

# Package ‘VIM’

April 19, 2009

**Version** 1.3

**Date** 2009-04-17

**Title** Visualization and Imputation of Missing Values

**Author** Matthias Templ, Andreas Alfons

**Maintainer** Matthias Templ <templ@tuwien.ac.at>

**Depends** car, graphics, grDevices, robustbase, stats, tcltk, tkrplot, sp, utils

**Imports** car, grDevices, robustbase, stats, tcltk, sp, utils

**Description** This package introduces new tools for the visualization of missing values in R, which can be used for exploring the data and the structure of the missing values. Depending on this structure, they may help to identify the mechanism generating the missings. A graphical user interface allows an easy handling of the implemented plot methods.

**License** GPL (>= 2)

**URL** <http://cran.r-project.org/package=VIM>

**Repository** CRAN

**Date/Publication** 2009-04-19 18:07:22

## R topics documented:

VIM-package . . . . .	2
aggr . . . . .	3
alphablend . . . . .	5
barMiss . . . . .	6
bgmap . . . . .	8
chorizonDL . . . . .	9
colormapMiss . . . . .	13
count . . . . .	15
growdotMiss . . . . .	15
histMiss . . . . .	17

initialise . . . . .	20
kola.background . . . . .	20
mapMiss . . . . .	21
marginmatrix . . . . .	22
marginplot . . . . .	23
matrixplot . . . . .	25
pairsVIM . . . . .	27
parcoordMiss . . . . .	29
pbox . . . . .	31
prepare . . . . .	33
print.aggr . . . . .	34
print.summary.aggr . . . . .	35
rugNA . . . . .	36
scattJitt . . . . .	37
scattmatrixMiss . . . . .	38
scattMiss . . . . .	40
sleep . . . . .	42
spineMiss . . . . .	43
summary.aggr . . . . .	45
tao . . . . .	46
vmGUImenu . . . . .	47
<b>Index</b>	<b>48</b>

**Description**

This package introduces new tools for the visualization of missing values in R, which can be used for exploring the data and the structure of the missing values. Depending on this structure, they may help to identify the mechanism generating the missing values. This knowledge is necessary for selecting an appropriate imputation method in order to reliably estimate the missing values. Thus the visualization tools should be applied before imputation.

Detecting missing values mechanisms is usually done by statistical tests or models. Visualization of missing values can support the test decision, but also reveals more details about the data structure. Most notably, statistical requirements for a test can be checked graphically, and problems like outliers or skewed data distributions can be discovered. Furthermore, the included plot methods may also be able to detect missing values mechanisms in the first place.

A graphical user interface allows an easy handling of the plot methods. In addition, VIM can be used for data from essentially any field.

Robust imputation methods and diagnostic tools will be included in future versions.

**Details**

Package: VIM  
 Type: Package  
 Version: 1.3  
 Date: 2009-04-17  
 License: GPL (>= 2)

**Author(s)**

Matthias Templ, Andreas Alfons  
 Maintainer: Matthias Templ <templ@statistik.tuwien.ac.at>

---

 aggr

*Aggregations for missing values*


---

**Description**

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

**Usage**

```
aggr(x, plot = TRUE, ...)
```

```
## S3 method for class 'aggr':
plot(x, col = c("skyblue", "red"), bars = TRUE,
     numbers = FALSE, prop = TRUE, border = par("fg"),
     sortVars = FALSE, sortCombs = TRUE, ylabs = NULL,
     axes = TRUE, labels = axes, cex.lab = 1.2,
     cex.axis = par("cex"), cex.numbers = par("cex"),
     gap = 4, ...)
```

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

**Arguments**

x	a vector, matrix or data.frame.
plot	a logical indicating whether the results should be plotted (the default is TRUE).
col	a vector of length two giving the colors to be used for observed and missing data. If only one color is supplied, it is used for missing data and observed data is transparent.
bars	a logical indicating whether a small barplot for the frequencies of the different combinations should be drawn.

<code>numbers</code>	a logical indicating whether the proportion or frequencies of the different combinations should be represented by numbers.
<code>prop</code>	a logical indicating whether the proportion of missing values and combinations should be used rather than the total amount.
<code>border</code>	the color to be used for the border of the bars and rectangles. Use <code>border=NA</code> to omit borders.
<code>sortVars</code>	a logical indicating whether the variables should be sorted by the number of missing values.
<code>sortCombs</code>	a logical indicating whether the combinations should be sorted by the frequency of occurrence.
<code>ylabs</code>	a character vector of length two giving the y-axis labels for the two plots.
<code>axes</code>	a logical indicating whether axes should be drawn.
<code>labels</code>	either a logical indicating whether labels should be plotted on the x-axis, or a character vector giving the labels.
<code>cex.lab</code>	the character expansion factor to be used for the axis labels.
<code>cex.axis</code>	the character expansion factor to be used for the axis annotation.
<code>cex.numbers</code>	the character expansion factor to be used for the proportion or frequencies of the different combinations
<code>gap</code>	a numeric value giving the distance between the two plots in margin lines.
<code>...</code>	for <code>aggr</code> and <code>TKRaggr</code> , further arguments and graphical parameters to be passed to <code>plot.aggr</code> . For <code>plot.aggr</code> , further graphical parameters to be passed down. <code>par("oma")</code> will be set appropriately unless supplied (see <code>par</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 combinations.
<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 <code>par</code> ).

## Details

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

The barplot on the left hand side shows the number of missing values in each variable. In the *aggregation plot* on the right hand side, all existing combinations of missing and non-missing values in the observations are visualized. Available and missing data are color coded as given by `col`. Additionally, the proportions or frequencies of the different combinations can be represented by a small bar plot and numbers. Variables may be sorted by the number of missing values and combinations by the frequency of occurrence to give more power to finding the structure of missing values.

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:

<code>x</code>	the data used.
<code>combinations</code>	a character vector representing the combinations of variables.
<code>count</code>	the frequencies of these combinations.
<code>percent</code>	the percentage of these combinations.
<code>missings</code>	a <code>data.frame</code> containing the amount of missing values in each variable.
<code>tabcomb</code>	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

**See Also**

`print.aggr`, `summary.aggr`

**Examples**

```
data(sleep, package="VIM")
a <- aggr(sleep)
a
summary(a)
```

---

alphablend

*Alphablending for colors*

---

**Description**

Convert colors to semitransparent colors.

**Usage**

```
alphablend(col, alpha = NULL, bg = NULL)
```

**Arguments**

col	a vector specifying colors.
alpha	a numeric vector containing the alpha values (between 0 and 1).
bg	the background color to be used for alphablending. This can be used as a workaround for graphics devices that do not support semitransparent colors.

**Value**

a vector containing the semitransparent colors.

**Author(s)**

Andreas Alfons

**Examples**

```
alphablend("red", 0.6)
```

---

barMiss

*Barplot with information about missing values*


---

**Description**

Barplot with highlighting of missing values in other variables by splitting each bar into two parts. Additionally, information about missing values in the variable of interest is shown on the right hand side.

**Usage**

```
barMiss(x, pos = 1, selection = c("any", "all"),
        col = c("skyblue", "red", "skyblue4", "red4"),
        border = NULL, main = NULL, sub = NULL,
        xlab = NULL, ylab = NULL, axes = TRUE,
        labels = axes, only.miss = TRUE,
        miss.labels = axes, interactive = TRUE, ...)
```

**Arguments**

x	a vector, matrix or data.frame.
pos	a numeric value giving the index of the variable of interest. Additional variables in x are used for highlighting.
selection	the selection method for highlighting missing values in multiple additional variables. Possible values are "any" (highlighting of missing values in <i>any</i> of the additional variables) and "all" (highlighting of missing values in <i>all</i> of the additional variables).

<code>col</code>	a vector of length four giving the colors to be used. If only one color is supplied, the bars are transparent and the supplied color is used for highlighting. Else if two colors are supplied, they are recycled.
<code>border</code>	the color to be used for the border of the bars. Use <code>border=NA</code> to omit borders.
<code>main, sub</code>	main and sub title.
<code>xlab, ylab</code>	axis labels.
<code>axes</code>	a logical indicating whether axes should be drawn on the plot.
<code>labels</code>	either a logical indicating whether labels should be plotted below each bar, or a character vector giving the labels.
<code>only.miss</code>	logical; if <code>TRUE</code> , the missing values in the variable of interest are visualized by a single bar. Otherwise, a small barplot is drawn on the right hand side (see ‘Details’).
<code>miss.labels</code>	either a logical indicating whether label(s) should be plotted below the bar(s) on the right hand side, or a character string or vector giving the label(s) (see ‘Details’).
<code>interactive</code>	a logical indicating whether variables can be switched interactively (see ‘Details’).
<code>...</code>	further graphical parameters to be passed to <code>title</code> and <code>axis</code> .

### Details

If more than one variable is supplied, the bars for the variable of interest are split according to missingness in the additional variables.

If `only.miss=TRUE`, the missing values in the variable of interest are visualized by one bar on the right hand side. If additional variables are supplied, this bar is again split into two parts according to missingness in the additional variables.

Otherwise, a small barplot consisting of two bars is drawn on the right hand side. The first bar corresponds to observed values in the variable of interest and the second bar to missing values. Since these two bars are not on the same scale as the main barplot, a second y-axis is plotted on the right (if `axes=TRUE`). Each of the two bars are again split into two parts according to missingness in the additional variables. Note that this display does not make sense if only one variable is supplied, therefore `only.miss` is ignored in that case.

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. When switching to a continuous variable, a histogram is plotted rather than a barplot.

### Value

a numeric vector giving the coordinates of the midpoints of the bars.

### 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 `axisnames`, `names.arg` and `names.miss` can still be supplied to `...` and are handled correctly. Nevertheless, they are deprecated and no longer documented. Use `labels` and `miss.labels` instead.

**Author(s)**

Andreas Alfons

**See Also**

[spineMiss](#), [histMiss](#)

**Examples**

```
data(sleep, package = "VIM")
x <- sleep[, c("Exp", "Sleep")]
barMiss(x)
barMiss(x, only.miss = FALSE)
```

---

bgmap

*Background map*

---

**Description**

Plot a background map.

**Usage**

```
bgmap(map, add=FALSE, ...)
```

**Arguments**

<code>map</code>	either a matrix or <code>data.frame</code> with two columns, a list with components <code>x</code> and <code>y</code> , or an object of any class that can be used for maps and provides its own plot method (e.g., <code>"SpatialPolygons"</code> from package <code>sp</code> ). A list of the previously mentioned types can also be provided.
<code>add</code>	a logical indicating whether <code>map</code> should be added to an already existing plot (the default is <code>FALSE</code> ).
<code>...</code>	further arguments and graphical parameters to be passed to <code>plot</code> and/or <code>lines</code> .

**Author(s)**

Andreas Alfons

**See Also**

[growdotMiss](#), [mapMiss](#)

## Examples

```
data(kola.background, package = "VIM")
bgmap(kola.background)
```

---

chorizonDL

*C-horizon of the Kola data with missing values*

---

## Description

This data set is the same as the `chorizon` data set in package `mvoutlier`, except that values below the detection limit are coded as NA.

## Usage

```
data(chorizonDL)
```

## Format

A data frame with 606 observations on the following 110 variables.

**\*ID** a numeric vector  
**XCOO** a numeric vector  
**YCOO** a numeric vector  
**Ag** a numeric vector  
**Ag\_INAA** a numeric vector  
**Al** a numeric vector  
**Al2O3** a numeric vector  
**As** a numeric vector  
**As\_INAA** a numeric vector  
**Au\_INAA** a numeric vector  
**B** a numeric vector  
**Ba** a numeric vector  
**Ba\_INAA** a numeric vector  
**Be** a numeric vector  
**Bi** a numeric vector  
**Br\_IC** a numeric vector  
**Br\_INAA** a numeric vector  
**Ca** a numeric vector  
**Ca\_INAA** a numeric vector  
**CaO** a numeric vector  
**Cd** a numeric vector

**Ce\_INAA** a numeric vector  
**Cl\_IC** a numeric vector  
**Co** a numeric vector  
**Co\_INAA** a numeric vector  
**EC** a numeric vector  
**Cr** a numeric vector  
**Cr\_INAA** a numeric vector  
**Cs\_INAA** a numeric vector  
**Cu** a numeric vector  
**Eu\_INAA** a numeric vector  
**F\_IC** a numeric vector  
**Fe** a numeric vector  
**Fe\_INAA** a numeric vector  
**Fe2O3** a numeric vector  
**Hf\_INAA** a numeric vector  
**Hg** a numeric vector  
**Hg\_INAA** a numeric vector  
**Ir\_INAA** a numeric vector  
**K** a numeric vector  
**K2O** a numeric vector  
**La** a numeric vector  
**La\_INAA** a numeric vector  
**Li** a numeric vector  
**LOI** a numeric vector  
**Lu\_INAA** a numeric vector  
**wt\_INAA** a numeric vector  
**Mg** a numeric vector  
**MgO** a numeric vector  
**Mn** a numeric vector  
**MnO** a numeric vector  
**Mo** a numeric vector  
**Mo\_INAA** a numeric vector  
**Na** a numeric vector  
**Na\_INAA** a numeric vector  
**Na2O** a numeric vector  
**Nd\_INAA** a numeric vector  
**Ni** a numeric vector

**Ni\_INAA** a numeric vector  
**NO3\_IC** a numeric vector  
**P** a numeric vector  
**P2O5** a numeric vector  
**Pb** a numeric vector  
**pH** a numeric vector  
**PO4\_IC** a numeric vector  
**Rb** a numeric vector  
**S** a numeric vector  
**Sb** a numeric vector  
**Sb\_INAA** a numeric vector  
**Sc** a numeric vector  
**Sc\_INAA** a numeric vector  
**Se** a numeric vector  
**Se\_INAA** a numeric vector  
**Si** a numeric vector  
**SiO2** a numeric vector  
**Sm\_INAA** a numeric vector  
**Sn\_INAA** a numeric vector  
**SO4\_IC** a numeric vector  
**Sr** a numeric vector  
**Sr\_INAA** a numeric vector  
**SUM\_XRF** a numeric vector  
**Ta\_INAA** a numeric vector  
**Tb\_INAA** a numeric vector  
**Te** a numeric vector  
**Th** a numeric vector  
**Th\_INAA** a numeric vector  
**Ti** a numeric vector  
**TiO2** a numeric vector  
**U\_INAA** a numeric vector  
**V** a numeric vector  
**W\_INAA** a numeric vector  
**Y** a numeric vector  
**Yb\_INAA** a numeric vector  
**Zn** a numeric vector  
**Zn\_INAA** a numeric vector

**ELEV** a numeric vector  
**\*COUN** a numeric vector  
**\*ASP** a numeric vector  
**TOPC** a numeric vector  
**LITO** a numeric vector  
**Al\_XRF** a numeric vector  
**Ca\_XRF** a numeric vector  
**Fe\_XRF** a numeric vector  
**K\_XRF** a numeric vector  
**Mg\_XRF** a numeric vector  
**Mn\_XRF** a numeric vector  
**Na\_XRF** a numeric vector  
**P\_XRF** a numeric vector  
**Si\_XRF** a numeric vector  
**Ti\_XRF** a numeric vector

### Note

For a more detailed description of this data set, see [chorizon](#) in package `mvoutlier`.

### Source

Kola Project (1993-1998)

### References

Reimann, C., Filzmoser, P., Garrett, R.G. and Dutter, R. (2008) *Statistical Data Analysis Explained: Applied Environmental Statistics with R*. Wiley, 2008.

### See Also

[chorizon](#)

### Examples

```
data(chorizonDL, package = "VIM")
summary(chorizonDL)
```

---

colormapMiss	<i>Colored map with information about missing values</i>
--------------	--

---

### Description

Colored map in which the proportion or amount of missing values in each region is coded according to a continuous *rgb* color scheme.

### Usage

```
colormapMiss(x, region, map, prop = TRUE,
             polysRegion = 1:length(x), col = "red",
             coords = NULL, numbers = TRUE, digits = 2,
             cex.numbers = 0.8, col.numbers = par("fg"),
             legend = TRUE, interactive = TRUE, ...)

colormapMissLegend(xleft, ybottom, xright, ytop, range,
                  prop = FALSE, col = "red", n = 1000,
                  horizontal = TRUE, digits = 2,
                  cex.numbers = 0.8, col.numbers = par("fg"),
                  ...)
```

### Arguments

<code>x</code>	a vector.
<code>region</code>	a vector or factor of the same length as <code>x</code> giving the regions.
<code>map</code>	an object of any class that contains polygons and provides its own plot method (e.g., "SpatialPolygons" from package <code>sp</code> ).
<code>prop</code>	a logical indicating whether the proportion of missing values should be used rather than the total amount. If <code>TRUE</code> , the full color value corresponds to the smallest integer not less than the largest proportion of missing values.
<code>polysRegion</code>	a numeric vector specifying the region that each polygon belongs to.
<code>col</code>	the color to be used in the continuous <i>rgb</i> color scheme.
<code>coords</code>	a matrix or <code>data.frame</code> with two columns giving the coordinates for the labels.
<code>numbers</code>	a logical indicating whether the corresponding proportions or numbers of missing values should be used as labels for the regions.
<code>digits</code>	the number of digits to be used in the labels (if <code>prop=TRUE</code> ).
<code>cex.numbers</code>	the character expansion factor to be used for the labels.
<code>col.numbers</code>	the color to be used for the labels.
<code>legend</code>	a logical indicating whether a legend should be plotted.
<code>interactive</code>	a logical indicating whether more detailed information about missing values should be displayed interactively (see 'Details').

<code>xleft</code>	left x position of the legend.
<code>ybottom</code>	bottom y position of the legend.
<code>xright</code>	right x position of the legend.
<code>ytopy</code>	top y position of the legend.
<code>range</code>	the range of the proportion or amount of missing values to be used in the legend.
<code>n</code>	the number of rectangles to be drawn in the legend. A reasonably large number makes it appear continuously.
<code>horizontal</code>	a logical indicating whether the legend should be drawn horizontally or vertically.
<code>...</code>	further arguments to be passed to <code>plot</code> .

### Details

Regions with higher proportions or numbers of missing values in `x` (as given by `prop`) receive a higher portion of `col`. Additionally, the proportions or numbers can be shown as labels in the regions.

If `interactive=TRUE`, clicking in a region displays more detailed information about missing values on the R console. Clicking outside the borders quits the interactive session.

### Value

`colormapMiss` returns a list with the following components:

<code>nmiss</code>	a vector containing the proportion or number of missing values in each region.
<code>colmax</code>	the amount or proportion of missing values corresponding to <code>col</code> .

### Note

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

### Author(s)

Andreas Alfons

### See Also

[growdotMiss](#), [mapMiss](#)

---

count	<i>Count number of infinite or missing values</i>
-------	---

---

**Description**

Count the number of infinite or missing values in a vector.

**Usage**

```
countInf(x)
```

```
countNA(x)
```

**Arguments**

`x` a vector.

**Value**

`countInf` returns the number of infinite values in `x`.

`countNA` returns the number of missing values in `x`.

**Author(s)**

Andreas Alfons

**Examples**

```
data(sleep, package="VIM")
countInf(log(sleep$Dream))
countNA(sleep$Dream)
```

---

growdotMiss	<i>Growing dot map with information about missing values</i>
-------------	--

---

**Description**

Map with dots whose sizes correspond to the values in a certain variable. Observations with missing values in additional variables are highlighted.

**Usage**

```
growdotMiss(x, coords, map, pos=1, selection = c("any", "all"),
            log = FALSE, col = c("skyblue", "red", "red4"),
            border = par("bg"), alpha = NULL, scale = NULL,
            size = NULL, exp = c(0, 0.95, 0.05),
            col.map = grey(0.5), legend = TRUE,
            legtitle = "Legend", cex.legtitle = par("cex"),
            cex.legtext = par("cex"), ncircles = 6, ndigits = 1,
            interactive = TRUE, ...)

bubbleMiss(...)
```

**Arguments**

<code>x</code>	a vector, matrix or <code>data.frame</code> .
<code>coords</code>	a matrix or <code>data.frame</code> with two columns giving the spatial coordinates of the observations.
<code>map</code>	a background map to be passed to <code>bgmap</code> .
<code>pos</code>	a numeric value giving the index of the variable determining the dot sizes.
<code>selection</code>	the selection method for highlighting missing values in multiple additional variables. Possible values are "any" (highlighting of missing values in <i>any</i> of the additional variables) and "all" (highlighting of missing values in <i>all</i> of the additional variables).
<code>log</code>	a logical indicating whether the variable given by <code>pos</code> should be log-transformed.
<code>col</code>	a vector of length three giving the colors to be used in the plot. The second and third color are used for highlighting. If only one color is supplied, it is used for the borders of non-highlighted dots and the surface area of highlighted dots.
<code>border</code>	the color to be used for the border of the growing dots. Use <code>border=NA</code> to omit borders. A vector of length three is also allowed, with the second and third color to be used for highlighted dots.
<code>alpha</code>	a numeric value between 0 and 1 giving the level of transparency of the colors, or <code>NULL</code> . This can be used to prevent overplotting.
<code>scale</code>	scaling factor of the map.
<code>size</code>	a vector of length two giving the sizes for the smallest and largest dots.
<code>exp</code>	a vector of length three giving the factors that define the shape of the exponential function (see 'Details').
<code>col.map</code>	the color to be used for the background map.
<code>legend</code>	a logical indicating whether a legend should be plotted.
<code>legtitle</code>	the title for the legend.
<code>cex.legtitle</code>	the character expansion factor to be used for the title of the legend.
<code>cex.legtext</code>	the character expansion factor to be used in the legend.
<code>ncircles</code>	the number of circles displayed in the legend.

<code>ndigits</code>	the number of digits displayed in the legend. Note that this is just a suggestion (see <a href="#">format</a> ).
<code>interactive</code>	a logical indicating whether information about certain observations can be displayed interactively (see ‘Details’).
<code>...</code>	for <code>growdotMiss</code> , further arguments and graphical parameters to be passed to <a href="#">bgmap</a> . For <code>bubbleMiss</code> , the arguments to be passed to <code>growdotMiss</code> .

### Details

The smallest dots correspond to the 10% quantile and the largest dots to the 99% quantile. In between, the dots grow exponentially, with `exp` defining the shape of the exponential function. Missings in the variable of interest will be drawn as rectangles.

If `interactive=TRUE`, detailed information for an observation can be printed on the console by clicking on the corresponding point. Clicking in a region that does not contain any points quits the interactive session.

### Note

The function was renamed to `growdotMiss` in version 1.3. `bubbleMiss` is a (deprecated) wrapper for `growdotMiss` for back compatibility with older versions. However, due to extended functionality, some of the argument positions have changed.

The code is based on [bubbleFIN](#) from package `StatDA`.

### Author(s)

Andreas Alfons

### See Also

[bgmap](#), [mapMiss](#), [colormapMiss](#)

### Examples

```
data(chorizonDL, package = "VIM")
data(kola.background, package = "VIM")
x <- chorizonDL[, c("Ca", "As", "Bi")]
coo <- chorizonDL[, c("XCOO", "YCOO")]
growdotMiss(x, coo, kola.background, border = "white")
```

---

histMiss

*Histogram with information about missing values*

---

### Description

Histogram with highlighting of missing values in other variables by splitting each bin into two parts. Additionally, information about missing values in the variable of interest is shown on the right hand side.

**Usage**

```
histMiss(x, pos = 1, selection = c("any", "all"),
         breaks = "Sturges", right = TRUE,
         col = c("skyblue", "red", "skyblue4", "red4"),
         border = NULL, main = NULL, sub = NULL,
         xlab = NULL, ylab = NULL, axes = TRUE,
         only.miss = TRUE, miss.labels = axes,
         interactive = TRUE, ...)
```

**Arguments**

<code>x</code>	a vector, matrix or <code>data.frame</code> .
<code>pos</code>	a numeric value giving the index of the variable of interest. Additional variables in <code>x</code> are used for highlighting.
<code>selection</code>	the selection method for highlighting missing values in multiple additional variables. Possible values are "any" (highlighting of missing values in <i>any</i> of the additional variables) and "all" (highlighting of missing values in <i>all</i> of the additional variables).
<code>breaks</code>	either a character string naming an algorithm to compute the breakpoints (see <a href="#">hist</a> ), or a numeric value giving the number of cells.
<code>right</code>	logical; if <code>TRUE</code> , the histogram cells are right-closed (left-open) intervals.
<code>col</code>	a vector of length four giving the colors to be used. If only one color is supplied, the bars are transparent and the supplied color is used for highlighting. Else if two colors are supplied, they are recycled.
<code>border</code>	the color to be used for the border of the cells. Use <code>border=NA</code> to omit borders.
<code>main, sub</code>	main and sub title.
<code>xlab, ylab</code>	axis labels.
<code>axes</code>	a logical indicating whether axes should be drawn on the plot.
<code>only.miss</code>	logical; if <code>TRUE</code> , the missing values in the first variable are visualized by a single bar. Otherwise, a small barplot is drawn on the right hand side (see 'Details').
<code>miss.labels</code>	either a logical indicating whether label(s) should be plotted below the bar(s) on the right hand side, or a character string or vector giving the label(s) (see 'Details').
<code>interactive</code>	a logical indicating whether the variables can be switched interactively (see 'Details').
<code>...</code>	further graphical parameters to be passed to <a href="#">title</a> and <a href="#">axis</a> .

**Details**

If more than one variable is supplied, the bins for the variable of interest will be split according to missingness in the additional variables.

If `only.miss=TRUE`, the missing values in the variable of interest are visualized by one bar on the right hand side. If additional variables are supplied, this bar is again split into two parts according to missingness in the additional variables.

Otherwise, a small barplot consisting of two bars is drawn on the right hand side. The first bar corresponds to observed values in the variable of interest and the second bar to missing values. Since these two bars are not on the same scale as the main barplot, a second y-axis is plotted on the right (if `axes=TRUE`). Each of the two bars are again split into two parts according to missingness in the additional variables. Note that this display does not make sense if only one variable is supplied, therefore only `.miss` is ignored in that case.

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. When switching to a categorical variable, a barplot is produced rather than a histogram.

### Value

a list with the following components:

<code>breaks</code>	the breakpoints.
<code>counts</code>	the number of observations in each cell.
<code>missings</code>	the number of highlighted observations in each cell.
<code>mids</code>	the cell midpoints.

### 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 `axisnames` and `names.miss` can still be supplied to `. . .` and are handled correctly. Nevertheless, they are deprecated and no longer documented. Use `miss.labels` instead.

### Author(s)

Andreas Alfons, Matthias Templ

### See Also

[spineMiss](#), [barMiss](#)

### Examples

```
data(tao, package = "VIM")
x <- tao[, c("Air.Temp", "Humidity")]
histMiss(x)
histMiss(x, only.miss = FALSE)
```

`initialise`                      *Initialization of missing values*

---

**Description**

Rough estimation of missing values in a vector according to its type.

**Usage**

```
initialise(x)
```

**Arguments**

`x`                      a vector.

**Details**

Missing values are imputed with the mean for vectors of class "numeric", with the median for vectors of class "integer", and with the mode for vectors of class "factor". Hence, `x` should be prepared in the following way: assign class "numeric" to numeric vectors, assign class "integer" to ordinal vectors, and assign class "factor" to nominal or binary vectors.

**Value**

the initialized vector.

**Note**

The function is used internally by some imputation algorithms.

**Author(s)**

Matthias Templ, modifications by Andreas Alfons

---

`kola.background`                      *Background map for the Kola project data*

---

**Description**

Coordinates of the Kola background map.

**Usage**

```
data(kola.background)
```

**Source**

Kola Project (1993-1998)

**References**

Reimann, C., Filzmoser, P., Garrett, R.G. and Dutter, R. (2008) *Statistical Data Analysis Explained: Applied Environmental Statistics with R*. Wiley, 2008.

**Examples**

```
data(kola.background, package = "VIM")
bgmap(kola.background)
```

---

mapMiss

*Map with information about missing values*


---

**Description**

Map of observed and missing values.

**Usage**

```
mapMiss(x, coords, map, selection = c("any", "all"),
        col = c("skyblue", "red"), alpha = NULL,
        pch = c(19, 15), col.map = grey(0.5), legend = TRUE,
        interactive = TRUE, ...)
```

**Arguments**

x	a vector, matrix or <code>data.frame</code> .
coords	a <code>data.frame</code> or matrix with two columns giving the spatial coordinates of the observations.
map	a background map to be passed to <code>bgmap</code> .
selection	the selection method for displaying missing values in the map. Possible values are "any" (display missing values in <i>any</i> variable) and "all" (display missing values in <i>all</i> variables).
col	a vector of length two giving the colors to be used for observed and missing values. If a single plot character is supplied, it will be used for both.
alpha	a numeric value between 0 and 1 giving the level of transparency of the colors, or <code>NULL</code> . This can be used to prevent overplotting.
pch	a vector of length two giving the plot characters to be used for observed and missing values. If a single plot character is supplied, it will be used for both.
col.map	the color to be used for the background map.
legend	a logical indicating whether a legend should be plotted.
interactive	a logical indicating whether information about selected observations can be displayed interactively (see ‘Details’).
...	further graphical parameters to be passed to <code>bgmap</code> and <code>points</code> .

**Details**

If `interactive=TRUE`, detailed information for an observation can be printed on the console by clicking on the corresponding point. Clicking in a region that does not contain any points quits the interactive session.

**Author(s)**

Matthias Templ, Andreas Alfons

**See Also**

[bgmap](#), [bubbleMiss](#), [colormapMiss](#)

**Examples**

```
data(chorizonDL, package = "VIM")
data(kola.background, package = "VIM")
x <- chorizonDL[, c("As", "Bi")]
coo <- chorizonDL[, c("XCOO", "YCOO")]
mapMiss(x, coo, kola.background)
```

---

marginmatrix

*Marginplot Matrix*


---

**Description**

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

**Usage**

```
marginmatrix(x, col = c("skyblue", "red", "red4"),
             alpha = NULL, ...)
```

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

**Arguments**

`x` a matrix or `data.frame`.

`col` a vector of length three 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 color for the univariate scatterplots and boxplots for the missing values in one variable, and the third color for the frequency of missing 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 values in one variable, whereas the boxplots for the available data are transparent. Else if two colors are supplied, the second one is recycled.

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.
...	further arguments and graphical parameters to be passed to <code>pairsVIM</code> and <code>marginplot</code> . <code>par("oma")</code> will be set appropriately unless supplied (see <code>par</code> ).
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 <code>par</code> ).

### 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

### See Also

`marginplot`, `pairsVIM`, `scattmatrixMiss`

### Examples

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

---

marginplot

*Scatterplot with additional information in the margins*

---

### Description

In addition to a standard scatterplot, information about missing values is shown in the plot margins.

**Usage**

```
marginplot(x, col = c("skyblue", "red", "red4"),
           alpha = NULL, pch = c(1, 16), cex = par("cex"),
           numbers = TRUE, cex.numbers = par("cex"),
           xlim = NULL, ylim = NULL, main = NULL,
           sub = NULL, xlab = NULL, ylab = NULL,
           ann = par("ann"), axes = TRUE,
           frame.plot = axes, ...)
```

**Arguments**

<code>x</code>	a matrix or data.frame with two columns.
<code>col</code>	a vector of length three giving the colors to be used in the plot. The first color is used for the scatterplot and the boxplots for the available data, the second color for the univariate scatterplots and boxplots for the missing values in one variable, and the third color for the frequency of missing 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 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 <code>NULL</code> . This can be used to prevent overplotting.
<code>pch</code>	a vector of length two giving the plot symbols to be used for the scatterplot and the univariate scatterplots. If a single plot character is supplied, it is used for the scatterplot and the default value will be used for the univariate scatterplots (see ‘Details’).
<code>cex</code>	the character expansion factor to be used for the bivariate and univariate scatterplots.
<code>numbers</code>	a logical indicating whether the frequencies of missing values should be displayed in the lower left of the plot (see ‘Details’).
<code>cex.numbers</code>	the character expansion factor to be used for the frequencies of the missing values.
<code>xlim, ylim</code>	axis limits.
<code>main, sub</code>	main and sub title.
<code>xlab, ylab</code>	axis labels.
<code>ann</code>	a logical indicating whether plot annotation ( <code>main</code> , <code>sub</code> , <code>xlab</code> , <code>ylab</code> ) should be displayed.
<code>axes</code>	a logical indicating whether both axes should be drawn on the plot. Use graphical parameter <code>"xaxt"</code> or <code>"yaxt"</code> to suppress just one of the axes.
<code>frame.plot</code>	a logical indicating whether a box should be drawn around the plot.
<code>...</code>	further graphical parameters to be passed down (see <code>par</code> ).

## Details

Boxplots for available and missing data, as well as univariate scatterplots for missing values in one variable are shown in the plot margins.

Furthermore, the frequencies of the missing values can be displayed by a number (lower left of the plot). The number in the lower left corner is the number of observations that are missing in both 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 argument `cex.text` can still be supplied to `...` and is handled correctly. Nevertheless, it is deprecated and no longer documented. Use `cex.numbers` instead.

## Author(s)

Andreas Alfons, Matthias Templ

## See Also

[scattMiss](#)

## Examples

```
data(tao, package = "VIM")
marginplot(tao[,c("Air.Temp", "Humidity")])
data(chorizonDL, package = "VIM")
marginplot(log10(chorizonDL[,c("CaO", "Bi")]))
```

---

matrixplot

*Matrix plot*

---

## Description

Create a matrix plot, in which all cells of a data matrix are visualized by rectangles.

## Usage

```
matrixplot(x, sortby = NULL, col = "red",
           xlim = NULL, ylim = NULL, main = NULL,
           sub = NULL, xlab = NULL, ylab = NULL,
           axes = TRUE, labels = axes, xpd = NULL,
           interactive = TRUE, ...)

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

```
iimagMiss(x, sortby = NULL, col = "red", main = NULL,
          sub = NULL, xlab = NULL, ylab = NULL,
          xlim = NULL, ylim = NULL, axes = TRUE,
          xaxlabels = NULL, las = 3, interactive = TRUE,
          ...)
```

### Arguments

<code>x</code>	a <code>matrix</code> or <code>data.frame</code> .
<code>sortby</code>	a numeric or character value specifying the variable to sort the data matrix by, or <code>NULL</code> to plot without sorting.
<code>col</code>	the color to be used to indicate missing values.
<code>xlim, ylim</code>	axis limits.
<code>main, sub</code>	main and sub title.
<code>xlab, ylab</code>	axis labels.
<code>axes</code>	a logical indicating whether axes should be drawn on the plot.
<code>labels</code>	either a logical indicating whether labels should be plotted below each column, or a character vector giving the labels.
<code>xpd</code>	a logical indicating whether the rectangles should be allowed to go outside the plot region. If <code>NULL</code> , it defaults to <code>TRUE</code> unless axis limits are specified.
<code>interactive</code>	a logical indicating whether a variable to be used for sorting can be selected interactively (see ‘Details’).
<code>xaxlabels</code>	a character vector containing the labels for the columns. If <code>NULL</code> , the column names of <code>x</code> will be used.
<code>las</code>	the style of axis labels (see <a href="#">par</a> ).
<code>...</code>	for <code>matrixplot</code> and <code>iimagMiss</code> , further graphical parameters to be passed to <code>plot.window</code> , <code>title</code> and <code>axis</code> . 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 <code>Tcl/Tk</code> 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 <code>Tcl/Tk</code> 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 <code>Tcl/Tk</code> window (see ‘Details’ and <a href="#">par</a> ).

### Details

In a *matrix plot*, all cells of a data matrix are visualized by rectangles. The color specified by `col` is used for missing values, and a grey scale for the available data. To determine the grey level, the variables are first scaled to the interval

$$[0, 1]$$

by subtracting the minimum and dividing by the range. Hence, small values are assigned a light grey and high values a dark grey. Note that `-Inf` and `Inf` are always assigned white and black, respectively.

Additionally, the observations can be sorted by the magnitude of a selected variable. If `interactive=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 package `dprep`.

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

### Author(s)

Andreas Alfons, Matthias Templ

### Examples

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

---

pairsVIM

*Scatterplot Matrices*

---

### Description

Create a scatterplot matrix.

### Usage

```
pairsVIM(x, ..., main = NULL, sub = NULL, panel = points,
         lower = panel, upper = panel, diagonal = NULL,
         labels = TRUE, pos.labels = NULL, cex.labels = NULL,
         font.labels = par("font"), layout = c("matrix", "graph"),
         gap = 1)
```

### Arguments

<code>x</code>	a <code>matrix</code> or <code>data.frame</code> .
<code>main</code> , <code>sub</code>	main and sub title.
<code>panel</code>	a function( <code>x</code> , <code>y</code> , ...), which is used to plot the contents of each off-diagonal panel of the display.
...	further arguments and graphical parameters to be passed down. <code>par("oma")</code> will be set appropriately unless supplied (see <code>par</code> ).

<code>lower</code> , <code>upper</code>	separate panel functions to be used below and above the diagonal, respectively.
<code>diagonal</code>	optional <code>function(x, ...)</code> to be applied on the diagonal panels.
<code>labels</code>	either a logical indicating whether labels should be plotted in the diagonal panels, or a character vector giving the labels.
<code>pos.labels</code>	the vertical position of the labels in the diagonal panels.
<code>cex.labels</code>	the character expansion factor to be used for the labels.
<code>font.labels</code>	the font to be used for the labels.
<code>layout</code>	a character string giving the layout of the scatterplot matrix. Possible values are "matrix" (a matrix-like layout with the first row on top) and "graph" (a graph-like layout with the first row at the bottom).
<code>gap</code>	a numeric value giving the distance between the panels in margin lines.

### Details

This function is the workhorse for `marginmatrix` and `scattmatrixMiss`.

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

A panel function should not attempt to start a new plot, since the coordinate system for each panel is set up by `pairsVIM`.

### Note

The code is based on `pairs`.

### Author(s)

Andreas Alfons

### See Also

`marginmatrix`, `scattmatrixMiss`

### Examples

```
data(sleep, package = "VIM")
x <- sleep[, -(8:10)]
x[, c(1, 2, 4, 6, 7)] <- log10(x[, c(1, 2, 4, 6, 7)])
pairsVIM(x)
```

---

parcoordMiss                      *Parallel coordinate plot with information about missing values*

---

## Description

Parallel coordinate plot in which observations with missing values in certain variables are highlighted.

## Usage

```
parcoordMiss(x, highlight = NULL, selection = c("any", "all"),
             plotvars = NULL, col = c("skyblue", "red"),
             alpha = NULL, lty = par("lty"), xlim = NULL,
             ylim = NULL, main = NULL, sub = NULL,
             xlab = NULL, ylab = NULL, labels = TRUE,
             xpd = NULL, interactive = TRUE, ...)
```

```
TKRparcoordMiss(x, highlight = NULL, selection = c("any", "all"),
                plotvars = NULL, col = c("skyblue", "red"),
                alpha = NULL, ..., hscale = NULL, vscale = 1,
                TKRpar = list())
```

## Arguments

x	a matrix or data.frame.
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 values in multiple highlight variables. Possible values are "any" (highlighting of missing values in <i>any</i> of the highlight variables) and "all" (highlighting of missing 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 two giving the colors to be used in the plot. The second color is used for highlighting.
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.
lty	a vector of length two giving the line types. The second line type is used for the highlighted observations. If a single value is supplied, it is used for both non-highlighted and highlighted observations.
xlim, ylim	axis limits.
main, sub	main and sub title.
xlab, ylab	axis labels.

labels	either a logical indicating whether labels should be plotted below each coordinate axis, or a character vector giving the labels.
xpd	a logical indicating whether the lines should be allowed to go outside the plot region. If NULL, it defaults to TRUE unless axis limits are specified.
interactive	a logical indicating whether the variables to be used for highlighting can be selected interactively (see ‘Details’).
...	for parcoordMiss, further graphical parameters to be passed down (see <a href="#">par</a> ). 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.
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 <a href="#">par</a> ).

### Details

In parallel coordinate plots, the variables are represented by parallel axes. Each observation of the scaled data is shown as a line.

If `interactive=TRUE`, the variables to be used for highlighting can be selected interactively. Observations with missing 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 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 version 1.3 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

### References

Wegman, E. J. (1990) Hyperdimensional data analysis using parallel coordinates. *Journal of the American Statistical Association* **85**, 664-675.

### See Also

[pbox](#)

**Examples**

```
data(chorizonDL, package = "VIM")
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"))
```

pbox

*Parallel boxplots with information about missing values***Description**

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

**Usage**

```
pbox(x, pos = 1, selection = c("none", "any", "all"),
  col = c("skyblue", "red", "red4"), numbers = TRUE,
  cex.numbers = par("cex"), xlim = NULL, ylim = NULL,
  main = NULL, sub = NULL, xlab = NULL, ylab = NULL,
  axes = TRUE, frame.plot = axes, labels = axes,
  interactive = TRUE, ...)
```

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

**Arguments**

<code>x</code>	a vector, matrix or <code>data.frame</code> .
<code>pos</code>	a numeric value giving the index of the variable of interest. Additional variables in <code>x</code> are used for grouping according to missingness.
<code>selection</code>	the selection method for grouping according to missingness in multiple additional variables. Possible values are "none" (grouping according to missingness in every other variable that contains missing values), "any" (grouping according to missingness in <i>any</i> of the additional variables) and "all" (grouping according to missingness in <i>all</i> of the additional variables).
<code>col</code>	a vector of length three giving the colors to be used in the plot. The first two colors are used for the boxplots of the available and missing data, respectively, and the third color for the frequencies of missing values in both variables (see 'Details'). If only one color is supplied, it is used for the boxplots for missing data, whereas the boxplots for the available data are transparent. Else if two colors are supplied, the second one is recycled.
<code>numbers</code>	a logical indicating whether the frequencies of missing values should be displayed (see 'Details').
<code>cex.numbers</code>	the character expansion factor to be used for the frequencies of the missing values.

<code>xlim, ylim</code>	axis limits.
<code>main, sub</code>	main and sub title.
<code>xlab, ylab</code>	axis labels.
<code>axes</code>	a logical indicating whether axes should be drawn on the plot.
<code>frame.plot</code>	a logical indicating whether a box should be drawn around the plot.
<code>labels</code>	either a logical indicating whether labels should be plotted below each box, or a character vector giving the labels.
<code>interactive</code>	a logical indicating whether variables can be switched interactively (see ‘Details’).
<code>...</code>	for <code>pbox</code> , further arguments and graphical parameters to be passed to <code>boxplot</code> and other functions. For <code>TKRpbox</code> , further arguments to be passed to <code>pbox</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 boxes to be drawn.
<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 <code>par</code> ).

### 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 values according to `selection` are produced for the variable of interest.

Additionally, the frequencies of the missing 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 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

**See Also**

[parcoordMiss](#)

**Examples**

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

---

```
prepare
```

*Transformation and standardization*

---

**Description**

This function is used by the VIM GUI for transformation and standardization of the data.

**Usage**

```
prepare(x, scaling = c("none", "classical", "MCD", "robust", "onestep"),
        transformation = c("none", "minus", "reciprocal", "logarithm",
                           "exponential", "boxcox", "clr", "ilr", "alr"),
        alpha = NULL, powers = NULL, start = 0, alrVar)
```

**Arguments**

<code>x</code>	a vector, matrix or <code>data.frame</code> .
<code>scaling</code>	the scaling to be applied to the data. Possible values are "none", "classical", "MCD", "robust" and "onestep".
<code>transformation</code>	the transformation of the data. Possible values are "none", "minus", "reciprocal", "logarithm", "exponential", "boxcox", "clr", "ilr" and "alr".
<code>alpha</code>	a numeric parameter controlling the size of the subset for the <i>MCD</i> (if <code>scaling="MCD"</code> ). See <a href="#">covMcd</a> .
<code>powers</code>	a numeric vector giving the powers to be used in the Box-Cox transformation (if <code>transformation="boxcox"</code> ). If <code>NULL</code> , the powers are calculated with function <a href="#">box.cox.powers</a> .
<code>start</code>	a constant to be added prior to Box-Cox transformation (if <code>transformation="boxcox"</code> ).
<code>alrVar</code>	variable to be used as denominator in the additive logratio transformation (if <code>transformation="alr"</code> ).

## Details

### Transformation:

"none": no transformation is used.

"logarithm": compute the the logarithm (to the base 10).

"boxcox": apply a Box-Cox transformation. Powers may be specified or calculated with the function `box.cox.powers`.

### Standardization:

"none": no standardization is used.

"classical": apply a z-Transformation on each variable by using function `scale`.

"robust": apply a robustified z-Transformation by using median and MAD.

## Value

Transformed and standardized data.

## Author(s)

Matthias Templ, modifications by Andreas Alfons

## See Also

`scale`, `box.cox.powers`

## Examples

```
data(sleep, package = "VIM")
x <- sleep[, c("BodyWgt", "BrainWgt")]
prepare(x, scaling = "robust", transformation = "logarithm")
```

---

print.aggr

*Print method for objects of class aggr*

---

## Description

Print method for objects of class "aggr".

## Usage

```
## S3 method for class 'aggr':
print(x, digits = NULL, ...)
```

## Arguments

`x` an object of class "aggr".

`digits` the minimum number of significant digits to be used (see `print.default`).

`...` further arguments (currently ignored).

**Author(s)**

Matthias Templ, modifications by Andreas Alfons

**See Also**

[aggr](#)

**Examples**

```
data(sleep, package = "VIM")
a <- aggr(sleep, plot=FALSE)
a
```

---

`print.summary.aggr` *Print method for objects of class `summary.aggr`*

---

**Description**

Print method for objects of class "summary.aggr".

**Usage**

```
## S3 method for class 'summary.aggr':
print(x, ...)
```

**Arguments**

`x` an object of class "summary.aggr".  
`...` further arguments (currently ignored).

**Author(s)**

Andreas Alfons

**See Also**

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

**Examples**

```
data(sleep, package = "VIM")
s <- summary(aggr(sleep, plot=FALSE))
s
```

---

`rugNA`*Rug representation of missing values*

---

**Description**

Add a rug representation of missing values in only one of the variables to scatterplots.

**Usage**

```
rugNA(x, y, ticksize = NULL, side = 1,  
      col = "red", alpha = NULL, lwd = 0.5, ...)
```

**Arguments**

<code>x, y</code>	numeric vectors.
<code>ticksize</code>	the length of the ticks. Positive lengths give inward ticks.
<code>side</code>	an integer giving the side of the plot to draw the rug representation.
<code>col</code>	the color to be used for the ticks.
<code>alpha</code>	the alpha value (between 0 and 1).
<code>lwd</code>	the line width to be used for the ticks.
<code>...</code>	further arguments to be passed to <a href="#">Axis</a> .

**Details**

If `side` is 1 or 3, the rug representation consists of values available in `x` but missing in `y`. Else if `side` is 2 or 4, it consists of values available in `y` but missing in `x`.

**Author(s)**

Andreas Alfons

**Examples**

```
data(tao, package = "VIM")  
x <- tao[, "Air.Temp"]  
y <- tao[, "Humidity"]  
plot(x, y)  
rugNA(x, y, side = 1)  
rugNA(x, y, side = 2)
```

---

 scattJitt

*Bivariate jitter plot*


---

### Description

Create a bivariate jitter plot.

### Usage

```
scattJitt(x, col = c("skyblue", "red", "red4"), alpha = NULL,
          cex = par("cex"), col.line = "lightgrey",
          lty = "dashed", lwd = par("lwd"),
          numbers = TRUE, cex.numbers = par("cex"),
          main = NULL, sub = NULL, xlab = NULL,
          ylab = NULL, axes = TRUE, frame.plot = axes,
          labels = c("observed", "missing"), ...)
```

### Arguments

<code>x</code>	a data.frame or matrix with two columns.
<code>col</code>	a vector of length three giving the colors to be used in the plot. The first color will be used for complete observations, the second color for missing values in only one variable, and the third color for missing values in both variables. If only one color is supplied, it is used for all. 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>cex</code>	the character expansion factor for the plot characters.
<code>col.line</code>	the color for the lines dividing the plot region.
<code>lty</code>	the line type for the lines dividing the plot region (see <a href="#">par</a> ).
<code>lwd</code>	the line width for the lines dividing the plot region.
<code>numbers</code>	a logical indicating whether the frequencies of observed/missing values should be displayed (see ‘Details’).
<code>cex.numbers</code>	the character expansion factor to be used for the frequencies of the observed/missing values.
<code>main, sub</code>	main and sub title.
<code>xlab, ylab</code>	axis labels.
<code>axes</code>	a logical indicating whether both axes should be drawn on the plot. Use graphical parameter <code>"xaxt"</code> or <code>"yaxt"</code> to suppress just one of the axes.
<code>frame.plot</code>	a logical indicating whether a box should be drawn around the plot.
<code>labels</code>	a vector of length two giving the axis labels for the regions for observed/missing values (see ‘Details’).
<code>...</code>	further graphical parameters to be passed down (see <a href="#">par</a> ).

**Details**

The amount of observed/missing values is visualized by jittered points. Thereby the plot region is divided into up to four regions according to the existence of missing values in one or both variables. In addition, the amount of observed/missing values can be represented by a number.

**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 argument `cex.text` can still be supplied to `...` and is handled correctly. Nevertheless, it is deprecated and no longer documented. Use `cex.numbers` instead.

**Author(s)**

Matthias Templ, modifications by Andreas Alfons

**Examples**

```
data(tao, package = "VIM")
scattJitt(tao[, c("Air.Temp", "Humidity")])
```

---

scattmatrixMiss      *Scatterplot matrix with information about missing values*

---

**Description**

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

**Usage**

```
scattmatrixMiss(x, highlight = NULL,
                selection = c("any", "all"), plotvars = NULL,
                col = c("skyblue", "red"), alpha = NULL,
                pch = c(1, 3), lty = par("lty"),
                diagonal = c("density", "none"),
                interactive = TRUE, ...)

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

**Arguments**

<code>x</code>	a <code>matrix</code> or <code>data.frame</code> .
<code>highlight</code>	a vector giving the variables to be used for highlighting. If <code>NULL</code> (the default), all variables are used for highlighting.
<code>selection</code>	the selection method for highlighting missing values in multiple <code>highlight</code> variables. Possible values are <code>"any"</code> (highlighting of missing values in <i>any</i> of the <code>highlight</code> variables) and <code>"all"</code> (highlighting of missing values in <i>all</i> of the <code>highlight</code> variables).
<code>plotvars</code>	a vector giving the variables to be plotted. If <code>NULL</code> (the default), all variables are plotted.
<code>col</code>	a vector of length two giving the colors to be used in the plot. The second color will be used for highlighting.
<code>alpha</code>	a numeric value between 0 and 1 giving the level of transparency of the colors, or <code>NULL</code> . This can be used to prevent overplotting.
<code>pch</code>	a vector of length two giving the plot characters. The second plot character will be used for the highlighted observations.
<code>lty</code>	a vector of length two giving the line types for the density plots in the diagonal panels (if <code>diagonal="density"</code> ). The second line type is used for the highlighted observations. If a single value is supplied, it is used for both non-highlighted and highlighted observations.
<code>diagonal</code>	a character string specifying the plot to be drawn in the diagonal panels. Possible values are <code>"density"</code> (density plots for non-highlighted and highlighted observations) and <code>"none"</code> .
<code>interactive</code>	a logical indicating whether the variables to be used for highlighting can be selected interactively (see ‘Details’).
<code>...</code>	for <code>scattmatrixMiss</code> , further arguments and graphical parameters to be passed to <code>pairsVIM</code> . <code>par("oma")</code> will be set appropriately unless supplied (see <code>par</code> ). For <code>TKRscattmatrixMiss</code> , further arguments to be passed to <code>scattmatrixMiss</code> .
<code>hscale</code>	horizontal scale factor for plot to be embedded in a <code>Tcl/Tk</code> 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 <code>Tcl/Tk</code> 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 <code>Tcl/Tk</code> window (see ‘Details’ and <code>par</code> ).

**Details**

`scattmatrixMiss` uses `pairsVIM` with a panel function that allows highlighting of missing values.

If `interactive=TRUE`, the variables to be used for highlighting can be selected interactively. Observations with missing 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

### See Also

`pairsVIM`, `marginmatrix`

### Examples

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

---

scattMiss

*Scatterplot with information about missing values*

---

### Description

In addition to a standard scatterplot, lines are plotted for the missing values in one variable.

### Usage

```
scattMiss(x, side = 1, col = c("skyblue", "red", "lightgrey"),
          alpha = NULL, lty = c("dashed", "dotted"),
          lwd = par("lwd"), quantiles = c(0.5, 0.975),
          inEllipse = FALSE, xlim = NULL, ylim = NULL,
          main = NULL, sub = NULL, xlab = NULL, ylab = NULL,
          interactive = TRUE, ...)
```

**Arguments**

<code>x</code>	a <code>matrix</code> or <code>data.frame</code> with two columns.
<code>side</code>	if <code>side=1</code> , a rug representation and vertical lines are plotted for the missing values in the second variable; if <code>side=2</code> , a rug representation and horizontal lines for the missing values in the first variable.
<code>col</code>	a vector of length three giving the colors to be used in the plot. The first color is used for the scatterplot, the second color for the rug representation and the lines, and the third color for the ellipses (see ‘Details’). If only one color is supplied, it is used for the scatterplot, the rug representation and the lines, whereas the default color is used for the ellipses. Else if a vector of length two is supplied, the default color is used for the ellipses as well.
<code>alpha</code>	a numeric value between 0 and 1 giving the level of transparency of the colors, or <code>NULL</code> . This can be used to prevent overplotting.
<code>lty</code>	a vector of length two giving the line types for the lines and ellipses. If a single value is supplied, it will be used for both.
<code>lwd</code>	a vector of length two giving the line widths for the lines and ellipses. If a single value is supplied, it will be used for both.
<code>quantiles</code>	a vector giving the quantiles of the chi-square distribution to be used for the tolerance ellipses, or <code>NULL</code> to suppress plotting ellipses (see ‘Details’).
<code>inEllipse</code>	plot lines only inside the largest ellipse. Ignored if <code>quantiles=NULL</code> .
<code>xlim, ylim</code>	axis limits.
<code>main, sub</code>	main and sub title.
<code>xlab, ylab</code>	axis labels.
<code>interactive</code>	a logical indicating whether the <code>side</code> argument can be changed interactively (see ‘Details’).
<code>...</code>	further graphical parameters to be passed down (see <a href="#">par</a> ).

**Details**

Information about missing values in one variable is included as vertical or horizontal lines, as determined by the `side` argument. The lines are thereby drawn at the observed x- or y-value. In addition, percentage coverage ellipses can be drawn to give a clue about the shape of the bivariate data distribution.

If `interactive=TRUE`, clicking in the bottom margin redraws the plot with information about missing values in the first variable and clicking in the left margin redraws the plot with information about missing values in the second variable. Clicking anywhere else in the plot quits the interactive session.

**Author(s)**

Andreas Alfons

**See Also**

[marginplot](#)

**Examples**

```
data(tao, package = "VIM")
scattMiss(tao[,c("Air.Temp", "Humidity")])
```

---

sleep

*sleep data*

---

**Description**

Sleep data with missing values.

**Usage**

```
data(sleep)
```

**Format**

A data frame with 62 observations on the following 10 variables.

**BodyWgt** a numeric vector

**BrainWgt** a numeric vector

**NonD** a numeric vector

**Dream** a numeric vector

**Sleep** a numeric vector

**Span** a numeric vector

**Gest** a numeric vector

**Pred** a numeric vector

**Exp** a numeric vector

**Danger** a numeric vector

**Details**

This data set is also known under the name `brain`, which is available in package `rrcov`. In addition to the `brain` data set, all observations that contain missing values are included.

**Source**

This data set was originally included in `GGobi`.

**Examples**

```
data(sleep, package = "VIM")
summary(sleep)
aggr(sleep)
```

---

 spineMiss

*Spineplot with information about missing values*


---

### Description

Spineplot or spinogram with highlighting of missing values in other variables by splitting each cell into two parts. Additionally, information about missing values in the variable of interest is shown on the right hand side.

### Usage

```
spineMiss(x, pos = 1, selection = c("any", "all"),
          breaks = "Sturges", right = TRUE,
          col = c("skyblue", "red", "skyblue4", "red4"),
          border = NULL, main = NULL, sub = NULL,
          xlab = NULL, ylab = NULL, axes = TRUE,
          labels = axes, only.miss = TRUE,
          miss.labels = axes, interactive = TRUE, ...)
```

### Arguments

<code>x</code>	a vector, matrix or <code>data.frame</code> .
<code>pos</code>	a numeric value giving the index of the variable of interest. Additional variables in <code>x</code> are used for highlighting.
<code>selection</code>	the selection method for highlighting missing values in multiple additional variables. Possible values are "any" (highlighting of missing values in <i>any</i> of the additional variables) and "all" (highlighting of missing values in <i>all</i> of the additional variables).
<code>breaks</code>	if the variable of interest is numeric, <code>breaks</code> controls the breakpoints (see <a href="#">hist</a> for possible values).
<code>right</code>	logical; if <code>TRUE</code> and the variable of interest is numeric, the spinogram cells are right-closed (left-open) intervals.
<code>col</code>	a vector of length four giving the colors to be used. If only one color is supplied, the bars are transparent and the supplied color is used for highlighting. Else if two colors are supplied, they are recycled.
<code>border</code>	the color to be used for the border of the cells. Use <code>border=NA</code> to omit borders.
<code>main, sub</code>	main and sub title.
<code>xlab, ylab</code>	axis labels.
<code>axes</code>	a logical indicating whether axes should be drawn on the plot.
<code>labels</code>	if the variable of interest is categorical, either a logical indicating whether labels should be plotted below each cell, or a character vector giving the labels. This is ignored if the variable of interest is numeric.

<code>only.miss</code>	logical; if TRUE, the missing values in the variable of interest are also visualized by a cell in the spineplot or spinogram. Otherwise, a small spineplot is drawn on the right hand side (see ‘Details’).
<code>miss.labels</code>	either a logical indicating whether label(s) should be plotted below the cell(s) on the right hand side, or a character string or vector giving the label(s) (see ‘Details’).
<code>interactive</code>	a logical indicating whether the variables can be switched interactively (see ‘Details’).
<code>...</code>	further graphical parameters to be passed to <code>title</code> and <code>axis</code> .

### Details

A spineplot is created if the variable of interest is categorial and a spinogram if it is numerical. The horizontal axis is scaled according to relative frequencies of the categories/classes. If more than one variable is supplied, the cells are split according to missingness in the additional variables. Thus the proportion of highlighted observations in each category/class is displayed on the vertical axis. Since the height of each cell corresponds to the proportion of highlighted observations, it is now possible to compare the proportions of missing values among the different categories/classes.

If `only.miss=TRUE`, the missing values in the variable of interest are also visualized by a cell in the spine plot or spinogram. If additional variables are supplied, this cell is again split into two parts according to missingness in the additional variables.

Otherwise, a small spineplot that visualizes missing values in the variable of interest is drawn on the right hand side. The first cell corresponds to observed values and the second cell to missing values. Each of the two cells is again split into two parts according to missingness in the additional variables. Note that this display does not make sense if only one variable is supplied, therefore `only.miss` is ignored in that case.

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.

### Value

a table containing the frequencies corresponding to the cells.

### 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 `xaxlabels` and `missaxlabels` can still be supplied to `...` and are handled correctly. Nevertheless, they are deprecated and no longer documented. Use `labels` and `miss.labels` instead.

The code is based on the function `spineplot` by Achim Zeileis.

### Author(s)

Andreas Alfons, Matthias Templ

**See Also**

[histMiss](#), [barMiss](#)

**Examples**

```
data(tao, package = "VIM")
spineMiss(tao[, c("Air.Temp", "Humidity")])
data(sleep, package = "VIM")
spineMiss(sleep[, c("Exp", "Sleep")])
```

---

summary.aggr

*Summary method for objects of class aggr*


---

**Description**

Summary method for objects of class "aggr".

**Usage**

```
## S3 method for class 'aggr':
summary(object, ...)
```

**Arguments**

object            an object of class "aggr".  
...                further arguments.

**Value**

a list of class "summary.aggr" containing the following components:

missings        a data.frame containing the amount of missing values in each variable.  
combinations   a data.frame containing a character vector representing the combinations of variables along with their frequencies and percentages.

**Author(s)**

Matthias Templ, modifications by Andreas Alfons

**See Also**

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

**Examples**

```
data(sleep, package = "VIM")
summary(aggr(sleep, plot=FALSE))
```

---

tao

*tao data*

---

## Description

A small subsample of the tropical atmosphere ocean (TAO) project data, derived from the GGOBI project.

## Usage

```
data(tao)
```

## Format

A data frame with 736 observations on the following 8 variables.

**Year** a numeric vector

**Latitude** a numeric vector

**Longitude** a numeric vector

**Sea.Surface.Temp** a numeric vector

**Air.Temp** a numeric vector

**Humidity** a numeric vector

**UWind** a numeric vector

**VWind** a numeric vector

## Details

All cases recorded for five locations and two time periods.

## Source

<http://www.pmel.noaa.gov/tao/>

## Examples

```
data(tao, package = "VIM")
summary(tao)
aggr(tao)
```

**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 R 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.

Last, but not least, the *Options* menu allows to set the colors and alpha channel 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.

Sections *Imputation* and *Diagonstics* are not yet implemented.

**Author(s)**

Andreas Alfons, based on an initial design by Matthias Templ

# Index

- \*Topic **color**
  - alphablend, 5
  - rugNA, 35
- \*Topic **datasets**
  - chorizonDL, 8
  - kola.background, 20
  - sleep, 41
  - tao, 45
- \*Topic **hplot**
  - aggr, 2
  - barMiss, 6
  - bgmap, 8
  - colormapMiss, 12
  - growdotMiss, 15
  - histMiss, 17
  - mapMiss, 20
  - marginmatrix, 22
  - marginplot, 23
  - matrixplot, 25
  - pairsVIM, 27
  - parcoordMiss, 28
  - pbox, 30
  - scattJitt, 36
  - scattmatrixMiss, 37
  - scattMiss, 40
  - spineMiss, 42
  - vmGUImenu, 46
- \*Topic **manip**
  - initialise, 19
  - prepare, 32
- \*Topic **multivariate**
  - vmGUImenu, 46
- \*Topic **package**
  - VIM-package, 2
- \*Topic **print**
  - print.aggr, 34
  - print.summary.aggr, 34
  - summary.aggr, 44
- \*Topic **utilities**
  - count, 14
  - aggr, 2, 34, 35, 45
  - alphablend, 5
  - Axis, 35
  - axis, 6, 18, 26, 43
  - barMiss, 6, 19, 44
  - bgmap, 8, 15, 16, 21
  - box.cox.powers, 33
  - boxplot, 31, 32
  - bubbleFIN, 16
  - bubbleMiss, 21
  - bubbleMiss (growdotMiss), 15
  - chorizon, 8, 12
  - chorizonDL, 8
  - colormapMiss, 12, 16, 21
  - colormapMissLegend (colormapMiss), 12
  - count, 14
  - countInf (count), 14
  - countNA (count), 14
  - covMcd, 33
  - format, 16
  - growdotMiss, 8, 14, 15
  - hist, 17, 42
  - histMiss, 7, 17, 44
  - iimagMiss (matrixplot), 25
  - imagMiss, 26
  - initialise, 19
  - kola.background, 20
  - lines, 8
  - mapMiss, 8, 14, 16, 20

marginmatrix, 22, 27, 28, 39  
marginplot, 22, 23, 23, 41  
matrixplot, 25

pairs, 28  
pairsVIM, 22, 23, 27, 38, 39  
par, 3, 4, 22, 24, 26, 27, 29, 31, 36–40  
parcoordMiss, 28, 32  
pbox, 30, 30  
plot.aggr, 3  
plot.aggr(aggr), 2  
plot.window, 26  
points, 21  
prepare, 32  
print.aggr, 4, 34  
print.default, 34  
print.summary.aggr, 34, 45

rugNA, 35

scale, 33  
scattJitt, 36  
scattmatrixMiss, 23, 27, 28, 37  
scattMiss, 24, 40  
sleep, 41  
spineMiss, 7, 19, 42  
spineplot, 44  
summary.aggr, 4, 35, 44

tao, 45  
title, 6, 18, 26, 43  
TKRaggr(aggr), 2  
TKRmarginmatrix(marginmatrix), 22  
TKRmatrixplot(matrixplot), 25  
TKRparcoordMiss(parcoordMiss), 28  
TKRpbox(pbox), 30  
tkrplot, 4, 22, 26, 29, 32, 39  
TKRscattmatrixMiss  
    (scattmatrixMiss), 37

VIM(VIM-package), 2  
VIM-package, 2  
vmGUImenu, 46