Package ‘janitor’

R functions (excerpt for Data Analysis):

adorn_ns	Add underlying Ns to a tabyl displaying percentages.

Description

This function adds back the underlying Ns to a tabyl whose percentages were calculated using adorn_percentages(), to display the Ns and percentages together. You can also call it on a non- tabyl data.frame with tabyl-like format to which you wish to append Ns.

Usage

```
adorn_ns(dat, position = "rear", ns = attr(dat, "core"))
```

Arguments

dat	a data.frame of class tabyl that has had adorn_percentages and/or adorn_pct_formatting called on it. If given a list of data.frames, this function will apply itself to each data.frame in the list (designed for 3-way tabyl lists).
position	should the N go in the front, or in the rear, of the percentage?
ns	the Ns to append. The default is the "core" attribute of the input tabyl dat, where the original Ns of a two-way tabyl are stored. However, if you need to modify the numbers, e.g., to format 4000 as 4,000 or 4k, you can do that separately and supply the formatted result here.

Value

a data.frame with Ns appended

Examples

```
mtcars
%>%
  tabyl(am, cyl) %>%
  adorn_percentages("col") %>%
  adorn_pct_formatting() %>%
  adorn_ns(position = "front")
```

adorn_pct_formatting   Format a data.frame of decimals as percentages.

Description

Numeric columns get multiplied by 100 and formatted as percentages according to user specifications. This function excludes the first column of the input data.frame, assuming that it contains a descriptive variable. Other non-numeric columns are also excluded.

Usage

```
adorn_pct_formatting(dat, digits = 1, rounding = "half to even", affix_sign
  = TRUE)
```

Arguments

dat   a data.frame with decimal values, typically the result of a call to adorn_percentages on a tabyl. If given a list of data.frames, this function will apply itself to each data.frame in the list (designed for 3-way tabyl lists).

digits   how many digits should be displayed after the decimal point?

rounding   method to use for rounding - either "half to even", the base R default method, or "half up", where 14.5 rounds up to 15.

affix_sign   should the % sign be affixed to the end?

Value

a data.frame with formatted percentages

Examples

```
mtcars
%>%
  tabyl(am, cyl) %>%
  adorn_percentages("col") %>%
  adorn_pct_formatting()
```

adorn_percentages Convert a data.frame of counts to percentages.

Description

This function excludes the first column of the input data.frame, assuming that it contains a descriptive variable. If the input data.frame is not a tabyl, it will convert to one in order to preserve the underlying values in the core attribute.

Usage

```
adorn_percentages(dat, denominator = "row", na.rm = TRUE)
```

Arguments

dat a tabyl or other data.frame with a tabyl-like layout. If given a list of data.frames, this function will apply itself to each data.frame in the list (designed for 3-way tabyl lists).

denominator the direction to use for calculating percentages. One of "row", "col", or "all".

na.rm should missing values (including NaN) be omitted from the calculations?

Value

Returns a data.frame of percentages, expressed as numeric values between 0 and 1.

Examples

```
mtcars %>%
  tabyl(am, cyl) %>%
  adorn_percentages("col")

# calculates correctly even with totals column and/or row:
mtcars %>%
  tabyl(am, cyl) %>%
  adorn_totals("row")
  %>%
  adorn_percentages()
```

tabyl   Generate a frequency table (1-, 2-, or 3-way).

Description

A fully-featured alternative to table(). Results are data.frames and can be formatted and enhanced with janitor's family of adorn_ functions.

Specify a data.frame and the one, two, or three unquoted column names you want to tabulate. Three variables generates a list of 2-way tabyls, split by the third variable.

Alternatively, you can tabulate a single variable that isn't in a data.frame by calling tabyl on a vector, e.g., tabyl(mtcars$gear).

Usage

```
tabyl(dat, ...)

## Default S3 method:
tabyl(dat, show_na = TRUE,
  show_missing_levels = TRUE, ...)

## S3 method for class Gdata.frameG
tabyl(dat, var1, var2, var3, show_na = TRUE,
  show_missing_levels = TRUE, ...)
```

Arguments

dat — a data.frame containing the variables you wish to count. Or, a vector you want to tabulate.

... — the arguments to tabyl (here just for the sake of documentation compliance, as all arguments are listed with the vector- and data.frame-specific methods)

show_na — should counts of NA values be displayed? In a one-way tabyl, the presence of NA values triggers an additional column showing valid percentages(calculated excluding NA values).

show_missing_levels — should counts of missing levels of factors be displayed? These will be rows and/or columns of zeroes. Useful for keeping consistent output dimensions even when certain factor levels may not be present in the data.

var1 — the column name of the first variable.

var2 — (optional) the column name of the second variable (the rows in a 2-way tabulation).

var3 — (optional) the column name of the third variable (the list in a 3-way tabulation).

Value

Returns a data.frame with frequencies and percentages of the tabulated variable(s). A 3-way tabulation returns a list of data.frames.