Classification II

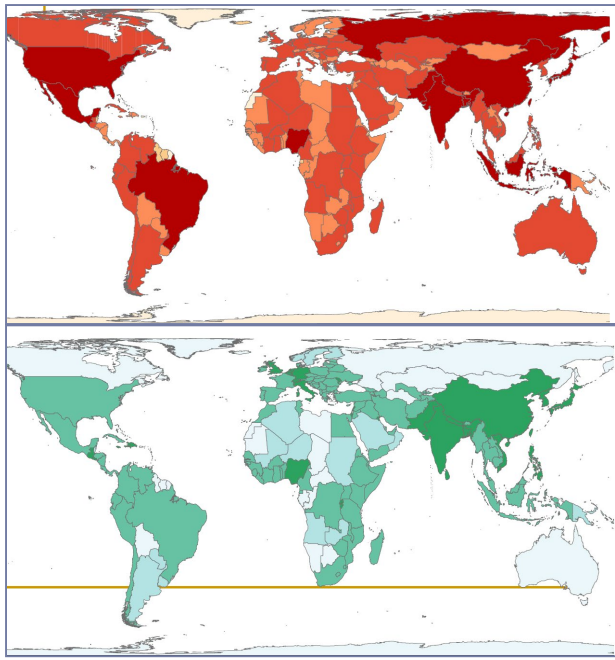
GEOG380 FA 2018

Contents

- Color schemes
- Legend design and symbology
- Classification methods

If you really want a choropleth map to represent classification of data…

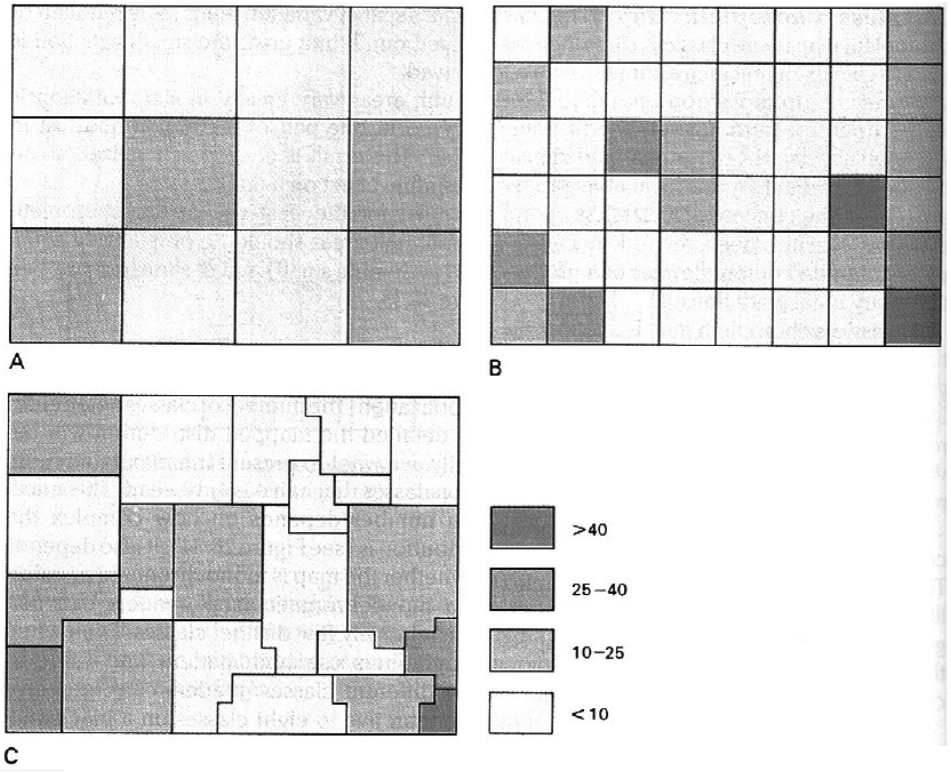
- Standardization (or normalization) is often necessary…
  - to make data in either ratio or rate form
  - to account for varying sizes of data unit
    - Divided by areas
      - Ex) yield in bushels per acres
    - Ratio of two totals
      - Ex) ratio of harvested to planted acres



- Also, typical statistical values for the units can be used
  - Mean, median, std. dev., …

Size and Shape of Unit Areas

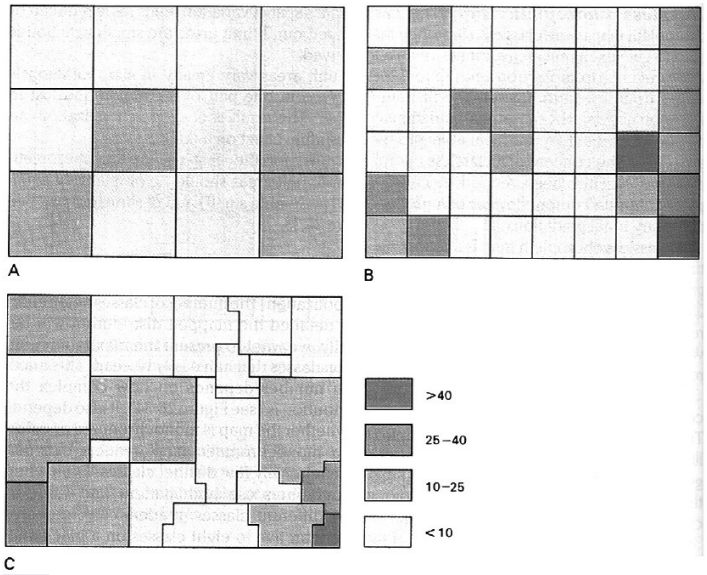
- Large unit areas tend to reduce (or average out) spatial variation
  - If unit areas vary greatly in size, spatial variation is preserved only in part of the region
  - It is hard to differentiate symbols of small units
- Ideally, small unit sizes and similar unit shapes are preferred
- ➔ Modifiable Areal Unit Problem (MAUP) (next slide)



Openshaw S, Alvandies S 1999 Applying geocomputation to the analysis of spatial distributions. In Longley P, Goodchild M, Maguire D, Rhind, D (eds) *Geograhpic Information Systems: Principles and Technical Issues*. Vol 1, 2nd ed. New York: John Wiley and Sons Inc.

MAUP (Modifiable Areal Unit Problem)

- Assumption: data is evenly distributed across space
- Maps draw attention to larger areas
- Maps may not represent actual underlying spatial phenomenon
  - Ex) distributions of individual observations, directions, densities…
- Can have a dramatic effect on trends or patterns based on size
  - Information may change or disappear
- The unit of the phenomena influences the spatial patterns on maps! (Openshaw and Alvanides, 1999; 1997)

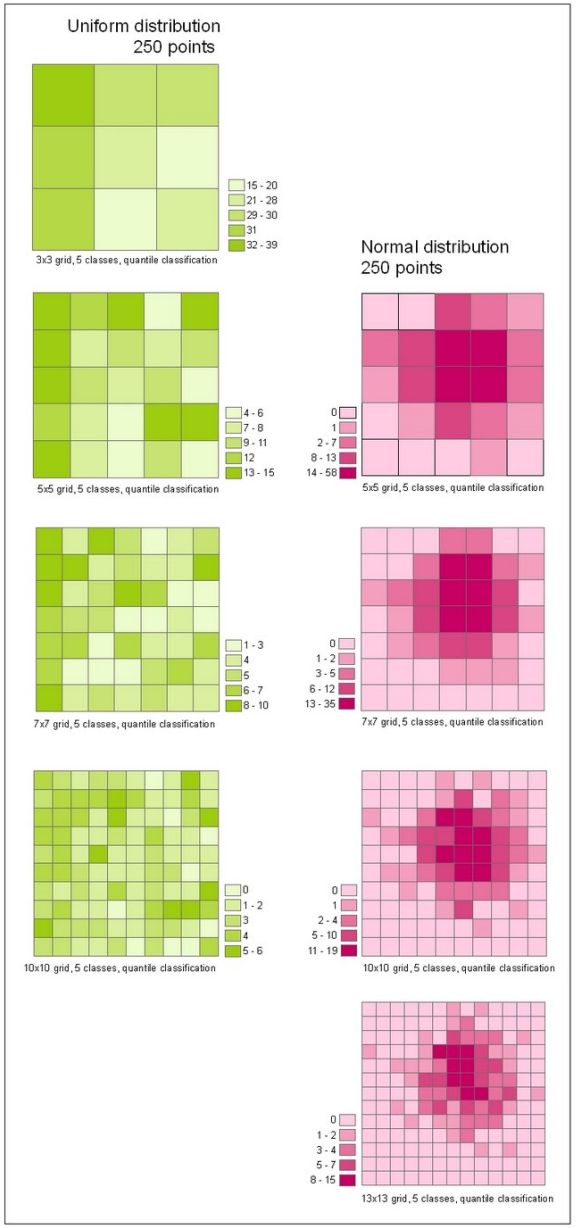


Openshaw S, Alvandies S 1999 Applying geocomputation to the analysis of spatial distributions. In Longley P, Goodchild M, Maguire D, Rhind, D (eds) *Geograhpic Information Systems: Principles and Technical Issues*. Vol 1, 2nd ed. New York: John Wiley and Sons Inc.

Openshaw,Stan & Alvanides, Seraphim; (1997);"Designing Zoning Systems for Representation of Socio-Economic Data".

MAUP (cont.)

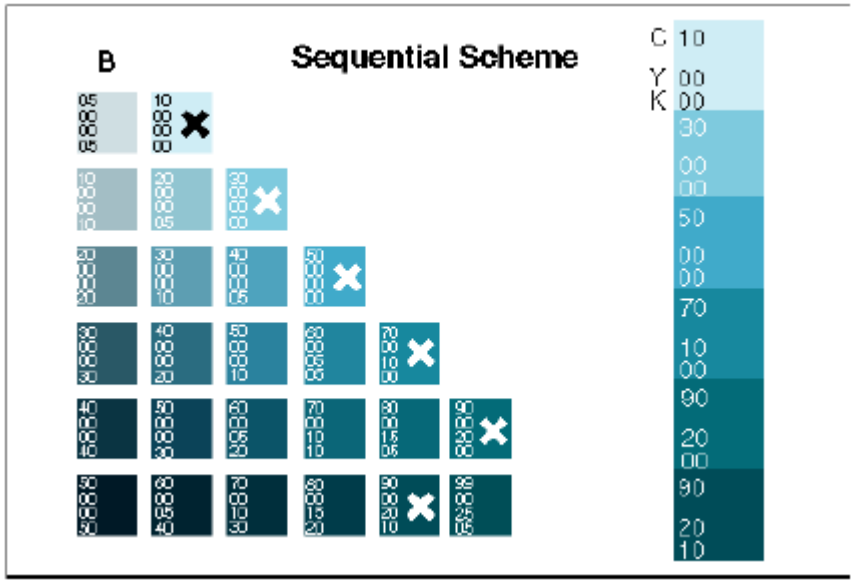
- Results from thematic mapping are sensitive to aggregation of boundaries
  - Census track, block, block group, county, state, country…
- Data assigned to these boundaries are affected by shape and size
- Purpose of these boundaries can also influence the analysis and interpretation of results in maps
  - Appropriate boundaries for the themes
  - Ex. Change of voting precincts



http://bushfiresandbsq.blogspot.com/2010_05_01_archive.html

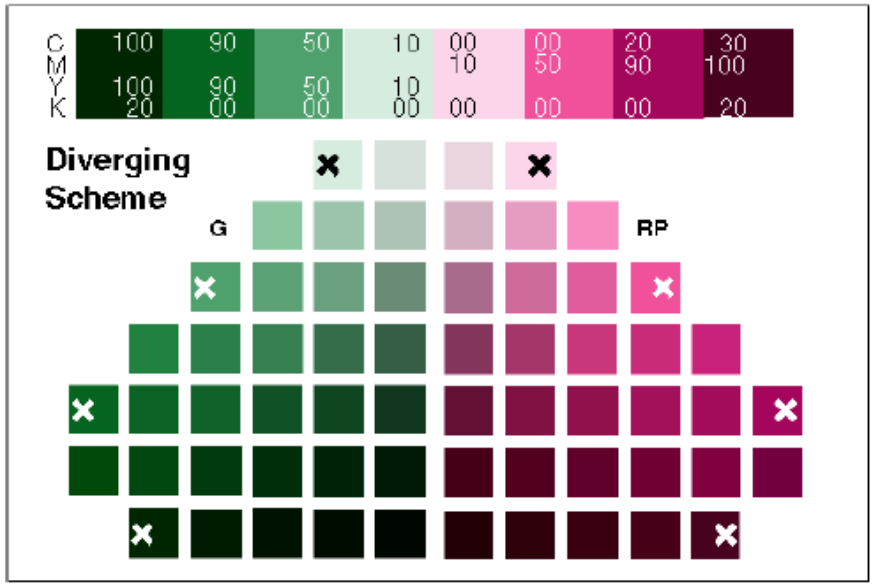
Color Scheme and Classification

- Sequential color scheme
  - Logically arranged from high to low, and should be represented by sequential lightness steps
  - Low data values are usually represented by light colors and high values by dark colors
  - Transitions between hues may be used in a sequential scheme, but the light-to-dark progression should dominate the scheme



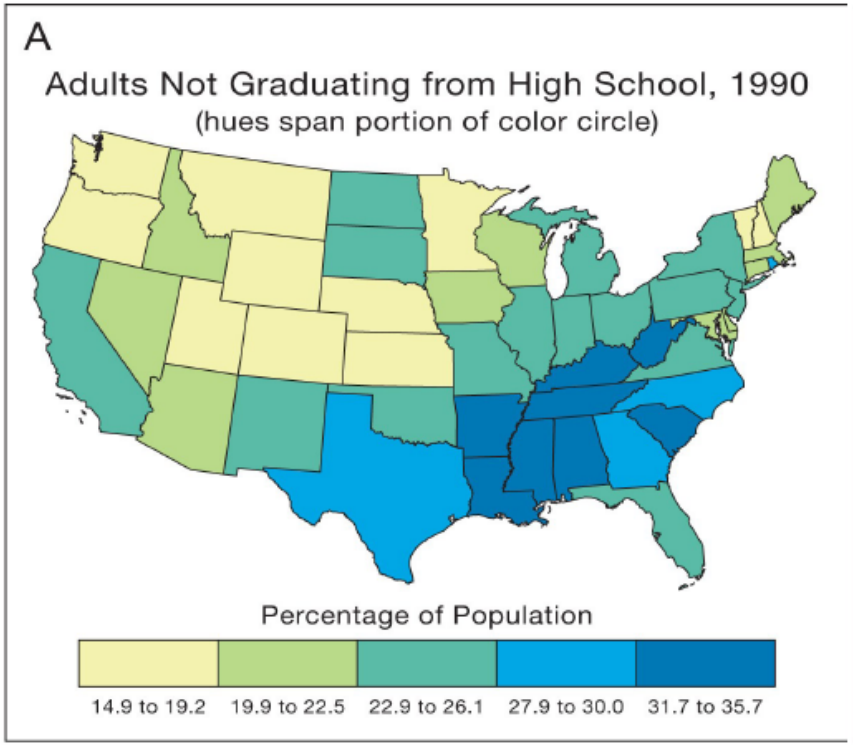
Color Scheme and Classification (cont.)

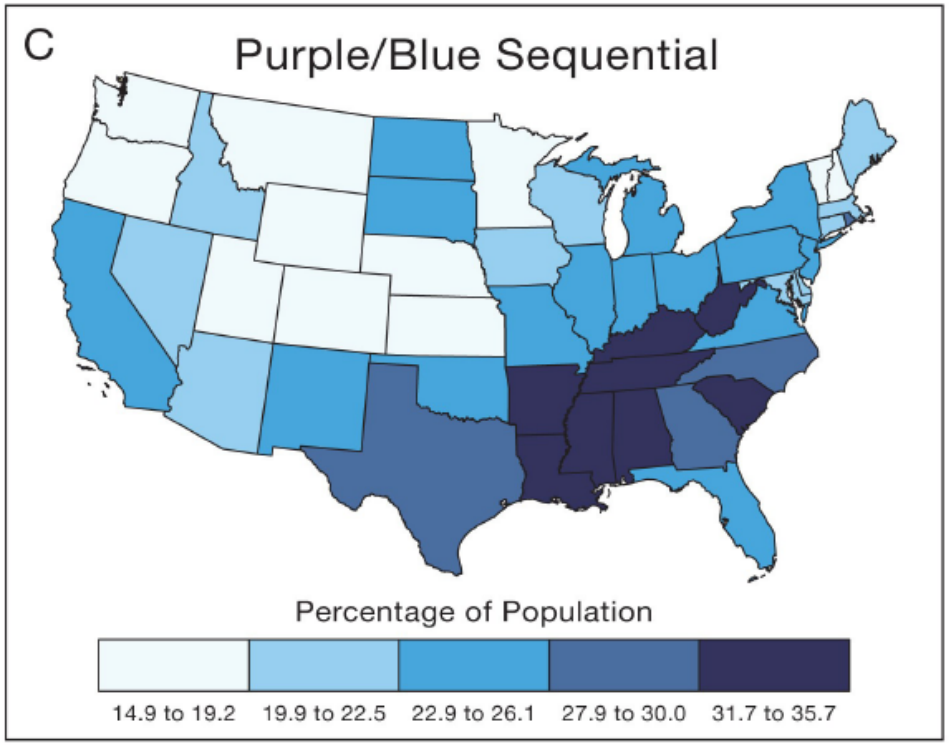
- Diverging color scheme
  - Emphasizes progressions outward from a critical-midpoint of the data range
    - Ex. Temperature, sea-level, standard deviation, …
  - Typically uses two different hues that diverge from a common light hue or gray, for the critical midpoint, toward dark colors of different hues at each extreme



Sequential Color Scheme

- It is acceptable if hues span a continuous portion of color wheel (but not the entire wheel)





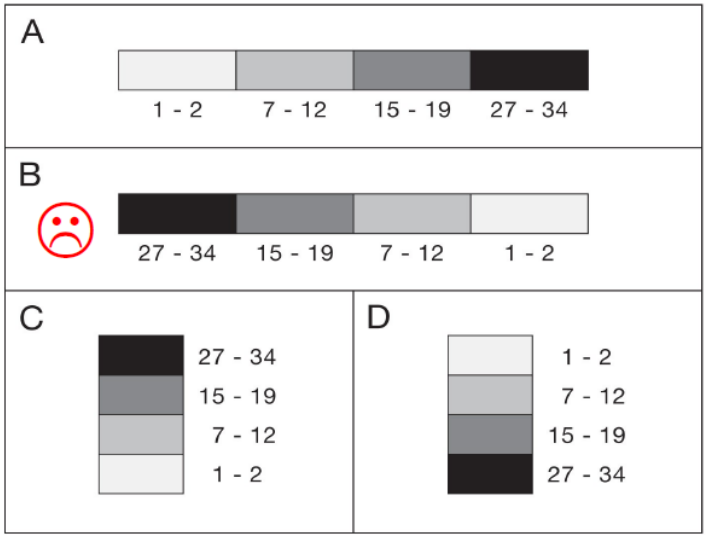
(source: Slocum et al. 2009, http://www.tigercolor.com)

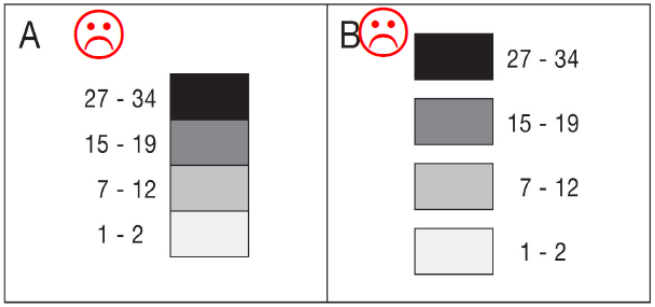
Diverging Color Schemes

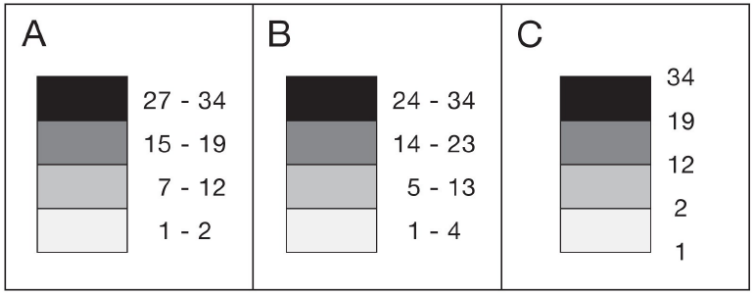
(source: Slocum et al. 2009 , http://www.tigercolor.com)

Legend Design

- Horizontal vs. vertical
- Vertical legend:
  - High values at the top or bottom
- Labels either at the bottom or to the right of boxes
- Contiguous vs. separated between class range values
- Gaps vs. no gaps between boxes

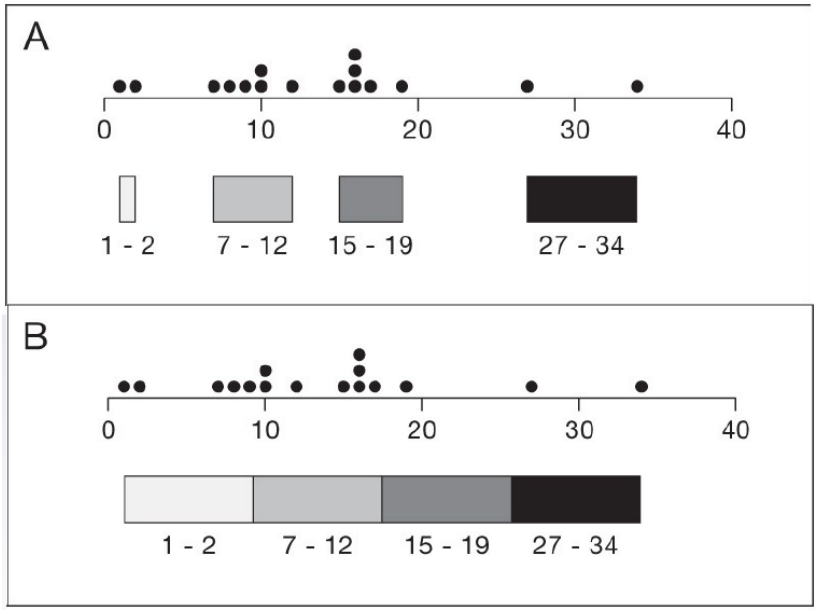


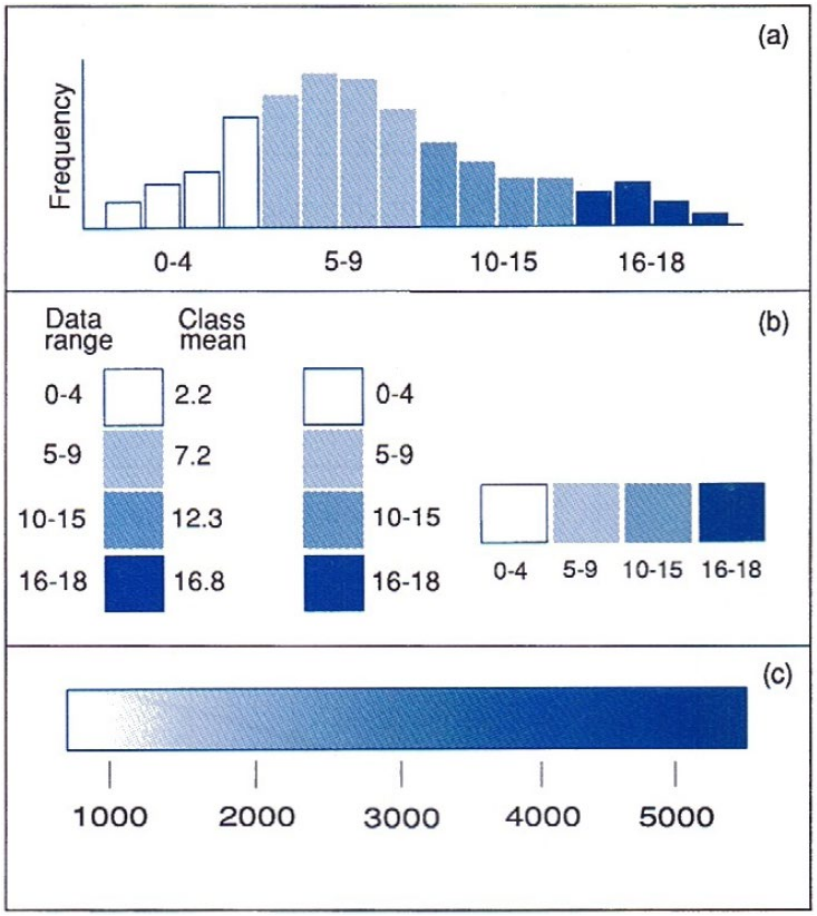




Additional graph designs

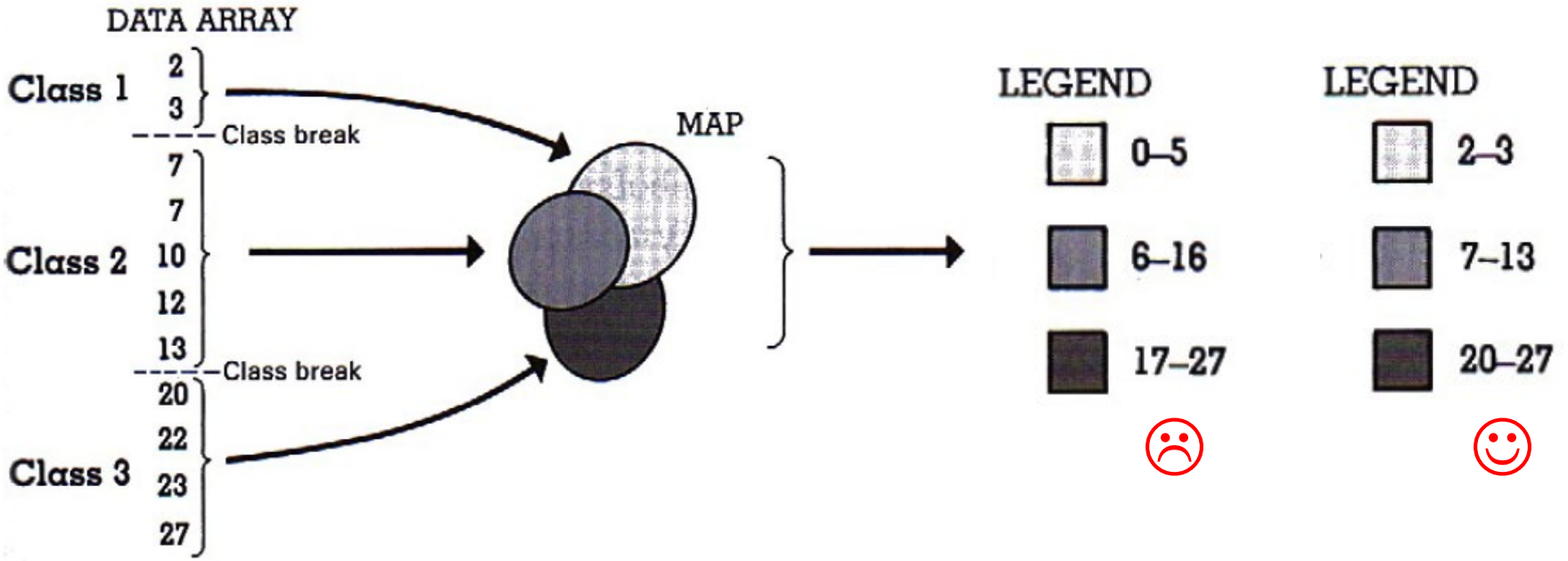
- Dispersion graph & Histogram





Continuous or Non-continuous?

- Interpretation errors may occur from the continuous legend design

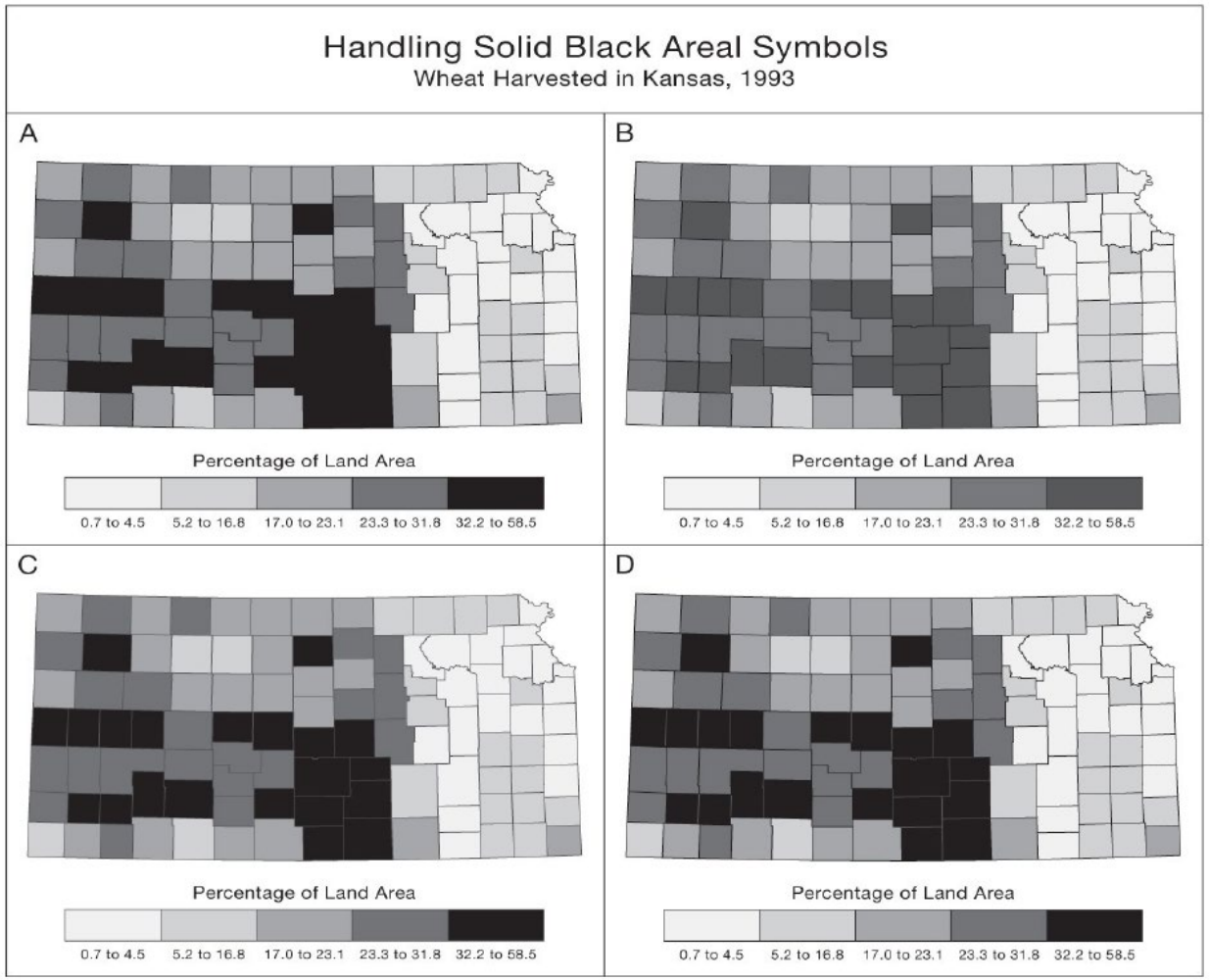


Handling Solid Black Area Symbols

Group Exercise

Q. Find the best usage of map symbology among A, B, C, and D

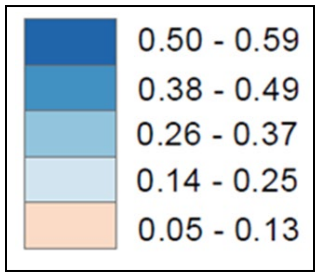
- A: black boundaries
- B: black boundaries with adjusted shades
- C: gray boundaries in between two darkest shades
- D: clearer gray boundaries than C

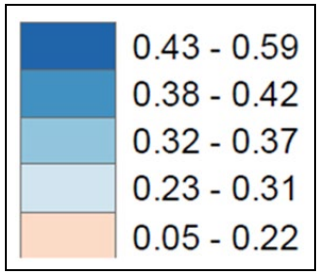


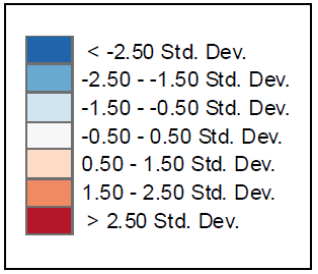
(source: Slocum et al. 2009)

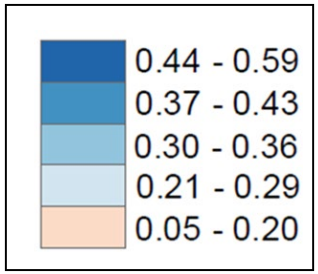
Classification Methods

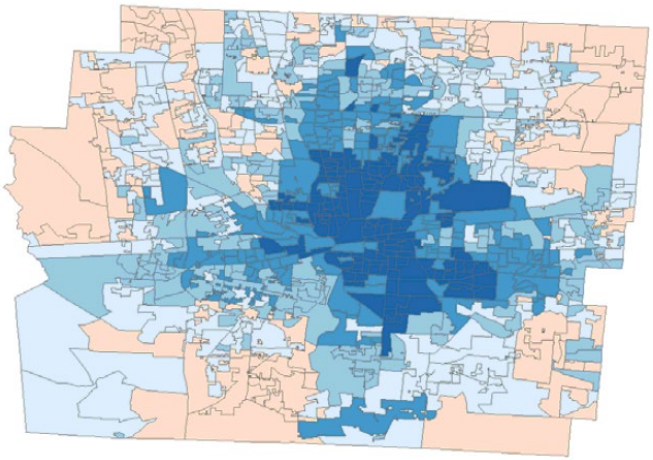
- Q. When you want to make a thematic map using numeric data, how would you classify the data?









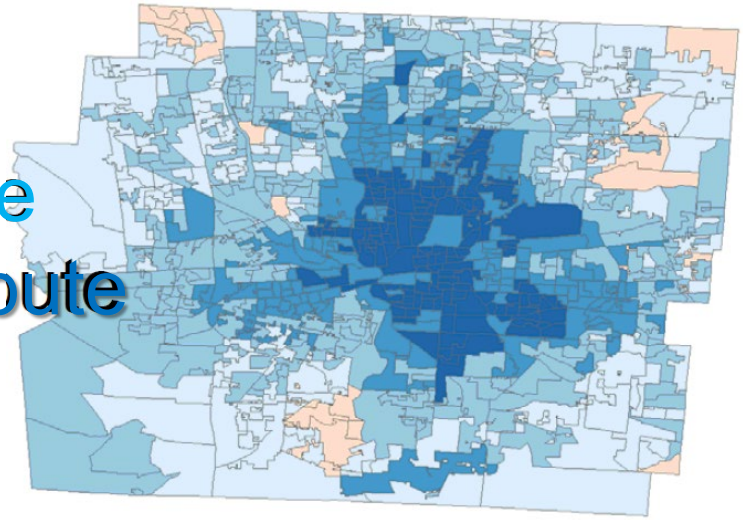


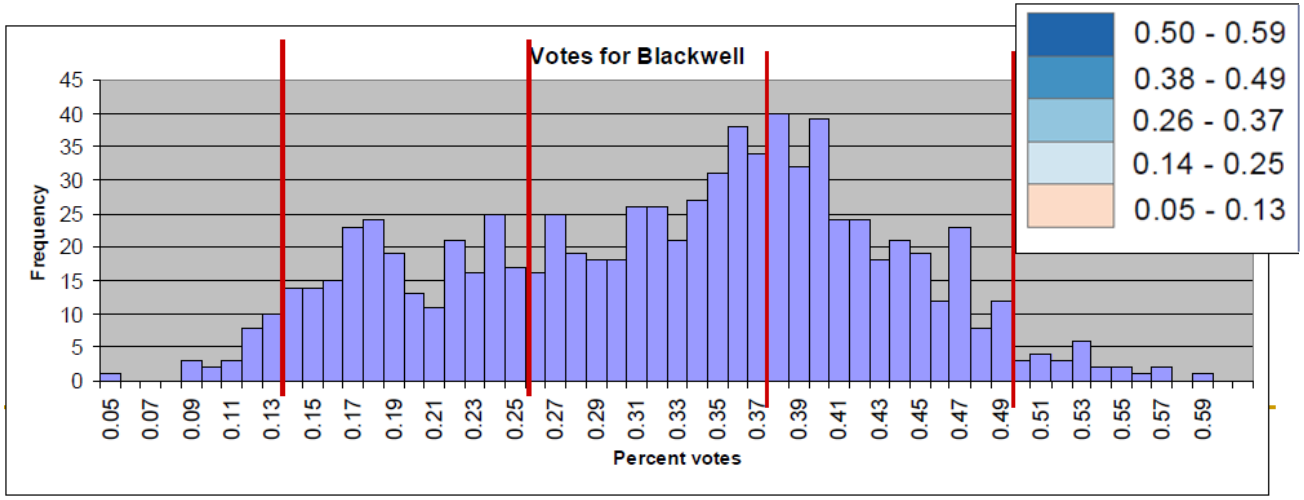
Mapping of Measurable Data

- Some common classification methods for choropleth mapping
  - Equal intervals
  - Quantiles
  - Standard deviation
  - Optimal classification

Equal interval

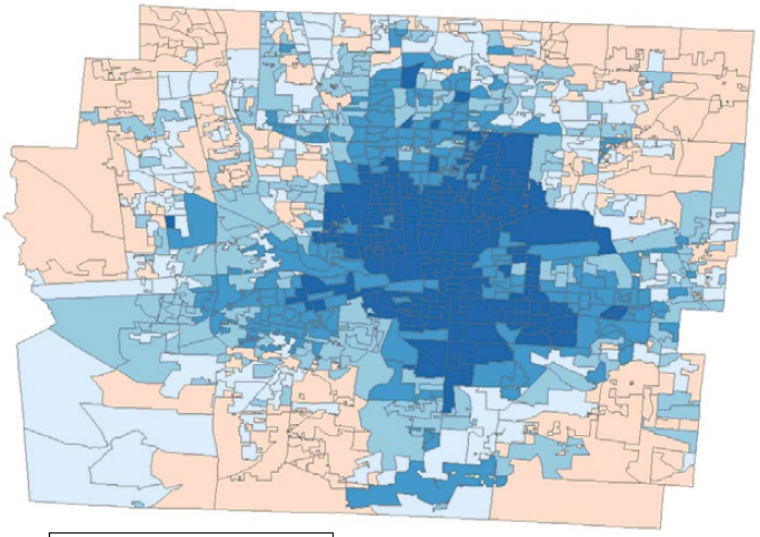
- Each class has the same size in terms of the attribute (size of value range)

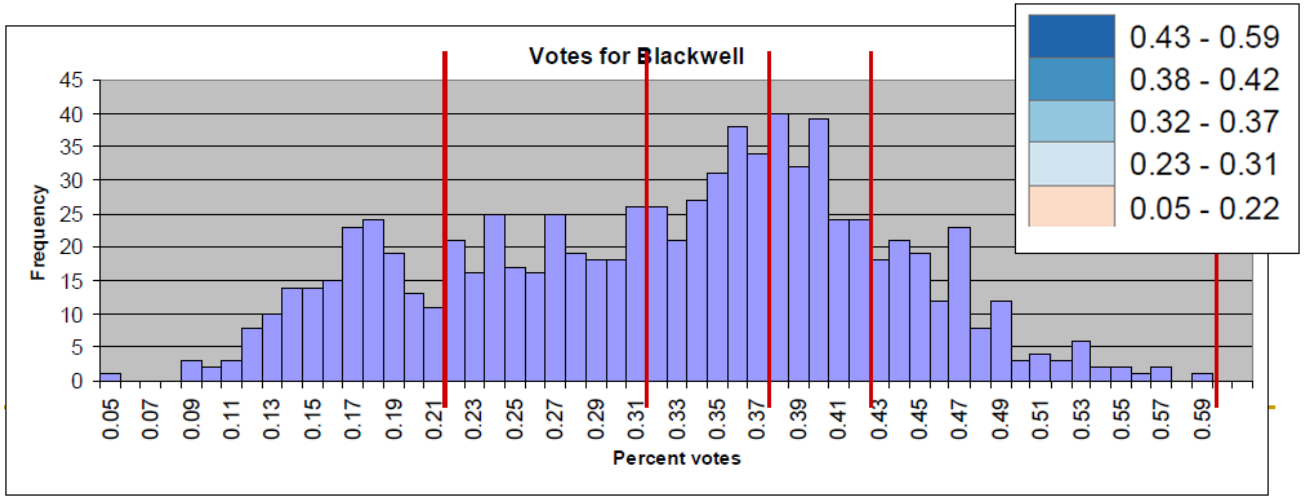




Quantile

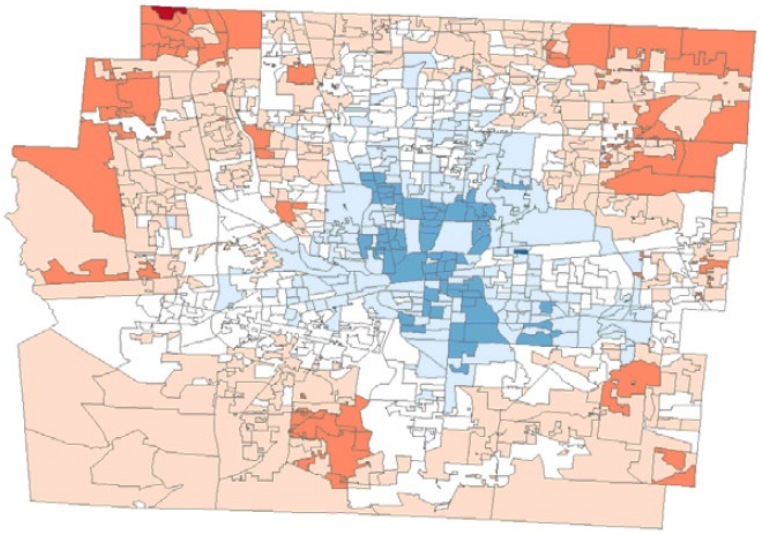
- Each class will contain equal number of objects (e.g., voting precincts)
  - quartiles (4 classes)
    quintiles (5 classes)
    deciles (10 classes)

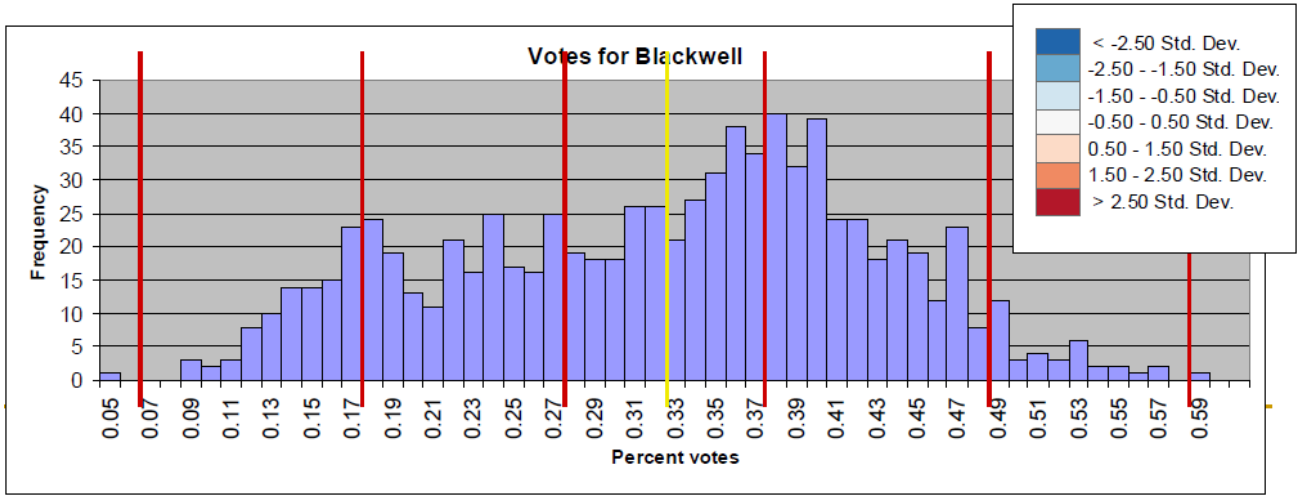




Mean-Standard Deviation

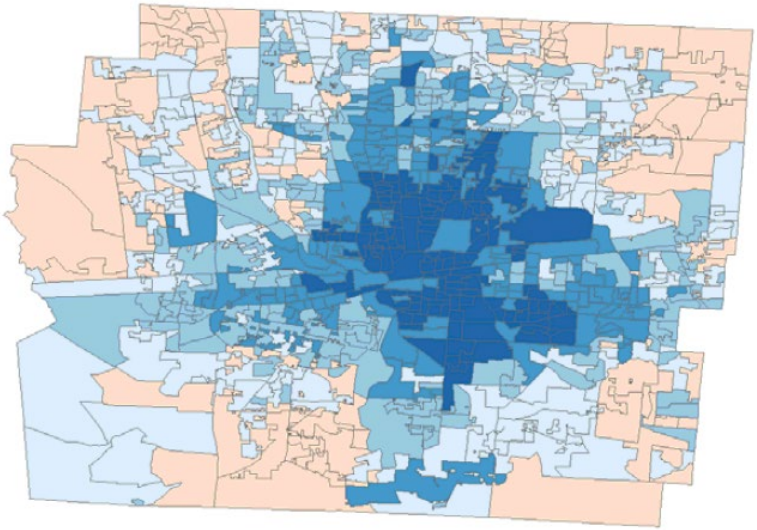
- Each class shows deviations from the mean

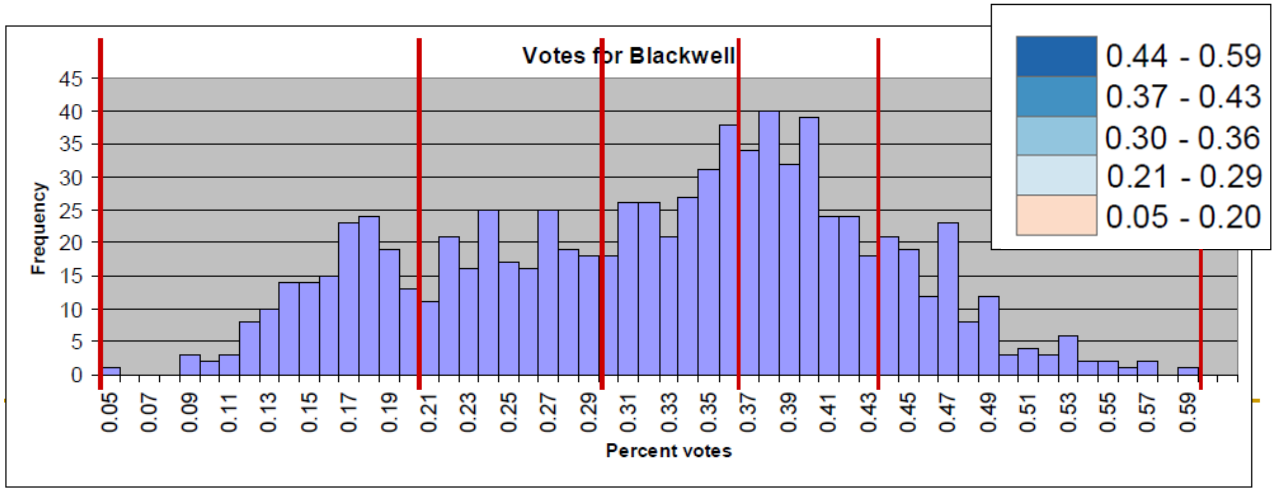




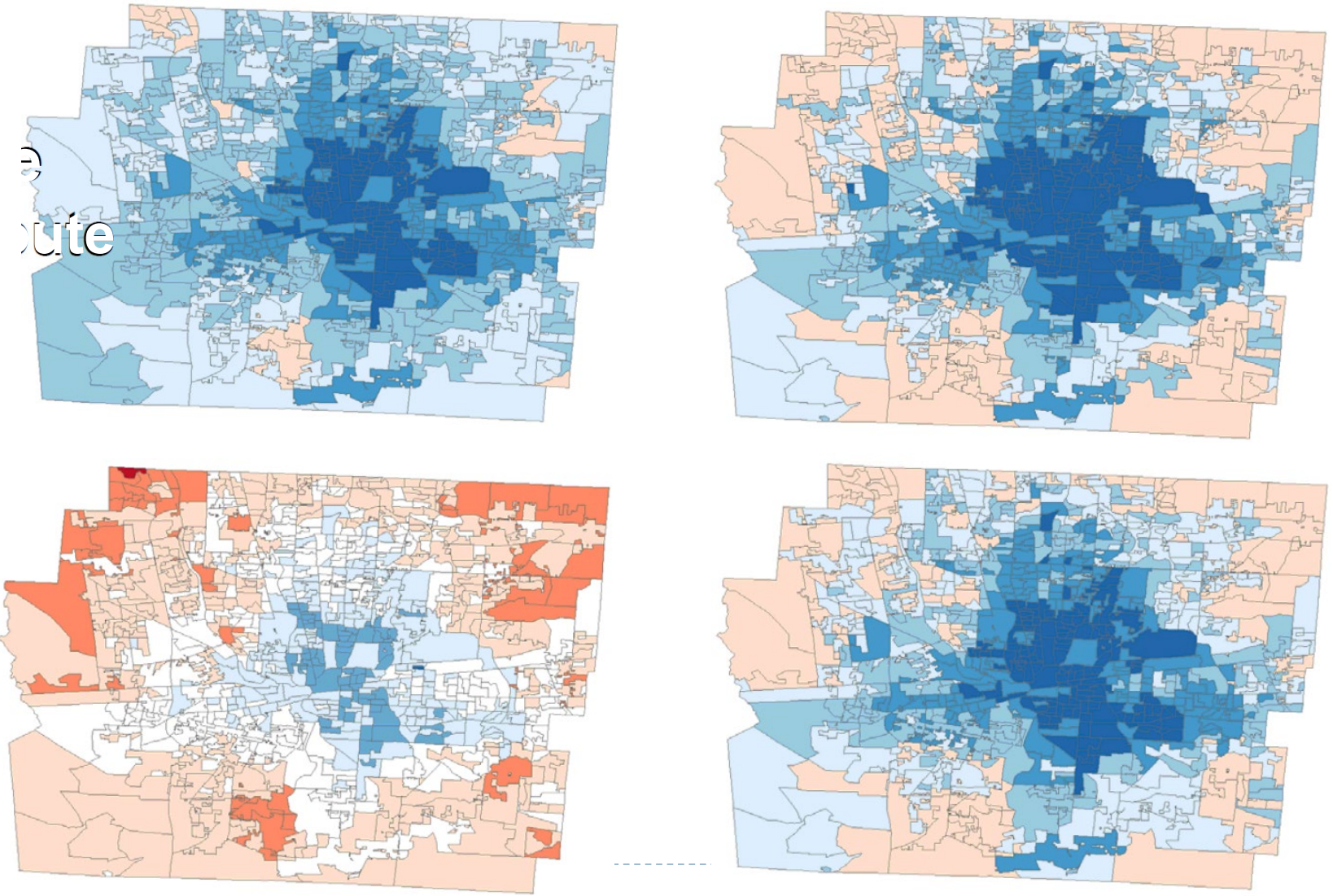
Optimal classification

- Minimize within-class variation
- Maximize between-class var.
  - Similar to clustering techniques





Which Classification Method to Choose?

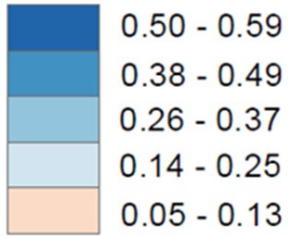


Some criteria for the selection

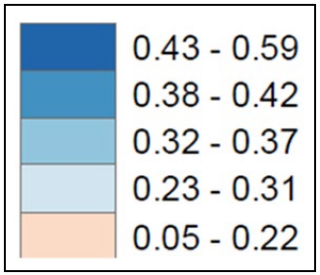
	Eq.Int.	Quantile	Std. dev.	Optimal.
Considers data distribution	P	P	G	VG
Easy to understand	VG	VG	VG	G
OK with ordinal data	U	A	U	U
Legend easy to understand	VG	P	G	P

See Figure 4.5
(P: poor, A: acceptable, U: unacceptable)

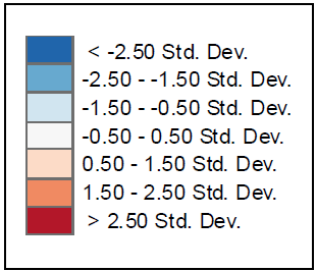
Equal interval



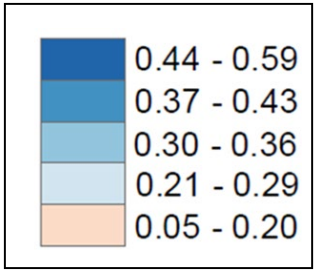
Quantile



Standard Deviation



Optimal



Demonstration

- Making classification of data using ArcGIS

Summary

- Statistical surfaces and a variety of visualization techniques for mapping
- Absolute numbers (dangerous!) and standardization for mapping
- Colors and legend design
- Classification methods

For next time…

- Readings
  - Ch.4 & 21

- Lab2 due today