

Package ‘VIMGUI’

January 2, 2014

Version 0.9.0

Date 2013-10-09

Title Visualization and Imputation of Missing Values

Author Daniel Schopfhauser, Matthias Templ, Andreas Alfons, Alexander Kowarik, Bernd Prantner

Maintainer Alexander Kowarik <alexander.kowarik@statistik.gv.at>

Depends R (>= 2.13), tcltk, tkrplot, VIM (>= 4.0.0), survey, gWidgetsRGtk2

Imports foreign, RGtk2, Cairo, Hmisc

Description This package introduces new tools for the visualization of missing and/or imputed values, which can be used for exploring the data and the structure of the missing and/or imputed values. Depending on this structure of the missing values, the corresponding methods may help to identify the mechanism generating the missings and allows to explore the data including missing values. In addition, the quality of imputation can be visually explored using various univariate, bivariate, multiple and multivariate plot methods. A graphical user interface allows an easy handling of the implemented plot methods.

LazyData TRUE

SystemRequirements BWidget

License GPL (>= 2)

URL <https://github.com/alexxowa/VIMGUI>

Repository CRAN

Collate 'helperGUI.R' 'TKRaggr.R' 'TKRdevice.R' 'TKRmarginmatrix.R'
'TKRmatrixplot.R' 'TKRparcoordMiss.R' 'TKRpbox.R'
'TKRscattmatrixMiss.R' 'utils.R' 'VIMGUI.R' 'vmGUImenu.R' 'zzz.R'

NeedsCompilation no

Date/Publication 2014-01-02 19:18:18

R topics documented:

TKRaggr	2
TKRmarginmatrix	4
TKRmatrixplot	6
TKRparcoordMiss	7
TKRpbox	10
TKRscattmatrixMiss	12
VIMGUI	14
vmGUImenu	14
Index	16

TKRaggr	<i>Aggregations for missing/imputed values</i>
---------	--

Description

Calculate or plot the amount of missing/imputed values in each variable and the amount of missing/imputed values in certain combinations of variables.

Usage

```
TKRaggr(x, ..., delimiter = NULL, hscale = NULL,
        vscale = NULL, TKRpar = list())
```

Arguments

x	a vector, matrix or data.frame.
delimiter	a character-vector to distinguish between variables and imputation-indices for imputed variables (therefore, x needs to have colnames). If given, it is used to determine the corresponding imputation-index for any imputed variable (a logical-vector indicating which values of the variable have been imputed). If such imputation-indices are found, they are used for highlighting and the colors are adjusted according to the given colors for imputed variables (see col).
...	for <code>aggr</code> and <code>TKRaggr</code> , further arguments and graphical parameters to be passed to plot.aggr . For <code>plot.aggr</code> , further graphical parameters to be passed down. <code>par("oma")</code> will be set appropriately unless supplied (see par).
hscale	horizontal scale factor for plot to be embedded in a <i>Tcl/Tk</i> window (see ‘Details’). The default value depends on the number of variables.
vscale	vertical scale factor for the plot to be embedded in a <i>Tcl/Tk</i> window (see ‘Details’). The default value depends on the number of combinations.
TKRpar	a list of graphical parameters to be set for the plot to be embedded in a <i>Tcl/Tk</i> window (see ‘Details’ and par).

Details

Often it is of interest how many missing/imputed values are contained in each variable. Even more interesting, there may be certain combinations of variables with a high number of missing/imputed values.

If `combined` is `FALSE`, two separate plots are drawn for the missing/imputed values in each variable and the combinations of missing/imputed and non-missing values. The barplot on the left hand side shows the amount of missing/imputed values in each variable. In the *aggregation plot* on the right hand side, all existing combinations of missing/imputed and non-missing values in the observations are visualized. Available, missing and imputed data are color coded as given by `col`. Additionally, there are two possibilities to represent the frequencies of occurrence of the different combinations. The first option is to visualize the proportions or frequencies by a small bar plot and/or numbers. The second option is to let the cell heights be given by the frequencies of the corresponding combinations. Furthermore, variables may be sorted by the number of missing/imputed values and combinations by the frequency of occurrence to give more power to finding the structure of missing/imputed values.

If `combined` is `TRUE`, a small version of the barplot showing the amount of missing/imputed values in each variable is drawn on top of the aggregation plot.

The graphical parameter `oma` will be set unless supplied as an argument.

TKRaggr behaves like `plot.aggr`, but uses `tkrplot` to embed the plot in a *Tcl/Tk* window. This is useful if the number of variables and/or combinations is large, because scrollbars allow to move from one part of the plot to another.

Value

for `aggr`, a list of class "aggr" containing the following components:

- `x` the data used.
- `combinations` a character vector representing the combinations of variables.
- `count` the frequencies of these combinations.
- `percent` the percentage of these combinations.
- `missings` a `data.frame` containing the amount of missing/imputed values in each variable.
- `tabcomb` the indicator matrix for the combinations of variables.

Note

Some of the argument names and positions have changed with version 1.3 due to extended functionality and for more consistency with other plot functions in VIM. For back compatibility, the arguments `labs` and `names.arg` can still be supplied to `...{}` and are handled correctly. Nevertheless, they are deprecated and no longer documented. Use `ylabs` and `labels` instead.

Author(s)

Andreas Alfons, Matthias Templ, modifications for displaying imputed values by Bernd Prantner

References

M. Templ, A. Alfons, P. Filzmoser (2012) Exploring incomplete data using visualization tools. *Journal of Advances in Data Analysis and Classification*, Online first. DOI: 10.1007/s11634-011-0102-y.

See Also

[print.aggr](#), [summary.aggr](#)

Examples

```
data(sleep, package="VIM")
## for missing values
a <- aggr(sleep)
a
summary(a)

## for imputed values
sleep_IMPUTED <- kNN(sleep)
a <- aggr(sleep_IMPUTED, delimiter="_imp")
a
summary(a)
```

TKRmarginmatrix

Marginplot Matrix

Description

Create a scatterplot matrix with information about missing/imputed values in the plot margins of each panel.

Usage

```
TKRmarginmatrix(x, delimiter = NULL,
  col = c("skyblue", "red", "red4", "orange", "orange4"),
  alpha = NULL, hscale = NULL, vscale = NULL,
  TKRpar = list(), ...)
```

Arguments

<code>x</code>	a matrix or data.frame.
<code>delimiter</code>	a character-vector to distinguish between variables and imputation-indices for imputed variables (therefore, <code>x</code> needs to have colnames). If given, it is used to determine the corresponding imputation-index for any imputed variable (a logical-vector indicating which values of the variable have been imputed). If such imputation-indices are found, they are used for highlighting and the colors are adjusted according to the given colors for imputed variables (see <code>col</code>).

<code>col</code>	a vector of length five giving the colors to be used in the marginplots in the off-diagonal panels. The first color is used for the scatterplot and the boxplots for the available data, the second/fourth color for the univariate scatterplots and boxplots for the missing/imputed values in one variable, and the third/fifth color for the frequency of missing/imputed values in both variables (see ‘Details’). If only one color is supplied, it is used for the bivariate and univariate scatterplots and the boxplots for missing/imputed values in one variable, whereas the boxplots for the available data are transparent. Else if two colors are supplied, the second one is recycled.
<code>alpha</code>	a numeric value between 0 and 1 giving the level of transparency of the colors, or NULL. This can be used to prevent overplotting.
<code>...</code>	further arguments and graphical parameters to be passed to pairsVIM and marginplot . <code>par("oma")</code> will be set appropriately unless supplied (see par).
<code>hscale</code>	horizontal scale factor for plot to be embedded in a <i>Tcl/Tk</i> window (see ‘Details’). The default value depends on the number of variables.
<code>vscale</code>	vertical scale factor for the plot to be embedded in a <i>Tcl/Tk</i> window (see ‘Details’). The default value depends on the number of variables.
<code>TKRpar</code>	a list of graphical parameters to be set for the plot to be embedded in a <i>Tcl/Tk</i> window (see ‘Details’ and par).

Details

`marginmatrix` uses [pairsVIM](#) with a panel function based on [marginplot](#).

The graphical parameter `oma` will be set unless supplied as an argument.

`TKRmarginmatrix` behaves like `marginmatrix`, but uses [tkrplot](#) to embed the plot in a *Tcl/Tk* window. This is useful if the number of variables is large, because scrollbars allow to move from one part of the plot to another.

Author(s)

Andreas Alfons, modifications by Bernd Prantner

References

M. Templ, A. Alfons, P. Filzmoser (2012) Exploring incomplete data using visualization tools. *Journal of Advances in Data Analysis and Classification*, Online first. DOI: 10.1007/s11634-011-0102-y.

See Also

[marginplot](#), [pairsVIM](#), [scattmatrixMiss](#)

Examples

```
data(sleep, package = "VIM")
## for missing values
x <- sleep[, 1:5]
x[,c(1,2,4)] <- log10(x[,c(1,2,4)])
```

```
marginmatrix(x)

## for imputed values
x_imp <- kNN(sleep[, 1:5])
x_imp[,c(1,2,4)] <- log10(x_imp[,c(1,2,4)])
marginmatrix(x_imp, delimiter = "_imp")
```

TKRmatrixplot

Matrix plot

Description

Create a matrix plot, in which all cells of a data matrix are visualized by rectangles. Available data is coded according to a continuous color scheme, while missing/imputed data is visualized by a clearly distinguishable color.

Usage

```
TKRmatrixplot(x, delimiter = NULL, hscale = NULL,
              vscale = NULL, TKRpar = list(), ...)
```

Arguments

<code>x</code>	a matrix or data.frame.
<code>delimiter</code>	a character-vector to distinguish between variables and imputation-indices for imputed variables (therefore, <code>x</code> needs to have colnames). If given, it is used to determine the corresponding imputation-index for any imputed variable (a logical-vector indicating which values of the variable have been imputed). If such imputation-indices are found, they are used for highlighting and the colors are adjusted according to the given colors for imputed variables (see <code>col</code>).
<code>...</code>	for <code>matrixplot</code> and <code>iimagMiss</code> , further graphical parameters to be passed to plot.window , title and axis . For <code>TKRmatrixplot</code> , further arguments to be passed to <code>matrixplot</code> .
<code>hscale</code>	horizontal scale factor for plot to be embedded in a <i>Tcl/Tk</i> window (see ‘Details’). The default value depends on the number of variables.
<code>vscale</code>	vertical scale factor for the plot to be embedded in a <i>Tcl/Tk</i> window (see ‘Details’). The default value depends on the number of observations.
<code>TKRpar</code>	a list of graphical parameters to be set for the plot to be embedded in a <i>Tcl/Tk</i> window (see ‘Details’ and par).

Details

In a *matrix plot*, all cells of a data matrix are visualized by rectangles. Available data is coded according to a continuous color scheme. To compute the colors via interpolation, the variables are first scaled to the interval

$$[0, 1]$$

. Missing/imputed values can then be visualized by a clearly distinguishable color. It is thereby possible to use colors in the *HCL* or *RGB* color space. A simple way of visualizing the magnitude of the available data is to apply a greyscale, which has the advantage that missing/imputed values can easily be distinguished by using a color such as red/orange. Note that $-\text{Inf}$ and Inf are always assigned the begin and end color, respectively, of the continuous color scheme.

Additionally, the observations can be sorted by the magnitude of a selected variable. If `interactive` is `TRUE`, clicking in a column redraws the plot with observations sorted by the corresponding variable. Clicking anywhere outside the plot region quits the interactive session.

`TKRmatrixplot` behaves like `matrixplot`, but uses `tkrplot` to embed the plot in a *Tcl/Tk* window. This is useful if the number of observations and/or variables is large, because scrollbars allow to move from one part of the plot to another.

Note

This is a much more powerful extension to the function `imagmiss` in the former CRAN package `dprep`.

`iimagMiss` is deprecated and may be omitted in future versions of `VIM`. Use `matrixplot` instead.

Author(s)

Andreas Alfons, Matthias Templ, modifications by Bernd Prantner

References

M. Templ, A. Alfons, P. Filzmoser (2012) Exploring incomplete data using visualization tools. *Journal of Advances in Data Analysis and Classification*, Online first. DOI: 10.1007/s11634-011-0102-y.

Examples

```
data(sleep, package = "VIM")
## for missing values
x <- sleep[, -(8:10)]
x[, c(1,2,4,6,7)] <- log10(x[, c(1,2,4,6,7)])
matrixplot(x, sortby = "BrainWgt")

## for imputed values
x_imp <- kNN(sleep[, -(8:10)])
x_imp[, c(1,2,4,6,7)] <- log10(x_imp[, c(1,2,4,6,7)])
matrixplot(x_imp, delimiter = "_imp", sortby = "BrainWgt")
```

Description

Parallel coordinate plot with adjustments for missing/imputed values. Missing values in the plotted variables may be represented by a point above the corresponding coordinate axis to prevent disconnected lines. In addition, observations with missing/imputed values in selected variables may be highlighted.

Usage

```
TKRparcoordMiss(x, delimiter = NULL, highlight = NULL,
  selection = c("any", "all"), plotvars = NULL,
  plotNA = TRUE,
  col = c("skyblue", "red", "skyblue4", "red4", "orange", "orange4"),
  alpha = NULL, hscale = NULL, vscale = 1,
  TKRpar = list(), ...)
```

Arguments

x	a matrix or data.frame.
delimiter	a character-vector to distinguish between variables and imputation-indices for imputed variables (therefore, x needs to have <code>colnames</code>). If given, it is used to determine the corresponding imputation-index for any imputed variable (a logical-vector indicating which values of the variable have been imputed). If such imputation-indices are found, they are used for highlighting and the colors are adjusted according to the given colors for imputed variables (see <code>col</code>).
highlight	a vector giving the variables to be used for highlighting. If NULL (the default), all variables are used for highlighting.
selection	the selection method for highlighting missing/imputed values in multiple highlight variables. Possible values are "any" (highlighting of missing/imputed values in <i>any</i> of the highlight variables) and "all" (highlighting of missing/imputed values in <i>all</i> of the highlight variables).
plotvars	a vector giving the variables to be plotted. If NULL (the default), all variables are plotted.
col	if plotNA is TRUE, a vector of length six giving the colors to be used for observations with different combinations of observed and missing/imputed values in the plot variables and highlight variables (vectors of length one or two are recycled). Otherwise, a vector of length two giving the colors for non-highlighted and highlighted observations (if a single color is supplied, it is used for both).
plotNA	a logical indicating whether missing values in the plot variables should be represented by a point above the corresponding coordinate axis to prevent disconnected lines.
alpha	a numeric value between 0 and 1 giving the level of transparency of the colors, or NULL. This can be used to prevent overplotting.
...	for parcoordMiss, further graphical parameters to be passed down (see <code>par</code>). For TKRparcoordMiss, further arguments to be passed to parcoordMiss.
hscale	horizontal scale factor for plot to be embedded in a <i>Tcl/Tk</i> window (see 'Details'). The default value depends on the number of variables.

<code>vscale</code>	vertical scale factor for the plot to be embedded in a <i>Tcl/Tk</i> window (see ‘Details’).
<code>TKRpar</code>	a list of graphical parameters to be set for the plot to be embedded in a <i>Tcl/Tk</i> window (see ‘Details’ and par).

Details

In parallel coordinate plots, the variables are represented by parallel axes. Each observation of the scaled data is shown as a line. Observations with missing/imputed values in selected variables may thereby be highlighted. However, plotting variables with missing values results in disconnected lines, making it impossible to trace the respective observations across the graph. As a remedy, missing values may be represented by a point above the corresponding coordinate axis, which is separated from the main plot by a small gap and a horizontal line, as determined by `plotNA`. Connected lines can then be drawn for all observations. Nevertheless, a caveat of this display is that it may draw attention away from the main relationships between the variables.

If `interactive` is `TRUE`, it is possible switch between this display and the standard display without the separate level for missing values by clicking in the top margin of the plot. In addition, the variables to be used for highlighting can be selected interactively. Observations with missing/imputed values in any or in all of the selected variables are highlighted (as determined by `selection`). A variable can be added to the selection by clicking on a coordinate axis. If a variable is already selected, clicking on its coordinate axis removes it from the selection. Clicking anywhere outside the plot region (except the top margin, if missing/imputed values exist) quits the interactive session.

`TKRparcoordMiss` behaves like `parcoordMiss`, but uses `tkrplot` to embed the plot in a *Tcl/Tk* window. This is useful if the number of variables is large, because scrollbars allow to move from one part of the plot to another.

Note

Some of the argument names and positions have changed with versions 1.3 and 1.4 due to extended functionality and for more consistency with other plot functions in `VIM`. For back compatibility, the arguments `colcomb` and `xaxlabels` can still be supplied to `...{}` and are handled correctly. Nevertheless, they are deprecated and no longer documented. Use `highlight` and `labels` instead.

Author(s)

Andreas Alfons, Matthias Templ, modifications by Bernd Prantner

References

- Wegman, E. J. (1990) Hyperdimensional data analysis using parallel coordinates. *Journal of the American Statistical Association* **85** (411), 664–675.
- M. Templ, A. Alfons, P. Filzmoser (2012) Exploring incomplete data using visualization tools. *Journal of Advances in Data Analysis and Classification*, Online first. DOI: 10.1007/s11634-011-0102-y.

See Also

[pbox](#)

Examples

```
data(chorizonDL, package = "VIM")
## for missing values
parcoordMiss(chorizonDL[,c(15,101:110)],
  plotvars=2:11, interactive = FALSE)
legend("top", col = c("skyblue", "red"), lwd = c(1,1),
  legend = c("observed in Bi", "missing in Bi"))

## for imputed values
parcoordMiss(kNN(chorizonDL[,c(15,101:110)]), delimiter = "_imp" ,
  plotvars=2:11, interactive = FALSE)
legend("top", col = c("skyblue", "orange"), lwd = c(1,1),
  legend = c("observed in Bi", "imputed in Bi"))
```

TKRpbox

Parallel boxplots with information about missing/imputed values

Description

Boxplot of one variable of interest plus information about missing/imputed values in other variables.

Usage

```
TKRpbox(x, pos = 1, delimiter = NULL, hscale = NULL,
  vscale = 1, TKRpar = list(), ...)
```

Arguments

x	a vector, matrix or data.frame.
delimiter	a character-vector to distinguish between variables and imputation-indices for imputed variables (therefore, x needs to have colnames). If given, it is used to determine the corresponding imputation-index for any imputed variable (a logical-vector indicating which values of the variable have been imputed). If such imputation-indices are found, they are used for highlighting and the colors are adjusted according to the given colors for imputed variables (see col).
pos	a numeric value giving the index of the variable of interest. Additional variables in x are used for grouping according to missingness/number of imputed missings.
...	for pbox, further arguments and graphical parameters to be passed to boxplot and other functions. For TKRpbox, further arguments to be passed to pbox.
hscale	horizontal scale factor for plot to be embedded in a <i>Tcl/Tk</i> window (see ‘Details’). The default value depends on the number of boxes to be drawn.
vscale	vertical scale factor for the plot to be embedded in a <i>Tcl/Tk</i> window (see ‘Details’).
TKRpar	a list of graphical parameters to be set for the plot to be embedded in a <i>Tcl/Tk</i> window (see ‘Details’ and par).

Details

This plot consists of several boxplots. First, a standard boxplot of the variable of interest is produced. Second, boxplots grouped by observed and missing/imputed values according to `selection` are produced for the variable of interest.

Additionally, the frequencies of the missing/imputed values can be represented by numbers. If so, the first line corresponds to the observed values of the variable of interest and their distribution in the different groups, the second line to the missing/imputed values.

If `interactive=TRUE`, clicking in the left margin of the plot results in switching to the previous variable and clicking in the right margin results in switching to the next variable. Clicking anywhere else on the graphics device quits the interactive session.

TKRpbox behaves like `pbox` with `selection="none"`, but uses `tkrplot` to embed the plot in a *Tcl/Tk* window. This is useful for drawing a large number of parallel boxes, because scrollbars allow to move from one part of the plot to another.

Value

a list as returned by `boxplot`.

Note

Some of the argument names and positions have changed with version 1.3 due to extended functionality and for more consistency with other plot functions in VIM. For back compatibility, the arguments names and `cex.text` can still be supplied to `...{}` and are handled correctly. Nevertheless, they are deprecated and no longer documented. Use `labels` and `cex.numbers` instead.

Author(s)

Andreas Alfons, Matthias Templ, modifications by Bernd Prantner

References

M. Templ, A. Alfons, P. Filzmoser (2012) Exploring incomplete data using visualization tools. *Journal of Advances in Data Analysis and Classification*, Online first. DOI: 10.1007/s11634-011-0102-y.

See Also

[parcoordMiss](#)

Examples

```
data(chorizonDL, package = "VIM")
## for missing values
pbox(log(chorizonDL[, c(4,5,8,10,11,16:17,19,25,29,37,38,40)]))

## for imputed values
pbox(kNN(log(chorizonDL[, c(4,8,10,11,17,19,25,29,37,38,40)])),
      delimiter = "_imp")
```

TKRscattmatrixMiss	<i>Scatterplot matrix with information about missing/imputed values</i>
--------------------	---

Description

Scatterplot matrix in which observations with missing/imputed values in certain variables are highlighted.

Usage

```
TKRscattmatrixMiss(x, delimiter = NULL, highlight = NULL,
  selection = c("any", "all"), plotvars = NULL,
  col = c("skyblue", "red", "orange"), alpha = NULL,
  hscale = NULL, vscale = NULL, TKRpar = list(), ...)
```

Arguments

x	a matrix or data.frame.
delimiter	a character-vector to distinguish between variables and imputation-indices for imputed variables (therefore, x needs to have colnames). If given, it is used to determine the corresponding imputation-index for any imputed variable (a logical-vector indicating which values of the variable have been imputed). If such imputation-indices are found, they are used for highlighting and the colors are adjusted according to the given colors for imputed variables (see col).
highlight	a vector giving the variables to be used for highlighting. If NULL (the default), all variables are used for highlighting.
selection	the selection method for highlighting missing/imputed values in multiple highlight variables. Possible values are "any" (highlighting of missing/imputed values in <i>any</i> of the highlight variables) and "all" (highlighting of missing/imputed values in <i>all</i> of the highlight variables).
plotvars	a vector giving the variables to be plotted. If NULL (the default), all variables are plotted.
col	a vector of length three giving the colors to be used in the plot. The second/third color will be used for highlighting missing/imputed values.
alpha	a numeric value between 0 and 1 giving the level of transparency of the colors, or NULL. This can be used to prevent overplotting.
...	for scattmatrixMiss, further arguments and graphical parameters to be passed to pairsVIM . <code>par("oma")</code> will be set appropriately unless supplied (see par). For TKRscattmatrixMiss, further arguments to be passed to scattmatrixMiss.
hscale	horizontal scale factor for plot to be embedded in a <i>Tcl/Tk</i> window (see ‘Details’). The default value depends on the number of variables.
vscale	vertical scale factor for the plot to be embedded in a <i>Tcl/Tk</i> window (see ‘Details’). The default value depends on the number of variables.
TKRpar	a list of graphical parameters to be set for the plot to be embedded in a <i>Tcl/Tk</i> window (see ‘Details’ and par).

Details

scattmatrixMiss uses [pairsVIM](#) with a panel function that allows highlighting of missing/imputed values.

If `interactive=TRUE`, the variables to be used for highlighting can be selected interactively. Observations with missing/imputed values in any or in all of the selected variables are highlighted (as determined by selection). A variable can be added to the selection by clicking in a diagonal panel. If a variable is already selected, clicking on the corresponding diagonal panel removes it from the selection. Clicking anywhere else quits the interactive session.

The graphical parameter `oma` will be set unless supplied as an argument.

TKRscattmatrixMiss behaves like `scattmatrixMiss`, but uses [tkrplot](#) to embed the plot in a *Tcl/Tk* window. This is useful if the number of variables is large, because scrollbars allow to move from one part of the plot to another.

Note

Some of the argument names and positions have changed with version 1.3 due to a re-implementation and for more consistency with other plot functions in VIM. For back compatibility, the argument `colcomb` can still be supplied to `...{}` and is handled correctly. Nevertheless, it is deprecated and no longer documented. Use `highlight` instead. The arguments `smooth`, `reg.line` and `legend.plot` are no longer used and ignored if supplied.

Author(s)

Andreas Alfons, Matthias Templ, modifications by Bernd Prantner

References

M. Templ, A. Alfons, P. Filzmoser (2012) Exploring incomplete data using visualization tools. *Journal of Advances in Data Analysis and Classification*, Online first. DOI: 10.1007/s11634-011-0102-y.

See Also

[pairsVIM](#), [marginmatrix](#)

Examples

```
data(sleep, package = "VIM")
## for missing values
x <- sleep[, 1:5]
x[,c(1,2,4)] <- log10(x[,c(1,2,4)])
scattmatrixMiss(x, highlight = "Dream")

## for imputed values
x_imp <- kNN(sleep[, 1:5])
x_imp[,c(1,2,4)] <- log10(x_imp[,c(1,2,4)])
scattmatrixMiss(x_imp, delimiter = "_imp", highlight = "Dream")
```

VIMGUI

GUI for Visualization and Imputation of Missing Values

Description

Graphical user interface for visualization and imputation of missing values.

Usage

```
VIMGUI(startupObject = NULL)
```

Arguments

startupObject Object loaded at the start of the GUI

Details

Details about handling survey objects follow soon.

Author(s)

Daniel Schopfhauser

vmGUImenu

GUI for Visualization and Imputation of Missing Values

Description

Graphical user interface for visualization and imputation of missing values.

Usage

```
vmGUImenu()
```

Details

The *Data* menu allows to select a data set from the workspace or load data into the workspace from RData files. Furthermore, it can be used to transform variables, which are then appended to the data set in use. Commonly used transformations in official statistics are available, e.g., the Box-Cox transformation and the log-transformation as an important special case of the Box-Cox transformation. In addition, several other transformations that are frequently used for compositional data are implemented. Background maps and coordinates for spatial data can be selected in the data menu as well.

After a data set was chosen, variables can be selected in the main menu, along with a method for scaling. An important feature is that the variables will be used in the same order as they were selected, which is especially useful for parallel coordinate plots. Variables for highlighting are

distinguished from the plot variables and can be selected separately. For more than one variable chosen for highlighting, it is possible to select whether observations with missing values in any or in all of these variables should be highlighted.

A plot method can be selected from the *Visualization* menu. Note that plots that are not applicable to the selected variables are disabled, for example, if only one plot variable is selected, multivariate plots cannot be chosen.

The *Imputation* menu offers robust imputation methods to impute variables of the data set.

The *Diagnostics* menu is similar to the *Visualization* menu, but is designed to verify the results after the imputation of missing values.

Last, but not least, the *Options* menu allows to set the colors, alpha channel and the delimiter for imputed variables to be used in the plots. In addition, it contains an option to embed multivariate plots in Tcl/Tk windows. This is useful if the number of observations and/or variables is large, because scrollbars allow to move from one part of the plot to another.

Internal information regarding the VIM GUI is stored in the environment `vmGUIenvir`.

Author(s)

Andreas Alfons, based on an initial design by Matthias Templ, modifications by Bernd Prantner

References

M. Templ, A. Alfons, P. Filzmoser (2012) Exploring incomplete data using visualization tools. *Journal of Advances in Data Analysis and Classification*, Online first. DOI: 10.1007/s11634-011-0102-y.

Index

*Topic **hplot**

- TKRaggr, [2](#)
- TKRmarginmatrix, [4](#)
- TKRmatrixplot, [6](#)
- TKRparcoordMiss, [7](#)
- TKRpbox, [10](#)
- TKRscattmatrixMiss, [12](#)
- VIMGUI, [14](#)
- vmGUImenu, [14](#)

*Topic **multivariate**

- VIMGUI, [14](#)
- vmGUImenu, [14](#)

axis, [6](#)

boxplot, [10](#), [11](#)

colnames, [2](#), [4](#), [6](#), [8](#), [10](#), [12](#)

marginmatrix, [13](#)

marginplot, [5](#)

pairsVIM, [5](#), [12](#), [13](#)

par, [2](#), [5](#), [6](#), [8–10](#), [12](#)

parcoordMiss, [11](#)

pbox, [9](#)

plot.aggr, [2](#)

plot.window, [6](#)

print.aggr, [4](#)

scattmatrixMiss, [5](#)

summary.aggr, [4](#)

title, [6](#)

TKRaggr, [2](#)

TKRmarginmatrix, [4](#)

TKRmatrixplot, [6](#)

TKRparcoordMiss, [7](#)

TKRpbox, [10](#)

tkrplot, [3](#), [5](#), [7](#), [9](#), [11](#), [13](#)

TKRscattmatrixMiss, [12](#)

VIMGUI, [14](#)

vmGUIenvir (vmGUImenu), [14](#)

vmGUImenu, [14](#)