Lecture 8

1

Jianguo (Jack) Ma

Department of Geographical Sciences
University of Maryland at College Park

Outline

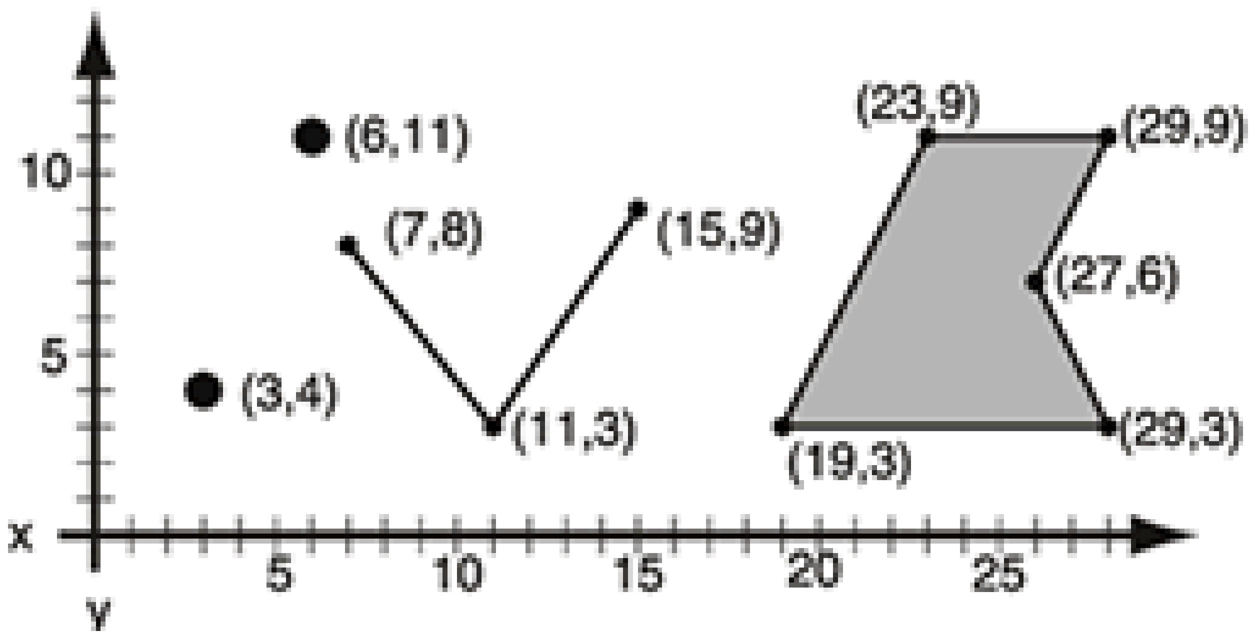
- Announcements & Updates
- Areal Analysis
  - Overview
  - Spatial Autocorrelation
  - Join Count Analysis
  - Indicators of Spatial Autocorrelation
    - Global Indicators
    - Local Indicators

Announcements

- Credit limit on ArcGIS Online
  - Lesson learned
- Software
  - Free, https://terpware.umd.edu/Windows
- Lab 3
  - An interesting animation by Michael Temchine
- Lab 4
  - Updates
- Lab 5
  - Heads-up

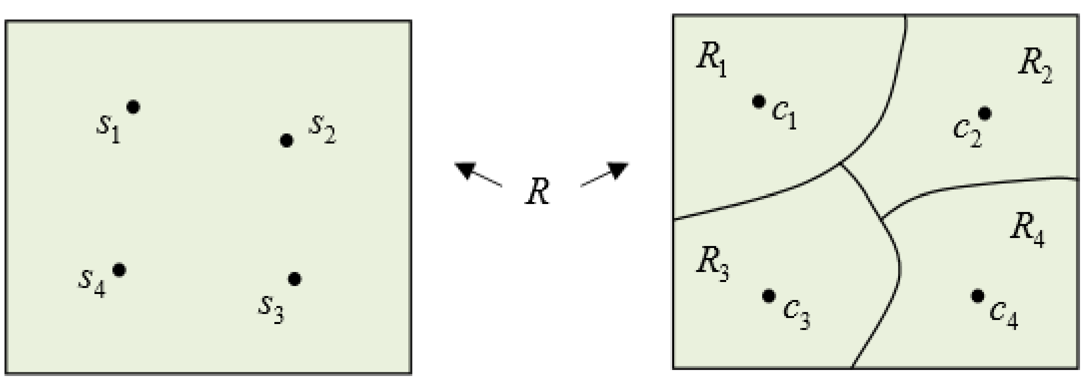
Areal Analysis

- Overview
  - – Areal analysis is used for identifying the spatial correlations of values that are represented by areas.
    - In vector data model, an area is represented as a polygon which is a closed set of lines.



Areal Analysis

- Overview
  - Point Samples vs. Areal Units
    - Point Samples ~ Point Pattern Analysis
      - Represents samples taken at various point locations from a continuous spatial distribution (e.g. groundwater well readings)
    - Areal Units ~ Areal Analysis
      - Represents aggregated quantities of a spatial variable over a defined region (e.g. average income within a census block)



Areal Analysis

- Overview
  - Areal analysis focus on the attribute values of the areas, instead of the location itself or the distance.
    - Point pattern analysis is different in that location is very important.
    - For discrete or non-contiguous areas (polygons), the analysis methods for points are also applicable.
      - Location and distance
    - For contiguous areas (polygons), the analysis is focused on attribute values associated with the features.

Areal Analysis

- Overview
  - The data is most likely not sampled.
  - Observations may not be random.
  - The data may represent the whole “population”.
  - The basic assumption of independence of observations in a sample may be violated.
  - The statistical results are susceptible to errors.
    - Type I error
    - Type II error

Areal Analysis

- Overview
  - The results should be used as only part of the information for decision making.
  - The results may be more useful when comparing with same statistics at different time periods.
    - Useful to study the trend.

Areal Analysis

- Overview
  - Types of Area Objects
    - Natural Areas
      - Self-defining – Boundaries defined by the phenomenon
      - Examples – Land use, soil profiles, water bodies, etc...
      - Issues:

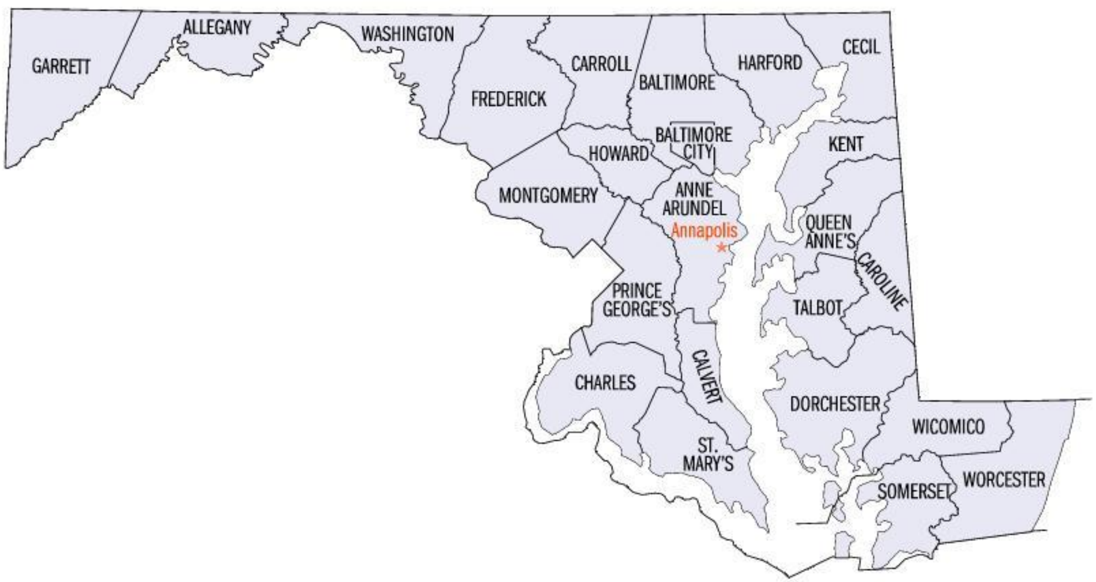
<u>Boundary Uncertainty</u>
(Land Use Data)

<u>Spatial Heterogeneity</u>
(SSURGO Soil Data)

Areal Analysis

- Overview
  - Types of Area Objects
    - Imposed Areas
      - Boundaries defined independent of any phenomenon
      - Examples – Political boundaries, census blocks, etc...
      - Issues:

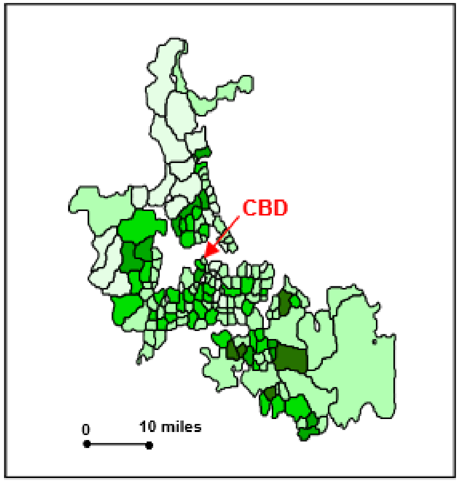
Modifiable Areal Unit Problem

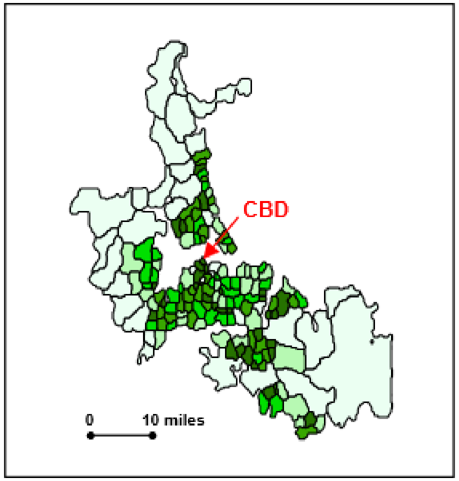


Areal Analysis

- Overview
  - Types of Area Objects
    - Extensive Data
      - Variables that are dependent on size (e.g. mass)
    - Intensive Data
      - Variables that are independent on size (e.g. density)
    - Example:
      - Population →
    - Usually best to use intensive data

11



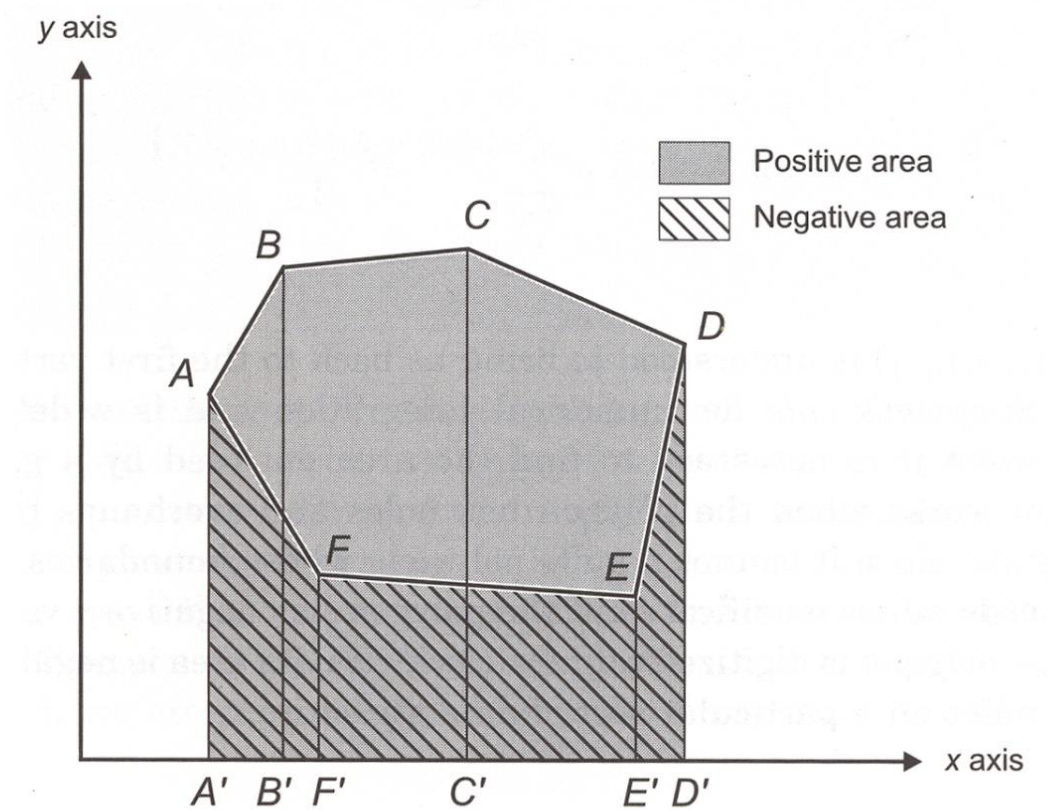


Areal Analysis

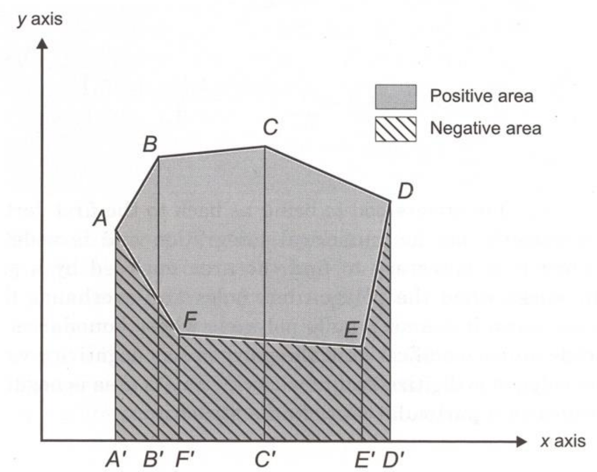
- Overview
  - Geometric Measures
    - Area
    - Perimeter
    - Centroid
    - Compactness
    - Cohesion Index

Areal Analysis

- Overview
  - Geometric Measures
    - Area
      - How to calculate the area of a polygon?

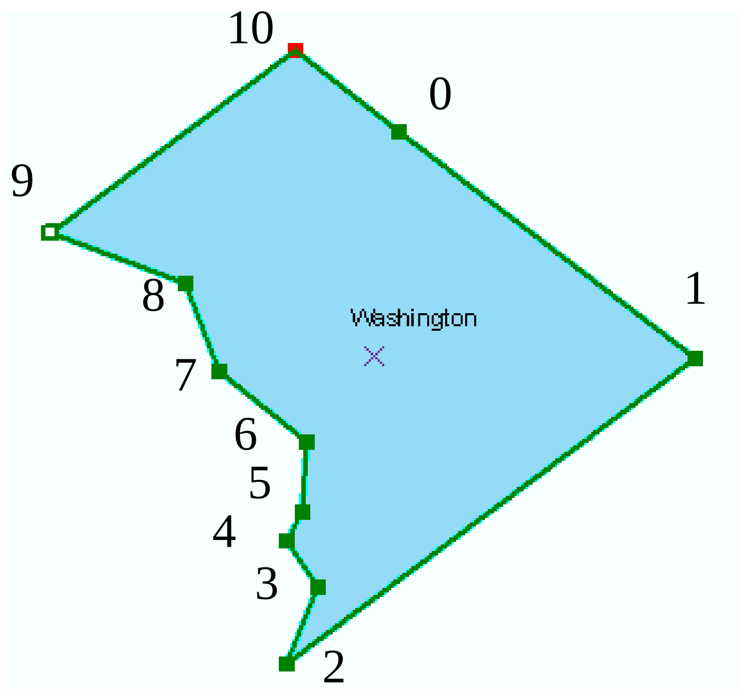


- Overview
  - Geometric Measures
    - Area
      - How to calculate the area of a polygon?
        - » The polygon can divided into a set of trapezoids.
        - » The area of a single trapezoid is: $A=\dfrac{(x_B-x_A)(y_B+y_A)}{2}$
        - » The area of the polygon is the sum of the area of each trapezoid minus the sum of the areas of the hashed trapezoids



Areal Analysis

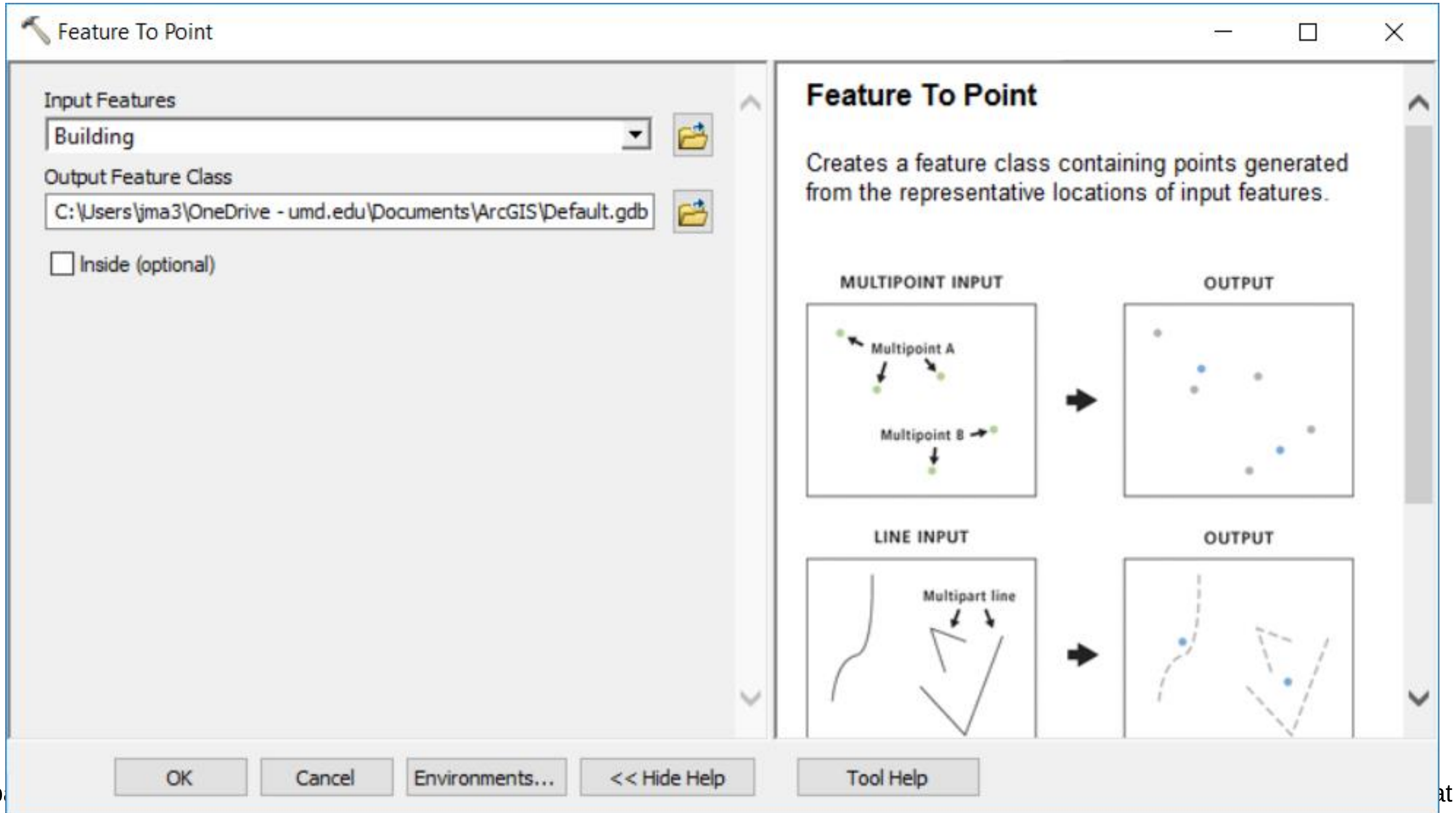
- Overview
  - Geometric Measures
    - Perimeter
      - The sum of the Euclidean lengths of the line segments that define the polygon boundary.



Areal Analysis

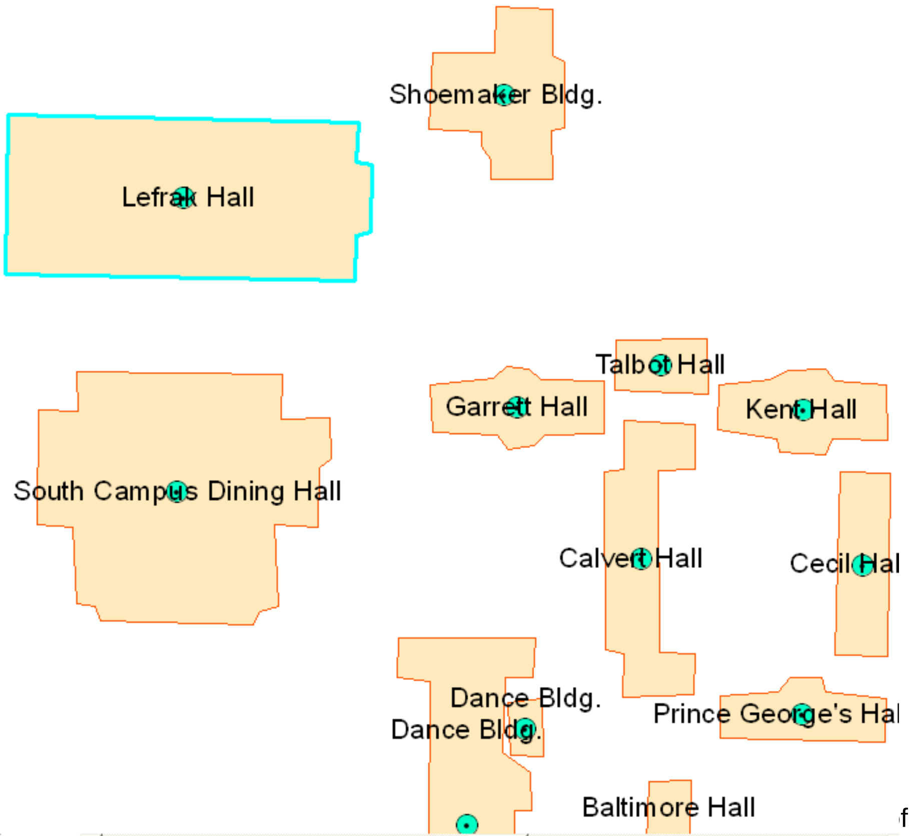
- Overview
  - Geometric Measures
    - Centroid
      - Similar to the concept of center of gravity, except that it applies to a two dimensional shape rather than an object.
      - For a given shape, the centroid location corresponds to the center of gravity for a thin flat plate of that shape, made from a homogeneous material.

- Overview
  - Geometric Measures
    - Centroid
      - ArcGIS tool – "Feature to Point"



Areal Analysis

- Overview
  - Geometric Measures
    - Centroid
      - Example: the centroids of those buildings (polygons) on UMD campus.



Areal Analysis

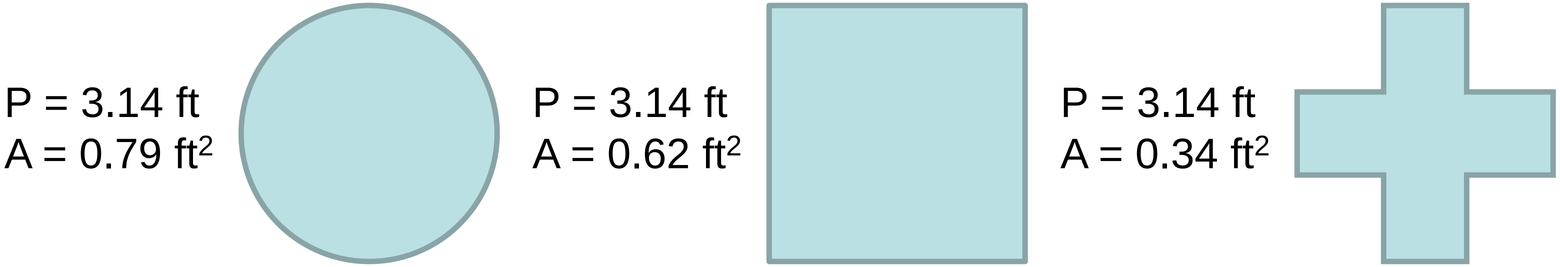
- Overview
  - Geometric Measures
    - Centroid
      - Not all centroids are located within the boundary of polygons.
      - Example: the centroid of Van Munching Hall on UMD campus.

Areal Analysis

- Overview
  - Geometric Measures
    - Centroid
      - The location of centrod is dependent on the shape of the polygon.
        - » That's why the location of an area is less important than that of a point.
      - Centroid is not always the best representative central location for a geographic area.

Areal Analysis

- Overview
  - Geometric Measures
    - Compactness
      - It is an indicator of shape for an areal object.
      - It measures how different the shape of an area is from a circle that has the same perimeter as the polygon.
        - » Given the same perimeter, the circle always has the largest area.



        - » Similarly: How about a sphere? Why does the rain drop or water drop tend to be like a sphere?

- Overview
  - Geometric Measures
    - Compactness

$$\text{Compactness ratio} = \sqrt{\frac{a}{a_2}}$$

a: area of the shape

a_2: area of a circle with the same perimeter

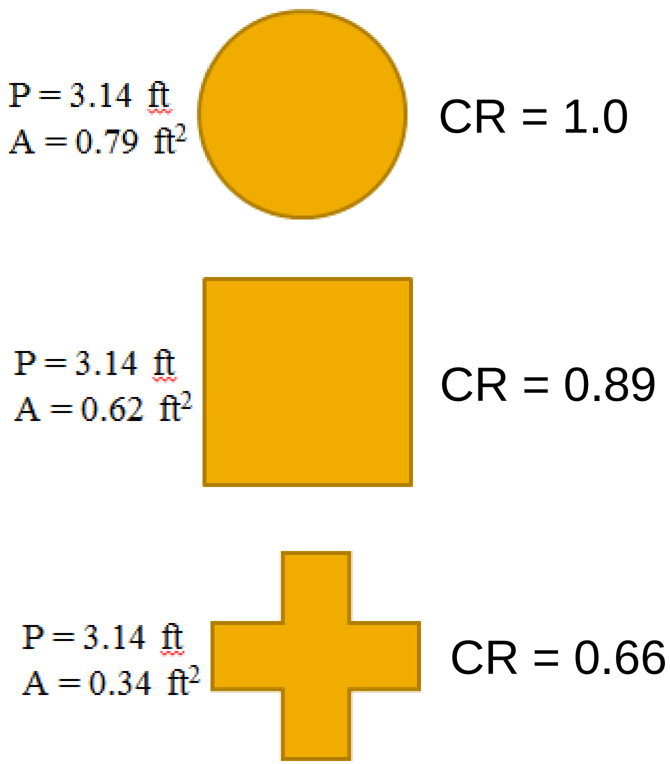
Compactness is a number between 0 and 1.

0: polygon is a line

1: polygon is a circle

Areal Analysis

- Overview
  - Geometric Measures
    - Compactness



Areal Analysis

- Overview
  - Geometric Measures
    - Compactness
      - A different compactness ratio:

$$C = \sqrt{4\pi \frac{A}{P^2}}$$

C: Compactness ratio

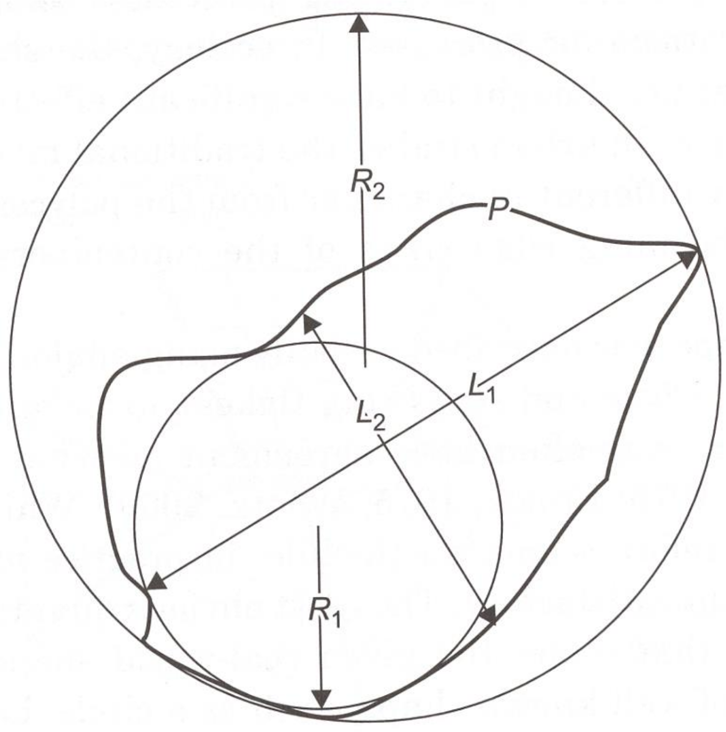
A: The area of the polygon.

P: Perimeter of the polygon.

Compactness is a number between 0 and 1.

0: polygon is a line

1: polygon is a circle



Areal Analysis

- Overview
  - Geometric Measures
    - Cohesion Index
      - It measures the connectivity of polygons with the same attribute.
      - It can be used to assess the fragmentation of habitat areas or wildlife protection zones.

- Overview
  - Geometric Measures
    - Cohesion Index

$$\text{Cohesion index} = \frac{1 - \dfrac{\sum_m p_i}{\sum_m p_i \sqrt{a_i}}}{1 - \sqrt{\dfrac{1}{A}}}$$

For m polygons, each polygon's perimeter is p_i and its area is a_i. The area of all polygons is A.

Areal Analysis

- Overview
  - Geometric Measures
    - Cohesion Index
      - The index is a number between 0 and 1.
      - If it is 0, there is only one polygon.
      - The cohesion increases as the polygon type becomes more connected (clustered).

Areal Analysis

- Spatial Autocorrelation
  - Spatial Autocorrelation indicates whether the distribution of values is dependent on the spatial distribution of the features.
    - Whether particular values are likely to occur in one location, or are equally likely to occur at any location.

Areal Analysis

- Spatial Autocorrelation
  - First Law of Geography
    - Everything is related to everything else, but near things are more related than distant things. - Waldo Tobler
      - Examples:
        - » Crime
        - » Rain
      - Exceptions:
        - » Barrier

Areal Analysis

- Spatial Autocorrelation
  - Two types of Spatial Autocorrelation
    - Positive Spatial Autocorrelation
      - Nearby features are more like each other than they are like more distant features.
      - More common in nature and human society.
    - Negative Spatial Autocorrelation
      - Nearby features are more unlike each other than they are like more distant features.

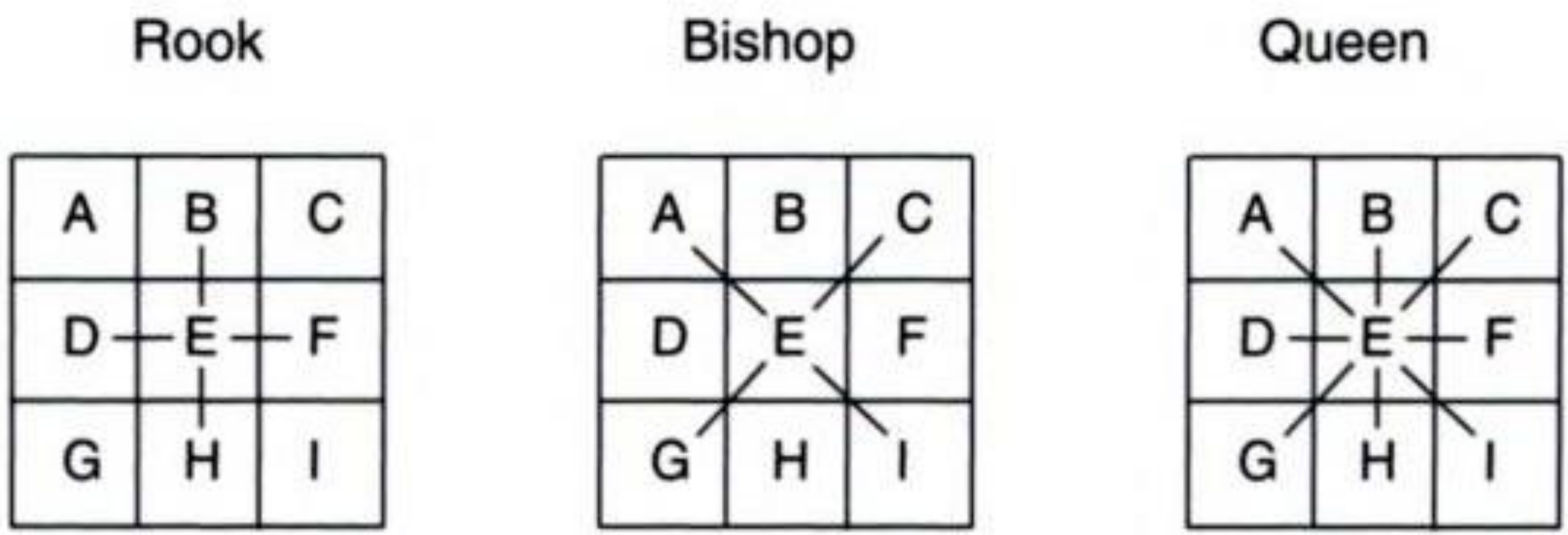
- Spatial Autocorrelation
  - Why spatial autocorrelation is important?
    - Most statistics are based on the assumption that the values of observations in each sample are independent of one another.
    - Positive spatial autocorrelation may violate this, if the samples were taken from nearby areas.
    - Goals of spatial autocorrelation
      - Measure the strength of spatial autocorrelation in a map
      - Test the assumption of independence or randomness

Areal Analysis

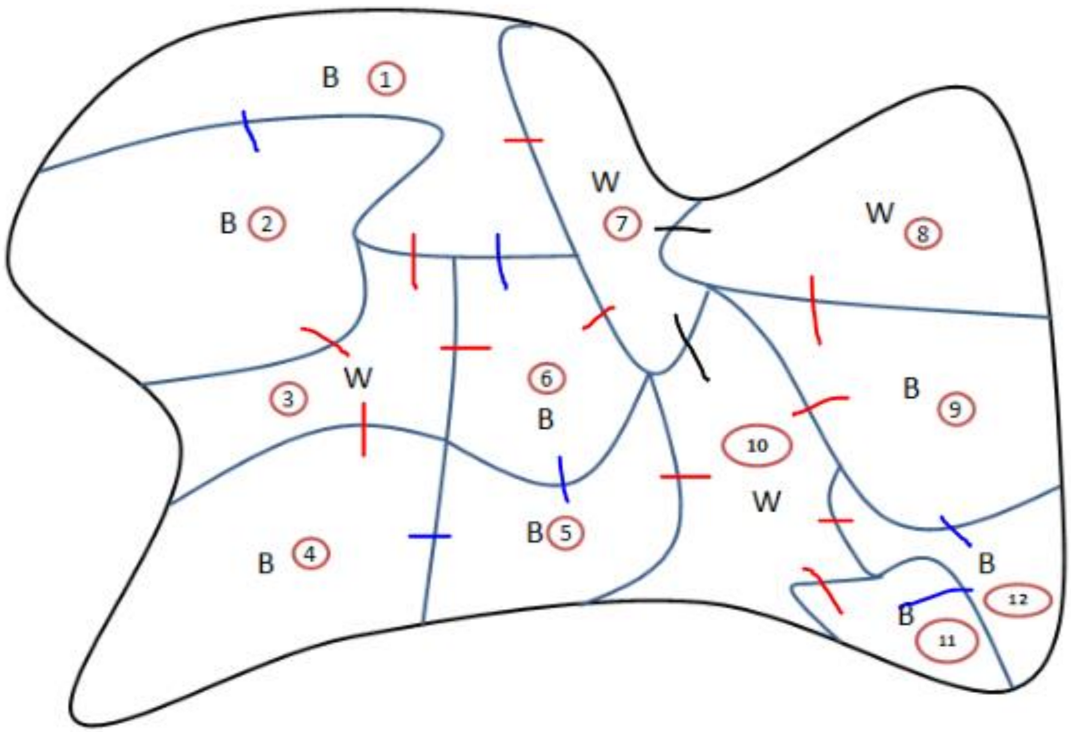
- Spatial Autocorrelation
  - Spatial Weight Matrix
    - Assume that we have *n* areal units in our study region. Then there are *nxn* pairs of relationships to be captured.
    - Usually we use matrix to store and organize spatial relationship among these areal units.
    - Spatial weight matrix, an n x n matrix, is composed of element representing neighborhood relationship between areal unit *i* and *j*.

Areal Analysis

- Spatial Autocorrelation
  - Neighborhood Definition
    - Each can produce significantly different statistical results.



- Spatial Autocorrelation
  - Neighborhood Definition
    - Each can produce significantly different statistical results.
      - Example: Join Count in Rook’s case



Areal Analysis

- Spatial Autocorrelation
  - Spatial Weight Matrix
    - This connectivity matrix is binary.

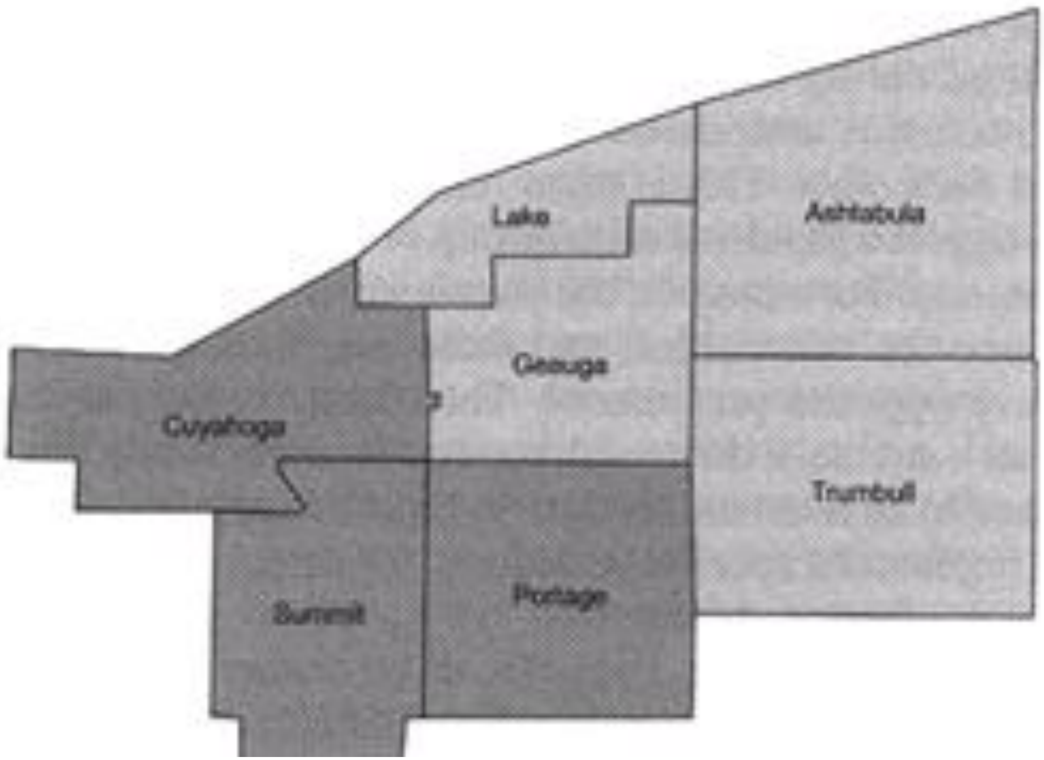
$$c_{ij} = \begin{cases} 1, \text{ when the } i\text{th polygon is adjacent to } j\text{th polygon} \\ 0, \text{ otherwise} \end{cases}$$

$$\text{Row sum}: c_{i.} = \sum_j c_{ij}$$

$$\text{Column sum}: c_{.j} = \sum_i c_{ij}$$

$$\text{Total cell value}: c_{..} = \sum_i \sum_j c_{ij}$$

- Spatial Autocorrelation
  - Spatial Weight Matrix
    - Example:

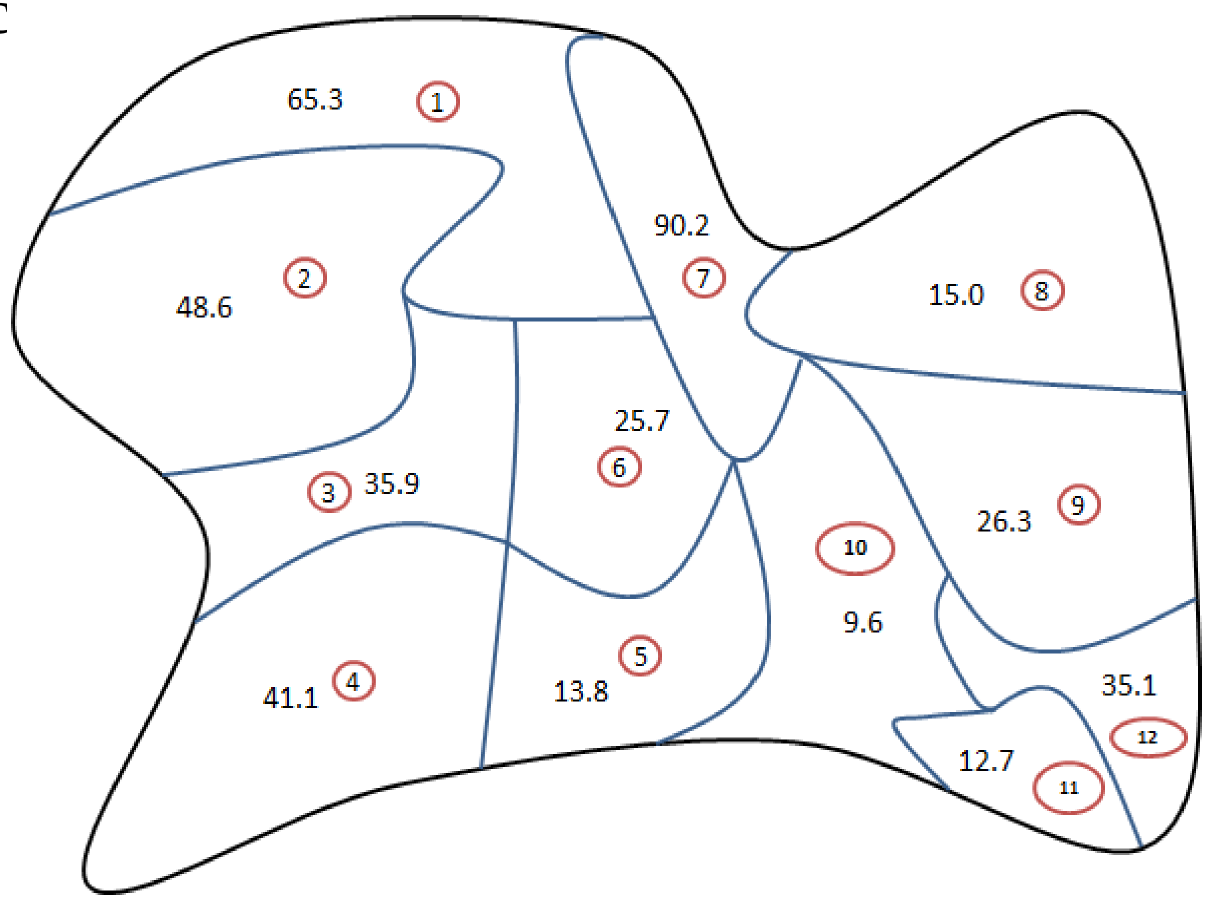


Areal Analysis

- Spatial Autocorrelation
  - Spatial Weight Matrix
    - Example: (assuming Rook's case)

ID	Geauga	Cuyahoga	Trumbull	Summit	Portage	Ashtabulla	Lake	Row sum
Geauga	0	1	1	1	1	1	1	6
Cuyahoga	1	0	0	1	1	0	1	4
Trumbull	1	0	0	0	1	1	0	3
Summit	1	1	0	0	1	0	0	3
Portage	1	1	1	1	0	0	0	4
Ashtabula	1	0	1	0	0	0	1	3
Lake	1	1	0	0	0	1	0	3
Column sum	6	4	3	3	4	3	3	26

- Spatial Autocorrelation
  - Spatial Weight Matrix
    - Different neighborhood definition can produce signific



Areal Analysis

- Spatial Autocorrelation
  - Spatial Weight Matrix
    - Different neighborhood definition can produce significantly different statistical results.
      - Example: spatial weight matrix in Rook's case

Polygon	1	2	3	4	5	6	7	8	9	10	11	12
1	0	1	1	0	0	1	1	0	0	0	0	0
2	1	0	1	0	0	0	0	0	0	0	0	0
3	1	1	0	1	0	1	0	0	0	0	0	0
4	0	0	1	0	1	0	0	0	0	0	0	0
5	0	0	0	1	0	1	0	0	0	1	0	0
6	1	0	1	0	1	0	1	0	0	0	0	0
7	1	0	0	0	0	1	0	1	0	1	0	0
8	0	0	0	0	0	0	1	0	1	0	0	0
9	0	0	0	0	0	0	0	1	0	1	0	1
10	0	0	0	0	1	0	1	0	1	0	1	1
11	0	0	0	0	0	0	0	0	0	1	0	1
12	0	0	0	0	0	0	0	0	1	1	1	0

Areal Analysis

- Spatial Autocorrelation
  - Spatial Weight Matrix
    - Different neighborhood definition can produce significantly different statistical results.
      - Example: spatial weight matrix in Queen's case

Polygon	1	2	3	4	5	6	7	8	9	10	11	12
1	0	1	1	0	0	1	1	0	0	0	0	0
2	1	0	1	0	0	0	0	0	0	0	0	0
3	1	1	0	1	1	1	0	0	0	0	0	0
4	0	0	1	0	1	1	0	0	0	0	0	0
5	0	0	1	1	0	1	1	0	0	1	0	0
6	1	0	1	1	1	0	1	0	0	1	0	0
7	1	0	0	0	1	1	0	1	1	1	0	0
8	0	0	0	0	0	0	1	0	1	1	0	0
9	0	0	0	0	0	0	1	1	0	1	0	1
10	0	0	0	0	1	1	1	1	1	0	1	1
11	0	0	0	0	0	0	0	0	0	1	0	1
12	0	0	0	0	0	0	0	0	1	1	1	0

Areal Analysis

- Spatial Autocorrelation
  - Spatial Weight Matrix
    - Different neighborhood definition can produce significantly different statistical results.
      - Example: spatial weight matrix in Bishop's case

Polygon	1	2	3	4	5	6	7	8	9	10	11	12
1	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	1	0	0	0	0	0	0	0
4	0	0	0	0	0	1	0	0	0	0	0	0
5	0	0	1	0	0	0	1	0	0	0	0	0
6	0	0	0	1	0	0	0	0	0	1	0	0
7	0	0	0	0	1	0	0	0	1	0	0	0
8	0	0	0	0	0	0	0	0	0	1	0	0
9	0	0	0	0	0	0	1	0	0	0	0	0
10	0	0	0	0	0	1	0	1	0	0	0	0
11	0	0	0	0	0	0	0	0	0	0	0	0
12	0	0	0	0	0	0	0	0	0	0	0	0

Areal Analysis

- Spatial Autocorrelation
  - Centroid Distance
    - Besides using adjacency as a measure to describe the spatial relationship between areal units, one can use distance between them.
    - Distance between polygons can be calculated as distance between their corresponding centroids.
    - Then, spatial weight between polygons can be defined as a function of distance between them.
    - Example:

$$w_{ij} = \frac{1}{d_{ij}^{\alpha}}$$

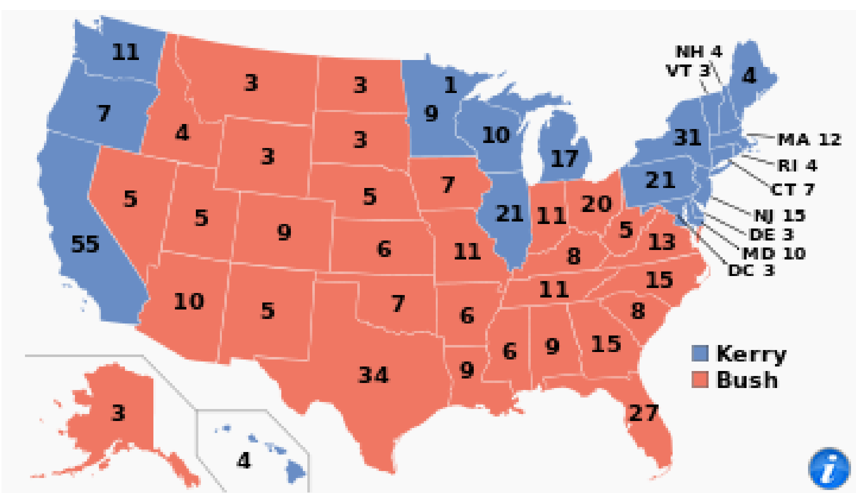
α- distance decay parameter

Areal Analysis

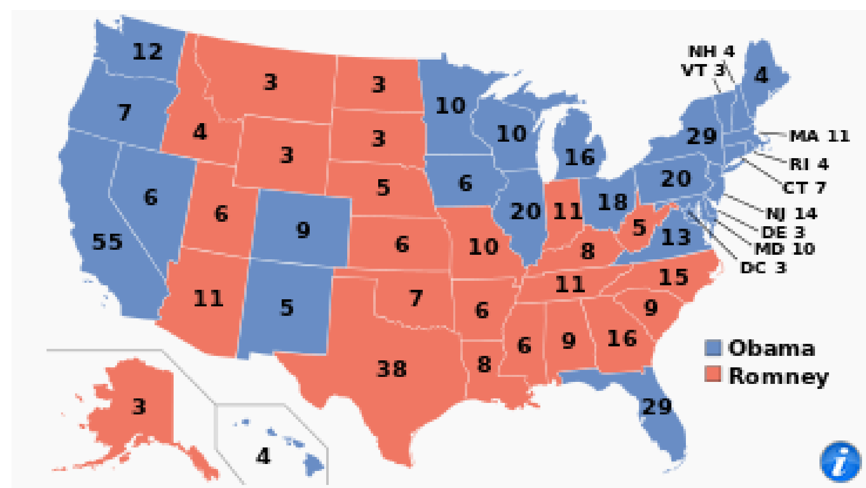
- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Different types of variables (attribute values) require different spatial autocorrelation measures.
      - Nominal values: categories (e.g. LULC)
      - Ordinal values: rank order (e.g. population density ranking)
      - Interval values: the magnitude of the difference between values (e.g. temperature)
      - Ratio values: the difference and the ratio between values can be calculated. (e.g. distance, precipitation)

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Join Count Analysis
      - Used for contiguous areas that have category attributes (i.e. nominal data).
        - » Example: election zones of “Yes” or “No”

2004 U.S. Presidential Election

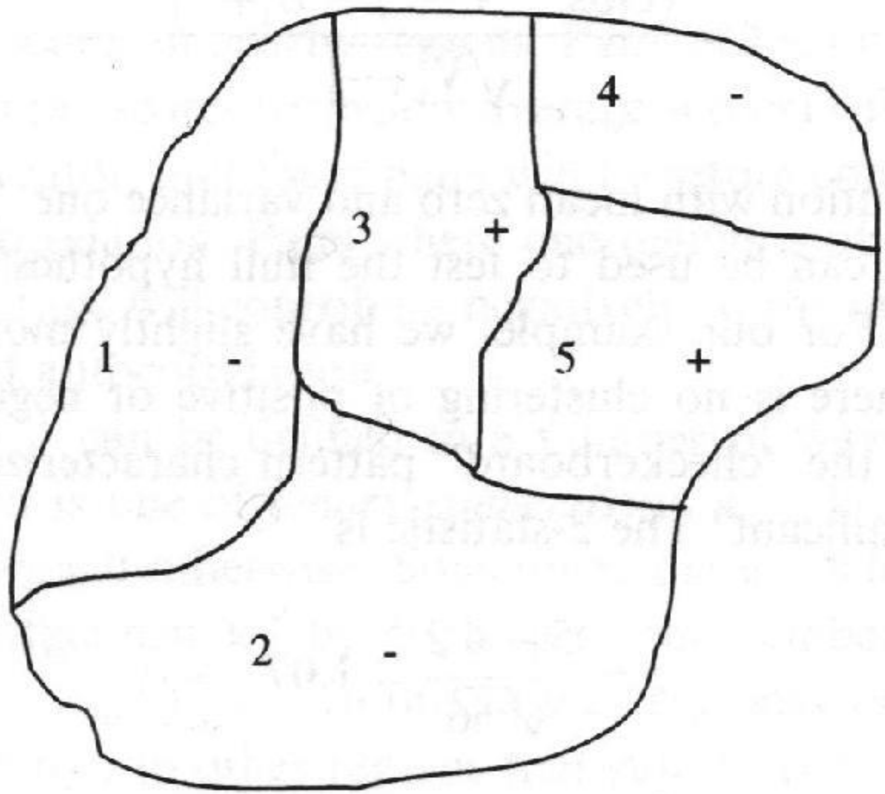


2012 U.S. Presidential Election



Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Join Count Analysis



Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Join Count Analysis
      - Joins: all shared borders between areas
      - Count: the number of joins for which the value on either side of the border is the same and the number for which the values are different.

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Join Count Analysis
      - How does it work?
        - » The number of joins of each type is compared with the expected number based on the probability of two adjacent areas having the same value by chance.
        - » If the counted number of joins for areas having the same value is greater than the expected number, that value is clustered.

Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Join Count Analysis
      - It deals with Binary nominal data such as 0 or 1, black or white, and positive or negative.
      - In this example: count the number of joins: PP(++), PN(+-), and NN(--).
        - » PP: (3,5)
        - » PN: (1,3), (2,3), (2,5), (3,4), (4,5)
        - » NN: (1,2)

Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Join Count Analysis
      - The join count statistic compares the observed number of PN joins with the number of PN joins that would be expected if no spatial autocorrelation were present.
      - Example: close to observed value (5).

$$E(PN) = \frac{2JPM}{N(N-1)} = \frac{2 \cdot 7 \cdot 2 \cdot 3}{5 \cdot 4} = 4.2$$

where J is the total number of joins,

P is the number of polygons with positive values,

M is the number of polygons with negative values,

N is the total number of polygons (= P+M)

Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Join Count Analysis
      - Significance test for PN joins

$$Z = \frac{(Observed\ PN) - E(PN)}{\sqrt{Var(PN)}}$$

Area *i*	PN joins
1	1
2	2
3	3
4	2
5	2
Var	0.50

Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Join Count Analysis
      - Significance test for PN joins

$$Z = \frac{(Observed\ PN) - E(PN)}{\sqrt{Var(PN)}}$$

$$= \frac{5 - 4.2}{\sqrt{0.50}} = 1.13 < 1.65$$ (90% confidence intervel)

      - Therefore, we cannot reject the null hypothesis of randomly placed positive-negative polygons.

Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Join Count Analysis
      - Join count analysis doesn't take into account the magnitude of the values in each area since it is based simply on whether the area falls into one category or the other.
        - » Example: In a two-candidate election, an area with 49.9% of voting "Yes" to Candidate A has the same category value as another area with 2% of voting "Yes" to the same candidate.
      - When the feature values are reclassified into two categories, there will be information loss.
        - » Reclassifying works best when there is a definite or natural break point between categories.

Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Join Count Analysis
      - Factors influencing the Join Count statistic results:
        - » There are less than 30 features in the study area.
        - » One of the category values occurs in less than 20% of the areas.
        - » The region is elongated. Most area have few joins.
          - » There are a couple of features with many joins while all other features have only one or two joins.
        - » Dependent on the connectivity rule (rook, bishop, queen)

Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Join Count Analysis
      - No corresponding tool in ArcGIS
      - Will introduce on how to use GeoDa to do Join Count Analysis

Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Interval or ratio data:
      - Moran's I
      - Geary's C
      - G-Statistic
    - These statistics measures and test how clustered/dispersed the point (or polygons) are with respect to their attributes.
    - These methods can be used for any type of geographic features (discrete or contiguous).

Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Moran's I
      - Moran's Index
      - Developed by Australian statistician Patrick Moran
      - The statistic compares the values for each feature in a neighboring features to the mean value for the dataset.
      - Scale of –1 (dispersed) to +1 (clustered).

Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Moran's I
      - In ArcGIS, it is the tool – **Spatial Autocorrelation (Global Moran's I)**
        - » Measures spatial autocorrelation based on feature locations and attribute values using the Global Moran's I statistic.

Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Geary's C
      - Developed by economist and statistician Robert Geary.
      - The statistic compares the values between any two neighboring features.
      - Scale of 0 (clustered) to 2 (dispersed)

Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Moran’s I
    - Geary’s C
    - For both methods, if the difference in values of nearby features is less than the difference in values among all features, like values are clustered.
    - These two methods only indicate that similar value cluster or not. However, they don’t tell us whether the clusters are composed of high values or low values.

Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - G-Statistic
      - Developed by Art Getis and Keith Ord.
      - This statistic measures within a specified distance how high or low the values are, and compares this to a measure of how high or low the values are over the entire study area.
        - » Unlike Moran's I and Geary's C, this statistic tells whether high values or low values are clustered.
      - It calculates a single statistic for the entire study area.

Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - G-Statistic
      - In ArcGIS, it is the tool - High/Low Clustering (Getis-Ord General G).
        - » Measures the degree of clustering for either high values or low values using the Getis-Ord General G statistic.

Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Generic form of spatial autocorrelation measures:

$$SAC \approx \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} s_{ij} w_{ij}}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}}$$

s_{ij} the similarity of point/polygon *i* and *j*'s attributes,

w_{ij} the proximity of point/polygon *i* and *j*'s location,

n the total number of points/polygons.

Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Global measures:
      - Global measures are used to describe the level of spatial autocorrelation for the entire study region.
      - Moran's I
      - Geary's C
      - General G statistics

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Global measures:
      - Moran's I

$$I = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} s_{ij}}{\sigma^2 \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} = \frac{n\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(x_i - \overline{x})(x_j - \overline{x})}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} \sum_{i=1}^{n}(x_i - \overline{x})^2}$$

Areal Analysis

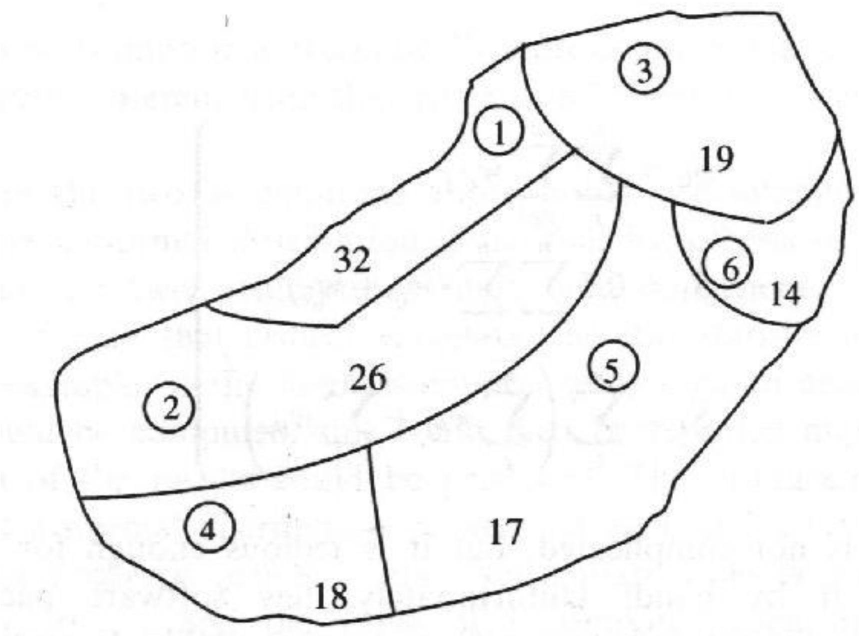
- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Global measures:
      - Geary’s C

$$C = \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} s_{ij}}{2\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij} S^2} = \frac{(n-1)\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(x_i - x_j)^2}{2\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

$$\text{where } S^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$$

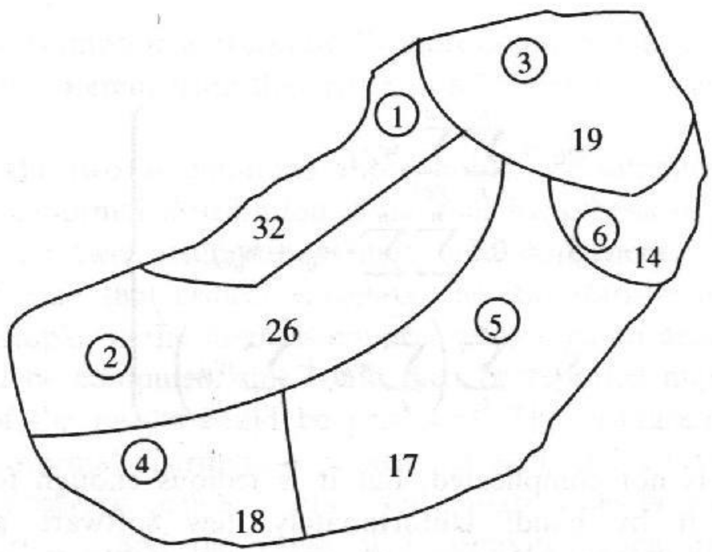
Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Global measures:
      - Moran's I
        - » Example:



Areal Analysis

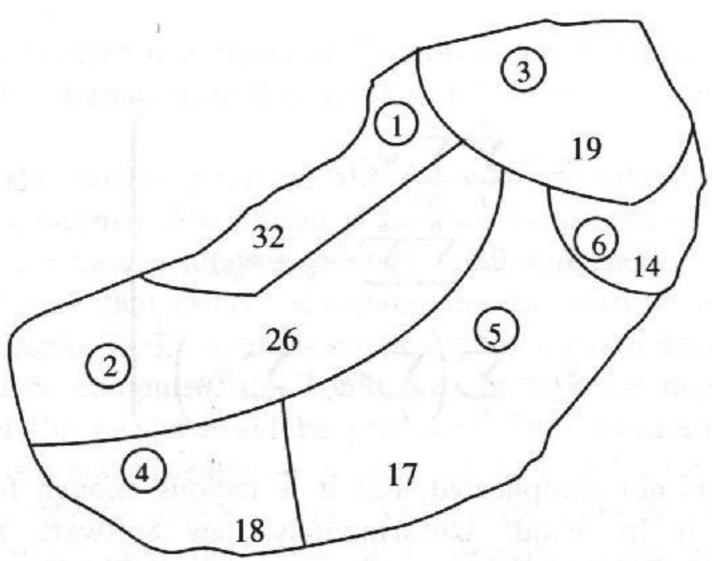
- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Global measures:
      - Moran's I
        - » Example: assuming Rook's case



Zone	1	2	3	4	5	6
1	0	1	1	0	0	0
2	1	0	1	1	1	0
3	1	1	0	0	1	1
4	0	1	0	0	1	0
5	0	1	1	1	0	1
6	0	0	1	0	1	0

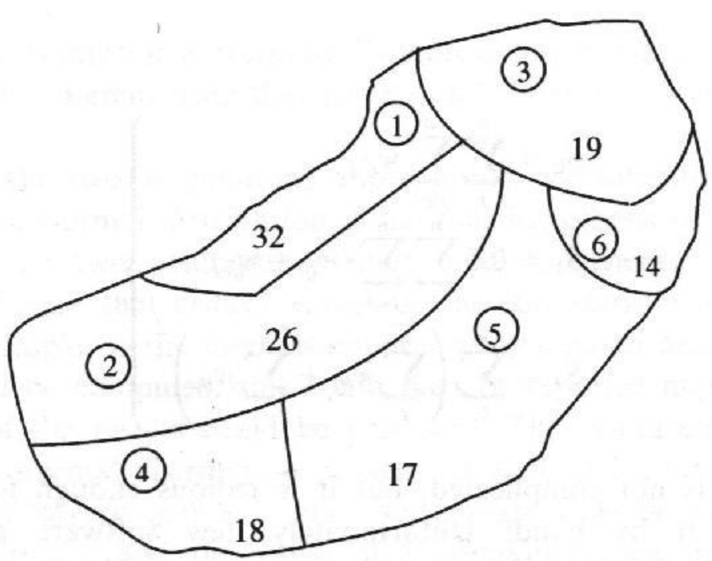
- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Global measures:
      - Moran's I
        - » Example:

$$\sum_i (x_i - \bar{x})^2$$



i/j	X	(X-Mean)	(X-Mean)^2
1	32	11	121
2	26	5	25
3	19	-2	4
4	18	-3	9
5	17	-4	16
6	14	-7	49
Mean	21		224

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Global measures:
      - Moran's I
        - » Example: $\text{Wij}(x_i - \bar{x})(x_j - \bar{x})$



Zone	1	2	3	4	5	6
1	0	55	-22	0	0	0
2	55	0	-10	-15	-20	0
3	-22	-10	0	0	8	14
4	0	-15	0	0	12	0
5	0	-20	8	12	0	28
6	0	0	14	0	28	0

Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Global measures:
      - Moran's I
        - » Example:

70

$$I = \frac{n\sum_i \sum_j w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{(\sum_i \sum_j w_{ij})\sum_i (x_i - \bar{x})^2} = \frac{6 \cdot 100}{18 \cdot 224} = 0.14881$$

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Global measures:
      - Moran's I
        - » If the number of areal units is large, the sampling distribution of I under the hypothesis of no spatial pattern, approaches a normal distribution, and the mean and variance of I can be used to create a Z-score for significance test.
        - » Where:
          - » I is the calculated Moran's I
          - » E(I) is the expected if random
          - » n is the number of observations
          - » √Var(I) is the standard error

$$E(I) = -\frac{1}{n-1}$$

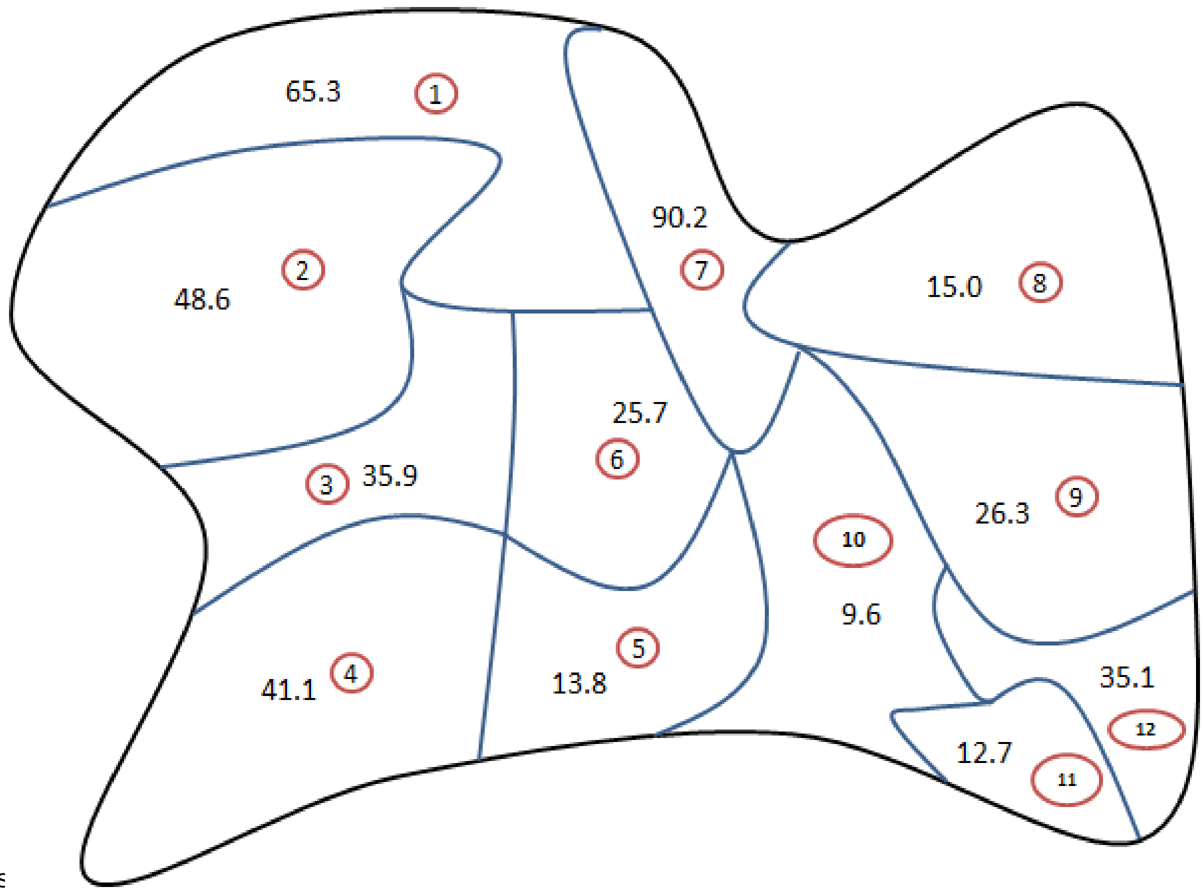
$$Z = \frac{I - E(I)}{\sqrt{Var(I)}}$$

Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Global measures:
      - Drawbacks of Moran's I and Geary's C:
        - » Neither is designed to assess localized spatial autocorrelation surrounding individual location or areal units.
        - » They are not effective in identifying different types of local clustering patterns: hotspots and cold spots.
        - » Hot spot refers to local cluster of high values, and cold spot refers to local cluster of low values.
        - » In Moran's I and Geary'C, both hot spots and cold spots will be indicated as high positive spatial autocorrelation.

Areal Analysis

- Spatial Autocorrelation
  - Exercise



- Spatial Autocorrelation
  - Exercise
    - Moran's I value
      - Spatial weight matrix

Polygon	1	2	3	4	5	6	7	8	9	10	11	12
1	0	1	1	0	0	1	1	0	0	0	0	0
2	1	0	1	0	0	0	0	0	0	0	0	0
3	1	1	0	1	0	1	0	0	0	0	0	0
4	0	0	1	0	1	0	0	0	0	0	0	0
5	0	0	0	1	0	1	0	0	0	1	0	0
6	1	0	1	0	1	0	1	0	0	0	0	0
7	1	0	0	0	0	1	0	1	0	1	0	0
8	0	0	0	0	0	0	1	0	1	0	0	0
9	0	0	0	0	0	0	0	1	0	1	0	1
10	0	0	0	0	1	0	1	0	1	0	1	1
11	0	0	0	0	0	0	0	0	0	1	0	1
12	0	0	0	0	0	0	0	0	1	1	1	0

- Spatial Autocorrelation
  - Exercise
    - Moran's I value

$$I = \frac{n\sum_i \sum_j w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{(\sum_i \sum_j w_{ij})\sum_i (x_i - \bar{x})^2}$$

Where:
$\boldsymbol{n}$ - the total number of points/polygons.
$\boldsymbol{w_{ij}}$ - the weight for the pair of polygon i and j's attributes,
$\boldsymbol{x_i}$, $\boldsymbol{x_j}$ - the value of polygon i and j,
$\bar{x}$ – the mean value

Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Global measures:
      - General G statistics

$$G(d) = \frac{\sum_i \sum_j w_{ij}(d) x_i x_j}{\sum_i \sum_j x_i x_j}, \quad \text{for } i \neq j$$

where

x_i is the attribute value of area i

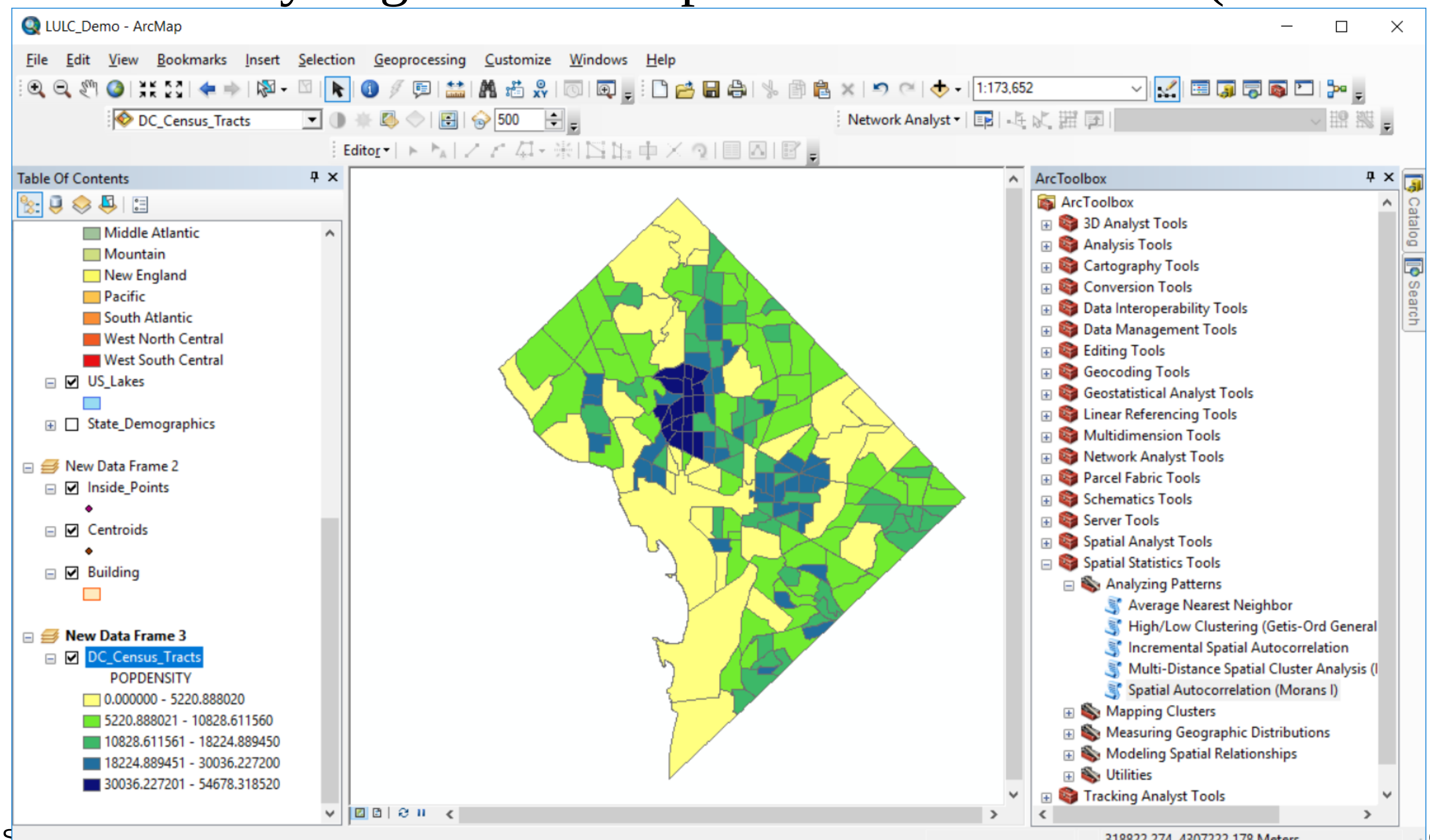
$$w_{ij} = \begin{cases} 1, & \text{if } j \text{ is within distance } d \text{ from } i \\ 0, & \text{otherwise} \end{cases}$$

Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Global measures:
      - General G statistics
        - » G(d) can be considered as the proportion of the spatial association accounted for within d distance units of each observation.
        - » Relatively large G(d) > E[G(d)] indicates presence of hotspots.
        - » Relatively low G(d) < E[G(d)] indicates presence of cold spots.
        - » As d increases, G(d) will also increase.

Areal Analysis

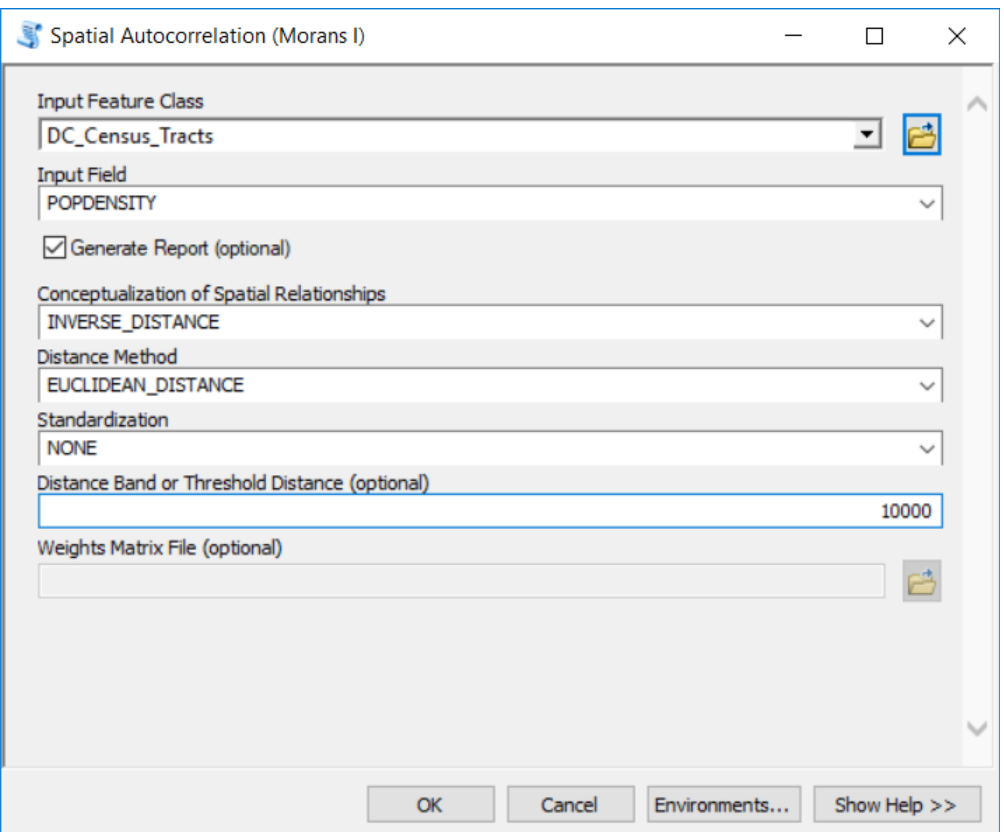
- Spatial Autocorrelation
  - Global Measures
    - Analyzing Pattern > Spatial Autocorrelation (Morans I)

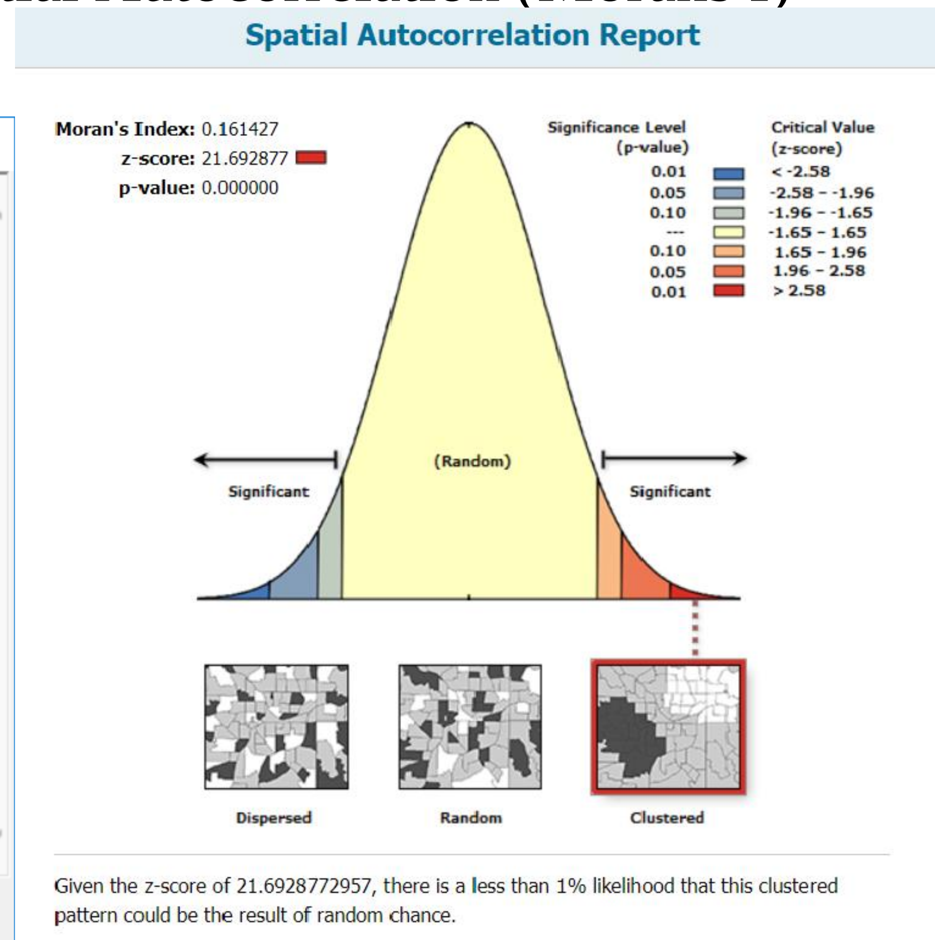


78

Areal Analysis

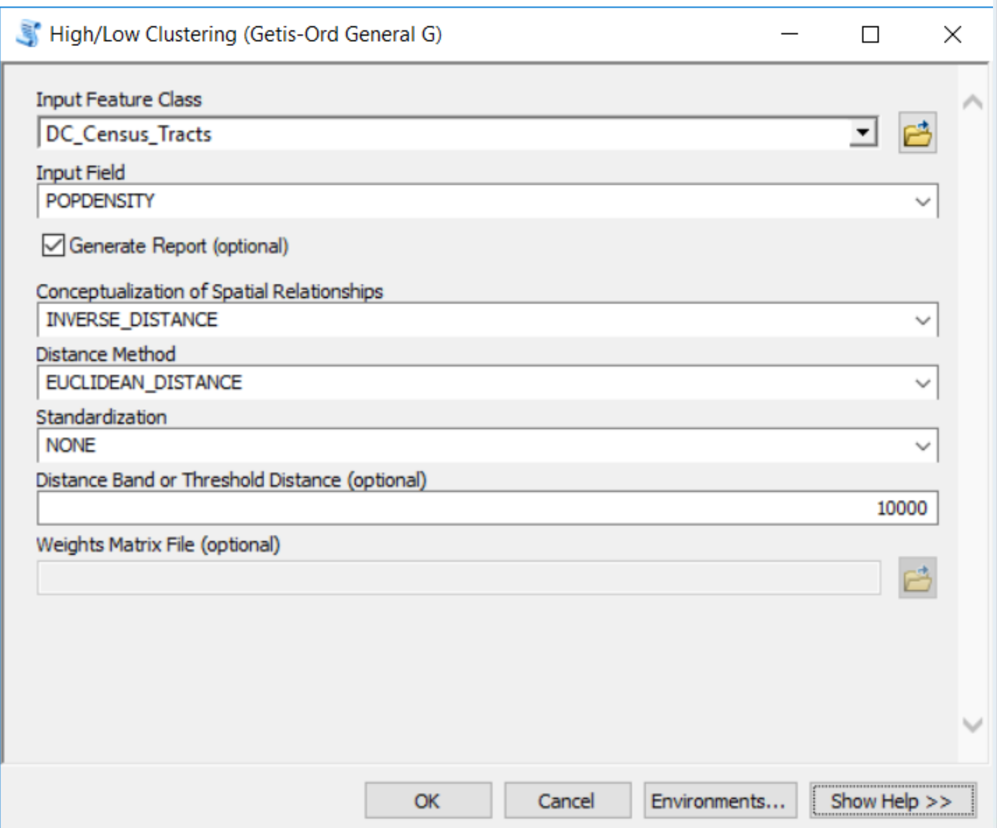
- Spatial Autocorrelation
  - Global Measures
    - Analyzing Pattern > Spatial Autocorrelation (Morans I)

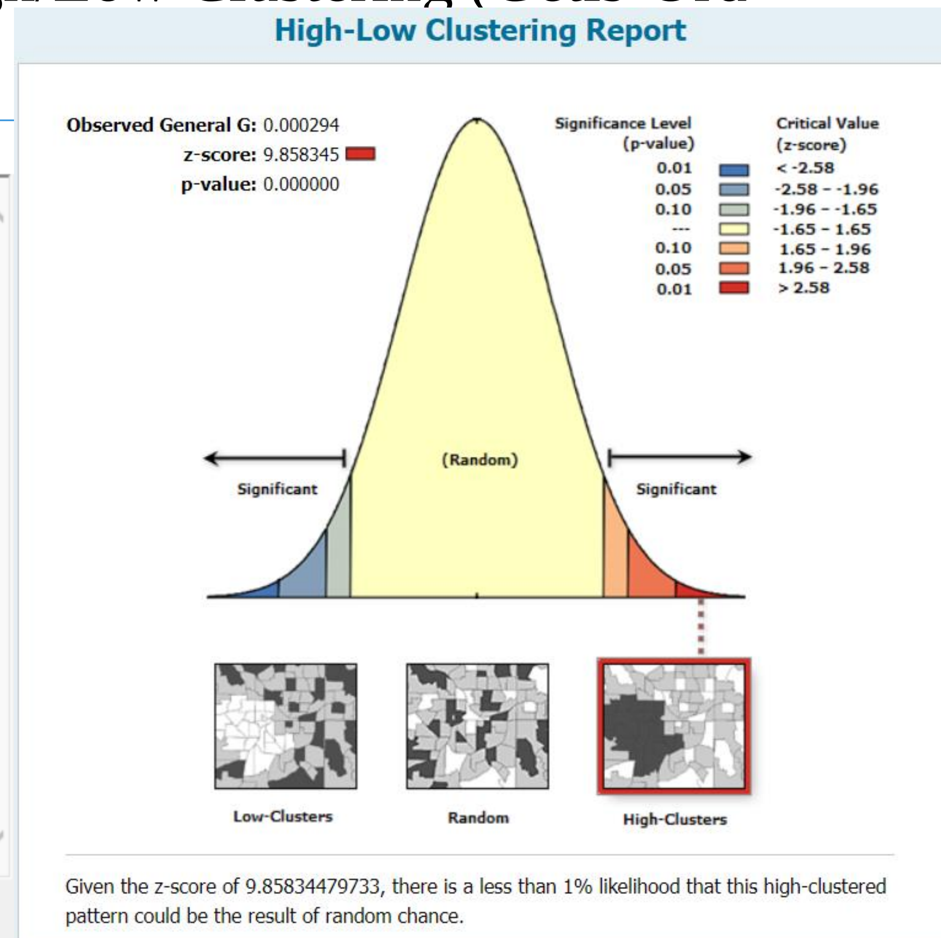




Areal Analysis

- Spatial Autocorrelation
  - Global Measures
    - Analyzing Pattern > High/Low Clustering (Getis-Ord General G)





Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Global measures:

	Data Type	Function	Strength	Limitations
Join Count	Categories	Whether values are clustered or dispersed	Straightforward way to identify patterns for areas	Only applies to categorical (nominal) data
Geary's C Moran's I	Continuous	Similarity of nearby features	Provides a single statistic summarizing the pattern	Doesn't indicate if clustering is for high values or low values
General G	Continuous	Concentration of high/low values	Indicates whether high or low values are clustered	Works best when either high or low values cluster

Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Global measures are useful to tell if there is clustering or not. However, they cannot tell you where. To answer these questions, we need local measures.

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Local measures:
      - Local Moran's I
      - Local Geary's C
      - Local G-statistic

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Local measures:
      - Local Moran's I

$$I_i = z_i \sum_j w_{ij} z_j$$

$$\text{where } z_i = \frac{x_i - \bar{x}}{\delta}$$

where $\bar{x}$ and δ are mean and standard deviation of x

» High values of local Moran's I indicates clustering of similar values.

» Low values of local Moran's I indicates clustering of dissimilar values.

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Local measures:
      - Local Moran's I

The Local Moran's I statistic of spatial association is given as:

$$I_i = \frac{x_i - \bar{X}}{S_i^2} \sum_{j=1, j \neq i}^{n} w_{i,j}(x_j - \bar{X}) \qquad (1)$$

where x_i is an attribute for feature i, $\bar{X}$ is the mean of the corresponding attribute, $w_{i,j}$ is the spatial weight between feature i and j, and:

$$S_i^2 = \frac{\sum_{j=1, j \neq i}^{n} (x_j - \bar{X})^2}{n - 1} \qquad (2)$$

with n equating to the total number of features.

The z_{I_i}-score for the statistics are computed as:

$$z_{I_i} = \frac{I_i - \mathrm{E}[I_i]}{\sqrt{\mathrm{V}[I_i]}} \qquad (3)$$

where:

$$\mathrm{E}[I_i] = -\frac{\sum_{j=1, j \neq i}^{n} w_{ij}}{n - 1} \qquad (4)$$

$$\mathrm{V}[I_i] = \mathrm{E}[I_i^2] - \mathrm{E}[I_i]^2 \qquad (5)$$

A detailed mathematic description

Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Local measures:
      - Local Moran’s I
        - » In ArcGIS, the tool - Cluster and Outlier Analysis (Anselin Local Moran's I).
        - » Given a set of weighted features, identifies statistically significant hot spots, cold spots, and spatial outliers using the Anselin Local Moran's I statistic.

Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Local measures:
      - Local Geary's C

$$c_i = \sum_j w_{ij} (z_i - z_j)^2$$

» Relatively low measure indicates clustering of similar values.

» Relatively high measure indicates clustering of dissimilar values.

Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Local measures:
      - In general, the results from local Moran's I are more useful than those from local Geary's C.
      - The significance test for Moran's I is also more reliable.

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Local measures:
      - Local G-Statistic

The Getis-Ord local statistic is given as:

$$G_i^* = \frac{\sum_{j=1}^{n} w_{i,j} x_j - \bar{X} \sum_{j=1}^{n} w_{i,j}}{S \sqrt{\frac{\left[n \sum_{j=1}^{n} w_{i,j}^2 - \left(\sum_{j=1}^{n} w_{i,j} \right)^2 \right]}{n-1}}} \qquad (1)$$

where x_j is the attribute value for feature j, $w_{i,j}$ is the spatial weight between feature i and j, n is equal to the total number of features and:

$$\bar{X} = \frac{\sum_{j=1}^{n} x_j}{n} \qquad (2)$$

$$S = \sqrt{\frac{\sum_{j=1}^{n} x_j^2}{n} - \left(\bar{X}\right)^2} \qquad (3)$$

The G_i^* statistic is a z-score so no further calculations are required.

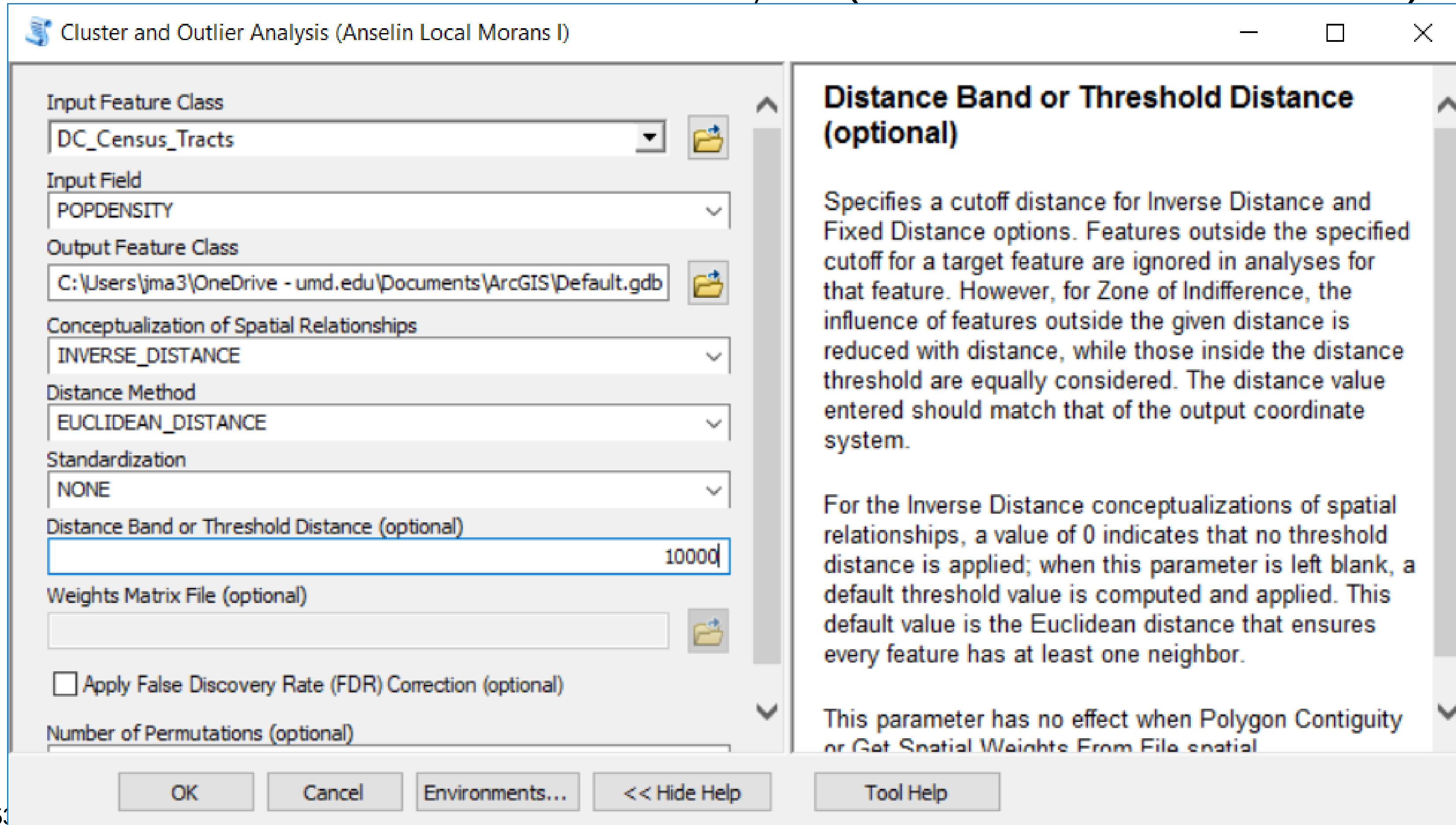
- Spatial Autocorrelation
  - – Measuring Spatial Autocorrelation
    - Local measures:
      - – Local G-Statistic
        - » A group of features with high Gi* values indicates a cluster or concentration of features with high attribute values.
        - » On the other hand, a group of feature with low Gi* values indicates a cold spot.
        - » A GI* value near 0 indicates there is no concentration of either high or low values surrounding the target feature.
          - » This occurs when the surrounding values are near the mean, or when the target feature is surrounded by a mix of high and low values.

Areal Analysis

- Spatial Autocorrelation
  - Measuring Spatial Autocorrelation
    - Local measures:
      - Local G-Statistic
        - » In ArcGIS, it is the tool - Hot Spot Analysis (Getis-Ord Gi*).
        - » Given a set of weighted features, identifies statistically significant hot spots and cold spots using the Getis-Ord Gi* statistic.

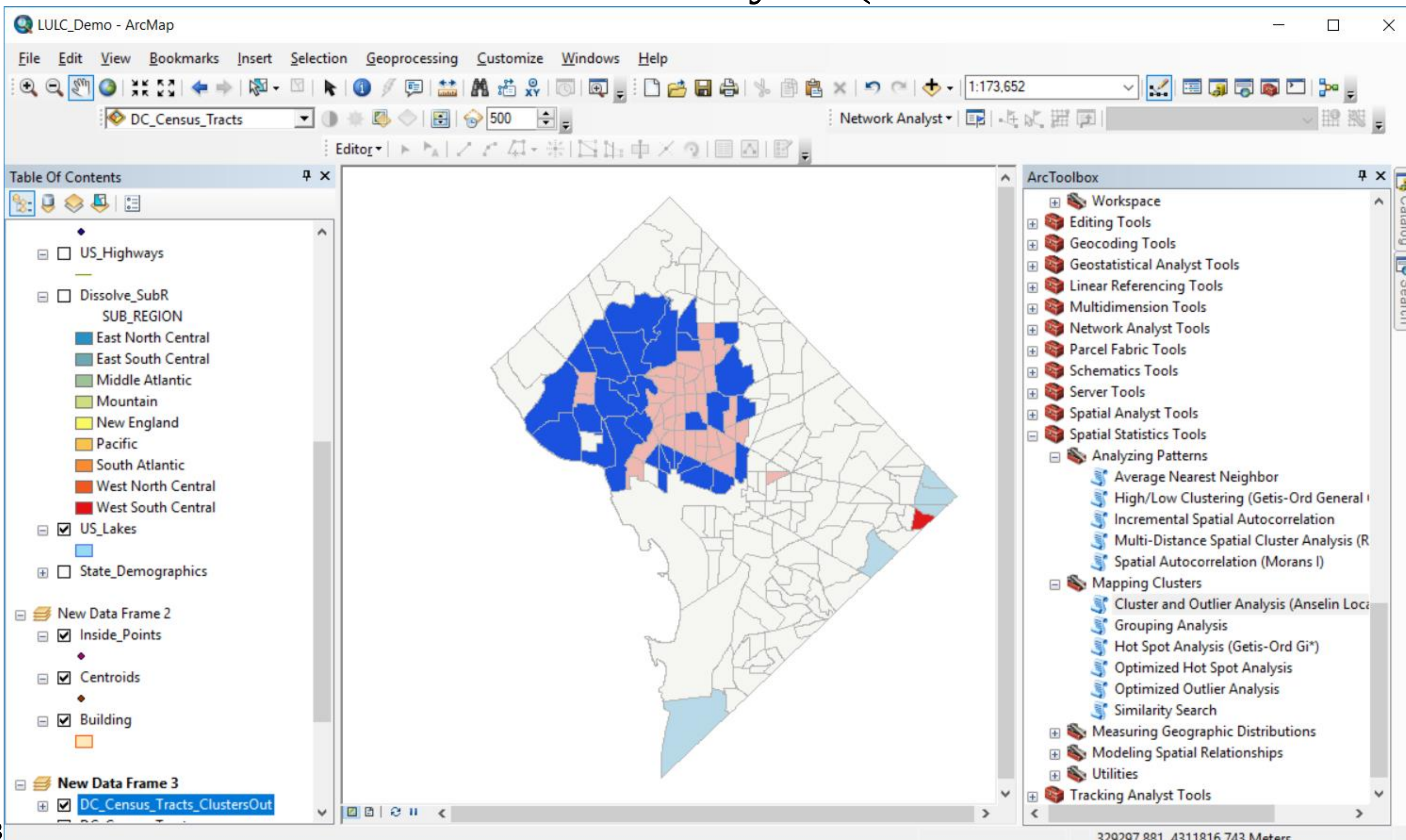
Areal Analysis

- Spatial Autocorrelation
  - Local Measures
    - Cluster and Outlier Analysis (Anselin Local Morans I)



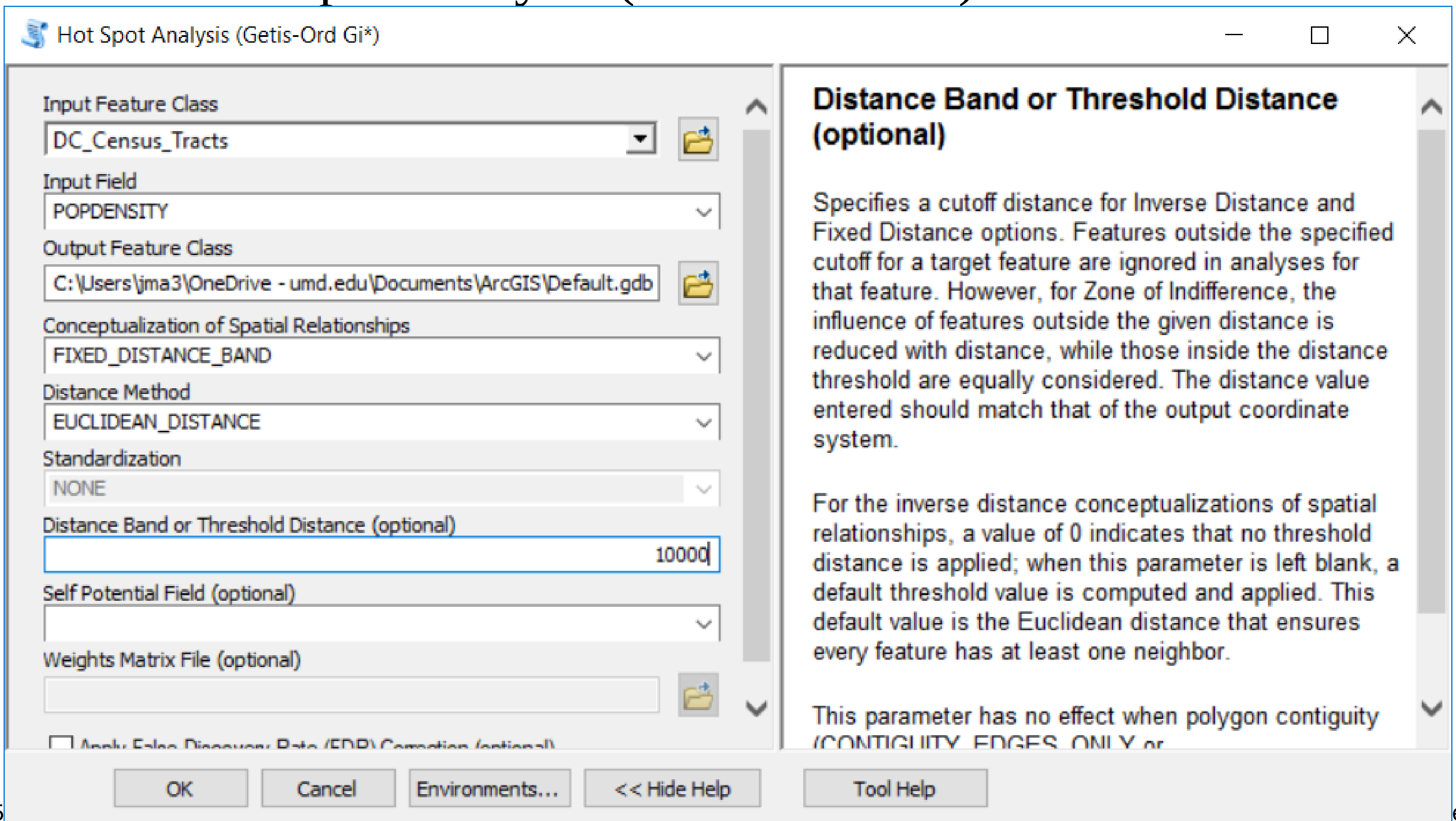
Areal Analysis

- Spatial Autocorrelation
  - – Local Measures
    - Cluster and Outlier Analysis (Anselin Local Morans I)



Areal Analysis

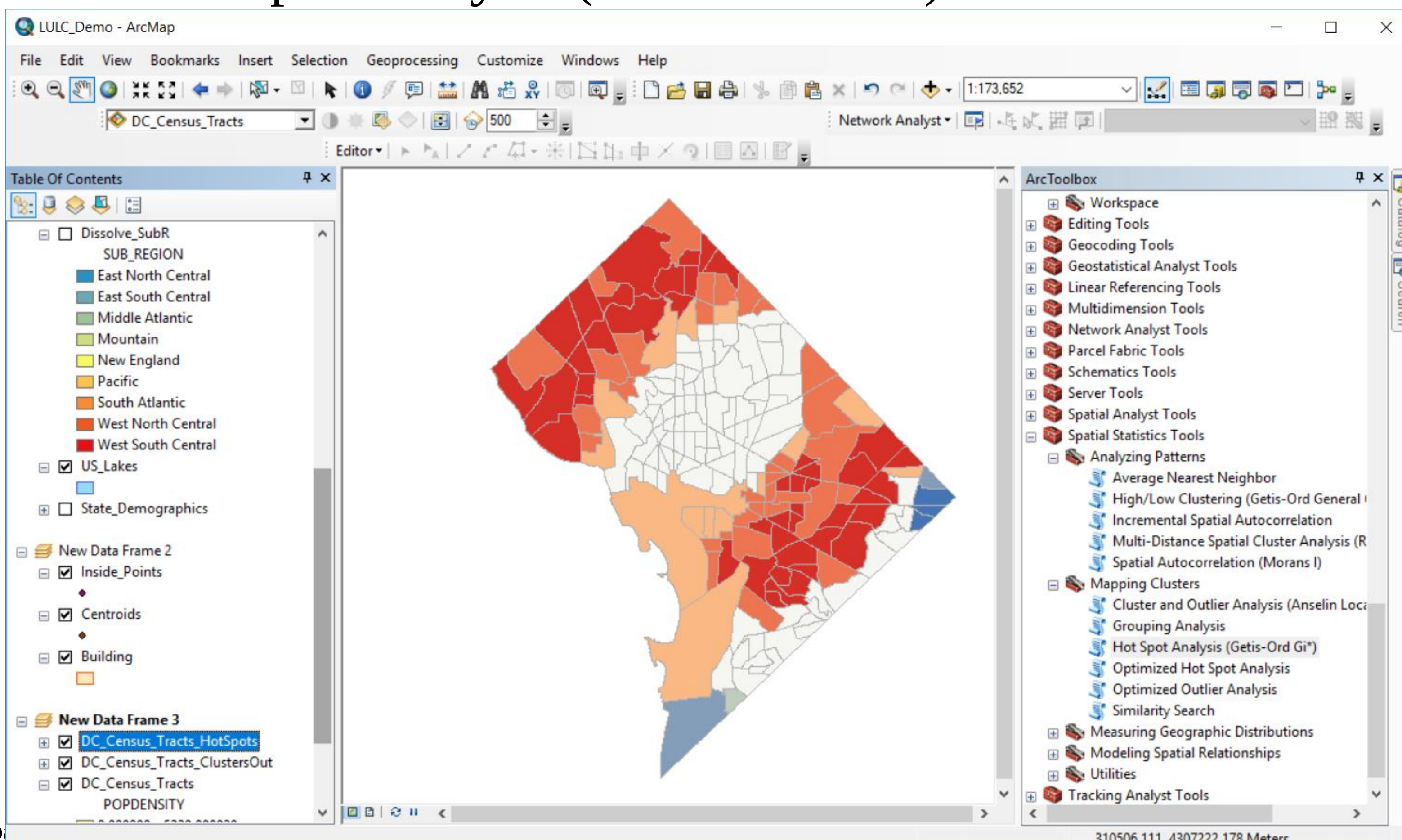
- Spatial Autocorrelation
  - Local Measures
    - Hot Spot Analysis (Getis-Ord Gi*)



Areal Analysis

- Spatial Autocorrelation
  - Local Measures
    - Hot Spot Analysis (Getis-Ord Gi*)

95



Areal Analysis

- Demos
  - Geometric measurements
  - Spatial Autocorrelation
    - Global Measures
      - Join Count Analysis
      - Moran's I
        - » Manual calculations
        - » Software verifying
      - General G Statistic
    - Local Measures
      - Moran's I
      - G Statistic
  - Similarities and differences (vs. Point Pattern Analysis)

THE END