmented data frame with the broom::glance data.frame and broom::tidy data.frame as 'broom\_glance' and 'broom\_tidy' attributes respectively. var\_x and var\_y variables are also added as attributes

## **Examples**

```
data(mtcars)
model <- stats::lm(mpg ~ wt + qsec + am, data = mtcars)
broomified_model <- broomify(model)
str(broomified_model)</pre>
```

eval\_data\_col

Evaluate data column

## **Description**

Evaluate data column

## Usage

```
eval_data_col(data, aes_col)
```

## Arguments

data set to evaluate the data with

aes\_col Single value from an ggplot2::aes(...) object

find\_plot\_type 9

## Value

Aes mapping with the x and y values switched

## **Examples**

```
mapping <- ggplot2::aes(Petal.Length)
eval_data_col(iris, mapping$x)</pre>
```

find\_plot\_type

Find Plot Types

## Description

Retrieves the type of plot for the specific columns

## Usage

```
find_plot_type(col1Name, col2Name, type1, type2, isAllNa, allowDiag)
```

## Arguments

```
col1Name x column name
col2Name y column name
type1 x column type
type2 y column type
isAllNa is.na(data)
```

allowDiag allow for diag values to be returned

## Author(s)

Barret Schloerke < schloerke@gmail.com>

flea

Historical data used for classification examples.

## Description

This data contains physical measurements on three species of flea beetles.

```
data(flea)
```

fn\_switch

#### **Format**

A data frame with 74 rows and 7 variables

#### **Details**

- species Ch. concinna, Ch. heptapotamica, Ch. heikertingeri
- tars1 width of the first joint of the first tarsus in microns
- tars2 width of the second joint of the first tarsus in microns
- head the maximal width of the head between the external edges of the eyes in 0.01 mm
- aede1 the maximal width of the aedeagus in the fore-part in microns
- aede2 the front angle of the aedeagus (1 unit = 7.5 degrees)
- aede3 the aedeagus width from the side in microns

#### References

Lubischew, A. A. (1962), On the Use of Discriminant Functions in Taxonomy, Biometrics 18:455-477.

fn switch

Function switch

## Description

Function that allows you to call different functions based upon an aesthetic variable value.

## Usage

```
fn_switch(types, mapping_val = "y")
```

## **Arguments**

types

list of functions that follow the ggmatrix function standard: function(data, mapping, . . .) {

#make ggplot2 object }. One key should be a 'default' key for a default switch

case.

mapping\_val

mapping value to switch on. Defaults to the 'y' variable of the aesthetics list.

```
ggnostic_continuous_fn <- fn_switch(list(
  default = ggally_points,
    .fitted = ggally_points,
    .se.fit = ggally_nostic_se_fit,
    .resid = ggally_nostic_resid,
    .hat = ggally_nostic_hat,
    .sigma = ggally_nostic_sigma,
    .cooksd = ggally_nostic_cooksd,</pre>
```

getPlot 11

```
.std.resid = ggally_nostic_std_resid
))

ggnostic_combo_fn <- fn_switch(list(
  default = ggally_box_no_facet,
  fitted = ggally_box_no_facet,
  .se.fit = ggally_nostic_se_fit,
  .resid = ggally_nostic_resid,
  .hat = ggally_nostic_hat,
  .sigma = ggally_nostic_sigma,
  .cooksd = ggally_nostic_cooksd,
  .std.resid = ggally_nostic_std_resid
))</pre>
```

getPlot

getPlot

## Description

Retrieves the ggplot object at the desired location.

## Usage

```
getPlot(pm, i, j)
## S3 method for class 'ggmatrix'
pm[i, j, ...]
```

## **Arguments**

```
pm ggmatrix object to select from
i row from the top
j column from the left
... ignored
```

## Author(s)

Barret Schloerke < schloerke@gmail.com>

```
data(tips, package = "reshape")
plotMatrix2 <- ggpairs(tips[, 3:2], upper = list(combo = "denstrip"))
plotMatrix2[1, 2]</pre>
```

12 ggally\_blank

ggally\_barDiag

Plots the Bar Plots by Using Diagonal

## **Description**

Plots the bar plots by using Diagonal.

## Usage

```
ggally_barDiag(data, mapping, ..., rescale = FALSE)
```

## Arguments

data data set using

mapping aesthetics being used

... other arguments are sent to geom\_bar

rescale boolean to decide whether or not to rescale the count output. Only applies to

numeric data

## Author(s)

Barret Schloerke <schloerke@gmail.com>

## **Examples**

```
data(tips, package = "reshape")
ggally_barDiag(tips, mapping = ggplot2::aes(x = day))
ggally_barDiag(tips, mapping = ggplot2::aes(x = tip), binwidth = 0.25)
```

ggally\_blank

Blank

## **Description**

Draws nothing.

## Usage

```
ggally_blank(...)
ggally_blankDiag(...)
```

## **Arguments**

... other arguments ignored

ggally\_box 13

## **Details**

Makes a "blank" ggplot object that will only draw white space

## Author(s)

Barret Schloerke < schloerke@gmail.com>

ggally\_box

Plots the Box Plot

## **Description**

Make a box plot with a given data set. ggally\_box\_no\_facet will be a single panel plot, while ggally\_box will be a faceted plot

## Usage

```
ggally_box(data, mapping, ...)
ggally_box_no_facet(data, mapping, ...)
```

## **Arguments**

```
data set using
mapping aesthetics being used
... other arguments being supplied to geom_boxplot
```

## Author(s)

Barret Schloerke < schloerke@gmail.com>

14 ggally\_cor

ggally\_cor

Correlation from the Scatter Plot

## **Description**

Estimate correlation from the given data.

## Usage

```
ggally_cor(
  data,
  mapping,
  alignPercent = 0.6,
  method = "pearson",
  use = "complete.obs",
  corAlignPercent = NULL,
  corMethod = NULL,
  corUse = NULL,
  displayGrid = TRUE,
  ...
)
```

## Arguments

data data set using

mapping aesthetics being used

alignPercent right align position of numbers. Default is 60 percent across the horizontal

method supplied to cor function

use use supplied to cor function

 ${\tt corAlignPercent}$ 

deprecated. Use parameter alignPercent

corMethod deprecated. Use parameter method corUse deprecated. Use parameter use

displayGrid if TRUE, display aligned panel gridlines

... other arguments being supplied to geom\_text

## Author(s)

Barret Schloerke < schloerke@gmail.com>

ggally\_density 15

## **Examples**

```
data(tips, package = "reshape")
ggally_cor(tips, mapping = ggplot2::aes_string(x = "total_bill", y = "tip"))
# display with no grid
ggally_cor(
  tips,
  mapping = ggplot2::aes_string(x = "total_bill", y = "tip"),
 displayGrid = FALSE
# change text attributes
ggally_cor(
  tips,
 mapping = ggplot2::aes(x = total_bill, y = tip),
 size = 15,
 colour = I("red")
# split by a variable
ggally_cor(
  tips,
 mapping = ggplot2::aes_string(x = "total_bill", y = "tip", color = "sex"),
  size = 5
)
```

ggally\_density

Plots the Scatter Density Plot

## **Description**

Make a scatter density plot from a given data.

## Usage

```
ggally_density(data, mapping, ...)
```

## **Arguments**

```
data data set using
mapping aesthetics being used
... parameters sent to either stat_density2d or geom_density2d
```

## **Details**

The aesthetic "fill" determines whether or not stat\_density2d (filled) or geom\_density2d (lines) is used.

## Author(s)

Barret Schloerke < schloerke@gmail.com>

16 ggally\_densityDiag

## **Examples**

```
data(tips, package = "reshape")
ggally_density(tips, mapping = ggplot2::aes(x = total_bill, y = tip))
ggally_density(tips, mapping = ggplot2::aes_string(x = "total_bill", y = "tip"))
ggally_density(
   tips,
   mapping = ggplot2::aes_string(x = "total_bill", y = "tip", fill = "..level..")
)
ggally_density(
   tips,
   mapping = ggplot2::aes_string(x = "total_bill", y = "tip", fill = "..level..")
) + ggplot2::scale_fill_gradient(breaks = c(0.05, 0.1, 0.15, 0.2))
```

ggally\_densityDiag

Plots the Density Plots by Using Diagonal

## **Description**

Plots the density plots by using Diagonal.

#### Usage

```
ggally_densityDiag(data, mapping, ..., rescale = FALSE)
```

## **Arguments**

data data set using
mapping aesthetics being used.

... other arguments sent to stat\_density

rescale boolean to decide whether or not to rescale the count output

## Author(s)

Barret Schloerke < schloerke@gmail.com>

```
data(tips, package = "reshape")
ggally_densityDiag(tips, mapping = ggplot2::aes(x = total_bill))
ggally_densityDiag(tips, mapping = ggplot2::aes(x = total_bill, color = day))
```

ggally\_denstrip 17

ggally\_denstrip

Plots a tile plot with facets

## Description

Make Tile Plot as densely as possible.

## Usage

```
ggally_denstrip(data, mapping, ...)
```

## **Arguments**

data data set using

mapping aesthetics being used

... other arguments being sent to stat\_bin

## Author(s)

Barret Schloerke < schloerke@gmail.com>

## **Examples**

```
data(tips, package = "reshape")
ggally_denstrip(tips, mapping = ggplot2::aes(x = total_bill, y = sex))
ggally_denstrip(tips, mapping = ggplot2::aes_string(x = "total_bill", y = "sex"))
ggally_denstrip(
   tips,
   mapping = ggplot2::aes_string(x = "sex", y = "tip", binwidth = "0.2")
) + ggplot2::scale_fill_gradient(low = "grey80", high = "black")
```

ggally\_diagAxis

Internal Axis Labeling Plot for ggpairs

## Description

This function is used when axisLabels == "internal".

18 ggally\_dot

## Usage

```
ggally_diagAxis(
  data,
  mapping,
  label = mapping$x,
  labelSize = 5,
  labelXPercent = 0.5,
  labelYPercent = 0.55,
  labelHJust = 0.5,
  labelVJust = 0.5,
  gridLabelSize = 4,
  ...
)
```

# **Arguments** data

data dataset being plotted

mapping aesthetics being used (x is the variable the plot will be made for)

label title to be displayed in the middle. Defaults to mapping\$x

labelSize size of variable label

labelXPercent percent of horizontal range

labelYPercent percent of vertical range

labelHJust hjust supplied to label

labelVJust vjust supplied to label

... other arguments for geom\_text

size of grid labels

## Author(s)

gridLabelSize

Jason Crowley <crowley.jason.s@gmail.com> and Barret Schloerke

## **Examples**

```
data(tips, package = "reshape")
ggally_diagAxis(tips, ggplot2::aes(x=tip))
ggally_diagAxis(tips, ggplot2::aes(x=sex))
```

ggally\_dot

Plots the Box Plot with Dot

#### **Description**

Add jittering with the box plot. ggally\_dot\_no\_facet will be a single panel plot, while ggally\_dot will be a faceted plot

ggally\_dot\_and\_box 19

## Usage

```
ggally_dot(data, mapping, ...)
ggally_dot_no_facet(data, mapping, ...)
```

## **Arguments**

```
data data set using
mapping aesthetics being used
... other arguments being supplied to geom_jitter
```

#### Author(s)

Barret Schloerke < schloerke@gmail.com>

## **Examples**

```
data(tips, package = "reshape")
ggally_dot(tips, mapping = ggplot2::aes(x = total_bill, y = sex))
ggally_dot(tips, mapping = ggplot2::aes_string(x = "total_bill", y = "sex"))
ggally_dot(
   tips,
   mapping = ggplot2::aes_string(y = "total_bill", x = "sex", color = "sex")
)
ggally_dot(
   tips,
   mapping = ggplot2::aes_string(y = "total_bill", x = "sex", color = "sex", shape = "sex")
) + ggplot2::scale_shape(solid=FALSE)
```

ggally\_dot\_and\_box

Plots either Box Plot or Dot Plots

## **Description**

Place box plots or dot plots on the graph

#### **Usage**

```
ggally_dot_and_box(data, mapping, ..., boxPlot = TRUE)
```

## **Arguments**

```
data data set using
mapping aesthetics being used
... parameters passed to either geom_jitter or geom_boxplot
boxPlot boolean to decide to plot either box plots (TRUE) or dot plots (FALSE)
```

20 ggally\_facetbar

## Author(s)

Barret Schloerke < schloerke@gmail.com>

## **Examples**

```
data(tips, package = "reshape")
ggally_dot_and_box(
  tips,
  mapping = ggplot2::aes(x = total_bill, y = sex, color = sex),
  boxPlot = TRUE
)
ggally_dot_and_box(
  tips,
  mapping = ggplot2::aes(x = total_bill, y = sex, color = sex),
  boxPlot = FALSE
)
```

ggally\_facetbar

Plots the Bar Plots Faceted by Conditional Variable

## Description

X variables are plotted using geom\_bar and faceted by the Y variable.

## Usage

```
ggally_facetbar(data, mapping, ...)
```

## **Arguments**

```
data data set using
mapping aesthetics being used
```

... other arguments are sent to geom\_bar

## Author(s)

Barret Schloerke < schloerke@gmail.com>

```
data(tips, package = "reshape")
ggally_facetbar(tips, ggplot2::aes(x = sex, y = smoker, fill = time))
ggally_facetbar(tips, ggplot2::aes(x = smoker, y = sex, fill = time))
```

ggally\_facetdensity 21

## **Description**

Make density plots by displaying subsets of the data in different panels.

## Usage

```
ggally_facetdensity(data, mapping, ...)
```

## Arguments

```
data data set using
mapping aesthetics being used
... other arguments being sent to stat_density
```

## Author(s)

Barret Schloerke < schloerke@gmail.com>

## **Examples**

```
data(tips, package = "reshape")
ggally_facetdensity(tips, mapping = ggplot2::aes(x = total_bill, y = sex))
ggally_facetdensity(
   tips,
   mapping = ggplot2::aes_string(y = "total_bill", x = "sex", color = "sex")
)
```

```
ggally_facetdensitystrip
```

Plots a density plot with facets or a tile plot with facets

## **Description**

Make Tile Plot as densely as possible.

```
ggally_facetdensitystrip(data, mapping, ..., den_strip = FALSE)
```

22 ggally\_facethist

## Arguments

data data set using

mapping aesthetics being used

... other arguments being sent to either geom\_histogram or stat\_density

den\_strip boolean to decide whether or not to plot a density strip(TRUE) or a facet den-

sity(FALSE) plot.

## Author(s)

Barret Schloerke <schloerke@gmail.com>

## **Examples**

```
example(ggally_facetdensity)
example(ggally_denstrip)
```

ggally\_facethist

Plots the Histograms by Faceting

## **Description**

Make histograms by displaying subsets of the data in different panels.

## Usage

```
ggally_facethist(data, mapping, ...)
```

## **Arguments**

data data set using

mapping aesthetics being used

... parameters sent to stat\_bin()

## Author(s)

Barret Schloerke < schloerke@gmail.com>

```
data(tips, package = "reshape")
ggally_facethist(tips, mapping = ggplot2::aes(x = tip, y = sex))
ggally_facethist(tips, mapping = ggplot2::aes_string(x = "tip", y = "sex"), binwidth = 0.1)
```

ggally\_na 23

ggally_na
-----------

## **Description**

Draws a large NA in the middle of the plotting area. This plot is useful when all X or Y data is NA

## Usage

```
ggally_na(data = NULL, mapping = NULL, size = 10, color = "grey20", ...)
ggally_naDiag(...)
```

## Arguments

```
data ignored
mapping ignored
size size of the geom_text 'NA'
color color of the geom_text 'NA'
... other arguments sent to geom_text
```

#### Author(s)

Barret Schloerke < schloerke@gmail.com>

```
ggally_nostic_cooksd ggnostic - Cook's distance
```

## **Description**

A function to display stats::cooks.distance.

```
ggally_nostic_cooksd(
  data,
  mapping,
  ...,
  linePosition = pf(0.5, length(attr(data, "var_x")), nrow(data) - length(attr(data,
        "var_x"))),
  lineColor = brew_colors("grey"),
  lineType = 2
)
```

24 ggally\_nostic\_hat

#### **Arguments**

#### **Details**

A line is added at  $F_p$ , n - p(0.5) to display the general cutoff point for Cook's Distance.

Reference: Michael H. Kutner, Christopher J. Nachtsheim, John Neter, and William Li. Applied linear statistical models. The McGraw-Hill / Irwin series operations and decision sciences. McGraw-Hill Irwin, 2005, p. 403

#### Value

```
ggplot2 plot object
```

#### See Also

```
stats::cooks.distance
```

## **Examples**

```
dt <- broomify(stats::lm(mpg ~ wt + qsec + am, data = mtcars))
ggally_nostic_cooksd(dt, ggplot2::aes(wt, .cooksd))</pre>
```

```
ggally_nostic_hat
```

ggnostic - leverage points

## **Description**

A function to display stats::influence's hat information against a given explanatory variable.

```
ggally_nostic_hat(
  data,
  mapping,
  ...,
  linePosition = 2 * sum(eval_data_col(data, mapping$y))/nrow(data),
  lineColor = brew_colors("grey"),
  lineSize = 0.5,
  lineAlpha = 1,
  lineType = 2,
  avgLinePosition = sum(eval_data_col(data, mapping$y))/nrow(data),
  avgLineColor = brew_colors("grey"),
  avgLineSize = lineSize,
  avgLineAlpha = lineAlpha,
  avgLineType = 1
)
```

ggally\_nostic\_line 25

## **Arguments**

## **Details**

As stated in stats::influence documentation:

hat: a vector containing the diagonal of the 'hat' matrix.

The diagonal elements of the 'hat' matrix describe the influence each response value has on the fitted value for that same observation.

A suggested "cutoff" line is added to the plot at a height of 2 \* p / n and an expected line at a height of p / n. If either linePosition or avgLinePosition is NULL, the respective line will not be drawn.

#### Value

```
ggplot2 plot object
```

#### See Also

```
stats::influence
```

## **Examples**

```
dt <- broomify(stats::lm(mpg ~ wt + qsec + am, data = mtcars))
ggally_nostic_hat(dt, ggplot2::aes(wt, .hat))</pre>
```

```
ggally_nostic_line
```

ggnostic -background line with geom

## **Description**

If a non-null linePosition value is given, a line will be drawn before the given continuous\_geom or combo\_geom is added to the plot.

```
ggally_nostic_line(
  data,
  mapping,
    ...,
  linePosition = NULL,
  lineColor = "red",
```

26 ggally\_nostic\_resid

```
lineSize = 0.5,
lineAlpha = 1,
lineType = 1,
continuous_geom = ggplot2::geom_point,
combo_geom = ggplot2::geom_boxplot,
mapColorToFill = TRUE
)
```

## **Arguments**

#### **Details**

Functions with a color in their name have different default color behavior.

## Value

```
ggplot2 plot object
```

```
ggally_nostic_resid ggnostic - residuals
```

## **Description**

If non-null pVal and sigma values are given, confidence interval lines will be added to the plot at the specified pVal percentiles of a N(0, sigma) distribution.

```
ggally_nostic_resid(
  data,
  mapping,
    ...,
  linePosition = 0,
  lineColor = brew_colors("grey"),
```

ggally\_nostic\_resid 27

```
lineSize = 0.5,
lineAlpha = 1,
lineType = 1,
lineConfColor = brew_colors("grey"),
lineConfSize = lineSize,
lineConfAlpha = lineAlpha,
lineConfType = 2,
pVal = c(0.025, 0.975),
sigma = attr(data, "broom_glance")$sigma,
se = TRUE,
method = "auto",
formula = y ~ x
```

## **Arguments**

```
data, mapping, ...

parameters supplied to ggally_nostic_line

linePosition, lineColor, lineSize, lineAlpha, lineType

parameters supplied to ggplot2::geom_line

lineConfColor, lineConfSize, lineConfAlpha, lineConfType

parameters supplied to the confidence interval lines

pVal percentiles of a N(0, sigma) distribution to be drawn

sigma sigma value for the pVal percentiles

se boolean to determine if the confidence intervals should be displayed

method, formula

parameters supplied to ggplot2::geom_smooth. Defaults to "auto" and "y ~ x"
```

## Value

ggplot2 plot object

#### See Also

```
stats::residuals
```

```
dt <- broomify(stats::lm(mpg ~ wt + qsec + am, data = mtcars))
ggally_nostic_resid(dt, ggplot2::aes(wt, .resid))</pre>
```

28 ggally\_nostic\_se\_fit

```
ggally_nostic_se_fit ggnostic - fitted value standard error
```

## **Description**

A function to display stats::predict's standard errors

## Usage

```
ggally_nostic_se_fit(
  data,
  mapping,
    ...,
  lineColor = brew_colors("grey"),
  linePosition = NULL
)
```

## **Arguments**

#### **Details**

As stated in stats::predict documentation:

If the logical 'se.fit' is 'TRUE', standard errors of the predictions are calculated. If the numeric argument 'scale' is set (with optional "df"), it is used as the residual standard deviation in the computation of the standard errors, otherwise this is extracted from the model fit.

Since the se.fit is TRUE and scale is unset by default, the standard errors are extracted from the model fit.

A base line of 0 is added to give reference to a perfect fit.

## Value

```
ggplot2 plot object
```

#### See Also

```
stats::influence
```

```
dt <- broomify(stats::lm(mpg ~ wt + qsec + am, data = mtcars))
ggally_nostic_se_fit(dt, ggplot2::aes(wt, .se.fit))</pre>
```

ggally\_nostic\_sigma 29

```
ggally_nostic_sigma
```

ggnostic - leave one out model sigma

## Description

A function to display stats::influence's sigma value.

## Usage

```
ggally_nostic_sigma(
  data,
  mapping,
  ...,
  lineColor = brew_colors("grey"),
  linePosition = attr(data, "broom_glance")$sigma
)
```

## **Arguments**

line Position

line that is drawn in the background of the plot. Defaults to the overall model's sigma value.

## **Details**

As stated in stats::influence documentation:

sigma: a vector whose i-th element contains the estimate of the residual standard deviation obtained when the i-th case is dropped from the regression. (The approximations needed for GLMs can result in this being 'NaN'.)

A line is added to display the overall model's sigma value. This gives a baseline for comparison

## Value

```
ggplot2 plot object
```

## See Also

```
stats::influence
```

```
dt <- broomify(stats::lm(mpg ~ wt + qsec + am, data = mtcars))
ggally_nostic_sigma(dt, ggplot2::aes(wt, .sigma))</pre>
```

30 ggally\_points

## Description

If non-null pVal and sigma values are given, confidence interval lines will be added to the plot at the specified pVal locations of a N(0, 1) distribution.

## Usage

```
ggally_nostic_std_resid(data, mapping, ..., sigma = 1)
```

## Arguments

```
data, mapping, ...

parameters supplied to ggally_nostic_resid

sigma sigma value for the pVal percentiles. Set to 1 for standardized residuals
```

#### Value

```
ggplot2 plot object
```

## See Also

```
stats::rstandard
```

## **Examples**

```
dt <- broomify(stats::lm(mpg ~ wt + qsec + am, data = mtcars))
ggally_nostic_std_resid(dt, ggplot2::aes(wt, .std.resid))</pre>
```

ggally\_points

Plots the Scatter Plot

## **Description**

Make a scatter plot with a given data set.

```
ggally_points(data, mapping, ...)
```

ggally\_ratio 31

## **Arguments**

```
data data set using
mapping aesthetics being used
... other arguments are sent to geom_point
```

## Author(s)

Barret Schloerke < schloerke@gmail.com>

## **Examples**

ggally\_ratio

Plots a mosaic plot

## **Description**

Plots the mosaic plot by using fluctuation.

## Usage

```
ggally_ratio(
  data,
  mapping = do.call(ggplot2::aes_string, as.list(colnames(data)[1:2])),
  ...,
  floor = 0,
  ceiling = NULL
)
```

## **Arguments**

```
data set using
mapping aesthetics being used. Only x and y will used and both are required
... passed to geom_tile(...)
floor don't display cells smaller than this value
ceiling max value to scale frequencies. If any frequency is larger than the ceiling, the
fill color is displayed darker than other rectangles
```

32 ggally\_smooth

## Author(s)

Barret Schloerke <schloerke@gmail.com>

## **Examples**

```
data(tips, package = "reshape")
ggally_ratio(tips, ggplot2::aes(sex, day))
ggally_ratio(tips, ggplot2::aes(sex, day)) + ggplot2::coord_equal()
# only plot tiles greater or equal to 20 and scale to a max of 50
ggally_ratio(
  tips, ggplot2::aes(sex, day),
  floor = 20, ceiling = 50
) + ggplot2::theme(aspect.ratio = 4/2)
```

ggally\_smooth

Plots the Scatter Plot with Smoothing

## **Description**

Add a smoothed condition mean with a given scatter plot.

## Usage

```
ggally_smooth(
  data,
  mapping,
  ...,
  method = "lm",
  formula = y ~ x,
  se = TRUE,
  shrink = TRUE
)

ggally_smooth_loess(data, mapping, ...)

ggally_smooth_lm(data, mapping, ...)
```

## Arguments

```
data set using
mapping aesthetics being used
method, se parameters supplied to geom_smooth
formula, ... other arguments to add to geom_smooth
shrink boolean to determine if y range is reduced to range of points or points and error ribbon
```

ggally\_text 33

## **Details**

Y limits are reduced to match original Y range with the goal of keeping the Y axis the same across plots.

## Author(s)

Barret Schloerke < schloerke@gmail.com>

## **Examples**

```
data(tips, package = "reshape")
ggally_smooth(tips, mapping = ggplot2::aes(x = total_bill, y = tip))
ggally_smooth(tips, mapping = ggplot2::aes_string(x = "total_bill", y = "tip"))
ggally_smooth(tips, mapping = ggplot2::aes_string(x = "total_bill", y = "tip", color = "sex"))
```

ggally\_text

Text Plot

## **Description**

Plot text for a plot.

## Usage

```
ggally_text(
  label,
  mapping = ggplot2::aes(color = "black"),
  xP = 0.5,
  yP = 0.5,
  xrange = c(0, 1),
  yrange = c(0, 1),
  ...
)
```

## **Arguments**

```
label text that you want to appear
mapping aesthetics that don't relate to position (such as color)

xP horizontal position percentage

yP vertical position percentage

xrange range of the data around it. Only nice to have if plotting in a matrix

yrange range of the data around it. Only nice to have if plotting in a matrix

other arguments for geom_text
```

## Author(s)

Barret Schloerke < schloerke@gmail.com>

34 ggcoef

## **Examples**

```
ggally_text("Example 1")
ggally_text("Example\nTwo", mapping = ggplot2::aes(size = 15), color = I("red"))
```

ggcoef

ggcoef - Plot Model Coefficients with broom and ggplot2

## **Description**

Plot the coefficients of a model with **broom** and **ggplot2**.

## Usage

```
ggcoef(
  Х,
 mapping = aes_string(y = "term", x = "estimate"),
 conf.int = TRUE,
  conf.level = 0.95,
  exponentiate = FALSE,
  exclude_intercept = FALSE,
  vline = TRUE,
  vline_intercept = "auto",
  vline_color = "gray50",
  vline_linetype = "dotted",
  vline_size = 1,
  errorbar_color = "gray25",
  errorbar_height = 0,
  errorbar_linetype = "solid",
 errorbar_size = 0.5,
  sort = c("none", "ascending", "decending"),
)
```

#### **Arguments**

```
a model object to be tidied with tidy or a data frame (see Details)
Х
                   default aesthetic mapping
mapping
conf.int
                   display confidence intervals as error bars?
conf.level
                  level of confidence intervals (passed to tidy if x is not a data frame)
                  if TRUE, x-axis will be logarithmic (also passed to tidy if x is not a data frame)
exponentiate
exclude_intercept
                   should the intercept be excluded from the plot?
                   print a vertical line?
vline
vline_intercept
                   xintercept for the vertical line. "auto" for x = 0 (or x = 1 if exponentiate is
                   TRUE)
```

ggcorr 35

```
color of the vertical line
vline_color
vline_linetype line type of the vertical line
vline_size
                  size of the vertical line
errorbar_color color of the error bars
errorbar_height
                  height of the error bars
errorbar_linetype
                  line type of the error bars
                  size of the error bars
errorbar_size
sort
                   "none" (default) do not sort, "ascending" sort by increasing coefficient value,
                  or "decending" sort by decreasing coefficient value
                   additional arguments sent to geom_point
. . .
```

#### **Examples**

```
library(broom)
reg <- lm(Sepal.Length ~ Sepal.Width + Petal.Length + Petal.Width, data = iris)
ggcoef(reg)
d <- as.data.frame(Titanic)
reg2 <- glm(Survived ~ Sex + Age + Class, family = binomial, data = d, weights = d$Freq)
ggcoef(reg2, exponentiate = TRUE)
ggcoef(
  reg2, exponentiate = TRUE, exclude_intercept = TRUE,
    errorbar_height = .2, color = "blue", sort = "ascending"
)</pre>
```

ggcorr

ggcorr - Plot a correlation matrix with ggplot2

## **Description**

Function for making a correlation matrix plot, using ggplot2. The function is directly inspired by Tian Zheng and Yu-Sung Su's corrplot function in the 'arm' package. Please visit <a href="http://github.com/briatte/ggcorr">http://github.com/briatte/ggcorr</a> for the latest version of ggcorr, and see the vignette at <a href="https://briatte.github.io/ggcorr">https://briatte.github.io/ggcorr</a> for many examples of how to use it.

```
ggcorr(
  data,
  method = c("pairwise", "pearson"),
  cor_matrix = NULL,
  nbreaks = NULL,
  digits = 2,
  name = "",
  low = "#3B9AB2",
```

36 ggcorr

```
mid = "#EEEEEE",
 high = "#F21A00",
  midpoint = 0,
  palette = NULL,
  geom = "tile",
 min_size = 2,
 max_size = 6,
  label = FALSE,
  label_alpha = FALSE,
  label_color = "black",
  label\_round = 1,
  label_size = 4,
  limits = c(-1, 1),
  drop = is.null(limits) || identical(limits, FALSE),
  layout.exp = 0,
  legend.position = "right",
  legend.size = 9,
)
```

## Arguments

method

data a data frame or matrix containing numeric (continuous) data. If any of the

columns contain non-numeric data, they will be dropped with a warning.

a vector of two character strings. The first value gives the method for computing covariances in the presence of missing values, and must be (an abbreviation of) one of "everything", "all.obs", "complete.obs", "na.or.complete" or "pairwise.complete.obs". The second value gives the type of correlation coefficient to compute, and must be one of "pearson", "kendall" or "spearman".

See cor for details. Defaults to c("pairwise", "pearson").

cor\_matrix the named correlation matrix to use for calculations. Defaults to the correlation

matrix of data when data is supplied.

nbreaks the number of breaks to apply to the correlation coefficients, which results in

a categorical color scale. See 'Note'. Defaults to NULL (no breaks, continuous

scaling).

digits the number of digits to show in the breaks of the correlation coefficients: see

cut for details. Defaults to 2.

name a character string for the legend that shows the colors of the correlation coeffi-

cients. Defaults to "" (no legend name).

low the lower color of the gradient for continuous scaling of the correlation coeffi-

cients. Defaults to "#3B9AB2" (blue).

mid the midpoint color of the gradient for continuous scaling of the correlation co-

efficients. Defaults to "#EEEEEE" (very light grey).

high the upper color of the gradient for continuous scaling of the correlation coeffi-

cients. Defaults to "#F21A00" (red).

midpoint the midpoint value for continuous scaling of the correlation coefficients. De-

faults to 0.

ggcorr 37

	palette	if nbreaks is used, a ColorBrewer palette to use instead of the colors specified by low, mid and high. Defaults to NULL.
	geom	the geom object to use. Accepts either "tile", "circle", "text" or "blank".
	min_size	when geom has been set to "circle", the minimum size of the circles. Defaults to 2.
	max_size	when ${\tt geom}$ has been set to "circle", the maximum size of the circles. Defaults to 6.
	label	whether to add correlation coefficients to the plot. Defaults to FALSE.
	label_alpha	whether to make the correlation coefficients increasingly transparent as they come close to 0. Also accepts any numeric value between 0 and 1, in which case the level of transparency is set to that fixed value. Defaults to FALSE (no transparency).
	label_color	the color of the correlation coefficients. Defaults to "grey75".
	label_round	the decimal rounding of the correlation coefficients. Defaults to 1.
	label_size	the size of the correlation coefficients. Defaults to 4.
	limits	bounding of color scaling for correlations, set limits = $\ensuremath{NULL}$ or $\ensuremath{FALSE}$ to remove
	drop	if using nbreaks, whether to drop unused breaks from the color scale. Defaults to ${\sf FALSE}$ (recommended).
	layout.exp	a multiplier to expand the horizontal axis to the left if variable names get clipped. Defaults to $\emptyset$ (no expansion).
legend.position		
		where to put the legend of the correlation coefficients: see theme for details. Defaults to "bottom".
	legend.size	the size of the legend title and labels, in points: see theme for details. Defaults to $9. $
		other arguments supplied to geom_text for the diagonal labels.

# Note

Recommended values for the nbreaks argument are 3 to 11, as values above 11 are visually difficult to separate and are not supported by diverging ColorBrewer palettes.

# Author(s)

Francois Briatte, with contributions from Amos B. Elberg and Barret Schloerke

# See Also

cor and corrplot in the arm package.

## **Examples**

```
# Basketball statistics provided by Nathan Yau at Flowing Data.
dt <- read.csv("http://datasets.flowingdata.com/ppg2008.csv")</pre>
# Default output.
ggcorr(dt[, -1])
# Labelled output, with coefficient transparency.
ggcorr(dt[, -1],
       label = TRUE,
       label_alpha = TRUE)
# Custom options.
ggcorr(
  dt[, -1],
  name = expression(rho),
  geom = "circle",
 max_size = 10,
 min_size = 2,
  size = 3,
  hjust = 0.75,
  nbreaks = 6,
  angle = -45,
  palette = "PuOr" # colorblind safe, photocopy-able
# Supply your own correlation matrix
ggcorr(
  data = NULL,
  cor_matrix = cor(dt[, -1], use = "pairwise")
)
```

ggduo

ggduo - A ggplot2 generalized pairs plot for two columns sets of a data.frame

## **Description**

Make a matrix of plots with a given data set with two different column sets

# Usage

```
ggduo(
  data,
  mapping = NULL,
  columnsX = 1:ncol(data),
  columnsY = 1:ncol(data),
  title = NULL,
  types = list(continuous = "smooth_loess", comboVertical = "box_no_facet",
```

```
comboHorizontal = "facethist", discrete = "ratio"),
axisLabels = c("show", "none"),
columnLabelsX = colnames(data[columnsX]),
columnLabelsY = colnames(data[columnsY]),
labeller = "label_value",
switch = NULL,
xlab = NULL,
ylab = NULL,
showStrips = NULL,
legend = NULL,
cardinality_threshold = 15,
progress = NULL,
legends = stop("deprecated")
)
```

#### **Arguments**

data data set using. Can have both numerical and categorical data.

mapping aesthetic mapping (besides x and y). See aes(). If mapping is numeric, columns

will be set to the mapping value and mapping will be set to NULL.

columnsX, columnsY

which columns are used to make plots. Defaults to all columns.

title, xlab, ylab

title, x label, and y label for the graph

types see Details

axisLabels either "show" to display axisLabels or "none" for no axis labels

columnLabelsX, columnLabelsY

label names to be displayed. Defaults to names of columns being used.

labeller labeller for facets. See labellers. Common values are "label\_value" (de-

fault) and "label\_parsed".

switch switch parameter for facet\_grid. See ggplot2::facet\_grid. By default, the

labels are displayed on the top and right of the plot. If "x", the top labels will be displayed to the bottom. If "y", the right-hand side labels will be displayed

to the left. Can also be set to "both"

showStrips boolean to determine if each plot's strips should be displayed. NULL will default

to the top and right side plots only. TRUE or FALSE will turn all strips on or off

respectively.

legend May be the two objects described below or the default NULL value. The legend

position can be moved by using ggplot2's theme element pm + theme(legend.position

= "bottom")

a numeric vector of length 2 provides the location of the plot to use the legend for the plot matrix's legend. Such as legend = c(3,5) which will use the legend from the plot in the third row and fifth column

a single numeric value provides the location of a plot according to the display order. Such as legend = 3 in a plot matrix with 2 rows and 5 columns displayed by column will return the plot in position c(1,2)

a object from grab\_legend() a predetermined plot legend that will be displayed directly

cardinality\_threshold

maximum number of levels allowed in a character / factor column. Set this value

to NULL to not check factor columns. Defaults to 15

progress NULL (default) for a progress bar in interactive sessions with more than 15 plots,

TRUE for a progress bar, FALSE for no progress bar, or a function that accepts at

least a plot matrix and returns a new progress::progress\_bar. See ggmatrix\_progress.

legends deprecated

#### **Details**

types is a list that may contain the variables 'continuous', 'combo', 'discrete', and 'na'. Each element of the list may be a function or a string. If a string is supplied, it must implement one of the following options:

**continuous** exactly one of ('points', 'smooth', 'smooth\_loess', 'density', 'cor', 'blank'). This option is used for continuous X and Y data.

**comboHorizontal** exactly one of ('box', 'box\_no\_facet', 'dot', 'dot\_no\_facet', 'facethist', 'facetdensity', 'denstrip', 'blank'). This option is used for either continuous X and categorical Y data or categorical X and continuous Y data.

**comboVertical** exactly one of ('box', 'box\_no\_facet', 'dot', 'dot\_no\_facet', 'facethist', 'facetdensity', 'denstrip', 'blank'). This option is used for either continuous X and categorical Y data or categorical X and continuous Y data.

**discrete** exactly one of ('facetbar', 'ratio', 'blank'). This option is used for categorical X and Y data.

**na** exactly one of ('na', 'blank'). This option is used when all X data is NA, all Y data is NA, or either all X or Y data is NA.

If 'blank' is ever chosen as an option, then ggduo will produce an empty plot.

If a function is supplied as an option, it should implement the function api of function(data, mapping,...){#make ggplot2 plot}. If a specific function needs its parameters set, wrap(fn,param1 = val1,param2 = val2) the function with its parameters.

## **Examples**

```
# small function to display plots only if it's interactive
p_ <- GGally::print_if_interactive

data(baseball, package = "plyr")

# Keep players from 1990-1995 with at least one at bat
# Add how many singles a player hit
# (must do in two steps as X1b is used in calculations)
dt <- transform(
   subset(baseball, year >= 1990 & year <= 1995 & ab > 0),
    X1b = h - X2b - X3b - hr
)
```

```
# Add
 # the player's batting average,
# the player's slugging percentage,
# and the player's on base percentage
 # Make factor a year, as each season is discrete
 dt <- transform(</pre>
  dt,
  batting_avg = h / ab,
  slug = (X1b + 2*X2b + 3*X3b + 4*hr) / ab,
  on_base = (h + bb + hbp) / (ab + bb + hbp),
  year = as.factor(year)
 pm <- ggduo(
  dt,
  c("year", "g", "ab", "lg"),
  c("batting_avg", "slug", "on_base"),
  mapping = ggplot2::aes(color = lg)
 # Prints, but
 # there is severe over plotting in the continuous plots
 # the labels could be better
 # want to add more hitting information
 p_{p}(pm)
 # address overplotting issues and add a title
 pm <- ggduo(
  dt,
  c("year", "g", "ab", "lg"),
  c("batting_avg", "slug", "on_base"),
  columnLabelsX = c("year", "player game count", "player at bat count", "league"),
  columnLabelsY = c("batting avg", "slug %", "on base %"),
   title = "Baseball Hitting Stats from 1990-1995",
  mapping = ggplot2::aes(color = lg),
   types = list(
     # change the shape and add some transparency to the points
    continuous = wrap("smooth_loess", alpha = 0.50, shape = "+")
  showStrips = FALSE
 );
p_{p}(pm)
# Example derived from:
## R Data Analysis Examples | Canonical Correlation Analysis. UCLA: Institute for Digital
## Research and Education.
## from http://www.stats.idre.ucla.edu/r/dae/canonical-correlation-analysis
## (accessed May 22, 2017).
# "Example 1. A researcher has collected data on three psychological variables, four
# academic variables (standardized test scores) and gender for 600 college freshman.
```

```
# She is interested in how the set of psychological variables relates to the academic
# variables and gender. In particular, the researcher is interested in how many
# dimensions (canonical variables) are necessary to understand the association between
# the two sets of variables."
data(psychademic)
summary(psychademic)
(psych_variables <- attr(psychademic, "psychology"))</pre>
(academic_variables <- attr(psychademic, "academic"))</pre>
## Within correlation
p_(ggpairs(psychademic, columns = psych_variables))
p_(ggpairs(psychademic, columns = academic_variables))
## Between correlation
loess_with_cor <- function(data, mapping, ..., method = "pearson") {</pre>
 x <- eval_data_col(data, mapping$x)</pre>
 y <- eval_data_col(data, mapping$y)</pre>
 cor <- cor(x, y, method = method)
 ggally_smooth_loess(data, mapping, ...) +
   ggplot2::geom_label(
      data = data.frame(
       x = min(x, na.rm = TRUE),
       y = max(y, na.rm = TRUE),
       lab = round(cor, digits = 3)
      mapping = ggplot2::aes(x = x, y = y, label = lab),
      hjust = 0, vjust = 1,
      size = 5, fontface = "bold",
      inherit.aes = FALSE # do not inherit anything from the ...
   )
}
pm <- ggduo(
 psychademic,
 rev(psych_variables), academic_variables,
 types = list(continuous = loess_with_cor),
 showStrips = FALSE
)
suppressWarnings(p_(pm)) # ignore warnings from loess
# add color according to sex
pm <- ggduo(
 psychademic,
 mapping = ggplot2::aes(color = sex),
 rev(psych_variables), academic_variables,
 types = list(continuous = loess_with_cor),
 showStrips = FALSE,
 legend = c(5,2)
)
suppressWarnings(p_(pm))
# add color according to sex
```

ggfacet 43

```
pm <- ggduo(
   psychademic,
   mapping = ggplot2::aes(color = motivation),
   rev(psych_variables), academic_variables,
   types = list(continuous = loess_with_cor),
   showStrips = FALSE,
   legend = c(5,2)
) +
   ggplot2::theme(legend.position = "bottom")
suppressWarnings(p_(pm))</pre>
```

ggfacet

ggfacet - single ggplot2 plot matrix with facet\_grid

# Description

ggfacet - single ggplot2 plot matrix with facet\_grid

# Usage

```
ggfacet(
  data,
  mapping = NULL,
  columnsX = 1:ncol(data),
  columnsY = 1:ncol(data),
  fn = ggally_points,
    ...,
  columnLabelsX = names(data[columnsX]),
  columnLabelsY = names(data[columnsY]),
  xlab = NULL,
  ylab = NULL,
  title = NULL,
  scales = "free"
)
```

#### **Arguments**

data frame that contains all columns to be displayed. This data will be melted before being passed into the function fn

mapping aesthetic mapping (besides x and y). See aes()

columnsX columns to be displayed in the plot matrix

columnsY rows to be displayed in the plot matrix

fn function to be executed. Similar to ggpairs and ggduo, the function may either be a string identifier or a real function that wrap understands.

... extra arguments passed directly to fn

columnLabelsX, columnLabelsY

column and row labels to display in the plot matrix

44 gglegend

# **Examples**

```
# Small function to display plots only if it's interactive
p_ <- GGally::print_if_interactive</pre>
if (requireNamespace("chemometrics", quietly = TRUE)) {
 data(NIR, package = "chemometrics")
 NIR_sub <- data.frame(NIR$yGlcEtOH, NIR$xNIR[,1:3])</pre>
 str(NIR_sub)
 x_cols <- c("X1115.0", "X1120.0", "X1125.0")</pre>
 y_cols <- c("Glucose", "Ethanol")</pre>
 # using ggduo directly
 p <- ggduo(NIR_sub, x_cols, y_cols, types = list(continuous = "points"))</pre>
 p_(p)
 # using ggfacet
 p <- ggfacet(NIR_sub, x_cols, y_cols)</pre>
 p_(p)
 # add a smoother
 p <- ggfacet(NIR_sub, x_cols, y_cols, fn = 'smooth_loess')</pre>
 p_(p)
 # same output
 p <- ggfacet(NIR_sub, x_cols, y_cols, fn = ggally_smooth_loess)</pre>
 # Change scales to be the same in for every row and for every column
 p <- ggfacet(NIR_sub, x_cols, y_cols, scales = "fixed")</pre>
 p_(p)
}
```

gglegend

Plot only legend of plot function

# **Description**

Plot only legend of plot function

# Usage

```
gglegend(fn)
```

#### **Arguments**

fn

this value is passed directly to an empty wrap call. Please see ?wrap for more details.

ggmatrix 45

## Value

a function that when called with arguments will produce the legend of the plotting function supplied.

## **Examples**

```
# display regular plot
ggally_points(iris, ggplot2::aes(Sepal.Length, Sepal.Width, color = Species))
# Make a function that will only print the legend
points_legend <- gglegend(ggally_points)</pre>
points_legend(iris, ggplot2::aes(Sepal.Length, Sepal.Width, color = Species))
# produce the sample legend plot, but supply a string that 'wrap' understands
same_points_legend <- gglegend("points")</pre>
identical(
  attr(attr(points_legend, "fn"), "original_fn"),
  attr(attr(same_points_legend, "fn"), "original_fn")
# Complicated examples
custom_legend <- wrap(gglegend("points"), size = 6)</pre>
custom_legend(iris, ggplot2::aes(Sepal.Length, Sepal.Width, color = Species))
# Use within ggpairs
pm <- ggpairs(</pre>
  iris, 1:2,
  mapping = ggplot2::aes(color = Species),
  upper = list(continuous = gglegend("points"))
)
# pm
# Place a legend in a specific location
pm <- ggpairs(iris, 1:2, mapping = ggplot2::aes(color = Species))</pre>
# Make the legend
pm[1,2] <- points_legend(iris, ggplot2::aes(Sepal.Width, Sepal.Length, color = Species))</pre>
```

ggmatrix

ggmatrix - A ggplot2 Matrix

## **Description**

Make a generic matrix of ggplot2 plots.

#### Usage

```
ggmatrix(
  plots,
  nrow,
```

46 ggmatrix

ncol,

```
xAxisLabels = NULL,
      vAxisLabels = NULL,
      title = NULL,
      xlab = NULL,
      ylab = NULL,
      byrow = TRUE,
      showStrips = NULL,
      showAxisPlotLabels = TRUE,
      showXAxisPlotLabels = TRUE,
      showYAxisPlotLabels = TRUE,
      labeller = NULL,
      switch = NULL,
      xProportions = NULL,
      yProportions = NULL,
      progress = NULL,
      data = NULL,
      gg = NULL,
      legend = NULL
Arguments
    plots
                      list of plots to be put into matrix
                      number of rows and columns
    nrow, ncol
    xAxisLabels, yAxisLabels
                      strip titles for the x and y axis respectively. Set to NULL to not be displayed
    title, xlab, ylab
                      title, x label, and y label for the graph. Set to NULL to not be displayed
                      boolean that determines whether the plots should be ordered by row or by col-
    byrow
                      umn
    showStrips
                      boolean to determine if each plot's strips should be displayed. NULL will default
                      to the top and right side plots only. TRUE or FALSE will turn all strips on or off
                      respectively.
    showAxisPlotLabels, showXAxisPlotLabels, showYAxisPlotLabels
                      booleans that determine if the plots axis labels are printed on the X (bottom) or Y
                      (left) part of the plot matrix. If showAxisPlotLabels is set, both showXAxisPlotLabels
                      and showYAxisPlotLabels will be set to the given value.
    labeller
                      labeller for facets. See labellers. Common values are "label_value" (de-
                      fault) and "label_parsed".
                      switch parameter for facet_grid. See ggplot2::facet_grid. By default, the
    switch
                      labels are displayed on the top and right of the plot. If "x", the top labels will
                      be displayed to the bottom. If "y", the right-hand side labels will be displayed
                      to the left. Can also be set to "both"
    xProportions, yProportions
                      Value to change how much area is given for each plot. Either NULL (default),
                      numeric value matching respective length, or grid::unit object with matching
```

respective length

ggmatrix 47

progress	NULL (default) for a progress bar in interactive sessions with more than 15 plots, TRUE for a progress bar, FALSE for no progress bar, or a function that accepts at least a plot matrix and returns a new progress::progress_bar. See ggmatrix_progress.
data	data set using. This is the data to be used in place of 'ggally_data' if the plot is a string to be evaluated at print time
gg	ggplot2 theme objects to be applied to every plot
legend	May be the two objects described below or the default NULL value. The legend position can be moved by using ggplot2's theme element pm + theme(legend.position = "bottom")
	a numeric vector of length 2 provides the location of the plot to use the legend for the plot matrix's legend. Such as legend = c(3,5) which will use the legend from the plot in the third row and fifth column
	a single numeric value provides the location of a plot according to the display order. Such as legend = 3 in a plot matrix with 2 rows and 5 columns displayed by column will return the plot in position c(1,2)
	<pre>a object from grab_legend() a predetermined plot legend that will be dis- played directly</pre>

# Memory usage

Now that the print.ggmatrix method uses a large gtable object, rather than print each plot independently, memory usage may be of concern. From small tests, memory usage flutters around object.size(data) \* 0.3 \* length(plots). So, for a 80Mb random noise dataset with 100 plots, about 2.4 Gb of memory needed to print. For the 3.46 Mb diamonds dataset with 100 plots, about 100 Mb of memory was needed to print. The benefits of using the ggplot2 format greatly outweigh the price of about 20

# Author(s)

Barret Schloerke < schloerke@gmail.com>

# **Examples**

```
# small function to display plots only if it's interactive
p_ <- GGally::print_if_interactive

plotList <- list()
for (i in 1:6) {
    plotList[[i]] <- ggally_text(paste("Plot #", i, sep = ""))
}
pm <- ggmatrix(
    plotList,
    2, 3,
    c("A", "B", "C"),
    c("D", "E"),
    byrow = TRUE
)
p_(pm)</pre>
```

48 ggmatrix\_gtable

```
pm <- ggmatrix(
  plotList,
  2, 3,
  xAxisLabels = c("A", "B", "C"),
  yAxisLabels = NULL,
  byrow = FALSE,
  showXAxisPlotLabels = FALSE
)
p_(pm)</pre>
```

ggmatrix\_gtable

Print ggmatrix object

# Description

Specialized method to print the ggmatrix object-

# Usage

```
ggmatrix_gtable(
  pm,
  ...,
  progress = NULL,
  progress_format = formals(ggmatrix_progress)$format
)
```

# **Arguments**

Please use the 'progress' parameter in your ggmatrix-like function. See ggmatrix\_progress for a few examples. These parameters will soon be deprecated.

## Author(s)

Barret Schloerke < schloerke@gmail.com>

# **Examples**

```
data(tips, package = "reshape")
pm <- ggpairs(tips, c(1,3,2), mapping = ggplot2::aes_string(color = "sex"))
ggmatrix_gtable(pm)</pre>
```

ggmatrix\_progress 49

ggmatrix\_progress

ggmatrix default progress bar

## **Description**

ggmatrix default progress bar

#### Usage

```
ggmatrix_progress(
  format = " plot: [:plot_i,:plot_j] [:bar]:percent est::eta ",
  clear = TRUE,
  show_after = 0,
  ...
)
```

# **Arguments**

#### Value

function that accepts a plot matrix as the first argument and . . . for future expansion. Internally, the plot matrix is used to determine the total number of plots for the progress bar.

# **Examples**

```
p_ <- GGally::print_if_interactive

pm <- ggpairs(iris, 1:2, progress = ggmatrix_progress())
p_(pm)

# does not clear after finishing
pm <- ggpairs(iris, 1:2, progress = ggmatrix_progress(clear = FALSE))
p_(pm)</pre>
```

ggnet

ggnet - Plot a network with ggplot2

## **Description**

Function for plotting network objects using ggplot2, now replaced by the ggnet2 function, which provides additional control over plotting parameters. Please visit <a href="http://github.com/briatte/ggnet">http://github.com/briatte/ggnet</a> for the latest version of ggnet2, and <a href="https://briatte.github.io/ggnet">https://github.io/ggnet</a> for a vignette that contains many examples and explanations.

# Usage

```
ggnet(
  net,
  mode = "fruchtermanreingold",
  layout.par = NULL,
  layout.exp = 0,
  size = 9,
  alpha = 1,
  weight = "none",
  weight.legend = NA,
  weight.method = weight,
  weight.min = NA,
  weight.max = NA,
  weight.cut = FALSE,
  group = NULL,
  group.legend = NA,
  node.group = group,
  node.color = NULL,
  node.alpha = alpha,
  segment.alpha = alpha,
  segment.color = "grey50",
  segment.label = NULL,
  segment.size = 0.25,
  arrow.size = 0,
  arrow.gap = 0,
  arrow.type = "closed",
  label = FALSE,
  label.nodes = label,
  label.size = size/2,
  label.trim = FALSE,
  legend.size = 9,
  legend.position = "right",
  names = c("", ""),
  quantize.weights = FALSE,
  subset.threshold = 0,
  top8.nodes = FALSE,
  trim.labels = FALSE,
)
```

## **Arguments**

net

an object of class network, or any object that can be coerced to this class, such as an adjacency or incidence matrix, or an edge list: see edgeset.constructors and network for details. If the object is of class igraph and the intergraph package is installed, it will be used to convert the object: see asNetwork for details.

mode

a placement method from those provided in the sna package: see gplot.layout

for details. Also accepts the names of two numeric vertex attributes of net, or a matrix of numeric coordinates, in which case the first two columns of the matrix are used. Defaults to the Fruchterman-Reingold force-directed algorithm.

layout.par options to be passed to the placement method, as listed in gplot.layout. Defaults

to NULL.

layout.exp a multiplier to expand the horizontal axis if node labels get clipped: see ex-

pand\_range for details. Defaults to 0 (no expansion).

size size of the network nodes. If the nodes are weighted, their area is proportionally

scaled up to the size set by size. Defaults to 9.

alpha a level of transparency for nodes, vertices and arrows. Defaults to 1.

weight the weighting method for the nodes, which might be a vertex attribute or a

vector of size values. Also accepts "indegree", "outdegree", "degree" or "freeman" to size the nodes by their unweighted degree centrality ("degree" and "freeman" are equivalent): see degree for details. All node weights must

be positive. Defaults to "none" (no weighting).

weight.legend the name to assign to the legend created by weight. Defaults to NA (no name).

weight.method see weight

weight.min whether to subset the network to nodes with a minimum size, based on the values

of weight. Defaults to NA (preserves all nodes).

weight.max whether to subset the network to nodes with a maximum size, based on the

values of weight. Defaults to NA (preserves all nodes).

weight.cut whether to cut the size of the nodes into a certain number of quantiles. Accepts

TRUE, which tries to cut the sizes into quartiles, or any positive numeric value, which tries to cut the sizes into that many quantiles. If the size of the nodes do not contain the specified number of distinct quantiles, the largest possible number is used. See quantile and cut for details. Defaults to FALSE (does

nothing).

group the groups of the nodes, either as a vector of values or as a vertex attribute. If

set to mode on a bipartite network, the nodes will be grouped as "actor" if they belong to the primary mode and "event" if they belong to the secondary mode.

group. legend the name to assign to the legend created by group.

node.group see group

node.color a vector of character strings to color the nodes with, holding as many colors as

there are levels in node.group. Defaults to NULL, which will assign grayscale

colors to each group.

node.alpha transparency of the nodes. Inherits from alpha.

segment.alpha the level of transparency of the edges. Defaults to alpha, which defaults to 1.

segment.color the color of the edges, as a color value, a vector of color values, or as an edge

attribute containing color values. Defaults to "grey50".

segment.label the labels to plot at the middle of the edges, as a single value, a vector of values,

or as an edge attribute. Defaults to NULL (no edge labels).

segment.size the size of the edges, in points, as a single numeric value, a vector of values, or

as an edge attribute. Defaults to 0.25.

arrow.size the size of the arrows for directed network edges, in points. See arrow for details. Defaults to 0 (no arrows). a setting aimed at improving the display of edge arrows by plotting slightly arrow.gap shorter edges. Accepts any value between 0 and 1, where a value of 0.05 will generally achieve good results when the size of the nodes is reasonably small. Defaults to 0 (no shortening). the type of the arrows for directed network edges. See arrow for details. Dearrow.type faults to "closed". label whether to label the nodes. If set to TRUE, nodes are labeled with their vertex names. If set to a vector that contains as many elements as there are nodes in net, nodes are labeled with these. If set to any other vector of values, the nodes are labeled only when their vertex name matches one of these values. Defaults to FALSE (no labels). label.nodes see label label.size the size of the node labels, in points, as a numeric value, a vector of numeric values, or as a vertex attribute containing numeric values. Defaults to size / 2 (half the maximum node size), which defaults to 6. label.trim whether to apply some trimming to the node labels. Accepts any function that can process a character vector, or a strictly positive numeric value, in which case the labels are trimmed to a fixed-length substring of that length: see substr for details. Defaults to FALSE (does nothing). legend.size the size of the legend symbols and text, in points. Defaults to 9. legend.position the location of the plot legend(s). Accepts all legend.position values supported by theme. Defaults to "right". deprecated: see group.legend and size.legend names quantize.weights deprecated: see weight.cut subset.threshold deprecated: see weight.min top8.nodes deprecated: this functionality was experimental and has been removed entirely from ggnet trim.labels deprecated: see label.trim other arguments passed to the geom\_text object that sets the node labels: see geom\_text for details.

## **Details**

The degree centrality measures that can be produced through the weight argument will take the directedness of the network into account, but will be unweighted. To compute weighted network measures, see the tnet package by Tore Opsahl (help("tnet", package = "tnet")).

#### Author(s)

Moritz Marbach and François Briatte, with help from Heike Hoffmann, Pedro Jordano and Ming-Yu Liu

#### See Also

ggnet2 in this package, gplot in the sna package, and plot.network in the network package

## **Examples**

```
library(network)
# random adjacency matrix
            <- 10
ndyads
            <-x * (x - 1)
density
            <- x / ndyads
            \leftarrow matrix(0, nrow = x, ncol = x)
dimnames(m) <- list(letters[ 1:x ], letters[ 1:x ])</pre>
m[ row(m) != col(m) ] <- runif(ndyads) < density</pre>
# random undirected network
n <- network::network(m, directed = FALSE)</pre>
ggnet(n, label = TRUE, alpha = 1, color = "white", segment.color = "black")
# random groups
g <- sample(letters[ 1:3 ], 10, replace = TRUE)</pre>
# color palette
p <- c("a" = "steelblue", "b" = "forestgreen", "c" = "tomato")</pre>
ggnet(n, node.group = g, node.color = p, label = TRUE, color = "white")
# edge arrows on a directed network
ggnet(network(m, directed = TRUE), arrow.gap = 0.05, arrow.size = 10)
```

ggnet2

ggnet2 - Plot a network with ggplot2

## **Description**

Function for plotting network objects using ggplot2, with additional control over graphical parameters that are not supported by the ggnet function. Please visit <a href="http://github.com/briatte/ggnet">http://github.com/briatte/ggnet</a> for the latest version of ggnet2, and <a href="https://briatte.github.io/ggnet">https://briatte.github.io/ggnet</a> for a vignette that contains many examples and explanations.

## Usage

```
ggnet2(
  net,
  mode = "fruchtermanreingold",
```

```
layout.par = NULL,
  layout.exp = 0,
  alpha = 1,
  color = "grey75",
  shape = 19,
  size = 9,
 max_size = 9,
  na.rm = NA,
  palette = NULL,
  alpha.palette = NULL,
  alpha.legend = NA,
  color.palette = palette,
  color.legend = NA,
  shape.palette = NULL,
  shape.legend = NA,
  size.palette = NULL,
  size.legend = NA,
  size.zero = FALSE,
  size.cut = FALSE,
  size.min = NA,
  size.max = NA,
  label = FALSE,
  label.alpha = 1,
  label.color = "black",
  label.size = max_size/2,
  label.trim = FALSE,
  node.alpha = alpha,
  node.color = color,
  node.label = label,
  node.shape = shape,
  node.size = size,
  edge.alpha = 1,
  edge.color = "grey50",
  edge.lty = "solid",
  edge.size = 0.25,
  edge.label = NULL,
  edge.label.alpha = 1,
  edge.label.color = label.color,
  edge.label.fill = "white",
  edge.label.size = max_size/2,
  arrow.size = 0,
  arrow.gap = 0,
  arrow.type = "closed",
  legend.size = 9,
  legend.position = "right",
)
```

#### **Arguments**

an object of class network, or any object that can be coerced to this class, such as an adjacency or incidence matrix, or an edge list: see edgeset.constructors and network for details. If the object is of class igraph and the intergraph package is installed, it will be used to convert the object: see asNetwork for

details.

mode a placement method from those provided in the sna package: see gplot.layout

for details. Also accepts the names of two numeric vertex attributes of net, or a matrix of numeric coordinates, in which case the first two columns of the matrix are used. Defaults to the Fruchterman-Reingold force-directed algorithm.

layout.par options to be passed to the placement method, as listed in gplot.layout. Defaults

to NULL.

layout.exp a multiplier to expand the horizontal axis if node labels get clipped: see ex-

pand\_range for details. Defaults to 0 (no expansion).

alpha the level of transparency of the edges and nodes, which might be a single value, a

vertex attribute, or a vector of values. Also accepts "mode" on bipartite networks

(see 'Details'). Defaults to 1 (no transparency).

color the color of the nodes, which might be a single value, a vertex attribute, or a

vector of values. Also accepts "mode" on bipartite networks (see 'Details').

Defaults to grey75.

shape the shape of the nodes, which might be a single value, a vertex attribute, or

a vector of values. Also accepts "mode" on bipartite networks (see 'Details').

Defaults to 19 (solid circle).

size the size of the nodes, in points, which might be a single value, a vertex attribute,

or a vector of values. Also accepts "indegree", "outdegree", "degree" or "freeman" to size the nodes by their unweighted degree centrality ("degree" and "freeman" are equivalent): see degree for details. All node sizes must be strictly positive. Also accepts "mode" on bipartite networks (see 'Details').

Defaults to 9.

max\_size the *maximum* size of the node when size produces nodes of different sizes, in

points. Defaults to 9.

na.rm whether to subset the network to nodes that are *not* missing a given vertex at-

tribute. If set to any vertex attribute of net, the nodes for which this attribute is

NA will be removed. Defaults to NA (does nothing).

palette the palette to color the nodes, when color is not a color value or a vector of color

values. Accepts named vectors of color values, or if RColorBrewer is installed, any ColorBrewer palette name: see brewer.pal and http://colorbrewer2.org/ for details. Defaults to NULL, which will create an array of grayscale color

values if color is not a color value or a vector of color values.

alpha.palette the palette to control the transparency levels of the nodes set by alpha when the

levels are not numeric values. Defaults to NULL, which will create an array of alpha transparency values if alpha is not a numeric value or a vector of numeric

values.

alpha.legend the name to assign to the legend created by alpha when its levels are not numeric

values. Defaults to NA (no name).

color.palette	see palette
color.legend	the name to assign to the legend created by palette. Defaults to NA (no name).
shape.palette	the palette to control the shapes of the nodes set by shape when the shapes are not numeric values. Defaults to NULL, which will create an array of shape values if shape is not a numeric value or a vector of numeric values.
shape.legend	the name to assign to the legend created by shape when its levels are not numeric values. Defaults to NA (no name).
size.palette	the palette to control the sizes of the nodes set by size when the sizes are not numeric values.
size.legend	the name to assign to the legend created by size. Defaults to NA (no name).
size.zero	whether to accept zero-sized nodes based on the value(s) of size. Defaults to FALSE, which ensures that zero-sized nodes are still shown in the plot and its size legend.
size.cut	whether to cut the size of the nodes into a certain number of quantiles. Accepts TRUE, which tries to cut the sizes into quartiles, or any positive numeric value, which tries to cut the sizes into that many quantiles. If the size of the nodes do not contain the specified number of distinct quantiles, the largest possible number is used. See quantile and cut for details. Defaults to FALSE (does nothing).
size.min	whether to subset the network to nodes with a minimum size, based on the values of size. Defaults to NA (preserves all nodes).
size.max	whether to subset the network to nodes with a maximum size, based on the values of size. Defaults to NA (preserves all nodes).
label	whether to label the nodes. If set to TRUE, nodes are labeled with their vertex names. If set to a vector that contains as many elements as there are nodes in net, nodes are labeled with these. If set to any other vector of values, the nodes are labeled only when their vertex name matches one of these values. Defaults to FALSE (no labels).
label.alpha	the level of transparency of the node labels, as a numeric value, a vector of numeric values, or as a vertex attribute containing numeric values. Defaults to 1 (no transparency).
label.color	the color of the node labels, as a color value, a vector of color values, or as a vertex attribute containing color values. Defaults to "black".
label.size	the size of the node labels, in points, as a numeric value, a vector of numeric values, or as a vertex attribute containing numeric values. Defaults to max_size / 2 (half the maximum node size), which defaults to 4.5.
label.trim	whether to apply some trimming to the node labels. Accepts any function that can process a character vector, or a strictly positive numeric value, in which case the labels are trimmed to a fixed-length substring of that length: see substr for details. Defaults to FALSE (does nothing).
node.alpha	see alpha
node.color	see color
node.label	see label

	node.shape	see shape
	node.size	see size
	edge.alpha	the level of transparency of the edges. Defaults to the value of alpha, which defaults to 1. $$
	edge.color	the color of the edges, as a color value, a vector of color values, or as an edge attribute containing color values. Defaults to "grey50".
	edge.lty	the linetype of the edges, as a linetype value, a vector of linetype values, or as an edge attribute containing linetype values. Defaults to "solid".
	edge.size	the size of the edges, in points, as a numeric value, a vector of numeric values, or as an edge attribute containing numeric values. All edge sizes must be strictly positive. Defaults to $\emptyset$ . 25.
	edge.label	the labels to plot at the middle of the edges, as a single value, a vector of values, or as an edge attribute. Defaults to NULL (no edge labels).
	edge.label.alph	na
		the level of transparency of the edge labels, as a numeric value, a vector of numeric values, or as an edge attribute containing numeric values. Defaults to 1 (no transparency).
	edge.label.cold	
		the color of the edge labels, as a color value, a vector of color values, or as an edge attribute containing color values. Defaults to label.color, which defaults to "black".
	edge.label.fill	
	odeo lobol oisa	the background color of the edge labels. Defaults to "white".
	edge.label.size	the size of the edge labels, in points, as a numeric value, a vector of numeric
		values, or as an edge attribute containing numeric values. All edge label sizes must be strictly positive. Defaults to max_size / 2 (half the maximum node size), which defaults to 4.5.
	arrow.size	the size of the arrows for directed network edges, in points. See arrow for details. Defaults to 0 (no arrows).
	arrow.gap	a setting aimed at improving the display of edge arrows by plotting slightly shorter edges. Accepts any value between 0 and 1, where a value of 0.05 will generally achieve good results when the size of the nodes is reasonably small. Defaults to 0 (no shortening).
	arrow.type	the type of the arrows for directed network edges. See ${\tt arrow}$ for details. Defaults to "closed".
	legend.size	the size of the legend symbols and text, in points. Defaults to 9.
legend.position		
		the location of the plot legend(s). Accepts all legend.position values supported by theme. Defaults to "right".

other arguments passed to the geom\_text object that sets the node labels: see

geom\_text for details.

#### **Details**

The degree centrality measures that can be produced through the size argument will take the directedness of the network into account, but will be unweighted. To compute weighted network measures, see the tnet package by Tore Opsahl (help("tnet",package = "tnet")).

The nodes of bipartite networks can be mapped to their mode by passing the "mode" argument to any of alpha, color, shape and size, in which case the nodes of the primary mode will be mapped as "actor", and the nodes of the secondary mode will be mapped as "event".

#### Author(s)

Moritz Marbach and François Briatte, with help from Heike Hoffmann, Pedro Jordano and Ming-Yu Liu

#### See Also

ggnet in this package, gplot in the sna package, and plot.network in the network package

## **Examples**

```
library(network)
# random adjacency matrix
           <- 10
           <- x * (x - 1)
ndyads
density
           <- x / ndyads
            \leftarrow matrix(0, nrow = x, ncol = x)
dimnames(m) <- list(letters[ 1:x ], letters[ 1:x ])</pre>
m[ row(m) != col(m) ] <- runif(ndyads) < density</pre>
# random undirected network
n <- network::network(m, directed = FALSE)</pre>
ggnet2(n, label = TRUE)
ggnet2(n, label = TRUE, shape = 15)
ggnet2(n, label = TRUE, shape = 15, color = "black", label.color = "white")
# add vertex attribute
x = network.vertex.names(n)
x = ifelse(x %in% c("a", "e", "i"), "vowel", "consonant")
n \%v\% "phono" = x
ggnet2(n, color = "phono")
ggnet2(n, color = "phono", palette = c("vowel" = "gold", "consonant" = "grey"))
ggnet2(n, shape = "phono", color = "phono")
if (require(RColorBrewer)) {
  # random groups
  n %v% "group" <- sample(LETTERS[1:3], 10, replace = TRUE)
```

```
ggnet2(n, color = "group", palette = "Set2")

# random weights
n %e% "weight" <- sample(1:3, network.edgecount(n), replace = TRUE)
ggnet2(n, edge.size = "weight", edge.label = "weight")

# edge arrows on a directed network
ggnet2(network(m, directed = TRUE), arrow.gap = 0.05, arrow.size = 10)

# Padgett's Florentine wedding data
data(flo, package = "network")
flo

ggnet2(flo, label = TRUE)
ggnet2(flo, label = TRUE, label.trim = 4, vjust = -1, size = 3, color = 1)
ggnet2(flo, label = TRUE, size = 12, color = "white")</pre>
```

ggnetworkmap

ggnetworkmap - Plot a network with ggplot2 suitable for overlay on a ggmap:: map ggplot, or other ggplot

# Description

This is a descendant of the original ggnet function. ggnet added the innovation of plotting the network geographically. However, ggnet needed to be the first object in the ggplot chain. ggnetworkmap does not. If passed a ggplot object as its first argument, such as output from ggmap, ggnetworkmap will plot on top of that chart, looking for vertex attributes lon and lat as coordinates. Otherwise, ggnetworkmap will generate coordinates using the Fruchterman-Reingold algorithm.

## Usage

```
ggnetworkmap(
   gg,
   net,
   size = 3,
   alpha = 0.75,
   weight,
   node.group,
   node.color = NULL,
   node.alpha = NULL,
   ring.group,
   segment.alpha = NULL,
   segment.color = "grey",
   great.circles = FALSE,
   segment.size = 0.25,
```

```
arrow.size = 0,
label.nodes = FALSE,
label.size = size/2,
...
)
```

# Arguments

gg	an object of class ggplot.
net	an object of class network, or any object that can be coerced to this class, such as an adjacency or incidence matrix, or an edge list: see edgeset.constructors and network for details. If the object is of class igraph and the intergraph package is installed, it will be used to convert the object: see asNetwork for details.
size	size of the network nodes. Defaults to 3. If the nodes are weighted, their area is proportionally scaled up to the size set by size.
alpha	a level of transparency for nodes, vertices and arrows. Defaults to 0.75.
weight	if present, the unquoted name of a vertex attribute in data. Otherwise nodes are unweighted.
node.group	NULL, the default, or the unquoted name of a vertex attribute that will be used to determine the color of each node.
node.color	If node group is null, a character string specifying a color.
node.alpha	transparency of the nodes. Inherits from alpha.
ring.group	if not NULL, the default, the unquoted name of a vertex attribute that will be used to determine the color of each node border.
segment.alpha	transparency of the vertex links. Inherits from alpha
segment.color	color of the vertex links. Defaults to "grey".
great.circles	whether to draw edges as great circles using the geosphere package. Defaults to FALSE
segment.size	size of the vertex links, as a vector of values or as a single value. Defaults to 0.25.
arrow.size	size of the vertex arrows for directed network plotting, in centimeters. Defaults to $0$ .
label.nodes	label nodes with their vertex names attribute. If set to TRUE, all nodes are labelled. Also accepts a vector of character strings to match with vertex names.
label.size	size of the labels. Defaults to size / 2.
	other arguments supplied to geom_text for the node labels. Arguments pertaining to the title or other items can be achieved through ggplot2 methods.

# **Details**

This is a function for plotting graphs generated by network or igraph in a more flexible and elegant manner than permitted by ggnet. The function does not need to be the first plot in the ggplot chain, so the graph can be plotted on top of a map or other chart. Segments can be straight lines, or plotted

as great circles. Note that the great circles feature can produce odd results with arrows and with vertices beyond the plot edges; this is a ggplot2 limitation and cannot yet be fixed. Nodes can have two color schemes, which are then plotted as the center and ring around the node. The color schemes are selected by adding scale\_fill\_ or scale\_color\_ just like any other ggplot2 plot. If there are no rings, scale\_color sets the color of the nodes. If there are rings, scale\_color sets the color of the rings, and scale\_fill sets the color of the centers. Note that additional arguments in the ... are passed to geom\_text for plotting labels.

#### Author(s)

## **Examples**

```
# small function to display plots only if it's interactive
p_ <- GGally::print_if_interactive</pre>
invisible(lapply(c("ggplot2", "maps", "network", "sna"), base::library, character.only = TRUE))
## Example showing great circles on a simple map of the USA
## http://flowingdata.com/2011/05/11/how-to-map-connections-with-great-circles/
airports <- read.csv("http://datasets.flowingdata.com/tuts/maparcs/airports.csv", header = TRUE)
rownames(airports) <- airports$iata</pre>
# select some random flights
set.seed(1234)
flights <- data.frame(</pre>
 origin = sample(airports[200:400, ]$iata, 200, replace = TRUE),
 destination = sample(airports[200:400, ]$iata, 200, replace = TRUE)
# convert to network
flights <- network(flights, directed = TRUE)
# add geographic coordinates
flights %v% "lat" <- airports[ network.vertex.names(flights), "lat" ]</pre>
flights %v% "lon" <- airports[ network.vertex.names(flights), "long" ]</pre>
# drop isolated airports
delete.vertices(flights, which(degree(flights) < 2))</pre>
# compute degree centrality
flights %v% "degree" <- degree(flights, gmode = "digraph")</pre>
# add random groups
flights %v% "mygroup" <- sample(letters[1:4], network.size(flights), replace = TRUE)
# create a map of the USA
usa <- ggplot(map_data("usa"), aes(x = long, y = lat)) +</pre>
 geom_polygon(aes(group = group), color = "grey65",
```

```
fill = "#f9f9f9", size = 0.2)
# overlay network data to map
p <- ggnetworkmap(</pre>
  usa, flights, size = 4, great.circles = TRUE,
  node.group = mygroup, segment.color = "steelblue",
  ring.group = degree, weight = degree
)
p_(p)
## Exploring a community of spambots found on Twitter
## Data by Amos Elberg: see ?twitter_spambots for details
data(twitter_spambots)
# create a world map
world <- fortify(map("world", plot = FALSE, fill = TRUE))</pre>
world <- ggplot(world, aes(x = long, y = lat)) +
  geom_polygon(aes(group = group), color = "grey65",
               fill = "#f9f9f9", size = 0.2)
# view global structure
p <- ggnetworkmap(world, twitter_spambots)</pre>
p_{-}(p)
# domestic distribution
p <- ggnetworkmap(net = twitter_spambots)</pre>
p_(p)
# topology
p <- ggnetworkmap(net = twitter_spambots, arrow.size = 0.5)</pre>
p_(p)
# compute indegree and outdegree centrality
twitter_spambots %v% "indegree" <- degree(twitter_spambots, cmode = "indegree")</pre>
twitter_spambots %v% "outdegree" <- degree(twitter_spambots, cmode = "outdegree")</pre>
p <- ggnetworkmap(</pre>
  net = twitter_spambots,
  arrow.size = 0.5,
  node.group = indegree,
  ring.group = outdegree, size = 4
  scale_fill_continuous("Indegree", high = "red", low = "yellow") +
  labs(color = "Outdegree")
p_{-}(p)
# show some vertex attributes associated with each account
p <- ggnetworkmap(</pre>
  net = twitter_spambots,
  arrow.size = 0.5,
  node.group = followers,
  ring.group = friends,
```

ggnostic 63

```
size = 4,
weight = indegree,
label.nodes = TRUE, vjust = -1.5
) +
scale_fill_continuous("Followers", high = "red", low = "yellow") +
labs(color = "Friends") +
scale_color_continuous(low = "lightgreen", high = "darkgreen")
p_(p)
```

ggnostic

ggnostic - Plot matrix of statistical model diagnostics

## **Description**

ggnostic - Plot matrix of statistical model diagnostics

#### Usage

```
ggnostic(
 model,
  columnsX = attr(data, "var_x"),
  columnsY = c(".resid", ".sigma", ".hat", ".cooksd"),
  columnLabelsX = attr(data, "var_x_label"),
  columnLabelsY = gsub("\\.", " ", gsub("^\\.", "", columnsY)),
  xlab = "explanatory variables",
  ylab = "diagnostics",
  title = paste(deparse(model$call, width.cutoff = 500L), collapse = "\n"),
  continuous = list(default = ggally_points, .fitted = ggally_points, .se.fit =
  ggally_nostic_se_fit, .resid = ggally_nostic_resid, .hat = ggally_nostic_hat, .sigma
    = ggally_nostic_sigma, .cooksd = ggally_nostic_cooksd, .std.resid =
    ggally_nostic_std_resid),
 combo = list(default = ggally_box_no_facet, fitted = ggally_box_no_facet, .se.fit =
  ggally_nostic_se_fit, .resid = ggally_nostic_resid, .hat = ggally_nostic_hat, .sigma
    = ggally_nostic_sigma, .cooksd = ggally_nostic_cooksd, .std.resid =
    ggally_nostic_std_resid),
  discrete = list(default = ggally_ratio, .fitted = ggally_ratio, .se.fit =
  ggally_ratio, .resid = ggally_ratio, .hat = ggally_ratio, .sigma = ggally_ratio,
    .cooksd = ggally_ratio, .std.resid = ggally_ratio),
  progress = NULL,
  data = broomify(model)
)
```

64 ggnostic

#### **Arguments**

model statistical model object such as output from stats::lm or stats::glm

... arguments passed directly to ggduo

columns to be displayed in the plot matrix. Defaults to the predictor columns of

the model

columnsY rows to be displayed in the plot matrix. Defaults to residuals, leave one out

sigma value, diagonal of the hat matrix, and Cook's Distance. The possible values are the response variables in the model and the added columns provided

by broom::augment(model). See details for more information.

columnLabelsX, columnLabelsY

column and row labels to display in the plot matrix

xlab, ylab, title

plot matrix labels passed directly to ggmatrix

continuous, combo, discrete

list of functions for each y variable. See details for more information.

progress NULL (default) for a progress bar in interactive sessions with more than 15 plots,

TRUE for a progress bar, FALSE for no progress bar, or a function that accepts at

least a plot matrix and returns a new progress::progress\_bar. See ggmatrix\_progress.

data defaults to a 'broomify'ed model object. This object will contain infor-

mation about the X variables, Y variables, and multiple broom outputs. See

broomify(model) for more information

#### 'columnsY'

broom::augment() collects data from the supplied model and returns a data.frame with the following columns (taken directly from broom documentation). These columns are the only allowed values in the columnsY parameter to ggnostic.

.resid Residuals

.hat Diagonal of the hat matrix

.sigma Estimate of residual standard deviation when corresponding observation is dropped from model

.cooksd Cooks distance, cooks.distance

.fitted Fitted values of model

.se.fit Standard errors of fitted values

.std.resid Standardized residuals

**response variable name** The response variable in the model may be added. Such as "mpg" in the model lm(mpg ~ .,data = mtcars)

## 'continuous', 'combo', 'discrete' types

Similar to ggduo and ggpairs, functions may be supplied to display the different column types. However, since the Y rows are fixed, each row has it's own corresponding function in each of the plot types: continuous, combo, and discrete. Each plot type list can have keys that correspond

to the broom::augment() output: ".fitted", ".resid", ".std.resid", ".sigma", ".se.fit", ".hat", ".cooksd". An extra key, "default", is used to plot the response variables of the model if they are included. Having a function for each diagnostic allows for very fine control over the diagnostics plot matrix. The functions for each type list are wrapped into a switch function that calls the function corresponding to the y variable being plotted. These switch functions are then passed directly to the types parameter in ggduo.

## **Examples**

```
# small function to display plots only if it's interactive
p_ <- GGally::print_if_interactive</pre>
data(mtcars)
# use mtcars dataset and alter the 'am' column to display actual name values
mtc$am <- c("0" = "automatic", "1" = "manual")[as.character(mtc$am)]</pre>
# step the complete model down to a smaller model
mod <- stats::step(stats::lm(mpg ~ ., data = mtc), trace = FALSE)</pre>
# display using defaults
pm <- ggnostic(mod)</pre>
p_{p}(pm)
# color by am value
pm <- ggnostic(mod, mapping = ggplot2::aes(color = am))</pre>
p_(pm)
# turn resid smooth error ribbon off
pm <- ggnostic(mod, continuous = list(.resid = wrap("nostic_resid", se = FALSE)))</pre>
p_{p}(pm)
## plot residuals vs fitted in a ggpairs plot matrix
dt <- broomify(mod)</pre>
pm <- ggpairs(</pre>
  dt, c(".fitted", ".resid"),
  columnLabels = c("fitted", "residuals"),
  lower = list(continuous = ggally_nostic_resid)
p_{p}(pm)
```

ggpairs

ggpairs - A ggplot2 generalized pairs plot

## Description

Make a matrix of plots with a given data set

## Usage

```
ggpairs(
  data,
 mapping = NULL,
 columns = 1:ncol(data),
  title = NULL,
 upper = list(continuous = "cor", combo = "box_no_facet", discrete = "facetbar", na =
    "na"),
 lower = list(continuous = "points", combo = "facethist", discrete = "facetbar", na =
    "na"),
  diag = list(continuous = "densityDiag", discrete = "barDiag", na = "naDiag"),
  params = NULL,
  . . . ,
  xlab = NULL,
  ylab = NULL,
  axisLabels = c("show", "internal", "none"),
  columnLabels = colnames(data[columns]),
  labeller = "label_value",
  switch = NULL,
  showStrips = NULL,
  legend = NULL,
  cardinality_threshold = 15,
  progress = NULL,
 legends = stop("deprecated")
)
```

# Arguments

data data set using. Can have both numerical and categorical data. mapping aesthetic mapping (besides x and y). See aes(). If mapping is numeric, columns will be set to the mapping value and mapping will be set to NULL. columns which columns are used to make plots. Defaults to all columns. title, xlab, ylab title, x label, and y label for the graph see Details upper see Details lower diag see Details params deprecated. Please see wrap\_fn\_with\_param\_arg deprecated. Please use mapping axisLabels either "show" to display axisLabels, "internal" for labels in the diagonal plots, or "none" for no axis labels columnLabels label names to be displayed. Defaults to names of columns being used. labeller for facets. See labellers. Common values are "label\_value" (delabeller fault) and "label\_parsed".

switch switch parameter for facet\_grid. See ggplot2::facet\_grid. By default, the labels are displayed on the top and right of the plot. If "x", the top labels will be displayed to the bottom. If "y", the right-hand side labels will be displayed to the left. Can also be set to "both"

boolean to determine if each plot's strips should be displayed. NULL will default

to the top and right side plots only. TRUE or FALSE will turn all strips on or off

respectively.

legend May be the two objects described below or the default NULL value. The legend

position can be moved by using ggplot2's theme element pm + theme(legend.position

= "bottom")

a numeric vector of length 2 provides the location of the plot to use the legend for the plot matrix's legend. Such as legend = c(3,5) which will use the legend from the plot in the third row and fifth column

a single numeric value provides the location of a plot according to the display order. Such as legend = 3 in a plot matrix with 2 rows and 5 columns displayed by column will return the plot in position c(1,2)

a object from grab\_legend() a predetermined plot legend that will be displayed directly

cardinality\_threshold

showStrips

maximum number of levels allowed in a character / factor column. Set this value

to NULL to not check factor columns. Defaults to 15

progress NULL (default) for a progress bar in interactive sessions with more than 15 plots,

TRUE for a progress bar, FALSE for no progress bar, or a function that accepts at

least a plot matrix and returns a new progress::progress\_bar. See ggmatrix\_progress.

legends deprecated

#### Details

upper and lower are lists that may contain the variables 'continuous', 'combo', 'discrete', and 'na'. Each element of the list may be a function or a string. If a string is supplied, it must implement one of the following options:

**continuous** exactly one of ('points', 'smooth', 'smooth\_loess', 'density', 'cor', 'blank'). This option is used for continuous X and Y data.

**combo** exactly one of ('box', 'box\_no\_facet', 'dot', 'dot\_no\_facet', 'facethist', 'facetdensity', 'denstrip', 'blank'). This option is used for either continuous X and categorical Y data or categorical X and continuous Y data.

**discrete** exactly one of ('facetbar', 'ratio', 'blank'). This option is used for categorical X and Y data

**na** exactly one of ('na', 'blank'). This option is used when all X data is NA, all Y data is NA, or either all X or Y data is NA.

diag is a list that may only contain the variables 'continuous', 'discrete', and 'na'. Each element of the diag list is a string implementing the following options:

continuous exactly one of ('densityDiag', 'barDiag', 'blankDiag'). This option is used for continuous X data.

**discrete** exactly one of ('barDiag', 'blankDiag'). This option is used for categorical X and Y data. **na** exactly one of ('naDiag', 'blankDiag'). This option is used when all X data is NA.

If 'blank' is ever chosen as an option, then ggpairs will produce an empty plot.

If a function is supplied as an option to upper, lower, or diag, it should implement the function api of function(data, mapping,...){#make ggplot2 plot}. If a specific function needs its parameters set, wrap(fn,param1 = val1,param2 = val2) the function with its parameters.

#### Value

ggmatrix object that if called, will print

#### Author(s)

Barret Schloerke <schloerke@gmail.com>, Jason Crowley <crowley.jason.s@gmail.com>, Di Cook <dicook@iastate.edu>, Heike Hofmann <hofmann@iastate.edu>, Hadley Wickham <h.wickham@gmail.com>

#### References

John W Emerson, Walton A Green, Barret Schloerke, Jason Crowley, Dianne Cook, Heike Hofmann, Hadley Wickham. The Generalized Pairs Plot. Journal of Computational and Graphical Statistics, vol. 22, no. 1, pp. 79-91, 2012.

#### See Also

wrap v1\_ggmatrix\_theme

# **Examples**

```
# small function to display plots only if it's interactive
 p_ <- GGally::print_if_interactive</pre>
## Quick example, with and without colour
data(flea)
ggpairs(flea, columns = 2:4)
pm <- ggpairs(flea, columns = 2:4, ggplot2::aes(colour=species))</pre>
p_{p}(pm)
# Note: colour should be categorical, else you will need to reset
# the upper triangle to use points instead of trying to compute corr
data(tips, package = "reshape")
pm <- ggpairs(tips[, 1:3])</pre>
p_{p}(pm)
pm <- ggpairs(tips, 1:3, columnLabels = c("Total Bill", "Tip", "Sex"))</pre>
p_{p}(pm)
pm <- ggpairs(tips, upper = "blank")</pre>
p_{p}(pm)
## Plot Types
# Change default plot behavior
```

```
pm <- ggpairs(</pre>
  tips[, c(1, 3, 4, 2)],
  upper = list(continuous = "density", combo = "box_no_facet"),
  lower = list(continuous = "points", combo = "dot_no_facet")
)
p_{p}(pm)
# Supply Raw Functions (may be user defined functions!)
pm <- ggpairs(</pre>
  tips[, c(1, 3, 4, 2)],
  upper = list(continuous = ggally_density, combo = ggally_box_no_facet),
  lower = list(continuous = ggally_points, combo = ggally_dot_no_facet)
p_{p}(pm)
# Use sample of the diamonds data
data(diamonds, package="ggplot2")
diamonds.samp <- diamonds[sample(1:dim(diamonds)[1], 1000), ]</pre>
# Different aesthetics for different plot sections and plot types
pm <- ggpairs(
diamonds.samp[, 1:5],
 mapping = ggplot2::aes(color = cut),
 upper = list(continuous = wrap("density", alpha = 0.5), combo = "box_no_facet"),
lower = list(continuous = wrap("points", alpha = 0.3), combo = wrap("dot_no_facet", alpha = 0.4)),
 title = "Diamonds"
)
p_{p}(pm)
## Axis Label Variations
# Only Variable Labels on the diagonal (no axis labels)
pm <- ggpairs(tips[, 1:3], axisLabels="internal")</pre>
p_{p}(pm)
# Only Variable Labels on the outside (no axis labels)
pm <- ggpairs(tips[, 1:3], axisLabels="none")</pre>
p_{p}(pm)
## Facet Label Variations
# Default:
df_x \leftarrow rnorm(100)
df_y \leftarrow df_x + rnorm(100, 0, 0.1)
df \leftarrow data.frame(x = df_x, y = df_y, c = sqrt(df_x^2 + df_y^2))
pm <- ggpairs(</pre>
  df,
  columnLabels = c("alpha[foo]", "alpha[bar]", "sqrt(alpha[foo]^2 + alpha[bar]^2)")
)
p_{p}(pm)
# Parsed labels:
pm <- ggpairs(</pre>
  columnLabels = c("alpha[foo]", "alpha[bar]", "sqrt(alpha[foo]^2 + alpha[bar]^2)"),
  labeller = "label_parsed"
p_{p}(pm)
```

70 ggparcoord

```
## Plot Insertion Example
custom_car <- ggpairs(mtcars[, c("mpg", "wt", "cyl")], upper = "blank", title = "Custom Example")</pre>
# ggplot example taken from example(geom_text)
  plot <- ggplot2::ggplot(mtcars, ggplot2::aes(x=wt, y=mpg, label=rownames(mtcars)))</pre>
  plot <- plot +
    ggplot2::geom_text(ggplot2::aes(colour=factor(cyl)), size = 3) +
    ggplot2::scale_colour_discrete(l=40)
custom_car[1, 2] <- plot</pre>
personal_plot <- ggally_text(</pre>
  "ggpairs allows you\nto put in your\nown plot.\nLike that one.\n <---"
)
custom_car[1, 3] <- personal_plot</pre>
p_(custom_car)
## Remove binwidth warning from ggplot2
# displays warning about picking a better binwidth
pm <- ggpairs(tips, 2:3)</pre>
p_(pm)
# no warning displayed
pm <- ggpairs(tips, 2:3, lower = list(combo = wrap("facethist", binwidth = 0.5)))</pre>
# no warning displayed with user supplied function
pm <- ggpairs(tips, 2:3, lower = list(combo = wrap(ggally_facethist, binwidth = 0.5)))</pre>
## Remove panel grid lines from correlation plots
pm <- ggpairs(</pre>
  flea, columns = 2:4,
  upper = list(continuous = wrap(ggally_cor, displayGrid = FALSE))
)
p_{p}(pm)
```

ggparcoord

ggparcoord - A ggplot2 Parallel Coordinate Plot

## **Description**

A function for plotting static parallel coordinate plots, utilizing the ggplot2 graphics package.

# Usage

```
ggparcoord(
  data,
  columns = 1:ncol(data),
  groupColumn = NULL,
  scale = "std",
  scaleSummary = "mean",
  centerObsID = 1,
  missing = "exclude",
```

ggparcoord 71

```
order = columns,
showPoints = FALSE,
splineFactor = FALSE,
alphaLines = 1,
boxplot = FALSE,
shadeBox = NULL,
mapping = NULL,
title = ""
```

## **Arguments**

data the dataset to plot a vector of variables (either names or indices) to be axes in the plot columns groupColumn a single variable to group (color) by scale method used to scale the variables (see Details) if scale=="center", summary statistic to univariately center each variable by scaleSummary centerObsID if scale=="centerObs", row number of case plot should univariately be centered on missing method used to handle missing values (see Details) method used to order the axes (see Details) order showPoints logical operator indicating whether points should be plotted or not splineFactor logical or numeric operator indicating whether spline interpolation should be used. Numeric values will multiplied by the number of columns, TRUE will default to cubic interpolation, AsIs to set the knot count directly and 0, FALSE, or non-numeric values will not use spline interpolation. alphaLines value of alpha scaler for the lines of the parcoord plot or a column name of the data boxplot logical operator indicating whether or not boxplots should underlay the distribution of each variable shadeBox color of underlaying box which extends from the min to the max for each variable (no box is plotted if shadeBox == NULL) mapping aes string to pass to ggplot object title character string denoting the title of the plot

#### **Details**

scale is a character string that denotes how to scale the variables in the parallel coordinate plot. Options:

- std: univariately, subtract mean and divide by standard deviation
- robust: univariately, subtract median and divide by median absolute deviation
- uniminmax: univariately, scale so the minimum of the variable is zero, and the maximum is one

72 ggparcoord

• globalminmax: no scaling is done; the range of the graphs is defined by the global minimum and the global maximum

- center: use uniminmax to standardize vertical height, then center each variable at a value specified by the scaleSummary param
- centerObs: use uniminmax to standardize vertical height, then center each variable at the value of the observation specified by the centerObsID param

missing is a character string that denotes how to handle missing missing values. Options:

- · exclude: remove all cases with missing values
- mean: set missing values to the mean of the variable
- median: set missing values to the median of the variable
- min10: set missing values to 10% below the minimum of the variable
- random: set missing values to value of randomly chosen observation on that variable

order is either a vector of indices or a character string that denotes how to order the axes (variables) of the parallel coordinate plot. Options:

- (default): order by the vector denoted by columns
- (given vector): order by the vector specified
- anyClass: order variables by their separation between any one class and the rest (as opposed to their overall variation between classes). This is accomplished by calculating the F-statistic for each class vs. the rest, for each axis variable. The axis variables are then ordered (decreasing) by their maximum of k F-statistics, where k is the number of classes.
- allClass: order variables by their overall F statistic (decreasing) from an ANOVA with groupColumn as the explanatory variable (note: it is required to specify a groupColumn with this ordering method). Basically, this method orders the variables by their variation between classes (most to least).
- skewness: order variables by their sample skewness (most skewed to least skewed)
- Outlying: order by the scagnostic measure, Outlying, as calculated by the package scagnostics.
   Other scagnostic measures available to order by are Skewed, Clumpy, Sparse, Striated,
   Convex, Skinny, Stringy, and Monotonic. Note: To use these methods of ordering, you
   must have the scagnostics package loaded.

#### Value

ggplot object that if called, will print

## Author(s)

Jason Crowley <crowley.jason.s@gmail.com>, Barret Schloerke <schloerke@gmail.com>, Di Cook <dicook@iastate.edu>, Heike Hofmann <hofmann@iastate.edu>, Hadley Wickham <h.wickham@gmail.com>

ggparcoord 73

```
# small function to display plots only if it's interactive
p_ <- GGally::print_if_interactive</pre>
# use sample of the diamonds data for illustrative purposes
data(diamonds, package="ggplot2")
diamonds.samp <- diamonds[sample(1:dim(diamonds)[1], 100), ]</pre>
# basic parallel coordinate plot, using default settings
p <- ggparcoord(data = diamonds.samp, columns = c(1, 5:10))</pre>
p_{-}(p)
# this time, color by diamond cut
p <- ggparcoord(data = diamonds.samp, columns = c(1, 5:10), groupColumn = 2)
p_{-}(p)
# underlay univariate boxplots, add title, use uniminmax scaling
p <- ggparcoord(data = diamonds.samp, columns = c(1, 5:10), groupColumn = 2,
 scale = "uniminmax", boxplot = TRUE, title = "Parallel Coord. Plot of Diamonds Data")
p_{p}(p)
# utilize ggplot2 aes to switch to thicker lines
p <- ggparcoord(data = diamonds.samp, columns = c(1, 5:10), groupColumn = 2,
 title ="Parallel Coord. Plot of Diamonds Data", mapping = ggplot2::aes(size = 1)) +
 ggplot2::scale_size_identity()
p_(p)
# basic parallel coord plot of the msleep data, using 'random' imputation and
# coloring by diet (can also use variable names in the columns and groupColumn
# arguments)
data(msleep, package="ggplot2")
p <- ggparcoord(data = msleep, columns = 6:11, groupColumn = "vore", missing =
  "random", scale = "uniminmax")
p_{-}(p)
# center each variable by its median, using the default missing value handler,
# 'exclude'
p <- ggparcoord(data = msleep, columns = 6:11, groupColumn = "vore", scale =</pre>
  "center", scaleSummary = "median")
p_{-}(p)
# with the iris data, order the axes by overall class (Species) separation using
# the anyClass option
p <- ggparcoord(data = iris, columns = 1:4, groupColumn = 5, order = "anyClass")</pre>
# add points to the plot, add a title, and use an alpha scalar to make the lines
# transparent
p <- ggparcoord(data = iris, columns = 1:4, groupColumn = 5, order = "anyClass",</pre>
 showPoints = TRUE, title = "Parallel Coordinate Plot for the Iris Data",
 alphaLines = 0.3)
p_{-}(p)
```

74 ggscatmat

```
# color according to a column
iris2 <- iris</pre>
iris2$alphaLevel <- c("setosa" = 0.2, "versicolor" = 0.3, "virginica" = 0)[iris2$Species]</pre>
p <- ggparcoord(data = iris2, columns = 1:4, groupColumn = 5, order = "anyClass",
  showPoints = TRUE, title = "Parallel Coordinate Plot for the Iris Data",
  alphaLines = "alphaLevel")
p_(p)
## Use splines on values, rather than lines (all produce the same result)
columns <- c(1, 5:10)
p <- ggparcoord(diamonds.samp, columns, groupColumn = 2, splineFactor = TRUE)</pre>
p_{-}(p)
p <- ggparcoord(diamonds.samp, columns, groupColumn = 2, splineFactor = 3)</pre>
p_(p)
splineFactor <- length(columns) * 3</pre>
p <- ggparcoord(diamonds.samp, columns, groupColumn = 2, splineFactor = I(splineFactor))</pre>
p_(p)
```

ggscatmat

ggscatmat - a traditional scatterplot matrix for purely quantitative variables

# **Description**

This function makes a scatterplot matrix for quantitative variables with density plots on the diagonal and correlation printed in the upper triangle.

## Usage

```
ggscatmat(
  data,
  columns = 1:ncol(data),
  color = NULL,
  alpha = 1,
  corMethod = "pearson"
)
```

# Arguments

data	a data matrix. Should contain numerical (continuous) data.
columns	an option to choose the column to be used in the raw dataset. Defaults to 1:ncol(data).
color	an option to group the dataset by the factor variable and color them by different colors. Defaults to NULL, i.e. no coloring. If supplied, it will be converted to a factor.
alpha	an option to set the transparency in scatterplots for large data. Defaults to 1.
corMethod	method argument supplied to cor

ggsurv 75

## Author(s)

Mengjia Ni, Di Cook <dicook@monash.edu>

## **Examples**

```
data(flea)
ggscatmat(flea, columns = 2:4)
ggscatmat(flea, columns = 2:4, color = "species")
```

ggsurv

Survival curves with ggplot2

# Description

This function produces Kaplan-Meier plots using ggplot2. As a first argument it needs a survfit object, created by the survival package. Default settings differ for single stratum and multiple strata objects.

# Usage

```
ggsurv(
 CI = "def",
 plot.cens = TRUE,
  surv.col = "gg.def",
  cens.col = "gg.def",
  lty.est = 1,
  lty.ci = 2,
  size.est = 0.5,
  size.ci = size.est,
  cens.size = 2,
  cens.shape = 3,
  back.white = FALSE,
  xlab = "Time",
 ylab = "Survival",
 main = "",
  order.legend = TRUE
)
```

#### **Arguments**

```
s an object of class survfit

CI should a confidence interval be plotted? Defaults to TRUE for single stratum objects and FALSE for multiple strata objects.

plot.cens mark the censored observations?
```

76 ggsurv

surv.col	colour of the survival estimate. Defaults to black for one stratum, and to the default ggplot2 colours for multiple strata. Length of vector with colour names should be either 1 or equal to the number of strata.
cens.col	colour of the points that mark censored observations.
lty.est	linetype of the survival curve(s). Vector length should be either 1 or equal to the number of strata.
lty.ci	linetype of the bounds that mark the 95% CI.
size.est	line width of the survival curve
size.ci	line width of the 95% CI
cens.size	point size of the censoring points
cens.shape	shape of the points that mark censored observations.
back.white	if TRUE the background will not be the default grey of ggplot2 but will be white with borders around the plot.
xlab	the label of the x-axis.
ylab	the label of the y-axis.
main	the plot label.
order.legend	boolean to determine if the legend display should be ordered by final survival time

#### Value

An object of class ggplot

# Author(s)

Edwin Thoen <edwinthoen@gmail.com>

```
if (require(survival) && require(scales)) {
   data(lung, package = "survival")
   sf.lung <- survival::survfit(Surv(time, status) ~ 1, data = lung)
   ggsurv(sf.lung)

# Multiple strata examples
   sf.sex <- survival::survfit(Surv(time, status) ~ sex, data = lung)
   pl.sex <- ggsurv(sf.sex)
   pl.sex

# Adjusting the legend of the ggsurv fit
   pl.sex +
        ggplot2::guides(linetype = FALSE) +
        ggplot2::scale_colour_discrete(
        name = 'Sex',
        breaks = c(1,2),
        labels = c('Male', 'Female')</pre>
```

ggts 77

```
)
 # Multiple factors
 lung2 <- plyr::mutate(lung, older = as.factor(age > 60))
 sf.sex2 <- survival::survfit(Surv(time, status) ~ sex + older, data = lung2)</pre>
 pl.sex2 <- ggsurv(sf.sex2)</pre>
 pl.sex2
 # Change legend title
 pl.sex2 + labs(color = "New Title", linetype = "New Title")
 # We can still adjust the plot after fitting
 data(kidney, package = "survival")
 sf.kid <- survival::survfit(Surv(time, status) ~ disease, data = kidney)</pre>
 pl.kid <- ggsurv(sf.kid, plot.cens = FALSE)</pre>
 pl.kid
 # Zoom in to first 80 days
 pl.kid + ggplot2::coord_cartesian(xlim = c(0, 80), ylim = c(0.45, 1))
 # Add the diseases names to the plot and remove legend
 pl.kid +
   ggplot2::annotate(
      "text",
      label = c("PKD", "Other", "GN", "AN"),
             = c(90, 125, 5, 60),
             = c(0.8, 0.65, 0.55, 0.30),
      size
            = 5,
      colour = scales::hue_pal(
       h
                  = c(0, 360) + 15,
       С
                  = 100,
       1
                  = 65,
       h.start = 0,
        direction = 1
     )(4)
   ) +
   ggplot2::guides(color = FALSE, linetype = FALSE)
}
```

ggts

Multiple Time Series

# **Description**

GGally implementation of ts.plot. Wraps around the ggduo function and removes the column strips

# Usage

```
ggts(..., columnLabelsX = NULL, xlab = "time")
```

78 glyphplot

#### **Arguments**

```
... supplied directly to ggduo

columnLabelsX remove top strips for the X axis by default

xlab defaults to "time"
```

#### Value

ggmatrix object

#### **Examples**

```
ggts(pigs, "time", c("gilts", "profit", "s_per_herdsz", "production", "herdsz"))
```

glyphplot

Glyph plot class

## **Description**

Glyph plot class

# Usage

```
glyphplot(data, width, height, polar, x_major, y_major)
is.glyphplot(x)

## S3 method for class 'glyphplot'
x[...]

## S3 method for class 'glyphplot'
print(x, ...)
```

# Arguments

data A data frame containing variables named in x\_major, x\_minor, y\_major and

y\_minor.

height, width The height and width of each glyph. Defaults to 95% of the resolution of the

data. Specify the width absolutely by supplying a numeric vector of length 1, or

relative to the

polar A logical of length 1, specifying whether the glyphs should be drawn in polar

coordinates. Defaults to FALSE.

x\_major, y\_major

The name of the variable (as a string) for the major x and y axes. Together, the

x glyphplot to be printed

... ignored

glyphs 79

#### Author(s)

Di Cook <dicook@monash.edu>, Heike Hofmann, Hadley Wickham

glyphs

Create the data needed to generate a glyph plot.

## **Description**

Create the data needed to generate a glyph plot.

## Usage

```
glyphs(
  data,
  x_major,
  x_minor,
  y_major,
  y_minor,
  polar = FALSE,
  height = ggplot2::rel(0.95),
  width = ggplot2::rel(0.95),
  y_scale = identity,
  x_scale = identity
)
```

#### **Arguments**

data A data frame containing variables named in x\_major, x\_minor, y\_major and

y\_minor.

x\_major, x\_minor, y\_major, y\_minor

The name of the variable (as a string) for the major and minor x and y axes.

Together, each unique

polar A logical of length 1, specifying whether the glyphs should be drawn in polar

coordinates. Defaults to FALSE.

height, width The height and width of each glyph. Defaults to 95% of the resolution of the

data. Specify the width absolutely by supplying a numeric vector of length 1, or

relative to the

y\_scale, x\_scale

The scaling function to be applied to each set of minor values within a grid cell. Defaults to identity so that no scaling is performed.

#### Author(s)

Di Cook <dicook@monash.edu>, Heike Hofmann, Hadley Wickham

grab\_legend

#### **Examples**

```
data(nasa)
nasaLate <- nasa[
  nasa$date >= as.POSIXct("1998-01-01") &
  nasa$lat >= 20 &
  nasa$lat <= 40 &
  nasa$long >= -80 &
  nasa$long <= -60
, ]
temp.gly <- glyphs(nasaLate, "long", "day", "lat", "surftemp", height=2.5)
ggplot2::ggplot(temp.gly, ggplot2::aes(gx, gy, group = gid)) +
  add_ref_lines(temp.gly, color = "grey90") +
  add_ref_boxes(temp.gly, color = "grey90") +
  ggplot2::geom_path() +
  ggplot2::theme_bw() +
  ggplot2::labs(x = "", y = "")</pre>
```

grab\_legend

Grab the legend and print it as a plot

# Description

Grab the legend and print it as a plot

#### Usage

```
grab_legend(p)
## S3 method for class 'legend_guide_box'
print(x, ..., plotNew = FALSE)
```

# **Arguments**

p ggplot2 plot object

x legend object that has been grabbed from a ggplot2 object

... ignored

plotNew boolean to determine if the 'grid.newpage()' command and a new blank rectangle should be printed

```
library(ggplot2)
histPlot <- qplot(
  x = Sepal.Length,
  data = iris,
  fill = Species,
  geom = "histogram",
  binwidth = 1/4</pre>
```

happy 81

```
(right <- histPlot)
(bottom <- histPlot + theme(legend.position = "bottom"))
(top <- histPlot + theme(legend.position = "top"))
(left <- histPlot + theme(legend.position = "left"))
grab_legend(right)
grab_legend(bottom)
grab_legend(top)
grab_legend(left)</pre>
```

happy

Data related to happiness from the General Social Survey, 1972-2006.

## **Description**

This data extract is taken from Hadley Wickham's productplots package. The original description follows, with minor edits.

# Usage

data(happy)

## Format

A data frame with 51020 rows and 10 variables

#### **Details**

The data is a small sample of variables related to happiness from the General Social Survey (GSS). The GSS is a yearly cross-sectional survey of Americans, run from 1972. We combine data for 25 years to yield 51,020 observations, and of the over 5,000 variables, we select nine related to happiness:

- age. age in years: 18–89.
- degree. highest education: It high school, high school, junior college, bachelor, graduate.
- finrela. relative financial status: far above, above average, average, below average, far below.
- happy. happiness: very happy, pretty happy, not too happy.
- health. health: excellent, good, fair, poor.
- marital. marital status: married, never married, divorced, widowed, separated.
- sex. sex: female, male.
- wtsall. probability weight. 0.43-6.43.

82 is\_horizontal

#### References

Smith, Tom W., Peter V. Marsden, Michael Hout, Jibum Kim. *General Social Surveys*, 1972-2006. [machine-readable data file]. Principal Investigator, Tom W. Smith; Co-Principal Investigators, Peter V. Marsden and Michael Hout, NORC ed. Chicago: National Opinion Research Center, producer, 2005; Storrs, CT: The Roper Center for Public Opinion Research, University of Connecticut, distributor. 1 data file (57,061 logical records) and 1 codebook (3,422 pp).

is\_horizontal

Check if plot is horizontal

# Description

Check if plot is horizontal

# Usage

```
is_horizontal(data, mapping, val = "y")
is_character_column(data, mapping, val = "y")
```

## **Arguments**

data used in ggplot2 plot
mapping ggplot2 aes() mapping
val key to retrieve from mapping

#### Value

Boolean determining if the data is a character-like data

```
is_horizontal(iris, ggplot2::aes(Sepal.Length, Species)) # TRUE
is_horizontal(iris, ggplot2::aes(Sepal.Length, Species), "x") # FALSE
is_horizontal(iris, ggplot2::aes(Sepal.Length, Sepal.Width)) # FALSE
```

lowertriangle 83

lowertriangle	lowertriangle - rearrange dataset as the preparation of ggscatmat function
---------------	----------------------------------------------------------------------------

# **Description**

function for making the melted dataset used to plot the lowertriangle scatterplots.

# Usage

```
lowertriangle(data, columns = 1:ncol(data), color = NULL)
```

# Arguments

data a data matrix. Should contain numerical (continuous) data.

columns an option to choose the column to be used in the raw dataset. Defaults to

1:ncol(data)

color an option to choose a factor variable to be grouped with. Defaults to (NULL)

#### Author(s)

Mengjia Ni, Di Cook <dicook@monash.edu>

# **Examples**

```
data(flea)
head(lowertriangle(flea, columns= 2:4))
head(lowertriangle(flea))
head(lowertriangle(flea, color="species"))
```

mapping\_string

Aes name

# **Description**

Aes name

# Usage

```
mapping_string(aes_col)
```

# **Arguments**

```
aes_col Single value from ggplot2::aes(...)
```

# Value

character string

# **Examples**

```
mapping <- ggplot2::aes(Petal.Length)
mapping_string(mapping$x)</pre>
```

mapping\_swap\_x\_y

Swap x and y mapping

# **Description**

Swap x and y mapping

# Usage

```
mapping_swap_x_y(mapping)
```

# Arguments

```
mapping output of ggplot2::aes(...)
```

# Value

Aes mapping with the x and y values switched

# **Examples**

```
mapping <- ggplot2::aes(Petal.Length, Sepal.Width)
mapping
mapping_swap_x_y(mapping)</pre>
```

```
model_response_variables
```

Model term names

# **Description**

Retrieve either the response variable names, the beta variable names, or beta variable names. If the model is an object of class 'lm', by default, the beta variable names will include anova significance stars.

nasa 85

#### Usage

```
model_response_variables(model, data = broom::augment(model))
model_beta_variables(model, data = broom::augment(model))
model_beta_label(model, data = broom::augment(model), lmStars = TRUE)
```

# **Arguments**

model model in question

1mStars boolean that determines if stars are added to labels

#### Value

character vector of names

nasa

Data from the Data Expo JSM 2006.

# Description

This data was provided by NASA for the competition.

# Usage

data(nasa)

#### **Format**

A data frame with 41472 rows and 17 variables

#### **Details**

The data shows 6 years of monthly measurements of a 24x24 spatial grid from Central America:

- time integer specifying temporal order of measurements
- x, y, lat, long spatial location of measurements.
- cloudhigh, cloudlow, cloudmid, ozone, pressure, surftemp, temperature are the various satellite measurements.
- date, day, month, year specifying the time of measurements.
- id unique ide for each spatial position.

#### References

Murrell, P. (2010) The 2006 Data Expo of the American Statistical Association. Computational Statistics, 25:551-554.

pigs pigs

pigs

United Kingdom Pig Production

# Description

This data contains about the United Kingdom Pig Production from the book 'Data' by Andrews and Herzberg. The original data can be on Statlib: http://lib.stat.cmu.edu/datasets/Andrews/T62.1

# Usage

```
data(pigs)
```

#### **Format**

A data frame with 48 rows and 8 variables

#### **Details**

The time variable has been added from a combination of year and quarter

- time year + (quarter 1)/4
- · year year of production
- quarter quarter of the year of production
- gilts number of sows giving birth for the first time
- profit ratio of price to an index of feed price
- s\_per\_herdsz ratio of the number of breeding pigs slaughtered to the total breeding herd size
- production number of pigs slaughtered that were reared for meat
- herdsz breeding herd size

#### References

Andrews, David F., and Agnes M. Herzberg. Data: a collection of problems from many fields for the student and research worker. Springer Science & Business Media, 2012.

print.ggmatrix 87

print.ggmatrix

Print ggmatrix object

# **Description**

Print method taken from ggplot2:::print.ggplot and altered for a ggmatrix object

# Usage

```
## S3 method for class 'ggmatrix'
print(x, newpage = is.null(vp), vp = NULL, ...)
```

# **Arguments**

```
x plot to display
```

newpage draw new (empty) page first?
vp viewport to draw plot in

... arguments passed onto ggmatrix\_gtable

#### Author(s)

Barret Schloerke

# **Examples**

```
data(tips, package = "reshape")
pMat <- ggpairs(tips, c(1,3,2), mapping = ggplot2::aes_string(color = "sex"))
pMat # calls print(pMat), which calls print.ggmatrix(pMat)</pre>
```

# **Description**

Small function to print a plot if the R session is interactive or in a travis build

# Usage

```
print_if_interactive(p)
```

# **Arguments**

p plot to be displayed

88 psychademic

psychademic

UCLA canonical correlation analysis data

# **Description**

This data contains 600 observations on eight variables

# Usage

data(psychademic)

#### **Format**

A data frame with 600 rows and 8 variables

#### **Details**

- locus\_of\_control psychological
- self\_concept psychological
- motivation psychological. Converted to four character groups
- read academic
- write academic
- math academic
- science academic
- female academic. Dropped from original source
- sex academic. Added as a character version of female column

#### References

R Data Analysis Examples | Canonical Correlation Analysis. UCLA: Institute for Digital Research and Education. from http://www.stats.idre.ucla.edu/r/dae/canonical-correlation-analysis (accessed May 22, 2017).

putPlot 89

putPlot Put Plot

#### **Description**

Function to place your own plot in the layout.

# Usage

```
putPlot(pm, value, i, j)
## S3 replacement method for class 'ggmatrix'
pm[i, j, ...] <- value</pre>
```

#### **Arguments**

```
pm ggally object to be altered value ggplot object to be placed i row from the top j column from the left ignored
```

### Author(s)

Barret Schloerke < schloerke@gmail.com>

```
custom_car <- ggpairs(mtcars[, c("mpg", "wt", "cyl")], upper = "blank", title = "Custom Example")</pre>
# ggplot example taken from example(geom_text)
  plot <- ggplot2::ggplot(mtcars, ggplot2::aes(x=wt, y=mpg, label=rownames(mtcars)))</pre>
  plot <- plot +
    ggplot2::geom_text(ggplot2::aes(colour=factor(cyl)), size = 3) +
    ggplot2::scale_colour_discrete(1=40)
custom_car[1, 2] <- plot</pre>
personal_plot <- ggally_text(</pre>
  "ggpairs allows you\nto put in your\nown plot.\nLike that one.\n <---"
custom_car[1, 3] <- personal_plot</pre>
# custom_car
# remove plots after creating a plot matrix
custom_car[2,1] <- NULL</pre>
custom_car[3,1] <- "blank" # the same as storing null</pre>
custom_car[3,2] <- NULL</pre>
custom_car
```

90 scag\_order

rescale01

Rescaling functions

# Description

Rescaling functions

# Usage

```
range01(x)
max1(x)
mean0(x)
min0(x)
rescale01(x, xlim = NULL)
rescale11(x, xlim = NULL)
```

# Arguments

x numeric vector xlim value used in range

scag\_order

Find order of variables

# Description

Find order of variables based on a specified scagnostic measure by maximizing the index values of that measure along the path.

# Usage

```
scag_order(scag, vars, measure)
```

# **Arguments**

scag scagnostics object

vars character vector of the variables to be ordered measure scagnostics measure to order according to

scatmat 91

# Value

character vector of variable ordered according to the given scagnostic measure

# Author(s)

Barret Schloerke

scatmat	scatmat - plot the lowertriangle plots and density plots of the scatter plot matrix.
	1

# Description

function for making scatterplots in the lower triangle and diagonal density plots.

# Usage

```
scatmat(data, columns = 1:ncol(data), color = NULL, alpha = 1)
```

# Arguments

data	a data matrix. Should contain numerical (continuous) data.
columns	an option to choose the column to be used in the raw dataset. Defaults to 1:ncol(data)
color	an option to group the dataset by the factor variable and color them by different colors. Defaults to $\ensuremath{NULL}$
alpha	an option to set the transparency in scatterplots for large data. Defaults to 1.

# Author(s)

Mengjia Ni, Di Cook <dicook@monash.edu>

```
data(flea)
scatmat(flea, columns=2:4)
scatmat(flea, columns= 2:4, color="species")
```

92 skewness

singleClassOrder	Order axis variables
011161C0140001 4C1	Order divis renteletes

# Description

Order axis variables by separation between one class and the rest (most separation to least).

#### Usage

```
singleClassOrder(classVar, axisVars, specClass = NULL)
```

# Arguments

class Variable (vector from original dataset)
axisVars variables to be plotted as axes (data frame)

specClass character string matching to level of classVar; instead of looking for separation

between any class and the rest, will only look for separation between this class

and the rest

#### Value

character vector of names of axisVars ordered such that the first variable has the most separation between one of the classes and the rest, and the last variable has the least (as measured by F-statistics from an ANOVA)

#### Author(s)

Jason Crowley < crowley.jason.s@gmail.com>

skewness	Sample skewness

## **Description**

Calculate the sample skewness of a vector while ignoring missing values.

# Usage

```
skewness(x)
```

# **Arguments**

x numeric vector

str.ggmatrix 93

# Value

```
sample skewness of x
```

# Author(s)

Jason Crowley <crowley.jason.s@gmail.com>

str.ggmatrix

ggmatrix structure

#### **Description**

View the condensed version of the ggmatrix object. The attribute "class" is ALWAYS altered to "\_class" to avoid recursion.

#### Usage

```
## S3 method for class 'ggmatrix'
str(object, ..., raw = FALSE)
```

# Arguments

object ggmatrix object to be viewed
... passed on to the default str method

raw boolean to determine if the plots should be converted to text or kept as original

objects

twitter\_spambots

Twitter spambots

# Description

A network of spambots found on Twitter as part of a data mining project.

### Usage

```
data(twitter_spambots)
```

#### **Format**

An object of class network with 120 edges and 94 vertices.

94 uppertriangle

#### **Details**

Each node of the network is identified by the Twitter screen name of the account and further carries five vertex attributes:

- location user's location, as provided by the user
- lat latitude, based on the user's location
- lon longitude, based on the user's location
- followers number of Twitter accounts that follow this account
- friends number of Twitter accounts followed by the account

# Author(s)

Amos Elberg

uppertriangle uppertriangle - rearrange dataset as the preparation of ggscatmat function

# Description

function for making the dataset used to plot the uppertriangle plots.

# Usage

```
uppertriangle(
  data,
  columns = 1:ncol(data),
  color = NULL,
  corMethod = "pearson"
)
```

# **Arguments**

data a data matrix. Should contain numerical (continuous) data.

columns an option to choose the column to be used in the raw dataset. Defaults to

1:ncol(data)

color an option to choose a factor variable to be grouped with. Defaults to (NULL)

corMethod method argument supplied to cor

#### Author(s)

Mengjia Ni, Di Cook <dicook@monash.edu>

v1\_ggmatrix\_theme 95

## **Examples**

```
data(flea)
head(uppertriangle(flea, columns=2:4))
head(uppertriangle(flea))
head(uppertriangle(flea, color="species"))
```

v1\_ggmatrix\_theme

Modify a ggmatrix object by adding an ggplot2 object to all plots

# Description

Modify a ggmatrix object by adding an ggplot2 object to all plots

# Usage

```
v1_ggmatrix_theme()
```

# **Examples**

```
ggpairs(iris, 1:2) + v1_ggmatrix_theme()
# move the column names to the left and bottom
ggpairs(iris, 1:2, switch = "both") + v1_ggmatrix_theme()
```

```
wrap_fn_with_param_arg
```

Wrap a function with different parameter values

#### **Description**

Wraps a function with the supplied parameters to force different default behavior. This is useful for functions that are supplied to ggpairs. It allows you to change the behavior of one function, rather than creating multiple functions with different parameter settings.

# Usage

```
wrap_fn_with_param_arg(
  funcVal,
  params = NULL,
  funcArgName = deparse(substitute(funcVal))
)
wrapp(funcVal, params = NULL, funcArgName = deparse(substitute(funcVal)))
wrap(funcVal, ..., funcArgName = deparse(substitute(funcVal)))
wrap_fn_with_params(funcVal, ..., funcArgName = deparse(substitute(funcVal)))
```

#### **Arguments**

funcVal function that the params will be applied to. The function should follow the api of function(data,mapping,...){}. funcVal is allowed to be a string of one of the ggally\_NAME functions, such as "points" for ggally\_points or "facetdensity" for ggally\_facetdensity.

params named vector or list of parameters to be applied to the funcVal name of function to be displayed
... named parameters to be supplied to wrap\_fn\_with\_param\_arg

#### **Details**

wrap is identical to wrap\_fn\_with\_params. These function take the new parameters as arguments. wrapp is identical to wrap\_fn\_with\_param\_arg. These functions take the new parameters as a single list.

The params and fn attributes are there for debugging purposes. If either attribute is altered, the function must be re-wrapped to have the changes take effect.

#### Value

a function(data, mapping,...){} that will wrap the original function with the parameters applied as arguments

```
# small function to display plots only if it's interactive
p_ <- GGally::print_if_interactive</pre>
# example function that prints 'val'
fn <- function(data, mapping, val = 2) {</pre>
 print(val)
fn(data = NULL, mapping = NULL) # 2
# wrap function to change default value 'val' to 5 instead of 2
wrapped_fn1 \leftarrow wrap(fn, val = 5)
wrapped_fn1(data = NULL, mapping = NULL) # 5
# you may still supply regular values
wrapped_fn1(data = NULL, mapping = NULL, val = 3) # 3
# wrap function to change 'val' to 5 using the arg list
wrapped_fn2 <- wrap_fn_with_param_arg(fn, params = list(val = 5))</pre>
wrapped_fn2(data = NULL, mapping = NULL) # 5
# change parameter settings in ggpairs for a particular function
## Goal output:
regularPlot <- ggally_points(</pre>
 iris,
 ggplot2::aes(Sepal.Length, Sepal.Width),
 size = 5, color = "red"
```

```
wrap_fn_with_param_arg
```

```
97
```

```
p_(regularPlot)

# Wrap ggally_points to have parameter values size = 5 and color = 'red'
w_ggally_points <- wrap(ggally_points, size = 5, color = "red")
wrappedPlot <- w_ggally_points(
    iris,
    ggplot2::aes(Sepal.Length, Sepal.Width)
)
p_(wrappedPlot)

# Double check the aes parameters are the same for the geom_point layer
identical(regularPlot$layers[[1]]$aes_params, wrappedPlot$layers[[1]]$aes_params)

# Use a wrapped function in ggpairs
pm <- ggpairs(iris, 1:3, lower = list(continuous = wrap(ggally_points, size = 5, color = "red")))
p_(pm)
pm <- ggpairs(iris, 1:3, lower = list(continuous = w_ggally_points))
p_(pm)</pre>
```

# **Index**

*Topic datasets	AsIs, 71
australia_PISA2012,6	asNetwork, 50, 55, 60
flea, 9	augment, 8, 64, 65
happy, 81	australia_PISA2012,6
nasa, <u>85</u>	_ /
pigs, 86	brew_colors, 7
psychademic, 88	brewer.pal,55
twitter_spambots, 93	broomify, 8, <i>64</i>
*Topic <b>hplot</b>	
getPlot, 11	cooks.distance, 23, 24, 64
ggally_barDiag, 12	cor, 36, 37, 74, 94
ggally_blank, 12	cut, <i>36</i> , <i>51</i> , <i>56</i>
ggally_box, 13	51.55
ggally_cor, 14	degree, <i>51</i> , <i>55</i>
ggally_density, 15	edgeset.constructors, 50, 55, 60
ggally_densityDiag, 16	eval_data_col, 8
ggally_denstrip, 17	expand_range, 51, 55
ggally_dot, 18	expand_1 ange, 31, 33
ggally_dot_and_box, 19	facet_grid, 39, 44, 46, 67
ggally_facetbar, 20	find_plot_type, 9
ggally_facetdensity, 21	flea, 9
ggally_facetdensitystrip, 21	fn_switch, 10
ggally_facethist, 22	
ggally_na, 23	geom_line, 5, 25-27
ggally_points, 30	geom_point, 35
ggally_ratio, 31	<pre>geom_rect, 5</pre>
ggally_smooth, 32	geom_smooth, 27, 32
ggally_text, 33	geom_text, 37, 52, 57
ggmatrix, 45	geom_tile, 31
ggpairs, 65	getPlot, 11
putPlot, 89	ggally_barDiag, 12
+.gg, 4, 4	ggally_blank, 12
[.ggmatrix(getPlot), 11	ggally_blankDiag(ggally_blank), 12
[.glyphplot(glyphplot), 78	ggally_box, 13
<pre>[<ggmatrix(putplot), 89<="" pre=""></ggmatrix(putplot),></pre>	<pre>ggally_box_no_facet (ggally_box), 13</pre>
	ggally_cor, 14
add_ref_boxes, 5	ggally_density, 15
add_ref_lines, 5	ggally_densityDiag, 16
aes, 8, 39, 43, 66, 83, 84	ggally_denstrip, 17
arrow, 52, 57	ggally_diagAxis, 17

INDEX 99

ggally_dot, 18	identity, 79
ggally_dot_and_box, 19	igraph, 50, 55, 60
ggally_dot_no_facet (ggally_dot), 18	influence, 25, 28, 29
ggally_facetbar, 20	intergraph, 50, 55, 60
ggally_facetdensity, 21	is.glyphplot(glyphplot),78
ggally_facetdensitystrip, 21	<pre>is_character_column(is_horizontal), 82</pre>
ggally_facethist, 22	is_horizontal, 82
ggally_na, 23	
ggally_naDiag (ggally_na), 23	labellers, <i>39</i> , <i>46</i> , <i>66</i>
ggally_nostic_cooksd, 23	lm, <i>64</i>
ggally_nostic_hat, 24	lowertriangle, 83
ggally_nostic_line, 24, 25, 25, 27-29	
ggally_nostic_resid, 26, 30	mapping_string, 83
ggally_nostic_se_fit, 28	mapping_swap_x_y, 84
ggally_nostic_sigma, 29	$\max 1 \text{ (rescale01)}, 90$
ggally_nostic_std_resid, 30	mean0 (rescale01), 90
ggally_points, 30	min0 (rescale01), $90$
ggally_ratio, 31	model_beta_label
ggally_smooth, 32	(model_response_variables), 84
ggally_smooth_lm (ggally_smooth), 32	<pre>model_beta_variables</pre>
ggally_smooth_loess(ggally_smooth), 32	(model_response_variables), 84
ggally_text, 33	model_response_variables, 84
ggcoef, 34	0.5
ggcorr, 35	nasa, 85
ggduo, 38, 43, 64, 65, 78	network, 50, 53, 55, 58, 60
ggfacet, 43	nigo 86
gglegend, 44	pigs, 86 plot.network, <i>53</i> , <i>58</i>
ggmatrix, 45, 64	predict, 28
ggmatrix_gtable, 48, 87	
ggmatrix_progress, 40, 47, 48, 49, 64, 67	<pre>print.ggmatrix, 87 print.glyphplot(glyphplot), 78</pre>
ggnet, 49, 53, 58	
ggnet2, 49, 53, 53	print.legend_guide_box(grab_legend), 80
ggnetworkmap, 59	print_if_interactive, 87 progress_bar, 40, 47, 49, 64, 67
ggnostic, 63	
ggpairs, 43, 64, 65	psychademic, 88
ggparcoord, 70	putPlot, 89
ggplot, 26	quantile, <i>51</i> , <i>56</i>
ggscatmat, 74	quantitie, 51, 50
ggsurv, 75	range01 (rescale01), 90
ggts, 77	RColorBrewer, 55
glance, 8	rescale01,90
glm, 64	rescale11 (rescale01), 90
glyphplot, 78	residuals, 27
glyphs, 79	resolution, 78, 79
gplot, 53, 58	rstandard, 30
gplot.layout, <i>50</i> , <i>51</i> , <i>55</i>	
grab_legend, 40, 47, 67, 80	scag_order, 90
-	scatmat, 91
happy, 81	singleClassOrder,92

100 INDEX

```
skewness, 92
sna, 50, 53, 55, 58
str.ggmatrix, 93
\operatorname{substr}, 52, 56
theme, 4, 37, 52, 57
tidy, 8, 34
{\tt twitter\_spambots}, \frac{93}{}
unit, 46
uppertriangle, 94
v1_ggmatrix_theme, 95
wrap, 40, 43, 44, 68
wrap (wrap_fn_with_param_arg), 95
wrap_fn_with_param_arg, 66, 95
wrap_fn_with_params
         (wrap_fn_with_param_arg), 95
wrapp (wrap_fn_with_param_arg), 95
```