

Esri International User Conference | San Diego, CA Technical Workshops | July 2011

Spatial Pattern Analysis: Mapping Trends and Clusters

Lauren M. Scott, PhD Lauren Rosenshein, MS Mark V. Janikas, PhD

Presentation Outline

- Spatial statistics overview
- Descriptive spatial statistics
- Inferential spatial statistics
 - Analyzing broad spatial patterns
 - Mapping clusters



DEMOS

- The spatial pattern of piracy
- Exploring Childhood Obesity using Hot Spot Analysis
- Zeroing in on the "Pocketman"

Kindly complete a course evaluation: www.esri.com/sessionevals

What are spatial statistics in ArcGIS?

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

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

What kinds of questions can you answer?



Where are gang territories overlapping in our city?

What kinds of questions can you answer?



Where are hot spots of foreclosures in the U.S.?

What kinds of questions can you answer?



Why are people dying young in South Dakota?

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.





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.

Charlen tring terett	Pythewww (015 py)
Die Die Des Destrants beer Des Die Die Die	Deletersenshabet
All Taking All Taking All Andrew Yook All Andrew	Construction of the second secon
Foreparture tray	

Spatial Statistics Tools		
Analyzing Patterns		
-3 Average Nearest Neighbor		
————————————————————————————————————		
————————————————————————————————————		
Spatial Autocorrelation (Morans I)		
😑 🗞 Mapping Clusters		
————————————————————————————————————		
— State of Analysis (Getis-Ord Gi*)		
Measuring Geographic Distributions		
-5 Central Feature		
————————————————————————————————————		
-3 Linear Directional Mean		
💲 Mean Center		
-3 Median Center		
-Standard Distance		
🖻 🗞 Modeling Spatial Relationships		
————————————————————————————————————		
————————————————————————————————————		
St Ordinary Least Squares		
🕀 🇞 Rendering		
🖻 🗞 Utilities		
S Calculate Areas		
— S Calculate Distance Band from Neighbor Count		
————————————————————————————————————		
—S Export Feature Attribute to ASCII		



Questions

- Which site is most accessible?
- Is there a directional trend or bias in sales ratios?
- What is the primary direction of urban growth?
- Where is the population center?
- Which gang has the broadest territory?



Finding the Center

• The Mean Center tool computes the average x and y coordinate, based on all features in the study area.









Measuring Distribution and Direction

(Standard Deviational Ellipse)

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



Measuring Distribution and Direction

(Standard Deviational Ellipse) Segregation Index = $\frac{E_1 \cap E_2 \cap E_3 \cap \ldots E_n}{E_1 \cup E_2 \cup E_3 \cup \ldots E_n}$ 2931680545.83 = 0.63 7994760004.92 Asian 2 Hispanic Black 🗉 🗹 White

Directional distribution



💐 Directional Distribution (Standard Deviation	nal Ellipse) 💶 🗖 🗙
 Input Feature Class 	
 Output Ellipse Feature Class 	
	<u> </u>
Ellipse Size	
1_STANDARD_DEVIATION	~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~~
2_STANDARD_DEVIATIONS	







The Spatial Pattern of Piracy

Demo



Inferential Statistics

- Start with a null hypothesis
 - 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



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

- P-values are probabilities
 - Small p-values, like 0.01, mean it is UNLIKELY the pattern is random
- Z-scores can be mapped to specific p-values
 - p-value 0.01 = z-score +/- 2.58



Given the 2-score of 5.73, there is a less than 1% likelihood that this dustered pattern could be the result of random chance.

Minimizing Subjectivity

Demo



Inferential statistics tools

- Analyzing broad spatial patterns
 - How intense is the clustering?

🖃 🗞 Analyzing Patterns

- 🦾 💐 Spatial Autocorrelation (Morans I)
- Mapping clusters
 - Where is the clustering?
 - 🖃 🚳 Mapping Clusters

 - 🏧 🎒 Hot Spot Analysis (Getis-Ord Gi*)



- 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 foreclosure rates?
- Are new AIDs cases remaining Geographically fixed or are they spreading to nearby counties?





- Which houses sold for much more than expected?
- Where do we find anomalous spending patterns?

- Where are the 911 call, crime, arson, disease... hot spots?
- Where do we see unexpectedly high rates of urban growth?



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











Pocket Man Analysis

- Potentially over 300 cases.
- Data from 1990 to 2006: 156 cases.
- Suspect apprehended in January, 2008.

Animation



Distinct spatial regimes

• Bergen

- Oslo
- Around Norway



Temporal patterns

- DOW analysis shows most incidents occur on Friday or Saturday
- Weekend incidents are more common outside of Bergen and Oslo





- Graphing incidents over time reveals 3 distinct time periods.
- Most of the incidents prior to Dec 99 are in Bergen

Space/Time Patterns

- Clustering is strongest prior to 1977 and the range of incidents is narrow.
- Incidents become increasingly random with each time period.





Hot Spot Analysis



Focusing on Bergen and Oslo

- Mean Center analysis shows the center of activities for each time period.
- Standard Deviational Ellipse analysis shows where activities overlap for each time period.







Resources for learning more...

- Spatial Pattern Analysis: Mapping Trends and Clusters
 - Tue 8:30 Rm 2; Wed 3:15 Rm 2
- Modeling Spatial Relationships using Regression Analysis
 - Tue 10:15 Rm 2; Thu 1:30 Rm 1A/B
- Spatial Statistics: Best Practices
 - Tue 3:15 Rm 2; Thu 3:15 Rm 1A/B
- Using R in ArcGIS
 - Wed 12:00 Rm 1A/B
- Road Ahead: Sharing of Analysis (ArcGIS 10.1)
 - Wed 11:05 Rm 6B

Resources for learning more...



QUESTIONS?

- www.esriurl.com/spatialstats
- Short videos
- Articles and blogs
- Online documentation
- Supplementary model and script tools
- Hot Spot, Regression, and ModelBuilder tutorials



<u>LScott@Esri.com</u> LRosenshein@Esri.com