Package 'percentiles'

April 8, 2021

Type Package

Title Calculate (Stratified) Percentiles

Version 0.2.2

Author J. Peter Marquardt <peter@kmarquardt.de> [aut, cre]

Maintainer J. Peter Marquardt <peter@kmarquardt.de>

Description Calculate (stratified) percentiles on a data.frame
Stratification will split the data.frame into subgroups and calculate percentiles for each independently.

Depends R (>= 4.0.0)

Imports dplyr, assertive.types, assertthat, R6

License GPL-3

Encoding UTF-8

RoxygenNote 7.1.1

NeedsCompilation no

Repository CRAN

Date/Publication 2021-04-08 21:50:08 UTC

R topics documented:

`calculate_percentiles` *Calculate percentiles*

Description

Calculate percentiles for values in a data.frame

Usage

```
calculate_percentiles(data, value_col)
```

Arguments

`data`	A data.frame
`value_col`	character name of column containing values

Value

A vector of numerics with percentile values of length of nrow(data)

Author(s)

Peter Marquardt

`calculate_stratified_percentiles`
Calculate stratified percentiles

Description

Calculate percentiles for values in a data.frame while stratifying for other characteritics in same df

Usage

```
calculate_stratified_percentiles(data, value_col, stratify_by, use.na = FALSE)
```

Arguments

`data`	A data frame
`value_col`	character name of column containing values
`stratify_by`	list or vector. Use a named list to specify column name as key and a value of type vector indicating accepted levels of the property stratified by to be included. If an unnamed list or vector is passed, all levels of indicated columns will be used
`use.na`	A logical indicating whether NA values should be used. If TRUE, NA values and non-included value levels will be grouped like a separate value level

Value

A vector of numerics with percentile values of length of nrow(data)

Author(s)

J. Peter Marquardt

Examples

```
data <- data.frame('values' = 100:1, 'group' = rep(c('A', 'B', NA, 'D'), 25))
calculate_stratified_percentiles(data, 'values', list(group = c('A', 'B', 'D')))
calculate_stratified_percentiles(data, 'values', c('group'), use.na = TRUE)
calculate_stratified_percentiles(data, 'values', list(group = c('A', 'C')), use.na=TRUE)
# The following example will result in NA values caused by NAs in 'group'.
# Therefore, it will return the percentile vector, but issue a warning.
calculate_stratified_percentiles(data, 'values', 'group')
```

Stratified_percentile_calculator_generator

R6 Class representing a compund of data and methods used to calculate stratified percentiles

---

Description

R6 Class representing a compund of data and methods used to calculate stratified percentiles

R6 Class representing a compund of data and methods used to calculate stratified percentiles

Details

A calculator has: - raw_data representing the data.frame passed in for calculation - result_data an environment containing the result data.frame $data, shared with - sub_results representing subordinate steps in recursive calculation process

Active bindings

`raw_data` Return the data.frame originally handed to the object

`result_data` Return the environment containing a data.frame (§data) containing results of current hierarchy

`sub_results` Return the named list with Stratified_percentile_calculator_generator objects for recursive stacking

Methods

Public methods:

- Stratified_percentile_calculator_generator$new()
- Stratified_percentile_calculator_generator$divide_and_calculate()
- Stratified_percentile_calculator_generator$clone()

Method `new()`**:** Create a new Stratified_percentile_calculator object.

Usage:

```
Stratified_percentile_calculator_generator$new(
  raw_data = NULL,
  result_data = new.env(),
  current_stratification_characteristic = NULL,
  remaining_stratification_characteristics = NULL,
  value_column = NULL,
  output_column = NULL,
  use.na = FALSE
)
```

Arguments:

`raw_data` data.frame to perform calculation/stratification on.

`result_data` environment containing $data, a data.frame with the current state of results.

`current_stratification_characteristic` named list with column name and levels of characteristic to stratify by.

`remaining_stratification_characteristics` named list with column names and levels of characteristics to stratify by.

`value_column` character column with values to calculate percentiles on

`output_column` character column to write calculated percentile values to

`use.na` logical indicating whether or not NA/non-listed stratification values should be included as a separate group

Returns: A new 'Stratified_percentile_calculator' object.

Method `divide_and_calculate()`**:** recursively calculate stratified percentiles on data.frame Updates following private fields: - ..result_data$data - ::sub_results - ..current_stratification_characteristic - ..remaining_stratification_characteristics

Usage:

```
Stratified_percentile_calculator_generator$divide_and_calculate()
```

Returns: void, but updates ..result_data field

Method `clone()`**:** The objects of this class are cloneable with this method.

Usage:

```
Stratified_percentile_calculator_generator$clone(deep = FALSE)
```

Arguments:

`deep` Whether to make a deep clone.

Index