Package 'ODMeans'

October 12, 2022

Type Package

Title OD-Means: k-Means for Origin-Destination

Version 0.1.0

Description OD-means is a hierarchical adaptive k-means algorithm based on origin-destination pairs. In the first layer of the hierarchy, the clusters are separated automatically based on the variation of the within-cluster distance of each cluster until convergence. The second layer of the hierarchy corresponds to the sub clustering process of small clusters based on the distance between the origin and destination of each cluster.

License GPL (>= 3)

Encoding UTF-8

LazyData true

RoxygenNote 7.1.2

Imports geosphere, ggplot2, stats

Collate 'ODMeansSampleData.R' 'dinamic_clusters.R' 'hierarchical_clusters.R' 'od_means.R'

Depends R (>= 2.10)

NeedsCompilation no

Author Sebastian Moreno [aut, cre], Cristobal Heredia [aut], Wilfredo Yushimito [ths]

Maintainer Sebastian Moreno <sebastian.moreno.araya@gmail.com>

Repository CRAN

Date/Publication 2022-02-01 09:20:13 UTC

R topics documented:

`dinamic_clusters` *Dinamic Clusters Function*

Description

Dinamic Clusters Function

Usage

```
dinamic_clusters(data, numK, limitsSeparation, maxDist)
```

Arguments

`data`	A data frame with four columns: Initial Latitude | Initial Longitude | Final Latitude | Final Longitude
`numK`	Initial number of clusters in the first call of K-Means.
`limitsSeparation`	Range to determine if a drastic change has happened between a cluster and its separation. A bigger value makes more difficult to separate a cluster.
`maxDist`	Maximum distance to join two points. This is based on the euclidean distance.

Value

Dinamic Clusters returns an object similar of class "kmeans". It is a list with at least the following components:

cluster A vector of integers (from 1:k) indicating the cluster to which each point is allocated. centers A matrix of cluster centres. totss The total sum of squares. withinss Vector of within-cluster sum of squares, one component per cluster. tot.withinss Total within-cluster sum of squares, i.e. sum(withinss). betweenss The between-cluster sum of squares, i.e. totss-tot.withinss. size The number of points in each cluster. level_hierarchy Corresponds of the hierarchy level of the cluster, can be "Global" or "Local"

Examples

```
data(ODMeansSampleData)
dinamic_clusters(ODMeansSampleData, 5, 200, 2500)
```

hierarchical_clusters *Hierarchical Clusters*

Description

Hierarchical Clusters

Usage

```
hierarchical_clusters(data, Kcluster, distHierarchical)
```

Arguments

data	A data frame with four columns: Initial Latitude | Initial Longitude | Final Latitude | Final Longitude
Kcluster	An ODMeans structure, result of function dinamic_clusters.
distHierarchical	Maximum distance to create a new hierarchy per cluster.

Value

Hierarchical Clusters returns an object similar of class "kmeans". It is a list with at least the following components:

cluster A vector of integers (from 1:k) indicating the cluster to which each point is allocated. centers A matrix of cluster centres. totss The total sum of squares. withinss Vector of within-cluster sum of squares, one component per cluster. tot.withinss Total within-cluster sum of squares, i.e. sum(withinss). betweenss The between-cluster sum of squares, i.e. totss-tot.withinss. size The number of points in each cluster. level_hierarchy Corresponds of the hierarchy level of the cluster, can be "Global" or "Local"

Examples

```
data(ODMeansSampleData)
hierarchical_clusters(ODMeansSampleData, dinamic_clusters(ODMeansSampleData, 5, 200, 2500), 500)
```

ODMeansSampleData *Origin-Destination points*

Description

A dataset containing 1700 Origin-Destination points

Usage

```
ODMeansSampleData
```

Format

A data frame with 10000 rows and 4 variables:

OriginLatitude Consists of the origin latitude dimension

OriginLongitude Consists of the origin longitude dimension

DestinationLatitude Consists of the destination latitude dimension

DestinationLongitude Consists of the destination longitude dimension

original_cluster Original cluster of the points when it was created ...

Source

Synthetic data

`od_means` *ODMeans*

Description

ODMeans

Usage

```
od_means(data, numK, limitsSeparation, maxDist, distHierarchical)
```

Arguments

`data`	A data frame with four columns: Initial Latitude | Initial Longitude | Final Latitude | Final Longitude
`numK`	Initial number of clusters in the first call of K-Means.
`limitsSeparation`	Range to determine if a drastic change has happened between a cluster and its separation. A bigger value makes more difficult to separate a cluster.
`maxDist`	Maximum distance to join two points. This is based on the euclidean distance.
`distHierarchical`	Maximum distance to create a new hierarchy per cluster

Value

Returns a structure that contains the final centers, clusters, sizes and hierarchy

Examples

```
data(ODMeansSampleData)
od_means(ODMeansSampleData, 5, 200, 2500, 500)
```

Index