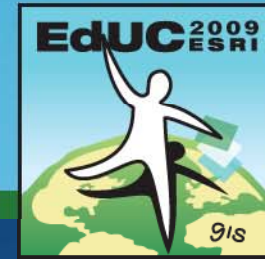


2009 ESRI Education User Conference
July 11–14, 2009



Please!
Turn OFF cell phones
and paging devices



Introduction to Spatial Statistics

Opportunities for Education

Lauren M. Scott, PhD
Mark V. Janikas, PhD
Lauren Rosenshein
Jorge Ruiz-Valdepeña

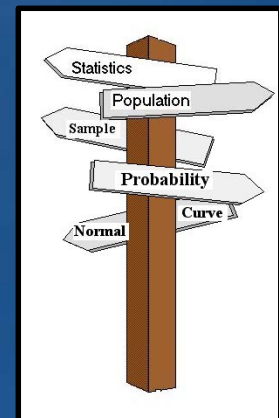
Objectives

- Define spatial statistics
- Introduce you to some of the core spatial statistics tools available in ArcGIS 9.3
- Present a variety of example applications
- Give you the opportunity for hands-on practice
- Interact with you and get feedback

DEMO

- The spatial pattern of piracy

Kindly complete an evaluation form before leaving.



What are spatial statistics in a GIS environment?

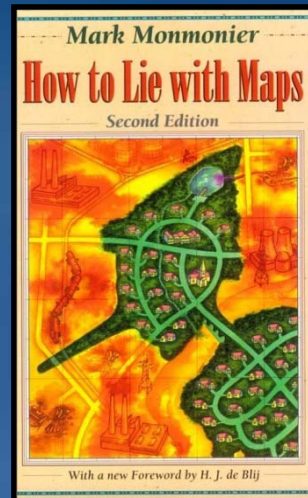
- Software-based tools, methods, and techniques developed specifically for use with geographic data.
- Spatial statistics:
 - Describe and model spatial distributions, spatial patterns, spatial processes, and spatial relationships.
 - Incorporate space (area, length, proximity, orientation, and/or spatial relationships) directly into their mathematics.

In many ways spatial statistics extend what the eyes and mind do intuitively to assess spatial patterns, trends and relationships.

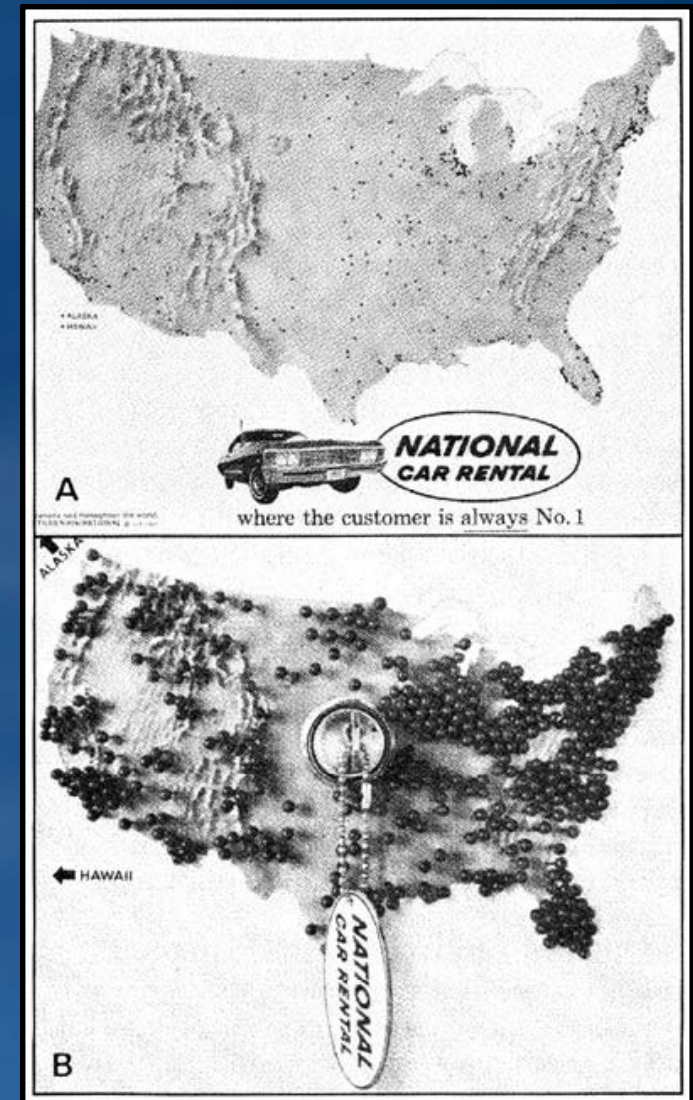
Why use spatial statistics?

Spatial Statistics help us assess:

- Patterns
- Relationships
- Trends



How we present our results (colors, class breaks, symbols...) can either enhance or obscure communication.



Two ads for National Car Rental. The lower map ad replaced the upper map ad a year later.

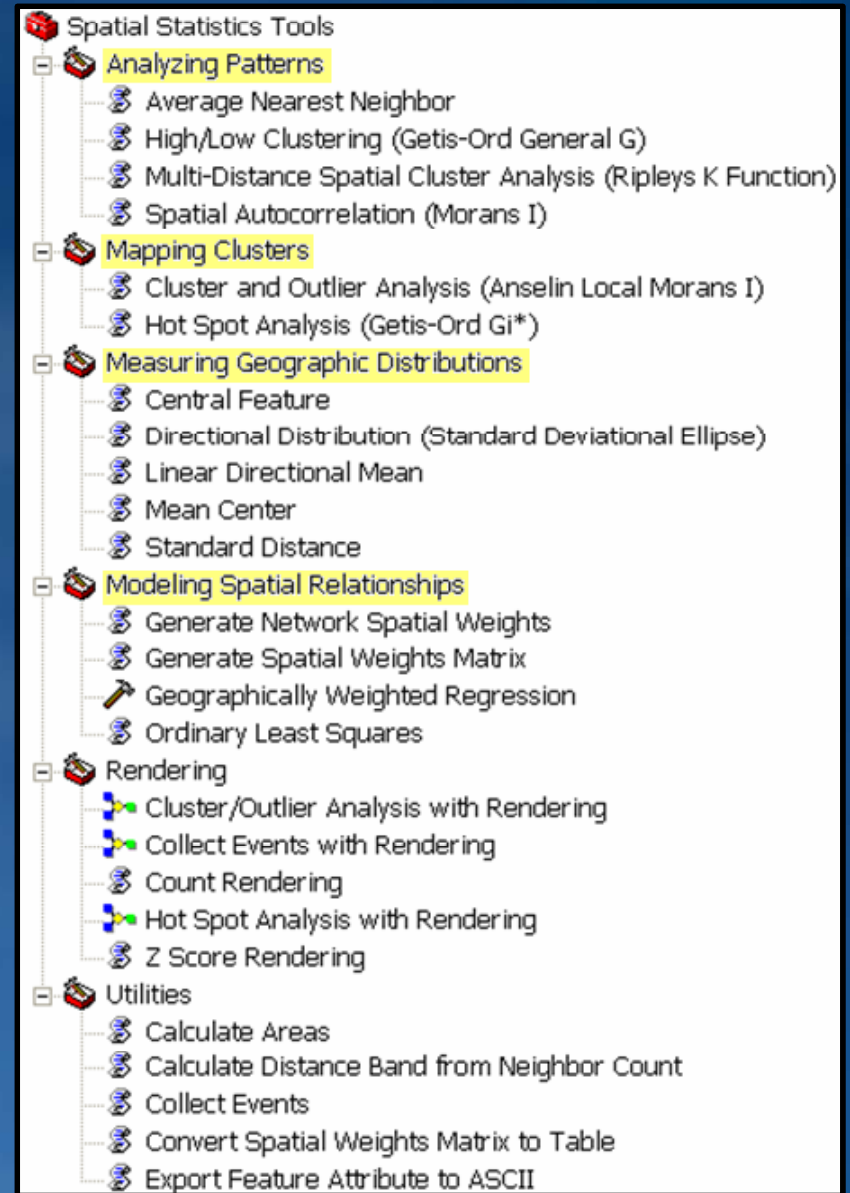
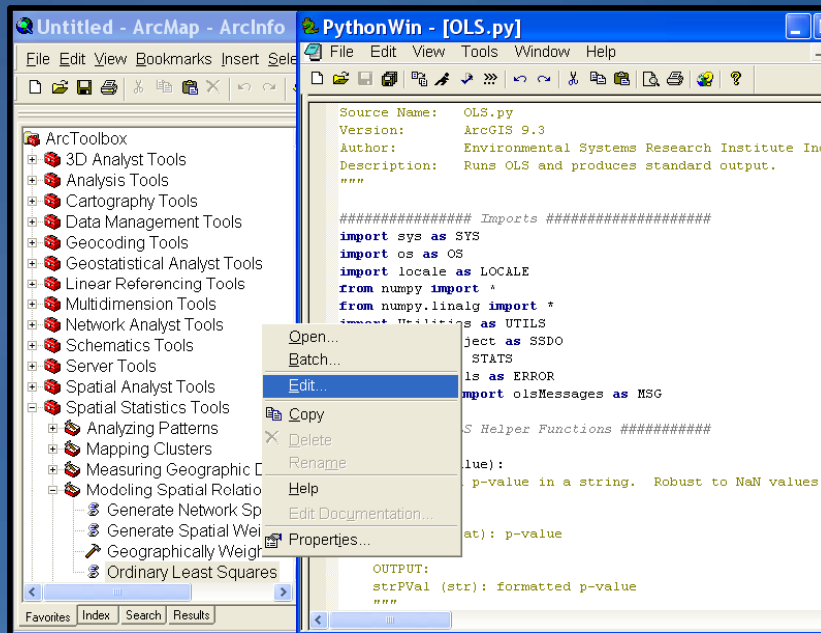
Spatial Statistics Do a Variety of Things

- Describe and summarize spatial distributions
- Quantify Shape
- Measure spatial dependency / spatial autocorrelation
- Assess spatial relationships
- Model surfaces

- Focus today is on Spatial Pattern Analysis:
 - Describing and summarizing spatial distributions:
 - Mean Center
 - Standard Deviational Ellipse
 - Assessing Global and Local Spatial Autocorrelation:
 - Average Nearest Neighbor
 - Global Moran's I
 - Hot Spot Analysis (Getis-Ord G_i^*)
 - Cluster/Outlier Analysis (Anselin Local Moran's I)

Spatial Statistics Toolbox in ArcGIS

- Core functionality with ArcGIS (not an extension).
- Most tools delivered with their source code.
- Most tools available at all license levels.



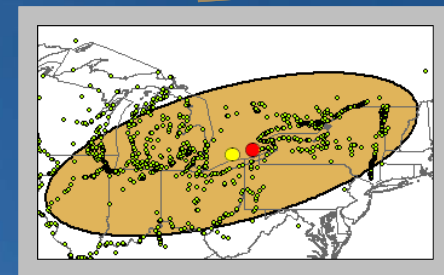
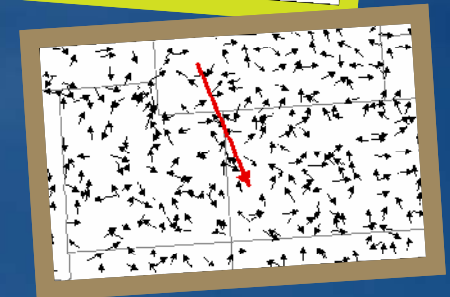
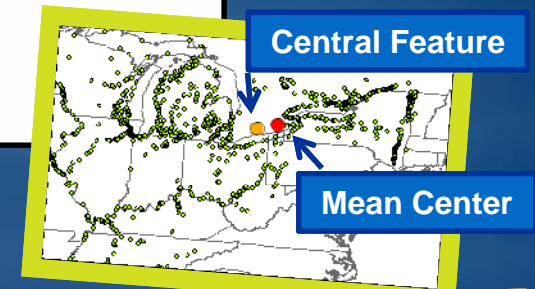
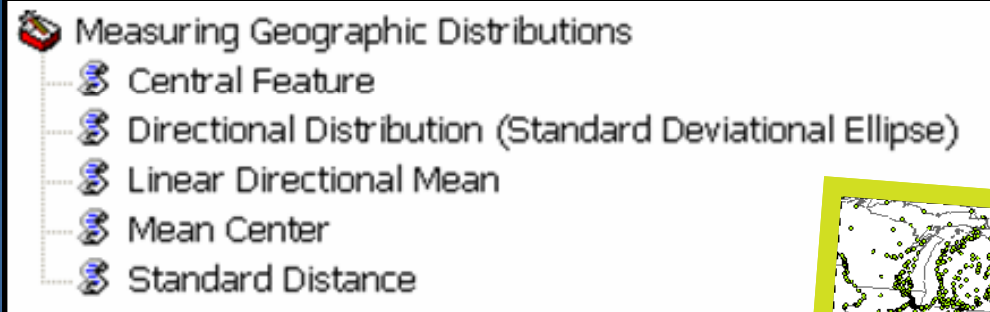
Spatial Statistics Toolbox

Measuring Geographic Distribution

Analyzing Patterns

Mapping Clusters

Modeling Spatial Relationships



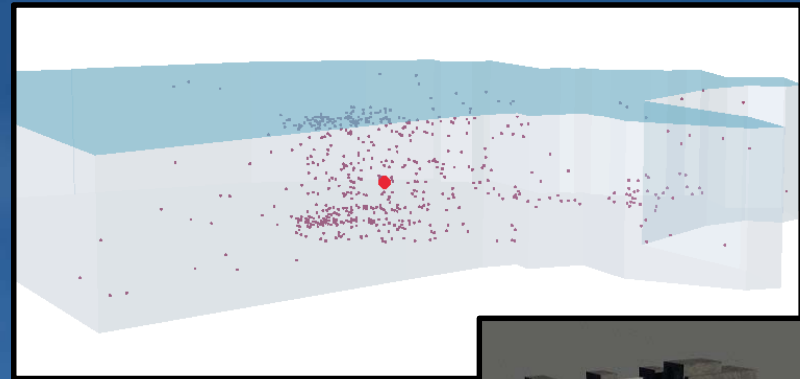
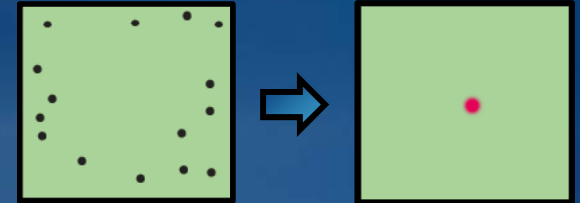
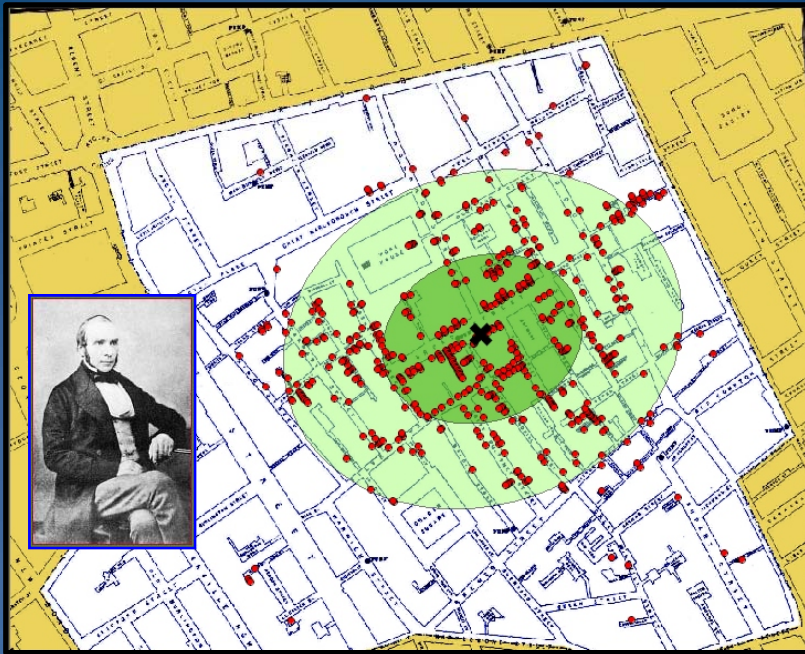
• Questions

- Which site is most accessible?
- Is there a directional trend to the spatial distribution of the incidents?
- What is the primary wind direction for this region in the winter?
- Where is the population center?
- Which gang has the broadest territory?

Mean Center

- Computes the average x and y coordinate, based on all features in the study area.

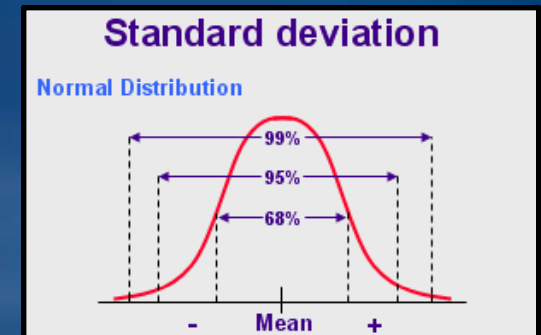
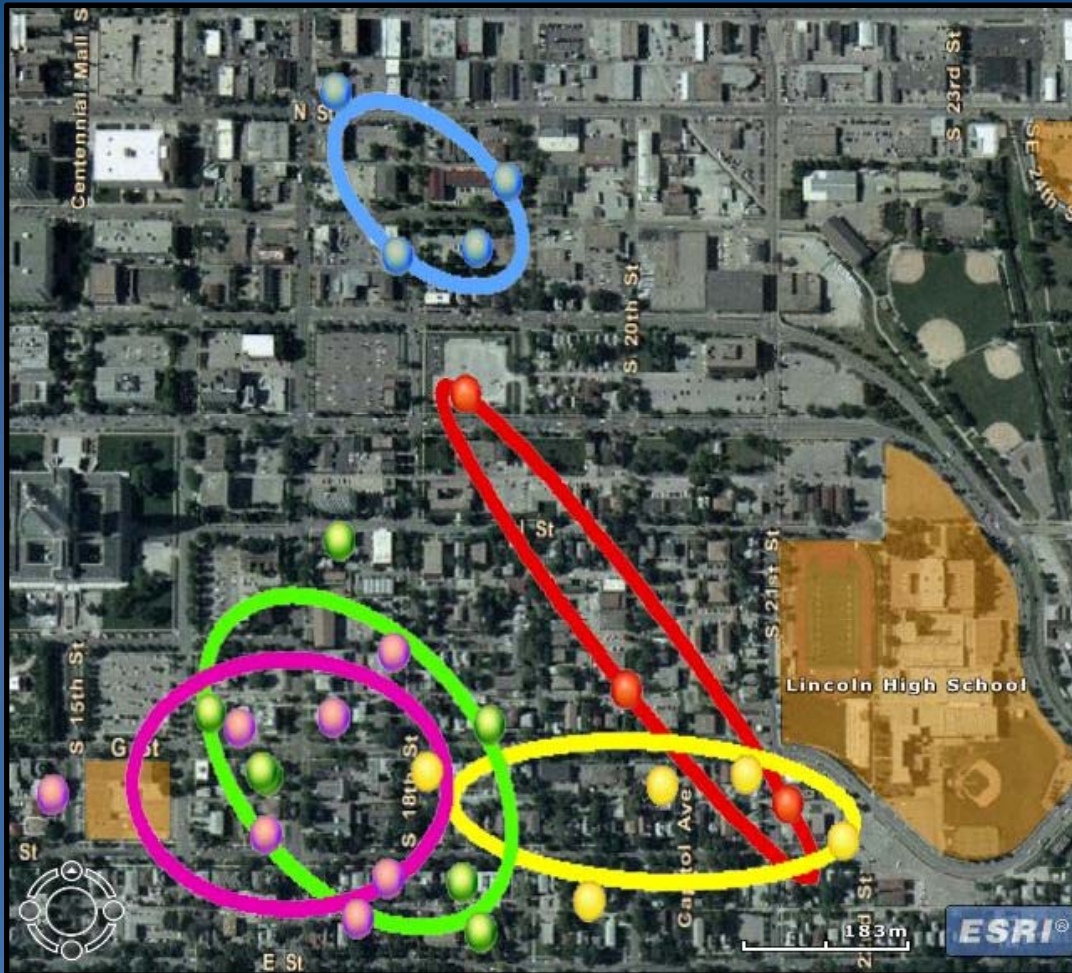
$$\bar{X} = \frac{\sum_{i=1}^n x_i}{n}, \quad \bar{Y} = \frac{\sum_{i=1}^n y_i}{n}$$



Mean Center Animation

Directional Distribution (Standard Deviational Ellipse)

- Abstracting spatial trends in a distribution of features
- Comparing distributions over time



- 1 = 68% of features
- 2 = 95% of features
- 3 = 99% of features

Standard Distance

Input Feature Class

Output Standard Distance Feature Class

Circle Size
1 Standard Deviation
2 Standard Deviations
3 Standard Deviations

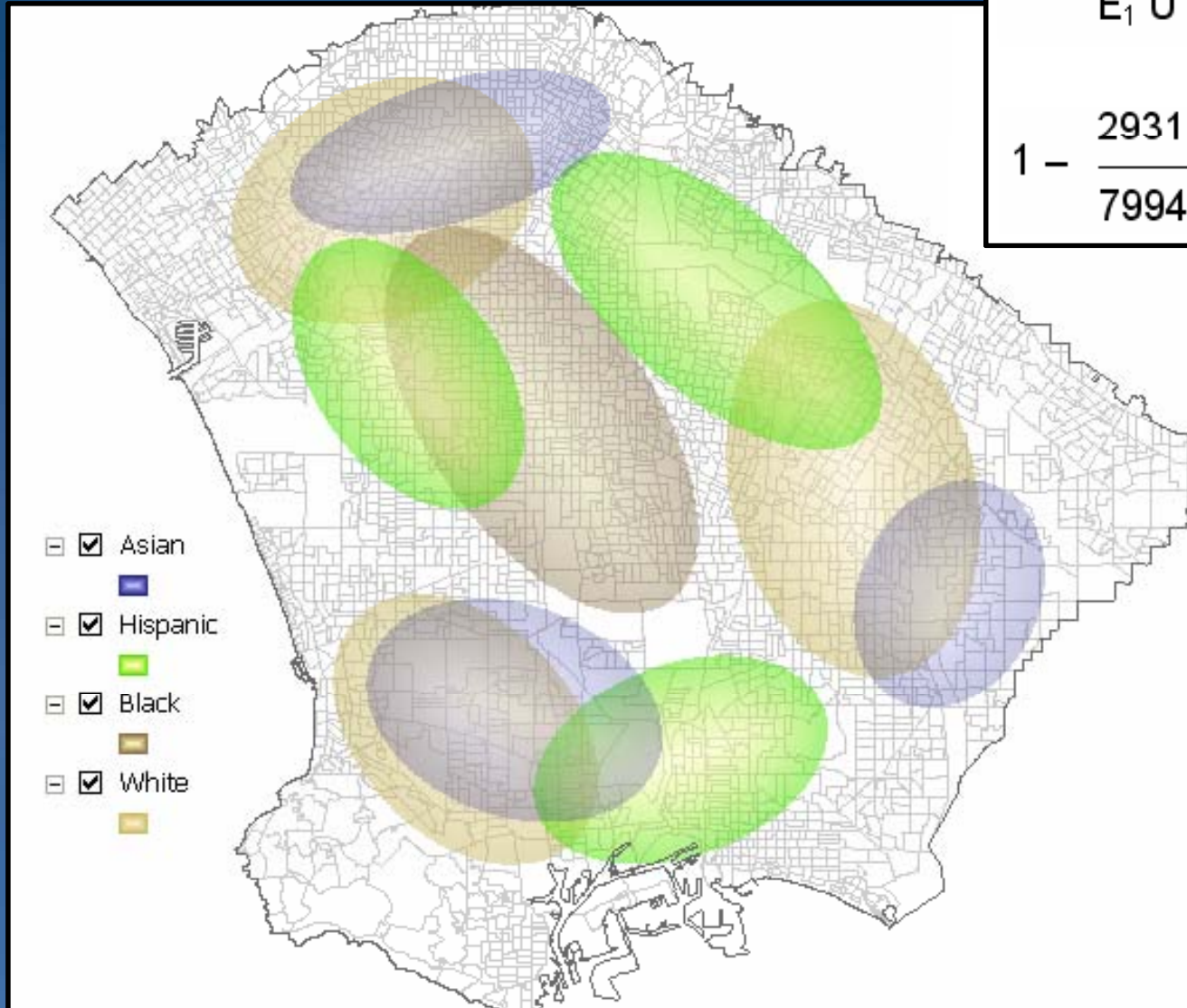
Case Field (optional)

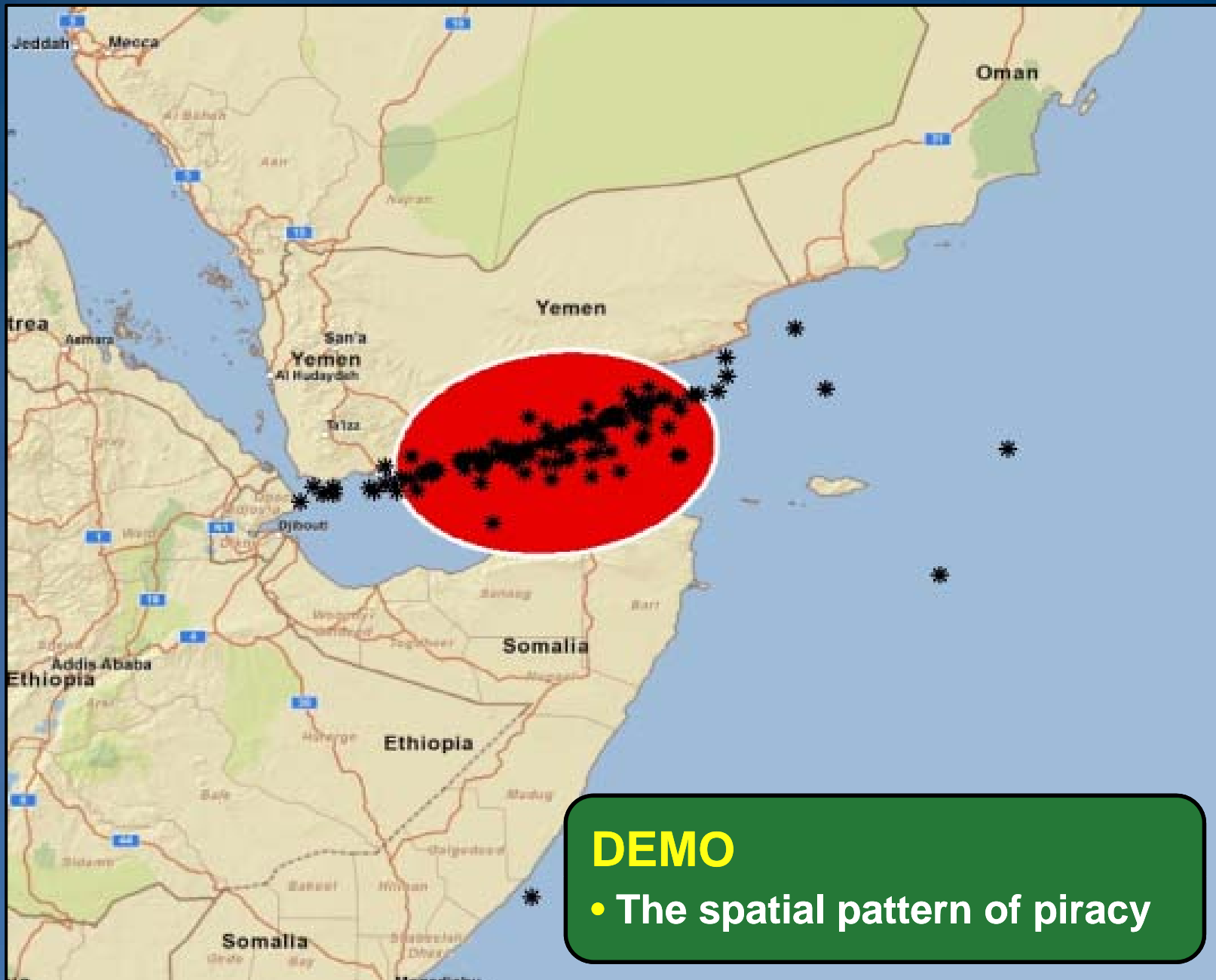
Directional Distribution (Standard Deviational Ellipse)

Segregation Index =

$$1 - \frac{E_1 \cap E_2 \cap E_3 \cap \dots \cap E_n}{E_1 \cup E_2 \cup E_3 \cup \dots \cup E_n} =$$

$$1 - \frac{2931680545.83}{7994760004.92} = 0.63$$





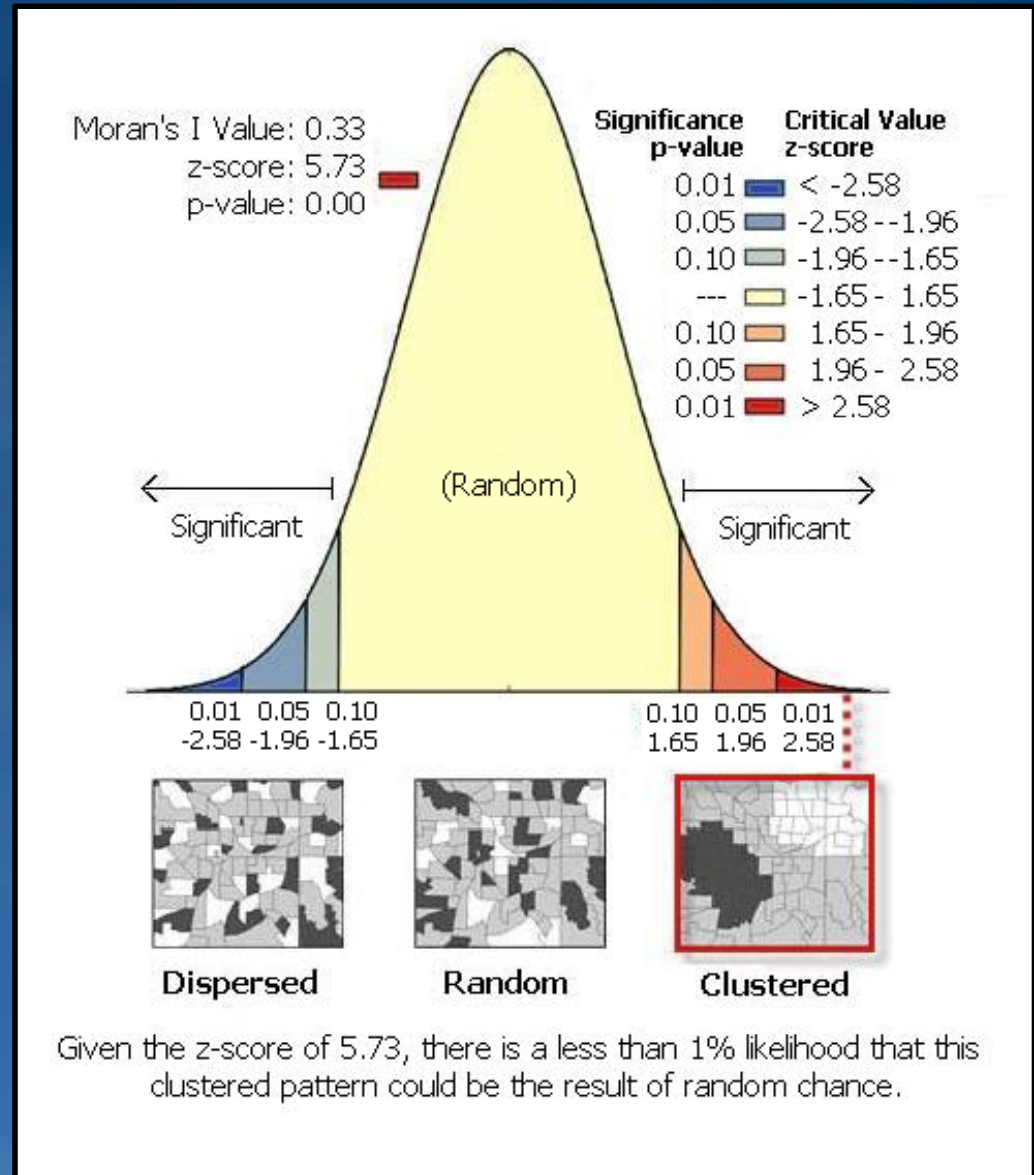
DEMO

- The spatial pattern of piracy

What is a z score? What is a p-value?

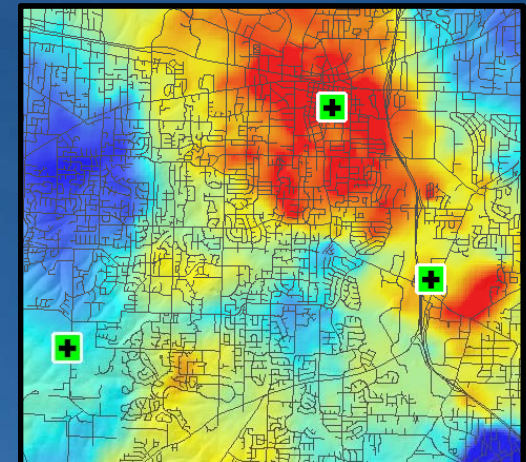
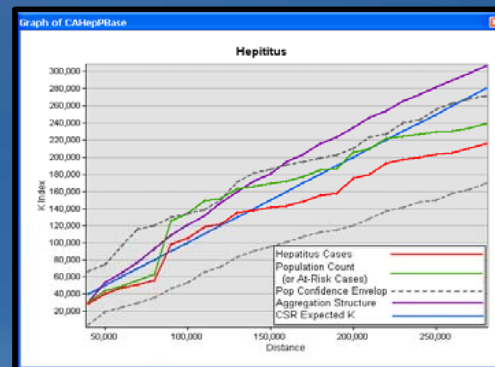
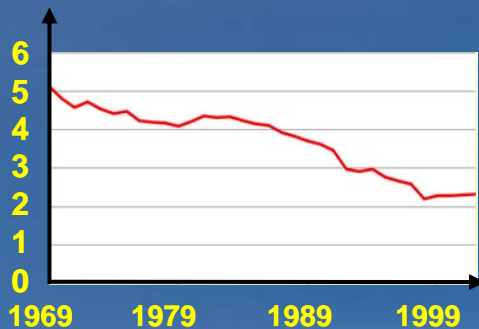
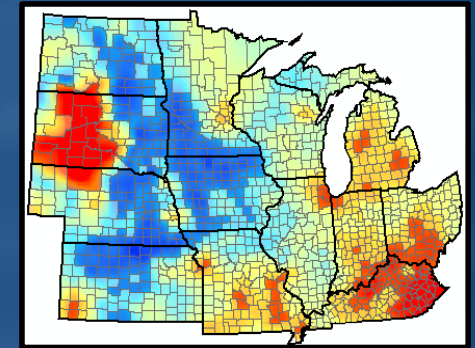
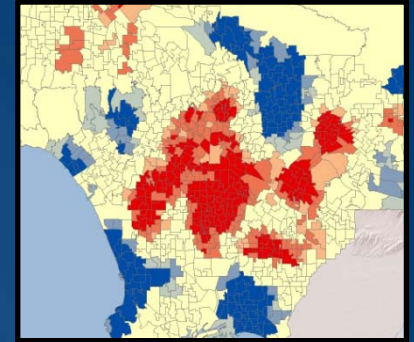
- The null hypothesis for the ArcGIS Spatial Pattern Analysis tools is CSR: Complete Spatial Randomness
- Reject the null hypothesis if the result (the p-value/z score) is statistically significant

Z Score (Standard Deviations)	P-Value (Probability)	Confidence Level
+/-1.65	0.10	90%
+/-1.96	0.05	95%
+/-2.58	0.01	99%



Is CSR useful?

- Raising the bar:
 - Normalize the analysis field to create a rate
 - Analyze average values
 - Compare z score magnitudes
 - Across space
 - Over time
 - Among control spatial distributions



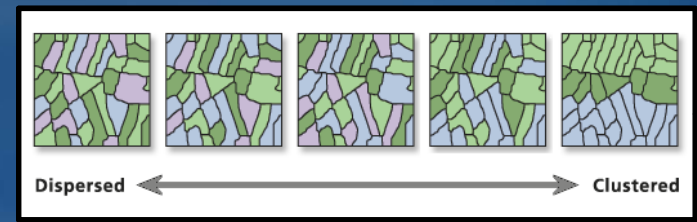
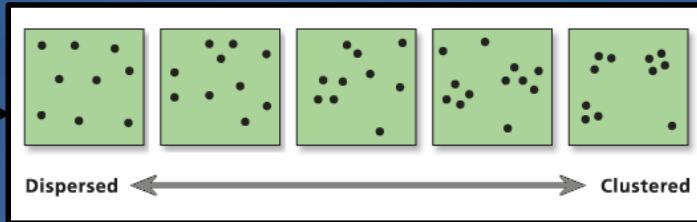
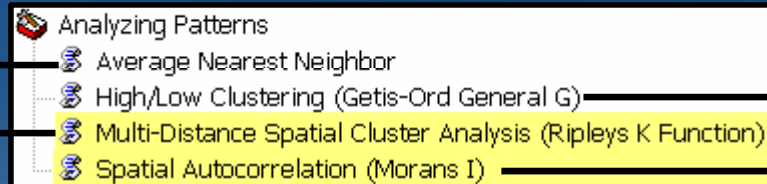
Spatial Statistics Toolbox

Measuring Geographic Distribution

Analyzing Patterns

Mapping Clusters

Modeling Spatial Relationships

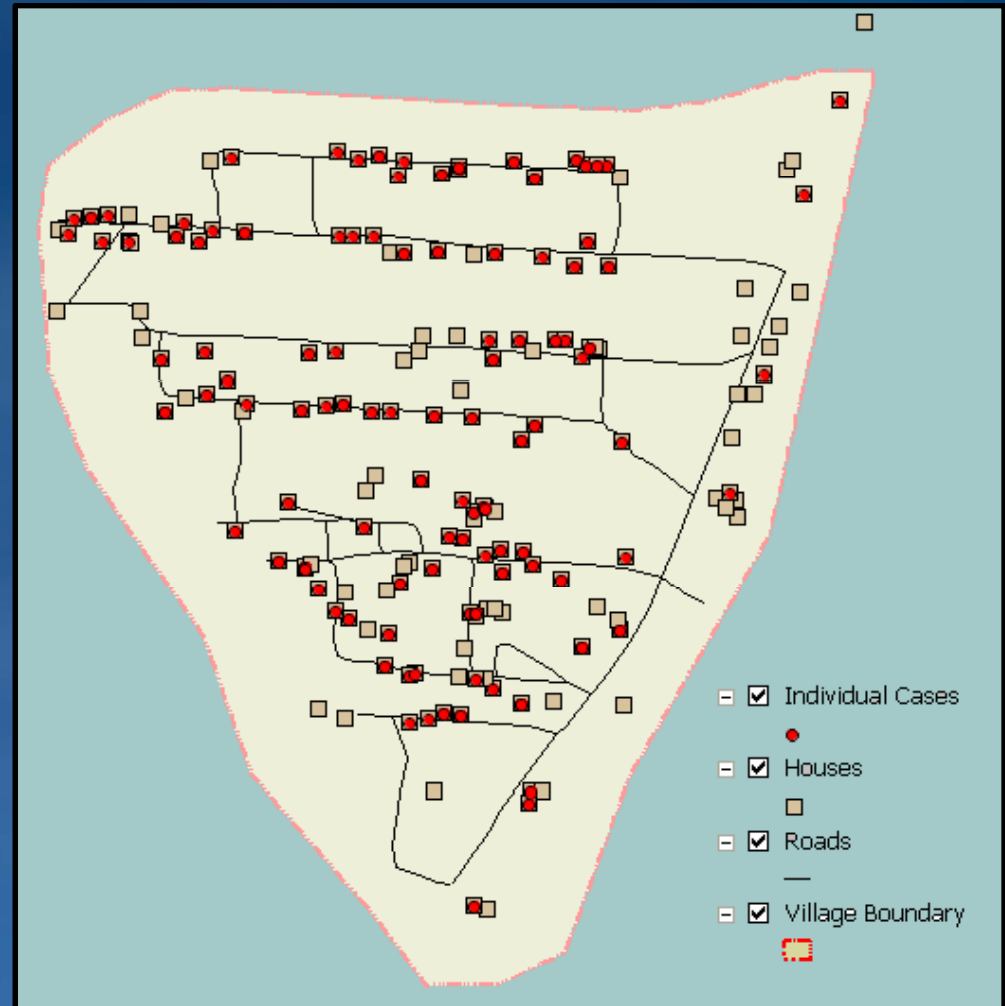
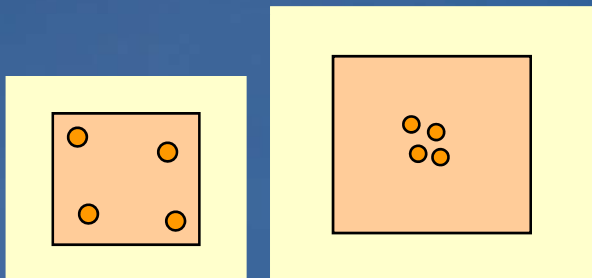


- Which plant species is most concentrated?
- Does the spatial pattern of the disease mirror the spatial pattern of the population at risk?

- Is there an unexpected spike in pharmaceutical purchases?
- Are new AIDs cases remaining Geographically fixed?

Average Nearest Neighbor

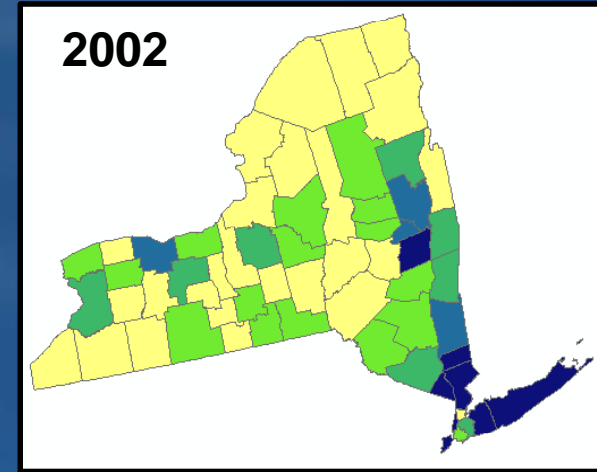
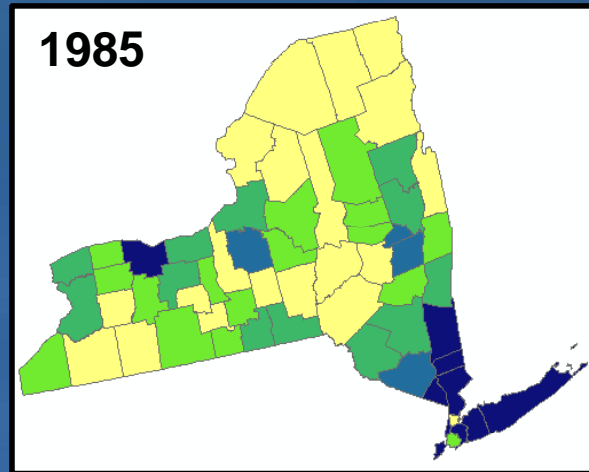
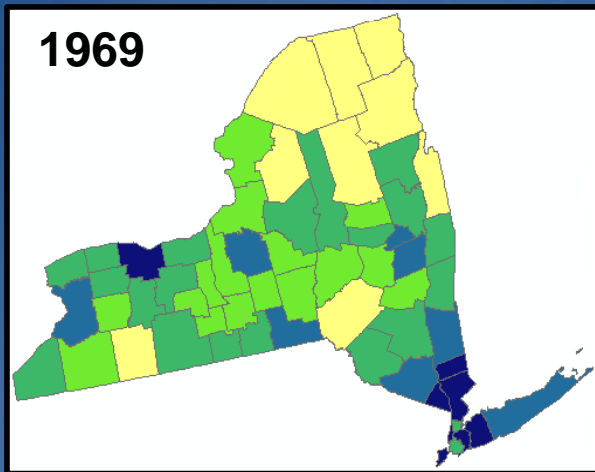
- Calculates the average distance between each feature
 - Considers arrangement and distance of features
- Based on **area**



Spatial Autocorrelation (Global Moran's I)

- This tool measures spatial clustering/dispersion
- Results are based on both feature locations and attributes

Thematic Maps showing Relative Per Capita Income for New York, 1969 to 2002

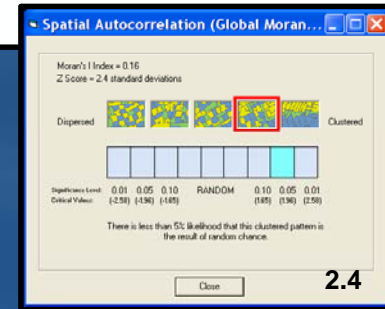
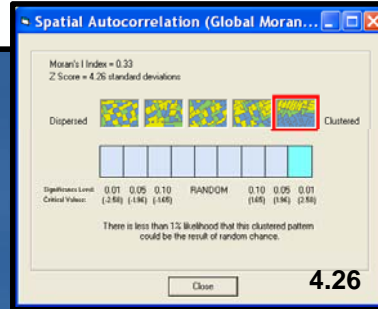
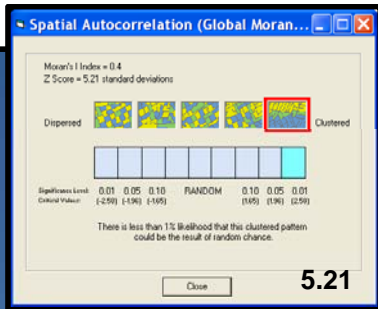
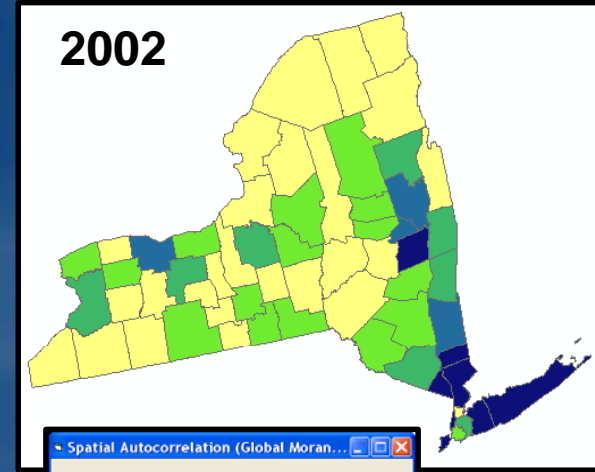
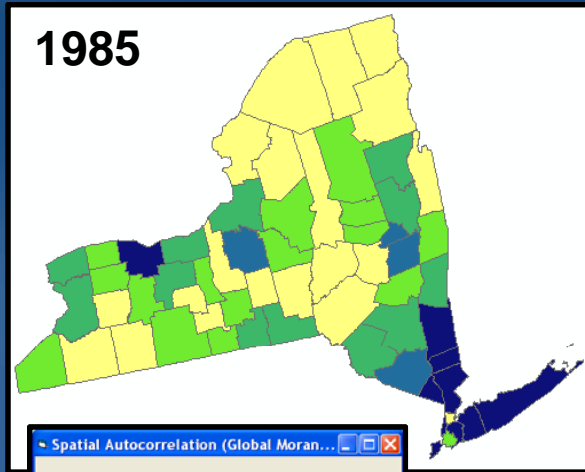
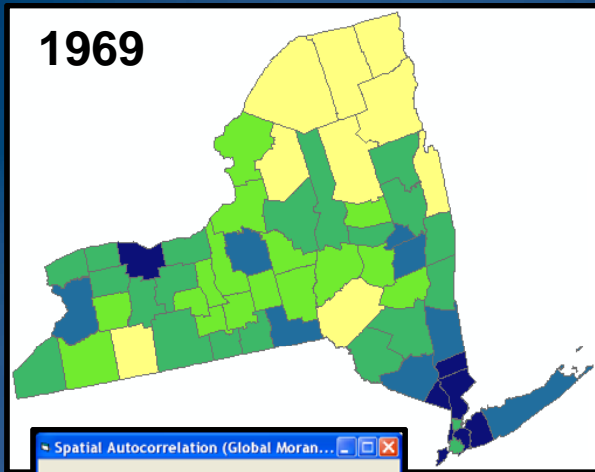


Are rich and poor becoming more or less spatially segregated?

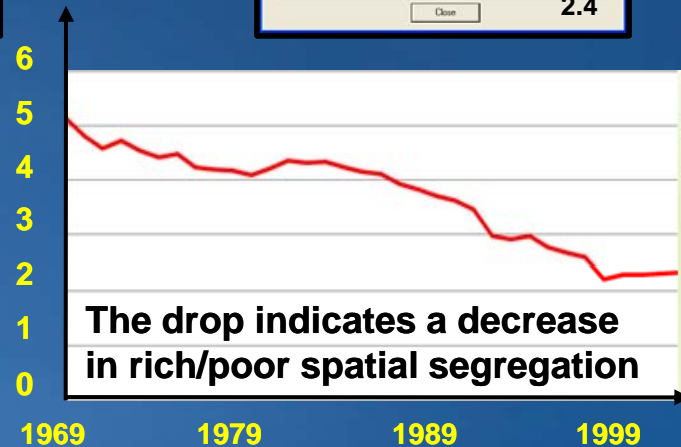
(It's difficult to answer this question looking at thematic maps alone).

Spatial Autocorrelation (Global Moran's I)

Relative Per Capita Income for New York, 1969 to 2002

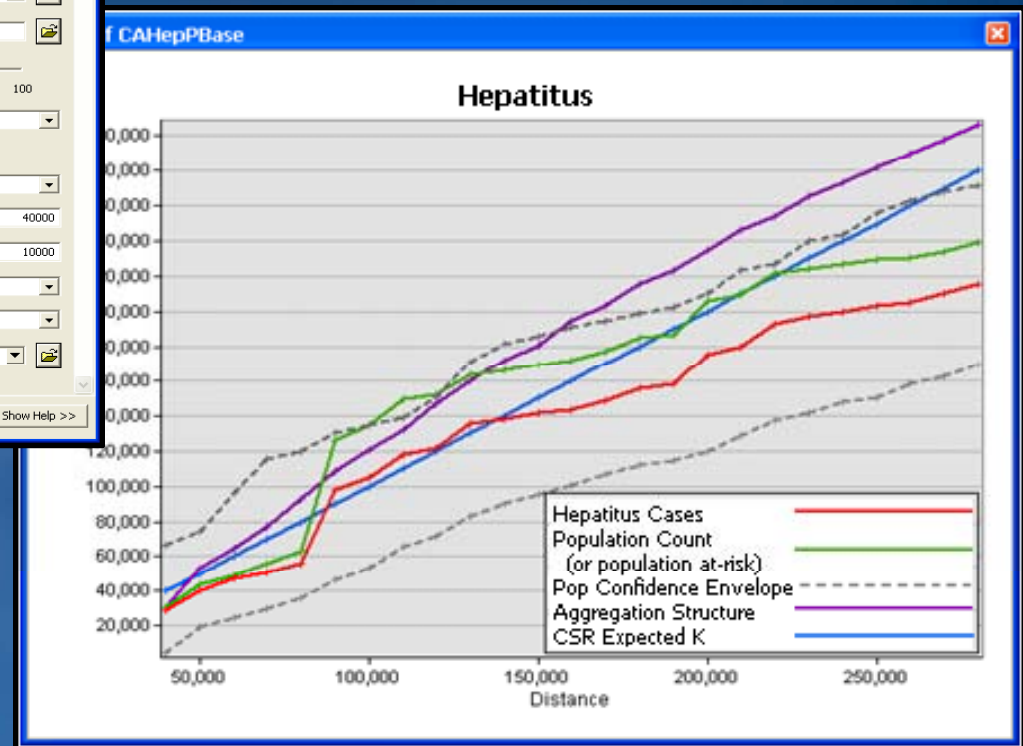
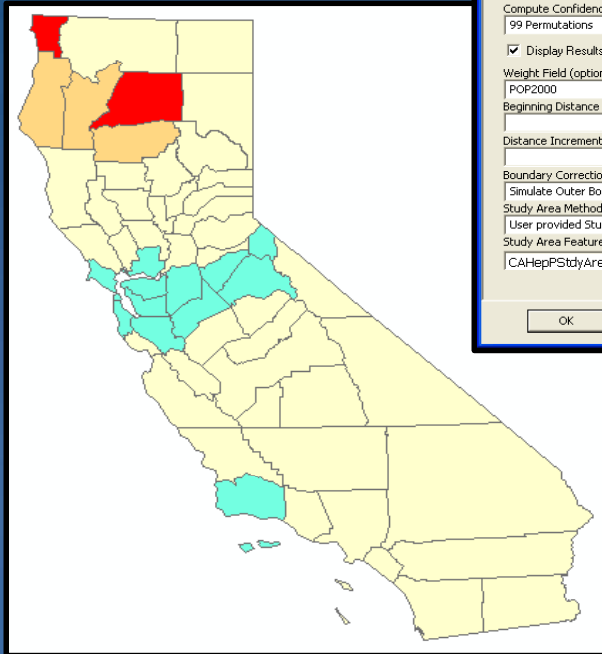
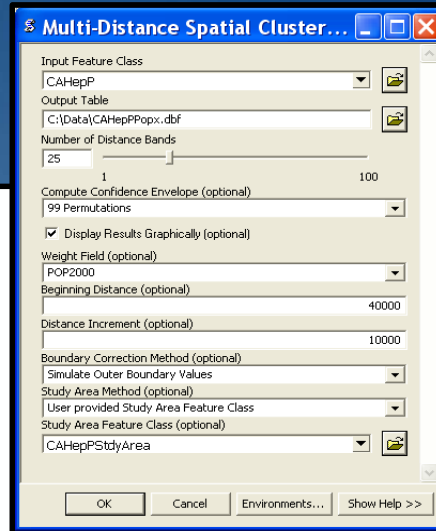
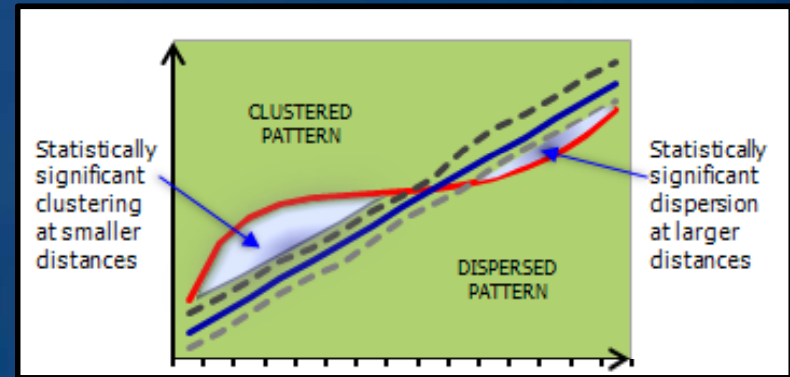


Plot the Z Score from the Global Spatial Autocorrelation tool to reveals broad trends over time.



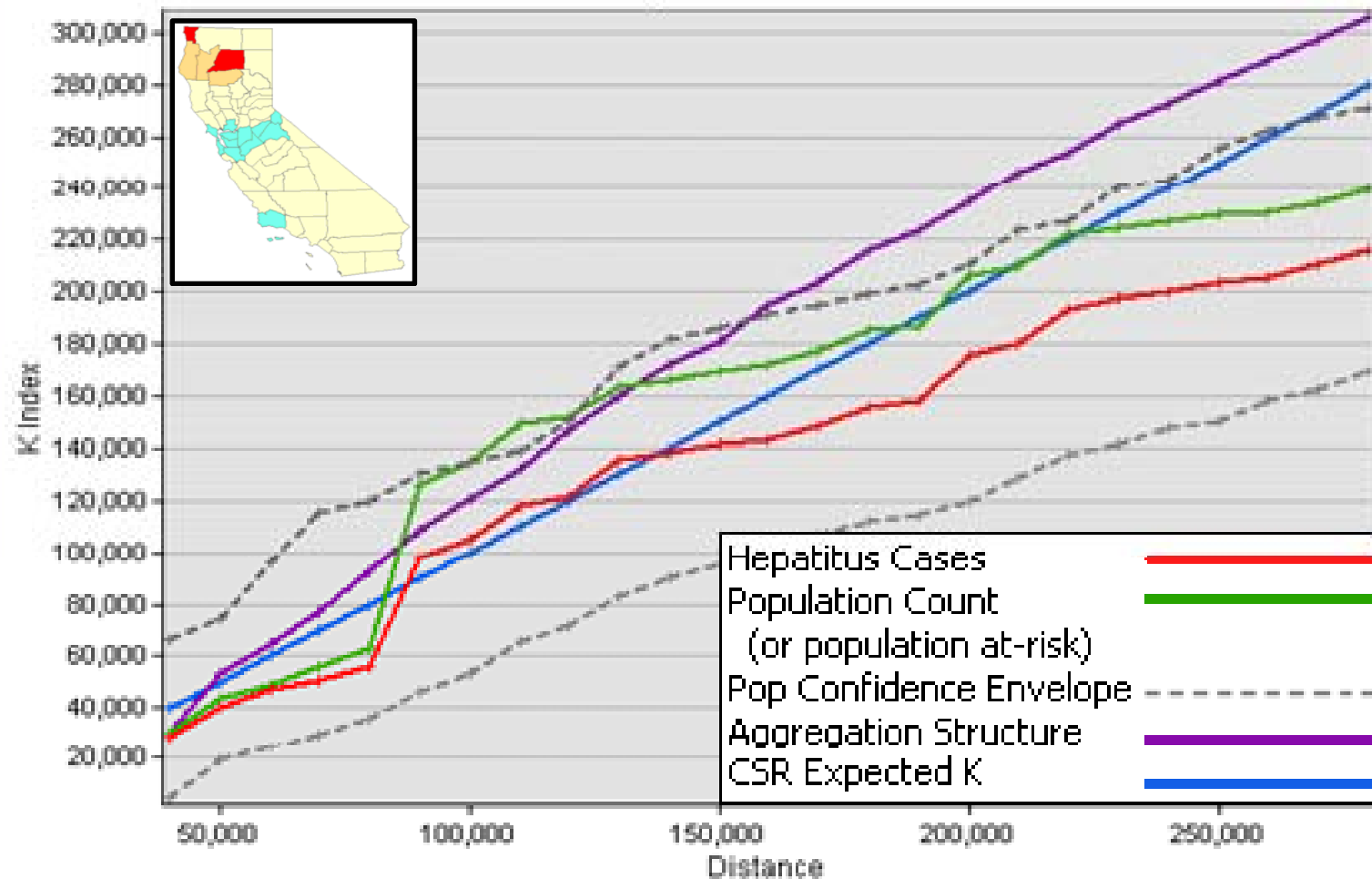
K Function

- Counts number of pairs within distance d of each feature.



The processes promoting hepatitis (red line) are strongly influenced by the spatial pattern of population (green line).

Hepatitis



The processes promoting hepatitis (red line) are strongly influenced by the spatial pattern of population (green line).

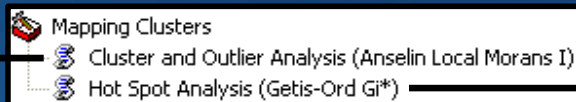
Spatial Statistics Toolbox

Measuring Geographic Distribution

Analyzing Patterns

Mapping Clusters

Modeling Spatial Relationships

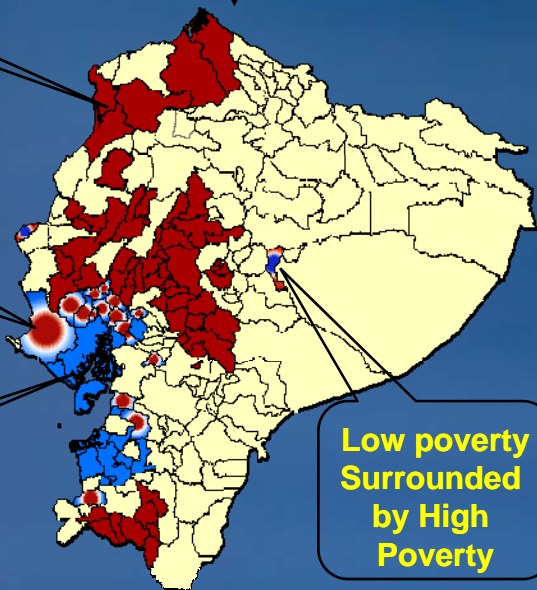


High Poverty

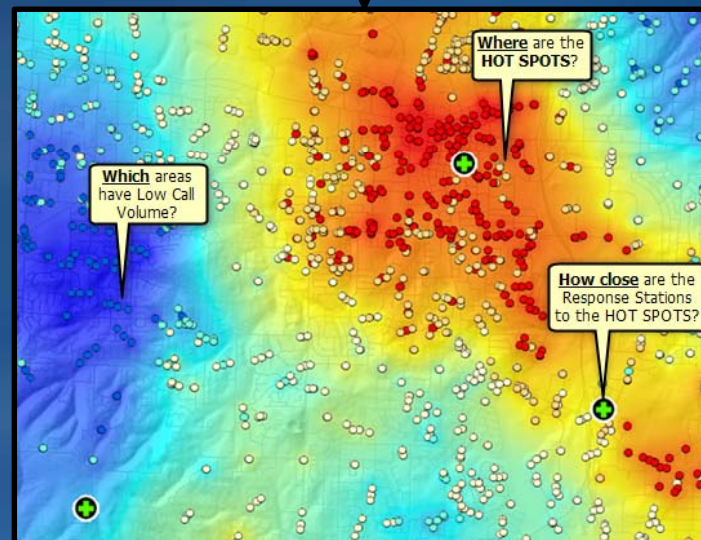
High Poverty Surrounded by Low poverty

Low Poverty

Low poverty Surrounded by High Poverty

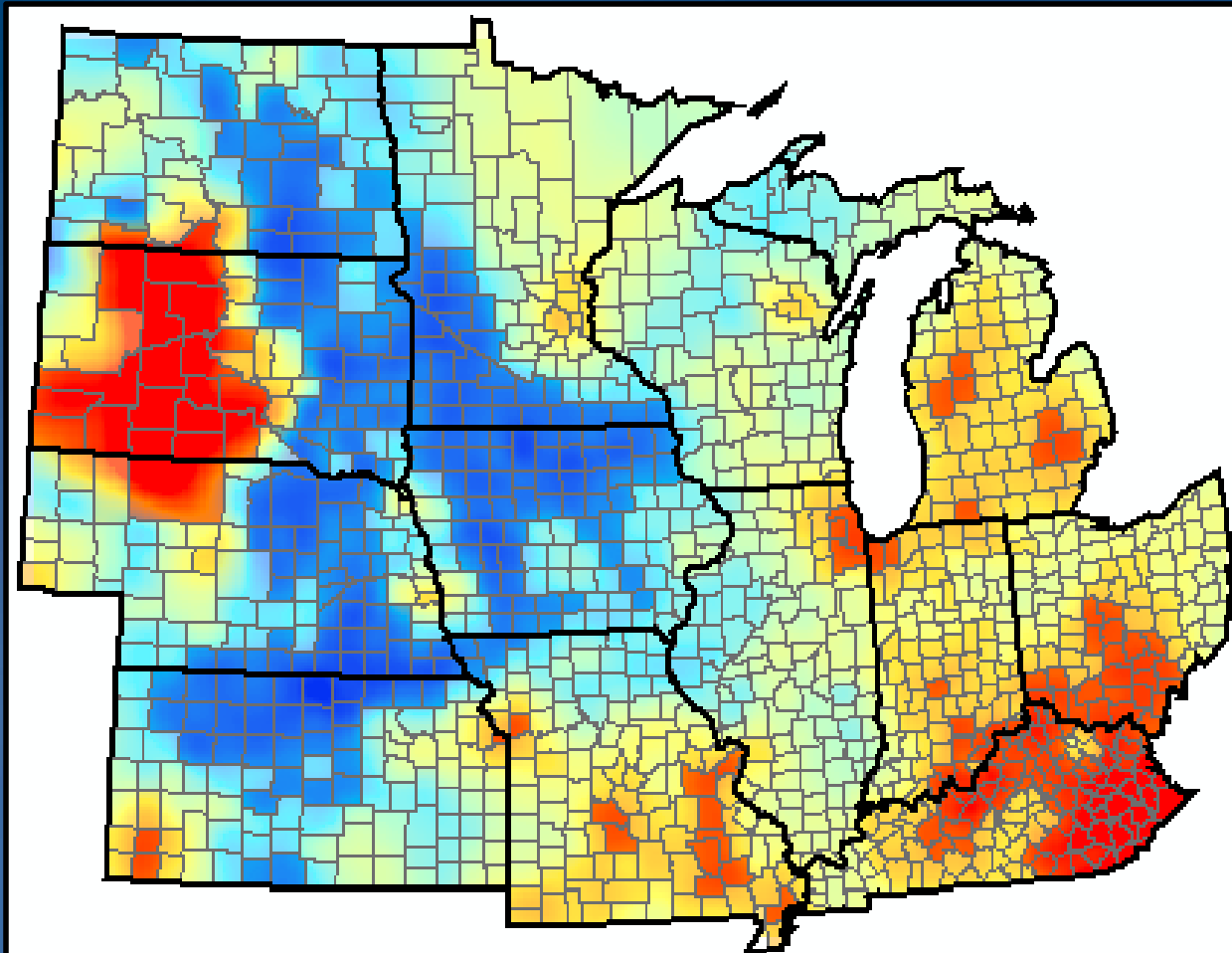


- Where are their sharp boundaries between affluence and poverty in Ecuador?
- Where do we find anomalous spending patterns in Los Angeles?



- Where are the 911 Call Hot Spots?
- Where do we see unexpectedly high rates of diabetes?

Average Age of Death Hot Spot Analysis



GiZScore	GiPValue
1.09548	0.273308
3.49006	0.000483
1.047	0.295098
1.159	0.246455
1.32213	0.186124
-2.30853	0.02097
-1.23803	0.215706
-1.59006	0.111821
-2.26428	0.023557
-0.665581	0.505679
-1.87229	0.061167
-2.25941	0.023858
-4.01195	0.00006
-2.85555	0.004296
-2.41466	0.01575

Spatial Statistic

Measuring Geographic Distribution

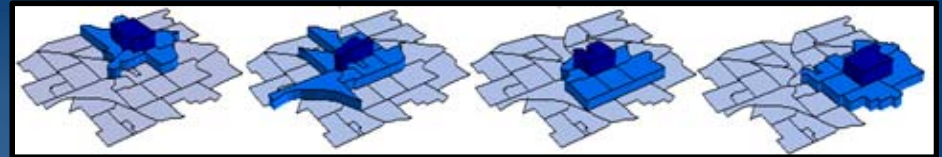
Analyzing Patterns

Mapping Clusters

Modeling Spatial Relationships

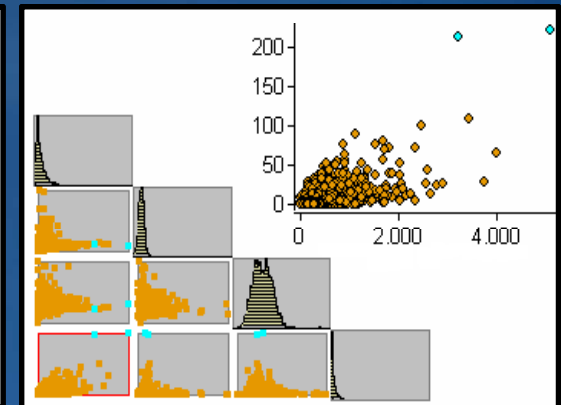
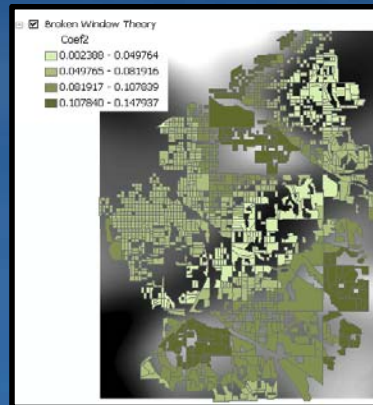
- Modeling Spatial Relationships
 - Generate Network Spatial Weights
 - Generate Spatial Weights Matrix
 - Geographically Weighted Regression
 - Ordinary Least Squares

- Can I model spatial relationships based on a real road network?
- Are spatial weights matrix files editable, sharable, re-usable?
- Can I create a custom spatial weights matrix file?



Construct spatial weights matrix files

- What is the relationship between educational attainment and income?
- Is there a relationship between income and public transportation usage? Is that relationship consistent across the study area?
- Where are real estate values likely to go up?



Ordinary Least Squares
Geographically Weighted Regression

Exercise

- **Create a simple model tool to:**
 - **Aggregate 911 incident data**
 - **Run Hot Spot Analysis**
- **Interpret analysis results**
- **Modify the model to use a spatial weights matrix file**
- **Explore alternate tool parameters**
- **Create a 911 hot spot surface using Interpolation tools**

Resources for learning more...

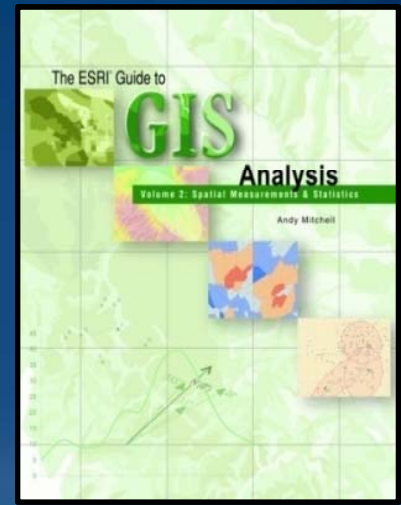
During the conference

- **Technical Workshops and Demo Theater Presentations**
 - **Using Spatial Statistics**
 - **Regression Analysis for Spatial Data**
 - **ArcGIS Spatial Analyst – Statistical Modeling**
 - **Geostatistical Analyst**
 - **Extending ArcGIS with R: Technical Briefing**
 - **Developing Python Scripts for Data Analysis Tips & Tricks**

Resources for learning more...

After the conference

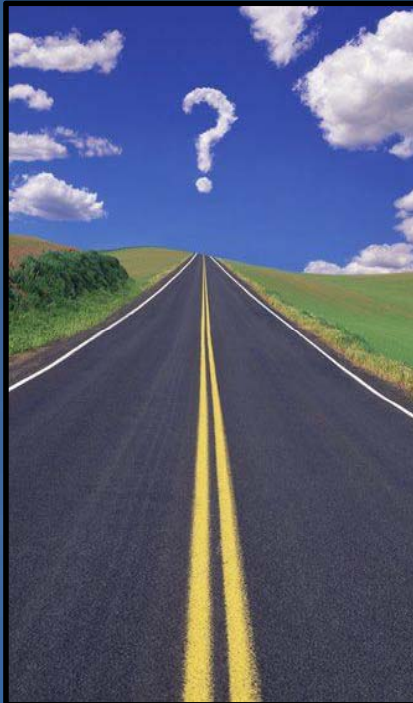
- Hot Spot and Regression Analysis Tutorials: <http://resources.esri.com/geoprocessing/>
- 911 emergency call analysis demo: <http://www.esri.com/software/arcgis/arcinfo/about/demos.html>
- Virtual campus free web seminars <http://campus.esri.com/>
- Articles (keyword search: “Spatial Statistics”) http://www.esri.com/news/arcuser/0405/ss_crimestats1of2.html
- Online help
- [The ESRI Guide to GIS Analysis, Volume 2 by Andy Mitchell](#)
- Instructor-led ESRI training



Kindly complete
an evaluation
form before you leave



QUESTIONS?



ESRI Resource Center for ArcGIS Desktop

Customer Care | ESRI.com | Careers

Home Content Web Help For Developers Support Other Resource Centers

What is Geoprocessing?

- Geoprocessing is how you [compute with data](#), connecting data to tools to derive new information
- You use geoprocessing for [automating GIS tasks](#)
- You use geoprocessing for [modeling and analysis tasks](#)
- There are two parts to geoprocessing; the suite of [tools](#) and the [framework](#)

Helpful Resources

- [Geoprocessing Tools quick reference guide](#) [PDF]
- [Geoprocessor scripting model diagram](#) [PDF]
- [Python - official web site](#)
- [What's new in 9.3?](#) [PDF]

Welcome to the Geoprocessing Resource Center

The Geoprocessing Resource Center is the place for you to:

- **Learn** how to solve problems using geoprocessing
- **Share** models and scripts
- **Communicate** with:
 - Other professionals like you
 - The development team

ModelBuilder Script tools Arc Toolbox Server Python Extensions Toolbox

learnsharecommunicate

Contact ESRI | Copyright © ESRI | PRIVACY

LScott@ESRI.com