**Abiotic_data_sample**

**Description**

Example dataset containing normalized cell numbers and abiotic parameters used for the NMDS plot.

**Usage**

```r
data(Abiotic_data_sample)
```

**Format**

Data frame with 10 observations of 44 variables.
**Cell_number_sample**

Dataset with relative measured cell numbers and sample names

**Description**

Example dataset of relative measured cell numbers with sample names.

**Usage**

```r
data(Cell_number_sample)
```

**Format**

Data frame with 10 observations of 31 variables.

**correlation**

Visualize the correlation

**Description**

A function to visualize the correlation of the relative or percental cell numbers and the abiotic parameters.

**Usage**

```r
## S4 method for signature 'data.frame'
correlation(x,cortype="spearman",exact=FALSE,colkey=bluered(21),Rowv=TRUE,Colv=TRUE,symm=TRUE,distfun=function(c) as.dist(1 - c),dendrogram="both",main="",verbose=FALSE,...)
```

```r
## S4 method for signature 'matrix'
correlation(x,cortype="spearman",exact=FALSE,colkey=bluered(21),Rowv=TRUE,Colv=TRUE,symm=TRUE,distfun=function(c) as.dist(1 - c),dendrogram="both",main="",verbose=FALSE,...)
```

**Arguments**

- `x`: Data frame or matrix of correlation values. Use one row per sample and one column per gate/abiotic parameter. Use the first column for the first gate/abiotic parameter. The names of the gates and abiotic parameters may not contain any whitespace. If the values contain commas they must be expressed as ".". Missing values or NA's are allowed but should be expressed as "NA" or empty entries within the table.

- `cortype`: Character string indicating which correlation coefficient is used. Choose between "spearman" (default), "kendall" and "pearson". See Details for more information.

- `exact`: logical (default=FALSE). Do not compute exact p-values. Used for Kendall’s tau and Spearman’s rho. Can be changed to TRUE if R can compute exact p-values without ties. Computing exact p-values depends on the dataset. For more details type "?cor.test" into R command line.

- `colkey`: Indicating the color key/palette used for visualization (default=bluered(21)). See Details for more information.
correlation

Rowv: logical (default=TRUE). The row dendrogram will be arranged according to similarity. FALSE will arrange the dendrogram according to the sequence of data e.g. according to sampling.

Colv: logical (default=TRUE). Determines how the column dendrogram should be reordered. Has the same options as the Rowv argument above.

symm: logical (default=TRUE). Indicates if x should be treated symmetrically. Can only be TRUE if x is a square matrix.

distfun: Function used for the computation of the distance (dissimilarity) between both rows and columns (default=function(c) as.dist(1 - c)).

dendrogram: Character string indicating whether dendrograms should be displayed (default="both"). Can be changed to "row", "column" or "none". Rowv/Colv has to be set TRUE to show a dendrogram for the rows/columns.

main: Character string used as title of the correlation heatmap (default=").

verbose: logical (default=FALSE). Change to TRUE to print estimates and additional information to the heatmap.

...: Additional heatmap parameters. For more details type "?heatmap.2" into R command line.

Details

Visualize the correlations between the cell number variations and abiotic parameters as a heatmap using the heatmap.2 function of the package "gplots" (see reference Warnes et al. 2013). Spearman's rank-order correlation coefficient (r_S) is used for this purpose, by default (see reference Koch et al. 2013). This value varies between -1 and 1. The stronger the correlation, the closer the value of the coefficient to 1 (for positive correlations) or -1 (for negative correlations). Kendall's tau rank correlation coefficient also assigns the values to ranks but considers the difference between the probability that any two points will agree on the relative ranks compared to the probability that they will disagree. Pearson's product-moment correlation coefficient uses the true values and can be chosen to depict linear relationships. You can choose the color key/palette and the gradations for visualization. By choosing the default color key positive correlations will be shown in shades of red, negative correlations will be shown in shades of blue and neutral correlations will be shown in white. Other possibilities are e.g. topo.colors(16), cm.colors(15) or heat.colors(11). The space within the brackets defines the gradations of the color key. For more details type "??Palettes" into R command line. It is also possible to define an own color palette. For more details on that type "?colorRampPalette" into R command line.

Author(s)

Joachim Schumann <joachim.schumann@ufz.de>, Christin Koch <christin.koch@ufz.de>, Susanne Günther <susanne.guenther@ufz.de>, Ingo Fetzer <info.fetzer@stockholmsilence.su.se>, Susann Müller <susann.mueller@ufz.de>

References


Examples

```r
require(gplots)
## Show the correlation saved in dataset Corr_data_sample
## Use default parameters
data(Corr_data_sample)
correlation(Corr_data_sample[,,-1])
## Change the title to "Correlation heatmap"
## Use palette heat.colors(11) as color key
data(Corr_data)
correlation(Corr_data_sample[,,-1],colkey=heat.colors(11),main="Correlation heatmap")
```

---

Corr_data_sample

**Dataset with correlation data and sample names**

Description

Example dataset containing relative measured cell numbers and abiotic parameters for correlation analysis.

Usage

```r
data(Corr_data_sample)
```

Format

Data frame with 10 observations of 44 variables.

---

cybar_plot

**Barcode of the normalized cell numbers and boxplot of relative or percental cell numbers**

Description

Show a barcode of the variation of normalized cell numbers (CyBar plot) and a boxplot of the relative or percental cell numbers.

Usage

```r
## S4 method for signature 'data.frame, data.frame'
cybar_plot(x,y,Rowv=FALSE,grad=21,trace="none",tracecol="black",dendrogram="col",from=0,to=2,na.order="original",boxmain="",verbose=FALSE,...)

## S4 method for signature 'matrix, matrix'
cybar_plot(x,y,Rowv=FALSE,grad=21,trace="none",tracecol="black",
```
cybar_plot

dendrogram="col",from=0,to=2,na.color="black",barmain="",labels="hor", order="original",boxmain="",verbose=FALSE,...)

## S4 method for signature 'data.frame,matrix'
cybar_plot(x,y,Rowv=FALSE,grad=21,trace="none",tracecol="black",dendrogram="col",from=0,to=2,na.color="black",barmain="",order="original",boxmain="",verbose=FALSE,...)

## S4 method for signature 'matrix,data.frame'
cybar_plot(x,y,Rowv=FALSE,grad=21,trace="none",tracecol="black",dendrogram="col",from=0,to=2,na.color="black",barmain="",order="original",boxmain="",verbose=FALSE,...)

## S4 method for signature 'data.frame,missing'
cybar_plot(x,y,Rowv=FALSE,grad=21,trace="none",tracecol="black",dendrogram="col",from=0,to=2,na.color="black",barmain="",verbose=FALSE,...)

## S4 method for signature 'matrix,missing'
cybar_plot(x,y,Rowv=FALSE,grad=21,trace="none",tracecol="black",dendrogram="col",from=0,to=2,na.color="black",barmain="",verbose=FALSE,...)

## S4 method for signature 'missing,data.frame'
cybar_plot(x,y,labels="hor",boxmain="",verbose=FALSE,...)

## S4 method for signature 'missing,matrix'
cybar_plot(x,y,labels="hor",boxmain="",verbose=FALSE,...)

Arguments

x Data frame or matrix of normalized cell numbers. Use one row per sample and one column per gate. Use the first column for the first gate. The names of the gates may not contain any whitespace. If the values contain commas they must be expressed as ".". Missing values or NA's are allowed but should be expressed as "NA" or empty entries within the table.

y Data frame or matrix of relative or percental cell numbers. Same requirements as described for x.

Rowv Keeps sequence of data in the barcode e.g. according to sampling (default=FALSE). If TRUE it will be arranged according to similarity.

grad Gradiations of the color key used in the barcode (default=21).

trace Display trace lines in the barcode (default="none"). For displaying trace lines choose between "column", "row" or "both".

tracecol Color of traces in the barcode if trace is active (default="black").

dendrogram Character string indicating whether dendrograms should be displayed in the barcode (default="column"). Can be changed to "row", "both" or "none". If you want to display a dendrogram for the rows and columns change to "both" and Rowv to TRUE.

from Beginning of the data range of the normalized cell numbers (default=0).

to End of the data range of the normalized cell numbers (default=2).

na.color Color used for missing values or NA (default="black").

barmain Character string used as title of the barcode (default="").

labels Character string indicating whether horizontal ("hor") data labels (default) or vertical ("vert") data labels should be plotted in the boxplot.
**cybar_plot**

order Character string indicating the order of the gates in the boxplot (default="original"). The boxes are ordered respective to the original order of the gates. Choose "sim" if the boxes should be ordered respective to the similarity calculated in the heatmap function. This can only be done if x is given and the barcode is created.

boxmain Character string used as title of the boxplot (default="").

verbose logical (default=FALSE). Change to TRUE to print additional information to the barcode and/or the boxplot.

... Additional plotting parameters used for the heatmap and the boxplot. For more details type "?heatmap.2" or "?boxplot" into R command line.

**Details**

This function generates a barcode of the normalized cell numbers for all measurements using the heatmap.2 function of the package "gplots" (see reference Warnes et al. 2013). The barcode displays the variation for each gate using a color key with 21 gradations. Very low values are shown in blue, intermediate values are shown in white and high values are sown in deep red. Missing values or NA's are shown in black. In addition a dendrogram to reveal clusters and a histogram of the values are shown. The distribution of the relative or percental cell numbers per gate is visualized as a vertical boxplot. One box consists of the median, the upper and lower quartil (top and bottom of the box) containing 50% of the data, whiskers showing the minimum (within 1.5-fold of the interquartile range of the lower quartile) and the maximum (within 1.5-fold of the interquartile range of the upper quartile) of the data and outliers. Choose between horizontal ("hor") and vertical ("vert") data labels that represent the gates. See reference Koch et al. 2013 for more details.

**Author(s)**

Joachim Schumann <joachim.schumann@ufz.de>, Christin Koch <christin.koch@ufz.de>, Susanne Günther <susanne.guenther@ufz.de>, Ingo Fetzer <info.fetzer@stockholmresilience.su.se>, Susann Müller <susann.mueller@ufz.de>

**References**


**Examples**

```R
require(gplots)

## Show a barcode of the normalized cell numbers
## and a boxplot with horizontal data labels of the percental cell numbers
## saved in dataset Cell_number_sample
## The boxes are ordered according to the original order

data(Cell_number_sample)
Normalized_mean<-normalize(Cell_number_sample[, -1], digits=2)
Normalized_mean<-data.frame(data.matrix(Normalized_mean))
```
cybar_plot(Normalized_mean, Cell_number_sample[, -1])

## Barcode and boxplot are titled according to the dataset names, respectively
## The orientation of the data labels is vertical

data(Cell_number_sample)
Normalized_mean <- normalize(Cell_number_sample[, -1], digits = 2)
Normalized_mean <- data.frame(data.matrix(Normalized_mean))
cybar_plot(Normalized_mean, Cell_number_sample[, -1],
barmain = "Barcode of normalized cell numbers", labels = "vert", boxmain = "Boxplot of cell numbers")

## Show a barcode of the normalized cell numbers using method "first"
## The end of the data range of the normalized cell numbers is set to 3.5
## The boxes are ordered according to the similarity calculated in the barcode

data(Cell_number_sample)
Normalized_first <- normalize(Cell_number_sample[, -1], method = "first", digits = 2)
Normalized_first <- data.frame(data.matrix(Normalized_first))
cybar_plot(Normalized_first, Cell_number_sample[, -1], to = 3.5,
barmain = "Barcode of normalized cell numbers", labels = "vert", order = "sim", boxmain = "Boxplot of cell numbers")

---

nmds

**nmds**

NMDS plot of cell numbers

**Description**

NMDS plot of the relative/percental or normalized cell numbers.

**Usage**

```r
## S4 method for signature 'data.frame'
nmds(x, distance = "bray", autotransform = FALSE, zerodist = "add", group, main = "", type = "p",
cex = 0.6, pos = 4, shrink = TRUE, legend_pos = "topleft", pch = 1, col = "black", abiotic, p.max = 0.05, col_abiotic =)

## S4 method for signature 'matrix'
nmds(x, distance = "bray", autotransform = FALSE, zerodist = "add", group, main = "", type = "p",
cex = 0.6, pos = 4, shrink = TRUE, legend_pos = "topleft", pch = 1, col = "black", abiotic, p.max = 0.05, col_abiotic =)
```

**Arguments**

- **x**: Data frame or matrix of relative/percental cell numbers or normalized cell numbers. Use one row per sample and one column per gate. Use the first column for the first gate. The names of the gates may not contain any whitespace. If the values contain commas they must be expressed as ".". Missing values or NA's are not allowed.
- **distance**: Dissimilarity index used in vegdist (default = "bray"). For more details type "?metaMDS" into R command line.
- **autotransform**: logical (default = FALSE). Determine the use of simple heuristics for a possible data transformation of typical community data. For more details type "?metaMDS" into R command line.
<table>
<thead>
<tr>
<th>Argument</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>zerodist</td>
<td>Handling of zero dissimilarities (default=&quot;add&quot;). For more details type &quot;?metaMDS&quot; into R command line.</td>
</tr>
<tr>
<td>group</td>
<td>Table with group assignments. Use only one column. Use the first line as header. Assign the samples to groups in the next lines. The order and the number of these lines has to be identical to the order and the number of the samples printed in R. Use only integer values in the range from 1 to 25.</td>
</tr>
<tr>
<td>main</td>
<td>Character string used as title of the NMDS plot (default=&quot;&quot;)</td>
</tr>
<tr>
<td>type</td>
<td>Type of the plot (default=&quot;p&quot;). The &quot;p&quot; indicates points. For more details type &quot;?points&quot; into R command line.</td>
</tr>
<tr>
<td>cex</td>
<td>numeric (default=0.6). Character expansion factor. Used for the final size of the characters.</td>
</tr>
<tr>
<td>pos</td>
<td>Position of the text (default=4). Values of 1, 2, 3 and 4 indicate positions below, to the left of, above and to the right of the specified coordinates, respectively.</td>
</tr>
<tr>
<td>shrink</td>
<td>logical (default=TRUE). Shrink back species scores if they were expanded originally.</td>
</tr>
<tr>
<td>legend_pos</td>
<td>Position of the legend (default=&quot;topleft&quot;). For more details type &quot;?legend&quot; into R command line.</td>
</tr>
<tr>
<td>pch</td>
<td>Plotting symbol (default=1) if group is FALSE. For more details type &quot;?points&quot; into R command line.</td>
</tr>
<tr>
<td>col</td>
<td>Color of the symbols and lines within the plot (default=&quot;black&quot;) if group is FALSE. Can be a color code or a name.</td>
</tr>
<tr>
<td>abiotic</td>
<td>Table with abiotic data. Should be a tab-delimited text file using '.' as decimal delimiter. Use one row for one sample and one column for one abiotic parameter. Use the first column for the first parameter and the first line as header. The order and the number of the lines have to be identical to the order and the number of the samples printed in R.</td>
</tr>
<tr>
<td>p.max</td>
<td>Decimal number defining the significance level of the abiotic parameters (default=0.05) if abiotic=TRUE. Only parameters less/equal this value are plotted.</td>
</tr>
<tr>
<td>col_abiotic</td>
<td>Color used for the plotted abiotic parameters (default=&quot;magenta&quot;).</td>
</tr>
<tr>
<td>verbose</td>
<td>logical (default=FALSE). Change to TRUE to print results of the metaMDS function and the p-values of the abiotic parameters.</td>
</tr>
</tbody>
</table>

Details

Visualizes the distance between samples based on their cytometrically measured cell abundance information using nonmetric multidimensional scaling (NMDS). NMDS is a mathematical technique to visualize the distances between objects (the samples, in this case). The distances within the plot are used to be in accordance with the dis-/similarities of the samples. The NMDS plot can be created using either the relative/percental cell numbers or the normalized values. If there are high abundance and low abundance gates the NMDS plot could be distorted. In this case it is better to use the normalized values. See reference Koch et al. 2013, Box 1 for more details. The NMDS analysis is performed by using the metaMDS function of the R package "vegan" (see reference Warnes et al. 2013).

Author(s)

Joachim Schumann <joachim.schumann@ufz.de>, Christin Koch <christin.koch@ufz.de>, Susanne Günther <susanne.guenther@ufz.de>, Ingo Fetzer <info.fetzer@stockholmresilience.su.se>, Susann Müller <susann.mueller@ufz.de>
References


Examples

```R
require(vegan)

## Show the NMDS plot of the normalized cell numbers

data(Cell_number_sample)
Normalized_mean<-normalize(Cell_number_sample[-1],digits=2)
Normalized_mean<-data.frame(data.matrix(Normalized_mean))
nmds(Normalized_mean)

## Change the title of the plot to "NMDS normalized", use triangles as plotting symbols, change the color to red and position the text below the triangles
## Use a dotted line (lty=3)

data(Cell_number_sample)
Normalized_mean<-normalize(Cell_number_sample[-1],digits=2)
Normalized_mean<-data.frame(data.matrix(Normalized_mean))
nmds(Normalized_mean,main="NMDS normalized",type="b",pos=1,pch=2,col="red",lty=3)

## Plot sample groups saved as data frame groups
## Print additional information

data(Cell_number_sample)
Normalized_mean<-normalize(Cell_number_sample[-1],digits=2)
Normalized_mean<-data.frame(data.matrix(Normalized_mean))
groups<-data.frame("groups"=c(1,1,1,1,2,3,3,3,3,3))
nmds(Normalized_mean,group=groups,main="NMDS normalized",verbose=TRUE)

## Plot additional gate information and
## abiotic parameters saved in dataset Abiotic_data_sample

data(Cell_number_sample)
Normalized_mean<-normalize(Cell_number_sample[-1],digits=2)
Normalized_mean<-data.frame(data.matrix(Normalized_mean))
groups<-data.frame("groups"=c(1,1,1,1,2,3,3,3,3,3))

data(Abiotic_data_sample)

nmds(Normalized_mean,group=groups,main="NMDS normalized",verbose=TRUE)
```

normalize

Normalize cell numbers
normalize

Description

Normalize the relative or percental cell numbers.

Usage

## S4 method for signature 'data.frame'
normalize(x, method = "mean", show_avg = TRUE, ...)

## S4 method for signature 'matrix'
normalize(x, method = "mean", show_avg = TRUE, ...)

Arguments

x

Data frame or matrix of all relative or percental cell numbers. Use one row per sample and one column per gate. Use the first column for the first gate. If your first column is used e.g. for sample names read the manual before you continue. The names of the gates may not contain any whitespace. If the values contain commas they must be expressed as ".". Missing values or NA's are allowed but should be expressed as "NA" or empty entries within the table.

method

Character string indicating which method is used for normalization (default = "mean"). Can be changed to "first". See Details for further information.

show_avg

logical (default = TRUE). Show average values. Can be changed to FALSE if the averages are not of interest. Averages are only shown if method = "mean".

...

Additional parameters used for printing and writing e.g. number of digits.

Details

This step should be done to ensure the comparability of gates with highly different (low and high) cell abundances. There are two different methods to do this. By choosing "mean" the individual cell number of each gate is normalized by dividing the relative value by the mean of the relative or percental cell numbers for that gate. By choosing "first" the individual cell number of each gate is normalized by dividing the relative value by the first value of that gate.

Author(s)

Joachim Schumann <joachim.schumann@ufz.de>, Christin Koch <christin.koch@ufz.de>, Susanne Günther <susanne.guenther@ufz.de>, Ingo Fetzer <info.fetzer@stockholmresilience.su.se>, Susann Müller <susann.mueller@ufz.de>

Examples

## Normalize the percental cell numbers saved in dataset Cell_number_sample
## Use the default method and show the averages with two digits
## Note that the averages are very different but the normalized values
## are in the same range

data(Cell_number_sample)
normalize(Cell_number_sample[, -1], digits = 2)

## Use the method "first"
## Print normalized values
## After normalization the first value of every gate is 1.00
data(Cell_number_sample)
normalize(Cell_number_sample[,-1], method="first", digits=2)
Index

*Topic barcode, normalize
  cybar_plot, 5

*Topic correlation
  correlation, 3

*Topic datasets
  Abiotic_data_sample, 2
  Cell_number_sample, 3
  Corr_data_sample, 5

*Topic file.show(file.path(R.home(doc),KEYWORDS))
  correlation, 3
  cybar_plot, 5
  nmds, 8
  normalize, 10

*Topic nmds, normalize
  nmds, 8

*Topic normalize
  normalize, 10

Abiotic_data_sample, 2

Cell_number_sample, 3
Corr_data_sample, 5
correlation, 3
correlation, data.frame-method
  (correlation), 3
correlation, matrix-method
  (correlation), 3
cybar_plot, 5
cybar_plot, data.frame-method
  (cybar_plot), 5
cybar_plot, data.frame, matrix-method
  (cybar_plot), 5
cybar_plot, data.frame, missing-method
  (cybar_plot), 5
cybar_plot, matrix-method
  (cybar_plot), 5
cybar_plot, matrix, matrix-method
  (cybar_plot), 5
cybar_plot, missing-method
  (cybar_plot), 5

cybar_plot, missing, data.frame-method
  (cybar_plot), 5

cybar_plot, missing, matrix-method
  (cybar_plot), 5

flowCyBar (flowCyBar-package), 2
flowCyBar-package, 2
nmds, 8
nmds, data.frame-method (nmds), 8
nmds, matrix-method (nmds), 8
normalize, 10
normalize, data.frame-method
  (normalize), 10
normalize, matrix-method (normalize), 10