Package ‘cartograflow’

May 10, 2019

Title  Filtering Matrix for Flow Mapping
Version  1.0.0
Description  Functions to prepare and filter an origin-destination matrix for thematic flow mapping purposes. This comes after Bahoken, Francoise (2016), Mapping flow matrix a contribution, PhD in Geography - Territorial sciences. See Bahoken (2017) <doi:10.4000/netcom.2565>.
Depends  R (>= 3.4.0)
License  GPL-3
Encoding  UTF-8
LazyData  true
RoxygenNote  6.1.1
Suggests  knitr, rmarkdown, testthat, cartography
VignetteBuilder  knitr
Imports  dplyr, g.data, graphics, ggplot2, maptools, plotly, rgdal, rgeos, reshape2, rlang, sp, utils
NeedsCompilation  no
Author  Sylvain Blondeau [aut, cre], Francoise Bahoken [aut]
Maintainer  Sylvain Blondeau <blondeau.sylvain@yahoo.fr>
Repository  CRAN
Date/Publication  2019-05-10 15:50:09 UTC

R topics documented:
cartograflow ................................................................. 2
flowanalysis ................................................................. 2
cartograflow  Cartograflow Package

Description

This package is designed to create the so-called flowmaps, by filtering origin-destination (OD) matrix. It is based on different functions that are mainly used to prepare the flow dataset. The spatial objects processing are those of sp or sf and the mapping elements are often those of Cartography except particular cases.

Details

To learn more about cartograflow, see the vignette cartograflow.html

Main functions :

- flowanalysis flowcarre flowcontig flowdist flowgini flowjointure flowmap flowreduct flowstructmat flowstructmat flowcontig

flowanalysis  Computation of a global concentration criterion of flows values or features

Description

Computation of a global selection criterion for thresholding flow information and/or features before mapping.
To be use after flowgini and before flowmap.

Usage

flowanalysis(tab, critflow, critlink, result)
Arguments

- **tab**  
  flow dataset from `flowgini`
- **critflow**  
  level of flow significativity. See Details.
- **critlink**  
  level of features density. See Details.
- **result**  
  resulting filtering criterion value. See Details.

Details

- **critflow** = desired level of flow’s information significativity (e.g. 80 of the total information);
- **critlink** = desired level of flow’s features density (e.g. 20 features that represents “more significant information.”)
- **result = "density"** returns the desired level of features density as a
- **result = "significativity"** returns the level of flow significativity as a

References


Examples

```r
library(cartograflow)
data(flowdata)
bkg <- system.file("shape/MGP_TER.shp", package="cartograflow", lib.loc = NULL, mustWork = TRUE)

#1/4: Computes Gini's coefficient
tab_gini <- flowgini(flows, format="L", origin="i", dest="j", valflow="Fij",
                     bkg, code="EPT_NUM", lorenz.plot = FALSE)
## [1] Gini's coefficient = 73.16 %

#2/4: Plot Lorenz curve
flowgini(tab_gini, format="L", origin="i", dest="j", valflow="ydata",
          bkg, code="EPT_NUM", lorenz.plot = TRUE)

#3/4: Compute critflow filtering parameter
#critflow = 0.8 #selected criterion
flowanalysis(tab_gini, critflow = 0.8, result = "signif")
## [1] "threshold = 11238  ---  flows = 80 %  ---  links = 22.94 %"

#4/4: Plot the flowmap
flowmap(flows, format="L", bkg, code = "EPT_NUM", filter = TRUE,
         threshold = 11238, taille = 8, a.head = 1, a.length = 0.11,
         a.angle = 30, a.col="#3f4247")
```
Create a square matrix from geographical ID

Usage

flowcarre(liste, tab, origin, dest, valflow, empty.sq, format, diagonale)

Arguments

liste list or all the geographical ID as a single vector flow dataset
tab the non squared input flow dataset with three columns: origin, destination, flow value
origin origin place
dest destination place
valflow flow value
empty.sq TRUE: to allows to have a matrix empty with only the ID of background map; FALSE or missing
format the squared flow dataset output format. See Details.
diagonale See Details.

Details

- format is "M" for matrix format
- format is "L" for long format
- diagonal is "TRUE" to zero the main diagonal

Examples

library(cartograflow)
data(flowdata)
var1<-geoid
var2<-flows

#1/2 Compute an empty square matrix with ID code, and sets the value to zero
#Example for matrix format (same procedure for the long format)
mat<-flowcarre(var1, var2, origin="i", dest="j", valflow="Fij",
format="M", empty.sq=TRUE)

#2/2 Fill in the matrix with external flow values
mat<-flowcarre(var1, var2, origin="i", dest="j", valflow="Fij",
format="M", empty.sq=FALSE)
flowcontig

#Square a matrice and zero the main diagonal
mat<-flowcarre(var1, var2, origin="i", dest="j", valflow="Fij", format="M",
  empty.sq=FALSE, diagonale = FALSE)

flowcontig Computes an ordinal distance matrices based on geographical background

Description
From a geographical background, compute an ordinal distance matrice based on a k-contiguity. The result is a neighbourhood graph that can be used for filtering flow values before flow mapping (flowmap)

Usage
flowcontig(fdc, code, ordre)

Arguments

  fdc is the map background file (ie. a shapefile of polygons)
  code identifiant
  ordre number of borders to cross between origin and destination place. See details.

Details
Contiguity is in terms of the (k=1,2,4) number of spatial boundaries to be crossed between a place of origin and a place of destination -ordre=1 is when the flow have to cross only 1 boundary -ordre=2 is when the origin-destinations places are distant from 2 borders -ordre=4 is when the origin-destinations places are distant from 4 borders

Value
a (k) contiguity matrice with the (k) contiguity measures

Examples
library(cartograflow)
data(flowdata)
bkg<- system.file("shape/MGP_TER.shp", package="cartograflow",
  lib.loc = NULL, mustWork = TRUE)
graph_ckij_1<-flowcontig(bkg,"EPT_NUM",ordre =1)
flowmap(graph_ckij_1,format="L",bkg,"EPT_NUM",
  filter = TRUE, taille = 0.5)
flowdist

Compute continuous distance matrix from geographical background

Description

From a geographical background computes (and threshold) a distance matrix.

Usage

flowdist(tab, dist.method, result)

Arguments

tab the input flow dataset.
dist.method euclidian calculation
result take the value "flowdist" or "dist" allows to parameter the resulting distance dataset (flows filtered by a distance criterion or not)

Details

– result = "dist" is the resulting tab of the distance
– result = "flowdist" with all the calculated parameters

Value

(1) A flowdata set with continuous euclidian distances calculations, see dist.method parameter
(2) A flowdata set with movement from euclidian distances calculations.
(3) A flowmap filtered by a global distance criterion.

Examples

library(cartograflow)
data(flowdata)
bkg<- system.file("shape/MGP_TER.shp", package="cartograflow",
lib.loc = NULL, mustWork = TRUE)
tab<-flowjointure(flows,bkg,"EPT_NUM")
#Format long with only origin, destination and distance parameters:
tab.distance<-flowdist(tab, dist.method = "euclidian",result = "dist")
#Format long with with all parameters: coordinates, distance, mouvement
tab.distance<-flowdist(tab, dist.method = "euclidian",result = "flowdist")
**flowgini**

Analysis of flow concentration (Gini coefficient)

**Description**

Calculates Gini coefficient, plot Lorenz curve and threshold the matrice according to a global concentration criterion for mapping flow intensity or flow density.

To be use before flowanalysis

**Usage**

`flowgini(tab, origin, dest, valflow, format, fdc, code, lorenz.plot)`

**Arguments**

- `tab`: dataset is a matrice or long format
- `origin`: origin place to be used with the long format
- `dest`: destination place to be used with the list format
- `valflow`: to be used with the list format
- `format`: is a variable that identify the data : matrice or long format
- `fdc`: is the map background file, ie. a shapefile.
- `code`: is the map background IDs code
- `lorenz.plot`: allows to plot the Lorenz curve associate to the gini coefficient

**Details**

`flowgini(...,lorenz.plot = TRUE)` for plotting Lorenz curve from cumulated flows and links.

**Value**

- plot Lorenz curve for the cumulated flow and links: `flowgini(...,gini.plot = TRUE)`
- warning: the function must be not assign a variable
- value of the Gini’s coefficient and the table: `table<-flowgini(...,missing(gini.plot) or gini.plot = FALSE)`

**References**


Examples

```r
library(cartograflow)
data(flowdata)
  bkg <- system.file("shape/MGP_TER.shp", package="cartograflow",
    lib.loc = NULL, mustWork = TRUE)
  # Computes Gini's coefficient
  tab_gini <- flowgini(flows, format = "L", origin = "i", dest = "j", valflow = "Fij",
    bkg, code = "EPT_NUM", lorenz.plot = FALSE)
  # Plot Lorenz curve
  flowgini(tab_gini, format = "L", origin = "i", dest = "j", valflow = "ydata",
    bkg, code = "EPT_NUM", lorenz.plot = TRUE)
  # See \link{flowanalysis} for viewing the tab_gini table
```

---

flowjointure

*Create a spatial join with flow*

Description

Create an attribute spatial join between a flow dataset table and a map background.

Usage

```r
flowjointure(tab, fdc, code)
```

Arguments

- **tab**: the input flow dataset table in long format
- **fdc**: the map background file, i.e., a shapefile
- **code**: the ID of the spatial units in the map background

Value

Resulting jointure table between flow dataset and map background

Examples

```r
library(cartograflow)
data(flowdata)
  bkg <- system.file("shape/MGP_TER.shp", package="cartograflow",
    lib.loc = NULL, mustWork = TRUE)
  tabflow <- flowjointure(flows, bkg, "EPT_NUM")
```
Description

Mapping a flow matrix origin-destination

Usage

flowmap(tab, format, fdc, code, filter, threshold, taille, a.head, 
a.length, a.angle, a.col)

Arguments

tab the input flow dataset .csv
format the flow dataset format : M=matrice or L=long.
fdc the geographical background file .shp
code is the column with the spatial units ID
filter allows you to filter (or not) the flow dataset. See details
threshold is the value of the threshold criterion used to filter the values. The default is 1.
taille is a graphical parameter for modifying the width of the feature
a.head integer code, determining the kind of arrows to be drawn. See Details
a.length length of the edges of the arrow head (in inches).
a.angle angle from the shaft of the arrow to the edge of the arrow head.
a.col color of the arrows

Details

The flow dataset must be converted to a dataframe for optimal performance (troubles remains with 
tibble format)

If filter = FALSE, all the matrice values are plot [(n*(n-1)] cells, i.e. all links out of the main 
diagonal. If filter = TRUE only non-zero values are plotted, i.e. existing links with or without 
threshold. The default threshold is set to 1.

a.head is for applying an arrow or not – code=0 : the link has no head - no arrow – code=1 : an 
arrow is draw at (x0[i], y0[i]). – code=2 : an arrow is draw at (x1[j], y1[j]) – code=3 : an arrow is 
draw at both nodes.

Value

a matrix or a list with the correct tabflow ID code

The resulting flowmap
Examples

```r
library(cartograflow)
data(flowdata)
bkg <- system.file("shape/MGP_TER.shp", package="cartograflow", lib.loc = NULL, mustWork = TRUE)
flowmap(flows, format = "L", bkg, code = "EPT_NUM", filter = TRUE,
        threshold = 20, taille = 5, a.head = 1, a.length = 0.05)
```

flowreduct

*Flow matrix reduction according to another matrix*

Description

Reduces a flow dataset according to an external matrix (e.g., distance travelled) Computes geographical movements (by weighting a flow dataset according to a distance criterion)

Usage

```r
flowreduct(tab, tab.metric, metric, select, d)
```

Arguments

- `tab` is the input flowdata set.
- `tab.metric` is the table of distance (continuous dataset) or contiguity (ordinal dataset)
- `metric` See Details.
- `select` is the continuous distance thresholding parameter. See Details.
- `d` distance thresholds criterion.

Details

- Metric is "continuous" for distance as euclidian, maximum, manhattan, etc. See `flowdist` for computing neighbourhood ordinal distance matrix.
- `select = dmin` for reducing flow dataset to values that are up or equal to the dmin distance parameter.
- `select = dmax` for reducing flow dataset to values that are less or equal to the dmax distance.
- Metric is "ordinal" for neighbourhood ordinal distance so-called k contiguity. See `flowcontig` for computing continuous distance matrix

Value

A flow dataset with continuous euclidian distances calculations
Examples

library(cartograflow)

data(flowdata)

bkg<- system.file("shape/MGP_TER.shp", package="cartograflow",
  lib.loc = NULL, mustWork = TRUE)

tab<-flowjointure(flows,bkg,"EPT_NUM")

#Example for reducing a flow matrice with a distance matrice, in long format (i,j, distance)
#1/2: Computes the matrice distances

tab.distance<-flowdist(tab, dist.method = "euclidian",result = "dist")

#2/2: Reduce the flow matrice

tab.flow<-flowreduct(tab,tab.distance, metric = "continous",
  select = "dmax", #maximum distance travelled criterion
  d = 8567) #maximum distance value

Description

Statistical dataset in .csv: Exraction of the french national census: "Mobilités professionnelles en 2015 :..."
- Base flux de mobilité - for the Greater Paris area.
Citation: INSEE - RP MOBPRO, 2015.
Variable (i) is the place of origin of the flow.
Variable (j) is the place of destination of the flow.
Variable (Fi) is the flow value for an (i,j) couple of origin-destination places.
Variable (count) is the frequency of the (i,j) couple of places.

Source

https://www.insee.fr/fr/statistiques/fichier/3566008/rp2015_mobpro_txt.zip

flowstructmat Structuring a matrix

Description

Fixes an ID shift in the flow matrix (to be use with flowjointure if necessary and flowtabmat)

Usage

flowstructmat(z)
### Arguments

- `z`  
  The flow dataset is in the matrice format where the first column is filled with the ID

### Value

A flowdataset with an usable format

### Examples

```r
library(cartograflow)
data(flowdata)

dim(mat_ex) # dimension fo the original matrice
### 10 11  # first column is fill with the ID

tab <- flowstructmat(mat_ex)
dim(tab)
### 10 10  # dimension fo the resulting matrice
```

### flowtabmat

**Changing the format of a flow dataset**

### Description

Transform a flow dataset from long to matrice format, and vice versa

### Usage

```r
flowtabmat(tab, matlist)
```

### Arguments

- `tab`  
  flow dataset, in matrice or long format
- `matlist`  
  choose matrice or long as the result format. See Details.

### Details

- From long to matrice format [n*n]: matlist="M"
- From matrice to long format [i,j,Fij]: matlist="L".

### Value

flow data in matrice or long format
Examples

library(cartograflow)
data(flowdata)
#1: From long to matrix format (n*m)
matFlow<-flowtabmat(flows,matlist="M")
#2: From matrix to long format [i,j,Fij]
listflow<-flowtabmat(matFlow,matlist="L")

flowtype

**Description**

Compute gross and net flows from initial asymmetric flow values

**Usage**

flowtype(tab, format, x)

**Arguments**

- **tab**: the input flow dataset
- **format**: specify the flow dataset format: M = square matrix [n*n] or L=lng [i,j.data]
- **x**: enter the computation type: "flux", "transpose", "bivolum" and "bisold".

**Details**

The matrix must be squared (if not, see flowcarre). This function compute for all (i,j) involved in an asymmetric flow matrix (Fij) several matrix. - x = "flux" for initial flow (Fij)
- x = "transpose" for reverse flow value (Fji)
- x = "bivolum" for bilateral gross flow Vij=(Fij+Fji)
- x = "bisold" for bilateral net flow Sij=(Fij-Fji)

**References**


**Examples**

library(cartograflow)
data(flowdata)
bkg<- system.file("shape/MGP_TER.shp", package="cartograflow",
lib.loc = NULL, mustWork = TRUE)

#1a:Computes flowtypes: Matrice format
```r
matflow<-flowtabmat(flows, matlist = "M")
m<-flowtype(matflow, format = "M", x = "flux")
m<-flowtype(matflow, format = "M", x = "transpose")
m<-flowtype(matflow, format = "M", x = "bivolum")
m<-flowtype(matflow, format = "M", x = "bisold")

#1b: Computes flowtypes: Long format
list<-flowtabmat(matflow, matlist="L")
colnames(list)<-c("i", "j", "Fij")
l_all<-flowtype(list, format = "L", x = "all")
l_sold<-flowtype(list, format = "L", x = "bisold")

#2: flowmapping: example of bisold
flowmap(l_sold, format = "L", bkg, code = "EPT_NUM",
        filter= TRUE, threshold= 20, taille = 5)
```

<table>
<thead>
<tr>
<th>geoid</th>
<th>Geographical ID</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Description**

One column dataframe in .csv. Variable (COD_GEO_EPT) is the geographical code of the territory. Citation: APUR, 2018

**Source**

[https://www.insee.fr/fr/statistiques/fichier/3566008/rp2015_mobpro_txt.zip](https://www.insee.fr/fr/statistiques/fichier/3566008/rp2015_mobpro_txt.zip)

---

**mat_ex**  
**Example of a small flow data matrix**

**Description**

Example of a small data, in format matrice
Index

cartograflow, 2
cartograflow-package (cartograflow), 2

flowanalysis, 2, 2, 7
flowcarre, 2, 4, 13
flowcontig, 2, 5, 10
flowdist, 2, 6, 10
flowgini, 2, 3, 7
flowjointure, 2, 8, 11
flowmap, 2, 5, 9
flowreduct, 2, 10
flows, 11
flowstructmat, 2, 11
flowtabmat, 11, 12
flowtype, 13

geoid, 14

mat_ex, 14