

Package: santoku (via r-universe)

August 16, 2026

Type Package

Title A Versatile Cutting Tool

Version 2.0.0

Maintainer David Hugh-Jones <davidhughjones@gmail.com>

Description A tool for cutting data into intervals. Allows singleton intervals. Always includes the whole range of data by default. Flexible labelling. Convenience functions for cutting by quantiles etc. Handles dates, times, units and other vectors.

License MIT + file LICENSE

Encoding UTF-8

Language en-GB

Roxygen list(markdown = TRUE)

Suggests bench, bit64, covr, haven, Hmisc, hms, knitr, lubridate, purrr, rmarkdown, scales, stringi, testthat (>= 3.2.0), units, withr, xts, zoo

Config/testthat/edition 3

LinkingTo Rcpp

Depends R (>= 3.5.0)

Imports Rcpp, assertthat, glue, lifecycle, rlang, vctrs

URL <https://github.com/hughjonesd/santoku>,
<https://hughjonesd.github.io/santoku/>

BugReports <https://github.com/hughjonesd/santoku/issues>

VignetteBuilder knitr

RdMacros lifecycle

Config/roxygen2/version 8.0.0

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

Date/Publication 2026-08-16 03:45:25 UTC

RemoteUrl <https://github.com/hughjonesd/santoku>

RemoteRef v2.0.0

RemoteSha f2c468d566c429729435802e6a3fa3fe9f27976f

Contents

santoku-package	2
breaks-class	4
brk_default	4
brk_manual	5
brk_width-for-datetime	6
chop	6
chop_equally	10
chop_evenly	11
chop_fn	12
chop_mean_sd	14
chop_n	15
chop_pretty	16
chop_proportions	17
chop_quantiles	18
chop_spikes	20
chop_width	21
dissect	22
exactly	24
fillet	24
lbl_dash	25
lbl_date	27
lbl_discrete	28
lbl_endpoints	30
lbl_glue	31
lbl_intervals	33
lbl_midpoints	35
lbl_seq	36
non-standard-types	37
percent	37
Index	38

santoku-package

A versatile cutting tool for R: package overview and options

Description

santoku is a tool for cutting data into intervals. It provides the function `chop()`, which is similar to base R's `cut()` or `Hmisc::cut2()`. `chop(x, breaks)` takes a vector `x` and returns a factor of the same length, coding which interval each element of `x` falls into.

Details

Here are some advantages of santoku:

- By default, `chop()` always covers the whole range of the data, so you won't get unexpected NA values.
- Unlike `cut()` or `cut2()`, `chop()` can handle single values as well as intervals. For example, `chop(x, breaks = c(1, 2, 2, 3))` will create a separate factor level for values exactly equal to 2.
- Flexible and easy labelling.
- Convenience functions for creating quantile intervals, evenly-spaced intervals or equal-sized groups.
- Convenience functions to quickly tabulate chopped data.
- Can chop numbers, dates, date-times and other objects.

These advantages make santoku especially useful for exploratory analysis, where you may not know the range of your data in advance.

To get started, read the vignette:

```
vignette("santoku")
```

For more details, start with the documentation for `chop()`.

Options

Santoku has two options:

- `options("santoku.infinity")` sets the symbol for infinity in breaks. The default is NULL, in which case the infinity symbol is used on platforms that support it, otherwise "Inf" is used.
- `options("santoku.warn_character")` warns if you try to chop a character vector. Set to FALSE to turn off this warning.

Author(s)

Maintainer: David Hugh-Jones <davidhughjones@gmail.com>

Authors:

- David Hugh-Jones <davidhughjones@gmail.com>

Other contributors:

- Daniel Possenriede <possenriede@gmail.com> [contributor]

See Also

Useful links:

- <https://github.com/hughjonesd/santoku>
- <https://hughjonesd.github.io/santoku/>
- Report bugs at <https://github.com/hughjonesd/santoku/issues>

breaks-class	<i>Class representing a set of intervals</i>
--------------	--

Description

Class representing a set of intervals

Usage

```
## S3 method for class 'breaks'
format(x, ...)

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

is.breaks(x, ...)
```

Arguments

x	A breaks object
...	Unused

brk_default	<i>Create a standard set of breaks</i>
-------------	--

Description

Create a standard set of breaks

Usage

```
brk_default(breaks)
```

Arguments

breaks	A numeric vector.
--------	-------------------

Value

A function which returns an object of class breaks.

Examples

```
chop(1:10, c(2, 5, 8))
chop(1:10, brk_default(c(2, 5, 8)))
```

brk_manual	Create a breaks object manually
------------	---------------------------------

Description

Create a breaks object manually

Usage

```
brk_manual(breaks, left_vec)
```

Arguments

breaks	A vector, which must be sorted.
left_vec	A logical vector, the same length as breaks. Specifies whether each break is left-closed or right-closed.

Details

All breaks must be closed on exactly one side, like $\dots, x)$ $[x, \dots$ (left-closed) or $\dots, x)$ $[x, \dots$ (right-closed).

For example, if `breaks = 1:3` and `left = c(TRUE, FALSE, TRUE)`, then the resulting intervals are

```
T      F      T
[ 1,  2 ] ( 2, 3 )
```

Singleton breaks are created by repeating a number in `breaks`. Singletons must be closed on both sides, so if there is a repeated number at indices $i, i+1$, `left[i]` *must* be `TRUE` and `left[i+1]` must be `FALSE`.

`brk_manual()` ignores `left` and `close_end` arguments passed in from `chop()`, since `left_vec` sets these manually. `extend` and `drop` arguments are respected as usual.

Value

A function which returns an object of class `breaks`.

Examples

```
lbrks <- brk_manual(1:3, rep(TRUE, 3))
chop(1:3, lbrks, extend = FALSE)
```

```
rbrks <- brk_manual(1:3, rep(FALSE, 3))
chop(1:3, rbrks, extend = FALSE)
```

```
brks_singleton <- brk_manual(
  c(1,  2,  2,  3),
  c(TRUE, TRUE, FALSE, TRUE))
```

```
chop(1:3, brks_singleton, extend = FALSE)
```

```
brk_width-for-datetime
```

Equal-width intervals for dates or datetimes

Description

`brk_width()` can be used with time interval classes from base R or the `lubridate` package.

Usage

```
## S3 method for class 'Duration'
brk_width(width, start)
```

Arguments

<code>width</code>	A scalar difftime , Period or Duration object.
<code>start</code>	A scalar of class Date or POSIXct . Can be omitted.

Details

If `width` is a `Period`, `lubridate::add_with_rollback()` is used to calculate the widths. This can be useful for e.g. calendar months.

Examples

```
if (requireNamespace("lubridate")) {
  year2001 <- as.Date("2001-01-01") + 0:364
  tab_width(year2001, months(1),
    labels = lbl_discrete(" to ", fmt = "%e %b %y"))
}
```

```
chop
```

Cut data into intervals

Description

`chop()` cuts `x` into intervals. It returns a [factor](#) of the same length as `x`, representing which interval contains each element of `x`. `kiru()` is an alias for `chop`. `tab()` calls `chop()` and returns a contingency [table](#) from the result.

Usage

```
chop(  
  x,  
  breaks,  
  labels = lbl_intervals(),  
  extend = NULL,  
  left = TRUE,  
  close_end = TRUE,  
  raw = NULL,  
  drop = TRUE  
)  
  
kiru(  
  x,  
  breaks,  
  labels = lbl_intervals(),  
  extend = NULL,  
  left = TRUE,  
  close_end = TRUE,  
  raw = NULL,  
  drop = TRUE  
)  
  
tab(  
  x,  
  breaks,  
  labels = lbl_intervals(),  
  extend = NULL,  
  left = TRUE,  
  close_end = TRUE,  
  raw = NULL,  
  drop = TRUE  
)
```

Arguments

<code>x</code>	A vector.
<code>breaks</code>	A numeric vector of cut-points, or a function to create cut-points from <code>x</code> .
<code>labels</code>	A character vector of labels or a function to create labels.
<code>extend</code>	Logical. If <code>TRUE</code> , always extend breaks to $\pm\infty$. If <code>NULL</code> , extend breaks to <code>min(x)</code> and/or <code>max(x)</code> only if necessary. If <code>FALSE</code> , never extend.
<code>left</code>	Logical. Left-closed or right-closed breaks?
<code>close_end</code>	Logical. Close last break at right? (If <code>left</code> is <code>FALSE</code> , close first break at left?)
<code>raw</code>	Logical. Use raw values in labels?
<code>drop</code>	Logical. Drop unused levels from the result?

Details

`x` may be a numeric vector, or more generally, any vector which can be compared with `<` and `==` (see [Ops](#)). In particular [Date](#) and [date-time](#) objects are supported. Character vectors are supported with a warning.

Breaks:

`breaks` may be a vector or a function.

If it is a vector, `breaks` gives the interval endpoints. Repeating a value creates a "singleton" interval, which contains only that value. For example `breaks = c(1, 3, 3, 5)` creates 3 intervals: `[1, 3)`, `{3}` and `(3, 5]`.

If `breaks` is a function, it is called with the `x`, `extend`, `left` and `close_end` arguments, and should return an object of class `breaks`. Use `brk_*` functions to create a variety of data-dependent breaks. Names of breaks may be used for labels. See "Labels" below.

Options for breaks:

By default, left-closed intervals are created. If `left` is `FALSE`, right-closed intervals are created.

If `close_end` is `TRUE` the final break (or first break if `left` is `FALSE`) will be closed at both ends. This guarantees that all values `x` with `min(breaks) <= x <= max(breaks)` are included in the intervals.

Before version 0.9.0, `close_end` was `FALSE` by default, and also behaved differently with respect to extended breaks: see "Extending intervals" below.

Using [mathematical set notation](#):

- If `left` is `TRUE` and `close_end` is `TRUE`, breaks will look like `[b1, b2)`, `[b2, b3)` ... `[b_(n-1), b_n]`.
- If `left` is `FALSE` and `close_end` is `TRUE`, breaks will look like `[b1, b2]`, `(b2, b3]` ... `(b_(n-1), b_n]`.
- If `left` is `TRUE` and `close_end` is `FALSE`, all breaks will look like ... `[b1, b2)`
- If `left` is `FALSE` and `close_end` is `FALSE`, all breaks will look like ... `(b1, b2]`

Extending intervals:

If `extend` is `TRUE`, intervals will be extended to `[-Inf, min(breaks))` and `(max(breaks), Inf]`.

If `extend` is `NULL` (the default), intervals will be extended to `[min(x), min(breaks))` and `(max(breaks), max(x)]`, only if necessary, i.e. only if elements of `x` would be outside the unextended breaks.

If `extend` is `FALSE`, intervals are never extended.

Note that even when `extend = TRUE`, extended intervals will be dropped from the factor levels if they contain no elements and `drop = TRUE`.

`close_end` is only relevant if intervals are not extended; extended intervals are always closed on the outside. This is a change from previous behaviour. Up to version 0.8.0, `close_end` was applied to the last user-specified interval, before any extended intervals were created.

Since 1.1.0, infinity is represented as ∞ in breaks on Unicode platforms. Set `options(santoku.infinity = "Inf")` to get the old behaviour.

Labels:

`labels` may be a character vector. It should have the same length as the (possibly extended) number of intervals. Alternatively, `labels` may be a `lbl_*` function such as `lbl_dash()`.

If `breaks` is a named vector, then names of breaks will be used as labels for the interval starting at the corresponding element. This overrides the `labels` argument (but unnamed breaks will still use `labels`). This feature is **[Experimental]**.

If `labels` is `NULL`, then integer codes will be returned instead of a factor.

If `raw` is `TRUE`, labels will show the actual interval endpoints, usually numbers. If `raw` is `FALSE` then labels may show other objects, such as quantiles for `chop_quantiles()` and friends, proportions of the range for `chop_proportions()`, or standard deviations for `chop_mean_sd()`.

If `raw` is `NULL` then `lbl_*` functions will use their default (usually `FALSE`). Otherwise, the `raw` argument to `chop()` overrides `raw` arguments passed into `lbl_*` functions directly.

Miscellaneous:

`NA` values in `x`, and values which are outside the extended endpoints, return `NA`.

`kiru()` is a synonym for `chop()`. If you load `{tidyr}`, you can use it to avoid confusion with `tidyr::chop()`.

Note that `chop()`, like all of R, uses binary arithmetic. Thus, numbers may not be exactly equal to what you think they should be. There is an example below.

Value

`chop()` returns a `factor` of the same length as `x`, representing the intervals containing the value of `x`.

`tab()` returns a contingency `table`.

See Also

`base::cut()`, `non-standard-types` for chopping objects that aren't numbers.

Other chopping functions: `chop_equally()`, `chop_evenly()`, `chop_fn()`, `chop_mean_sd()`, `chop_n()`, `chop_proportions()`, `chop_quantiles()`, `chop_spikes()`, `chop_width()`, `fillet()`

Examples

```
chop(1:7, c(2, 4, 6))

chop(1:7, c(2, 4, 6), extend = FALSE)

# Repeat a number for a singleton break:
chop(1:7, c(2, 4, 4, 6))

chop(1:7, c(2, 4, 6), left = FALSE)

chop(1:7, c(2, 4, 6), close_end = FALSE)

chop(1:7, brk_quantiles(c(0.25, 0.75)))

# A single break is fine if `extend` is not `FALSE`:
chop(1:7, 4)

# Floating point inaccuracy:
chop(0.3/3, c(0, 0.1, 0.1, 1), labels = c("< 0.1", "0.1", "> 0.1"))
```

```
# -- Labels --

chop(1:7, c(Lowest = 1, Low = 2, Mid = 4, High = 6))

chop(1:7, c(2, 4, 6), labels = c("Lowest", "Low", "Mid", "High"))

chop(1:7, c(2, 4, 6), labels = lbl_dash())

# Mixing names and other labels:
chop(1:7, c("<2" = 1, 2, 4, ">=6" = 6), labels = lbl_dash())

# -- Non-standard types --

chop(as.Date("2001-01-01") + 1:7, as.Date("2001-01-04"))

suppressWarnings(chop(LETTERS[1:7], "D"))

tab(1:10, c(2, 5, 8))
```

chop_equally

Chop equal-sized groups

Description

chop_equally() chops x into groups with an equal number of elements.

Usage

```
chop_equally(
  x,
  groups,
  ...,
  labels = lbl_intervals(),
  left = is.numeric(x),
  raw = TRUE
)

brk_equally(groups)

tab_equally(x, groups, ..., left = is.numeric(x), raw = TRUE)
```

Arguments

x	A vector.
groups	Number of groups.

...	Passed to chop() .
labels	A character vector of labels or a function to create labels.
left	Logical. Left-closed or right-closed breaks?
raw	Logical. Use raw values in labels?

Details

`chop_equally()` uses [brk_quantiles\(\)](#) under the hood. If `x` has duplicate elements, you may get fewer groups than requested. If so, a warning will be emitted. See the examples.

Value

`chop_*` functions return a [factor](#) of the same length as `x`.
`brk_*` functions return a [function](#) to create breaks.
`tab_*` functions return a contingency [table](#).

See Also

Other chopping functions: [chop\(\)](#), [chop_evenly\(\)](#), [chop_fn\(\)](#), [chop_mean_sd\(\)](#), [chop_n\(\)](#), [chop_proportions\(\)](#), [chop_quantiles\(\)](#), [chop_spikes\(\)](#), [chop_width\(\)](#), [fillet\(\)](#)

Examples

```
chop_equally(1:10, 5)

# You can't always guarantee equal-sized groups:
dupes <- c(1, 1, 1, 2, 3, 4, 4, 4)
quantile(dupes, 0:4/4)
chop_equally(dupes, 4)
# Or as many groups as you ask for:
chop_equally(c(1, 1, 2, 2), 3)
```

chop_evenly

Chop into equal-width intervals

Description

`chop_evenly()` chops `x` into intervals intervals of equal width.

Usage

```
chop_evenly(x, intervals, ...)
```

```
brk_evenly(intervals)
```

```
tab_evenly(x, intervals, ...)
```

Arguments

`x` A vector.
`intervals` Integer: number of intervals to create.
`...` Passed to `chop()`.

Value

`chop_*` functions return a [factor](#) of the same length as `x`.
`brk_*` functions return a [function](#) to create breaks.
`tab_*` functions return a contingency [table](#).

See Also

Other chopping functions: [chop\(\)](#), [chop_equally\(\)](#), [chop_fn\(\)](#), [chop_mean_sd\(\)](#), [chop_n\(\)](#), [chop_proportions\(\)](#), [chop_quantiles\(\)](#), [chop_spikes\(\)](#), [chop_width\(\)](#), [fillet\(\)](#)

Examples

```
chop_evenly(0:10, 5)
```

 chop_fn

Chop using an existing function

Description

`chop_fn()` is a convenience wrapper: `chop_fn(x, foo, ...)` is the same as `chop(x, foo(x, ...))`.

Usage

```
chop_fn(
  x,
  fn,
  ...,
  extend = NULL,
  left = TRUE,
  close_end = TRUE,
  raw = NULL,
  drop = TRUE
)
```

```
brk_fn(fn, ...)
```

```
tab_fn(
  x,
  fn,
```

```

    ...,
    extend = NULL,
    left = TRUE,
    close_end = TRUE,
    raw = NULL,
    drop = TRUE
  )

```

Arguments

<code>x</code>	A vector.
<code>fn</code>	A function which returns a numeric vector of breaks.
<code>...</code>	Further arguments to <code>fn</code>
<code>extend</code>	Logical. If TRUE, always extend breaks to +/-Inf. If NULL, extend breaks to <code>min(x)</code> and/or <code>max(x)</code> only if necessary. If FALSE, never extend.
<code>left</code>	Logical. Left-closed or right-closed breaks?
<code>close_end</code>	Logical. Close last break at right? (If <code>left</code> is FALSE, close first break at left?)
<code>raw</code>	Logical. Use raw values in labels?
<code>drop</code>	Logical. Drop unused levels from the result?

Value

`chop_*` functions return a [factor](#) of the same length as `x`.
`brk_*` functions return a [function](#) to create breaks.
`tab_*` functions return a contingency [table](#).

See Also

Other chopping functions: [chop\(\)](#), [chop_equally\(\)](#), [chop_evenly\(\)](#), [chop_mean_sd\(\)](#), [chop_n\(\)](#), [chop_proportions\(\)](#), [chop_quantiles\(\)](#), [chop_spikes\(\)](#), [chop_width\(\)](#), [fillet\(\)](#)

Examples

```

if (requireNamespace("scales")) {
  chop_fn(rlnorm(10), scales::breaks_log(5))
  # same as
  # x <- rlnorm(10)
  # chop(x, scales::breaks_log(5)(x))
}

```

chop_mean_sd	<i>Chop by standard deviations</i>
--------------	------------------------------------

Description

Intervals are measured in standard deviations on either side of the mean.

Usage

```
chop_mean_sd(x, sds = 1:3, ..., raw = FALSE, weights = NULL, sd = deprecated())
```

```
brk_mean_sd(sds = 1:3, weights = NULL, sd = deprecated())
```

```
tab_mean_sd(x, sds = 1:3, ..., raw = FALSE)
```

Arguments

x	A vector.
sds	Positive numeric vector of standard deviations.
...	Passed to chop() .
raw	Logical. Use raw values in labels?
weights	Optional weights for the mean/sd calculation. Weights are normalized to sum to <code>length(x)</code> .
sd	[Deprecated]

Details

In version 0.7.0, these functions changed to specifying `sds` as a vector. To chop 1, 2 and 3 standard deviations around the mean, write `chop_mean_sd(x, sds = 1:3)` instead of `chop_mean_sd(x, sd = 3)`.

Value

`chop_*` functions return a [factor](#) of the same length as `x`.
`brk_*` functions return a [function](#) to create breaks.
`tab_*` functions return a contingency [table](#).

See Also

Other chopping functions: [chop\(\)](#), [chop_equally\(\)](#), [chop_evenly\(\)](#), [chop_fn\(\)](#), [chop_n\(\)](#), [chop_proportions\(\)](#), [chop_quantiles\(\)](#), [chop_spikes\(\)](#), [chop_width\(\)](#), [fillet\(\)](#)

Examples

```
chop_mean_sd(1:10)

chop_mean_sd(1:4, weights = 1:4)

chop(1:10, brk_mean_sd())

tab_mean_sd(1:10)
```

chop_n	<i>Chop into fixed-sized groups</i>
--------	-------------------------------------

Description

chop_n() creates intervals containing a fixed number of elements.

Usage

```
chop_n(x, n, ..., tail = "split")

brk_n(n, tail = "split")

tab_n(x, n, ..., tail = "split")
```

Arguments

x	A vector.
n	Integer. Number of elements in each interval.
...	Passed to chop() .
tail	String. What to do if the final interval has fewer than n elements? "split" to keep it separate. "merge" to merge it with the neighbouring interval.

Details

The algorithm guarantees that intervals contain no more than n elements, so long as there are no duplicates in x and tail = "split". It also guarantees that intervals contain no fewer than n elements, except possibly the last interval (or first interval if left is FALSE).

To ensure that all intervals contain at least n elements (so long as there are at least n elements in x!) set tail = "merge".

If tail = "split" and there are intervals containing duplicates with more than n elements, a warning is given.

Value

chop_* functions return a [factor](#) of the same length as x.
 brk_* functions return a [function](#) to create breaks.
 tab_* functions return a contingency [table](#).

See Also

Other chopping functions: [chop\(\)](#), [chop_equally\(\)](#), [chop_evenly\(\)](#), [chop_fn\(\)](#), [chop_mean_sd\(\)](#), [chop_proportions\(\)](#), [chop_quantiles\(\)](#), [chop_spikes\(\)](#), [chop_width\(\)](#), [fillet\(\)](#)

Examples

```
chop_n(1:10, 5)

chop_n(1:5, 2)
chop_n(1:5, 2, tail = "merge")

# too many duplicates
x <- rep(1:2, each = 3)
chop_n(x, 2)

tab_n(1:10, 5)

# fewer elements in one group
tab_n(1:10, 4)
```

chop_pretty

Chop using pretty breakpoints

Description

`chop_pretty()` uses [base::pretty\(\)](#) to calculate breakpoints which are 1, 2 or 5 times a power of 10. These look nice in graphs.

Usage

```
chop_pretty(x, n = 5, ...)

brk_pretty(n = 5, ...)

tab_pretty(x, n = 5, ...)
```

Arguments

<code>x</code>	A vector.
<code>n</code>	Positive integer passed to base::pretty() . How many intervals to chop into?
<code>...</code>	Passed to chop() by <code>chop_pretty()</code> and <code>tab_pretty()</code> ; passed to base::pretty() by <code>brk_pretty()</code> .

Details

[base::pretty\(\)](#) tries to return $n+1$ breakpoints, i.e. n intervals, but note that this is not guaranteed. There are methods for Date and POSIXct objects.

For fine-grained control over [base::pretty\(\)](#) parameters, use `chop(x, brk_pretty(...))`.

Value

chop_* functions return a [factor](#) of the same length as x.
 brk_* functions return a [function](#) to create breaks.
 tab_* functions return a contingency [table](#).

Examples

```
chop_pretty(1:10)

chop(1:10, brk_pretty(n = 5, high.u.bias = 0))

tab_pretty(1:10)
```

chop_proportions	<i>Chop into proportions of the range of x</i>
------------------	--

Description

chop_proportions() chops x into proportions of its range, excluding infinite values.

Usage

```
chop_proportions(x, proportions, ..., raw = TRUE)

brk_proportions(proportions)

tab_proportions(x, proportions, ..., raw = TRUE)
```

Arguments

x	A vector.
proportions	Numeric vector between 0 and 1: proportions of x's range. If proportions has names, these will be used for labels.
...	Passed to chop() .
raw	Logical. Use raw values in labels?

Details

By default, labels show the raw numeric endpoints. To label intervals by the proportions, use raw = FALSE.

Value

chop_* functions return a [factor](#) of the same length as x.
 brk_* functions return a [function](#) to create breaks.
 tab_* functions return a contingency [table](#).

See Also

Other chopping functions: [chop\(\)](#), [chop_equally\(\)](#), [chop_evenly\(\)](#), [chop_fn\(\)](#), [chop_mean_sd\(\)](#), [chop_n\(\)](#), [chop_quantiles\(\)](#), [chop_spikes\(\)](#), [chop_width\(\)](#), [fillet\(\)](#)

Examples

```
chop_proportions(0:10, c(0.2, 0.8))
chop_proportions(0:10, c(Low = 0, Mid = 0.2, High = 0.8))
```

chop_quantiles	<i>Chop by quantiles</i>
----------------	--------------------------

Description

`chop_quantiles()` chops data by quantiles. `chop_deciles()` is a convenience function which chops into deciles.

Usage

```
chop_quantiles(
  x,
  probs,
  ...,
  labels = if (raw) lbl_intervals() else lbl_intervals(single = NULL),
  left = is.numeric(x),
  raw = FALSE,
  weights = NULL,
  recalc_probs = FALSE
)

chop_deciles(x, ...)

brk_quantiles(probs, ..., weights = NULL, recalc_probs = FALSE)

tab_quantiles(x, probs, ..., left = is.numeric(x), raw = FALSE)

tab_deciles(x, ...)
```

Arguments

<code>x</code>	A vector.
<code>probs</code>	A vector of probabilities for the quantiles. If <code>probs</code> has names, these will be used for labels.
<code>...</code>	For <code>chop_quantiles</code> , passed to chop() . For <code>brk_quantiles()</code> , passed to stats::quantile() or Hmisc::wtd.quantile() .

labels	A character vector of labels or a function to create labels.
left	Logical. Left-closed or right-closed breaks?
raw	Logical. Use raw values in labels?
weights	Optional weights for the quantile calculation. Weights are normalized to sum to <code>length(x)</code> .
recalc_probs	Logical. Recalculate probabilities of quantiles using <code>ecdf(x)</code> ? See below.

Details

For non-numeric `x`, `left` is set to `FALSE` by default. This works better for calculating "type 1" quantiles, since they round down. See [stats::quantile\(\)](#).

By default, `chop_quantiles()` shows the requested probabilities in the labels. To show the numeric quantiles themselves, set `raw = TRUE`.

When `x` contains duplicates, consecutive quantiles may be the same number. If so, interval labels may be misleading, and if `recalc_probs = FALSE` a warning is emitted. Set `recalc_probs = TRUE` to recalculate the probabilities of the quantiles using the [empirical cumulative distribution function](#) of `x`. Doing so may give you different labels from what you expect, and will remove any names from probs, but it never changes the actual quantiles used for breaks. At present, `recalc_probs = TRUE` is incompatible with non-null weights. See the example below.

Value

`chop_*` functions return a [factor](#) of the same length as `x`.

`brk_*` functions return a [function](#) to create breaks.

`tab_*` functions return a contingency [table](#).

See Also

Other chopping functions: [chop\(\)](#), [chop_equally\(\)](#), [chop_evenly\(\)](#), [chop_fn\(\)](#), [chop_mean_sd\(\)](#), [chop_n\(\)](#), [chop_proportions\(\)](#), [chop_spikes\(\)](#), [chop_width\(\)](#), [fillet\(\)](#)

Examples

```
chop_quantiles(1:10, 1:3/4)

chop_quantiles(1:10, c(Q1 = 0, Q2 = 0.25, Q3 = 0.5, Q4 = 0.75))

chop(1:10, brk_quantiles(1:3/4))

chop_deciles(1:10)

# to label by the quantiles themselves:
chop_quantiles(1:10, 1:3/4, raw = TRUE)

# duplicate quantiles:
x <- c(1, 1, 1, 2, 3)
quantile(x, 1:5/5)
tab_quantiles(x, 1:5/5)
tab_quantiles(x, 1:5/5, recalc_probs = TRUE)
```

```
set.seed(42)
tab_quantiles(rnorm(100), probs = 1:3/4, raw = TRUE)
```

chop_spikes

Chop common values into singleton intervals

Description

chop_spikes() lets you chop common values of x into their own singleton intervals. This can help make unusual values visible.

Usage

```
chop_spikes(x, breaks, ..., n = NULL, prop = NULL)
```

```
brk_spikes(breaks, n = NULL, prop = NULL)
```

```
tab_spikes(x, breaks, ..., n = NULL, prop = NULL)
```

Arguments

x	A vector.
breaks	A numeric vector of cut-points or a call to a brk_* function. The resulting breaks object will be modified to add singleton breaks.
...	Passed to chop() .
n, prop	Scalar. Provide either n, a number of values, or prop, a proportion of length(x). Values of x which occur at least this often will get their own singleton break.

Details

This function is **[Experimental]**.

Value

chop_* functions return a [factor](#) of the same length as x.
brk_* functions return a [function](#) to create breaks.
tab_* functions return a contingency [table](#).

See Also

[dissect\(\)](#) for a different approach.

Other chopping functions: [chop\(\)](#), [chop_equally\(\)](#), [chop_evenly\(\)](#), [chop_fn\(\)](#), [chop_mean_sd\(\)](#), [chop_n\(\)](#), [chop_proportions\(\)](#), [chop_quantiles\(\)](#), [chop_width\(\)](#), [fillet\(\)](#)

Examples

```
x <- c(1:4, rep(5, 5), 6:10)
chop_spikes(x, c(2, 7), n = 5)
chop_spikes(x, c(2, 7), prop = 0.25)
chop_spikes(x, brk_width(5), n = 5)

set.seed(42)
x <- runif(40, 0, 10)
x <- sample(x, 200, replace = TRUE)
tab_spikes(x, brk_width(2, 0), prop = 0.05)
```

chop_width	<i>Chop into fixed-width intervals</i>
------------	--

Description

chop_width() chops x into intervals of fixed width.

Usage

```
chop_width(x, width, start, ..., left = sign(width) > 0)

brk_width(width, start)

## Default S3 method:
brk_width(width, start)

tab_width(x, width, start, ..., left = sign(width) > 0)
```

Arguments

x	A vector.
width	Width of intervals.
start	Starting point for intervals. By default the smallest finite x (largest if width is negative).
...	Passed to chop() .
left	Logical. Left-closed or right-closed breaks?

Details

If width is negative, chop_width() sets left = FALSE and intervals will go downwards from start.

Value

chop_* functions return a [factor](#) of the same length as x.
 brk_* functions return a [function](#) to create breaks.
 tab_* functions return a contingency [table](#).

See Also

[brk_width-for-datetime](#)

Other chopping functions: [chop\(\)](#), [chop_equally\(\)](#), [chop_evenly\(\)](#), [chop_fn\(\)](#), [chop_mean_sd\(\)](#), [chop_n\(\)](#), [chop_proportions\(\)](#), [chop_quantiles\(\)](#), [chop_spikes\(\)](#), [fillet\(\)](#)

Examples

```
chop_width(1:10, 2)

chop_width(1:10, 2, start = 0)

chop_width(1:9, -2)

chop(1:10, brk_width(2, 0))

tab_width(1:10, 2, start = 0)
```

dissect

Cut data into intervals, separating out common values

Description

Sometimes it's useful to separate out common elements of `x`. `dissect()` chops `x`, but puts common elements of `x` ("spikes") into separate categories.

Usage

```
dissect(
  x,
  breaks,
  ...,
  n = NULL,
  prop = NULL,
  spike_labels = "{{{1}}}",
  exclude_spikes = FALSE
)

tab_dissect(x, breaks, ..., n = NULL, prop = NULL)
```

Arguments

<code>x, breaks, ...</code>	Passed to chop() .
<code>n, prop</code>	Scalar. Provide either <code>n</code> , a number of values, or <code>prop</code> , a proportion of <code>length(x)</code> . Values of <code>x</code> which occur at least this often will get their own singleton break.
<code>spike_labels</code>	Blue string for spike labels. Use "{1}" for the spike value.
<code>exclude_spikes</code>	Logical. Exclude spikes before chopping <code>x</code> ? This can affect the location of data-dependent breaks.

Details

Unlike `chop_spikes()`, `dissect()` doesn't break up intervals which contain a spike. As a result, unlike `chop_*` functions, `dissect()` does not chop `x` into disjoint intervals. See the examples.

If breaks are data-dependent, their labels may be misleading after common elements have been removed. See the example below. To get round this, set `exclude_spikes` to `TRUE`. Then breaks will be calculated after removing spikes from the data.

Levels of the result are ordered by the minimum element in each level. As a result, if `drop = FALSE`, empty levels will be placed last.

This function is **[Experimental]**.

Value

`dissect()` returns the result of `chop()`, but with common values put into separate factor levels.

`tab_dissect()` returns a contingency `table()`.

See Also

`chop_spikes()` for a different approach.

Examples

```
x <- c(2, 3, 3, 3, 4)
dissect(x, c(2, 4), n = 3)
dissect(x, brk_width(2), prop = 0.5)

set.seed(42)
x <- runif(40, 0, 10)
x <- sample(x, 200, replace = TRUE)
# Compare:
table(dissect(x, brk_width(2, 0), prop = 0.05))
# Versus:
tab_spikes(x, brk_width(2, 0), prop = 0.05)

# Potentially confusing data-dependent breaks:
set.seed(42)
x <- rnorm(99)
x[1:9] <- x[1]
tab_quantiles(x, 1:2/3)
tab_dissect(x, brk_quantiles(1:2/3), n = 9)
# Calculate quantiles excluding spikes:
tab_dissect(x, brk_quantiles(1:2/3), n = 9, exclude_spikes = TRUE)
```

 exactly

Define singleton intervals explicitly

Description

`exactly()` duplicates its input. It lets you define singleton intervals like this: `chop(x, c(1, exactly(2), 3))`. This is the same as `chop(x, c(1, 2, 2, 3))` but conveys your intent more clearly.

Usage

```
exactly(x)
```

Arguments

`x` A numeric vector.

Value

The same as `rep(x, each = 2)`.

Examples

```
chop(1:10, c(2, exactly(5), 8))

# same:
chop(1:10, c(2, 5, 5, 8))
```

 fillet

Chop data precisely (for programmers)

Description

`fillet()` calls `chop()` with `extend = FALSE` and `drop = FALSE`. This ensures that you get only the breaks and labels you ask for. When programming, consider using `fillet()` instead of `chop()`.

Usage

```
fillet(
  x,
  breaks,
  labels = lbl_intervals(),
  left = TRUE,
  close_end = TRUE,
  raw = NULL
)
```

Arguments

<code>x</code>	A vector.
<code>breaks</code>	A numeric vector of cut-points, or a function to create cut-points from <code>x</code> .
<code>labels</code>	A character vector of labels or a function to create labels.
<code>left</code>	Logical. Left-closed or right-closed breaks?
<code>close_end</code>	Logical. Close last break at right? (If <code>left</code> is FALSE, close first break at left?)
<code>raw</code>	Logical. Use raw values in labels?

Value

`fillet()` returns a [factor](#) of the same length as `x`, representing the intervals containing the value of `x`.

See Also

Other chopping functions: [chop\(\)](#), [chop_equally\(\)](#), [chop_evenly\(\)](#), [chop_fn\(\)](#), [chop_mean_sd\(\)](#), [chop_n\(\)](#), [chop_proportions\(\)](#), [chop_quantiles\(\)](#), [chop_spikes\(\)](#), [chop_width\(\)](#)

Examples

```
fillet(1:10, c(2, 5, 8))
```

lbl_dash	<i>Label chopped intervals like 1-4, 4-5, ...</i>
----------	---

Description

This label style is user-friendly, but doesn't distinguish between left- and right-closed intervals. It's good for continuous data where you don't expect points to be exactly on the breaks.

Usage

```
lbl_dash(
  symbol = em_dash(),
  fmt = NULL,
  single = "{l}",
  first = NULL,
  last = NULL,
  raw = deprecated()
)
```

Arguments

symbol	String: symbol to use for the dash.
fmt	String, list or function. A format for break endpoints.
single	Glue string: label for singleton intervals. See lbl_glue() for details. If NULL, singleton intervals will be labelled the same way as other intervals.
first	Glue string: override label for the first category. Write e.g. <code>first = "<{r}"</code> to create a label like <code>"<18"</code> . See lbl_glue() for details.
last	String: override label for the last category. Write e.g. <code>last = ">{l}"</code> to create a label like <code>">65"</code> . See lbl_glue() for details.
raw	[Deprecated] . Throws an error. Use the raw argument to chop() instead.

Details

If you don't want Unicode output, use `lbl_dash("-")`.

Value

A function that creates a vector of labels.

Formatting endpoints

If `fmt` is not NULL then it is used to format the endpoints.

- If `fmt` is a string, then numeric endpoints will be formatted by `sprintf(fmt, breaks)`; other endpoints, e.g. [Date](#) objects, will be formatted by `format(breaks, fmt)`.
- If `fmt` is a list, then it will be used as arguments to [format](#).
- If `fmt` is a function, it should take a vector of numbers (or other objects that can be used as breaks) and return a character vector. It may be helpful to use functions from the `{scales}` package, e.g. [scales::label_comma\(\)](#).

See Also

Other labelling functions: [lbl_date\(\)](#), [lbl_discrete\(\)](#), [lbl_endpoints\(\)](#), [lbl_glue\(\)](#), [lbl_intervals\(\)](#), [lbl_manual\(\)](#), [lbl_midpoints\(\)](#), [lbl_seq\(\)](#)

Examples

```
chop(1:10, c(2, 5, 8), lbl_dash())

chop(1:10, c(2, 5, 8), lbl_dash(" to ", fmt = "%.1f"))

chop(1:10, c(2, 5, 8), lbl_dash(first = "<{r}"))

pretty <- function(x) prettyNum(x, big.mark = ",", digits = 1)
chop(runif(10) * 10000, c(3000, 7000), lbl_dash(" to ", fmt = pretty))
```

lbl_date	<i>Label dates and datetimes</i>
----------	----------------------------------

Description

[Experimental]

lbl_date() and lbl_datetime() produce nice labels for dates and datetimes. Where possible ranges are simplified, like like "13-14 Jul 2026" or "11:15-12:15 1 Dec 2025".

Usage

```
lbl_date(
  fmt = "%e %b %Y",
  symbol = "-",
  unit = as.difftime(1, units = "days"),
  single = "{l}",
  first = NULL,
  last = NULL
)

lbl_datetime(
  fmt = "%H:%M:%S %b %e %Y",
  symbol = "-",
  unit = NULL,
  single = "{l}",
  first = NULL,
  last = NULL
)
```

Arguments

fmt	String, list or function. A format for break endpoints.
symbol	String: separator to use for full ranges.
unit	Optional interval unit for non-overlapping labels. If not NULL, . endpoints are adjusted in the style of lbl_discrete() .
single	Glue string: label for singleton intervals. See lbl_glue() for details. If NULL, singleton intervals will be labelled the same way as other intervals.
first	Glue string: override label for the first category. Write e.g. first = "<{r}" to create a label like "<18". See lbl_glue() for details.
last	String: override label for the last category. Write e.g. last = ">{l}" to create a label like ">65". See lbl_glue() for details.

Value

A function that creates a vector of labels.

Formatting endpoints

If `fmt` is not `NULL` then it is used to format the endpoints.

- If `fmt` is a string, then numeric endpoints will be formatted by `sprintf(fmt, breaks)`; other endpoints, e.g. [Date](#) objects, will be formatted by `format(breaks, fmt)`.
- If `fmt` is a list, then it will be used as arguments to [format](#).
- If `fmt` is a function, it should take a vector of numbers (or other objects that can be used as breaks) and return a character vector. It may be helpful to use functions from the `{scales}` package, e.g. [scales::label_comma\(\)](#).

See Also

Other labelling functions: [lbl_dash\(\)](#), [lbl_discrete\(\)](#), [lbl_endpoints\(\)](#), [lbl_glue\(\)](#), [lbl_intervals\(\)](#), [lbl_manual\(\)](#), [lbl_midpoints\(\)](#), [lbl_seq\(\)](#)

Examples

```
winter <- as.Date("2025-12-01") + 0:89
tab(winter, as.Date(c("2025-12-25", "2026-01-06")),
    labels = lbl_date())
new_year <- as.POSIXct("2025-12-31 23:00") + 0:120 * 60
round_midnight <- as.POSIXct(c("2025-12-31 23:59", "2026-01-01 00:05"))
tab(new_year, round_midnight,
    labels = lbl_datetime())
tab(new_year, round_midnight,
    labels = lbl_datetime(unit = as.difftime(1, units = "mins")))
```

lbl_discrete

Label discrete data

Description

`lbl_discrete()` creates labels for discrete data, such as integers. For example, breaks `c(1, 3, 4, 6, 7)` are labelled: "1-2", "3", "4-5", "6-7".

Usage

```
lbl_discrete(
  symbol = em_dash(),
  unit = 1L,
  fmt = NULL,
  single = NULL,
  first = NULL,
  last = NULL
)
```

Arguments

symbol	String: symbol to use for the dash.
unit	Minimum difference between distinct values of data. For integers, 1.
fmt	String, list or function. A format for break endpoints.
single	Glue string: label for singleton intervals. See lbl_glue() for details. If NULL, singleton intervals will be labelled the same way as other intervals.
first	Glue string: override label for the first category. Write e.g. first = "<{r}" to create a label like "<18". See lbl_glue() for details.
last	String: override label for the last category. Write e.g. last = ">{1}" to create a label like ">65". See lbl_glue() for details.

Details

No check is done that the data are discrete-valued. If they are not, then these labels may be misleading. Here, discrete-valued means that if $x < y$, then $x \leq y - \text{unit}$.

Be aware that Date objects may have non-integer values. See [Date](#).

Value

A function that creates a vector of labels.

Formatting endpoints

If `fmt` is not NULL then it is used to format the endpoints.

- If `fmt` is a string, then numeric endpoints will be formatted by `sprintf(fmt, breaks)`; other endpoints, e.g. [Date](#) objects, will be formatted by `format(breaks, fmt)`.
- If `fmt` is a list, then it will be used as arguments to [format](#).
- If `fmt` is a function, it should take a vector of numbers (or other objects that can be used as breaks) and return a character vector. It may be helpful to use functions from the `{scales}` package, e.g. [scales::label_comma\(\)](#).

See Also

Other labelling functions: [lbl_dash\(\)](#), [lbl_date\(\)](#), [lbl_endpoints\(\)](#), [lbl_glue\(\)](#), [lbl_intervals\(\)](#), [lbl_manual\(\)](#), [lbl_midpoints\(\)](#), [lbl_seq\(\)](#)

Examples

```
tab(1:7, c(1, 3, 5), lbl_discrete())

tab(1:7, c(3, 5), lbl_discrete(first = "<= {r}"))

tab(1:7 * 1000, c(1, 3, 5) * 1000, lbl_discrete(unit = 1000))

# Misleading labels for non-integer data
chop(2.5, c(1, 3, 5), lbl_discrete())
```

lbl_endpoints	<i>Label chopped intervals by their left or right endpoints</i>
---------------	---

Description

This is useful when the left endpoint unambiguously indicates the interval. In other cases it may give errors due to duplicate labels.

Usage

```
lbl_endpoints(
  left = TRUE,
  fmt = NULL,
  single = NULL,
  first = NULL,
  last = NULL,
  raw = deprecated()
)
```

```
lbl_endpoint(fmt = NULL, raw = FALSE, left = TRUE)
```

Arguments

left	Flag. Use left endpoint or right endpoint?
fmt	String, list or function. A format for break endpoints.
single	Glue string: label for singleton intervals. See lbl_glue() for details. If NULL, singleton intervals will be labelled the same way as other intervals.
first	Glue string: override label for the first category. Write e.g. first = "<{r}" to create a label like "<18". See lbl_glue() for details.
last	String: override label for the last category. Write e.g. last = ">{1}" to create a label like ">65". See lbl_glue() for details.
raw	[Deprecated] . Throws an error. Use the raw argument to chop() instead.

Details

lbl_endpoint() is **[Defunct]** and gives an error since santoku 1.0.0.

Value

A function that creates a vector of labels.

Formatting endpoints

If `fmt` is not `NULL` then it is used to format the endpoints.

- If `fmt` is a string, then numeric endpoints will be formatted by `sprintf(fmt, breaks)`; other endpoints, e.g. [Date](#) objects, will be formatted by `format(breaks, fmt)`.
- If `fmt` is a list, then it will be used as arguments to [format](#).
- If `fmt` is a function, it should take a vector of numbers (or other objects that can be used as breaks) and return a character vector. It may be helpful to use functions from the `{scales}` package, e.g. `scales::label_comma()`.

See Also

Other labelling functions: [lbl_dash\(\)](#), [lbl_date\(\)](#), [lbl_discrete\(\)](#), [lbl_glue\(\)](#), [lbl_intervals\(\)](#), [lbl_manual\(\)](#), [lbl_midpoints\(\)](#), [lbl_seq\(\)](#)

Examples

```
chop(1:10, c(2, 5, 8), lbl_endpoints(left = TRUE))
chop(1:10, c(2, 5, 8), lbl_endpoints(left = FALSE))
if (requireNamespace("lubridate")) {
  tab_width(
    as.Date("2000-01-01") + 0:365,
    months(1),
    labels = lbl_endpoints(fmt = "%b")
  )
}

## Not run:
# This gives breaks `[1, 2) [2, 3) {3}` which lead to
# duplicate labels `"2", "3", "3"`:
chop(1:3, 1:3, lbl_endpoints(left = FALSE))

## End(Not run)
```

lbl_glue

Label chopped intervals using the glue package

Description

Use `"{l}"` and `"{r}"` to show the left and right endpoints of the intervals.

Usage

```
lbl_glue(
  label,
  fmt = NULL,
  single = NULL,
  first = NULL,
```

```

    last = NULL,
    raw = deprecated(),
    ...
)

```

Arguments

label	A glue string passed to <code>glue::glue()</code> .
fmt	String, list or function. A format for break endpoints.
single	Glue string: label for singleton intervals. See <code>lbl_glue()</code> for details. If NULL, singleton intervals will be labelled the same way as other intervals.
first	Glue string: override label for the first category. Write e.g. <code>first = "<{r}"</code> to create a label like <code>"<18"</code> . See <code>lbl_glue()</code> for details.
last	String: override label for the last category. Write e.g. <code>last = ">{1}"</code> to create a label like <code>">65"</code> . See <code>lbl_glue()</code> for details.
raw	[Deprecated] . Throws an error. Use the raw argument to <code>chop()</code> instead.
...	Further arguments passed to <code>glue::glue()</code> .

Details

The following variables are available in the glue string:

- `l` is a character vector of left endpoints of intervals.
- `r` is a character vector of right endpoints of intervals.
- `l_closed` is a logical vector. Elements are TRUE when the left endpoint is closed.
- `r_closed` is a logical vector, TRUE when the right endpoint is closed.

Endpoints will be formatted by `fmt` before being passed to `glue()`.

Value

A function that creates a vector of labels.

Formatting endpoints

If `fmt` is not NULL then it is used to format the endpoints.

- If `fmt` is a string, then numeric endpoints will be formatted by `sprintf(fmt, breaks)`; other endpoints, e.g. `Date` objects, will be formatted by `format(breaks, fmt)`.
- If `fmt` is a list, then it will be used as arguments to `format`.
- If `fmt` is a function, it should take a vector of numbers (or other objects that can be used as breaks) and return a character vector. It may be helpful to use functions from the `{scales}` package, e.g. `scales::label_comma()`.

See Also

Other labelling functions: `lbl_dash()`, `lbl_date()`, `lbl_discrete()`, `lbl_endpoints()`, `lbl_intervals()`, `lbl_manual()`, `lbl_midpoints()`, `lbl_seq()`

Examples

```

tab(1:10, c(1, 3, 3, 7),
    labels = lbl_glue("{l} to {r}", single = "Exactly {l}"))

tab(1:10 * 1000, c(1, 3, 5, 7) * 1000,
    labels = lbl_glue("{l}-{r}",
                      fmt = function(x) prettyNum(x, big.mark=','))

# reproducing lbl_intervals():
interval_left <- "{ifelse(l_closed, '[', '(')}"
interval_right <- "{ifelse(r_closed, ']', ')')}"
glue_string <- paste0(interval_left, "{l}", ", ", "{r}", interval_right)
tab(1:10, c(1, 3, 3, 7), labels = lbl_glue(glue_string, single = "{{{l}}}")

```

lbl_intervals	<i>Label chopped intervals using set notation</i>
---------------	---

Description

These labels are the most exact, since they show you whether intervals are "closed" or "open", i.e. whether they include their endpoints.

Usage

```

lbl_intervals(
  fmt = NULL,
  single = "{{{l}}}",
  first = NULL,
  last = NULL,
  raw = deprecated()
)

```

Arguments

fmt	String, list or function. A format for break endpoints.
single	Glue string: label for singleton intervals. See lbl_glue() for details. If NULL, singleton intervals will be labelled the same way as other intervals.
first	Glue string: override label for the first category. Write e.g. first = "<{r}" to create a label like "<18". See lbl_glue() for details.
last	String: override label for the last category. Write e.g. last = ">{l}" to create a label like ">65". See lbl_glue() for details.
raw	[Deprecated] . Throws an error. Use the raw argument to chop() instead.

Details

Mathematical set notation looks like this:

- $[a, b]$: all numbers x where $a \leq x \leq b$;
- (a, b) : all numbers where $a < x < b$;
- $[a, b)$: all numbers where $a \leq x < b$;
- $(a, b]$: all numbers where $a < x \leq b$;
- $\{a\}$: just the number a exactly.

Value

A function that creates a vector of labels.

Formatting endpoints

If `fmt` is not `NULL` then it is used to format the endpoints.

- If `fmt` is a string, then numeric endpoints will be formatted by `sprintf(fmt, breaks)`; other endpoints, e.g. [Date](#) objects, will be formatted by `format(breaks, fmt)`.
- If `fmt` is a list, then it will be used as arguments to [format](#).
- If `fmt` is a function, it should take a vector of numbers (or other objects that can be used as breaks) and return a character vector. It may be helpful to use functions from the `{scales}` package, e.g. [scales::label_comma\(\)](#).

See Also

Other labelling functions: [lbl_dash\(\)](#), [lbl_date\(\)](#), [lbl_discrete\(\)](#), [lbl_endpoints\(\)](#), [lbl_glue\(\)](#), [lbl_manual\(\)](#), [lbl_midpoints\(\)](#), [lbl_seq\(\)](#)

Examples

```
tab(-10:10, c(-3, 0, 0, 3),
  labels = lbl_intervals())

tab(-10:10, c(-3, 0, 0, 3),
  labels = lbl_intervals(fmt = list(nsmall = 1)))

tab_evenly(runif(20), 10,
  labels = lbl_intervals(fmt = percent))
```

lbl_midpoints	<i>Label chopped intervals by their midpoints</i>
---------------	---

Description

This uses the midpoint of each interval for its label.

Usage

```
lbl_midpoints(
  fmt = NULL,
  single = NULL,
  first = NULL,
  last = NULL,
  raw = deprecated()
)
```

Arguments

fmt	String, list or function. A format for break endpoints.
single	Glue string: label for singleton intervals. See lbl_glue() for details. If NULL, singleton intervals will be labelled the same way as other intervals.
first	Glue string: override label for the first category. Write e.g. first = "<{r}" to create a label like "<18". See lbl_glue() for details.
last	String: override label for the last category. Write e.g. last = ">{l}" to create a label like ">65". See lbl_glue() for details.
raw	[Deprecated] . Throws an error. Use the raw argument to chop() instead.

Value

A function that creates a vector of labels.

Formatting endpoints

If fmt is not NULL then it is used to format the endpoints.

- If fmt is a string, then numeric endpoints will be formatted by `sprintf(fmt, breaks)`; other endpoints, e.g. [Date](#) objects, will be formatted by `format(breaks, fmt)`.
- If fmt is a list, then it will be used as arguments to [format](#).
- If fmt is a function, it should take a vector of numbers (or other objects that can be used as breaks) and return a character vector. It may be helpful to use functions from the `{scales}` package, e.g. [scales::label_comma\(\)](#).

See Also

Other labelling functions: [lbl_dash\(\)](#), [lbl_date\(\)](#), [lbl_discrete\(\)](#), [lbl_endpoints\(\)](#), [lbl_glue\(\)](#), [lbl_intervals\(\)](#), [lbl_manual\(\)](#), [lbl_seq\(\)](#)

Examples

```
chop(1:10, c(2, 5, 8), lbl_midpoints())
```

lbl_seq
Label chopped intervals in sequence

Description

lbl_seq() labels intervals sequentially, using numbers or letters.

Usage

```
lbl_seq(start = "a")
```

Arguments

start String. A template for the sequence. See below.

Details

start shows the first element of the sequence. It must contain exactly *one* character out of the set "a", "A", "i", "I" or "1". For later elements:

- "a" will be replaced by "a", "b", "c", ...
- "A" will be replaced by "A", "B", "C", ...
- "i" will be replaced by lower-case Roman numerals "i", "ii", "iii", ...
- "I" will be replaced by upper-case Roman numerals "I", "II", "III", ...
- "1" will be replaced by numbers "1", "2", "3", ...

Other characters will be retained as-is.

Value

A function that creates a vector of labels.

See Also

Other labelling functions: [lbl_dash\(\)](#), [lbl_date\(\)](#), [lbl_discrete\(\)](#), [lbl_endpoints\(\)](#), [lbl_glue\(\)](#), [lbl_intervals\(\)](#), [lbl_manual\(\)](#), [lbl_midpoints\(\)](#)

Examples

```
chop(1:10, c(2, 5, 8), lbl_seq())
```

```
chop(1:10, c(2, 5, 8), lbl_seq("i."))
```

```
chop(1:10, c(2, 5, 8), lbl_seq("(A)"))
```

non-standard-types	<i>Tips for chopping non-standard types</i>
--------------------	---

Description

Santoku can handle many non-standard types.

Details

- If objects can be compared using `<`, `==` etc. then they should be choppable.
- Objects which can't be converted to numeric are handled within R code, which may be slower.
- Character `x` and breaks are chopped with a warning.
- If `x` and breaks are not the same type, they should be able to be cast to the same type, usually using `vctrs::vec_cast_common()`.
- Not all chopping operations make sense, for example, `chop_mean_sd()` on a character vector.
- For indexed objects such as `stats::ts()` objects, indices will be dropped from the result.
- If you get errors, try setting `extend = FALSE` (but also file a bug report).
- To request support for a type, open an issue on GitHub.

See Also

`brk-width-for-Datetime`

percent	<i>Simple percentage formatter</i>
---------	------------------------------------

Description

`percent()` formats `x` as a percentage. For a wider range of formatters, consider the [scales package](#).

Usage

```
percent(x)
```

Arguments

<code>x</code>	Numeric values.
----------------	-----------------

Value

`x` formatted as a percent.

Examples

```
percent(0.5)
```

Index

* chopping functions

- chop, 6
- chop_equally, 10
- chop_evenly, 11
- chop_fn, 12
- chop_mean_sd, 14
- chop_n, 15
- chop_proportions, 17
- chop_quantiles, 18
- chop_spikes, 20
- chop_width, 21
- fillet, 24

* labelling functions

- lbl_dash, 25
- lbl_date, 27
- lbl_discrete, 28
- lbl_endpoints, 30
- lbl_glue, 31
- lbl_intervals, 33
- lbl_midpoints, 35
- lbl_seq, 36

- base::cut(), 9
- base::pretty(), 16
- breaks, 20
- breaks-class, 4
- brk_default, 4
- brk_equally (chop_equally), 10
- brk_evenly (chop_evenly), 11
- brk_fn (chop_fn), 12
- brk_manual, 5
- brk_mean_sd (chop_mean_sd), 14
- brk_n (chop_n), 15
- brk_pretty (chop_pretty), 16
- brk_proportions (chop_proportions), 17
- brk_quantiles (chop_quantiles), 18
- brk_quantiles(), 11
- brk_spikes (chop_spikes), 20
- brk_width (chop_width), 21
- brk_width-for-datetime, 6, 22

- brk_width.Duration
 - (brk_width-for-datetime), 6

- chop, 6
- chop(), 2, 3, 5, 11–26, 30, 32, 33, 35
- chop_deciles (chop_quantiles), 18
- chop_equally, 10
- chop_equally(), 9, 12–14, 16, 18–20, 22, 25
- chop_evenly, 11
- chop_evenly(), 9, 11, 13, 14, 16, 18–20, 22, 25
- chop_fn, 12
- chop_fn(), 9, 11, 12, 14, 16, 18–20, 22, 25
- chop_mean_sd, 14
- chop_mean_sd(), 9, 11–13, 16, 18–20, 22, 25, 37
- chop_n, 15
- chop_n(), 9, 11–14, 18–20, 22, 25
- chop_pretty, 16
- chop_proportions, 17
- chop_proportions(), 9, 11–14, 16, 19, 20, 22, 25
- chop_quantiles, 18
- chop_quantiles(), 9, 11–14, 16, 18, 20, 22, 25
- chop_spikes, 20
- chop_spikes(), 9, 11–14, 16, 18, 19, 22, 23, 25
- chop_width, 21
- chop_width(), 9, 11–14, 16, 18–20, 25
- cut(), 2
- Date, 6, 8, 26, 28, 29, 31, 32, 34, 35
- date-time, 8
- difftime, 6
- dissect, 22
- dissect(), 20
- Duration, 6
- ecdf(x), 19

empirical cumulative distribution
 function, 19
 exactly, 24

 factor, 6, 9, 11–15, 17, 19–21, 25
 fillet, 24
 fillet(), 9, 11–14, 16, 18–20, 22
 format, 26, 28, 29, 31, 32, 34, 35
 format.breaks (breaks-class), 4
 function, 11–15, 17, 19–21

 Glue, 22
 glue::glue(), 32

 Hmisc::wtd.quantile(), 18

 is.breaks (breaks-class), 4

 kiru (chop), 6

 lbl_dash, 25
 lbl_dash(), 8, 28, 29, 31, 32, 34–36
 lbl_date, 27
 lbl_date(), 26, 29, 31, 32, 34–36
 lbl_datetime (lbl_date), 27
 lbl_discrete, 28
 lbl_discrete(), 26–28, 31, 32, 34–36
 lbl_endpoint (lbl_endpoints), 30
 lbl_endpoints, 30
 lbl_endpoints(), 26, 28, 29, 32, 34–36
 lbl_glue, 31
 lbl_glue(), 26–36
 lbl_intervals, 33
 lbl_intervals(), 26, 28, 29, 31, 32, 35, 36
 lbl_manual(), 26, 28, 29, 31, 32, 34–36
 lbl_midpoints, 35
 lbl_midpoints(), 26, 28, 29, 31, 32, 34, 36
 lbl_seq, 36
 lbl_seq(), 26, 28, 29, 31, 32, 34, 35
 lubridate::add_with_rollback(), 6

 mathematical set notation, 8

 non-standard-types, 37

 Ops, 8

 percent, 37
 Period, 6
 POSIXct, 6

 print.breaks (breaks-class), 4

 santoku (santoku-package), 2
 santoku-package, 2
 scales::label_comma(), 26, 28, 29, 31, 32,
 34, 35
 stats::quantile(), 18, 19
 stats::ts(), 37

 tab (chop), 6
 tab_deciles (chop_quantiles), 18
 tab_dissect (dissect), 22
 tab_equally (chop_equally), 10
 tab_evenly (chop_evenly), 11
 tab_fn (chop_fn), 12
 tab_mean_sd (chop_mean_sd), 14
 tab_n (chop_n), 15
 tab_pretty (chop_pretty), 16
 tab_proportions (chop_proportions), 17
 tab_quantiles (chop_quantiles), 18
 tab_spikes (chop_spikes), 20
 tab_width (chop_width), 21
 table, 6, 9, 11–15, 17, 19–21
 table(), 23

 vctrs::vec_cast_common(), 37