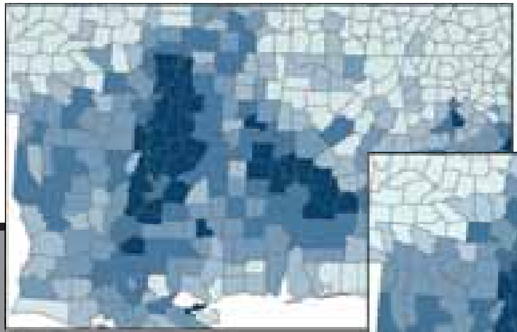
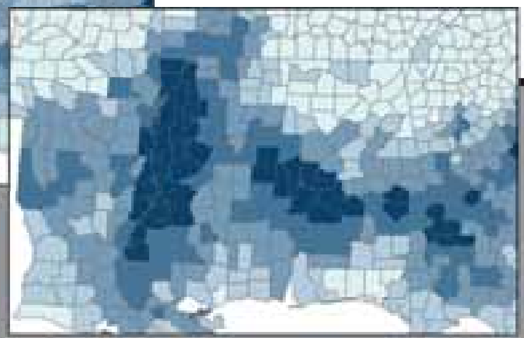


Jenks Classification



Autocorrelation based Regioclassification

# Spatially aware Data Classification for Choropleth Maps

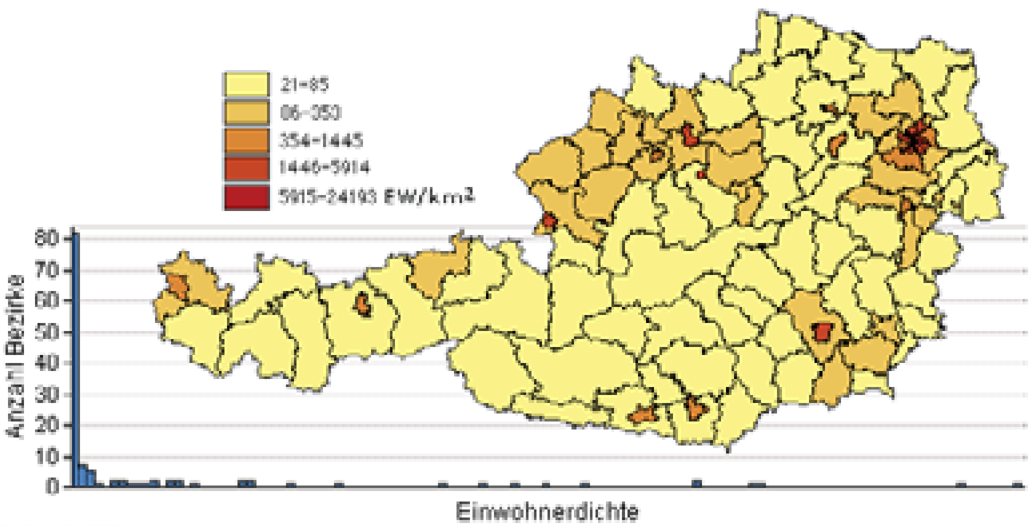
Esri International User Conference, July 15th 2014

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.**

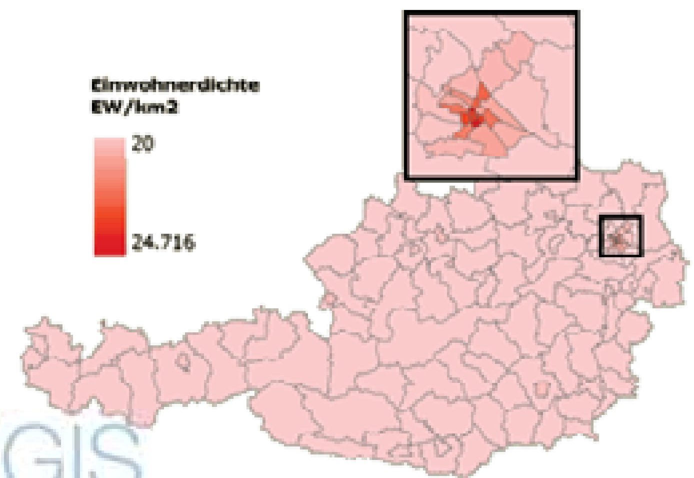


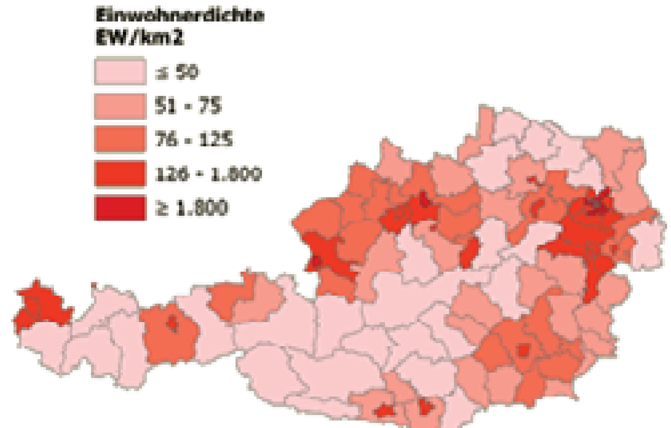

| BEZIRK | Dichte |
| --- | --- |
| AMSTETTEN <305> | 351 |
| BADEN <306> | 167 |
| BLUDENZ <801> | 47 |
| BRAUNAU AM INN <404> | 91 |
| BREGENZ <802> | 140 |
| BRUCK AN DER LEITHA <307> | 80 |
| BRUCK AN DER MUR <602> | 49 |
| DEUTSCHLANDSBERG <603> | 71 |
| DORNBIRN <803> | 437 |
| EFERDING <405> | 118 |

# 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

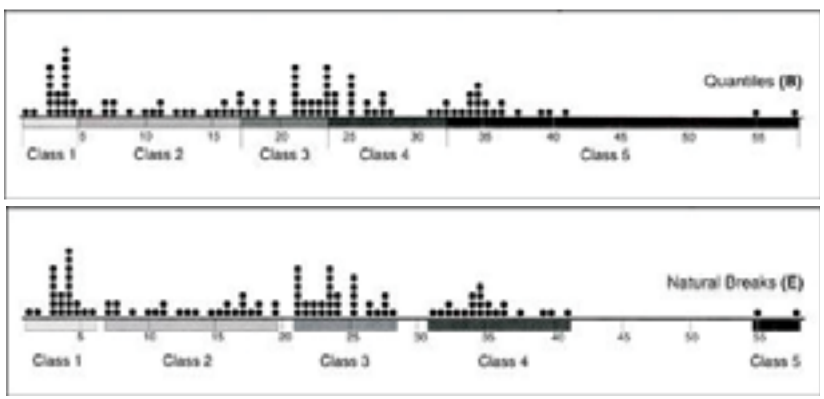


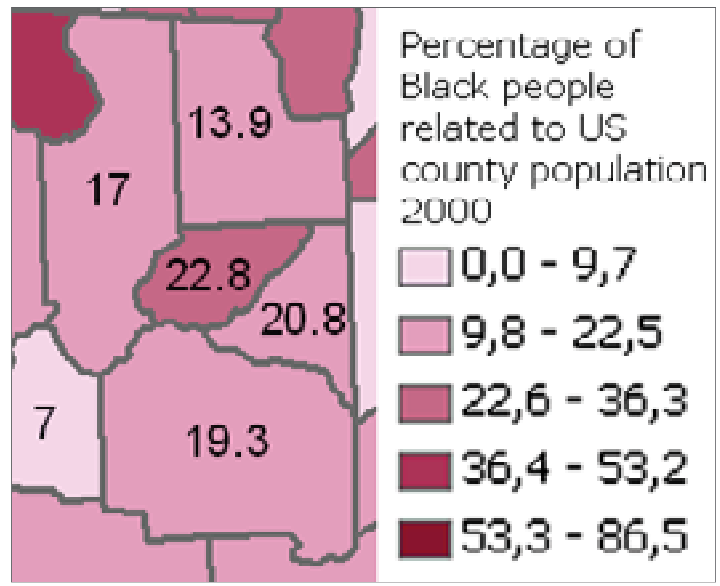



# What data classification **does**…

Classification routines implemented in GIS and mapping software

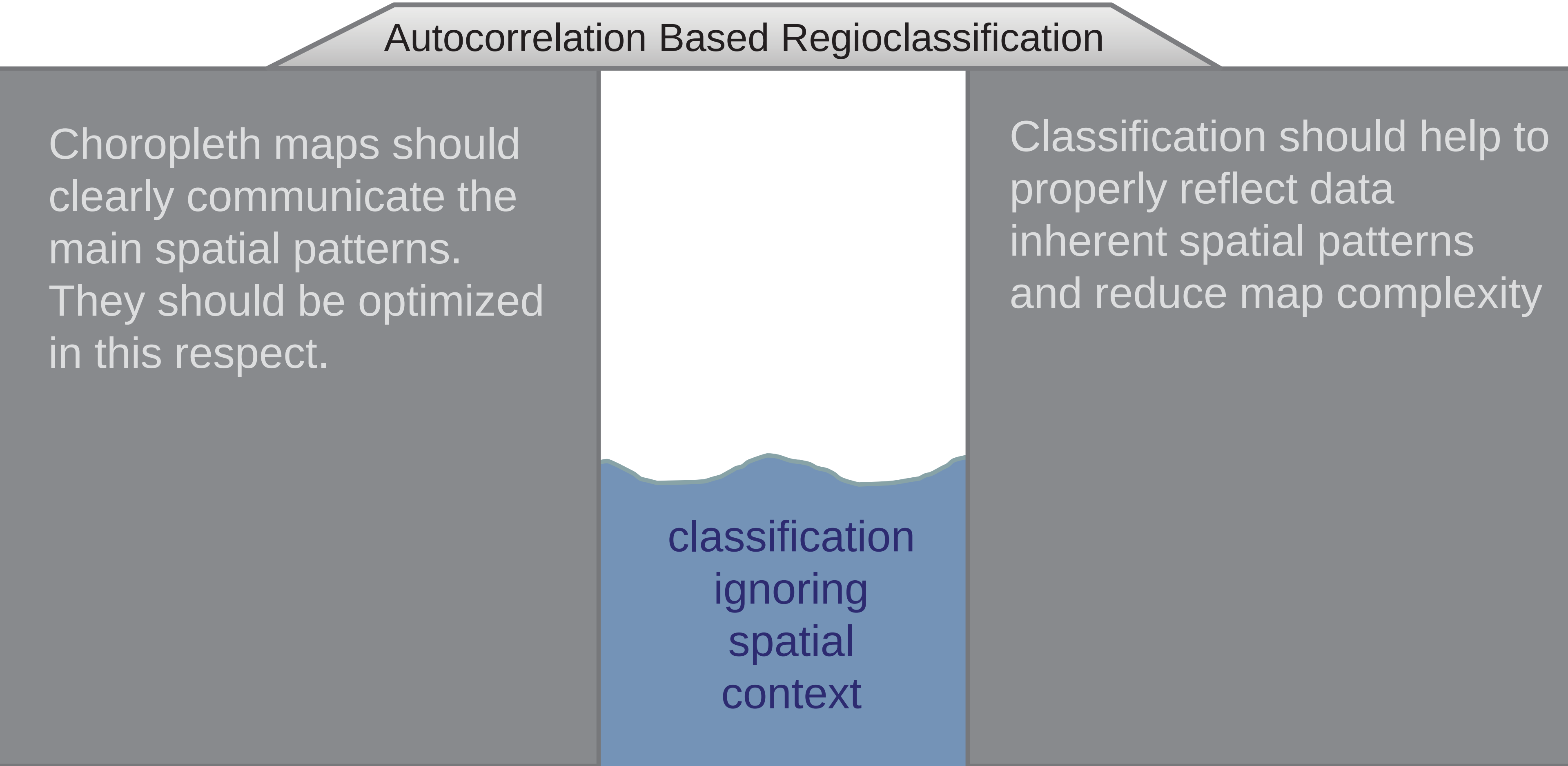
- group data into classes based on the distribution of attribute values along the number line
  - Equal Interval
  - Quantile
  - Optimal (natural breaks)
  - …




- …and completely ignore the spatial context of data.

# There is a gap!

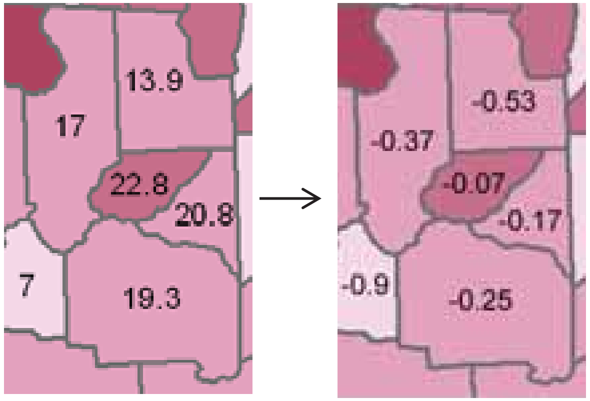


# Autocorrelation based Regioclassification – Step 1

[Moran's I Scatterplot: see Luc Anselin 1995]

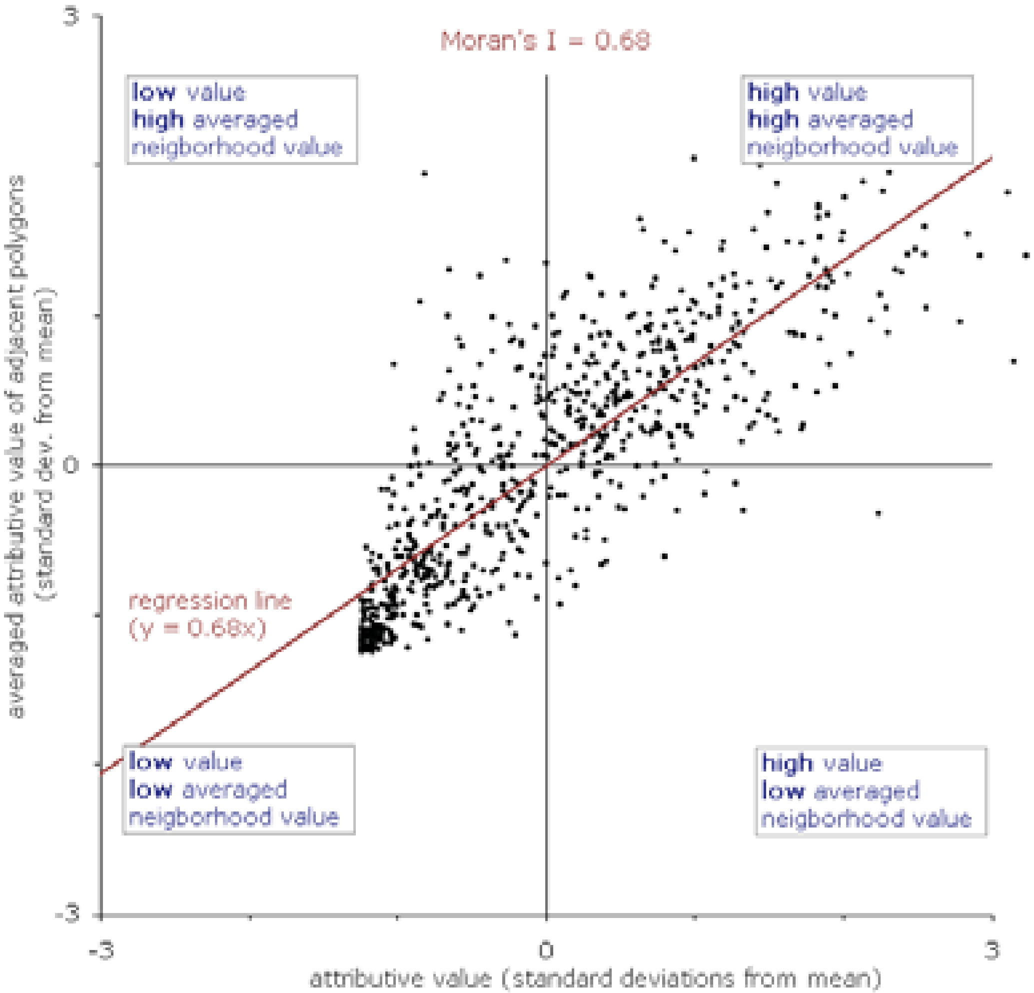
1. z-standardizing all values:
   $z_i = (x_i$ – mean)/standard deviation

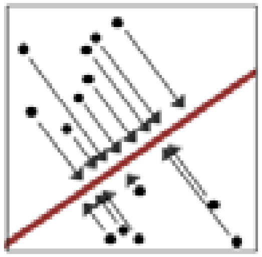
   Overall Mean (n= 755): **24.1**
   Overall Std.Dev.: **19.3**



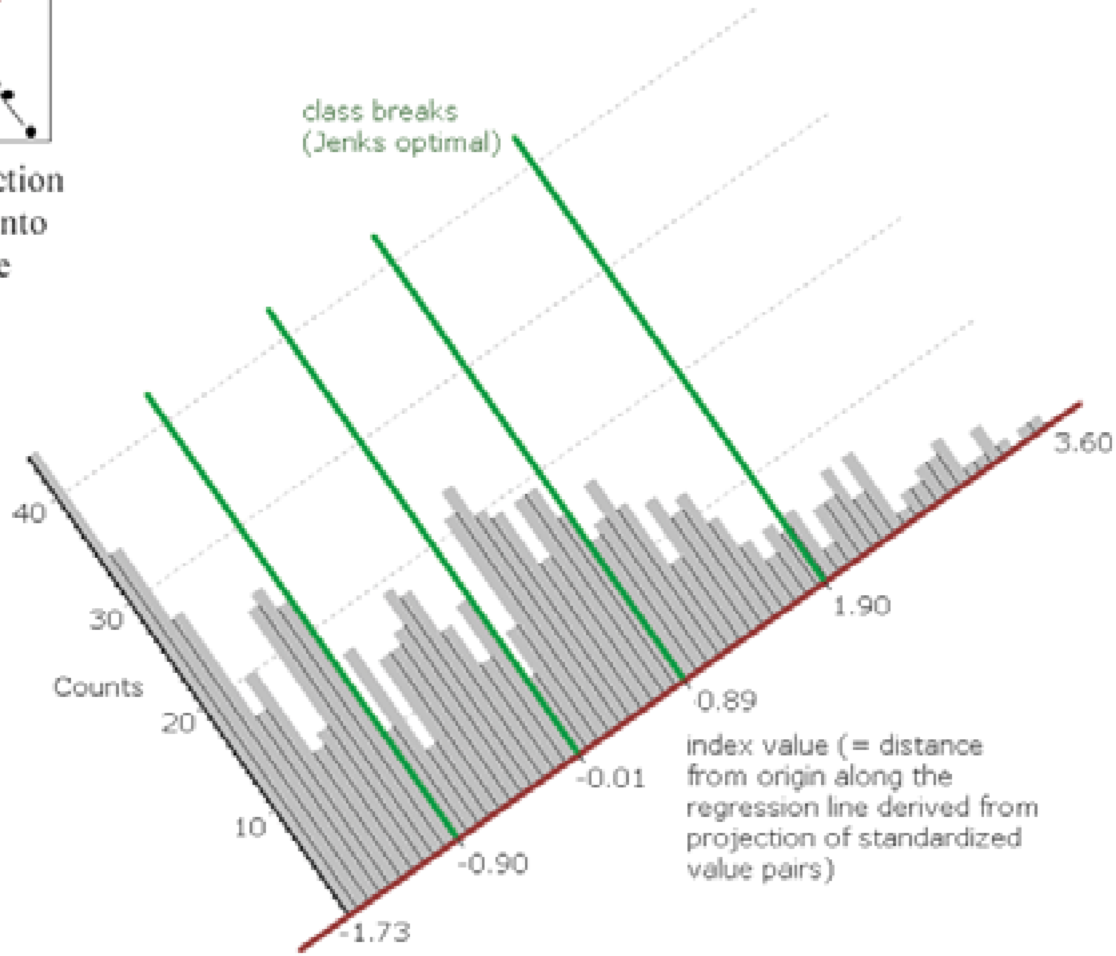

2. Calculate mean value of local neigborhood (adjacent polygons):

   (-0.37+ -0.53 + -0.17 + -0.25)/4 = **-0.33**
   (Example for polygon z = -0.07)



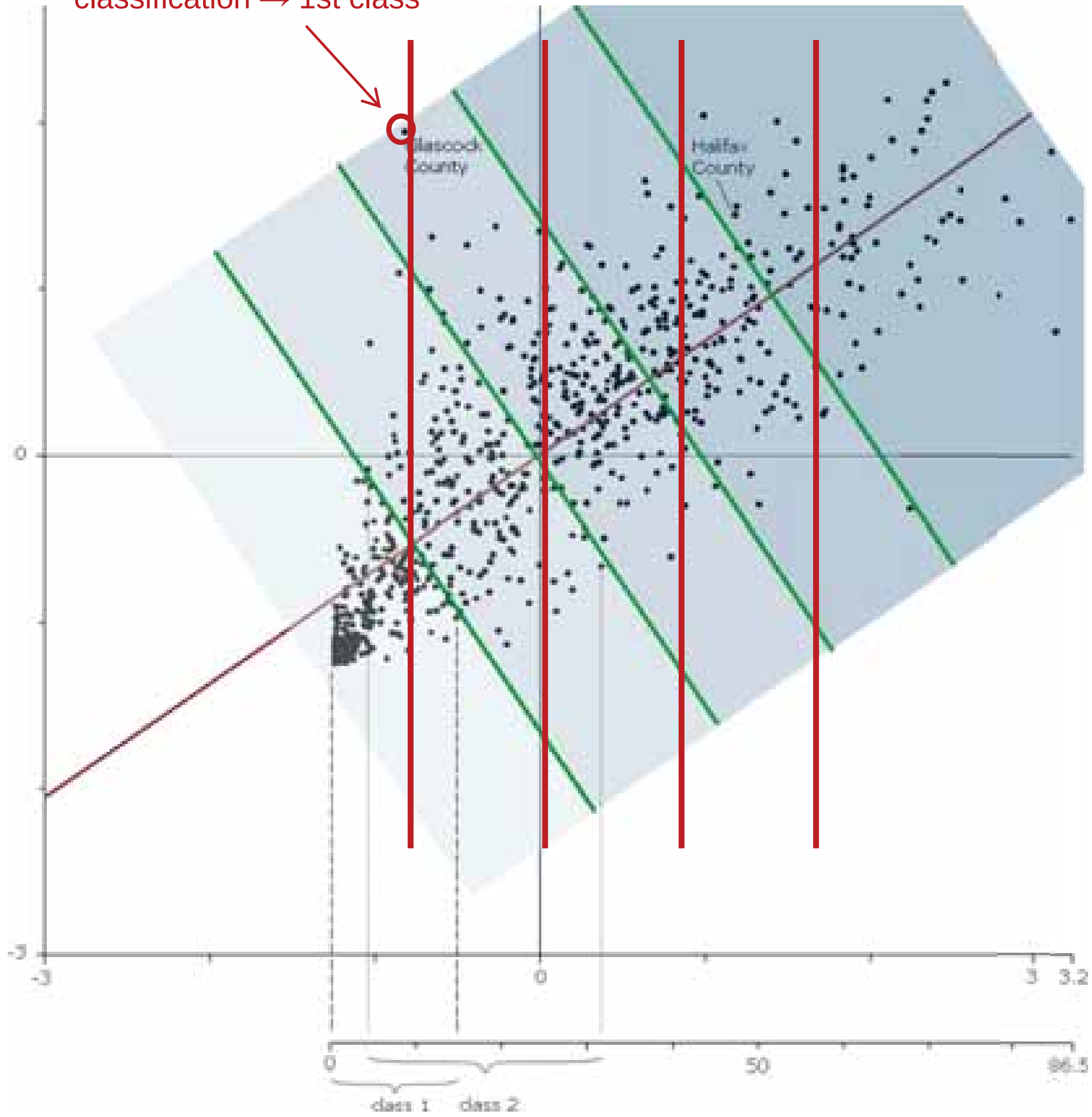


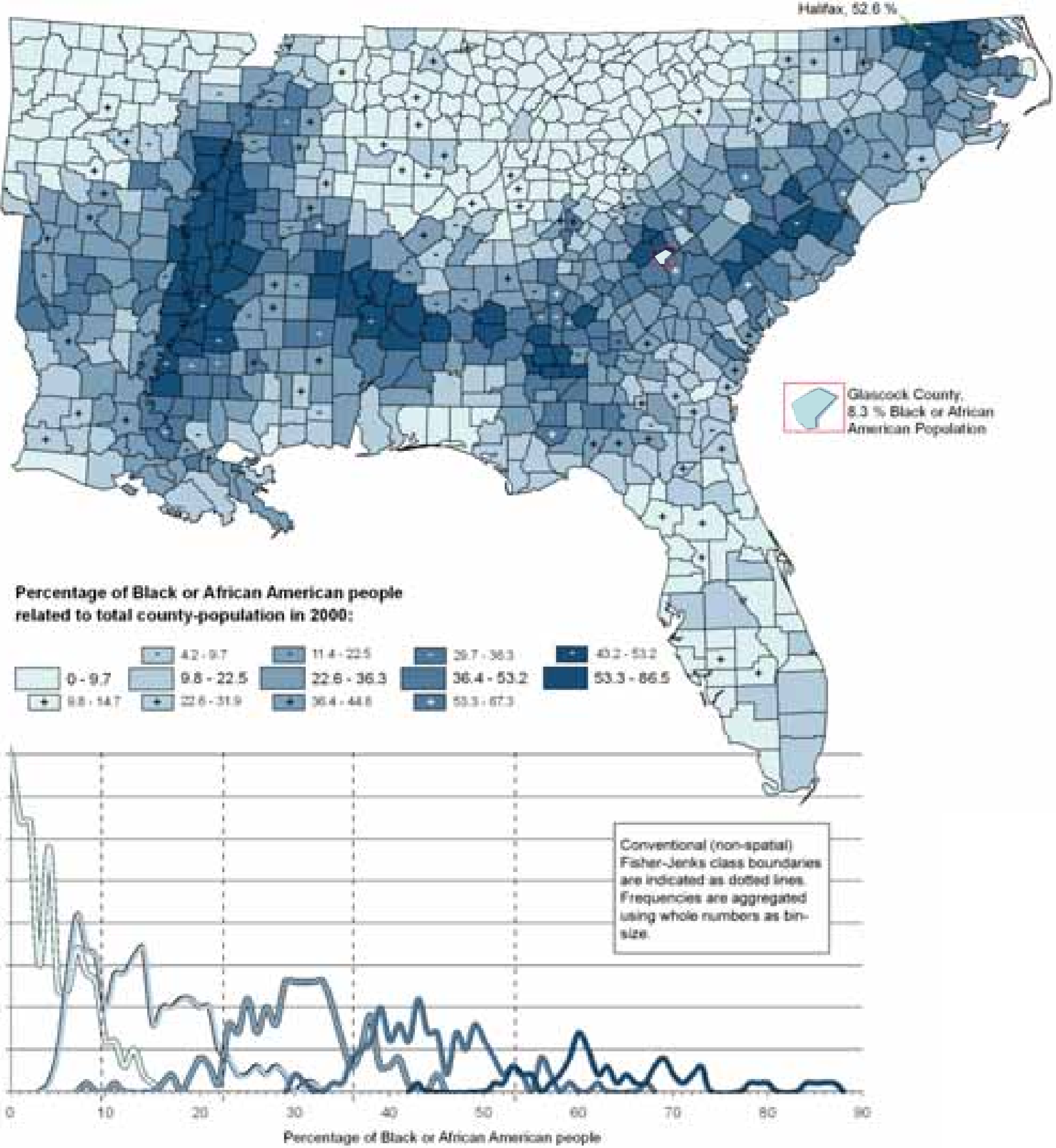
**Step 2:** Projection of all points onto regression line (detail view)

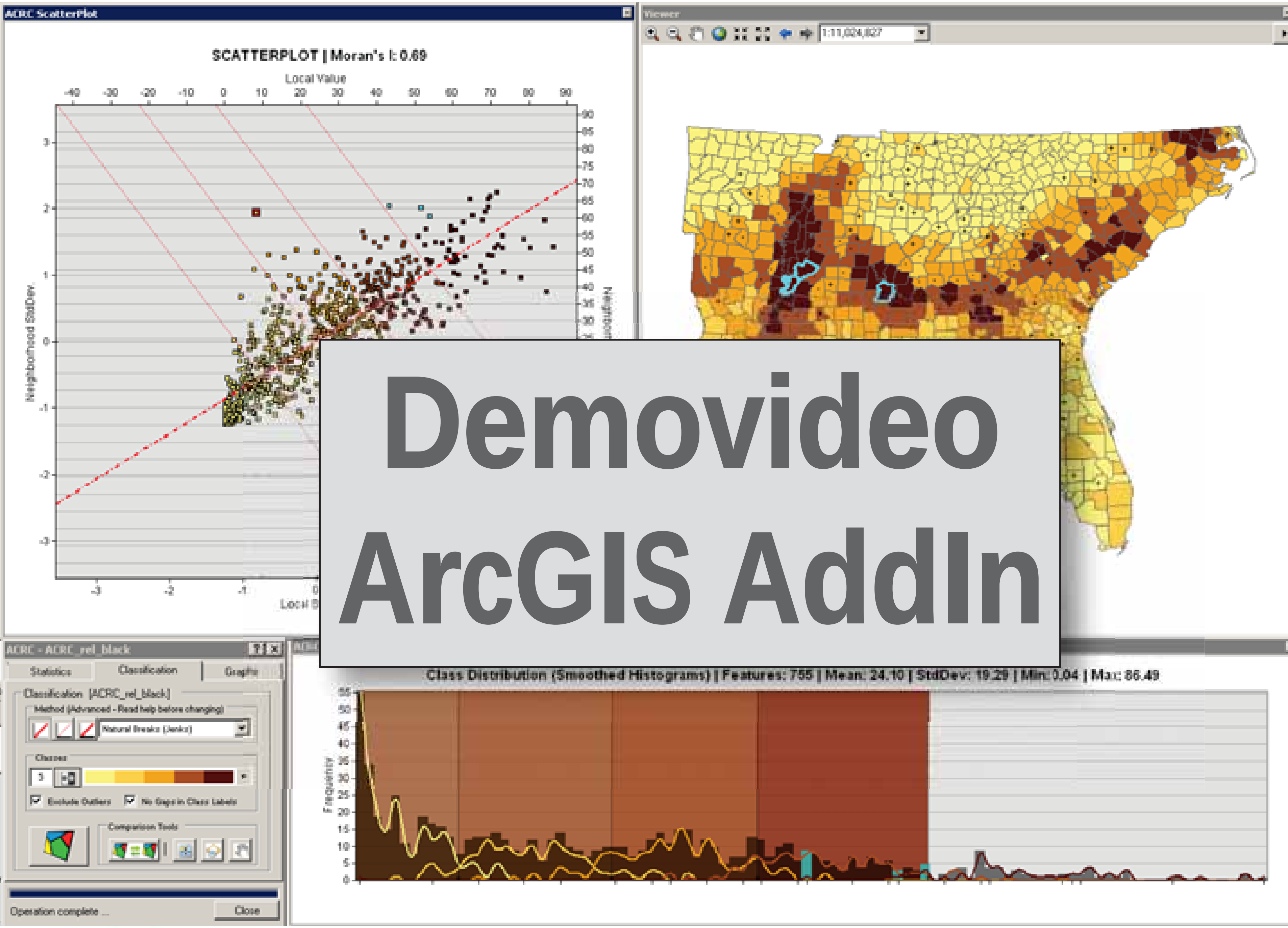



**Step 3:** Determination of class breaks based on the distribution of projected, standardized value pairs

LISA outlier → conventional classification → 1st class

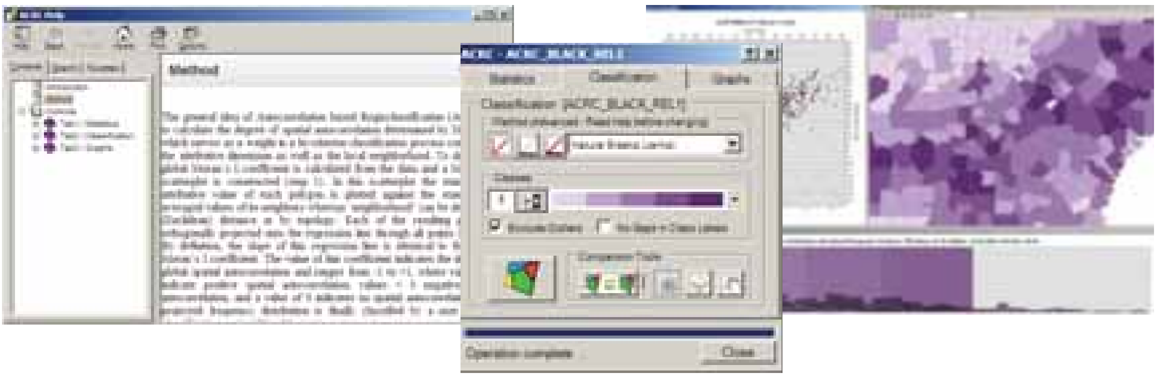



Halifax, 52.6 %
Glascock County, 8.3 % Black or African American Population
Percentage of Black or African American people related to total county-population in 2000:
4.2 - 9.7
11.4 - 22.5
29.7 - 36.3
43.2 - 53.2
0 - 9.7
9.8 - 22.5
22.6 - 36.3
36.4 - 53.2
53.3 - 86.5
9.8 - 14.7
22.6 - 31.9
36.4 - 44.8
53.3 - 67.3
Frequency per class (smoothed)
0
5
10
15
20
25
30
35
40
Conventional (non-spatial) Fisher-Jenks class boundaries are indicated as dotted lines. Frequencies are aggregated using whole numbers as bin-size.
0
10
20
30
40
50
60
70
80
90
Percentage of Black or African American people


Demovideo
ArcGIS AddIn
ACRC ScatterPlot
SCATTERPLOT | Moran's I: 0.69
Local Value
Neighborhood StdDev.
Viewer
ACRC - ACRC_rel_black
Statistics
Classification
Graphs
Classification [ACRC_rel_black]
Method (Advanced - Read help before changing)
Natural Breaks (Jenks)
Classes
Exclude Outliers
No Gaps in Class Labels
Comparison Tools
Operation complete ...
Close
Class Distribution (Smoothed Histograms) | Features: 755 | Mean: 24.10 | StdDev: 19.29 | Min: 0.04 | Max: 86.49
Frequency

# Ressources and further research

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



- Publications:


  - Traun, C. & Loidl, M. (2012): Autocorrelation-Based Regioclassification - A self-calibrating classification approach for choropleth maps explicitly considering spatial autocorrelation. In: International Journal of Geographical Information Science 26(5), p.923-939.
  - Loidl, M. & Traun, C. (2013): The Effect of ACRC on the Results of Cartographic Classification Depending on Spatial Autocorrelation. In: International Journal of Geoinformatics 9(2), p. 29-36.


- 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)

Z_GIS