Package 'md.tools'

May 21, 2022

Type Package

Title Masked data tools

Version 0.1.0

Maintainer Alex Towell <lex@metafunctor.com>

Description A set of tools for reading, writing, and manipulating masked data. Since masked data must often deal with matrices, particularly Boolean matrices, it also provides auxilliary tools designed to make working with them easier.

License MIT + file LICENSE

Encoding UTF-8

LazyData true

RoxygenNote 7.1.2

URL https://queelius.github.io/md.tools/,

https://github.com/queelius/md.tools

Imports purrr,
stats,
dplyr,
tibble,
jsonlite,
readr,
stringr

BugReports https://github.com/queelius/md.tools/issues

Depends R (>= 2.10)

R topics documented:

1

---

exp_series_data_1	*Masked data for a series system with exponentially distributed nodes.*

Description

Masked data containing the system lifetime and other attributes of n=1000 series system with parameter value theta=c(3,4,5) and candidate model in which conditions 1, 2, and 3 are met.

Usage

```
exp_series_data_1
```

Format

A data frame with 1000 rows and 9 variables:

s Real observable variable, series system lifetime

k Integer latent variable, the assigned index of the failed component

t1 Real latent variable, lifetime of component 1

t2 Real latent variable, lifetime of component 2

t3 Real latent variable, lifetime of component 3

x1 Boolean observable variable, TRUE indicates component assigned index 1 is in candidate set

x2 Boolean observable variable, TRUE indicates component assigned index 2 is in candidate set

x3 Boolean observable variable, TRUE indicates component assigned index 3 is in candidate set

Source

https://github.com/queelius/masked.data/blob/master/data-raw/exp_series_data_1_gen.R

---

is_md	*Test whether an object is a masked data (*tbl_md*).*

Description

Test whether an object is a masked data (tbl_md).

Usage

```
is_md(x)
```

Arguments

x   object to determine if masked data

---

md	*Constructor for masked data.*

Description

Takes an object x and converts it to masked data, a tibble (data frame) with some extra attributes, e.g., latent attribute to specify which variables are latent in the model.

Usage

```
md(x)
```

Arguments

x   object to convert to masked data.

---

md_boolean_matrix_to_list	
	Map Boolean matrix mat *to a list of integer vectors.*

Description

Map Boolean matrix mat to a list of integer vectors.

Usage

```
md_boolean_matrix_to_list(mat)
```

Arguments

mat   Boolean matrix

`md_cand_contains`	*Decorates masked data with whether candidate set contains failed component.*

Description

Takes masked data frame `md` with candidate set encoded as `x` and a column `k` denoting component cause of failure and returns a decorated masked data frame with a column `contains` that denotes whether the candidate set contains the component cause of failure.

Usage

```
md_cand_contains(md, var = "x")
```

Arguments

md	masked data frame
var	column prefix for matrix encoding

`md_cand_sizes`	*Decorates masked data with candidate sizes*

Description

Takes masked data frame `md` with candidate set encoded as `x` and returns a decorated masked data frame with a column `w` that denotes the size of the candidate sets. No new information is added, it just counts the number of times that a row element of `x` is `TRUE`.

Usage

```
md_cand_sizes(md, var = "x")
```

Arguments

md	masked data frame
var	column prefix for matrix encoding

---

md_decode_matrix	*Decodes a matrix from specified columns in a data frame.*

Description

An `nrow(df)`-by-p matrix `var` is encoded in data frame `df` with the columns `var.1`,...,`var.p` or `var1`,...,`varp`.

Usage

```
md_decode_matrix(df, var)
```

Arguments

df	data frame that contains the matrix
var	the symbolic name of the matrix

Details

A matrix will be returned with the appropriate ordering denoted by the index, e.g., `a.2` will come before `a.4`. There should be no gaps in the matrix indexes, e.g., if there is `a.4` then there must also be `a.1`,`a.2`,`a.3`.

Value

a matrix

---

md_encode_matrix	*Encodes a matrix as a data frame with specified columns.*

Description

Encodes a matrix as a data frame with specified columns.

Usage

```
md_encode_matrix(mat, var)
```

Arguments

mat	matrix
var	the symbolic name of the matrix (prefix of column names)

Value

a tibble (data frame) encoding of a matrix

---

md_latent   *Obtain a list of latent variables from masked data.*

Description

Obtain a list of latent variables from masked data.

Usage

```
md_latent(md)
```

Arguments

md   masked data to retrieve latent variables from.

---

md_list_to_boolean_matrix

Map list of integer vectors to Boolean matrix.

Description

Map list of integer vectors to Boolean matrix.

Usage

```
md_list_to_boolean_matrix(xs)
```

Arguments

xs   List of integer vectors.

---

md_mark_latent   *Mark a variable in a masked data frame as latent.*

Description

Mark a variable in a masked data frame as latent.

Usage

```
md_mark_latent(md, vars)
```

Arguments

md   masked data to modify
vars   variables to mark as latent.

| md_par_ttf | *Generates time-to-failure (ttf) and component cause of failure for a parallel system with the given data frame of component times-to-failure encoded by the matrix columns prefixed with* var *in masked data frame* md. |

Description

Generates time-to-failure (ttf) and component cause of failure for a parallel system with the given data frame of component times-to-failure encoded by the matrix columns prefixed with var in masked data frame md.

Usage

```
md_par_ttf(md, tau = NULL, var = "t")
```

Arguments

md	a data frame of component failure times
tau	suspension time
var	component symbolic variable, defaults to t

| md_read_csv_with_meta | *Read a (masked) data frame table from a connection (e.g., url or filename).* |

Description

Read a (masked) data frame table from a connection (e.g., url or filename).

Usage

```
md_read_csv_with_meta(
  file,
  read_meta = T,
  comment = "#",
  max_meta_lns = 1000,
  ...
)
```

Arguments

`file`	a path to a file, a connection, or literal data
`read_meta`	whether to read in metadata to populate attributes
`comment`	comment indicator, defaults to #
`max_meta_lns`	limit metadata search to the indicated number of lines
`...`	additional arguments to pass, like `skip`

`md_series_ttf`	*Generates time-to-failure (ttf) and component cause of failure for a series system with the given data frame of component times-to-failure encoded by the matrix columns prefixed with* `var` *in masked data frame* `md`.

Description

Generates time-to-failure (ttf) and component cause of failure for a series system with the given data frame of component times-to-failure encoded by the matrix columns prefixed with `var` in masked data frame `md`.

Usage

```
md_series_ttf(md, tau = NULL, var = "t")
```

Arguments

`md`	a data frame with the indicated component times-to-failure
`tau`	suspension time
`var`	component symbolic variable, defaults to `t`

`md_unmark_latent`	*Mark a variable in a masked data frame as latent.*

Description

Mark a variable in a masked data frame as latent.

Usage

```
md_unmark_latent(md, vars)
```

Arguments

`md`	masked data to modify
`vars`	variables to mark as latent.

md_write_csv_with_meta

Write data frame object to a CSV (comma separated file), optionally with associated attribute data (stored as JSON in comments)

Description

Write data frame object to a CSV (comma separated file), optionally with associated attribute data (stored as JSON in comments)

Usage

```
md_write_csv_with_meta(df, file, comment = "#", ...)
```

Arguments

df	a data frame with attributes to write (like a masked data frame)
file	filename for csv
comment	denotes a comment block
...	additional arguments to pass

print.tbl_md   *Print method for masked data (*tbl_md*).*

Description

Print method for masked data (tbl_md).

Usage

```
## S3 method for class 'tbl_md'
print(x, drop_latent = F, ...)
```

Arguments

x	masked data to print
drop_latent	If TRUE, drop the latent random variables
...	additional arguments to pass

Index