Package 'Cluster.OBeu'

June 7, 2017

Type Package

Title Cluster Analysis OpenBudgets.eu

Version 1.1.3

Date 2016-08-09

Author Kleanthis Koupidis

Maintainer Kleanthis Koupidis <koupidis.okfgr@gmail.com>

Description Cluster Analysis for OBEU datasets.

URL https://github.com/okgreece/Cluster.OBeu

BugReports https://github.com/okgreece/Cluster.OBeu/issues

License GPL-2 | file LICENSE

LazyData TRUE

Imports car, cluster, clValid, data.tree, dendextend, graphics, jsonlite, mclust, RCurl, reshape, stringr, utils, methods,

RoxygenNote 6.0.1

R topics documented:

`cl.analysis` *Cluster OBEU*

Description

Clustering Analysis for OBEU datasets.

Usage

```
cl.analysis(cl.data, cl_feature=NULL, amount=NULL, cl.aggregate="sum",
cl.meth=NULL, clust.numb=NULL, dist="euclidean")
```

Arguments

cl.data	The input data
cl_feature	Select a subset of the input data
amount	The numeric
cl.aggregate	Select a different aggregation in case of filtering the inpupt data
cl.meth	The clustering method algorithm
clust.numb	The number of clusters
dist	The distance metric

Details

There are different clustering models to be selected through an evaluation process. The user should define the dimensions, measured.dim and amount parameters to form the structure of cluster data. The clustering algorithm, the number of clusters and the distance metric of the clustering model are set to the best selection using internal and stability measures. The end user can also interact with the cluster analysis and these parameters by specifying the cl.method, cl.num and cl.dist parameters respectively.

Value

The final returns are the parameters needed for visualizing the cluster data depending on the selected algorithm and the specification parameters, as long as some comparison measure matrices.

- cl.meth - Label of the clustering algorithm
- clust.numb - The number of clusters
- data.pca - The principal components to visualize the input data
- modelparam - The results of this parameter depend of the selected clustering model

Author(s)

Kleanthis Koupidis, Jaroslav Kuchar

See Also

cl.features, clValid, diana, agnes, pam, clara, fanny, Mclust

cl.features *Clustering features*

Description

Select clustering characteristic of OBEU datasets.

Usage

```
cl.features(data, features=NULL, amounts=NULL, aggregate="sum")
```

Arguments

data	The input data
features	The clustering features
amounts	The amount measures of the dataset
aggregate	The function to aggregate

Details

This function adapts the dataset according to the selected dimension of the dataset and the aggregation function.

Value

This function returns the dataset for cluster analysis adapted to the desired features.

Author(s)

Kleanthis Koupidis

See Also

cl.analysis

cl.plot *Clustering model plotting*

Description

cl.plot function plots the clustering model constructed by the cl.analysis function.

Usage

```
cl.plot(clustering.model, parameters = list())
```

Arguments

clustering.model	Object returned by the cl.analysis function.
parameters	List of parameters to indicate plotting of ellipses or convex hulls. Default values: list(ellipses=FALSE, convex.hulls=FALSE).

Author(s)

Jaroslav Kuchar <https://github.com/jaroslav-kuchar>

See Also

cl.analysis

Examples

```
#data("iris")
#inputs.data <- scale(iris[,1:4])
#inputs.clustering <- cl.analysis(inputs.data, cl.meth="kmeans", clust.numb=3)
#cl.plot(inputs.clustering, parameters = list(convex.hulls=TRUE))
```

---

cl.summary — *Extract the proposed clustering method and the number of clusters from clvalid method*

Description

Extract the most frequent

Usage

```
cl.summary(clv)
```

Arguments

clv	A clValid object

Details

This function returns the proposed method or number of clusters or both according to the majority clustering indices of a clValid process

Value

A value that indicates the proposed method and number of clusters.

Author(s)

Kleanthis Koupidis

nums — *Select the numeric columns of a given dataset*

Description

Extract and return a data frame with the columns that include only numeric values

Usage

```
nums(data)
```

Arguments

data	The input data frame

Value

This function returns a data frame with the numeric columns of the input dataset.

Author(s)

Kleanthis Koupidis

open_spending.cl — *Read and Calculate the Basic Information for Cluster Analysis Tasks from Open Spending API*

Description

Extract and analyze the input data provided from Open Spending API, using the `cl.analysis` function.

Usage

```
open_spending.cl(json_data, dimensions=NULL, amounts=NULL, measured.dim=NULL,
cl.aggregate="sum", cl.method=NULL, cl.num=NULL, cl.dist="euclidean")
```

Arguments

json_data	The json string, URL or file from Open Spending API
dimensions	The dimensions/feature of the input data
amounts	The measures of the input data
measured.dim	The dimensions to which correspond amount/numeric variables
cl.aggregate	The desired aggregation of the input data
cl.method	The clustering method algorithm
cl.num	The number of clusters
cl.dist	The distance metric

Details

This function is used to read data in json format from Open Spending API, in order to implement cluster analysis through `cl.analysis` function.

Value

A json string with the resulted parameters of the `cl.analysis` function.

Author(s)

Kleanthis Koupidis

See Also

`cl.analysis`

Index