

Package ‘cmahalanobis’

September 14, 2025

Type Package

Title Calculate Distance Measures for DataFrames

Version 1.0.0

Date 2025-09-13

Maintainer Flavio Gioia <flaviogioia.fg@gmail.com>

Description It provides functions that calculate Mahalanobis distance, Euclidean distance, Manhattan distance, Chebyshev distance, Hamming distance, Canberra distance, Minkowski dissimilarity (distance defined for $p \geq 1$), Cosine dissimilarity, Bhattacharyya dissimilarity, Jaccard distance, Hellinger distance, Bray-Curtis dissimilarity, Sorensen-Dice dissimilarity between each pair of species in a list of data frames. These statistics are fundamental in various fields, such as cluster analysis, classification, and other applications of machine learning and data mining, where assessing similarity or dissimilarity between data is crucial. The package is designed to be flexible and easily integrated into data analysis workflows, providing reliable tools for evaluating distances in multidimensional contexts.

License GPL-3

Encoding UTF-8

Imports stats, ggplot2, reshape2, gridExtra, matrixStats

Suggests rmarkdown, testthat (>= 3.0.0)

NeedsCompilation no

RoxygenNote 7.3.3

Author Flavio Gioia [aut, cre] (ORCID:
<<https://orcid.org/0009-0000-0326-3840>>)

Repository CRAN

Date/Publication 2025-09-14 11:30:19 UTC

Contents

cbhattacharyya	2
cbraycurtis	4
ccanberra	5
cchebyshev	6

ccosine	8
ceulide	9
chamming	11
chellinger	12
cjaccard	13
cmahalanobis	15
cmanhattan	16
cminkowski	18
csorensendice	19
generate_report_cbhattacharyya	21
generate_report_cbraycurtis	22
generate_report_ccanberra	23
generate_report_cchebyshev	24
generate_report_ccosine	25
generate_report_ceulide	27
generate_report_chamming	28
generate_report_chellinger	29
generate_report_cjaccard	30
generate_report_cmahalanobis	31
generate_report_cmanhattan	32
generate_report_cminkowski	34
generate_report_csorensendice	35
pvaluescbatt	36
pvaluescbreu	37
pvaluesccanb	39
pvaluesccheb	40
pvaluesccosi	41
pvaluesceucl	42
pvalueschamm	43
pvalueschell	45
pvaluescjacc	46
pvaluescmaha	47
pvaluescmanh	48
pvaluescmink	49
pvaluescsore	51

Index 53

cbhattacharyya	<i>Calculate the Bhattacharyya dissimilarities for each pair of factors or for the index.</i>
----------------	---

Description

This function takes a dataframe and a variable or variables (two or more) in input, and returns a matrix or matrices (two or more) with the Bhattacharyya dissimilarities about the factors inside them. You can also select "index" to calculate the Bhattacharyya dissimilarities between each row.

Usage

```
cbhattacharyya(  
  dataset,  
  formula,  
  plot = TRUE,  
  plot_title = "Bhattacharyya Dissimilarity Between Groups",  
  min_group_size = 3  
)
```

Arguments

dataset	A dataframe.
formula	The index of the dataframe, otherwise a variable or variables (two or more) with factors which you want to calculate the Bhattacharyya dissimilarities matrix or matrices (two or more).
plot	Logical, if TRUE, a plot or plots (two or more) of the Bhattacharyya dissimilarities matrix or matrices about factors (two or more) are displayed.
plot_title	If plot is TRUE, the title to be used for plot or plots about factors. The default value is TRUE.
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded. For "index", this value is always 1.

Value

According to the option chosen in formula, with "index" the Bhattacharyya dissimilarities matrix will be printed; instead, by specifying variables, the Bhattacharyya dissimilarities matrix or matrices (two or more) between each pair of groups and, optionally, the plot or plots (two or more) will be printed.

Note

If "index" is selected with variables, only dissimilarities between rows are calculated. Therefore, this snippet: "cbhattacharyya(mtcars, ~am + carb + index)" will print dissimilarities only considering "index". Rows with NA values are omitted.

Examples

```
# Example with the iris dataset  
  
data(iris)  
  
cbhattacharyya(iris, ~Species, plot = TRUE,  
plot_title = "Bhattacharyya Dissimilarity Between Groups")  
  
# Example with the mtcars dataset  
  
data(mtcars)
```

```
cbhattacharyya(mtcars, ~am, plot = TRUE,
plot_title = "Bhattacharyya Dissimilarity Between Groups")

# Calculate Bhattacharyya distance for index
res <- cbhattacharyya(mtcars, ~index)
```

cbraycurtis	<i>Calculate the Bray-Curtis dissimilarities for each pair of factors or for the index.</i>
-------------	---

Description

This function takes a dataframe and a variable or variables (two or more) in input, and returns a matrix or matrices (two or more) with the Bray-Curtis dissimilarities about the factors inside them. You can also select "index" to calculate the Bray-Curtis dissimilarities between each row.

Usage

```
cbraycurtis(
  dataset,
  formula,
  plot = TRUE,
  plot_title = "Bray-Curtis Dissimilarity Between Groups",
  min_group_size = 3
)
```

Arguments

dataset	A dataframe.
formula	The index of the dataframe, otherwise a variable or variables (two or more) with factors which you want to calculate the Bray-Curtis dissimilarities matrix or matrices (two or more).
plot	Logical, if TRUE, a plot or plots (two or more) of the Bray-Curtis dissimilarities matrix or matrices about factors (two or more) are displayed.
plot_title	If plot is TRUE, the title to be used for plot or plots about factors. The default value is TRUE.
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded. For "index", this value is always 1.

Value

According to the option chosen in formula, with "index" the Bray-Curtis dissimilarities matrix will be printed; instead, by specifying variables, the Bray-Curtis dissimilarities matrix or matrices (two or more) between each pair of factors and, optionally, the plot or plots (two or more) will be printed.

Note

If "index" is selected with variables, only dissimilarities between rows are calculated. Therefore, this snippet: "cbraycurtis(mtcars, ~am + carb + index)" will print dissimilarities only considering "index". Rows with NA values are omitted.

Examples

```
# Example with the iris dataset
data(iris)

cbraycurtis(iris, ~Species, plot = TRUE,
plot_title = "Bray-Curtis Dissimilarity Between Groups")

# Example with mtcars dataset
data(mtcars)

# Example with the mtcars dataset
cbraycurtis(mtcars, ~am,
plot = TRUE, plot_title = "Bray-Curtis Dissimilarity Between Groups")

# Calculate the Bray-Curtis dissimilarity for 32 car models in "mtcars" dataset
res <- cbraycurtis(mtcars, ~index)
```

ccanberra

Calculate the Canberra distances for each pair of factors or for the index.

Description

This function takes a dataframe and a variable or variables (two or more) in input, and returns a matrix or matrices (two or more) with the Canberra distances about the factors inside them. You can also select "index" to calculate the Canberra distances between each row.

Usage

```
ccanberra(
  dataset,
  formula,
  plot = TRUE,
  plot_title = "Canberra Distance Between Groups",
  min_group_size = 3
)
```

Arguments

dataset A dataframe.

formula	The index of the dataframe, otherwise a variable or variables (two or more) with factors which you want to calculate the Canberra distances matrix or matrices (two or more).
plot	Logical, if TRUE, a plot or plots (two or more) of the Canberra distances matrix or matrices about factors (two or more) are displayed.
plot_title	If plot is TRUE, the title to be used for plot or plots about factors. The default value is TRUE.
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded. For "index", this value is always 1.

Value

According to the option chosen in formula, with "index" the Canberra distances matrix will be printed; instead, by specifying variables, the Canberra distances matrix or matrices (two or more) between each pair of groups and, optionally, the plot or plots (two or more) will be printed.

Note

If "index" is selected with variables, only distances between rows are calculated. Therefore, this snippet: `"ccanberra(mtcars, ~am + carb + index)"` will print distances only considering "index". Rows with NA values are omitted.

Examples

```
# Example with the iris dataset

data(iris)

ccanberra(iris, ~Species, plot = TRUE,
plot_title = "Canberra Distance Between Groups")

# Example with the mtcars dataset

data(mtcars)

ccanberra(mtcars, ~am, plot = TRUE,
plot_title = "Canberra Distance Between Groups")

# Calculate the Canberra distance for 32 car models in "mtcars" dataset
res <- ccanberra(mtcars, ~index)
```

Description

This function takes a dataframe and a variable or variables (two or more) in input, and returns a matrix or matrices (two or more) with the Chebyshev distances about the factors inside them. You can also select "index" to calculate the Chebyshev distances between each row.

Usage

```
cchebyshev(  
  dataset,  
  formula,  
  plot = TRUE,  
  plot_title = "Chebyshev Distance Between Groups",  
  min_group_size = 3  
)
```

Arguments

dataset	A dataframe.
formula	The index of the dataframe, otherwise a variable or variables (two or more) with factors which you want to calculate the Chebyshev distances matrix or matrices (two or more).
plot	Logical, if TRUE, a plot or plots (two or more) of the Chebyshev distances matrix or matrices about factors (two or more) are displayed.
plot_title	If plot is TRUE, the title to be used for plot or plots about factors. The default value is TRUE.
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded. For "index", this value is always 1.

Value

According to the option chosen in formula, with "index" the Chebyshev distances matrix will be printed; instead, by specifying variables, the Chebyshev distances matrix or matrices (two or more) between each pair of groups and, optionally, the plot or plots (two or more) will be printed.

Note

If "index" is selected with variables, only distances between rows are calculated. Therefore, this snippet: "cchebyshev(mtcars, ~am + carb + index)" will print distances only considering "index". Rows with NA values are omitted.

Examples

```
# Example with iris dataset  
  
data(iris)  
  
cchebyshev(iris, ~Species, plot = TRUE,
```

```

plot_title = "Chebyshev Distance Between Groups")

# Example with mtcars dataset

data(mtcars)

cchebyshev(mtcars, ~am, plot = TRUE,
plot_title = "Chebyshev Distance Between Groups")

# Calculate the Chebyshev distance for 32 car models in "mtcars" dataset
res <- cchebyshev(mtcars, ~index)

```

ccosine	<i>Calculate the Cosine dissimilarities for each pair of factors or for the index.</i>
---------	--

Description

This function takes a dataframe and a variable or variables (two or more) in input, and returns a matrix or matrices (two or more) with the Cosine dissimilarities about the factors inside them. You can also select "index" to calculate the Cosine dissimilarities between each row.

Usage

```

ccosine(
  dataset,
  formula,
  plot = TRUE,
  plot_title = "Cosine Dissimilarity Between Groups",
  min_group_size = 3
)

```

Arguments

dataset	A dataframe.
formula	The index of the dataframe, otherwise a variable or variables (two or more) with factors which you want to calculate the Cosine dissimilarities matrix or matrices (two or more).
plot	Logical, if TRUE, a plot or plots (two or more) of the Cosine dissimilarities matrix or matrices about factors (two or more) are displayed.
plot_title	If plot is TRUE, the title to be used for plot or plots about factors. The default value is TRUE.
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded. For "index", this value is always 1.

Value

According to the option chosen in formula, with "index" the Cosine dissimilarities matrix will be printed; instead, by specifying variables, the Cosine dissimilarities matrix or matrices (two or more) between each pair of groups and, optionally, the plot or plots (two or more) will be printed.

Note

If "index" is selected with variables, only dissimilarities between rows are calculated. Therefore, this snippet: "ccosine(mtcars, ~am + carb + index)" will print dissimilarities only considering "index". Rows with NA values are omitted.

Examples

```
# Example with iris dataset

data(iris)

ccosine(iris, ~Species, plot = TRUE,
plot_title = "Cosine Dissimilarity Between Groups")

# Example with mtcars dataset

data(mtcars)

ccosine(mtcars, ~am, plot = TRUE,
plot_title = "Cosine Dissimilarity Between Groups")

# Calculate the Cosine dissimilarity for 32 car models in "mtcars" dataset
res <- ccosine(mtcars, ~index)
```

ceulide	<i>Calculate the Euclidean distances for each pair of factors or for the index.</i>
---------	---

Description

This function takes a dataframe and a variable or variables (two or more) in input, and returns a matrix or matrices (two or more) with the Euclidean distances about each pair of factors inside them. You can also select "index" to calculate the Euclidean distances between each row.

Usage

```
ceulide(
  dataset,
  formula,
  plot = TRUE,
  plot_title = "Euclidean Distance Between Groups",
  min_group_size = 3
)
```

Arguments

dataset	A dataframe.
formula	The index of the dataframe, otherwise a variable or variables (two or more) with factors which you want to calculate the Euclidean distances matrix or matrices (two or more).
plot	Logical, if TRUE, a plot or plots (two or more) of the Euclidean distances matrix or matrices about factors (two or more) are displayed.
plot_title	If plot is TRUE, the title to be used for plot or plots about factors. The default value is TRUE.
min_group_size	Minimum group size to maintain. The default value is 3, therefore factors, inside variables, with less than 3 observations will be discarded. For "index", this value is always 1.

Value

According to the option chosen in formula, with "index" the Euclidean distance matrix will be printed; instead, by specifying variables, the Euclidean distances matrix or matrices (two or more) between each pair of groups and, optionally, the plot or plots (two or more) will be printed.

Note

If "index" is selected with variables, only distances between rows are calculated. Therefore, this snippet: "ceulide(mtcars, ~am + carb + index)" will print distances only considering "index". Rows with NA values are omitted.

Examples

```
# Example with iris dataset

data(iris)

ceulide(iris, ~Species, plot = TRUE,
plot_title = "Euclidean Distance Between Groups", min_group_size = 2)

# Example with mtcars dataset

data(mtcars)

ceulide(mtcars, ~am + carb, plot = TRUE,
plot_title = "Euclidean Distance Between Groups", min_group_size = 3)

# Calculate ceulide for index
res <- ceulide(mtcars, ~index,
min_group_size = 3)
```

chamming	<i>Calculate the Hamming distances for each pair of factors or for the index.</i>
----------	---

Description

This function takes a dataframe and a variable or variables (two or more) in input, and returns a matrix or matrices (two or more) with the Hamming distances about the factors inside them. You can also select "index" to calculate the Hamming distances between each row.

Usage

```
chamming(  
  dataset,  
  formula,  
  plot = TRUE,  
  plot_title = "Hamming Distance Between Groups",  
  min_group_size = 3  
)
```

Arguments

dataset	A dataframe.
formula	The index of the dataframe, otherwise a variable or variables (two or more) with factors which you want to calculate the Hamming distances matrix or matrices (two or more).
plot	Logical, if TRUE, a plot or plots (two or more) of the Hamming distances matrix or matrices about factors (two or more) are displayed.
plot_title	If plot is TRUE, the title to be used for plot or plots about factors. The default value is TRUE.
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded. For "index", this value is always 1.

Value

According to the option chosen in formula, with "index" the Hamming distances matrix will be printed; instead, by specifying variables, the Hamming distances matrix or matrices (two or more) between each pair of factors and, optionally, the plot or plots (two or more) will be printed.

Note

If "index" is selected with variables, only distances between rows are calculated. Therefore, this snippet: "chamming(mtcars, ~am + carb + index)" will print the distances only considering "index". Rows with NA values are omitted.

Examples

```
# Example with iris dataset

data(iris)

chamming(iris, ~Species, plot = TRUE,
plot_title = "Hamming Distance Between Groups")

# Example with mtcars dataset

data(mtcars)

chamming(mtcars, ~am, plot = TRUE,
plot_title = "Hamming Distance Between Groups")

# Calculate the Hamming distance for 32 car models in "mtcars" dataset
res <- chamming(mtcars, ~index)
```

chellinger

Calculate the Hellinger distances for each pair of factors or for the index.

Description

This function takes a dataframe and a variable or variables (two or more) in input, and returns a matrix or matrices (two or more) with the Hellinger distances about the factors inside them. You can also select "index" to calculate the Hellinger distances between each row.

Usage

```
chellinger(
  dataset,
  formula,
  plot = TRUE,
  plot_title = "Hellinger Distance Between Groups",
  min_group_size = 3
)
```

Arguments

dataset	A dataframe.
formula	The index of the dataframe, otherwise a variable or variables (two or more) with factors which you want to calculate the Hellinger distances matrix or matrices (two or more).
plot	Logical, if TRUE, a plot or plots (two or more) of the Hellinger distances matrix or matrices about factors (two or more) are displayed.

<code>plot_title</code>	If plot is TRUE, the title to be used for plot or plots about factors. The default value is TRUE.
<code>min_group_size</code>	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded. For "index", this value is always 1.

Value

According to the option chosen in formula, with "index" the Hellinger distances matrix will be printed; instead, by specifying variables, the Hellinger distances matrix or matrices (two or more) between each pair of groups and, optionally, the plot or plots (two or more) will be printed.

Note

If "index" is selected with variables, only distances between rows are calculated. Therefore, this snippet: `chellinger(mtcars, ~am + carb + index)` will print the distances only considering "index". Rows with NA values are omitted.

Examples

```
# Example with the iris dataset

data(iris)

chellinger(iris, ~Species, plot = TRUE,
plot_title = "Hellinger Distance Between Groups")

# Example with the mtcars dataset

data(mtcars)

chellinger(mtcars, ~am, plot = TRUE,
plot_title = "Hellinger Distance Between Groups")

res <- kellinger(mtcars, ~index)
```

cjaccard

Calculate the Jaccard distances for each pair of factors or for the index.

Description

This function takes a dataframe and a variable or variables (two or more) in input, and returns a matrix or matrices (two or more) with the Jaccard distances about the factors inside them. You can also select "index" to calculate the Jaccard distances between each row.

Usage

```

cjaccard(
  dataset,
  formula,
  plot = TRUE,
  plot_title = "Jaccard Distance Between Groups",
  min_group_size = 3
)

```

Arguments

dataset	A dataframe.
formula	The index of the dataframe, otherwise a variable or variables (two or more) with factors which you want to calculate the Jaccard distances matrix or matrices (two or more).
plot	Logical, if TRUE, a plot or plots (two or more) of the Jaccard distances matrix or matrices about factors (two or more) are displayed.
plot_title	If plot is TRUE, the title to be used for plot or plots about factors. The default value is TRUE.
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded. For "index", this value is always 1.

Value

According to the option chosen in formula, with "index" the Jaccard distances matrix will be printed; instead, by specifying variables, the Jaccard distances matrix or matrices (two or more) between each pair of groups and, optionally, the plot or plots (two or more) will be printed.

Note

If "index" is selected with variables, only distances between rows are calculated. Therefore, this snippet: "cjaccard(mtcars, ~am + carb + index)" will print distances only considering "index". Rows with NA values are omitted.

Examples

```

# Example with the iris dataset

data(iris)

cjaccard(iris, ~Species, plot = TRUE,
plot_title = "Jaccard Distance Between Groups")

# Example with the mtcars dataset

data(mtcars)

cjaccard(mtcars, ~am,

```

```

plot = TRUE, plot_title = "Jaccard Distance Between Groups")

res <- cjaccard(mtcars, ~index,
plot = TRUE)

```

cmahalanobis	<i>Calculate the Mahalanobis distances for each pair of factors or for the index.</i>
--------------	---

Description

This function takes a dataframe and a variable or variables (two or more) in input, and returns a matrix or matrices (two or more) with the Mahalanobis distances about each pair of factors inside them. You can also select "index" to calculate the Mahalanobis distances between each row.

Usage

```

cmahalanobis(
  dataset,
  formula,
  plot = TRUE,
  plot_title = "Mahalanobis Distance Between Groups",
  min_group_size = 3,
  pvalues_chisq = FALSE
)

```

Arguments

dataset	A dataframe.
formula	The index of the dataframe, otherwise a variable or variables (two or more) with factors which you want to calculate the Mahalanobis distances matrix or matrices (two or more).
plot	Logical, if TRUE, a plot or plots (two or more) of the Mahalanobis distances matrix or matrices about factors (two or more) are displayed.
plot_title	If plot is TRUE, the title to be used for plot or plots about factors. The default value is TRUE.
min_group_size	Minimum group size to maintain. The default value is 3, therefore factors, inside variables, with less than 3 observations will be discarded. For "index", this value is always 1.
pvalues_chisq	If TRUE, print the result of the chi-squared test on squared distances. The distances with "pvalues_chisq = FALSE" are not squared; instead, with "pvalues_chisq = TRUE", the squared Mahalanobis distances with corresponding p_values will be printed. Default is FALSE.

Value

According to the option chosen in formula and in `pvalues_chisq`, with "index" and "pvalues_chisq = TRUE" the squared Mahalanobis distance matrix will be printed with corresponding pvalues; instead, with "index" and "pvalues_chisq = FALSE", only the Mahalanobis distances (not squared) will be printed. By specifying variables, the Mahalanobis distances matrix or matrices (two or more) between each pair of factors and, optionally, the plot or plots (two or more) will be printed.

Note

If "index" is selected with variables, only distances between rows are calculated. Therefore, this snippet: `"cmahalanobis(mtcars, ~am + carb + index)"` will print distances and plot only considering "index". Rows with NA values are omitted.

Examples

```
# Example with the iris dataset

data(iris)

# Calculate the Mahalanobis distance for "Species" groups in "iris" dataset
cmahalanobis(iris, ~Species, plot = TRUE,
plot_title = "Mahalanobis Distance Between Groups", min_group_size = 3)

# Example with the mtcars dataset
data(mtcars)

# Calculate the Mahalanobis distance for two factors in "mtcars" dataset
cmahalanobis(mtcars, ~am + vs,
plot = TRUE, plot_title = "Mahalanobis Distance Between Groups",
min_group_size = 2, pvalues_chisq = TRUE)

# Calculate the Mahalanobis distance for "index" in mtcars
cmahalanobis(mtcars, ~index, pvalues_chisq = TRUE)
```

cmanhattan

Calculate the Manhattan distances for each pair of factors or for the index.

Description

This function takes a dataframe and a variable or variables (two or more) in input, and returns a matrix or matrices (two or more) with the Manhattan distances about the factors inside them. You can also select "index" to calculate the Manhattan distances between each row.

Usage

```
cmanhattan(  
  dataset,  
  formula,  
  plot = TRUE,  
  plot_title = "Manhattan Distance Between Groups",  
  min_group_size = 3  
)
```

Arguments

dataset	A dataframe.
formula	The index of the dataframe, otherwise a variable or variables (two or more) with factors which you want to calculate the Manhattan distances matrix or matrices (two or more).
plot	Logical, if TRUE, a plot or plots (two or more) of the Manhattan distances matrix or matrices about factors (two or more) are displayed.
plot_title	If plot is TRUE, the title to be used for plot or plots about factors. The default value is TRUE.
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded. For "index", this value is always 1.

Value

According to the option chosen in formula, with "index" the Manhattan distances matrix will be printed; instead, by specifying variables, the Manhattan distances matrix or matrices (two or more) between each pair of groups and, optionally, the plot or plots (two or more) will be printed.

Note

If "index" is selected with variables, only distances between rows are calculated. Therefore, this snippet: "cmanhattan(mtcars, ~am + carb + index)" will print the distances only considering "index". Rows with NA values are omitted.

Examples

```
# Example with iris dataset  
  
data(iris)  
  
cmanhattan(iris, ~Species, plot = TRUE,  
plot_title = "Manhattan Distance Between Groups", min_group_size = 3)  
  
# Example with mtcars dataset  
  
data(mtcars)  
  
cmanhattan(mtcars, ~am + vs, plot = TRUE,
```

```
plot_title = "Manhattan Distance Between Groups", min_group_size = 3)

# Calculate the Manhattan distance for 32 car models in "mtcars" dataset
res <- cmanhattan(mtcars, ~index, min_group_size = 3)
```

cminkowski

Calculate the Minkowski distances for each pair of factors or for the index.

Description

This function takes a dataframe and a variable or variables (two or more) in input, and returns a matrix or matrices (two or more) with the Minkowski distances about the factors inside them. You can also select "index" to calculate the Minkowski distances between each row.

Usage

```
cminkowski(
  dataset,
  formula,
  p = 3,
  plot = TRUE,
  plot_title = "Minkowski Distance Between Groups",
  min_group_size = 3
)
```

Arguments

dataset	A dataframe.
formula	The index of the dataframe, otherwise a variable or variables (two or more) with factors which you want to calculate the Minkowski distances matrix or matrices (two or more).
p	Order of the Minkowski distance.
plot	Logical, if TRUE, a plot or plots (two or more) of the Minkowski distances matrix or matrices about factors (two or more) are displayed.
plot_title	If plot is TRUE, the title to be used for plot or plots about factors. The default value is TRUE.
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded. For "index", this value is always 1.

Value

According to the option chosen in formula, with "index" the Minkowski distances matrix will be printed; instead, by specifying variables, the Minkowski distances matrix or matrices (two or more) between each pair of groups and, optionally, the plot or plots (two or more) will be printed.

Note

When $p < 1$, the Minkowski distance is a "dissimilarity" measure. When $p \geq 1$, the triangle inequality property is satisfied and we say "Minkowski distance". If "index" is selected with variables, only distances between rows are calculated. Therefore, this snippet: "cminkowski(mtcars, ~am + carb + index)" will print distances only considering "index". Rows with NA values are omitted.

Examples

```
# Example with iris dataset

data(iris)

cminkowski(iris, ~Species, p = 3, plot = TRUE,
plot_title = "Minkowski Distance Between Groups")

# Example with mtcars dataset

data(mtcars)

cminkowski(mtcars, ~am, p = 3, plot = TRUE,
plot_title = "Minkowski Distance Between Groups")

# Calculate the Minkowski distance for 32 car models in "mtcars" dataset
res <- cminkowski(mtcars, ~index, p = 2, plot = TRUE)
```

csorensendice

Calculate the Sorensen-Dice dissimilarities for each pair of factors or for the index.

Description

This function takes a dataframe and a variable or variables (two or more) in input, and returns a matrix or matrices (two or more) with the Sorensen-Dice dissimilarities about the factors inside them. You can also select "index" to calculate the Sorensen-Dice dissimilarities between each row.

Usage

```
csorensendice(
  dataset,
  formula,
  plot = TRUE,
  plot_title = "Sorensen-Dice Dissimilarity Between Groups",
  min_group_size = 3
)
```

Arguments

dataset	A dataframe.
formula	The index of the dataframe, otherwise a variable or variables (two or more) with factors which you want to calculate the Sorensen-Dice dissimilarities matrix or matrices (two or more).
plot	Logical, if TRUE, a plot or plots (two or more) of the Sorensen-Dice dissimilarities matrix or matrices about factors (two or more) are displayed.
plot_title	If plot is TRUE, the title to be used for plot or plots about factors. The default value is TRUE.
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded. For "index", this value is always 1.

Value

According to the option chosen in formula, with "index" the Sorensen-Dice dissimilarities matrix will be printed; instead, by specifying variables, the Sorensen-Dice dissimilarities matrix or matrices (two or more) between each pair of groups and, optionally, the plot or plots (two or more) will be printed.

Note

If "index" is selected with variables, only dissimilarities between rows are calculated. Therefore, this snippet: "csorensendice(mtcars, ~am + carb + index)" will print dissimilarities only considering "index". Rows with NA values are omitted.

Examples

```
# Example with the iris dataset
data(iris)

csorensendice(iris, ~Species,
plot = TRUE, plot_title = "Sorensen-Dice Dissimilarity Between Groups")

# Example with mtcars dataset
data(mtcars)

# Example with the mtcars dataset
csorensendice(mtcars, ~am, plot = TRUE,
plot_title = "Sorensen-Dice Dissimilarity Between Groups")

# Calculate the Sorensen-Dice dissimilarity for 32 car models in "mtcars" dataset
res <- csorensendice(mtcars, ~index)
```

```
generate_report_cbhattacharyya
```

Generate a Microsoft Word document about the Bhattacharyya dissimilarities matrix or matrices (two or more) and the p-values matrix or matrices (two or more).

Description

This function takes a dataframe, a factor or factors (two or more) and returns a Microsoft Word document about the Bhattacharyya dissimilarities matrix or matrices (two or more) and the p-values matrix or matrices (two or more).

Usage

```
generate_report_cbhattacharyya(
  dataset,
  formula,
  pvalue.method = "permutation",
  seed = NULL,
  min_group_size = 3
)
```

Arguments

dataset	A dataframe.
formula	A variable or variables (two or more) with factors which you want to calculate the Bhattacharyya dissimilarities matrix or matrices (two or more) and the p_values matrix or matrices (two or more).
pvalue.method	A p_value method used to calculate the matrix or matrices (two or more), the default value is "permutation". Another method is "bootstrap".
seed	Optionally, set a seed for "bootstrap" or "permutation".
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded.

Value

A Microsoft Word document about the Bhattacharyya dissimilarities matrix or matrices (two or more) and the p_values matrix or matrices (two or more).

Examples

```
# Example with iris dataset
data(iris)

# Generate a report about "Species" factor in iris dataset
generate_report_cbhattacharyya(iris, ~Species,
```

```

pvalue.method = "permutation")

# Example with mtcars dataset
data(mtcars)

# Generate a report about "am" factor in mtcars dataset
generate_report_cbhattacharyya(mtcars, ~am,
pvalue.method = "bootstrap", seed = 123)

```

```
generate_report_cbraycurtis
```

Generate a Microsoft Word document about the Bray-Curtis dissimilarities matrix or matrices (two or more) and the p-values matrix or matrices (two or more).

Description

This function takes a dataframe, a factor or factors (two or more) and returns a Microsoft Word document about the Bray-Curtis dissimilarities matrix or matrices (two or more) and the p-values matrix or matrices (two or more).

Usage

```

generate_report_cbraycurtis(
  dataset,
  formula,
  pvalue.method = "permutation",
  seed = NULL,
  min_group_size = 3
)

```

Arguments

dataset	A dataframe.
formula	A variable or variables (two or more) with factors which you want to calculate the Bray-Curtis dissimilarities matrix or matrices (two or more) and the p_values matrix or matrices (two or more).
pvalue.method	A p_value method used to calculate the matrix or matrices (two or more), the default value is "permutation". Another method is "bootstrap".
seed	Optionally, set a seed for 'bootstrap' or 'permutation'.
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded.

Value

A Microsoft Word document about the Bray-Curtis dissimilarities matrix or matrices (two or more) and the p_values matrix or matrices (two or more).

Examples

```
# Example with iris dataset
data(iris)

# Generate a report about "Species" factor in iris dataset
generate_report_cbraycurtis(iris, ~Species,
pvalue.method = "permutations")

# Example with mtcars dataset
data(mtcars)

# Generate a report about "am" factor in mtcars dataset
generate_report_cbraycurtis(mtcars, ~am,
pvalue.method = 'bootstrap', seed = 124)
```

```
generate_report_ccanberra
```

Generate a Microsoft Word document about the Canberra distances matrix or matrices (two or more) and the p-values matrix or matrices (two or more).

Description

This function takes a dataframe, a factor or factors (two or more) and returns a Microsoft Word document about the Canberra distances matrix or matrices (two or more) and the p-values matrix or matrices (two or more).

Usage

```
generate_report_ccanberra(
  dataset,
  formula,
  pvalue.method = "permutation",
  seed = NULL,
  min_group_size = 3
)
```

Arguments

dataset	A dataframe.
formula	A variable or variables (two or more) with factors which you want to calculate the Canberra distances matrix or matrices (two or more) and the p_values matrix or matrices (two or more).
pvalue.method	A p_value method used to calculate the matrix or matrices (two or more), the default value is "permutation". Another method is "bootstrap".
seed	Optionally, set a seed for "bootstrap" and "permutation".

`min_group_size` Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded.

Value

A Microsoft Word document about the Canberra distances matrix or matrices (two or more) and the `p_values` matrix or matrices (two or more).

Examples

```
# Example with iris dataset
data(iris)

# Generate a report about "Species" factor in iris dataset
generate_report_ccanberra(iris, ~Species,
pvalue.method = "permutation")

# Example with mtcars dataset
data(mtcars)

# Generate a report about "am" factor in mtcars dataset
generate_report_ccanberra(mtcars, ~am,
pvalue.method = "bootstrap", seed = 123)

# Generate a report for 32 car models in "mtcars" dataset,
# using "bootstrap" method
generate_report_ccanberra(mtcars, ~am, pvalue.method = "bootstrap")
```

```
generate_report_cchebyshev
```

Generate a Microsoft Word document about the Chebyshev distances matrix or matrices (two or more) and the p-values matrix or matrices (two or more).

Description

This function takes a dataframe, a factor or factors (two or more) and returns a Microsoft Word document about the Chebyshev distances matrix or matrices (two or more) and the p-values matrix or matrices (two or more).

Usage

```
generate_report_cchebyshev(
  dataset,
  formula,
  pvalue.method = "permutation",
  seed = NULL,
  min_group_size = 3
)
```


Arguments

dataset	A dataframe.
formula	A variable or variables (two or more) with factors which you want to calculate the Chebyshev distances matrix or matrices (two or more) and the p_values matrix or matrices (two or more).
pvalue.method	A p_value method used to calculate the matrix or matrices (two or more), the default value is "permutation". Another method is "bootstrap".
seed	Optionally, set a seed for "bootstrap" and "permutation".
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded.

Value

A Microsoft Word document about the Chebyshev distances matrix or matrices (two or more) and the p_values matrix or matrices (two or more).

Examples

```
# Example with iris dataset
data(iris)

# Generate a report about "Species" factor in iris dataset
generate_report_cchebyshev(iris, ~Species, pvalue.method = "permutation")

# Example with mtcars dataset
data(mtcars)

# Generate a report about "am" factor in mtcars dataset
generate_report_cchebyshev(mtcars, ~am,
pvalue.method = "bootstrap", seed = 100)
```

```
generate_report_ccosine
```

Generate a Microsoft Word document about the Cosine dissimilarities matrix or matrices (two or more) and the p-values matrix or matrices (two or more).

Description

This function takes a dataframe, a factor or factors (two or more) and returns a Microsoft Word document about the Cosine dissimilarities matrix or matrices (two or more) and the p-values matrix or matrices (two or more).

Usage

```
generate_report_ccosine(  
  dataset,  
  formula,  
  pvalue.method = "permutation",  
  seed = NULL,  
  min_group_size = 3  
)
```

Arguments

dataset	A dataframe.
formula	A variable or variables (two or more) with factors which you want to calculate the Cosine dissimilarities matrix or matrices (two or more) and the p_values matrix or matrices (two or more).
pvalue.method	A p_value method used to calculate the matrix or matrices (two or more), the default value is "permutation". Another method is "bootstrap".
seed	Optionally, set a seed for "bootstrap" or "permutation".
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded.

Value

A Microsoft Word document about the Cosine dissimilarities matrix or matrices (two or more) and the p_values matrix or matrices (two or more).

Examples

```
# Example with iris dataset  
data(iris)  
  
# Generate a report about "Species" factor in iris dataset  
generate_report_ccosine(iris, ~Species, pvalue.method = "permutation")  
  
# Example with mtcars dataset  
data(mtcars)  
  
# Generate a report about "am" factor in mtcars dataset  
generate_report_ccosine(mtcars, ~am,  
  pvalue.method = "bootstrap", seed = 123)
```

generate_report_ceuclide

Generate a Microsoft Word document about the Euclidean distances matrix or matrices and the p-values matrix or matrices.

Description

This function takes a dataframe, a factor or factors (two or more) and returns a Microsoft Word document about the Euclidean distances matrix or matrices (two or more) and the p-values matrix or matrices (two or more).

Usage

```
generate_report_ceuclide(
  dataset,
  formula,
  pvalue.method = "permutation",
  seed = NULL,
  min_group_size = 3
)
```

Arguments

dataset	A dataframe.
formula	A variable or variables (two or more) with factors which you want to calculate the Euclidean distances matrix or matrices (two or more) and the p_values matrix or matrices (two or more).
pvalue.method	A p_value method used to calculate the matrix or matrices (two or more), the default value is "permutation". Another method is "bootstrap".
seed	Optionally, set a seed for "bootstrap" or "permutation".
min_group_size	Minimum group size to maintain. The default value is 3, therefore factors, inside variables, with less than 3 observations will be discarded.

Value

A Microsoft Word document about the Euclidean distances matrix or matrices (two or more) and the p_values matrix or matrices (two or more).

Examples

```
# Example with iris dataset
data(airquality)

# Generate a report about "Species" factor in iris dataset
generate_report_ceuclide(airquality, ~Month, pvalue.method = 'bootstrap',
min_group_size = 3)
```

```
# Example with mtcars dataset
data(mtcars)

# Generate a report about "am" and "vs" factors in mtcars dataset
generate_report_ceuclide(mtcars, ~am + vs,
  pvalue.method = 'bootstrap', seed = 100, min_group_size = 3)
```

```
generate_report_chamming
```

Generate a Microsoft Word document about the Hamming distances matrix or matrices (two or more) and the p-values matrix or matrices (two or more).

Description

This function takes a dataframe, a factor or factors (two or more) and returns a Microsoft Word document about the Hamming distances matrix or matrices (two or more) and the p-values matrix or matrices (two or more).

Usage

```
generate_report_chamming(
  dataset,
  formula,
  pvalue.method = "permutation",
  seed = NULL,
  min_group_size = 3
)
```

Arguments

dataset	A dataframe.
formula	A variable or variables (two or more) with factors which you want to calculate the Hamming distances matrix or matrices (two or more) and the p_values matrix or matrices (two or more).
pvalue.method	A p_value method used to calculate the matrix or matrices (two or more), the default value is "permutation". Another method is "bootstrap".
seed	Optionally, set a seed for "bootstrap" and "permutation".
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded.

Value

A Microsoft Word document about the Hamming distances matrix or matrices (two or more) and the p_values matrix or matrices (two or more).

Examples

```
# Example with iris dataset
data(iris)

# Generate a report about "Species" factor in iris dataset
generate_report_chamming(iris, ~Species)

# Example with mtcars dataset
data(mtcars)

# Generate a report about "am" factor in mtcars dataset
generate_report_chamming(mtcars, ~am,
pvalue.method = "bootstrap", seed = 124)
```

```
generate_report_chellinger
```

Generate a Microsoft Word document about the Hellinger distances matrix or matrices (two or more) and the p-values matrix or matrices (two or more).

Description

This function takes a dataframe, a factor or factors (two or more) and returns a Microsoft Word document about the Hellinger distances matrix or matrices (two or more) and the p-values matrix or matrices (two or more).

Usage

```
generate_report_chellinger(
  dataset,
  formula,
  pvalue.method = "permutation",
  seed = NULL,
  min_group_size = 3
)
```

Arguments

dataset	A dataframe.
formula	A variable or variables (two or more) with factors which you want to calculate the Hellinger distances matrix or matrices (two or more) and the p_values matrix or matrices (two or more).
pvalue.method	A p_value method used to calculate the matrix or matrices (two or more), the default value is "permutation". Another method is "bootstrap".
seed	Optionally, set a seed for 'bootstrap' and 'permutation'.
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded.

Value

A Microsoft Word document about the Hellinger distances matrix or matrices (two or more) and the p_values matrix or matrices (two or more).

Examples

```
# Example with iris dataset
data(iris)

# Generate a report about "Species" factor in iris dataset
generate_report_chellinger(iris, ~Species,
pvalue.method = "bootstrap")

# Example with mtcars dataset
data(mtcars)

# Generate a report about "am" factor in mtcars dataset
generate_report_chellinger(mtcars, ~am,
pvalue.method = "bootstrap", seed = 100)
```

```
generate_report_cjaccard
```

Generate a Microsoft Word document about the Jaccard distances matrix or matrices (two or more) and the p-values matrix or matrices (two or more).

Description

This function takes a dataframe, a factor and returns a Microsoft Word document about the Jaccard distances matrix or matrices and the p-values matrix or matrices.

Usage

```
generate_report_cjaccard(
  dataset,
  formula,
  pvalue.method = "permutation",
  seed = NULL,
  min_group_size = 3
)
```

Arguments

dataset	A dataframe.
formula	A variable or variables (two or more) with factors which you want to calculate the Jaccard distances matrix or matrices and the p_values matrix or matrices.

`pvalue.method` A p_value method used to calculate the matrix, the default value is "permutation". Another method is "bootstrap".

`seed` Optionally, set a seed for "bootstrap" or "permutation".

`min_group_size` Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded.

Value

A Microsoft Word document about the Jaccard distance matrix or matrices and the p_values matrix or matrices.

Examples

```
# Example with iris dataset
data(iris)

# Generate a report about "Species" factor in iris dataset
generate_report_cjaccard(iris, ~Species,
pvalue.method = "permutation")

# Example with mtcars dataset
data(mtcars)

# Generate a report about "am" factor in mtcars dataset
generate_report_cjaccard(mtcars, ~am,
pvalue.method = "bootstrap", seed = 223)
```

```
generate_report_cmahalanobis
```

Generate a Microsoft Word document about the Mahalanobis distances matrix or matrices and the p-values matrix or matrices.

Description

This function takes a dataframe, a factor or factors (two or more) and returns a Microsoft Word document about the Mahalanobis distances matrix or matrices (two or more) and the p-values matrix or matrices.

Usage

```
generate_report_cmahalanobis(
  dataset,
  formula,
  pvalue.method = "permutation",
  seed = NULL,
  min_group_size = 3,
  pvalues_chisq = FALSE
)
```

Arguments

dataset	A dataframe.
formula	A variable or variables (two or more) with factors which you want to calculate the Mahalanobis distances matrix or matrices (two or more) and the p_values matrix or matrices (two or more).
pvalue.method	A p_value method used to calculate the matrix or matrices (two or more), the default value is "permutation". Another method is "bootstrap".
seed	Optionally, set a seed for "bootstrap" or "permutation".
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded.
pvalues_chisq	If TRUE, print the result of the chi-squared test on squared distances. The resulting distances with "pvalues_chisq = FALSE" are not squared; instead, with "pvalues_chisq = TRUE", the squared Mahalanobis distance matrix with corresponding p_values will be printed. Default is FALSE.

Value

A Microsoft Word document about the Mahalanobis distances matrix or matrices (two or more) and the p_values matrix or matrices (two or more).

Examples

```
# Example with iris dataset
data(iris)

# Generate a report about the "Species"
# factor in the iris dataset using the "permutation" method.
generate_report_cmahalanobis(iris, ~Species, min_group_size = 3)

# Example with mtcars dataset
data(mtcars)

# Generate a report about the "am" and "vs" in mtcars using "bootstrap" method.
generate_report_cmahalanobis(mtcars, ~am + vs,
  pvalue.method = "bootstrap",
  seed = 100, min_group_size = 2)
```

generate_report_cmanhattan

Generate a Microsoft Word document about the Manhattan distances matrix or matrices (two or more) and the p-values matrix or matrices (two or more).

Description

This function takes a dataframe, a factor or factors (two or more) and returns a Microsoft Word document about the Manhattan distances matrix or matrices (two or more) and the p-values matrix or matrices (two or more).

Usage

```
generate_report_cmanhattan(  
  dataset,  
  formula,  
  pvalue.method = "permutation",  
  seed = NULL,  
  min_group_size = 3  
)
```

Arguments

dataset	A dataframe.
formula	A variable or variables (two or more) with factors which you want to calculate the Manhattan distances matrix or matrices (two or more) and the p_values matrix or matrices (two or more).
pvalue.method	A p_value method used to calculate the matrix or matrices (two or more), the default value is "permutation". Another method is "bootstrap".
seed	Optionally, set a seed for "bootstrap" and "permutation".
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded.

Value

A Microsoft Word document about the Manhattan distances matrix or matrices (two or more) and the p_values matrix or matrices (two or more).

Examples

```
# Example with iris dataset  
data(iris)  
  
# Generate a report about "Species" factor in iris dataset  
generate_report_cmanhattan(iris, ~Species, pvalue.method = "permutation")  
  
# Example with mtcars dataset  
data(mtcars)  
  
# Generate a report about "am" factor in mtcars dataset  
generate_report_cmanhattan(mtcars, ~am,  
  pvalue.method = 'bootstrap', seed = 123)
```

```
generate_report_cminkowski
```

Generate a Microsoft Word document about the Minkowski dissimilarities/distances matrix or matrices (two or more) and the p-values matrix or matrices (two or more).

Description

This function takes a dataframe, a factor or factors (two or more) and returns a Microsoft Word document about the Minkowski dissimilarities/distances matrix or matrices (two or more) and the p-values matrix or matrices (two or more).

Usage

```
generate_report_cminkowski(
  dataset,
  formula,
  p = 3,
  pvalue.method = "permutation",
  seed = NULL,
  min_group_size = 3
)
```

Arguments

dataset	A dataframe.
formula	A variable or variables (two or more) with factors which you want to calculate the Minkowski dissimilarities/distances matrix or matrices (two or more) and the p_values matrix or matrices (two or more).
p	Order of the Minkowski dissimilarities/distances. The default value is 3.
pvalue.method	A p_value method used to calculate the matrix or matrices (two or more), the default value is "permutation". Another method is "bootstrap".
seed	Optionally, set a seed for "bootstrap" and "permutation".
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded.

Details

When $p < 1$, the Minkowski distance is a "dissimilarity" measure. When $p \geq 1$, the triangle inequality property is satisfied and we say "Minkowski distance".

Value

A Microsoft Word document about the Minkowski dissimilarities/distances matrix or matrices (two or more) and the p_values matrix or matrices (two or more).

Examples

```
# Example with iris dataset
data(iris)

# Generate a report about "Species" factor in iris dataset
generate_report_cminkowski(iris, ~Species, p = 3,
pvalue.method = "permutation")

# Example with mtcars dataset
data(mtcars)

# Generate a report about "am" factor in mtcars dataset
generate_report_cminkowski(mtcars, ~am,
p = 3, pvalue.method = 'permutation', seed = 234)
```

```
generate_report_csorensendice
```

Generate a Microsoft Word document about the Sorensen-Dice dissimilarity matrix or matrices (two or more) and the p-values matrix or matrices (two or more).

Description

This function takes a dataframe, a factor or factors (two or more) and returns a Microsoft Word document about the Sorensen-Dice dissimilarities matrix or matrices (two or more) and the p-values matrix or matrices (two or more).

Usage

```
generate_report_csorensendice(
  dataset,
  formula,
  pvalue.method = "permutation",
  seed = NULL,
  min_group_size = 3
)
```

Arguments

dataset	A dataframe.
formula	A variable or variables (two or more) with factors which you want to calculate the Sorensen-Dice dissimilarities matrix or matrices (two or more) and the p_values matrix or matrices (two or more).
pvalue.method	A p_value method used to calculate the matrix or matrices (two or more), the default value is "permutation". Another method is "bootstrap".
seed	Optionally, set a seed for "bootstrap" or "permutation".

`min_group_size` Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded.

Value

A Microsoft Word document about the Sorensen-Dice dissimilarities matrix or matrices (two or more) and the `p_values` matrix or matrices (two or more).

Examples

```
# Example with iris dataset
data(iris)

# Generate a report about "Species" factor in iris dataset
generate_report_csorensendice(iris, ~Species,
pvalue.method = 'permutation')

# Example with mtcars dataset
data(mtcars)

# Generate a report about "am" factor in mtcars dataset
generate_report_csorensendice(mtcars, ~am,
pvalue.method = "bootstrap", seed = 123)
```

<code>pvaluescbatt</code>	<i>Calculate the <code>p_values</code> matrix or matrices (two or more) for each factor inside variable or variables (two or more), using Bhattacharyya dissimilarities as a base.</i>
---------------------------	--

Description

Using the Bhattacharyya dissimilarity for the dissimilarities calculation, this function takes a dataframe, a variable or variables (two or more), a `p_value` method such as "bootstrap" and "permutation" and returns the `p_values` matrix or matrices (two or more) between each pair of factors and a plot or plots (two or more) if the user select TRUE or leaves the parameter without argument.

Usage

```
pvaluescbatt(
  dataset,
  formula,
  pvalue.method = "permutation",
  plot = TRUE,
  seed = NULL,
  min_group_size = 3
)
```

Arguments

dataset	A dataframe.
formula	A variable or variables (two or more) with factors which you want to calculate Bhattacharyya dissimilarities matrix or matrices (two or more).
pvalue.method	A p_value method used to calculate the matrix or matrices (two or more), the default value is "permutation". Another method is "bootstrap".
plot	if TRUE, plot the p_values heatmap or heatmaps (two or more). The default value is TRUE.
seed	Optionally, set a seed for "bootstrap" or "permutation".
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded.

Value

A list containing a matrix or matrices (two or more) of p_values and, optionally, the plot.

Examples

```
# Example with iris dataset
data(iris)

# Calculate p_values of "Species" variable in iris dataset
pvaluescbatt(iris, ~Species, pvalue.method = "permutation")

# Example with mtcars dataset
data(mtcars)

# Calculate p_values of "am" variable in mtcars dataset
pvaluescbatt(mtcars, ~am,
pvalue.method = "bootstrap", seed = 123)
```

pvaluescbrcu	<i>Calculate the p_values matrix or matrices (two or more) for each factor inside variable or variables (two or more), using Bray-Curtis dissimilarity as a base.</i>
--------------	---

Description

Using the Bray-Curtis dissimilarity for the dissimilarities calculation, this function takes a dataframe, a variable or variables (two or more), a p_value method such as "bootstrap" and "permutation" and returns the p_values matrix or matrices (two or more) between each pair of factors and a plot or plots (two or more) if the user select TRUE or leaves the parameter without argument.

Usage

```
pvaluescbrcu(  
  dataset,  
  formula,  
  pvalue.method = "permutation",  
  plot = TRUE,  
  seed = NULL,  
  min_group_size = 3  
)
```

Arguments

dataset	A dataframe.
formula	A variable or variables (two or more) with factors which you want to calculate the Bray-Curtis dissimilarities matrix or matrices (two or more).
pvalue.method	A p_value method used to calculate the matrix or matrices (two or more), the default value is "permutation". Another method is "bootstrap".
plot	if TRUE, plot the p_values heatmap or heatmaps (two or more). The default value is TRUE.
seed	Optionally, set a seed for 'bootstrap' or 'permutation'.
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded.

Value

A list containing a matrix or matrices (two or more) of p_values and, optionally, the plot.

Examples

```
# Example with iris dataset  
data(iris)  
  
# Calculate p_values of "Species" variable in iris dataset  
pvaluescbrcu(iris, ~Species, pvalue.method = "bootstrap")  
  
# Example with mtcars dataset  
data(mtcars)  
  
# Calculate p_values of "am" variable in mtcars dataset  
pvaluescbrcu(mtcars, ~am,  
  pvalue.method = "permutation", seed = 111)
```

pvaluesccanb	<i>Calculate the p_values matrix or matrices (two or more) for each factor inside variable or variables (two or more), using Canberra distance as a base.</i>
--------------	---

Description

Using the Canberra distance for the distances calculation, this function takes a dataframe, a variable or variables (two or more), a p_value method such as "bootstrap" and "permutation" and returns the p_values matrix or matrices (two or more) between each pair of factors and a plot or plots (two or more) if the user select TRUE or leaves the parameter without argument.

Usage

```
pvaluesccanb(
  dataset,
  formula,
  pvalue.method = "permutation",
  plot = TRUE,
  seed = NULL,
  min_group_size = 3
)
```

Arguments

dataset	A dataframe.
formula	A variable or variables (two or more) with factors which you want to calculate the Canberra distances matrix or matrices (two or more).
pvalue.method	A p_value method used to calculate the matrix or matrices (two or more), the default value is "permutation". Another method is "bootstrap".
plot	if TRUE, plot the p_values heatmap or heatmaps (two or more). The default value is TRUE.
seed	Optionally, set a seed for "bootstrap" and "permutation".
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded.

Value

A list containing a matrix or matrices (two or more) of p_values and, optionally, the plot.

Examples

```
# Example with iris dataset
data(iris)

# Calculate p_values of "Species" variable in iris dataset
```

```

pvaluesccanb(iris,~Species, pvalue.method = "permutation")

# Example with mtcars dataset
data(mtcars)

# Calculate p_values of "am" variable in mtcars dataset
pvaluesccanb(mtcars,~am + vs,
pvalue.method = "permutation", seed = 100)

```

pvaluesccheb	<i>Calculate the p_values matrix or matrices (two or more) for each factor inside variable or variables (two or more), using Chebyshev distance as a base.</i>
--------------	--

Description

Using the Chebyshev distance for the distances calculation, this function takes a dataframe, a variable or variables (two or more), a p_value method such as "bootstrap" and "permutation" and returns the p_values matrix or matrices (two or more) between each pair of factors and a plot or plots (two or more) if the user select TRUE or leaves the parameter without argument.

Usage

```

pvaluesccheb(
  dataset,
  formula,
  pvalue.method = "permutation",
  plot = TRUE,
  seed = NULL,
  min_group_size = 3
)

```

Arguments

dataset	A dataframe.
formula	A variable or variables (two or more) with factors which you want to calculate the Chebyshev distances matrix or matrices (two or more).
pvalue.method	A p_value method used to calculate the matrix or matrices (two or more), the default value is "permutation". Another method is "bootstrap".
plot	if TRUE, plot the p_values heatmap or heatmaps (two or more). The default value is TRUE.
seed	Optionally, set a seed for "bootstrap" and "permutation".
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded.

Value

A list containing a matrix or matrices (two or more) of p_values and, optionally, the plot.

Examples

```
# Example with "iris" dataset
data(iris)

# Calculate p_values of "Species" variable in iris dataset
pvaluesccheb(iris,~Species, pvalue.method = "permutation")

# Example with "mtcars" dataset
data(mtcars)

# Calculate p_values of "am" variable in mtcars dataset
pvaluesccheb(mtcars,~am,
pvalue.method = "bootstrap", seed = 100)
```

pvaluesccosi	<i>Calculate the p_values matrix or matrices (two or more) for each factor inside variable or variables (two or more), using Cosine dissimilarity as a base.</i>
--------------	--

Description

Using the Cosine dissimilarity for the dissimilarities calculation, this function takes a dataframe, a variable or variables (two or more), a p_value method such as "bootstrap" and "permutation" and returns the p_values matrix or matrices (two or more) between each pair of factors and a plot or plots (two or more) if the user select TRUE or leaves the parameter without argument.

Usage

```
pvaluesccosi(
  dataset,
  formula,
  pvalue.method = "permutation",
  plot = TRUE,
  seed = NULL,
  min_group_size = 3
)
```

Arguments

dataset	A dataframe.
formula	A variable or variables (two or more) with factors which you want to calculate the Cosine dissimilarities matrix or matrices (two or more).

pvalue.method	A p_value method used to calculate the matrix or matrices (two or more), the default value is "permutation". Another method is "bootstrap".
plot	if TRUE, plot the p_values heatmap or heatmaps (two or more). The default value is TRUE.
seed	Optionally, set a seed for "bootstrap" or "permutation".
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded.

Value

A list containing a matrix or matrices (two or more) of p_values and, optionally, the plot.

Examples

```
# Example with iris dataset
data(iris)

# Calculate p_values of "Species" variable in iris dataset
pvaluesccosi(iris,~Species, pvalue.method = "permutation")

# Example with mtcars
data(mtcars)

# Calculate p_values of "am" variable in mtcars dataset
pvaluesccosi(mtcars,~am,
pvalue.method = "permutation", seed = 123)
```

pvaluesceuc1	<i>Calculate the p_values matrix or matrices (two or more) for each pair of factors inside variable or variables (two or more), using Euclidean distance as a base.</i>
--------------	---

Description

Using the Euclidean distance for the distances calculation, this function takes a dataframe, a variable or variables (two or more), a p_value method such as "bootstrap" and "permutation" and returns the p_values matrix or matrices (two or more) between each pair of factors and a plot or plots (two or more) if the user select TRUE or leaves the parameter without argument.

Usage

```
pvaluesceuc1(
  dataset,
  formula,
  pvalue.method = "permutation",
  plot = TRUE,
  seed = NULL,
  min_group_size = 3
)
```

Arguments

dataset	A dataframe.
formula	A variable or variables (two or more) with factors which you want to calculate the Euclidean distances matrix or matrices (two or more).
pvalue.method	A p_value method used to calculate the matrix or matrices (two or more), the default value is "permutation". Another method is "bootstrap".
plot	if TRUE, plot the p_values heatmap or heatmaps (two or more). The default value is TRUE.
seed	Optionally, set a seed for "bootstrap" and "permutation" methods.
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded.

Value

A list containing a matrix or matrices (two or more) of p_values and, optionally, the plot.

Examples

```
# Example with iris dataset

# Calculate p_values of "Species" variable in iris dataset
pvaluesceucl(iris,~Species, pvalue.method = "permutation"
, min_group_size = 3)

# Example with mtcars dataset

# Calculate p_values of "am" variable in mtcars dataset
pvaluesceucl(mtcars,~am + carb,
pvalue.method = "bootstrap",
seed = 100, min_group_size = 2)
```

pvalueschamm	<i>Calculate the p_values matrix or matrices (two or more) for each factor inside variable or variables (two or more), using Hamming distance as a base.</i>
--------------	--

Description

Using the Hamming distance for the distances calculation, this function takes a dataframe, a variable or variables (two or more), a p_value method such as "bootstrap" and "permutation" and returns the p_values matrix or matrices (two or more) between each pair of factors and a plot or plots (two or more) if the user select TRUE or leaves the parameter without argument.

Usage

```
pvalueschamm(
  dataset,
  formula,
  pvalue.method = "permutation",
  plot = TRUE,
  seed = NULL,
  min_group_size = 3
)
```

Arguments

dataset	A dataframe.
formula	A variable or variables (two or more) with factors which you want to calculate the Hamming distances matrix or matrices (two or more).
pvalue.method	A p_value method used to calculate the matrix or matrices (two or more), the default value is "permutation". Another method is "bootstrap".
plot	if TRUE, plot the p_values heatmap or heatmaps (two or more). The default value is TRUE.
seed	Optionally, set a seed for "bootstrap" and "permutation".
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded.

Value

A list containing a matrix or matrices (two or more) of p_values and, optionally, the plot.

Examples

```
# Example with "iris" dataset
data(iris)

# Calculate p_values of "Species" variable in "iris" dataset
pvalueschamm(iris, ~Species, pvalue.method = "bootstrap")

# Example with mtcars dataset
data(mtcars)

# Calculate p_values of "am" variable in mtcars dataset
pvalueschamm(mtcars, ~am,
  pvalue.method = "permutation", seed = 100)
```

pvalueshell	<i>Calculate the p_values matrix or matrices (two or more) for each factor inside variable or variables (two or more), using Hellinger distances as a base.</i>
-------------	---

Description

Using the Hellinger distance for the distances calculation, this function takes a dataframe, a variable or variables (two or more), a p_value method such as "bootstrap" and "permutation" and returns the p_values matrix or matrices (two or more) between each pair of factors and a plot or plots (two or more) if the user select TRUE or leaves the parameter without argument.

Usage

```
pvalueshell(  
  dataset,  
  formula,  
  pvalue.method = "permutation",  
  plot = TRUE,  
  seed = NULL,  
  min_group_size = 3  
)
```

Arguments

dataset	A dataframe.
formula	A variable or variables (two or more) with factors which you want to calculate the Hellinger distances matrix or matrices (two or more).
pvalue.method	A p_value method used to calculate the matrix or matrices (two or more), the default value is "permutation". Another method is "bootstrap".
plot	if TRUE, plot the p_values heatmap or heatmaps (two or more). The default value is TRUE.
seed	Optionally, set a seed for 'bootstrap' and 'permutation'.
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded.

Value

A list containing a matrix or matrices (two or more) of p_values and, optionally, the plot.

Examples

```
# Example with iris dataset  
data(iris)  
  
# Calculate p_values of "Species" variable in iris dataset
```

```

pvalueschell(iris,~Species, pvalue.method = "bootstrap")

# Example with mtcars dataset
data(mtcars)

# Calculate p_values of "am" variable in mtcars dataset
pvalueschell(mtcars,~am,
pvalue.method = "permutation", seed = 122)

```

pvaluescjacc	<i>Calculate the p_values matrix or matrices (two or more) for each factor inside variable or variables (two or more), using Jaccard distance as a base.</i>
--------------	--

Description

Using the Jaccard distance for the distances calculation, this function takes a dataframe, a variable or variables (two or more), a p_value method such as "bootstrap" and "permutation" and returns the p_values matrix or matrices (two or more) between each pair of factors and a plot or plots (two or more) if the user select TRUE or leaves the parameter without argument.

Usage

```

pvaluescjacc(
  dataset,
  formula,
  pvalue.method = "permutation",
  plot = TRUE,
  seed = NULL,
  min_group_size = 3
)

```

Arguments

dataset	A dataframe.
formula	A variable or variables (two or more) with factors which you want to calculate the Jaccard distances matrix or matrices (two or more).
pvalue.method	A p_value method used to calculate the matrix or matrices (two or more), the default value is "permutation". Another method is "bootstrap".
plot	if TRUE, plot the p_values heatmap or heatmaps (two or more). The default value is TRUE.
seed	Optionally, set a seed for "bootstrap" or "permutation".
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded.

Value

A list containing a matrix or matrices (two or more) of p_values and, optionally, the plot.

Examples

```
# Example with the iris dataset
data(iris)

# Calculate p_values of "Species" variable in iris dataset
pvaluescjacc(iris,~Species, pvalue.method = "bootstrap")

# Example with the mtcars dataset
data(mtcars)

# Calculate p_values of "am" variable in mtcars dataset
pvaluescjacc(mtcars,~am,
pvalue.method = "permutation",
seed = 122)
```

pvaluescmaha	<i>Calculate the p_values matrix or matrices (two or more) for each pair of factors inside variable or variables (two or more), using Mahalanobis distance as a base.</i>
--------------	---

Description

Using the Mahalanobis distance for the distances calculation, this function takes a dataframe, a variable or variables (two or more), a p_value method such as "bootstrap" and "permutation" and returns the p_values matrix or matrices (two or more) between each pair of factors and a plot or plots (two or more) if the user select TRUE or leaves the parameter without argument.

Usage

```
pvaluescmaha(
  dataset,
  formula,
  pvalue.method = "permutation",
  plot = TRUE,
  seed = NULL,
  min_group_size = 3
)
```

Arguments

dataset	A dataframe.
formula	A variable or variables (two or more) with factors which you want to calculate the Mahalanobis distances matrix or matrices (two or more).

pvalue.method	A p_value method used to calculate the matrix or matrices (two or more), the default value is "permutation". Another method is "bootstrap".
plot	if TRUE, plot the p_values heatmap or heatmaps (two or more). The default value is TRUE.
seed	Optionally, set a seed for "bootstrap" and "permutation" methods.
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded.

Value

A list containing a matrix or matrices (two or more) of p_values and, optionally, the plot.

Examples

```
# Example with "airquality" dataset
data(airquality)

# Calculate p_values of "Month" variable in "airquality" dataset
pvaluescmaha(airquality, ~Month, pvalue.method = "permutation", seed = 12,
min_group_size = 3)

# Example with "mtcars" dataset
data(mtcars)

# Calculate p_values of "am" and "carb" variable in mtcars dataset
pvaluescmaha(mtcars, ~am + carb,
pvalue.method = "permutation", seed = 100, min_group_size = 2)
```

pvaluescmanh	<i>Calculate the p_values matrix or matrices (two or more) for each factor inside variable or variables (two or more), using Manhattan distance as a base.</i>
--------------	--

Description

Using the Manhattan distance for the distances calculation, this function takes a dataframe, a variable or variables (two or more), a p_value method such as "bootstrap" and "permutation" and returns the p_values matrix or matrices (two or more) between each pair of factors and a plot or plots (two or more) if the user select TRUE or leaves the parameter without argument.

Usage

```
pvaluescmanh(
  dataset,
  formula,
  pvalue.method = "permutation",
  plot = TRUE,
```



```

    seed = NULL,
    min_group_size = 3
  )

```

Arguments

dataset	A dataframe.
formula	A variable or variables (two or more) with factors which you want to calculate the Manhattan distances matrix or matrices (two or more).
pvalue.method	A p_value method used to calculate the matrix or matrices (two or more), the default value is "permutation". Another method is "bootstrap".
plot	if TRUE, plot the p_values heatmap or heatmaps (two or more). The default value is TRUE.
seed	Optionally, set a seed for "bootstrap" and "permutation".
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded.

Value

A list containing a matrix or matrices (two or more) of p_values and, optionally, the plot.

Examples

```

# Example with iris dataset
data(iris)

# Calculate p_values of "Species" variable in iris dataset
pvaluescmank(iris,~Species, pvalue.method = "bootstrap")

# Example with mtcars dataset
data(mtcars)

# Calculate p_values of "am" variable in mtcars dataset
pvaluescmank(mtcars,~am,
pvalue.method = "permutation", seed = 123)

```

pvaluescmink	<i>Calculate the p_values matrix or matrices (two or more) for each factor inside variable or variables (two or more), using Minkowski dissimilarity/distance as a base.</i>
--------------	--

Description

Using the Minkowski dissimilarity/distance for the dissimilarities/distances calculation, this function takes a dataframe, a variable or variables (two or more), a p_value method such as "bootstrap" and "permutation" and returns the p_values matrix or matrices (two or more) between each pair of factors and a plot or plots (two or more) if the user select TRUE or leaves the parameter without argument.

Usage

```
pvaluescmink(
  dataset,
  formula,
  pvalue.method = "permutation",
  p = 3,
  plot = TRUE,
  seed = NULL,
  min_group_size = 3
)
```

Arguments

dataset	A dataframe.
formula	A variable or variables (two or more) with factors which you want to calculate the Minkowski dissimilarities/distances matrix or matrices (two or more).
pvalue.method	A p_value method used to calculate the matrix or matrices (two or more), the default value is "permutation". Another method is "bootstrap".
p	Order of the Minkowski dissimilarities/distances. The default value is 3.
plot	if TRUE, plot the p_values heatmap or heatmaps (two or more). The default value is TRUE.
seed	Optionally, set a seed for "bootstrap" and "permutation".
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded.

Value

A list containing a matrix of p_values and, optionally, the plot.

Note

When $p < 1$, the Minkowski distance is a "dissimilarity" measure. When $p \geq 1$, the triangle inequality property is satisfied and we say "Minkowski distance".

Examples

```
# Example with iris dataset

# data(iris)

# Calculate p_values of "Species" variable in iris dataset

pvaluescmink(iris, ~Species, p = 3, pvalue.method = "permutation")

# Example with mtcars dataset
data(mtcars)

# Calculate p_values of "am" variable in mtcars dataset
```

```
pvaluescmink(mtcars, ~am, p = 3,
pvalue.method = "permutation", seed = 100)
```

pvaluescore	<i>Calculate the p_values matrix or matrices (two or more) for each factor inside variable or variables (two or more), using Sorensen-Dice dissimilarity as a base.</i>
-------------	---

Description

Using the Sorensen-Dice dissimilarity for the dissimilarities calculation, this function takes a dataframe, a variable or variables (two or more), a p_value method such as "bootstrap" and "permutation" and returns the p_values matrix or matrices (two or more) between each pair of factors and a plot or plots (two or more) if the user select TRUE or leaves the parameter without argument.

Usage

```
pvaluescore(
  dataset,
  formula,
  pvalue.method = "permutation",
  plot = TRUE,
  seed = NULL,
  min_group_size = 3
)
```

Arguments

dataset	A dataframe.
formula	A variable or variables (two or more) with factors which you want to calculate the Sorensen-Dice dissimilarities matrix or matrices (two or more).
pvalue.method	A p_value method used to calculate the matrix or matrices (two or more), the default value is "permutation". Another method is "bootstrap".
plot	if TRUE, plot the p_values heatmap or heatmaps (two or more). The default value is TRUE.
seed	Optionally, set a seed for "bootstrap" and "permutation".
min_group_size	Minimum group size to maintain. The default value is 3, therefore groups, inside variables, with less than 3 observations will be discarded.

Value

A list containing a matrix or matrices (two or more) of p_values and, optionally, the plot.

Examples

```
# Example with the iris dataset
data(iris)

# Calculate p_values of "Species" variable in iris dataset
pvaluescore(iris, ~Species, pvalue.method = "bootstrap")

# Example with mtcars dataset
data(mtcars)

# Calculate p_values of "am" variable in mtcars dataset
pvaluescore(mtcars, ~am,
pvalue.method = "permutation", seed = 134)
```

Index

cbhattacharyya, [2](#)
cbraycurtis, [4](#)
ccanberra, [5](#)
cchebyshev, [6](#)
ccosine, [8](#)
ceuklide, [9](#)
chamming, [11](#)
chellinger, [12](#)
cjaccard, [13](#)
cmahalanobis, [15](#)
cmanhattan, [16](#)
cminkowski, [18](#)
csorensendice, [19](#)

generate_report_cbhattacharyya, [21](#)
generate_report_cbraycurtis, [22](#)
generate_report_ccanberra, [23](#)
generate_report_cchebyshev, [24](#)
generate_report_ccosine, [25](#)
generate_report_ceuklide, [27](#)
generate_report_chamming, [28](#)
generate_report_chellinger, [29](#)
generate_report_cjaccard, [30](#)
generate_report_cmahalanobis, [31](#)
generate_report_cmanhattan, [32](#)
generate_report_cminkowski, [34](#)
generate_report_csorensendice, [35](#)

pvaluescbatt, [36](#)
pvaluescbrcu, [37](#)
pvaluesccanb, [39](#)
pvaluesccheb, [40](#)
pvaluesccosi, [41](#)
pvaluesceukl, [42](#)
pvalueschamm, [43](#)
pvalueschell, [45](#)
pvaluescjacc, [46](#)
pvaluescmaha, [47](#)
pvaluescmanh, [48](#)
pvaluescmink, [49](#)
pvaluescsore, [51](#)