Lecture 2:
Spatial Analysis

GEOG413/613

Dr. Anthony Jjumba

Spatial Analysis

- **Spatial analysis** or **Spatial data analysis**
  - Definitions
  - Some central concepts
  - Spatial statistics

Spatial Analysis

- Definitions:
  - Haining 1994: “a collection of techniques for analysing geographical events where the results of analysis depend on the spatial arrangement of the events”
  - Goodchild and Haining, 2003: “it represents a collection of techniques and models that explicitly use the spatial referencing of each data case”.
  - ESRI: The process of examining the locations, attributes, and relationships of features in spatial data... Spatial analysis extracts or creates new information from spatial data.

Spatial Analysis

- Definitions:

- How do we define Spatial Analysis?

  - Spatial analysis
    - It involves the application of techniques to extract information from spatial data. We consider the spatial context of the problem and the spatial arrangement of the data.
  - Spatial makes spatial data special

Spatial Data

- Is spatial data special so that we can exalt ourselves (the GIS tribe)?

Spatial Data

- Type
  - Primary/secondary
- Attribute/Variable
  - Continuous/discrete; Qualitative/quantitative
- Levels
  - Collection
    - Individual/Aggregated
  - Measurement
    - Nominal, Ordinal, Interval, Ratio

Spatial Statistics

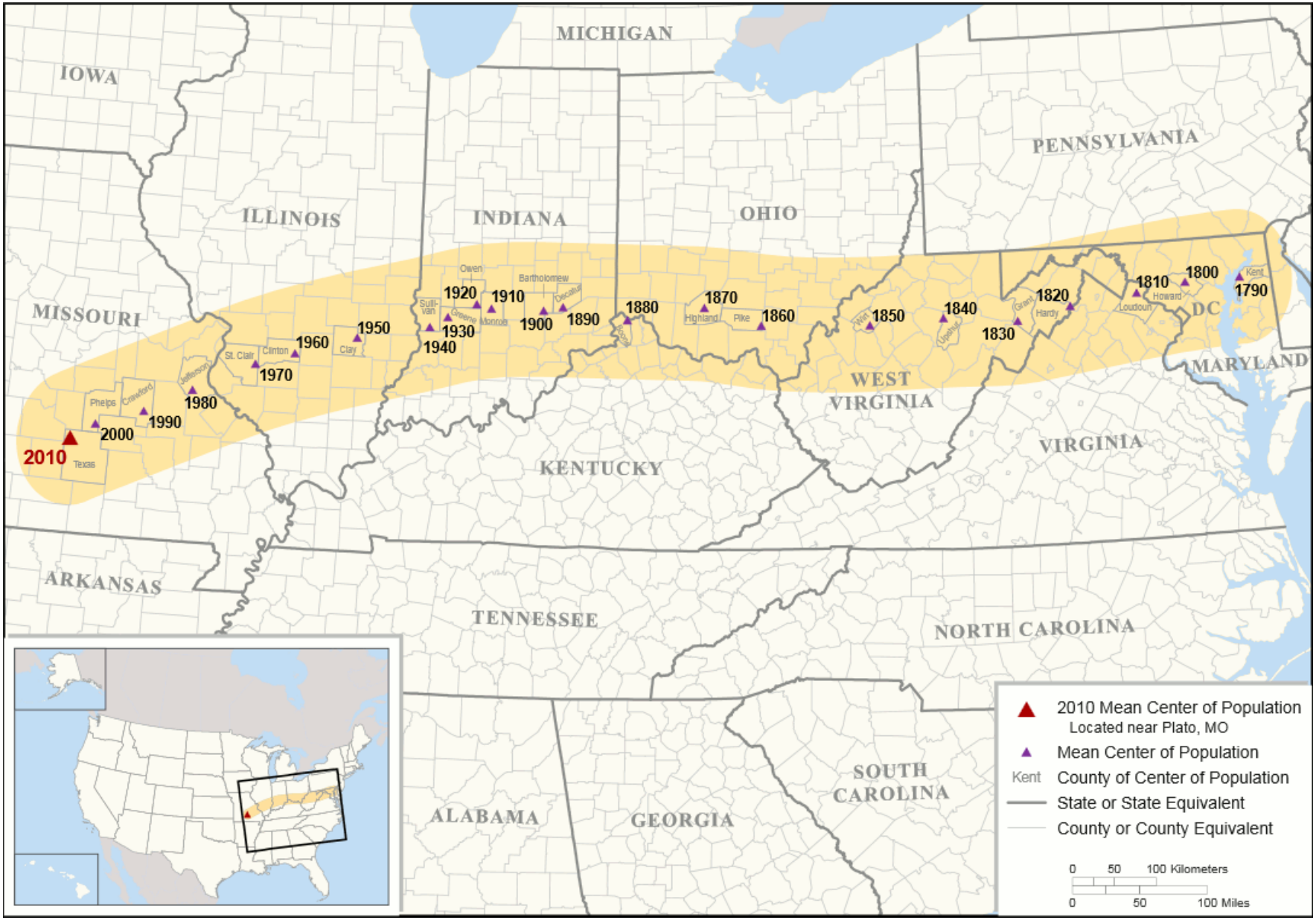
- Measures of Central Tendency
  - Mean Center
    - Arithmetic mean of a set of spatial objects (Centroid)
    - Mean Center $(\bar{x}, \bar{y})$

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad \bar{y} = \frac{\sum_{i=1}^{n} y_i}{n}$$

Spatial Statistics

- Measures of Central Tendency
  - Weighed Mean Center
    - Mean affected by a weight factor (e.g. frequency, population)
    - Represents the centre of gravity $(\bar{x}, \bar{y})$

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i w_i}{n} \qquad \bar{y} = \frac{\sum_{i=1}^{n} y_i w_i}{n}$$



Map showing changes to the **mean center of population** for the United States, 1790–2010 (US Census Bureau)

Spatial Statistics

- Measures of Central Tendency
  - Median Center/Euclidean Median
    - Center of minimum travel

Spatial Statistics

- Measures of Central Tendency
  - Manhattan Median
    - The point for which
      - half of the distribution is to the west the other half to the east
      - And half to the north and the other half to the south
    - The solution changes upon rotating the axes
    - For an even number of points, no exact solution

Spatial Statistics

- Measures of Dispersion
  - Standard Distance
    - Absolute spread of points around the mean center
    - Analogous to standard deviation

$$SD = \sqrt{\frac{\sum_{i=1}^{n}(X_i - \overline{X})^2}{n} + \frac{\sum_{i=1}^{n}(Y_i - \overline{Y})^2}{n}}$$

Spatial Statistics

- Measures of Dispersion
  - Relative Distance
    - Relative dispersion around a mean center
    - relates the standard distance to the size of the study area

    - RD = SD/r
      - RD - Relative Distance
      - SD - Standard Distance
      - r – radius of the cicle with the same area as the study area

Spatial Data

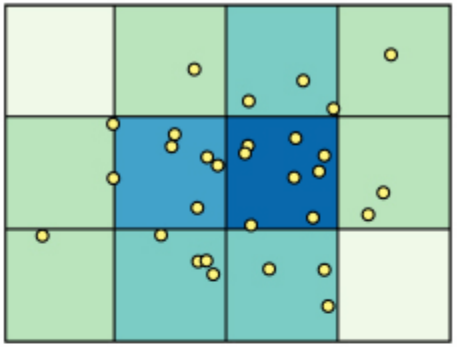
- Some pitfalls of spatial analysis:
  - Spatial autocorrelation
    - Implies that you can't assume a phenomenon is distributed randomly
    - Understanding it's nature is of primary importance
  - The Modifiable Area Unit Problem
    - "a problem arising from the imposition of artificial units of spatial reporting on continuous geographical phenomena resulting in the generation of artificial spatial patterns" (Heywood, 1988).

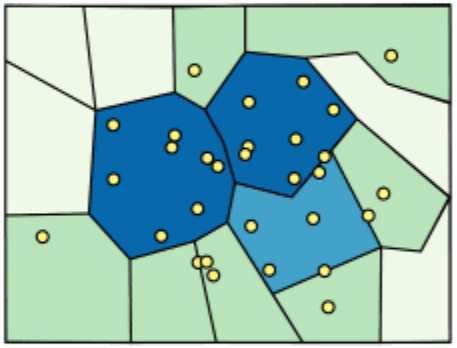
Spatial Data

- Some pitfalls of spatial analysis:
  - The Modifiable Area Unit Problem
    - Scale
    - Zones (see image)

"the areal units (zonal objects) used in many geographical studies are arbitrary, modifiable, and subject to the whims and fancies of whoever is doing, or did, the aggregating" Openshaw, 1983

Thus, statistics should be interpreted and evaluated by acknowledging the particular boundary scheme used in the study or experiment (McGrew and Monroe 2000)





source: gispopsci.org

Let’s assume a phenomenon such as average precipitation by zone

4.6	9.5	9.2	9.3
7.0	1.4	7.2	9.9
6.5	8.1	7.2	4.1
5.9	2.6	7.7	1.6

Mean: 6.3625
Standard Deviation: 2.7758

8.9	5.628
5.15	5.775

Mean: 6.3625
Standard Deviation: 1.7121

How do the statistics differ if the zones are configured differently?

Spatial Data

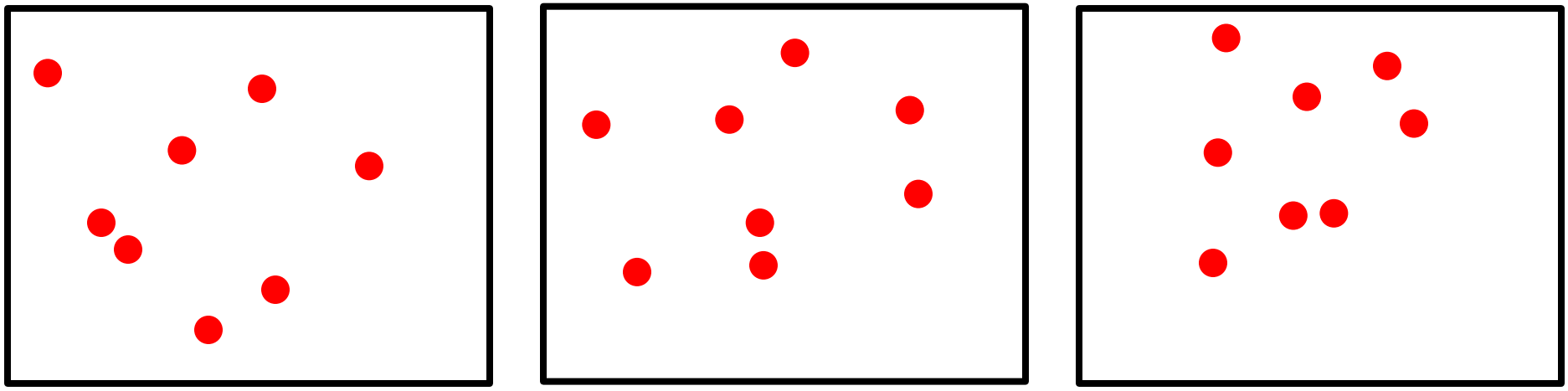
- Some pitfalls in spatial analysis:
  - Ecological Fallacy
    - Results Aggregated from data cannot be applied to individuals
  - Scale
    - Results depend on scale at which data was collected
  - Non-uniformity of space
    - Patterns can be random, clustered, uniform
  - Edge or Boundary Effects
    - No data beyond the study region
    - However, due to spatial autocorrelation, outside data could be affecting your study region

Point Pattern Analysis

- Point pattern analysis: the examination or evaluation of the pattern of distribution (spatial arrangement) for set of points.

Point Pattern Analysis

- Can you tell which is one Random, Clustered or Uniformly (dispersed) distributed?
- Always test



Point Pattern Analysis

- Next Week
  - Methods
    - Nearest Neighbor Analysis
    - Quadrat Analysis
    - K-Function
    - Kernel estimation
  - Spatial Autocorrelation

References

- David O'Sullivan, David Unwin, 2010 *Geographic Information Analysis*. Hoboken, NJ, John Wiley
- David W. S. Wong, Jay Lee, 2005. *Statistical Analysis of Geographic Information with ArcView GIS and ArcGIS*. Hoboken, NJ, John Wiley