

Package: marimekko (via r-universe)

May 17, 2026

Title Marimekko Plots for 'ggplot2'

Version 0.1.0

Description Create marimekko (mosaic) plots as a 'ggplot2' layer.
Column widths encode marginal proportions of one categorical variable and segment heights encode conditional proportions of a second categorical variable. Based on the mosaic display method by Hartigan and Kleiner (1981)
<doi:10.1007/978-1-4613-9464-8_37>.

License MIT + file LICENSE

Encoding UTF-8

Roxygen list(markdown = TRUE)

RoxygenNote 7.3.3

Depends R (>= 4.1.0)

Imports ggplot2 (>= 3.5.0)

Suggests knitr, pkgdown, plotly, rmarkdown, testthat (>= 3.0.0),
vdiff (>= 1.0.0)

VignetteBuilder knitr

Config/testthat/edition 3

Collate 'marimekko-package.R' 'geom-marimekko.R' 'fortify-marimekko.R'
'stat-marimekko-tiles.R' 'theme-marimekko.R'

Repository <https://gogonzo.r-universe.dev>

Date/Publication 2026-03-18 07:00:51 UTC

RemoteUrl <https://github.com/gogonzo/marimekko>

RemoteRef HEAD

RemoteSha ac1cc967973fc7c9b939a7278f6e59704dd963ca

Contents

fortify_marimekko	2
geom_marimekko	3

geom_marimekko_label	6
geom_marimekko_text	7
marimekko_pal	8
StatMarimekkoTiles	9
theme_marimekko	10

Index	12
--------------	-----------

fortify_marimekko	<i>Compute marimekko tile rectangles as a data frame</i>
-------------------	--

Description

Compute marimekko tile rectangles as a data frame

Usage

```
fortify_marimekko(
  data,
  formula,
  weight = NULL,
  gap = 0.01,
  gap_x = NULL,
  gap_y = NULL,
  standardize = FALSE
)
```

Arguments

data	A data frame.
formula	A one-sided formula specifying the mosaic hierarchy, using the same syntax as geom_marimekko() . Example: <code>~ Class Survived</code> .
weight	Name of the weight variable (unquoted or string), or NULL for unweighted counts. Default NULL.
gap	Numeric. Size of gap between tiles. Default 0.01.
gap_x	Numeric. Horizontal gap. Overrides gap for x. Default NULL.
gap_y	Numeric. Vertical gap. Overrides gap for y. Default NULL.
standardize	Logical. Equal-width columns. Default FALSE.

Value

A data frame with columns for each formula variable, plus `fill`, `colour`, `xmin`, `xmax`, `ymin`, `ymax`, `x`, `y`, `weight`, `.proportion`, `.marginal`, and `.residuals`.

Examples

```
titanic <- as.data.frame(Titanic)
fortify_marimekko(titanic, formula = ~ Class | Survived, weight = Freq)

# 3-variable formula
fortify_marimekko(titanic, formula = ~ Class | Survived | Sex, weight = Freq)
```

geom_marimekko

*Generalized mosaic plot with formula-based variable nesting***Description**

Generalized mosaic plot with formula-based variable nesting

Usage

```
geom_marimekko(
  mapping = NULL,
  data = NULL,
  formula = NULL,
  gap = 0.01,
  gap_x = NULL,
  gap_y = NULL,
  colour = NULL,
  alpha = 0.9,
  show_percentages = FALSE,
  na.rm = FALSE,
  show.legend = NA,
  inherit.aes = TRUE,
  ...
)
```

Arguments

mapping	Aesthetic mapping. Optionally accepts <code>fill</code> and <code>weight</code> for pre-aggregated data. If <code>fill</code> is not specified, it defaults to the last variable in the formula. The <code>fill</code> variable controls tile colour and does not need to appear in the formula.
data	A data frame.
formula	A one-sided formula specifying the mosaic hierarchy. See the sections above for a detailed explanation. Quick reference: <ul style="list-style-type: none"> • <code>~ a b</code> — <code>h(a)</code>, <code>v(b)</code>. Standard mosaic. • <code>~ a b c</code> — <code>h(a)</code>, <code>v(b)</code>, <code>h(c)</code>. Alternating mosaic. • <code>~ a + b c</code> — <code>h(a)</code>, <code>h(b)</code>, <code>v(c)</code>. Double decker.

	• $\sim a b + c$ — $h(a), v(b), v(c)$. Multiple vertical variables.
gap	Numeric. Gap between tiles as fraction of plot area. Default 0.01 .
gap_x	Numeric. Horizontal gap override. Default NULL (uses gap).
gap_y	Numeric. Vertical gap override. Default NULL (uses gap).
colour	Tile border colour. Default NULL (no border). Can also be mapped via <code>aes(colour = variable)</code> .
alpha	Tile transparency. Default 0.9 .
show_percentages	Logical. If TRUE, appends marginal percentage to each x-axis label. Default FALSE.
na.rm	Logical. Remove missing values. Default FALSE.
show.legend	Logical. Show legend. Default NA.
inherit.aes	Logical. Inherit aesthetics from <code>ggplot()</code> . Default TRUE.
...	Additional arguments passed to the layer.

Value

A list of `ggplot2` layers (geom + axis scales).

How the formula works

The formula uses two operators to encode the full partitioning hierarchy in a single expression:

- | (**pipe**) Separates nesting levels. Each | switches the splitting direction, alternating horizontal, vertical, horizontal, vertical, and so on. The first variable (or group) listed is the **outermost** split — it partitions the entire plot area. Each subsequent level partitions the tiles created by the previous level.
- + (**plus**) Groups variables at the **same** nesting level. All variables joined by + share the same splitting direction and are applied sequentially within that level. The first + variable partitions the current tiles, then the second + variable further subdivides those tiles, still in the same direction.

Reading order — outermost to innermost

The formula is read left to right, from the coarsest (outermost) partition to the finest (innermost):

- $\sim a | b$ First split the plot horizontally by a (columns whose widths reflect marginal proportions of a). Then, within each column, split vertically by b (rows whose heights reflect conditional proportions of b given a). This is the classic two-variable marimekko / mosaic plot.
- $\sim a | b | c$ Horizontal by a , then vertical by b , then horizontal again by c . Three levels of nesting with alternating directions ($h \rightarrow v \rightarrow h$).
- $\sim a + b | c$ Horizontal by a , then horizontal again by b (same direction because + groups them), then vertical by c . This is the **double decker** pattern — all horizontal splits first, with a single vertical split at the end.
- $\sim a | b + c$ Horizontal by a , then vertical by b , then vertical again by c . Two vertical variables nested within each column.

Computed variables

The stat computes the following variables that can be accessed with `ggplot2::after_stat()`:

- .proportion Conditional proportion of the tile within its immediate parent. For a formula $\sim a \mid b$, this is the proportion of b within each level of a , i.e. $P(b \mid a)$. Values sum to 1 within each parent tile. Useful for mapping to alpha to fade tiles by their local share: `aes(alpha = after_stat(.proportion))`.
- .marginal Joint (marginal) proportion of the tile relative to the whole dataset, i.e. n_{cell}/N . Values sum to 1 across all tiles. Used internally for x-axis percentage labels when `show_percentages = TRUE`, and can be mapped to aesthetics to emphasise cells by overall frequency.
- .residuals Pearson residual measuring departure from statistical independence between the horizontal and vertical variable groups. Computed as $(O - E)/\sqrt{E}$, where O is the observed cell count and E is the count expected under independence. Positive values indicate the cell is **more** frequent than expected; negative values indicate **less** frequent. When only one direction (all horizontal or all vertical) is present, `.residuals` is set to 0. Map to alpha or fill to highlight deviations: `aes(alpha = after_stat(abs(.residuals)))`.

Examples

```
library(ggplot2)

titanic <- as.data.frame(Titanic)

# 2-variable mosaic
ggplot(titanic) +
  geom_marimekko(
    aes(fill = Survived, weight = Freq),
    formula = ~ Class | Survived
  )

# 3-variable mosaic (h -> v -> h)
ggplot(titanic) +
  geom_marimekko(
    aes(fill = Survived, weight = Freq),
    formula = ~ Class | Survived | Sex
  )

# Multi-variable fill with interaction()
ggplot(titanic) +
  geom_marimekko(
    aes(fill = interaction(Sex, Survived), weight = Freq),
    formula = ~ Class | Sex + Survived
  )

# Fade tiles by conditional proportion
ggplot(titanic) +
  geom_marimekko(
    aes(fill = Survived, alpha = after_stat(.proportion), weight = Freq),
    formula = ~ Class | Survived
  ) +
```

```

guides(alpha = "none")

# Highlight cells that deviate from independence
ggplot(titanic) +
  geom_marimekko(
    aes(fill = Survived, alpha = after_stat(abs(.residuals)), weight = Freq),
    formula = ~ Class | Survived
  ) +
  guides(alpha = "none")

```

geom_marimekko_label *Add labels with background to a marimekko plot*

Description

Add labels with background to a marimekko plot

Usage

```

geom_marimekko_label(
  mapping = NULL,
  data = NULL,
  position = "identity",
  ...,
  size = 3.5,
  colour = "black",
  fill = alpha("white", 0.7),
  label.padding = unit(0.15, "lines"),
  na.rm = FALSE,
  show.legend = FALSE,
  inherit.aes = FALSE
)

```

Arguments

mapping	Set of aesthetic mappings. Only label is required. Use <code>ggplot2::after_stat()</code> for computed variables.
data	A data frame. Default NULL (uses plot data; tile positions come from <code>geom_marimekko()</code>).
position	Position adjustment. Default "identity".
...	Additional arguments passed to the layer.
size	Text size. Default 3.5.
colour	Text colour. Default "white" for text, "black" for labels.
fill	Label background colour. Default <code>alpha("white", 0.7)</code> .
label.padding	Amount of padding around label. Default <code>ggplot2::unit(0.15, "lines")</code> .
na.rm	Logical. Remove missing values. Default FALSE.
show.legend	Logical. Show legend. Default FALSE.
inherit.aes	Logical. Inherit aesthetics. Default FALSE.

Value

A ggplot2 layer.

Examples

```
library(ggplot2)

titanic <- as.data.frame(Titanic)
ggplot(titanic) +
  geom_marimekko(
    aes(fill = Survived, weight = Freq),
    formula = ~ Class | Survived
  ) +
  geom_marimekko_label(aes(label = after_stat(weight)))
```

geom_marimekko_text *Add text labels to a marimekko plot*

Description

Add text labels to a marimekko plot

Usage

```
geom_marimekko_text(
  mapping = NULL,
  data = NULL,
  position = "identity",
  ...,
  size = 3.5,
  colour = "white",
  na.rm = FALSE,
  show.legend = FALSE,
  inherit.aes = FALSE
)
```

Arguments

mapping	Set of aesthetic mappings. Only label is required. Use <code>ggplot2::after_stat()</code> for computed variables.
data	A data frame. Default NULL (uses plot data; tile positions come from <code>geom_marimekko()</code>).
position	Position adjustment. Default "identity".
...	Additional arguments passed to the layer.
size	Text size. Default 3.5.
colour	Text colour. Default "white" for text, "black" for labels.

na.rm	Logical. Remove missing values. Default FALSE.
show.legend	Logical. Show legend. Default FALSE.
inherit.aes	Logical. Inherit aesthetics. Default FALSE.

Value

A ggplot2 layer.

Examples

```
library(ggplot2)

titanic <- as.data.frame(Titanic)
ggplot(titanic) +
  geom_marimekko(
    aes(fill = Survived, weight = Freq),
    formula = ~ Class | Survived
  ) +
  geom_marimekko_text(aes(label = after_stat(weight)))
```

marimekko_pal

Unikko-inspired colour palette

Description

A character vector of 8 bold colours inspired by Marimekko's iconic Unikko poppy pattern. Vibrant, high-contrast tones suited for categorical data visualisation.

Usage

```
marimekko_pal
```

Format

An object of class character of length 8.

StatMarimekkoTiles	<i>Retrieve computed tile positions from a marimekko layer</i>
--------------------	--

Description

Retrieve computed tile positions from a marimekko layer

Usage

```
StatMarimekkoTiles
```

Format

An object of class StatMarimekkoTiles (inherits from Stat, ggproto, gg) of length 2.

Usage with custom geoms

Use StatMarimekkoTiles as the `stat` argument in `ggplot2::layer()` to pair the tile data with any geom. The only requirement is that `geom_marimekko()` must appear **before** the custom layer so that tile positions are computed first.

See Also

[geom_marimekko\(\)](#), [geom_marimekko_text\(\)](#), [geom_marimekko_label\(\)](#), [fortify_marimekko\(\)](#)

Examples

```
library(ggplot2)

titanic <- as.data.frame(Titanic)

# Bubble overlay - point size encodes tile count
ggplot(titanic) +
  geom_marimekko(
    aes(fill = Survived, weight = Freq),
    formula = ~ Class | Survived, alpha = 0.4
  ) +
  layer(
    stat = StatMarimekkoTiles,
    geom = GeomPoint,
    mapping = aes(size = after_stat(weight)),
    data = titanic,
    position = "identity",
    show.legend = FALSE,
    inherit.aes = FALSE,
    params = list(colour = "white", alpha = 0.7)
  ) +
  scale_size_area(max_size = 12)
```

```

# Residual markers – colour and size show deviation from independence
ggplot(titanic) +
  geom_marimekko(
    aes(fill = Survived, weight = Freq),
    formula = ~ Class | Survived
  ) +
  layer(
    stat = StatMarimekkoTiles,
    geom = GeomPoint,
    mapping = aes(
      size = after_stat(abs(.residuals)),
      colour = after_stat(ifelse(.residuals > 0, "over", "under"))
    ),
    data = titanic,
    position = "identity",
    show.legend = TRUE,
    inherit.aes = FALSE,
    params = list(alpha = 0.8)
  ) +
  scale_colour_manual(
    values = c(over = "tomato", under = "steelblue"),
    name = "Deviation"
  ) +
  scale_size_continuous(range = c(1, 8), name = "|Residual|")

```

 theme_marimekko

Minimal theme for marimekko plots

Description

Removes x-axis gridlines and adjusts spacing for mosaic plots. Also applies the [marimekko_pal](#) fill scale.

Usage

```
theme_marimekko(base_size = 12, ...)
```

Arguments

base_size	Base font size. Default 12.
...	Arguments passed to <code>ggplot2::theme_minimal()</code> .

Value

A ggplot2 theme.

Examples

```
library(ggplot2)

titanic <- as.data.frame(Titanic)
ggplot(titanic) +
  geom_marimekko(
    aes(fill = Survived, weight = Freq),
    formula = ~ Class | Survived
  ) +
  theme_marimekko()
```

Index

* datasets

marimekko_pal, 8
StatMarimekkoTiles, 9

fortify_marimekko, 2
fortify_marimekko(), 9

geom_marimekko, 3
geom_marimekko(), 2, 6, 7, 9
geom_marimekko_label, 6
geom_marimekko_label(), 9
geom_marimekko_text, 7
geom_marimekko_text(), 9
ggplot2::after_stat(), 5–7
ggplot2::layer(), 9
ggplot2::theme_minimal(), 10

marimekko_pal, 8, 10

StatMarimekkoTiles, 9

theme_marimekko, 10