Using the Spatial Statistics Toolbox in ArcGIS 9 for Crime Analysis

Lauren M. Scott, PhD
Nathan Warmerdam, MSc
ESRI
Geoprocessing, Spatial Statistics

Objectives

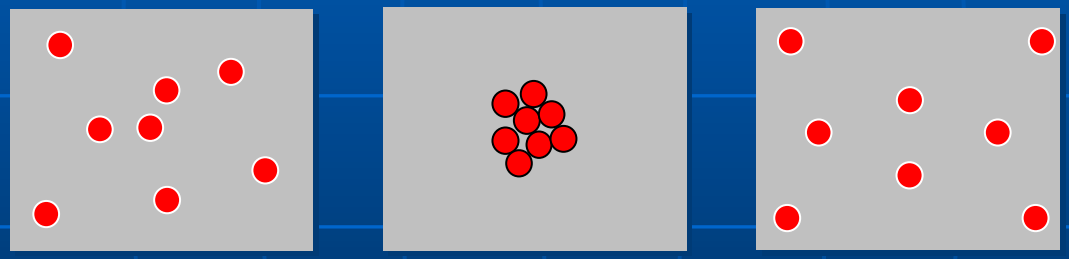
- Introduce core spatial statistics tools available in ArcGIS 9
- Provide examples of how these tools might be used for crime analysis
- Demonstrate custom tool development capabilities available through the Geoprocessing Framework of ArcGIS 9

Presentation

- Basics: What are spatial statistics and why do we care?
- Overview of Spatial Statistics Toolbox
  - Analyzing hot spots
  - Assessing risk
  - Looking for clues in spatial relationships
  - Crime forecasting
- Creating custom tools and models

What are spatial statistics?

- A measure of what's going on spatially
- Measure characteristics of a distribution
- Quantify geographic patterns



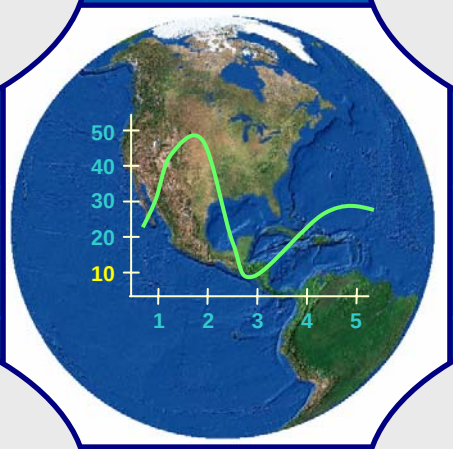
Random Clustered Dispersed

With spatial statistics, space and spatial relationships are an integral component of analysis

Why use spatial statistics?

Better understand behavior of geographic phenomena

Make decisions with higher level of confidence



Pinpoint causes of specific geographic patterns

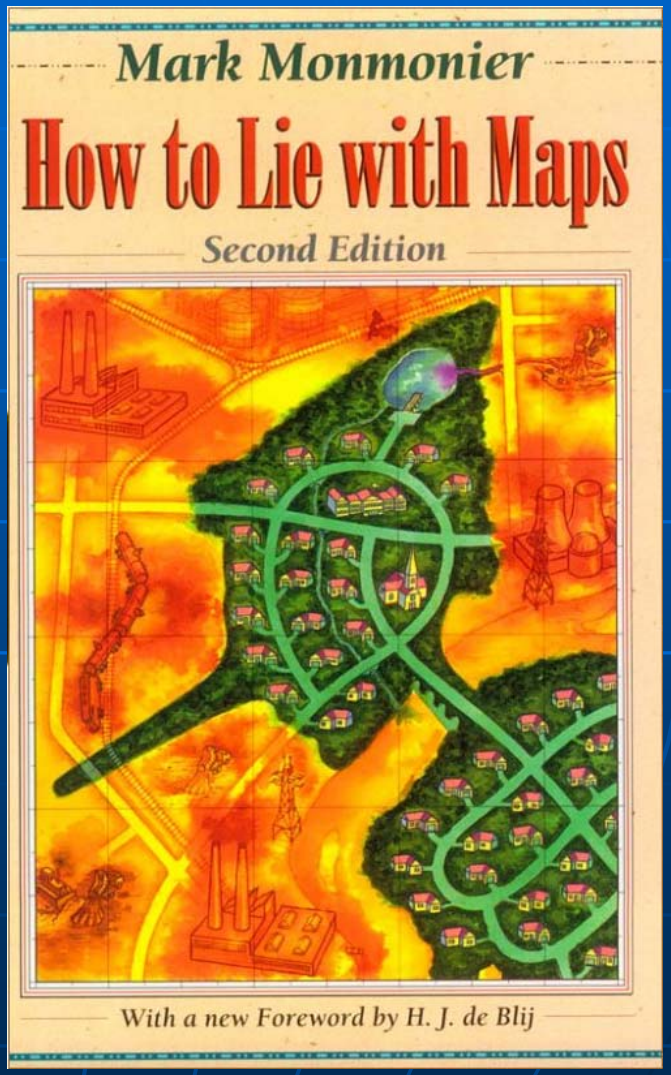
Summarize the distribution in a single number

Why Use Spatial Statistics?

Spatial Statistics help us assess:

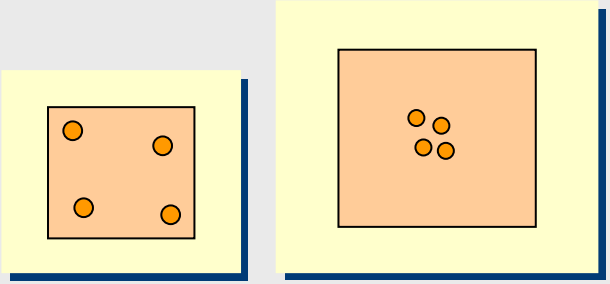
- **Patterns**
- **Relationships**
- **Trends**

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

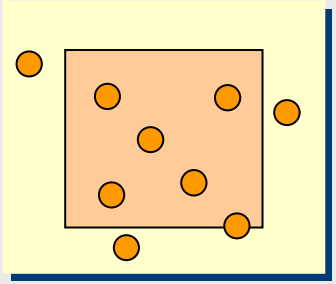


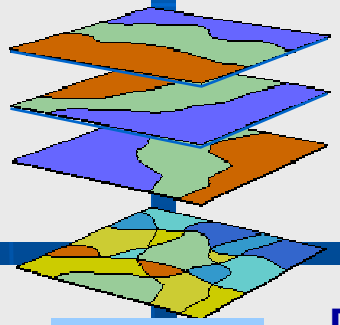
Factors that affect results

Geographic scale



Study area boundary





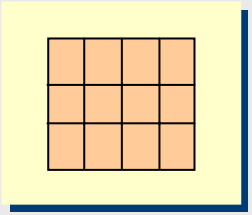
Results

Time period

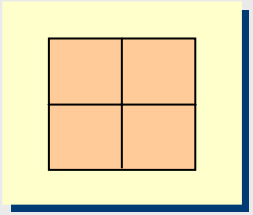
Week 1
Week 2
Week 3

Year 1
Year 2
Year 3

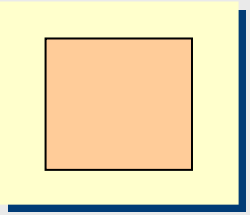
Data summarized by areas



Census block group



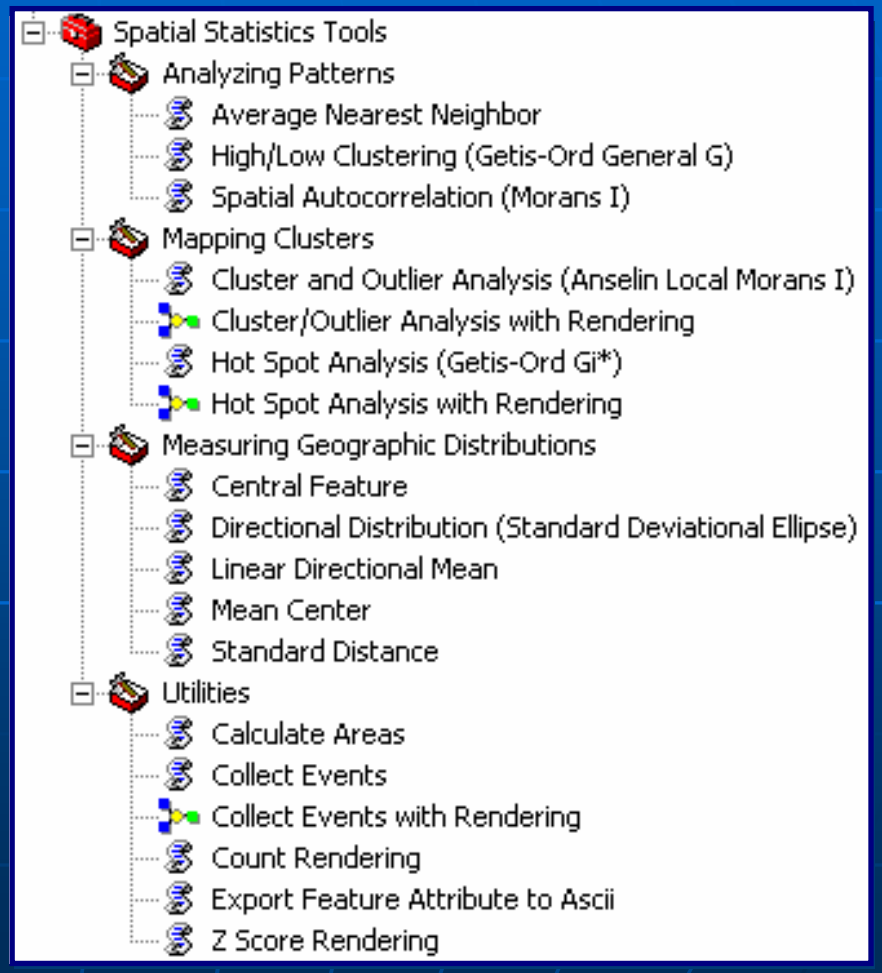
Census tract



County

Toolsets and tools

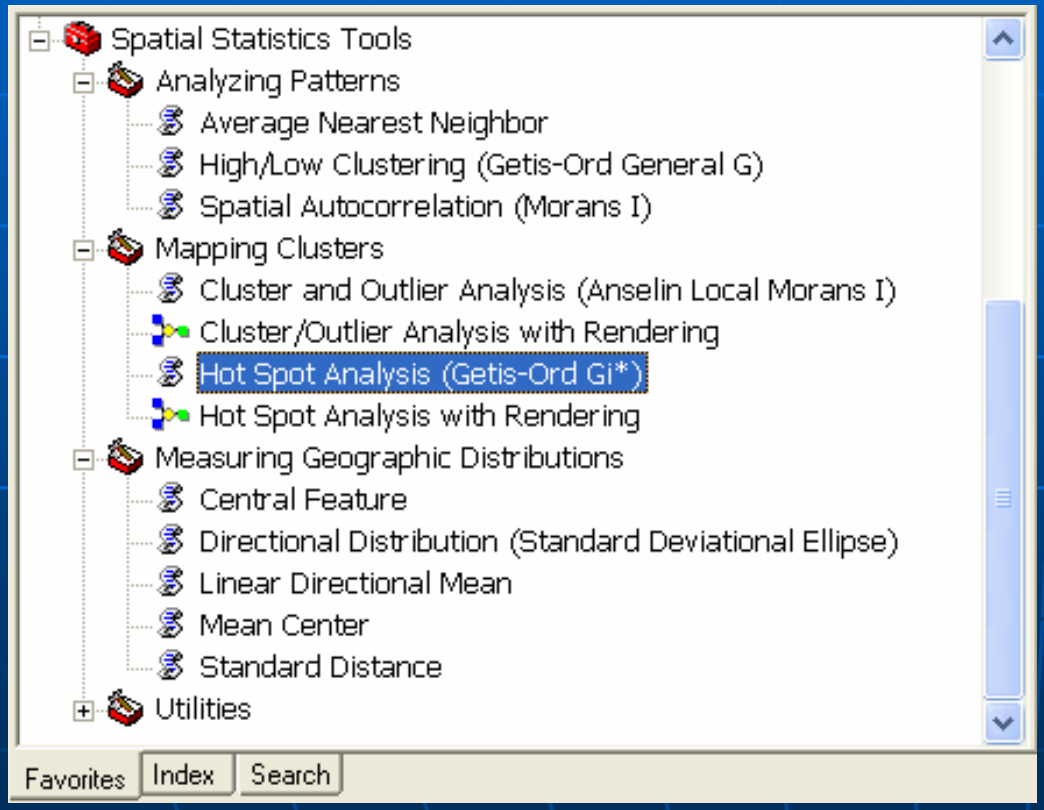
- Core functionality with ArcGIS 9 (not an extension)
- Available with all ArcGIS licenses
- Source code provided
- For ArcGIS 9.0, please use SP2 or SP3
- *The ESRI Guide to GIS Analysis, Volume 2*



911 Emergency Call Analysis

Demo

Spatial Statistics Toolbox



- Analyze broad regional patterns.
- Map hot spots and local outliers
- Find the geographic center or identify distributional trends
- Utilities

Measuring Geographic Distributions toolset

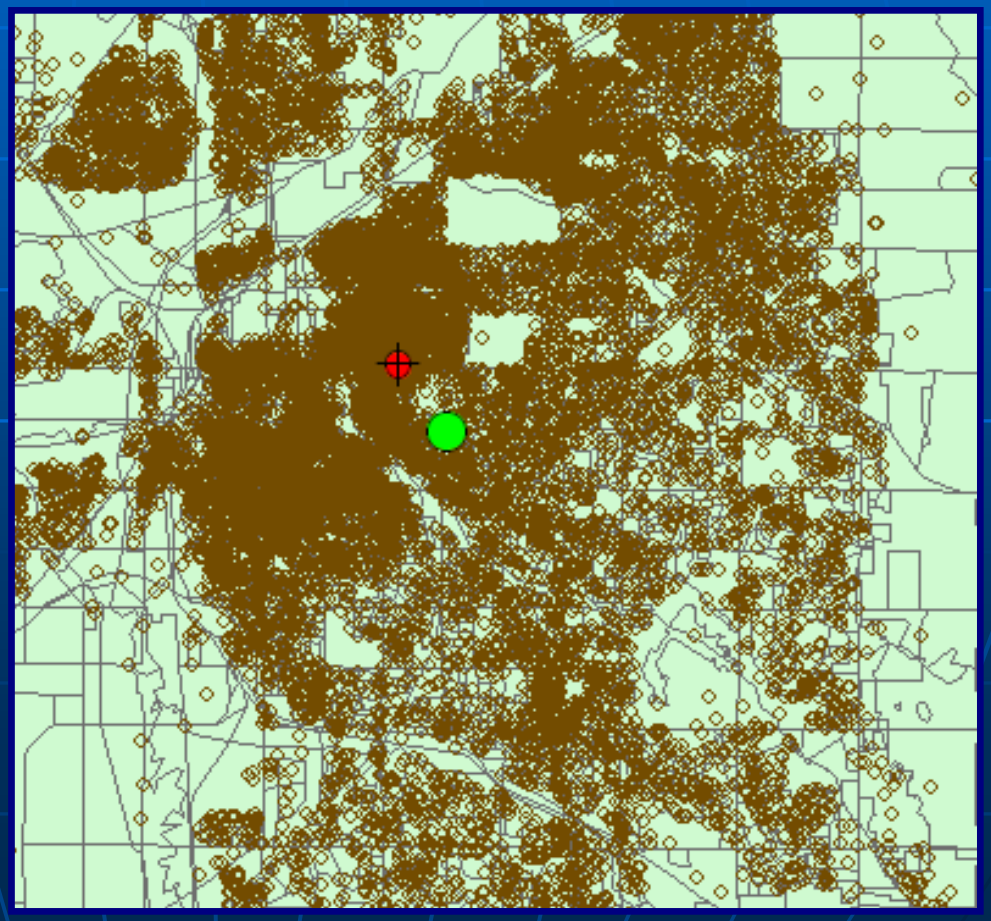
- Identify spatial characteristics of a distribution
  - Where is the center?
  - What is most central?
  - How dispersed are features around the center?
  - Are there any directional trends?
- Often used to:
  - Compare different feature categories
  - Examine changes over time

1. **Central Feature**
2. **Mean Center**
3. **Linear Directional Mean**
4. **Standard Distance**
5. **Standard Deviational Ellipse**

Mean Center

Computes the average x and y coordinates, based on all features in the study area

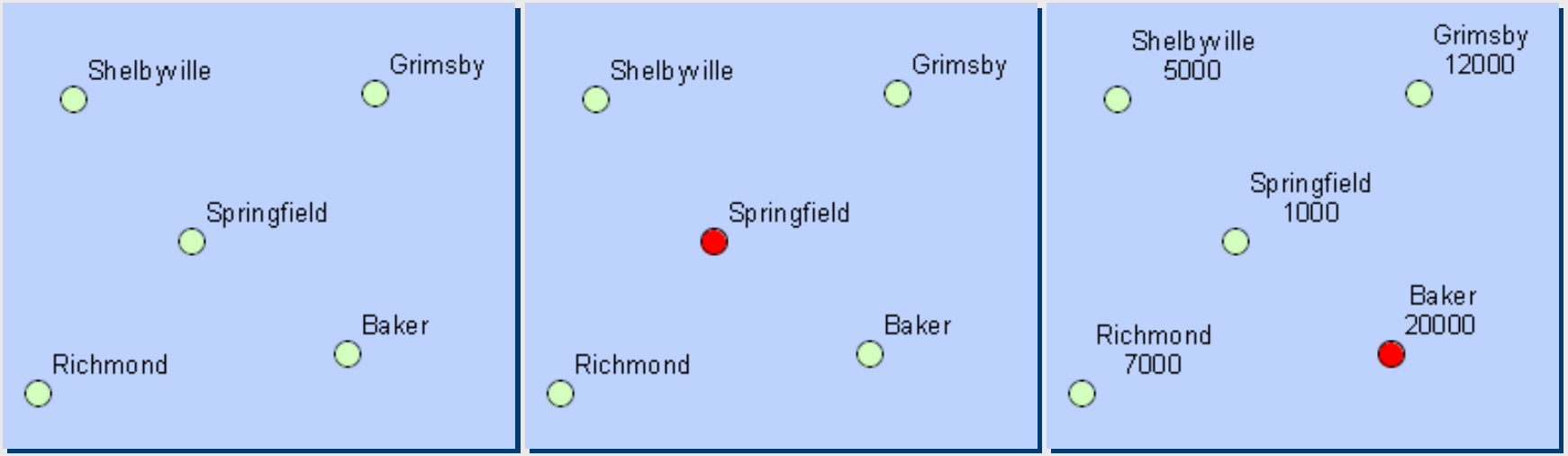
Mean Center 2002:
Crime (red) vs.
Population (green)



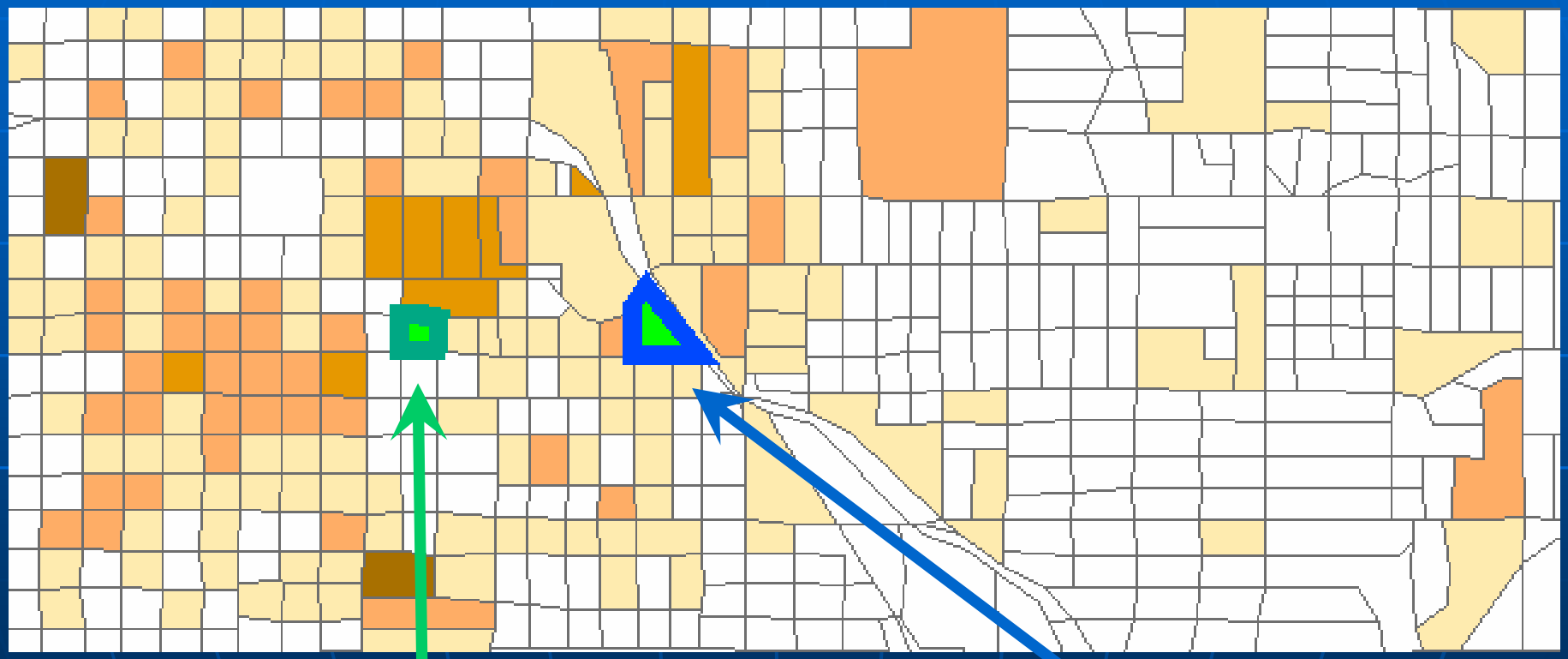
Central Feature

- Identifies the most centrally located feature
  - Feature with the lowest total distance to all other features
- Uses:
  - Finding the most accessible existing feature

What is the best location for a new service center?



Locating services for victims of domestic crime



New Service

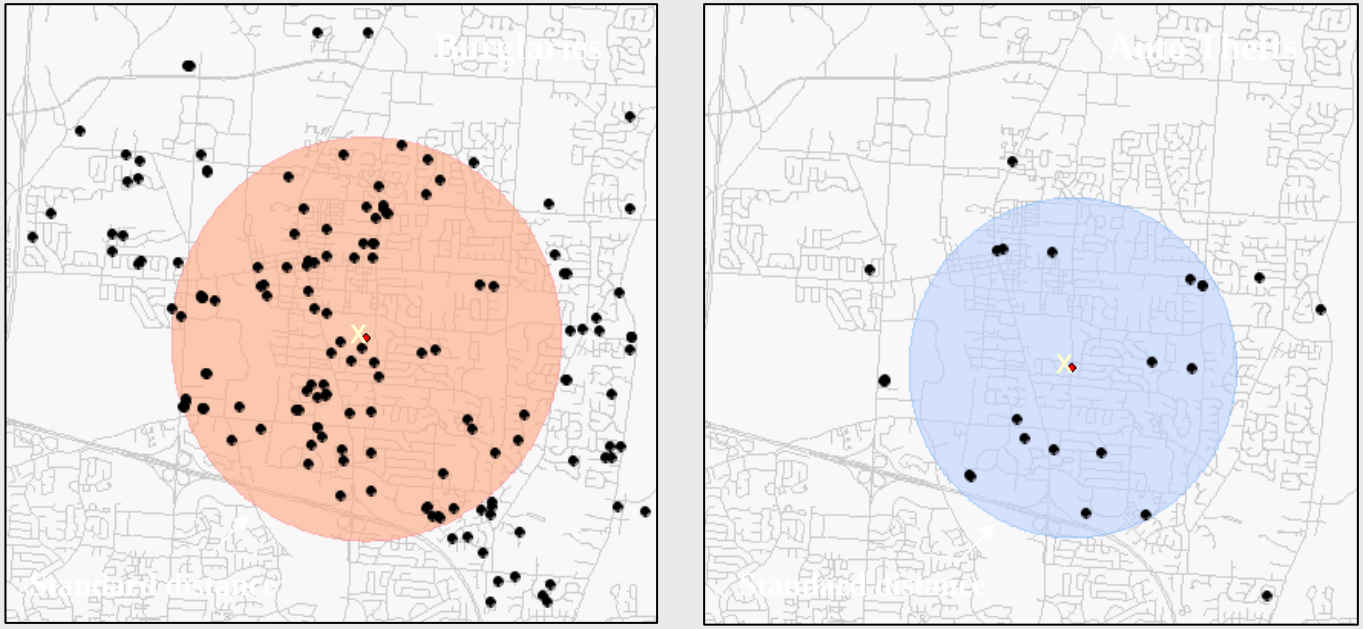
(most central with weighting)

Most Central Tract

Standard Distance

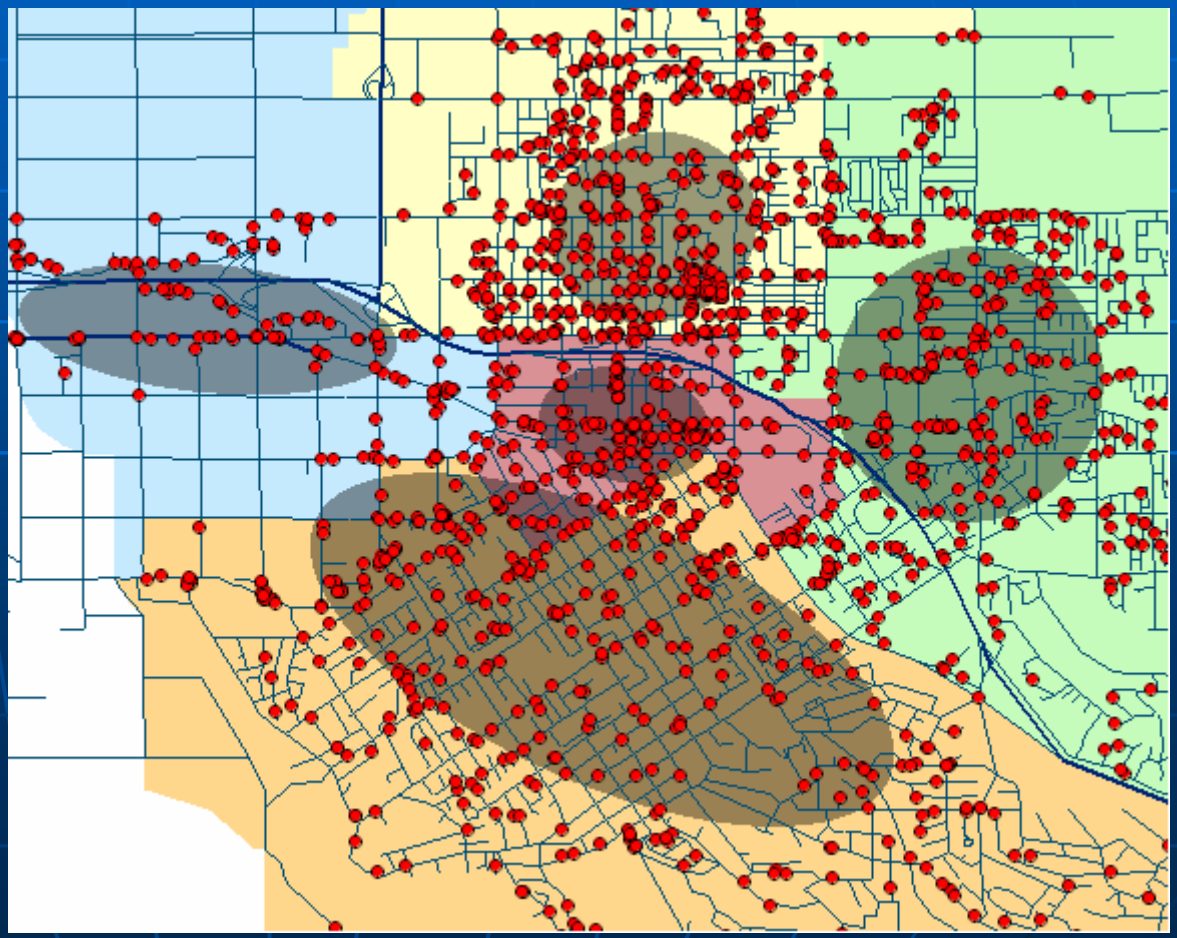
- Measures dispersion of features around the mean
  - Result is a summary statistic representing distance
- Sensitive to outliers

Westerville, Ohio
Distribution of burglaries and auto theft incidents, 2001

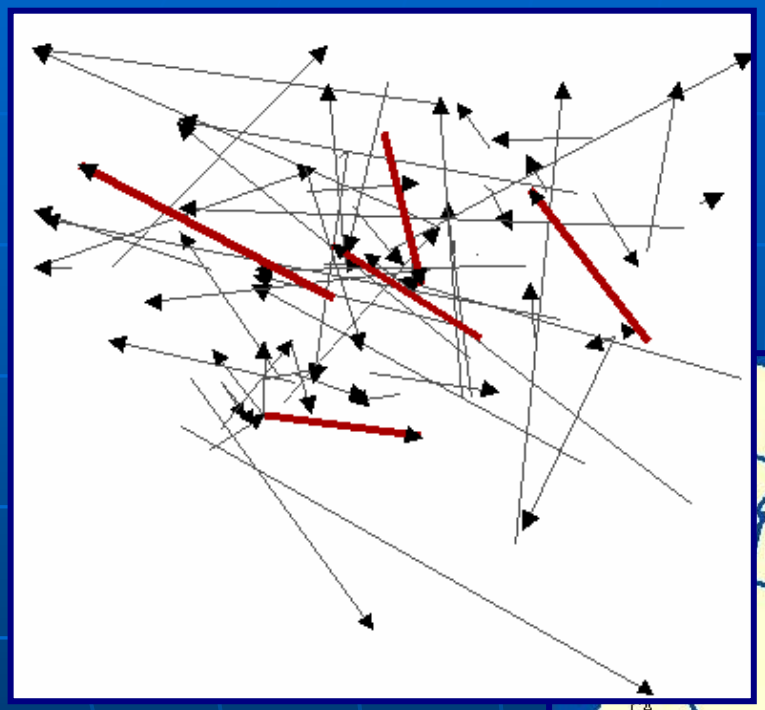


Directional Distribution
(Standard Deviational Ellipse)

- Abstracting spatial trends in a distribution of features
- Uses:
  - Comparing distributions
  - Examining different time periods
  - Showing compactness and orientation



Linear Directional Mean

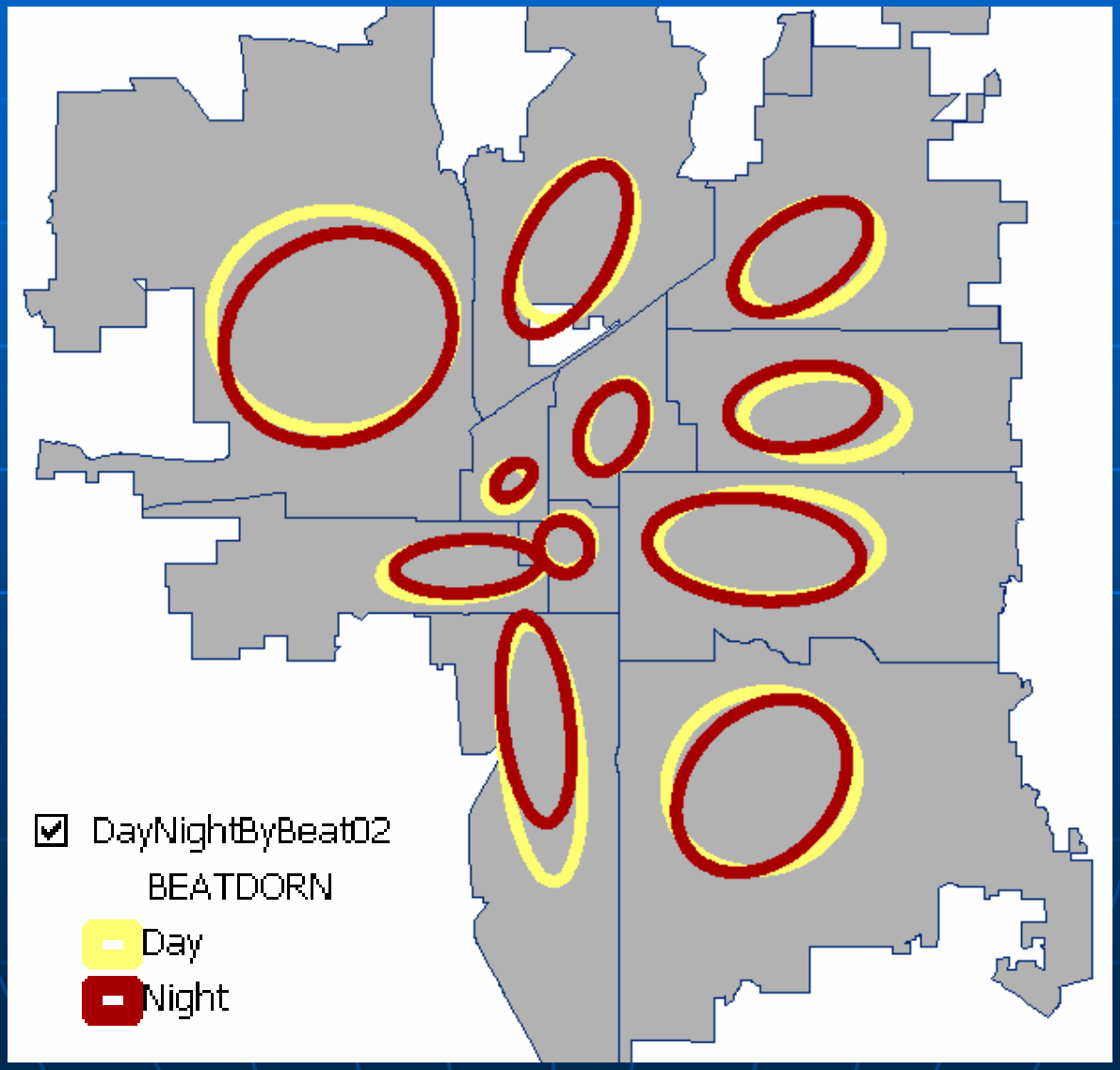


Angle of a line representing mean direction (or orientation) based on all lines in a dataset

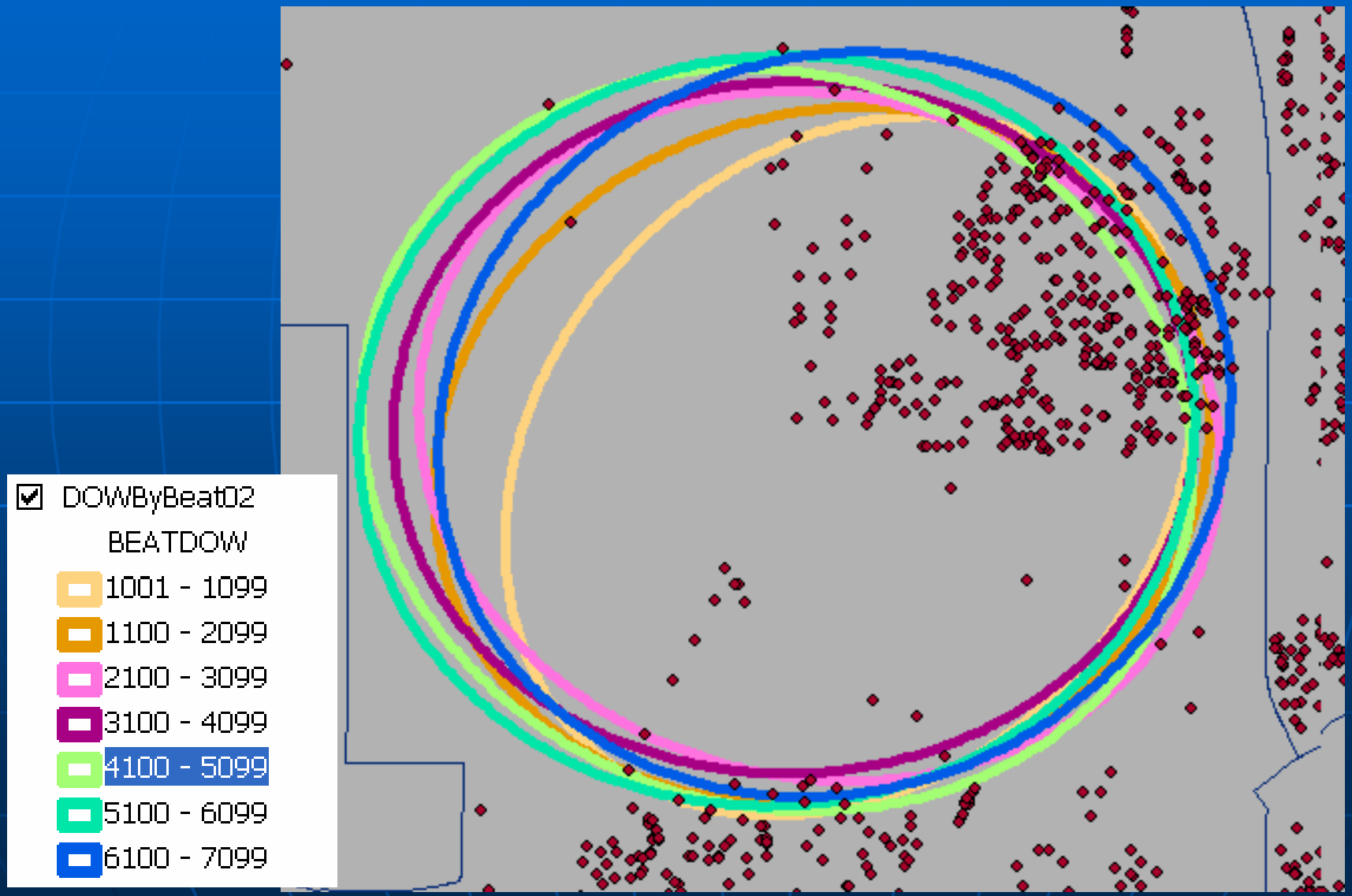
Daytime vs. Nighttime Crime Analysis

Demo

Day vs. Night Crime Distributions



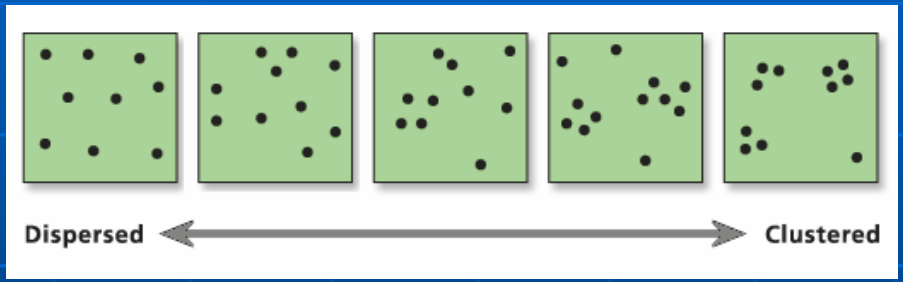
Day of Week Analysis by Beat

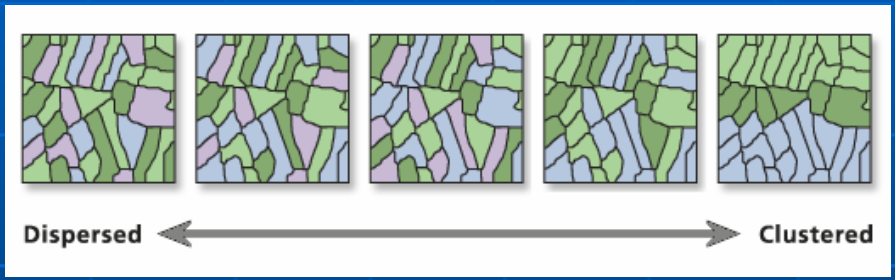


Analyzing Patterns toolset

3 Tools

- Determine the extent to which features are clustered, dispersed, or random across study area

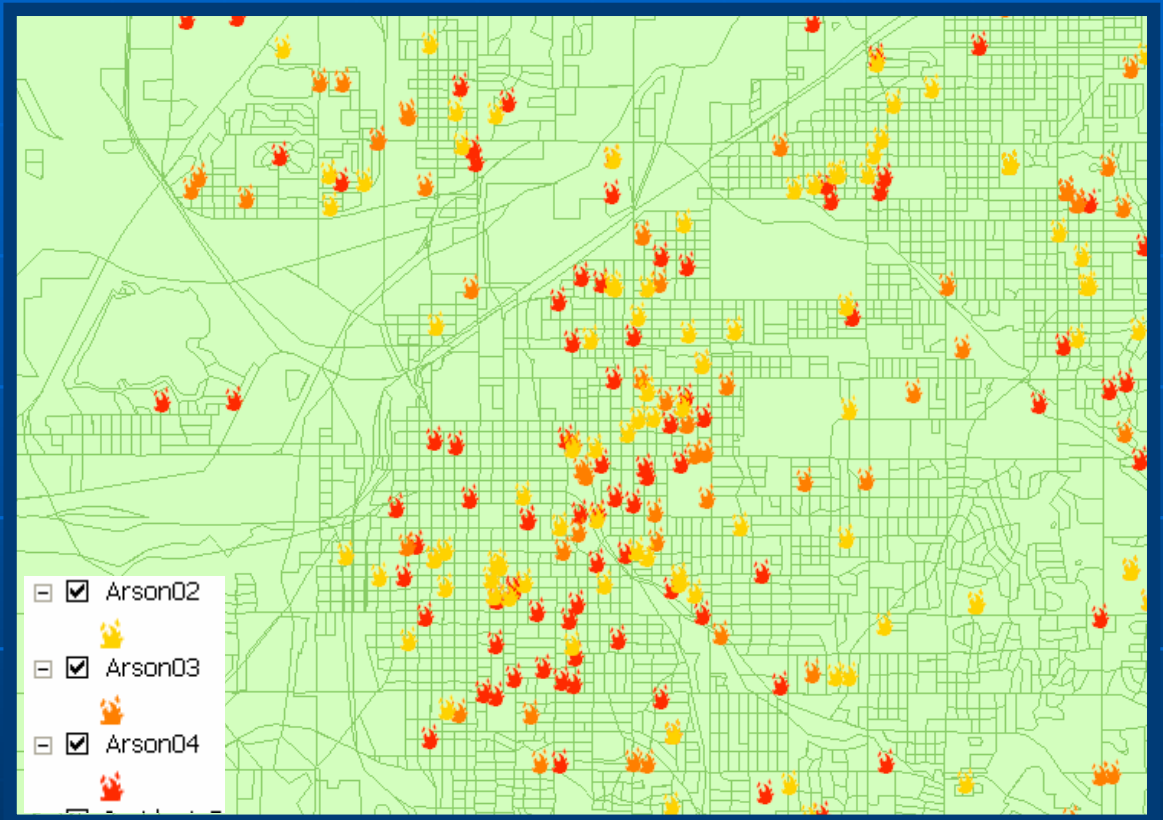




- Global calculations that quantify geographic patterns
- Useful for:
  - Gaining a better understanding of feature distribution
    - Degree of clustering or dispersion across study area
  - Comparing different sets of features
  - Tracking changes in patterns (over time)

Average Nearest Neighbor

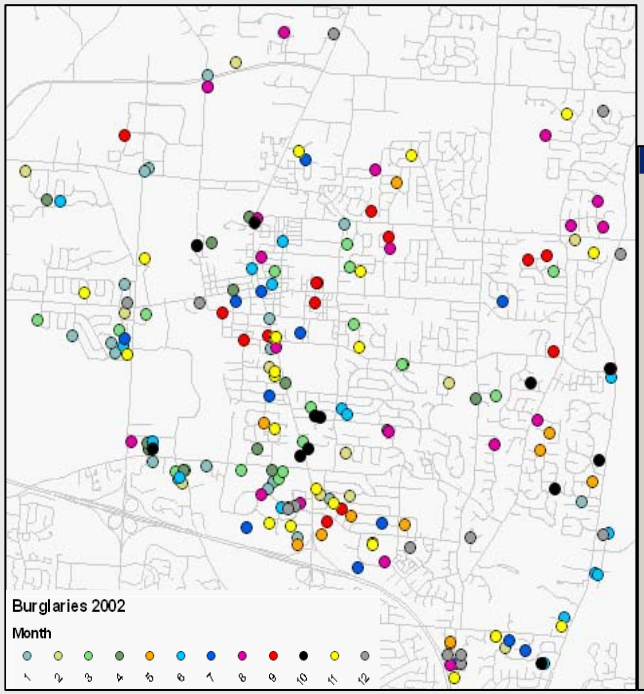
- Calculates the average distance between each feature
  - Considers arrangement and distance of features
- Based on area



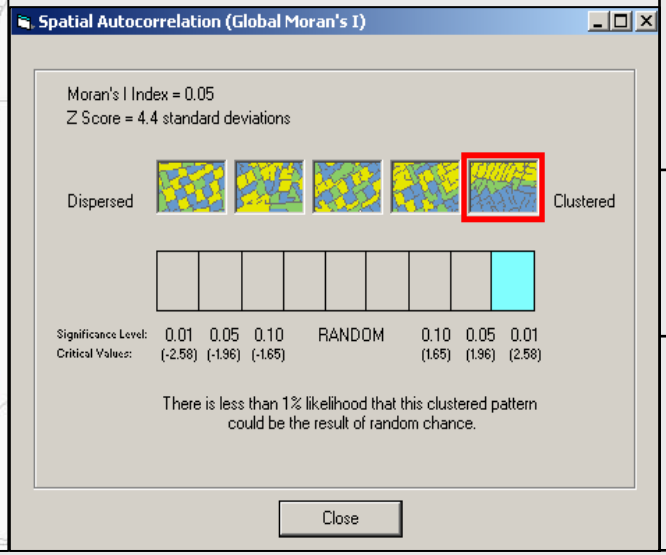
2002	2003	2004
Z = -6.8	Z = -6.6	Z = -6.4

Spatial Autocorrelation (Moran's I)

- Measures similarity of neighboring features
  - Things closer are more alike than things that are not
- Based on both feature locations and attribute values



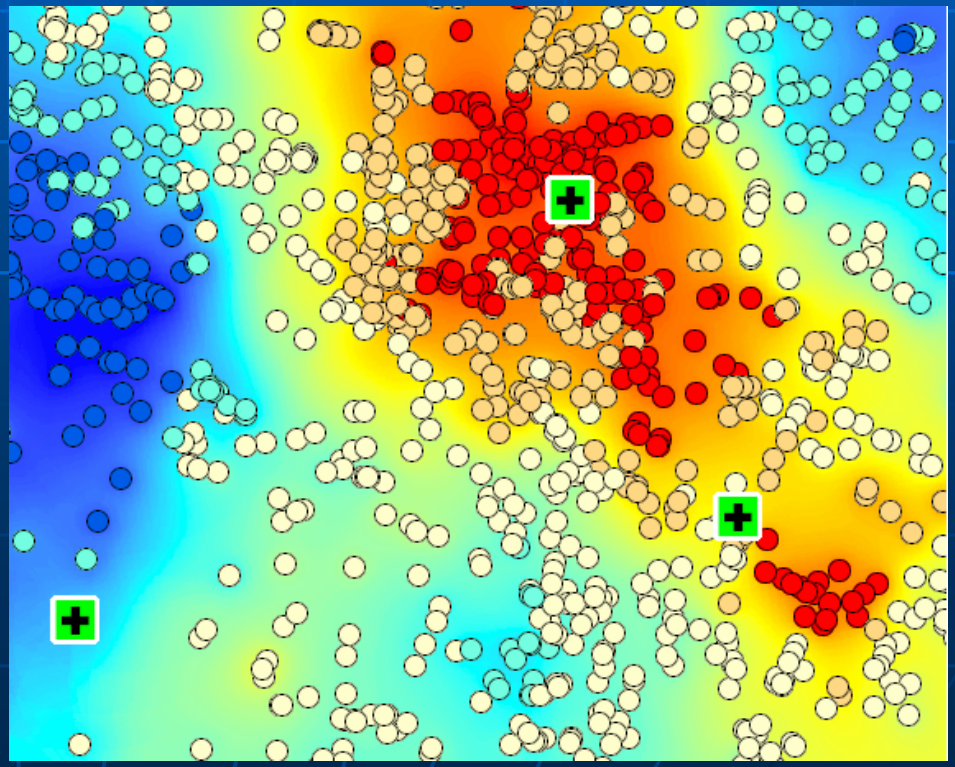
Do burglary incidents exhibit any seasonal patterns?



Moran I	Pattern
I > 0	Clustered *similar values are found together*
I = 0	Random *no apparent pattern*
I < 0	Dispersed *high and low values are interspersed*

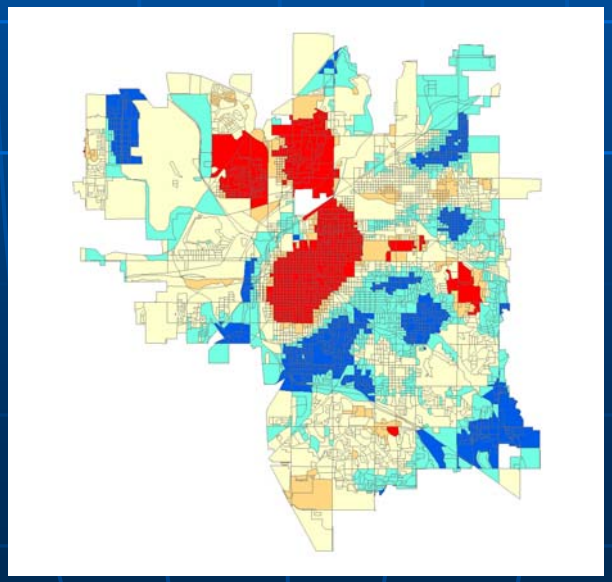
Mapping Clusters toolset

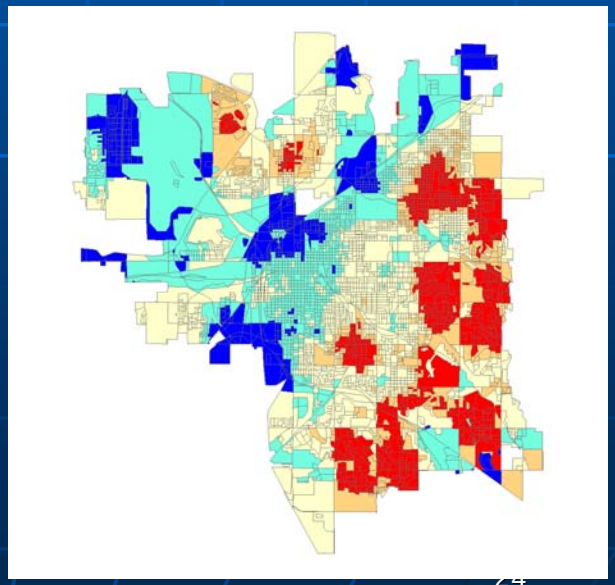
- Comparing clusters to locations of other features can help to better understand why clusters occur
- Reveal clustering and/or outliers within a study area
- Local calculations
  - Considers each feature
  - See differences across the study area
  - Explore how patterns change with changes in spatial scale



Hot Spot Analysis (Getis-Ord Gi*)

- Indicates the extent to which each feature is surrounded by similarly high or low values
  - Identifies where clusters of high and low values exist in study area
- Includes target feature in calculation
- Calculates a z score for each feature
- Uses:
  - Crime prevention
  - Locating target market
  - Finding source of an epidemic
  - Map insets



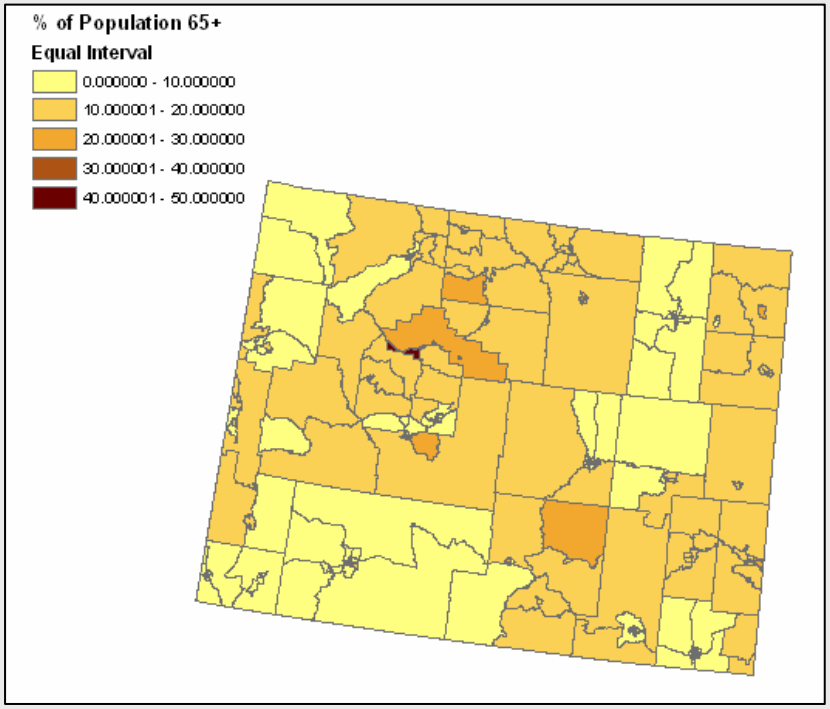


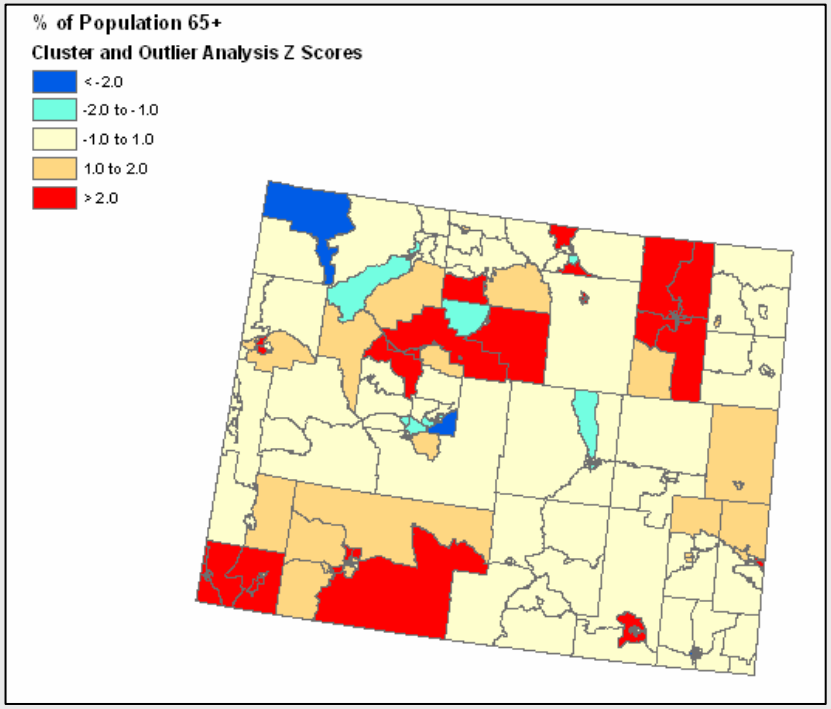
Cluster and Outlier Analysis

(Anselin Local Moran's I)

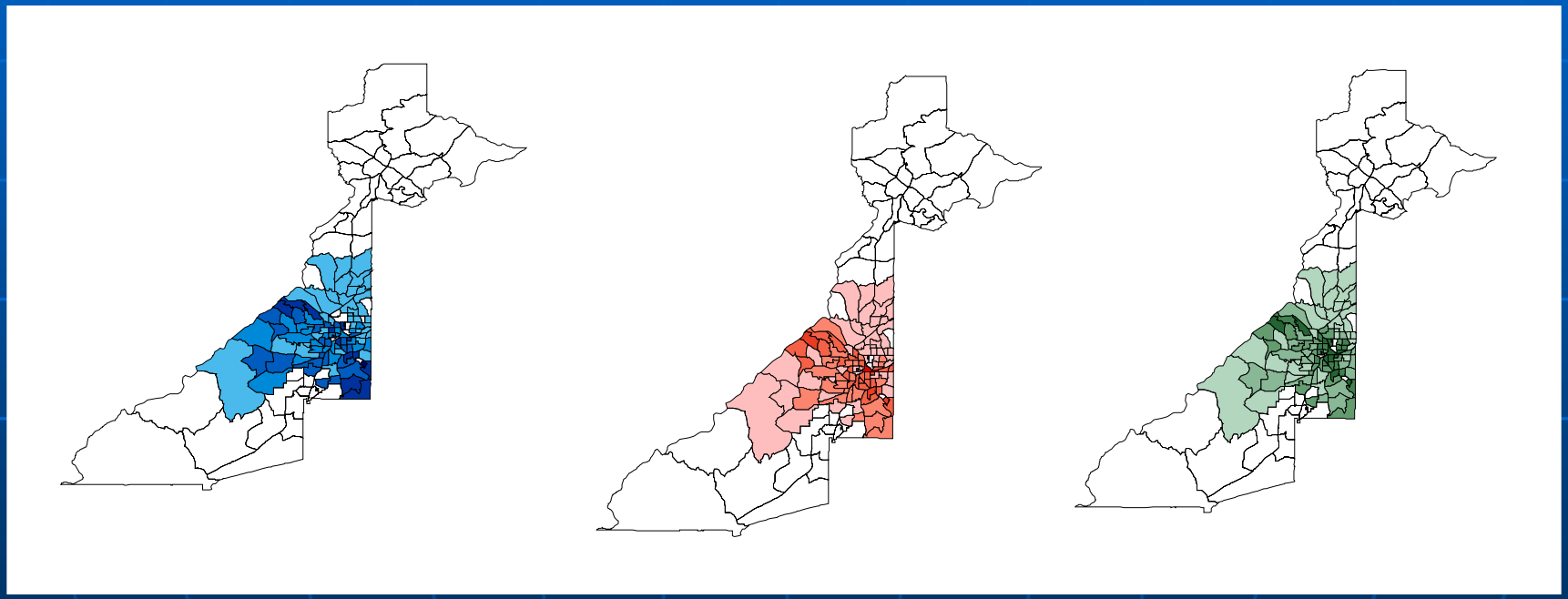
- Emphasizes how feature values differ
- Helps identify outliers and errors in data
- Calculates index value and z score

Wyoming Census Block Groups, 2000





Comparison of Neighborhood Risk Factors and Neighborhood Rates of Recurrent Domestic Violence for Atlanta (Fulton Co.)*



% Female Householders with Children (1990 US Census)

Neighborhood Recurrent Rates of Domestic Violence (1997)

% Below Poverty Line (1990 US Census)

*** Darker color = higher rates/ percentage**

Maps Created by Ann C. McClellan

Compare Gi* to Li*

- Hot Spot Analysis (Gi*)

H H → Positive Z
L L → Negative Z

H L → 0 ?
L H → 0 ?

- Local Moran's I (Li*)

H H → Positive Z
L L → Positive Z

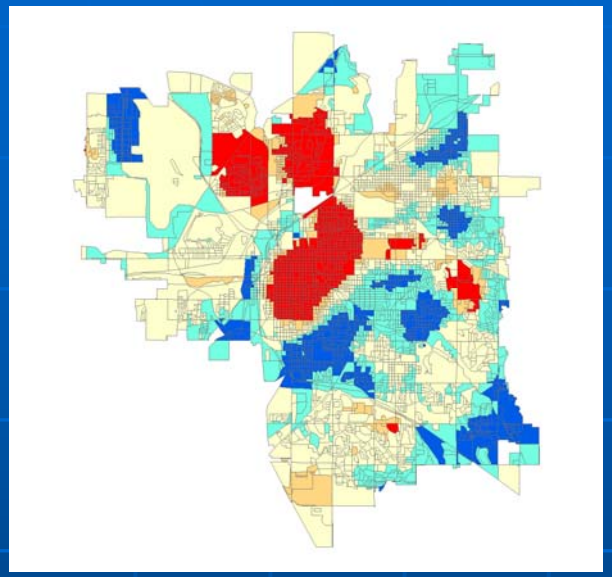
H L → Negative Z
L H → Negative Z

Hot Spot Analysis
Demo

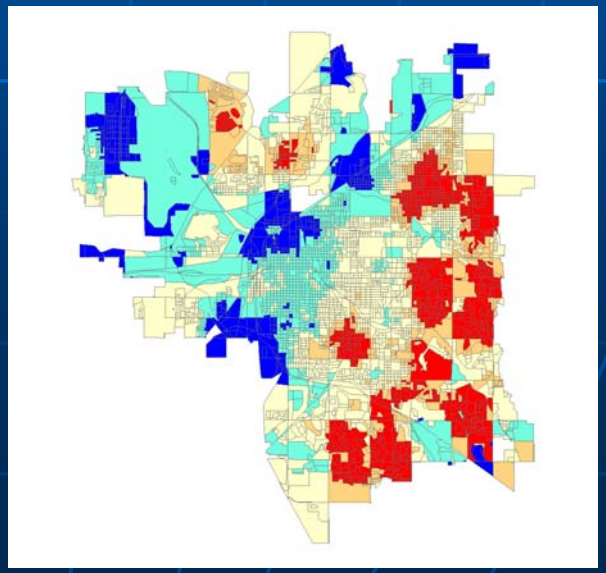
- Defining “Hot Spot”
- Considering population at risk: controls
  - Population densities
  - Regional variations in crime rates

Crime Example Using “Controls”

- Before controlling for population and bad neighborhoods, vandalism hot spots are just where we expect: downtown and high population areas



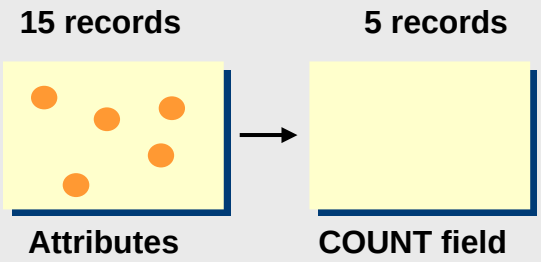
- After normalizing by ALL Crime values, vandalism is found to be primarily a suburban problem



Utilities toolset

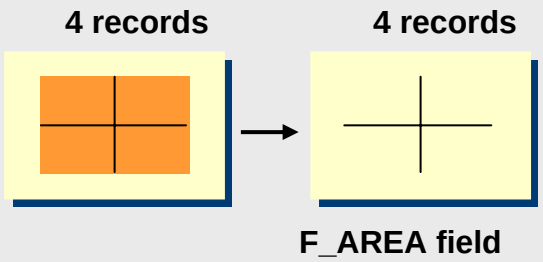
Collect Events

- Converts event data into weighted point data



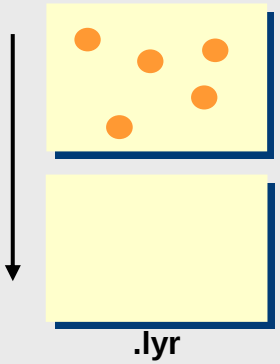
Calculate Areas

- Calculates AREA for polygon features



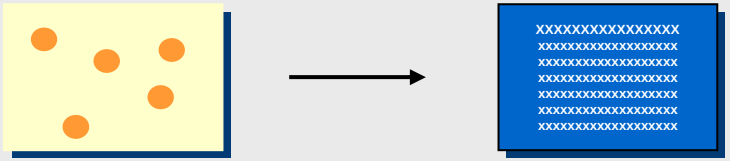
Count Rendering

- Applies graduated circle rendering to a point feature class with COUNT field



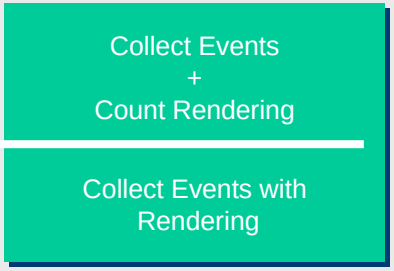
Export Attributes to ASCII

- Exports coordinates and attributes to ASCII text file



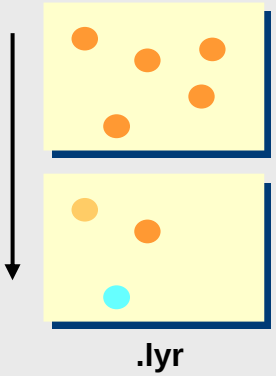
Collect Events with Rendering

- Converts event data into weighted point data
- Applies graduated circle rendering



Z Score Rendering

- Applies graduated hot and cold color rendering to point feature class with z score values



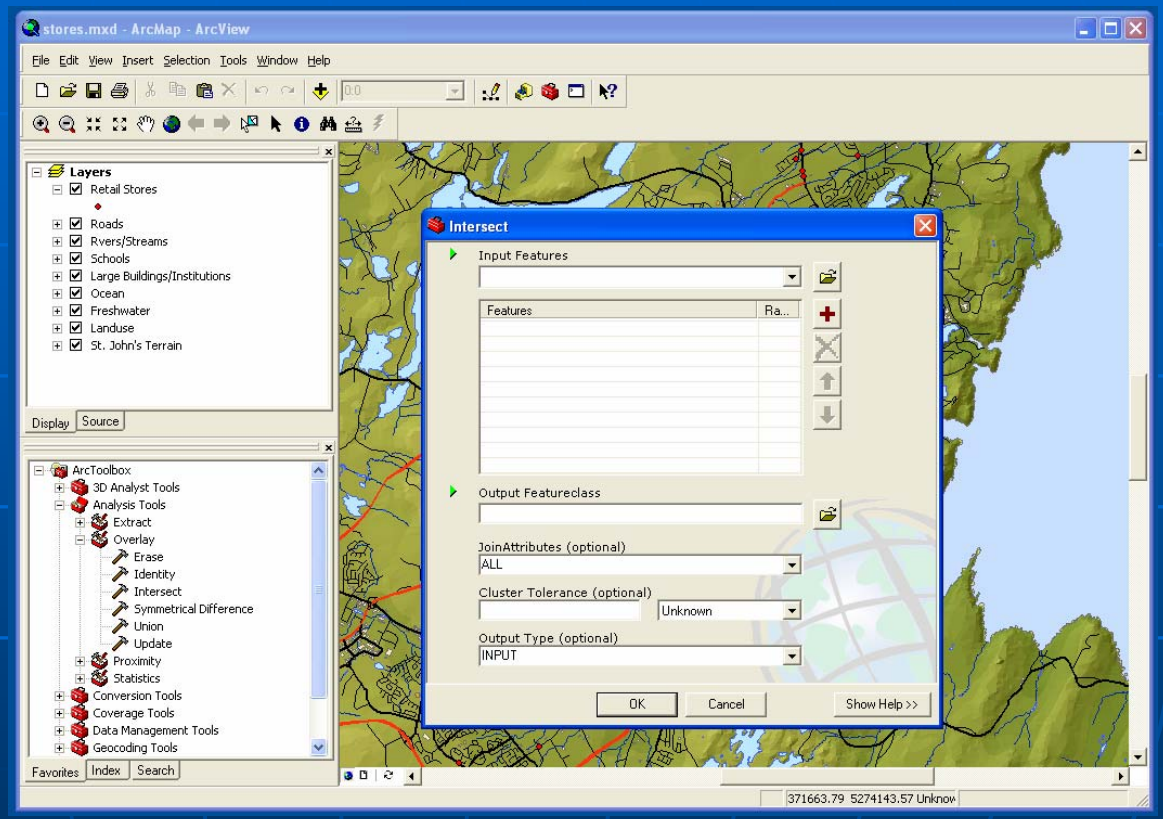
Intermission

1. What are the typical analyses I've missed:
   - Hot Spot Analysis
   - Risk Assessment
   - Crime distribution analysis
2. Will you be able to use the tools presented so far? How? What specific analyses?
3. Are there specific analytical methods that are missing from ArcGIS?
4. Other questions or comments?

Delivering Source Code

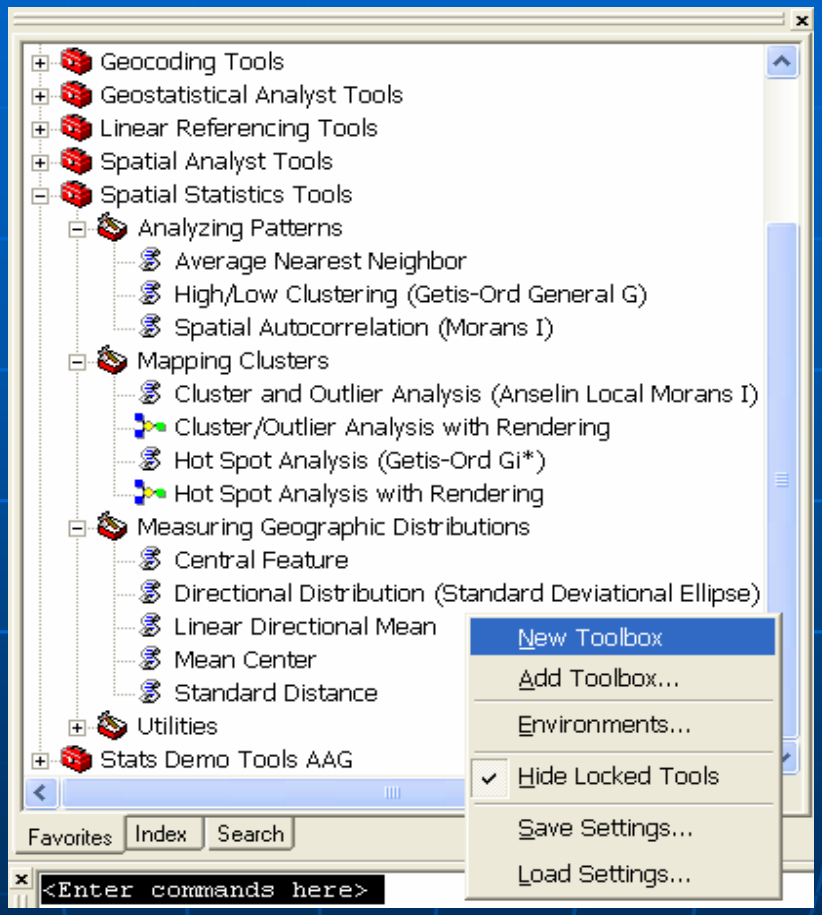
- Samples/templates may help you develop your own custom tools.
- There are still many unresolved issues associated with spatial statistics.
- Sharing tools and ideas may lead to better tools, new tools, and even new questions/approaches.

The Geoprocessing Framework



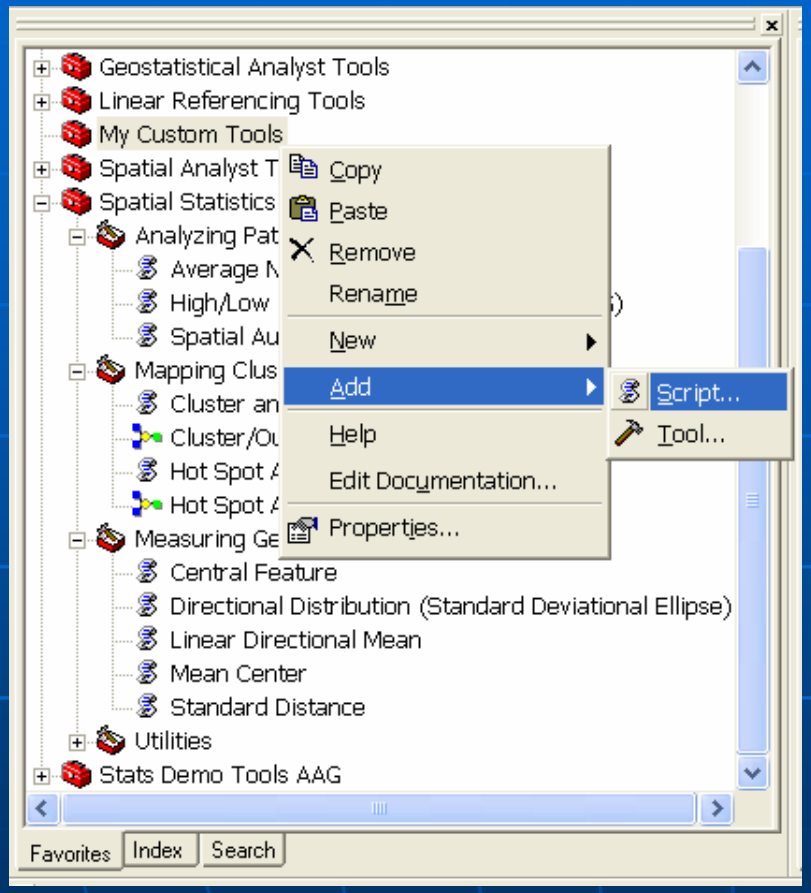
- **Core Tools**
- **Extension Tools**
- **A *Framework* for creating Custom tools and utilities !!**

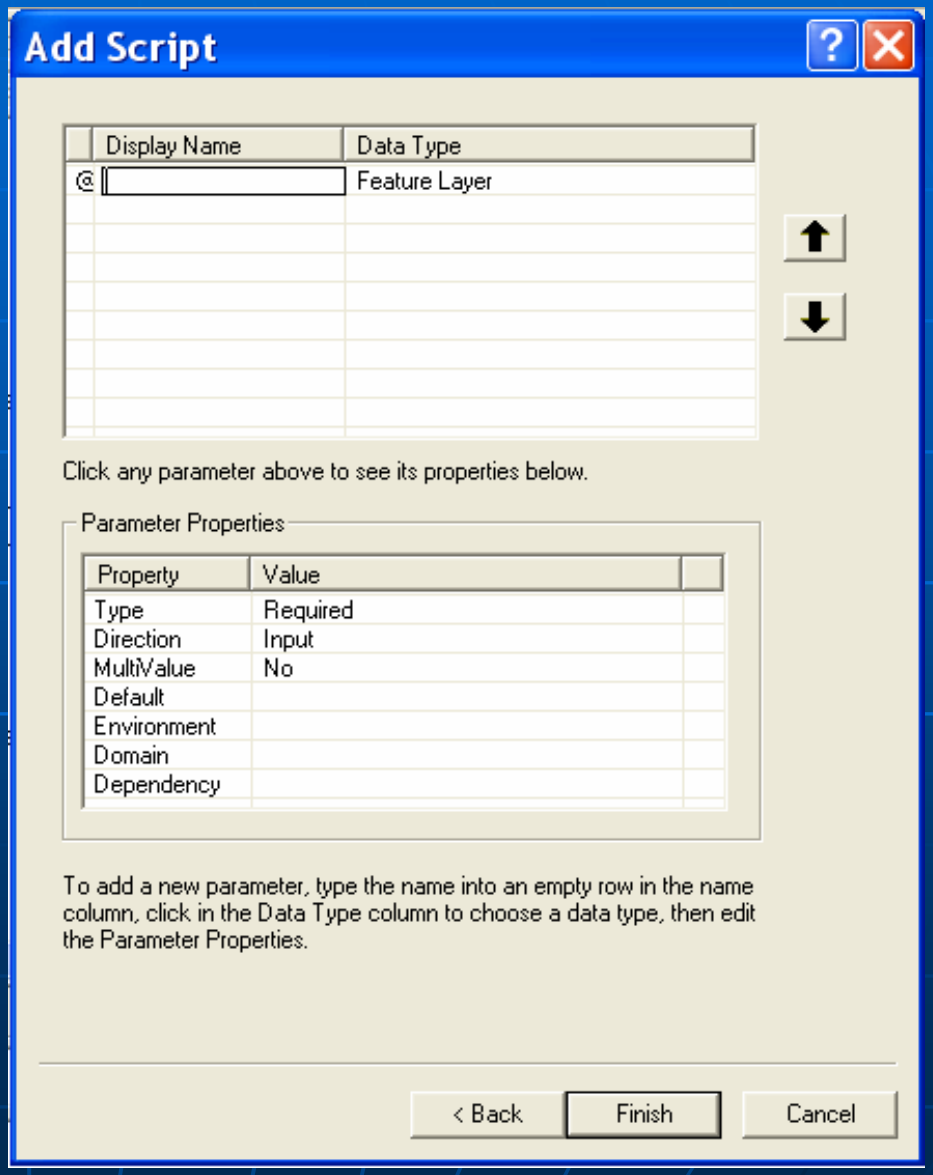
Creating a new toolbox:



Right click anywhere in the ArcToolbox pane to create a new toolbox.

Adding new tools:





Right click on your new toolbox to add a script or model tool. Then define your tool parameters.

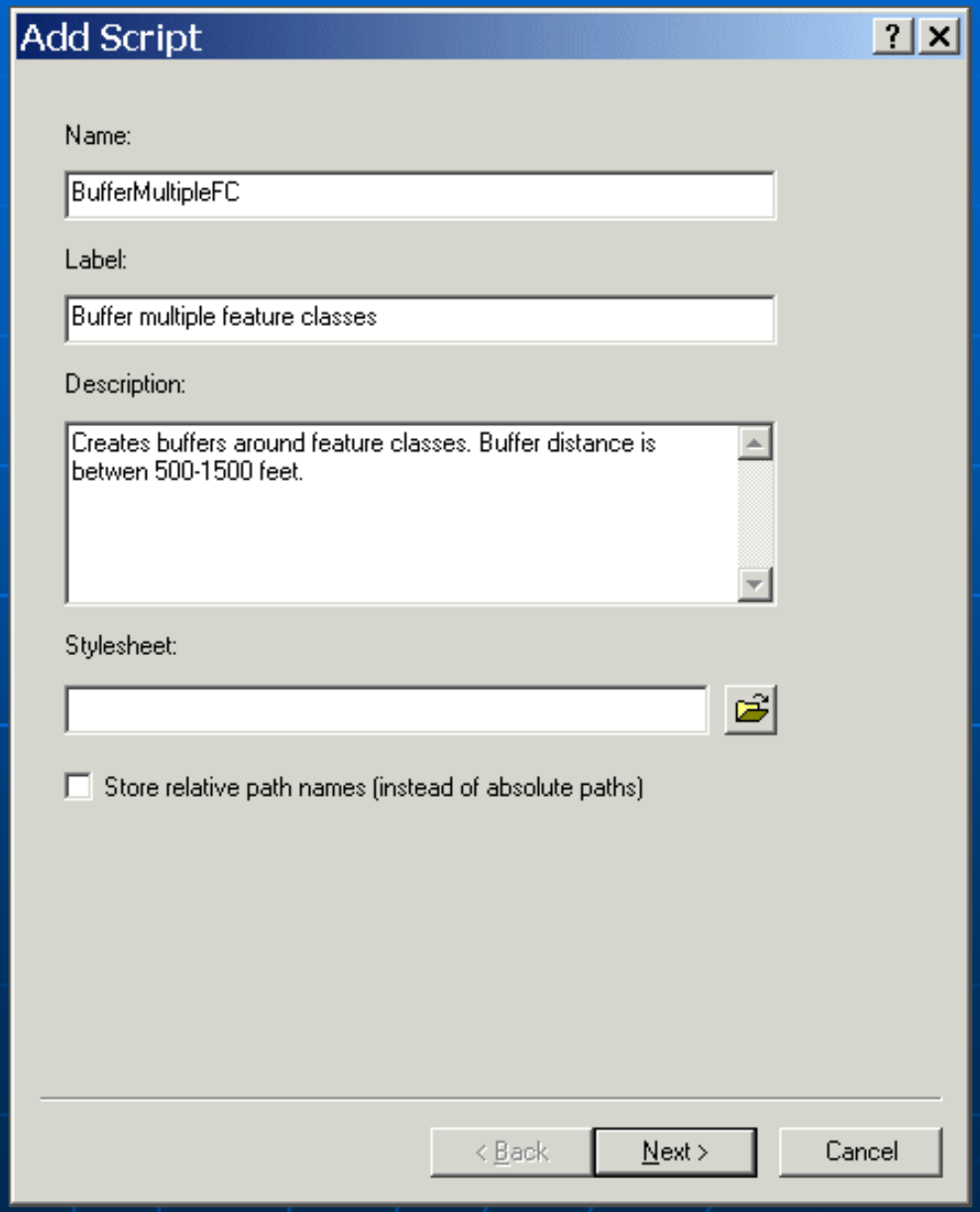
Tool Wizard:

Name →

Label →

Description →

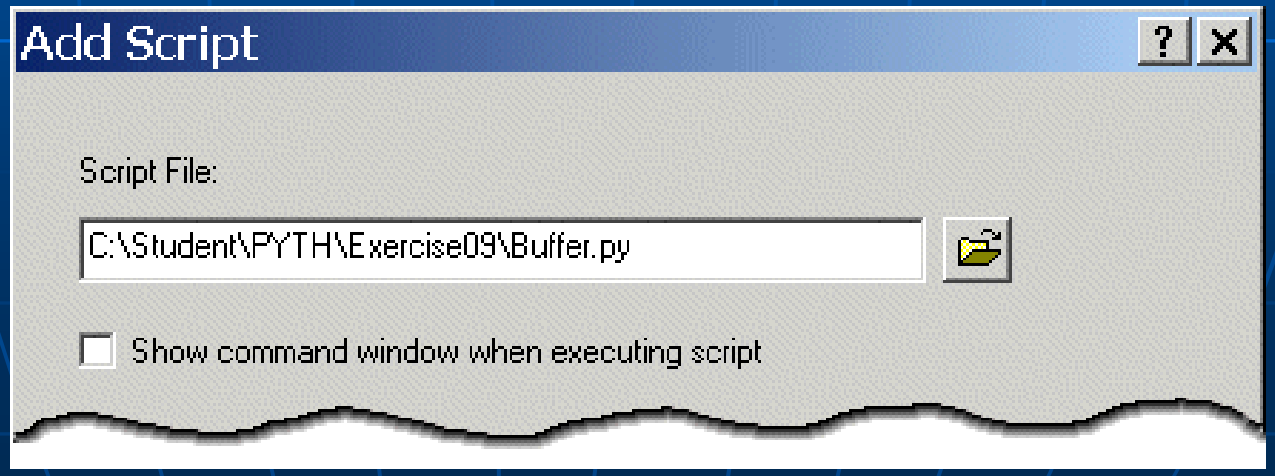
Stylesheet →



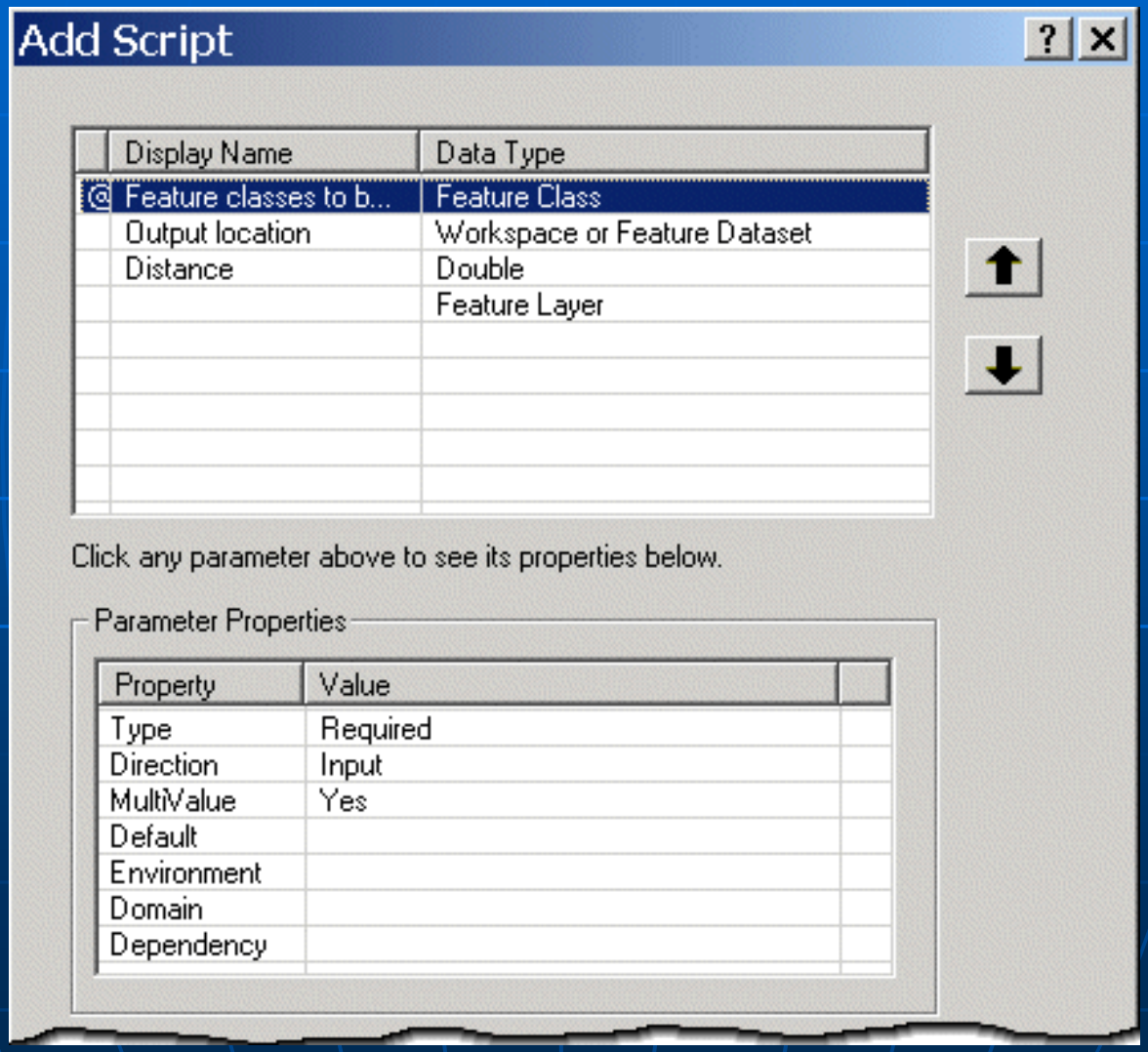
Setting the script source

- Script File
  - Add the script that will be executed from the tool
  - Can add Python, VBScript, JScript, AML, and EXE

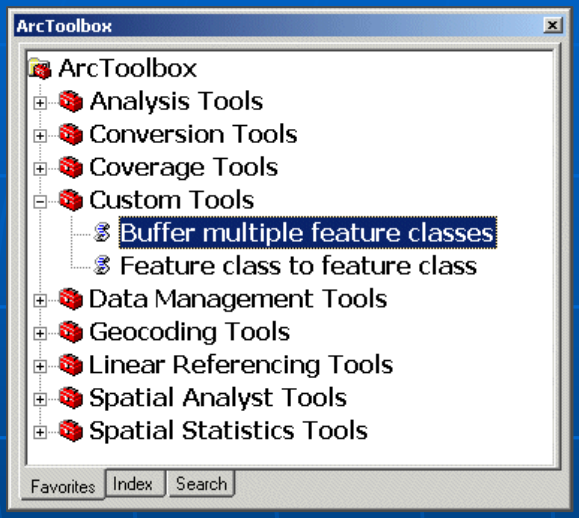
Script File

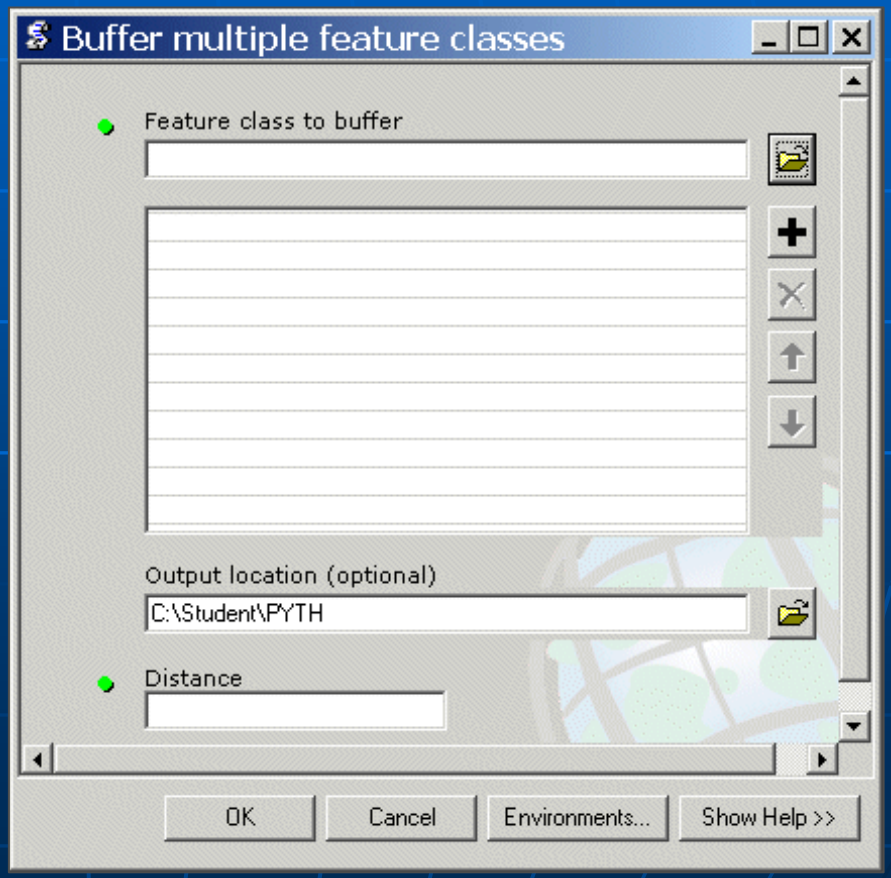


Parameter properties:

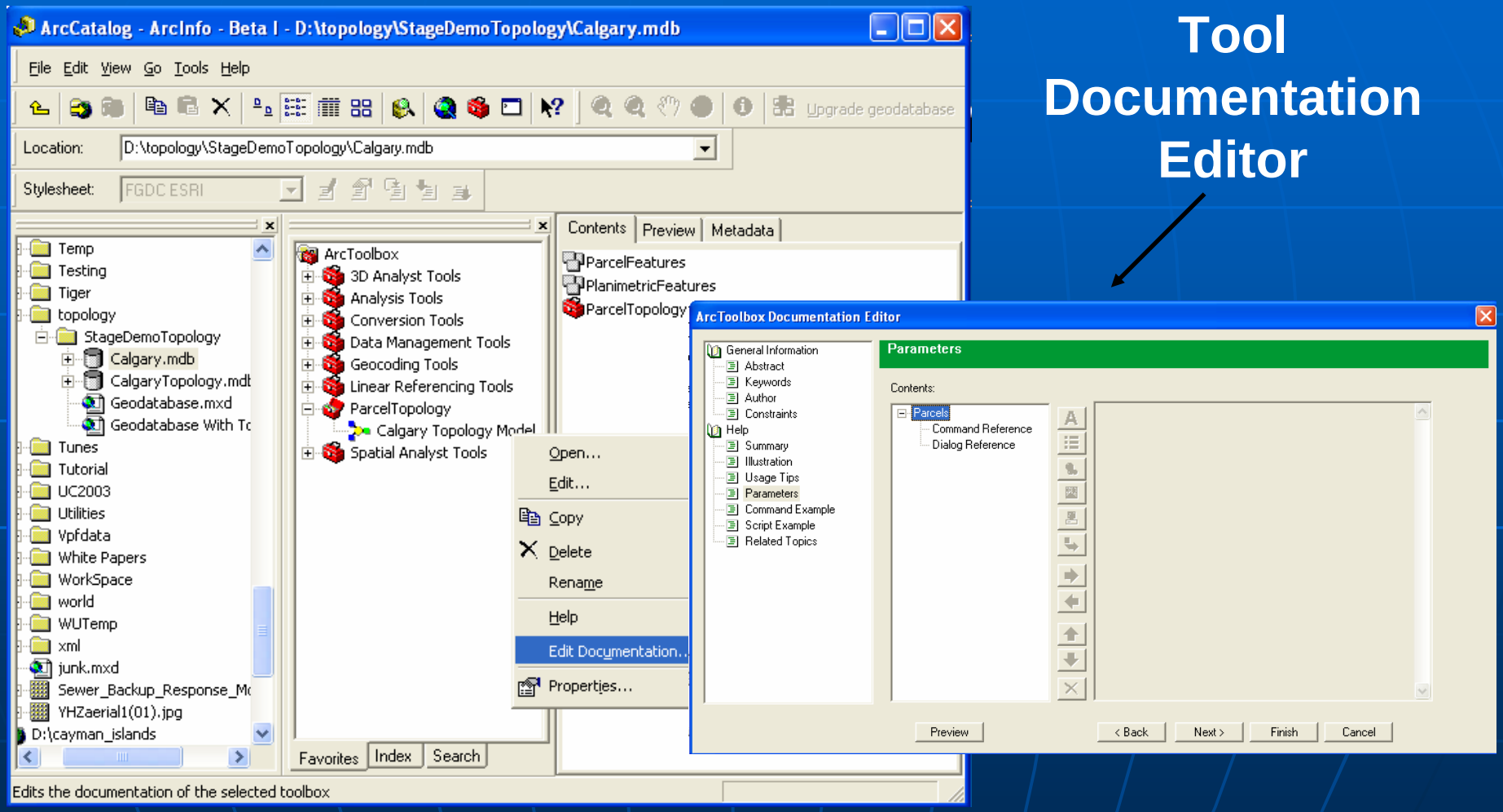


Tool User Interface is created for you:





Documentation:

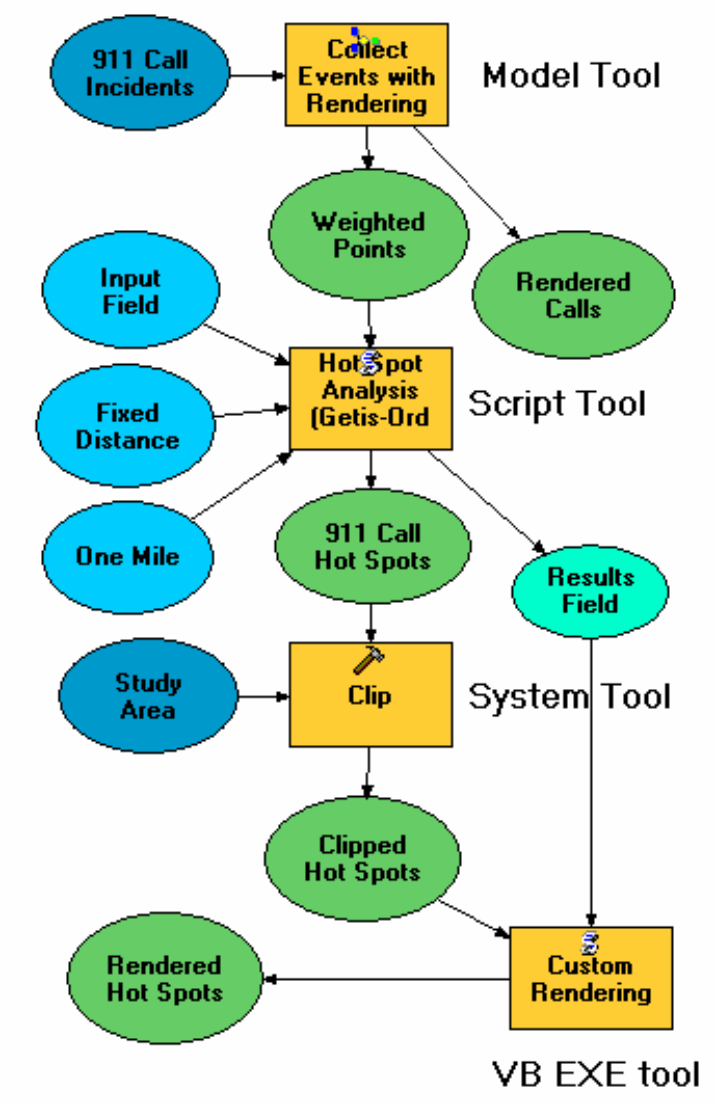


Tool Documentation Editor

“How do I get started?”

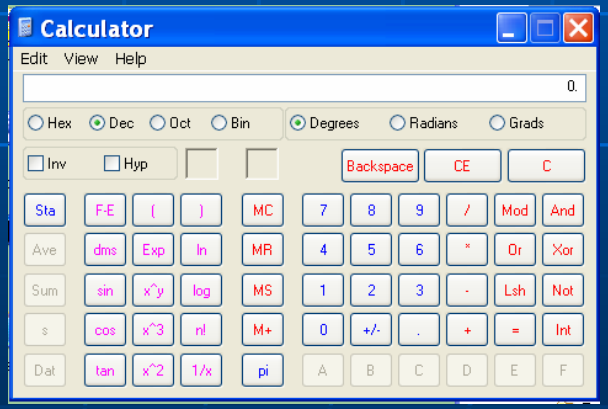
Option1: **Create a New Model Tool:**

- Drag and drop
- Existing functionality
  - System tools
    - Clip
    - Buffer
    - Select
- Custom functionality
  - Custom models
  - Custom scripts
  - Proprietary software



Option 2: **Integrate existing software:**

- Any software package that can be launched from the command line, is a custom script tool candidate:
  - In-house software packages
  - 3rd Party Software
  - Script tools and utilities



Microsoft Office Online

sas

Option 3: **Modify existing scripts:**

- Copy script tool (.py file)
- Add copy to custom toolbox
- Set up parameters
- Right click tool
- Select EDIT

Calculate Area

```
while pRow <> None:
    pGeometry = pRow.GetValue (properties.sShapeField)
    sArea = pGeometry.Area
    pRow.SetValue(sAreaField,sArea)
```

Calculate Perimeter

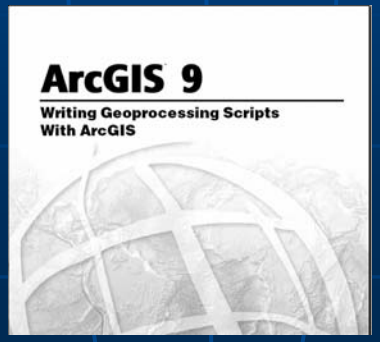
```
while pRow <> None:
    pGeometry = pRow.GetValue (properties.sShapeField)
    sLength = pGeometry.Length
    pRow.SetValue(sLengthField,sLength)
```

Option 4: **Learn Python**:

- *Learning Python (2nd Edition), Mark Lutz & David Asher*
- *Writing Geoprocessing Scripts with ArcGIS, ESRI*

  http://support.esri.com/geoprocessing

  see: "ArcGIS Desktop product documentation" (Login required)
- *Python Essential Reference (2nd Edition), David M. Beasley*
- *ESRI Instructor-Led Course*
- *www.python.org*



Building Custom Analysis Tools

Based on Research by Bryan Hill

Demo

Resources and Support

http://support.esri.com/...

Virtual Campus Free Seminars

- – "Integrating SAS® Software with ArcGIS"
- – "Understanding Spatial Statistics" ← *coming soon*

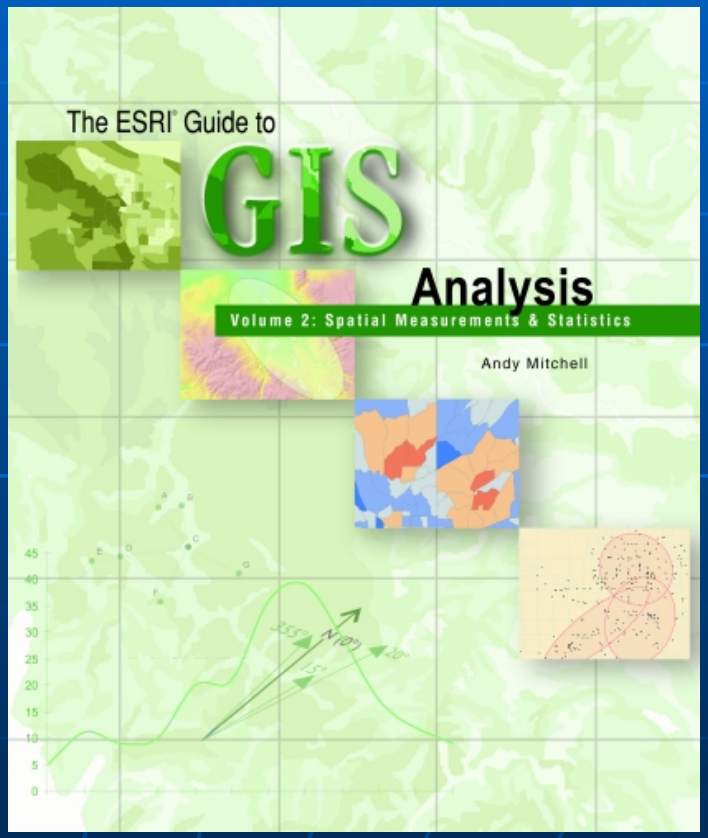
The ESRI Guide to GIS Analysis
Volume 2

Chapter 2: Measuring Geographic Distributions

Chapter 3: Analyzing patterns

Chapter 4: Mapping clusters

Chapter 5: Analyzing relationships



Questions?

Lauren M. Scott, Ph.D.
LScott@ESRI.com