

# Package: epivis (via r-universe)

October 25, 2024

**Title** Tools for visualising epidemiological data  
**Version** 0.0.0.9000  
**Description** Static visualisation of patient level linelist data.  
**License** MIT + file LICENSE  
**Encoding** UTF-8  
**Roxygen** list(markdown = TRUE)  
**LazyData** true  
**Imports** cli, rlang, ggplot2, dplyr, purrr, tidyr, stringr, magrittr, scales, glue, gt, ggthemes, lubridate, classInt, aweek  
**RoxygenNote** 7.2.3  
**Suggests** knitr, rmarkdown, outbreaks, forcats, hrbrthemes  
**VignetteBuilder** knitr  
**URL** <https://github.com/epicentre-msf/epivis>  
**BugReports** <https://github.com/epicentre-msf/epivis/issues>  
**Depends** R (>= 2.10)  
**Repository** <https://epicentre-msf.r-universe.dev>  
**RemoteUrl** <https://github.com/epicentre-msf/epivis>  
**RemoteRef** HEAD  
**RemoteSha** 647687117aaa011da9fd66fd0f5636c5110ed9fd

## Contents

dodge_x_labs . . . . .	2
label_breaks . . . . .	2
moissala_measles . . . . .	3
plot_epicurve . . . . .	4
plot_miss_vis . . . . .	6
plot_pyramid . . . . .	7
plot_stacked_bar . . . . .	9

<b>Index</b>	<b>11</b>
--------------	-----------

---

dodge_x_labs	<i>Wrapper function to dodge xAxis labels</i>
--------------	---

---

**Description**

Useful when you have overlapping labels on the xAxis.

**Usage**

```
dodge_x_labs(n.dodge = 2)
```

**Arguments**

n.dodge            passed to [ggplot2::guide\\_axis](#)

---

label_breaks	<i>Format break labels</i>
--------------	----------------------------

---

**Description**

Format break labels

**Usage**

```
label_breaks(breaks, lab_accuracy = 0.1, replace_Inf = TRUE)
```

**Arguments**

breaks            numeric vector of breaks  
lab\_accuracy      accuracy of labels, passed to [scales::number](#)  
replace\_Inf       if Inf is your final break, replace with a + sign in the label?

---

moissala_measles	<i>Simulated Measles outbreak in Moïssala (Chad)</i>
------------------	--

---

### Description

This linelist replicates a measles outbreak over several regions of Southern Chad. It uses realistic distributions and parameters from a measles outbreak.

### Usage

```
moissala_measles
```

### Format

```
moissala_measles:
```

A data frame with 5,028 rows and 30 columns:

**id** epi id

**site** Site name

**case\_name** Case name

**sex** Case sex

**age** Case age

**age\_unit** Age units

**age\_group** Age group

**region** Case region of residence

**sub\_prefecture** Case sub-prefecture of residence

**date\_onset** Date of symptoms onset

**hospitalisation** Hospitalisation status

**date\_admission** Date of hospital admission

**ct\_value** Ct value of RT-PCR

**malaria\_rdt** Malaria RDT result

**fever** presence of fever upon admission

**rash** presence of rash upon admission

**cough** presence of cough upon admission

**red\_eye** presence of red eyes upon admission

**pneumonia** presence of pneumonia upon admission

**encephalitis** presence of encephalitis upon admission

**muac** Middle Upper Arm Circumference (MUAC) upon admission

**muac\_cat** MUAC category

**vacc\_status** Vaccination status

**vacc\_doses** Vaccine doses received

**outcome** Outcome

**date\_death** Date of death

**date\_exit** Date of hospital exit

**epi\_classification** Epidemiological classification

**Source**

<https://epicentre-msf.github.io/gallery/>

---

plot\_epicurve

*Plot incidence over time from patient level data*

---

**Description**

Helper function to plot epidemic curves with ggplot2 with options for grouping data, facets and proportion lines.

**Usage**

```
plot_epicurve(  
  df,  
  date_col,  
  group_col = NULL,  
  facet_col = NULL,  
  prop_col = NULL,  
  prop_numer = NULL,  
  prop_denom = "non_missing",  
  prop_line_colour = "black",  
  prop_line_size = 0.8,  
  floor_date_week = FALSE,  
  label_weeks = FALSE,  
  week_start = 1,  
  date_breaks,  
  date_labels = waiver(),  
  date_max = NULL,  
  sec_date_axis = FALSE,  
  facet_nrow = NULL,  
  facet_ncol = NULL,  
  facet_scales = "fixed",  
  facet_labs = ggplot2::label_wrap_gen(width = 25),  
  facet_lab_pos = "top",  
  group_na_colour = "grey",  
  title = waiver(),  
  subtitle = waiver(),  
  date_lab = waiver(),  
  y_lab = waiver(),  
  group_lab = waiver(),  
  prop_lab = NULL  
)
```

**Arguments**

df	un-aggregated dataframe with a minimum of a date column with a date or POSIX class
date_col	date variable to plot incidence with. Must be provided.
group_col	optional grouping variable to be applied to the fill aesthetic of columns
facet_col	optional faceting variable to split chart into small multiples
prop_col	optional variable to be used to plot a proportion line on top of the epicurve
prop_numer	value(s) in the prop_col variable as a single value or vector to be used to calculate the numerator of the proportion calculation
prop_denom	value(s) in the prop_col variable as a single value or vector to be used to calculate the denominator of the proportion calculation. default "non_missing" will take the sum of all non-missing values in the column.
prop_line_colour	colour of the proportion line. defaults to "red"
prop_line_size	width of the proportion line. defaults to 0.8
floor_date_week	should date_col dates be floored to the Monday of the ISO week they fall in? defaults to TRUE
label_weeks	label primary date axis with week numbers? defaults to TRUE
week_start	day of week defined as the start of the week as integer 1-7 (Monday = 1, Sunday = 7). defaults to 1 (ISO week standard)
date_breaks	date break intervals passed to <code>ggplot2::scale_x_date</code> . defaults to "2 weeks"
date_labels	<code>base::strptime</code> date label code passed to <code>ggplot2::scale_x_date</code> . defaults to "%V" (ISO Week)
date_max	force a date axis max date. Useful for when a week has passed with no incidence and you want to show this on the plot. Setting date_max to the current week will force the date axis to show this week with no incidence.
sec_date_axis	plot a secondary date axis using default calculated ggplot2 date breaks and labels? defaults to FALSE
facet_nrow	nrow argument passed to <code>ggplot2::facet_wrap</code>
facet_ncol	ncol argument passed to <code>ggplot2::facet_wrap</code>
facet_scales	value for the scales argument passed to <code>ggplot2::facet_wrap</code> . Default to fixed.
facet_labs	facet labeller argument passed to <code>ggplot2::facet_wrap</code> . Defaults to <code>ggplot2::label_wrap_gen(width = 25)</code> .
facet_lab_pos	facet label position argument passed to <code>strip.position</code> in <code>ggplot2::facet_wrap</code> . defaults to "top". Options are <code>c("top", "bottom", "left", "right")</code>
group_na_colour	colour for missing values in group_col. defaults to "grey"
title	optional title for the plot
subtitle	optional subtitle for the plot

date_lab	optional label for the date axis. defaults to date_col name if not provided
y_lab	optional label for the Y axis. defaults to n if not provided
group_lab	optional label for the group legend. defaults to group_col name if not provided
prop_lab	label for the proportion line. There is no default so this should be provided when plotting proportion lines

**Value**

a ggplot object

**Examples**

```
library(dplyr)
df_ebola <- dplyr::as_tibble(outbreaks::ebola_sim_clean$linelist)

df_ebola |>
  dplyr::mutate(outcome = forcats::fct_explicit_na(outcome, "Unknown")) |>
  plot_epicurve(
    date_col = date_of_onset,
    group_col = outcome,
    prop_col = outcome,
    prop_numer = "Death",
    prop_denom = c("Death", "Recover"),
    floor_date_week = TRUE,
    date_breaks = "2 weeks",
    sec_date_axis = TRUE,
    date_lab = "Week of onset",
    y_lab = "Incidence",
    group_lab = "Outcome",
    prop_lab = "CFR"
  )
```

---

plot\_miss\_vis

*Missing data visualisation*

---

**Description**

Function to generate a tile plot exploring the missing values for all observations across all variables of a dataframe.

**Usage**

```
plot_miss_vis(
  x,
  facet = NULL,
  col_vec = c("#6a040f", "#cce3de"),
  y_axis_text_size = 8
)
```

**Arguments**

**x** a dataframe  
**facet** a character value of variable to facet the graph  
**col\_vec** a vector of length 2 specifying the color for Missing and Present values respectively  
**y\_axis\_text\_size** a numeric value for the size of the y axis text

**Value**

a ggplot tile graph displaying the missing/present values for all variables of the dataframe

**Examples**

```

# Use simulated measles data

suppressMessages(library(dplyr))

epivis::moissala_measles |>
  filter(site %in% c("Moissala Hospital", "Bouna Hospital")) |>
  plot_miss_vis(facet = "site")
  
```

---

plot_pyramid	<i>Plot Age/Sex Pyramids</i>
--------------	------------------------------

---

**Description**

Plot Age/Sex Pyramids

**Usage**

```

plot_pyramid(
  df,
  age_col,
  gender_col,
  gender_levels,
  facet_col = NULL,
  make_age_groups = TRUE,
  age_breaks = c(seq(0, 80, 10), Inf),
  age_labels = label_breaks(age_breaks),
  drop_age_levels = FALSE,
  gender_labs = NULL,
  x_lab = waiver(),
  y_lab = waiver(),
  colours = c("#486090FF", "#7890A8FF"),
  show_data_labs = FALSE,
  lab_size = 4,
)
  
```

```

lab_in_col = "white",
lab_out_col = "grey30",
lab_nudge_factor = 5,
facet_nrow = NULL,
facet_ncol = NULL,
facet_scales = "fixed",
facet_labs = label_wrap_gen(width = 25),
facet_lab_pos = "top",
add_missing_cap = TRUE
)

```

### Arguments

<code>df</code>	un-aggregated dataframe with a minimum of age and gender variables.
<code>age_col</code>	age variable name in <code>df</code> . Can be either a numeric vector of ages or a character/factor vector of age groups.
<code>gender_col</code>	gender variable name in <code>df</code> with levels indicating male or female.
<code>gender_levels</code>	length 2 character vector with male and female level in <code>gender_col</code> , respectively.
<code>facet_col</code>	optional faceting variable name to split chart into small multiples.
<code>make_age_groups</code>	set to <code>TRUE</code> (default) if <code>age_col</code> is numeric and needs to be binned into groups.
<code>age_breaks</code>	breaks to be used for binning a numerical <code>age_col</code> .
<code>age_labels</code>	break labels to accompany <code>age_breaks</code> . Defaults to <code>epivis::label_breaks(age_breaks)</code> .
<code>drop_age_levels</code>	should age groups with no observations be removed from the chart? Defaults to <code>FALSE</code> .
<code>gender_labs</code>	optional labels for <code>gender_levels</code>
<code>x_lab</code>	optional label for the X axis.
<code>y_lab</code>	optional label for the Y axis.
<code>colours</code>	length 2 character vector of colours used for male and female, respectively.
<code>show_data_labs</code>	show data labels on chart? Defaults to <code>FALSE</code> .
<code>lab_size</code>	data labels size.
<code>lab_in_col</code>	data label colour when placed inside a bar.
<code>lab_out_col</code>	data label colour when placed outside a bar.
<code>lab_nudge_factor</code>	threshold for moving a data label outside a bar. Defaults to 5. Increasing the number increases the distance from the max value required to move a label outside the bar.
<code>facet_nrow</code>	<code>nrow</code> argument passed to <code>ggplot2::facet_wrap</code> .
<code>facet_ncol</code>	<code>ncol</code> argument passed to <code>ggplot2::facet_wrap</code> .
<code>facet_scales</code>	facet scales argument passed to <code>ggplot2::facet_wrap</code> . Should scales be fixed ("fixed", the default), free ("free"), or free in one dimension ("free_x", "free_y")?



**facet\_labs** facet labeller argument passed to `ggplot2::facet_wrap`. Defaults to `ggplot2::label_wrap_gen(width = 25)`.  
**facet\_lab\_pos** facet label position argument passed to `strip.position` in `ggplot2::facet_wrap`. Defaults to "top". Options are `c("top", "bottom", "left", "right")`.  
**add\_missing\_cap** show missing data counts for `age_col` and `gender_col`? Defaults to TRUE.

## Value

a ggplot object

## Examples

```

suppressMessages(library(dplyr))
df_flu <- outbreaks::fluH7N9_china_2013

plot_pyramid(
  df = df_flu,
  age_col = age,
  gender_col = gender,
  gender_levels = c("m", "f")
)

```

---

plot_stacked_bar	<i>Stacked barplot</i>
------------------	------------------------

---

## Description

Stacked barplot

## Usage

```

plot_stacked_bar(
  df,
  cols,
  levels_value,
  keep_na = TRUE,
  use_counts = TRUE,
  flip = FALSE,
  x_lab = waiver(),
  caption = TRUE
)

```

**Arguments**

<code>df</code>	un-aggregated dataframe (linelist).
<code>cols</code>	vector of character/factor variables names in <code>df</code> to be displayed in the barplot.
<code>levels_value</code>	vector of level values to be used for the plotting.
<code>keep_na</code>	logical, default = TRUE. Keep NAs in the graphs and the proportions ?
<code>use_counts</code>	logical, default = TRUE. Use counts or proportion in y axis ?
<code>flip</code>	logical, default = FALSE. Flip the barplot ?
<code>x_lab</code>	character name for the x axis
<code>caption</code>	logical, default = TRUE. Display the plot caption summarising the number of cases ?

**Value**

a ggplot object

**Examples**

```
# Use fake data from Epidemiologist R handbook
suppressMessages(library(dplyr))

epivis::moissala_measles |>
  mutate(across(
    c(
      fever,
      rash,
      cough,
      red_eye,
      pneumonia,
      encephalitis
    ),
    ~ as.character(.x)
  )) |>
  plot_stacked_bar(
    cols = c("fever", "rash", "cough", "red_eye", "pneumonia", "encephalitis"),
    levels_value = c(0, 1),
    keep_na = FALSE,
    use_counts = FALSE,
    flip = TRUE
  )
```

# Index

## \* datasets

moissala\_measles, 3

base::strptime, 5

dodge\_x\_labs, 2

ggplot2::facet\_wrap, 5, 8, 9

ggplot2::guide\_axis, 2

ggplot2::scale\_x\_date, 5

label\_breaks, 2

moissala\_measles, 3

plot\_epicurve, 4

plot\_miss\_vis, 6

plot\_pyramid, 7

plot\_stacked\_bar, 9

scales::number, 2