Package ‘gdverse’

February 17, 2025

Title Analysis of Spatial Stratified Heterogeneity

Version 1.3-2

Description Analyzing spatial factors and exploring spatial associations based on the concept of spatial stratified heterogeneity, while also taking into account local spatial dependencies, spatial interpretability, complex spatial interactions, and robust spatial stratification. Additionally, it supports the spatial stratified heterogeneity family established in academic literature.

License GPL-3

Encoding UTF-8

RoxygenNote 7.3.2

URL https://stscl.github.io/gdverse/, https://github.com/stscl/gdverse

BugReports https://github.com/stscl/gdverse/issues

Depends R (>= 4.1.0)

Imports dplyr, forcats, ggplot2, magrittr, parallel, patchwork, purrr, reticulate, rpart, scatterpie, sdsfun (>= 0.7.0), sf, stats, tibble, tidyr, utils

Suggests cowplot, itmsa, knitr, Rcpp, readr, rmarkdown, terra, testthat (>= 3.0.0)

LinkingTo Rcpp

VignetteBuilder knitr

LazyData true

Config/testthat/edition 3

NeedsCompilation yes

Author Wenbo Lv [aut, cre, cph] (<https://orcid.org/0009-0002-6003-3800>),
Yangyang Lei [aut] (<https://orcid.org/0009-0002-6518-3613>),
Fangmei Liu [aut] (<https://orcid.org/0009-0008-8634-7588>),
Jianwu Yan [aut] (<https://orcid.org/0000-0001-9411-8453>),
Yongze Song [aut] (<https://orcid.org/0000-0003-3420-9622>),
Wufan Zhao [aut] (<https://orcid.org/0000-0002-0265-3465>)

Maintainer Wenbo Lv <lyu.geosocial@gmail.com>

Repository CRAN

Date/Publication 2025-02-17 09:00:02 UTC

Contents

`all2int` *convert all discretized vectors to integer*

Description

convert all discretized vectors to integer

Usage

```
all2int(x)
```

Arguments

x	A discretized vector.

Value

An integer vector.

Examples

```
all2int(factor(letters[1:3],levels = c('b','a','c')))
all2int(letters[1:3])
```

---

cpsd_disc *optimal spatial data discretization based on SPADE q-statistics*

Description

Function for determining the optimal spatial data discretization based on SPADE q-statistics.

Usage

```
cpsd_disc(
  formula,
  data,
  wt,
  discnum = 3:8,
  discmethod = "quantile",
  strategy = 2L,
  increase_rate = 0.05,
  cores = 1,
  seed = 123456789,
  ...
)
```

Arguments

formula	A formula of optimal spatial data discretization.
data	A `data.frame`, `tibble` or `sf` object of observation data.
wt	The spatial weight matrix.
discnum	(optional) A vector of number of classes for discretization. Default is `3:8`.
discmethod	(optional) The discretization methods. Default all use `quantile`. Noted that `rpart` will use `rpart_disc()`; Others use `sdsfun::discretize_vector()`.
strategy	(optional) Discretization strategy. When `strategy` is `1L`, choose the highest SPADE model q-statistics to determinate optimal spatial data discretization parameters. When `strategy` is `2L`, The optimal discrete parameters of spatial data are selected by combining LOESS model.

`increase_rate`	(optional) The critical increase rate of the number of discretization. Default is `5%`.
`cores`	(optional) Positive integer (default is 1). When cores are greater than 1, use multi-core parallel computing.
`seed`	(optional) Random seed number, default is `123456789`.
`...`	(optional) Other arguments passed to `sdsfun::discretize_vector()` or `rpart_disc()`.

Value

A list.

`x` discretization variable name

`k` optimal number of spatial data discreteization

`method` optimal spatial data discretization method

`disc` the result of optimal spatial data discretization

Note

When the `discmethod` is configured to `robust`, it will operate at a significantly reduced speed. Consequently, the use of robust discretization is not advised.

Author(s)

Wenbo Lv <`lyu.geosocial@gmail.com`>

References

Yongze Song & Peng Wu (2021) An interactive detector for spatial associations, International Journal of Geographical Information Science, 35:8, 1676-1701, DOI:10.1080/13658816.2021.1882680

Examples

```
data('sim')
wt = sdsfun::inverse_distance_swm(sf::st_as_sf(sim,coords = c('lo','la')))
cpsd_disc(y ~ xa + xb + xc, data = sim, wt = wt)
```

`cpsd_spade` *compensated power of spatial determinant(CPSD)*

Description

Function for calculate compensated power of spatial determinant `Q_s`.

Usage

```
cpsd_spade(yobs, xobs, xdisc, wt)
```

Arguments

yobs	Variable Y
xobs	The original undiscretized covariable X.
xdisc	The discretized covariable X.
wt	The spatial weight matrix.

Details

The power of compensated spatial determinant formula is

$$Q_s = \frac{q_s}{q_{s_{inforkep}}} = \frac{1 - \frac{\sum_{h=1}^{L} N_h \Gamma_{kdep}}{N \Gamma_{totaldep}}}{1 - \frac{\sum_{h=1}^{L} N_h \Gamma_{hind}}{N \Gamma_{totalind}}}$$

Value

A value of compensated power of spatial determinant `Q_s`.

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

References

Xuezhi Cang & Wei Luo (2018) Spatial association detector (SPADE),International Journal of Geographical Information Science, 32:10, 2055-2075, DOI: 10.1080/13658816.2018.1476693

Examples

```
data('sim')
wt = sdsfun::inverse_distance_swm(sf::st_as_sf(sim,coords = c('lo','la')))
xa = sim$xa
xa_disc = sdsfun::discretize_vector(xa,5)
cpsd_spade(sim$y,xa,xa_disc,wt)
```

---

ecological_detector *ecological detector*

Description

Compare the effects of two factors X_1 and X_2 on the spatial distribution of the attribute Y.

Usage

```
ecological_detector(y, x1, x2, alpha = 0.95)
```

Arguments

y	Dependent variable, continuous numeric vector.
x1	Covariate X_1, factor, character or discrete numeric.
x2	Covariate X_2, factor, character or discrete numeric.
alpha	(optional) Confidence level of the interval,default is 0.95.

Value

A list.

F-statistic the result of F statistic for ecological detector

P-value the result of P value for ecological detector

Ecological is there a significant difference between the two factors X_1 and X_2 on the spatial distribution of the attribute Y

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

Examples

```
ecological_detector(y = 1:7,
                    x1 = c('x',rep('y',3),rep('z',3)),
                    x2 = c(rep('a',2),rep('b',2),rep('c',3)))
```

---

factor_detector *factor detector*

Description

The factor detector q-statistic measures the spatial stratified heterogeneity of a variable Y, or the determinant power of a covariate X of Y.

Usage

```
factor_detector(y, x, confintv = FALSE, alpha = 0.95)
```

Arguments

y	Variable Y, continuous numeric vector.
x	Covariate X, factor, character or discrete numeric.
confintv	(optional) Whether to compute the confidence interval for the q statistic, default is FALSE.
alpha	(optional) Confidence level of the interval, default is 0.95.

Value

A list.

`Q-statistic` the q statistic for factor detector

`P-value` the p value for factor detector

`CIL` the confidence interval lower bound

`CIU` the confidence interval upper bound

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

Examples

```
factor_detector(y = 1:7,x = c('x',rep('y',3),rep('z',3)))
```

F_informationloss *measure information loss by information entropy*

Description

Function for measure information loss by shannon information entropy.

Usage

```
F_informationloss(xvar, xdisc)
```

Arguments

`xvar`	The original undiscretized vector.
`xdisc`	The discretized vector.

Details

The information loss measured by information entropy formula is $F = -\sum\limits_{i=1}^{N} p_{(i)} \log_2 p_{(i)} - \left(-\sum\limits_{h=1}^{L} p_{(h)} \log_2 p_{(h)}\right)$

Value

A numeric value of information loss measured by information entropy.

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

Examples

```
F_informationloss(1:7,c('x',rep('y',3),rep('z',3)))
```

---

gd   *native geographical detector(GD) model*

Description

native geographical detector(GD) model

Usage

```
gd(formula, data, type = "factor", alpha = 0.95)
```

Arguments

`formula`	A formula of geographical detector model.
`data`	A `data.frame`, `tibble` or `sf` object of observation data.
`type`	(optional) The type of geographical detector, which must be one of `factor`(default), `interaction`, `risk`, `ecological`. You can run one or more types at one time.
`alpha`	(optional) Specifies the size of the alpha (confidence level). Default is `0.95`.

Value

A list.

`factor` the result of factor detector

`interaction` the result of interaction detector

`risk` the result of risk detector

`ecological` the result of ecological detector

Author(s)

Wenbo Lv <`lyu.geosocial@gmail.com`>

References

Jin-Feng Wang, Xin-Hu Li, George Christakos, Yi-Lan Liao, Tin Zhang, XueGu & Xiao-Ying Zheng (2010) Geographical Detectors-Based Health Risk Assessment and its Application in the Neural Tube Defects Study of the Heshun Region, China, International Journal of Geographical Information Science, 24:1, 107-127, DOI: 10.1080/13658810802443457

Examples

```
data("NTDs")
g = gd(incidence ~ watershed + elevation + soiltype,
       data = NTDs,type = c('factor','interaction'))
g
```

---

gd_optunidisc *optimal univariate discretization based on geodetector q-statistic*

Description

optimal univariate discretization based on geodetector q-statistic

Usage

```
gd_optunidisc(
  formula,
  data,
  discnum = 3:8,
  discmethod = c("sd", "equal", "geometric", "quantile", "natural"),
  cores = 1,
  seed = 123456789,
  ...
)
```

Arguments

formula	A formula.
data	A data.frame, tibble or sf object of observation data.
discnum	(optional) A vector of numbers of discretization. Default is 3:8.
discmethod	(optional) A vector of methods for discretization, default is using c("sd","equal","geometric","quan by invoking sdsfun.
cores	(optional) Positive integer (default is 1). When cores are greater than 1, use multi-core parallel computing.
seed	(optional) Random seed number, default is 123456789.
...	(optional) Other arguments passed to sdsfun::discretize_vector().

Value

A list.

x the name of the variable that needs to be discretized

k optimal discretization number

method optimal discretization method

qstatistic optimal q-statistic

disc optimal discretization results

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

Examples

```
data('sim')
gd_optunidisc(y ~ xa + xb + xc,
              data = sim,
              discnum = 3:6)
```

`gen_permutations` *generate permutations*

Description

generate permutations

Usage

```
gen_permutations(x, seed = 123456789)
```

Arguments

x	A vector.
seed	(optional) Random seed number. Default is 123456789.

Value

A permutations vector.

Examples

```
gen_permutations(1:100,42)
```

geodetector	*geographical detector*

Description

geographical detector

Usage

```
geodetector(formula, data, type = "factor", alpha = 0.95)
```

Arguments

`formula`	A formula of geographical detector model.
`data`	A data.frame or tibble of observation data.
`type`	(optional) The type of geographical detector, which must be one of `factor`(default), `interaction`, `risk`, `ecological`.
`alpha`	(optional) Specifies the size of the alpha (confidence level). Default is `0.95`.

Value

A list.

`factor` the result of factor detector

`interaction` the result of interaction detector

`risk` the result of risk detector

`ecological` the result of ecological detector

Note

Note that only one type of geodetector is supported at a time in `geodetector()`.

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

Examples

```
geodetector(y ~ x1 + x2,
   tibble::tibble(y = 1:7,
                  x1 = c('x',rep('y',3),rep('z',3)),
                  x2 = c(rep('a',2),rep('b',2),rep('c',3))))

geodetector(y ~ x1 + x2,
   tibble::tibble(y = 1:7,
                  x1 = c('x',rep('y',3),rep('z',3)),
                  x2 = c(rep('a',2),rep('b',2),rep('c',3))),
```

```
    type = 'interaction')

  geodetector(y ~ x1 + x2,
    tibble::tibble(y = 1:7,
                   x1 = c('x',rep('y',3),rep('z',3)),
                   x2 = c(rep('a',2),rep('b',2),rep('c',3))),
    type = 'risk',alpha = 0.95)

  geodetector(y ~ x1 + x2,
    tibble::tibble(y = 1:7,
                   x1 = c('x',rep('y',3),rep('z',3)),
                   x2 = c(rep('a',2),rep('b',2),rep('c',3))),
    type = 'ecological',alpha = 0.95)
```

---

gozh *geographically optimal zones-based heterogeneity(GOZH) model*

Description

geographically optimal zones-based heterogeneity(GOZH) model

Usage

```
gozh(formula, data, cores = 1, type = "factor", alpha = 0.95, ...)
```

Arguments

`formula`	A formula of GOZH model.
`data`	A `data.frame`, `tibble` or `sf` object of observation data.
`cores`	(optional) Positive integer (default is 1). When cores are greater than 1, use multi-core parallel computing.
`type`	(optional) The type of geographical detector, which must be `factor`(default), `interaction`, `risk`, `ecological`. You can run one or more types at one time.
`alpha`	(optional) Specifies the size of confidence level. Default is `0.95`.
`...`	(optional) Other arguments passed to `rpart_disc()`.

Value

A list.

`factor` the result of factor detector

`interaction` the result of interaction detector

`risk` the result of risk detector

`ecological` the result of ecological detector

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

References

Luo, P., Song, Y., Huang, X., Ma, H., Liu, J., Yao, Y., & Meng, L. (2022). Identifying determinants of spatio-temporal disparities in soil moisture of the Northern Hemisphere using a geographically optimal zones-based heterogeneity model. ISPRS Journal of Photogrammetry and Remote Sensing: Official Publication of the International Society for Photogrammetry and Remote Sensing (ISPRS), 185, 111–128. https://doi.org/10.1016/j.isprsjprs.2022.01.009

Examples

```
data('ndvi')
g = gozh(NDVIchange ~ ., data = ndvi)
g
```

`gozh_detector` *geographically optimal zones-based heterogeneity detector*

Description

Function for geographically optimal zones-based heterogeneity detector.

Usage

```
gozh_detector(formula, data, cores = 1, type = "factor", alpha = 0.95, ...)
```

Arguments

`formula`	A formula of GOZH detector.
`data`	A data.frame or tibble of observation data.
`cores`	(optional) Positive integer (default is 1). When cores are greater than 1, use multi-core parallel computing.
`type`	(optional) The type of geographical detector,which must be one of `factor`(default), `interaction`, `risk`, `ecological`.
`alpha`	(optional) Confidence level of the interval,default is `0.95`.
`...`	(optional) Other arguments passed to `rpart_disc()`.

Value

A list of tibble with the corresponding result under different detector types.

`factor` the result of factor detector

`interaction` the result of interaction detector

`risk` the result of risk detector

`ecological` the result of ecological detector

Note

Only one type of detector is supported in a `gozh_detector()` run at a time.

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

References

Luo, P., Song, Y., Huang, X., Ma, H., Liu, J., Yao, Y., & Meng, L. (2022). Identifying determinants of spatio-temporal disparities in soil moisture of the Northern Hemisphere using a geographically optimal zones-based heterogeneity model. ISPRS Journal of Photogrammetry and Remote Sensing: Official Publication of the International Society for Photogrammetry and Remote Sensing (ISPRS), 185, 111–128. https://doi.org/10.1016/j.isprsjprs.2022.01.009

Examples

```
data('ndvi')
g = gozh_detector(NDVIchange ~ ., data = ndvi)
g
```

---

idsa — *interactive detector for spatial associations(IDSA) model*

Description

interactive detector for spatial associations(IDSA) model

Usage

```
idsa(
  formula,
  data,
  wt = NULL,
  discnum = 3:8,
  discmethod = "quantile",
  overlay = "and",
  strategy = 2L,
  increase_rate = 0.05,
  cores = 1,
  seed = 123456789,
  alpha = 0.95,
  ...
)
```

Arguments

- `formula` A formula of IDSA model.
- `data` A `data.frame`, `tibble` or `sf` object of observation data.
- `wt` (optional) The spatial weight matrix. When `data` is not an `sf` object, must provide `wt`.
- `discnum` (optional) Number of multilevel discretization. Default will use `3:8`.
- `discmethod` (optional) The discretization methods. Default all use `quantile`. Noted that `rpart` will use `rpart_disc()`; Others use `sdsfun::discretize_vector()`.
- `overlay` (optional) Spatial overlay method. One of `and`, `or`, `intersection`. Default is `and`.
- `strategy` (optional) Discretization strategy. When `strategy` is `1L`, choose the highest SPADE model q-statistics to determinate optimal spatial data discretization parameters. When `strategy` is `2L`, The optimal discrete parameters of spatial data are selected by combining LOESS model.
- `increase_rate` (optional) The critical increase rate of the number of discretization. Default is `5%`.
- `cores` (optional) Positive integer (default is 1). When cores are greater than 1, use multi-core parallel computing.
- `seed` (optional) Random number seed, default is `123456789`.
- `alpha` (optional) Specifies the size of confidence level. Default is `0.95`.
- `...` (optional) Other arguments passed to `cpsd_disc()`.

Value

A list.

- `interaction` the interaction result of IDSA model
- `risk` whether values of the response variable between a pair of overlay zones are significantly different
- `number_individual_explanatory_variables` the number of individual explanatory variables used for examining the interaction effects
- `number_overlay_zones` the number of overlay zones
- `percentage_finely_divided_zones` the percentage of finely divided zones that are determined by the interaction of variables

Note

Please note that all variables in the IDSA model need to be continuous data.

The IDSA model requires at least $2^n - 1$ calculations when has n explanatory variables. When there are more than 10 explanatory variables, carefully consider the computational burden of this model. When there are a large number of explanatory variables, the data dimensionality reduction method can be used to ensure the trade-off between analysis results and calculation speed.

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

References

Yongze Song & Peng Wu (2021) An interactive detector for spatial associations, International Journal of Geographical Information Science, 35:8, 1676-1701, DOI:10.1080/13658816.2021.1882680

Examples

```
data('sim')
sim1 = sf::st_as_sf(sim,coords = c('lo','la'))
g = idsa(y ~ ., data = sim1)
g
```

---

interaction_detector *interaction detector*

Description

Identify the interaction between different risk factors, that is, assess whether factors X1 and X2 together increase or decrease the explanatory power of the dependent variable Y, or whether the effects of these factors on Y are independent of each other.

Usage

```
interaction_detector(y, x1, x2)
```

Arguments

y	Dependent variable, continuous numeric vector.
x1	Covariate X_1, factor, character or discrete numeric.
x2	Covariate X_2, factor, character or discrete numeric.

Value

A list.

`Variable1 Q-statistics` Q-statistics for variable1

`Variable2 Q-statistics` Q-statistics for variable2

`Variable1 and Variable2 interact Q-statistics` Q-statistics for variable1 and variable2 interact

`Interaction` the interact result type

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

Examples

```
interaction_detector(y = 1:7,
                     x1 = c('x',rep('y',3),rep('z',3)),
                     x2 = c(rep('a',2),rep('b',2),rep('c',3)))
```

`lesh` *locally explained stratified heterogeneity(LESH) model*

Description

locally explained stratified heterogeneity(LESH) model

Usage

```
lesh(formula, data, cores = 1, ...)
```

Arguments

`formula`	A formula of LESH model.
`data`	A `data.frame`, `tibble` or `sf` object of observation data.
`cores`	(optional) Positive integer (default is 1). When cores are greater than 1, use multi-core parallel computing.
`...`	(optional) Other arguments passed to `rpart_disc()`.

Value

A list.

`interaction` the interaction result of LESH model

`spd_lesh` a tibble of the shap power of determinants

Note

The LESH model requires at least $2^n - 1$ calculations when has n explanatory variables. When there are more than 10 explanatory variables, carefully consider the computational burden of this model. When there are a large number of explanatory variables, the data dimensionality reduction method can be used to ensure the trade-off between analysis results and calculation speed.

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

References

Li, Y., Luo, P., Song, Y., Zhang, L., Qu, Y., & Hou, Z. (2023). A locally explained heterogeneity model for examining wetland disparity. International Journal of Digital Earth, 16(2), 4533–4552. https://doi.org/10.1080/17538947.2023.2271883

Examples

```
data('ndvi')
g = lesh(NDVIchange ~ ., data = ndvi)
g
```

---

loess_optscale *determine optimal spatial data analysis scale*

Description

Function for determining optimal spatial data analysis scale based on locally estimated scatter plot smoothing (LOESS) model.

Usage

```
loess_optscale(qvec, spscalevec, increase_rate = 0.05)
```

Arguments

qvec	A numeric vector of q statistics.
spscalevec	A numeric vector of spatial scales corresponding to qvec.
increase_rate	(optional) The critical increase rate of the number of discretization. Default is 5%.

Value

A numeric vector about optimal number of spatial scale and the critical increase rate of q value.

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

Examples

```
## Not run:
## The following code takes a long time to run:
library(tidyverse)
fvcpath = "https://github.com/SpatLyu/rdevdata/raw/main/FVC.tif"
fvc = terra::rast(paste0("/vsicurl/",fvcpath))
fvc1000 = fvc %>%
  terra::as.data.frame(na.rm = T) %>%
```

```
  as_tibble()
fvc5000 = fvc %>%
  terra::aggregate(fact = 5) %>%
  terra::as.data.frame(na.rm = T) %>%
  as_tibble()
qv1000 = factor_detector(fvc1000$fvc,
                         sdsfun::discretize_vector(fvc1000$premax,10,'quantile'))[[1]]
qv5000 = factor_detector(fvc5000$fvc,
                         sdsfun::discretize_vector(fvc5000$premax,10,'quantile'))[[1]]
loess_optscale(c(qv1000,qv5000),c(1000,5000))

## End(Not run)
```

---

ndvi *dataset of NDVI changes and its influencing factors*

Description

dataset of NDVI changes and its influencing factors, modified from GD package.

Usage

```
ndvi
```

Format

ndvi: A tibble with 713 rows and 7 variables

Author(s)

Yongze Song <yongze.song@outlook.com>

NTDs *NTDs data*

Description

The data were obtained by preprocessing use sf and tidyverse.

Usage

```
NTDs
```

Format

NTDs: A tibble with 185 rows and 4 variable columns and 2 location columns, modified from geodetector package.

opgd *optimal parameters-based geographical detector(OPGD) model*

Description

optimal parameters-based geographical detector(OPGD) model

Usage

```
opgd(
  formula,
  data,
  discvar = NULL,
  discnum = 3:8,
  discmethod = c("sd", "equal", "geometric", "quantile", "natural"),
  cores = 1,
  type = "factor",
  alpha = 0.95,
  ...
)
```

Arguments

`formula`	A formula of OPGD model.
`data`	A `data.frame`, `tibble` or `sf` object of observation data.
`discvar`	Name of continuous variable columns that need to be discretized. Noted that when `formula` has `discvar`, `data` must have these columns. By default, all independent variables are used as `discvar`.
`discnum`	(optional) A vector of number of classes for discretization. Default is `3:8`.
`discmethod`	(optional) A vector of methods for discretization, default is using `c("sd","equal","geometric","quan` by invoking `sdsfun`.
`cores`	(optional) Positive integer (default is 1). When cores are greater than 1, use multi-core parallel computing.
`type`	(optional) The type of geographical detector, which must be `factor`(default), `interaction`, `risk`, `ecological`. You can run one or more types at one time.
`alpha`	(optional) Specifies the size of confidence level. Default is `0.95`.
`...`	(optional) Other arguments passed to `gd_optunidisc()`. A useful parameter is `seed`, which is used to set the random number seed.

Value

A list.

`opt_param` optimal discretization parameter

`factor` the result of factor detector

`interaction` the result of interaction detector

`risk` the result of risk detector

`ecological` the result of ecological detector

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

References

Song, Y., Wang, J., Ge, Y. & Xu, C. (2020) An optimal parameters-based geographical detector model enhances geographic characteristics of explanatory variables for spatial heterogeneity analysis: Cases with different types of spatial data, GIScience & Remote Sensing, 57(5), 593-610. doi: 10.1080/15481603.2020.1760434.

Examples

```
data('sim')
opgd(y ~ xa + xb + xc, data = sim,
     discvar = paste0('x',letters[1:3]),
     discnum = 3:6)
```

`pid_idsa` *IDSA Q-saistics* `PID`

Description

IDSA Q-saistics `PID`

Usage

```
pid_idsa(formula, rawdata, discdata, wt, overlaymethod = "and")
```

Arguments

`formula`	A formula for IDSA Q-saistics
`rawdata`	Raw observation data
`discdata`	Observed data with discrete explanatory variables
`wt`	Spatial weight matrix
`overlaymethod`	(optional) Spatial overlay method. One of `and`, `or`, `intersection`. Default is `and`.

Details

$Q_{IDSA} = \frac{\theta_r}{\phi}$

Value

The value of IDSA Q-saistics `PID`.

Examples

```
data('sim')
wt = sdsfun::inverse_distance_swm(sf::st_as_sf(sim,coords = c('lo','la')))
sim1 = dplyr::mutate(sim,dplyr::across(xa:xc,\(.x) sdsfun::discretize_vector(.x,5)))
pid_idsa(y ~ xa + xb + xc, rawdata = sim,
         discdata = sim1, wt = wt)
```

---

plot.ecological_detector

plot ecological detector

Description

S3 method to plot output for ecological detector in `geodetector()`.

Usage

```
## S3 method for class 'ecological_detector'
plot(x, ...)
```

Arguments

x	Return by `geodetector()`.
...	(optional) Other arguments passed to `ggplot2::theme()`.

Value

A ggplot2 layer

Author(s)

Wenbo Lv <`lyu.geosocial@gmail.com`>

plot.factor_detector *plot factor detector result*

Description

S3 method to plot output for factor detector in `geodetector()`.

Usage

```
## S3 method for class 'factor_detector'
plot(x, slicenum = 2, alpha = 0.95, keep = TRUE, qlabelsize = 3.88, ...)
```

Arguments

x	Return by `geodetector()`.
slicenum	(optional) The number of labels facing inward. Default is `2`.
alpha	(optional) Confidence level. Default is `0.95`.
keep	(optional) Whether to keep Q-value results for insignificant variables, default is `TRUE`.
qlabelsize	(optional) Set the font size of the q-value text labels in the plot.
...	(optional) Other arguments passed to `ggplot2::theme()`.

Value

A ggplot2 layer.

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

plot.gd_result *plot GD result*

Description

S3 method to plot output for GD model result in `gd()`.

Usage

```
## S3 method for class 'gd_result'
plot(x, ...)
```

Arguments

`x`	Return by `gd()`.
`...`	(optional) Other arguments.

Value

A ggplot2 layer

Author(s)

Wenbo Lv <`lyu.geosocial@gmail.com`>

`plot.gozh_result` *plot GOZH result*

Description

S3 method to plot output for GOZH model result in `gozh()`.

Usage

```
## S3 method for class 'gozh_result'
plot(x, ...)
```

Arguments

`x`	Return by `gozh()`.
`...`	(optional) Other arguments.

Value

A ggplot2 layer

Author(s)

Wenbo Lv <`lyu.geosocial@gmail.com`>

plot.idsa_result *plot IDSA risk result*

Description

S3 method to plot output for IDSA risk result in `idsa()`.

Usage

```
## S3 method for class 'idsa_result'
plot(x, ...)
```

Arguments

`x`	Return by `idsa()`.
`...`	(optional) Other arguments passed to `ggplot2::theme()`.

Value

A ggplot2 layer

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

plot.interaction_detector *plot interaction detector result*

Description

S3 method to plot output for interaction detector in `geodetector()`.

Usage

```
## S3 method for class 'interaction_detector'
plot(x, alpha = 1, ...)
```

Arguments

`x`	Return by `geodetector()`.
`alpha`	(optional) Picture transparency. Default is `1`.
`...`	(optional) Other arguments passed to `ggplot2::theme()`.

Value

A ggplot2 layer

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

`plot.lesh_result` *plot LESH model result*

Description

S3 method to plot output for LESH model interaction result in `lesh()`.

Usage

```
## S3 method for class 'lesh_result'
plot(
  x,
  pie = TRUE,
  scatter = FALSE,
  scatter_alpha = 1,
  pieradius_factor = 15,
  pielegend_x = 0.99,
  pielegend_y = 0.1,
  pielegend_num = 3,
  ...
)
```

Arguments

`x`	x Return by `lesh()`.
`pie`	(optional) Whether to draw the interaction contributions. Default is `TRUE`.
`scatter`	(optional) Whether to draw the interaction direction diagram. Default is `FALSE`.
`scatter_alpha`	(optional) Picture transparency. Default is `1`.
`pieradius_factor`	(optional) The radius expansion factor of interaction contributions pie plot. Default is `15`.
`pielegend_x`	(optional) The X-axis relative position of interaction contributions pie plot legend. Default is `0.99`.
`pielegend_y`	(optional) The Y-axis relative position of interaction contributions pie plot legend. Default is `0.1`.
`pielegend_num`	(optional) The number of interaction contributions pie plot legend. Default is `3`.
`...`	(optional) Other arguments passed to `ggplot2::theme()`.

Value

A ggplot2 layer.

Note

When both `scatter` and `pie` are set to `TRUE` in RStudio, enlarge the drawing frame for normal display.

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

---

plot.opgd_result *plot OPGD result*

Description

S3 method to plot output for OPGD model result in `opgd()`.

Usage

```
## S3 method for class 'opgd_result'
plot(x, ...)
```

Arguments

x	Return by `opgd()`.
...	(optional) Other arguments.

Value

A ggplot2 layer

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

plot.rgd_result *plot RGD result*

Description

S3 method to plot output for RGD model result in `rgd()`.

Usage

```
## S3 method for class 'rgd_result'
plot(x, slicenum = 2, alpha = 0.95, keep = TRUE, ...)
```

Arguments

`x`	Return by `rgd()`.
`slicenum`	(optional) The number of labels facing inward. Default is `2`.
`alpha`	(optional) Confidence level. Default is `0.95`.
`keep`	(optional) Whether to keep Q-value results for insignificant variables, default is `TRUE`.
`...`	(optional) Other arguments passed to `ggplot2::theme()`.

Value

A ggplot2 layer

Author(s)

Wenbo Lv <`lyu.geosocial@gmail.com`>

plot.rid_result *plot RID result*

Description

S3 method to plot output for RID model from `rid()`.

Usage

```
## S3 method for class 'rid_result'
plot(x, alpha = 1, ...)
```

Arguments

`x`	Return by `rid()`.
`alpha`	(optional) Picture transparency. Default is `1`.
`...`	(optional) Other arguments passed to `ggplot2::theme()`.

Value

A ggplot2 layer

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

---

plot.risk_detector *plot risk detector*

Description

S3 method to plot output for risk detector in `geodetector()`.

Usage

```
## S3 method for class 'risk_detector'
plot(x, ...)
```

Arguments

`x`	Return by `geodetector()`.
`...`	(optional) Other arguments passed to `ggplot2::theme()`.

Value

A ggplot2 layer

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

---

plot.sesu_gozh *plot gozh sesu*

Description

S3 method to plot output for gozh sesu in `sesu_gozh()`.

Usage

```
## S3 method for class 'sesu_gozh'
plot(x, ...)
```

Arguments

`x`	Return by `sesu_gozh()`.
`...`	(optional) Other arguments passed to `ggplot2::theme()`.

Value

A ggplot2 layer.

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

---

plot.sesu_opgd *plot opgd sesu*

Description

S3 method to plot output for opgd sesu in `sesu_opgd()`.

Usage

```
## S3 method for class 'sesu_opgd'
plot(x, ...)
```

Arguments

`x`	Return by `sesu_opgd()`.
`...`	(optional) Other arguments passed to `ggplot2::theme()`.

Value

A ggplot2 layer.

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

plot.spade_result *plot SPADE power of spatial and multilevel discretization determinant*

Description

S3 method to plot output for SPADE power of spatial and multilevel discretization determinant from `spade()`.

Usage

```
## S3 method for class 'spade_result'
plot(x, slicenum = 2, alpha = 0.95, keep = TRUE, ...)
```

Arguments

`x`	Return by `spade()`.The number of labels facing inward.
`slicenum`	(optional) The number of labels facing inward. Default is `2`.
`alpha`	(optional) Confidence level.Default is `0.95`.
`keep`	(optional) Whether to keep Q-value results for insignificant variables, default is `TRUE`.
`...`	(optional) Other arguments passed to `ggplot2::theme()`.

Value

A ggplot2 layer.

Author(s)

Wenbo Lv <`lyu.geosocial@gmail.com`>

plot.srsgd_result *plot SRSGD result*

Description

S3 method to plot output for SRSGD model result in `srsgd()`.

Usage

```
## S3 method for class 'srsgd_result'
plot(x, ...)
```

Arguments

`x`	Return by `srsgd()`.
`...`	(optional) Other arguments.

Value

A ggplot2 layer

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

---

plot.srs_ecological_detector

plot spatial rough set-based ecological detector

Description

S3 method to plot output for spatial rough set-based ecological detector in `srsgd()`.

Usage

```
## S3 method for class 'srs_ecological_detector'
plot(x, ...)
```

Arguments

`x`	Return by `srsgd()`.
`...`	(optional) Other arguments passed to `ggplot2::theme()`.

Value

A ggplot2 layer

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

---

plot.srs_factor_detector

plot spatial rough set-based factor detector result

Description

S3 method to plot output for spatial rough set-based factor detector in `srsgd()`.

Usage

```
## S3 method for class 'srs_factor_detector'
plot(x, slicenum = 2, ...)
```

Arguments

x	Return by srsgd().
slicenum	(optional) The number of labels facing inward. Default is 2.
...	(optional) Other arguments passed to ggplot2::theme().

Value

A ggplot2 layer.

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

---

plot.srs_interaction_detector

plot spatial rough set-based interaction detector result

Description

S3 method to plot output for spatial rough set-based interaction detector in srsgd().

Usage

```
## S3 method for class 'srs_interaction_detector'
plot(x, alpha = 1, ...)
```

Arguments

x	Return by srsgd().
alpha	(optional) Picture transparency. Default is 1.
...	(optional) Other arguments passed to ggplot2::theme().

Value

A ggplot2 layer

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

print.ecological_detector *print ecological detector*

Description

S3 method to format output for ecological detector in `geodetector()`.

Usage

```
## S3 method for class 'ecological_detector'
print(x, ...)
```

Arguments

`x`	Return by `geodetector()`.
`...`	(optional) Other arguments passed to `knitr::kable()`.

Value

Formatted string output

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

print.factor_detector *print factor detector*

Description

S3 method to format output for factor detector in `geodetector()`.

Usage

```
## S3 method for class 'factor_detector'
print(x, ...)
```

Arguments

`x`	Return by `geodetector()`.
`...`	(optional) Other arguments passed to `knitr::kable()`.

Value

Formatted string output

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

---

print.gd_result   *print GD result*

Description

S3 method to format output for GD model from `gd()`.

Usage

```
## S3 method for class 'gd_result'
print(x, ...)
```

Arguments

x	Return by `gd()`.
...	(optional) Other arguments passed to `knitr::kable()`.

Value

Formatted string output

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

---

print.gozh_result   *print GOZH result*

Description

S3 method to format output for GOZH model from `gozh()`.

Usage

```
## S3 method for class 'gozh_result'
print(x, ...)
```

Arguments

x	Return by `gozh()`.
...	(optional) Other arguments passed to `knitr::kable()`.

Value

Formatted string output

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

`print.idsa_result` *print IDSA result*

Description

S3 method to format output for IDSA model from `idsa()`.

Usage

```
## S3 method for class 'idsa_result'
print(x, ...)
```

Arguments

`x`	Return by `idsa()`.
`...`	(optional) Other arguments passed to `knitr::kable()`.

Value

Formatted string output

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

`print.interaction_detector` *print interaction detector*

Description

S3 method to format output for interaction detector in `geodetector()`.

Usage

```
## S3 method for class 'interaction_detector'
print(x, ...)
```

Arguments

x	Return by `geodetector()`.
...	(optional) Other arguments passed to `knitr::kable()`.

Value

Formatted string output

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

---

print.lesh_result — *print LESH model interaction result*

Description

S3 method to format output for LESH model interaction result in `lesh()`.

Usage

```
## S3 method for class 'lesh_result'
print(x, ...)
```

Arguments

x	Return by `lesh()`.
...	(optional) Other arguments passed to `knitr::kable()`.

Value

Formatted string output

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

print.opgd_result *print OPGD result*

Description

S3 method to format output for OPGD model from `opgd()`.

Usage

```
## S3 method for class 'opgd_result'
print(x, ...)
```

Arguments

`x`	Return by `opgd()`.
`...`	(optional) Other arguments passed to `knitr::kable()`.

Value

Formatted string output

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

print.rgd_result *print RGD result*

Description

S3 method to format output for RGD model from `rgd()`.

Usage

```
## S3 method for class 'rgd_result'
print(x, ...)
```

Arguments

`x`	Return by `rgd()`.
`...`	(optional) Other arguments passed to `knitr::kable()`.

Value

Formatted string output

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

`print.rid_result` *print RID result*

Description

S3 method to format output for RID model from `rid()`.

Usage

```
## S3 method for class 'rid_result'
print(x, ...)
```

Arguments

x	Return by `rid()`.
...	(optional) Other arguments passed to `knitr::kable()`.

Value

Formatted string output

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

`print.risk_detector` *print risk detector*

Description

S3 method to format output for risk detector in `geodetector()`.

Usage

```
## S3 method for class 'risk_detector'
print(x, ...)
```

Arguments

x	Return by `geodetector()`.
...	(optional) Other arguments passed to `knitr::kable()`.

Value

Formatted string output

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

---

print.sesu_gozh *print gozh sesu*

Description

S3 method to format output for gozh sesu from `sesu_gozh()`.

Usage

```
## S3 method for class 'sesu_gozh'
print(x, ...)
```

Arguments

x	Return by `sesu_gozh()`.
...	(optional) Other arguments passed to `knitr::kable()`.

Value

Formatted string output

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

---

print.sesu_opgd *print opgd sesu*

Description

S3 method to format output for opgd sesu from `sesu_opgd()`.

Usage

```
## S3 method for class 'sesu_opgd'
print(x, ...)
```

Arguments

`x`	Return by `sesu_opgd()`.
`...`	(optional) Other arguments passed to `knitr::kable()`.

Value

Formatted string output

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

`print.spade_result` *print SPADE power of spatial and multilevel discretization determinant*

Description

S3 method to format output for SPADE power of spatial and multilevel discretization determinant from `spade()`.

Usage

```
## S3 method for class 'spade_result'
print(x, ...)
```

Arguments

`x`	Return by `spade()`.
`...`	Other arguments.

Value

Formatted string output

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

print.srsgd_result *print SRSGD result*

Description

S3 method to format output for SRSGD model from `srsgd()`.

Usage

```
## S3 method for class 'srsgd_result'
print(x, ...)
```

Arguments

`x`	Return by `srsgd()`.
`...`	(optional) Other arguments passed to `knitr::kable()`.

Value

Formatted string output

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

print.srs_ecological_detector *print spatial rough set-based ecological detector*

Description

S3 method to format output for spatial rough set-based ecological detector in `srsgd()`.

Usage

```
## S3 method for class 'srs_ecological_detector'
print(x, ...)
```

Arguments

`x`	Return by `srsgd()`.
`...`	(optional) Other arguments passed to `knitr::kable()`.

Value

Formatted string output

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

---

print.srs_factor_detector

print spatial rough set-based factor detector

Description

S3 method to format output for spatial rough set-based factor detector in `srsgd()`.

Usage

```
## S3 method for class 'srs_factor_detector'
print(x, ...)
```

Arguments

x	Return by `srsgd()`.
...	(optional) Other arguments passed to `knitr::kable()`.

Value

Formatted string output

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

---

print.srs_interaction_detector

print spatial rough set-based interaction detector

Description

S3 method to format output for spatial rough set-based interaction detector in `srsgd()`.

Usage

```
## S3 method for class 'srs_interaction_detector'
print(x, ...)
```

Arguments

x	Return by `srsgd()`.
...	(optional) Other arguments passed to `knitr::kable()`.

Value

Formatted string output

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

`psd_iev` *PSD of an interaction of explanatory variables (PSD-IEV)*

Description

PSD of an interaction of explanatory variables (PSD-IEV)

Usage

```
psd_iev(discdata, spzone, wt)
```

Arguments

`discdata`	Observed data with discrete explanatory variables. A `tibble` or `data.frame` .
`spzone`	Fuzzy overlay spatial zones. Returned from `st_fuzzyoverlay()`.
`wt`	Spatial weight matrix

Details

$$\phi = 1 - \frac{\sum_{i=1}^{m} \sum_{k=1}^{n_i} N_{i,k} \tau_{i,k}}{\sum_{i=1}^{m} N_i \tau_i}$$

Value

The Value of `PSD-IEV`

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

References

Yongze Song & Peng Wu (2021) An interactive detector for spatial associations, International Journal of Geographical Information Science, 35:8, 1676-1701, DOI:10.1080/13658816.2021.1882680

Examples

```
data('sim')
wt = sdsfun::inverse_distance_swm(sf::st_as_sf(sim,coords = c('lo','la')))
sim1 = dplyr::mutate(sim,dplyr::across(xa:xc,\(.x) sdsfun::discretize_vector(.x,5)))
sz = sdsfun::fuzzyoverlay(y ~ xa + xb + xc, data = sim1)
psd_iev(dplyr::select(sim1,xa:xc),sz,wt)
```

psd_pseudop *calculate power of spatial determinant(PSD) and the corresponding pseudo-p value*

Description

Function for calculate power of spatial determinant q_s.

Usage

```
psd_pseudop(y, x, wt, cores = 1, seed = 123456789, permutations = 0)
```

Arguments

y	Variable Y, continuous numeric vector.
x	Covariable X, factor, character or discrete numeric.
wt	The spatial weight matrix.
cores	(optional) A positive integer(default is 1). If cores > 1, use parallel computation.
seed	(optional) Random seed number, default is 123456789.
permutations	(optional) The number of permutations for the PSD computation. Default is 0, which means no pseudo-p values are calculated.

Details

The power of spatial determinant formula is $q_s = 1 - \frac{\sum_{h=1}^{L} N_h \Gamma_h}{N \Gamma}$

Value

A tibble of power of spatial determinant and the corresponding pseudo-p value.

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

References

Xuezhi Cang & Wei Luo (2018) Spatial association detector (SPADE),International Journal of Geographical Information Science, 32:10, 2055-2075, DOI: 10.1080/13658816.2018.1476693

Examples

```
data('sim')
wt = sdsfun::inverse_distance_swm(sf::st_as_sf(sim,coords = c('lo','la')),
                                  power = 2)
psd_pseudop(sim$y,sdsfun::discretize_vector(sim$xa,5),wt)
```

psd_spade *power of spatial determinant(PSD)*

Description

Function for calculate power of spatial determinant q_s

Usage

```
psd_spade(y, x, wt)
```

Arguments

y	Variable Y, continuous numeric vector.
x	Covariable X, factor, character or discrete numeric.
wt	The spatial weight matrix.

Details

The power of spatial determinant formula is

$q_s = 1 - \frac{\sum_{h=1}^{L} N_h \Gamma_h}{N \Gamma}$

Value

A value of power of spatial determinant q_s.

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

References

Xuezhi Cang & Wei Luo (2018) Spatial association detector (SPADE),International Journal of Geographical Information Science, 32:10, 2055-2075, DOI: 10.1080/13658816.2018.1476693

Examples

```
data('sim')
wt = sdsfun::inverse_distance_swm(sf::st_as_sf(sim,coords = c('lo','la')),
                                  power = 2)
psd_spade(sim$y,sdsfun::discretize_vector(sim$xa,5),wt)
```

psmd_pseudop *power of spatial and multilevel discretization determinant(PSMD) and the corresponding pseudo-p value*

Description

Function for calculate power of spatial and multilevel discretization determinant and the corresponding pseudo-p value.

Usage

```
psmd_pseudop(
  yobs,
  xobs,
  wt,
  discnum = 3:8,
  discmethod = "quantile",
  cores = 1,
  seed = 123456789,
  permutations = 0,
  ...
)
```

Arguments

yobs	Variable Y
xobs	The original undiscretized covariable X.
wt	The spatial weight matrix.
discnum	(optional) Number of multilevel discretization. Default will use `3:8`.
discmethod	(optional) The discretization methods. Default will use `quantile`. If `discmethod` is set to `robust`, the function `robust_disc()` will be used. Conversely, if `discmethod` is set to `rpart`, the `rpart_disc()` function will be used. Others use `sdsfun::discretize_vector()`. Currently, only one `discmethod` can be used at a time.
cores	(optional) A positive integer(default is 1). If cores > 1, use parallel computation.
seed	(optional) Random seed number, default is `123456789`.
permutations	(optional) The number of permutations for the PSD computation. Default is `0`, which means no pseudo-p values are calculated.
...	(optional) Other arguments passed to `sdsfun::discretize_vector()`,`robust_disc()` or `rpart_disc()`.

Details

The power of spatial and multilevel discretization determinant formula is $PSMDQ_s = MEAN(Q_s)$

Value

A tibble of power of spatial and multilevel discretization determinant and the corresponding pseudo-p value.

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

References

Xuezhi Cang & Wei Luo (2018) Spatial association detector (SPADE),International Journal of Geographical Information Science, 32:10, 2055-2075, DOI: 10.1080/13658816.2018.1476693

Examples

```
data('sim')
wt = sdsfun::inverse_distance_swm(sf::st_as_sf(sim,coords = c('lo','la')))
psmd_pseudop(sim$y,sim$xa,wt)
```

---

psmd_spade *power of spatial and multilevel discretization determinant(PSMD)*

Description

Function for calculate power of spatial and multilevel discretization determinant PSMDQ_s.

Usage

```
psmd_spade(
  yobs,
  xobs,
  wt,
  discnum = 3:8,
  discmethod = "quantile",
  cores = 1,
  seed = 123456789,
  ...
)
```

Arguments

yobs	Variable Y
xobs	The original continuous covariable X.
wt	The spatial weight matrix.
discnum	(optional) Number of multilevel discretizations. Default will use 3:8.

discmethod	(optional) The discretize methods. Default will use quantile. If discmethod is set to robust, the function robust_disc() will be used. Conversely, if discmethod is set to rpart, the rpart_disc() function will be used. Others use sdsfun::discretize_vector(). Currently, only one discmethod can be used at a time.
cores	(optional) A positive integer(default is 1). If cores > 1, use parallel computation.
seed	(optional) Random seed number, default is 123456789.
...	(optional) Other arguments passed to sdsfun::discretize_vector(),robust_disc() or rpart_disc().

Details

The power of spatial and multilevel discretization determinant formula is $PSMDQ_s = MEAN(Q_s)$

Value

A value of power of spatial and multilevel discretization determinant PSMDQ_s.

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

References

Xuezhi Cang & Wei Luo (2018) Spatial association detector (SPADE),International Journal of Geographical Information Science, 32:10, 2055-2075, DOI: 10.1080/13658816.2018.1476693

Examples

```
data('sim')
wt = sdsfun::inverse_distance_swm(sf::st_as_sf(sim,coords = c('lo','la')))
psmd_spade(sim$y,sim$xa,wt)
```

---

rgd *robust geographical detector(RGD) model*

Description

robust geographical detector(RGD) model

Usage

```
rgd(
  formula,
  data,
  discvar = NULL,
  discnum = 3:8,
  minsize = 1,
  strategy = 2L,
  increase_rate = 0.05,
  cores = 1
)
```

Arguments

`formula`	A formula of RGD model.
`data`	A `data.frame`, `tibble` or `sf` object of observation data.
`discvar`	Name of continuous variable columns that need to be discretized. Noted that when `formula` has `discvar`, `data` must have these columns. By default, all independent variables are used as `discvar`.
`discnum`	A numeric vector of discretized classes of columns that need to be discretized. Default all `discvar` use `3:8`.
`minsize`	(optional) The min size of each discretization group. Default all use `1`.
`strategy`	(optional) Optimal discretization strategy. When `strategy` is `1L`, choose the highest q-statistics to determinate optimal spatial data discretization parameters. When `strategy` is `2L`, The optimal discrete parameters of spatial data are selected by combining LOESS model.
`increase_rate`	(optional) The critical increase rate of the number of discretization. Default is `5%`.
`cores`	(optional) Positive integer (default is 1). When cores are greater than 1, use multi-core parallel computing.

Value

A list.

`factor` robust power of determinant

`opt_disc` optimal robust discrete results

`allfactor` factor detection results corresponding to different number of robust discreteizations

`alldisc` all robust discrete results

Note

Please set up python dependence and configure `GDVERSE_PYTHON` environment variable if you want to run `rgd()`. See `vignette('rgdrid',package = 'gdverse')` for more details.

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

References

Zhang, Z., Song, Y.*, & Wu, P., 2022. Robust geographical detector. International Journal of Applied Earth Observation and Geoinformation. 109, 102782. DOI: 10.1016/j.jag.2022.102782.

Examples

```
## Not run:
## The following code needs to configure the Python environment to run:
data('sim')
g = rgd(y ~ .,
        data = dplyr::select(sim,-dplyr::any_of(c('lo','la'))),
        discnum = 3:6, cores = 1)
g

## End(Not run)
```

---

rid *robust interaction detector(RID) model*

Description

robust interaction detector(RID) model

Usage

```
rid(
  formula,
  data,
  discvar = NULL,
  discnum = 3:8,
  minsize = 1,
  strategy = 2L,
  increase_rate = 0.05,
  cores = 1
)
```

Arguments

`formula`	A formula of RGD model.
`data`	A `data.frame`, `tibble` or `sf` object of observation data.
`discvar`	Name of continuous variable columns that need to be discretized. Noted that when `formula` has `discvar`, `data` must have these columns. By default, all independent variables are used as `discvar`.

discnum	A numeric vector of discretized classes of columns that need to be discretized. Default all discvar use 3:8.
minsize	(optional) The min size of each discretization group. Default all use 1.
strategy	(optional) Optimal discretization strategy. When strategy is 1L, choose the highest q-statistics to determinate optimal spatial data discretization parameters. When strategy is 2L, The optimal discrete parameters of spatial data are selected by combining LOESS model.
increase_rate	(optional) The critical increase rate of the number of discretization. Default is 5%.
cores	(optional) Positive integer (default is 1). When cores are greater than 1, use multi-core parallel computing.

Value

A list.

interaction the result of RID model

Note

Please set up python dependence and configure GDVERSE_PYTHON environment variable if you want to run rid(). See vignette('rgdrid',package = 'gdverse') for more details.

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

References

Zhang, Z., Song, Y., Karunaratne, L., & Wu, P. (2024). Robust interaction detector: A case of road life expectancy analysis. Spatial Statistics, 59(100814), 100814. https://doi.org/10.1016/j.spasta.2024.100814

Examples

```
## Not run:
## The following code needs to configure the Python environment to run:
data('sim')
g = rid(y ~ .,
        data = dplyr::select(sim,-dplyr::any_of(c('lo','la'))),
        discnum = 3:6, cores = 1)
g

## End(Not run)
```

risk_detector *risk detector*

Description

Determine whether there is a significant difference between the attribute means of two sub regions.

Usage

```
risk_detector(y, x, alpha = 0.95)
```

Arguments

y	Variable Y, continuous numeric vector.
x	Covariate X, `factor`, `character` or `discrete numeric`.
alpha	(optional) Confidence level of the interval, default is `0.95`.

Value

A tibble.

Author(s)

Wenbo Lv <`lyu.geosocial@gmail.com`>

Examples

```
risk_detector(y = 1:7,
              x = c('x',rep('y',3),rep('z',3)))
```

robust_disc *univariate discretization based on offline change point detection*

Description

Determines discretization interval breaks using an optimization algorithm for variance-based change point detection.

Usage

```
robust_disc(formula, data, discnum, minsize = 1, cores = 1)
```

Arguments

formula	A formula of univariate discretization.
data	A data.frame or tibble of observation data.
discnum	A numeric vector of discretized classes of columns that need to be discretized.
minsize	(optional) The min size of each discretization group. Default all use 1.
cores	(optional) A positive integer(default is 1). If cores > 1, use python joblib package to parallel computation.

Value

A tibble.

Note

Please set up python dependence and configure GDVERSE_PYTHON environment variable if you want to run robust_disc(). See vignette('rgdrid',package = 'gdverse') for more details.

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

Examples

```
## Not run:
## The following code needs to configure the Python environment to run:
data('sim')
robust_disc(y ~ xa, data = sim, discnum = 5)
robust_disc(y ~ .,
            data = dplyr::select(sim,-dplyr::any_of(c('lo','la'))),
            discnum = 5, cores = 3)

## End(Not run)
```

---

rpart_disc *discretization of variables based on recursive partitioning*

Description

discretization of variables based on recursive partitioning

Usage

```
rpart_disc(formula, data, ...)
```

Arguments

formula	A formula.
data	A data.frame or tibble of observation data.
...	(optional) Other arguments passed to rpart::rpart().

Value

A vector that being discretized.

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

References

Luo, P., Song, Y., Huang, X., Ma, H., Liu, J., Yao, Y., & Meng, L. (2022). Identifying determinants of spatio-temporal disparities in soil moisture of the Northern Hemisphere using a geographically optimal zones-based heterogeneity model. ISPRS Journal of Photogrammetry and Remote Sensing: Official Publication of the International Society for Photogrammetry and Remote Sensing (ISPRS), 185, 111–128. https://doi.org/10.1016/j.isprsjprs.2022.01.009

Examples

```
data('ndvi')
rpart_disc(NDVIchange ~ ., data = ndvi)
```

---

sesu_gozh — *comparison of size effects of spatial units based on GOZH*

Description

comparison of size effects of spatial units based on GOZH

Usage

```
sesu_gozh(
  formula,
  datalist,
  su,
  cores = 1,
  strategy = 2L,
  increase_rate = 0.05,
  alpha = 0.95,
  ...
)
```

Arguments

`formula`	A formula of comparison of size effects of spatial units.
`datalist`	A list of `data.frame` or `tibble`.
`su`	A vector of sizes of spatial units.
`cores`	(optional) Positive integer (default is 1). When cores are greater than 1, use multi-core parallel computing.
`strategy`	(optional) Calculation strategies of Q statistics at different scales. Default is `2L`, see `details` for more contents.
`increase_rate`	(optional) The critical increase rate of the number of discretization. Default is `5%`.
`alpha`	(optional) Specifies the size of confidence level. Default is `0.95`.
`...`	(optional) Other arguments passed to `rpart_disc()`.

Details

When `strategy` is `1`, use the same process as `sesu_opgd()`.If not, all explanatory variables are used to generate a unique Q statistic corresponding to the data in the datalist based on `rpart_disc()` and `gd()`, and then `loess_optscale()`is used to determine the optimal analysis scale.

Value

A list.

`sesu` a tibble representing size effects of spatial units

`optsu` optimal spatial unit

`strategy` the optimal analytical scale selection strategy

`increase_rate` the critical increase rate of q value

Author(s)

Wenbo Lv `<lyu.geosocial@gmail.com>`

References

Song, Y., Wang, J., Ge, Y. & Xu, C. (2020) An optimal parameters-based geographical detector model enhances geographic characteristics of explanatory variables for spatial heterogeneity analysis: Cases with different types of spatial data, GIScience & Remote Sensing, 57(5), 593-610. doi: 10.1080/15481603.2020.1760434.

Luo, P., Song, Y., Huang, X., Ma, H., Liu, J., Yao, Y., & Meng, L. (2022). Identifying determinants of spatio-temporal disparities in soil moisture of the Northern Hemisphere using a geographically optimal zones-based heterogeneity model. ISPRS Journal of Photogrammetry and Remote Sensing: Official Publication of the International Society for Photogrammetry and Remote Sensing (ISPRS), 185, 111–128. https://doi.org/10.1016/j.isprsjprs.2022.01.009

Examples

```
## Not run:
## The following code takes a long time to run:
library(tidyverse)
fvcpath = "https://github.com/SpatLyu/rdevdata/raw/main/FVC.tif"
fvc = terra::rast(paste0("/vsicurl/",fvcpath))
fvc1000 = fvc %>%
  terra::as.data.frame(na.rm = T) %>%
  as_tibble()
fvc5000 = fvc %>%
  terra::aggregate(fact = 5) %>%
  terra::as.data.frame(na.rm = T) %>%
  as_tibble()
sesu_gozh(fvc ~ .,
          datalist = list(fvc1000,fvc5000),
          su = c(1000,5000),
          cores = 6)

## End(Not run)
```

---

sesu_opgd *comparison of size effects of spatial units based on OPGD*

Description

comparison of size effects of spatial units based on OPGD

Usage

```
sesu_opgd(
  formula,
  datalist,
  su,
  discvar,
  discnum = 3:8,
  discmethod = c("sd", "equal", "geometric", "quantile", "natural"),
  cores = 1,
  increase_rate = 0.05,
  alpha = 0.95,
  ...
)
```

Arguments

formula	A formula of comparison of size effects of spatial units.
datalist	A list of data.frame or tibble.
su	A vector of sizes of spatial units.

discvar	Name of continuous variable columns that need to be discretized.Noted that when formula has discvar, data must have these columns.
discnum	(optional) A vector of number of classes for discretization. Default is 3:8.
discmethod	(optional) A vector of methods for discretization, default is using c("sd","equal","geometric","quan by invoking sdsfun.
cores	(optional) Positive integer (default is 1). When cores are greater than 1, use multi-core parallel computing.
increase_rate	(optional) The critical increase rate of the number of discretization. Default is 5%.
alpha	(optional) Specifies the size of confidence level. Default is 0.95.
...	(optional) Other arguments passed to gd_bestunidisc().

Details

Firstly, the OPGD model is executed for each data in the datalist (all significant Q statistic of each data are averaged to represent the spatial association strength under this spatial unit), and then the loess_optscale function is used to select the optimal spatial analysis scale.

Value

A list.

sesu a tibble representing size effects of spatial units

optsu optimal spatial unit

increase_rate the critical increase rate of q value

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

References

Song, Y., Wang, J., Ge, Y. & Xu, C. (2020) An optimal parameters-based geographical detector model enhances geographic characteristics of explanatory variables for spatial heterogeneity analysis: Cases with different types of spatial data, GIScience & Remote Sensing, 57(5), 593-610. doi: 10.1080/15481603.2020.1760434.

Examples

```
## Not run:
## The following code takes a long time to run:
library(tidyverse)
fvcpath = "https://github.com/SpatLyu/rdevdata/raw/main/FVC.tif"
fvc = terra::rast(paste0("/vsicurl/",fvcpath))
fvc1000 = fvc %>%
  terra::as.data.frame(na.rm = T) %>%
  as_tibble()
fvc5000 = fvc %>%
```

```
  terra::aggregate(fact = 5) %>%
  terra::as.data.frame(na.rm = T) %>%
  as_tibble()
sesu_opgd(fvc ~ .,
          datalist = list(fvc1000,fvc5000),
          su = c(1000,5000),
          discvar = names(select(fvc5000,-c(fvc,lulc))),
          cores = 6)

## End(Not run)
```

---

sim *Simulation data.*

Description

Simulation data.

Usage

```
sim
```

Format

`sim`: A tibble with 80 rows and 6 variables, modified from `IDSA` package.

Author(s)

Yongze Song <yongze.song@outlook.com>

---

spade *spatial association detector (SPADE) model*

Description

spatial association detector (SPADE) model

Usage

```
spade(
  formula,
  data,
  wt = NULL,
  discvar = NULL,
  discnum = 3:8,
  discmethod = "quantile",
  cores = 1,
```

```
  seed = 123456789,
  permutations = 0,
  ...
)
```

Arguments

formula	A formula of spatial association detector (SPADE) model.
data	A data.frame, tibble or sf object of observation data.
wt	(optional) The spatial weight matrix. When data is not an sf object, must provide wt.
discvar	(optional) Name of continuous variable columns that need to be discretized. Noted that when formula has discvar, data must have these columns. By default, all independent variables are used as discvar.
discnum	(optional) Number of multilevel discretization. Default will use 3:8.
discmethod	(optional) The discretization methods. Default all use quantile. Note that when using different discmethod for discvar, please ensure that the lengths of both are consistent. Noted that robust will use robust_disc(); rpart will use rpart_disc(); Others use sdsfun::discretize_vector().
cores	(optional) Positive integer (default is 1). When cores are greater than 1, use multi-core parallel computing.
seed	(optional) Random number seed, default is 123456789.
permutations	(optional) The number of permutations for the PSD computation. Default is 0, which means no pseudo-p values are calculated.
...	(optional) Other arguments passed to sdsfun::discretize_vector(),robust_disc() or rpart_disc().

Value

A list.

factor the result of SPADE model

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

References

Xuezhi Cang & Wei Luo (2018) Spatial association detector (SPADE),International Journal of Geographical Information Science, 32:10, 2055-2075, DOI: 10.1080/13658816.2018.1476693

Examples

```
data('sim')
sim1 = sf::st_as_sf(sim,coords = c('lo','la'))
g = spade(y ~ ., data = sim1)
g
```

spd_lesh *shap power of determinants*

Description

Function for calculate shap power of determinants SPD.

Usage

```
spd_lesh(formula, data, cores = 1, ...)
```

Arguments

`formula`	A formula of calculate shap power of determinants.
`data`	A data.frame or tibble of observation data.
`cores`	(optional) Positive integer (default is 1). When cores are greater than 1, use multi-core parallel computing.
`...`	(optional) Other arguments passed to `rpart_disc()`.

Details

The power of shap power of determinants formula is

$\theta_{x_j}(S) = \sum\limits_{s \in M \setminus \{x_j\}} \frac{|S|!(|M|-|S|-1)!}{|M|!}(v(S \cup \{x_j\}) - v(S))$.

shap power of determinants (SPD) is the contribution of variable x_j to the power of determinants.

Value

A tibble with variable and its corresponding SPD value.

Note

The shap power of determinants (SPD) requires at least $2^n - 1$ calculations when has n explanatory variables. When there are more than 10 explanatory variables, carefully consider the computational burden of this model. When there are a large number of explanatory variables, the data dimensionality reduction method can be used to ensure the trade-off between analysis results and calculation speed.

Author(s)

Wenbo Lv <`lyu.geosocial@gmail.com`>

References

Li, Y., Luo, P., Song, Y., Zhang, L., Qu, Y., & Hou, Z. (2023). A locally explained heterogeneity model for examining wetland disparity. International Journal of Digital Earth, 16(2), 4533–4552. https://doi.org/10.1080/17538947.2023.2271883

Examples

```
data('ndvi')
g = spd_lesh(NDVIchange ~ ., data = ndvi)
g
```

---

srsgd *spatial rough set-based geographical detector(SRSGD) model*

Description

spatial rough set-based geographical detector(SRSGD) model

Usage

```
srsgd(formula, data, wt = NULL, type = "factor", alpha = 0.95)
```

Arguments

`formula`	A formula of spatial rough set-based geographical detector model.
`data`	A `data.frame`, `tibble` or `sf` object of observation data.
`wt`	Spatial adjacency matrix. If `data` is a `sf` polygon object, the queen adjacency matrix is used when no `wt` object is provided. In other cases, you must provide a `wt` object.
`type`	(optional) The type of geographical detector, which must be one of `factor`(default), `interaction` and `ecological`.
`alpha`	(optional) Specifies the size of the alpha (confidence level). Default is `0.95`.

Value

A list.

`factor` the result of spatial rough set-based factor detector

`interaction` the result of spatial rough set-based interaction detector

`ecological` the result of spatial rough set-based ecological detector

Note

The Spatial Rough Set-based Geographical Detector Model (SRSGD) conducts spatial hierarchical heterogeneity analysis utilizing a geographical detector for data where *the dependent variable* is *discrete*. Given the complementary relationship between SRSGD and the native version of geographical detector, I strive to maintain consistency with `gd()` function when establishing `srsgd()` function. This implies that all input variable data in srsgd must *be discretized prior to use*.

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

References

Bai, H., Li, D., Ge, Y., Wang, J., & Cao, F. (2022). Spatial rough set-based geographical detectors for nominal target variables. Information Sciences, 586, 525–539. https://doi.org/10.1016/j.ins.2021.12.019

Examples

```
data('srs_table')
data('srs_wt')
srsgd(d ~ a1 + a2 + a3, data = srs_table, wt = srs_wt,
      type = c('factor','interaction','ecological'))
```

---

srs_ecological_detector

spatial rough set-based ecological detector

Description

spatial rough set-based ecological detector

Usage

```
srs_ecological_detector(y, x1, x2, wt, alpha = 0.95)
```

Arguments

y	Dependent variable, `factor`, `character` or `discrete` `numeric`.
x1	Covariate X_1, `factor`, `character` or `discrete` `numeric`.
x2	Covariate X_2, `factor`, `character` or `discrete` `numeric`.
wt	Spatial adjacency matrix.
alpha	(optional) Confidence level of the interval,default is `0.95`.

Value

A list.

`T-statistic` the result of T statistic for spatial rough set-based ecological detector

`P-value` the result of P value for spatial rough set-based ecological detector

`Ecological` does one spatial feature X_1 play a more important role than X_2

Author(s)

Wenbo Lv `<lyu.geosocial@gmail.com>`

References

Bai, H., Li, D., Ge, Y., Wang, J., & Cao, F. (2022). Spatial rough set-based geographical detectors for nominal target variables. Information Sciences, 586, 525–539. https://doi.org/10.1016/j.ins.2021.12.019

Examples

```
data('srs_table')
data('srs_wt')
srs_ecological_detector(srs_table$d,srs_table$a1,srs_table$a2,srs_wt)
```

---

srs_factor_detector   *spatial rough set-based factor detector*

Description

spatial rough set-based factor detector

Usage

```
srs_factor_detector(y, x, wt)
```

Arguments

y	Variable Y, `factor`, `character` or `discrete` `numeric`.
x	Covariate X, `factor`, `character` or `discrete` `numeric`.
wt	Spatial adjacency matrix.

Value

A list.

`PD` the average local explanatory power

`SE_PD` the degree of spatial heterogeneity of the local explanatory power

Author(s)

Wenbo Lv <`lyu.geosocial@gmail.com`>

References

Bai, H., Li, D., Ge, Y., Wang, J., & Cao, F. (2022). Spatial rough set-based geographical detectors for nominal target variables. Information Sciences, 586, 525–539. https://doi.org/10.1016/j.ins.2021.12.019

Examples

```
data('srs_table')
data('srs_wt')
srs_factor_detector(srs_table$d,srs_table$a1,srs_wt)
```

srs_geodetector *spatial rough set-based geographical detector*

Description

spatial rough set-based geographical detector

Usage

```
srs_geodetector(formula, data, wt = NULL, type = "factor", alpha = 0.95)
```

Arguments

`formula`	A formula of spatial rough set-based geographical detector model.
`data`	A data.frame, tibble or sf object of observation data.
`wt`	Spatial adjacency matrix. If `data` is a `sf` polygon object, the queen adjacency matrix is used when no `wt` object is provided. In other cases, you must provide a `wt` object.
`type`	(optional) The type of geographical detector, which must be one of `factor`(default), `interaction` and `ecological`.
`alpha`	(optional) Specifies the size of the alpha (confidence level). Default is `0.95`.

Value

A list.

`factor` the result of spatial rough set-based factor detector

`interaction` the result of spatial rough set-based interaction detector

`ecological` the result of spatial rough set-based ecological detector

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

Examples

```
data('srs_table')
data('srs_wt')
srs_geodetector(d ~ a1 + a2 + a3, data = srs_table, wt = srs_wt)
srs_geodetector(d ~ a1 + a2 + a3, data = srs_table,
                wt = srs_wt, type = 'interaction')
srs_geodetector(d ~ a1 + a2 + a3, data = srs_table,
                wt = srs_wt, type = 'ecological')
```

`srs_interaction_detector` *spatial rough set-based interaction detector*

Description

spatial rough set-based interaction detector

Usage

```
srs_interaction_detector(y, x1, x2, wt)
```

Arguments

`y`	Dependent variable, `factor`, `character` or `discrete` `numeric`.
`x1`	Covariate X_1, `factor`, `character` or `discrete` `numeric`.
`x2`	Covariate X_2, `factor`, `character` or `discrete` `numeric`.
`wt`	Spatial adjacency matrix.

Value

A list.

`Variable1 PD` the average local explanatory power for variable1

`Variable2 PD` the average local explanatory power for variable2

`Variable1 and Variable2 interact PD` the average local explanatory power for variable1 and variable2 interact

`Variable1 SE_PD` the degree of spatial heterogeneity of the local explanatory power for variable1

`Variable2 SE_PD` the degree of spatial heterogeneity of the local explanatory power for variable2

`Variable1 and Variable2 SE_PD` the degree of spatial heterogeneity of the local explanatory power for variable1 and variable2 interact

`Interaction` the interact result type

Author(s)

Wenbo Lv <lyu.geosocial@gmail.com>

References

Bai, H., Li, D., Ge, Y., Wang, J., & Cao, F. (2022). Spatial rough set-based geographical detectors for nominal target variables. Information Sciences, 586, 525–539. https://doi.org/10.1016/j.ins.2021.12.019

Examples

```
data('srs_table')
data('srs_wt')
srs_interaction_detector(srs_table$d,srs_table$a1,srs_table$a2,srs_wt)
```

`srs_table` *example of spatial information system table*

Description

example of spatial information system table

Usage

```
srs_table
```

Format

`srs_table`: A tibble with 11 rows and 5 variables(one `ID` column).

`srs_wt` *example of spatial information system spatial adjacency matrix*

Description

example of spatial information system spatial adjacency matrix

Usage

```
srs_wt
```

Format

`srs_wt`: A matrix with 11rows and 11cols.

`weight_assign` *assign values by weight*

Description

assign values by weight

Usage

```
weight_assign(x, w, list = FALSE)
```

Arguments

`x`	A numeric value
`w`	A weight vector
`list`	(optional) Return list or not. if `list` is `TRUE`, return a list, otherwise return a vector. Default is `FALSE`.

Value

A numeric Vector.

Examples

```
weight_assign(0.875,1:3)
```

Index