SPATIAL MEASURES FOR ARCGIS 10.1/2

User's guide

Ostrava 2015

Contents

Introduction

ArcGIS toolbox **Spatial measures** was developed under VŠB - Technical university of Ostrava, at Institute of Geoinformatics. This document is also part of thesis.

The following chapters describe implemented indices, measures and installation of created toolbox.

For all indices and measures is there only **ONE NUMBER** result for one input feature, except of location quotient (there is unique number for every point or polygon, in input feature).

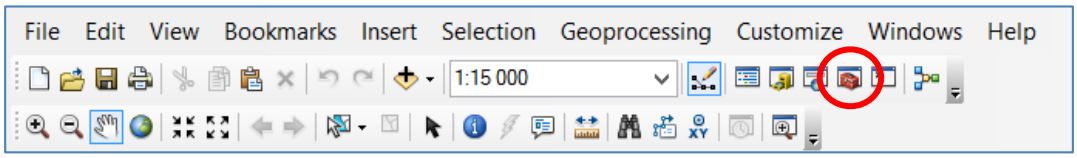
You can also have **multiple input features** for computation, but the features must be added to ArcGIS before that and you must adding those input features one by one, otherwise the computation will fail.

1 Installation

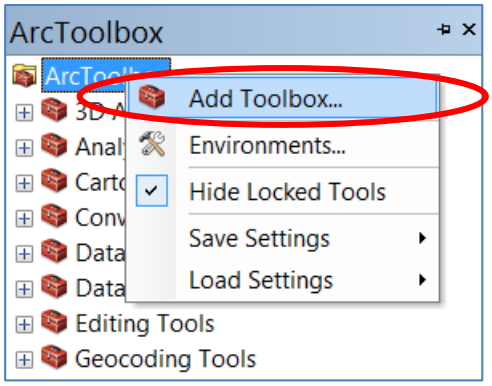
The installation takes like one minute, so it's pretty simple.

This can be done in the following steps:

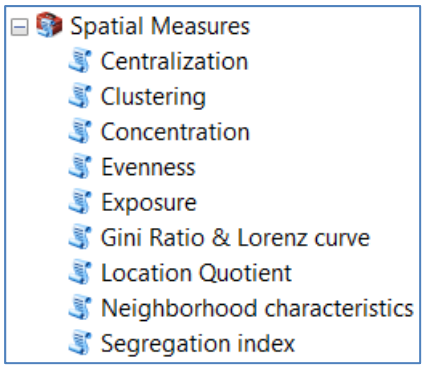
1. Download ***Spatial Measures toolbox*** to your computer from: http://spatial.vsb.cz/
2. Unzip downloaded content to any location you want.
3. Launch your ArcGIS 10.1/2 software.
4. Open window for *ArcToolbox*.



5. Right click on *ArcToolbox* in *ArcToolbox* window and then click on "Add Toolbox...".

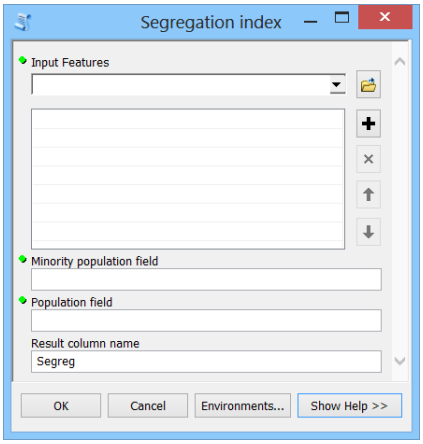


6. Find unzipped ***Spatial measures toolbox*** and then add/open "**Spatial_Measures.pyt**".
7. Then enjoy it!



2 Measures and indices

2.1 Segregation index (Temelová, Sýkora, 2005)



Indicates the relative difference between the spatial distribution of one minority group population and the rest of the population.

$$IS = \frac{D_{XY}}{\left[1 - \frac{X}{Y}\right]}$$

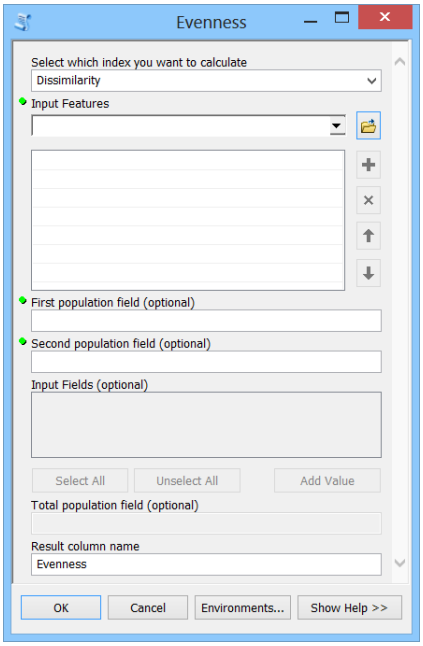
where D_{XY} is index of dissimilarity between minority group and total population (defined in chapter 2.2), X the total minority population of the entire area, Y the total population of the entire area.

The index ranges from **0.0** (no segregation) to **1.0** (complete segregation).

In the user interface (*)

X	"*Minority population field*"
Y	"*Population field*"

2.2 Evenness (Massey, Denton, 1988)



This dimension measures the spatial segregation of various groups. Segregation is lowest when each area reflects the overall population share, considering minority and majority groups.

2.2.1 Index of Dissimilarity

This index measures the percentage of one group that would have to change residence in order to produce an even distribution of two groups among areas.

$$D = \frac{1}{2}\sum_{i=1}^{N}\left|\frac{P_{ig}}{P_g} - \frac{P_{ih}}{P_h}\right| = \frac{1}{2}\sum_{i=1}^{N}|x_i - y_i|$$

where x_i and y_i represents proportions of the population in each group in each subunit *i* out of the area. So P_{ig}, P_{ih} are minority (majority) populations in area subunit *i*. P_g, P_h are total minority (majority) populations in area. N is the number of subunits in area.

The index ranges from **0.0** (no segregation) to **1.0** (complete segregation).

In the user interface (*)

P_g	"*First population field*"
P_h	"*Second population field*"

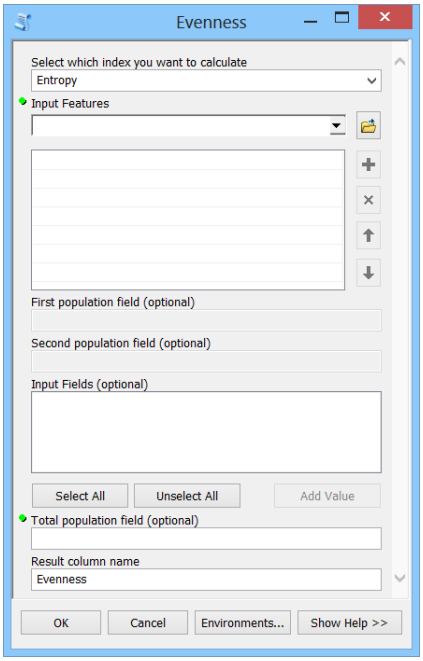
* - You may only choose number values (Double, Long, etc.)

2.2.2 Entropy Index (Theil index)

Entropy index also measures the differences in the distributions of groups within a geographical area. Unlike the *index of dissimilarity*, however, it allows for the calculation of measures for **multiple groups** simultaneously.

First it measures diversity of the total area:

$$E = \sum_{j=1}^{Z} (X_j)\, ln(1/X_j)$$



where X_J is the share for the population of the entire area in each category of variable studied and Z is number of categories.

Then it measures the individual entropy score:

$$E_i = \sum_{j=1}^{Z} (X_j) ln(1/X_j)$$

where X_j is the share of the total in each category of the variable studied, in the area subunit *i*.

The result of this index is:

$$\boldsymbol{H = \sum_{i=1}^{N} [t_i(E - E_i)]/ET}$$

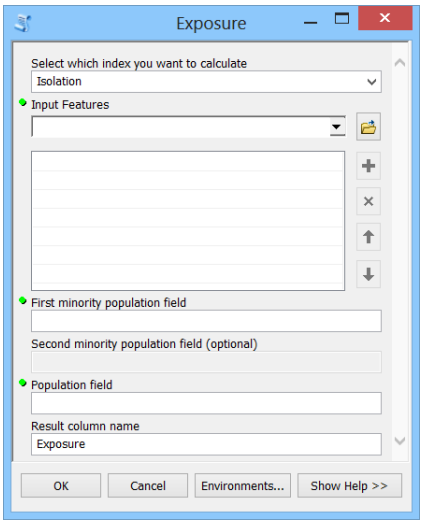
where t_i represents total population of subunit *i* and T is the total area population.

The measure varies between **0.0** (all subunits have the same composition as the overall area) and **1.0** (all subunits contain only one group).

In the user interface (*)

Z	"*Input Fields*"
T	"*Total population field* "

2.3 Exposure (Massey, Denton, 1988)



These indices measure the extent of possible contract between group members. The following two indices ranges from **0.0** (no segregation) to **1.0** (complete segregation).

2.3.1 Isolation

This index measures likelihood that a randomly chosen member of one group will meet another member of the same group.

$$IS_{xx} = \sum_{i=1}^{N} \left[\left(\frac{x_i}{X}\right) \left(\frac{x_i}{t_i}\right) \right]$$

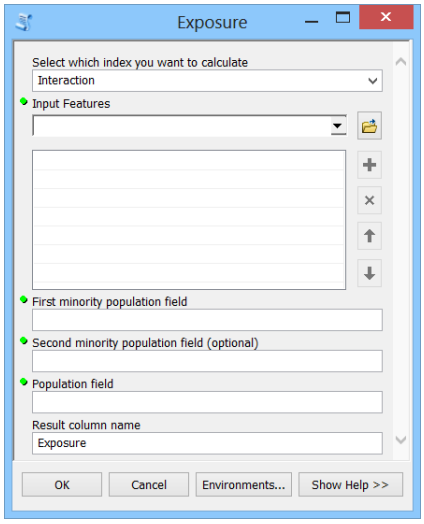
where x_i represents the minority population in area subunit *i*, X the total minority population of the entire area, t_i the total population in area subunit *i*.

In the user interface (*)

X	"*First minority population field*"
t_i	"*Population field* "

* - You may only choose number values (Double, Long, etc.)

2.3.2 Interaction



This index measures the probability that a member of one group will meet a member of another group.

$$IN_{xy} = \sum_{i=1}^{N} \left[\left(\frac{x_i}{X} \right) \left(\frac{y_i}{t_i} \right) \right]$$

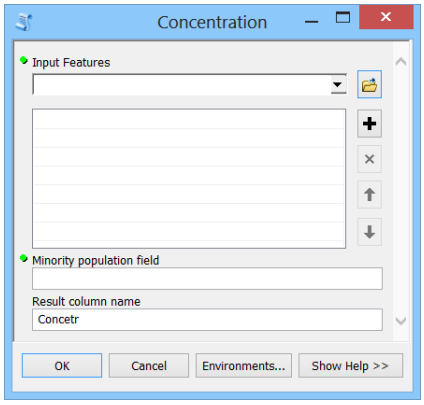
where x_i, X and t_i are defined as before and y_i the total population of the second group in the area subunit *i*.

In the user interface (*)

X	"*First minority population field*"
t_i	"*Population field* "

2.4 Concentration (Massey, Denton, 1988)

The indices categorized as concentration measures introduce the idea of physical space. If groups have equal population size but occupy different amounts of space, the area would be considered as segregated.



Delta (Concentration) index is interpreted as the share of minority members that would have to shift units to achieve a uniform density of minority members over all units.

$$DEL = \frac{1}{2} \sum_{i=1}^{N} \left| \frac{x_i}{X} - \frac{a_i}{A} \right|$$

where x_i and X are defined as before, a_i the land area of area subunits and A the total land area of the entire geographical unit.

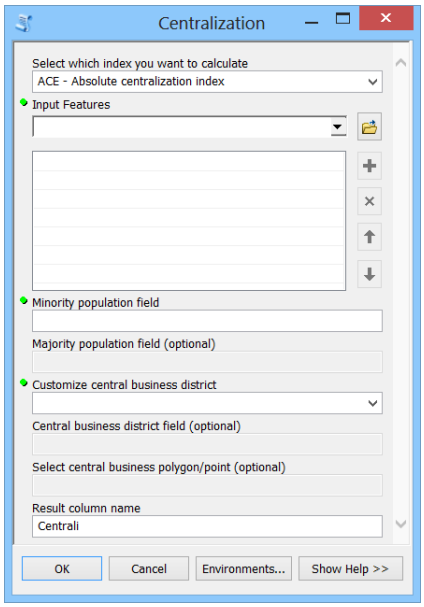
The index varies from **0.0** to **1.0**, where a score of **1.0** means that a group has achieved the maximum spatial concentration possible (all members live in the smallest areal units), and a score of **0.0** indicates the maximum deconcentration possible (all X members live in the largest areal units).

In the user interface (*)

X	"M*inority population field*"
A	Is computed by ArcGIS

* - You may only choose number values (Double, Long, etc.)

2.5 Centralization (Massey, Denton, 1988)



In this category, the concern is the degree to which a group is near the center of the geographical unit. The nearness to the center of the area can be examined with absolute or relative measures. They have range of **-1.0** to **+1.0**. A negative score means a tendency for the minority group to live in the outlying areas, a positive score represents a tendency for minority members to live near the city center and score of 0 indicates that group has a uniform distribution throughout the geographical area.

2.5.1 Absolute Centralization Index

This index measures the distribution of the minority group around the center of the geographical unit.

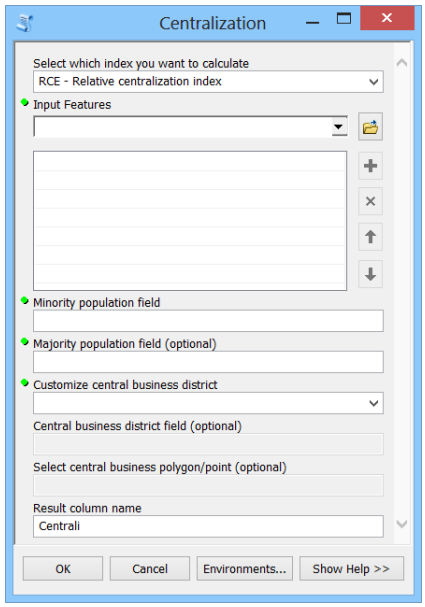
$$ACE = \sum_{i=1}^{N}(C_{i-1}A_i) - \sum_{i=1}^{N}(C_iA_{i-1})$$

where the N area subunits are ordered by increasing distance from the central business district, C is the cumulative proportion of the minority population up through subunit *i*, and A is the cumulative proportion of land area up through subunit *i*.

In the user interface (*)

C	"*Minority population field*"
A	Is computed by ArcGIS

2.5.2 Relative Centralization Index



This index measures the area profile of the minority and majority groups. It represents the relative share of one group's population that would have to change their residences to match the centralization distribution of the other group.

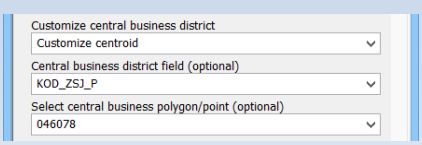
$$RCE = \sum_{i=1}^{N}(x_{i-1}y_i) - \sum_{i=1}^{N}(x_iy_{i-1})$$

where N is defined as before, x_i represents the cumulative proportion of the minority population in subunit *i* and y_i the cumulative proportion of the majority population in subunit *i*.

In the user interface (*)

x	"*Minority population field*"
y	"*Majority population field*"

There is also an option "***Customize central business district***", where you can choose "*Whole area centroid*" or "*Customize centroid*". When you choose "*Whole area centroid*", the input layer acts like one whole polygon.



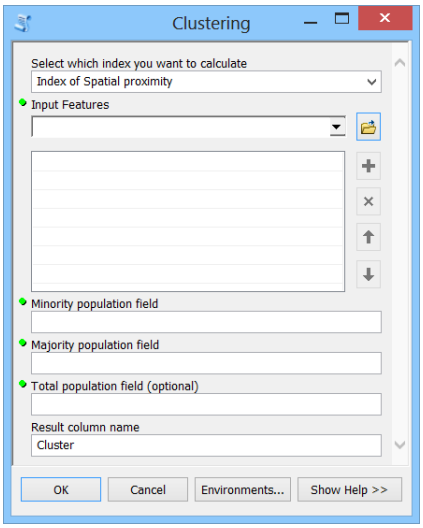
But when you select "*Customize centroid*" (**this option is available only for ONE input feature**), then you can select any polygon, in whole area, to be your central business district! Also two new boxes in interface will become available. The first box ("*Central business district filed*") is used to define the column ("KOD_ZSJ_P"), by which the desired polygon (or point) will be defined. And in the second box ("*Select central business polygon/point*") you will select that desired polygon (or point).

* - You may only choose number values (Double, Long, etc.)

2.6 Clustering (Massey, Denton, 1988)

Clustering measures the extent to which the area subunits with minority members are grouped together or clustered. A high degree of clustering indicates a racial community. To measure this dimension adequately requires a two-step process.

2.6.1 Index of Spatial Proximity



This measure is the average proximity between members of the same group and members of different groups.

The average proximity between members of the same groups is calculated by:

$$P_{xx} = \sum_{i=1}^{N}\sum_{j=1}^{N}\frac{x_i x_j c_{ij}}{X^2}$$

and the average proximity between members of different groups is calculated by:

$$P_{xy} = \sum_{i=1}^{N}\sum_{j=1}^{N}\frac{x_i y_j c_{ij}}{XY}$$

where c_{ij} represents a negative exponential of distance between area *i* and *j*, x_i the minority population in area subunit *i*, x_j the majority population of area subunit *j*, X and Y the total minority (majority) population of the entire area and N the total number of census tracts within the entire area.

$$SP = (XP_{xx} + YP_{yy})/TP_{tt}$$

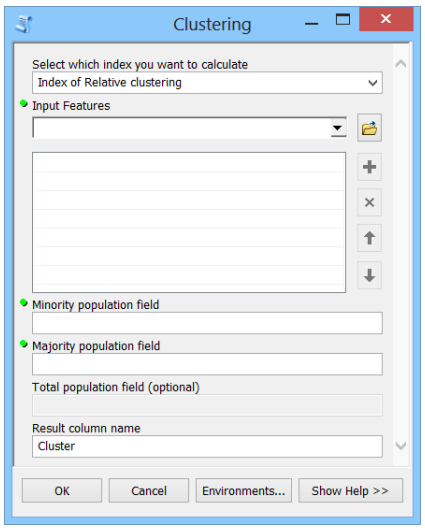
where T represents the total population and P_{tt} the proportion of the population that is minority.

In the user interface (*)

x	"*Minority population field*"
y	"*Majority population field*"
t	"*Total population field*"

If there is no differential clustering between X and Y, the index is **1.0**. The larger the number, the nearer the members of the same group live to each other.

2.6.2 Index of Relative Clustering



Using the results from the calculations for the index of spatial proximity, the following formula is applied to compare the average distance between the minority and majority members.

$$RLC = \frac{P_{xx}}{P_{yy}} - 1$$

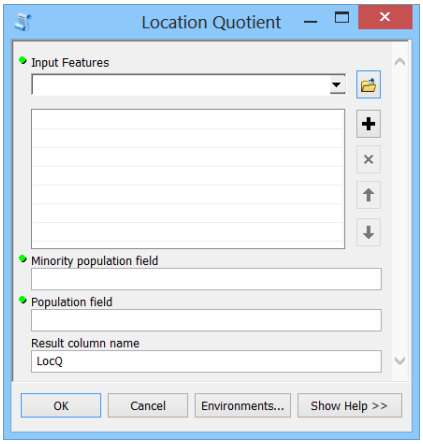
where P_{xx} and P_{yy} are defined as before.

When both groups have the same amount of clustering, the score will be **0.0**. A negative score indicates less clustering of the minority group as compared to the majority group, while a positive score means more clustering of minority group.

* - You may only choose number values (Double, Long, etc.)

2.7 Location quotient (Žítek, 2014)

Location quotient (LQ) is basically a way of quantifying how concentrated a particular demographic group is in a region as compared to the nation. It can reveal what makes a particular region "unique" in comparison to the national average.



$$LQ = \frac{x_i}{X} \Big/ \frac{y_i}{Y}$$

where x_i is the minority population in subunit *i*, X the total minority population in the entire area, y_i the total population in subunit *i*, Y the total population in the entire area.

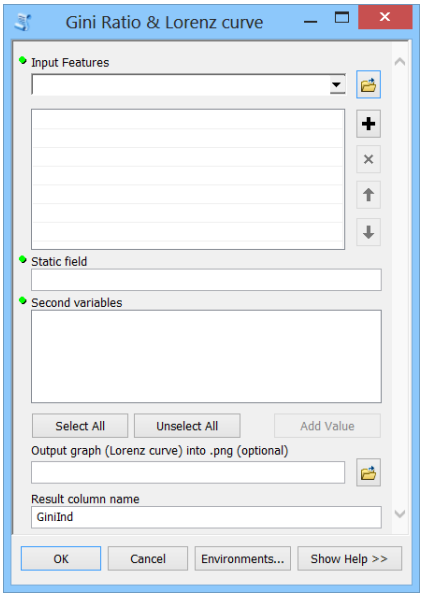
LQ of **1.0** means that the subunit and the nation are equal in the particular demographic group, while an LQ of more than **1.0** means that the subunit has a higher concentration of particular demographic group that the nation.

In the user interface (*)

x	"*Minority population field*"
y	"*Population field*"

2.8 Gini index and Lorenz curve (Siegel, Swanson, 2004)

The Gini concentration ratio (Gini index) measures the degree of inequality or the size of the gap. When this ratio is **1.0**, then it indicates complete inequality, with all population located in one part of a country and no population in the remaining areas. Ratio of **0.0** indicates a perfect distribution of population in all areas of the country.



$$GI = \left(\sum_{i=1} X_i Y_{i+1}\right) - \left(\sum_{i=1} X_{i+1} Y_i\right)$$

where X_i is the proportion of population (static) in the area and Y_i is the proportion of population (multi-variable) in the area.

Lorenz curve is a graphic device for representing the inequality of two distributions. Such inequality of distributions is represented by a diagonal line. This diagonal line is compared to the actual distribution and the gap between the ideal and actual lines is interpreted as the degree of inequality. You can generate Lorenz curve, by using last box in interface.

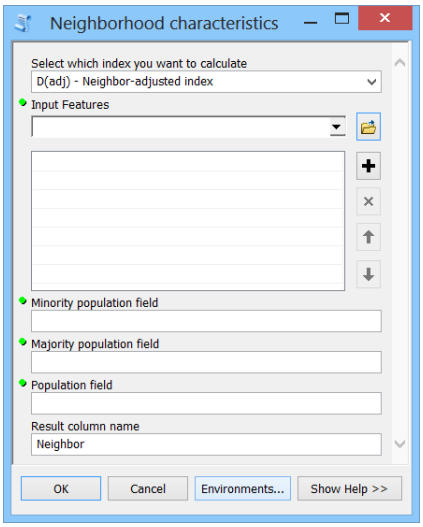
In the user interface (*)

x	"*Static field*"
y	"*Second variables*"

* - You may only choose number values (Double, Long, etc.)

2.9 Neighborhood characteristic (Dissimilarity index adjusted) (Wong, 2003)

It is true that D (Dissimilarity) is effective to capture the evenness of population, but only to the extent that the spatial arrangement of population is not considered. When one group is dominating in each subunit in study area, then D (Dissimilarity) index will return a **1.0**, indicating a perfect segregated study area, even if some adjacent subunits are occupied by different groups.



2.9.1 Neighbor-adjusted D(adj) index

This index introduced by Morrill (1991) is the original dissimilarity index less the amount potential interaction between different groups across subunit boundaries. The level of potential interaction between any pair of neighboring units is then determined by the differences in the racial mixes of neighboring subunits.

$$D(adj) = D - \frac{\sum_i \sum_j \left| c_{ij} * \left(\frac{x_i}{t_i} - \frac{x_j}{t_j} \right) \right|}{\sum_i \sum_j c_{ij}}$$

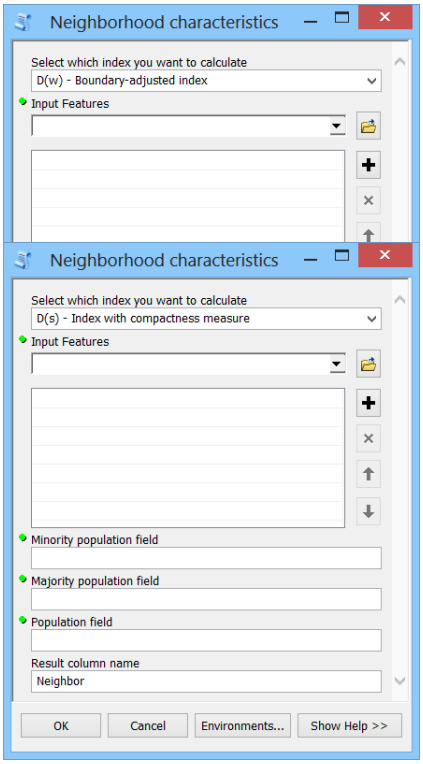
where D is defined as before in chapter 2.2, x_i and x_j are minorities of population in area subunits *i* and *j*, t_i and t_j the total populations in subunits *i* and *j*, while c_{ij} will be zero if *i* and *j* are not neighbors and one if they are.

In the user interface (*)

z_i	"*Minority population field*"
z_j	"*Majority population field*"

2.9.2 Boundary-adjusted D(w) index

Based upon the premise that the intensity of interactions across a boundary is not a simple function of adjacency, but likely the length of the shared boundary, the D(adj) index is slightly rewritten to incorporate a boundary-length component to moderate the interactions across subunits in area (Wong, 1993).



$$D(w) = D - \frac{1}{2} \sum_i \sum_j w_{ij} \left| \frac{x_i}{t_i} - \frac{x_j}{t_j} \right|$$

where all terms are defined as before and $w_{ij} = \frac{d_{ij}}{\sum_j d_{ij}}$

where d_{ij} is the length of the shared boundary between areal units *i* and *j*, and the denominator is basically the total length of boundary for subunit *i*.

2.9.3 Index D(s) with compactness measures

To incorporate the geometric characteristics of areal units into the evaluation of segregation, a compactness measure based upon the perimeter-area ratio was used.

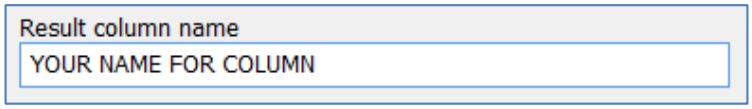
$$D(w) = D - \frac{1}{2} \sum_i \sum_j w_{ij} \left| \frac{x_i}{t_i} - \frac{x_j}{t_j} \right| * \frac{\frac{1}{2}\left[(P_i/A_i) + (P_j/A_j) \right]}{MAX(P/A)}$$

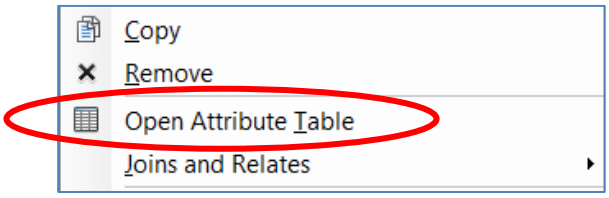
where Pi/Ai is the perimeter-area ratio for areal unit i, and MAX(P/A) is the maximum perimeter-area ratio among all the subunits in the study region.

* - You may only choose number values (Double, Long, etc.)

3 Where you can find the results

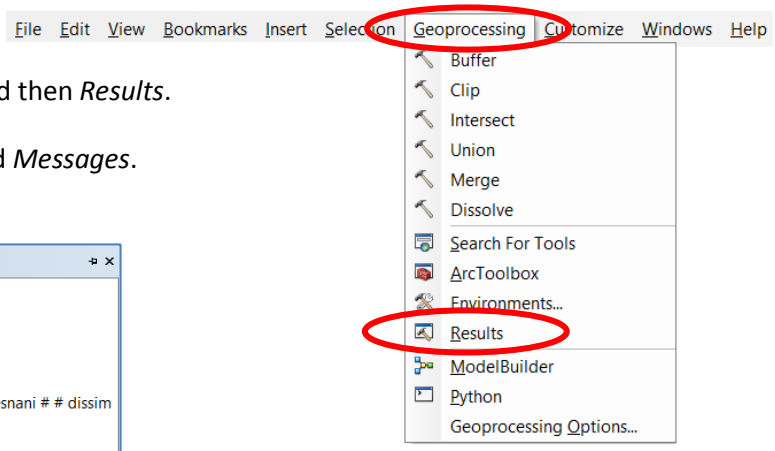
3.1 Attribute table





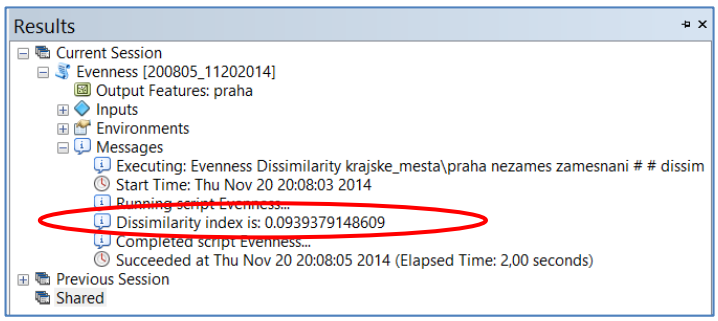
Simply open attribute table of calculated layer and find your column that you named in interface of selected index ("*Result column name*").

3.2 The result window



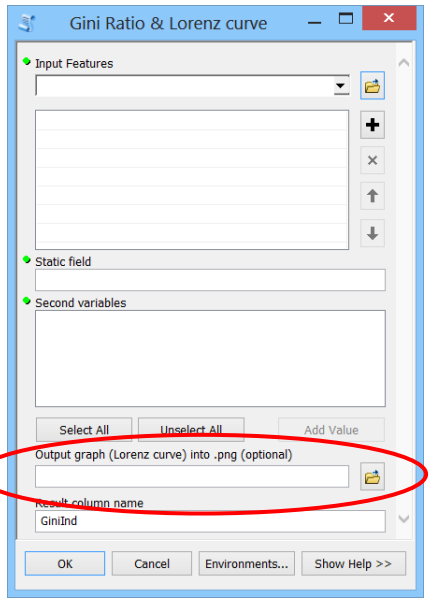
You can find this window in *Geoprocessing* and then *Results*.

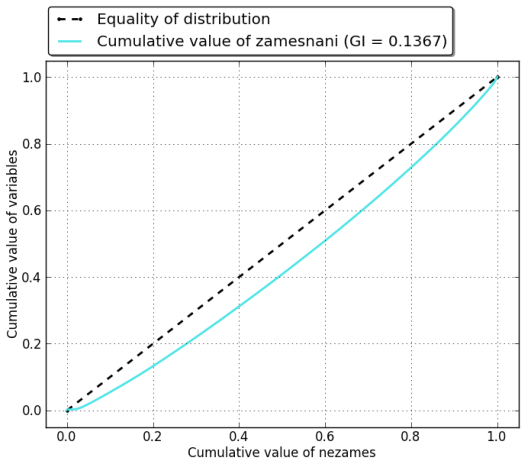
Your result is hiding under *Current Session* and *Messages*.



3.3 Gini index and lorenz curve

Because you can have multiple results in the result window (see chapter before) or if you fill "*Output graph*", you can find Gini index and also Lorenz curve in the .png format, in the export location that you choose.





This is the resulting graph of Lorenz curve and in the brackets you can also see the Gini index (GI=0.1367).

References

MASSEY, D. S., DENTON, N. A. (1988). The dimensions of residential segregation. Social Forces, 67, 281–315.

SIEGEL, J., SWANSON, D., (2004). The methods and materials of demography (2nd edition). Elsevier Academic Press, London., 114-123

SÝKORA, L., TEMELOVÁ, J., eds (2005). Prevence prostorové segregace. Praha Univerzita Karlova v Praze, Přírodovědecká fakulta, katedra sociální geografie a regionálního rozvoje, Centrum pro výzkum měst a regionů a Ministerstvo pro místní rozvoj., 5-17

WONG, D. W. S. (2003). Implementing spatial segregation measures in GIS. Computers, Environment and Urban Systems, Vol. 27, 53-70.

ŽÍTEK, V. (2014). Specializace, prostorová koncentrace a diverzifikace v českých a slovenských krajích, XVII. mezinárodní kolokvium o regionálních vědách, Sborník příspěvků, Masarykova univerzita, 326-331, doi:10.5817/CZ.MUNI.P210-6840-2014-40.

Bc. David BAROŠ

Institute of Geoinformatics

Faculty of Mining and Geology

VŠB - Technical University

17. Listopadu 15/2172

708 33 Ostrava - Poruba

Czech Republic