Spatial Pattern Analysis: Mapping Trends and Clusters

Lauren M. Scott, PhD
Lauren Rosenshein Bennett, MS
Presentation Outline

- Spatial problem solving
- Spatial statistics overview
- Describing spatial patterns
- Quantifying spatial patterns
- Mapping spatial clusters
  - Hot and cold spots
  - Spatial outliers
  - Similar features

DEMO
- Exploring regional variations in health care.

Kindly complete a course evaluation: www.esri.com/ucsessionssurveys
The Pocket Man
The Pocket Man
All Incidents
The Pocket Man
Spatial Clusters
The Pocket Man
Time (early events)
The Pocket Man

Time (mid events)
The Pocket Man

Time (late events)
The Pocket Man
Day of the week
The Pocket Man
Hot Spot Analysis

- Bergen Hot Spot
- Oslo Hot Spot
- Oslo Cold Spot
The Pocket Man
Caught

Norwegian Police: 'Pocket Man' Arrested in Sex Abuse of Hundreds of Boys

Associated Press

OSLO, Norway — One of the largest, highest-profile manhunts in Norwegian history may have succeeded Friday when police arrested a suspect they call "The Pocket Man" in the sexual abuse of 300 to 400 young boys over three decades.

Police said the suspect, identified only as a 55-year-old businessman from the western city of Bergen, was linked to five of the assaults through his DNA and that they expected to prove more cases.
What are Spatial Statistics in ArcGIS?

• A set of exploratory methods and techniques
• Specifically developed for use with geographic data
  - They incorporate space (area, length, proximity, orientation) directly into their mathematics
• They describe and model:
  - Spatial Distributions
  - Spatial Patterns
  - Spatial Processes
  - Spatial Relationships

*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.
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.
Describing Spatial Data

• **Questions**
  - Which site is most accessible?
  - What is the primary wind direction in the winter months?
  - Where is the population center and how it is changing over time?
  - Which species has the broadest spatial distribution? Which are the most spatially integrated?
Finding the center

- The Mean Center tool computes the average x and y coordinate, based on all features in the study area.
Distribution and Direction
(Standard Deviational Ellipse)

- Abstracting spatial and temporal trends in a distribution of features
Ellipse size

Normal distribution

- Mean

99%

95%

68%

1 = 68% of features

2 = 95% of features

3 = 99% of features
Analyzing territories
(Standard Deviational Ellipse)

- Center
- Dispersion
- Orientation
Measuring segregation and integration
(Standard Deviational Ellipse)

Segregation Index =

\[
1 - \frac{E_1 \cap E_2 \cap E_3 \cap \ldots \cap E_n}{E_1 \cup E_2 \cup E_3 \cup \ldots \cup E_n}
\]

\[
1 - \frac{2931680545.83}{7994760004.92} = 0.63
\]
Quantifying Spatial Patterns

Which pattern is most clustered?
Quantifying Spatial Patterns

Which pattern is most clustered?

Natural Breaks
Equal Interval
Quantile
Quantifying Spatial Patterns

You say it’s clustered

- Compared to what?
- Says who?
Compared to what?
Complete Spatial Randomness

- Inferential statistics start with a null hypothesis
  - Random distribution of features
  - Random distribution of values within fixed features
Says who?
Probability theory

- Inferential statistics report a p-value and z-score
  - z-scores are standard deviations
  - p-values are probabilities

- z-scores can be mapped to specific p-values
  - p-value 0.01 = z-score +/- 2.58
  - p-value 0.05 = z-score +/- 1.96
  - p-value 0.10 = z-score +/- 1.65
Quantifying spatial patterns over time
Spatial Autocorrelation (Global Moran’s I)

Thematic maps showing relative per capita Income for New York by county, 1969 to 2002

Is the spatial “gap” between rich and poor increasing or dissipating?
Quantifying spatial patterns over time

Spatial Autocorrelation (Global Moran’s I)

- 1969: 5.21
- 1985: 4.26
- 2002: 2.4

Analyzing Patterns:
- Average Nearest Neighbor
- High/Low Clustering (Getis-Ord General G)
- Incremental Spatial Autocorrelation
- Multi-Distance Spatial Cluster Analysis (Ripley’s K Function)
- Spatial Autocorrelation (Moran’s I)
Mapping Spatial Clusters

- Some of the questions you can answer:
  - Where do we find anomalous spending patterns?
  - Where do we see unexpectedly high rates of suicide?
  - Do burglaries committed during the day have the same spatial pattern as those occurring at night?
  - Which disease incidents are likely part of the same outbreak?

Which homes sold for more than expected?

Where are the 911 call hot spots?

Which countries face similar challenges?
Exploring regional variations in health care

Lauren Rosenshein Bennett, MS
Examining the spatial pattern of poverty in Kenya

(Cluster and Outlier Analysis: Anselin Local Moran’s I)
Grouping Analysis

Provide:
- Number of groups
- Variables to use for grouping
- Spatial-temporal constraints (if any)
- Option to evaluate optimal number of groups
Grouping Analysis

Group 1: low income, low education, high unemployment, high crime

Group 2: middle values for all variables

Group 3: higher income and education, lower unemployment and crime
Grouping Analysis Reports

### Overall Variable Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
<th>R2</th>
</tr>
</thead>
<tbody>
<tr>
<td>CRIMECNT</td>
<td>42.4021</td>
<td>59.1462</td>
<td>0.0000</td>
<td>593.0000</td>
<td>0.2852</td>
</tr>
<tr>
<td>UNEMPRAPE</td>
<td>0.1982</td>
<td>0.1273</td>
<td>0.0000</td>
<td>0.8000</td>
<td>0.1887</td>
</tr>
<tr>
<td>MEDHHINC</td>
<td>49379.4271</td>
<td>22747.1008</td>
<td>0.0000</td>
<td>225000.0000</td>
<td>0.1296</td>
</tr>
<tr>
<td>HSDROPRATE</td>
<td>0.2393</td>
<td>0.1547</td>
<td>0.0000</td>
<td>1.0000</td>
<td>0.0762</td>
</tr>
</tbody>
</table>

### UNEMPRAPE: R2 = 0.19

<table>
<thead>
<tr>
<th>Group</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
<th>Share</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.3057</td>
<td>0.1298</td>
<td>0.0000</td>
<td>0.5883</td>
<td>0.7354</td>
</tr>
<tr>
<td>2</td>
<td>0.2407</td>
<td>0.1287</td>
<td>0.0000</td>
<td>0.8000</td>
<td>1.0000</td>
</tr>
<tr>
<td>3</td>
<td>0.1403</td>
<td>0.0954</td>
<td>0.0000</td>
<td>0.5962</td>
<td>0.7452</td>
</tr>
<tr>
<td>Total</td>
<td>0.1982</td>
<td>0.1273</td>
<td>0.0000</td>
<td>0.8000</td>
<td>1.0000</td>
</tr>
</tbody>
</table>
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
Resources for learning more…

During the conference

<table>
<thead>
<tr>
<th>Technical Workshops</th>
<th>Room 5A/B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modeling Spatial Relationships Using Regression</td>
<td>T 10:15</td>
</tr>
<tr>
<td>What’s New in Spatial Statistics</td>
<td>T 3:15</td>
</tr>
<tr>
<td>Spatial Pattern Analysis</td>
<td>W 1:30</td>
</tr>
<tr>
<td>Modeling Spatial Relationships Using Regression</td>
<td>W 3:15</td>
</tr>
<tr>
<td>Spatial Statistics Best Practices</td>
<td>Th 1:30</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Demos</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Hot Spot Maps: Minimizing Subjectivity</td>
<td>W 5:30</td>
</tr>
<tr>
<td>Geoprocessing and Analysis Demo Theater</td>
<td></td>
</tr>
<tr>
<td>Crime Mapping: Using Spatial Statistics</td>
<td>Th 11:00</td>
</tr>
<tr>
<td>Public Safety Showcase Demo Theater</td>
<td></td>
</tr>
</tbody>
</table>

- Integrating open-source statistical packages Rm 1A 12:00
- Working with Temporal Data Rm 28D Th 10:15AM
Resources for learning more…

After the conference

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

QUESTIONS?

LRosenshein@Esri.com
LScott@Esri.com

Esri.com/ucsessionssurveys  Session ID: 117/594
Spatial Pattern Analysis:
Mapping Trends and Clusters

Lauren M. Scott, PhD
Lauren Rosenshein Bennett, MS
Resources for learning more…
During the conference

<table>
<thead>
<tr>
<th>Technical Workshops</th>
<th>Room 5A/B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Modeling Spatial Relationships Using Regression</td>
<td>W 3:15</td>
</tr>
<tr>
<td>Spatial Statistics Best Practices</td>
<td>Th 1:30</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Demos</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hot Spot Maps: Minimizing Subjectivity</td>
</tr>
<tr>
<td>Geoprocessing and Analysis Demo Theater</td>
</tr>
<tr>
<td>Crime Mapping: Using Spatial Statistics</td>
</tr>
<tr>
<td>Public Safety Showcase Demo Theater</td>
</tr>
</tbody>
</table>

- Integrating open-source statistical packages Rm 1A 12:00
- Working with Temporal Data Rm 28D Th 10:15AM
Resources for learning more…
After the conference

QUESTIONS?

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

- resources.arcgis.com

LRosenshein@Esri.com
LScott@Esri.com

Esri.com/ucsessionsurveys
Session ID: 117/739