Package 'sabre'

September 17, 2018

Version 0.2.1

Title Spatial Association Between Regionalizations

Description Calculates a degree of spatial association between regionalizations or categorical maps using the information-theoretical V-measure (Nowosad and Stepinski (2018) <doi:10.1080/13658816.2018.1511794>). It also offers an R implementation of the MapCurve method (Hargrove et al. (2006) <doi:10.1007/s10109-006-0025-x>).

License MIT + file LICENSE

Encoding UTF-8

LazyData true

ByteCompile true

Suggests testthat, covr, knitr, rmarkdown

RoxygenNote 6.1.0

Depends R (>= 3.3.0)

Imports sf, entropy, tidyr, dplyr, rlang, tibble

VignetteBuilder knitr

URL https://github.com/Nowosad/sabre

BugReports https://github.com/Nowosad/sabre/issues

NeedsCompilation no

Author Jakub Nowosad [aut, cre] (<https://orcid.org/0000-0002-1057-3721>),
Tomasz Stepinski [aut],
Space Informatics Lab [cph]

Maintainer Jakub Nowosad <nowosad.jakub@gmail.com>

Repository CRAN

Date/Publication 2018-09-17 14:10:03 UTC

R topics documented:

`eco_us` *Ecoregions of the United States*

Description

Bailey's Ecoregions of the Conterminous United States

Usage

```
eco_us
```

Format

An object of class `sf` (inherits from `data.frame`) with 330 rows and 5 columns.

Source

https://www.sciencebase.gov/catalog/item/54244abde4b037b608f9e23d

---

`mapcurves` *Mapcurves*

Description

Mapcurves: a quantitative method for comparing categorical maps.

Usage

```
mapcurves(x, y, z = NULL)
```

Arguments

x	A numeric vector, representing a categorical values.
y	A numeric vector, representing a categorical values.
z	A numeric matrix. The goodness of fit (GOF) value fir each pair of classes in x and y. By default this argument is set to NULL, and the value of z is calculated based on x and y.

Value

A list with two elements:

- "ref_map" - the map to be used as reference ("x" or "y")
- "gof" - the Mapcurves's goodness of fit value

References

Hargrove, William W., Forrest M. Hoffman, and Paul F. Hessburg. "Mapcurves: a quantitative method for comparing categorical maps." Journal of Geographical Systems 8.2 (2006): 187.

Examples

```
set.seed(2018-03-21)
A = floor(matrix(runif(100, 0, 9), 10))
B = floor(matrix(runif(100, 0, 9), 10))
mapcurves(A, B)
```

---

mapcurves_calc *Mapcurves calculation*

Description

It calculates the Mapcurves's goodness-of-fit (GOF)

Usage

```
mapcurves_calc(x, x_name, y, y_name, precision = NULL)
```

Arguments

x	An object of class sf with a POLYGON or MULTIPOLYGON geometry type.
x_name	A name of the column with regions/clusters names.
y	An object of class sf with a POLYGON or MULTIPOLYGON geometry type.
y_name	A name of the column with regions/clusters names.
precision	numeric; see st_as_binary for how to do this.

Value

A list with four elements:

- "map1" - the sf object containing the first map used for calculation of GOF
- "map2" - the sf object containing the second map used for calculation of GOF
- "ref_map" - the map used as a reference ("x" or "y")
- "gof" - the Mapcurves's goodness of fit value

References

Hargrove, William W., Forrest M. Hoffman, and Paul F. Hessburg. "Mapcurves: a quantitative method for comparing categorical maps." Journal of Geographical Systems 8.2 (2006): 187.

Examples

```
library(sf)
data("regions1")
data("regions2")

mc = mapcurves_calc(regions1, z, regions2, z)
mc

plot(mc$map1)
plot(mc$map2)
```

---

regions1 *Red regionalization*

Description

Data of the red regionalization used in Figure 1 of Stepinski and Nowosad (2018)

Usage

```
regions1
```

Format

An object of class `sf` (inherits from `data.frame`) with 4 rows and 2 columns.

References

Stepinski, Tomasz, and Jakub Nowosad. "Assessing a degree of spatial association between regionalizations or categorical maps using the information-theoretical V-measure."

regions2 *Blue regionalization*

Description

Data of the blue regionalization used in Figure 1 of Stepinski and Nowosad (2018)

Usage

```
regions2
```

Format

An object of class `sf` (inherits from `data.frame`) with 3 rows and 2 columns.

References

Stepinski, Tomasz, and Jakub Nowosad. "Assessing a degree of spatial association between regionalizations or categorical maps using the information-theoretical V-measure."

vmeasure *V-measure*

Description

A conditional entropy-based external cluster evaluation measure.

Usage

```
vmeasure(x, y, z = NULL, B = 1)
```

Arguments

`x`	A numeric vector, representing a categorical values.
`y`	A numeric vector, representing a categorical values.
`z`	A numeric matrix. A contingency table of the counts at each combination of categorical levels. By default this argument is set to `NULL`, and the value of `z` is calculated based on `x` and `y`.
`B`	A numeric value. If `B` > 1 then completeness is weighted more strongly than homogeneity, and if `B` < 1 then homogeneity is weighted more strongly than completeness. By default this value is 1.

Value

A list with three elements:

- "v_measure"
- "homogeneity"
- "completeness"

References

Rosenberg, Andrew, and Julia Hirschberg. "V-measure: A conditional entropy-based external cluster evaluation measure." Proceedings of the 2007 joint conference on empirical methods in natural language processing and computational natural language learning (EMNLP-CoNLL). 2007.

Examples

```
x = c(1, 1, 1, 2, 2, 3, 3, 3, 1, 1, 2, 2, 2, 3, 3)
y = c(rep(1, 5), rep(2, 5), rep(3, 5))
vmeasure(x, y)
```

---

vmeasure_calc *V-measure calculation*

Description

It calculates a degree of spatial association between regionalizations using an information-theoretical measure called the V-measure

Usage

```
vmeasure_calc(x, x_name, y, y_name, B = 1, precision = NULL)
```

Arguments

x	An object of class sf with a POLYGON or MULTIPOLYGON geometry type.
x_name	A name of the column with regions/clusters names.
y	An object of class sf with a POLYGON or MULTIPOLYGON geometry type.
y_name	A name of the column with regions/clusters names.
B	A numeric value. If B > 1 then completeness is weighted more strongly than homogeneity, and if B < 1 then homogeneity is weighted more strongly than completeness. By default this value is 1.
precision	numeric; see st_as_binary for how to do this.

Value

A list with five elements:

- "map1" - the sf object containing the first preprocessed map used for calculation of GOF with two attributes - `map1` (name of the category) and `rih` (region inhomogeneity)
- "map2" - the sf object containing the second preprocessed map used for calculation of GOF with two attributes - `map1` (name of the category) and `rih` (region inhomogeneity)
- "v_measure"
- "homogeneity"
- "completeness"

References

Nowosad, Jakub, and Tomasz F. Stepinski. "Spatial association between regionalizations using the information-theoretical V-measure." International Journal of Geographical Information Science (2018). https://doi.org/10.1080/13658816.2018.1511794

Rosenberg, Andrew, and Julia Hirschberg. "V-measure: A conditional entropy-based external cluster evaluation measure." Proceedings of the 2007 joint conference on empirical methods in natural language processing and computational natural language learning (EMNLP-CoNLL). 2007.

Examples

```
library(sf)
data("regions1")
data("regions2")
vm = vmeasure_calc(regions1, z, regions2, z)
vm

plot(vm$map1["rih"])
plot(vm$map2["rih"])
```

Index