

Package: dsmisc (via r-universe)

August 22, 2024

Type Package

Title Data Science Box of Pandora Miscellaneous

Version 0.4.0

Date 2020-10-15

Description Tool collection for common and not so common data science use cases. This includes custom made algorithms for data management as well as value calculations that are hard to find elsewhere because of their specificity but would be a waste to get lost nonetheless. Currently available functionality: find sub-graphs in an edge list data.frame, find mode or modes in a vector of values, extract (a) specific regular expression group(s), generate ISO time stamps that play well with file names, or generate URL parameter lists by expanding value combinations.

License GPL (>= 2)

Imports Rcpp (>= 1.0.1), stringr

LinkingTo Rcpp

RoxygenNote 7.1.1

Encoding UTF-8

SystemRequirements C++11

Suggests covr, testthat, spelling

Language en-US

BugReports <https://github.com/petermeissner/dsmisc/issues>

URL <https://github.com/petermeissner/dsmisc>

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

RemoteUrl <https://github.com/petermeissner/dsmisc>

RemoteRef HEAD

RemoteSha ed730b2d5f51cfc857e9830e24130e7969dcddc9

Contents

df_defactorize	2
graphs_find_subgraphs	3
stats_mode	3
stats_mode_multi	4
str_group_extract	4
time_stamp	5
tool_i_fit_index	6
tool_i_fit_obj	6
web_gen_param_list_expand	7

Index	8
--------------	----------

df_defactorize	<i>df_defactorize</i>
----------------	-----------------------

Description

df_defactorize

Usage

```
df_defactorize(df)
```

Arguments

df a data.frame like object

Value

returns the same data.frame except that factor columns have been transformed into character columns

Examples

```
df <-
  data.frame(
    a = 1:2,
    b = factor(c("a", "b")),
    c = as.character(letters[3:4]),
    stringsAsFactors = FALSE
  )
vapply(df, class, "")

df_df <- df_defactorize(df)
vapply(df_df, class, "")
```

 graphs_find_subgraphs *Subgraphs in Undirected Graphs/Networks*

Description

Finding and indexing subgraphs in undirected graph.

Usage

```
graphs_find_subgraphs(id_1, id_2, verbose = 1L)
```

Arguments

id_1	vector of integers indicating ids
id_2	vector of integers indicating ids
verbose	in integer indicating the amount of verbosity; good for long running tasks or to get more information about the workings of the algorithm; currently accepted values: 0, 1, 2

Details

Input is given as two vectors where each pair of node ids 'id_1[i]' - 'id_2[i]' indicates an edge between two nodes.

Value

An integer vector with subgraph ids such that each distinct subgraph - i.e. all nodes are reachable within the graph and no node outside the subgraph is reachable - gets a distinct integer value. Integer values are assigned via

Examples

```
graphs_find_subgraphs(c(1,2,1,5,6,6), c(2,3,3,4,5,4), verbose = 0)
graphs_find_subgraphs(c(1,2,1,5,6,6), c(2,3,3,4,5,4), verbose = 2)
```

stats_mode	<i>Mode</i>
------------	-------------

Description

Function calculating the mode.

Usage

```
stats_mode(x, multimodal = FALSE, warn = TRUE)
```

Arguments

x	vector to get mode for
multimodal	wether or not all modes should be returned in case of more than one
warn	should the function warn about multimodal outcomes?

Value

vector of mode or modes

stats_mode_multi *Mode Allowing for Multi Modal Mode*

Description

Function calculating the mode, allowing for multiple modes in case of equal frequencies.

Usage

```
stats_mode_multi(x)
```

Arguments

x	vector to get mode for
---	------------------------

Value

vector with all modes

str_group_extract *Extract Regular Expression Groups*

Description

Extract Regular Expression Groups

Usage

```
str_group_extract(string, pattern, group = NULL, nas = TRUE)
```

Arguments

string	string to extract from
pattern	pattern with groups to match
group	groups to extract
nas	return NA values (TRUE) or filter them out (FALSE)

Value

string vector or string matrix

Examples

```
strings <- paste(LETTERS, seq_along(LETTERS), sep = "_")
str_group_extract(strings, "[\\w]_(\\d+)")
str_group_extract(strings, "[\\w]_(\\d+)", 1)
str_group_extract(strings, "[\\w]_(\\d+)", 2)
```

time_stamp

Time Stamps for File Names

Description

Generating file name ready iso time stamps.

Usage

```
time_stamp(ts = Sys.time(), sep = c("-", "_", "_"))
```

Arguments

ts	one or more POSIX time stamp
sep	separators to be used for formatting

Value

Returns timestamp string in format yyyy-mm-dd_HH_MM_SS ready to be used safely in file names on various operating systems.

Examples

```
time_stamp()
time_stamp( Sys.time() - 10000 )
```

tool_i_fit_index	<i>Fit number into length of index</i>
------------------	--

Description

Fit number into length of index

Usage

```
tool_i_fit_index(i, index)
```

Arguments

i	number to make sure to fit into index range
index	size of index

Value

number that fits into index by circularly mapping i to range of 1 to size of index

See Also

[tool_i_fit_obj](#)

Examples

```
tool_i_fit_index(-2:6, 3)
```

tool_i_fit_obj	<i>Subset object even if index is out of range</i>
----------------	--

Description

Subset object even if index is out of range

Usage

```
tool_i_fit_obj(i, obj)
```

Arguments

i	number to make sure to fit into index range
obj	object to subset data from

Value

elements of object circularly indexed by mapping i to range of 1 to size of object

See Also

[tool_i_fit_index](#)

Examples

```
tool_i_fit_obj(-2:6, 3)
```

web_gen_param_list_expand

URL Parameter Combinations

Description

Generate URL parameter combinations from sets of parameter values.

Usage

```
web_gen_param_list_expand(..., sep_1 = "=", sep_2 = "&")
```

Arguments

...	multiple vectors passed on as named arguments or a single list or a data.frame
sep_1	first separator to use between key and value
sep_2	second separator to use between key-value pairs

Value

string vector with assembled query string parameter combinations

Examples

```
web_gen_param_list_expand(q = "beluga", lang = c("de", "en"))
```

Index

df_defactorize, [2](#)
graphs_find_subgraphs, [3](#)
stats_mode, [3](#)
stats_mode_multi, [4](#)
str_group_extract, [4](#)
time_stamp, [5](#)
tool_i_fit_index, [6](#), [7](#)
tool_i_fit_obj, [6](#), [6](#)
web_gen_param_list_expand, [7](#)