Package ‘REAT’

June 18, 2016

Type Package

Title Regional Economic Analysis Tools

Version 1.1.0

Date 2016-06-17

Author Thomas Wieland

Maintainer Thomas Wieland <thomas.wieland.geo@googlemail.com>

Description Collection of analysis methods used in regional and urban economics and (quantitative) economic geography, e.g. measures of inequality, regional disparities, regional specialization and spatial interaction models.

License GPL (>= 2)

NeedsCompilation no

Repository CRAN

Date/Publication 2016-06-18 16:10:04

R topics documented:

`REAT-package` *Regional Economic Analysis Tools*

Description

This package contains a collection of models and analysis methods used in regional and urban economics and (quantitative) economic geography. The functions in this package can be divided in six groups: (1) analysis of regional disparities and inequality (including *Gini coefficient* and *Lorenz curve*), (2) specialization of regions (including *spatial Gini coefficient* and *Krugman coefficient for regional specialization*), (3) spatial concentration of industries (including *spatial Gini coefficient for industry concentration*), (4) regional growth (including traditional *shift-share analysis*), spatial interaction models (including *Huff model* and *Hansen accessibility*) and (6) additional tools for data preparation und visualization. The package also contains data examples.

Author(s)

Thomas Wieland

Maintainer: Thomas Wieland <thomas.wieland.geo@googlemail.com>

References

Dinc, M. (2002): "Regional and Local Economic Analysis Tools". Prepared for the Public Finance, Decentralization and Poverty Reduction Program (World Bank Institute). Washington, D.C. : The World Bank.

Doersam, P. (2004): "Wirtschaftsstatistik anschaulich dargestellt". Heidenau : PD-Verlag.

Farhauer, O./Kroell, A. (2013): "Standorttheorien: Regional- und Stadtoekonomik in Theorie und Praxis". Wiesbaden : Springer.

Guessefeldt, J. (1999): "Regionalanalyse". Muenchen : Oldenbourg.

Lessmann, C. (2005): "Regionale Disparitaeten in Deutschland und ausgesuchten OECD-Staaten im Vergleich". *ifo Dresden berichtet*, **3/2005**.

Loeffler, G. (1998): "Market areas - a methodological reflection on their boundaries". In: *GeoJournal*, **45**, 4, p. 265-272

`converse` *Breaking point formula by Converse*

Description

Calculating the breaking point between two cities or retail locations

Usage

```
converse(P_a, P_b, D_ab)
```

Arguments

`P_a`	a single numeric value of attractivity/population size of location/city a
`P_b`	a single numeric value of attractivity/population size of location/city b
`D_ab`	a single numeric value of the transport costs (e.g. distance) between a and b

Details

The *breaking point formula* by Converse (1949) is a modification of the *law of retail gravitation* by Reilly (1929, 1931) (see the functions `reilly` and `reilly.lambda`). The aim of the calculation is to determine the boundaries of the market areas between two locations/cities in consideration of their attractivity/population size and the transport costs (e.g. distance) between them. The models by Reilly and Converse are simple *spatial interaction models* and are considered as *deterministic market area models* due to their exact allocation of demand origins to locations. A probabilistic approach including a theoretical framework was developed by Huff (1962) (see the function `huff`).

Value

a list with two values (`B_a`: distance from location a to breaking point, `B_b`: distance from location b to breaking point)

Author(s)

Thomas Wieland

References

Converse, P. D. (1949): "New Laws of Retail Gravitation". In: *Journal of Marketing*, **14**, 3, p. 379-384.

Huff, D. L. (1962): "Determination of Intra-Urban Retail Trade Areas". Los Angeles : University of California.

Loeffler, G. (1998): "Market areas - a methodological reflection on their boundaries". In: *GeoJournal*, **45**, 4, p. 265-272

Reilly, W. J. (1929): "Methods for the Study of Retail Relationships". *Studies in Marketing*, **4**. Austin : Bureau of Business Research, The University of Texas.

Reilly, W. J. (1931): "The Law of Retail Gravitation". New York.

See Also

huff, reilly, reilly.lambda

Examples

```
# Example from Huff (1962):
converse (400000, 200000, 80)
# two cities (population 400.000 and 200.000 with a distance separating them of 80 miles)
```

---

cv *Coefficient of variation*

Description

Calculating the coefficient of variation (cv), standardized and non-standardized

Usage

```
cv(x, norm = FALSE)
```

Arguments

x	a numeric vector
norm	logical argument that indicates if the function output is the standardized cv or not (default: norm = FALSE)

Details

The coefficient of variation v is a simple standardized measure of distribution and is used especially in the analysis of regional economic disparities (e.g. disparities in regional GDP per capita) or disparities in supply (e.g. density of physicians or grocery stores). The function returns the non-standardized cv ($0 < v < \infty$) or the standardized cv ($0 < v* < 1$).

Value

Single numeric value. If norm = FALSE the function returns the non-standardized cv ($0 < v < \infty$). If norm = TRUE the standardized cv ($0 < v* < 1$) is returned.

Author(s)

Thomas Wieland

References

Lessmann, C. (2005): "Regionale Disparitaeten in Deutschland und ausgesuchten OECD-Staaten im Vergleich". *ifo Dresden berichtet*, **3/2005**. https://www.cesifo-group.de/link/ifodb_2005_3_25-33.pdf.

Examples

```
data(Goettingen)
# Loads the data
# (Grocery stores in Goettingen, Germany, at the level of 69 statistical districts:
# number and sales area of stores, absolute values and per 1.000 inhabitants)
cv(Goettingen$stores_p1000p)
# cv of stores per 1.000 inhabitants
cv(Goettingen$salesarea_p1000p)
# cv of sales area per 1.000 inhabitants
cv(Goettingen$salesarea_p1000p, norm=TRUE)
# standardized cv of sales area per 1.000 inhabitants
```

---

data.index *Creating index values*

Description

Standardizing a variable as an index

Usage

```
data.index(dataset, col_index, col_ref, value_ref)
```

Arguments

dataset	regarded data.frame
col_index	column to be converted to index values
col_ref	column with the reference values of the index (e.g. years or months)
value_ref	value from the reference column which is the reference value (=100)

Value

a numeric vector consisting of the indexed values of col_index

Author(s)

Thomas Wieland

Examples

```
# Creating test data
year <- 2010:2015
values <- c(20,24,21,28,27,29)
timeseries <- data.frame(year, values)
data.index(timeseries, "values", "year", "2012")
# returns index values
valuesindex <- data.index(timeseries, "values", "year", "2012")
timeseries2 <- data.frame(timeseries, valuesindex)
# add index values to data
```

`disp` *Coefficients for disparities*

Description

Calculating the Gini coefficient (non-standardized and standardized), the Herfindahl-Hirschman coefficient (non-standardized and standardized) and the Herfindahl-Hirschman equivalent number

Usage

```
disp(x)
```

Arguments

`x`	a `numeric vector` containing the regarded objects

Details

The *Gini coefficient* and the *Herfindahl-Hirschman coefficient* are measures of the degree of a concentration (e.g. household income, sales or market shares of firms in an industry, distribution of facilities in regions). This function returns both coefficients as non-standardized (G, HHI) and standardized values ($G*$, $HHI*$) and the HHI equivalent number (HHI_{eq}). For more information about the coefficients, see the single function documentations (*gini*, *herf*, *herf.eq*).

Value

a `numeric vector` with the five result values (`HHI, HHI*, HHI_{eq}, G, G*`)

Author(s)

Thomas Wieland

References

Doersam, P. (2004): “Wirtschaftsstatistik anschaulich dargestellt”. Heidenau : PD-Verlag.

Lessmann, C. (2005): “Regionale Disparitaeten in Deutschland und ausgesuchten OECD-Staaten im Vergleich”. *ifo Dresden berichtet*, **3/2005**. https://www.cesifo-group.de/link/ifodb_2005_3_25-33.pdf.

See Also

`gini`, `herf`, `herf.eq`

Examples

```
# Example from Doersam (2004)
# (Sales of four car manufacturing firms)
sales <- c(20,50,20,10)
disp(sales)
```

Freiburg *Employment data in Freiburg and Germany*

Description

Dataset with industry-specific employment in Freiburg and Germany in the years 2008 and 2014

Usage

```
data("Freiburg")
```

Format

A data frame with 9 observations on the following 8 variables.

- `industry` a factor with levels for the regarded industry based on the German official economic statistics (WZ2008)
- `e_Freiburg2008` a numeric vector with industry-specific employment in Freiburg 2008
- `e_Freiburg2014` a numeric vector with industry-specific employment in Freiburg 2014
- `e_g_Freiburg_0814` a numeric vector containing the growth of industry-specific employment in Freiburg 2008-2014, percentage
- `e_Germany2008` a numeric vector with industry-specific employment in Germany 2008
- `e_Germany2014` a numeric vector with industry-specific employment in Germany 2014
- `e_g_Germany_0814` a numeric vector containing the growth of industry-specific employment in Germany 2008-2014, percentage
- `color` a factor containg colors (`blue`, `brown`, ...)

Source

Statistische Aemter des Bundes und der Laender: Regionaldatenbank Deutschland, Tab. 254-74-4, own calculations

Examples

```
data(Freiburg)
# Loads the data
industries <- Freiburg$industry
x <- Freiburg$e_g_Freiburg_0814
y <- Freiburg$e_g_Germany_0814
z <- Freiburg$e_Freiburg2014
portfolio(x,y,z, "Freiburg", "Germany", "Growth portfolio Freiburg and Germany",
pcol="given", colsp=Freiburg$color, leg=1, leg_vec=industries, leg_fsize=0.6)
# Creates a portfolio comparing the industry growth in Freiburg and Germany
```

Freiburg1 *Distance matrix for grocery stores in Freiburg*

Description

Preliminary stage of an interaction matrix: Distance matrix for all statistical 42 districts and all 63 grocery stores (i = 42 submarkets x j = 63 suppliers) in Freiburg (Germany) including the size of the grocery stores.

Usage

```
data("Freiburg1")
```

Format

A data frame with 2646 observations on the following 4 variables.

`district` a numeric vector representing the 42 statistical districts of Freiburg

`store` a numeric vector identifying the store code of the mentioned grocery store in the study area

`salesarea` a numeric vector for the sales area of the grocery stores in sqm

`distance` a numeric vector for the distance from the places of residence (statistical districts) to the grocery stores in km

Source

Wieland, T. (2015): "Nahversorgung im Kontext raumoekonomischer Entwicklungen im Lebensmitteleinzelhandel - Konzeption und Durchfuehrung einer GIS-gestuetzten Analyse der Strukturen des Lebensmitteleinzelhandels und der Nahversorgung in Freiburg im Breisgau". Projektbericht. Goettingen : GOEDOC, Dokumenten- und Publikationsserver der Georg-August-Universitaet Goettingen. http://webdoc.sub.gwdg.de/pub/mon/2015/5-wieland.pdf

References

Wieland, T. (2015): "Nahversorgung im Kontext raumoekonomischer Entwicklungen im Lebensmitteleinzelhandel - Konzeption und Durchfuehrung einer GIS-gestuetzten Analyse der Strukturen des Lebensmitteleinzelhandels und der Nahversorgung in Freiburg im Breisgau". Projektbericht. Goettingen : GOEDOC, Dokumenten- und Publikationsserver der Georg-August-Universitaet Goettingen. http://webdoc.sub.gwdg.de/pub/mon/2015/5-wieland.pdf

Examples

```
data(Freiburg1)
data(Freiburg2)
# Loads the data

huff (Freiburg1, "district", "store", "salesarea", "distance")
# Huff probabilities with standard values (gamma = 1, lambda = -2)
```

```
huff (Freiburg1, "district", "store", "salesarea", "distance", gamma=1, lambda=-2, atype="pow",
dtype="pow", gamma2 = NULL, lambda2=NULL, output_total = TRUE, Freiburg2, "ppower")
# Calculation of total sales using district data in Freiburg2
```

Freiburg2 *Statistical districts of Freiburg*

Description

The 42 statistical districts of Freiburg (Germany) and the estimated annual purchasing power for groceries, based on average expenditures and population.

Usage

```
data("Freiburg2")
```

Format

A data frame with 42 observations on the following 2 variables.

`district` a numeric vector representing the 42 statistical districts of Freiburg

`ppower` a numeric vector containing the estimated absolute value of annual purchasing power for groceries in the district

Source

Wieland, T. (2015): "Nahversorgung im Kontext raumoekonomischer Entwicklungen im Lebensmitteleinzelhandel - Konzeption und Durchfuehrung einer GIS-gestuetzten Analyse der Strukturen des Lebensmitteleinzelhandels und der Nahversorgung in Freiburg im Breisgau". Projektbericht. Goettingen : GOEDOC, Dokumenten- und Publikationsserver der Georg-August-Universitaet Goettingen. http://webdoc.sub.gwdg.de/pub/mon/2015/5-wieland.pdf

References

Wieland, T. (2015): "Nahversorgung im Kontext raumoekonomischer Entwicklungen im Lebensmitteleinzelhandel - Konzeption und Durchfuehrung einer GIS-gestuetzten Analyse der Strukturen des Lebensmitteleinzelhandels und der Nahversorgung in Freiburg im Breisgau". Projektbericht. Goettingen : GOEDOC, Dokumenten- und Publikationsserver der Georg-August-Universitaet Goettingen. http://webdoc.sub.gwdg.de/pub/mon/2015/5-wieland.pdf

Examples

```
data(Freiburg1)
data(Freiburg2)
# Loads the data

huff (Freiburg1, "district", "store", "salesarea", "distance")
# Huff probabilities with standard values (gamma = 1, lambda = -2)
```

```
huff (Freiburg1, "district", "store", "salesarea", "distance", gamma=1, lambda=-2, atype="pow",
dtype="pow", gamma2 = NULL, lambda2=NULL, output_total = TRUE, Freiburg2, "ppower")
# Calculation of total sales using district data in Freiburg2
```

---

gini *Gini coefficient and Lorenz curve*

Description

Calculating the Gini coefficient of inequality (or concentration), standardized and non-standardized, and plotting the Lorenz curve

Usage

```
gini(x, norm = FALSE, lc = FALSE, lcx = "% of objects",
lcy = "% of regarded variable", lctitle = "Lorenz curve", lcg = FALSE,
lcgn = FALSE)
```

Arguments

`x`	A numeric vector (e.g. dataset of household income, sales turnover or supply)
`norm`	logical argument that indicates if the function output is the non-standardized or the standardized Gini coefficient (default: `norm = FALSE`, that means the non-standardized Gini coefficient is returned)
`lc`	logical argument that indicates if the Lorenz curve is plotted additionally (default: `lc = FALSE`, so no Lorenz curve is displayed)
`lcx`	if `lc = TRUE` (plot of Lorenz curve), `lcx` defines the x axis label
`lcy`	if `lc = TRUE` (plot of Lorenz curve), `lcy` defines the y axis label
`lctitle`	if `lc = TRUE` (plot of Lorenz curve), `lctitle` defines the overall title of the Lorenz curve plot
`lcg`	if `lc = TRUE` (plot of Lorenz curve), the logical argument `lcg` defines if the non-standardized Gini coefficient is displayed in the Lorenz curve plot
`lcgn`	if `lc = TRUE` (plot of Lorenz curve), the logical argument `lcgn` defines if the standardized Gini coefficient is displayed in the Lorenz curve plot

Details

The *Gini coefficient* (Gini 1912) is a popular measure of statistical dispersion, especially used for analyzing inequality or concentration of income, wealth or sales turnover of competing firms. The coefficient (G) varies between 0 (no inequality/concentration) and 1 (complete inequality/concentration). The *Lorenz curve* (Lorenz 1905), though developed independently, can be regarded as a graphical representation of the degree of inequality calculated by the *Gini coefficient* and can also be used for additional interpretations of G. The curve displays the deviations of the empirical distribution from a perfectly equal distribution as the difference between two graphs (the distribution curve and a diagonal line of perfect equality). This function calculates G and plots the *Lorenz curve* optionally. As there are several ways to calculate the *Gini coefficient*, this function uses the formula given in Doersam (2004). Because the maximum of G is not equal to 1, also a standardized coefficient ($G*$) with a maximum equal to 1 can be calculated alternatively.

Value

A single numeric value of the *Gini coefficient* ($0 < G < 1$) or the *standardized Gini coefficient* ($0 < G* < 1$) and, optionally, a plot of the *Lorenz curve*.

Author(s)

Thomas Wieland

References

Cerlani, L./Verme, P. (2012): “The origins of the Gini index: extracts from Variabilita e Mutabilita (1912) by Corrado Gini”. In: *The Journal of Economic Inequality*, **10**, 3, p. 421-443.

Doersam, P. (2004): “Wirtschaftsstatistik anschaulich dargestellt”. Heidenau : PD-Verlag.

Gini, C. (1912): “Variabilita e Mutabilita”. Contributo allo Studio delle Distribuzioni e delle Relazioni Statistiche. Bologna : Cuppini.

Lessmann, C. (2005): “Regionale Disparitaeten in Deutschland und ausgesuchten OECD-Staaten im Vergleich”. *ifo Dresden berichtet*, **3/2005**. https://www.cesifo-group.de/link/ifodb_2005_3_25-33.pdf.

Lorenz, M. O. (1905): “Methods of Measuring the Concentration of Wealth”. In: *Publications of the American Statistical Association*, **9**, 70, p. 209-219.

See Also

cv, gini.conc, gini.spec, herf

Examples

```
# Example from Doersam (2004):
sales <- c(20,50,20,10)
# sales turnover of four car manufacturing companies
gini(sales, lc=TRUE, lcx="percentage of companies", lcy="percentrage of sales",
lctitle="Lorenz curve of sales", lcg=TRUE, lcgn=TRUE)
# returs the non-standardized Gini coefficient (0.3) and
# plots the Lorenz curve with user-defined title and labels
gini(sales, norm=TRUE)
# returns the standardized Gini coefficient (0.4)
```

---

gini.conc *Gini coefficient of spatial industry concentration*

Description

Calculating the Gini coefficient of spatial industry concentration based on regional industry data (normally employment data)

Usage

```
gini.conc(e_ij, e_j)
```

Arguments

e_ij	a numeric vector with the employment of the industry i in region j
e_j	a numeric vector with the employment in region j

Details

The *Gini coefficient of spatial industry concentration* (G_i) is a special spatial modification of the Gini coefficient of inequality (see the function `gini()`). It represents the rate of spatial concentration of the industry i referring to j regions (e.g. cities, counties, states). The coefficient G_i varies between 0 (perfect distribution, respectively no concentration) and 1 (complete concentration in one region).

Value

A single numeric value ($0 < G_i < 1$)

Author(s)

Thomas Wieland

References

Farhauer, O./Kroell, A. (2013): "Standorttheorien: Regional- und Stadtoekonomik in Theorie und Praxis". Wiesbaden : Springer.

See Also

gini, gini.spec

Examples

```
# Example from Farhauer/Kroell (2013):
E_ij <- c(500,500,1000,7000,1000)
# employment of the industry in five regions
E_j <- c(20000,15000,20000,40000,5000)
# employment in the five regions
gini.conc (E_ij, E_j)
# Returns the Gini coefficient of industry concentration (0.4068966)
```

gini.spec — *Gini coefficient of regional specialization*

Description

Calculating the Gini coefficient of regional specialization based on regional industry data (normally employment data)

Usage

```
gini.spec(e_ij, e_i)
```

Arguments

`e_ij`	a numeric vector with the employment of the industries i in region j
`e_i`	a numeric vector with the employment in the industries i

Details

The *Gini coefficient of regional specialization* (G_j) is a special spatial modification of the *Gini coefficient* of inequality (see the function `gini()`). It represents the degree of regional specialization of the region j referring to i industries. The coefficient G_j varies between 0 (no specialization) and 1 (complete specialization).

Value

A single numeric value ($0 < G_j < 1$)

Author(s)

Thomas Wieland

References

Farhauer, O./Kroell, A. (2013): “Standorttheorien: Regional- und Stadtoekonomik in Theorie und Praxis”. Wiesbaden : Springer.

See Also

`gini`, `gini.conc`

Examples

```
# Example from Farhauer/Kroell (2013):
E_ij <- c(700,600,500,10000,40000)
# employment of five industries in the region
E_i <- c(30000,15000,10000,60000,50000)
# over-all employment in the five industries
gini.spec (E_ij, E_i)
# Returns the Gini coefficient of regional specialization (0.6222222)

# Example Freiburg
data(Freiburg)
# Loads the data
E_ij <- Freiburg$e_Freiburg2014
# industry-specific employment in Freiburg 2014
E_i <- Freiburg$e_Germany2014
# industry-specific employment in Germany 2014
gini.spec (E_ij, E_i)
# Returns the Gini coefficient of regional specialization (0.2089009)
```

Goettingen *Grocery store supply in Goettingen*

Description

Grocery stores and their sales areas in Goettingen (Germany) on the level of 69 statistical districts, data from Wieland (2011)

Usage

```
data("Goettingen")
```

Format

A data frame with 69 observations on the following 5 variables.

- `statdist` a factor with 69 levels containing the codes of the statistical districts (`010`, `011`, ..., `180`)
- `stores_count` a numeric vector containing the number of grocery stores in the regarded district
- `salesarea_sqm` a numeric vector containing the sales area (in sqm) of the grocery stores in the regarded district
- `stores_p1000p` a numeric vector containing the density of grocery stores in the regarded district (stores per 1.000 inhabitants)
- `salesarea_p1000p` a numeric vector containing the density of grocery sales area in the regarded district (sales area in sqm per 1.000 inhabitants)

Details

The grocery store dataset contains all grocery stores (supermarkets and discounters) in Goettingen (Germany) as of march 2011.

Source

Wieland, T. (2011): “Nahversorgung mit Lebensmitteln in Goettingen 2011 - Eine Analyse der Angebotssituation im Goettinger Lebensmitteleinzelhandel unter besonderer Beruecksichtigung der Versorgungsqualitaet”. *Goettinger Statistik Aktuell*, **35**. Goettingen. http://www.goesis.goettingen.de/pdf/Aktuell35.pdf.

Examples

```
data(Goettingen)
# Loads the data
cv(Goettingen$stores_p1000p)
# cv of stores per 1.000 inhabitants
cv(Goettingen$salesarea_p1000p)
# cv of sales area per 1.000 inhabitants
cv(Goettingen$salesarea_p1000p, norm=TRUE)
# standardized cv of sales area per 1.000 inhabitants
```

---

hansen — *Hansen accessibility*

Description

Calculating the Hansen accessibility for given origins and destinations

Usage

```
hansen(od_dataset, origins, destinations, attrac, dist, gamma = 1, lambda = -2,
atype = "pow", dtype = "pow", gamma2 = NULL, lambda2 = NULL)
```

Arguments

`od_dataset`	an interaction matrix which is a `data.frame` containing the origins, destinations, the distances between them and a size variable for the opportunities of the destinations
`origins`	the column in the interaction matrix `od_dataset` containing the origins
`destinations`	the column in the interaction matrix `od_dataset` containing the destinations
`attrac`	the column in the interaction matrix `od_dataset` containing the "attractivity" variable of the destinations (e.g. no. of opportunities)
`dist`	the column in the interaction matrix `od_dataset` containing the transport costs (e.g. travelling time)
`gamma`	a single numeric value for the exponential weighting of size (default: 1)
`lambda`	a single numeric value for the exponential weighting of distance (transport costs, default: -2)
`atype`	Type of attractivity weighting function: `atype = "pow"` (power function), `atype = "exp"` (exponential function) or `atype = "logistic"` (default: `atype = "pow"`)

dtype — Type of distance weighting function: `dtype = "pow"` (power function), `dtype = "exp"` (exponential function) or `dtype = "logistic"` (default: `dtype = "pow"`)

gamma2 — if `atype = "logistic"` a second γ parameter is needed

lambda2 — if `dtype = "logistic"` a second λ parameter is needed

Details

The *Hansen accessibility* (Hansen 1959) can be regarded as a *potential model* of *spatial interaction*. From a theoretical perspective, the accessibility (A) can be seen as the sum of all utilities of every opportunity outgoing from given starting points (O_j) and weighted by the distance to them (d_{ij}) (Orpana/Lampinen 2003). Originally the weighting function of distance is not explicitly given and the "attractivities" (e.g. size of the activity at the destinations) is not weighted($A_{ij} = \sum_j O_j f(d_{ij})$). This specification is relaxed is this case, so both variables can be weighted by a power (e.g. $d_{ij}^{-}\lambda$), exponential or logistic function.

Value

Returns a `data frame` with the origins and the accessibility values (column `accessibility`).

Author(s)

Thomas Wieland

References

Hansen, W. G. (1959): “How Accessibility Shapes Land Use”. In: *Journal of the American Institute of Planners*, **25**, 2, p. 73-76.

Orpana, T./Lampinen, J. (2003): “Building spatial choice models from aggregate data”. In: *Journal of Regional Science*, **43**, 2, p. 319-347.

See Also

huff

Examples

```
data(Freiburg1)
data(Freiburg2)
# Loads the data

hansen (Freiburg1, "district", "store", "salesarea", "distance", gamma = 1, lambda = -2)
# Computes the Hansen accessibility from the statistical districts of Freiburg
# regarding the grocery stores and their sizes
```

`herf` *Herfindahl-Hirschman coefficient*

Description

Calculating the Herfindahl-Hirschman coefficient of concentration, standardized and non-standardized

Usage

```
herf(x, norm = FALSE)
```

Arguments

`x`	A numeric vector (e.g. dataset of sales turnover or size of firms)
`norm`	logical argument that indicates if the function output is the non-standardized or the standardized Herfindahl-Hirschman coefficient (default: `norm = FALSE`, that means the non-standardized Herfindahl-Hirschman coefficient is returned)

Details

The *Herfindahl-Hirschman coefficient* is a popular measure of statistical dispersion, especially used for analyzing concentration in markets, regarding sales turnovers or sizes of n competing firms in an industry. This indicator is e.g. used as a measure of market power and distortions of competition in the governmental competition policy. The coefficient (HHI) varies between $\frac{1}{n}$ (parity resp. no concentration) and 1 (complete concentration). Because the minimum of HHI is not equal to 0, also a standardized coefficient ($HHI*$) with a minimum equal to 0 can be calculated alternatively.

Value

A single numeric value of the *Herfindahl-Hirschman coefficient* ($\frac{1}{n} < HHI < 1$) or the *standardized Herfindahl-Hirschman coefficient* ($0 < HHI* < 1$).

Author(s)

Thomas Wieland

References

Doersam, P. (2004): “Wirtschaftsstatistik anschaulich dargestellt”. Heidenau : PD-Verlag.

Lessmann, C. (2005): “Regionale Disparitaeten in Deutschland und ausgesuchten OECD-Staaten im Vergleich”. *ifo Dresden berichtet*, **3/2005**. https://www.cesifo-group.de/link/ifodb_2005_3_25-33.pdf.

See Also

`cv`, `gini`, `herf.eq`

Examples

```
# Example from Doersam (2004):
sales <- c(20,50,20,10)
# sales turnover of four car manufacturing companies
herf(sales)
# returns the non-standardized HHI (0.34)
herf(sales, norm=TRUE)
# returns the standardized HHI (0.12)
```

---

herf.eq — *Herfindahl-Hirschman coefficient equivalent number*

Description

Calculating the Herfindahl-Hirschman coefficient equivalent number of objects

Usage

```
herf.eq(x)
```

Arguments

x	A numeric vector (e.g. dataset of sales turnover or size of firms)

Details

The *equivalent number* referring to the *Herfindahl-Hirschman coefficient* (see the function `herf()`) reflects the theoretical number of economic objects (normally firms) where a calculated coefficient is $\frac{1}{n}$ (parity).

Value

A single numeric value of the *Herfindahl-Hirschman coefficient equivalent number*.

Author(s)

Thomas Wieland

References

Doersam, P. (2004): "Wirtschaftsstatistik anschaulich dargestellt". Heidenau : PD-Verlag.

See Also

cv, gini, herf

Examples

```
# Example from Doersam (2004):
sales <- c(20,50,20,10)
# sales turnover of four car manufacturing companies
herf(sales)
# returns the non-standardized HHI (0.34)
herf.eq(sales)
# returns the HHI equivalent number (2.941176)
```

---

huff *Huff model*

Description

Calculating market areas using the probabilistic market area model by Huff

Usage

```
huff(huffdataset, origins, locations, attrac, dist, gamma = 1, lambda = -2,
atype = "pow", dtype = "pow", gamma2 = NULL, lambda2 = NULL,
output_total = FALSE, origindataset, od_localmarket)
```

Arguments

huffdataset	an interaction matrix which is a `data.frame` containing the origins, locations and the explanatory variables
origins	the column in the interaction matrix `huffdataset` containing the origins (e.g. ZIP codes)
locations	the column in the interaction matrix `huffdataset` containing the locations (e.g. store codes)
attrac	the column in the interaction matrix `huffdataset` containing the attractivity variable (e.g. sales area)
dist	the column in the interaction matrix `huffdataset` containing the transport costs (e.g. travelling time)
gamma	a single numeric value for the exponential weighting of size (default: 1)
lambda	a single numeric value for the exponential weighting of distance (transport costs, default: -2)
atype	Type of attractivity weighting function: `atype = "pow"` (power function), `atype = "exp"` (exponential function) or `atype = "logistic"` (default: `atype = "pow"`)
dtype	Type of distance weighting function: `dtype = "pow"` (power function), `dtype = "exp"` (exponential function) or `dtype = "logistic"` (default: `dtype = "pow"`)
gamma2	if `atype = "logistic"` a second γ parameter is needed
lambda2	if `dtype = "logistic"` a second λ parameter is needed

`output_total` logical argument that indicates if the total sales and total shares of the j locations is the output of the function (default: `output_total = FALSE`)

`origindataset` a `data.frame` containing the origins and additional data about their local market potential (e.g. purchasing power, population, households, ...)

`od_localmarket` a column in `origindataset` containing the data about the local market potential (e.g. purchasing power, population, households, ...)

Details

The *Huff model* (Huff 1962, 1963, 1964) is the most popular *spatial interaction model* for retailing and services and belongs to the family of *probabilistic market area models*. The basic idea of the model is that consumer decisions are not deterministic but probabilistic, so the decision of customers for a shopping location in a competitive environment cannot be predicted exactly. The results of the model are probabilities for these decisions, which can be interpreted as market shares of the regarded locations (j) in the customer origins (i), p_{ij}, which can be regarded as an equilibrium solution with logically consistent market shares (0 < p_{ij} < 1, $\sum_{j=1}^{n} p_{ij} = 1$). From a theoretical perspective, the model is based on an utility function with two explanatory variables ("attractivity" of the locations, transport costs between origins and locations), which are weighted by an exponent: $U_{ij} = A_j^{\gamma} d_{ij}^{-} \lambda$. But the weighting functions can also be exponential or logistic. This function computes the market shares from a given interaction matrix and given weighting parameters. If `output_total = TRUE` you need local market information about the origins (e.g. purchasing power, population size etc.) filed in another `data.frame`, so the function results are the total sales/shares of the given stores/locations.

Value

Returns either the input interaction matrix including the calculated shares (`p_ij`) (if `output_total = FALSE`) or the total sales (`sum_E_j`) and total shares (`share_j`) of the stores locations (if `output_total = TRUE`). Both results are `data.frame`.

Author(s)

Thomas Wieland

References

Huff, D. L. (1962): “Determination of Intra-Urban Retail Trade Areas”. Los Angeles : University of California.

Huff, D. L. (1963): “A Probabilistic Analysis of Shopping Center Trade Areas”. In: *Land Economics*, **39**, 1, p. 81-90.

Huff, D. L. (1964): “Defining and Estimating a Trading Area”. In: *Journal of Marketing*, **28**, 4, p. 34-38.

Loeffler, G. (1998): “Market areas - a methodological reflection on their boundaries”. In: *GeoJournal*, **45**, 4, p. 265-272.

Wieland, T. (2015): “Nahversorgung im Kontext raumoekonomischer Entwicklungen im Lebensmitteleinzelhandel - Konzeption und Durchfuehrung einer GIS-gestuetzten Analyse der Strukturen des Lebensmitteleinzelhandels und der Nahversorgung in Freiburg im Breisgau”. Projektbericht.

Goettingen : GOEDOC, Dokumenten- und Publikationsserver der Georg-August-Universitaet Goettingen. http://webdoc.sub.gwdg.de/pub/mon/2015/5-wieland.pdf

See Also

converse, reilly, reilly.lambda

Examples

```
data(Freiburg1)
data(Freiburg2)
# Loads the data

huff (Freiburg1, "district", "store", "salesarea", "distance")
# Standard weighting (power function with gamma=1 and lambda=-2)

huff (Freiburg1, "district", "store", "salesarea", "distance", gamma = 1, lambda = -2,
atype="pow", dtype="pow", gamma2 = NULL, lambda2 = NULL, output_total = TRUE, Freiburg2, "ppower")
# Calculating total sales of the stores
```

---

krugman.spec *Krugman coefficient of regional specialization*

Description

Calculating the Krugman coefficient for the specialization of two regions based on regional industry data (normally employment data)

Usage

```
krugman.spec(e_ij, e_il, e_j, e_l)
```

Arguments

e_ij	a numeric vector with the employment of the industries i in region j
e_il	a numeric vector with the employment of the industries i in region l
e_j	a single numeric value reflecting the employment in region j
e_l	a single numeric value reflecting the employment in region l

Details

The *Krugman coefficient of regional specialization* (K_{jl}) is a measure for the dissimilarity of the industrial structure of two regions (j and l) regarding the employment in the i industries in these regions. The coefficient K_{jl} varies between 0 (no specialization/same structure) and 2 (maximum difference, that means there is no single industry localized in both regions).

Value

A single numeric value ($0 < K_{jl} < 2$)

Author(s)

Thomas Wieland

References

Farhauer, O./Kroell, A. (2013): "Standorttheorien: Regional- und Stadtoekonomik in Theorie und Praxis". Wiesbaden : Springer.

See Also

gini.conc, gini.spec

Examples

```
# Example from Farhauer/Kroell (2013), modified:
E_ij <- c(20,10,70,0,0)
# employment of five industries in region j
E_il <- c(0,0,0,60,40)
# employment of five industries in region l
E_j <- 100
E_l <- 100
# over-all employment in regions j and l
krugman.spec(E_ij, E_il, E_j, E_l)
# results the specialization coefficient (2)
```

---

locq *Location quotient*

Description

Calculating the *location quotient*

Usage

```
locq(e_ij, e_j, e_i, e)
```

Arguments

e_ij	a single numeric value with the employment of industry i in region j
e_j	a single numeric value with the over-all employment in region j
e_i	a single numeric value with the over-all employment in industry i
e	a single numeric value with the over-all employment in all regions

Details

The *location quotient* is a simple measure for the concentration of an industry (i) in a region (j) and is also the mathematical basis for other related indicators in regional economics (e.g. `gini.conc()`). The function returns the value LQ which is equal to 1 if the concentration of the regarded industry is exactly the same as the over-all concentration (that means, it is proportionally represented in region j). If the value of LQ is smaller (bigger) than 1, the industry is underrepresented (overrepresented). The function checks the input values for errors (i.e. if employment in a region is bigger than over-all employment).

Value

A single numeric value (LQ)

Author(s)

Thomas Wieland

References

Farhauer, O./Kroell, A. (2013): “Standorttheorien: Regional- und Stadtoekonomik in Theorie und Praxis”. Wiesbaden : Springer.

See Also

gini.conc, gini.spec

Examples

```
# Example from Farhauer/Kroell (2013):
locq (1714, 79006, 879213, 15593224)
# returns the location quotient (0.3847623)
```

---

portfolio *Portfolio matrix*

Description

Portfolio matrix plot comparing two numeric vectors

Usage

```
portfolio(x, y, z, label_x = "X", label_y = "Y", heading = "Portfolio",
pcol = "given", colsp = 0, leg = FALSE, leg_vec = 0, leg_fsize = 1,
leg_x = -max_val, leg_y = -max_val/2)
```

Arguments

`x`	A numeric vector representing the values for the x axis
`y`	A numeric vector representing the values for the y axis
`z`	A numeric vector representing the size of the points/bubbles
`label_x`	Label for the x axis
`label_y`	Label for the y axis
`heading`	Heading for the plot
`pcol`	indicates if the colors of the points are given by the user (`pcol = "given"`) and defined by the vector `colsp` or set by random (`pcol = "random"`)
`colsp`	a vector representing the user-defined colors of the points
`leg`	logical argument that indicates if the plot has a legend or not (default: `leg = FALSE`)
`leg_vec`	if `leg = TRUE`, this vector defines the values for the plot legend
`leg_fsize`	if `leg = TRUE`, this value defines the font size of the legend
`leg_x`	if `leg = TRUE`: x coordinate for the legend (default: `leg_x=-max_val`, where `max_val` is the maximum value of all values in the dataset)
`leg_y`	if `leg = TRUE`: y coordinate for the legend (default: `leg_y=-max_val/2`, where `max_val` is the maximum value of all values in the dataset)

Details

The *portfolio matrix* is a graphic tool displaying the development of one variable compared to another variable. The plot shows the regarded variable on the x axis and a variable with which it is confronted on the y axis while the graph is divided in four quadrants. Originally, the *portfolio matrix* was developed by the *Boston Consulting Group* to analyze the performance of product lines in marketing, also known as the *growth-share matrix*. The quadrants show the performace of the regarded objects (stars, cash cows, question marks, dogs) (Henderson 1973). But the *portfolio matrix* can also be used to analyze/illustrate the world market integration of a region or a national economy by confronting e.g. the increase in world market share (x axis) and the world trade growth (y axis) (Baker et al. 2002). Another option is to analyze/illustrate the economic performance of a region (Howard 2007). E.g. it is possible to confront the growth of industries in a region with the all-over growth of these industries in the national economy.

Value

A plot of the portfolio matrix

Author(s)

Thomas Wieland

References

Baker, P./von Kirchbach, F./Mimouni, M./Pasteels, J.-M. (2002): “Analytical tools for enhancing the participation of developing countries in the Multilateral Trading System in the context of the Doha Development Agenda”. In: *Aussenwirtschaft*, **57**, 3, p. 343-372.

Howard, D. (2007): “A regional economic performance matrix - an aid to regional economic policy development”. In: *Journal of Economic and Social Policy*, **11**, 2, Art. 4.

Henderson, B. D. (1973): “The Experience Curve - Reviewed, IV. The Growth Share Matrix or The Product Portfolio”. The Boston Consulting Group (BCG).

See Also

shift

Examples

```
data(Freiburg)
# Loads the data
industries <- Freiburg$industry
x <- Freiburg$e_g_Freiburg_0814
y <- Freiburg$e_g_Germany_0814
z <- Freiburg$e_Freiburg2014
portfolio(x,y,z, "Freiburg", "Germany", "Growth portfolio Freiburg and Germany",
pcol="given", colsp=Freiburg$color, leg=1, leg_vec=industries, leg_fsize=0.6)
# Creates a portfolio comparing the industry growth in Freiburg and Germany
```

---

reilly *Law of retail gravitation by Reilly*

Description

Calculating the proportion of sales from an intermediate town between two cities or retail locations

Usage

```
reilly(P_a, P_b, D_a, D_b, gamma = 1, lambda = 2)
```

Arguments

P_a	a single numeric value of attractivity/population size of location/city a
P_b	a single numeric value of attractivity/population size of location/city b
D_a	a single numeric value of the distance from the intermediate town to location/city a
D_b	a single numeric value of the distance from the intermediate town to location/city b
gamma	a single numeric value for the exponential weighting of size (default: 1)
lambda	a single numeric value for the exponential weighting of distance (transport costs, default: -2)

Details

The *law of retail gravitation* by Reilly (1929, 1931) was the first *spatial interaction model* for retailing and services. This "law" states that two cities/locations attract customers from an intermediate town proportionally to the attractivity/population size of the two cities/locations and in inverse proportion to the squares of the transport costs (e.g. distance, travelling time) from these two locations to the intermediate town. But both variables can be weighted by exponents. The distance exponent can also be derived from empirical data (see the function `reilly.lambda`). The *breaking point formula* by Converse (1949) is a separate transformation of Reilly's law (see the function `converse`). The models by Reilly and Converse are simple *spatial interaction models* and are considered as *deterministic market area models* due to their exact allocation of demand origins to locations. A probabilistic approach including a theoretical framework was developed by Huff (1962) (see the function `huff`).

Value

a list with three values (`relation_AB`: relation of trade between cities/locations a and b, `prop_A`: proportion of city/location a, `prop_B`: proportion of city/location b)

Author(s)

Thomas Wieland

References

Converse, P. D. (1949): “New Laws of Retail Gravitation”. In: *Journal of Marketing*, **14**, 3, p. 379-384.

Huff, D. L. (1962): “Determination of Intra-Urban Retail Trade Areas”. Los Angeles : University of California.

Loeffler, G. (1998): “Market areas - a methodological reflection on their boundaries”. In: *GeoJournal*, **45**, 4, p. 265-272

Reilly, W. J. (1929): “Methods for the Study of Retail Relationships”. *Studies in Marketing*, **4**. Austin : Bureau of Business Research, The University of Texas.

Reilly, W. J. (1931): “The Law of Retail Gravitation”. New York.

See Also

`huff`, `converse`, `reilly.lambda`

Examples

```
# Example from Converse (1949):
reilly (39851, 37366, 27, 25)
# two cities (pop. size 39.851 and 37.366)
# with distances of 27 and 25 miles to intermediate town

myresults <- reilly (39851, 37366, 27, 25)
myresults$prop_A
# proportion of location a
```

reilly.lambda *Distance exponent in law of retail gravitation by Reilly*

Description

Deriving the distance exponent in Reilly's law of retail gravitation from empirical data

Usage

```
reilly.lambda(P_a, P_b, D_a, D_b, B_a_div_B_b)
```

Arguments

`P_a`	a single numeric value of attractivity/population size of location/city a
`P_b`	a single numeric value of attractivity/population size of location/city b
`D_a`	a single numeric value of the distance from the intermediate town to location/city a
`D_b`	a single numeric value of the distance from the intermediate town to location/city b
`B_a_div_B_b`	relation of trade between cities/locations a and b

Details

The *law of retail gravitation* by Reilly (1929, 1931) was the first *spatial interaction model* for retailing and services. This "law" states that two cities/locations attract customers from an intermediate town proportionally to the attractivity/population size of the two cities/locations and in inverse proportion to the squares of the transport costs (e.g. distance, travelling time) from these two locations to the intermediate town. But both variables can be weighted by exponents (see the function `reilly`). The distance exponent (lambda) can also be derived from empirical data. The *breaking point formula* by Converse (1949) is a separate transformation of Reilly's law (see the function `converse`). The models by Reilly and Converse are simple *spatial interaction models* and are considered as *deterministic market area models* due to their exact allocation of demand origins to locations. A probabilistic approach including a theoretical framework was developed by Huff (1962) (see the function `huff`).

Value

a single numeric value (lambda)

Author(s)

Thomas Wieland

References

Converse, P. D. (1949): "New Laws of Retail Gravitation". In: *Journal of Marketing*, **14**, 3, p. 379-384.

Huff, D. L. (1962): "Determination of Intra-Urban Retail Trade Areas". Los Angeles : University of California.

Loeffler, G. (1998): "Market areas - a methodological reflection on their boundaries". In: *GeoJournal*, **45**, 4, p. 265-272

Reilly, W. J. (1929): "Methods for the Study of Retail Relationships". *Studies in Marketing*, **4**. Austin : Bureau of Business Research, The University of Texas.

Reilly, W. J. (1931): "The Law of Retail Gravitation". New York.

See Also

reilly, huff, converse

Examples

```
# Example from Converse (1949):
reilly.lambda (39851, 37366, 27, 25, 0.9143555)
# returns lambda of about 2
```

---

shift *Shift-share analysis*

Description

Analyzing regional growth with the shift-share analysis

Usage

```
shift(region_t, region_t1, nation_t, nation_t1)
```

Arguments

region_t	a numeric vector with i values containing the employment in i industries in a region at time t
region_t1	a numeric vector with i values containing the employment in i industries in a region at time $t + 1$
nation_t	a numeric vector with i values containing the employment in i industries in the national economy at time t
nation_t1	a numeric vector with i values containing the employment in i industries in the national economy at time $t + 1$

Details

The *shift-share analysis* (Dunn 1960) adresses the regional growth (or decline) regarding the overall development in the national economy. The aim of this analysis model is to identify which parts of the regional economic development can be traced back to national trends, effects of the regional industry structure and (positive) regional factors. As there is more than one way to calculate a *shift-share analysis* and the terms are not used consequently in the regional economic literature, this function and the documentation use the formulas and terms given in Farhauer/Kroell (2013). This function calculates the *net proportionality shift* (nps), the *net differential shift* (nds) and the *net total shift* (nts) where the last one represents the residuum of (positive) regional factors.

Value

a list with three values (nps, nds, nts)

Author(s)

Thomas Wieland

References

Casler, S. D. (1989): “A Theoretical Context for Shift and Share Analysis”. In: *Regional Studies*, **23**, 1, p. 43-48.

Dunn, E. S. Jr. (1960): “A statistical and analytical technique for regional analysis”. In: *Papers and Proceedings of the Regional Science Association*, **6**, p. 97-112.

Farhauer, O./Kroell, A. (2013): “Standorttheorien: Regional- und Stadtoekonomik in Theorie und Praxis”. Wiesbaden : Springer.

See Also

portfolio

Examples

```
# Example from Farhauer/Kroell (2013):
region_A_t <- c(90,20,10,60)
region_A_t1 <- c(100,40,10,55)
# data for region A (time t and t+1)
nation_X_t <- c(400,150,150,400)
nation_X_t1 <- c(440,210,135,480)
# data for the national economy (time t and t+1)
resultsA <- shift(region_A_t, region_A_t1, nation_X_t, nation_X_t1)
# results for region A
region_B_t <- c(60,30,30,40)
region_B_t1 <- c(85,55,40,35)
# data for region B (time t and t+1)
resultsB <- shift(region_B_t, region_B_t1, nation_X_t, nation_X_t1)
# results for region B
region_C_t <- c(250,100,110,300)
region_C_t1 <- c(255,115,85,390)
# data for region C (time t and t+1)
```

```
resultsC <- shift(region_C_t, region_C_t1, nation_X_t, nation_X_t1)
# results for region C

# Example Freiburg dataset
data(Freiburg)
# Loads the data
shift(Freiburg$e_Freiburg2008, Freiburg$e_Freiburg2014, Freiburg$e_Germany2008,
Freiburg$e_Germany2014)
# results for Freiburg and Germany (2008 vs. 2014)
```

Index