

Package ‘discover’

October 13, 2022

Title Exploratory Data Analysis System

Version 2.2.6

Description Performs an exploratory data analysis through a 'shiny' interface. It includes basic methods such as the mean, median, mode, normality test, among others. It also includes clustering techniques such as Principal Components Analysis, Hierarchical Clustering and the K-Means Method.

License GPL (>= 2)

Imports DT, rlang, golem, shiny (>= 1.7.1), config, plotly, cluster, ggplot2, shinyjs, shinyAce, ggdendro, echarts4r, htmltools, htmlwidgets, colourpicker, shinydashboard, shinycustomloader, shinydashboardPlus (>= 2.0.0)

Depends R (>= 4.0)

Encoding UTF-8

URL <https://www.promidat.com>

RoxygenNote 7.1.2

NeedsCompilation no

Author Oldemar Rodriguez [aut, cre],
Diego Jiménez [ctb, prg]

Maintainer Oldemar Rodriguez <oldemar.rodriguez@ucr.ac.cr>

Repository CRAN

Date/Publication 2021-10-11 17:30:02 UTC

R topics documented:

BP	2
calc.centros	3
datos.disyuntivos	4
dfnormal	4
discoverR	5
e_cat	5
e_cor	6
e_histboxplot	7

e_histnormal	8
e_horiz	9
e_inercia	9
e_jambu	10
e_mapa	11
e_mapa_3D	11
e_pcabi	12
e_pcabi_3D	13
e_pcaind	14
e_pcaind_3D	15
e_pcavar	16
e_pcavar_3D	17
e_qq	18
e_radar	19
e_silhouette	20
e_vert	20
gg_dendrograma	21
inercia.total	22
run_app	22
var.categoricas	23
var.numericas	23
WP	24

Index**25**

BP *Calculate inter-class inertia*

Description

Calculate inter-class inertia

Usage

BP(DF, clusters)

Arguments

DF a data.frame object.
clusters a vector specifying the cluster of each individual.

Value

numeric

Author(s)

Diego Jimenez <diego.jimenez@promidat.com>

Examples

```
m <- hclust(dist(iris[, -5]))
BP(iris[, -5], cutree(m, 3))
```

calc.centros	<i>Calculation of the center of clusters</i>
--------------	----------------------------------------------

Description

Calculation of the center of clusters

Usage

```
calc.centros(data, clusters)
```

Arguments

data a data.frame object.
clusters a vector specifying the cluster of each individual.

Value

list

Author(s)

Diego Jimenez <diego.jimenez@promidat.com>

Examples

```
clusters <- factor(kmeans(iris[, -5], 3)$cluster)
calc.centros(iris[, -5], clusters)
```

`datos.disyuntivos` *Create disjunctive columns to a data.frame.*

Description

Create disjunctive columns to a data.frame.

Usage

```
datos.disyuntivos(data, var)
```

Arguments

`data` a data.frame object.
`var` the column name to apply disjunctive code.

Value

data.frame

Author(s)

Diego Jimenez <diego.jimenez@promidat.com>

Examples

```
datos.disyuntivos(iris, "Species")
```

`dfnormal` *Data.frame with normal test*

Description

Data.frame with normal test

Usage

```
dfnormal(data)
```

Arguments

`data` a data.frame object only with the numeric columns.

Value

data.frame

Author(s)

Diego Jimenez <diego.jimenez@promidat.com>

Examples

```
dfnormal(iris[, -5])
```

discoverR

Exploratory Data Analysis System

Description

Performs an exploratory data analysis through a 'shiny' interface. It includes basic methods such as the mean, median, mode, normality test, among others. It also includes clustering techniques such as Principal Components Analysis, Hierarchical Clustering and the K-Means Method.

Details

Package: discoverR
Type: Package
Version: 2.2.6
Date: 2021-10-08
License: GPL (>=2)

Author(s)

Oldemar Rodriguez Rojas
Maintainer: Oldemar Rodriguez Rojas <oldemar.rodriguez@ucr.ac.cr>

e_cat

Barplot for categoric variable by clusters.

Description

Barplot for categoric variable by clusters.

Usage

```
e_cat(clusters, var, colores = NULL, escalar = T)
```

Arguments

clusters	a vector specifying the cluster of each individual.
var	a factor column of a data.frame.
colores	a vector of color for each cluster.
escalar	a boolean value specifying if use percentage or real values.

Value

echarts4r plot

Author(s)

Diego Jimenez <diego.jimenez@promidat.com>

Examples

```
clusters <- factor(kmeans(iris[, -5], 3)$cluster)
e_cat(clusters, iris[, 5], colores = c("steelblue", "pink", "forestgreen"))
```

e_cor

Correlation plot

Description

Correlation plot

Usage

```
e_cor(x, colors = c("#FF5733", "#F8F5F5", "#2E86C1"))
```

Arguments

x	a data.frame with correlation values.
colors	a vector of length 3 with color values.

Value

echarts4r plot

Author(s)

Diego Jimenez <diego.jimenez@promidat.com>

Examples

```
p <- round(cor(iris[, -5]), 3)
e_cor(p)
```

e_histboxplot	<i>Histogram + boxplot</i>
---------------	----------------------------

Description

Histogram + boxplot

Usage

```
e_histboxplot(  
  data,  
  var.name,  
  colorBar = "steelblue",  
  colorPoint = "red",  
  titulos = c("Minimo", "Primer Cuartil", "Mediana", "Tercer Cuartil", "Maximo")  
)
```

Arguments

data	a numeric column of a data.frame.
var.name	a character value specifying the name of the variable.
colorBar	a color for the bars.
colorPoint	a color for the points.
titulos	a character vector of length 5 specifying the titles to use on legend.

Value

echarts4r plot

Author(s)

Diego Jimenez <diego.jimenez@promidat.com>

Examples

```
e_histboxplot(iris$Sepal.Width, "Sepal.Width")
```

e_histnormal	<i>Normal plot</i>
--------------	--------------------

Description

Normal plot

Usage

```
e_histnormal(  
  data,  
  colorbar = "steelblue",  
  colorline = "gray",  
  nombres = c("Histograma", "Curva Normal")  
)
```

Arguments

data	a numeric column of a data.frame.
colorbar	a color for the bars.
colorline	a color for the line.
nombres	a character vector of length 2 specifying the titles to use on legend.

Value

echarts4r plot

Author(s)

Diego Jimenez <diego.jimenez@promidat.com>

Examples

```
e_histnormal(iris$Sepal.Length)
```

e_horiz	<i>Horizontal representation for centers of clusters.</i>
---------	-----------------------------------------------------------

Description

Horizontal representation for centers of clusters.

Usage

```
e_horiz(centros, colores = NULL)
```

Arguments

centros	a data.frame object with the centers of the clusters.
colores	a vector of color for each cluster.

Value

echarts4r plot

Author(s)

Diego Jimenez <diego.jimenez@promidat.com>

Examples

```
clusters <- factor(kmeans(iris[, -5], 3)$cluster)
c <- calc.centros(iris[, -5], clusters)
e_horiz(c$real, c("steelblue", "pink", "forestgreen"))
```

e_inercia	<i>Inertia plot of clusterization</i>
-----------	---------------------------------------

Description

Inertia plot of clusterization

Usage

```
e_inercia(
  data,
  titulos = c("Inercia", "Inercia Inter-Clase", "Inercia Inter-Clase")
)
```

Arguments

`data` a data.frame object with the inertia values.
`titulos` a character vector of length 3 specifying the titles to use on legend.

Value

echarts4r plot

Author(s)

Diego Jimenez <diego.jimenez@promidat.com>

e_jambu

Jambu Elbow plot

Description

Jambu Elbow plot

Usage

```
e_jambu(data, max.clusters)
```

Arguments

`data` a data.frame object.
`max.clusters` a numeric value specifying the number of times to generate the model.

Value

echarts4r plot

Author(s)

Diego Jimenez <diego.jimenez@promidat.com>

Examples

```
e_jambu(iris[, -5], 10)
```

e_mapa *PCA plot of individuals colored by clusters*

Description

PCA plot of individuals colored by clusters

Usage

```
e_mapa(pca.model, clusters, colores = NULL, ejes = c(1, 2))
```

Arguments

`pca.model` an object of class PCA [FactoMineR].
`clusters` a vector specifying the cluster of each individual.
`colores` a vector of color for each cluster.
`ejes` a numeric vector of length 2 specifying the dimensions to be plotted.

Value

echarts4r plot

Author(s)

Diego Jimenez <diego.jimenez@promidat.com>

Examples

```
p <- discoverR::PCA(iris[, -5], graph = FALSE)
clusters <- factor(kmeans(iris[, -5], 3)$cluster)
e_mapa(p, clusters, c("steelblue", "pink", "forestgreen"))
```

e_mapa_3D *PCA plot of individuals colored by clusters*

Description

PCA plot of individuals colored by clusters

Usage

```
e_mapa_3D(pca.model, clusters, colores = NULL, ejes = c(1, 2, 3))
```

Arguments

`pca.modelo` an object of class PCA [FactoMineR].
`clusters` a vector specifying the cluster of each individual.
`colores` a vector of color for each cluster.
`ejes` a numeric vector of length 3 specifying the dimensions to be plotted.

Value

echarts4r plot

Author(s)

Diego Jimenez <diego.jimenez@promidat.com>

Examples

```
p <- discoverR::PCA(iris[, -5], graph = FALSE)
clusters <- factor(kmeans(iris[, -5], 3)$cluster)
e_mapa_3D(p, clusters, c("steelblue", "pink", "forestgreen"))
```

e_pcabi

PCA biplot

Description

PCA biplot

Usage

```
e_pcabi(
  modelo,
  axes = c(1, 2),
  colorInd = "steelblue",
  colorVar = "forestgreen",
  cos2Ind = 0,
  cos2Var = 0,
  colorIndCos = "firebrick",
  colorVarCos = "darkorchid",
  titulos = c("Bien Representados", "Mal Representados")
)
```

Arguments

modelo	an object of class PCA [FactoMineR].
axes	a numeric vector of length 2 specifying the dimensions to be plotted.
colorInd	a color for the individuals well represented.
colorVar	a color for the variables well represented.
cos2Ind	a numeric value from 0 to 1 specifying the quality of the individuals.
cos2Var	a numeric value from 0 to 1 specifying the quality of the variables.
colorIndCos	a color for the individuals badly represented.
colorVarCos	a color for the variables badly represented.
titulos	a character vector of length 2 specifying the titles to use on legend.

Value

echarts4r plot

Author(s)

Diego Jimenez <diego.jimenez@promidat.com>

Examples

```
p <- discoverR:::PCA(iris[, -5], graph = FALSE)
e_pcabi(p)
```

e_pcabi_3D

PCA biplot in 3D

Description

PCA biplot in 3D

Usage

```
e_pcabi_3D(
  modelo,
  axes = c(1, 2, 3),
  colorInd = "steelblue",
  colorVar = "forestgreen",
  cos2Ind = 0,
  cos2Var = 0,
  colorIndCos = "firebrick",
  colorVarCos = "darkorchid",
  titulos = c("Bien Representados", "Mal Representados")
)
```

Arguments

modelo	an object of class PCA [FactoMineR].
axes	a numeric vector of length 3 specifying the dimensions to be plotted.
colorInd	a color for the individuals well represented.
colorVar	a color for the variables well represented.
cos2Ind	a numeric value from 0 to 1 specifying the quality of the individuals.
cos2Var	a numeric value from 0 to 1 specifying the quality of the variables.
colorIndCos	a color for individuals badly represented.
colorVarCos	a color for variables badly represented.
titulos	a character vector of length 2 specifying the titles to use on legend.

Value

echarts4r plot

Author(s)

Diego Jimenez <diego.jimenez@promidat.com>

Examples

```
p <- discoverR:::PCA(iris[, -5], graph = FALSE)
e_pcabi_3D(p)
```

e_pcaind

PCA plot of individuals

Description

PCA plot of individuals

Usage

```
e_pcaind(
  modelo,
  axes = c(1, 2),
  colorInd = "steelblue",
  cos2 = 0,
  colorCos = "firebrick",
  titulos = c("Bien Representados", "Mal Representados")
)
```

Arguments

modelo	an object of class PCA [FactoMineR].
axes	a numeric vector of length 2 specifying the dimensions to be plotted.
colorInd	a color for the individuals well represented.
cos2	a numeric value from 0 to 1 specifying the quality of the individuals.
colorCos	a color for individuals badly represented.
titulos	a character vector of length 2 specifying the titles to use on legend.

Value

echarts4r plot

Author(s)

Diego Jimenez <diego.jimenez@promidat.com>

Examples

```
p <- discover::PCA(iris[, -5], graph = FALSE)
e_pcaind(p)
```

e_pcaind_3D

PCA plot of individuals in 3D

Description

PCA plot of individuals in 3D

Usage

```
e_pcaind_3D(
  modelo,
  axes = c(1, 2, 3),
  colorInd = "steelblue",
  cos2 = 0,
  colorCos = "firebrick",
  titulos = c("Bien Representados", "Mal Representados")
)
```

Arguments

modelo	an object of class PCA [FactoMineR].
axes	a numeric vector of length 3 specifying the dimensions to be plotted.
colorInd	a color for the individuals well represented.
cos2	a numeric value from 0 to 1 specifying the quality of the individuals.
colorCos	a color for individuals badly represented.
titulos	a character vector of length 2 specifying the titles to use on legend.

Value

echarts4r plot

Author(s)

Diego Jimenez <diego.jimenez@promidat.com>

Examples

```
p <- discoverR::PCA(iris[, -5], graph = FALSE)
e_pcaind_3D(p)
```

e_pcavar

PCA plot of variables

Description

PCA plot of variables

Usage

```
e_pcavar(
  modelo,
  axes = c(1, 2),
  colorVar = "forestgreen",
  cos2 = 0,
  colorCos = "darkorchid",
  titulos = c("Bien Representados", "Mal Representados")
)
```


Arguments

modelo	an object of class PCA [FactoMineR].
axes	a numeric vector of length 2 specifying the dimensions to be plotted.
colorVar	a color for the variables well represented.
cos2	a numeric value from 0 to 1 specifying the quality of the variables.
colorCos	a color for the variables badly represented.
titulos	a character vector of length 2 specifying the titles to use on legend.

Value

echarts4r plot

Author(s)

Diego Jimenez <diego.jimenez@promidat.com>

Examples

```
p <- discoverR::PCA(iris[, -5], graph = FALSE)
e_pcavar(p)
```

e_pcavar_3D

PCA plot of variables in 3D

Description

PCA plot of variables in 3D

Usage

```
e_pcavar_3D(
  modelo,
  axes = c(1, 2, 3),
  colorVar = "forestgreen",
  cos2 = 0,
  colorCos = "darkorchid",
  titulos = c("Bien Representados", "Mal Representados")
)
```

Arguments

modelo	an object of class PCA [FactoMineR].
axes	a numeric vector of length 3 specifying the dimensions to be plotted.
colorVar	a color for the variables well represented.
cos2	a numeric value from 0 to 1 specifying the quality of the variables.
colorCos	a color for variables badly represented.
titulos	a character vector of length 2 specifying the titles to use on legend.

Value

echarts4r plot

Author(s)

Diego Jimenez <diego.jimenez@promidat.com>

Examples

```
p <- discoverR::PCA(iris[, -5], graph = FALSE)
e_pcavar_3D(p)
```

e_qq

Qplot + Qline

Description

Qplot + Qline

Usage

```
e_qq(data, colorpoint = "steelblue", colorline = "gray")
```

Arguments

data	a numeric column of a data.frame.
colorpoint	a color for the points.
colorline	a color for the line.

Value

echarts4r plot

Author(s)

Diego Jimenez <diego.jimenez@promidat.com>

Examples

```
e_qq(iris$Sepal.Length)
```

e_radar

Radar representation for centers of clusters.

Description

Radar representation for centers of clusters.

Usage

```
e_radar(centros, colores = NULL)
```

Arguments

centros a data.frame object with the centers of the clusters.
colores a vector of color for each cluster.

Value

echarts4r plot

Author(s)

Diego Jimenez <diego.jimenez@promidat.com>

Examples

```
clusters <- factor(kmeans(iris[, -5], 3)$cluster)
c <- calc.centros(iris[, -5], clusters)
e_radar(c$porcentual, c("steelblue", "pink", "forestgreen"))
```

e_silhouette	<i>Silhouette plot</i>
--------------	------------------------

Description

Silhouette plot

Usage

```
e_silhouette(data, max.clusters)
```

Arguments

data a data.frame object.
max.clusters a numeric value specifying the number of times to generate the model.

Value

echarts4r plot

Author(s)

Diego Jimenez <diego.jimenez@promidat.com>

Examples

```
e_silhouette(iris[, -5], 10)
```

e_vert	<i>Vertical representation for centers of clusters.</i>
--------	---------------------------------------------------------

Description

Vertical representation for centers of clusters.

Usage

```
e_vert(centros, colores = NULL)
```

Arguments

centros a data.frame object with the centers of the clusters.
colores a vector of color for each cluster.

Value

echarts4r plot

Author(s)

Diego Jimenez <diego.jimenez@promidat.com>

Examples

```
clusters <- factor(kmeans(iris[, -5], 3)$cluster)
c <- calc.centros(iris[, -5], clusters)
e_vert(c$real, c("steelblue", "pink", "forestgreen"))
```

gg_dendrograma

Dendrogram plot

Description

Dendrogram plot

Usage

```
gg_dendrograma(model, k, colors = NULL)
```

Arguments

model	an object of class hclust.
k	a vector specifying the cluster of each individual.
colors	a vector of color for each cluster.

Value

ggplot

Author(s)

Diego Jimenez <diego.jimenez@promidat.com>

<code>inercia.total</code>	<i>Calculate total inertia</i>
----------------------------	--------------------------------

Description

Calculate total inertia

Usage

```
inercia.total(DF)
```

Arguments

DF a data.frame object.

Value

numeric

Author(s)

Diego Jimenez <diego.jimenez@promidat.com>

<code>run_app</code>	<i>Run the Shiny Application</i>
----------------------	----------------------------------

Description

Run the Shiny Application

Usage

```
run_app(...)
```

Arguments

... A series of options to be used inside the app.

Examples

```
if(interactive()) {  
  run_app()  
}
```

var.categoricas *Filter category variables of a data.frame*

Description

Filter category variables of a data.frame

Usage

```
var.categoricas(data)
```

Arguments

data a data.frame object.

Value

data.frame

Author(s)

Diego Jimenez <diego.jimenez@promidat.com>

Examples

```
var.categoricas(iris)
```

var.numericas *Filter numeric variables of a data.frame*

Description

Filter numeric variables of a data.frame

Usage

```
var.numericas(data)
```

Arguments

data a data.frame object.

Value

data.frame

Author(s)

Diego Jimenez <diego.jimenez@promidat.com>

Examples

```
var.numericas(iris)
```

WP

Calculate intra-class inertia

Description

Calculate intra-class inertia

Usage

```
WP(DF, clusters)
```

Arguments

DF a data.frame object.
clusters a vector specifying the cluster of each individual.

Value

numeric

Author(s)

Diego Jimenez <diego.jimenez@promidat.com>

Examples

```
m <- hclust(dist(iris[, -5]))  
WP(iris[, -5], cutree(m, 3))
```


Index

* package

discoverR, 5

BP, 2

calc.centros, 3

datos.disyuntivos, 4

dfnormal, 4

discoverR, 5

e_cat, 5

e_cor, 6

e_histboxplot, 7

e_histnormal, 8

e_horiz, 9

e_inercia, 9

e_jambu, 10

e_mapa, 11

e_mapa_3D, 11

e_pcabi, 12

e_pcabi_3D, 13

e_pcaind, 14

e_pcaind_3D, 15

e_pcavar, 16

e_pcavar_3D, 17

e_qq, 18

e_radar, 19

e_silhouette, 20

e_vert, 20

gg_dendrograma, 21

inercia.total, 22

run_app, 22

var.categoricas, 23

var.numericas, 23

WP, 24