

Package ‘dfexplore’

February 19, 2015

Type Package

Title Explore data.frames by plotting NA and classes of each variable

Version 0.2.1

Date 2013-12-06

Author Joris Muller

Maintainer Joris Muller <joris.muller@etu.unistra.fr>

Description Quickly and graphically show missing values and classes of each variable and each observation of a data.frame. The aim is to show pattern of missing values and if there is wrong class attribution.

License GPL

Depends ggplot2, methods

URL <https://github.com/jomuller/dfexplore>

Encoding UTF-8

LazyData true

NeedsCompilation no

Repository CRAN

Date/Publication 2014-01-16 09:27:47

R topics documented:

dfexplore-package	2
dfplot	2
example_df	4
expand_dfmatrix	5
simulate_dataframe	5
Index	7

dfexplore-package *Data-frame explorer*

Description

This package aim to explore quickly a data.frame. For example, find visually NA and classes of data.

Details

Package: dfexplore
Type: Package
Version: 0.2
Date: 2013-12-06
License: GPL

Author(s)

Joris Muller

Maintainer: Joris Muller <joris.muller@etu.unistra.fr>

Testing: Maël Barthoulot

See Also

[dfplot](#)

Examples

```
dfplot(example_df)
```

dfplot *Data Frame plot to explore NA and classes*

Description

Create a plot showing for each variable and observation of a data.frame the classe and if there is NA.

Usage

```
dfplot(dfdescription,title=NULL)
```

Arguments

`dfdescription` a data.frame or a data.frame.description object to plot.

`title` a character vector of length 1. Title of the data.frame. By default the name of the the object data.frame.

Value

Return a ggplot2 object. Directly plotted if not assign to a variable. Because it's a ggplot object, every layer could be changed (see example)

Author(s)

Joris Muller

See Also

[ggplot](#)

Examples

```
# Plot quickly a representation of the "example_df" data.frame
dfplot(example_df)

# Plot it with some changes using ggplot layers
graph_data.frame <- dfplot(example_df)

# Change title
with_title<-graph_data.frame + ggtitle("An example of dfplot() with example_df")
with_title

# Change text orientation
horizontal_text<-with_title + theme(axis.text.x = element_text(angle = 0))
horizontal_text

# Add subject number in the column subject
# May be useful if you want to find quickly the number of a subject
nb_obs <- nrow(example_df)
with_subject_number<-horizontal_text +
  geom_text(data =example_df,
            aes(y=1:nb_obs,
                x=rep(x=c(0.7,1,1.3), length.out=nb_obs),
                label =example_df$subject),
            size=3)

with_subject_number
```

`example_df`*Simulated data to test [dfplot](#).*

Description

A small data set randomly generated to simulate 100 observation on a survey with 10 questions.

Usage

```
data(example_df)
```

Format

A data frame with 100 observations on the following 18 variables.

`subject` a numeric vector : Unique number of the subject

`initial` a character vector : Initials of the subject

`birth` Birthdate

`sex` a factor with levels male female

`study_level` an ordered factor with levels primary < secondary < superior

`heigh` a numeric vector

`weight` a numeric vector

`siblings` a numeric vector

`Q1` a numeric vector : question 1

`Q2` a numeric vector

`Q3` a numeric vector

`Q4` a numeric vector

`Q5` a numeric vector

`Q6` a numeric vector

`Q7` a numeric vector

`Q8` a numeric vector

`Q9` a numeric vector

`Q10` a numeric vector

Examples

```
dfplot(example_df)
```

expand_dfmatrix	<i>Transform a data frame containing matrix to a data frame without matrix</i>
-----------------	--

Description

Data frames may have components matrices components. This is unusual, but technically allowed and necessary in some cases. Methods of dfexplore package have to deal with this.

Usage

```
expand_dfmatrix( dataframe, matrix_var = NA)
```

Arguments

dataframe	data.frame - a data frame with a matrix included
matrix_var	integer - position of the column containing matrix. If no argument are given, calculate these positions

Value

A data.frame with all the data. The column wich are not matrix keep unchanged and column with matrix are transform to a data.frame and combined.

Author(s)

Joris Muller

Examples

```
dataframe_with_matrix <- simulate_dataframe( includeMatrix=TRUE)
str(dataframe_with_matrix)
dataframe_without_matrix <- expand_dfmatrix(dataframe_with_matrix)
str(dataframe_without_matrix)
```

simulate_dataframe	<i>Simulate data frames with missing values</i>
--------------------	---

Description

Simulate a data frame representing a questionnaire randomly generated with n observation on a survey with questions including missing values. The aim of this data is to test the methods of dfexplore package based on different kind of data but should be used

Usage

```
simulate_dataframe(nsubjects = 100, nquestions = 10,  
                  includeMatrix = FALSE)
```

Arguments

nsubjects	integer - number of subjects in the data frame
nquestions	integer - number of questions in the data frame
includeMatrix	boolean - does the answer of the questions should be included as a matrix in the final data frame (see details)

Details

Data frames may have components matrices components. This is unusual, but technically allowed and necessary in some cases. Methods of `dfexplore` package have to deal with this.

Value

A `data.frame` with `nquestions + 8` columns and `nsubjects` observations with various data classes. It should contain a matrix component if `includeMatrix = TRUE`.

Author(s)

Joris Muller

See Also

There is already some simulated data frames included as [example_df](#)

Examples

```
simulated <- simulate_dataframe(nsubjects=200)  
str(simulated)  
dfplot(simulated)
```

Index

*Topic **datasets**

example_df, [4](#)

*Topic **package**

dfexplore-package, [2](#)

dfexplore (dfexplore-package), [2](#)

dfexplore-package, [2](#)

dfplot, [2](#), [2](#), [4](#)

dfplot, data.frame-method (dfplot), [2](#)

dfplot, Data.frame.description-method
(dfplot), [2](#)

example_df, [4](#), [6](#)

expand_dfmatrix, [5](#)

ggplot, [3](#)

simulate_dataframe, [5](#)