Package 'dcmodify'

October 6, 2017

Maintainer Mark van der Loo <mark.vanderloo@gmail.com>

License GPL-3

Title Modify Data Using Externally Defined Modification Rules

LazyData no

Type Package

LazyLoad yes

Description Data cleaning scripts typically contain a lot of 'if this change that' type of statements. Such statements are typically condensed expert knowledge. With this package, such 'data modifying rules' are taken out of the code and become in stead parameters to the work flow. This allows one to maintain, document, and reason about data modification rules separately from the workflow.

Version 0.1.1

Depends methods

URL https://github.com/data-cleaning/dcmodify

BugReports https://github.com/data-cleaning/dcmodify/issues

Date 2017-10-06

Imports yaml, validate (>= 0.1.5), settings, utils

Suggests testthat, knitr, lumberjack

VignetteBuilder knitr

Collate 'guard.R' 'modifier.R' 'modify.R'

RoxygenNote 6.0.1

NeedsCompilation no

Author Mark van der Loo [cre, aut],
Edwin de Jonge [aut]

Repository CRAN

Date/Publication 2017-10-06 16:49:13 UTC

R topics documented:

`dcmodify` *Data Modification By Modifying Rules*

Description

Data Modification By Modifying Rules

Introduction

Data often contain errors and missing data. Experts can often correct commonly occuring errors based on simple conditional rules. This package facilitates the expression, management, and application of such rules on data sets.

The general workflow in `dcmodify` follows the following pattern.

- Define or read a set of rules with `modifier`.
- `modify` data with the modification rules.
- Examine the results either graphically or by summary.

There are several convenience functions that allow one to define modification rules from the commandline, through a (freeform or yaml) file and to investigate and maintain the rules themselves. Please have a look at the introductory vignette

```
vignette("introduction",package="dcmodify")
```

`modifier` *Create or read a set of data modification rules*

Description

Create or read a set of data modification rules

Usage

```
modifier(..., .file)
```

Arguments

`...`	A comma-separated list of modification rules.
`.file`	A character vector of file locations.

Value

An object of class `modifier`.

Examples

```
m <- modifier( if (height < mean(height)) height <- 2*height
, if ( weight > mean(weight) ) weight <- weight/2  )
modify(women,m)
```

`modify` *Modify a data set*

Description

Modify a data set

Usage

```
modify(dat, x, ...)

## S4 method for signature 'data.frame,modifier'
modify(dat, x, ...)
```

Arguments

`dat`	An R object carrying data
`x`	An R object carrying modififying rules
`...`	Options.

Examples

```
m <- modifier( if (height < mean(height)) height <- 2*height
, if ( weight > mean(weight) ) weight <- weight/2  )
modify(women,m)
```

`modify_so` *Shortcut to modify data*

Description

Shortcut to modify data

Usage

```
modify_so(dat, ...)
```

Arguments

`dat`	A `data.frame`
`...`	A comma-separated list of modifying rules.

Index