Spatially aware Data Classification for Choropleth Maps

Esri International User Conference, July 15th 2014
Christoph Traun
What a choropleth map should do…

Choropleth maps should

- provide an overview of the statistical distribution of the data.
- allow to identify approximate attribute values for each location, and
- provide an image of the spatial configuration of the mapped phenomenon by focusing on boundary lines between different classes/shadings.
What data-classification should do...

Data classification should

- reduce (map) complexity to enhance readability/perception
- enable the unambiguous visual relation between a map feature shading and the corresponding value (range).
- enable the comparison of multiple datasets (time series, ...), and
- help to properly reflect data inherent spatial patterns
What data classification does...

Classification routines implemented in GIS and mapping software

- group data into classes based on the distribution of attribute values along the number line
  - Equal Interval
  - Quantile
  - Optimal (natural breaks)
  - ...

- ...and completely ignore the spatial context of data.
There is a gap!

Choropleth maps should clearly communicate the main spatial patterns. They should be optimized in this respect.

Classification should help to properly reflect data inherent spatial patterns and reduce map complexity.

Classification ignoring spatial context
Autocorrelation based Regioclassification – Step 1

1. Z-standardizing all values:

   \[ z_i = \frac{x_i - \text{mean}}{\text{standard deviation}} \]

   Overall Mean (n= 755): \textbf{24.1}
   Overall Std.Dev.: \textbf{19.3}

2. Calculate mean value of local neighborhood (adjacent polygons):

   \[ (-0.37 + -0.53 + -0.17 + -0.25)/4 = -0.33 \]

   (Example for polygon \( z = -0.07 \))
Step 2: Projection of all points onto regression line (detail view)

Step 3: Determination of class breaks based on the distribution of projected, standardized value pairs

index value (= distance from origin along the regression line derived from projection of standardized value pairs)
LISA outlier → conventional classification → 1st class
Ressources and further research

- **Download ArcGIS AddIn** (Author: Christoph Mayrhofer, University of Salzburg)

- **Publications:**

- **Current research** (publication in finalization): Use of autocorrelation in space and time to reduce spatial and transitional complexity in visualisation of time series data (small multiples & map animation)