



Center for Spatially Integrated Social Science

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

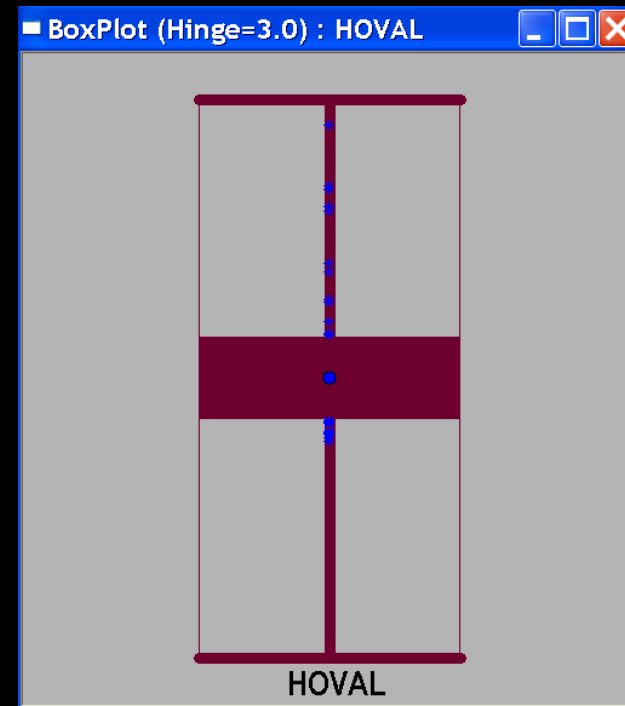
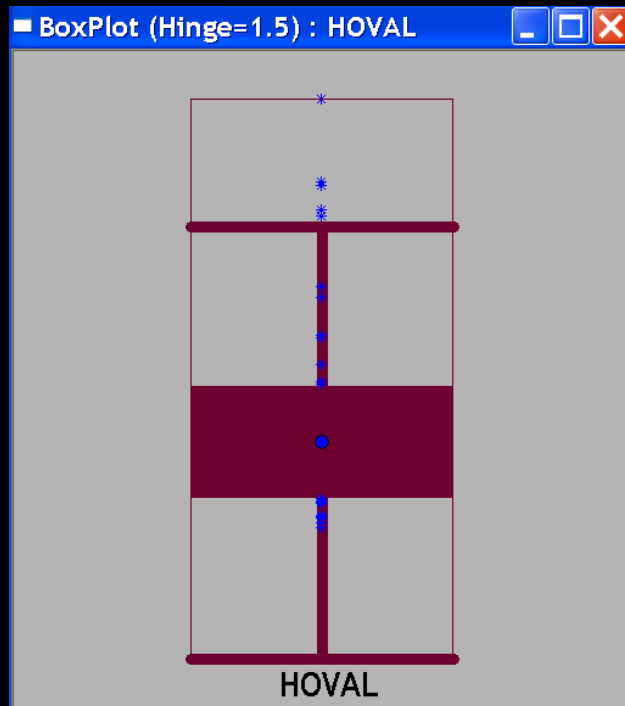
ESDA Tools

(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

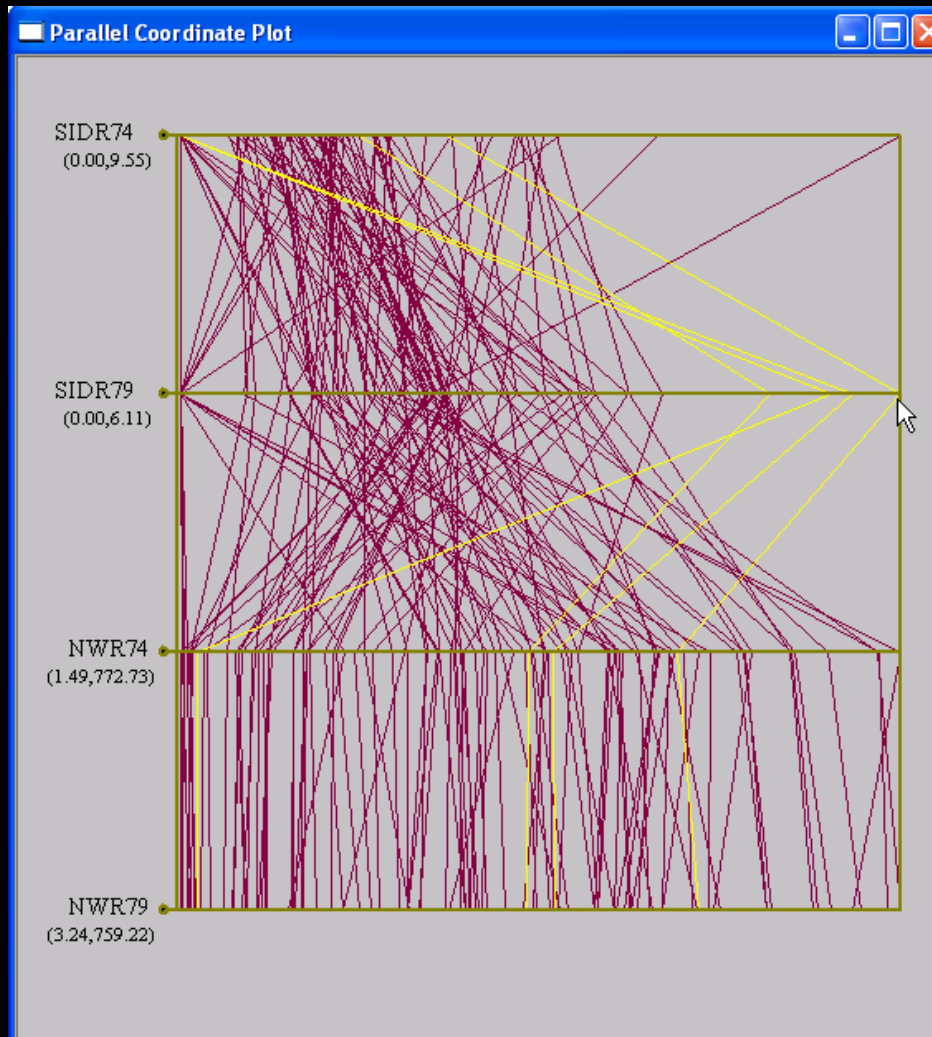
Outliers in Box Plot



outlier = outside "fence"

fence = $Q3 + 1.5 \text{ times IQR}$ inter-quartile range 75%-25%)

Parallel Coordinate Plot (PCP)



Scatterplot Matrix

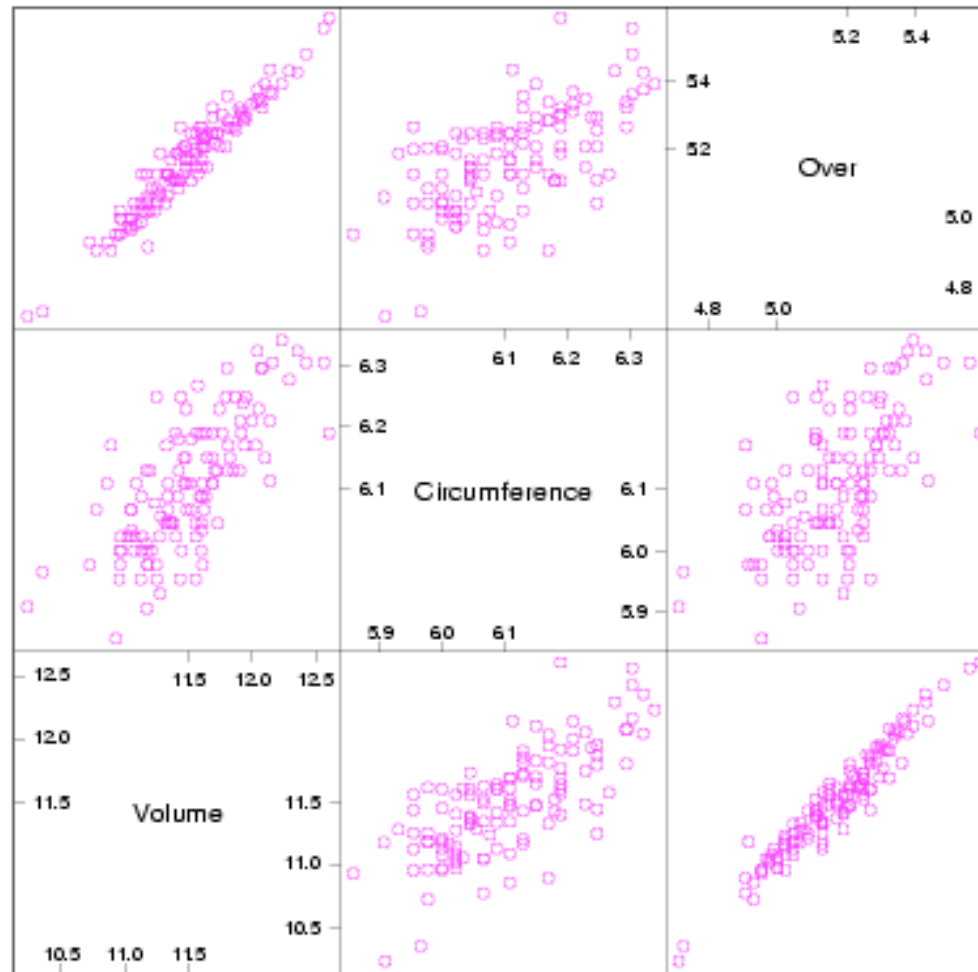
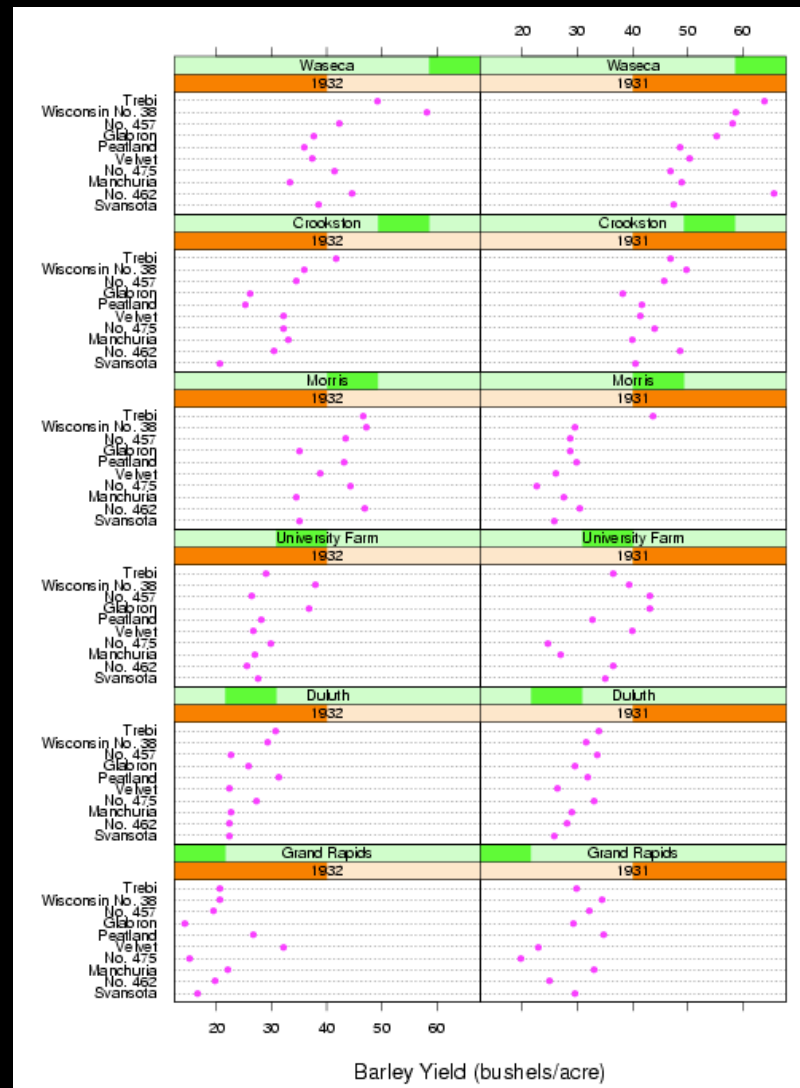


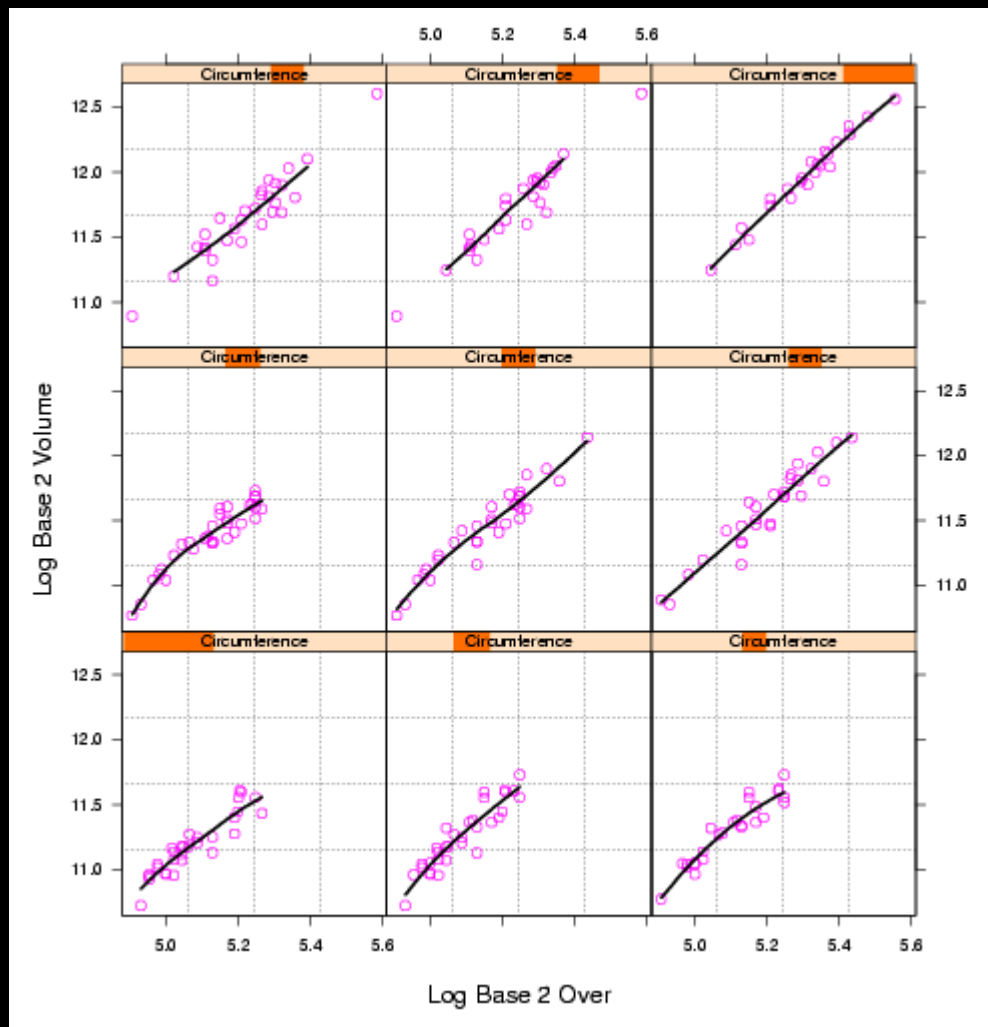
Figure 18: Scatterplot matrix of log base 2 of the haystack data.

Trellis Display

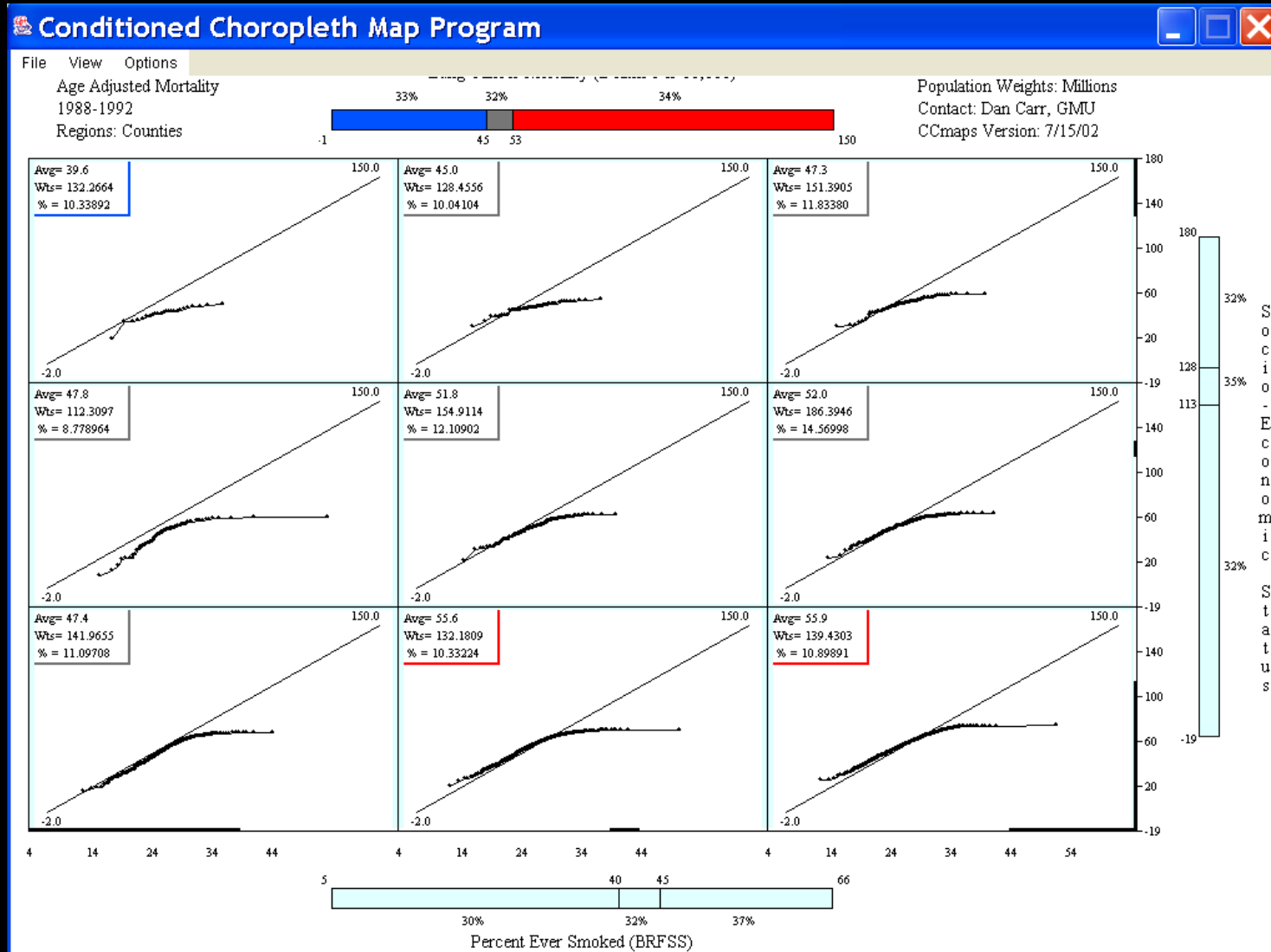
or Conditional Plot



Conditional Scatterplots



Conditional QQ-Plots



Linking and Brushing

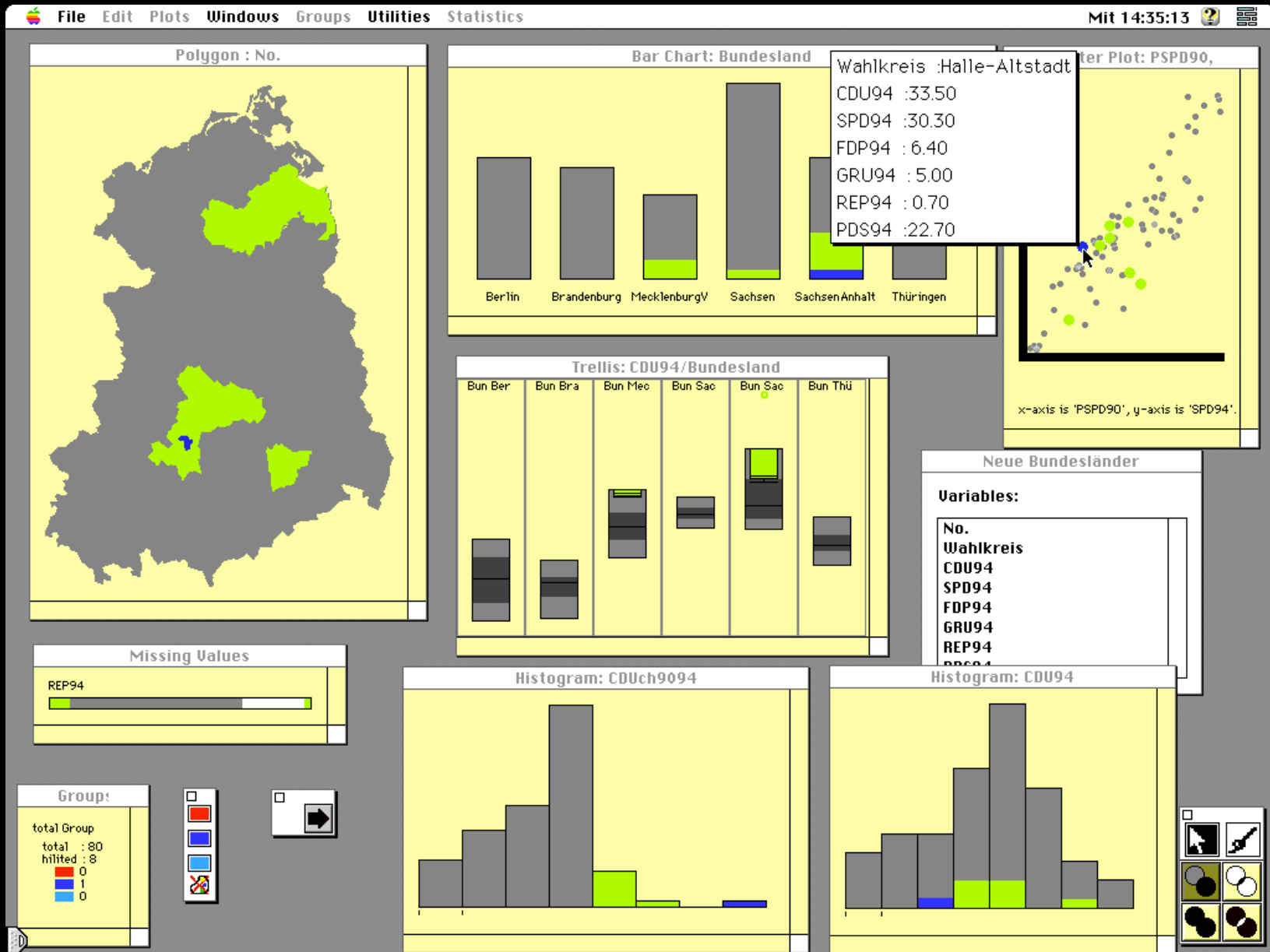
Linking

➤ Views

- different "views" of data
 - statistical graphs: histogram, box plot, scatterplot
 - map
 - Table (list)

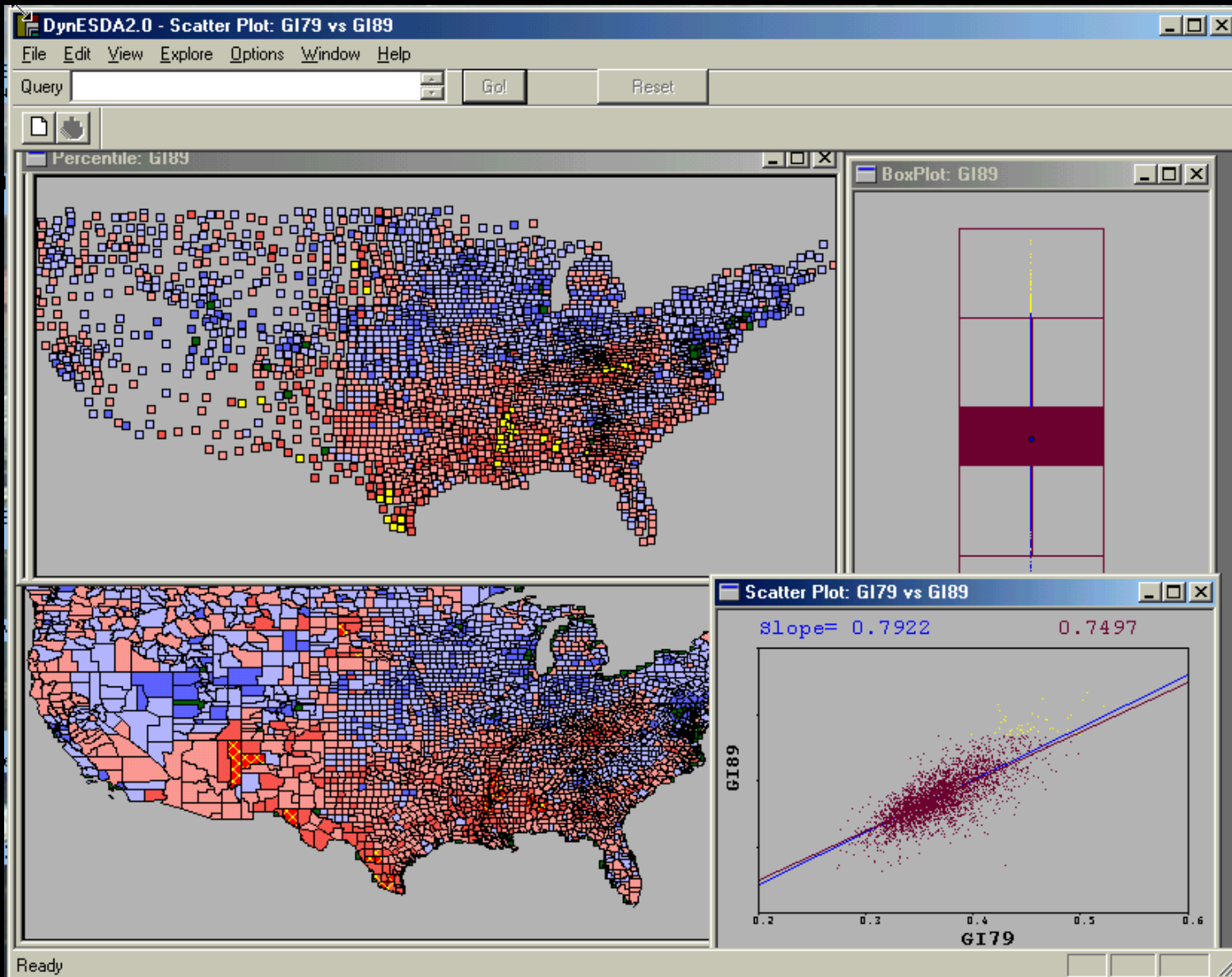
➤ Dynamic Linking

- 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

The screenshot displays the panoraMap software interface with several windows open:

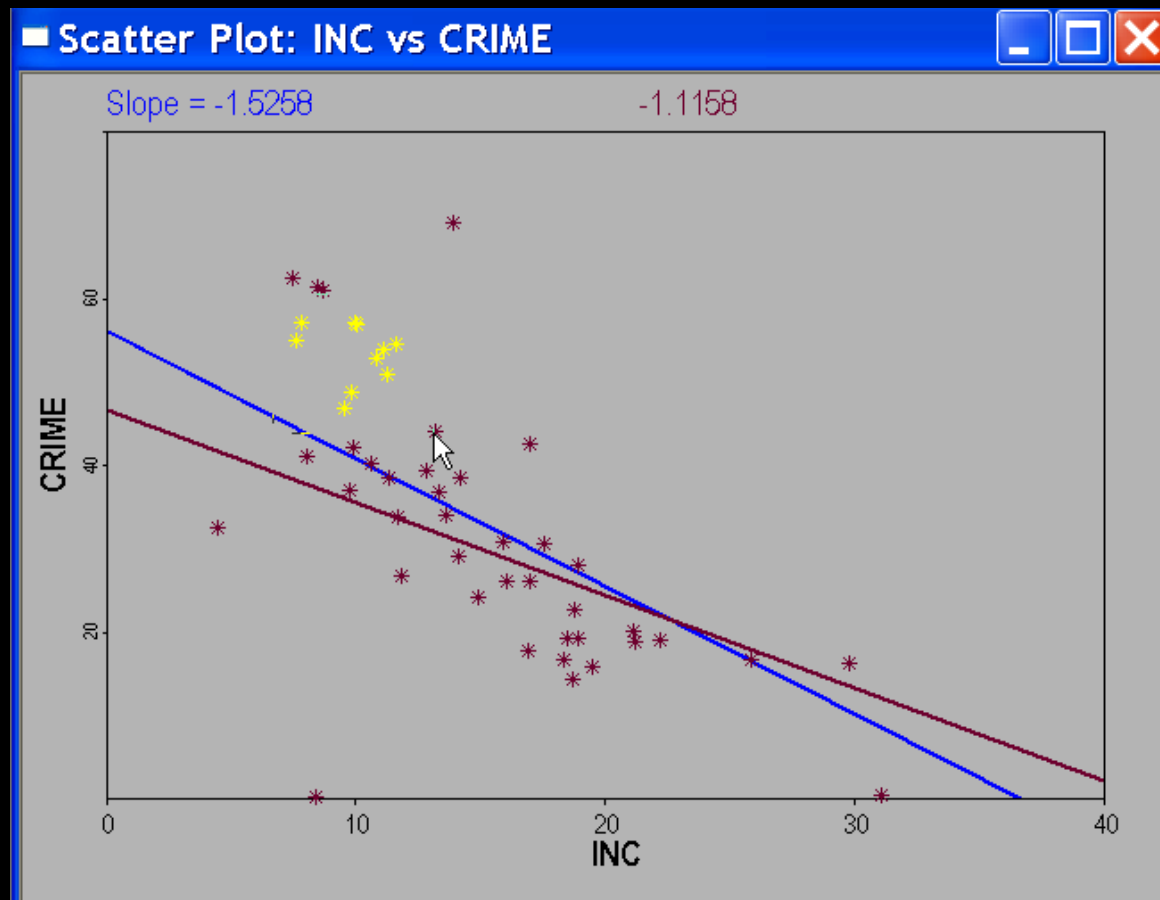
- VFC Menu Window:** Contains menu options (File, Options, Maps, GPS, Outlines, Areas, Help) and the VFC panoraMap logo. The current area is labeled "RetailArea21".
- Map Legend:** A color-coded legend for counts, ranging from 0 (magenta) to 9 (white).
- Exeter-1995: Area Map:** A large heatmap showing spatial data distribution. A green line and an orange arrow point from a specific location on the heatmap to the street view window.
- VFC panoraMap:** A window showing a street view image of a residential street.
- Select Variables:** A dialog box titled "Graph from StudentData.txt" with a list of variables: Name, Council, Sunday, Monday, Tuesday, Mon-Sun, Tue-Sun, Tue-Mon, FootCount, HighStreet, Traffic, Wednesday, Thursday, FootCount, Visited, CarsParked, and NumShops. The "FootCount" variable is checked.
- Registry List (2):** A table listing various data files and their types and timestamps.
- 600:** A window showing a list of variables: Aspect, Elevation, Exeter-1890, Exeter-1911, Exeter-1933, Exeter-1946, Exeter-1995, Landsat123, Population, ShadedDEM, Slope, and SpotImage. The "Population" variable is checked.
- Image - OxfordStreet.gif:** A window showing a street view image of a residential street with a yellow box highlighting a specific area.
- Parallel Plot:** A window showing a parallel plot visualization with a red line connecting data points across multiple vertical axes.

File Name	Type	Time
TheWeir.gif	Image	15:53
NewFlats.gif	Image	15:54
TerracedHousing.gif	Image	15:55
OxfordStreet.gif	Image	15:55
LethbridgeRoad.gif	Panorama	15:42
NewHouses.gif	Panorama	16:20
Scum.gif	Image	15:20

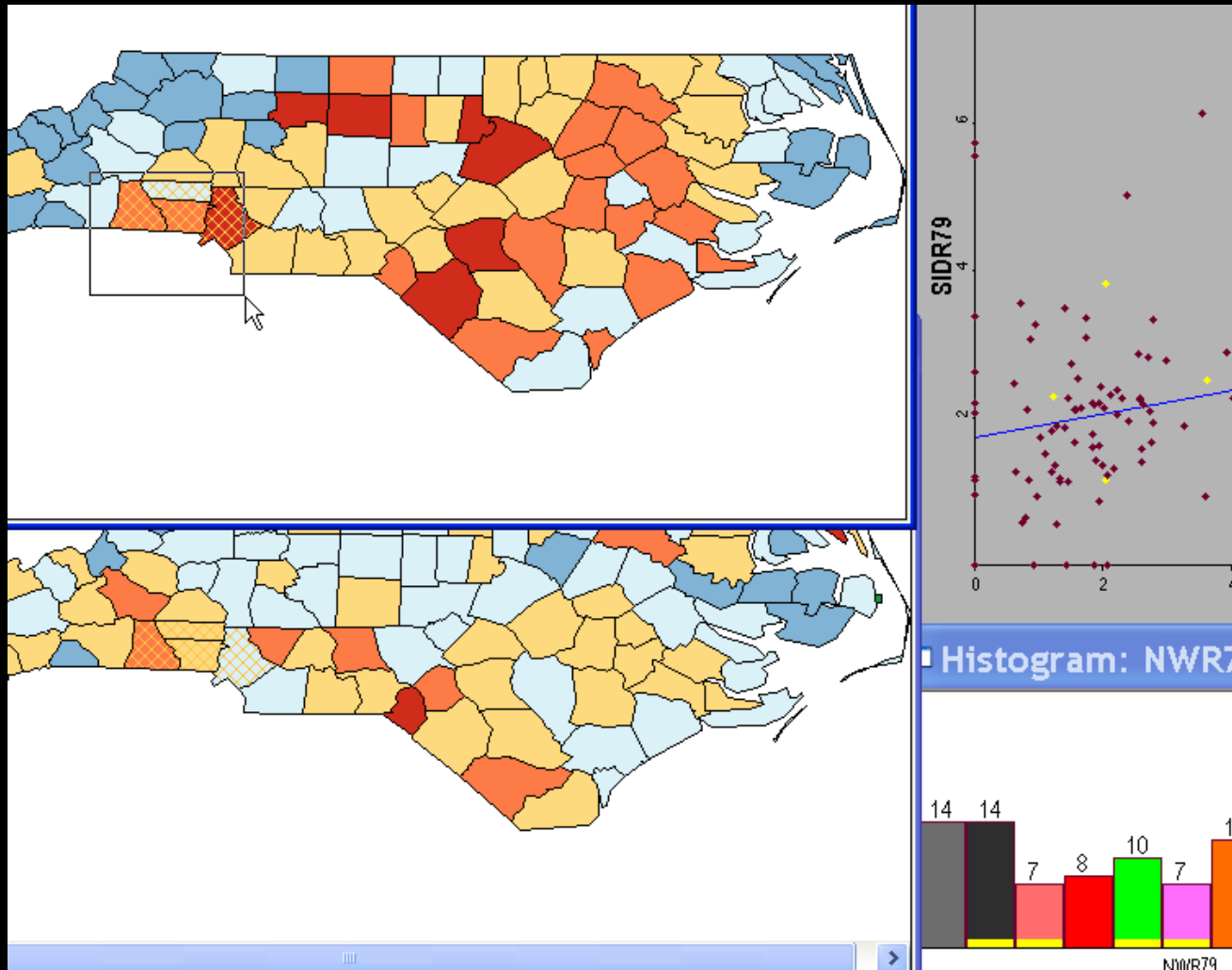
Brushing

- 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



Map Brushing in GeoDa



Statistical Maps

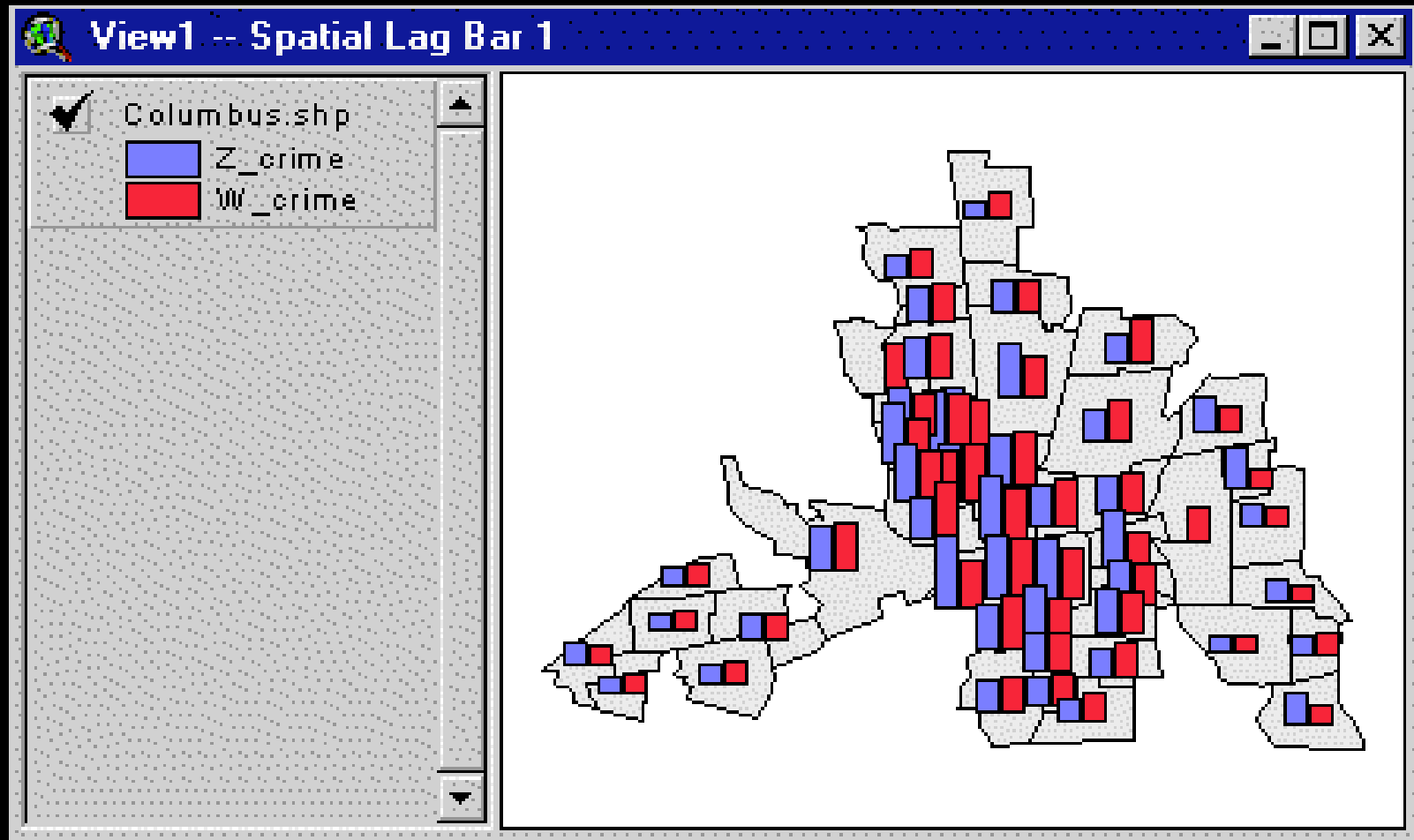
Visualizing Spatial Distributions

➤ 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



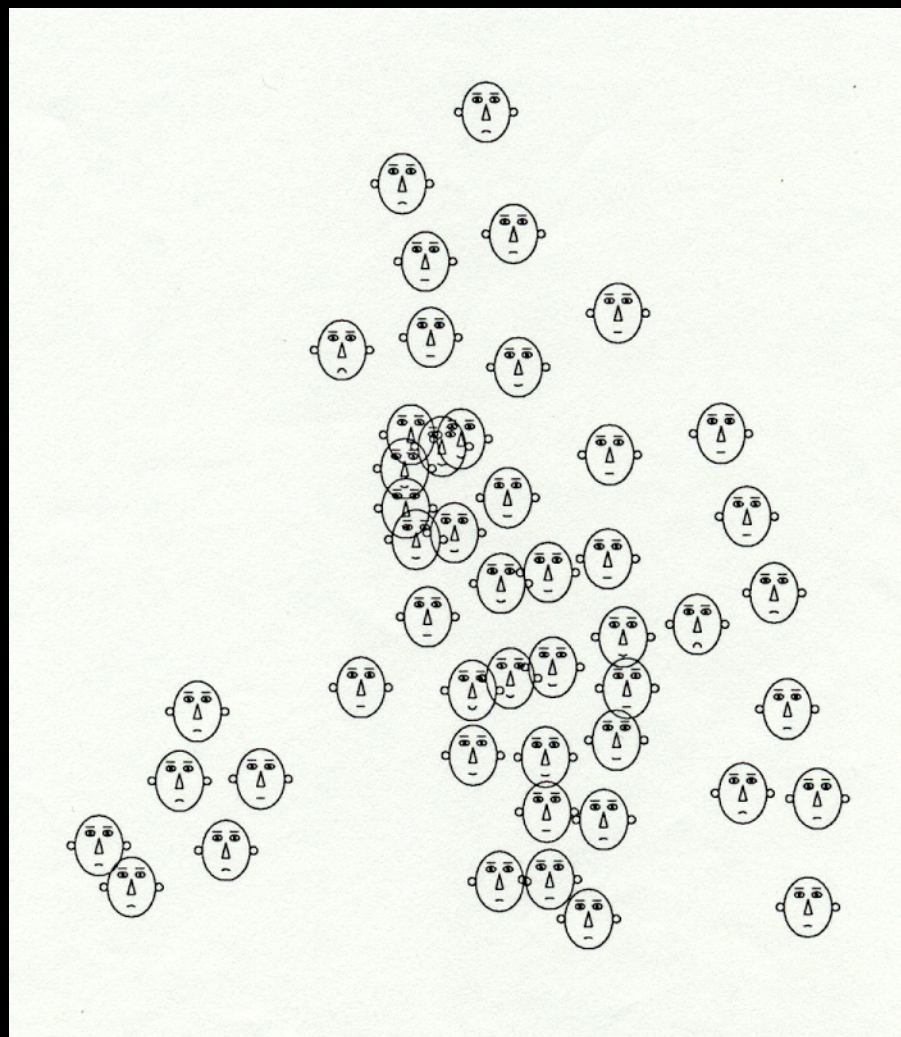
spatial lag bar chart

blue = crime at i, red = spatial lag, average crime for neighbors

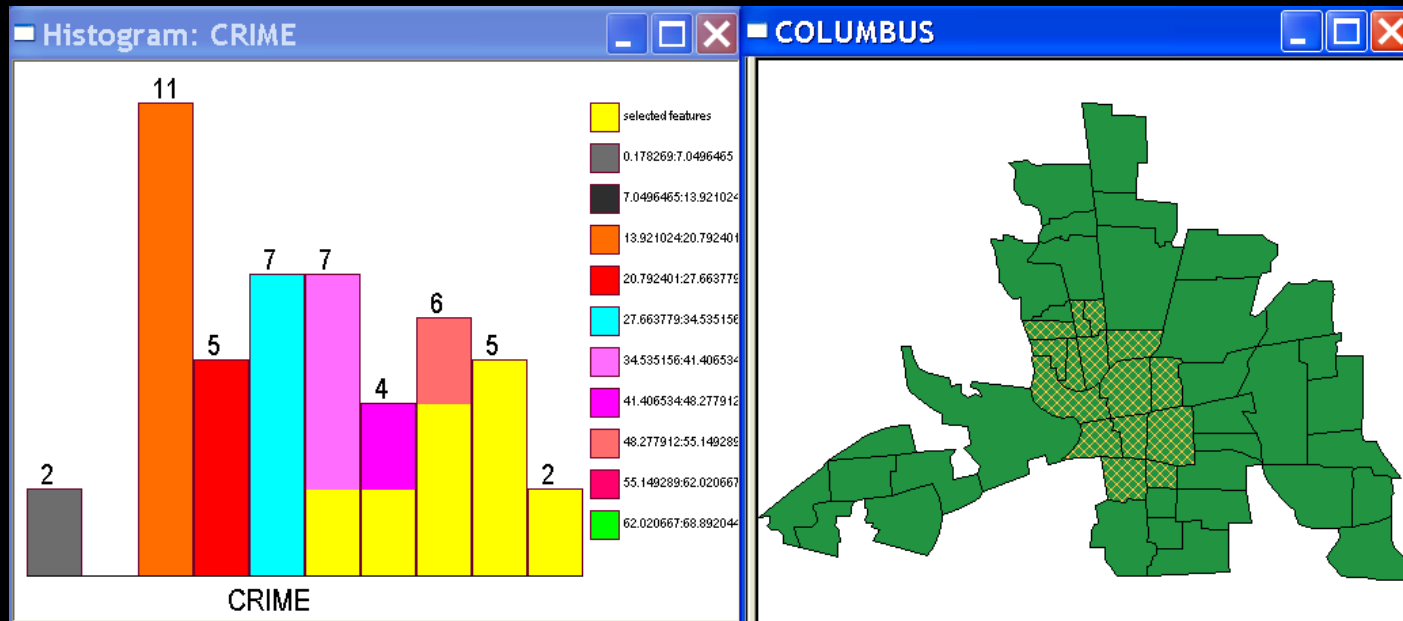
Spatialized EDA

Spatial Chernoff Faces

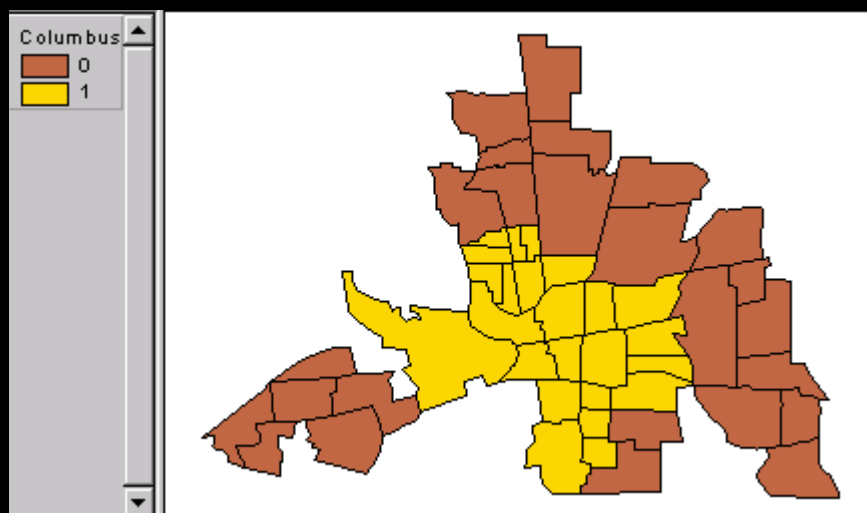
the burglar's view
of crime clusters in
Columbus



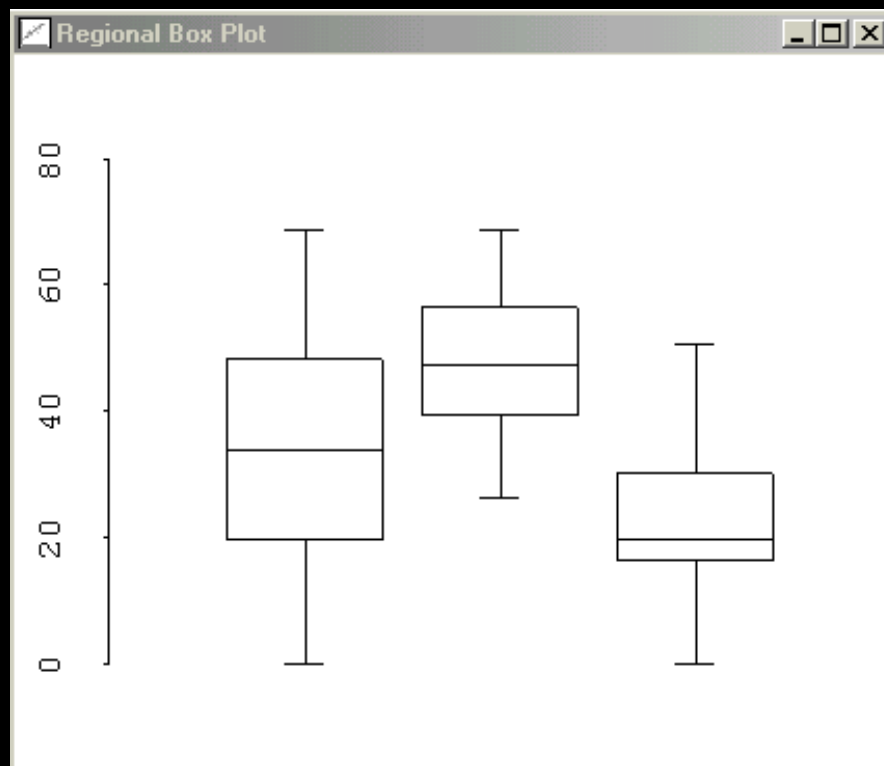
Regional Histogram



Regional Box Plot



Columbus crime: core vs periphery

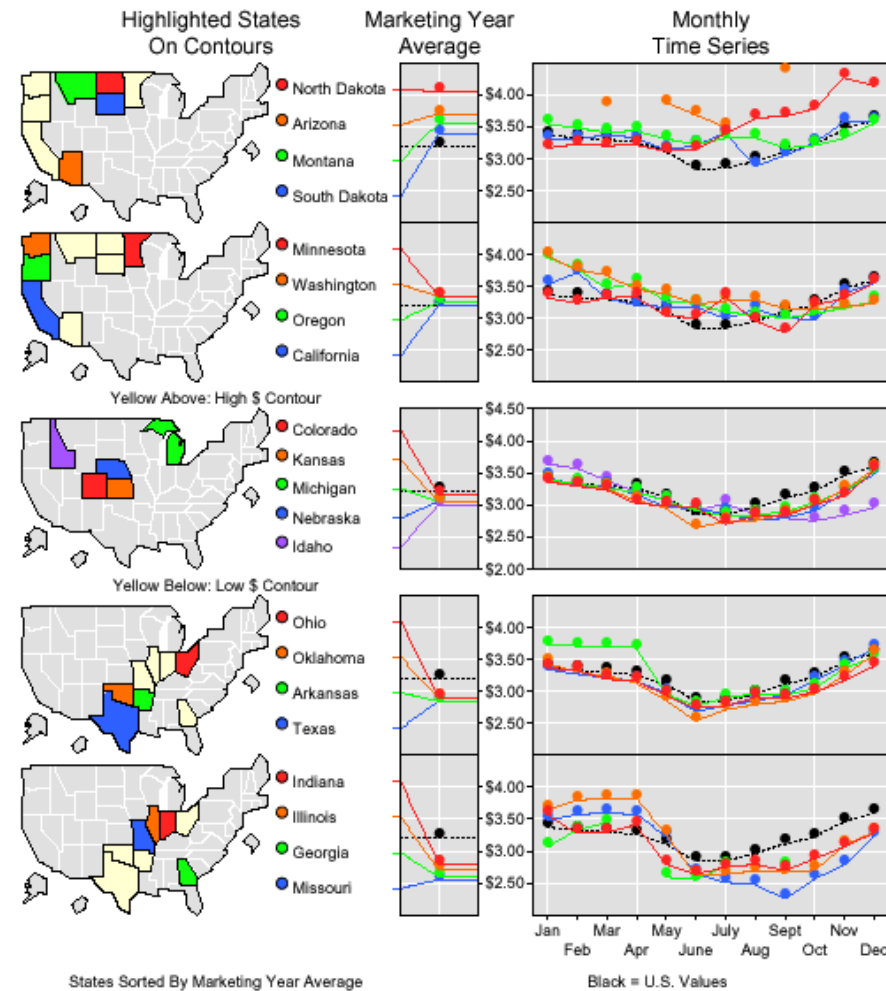


Micromaps (Carr)

- Linked Micromap Plots - LM plots
 - a micromap for each quantile
 - micromaps **linked** to other statistical graphs
- Conditioned Choropleth Maps
cc maps
 - choropleth maps on dependent variable
 - micromap matrix
 - **conditioning** along two dimensions

Linked Micromap Plots

Wheat Prices Received By State: 1993
Dollars Per Bushel



Conditioned Choropleth Map

