EDA and ESDA

Luc Anselin
Spatial Analysis Laboratory
Dept. Agricultural and Consumer Economics
University of Illinois, Urbana-Champaign
http://sal.agecon.uiuc.edu
Outline

- Concepts
- Statistical Graphs
- Linking and Brushing
- Statistical Maps
Concepts
EDA

- EDA
  - Discover Potentially Explicable Patterns (Good)

- Data Visualization (Buja)
  - View = graphical representation and/or summary
  - Interactive View Manipulation
    - focusing individual views
    - linking multiple views
    - arranging many views

- No Role for Explicit Treatment of Space in traditional EDA
Exploratory Spatial Data Analysis

- EDA +
- Describe Spatial Distributions
  - spatial trends, spatial means
- Identify Atypical Observations
  - spatial outliers
- Discover Patterns of Spatial Association
  - spatial clusters
- Suggest Spatial Regimes
  - spatial non-stationarity
ESDA Tools
(areal data)

- **Visualize Spatial Distributions**
  - augmenting the information in the map with formal statistical graphs or glyphs

- **Dynamic Graphics**
  - linking and brushing statistical plots and map

- **Visualize Spatial Autocorrelation**
  - providing a graphical representation of magnitude and significance of global and local measures of spatial autocorrelation
Statistical Graphs
Outliers in Box Plot

outlier = outside “fence”
fence = Q3 + 1.5 times IQR inter-quartile range 75%-25%
Parallel Coordinate Plot (PCP)
Figure 18: Scatterplot matrix of log base 2 of the haystack data.
Trellis Display

or Conditional Plot
Conditional Scatterplots
Conditional QQ-Plots
Linking and Brushing
Linking

- **Views**
  - different “views” of data
    - statistical graphs: histogram, box plot, scatterplot
    - map
    - Table (list)

- **Dynamic Linking**
  - views dynamically linked
    - click on one view and corresponding observations (points, areas) on other views are highlighted
linking views - same Kreise are highlighted in all (Regard/Manet Unwin et al.)
Linking Point and Polygon Maps
Dynamic Linking and Multimedia - panoraMap
Brushing

- Brushing
  - moving “brush” over map or graph highlights matching observations in other statistical graphs and vice versa

- Brushing Scatterplots
  - recalculates slope of regression line

- Geographic Brushing
  - simultaneous selecting on multiple maps
Selection in Scatterplot

Slope = -1.5258

CRIME

INC

© 1999-2003 Luc Anselin, All Rights Reserved
Map Brushing in GeoDa
Statistical Maps
Visualizing Spatial Distributions

- **Spatialized EDA**
  - icons and glyphs matching locations
  - special case of symbol maps

- **Regional Plots**
  - spatial heterogeneity
  - different distributions in spatial subsets
  - regional histogram, regional box plot
spatial lag bar chart
blue = crime at i, red = spatial lag, average crime for neighbors
Spatialized EDA

Spatial Chernoff Faces

the burglar’s view
of crime clusters in
Columbus
Regional Histogram
Regional Box Plot

Columbus crime: core vs periphery
Micromaps (Carr)

- Linked Micromap Plots - LM plots
  - a micromap for each quantile
  - micromaps linked to other statistical graphs

- Conditioned Choropleth Maps
  - choropleth maps on dependent variable
  - micromap matrix
  - conditioning along two dimensions