

# Package ‘ggupset’

May 5, 2020

**Type** Package

**Title** Combination Matrix Axis for 'ggplot2' to Create 'UpSet' Plots

**Version** 0.3.0

**URL** <https://github.com/const-ae/ggupset>

**BugReports** <https://github.com/const-ae/ggupset/issues>

**Description** Replace the standard x-axis in 'ggplots' with a combination matrix to visualize complex set overlaps. 'UpSet' has introduced a new way to visualize the overlap of sets as an alternative to Venn diagrams.

This package provides a simple way to produce such plots using 'ggplot2'. In addition it can convert any categorical axis into a combination matrix axis.

**License** GPL-3

**Encoding** UTF-8

**LazyData** true

**RoxygenNote** 7.0.2

**Depends** R (>= 2.10)

**Suggests** testthat

**Imports** ggplot2 (>= 3.3.0), gtable, grid, tibble, rlang, scales

**NeedsCompilation** no

**Author** Constantin Ahlmann-Eltze [aut, cre]  
(<<https://orcid.org/0000-0002-3762-068X>>)

**Maintainer** Constantin Ahlmann-Eltze <[artjom31415@googlemail.com](mailto:artjom31415@googlemail.com)>

**Repository** CRAN

**Date/Publication** 2020-05-05 10:40:03 UTC

## R topics documented:

axis_combmatrix . . . . .	2
df_complex_conditions . . . . .	3

gene_pathway_membership . . . . .	4
scale_x_mergelist . . . . .	4
scale_x_upset . . . . .	5
theme_combmatrix . . . . .	6
tidy_movies . . . . .	8

<b>Index</b>	<b>9</b>
--------------	----------

---

axis_combmatrix	<i>Convert delimited text labels into a combination matrix axis</i>
-----------------	---

---

## Description

The function splits the text based on the sep argument and views each occurring element as potential set.

## Usage

```
axis_combmatrix(
  sep = "[^[:alnum:]]+",
  levels = NULL,
  xlim = NULL,
  ylim = NULL,
  expand = TRUE,
  clip = "on",
  ytrans = "identity"
)
```

## Arguments

sep	The separator that is used to split the string labels. Can be a regex. Default: "[^[:alnum:]]+"
levels	The selection of string elements that are displayed in the combination matrix axis. Default: NULL, which means simply all elements in the text labels are used
xlim, ylim	The limits for the x and y axes
expand	Boolean with the same effect as in <code>ggplot2::coord_cartesian()</code> . Default: TRUE
clip	String with the same effect as in <code>ggplot2::coord_cartesian()</code> . Default: "on"
ytrans	transformers for y axis. For more information see <code>ggplot2::coord_trans()</code> . Default: "identity"

## Details

Technically the function appends a coord system to the ggplot object. To maintain compatibility additional arguments like `ytrans`, `ylim`, and `clip` are forwarded to `coord_trans()`.

*Note:* make sure that the argument to the 'x' aesthetic is character vector that contains the sep sequence. The only exception is if `axis_combmatrix()` is combined with a `scale_x_mergelist()`. This pattern works because in the first step `scale_x_mergelist()` turns a list argument to 'x' into a character vector that `axis_combmatrix()` can work with.

## Examples

```
library(ggplot2)
mtcars$combined <- paste0("Cyl: ", mtcars$cyl, "_Gears: ", mtcars$gear)
head(mtcars)
ggplot(mtcars, aes(x=combined)) +
  geom_bar() +
  axis_combmatrix(sep = "_")
```

---

df\_complex\_conditions *A fictional biological dataset with a complex experimental design*

---

## Description

A fictional biological dataset with a complex experimental design

## Usage

```
df_complex_conditions
```

## Format

a data frame with 360 rows and 4 variables

- `KO`. Boolean value if the sample had a knock out.
- `DrugA`. character vector with "Yes" and "No" elements indicating if the sample was treated with drug A.
- `Timepoint`. Numeric vector with elements 8, 24, and 48 indicating the time of measurement since the beginning of the experiment.
- `response`. Numeric vector with the response of the sample to the treatment conditions. Could for example be the concentration of a metabolite.

## Examples

```
dim(df_complex_conditions)
head(df_complex_conditions)
```

---

gene\_pathway\_membership

*A fictional dataset describing which genes belong to certain pathways*

---

### Description

A fictional dataset describing which genes belong to certain pathways

### Usage

```
gene_pathway_membership
```

### Format

a matrix with 6 rows and 37 columns. Each row is one pathway, with its name given as ‘rownames’ and each column is a gene. The values in the matrix are Boolean indicators if the gene is a member of the pathway.

### Examples

```
dim(gene_pathway_membership)
gene_pathway_membership[, 1:15]
```

---

scale\_x\_mergelist

*Merge list columns into character vectors*

---

### Description

The function handles list columns by collapsing them into delimited strings using the sep argument. This is useful to show sets and in combination with the axis\_combmatrix() function.

### Usage

```
scale_x_mergelist(sep = "-", ..., position = "bottom")
```

### Arguments

sep	String the is used to delimit the elements in each list entry. Default: "-".
...	additional arguments that are passed on to ggplot2::scale_x_discrete
position	either "top" or "bottom" to specify where the x axis drawn. Default: "bottom"

### See Also

[discrete\\_scale](#)

**Examples**

```
library(ggplot2)
ggplot(tidy_movies[1:100, ], aes(x=Genres)) +
  geom_bar() +
  scale_x_mergelist() +
  theme(axis.text.x = element_text(angle = 90, hjust=1, vjust = 0.5))

ggplot(tidy_movies[1:100, ], aes(x=Genres)) +
  geom_bar() +
  scale_x_mergelist(sep = " & ", name = "Merged Movie Genres", position = "top") +
  theme(axis.text.x = element_text(angle = 90, hjust=0, vjust = 0.5))
```

---

scale\_x\_upset

*Scale to make UpSet plots*


---

**Description**

This function takes a list column and turns it into a combination matrix axis. It internally wraps the call to `scale_x_mergelist()` and `axis_combmatrix()` and makes sure that the elements are sorted by size.

**Usage**

```
scale_x_upset(
  order_by = c("freq", "degree"),
  n_sets = Inf,
  n_intersections = Inf,
  sets = NULL,
  intersections = NULL,
  reverse = FALSE,
  ytrans = "identity",
  ...,
  position = "bottom"
)
```

**Arguments**

<code>order_by</code>	either "freq" or "degree". Default: "freq"
<code>n_sets</code>	maximum number of sets that are displayed. Default: Inf
<code>n_intersections</code>	maximum number of intersections that are displayed. Default: Inf
<code>sets</code>	character vector that specifies which sets are displayed
<code>intersections</code>	a list of character vectors that specifies which intersections are displayed
<code>reverse</code>	boolean if the order of the intersections is reversed. Default: FALSE

ytrans	transformers for y axis. For more information see axis_combmatrix(). Default: "identity"
...	additional parameters for ggplot2::discrete_scale()
position	either "top" or "bottom" to specify where the combination matrix is drawn. Default: "bottom"

## Examples

```
library(ggplot2)
ggplot(tidy_movies[1:100, ], aes(x=Genres)) +
  geom_bar() +
  scale_x_upset(reverse = TRUE, sets=c("Drama", "Action"))

ggplot(tidy_movies[1:100, ], aes(x=Genres)) +
  geom_bar() +
  scale_x_upset(n_intersections = 5, ytrans="sqrt")

ggplot(tidy_movies[1:100, ], aes(x=Genres, y=year)) +
  geom_boxplot() +
  scale_x_upset(intersections = list(c("Drama", "Comedy"), c("Short"), c("Short", "Animation")),
               sets = c("Drama", "Comedy", "Short", "Animation", "Horror"))
```

---

theme_combmatrix	<i>Theme for the combination matrix</i>
------------------	---

---

## Description

This theme sets the default styling for the combination matrix axis by extending the default ggplot2 theme().

## Usage

```
theme_combmatrix(
  combmatrix.label.make_space = TRUE,
  combmatrix.label.width = NULL,
  combmatrix.label.height = NULL,
  combmatrix.label.extra_spacing = 3,
  combmatrix.label.total_extra_spacing = unit(10, "pt"),
  combmatrix.label.text = NULL,
  combmatrix.panel.margin = unit(c(1.5, 1.5), "pt"),
  combmatrix.panel.striped_background = TRUE,
  combmatrix.panel.striped_background.color.one = "white",
  combmatrix.panel.striped_background.color.two = "#F7F7F7",
  combmatrix.panel.point.size = 3,
  combmatrix.panel.line.size = 1.2,
  combmatrix.panel.point.color.fill = "black",
  combmatrix.panel.point.color.empty = "#E0E0E0",
  ...
)
```

**Arguments**

`combmatrix.label.make_space`  
 Boolean indicator if the y-axis label is moved so far to the left to make enough space for the combination matrix labels. Default: TRUE

`combmatrix.label.width`  
 A unit that specifies how much space to make for the labels of the combination matrix. Default: NULL, which means the width of the label text is used

`combmatrix.label.height`  
 A unit that specifies how high the combination matrix should be. Default: NULL, which means that the height of the label text + `combmatrix.label.total_extra_spacing` + `#rows * combmatrix.label.extra_spacing` is used. Default: 3

`combmatrix.label.extra_spacing`  
 A single number for the additional height per row. Default: `unit(10,"pt")`

`combmatrix.label.total_extra_spacing`  
 A unit that specifies the total offset for the height of the combination matrix

`combmatrix.label.text`  
 A `element_text()` to style the label text of the combination matrix. Default: NULL, which means the style of `axis.text.y` is used.

`combmatrix.panel.margin`  
 A two element unit vector to specify top and bottom margin around the combination matrix. Default: `unit(c(1.5,1.5),"pt")`

`combmatrix.panel.stripped_background`  
 Boolean to indicate if the background of the plot is striped. Default: TRUE

`combmatrix.panel.stripped_background.color.one`  
 Color of the first kind of stripes. Default: "white"

`combmatrix.panel.stripped_background.color.two`  
 Color of the second kind of stripes. Default: "#F7F7F7"

`combmatrix.panel.point.size`  
 Number to specify the size of the points in the combination matrix. Default: 3

`combmatrix.panel.line.size`  
 Number to specify the size of the lines connecting the points. Default: 1.2

`combmatrix.panel.point.color.fill`  
 Color of the filled points. Default: "black"

`combmatrix.panel.point.color.empty`  
 Color of the empty points. Default: "#E0E0E0"

... additional arguments that are passed to `theme()`

**Examples**

```
library(ggplot2)
# Ensure that the y-axis label is next to the axis by setting
# combmatrix.label.make_space to FALSE
ggplot(tidy_movies[1:100, ], aes(x=Genres)) +
  geom_bar() +
  scale_x_upset() +
  theme_combmatrix(combmatrix.label.text = element_text(color = "black", size=15),
```

```

      combmatrix.label.make_space = FALSE,
      plot.margin = unit(c(1.5, 1.5, 1.5, 65), "pt"))

# Change the color of the background stripes
ggplot(tidy_movies[1:100, ], aes(x=Genres)) +
  geom_bar() +
  scale_x_upset() +
  theme_combmatrix(combmatrix.panel.striped_background = TRUE,
                  combmatrix.panel.striped_background.color.one = "grey")

```

---

tidy\_movies

*Tidy version of the movies dataset from the ggplot2 package*


---

## Description

The original `ggplot2movies::movies` dataset has 7 columns that contain indicators if a movies belongs to a certain genre. In this version the 7 columns are collapsed to a single list column to create a tidy dataset. It also has information on only 5,000 movies to reduce the size of the dataset. Furthermore each star rating is in its own row.

## Usage

```
tidy_movies
```

## Format

a data frame with 50,000 rows and 10 columns

- title. The title of the movie.
- year. Year of release.
- budget. Total budget (if known) in US dollars.
- length. Length in minutes.
- rating. Average IMDB user rating.
- votes. Number of IMDB user who rated this movie.
- mpaa. MPAA rating
- Genres. List column with all genres the movie belongs to
- stars, percent\_rating. The number of stars and the corresponding percentage of people rating the movie with this many stars.

## Examples

```
dim(tidy_movies)
head(tidy_movies)
```



# Index

## \*Topic **datasets**

- df\_complex\_conditions, [3](#)
- gene\_pathway\_membership, [4](#)
- tidy\_movies, [8](#)

axis\_combmatrix, [2](#)

df\_complex\_conditions, [3](#)  
discrete\_scale, [4](#)

gene\_pathway\_membership, [4](#)

scale\_x\_mergelist, [4](#)  
scale\_x\_upset, [5](#)

theme\_combmatrix, [6](#)  
tidy\_movies, [8](#)