Introduction to Geographic Information Analysis

SERGIO REY

GPH 483/598
Geographic Information Analysis
School of Geographical Sciences and Urban Planning
Arizona State University
Fall 2010

Outline

Course Objectives

- Introduction to fundamentals of ESDA
- Conceptual background
- Hands-on

Components

Four Sections

- Introduction and Background
- Point Patterns
- Spatial Autocorrelation
- Geostatistics

Part I: Introduction

Month	Date	Topic/Readings	Out	Due
Jan	20	Introduction to Geographic Information Analysis GIA 1, GA 1		
	27	Spatial Data GIA 2,3, GA 2,4	Exercise 1	

Part II: Point Patterns

Month	Date	Topic/Readings	Out	Due
Feb	3	Point Pattern Basics GIA 4	Project	
	10	Point Pattern Processes GIA 5	Exercise 2	Exercise 1
	17	Quadrat and Distance Based Methods GA p 211-221		
	24	Advanced Point Patterns ASD 7	Data Proposals	Exercise 2

Part III: Lattice Data

Month	Date	Topic/Readings	Out	Due
Mar	3	Lattice Data Basics GIA 7	Exercise 3	
	10	Spatial Weights Anselin and Rey (2010)		Data Approval
	17	Spring Break		
	24	Global Autocorrelation GA p 222-234	Exercise 4	Exercise 3
	31	Local Autocorrelation GIA p 203-208, GA p 234-238		

Part IV: Geostatistical Data

Month	Date	Topic/Readings	Out	Due
Apr	7	Geostatistics Basics GIA 8		Exercise 4
	14	Variography GA 6, ASD p 191-208	Exercise 5	
	21	Kriging GIA 9, ASD p 208-226		
	28	Future Directions		Exercise 5
May	5			Projects
	12			Presentations

Grading

Grading will be assigned on the following scale: A: 90-100, B: 80-89, C:70-79, D: 60-69, F below 60. There will be no curves and no extra credit. I will assign +/- on an individual basis. Points are assigned as follows:

	Points	
Component	Undergraduate	Graduate
Exercises	40	30
Project	40	40
Quizzes	20	20
Presentations		10
TOTAL	100	100

Grading 483 vs. 598

Undergraduate

For your project you can either:

- select your own data
- use data I give you

Graduate

- you must select your own data for your project
- present article synopsis
- present final project

Prerequisites

All participants are expected to have working knowledge of spatial analysis concepts and to be familiar with multivariate statistics. No extensive GIS background beyond ArcGIS basics is needed.

Course Organization

The course will meet in CoorL 1-18. The first part of each weekly meeting will focus on background and theoretical material related to the particular type of spatial data analysis. Following a short break, the attention will shift to introduce various software packages for the analysis of spatial data.

Ground Rules

Classroom etiquette

To ensure a productive learning environment, all participants are expected to abide by the following rules:

- Because we are meeting in a computer laboratory food is strictly forbidden in the class meetings.
- Use of the classroom computers should only be in support of the course. Extracurricular use (i.e., browsing non-course materials, using social networks, checking email, chat sessions, etc) during class meetings is disruptive for your colleagues and disrespectful. Individuals violating this rule will be asked to leave for that course session.

Readings

Textbooks

Assigned readings are listed in the schedule and are to be done prior to class meeting. The majority of these are taken from the following texts:

GIA David O'Sullivan and David J. Unwin (2006) *Geographic Information Analysis*, Wiley: Hoboken

GA Michael de Smith, Michael F. Goodchild, and Paul A. Longley (2007) *Geospatial Analysis*, Winchelsea Press: Leicester

ASD Roger S. Bivand, Edzer J. Pebesma and Virgilio Gómez-Rubio (2008) *Applied Spatial Data Analysis with R*, Springer: New York.

Additional readings will be assigned by topic.

Software

GeoDa

`http://geodacenter.asu.edu/software/downloads`

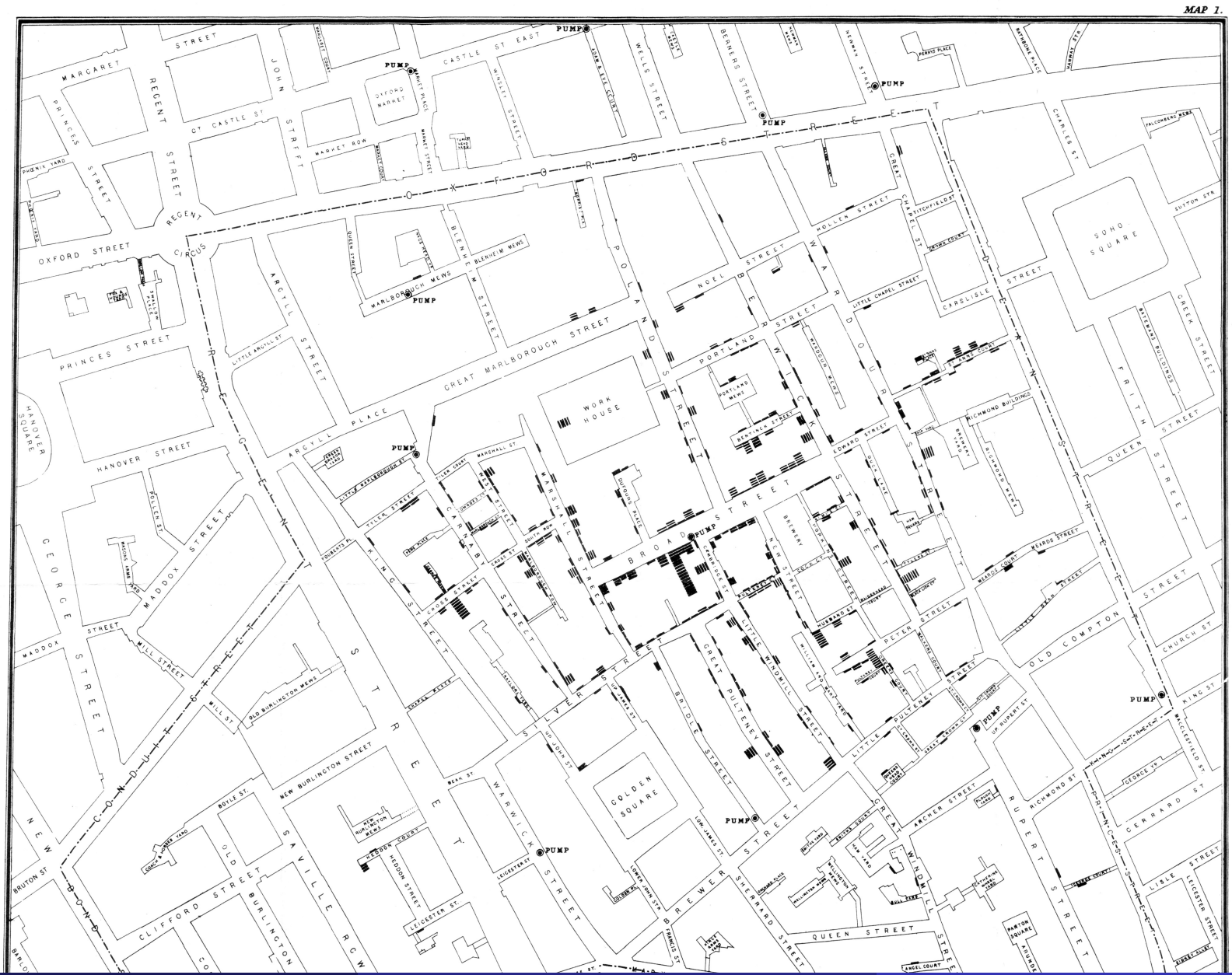
R

`http://cran.r-project.org/`

STARS

`http://regionalanalysislab.org/index.php/Main/STARS`

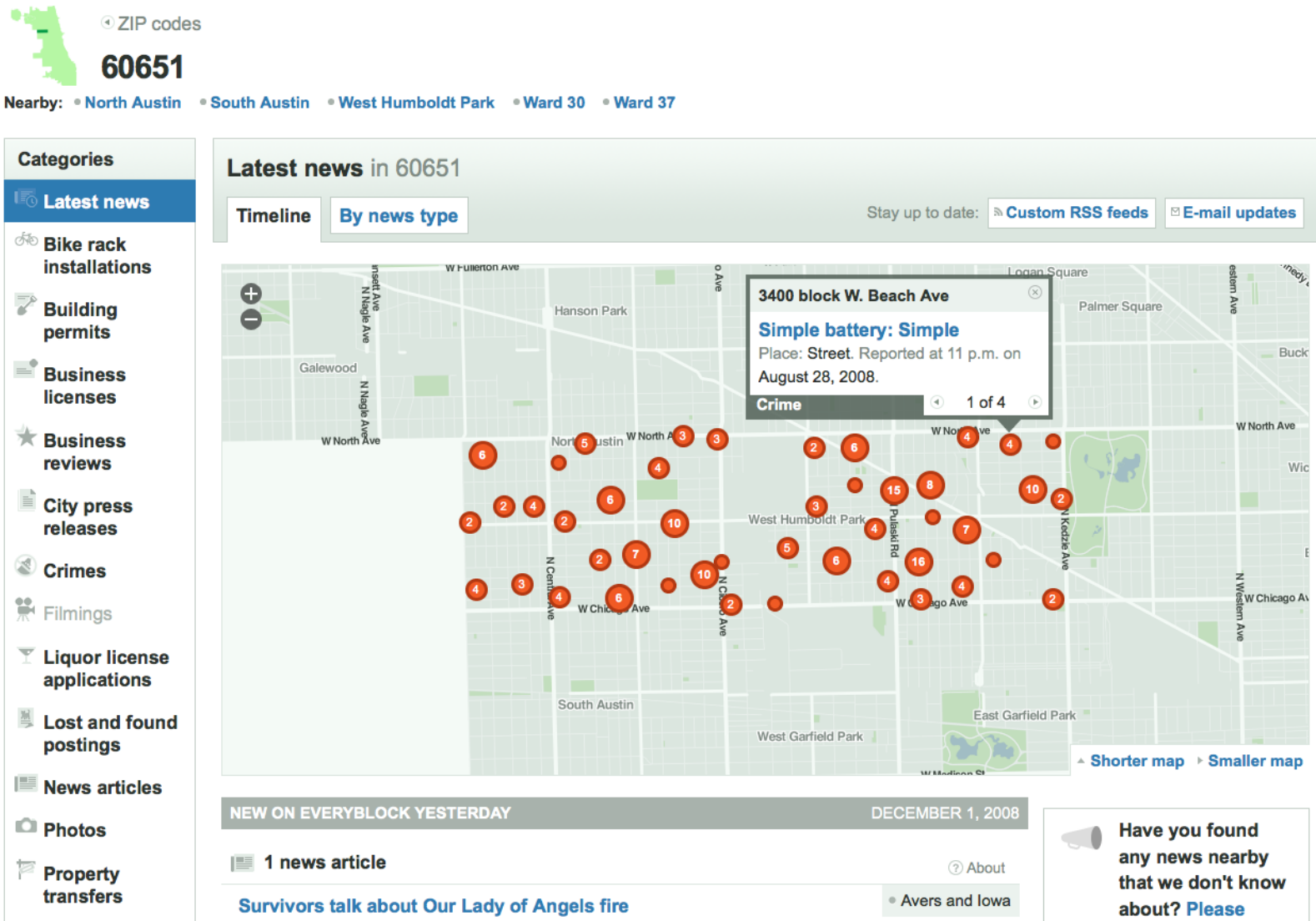
GIS Then


MAP 1.
OXFORD STREET
REGENT STREET
PRINCES STREET
HANOVER STREET
GREAT MARLBOROUGH STREET
ARGYLL PLACE
POLAND STREET
BROAD STREET
GOLDEN SQUARE
SOHO SQUARE
WORK HOUSE
PUMP

GIS Then


WORK HOUSE
BENTINCK STREET
EDWARD ST
MARSHALL STREET
DUFOURS PLACE
BROAD STREET
PUMP
SOUTH ROW
HARLBOROUGH ROW
CAMBRIDGE ST
BREWERY
NEW STREET
HOPKINS STREET
COCK CT
FULTENEY
HUSBAND ST
LITTLE
WILLIAM

GIS Now


ZIP codes
60651
Nearby: North Austin South Austin West Humboldt Park Ward 30 Ward 37
Categories
Latest news
Bike rack installations
Building permits
Business licenses
Business reviews
City press releases
Crimes
Filmings
Liquor license applications
Lost and found postings
News articles
Photos
Property transfers
Latest news in 60651
Timeline
By news type
Stay up to date:
Custom RSS feeds
E-mail updates
3400 block W. Beach Ave
Simple battery: Simple
Place: Street. Reported at 11 p.m. on August 28, 2008.
Crime
1 of 4
Hanson Park
Galewood
Logan Square
Palmer Square
West Humboldt Park
South Austin
East Garfield Park
West Garfield Park
W North Ave
Shorter map
Smaller map
NEW ON EVERYBLOCK YESTERDAY
DECEMBER 1, 2008
1 news article
About
Survivors talk about Our Lady of Angels fire
Avers and Iowa
Have you found any news nearby that we don't know about? Please

GIS Functions

Anselin-Getis (1992) Taxonomy

- Input
- Storage
- Analysis
- Output

Many other taxonomies

GIScience

Goodchild (1992)

- cross-disciplinary
- central role for spatial analysis
- scientific glue

What is Spatial Analysis?

From Data to Information

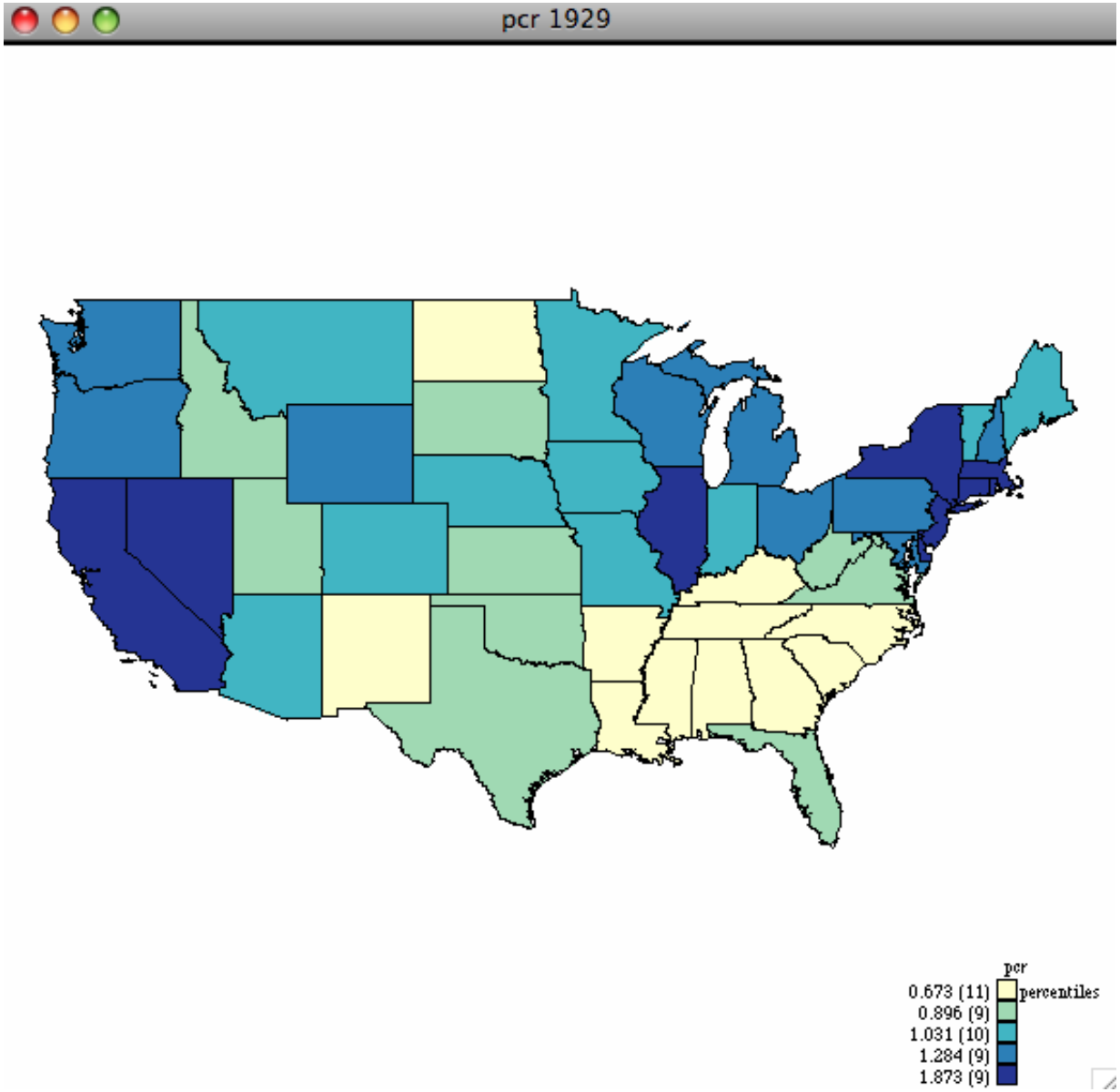
- Beyond mapping
- added value
- transformations, manipulations and application of analytical methods to spatial (geographic data)

Locational Invariance

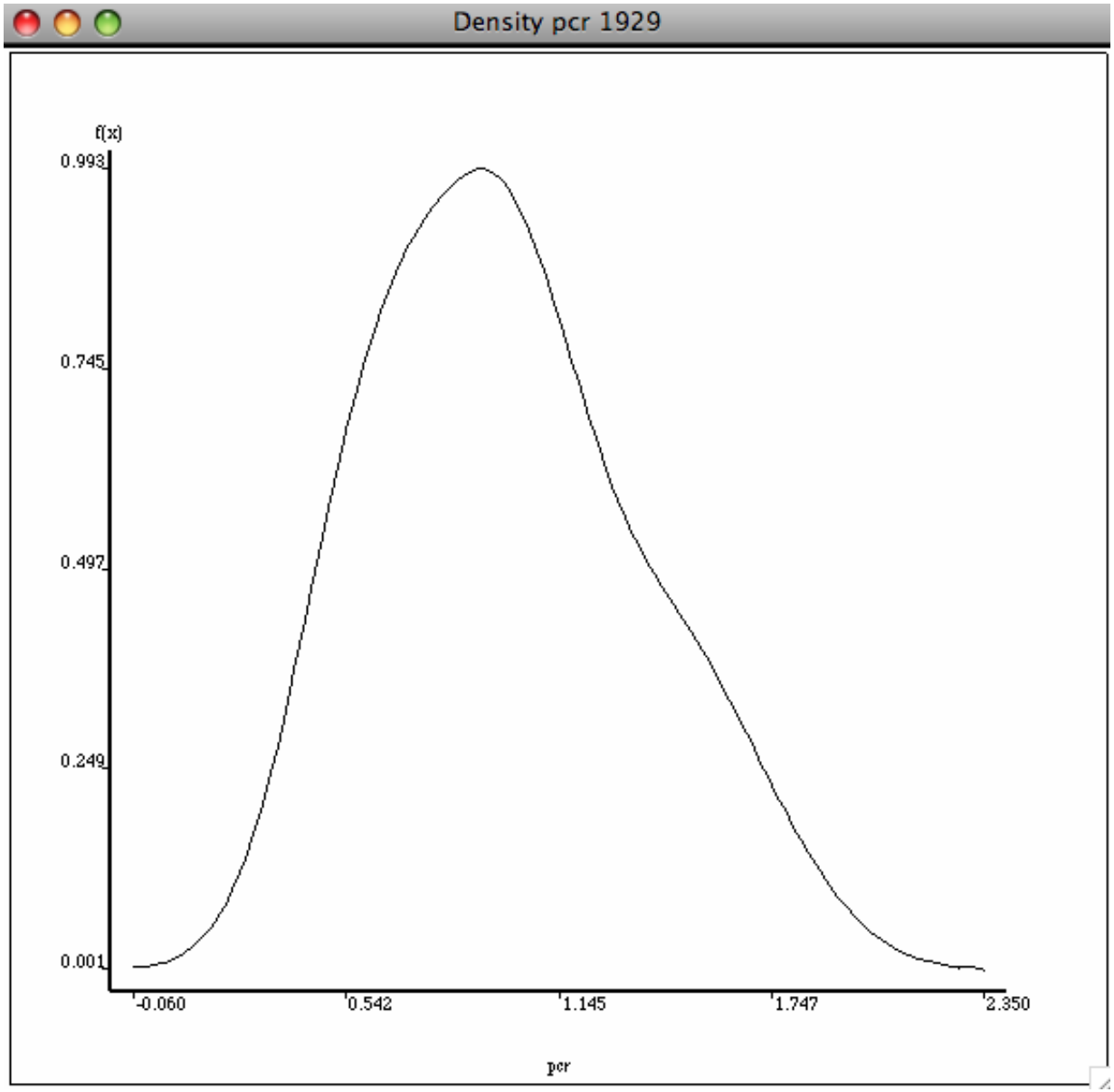
How Insights Change with location

- spatial analysis is not locationally invariant
- the results change when the locations of the study objects change
- where matters

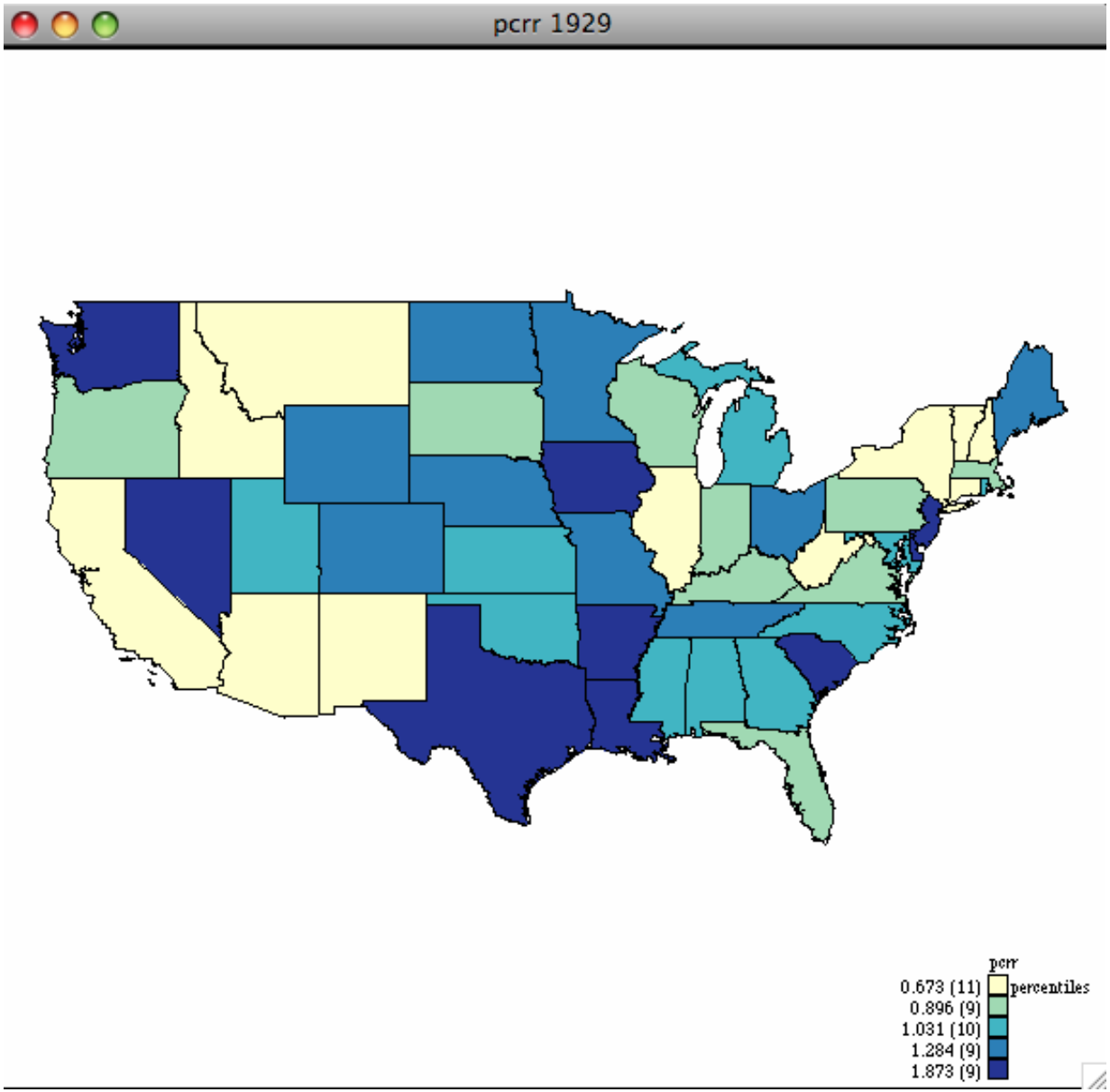
State Income Distributions 1929



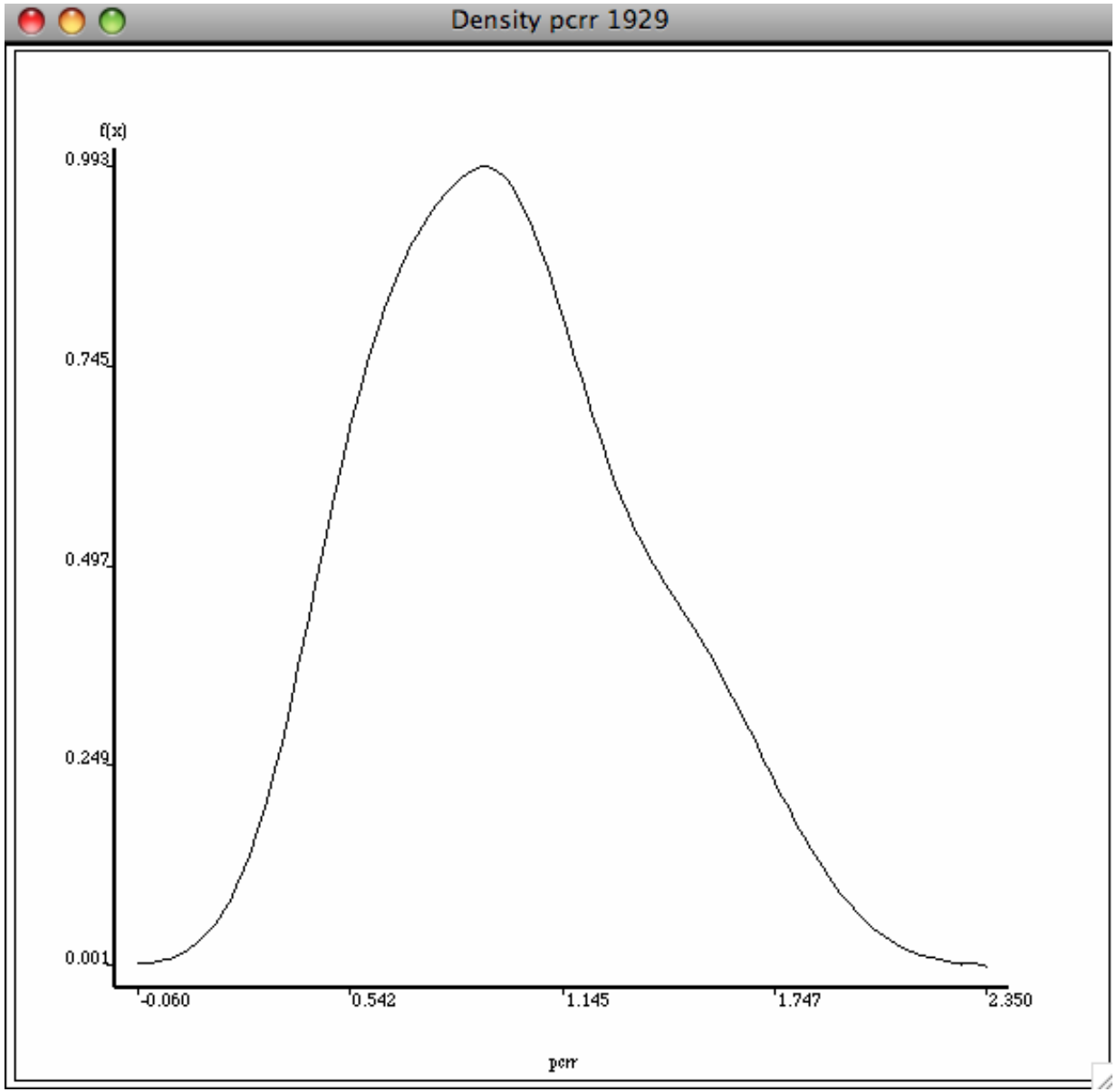
State Income Distributions 1929



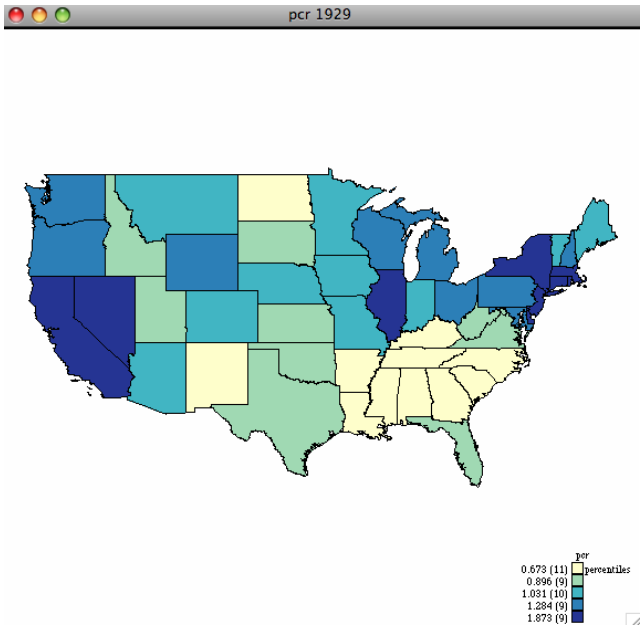
Randomized Income Distribution 1929

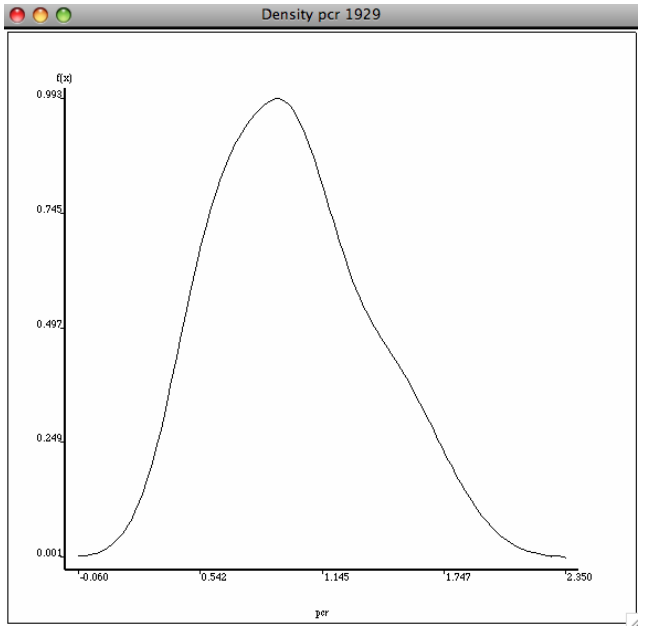


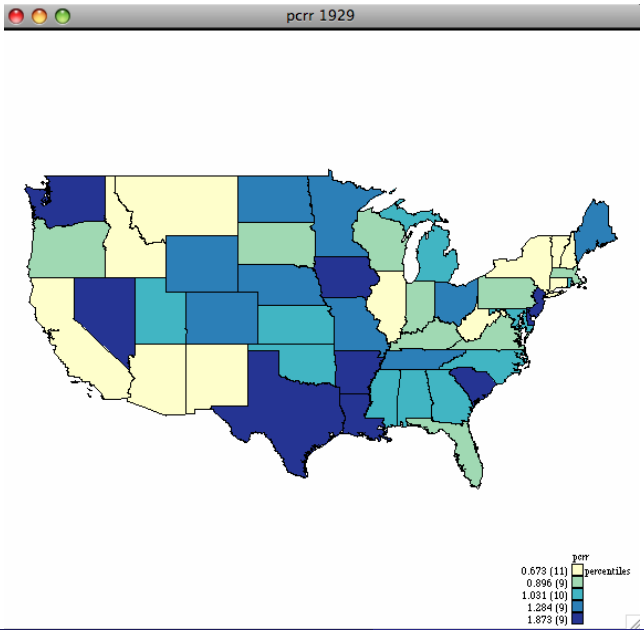
Randomized Income Density 1929

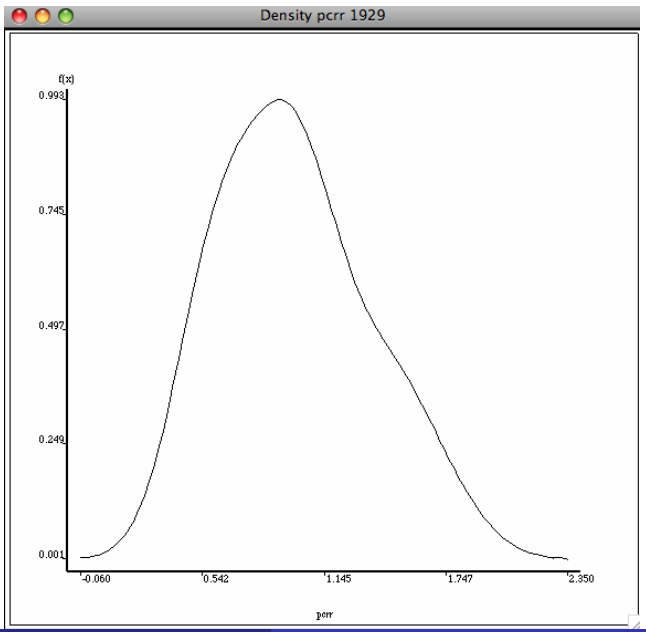


Locational Invariance

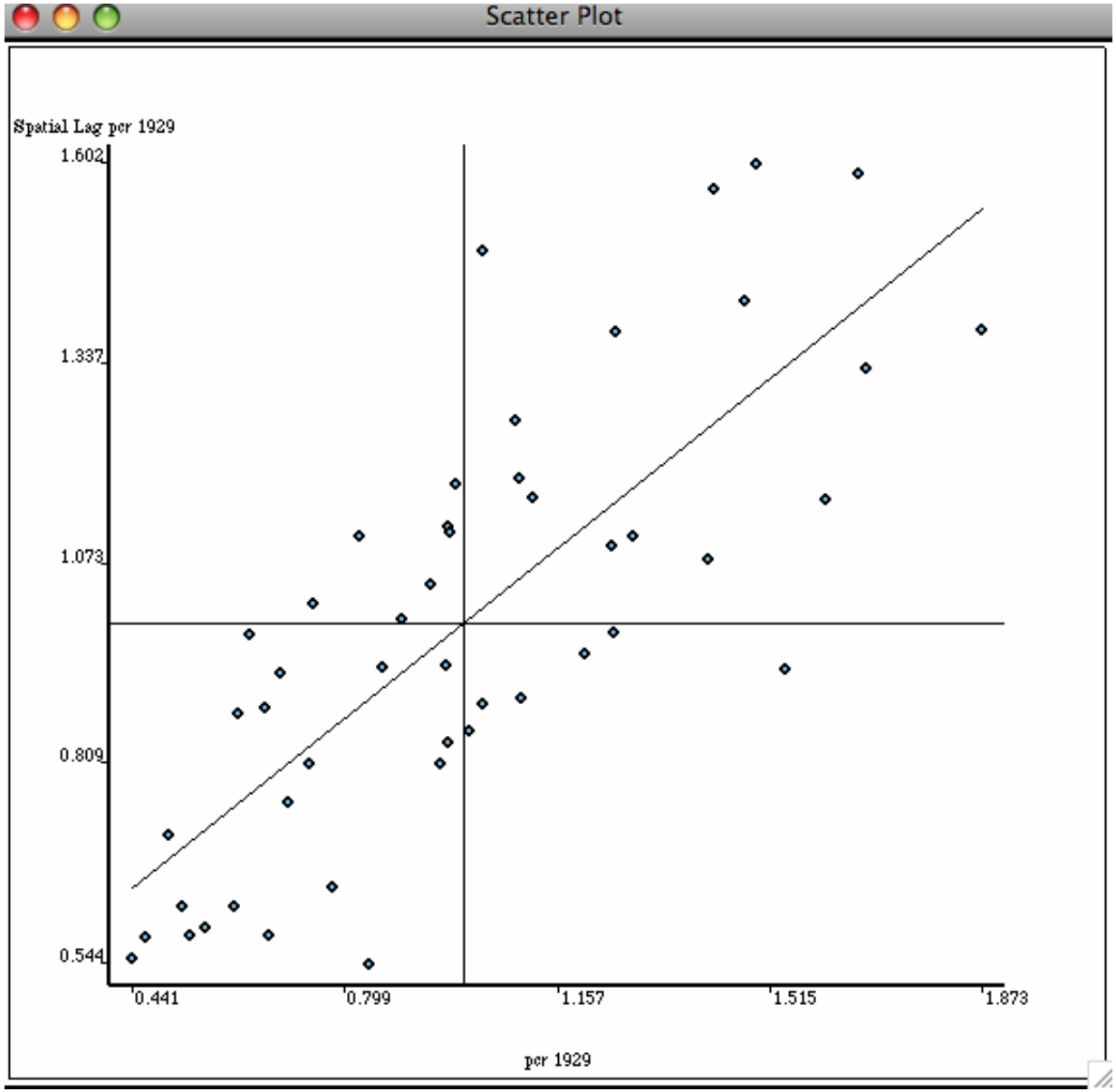




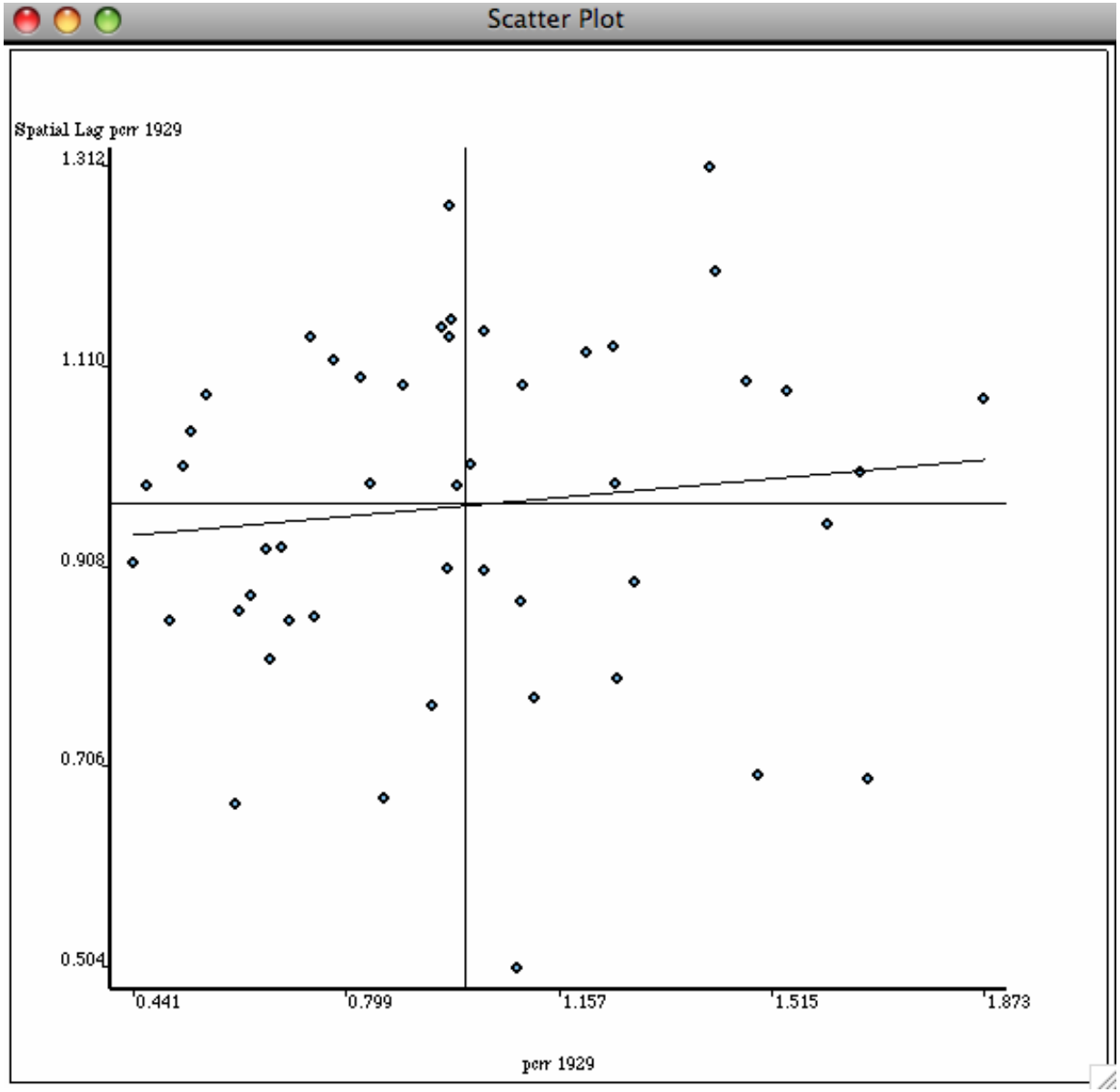




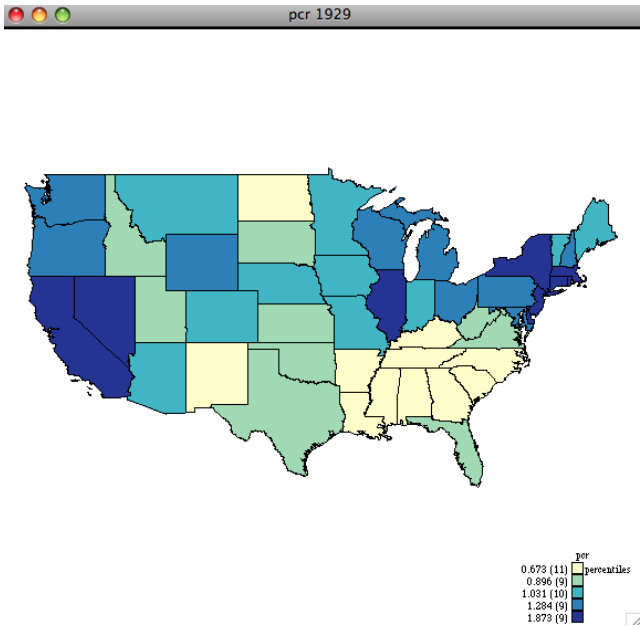
Spatial Autocorrelation Income 1929

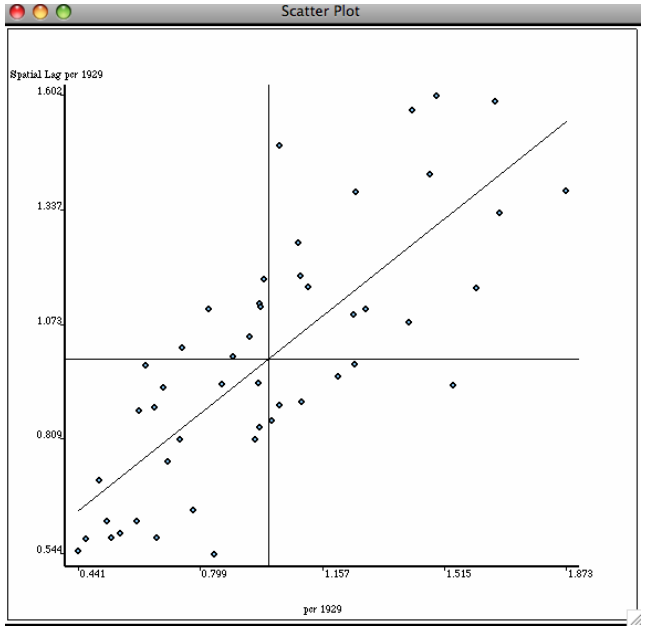


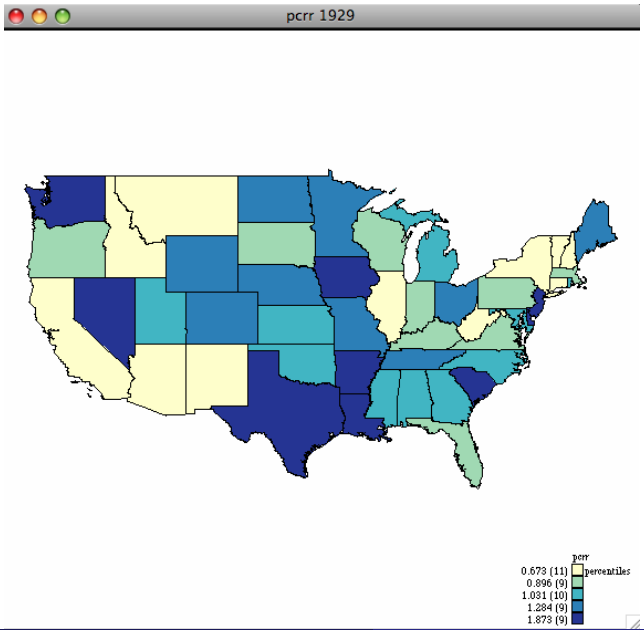
Spatial Autocorrelation Randomized Income 1929

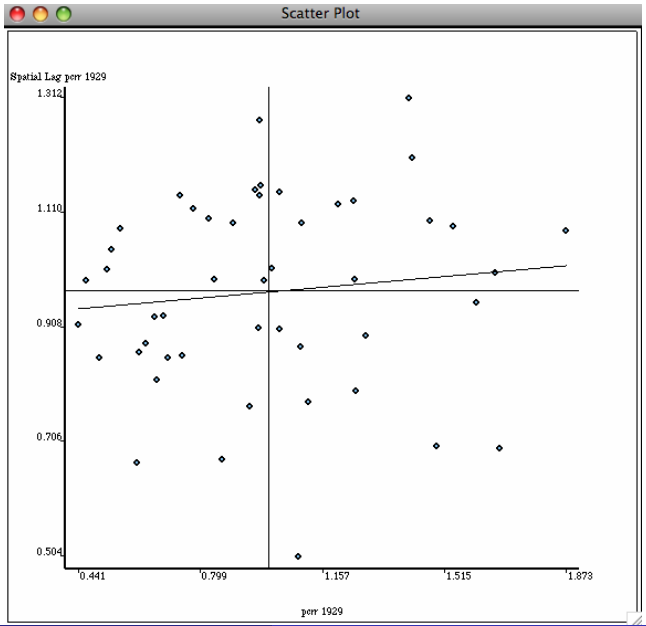


Locational Invariance









Components of Spatial Analysis

Mapping and Geovisualization

showing interesting patterns

Exploratory Spatial Data Analysis

discovering interesting patterns

Spatial Modeling

explaining interesting patterns

Summary: Spatial Analysis

Beyond Mapping

Central role for analysis

Distinguished by Locational Variance

Location matters

Components

Showing, discovering, explaining

What is EDA?

Philosophy

EDA is an approach, not simply a set of techniques, but an attitude/philosophy about how a data analysis should be carried out. Postpones the usual assumptions about what kind of model the data follow

Origins

Tukey, J. (1977) *Exploratory Data Analysis*. Addison, Wesely

Components

Set of techniques to

- maximize insight into a data set
- uncover underlying structures
- extract important variables
- detect outliers and anonalies
- test underlying assumptions
- suggest hypotheses
- develop parsimonious models

EDA Techniqes

Statistical Graphics

- EDA relies heavily on statistical graphics
- EDA is not identical to statistical graphics
- Graphics support pattern recognition and open-minded exploration
- Interactive graphcs push this even further

Quantitiatve Methods

Although heavily graphic in orientation, there are also a number of numerical techniques in EDA.

EDA Versus Confirmatory Analysis

Confirmatory Analysis (e.g. regression)

Problem $\longrightarrow$ Theory $\longrightarrow$ Model $\longrightarrow$ Data $\longrightarrow$ Conclusion

Exploratory Analysis

Problem $\longrightarrow$ Data $\longrightarrow$ Analysis $\longrightarrow$ Model

What is ESDA?

Definitions

- Type of EDA
- Extended to include spatial attributes of the data

Crossfertilization

- Applying classic EDA to spatial data
- Developing new EDA methods for spatial data
- Interactions between EDA and ESDA

How does ESDA fit in spatial analysis?

Spatial Modeling?

- Modeling based on assumptionss
- ESDA largely model free
- Matter of degree (e.g., clustering)

Mapping?

- Maps play a critical role in ESDA
- Does a map = ESDA?
- No. ESDA = map, manipulation + visualization

Geovisualization

Beyond Mapping

- Combing map and scientific visualization methods
- Exploit human pattern recognition capabilities

Statistical Maps

- innovative map devices

Mapping Issues

How to Lie with Maps

- Monmonnier (1996)
- many design issues
- projects
- human perception can be tricked

Visual Analytics

The Science of Analytical Reasoning Facilitated by Interactive Visual Interfaces

- NVAC 2005
- science of analytical reasoning
- visual representation and interaction
- data representation and transformations
- production, presentation and dissemination

Visual Analysis

Tools

- synthesize inforation
- derive insights
- detect the expected and discover the unexpected
- understandable assessments
- communicate effectively
- focused on policy actions

Visual Explanations

Tufte (1997)

Reasoning about Evidence and Design of Graphics

- documenting sources (metadata)
- appropriate comparisons
- quantify and show cause and effect
- multivariate nature of analytic problems
- evaluate alternative explanations

Choropleth Map

Map Counterpart of Histogram

- values for discrete spatial uits
- choro from choros (region) NOT chloro

Discrete Approximations

- intervals
- continuous shading

Map Design Issues

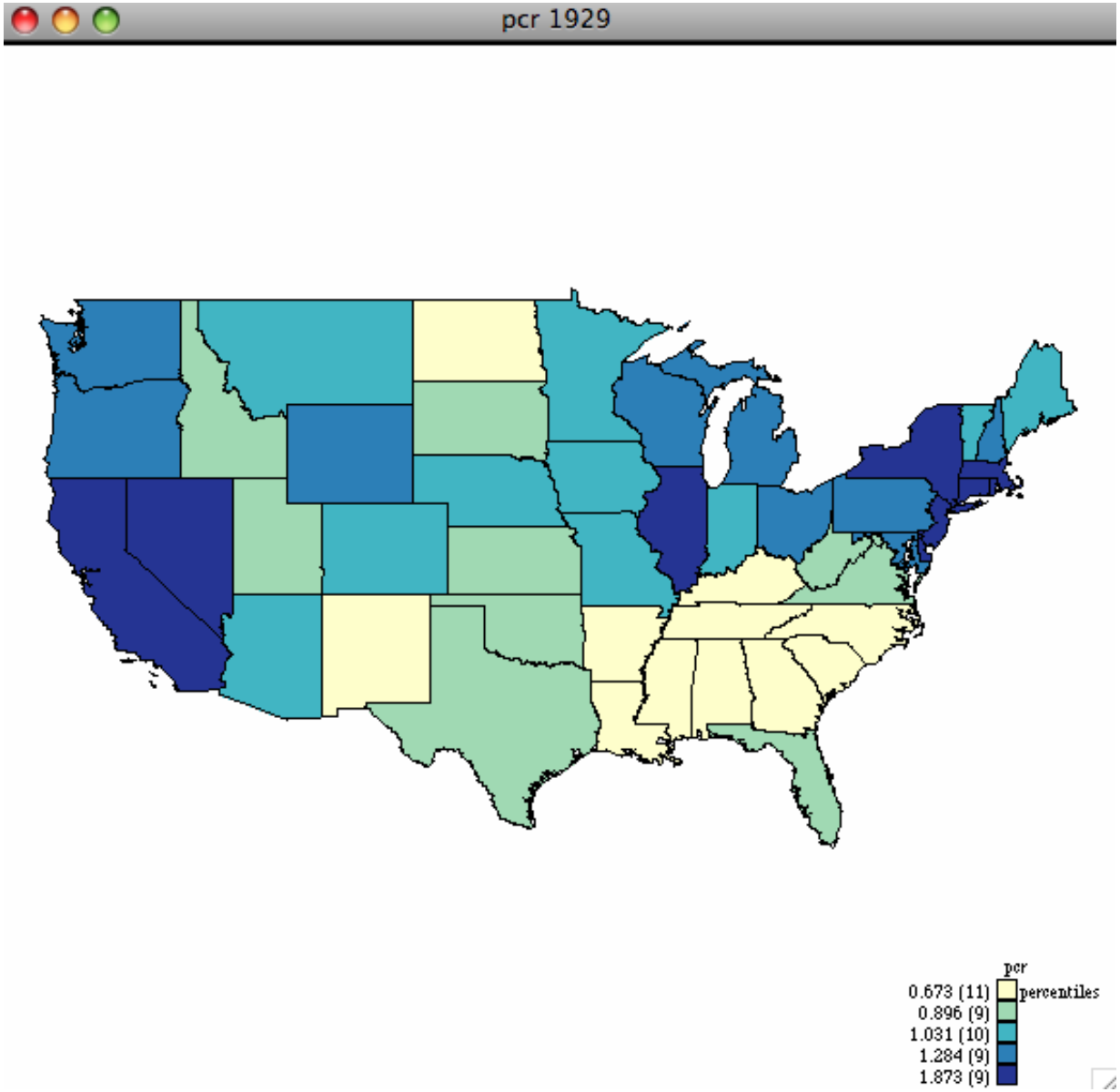
Choice of Intervals

- cut points: equal interval, natural breaks
- statistical criteria: equal area (quantile)

Choice of Colors

- important for perception of pattern

Income Quintiles

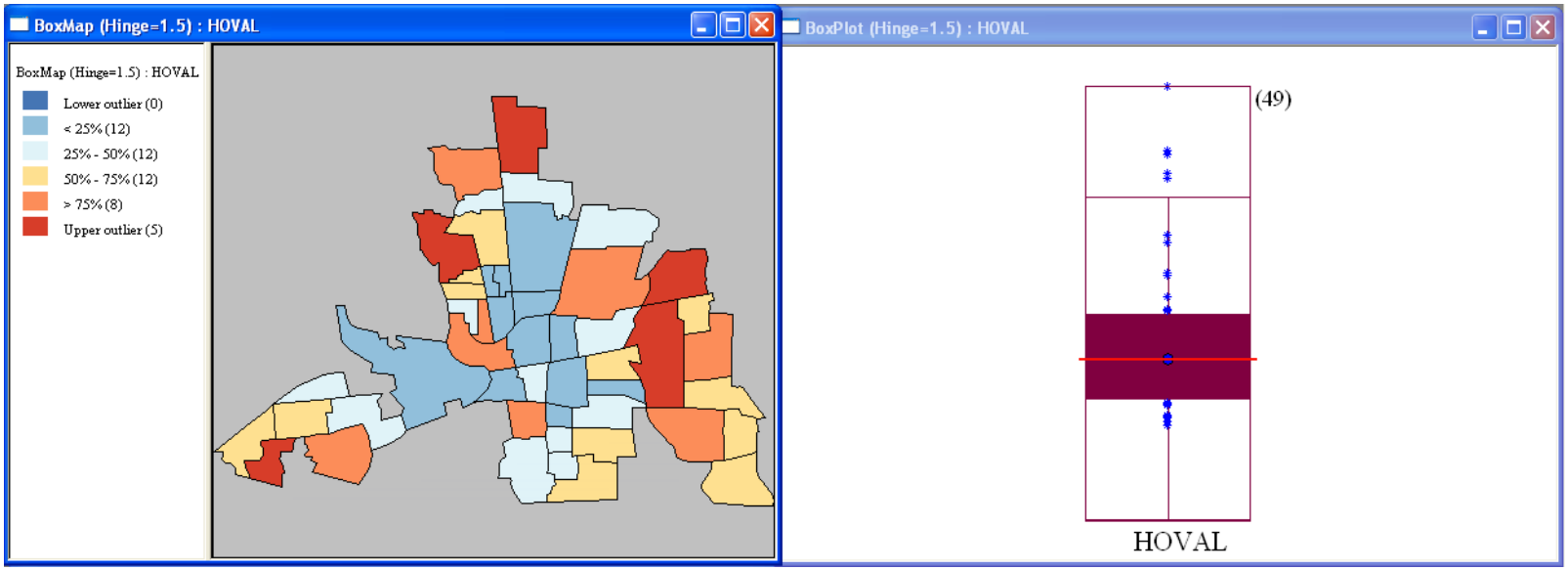


Outlier Map

Box Map

- Special Quartile Map
- Outliers Highighlited
  - same criteria as a box plot
  - outliers added as extra categories
  - six instead of four categories
- Both Magnitude and Location

Box Map


BoxMap (Hinge=1.5) : HOVAL
BoxMap (Hinge=1.5) : HOVAL
Lower outlier (0)
< 25% (12)
25% - 50% (12)
50% - 75% (12)
> 75% (8)
Upper outlier (5)
BoxPlot (Hinge=1.5) : HOVAL
(49)
HOVAL

Special Maps

- Cartogram
- Conditional Maps
- Map Animation

Cartogram

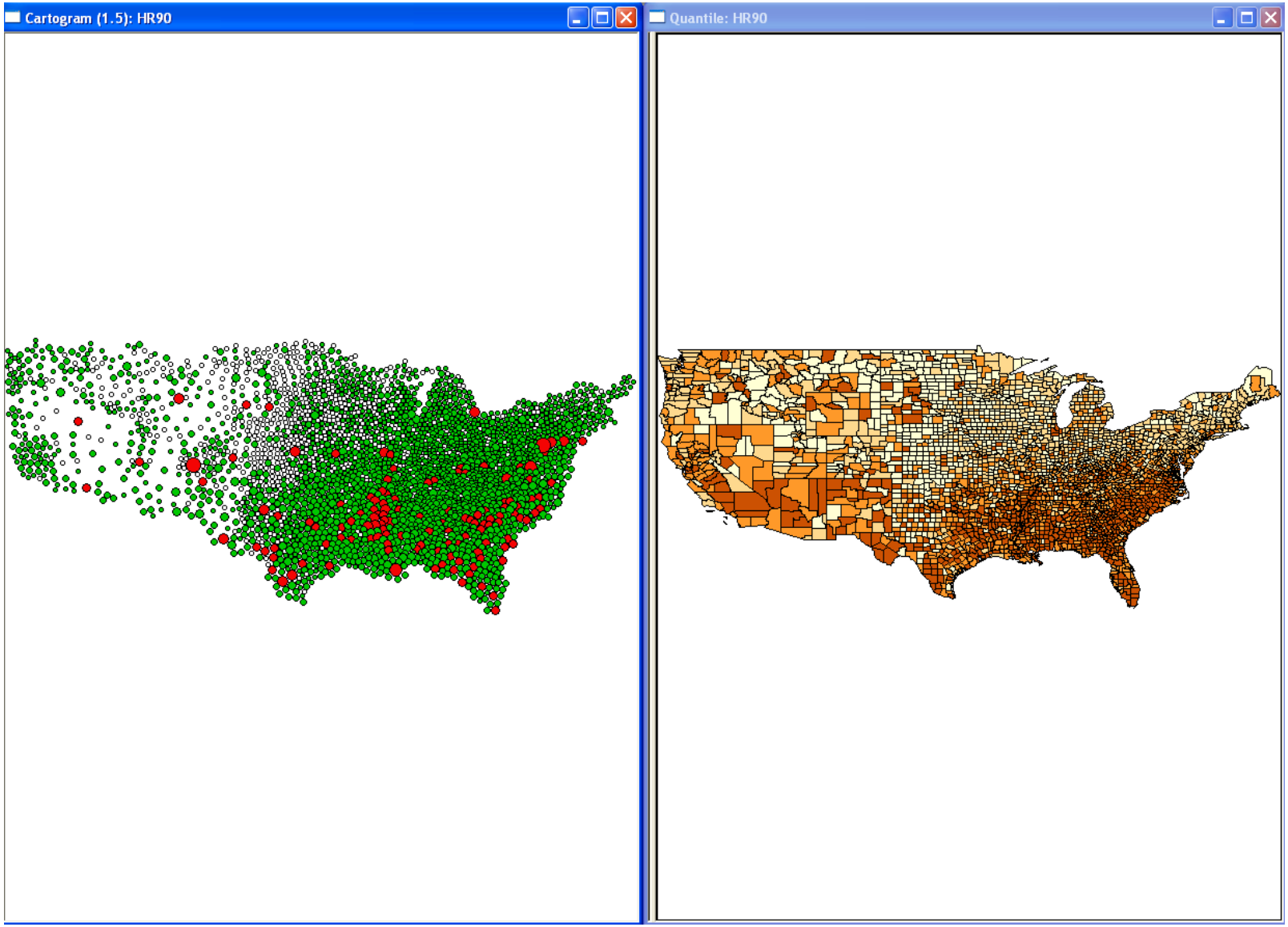
Objectives

- Correct for misleading effect of area
  - larger area units draw attention
  - change layout to reflect size other than area
- Respect topology

Circular Cartogram

- variable maps to area/radius of circle

Cartogram



Conditional Maps

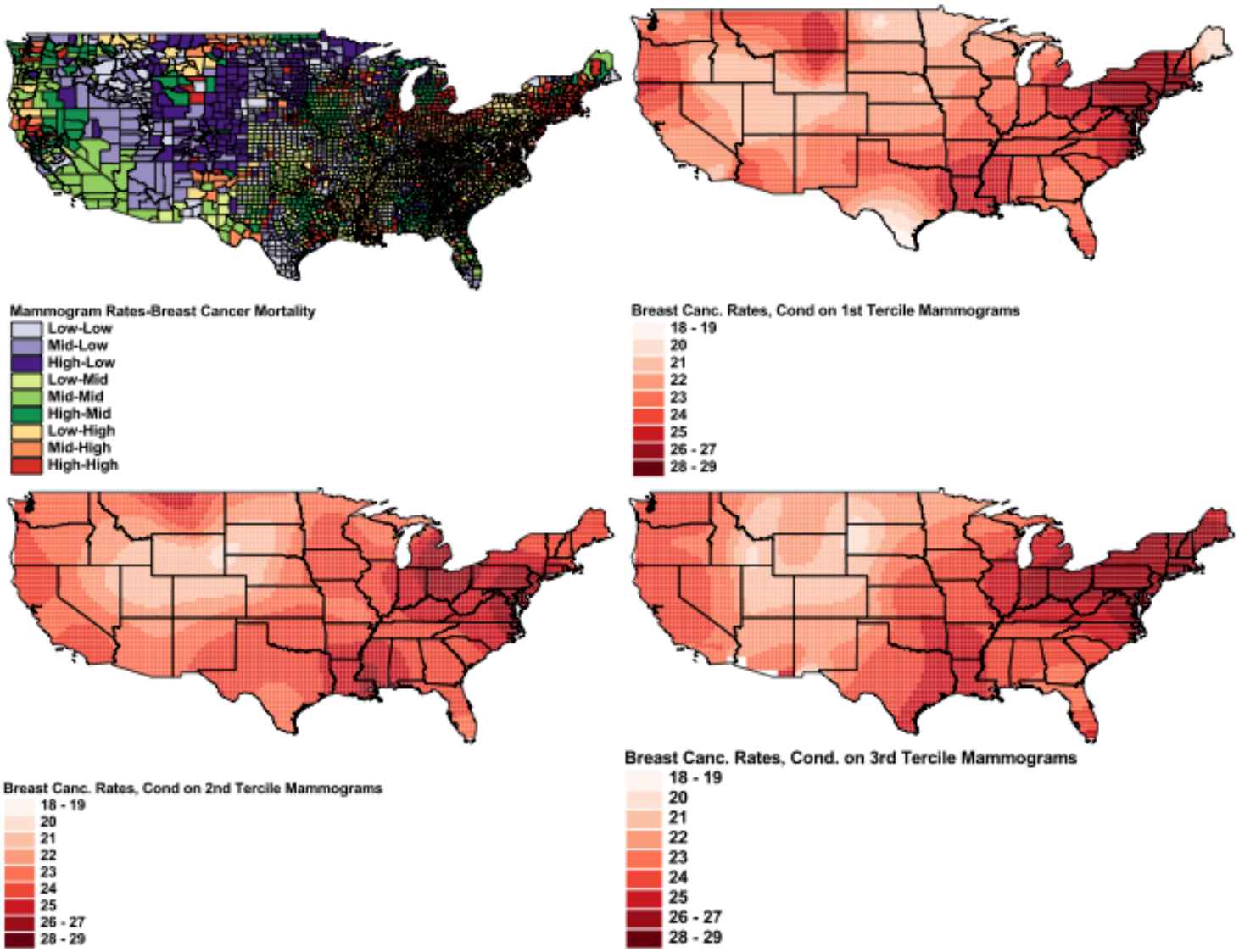
Conditional Choropleth Map (Carr)

- Special case of conditional plots
- trellis graphs

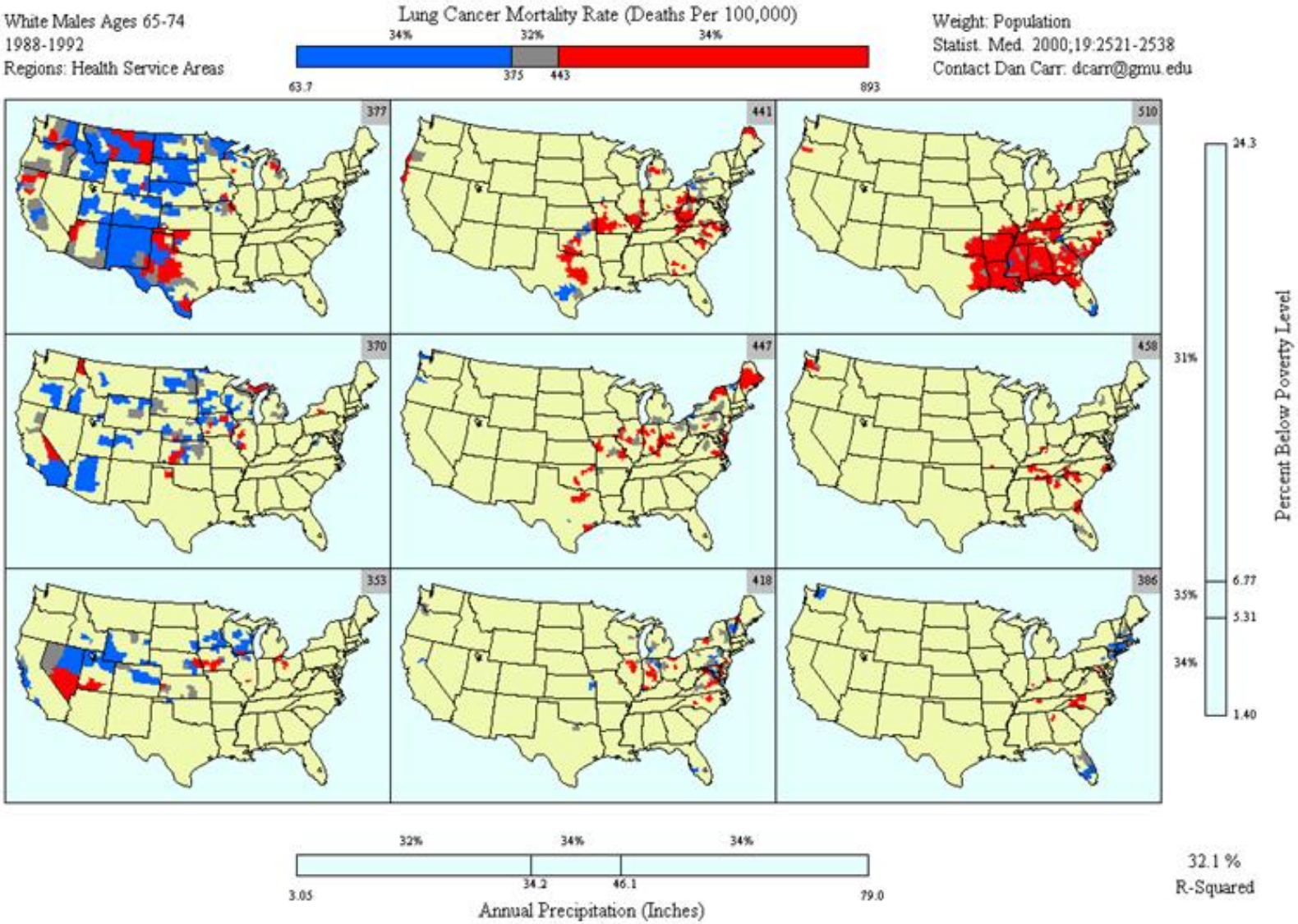
Conditioning

- along two dimensions (variables)
- micromap matrix
- choropleth map on dependent variable

Conditional Choropleth: Univariate Conditioning



Conditional Choropleth: Bivariate Conditioning



Map Animation

Map Movie

- location highlighted in turn
- from low value to high value

Looking for pattern

- spatial heterogeneity
- systematic movements/locations

Map Animation

Demo

Interactive Graphics

Interactive View Manipulation

- the analyst interacts with the data
- dynamic graphics
- no longer passive

Linking and Brushing

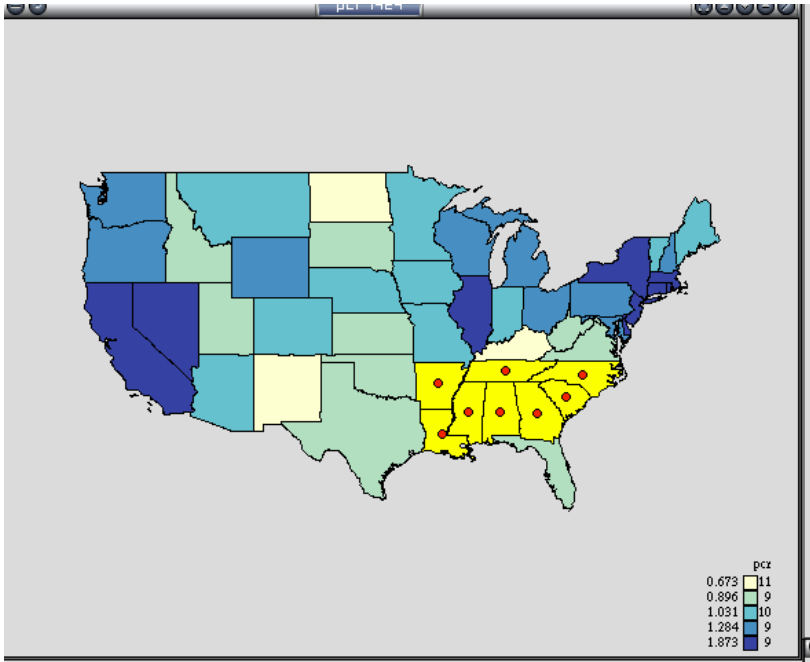
Linking

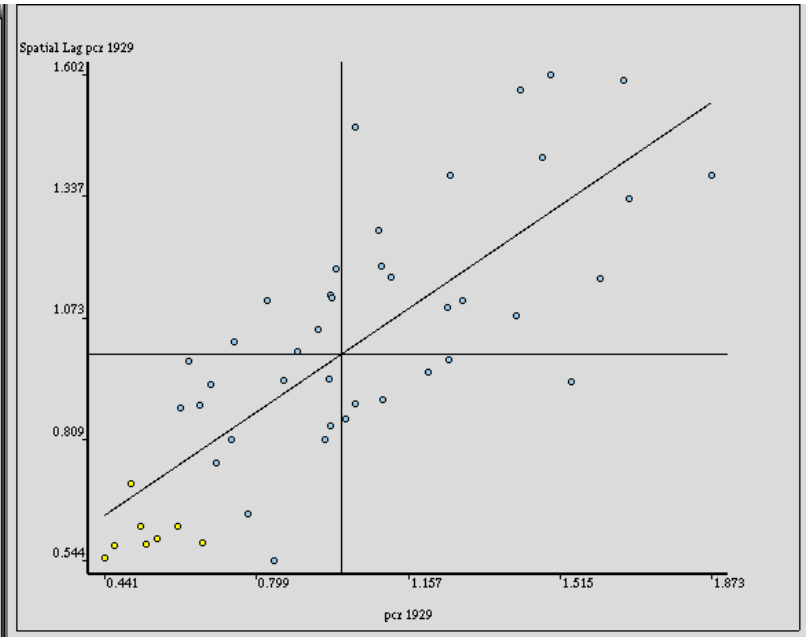
- selection in one graph is simultaneously selected in all graphs

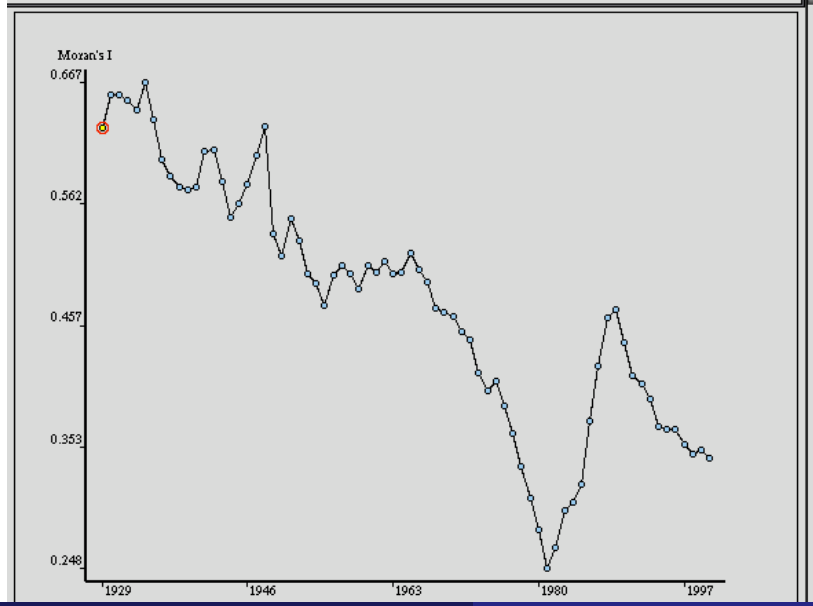
Brushing

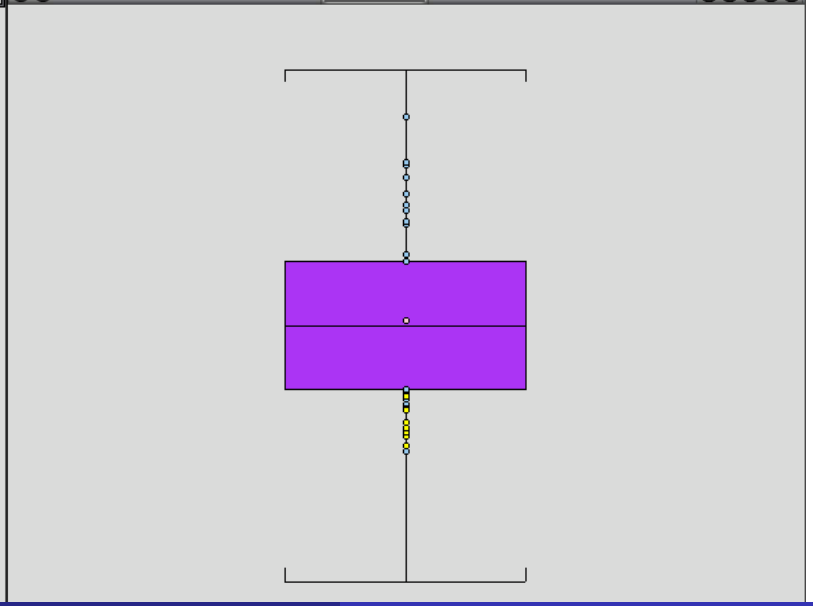
- changing the selection set is dynamically updated in all graphs and maps

Linking

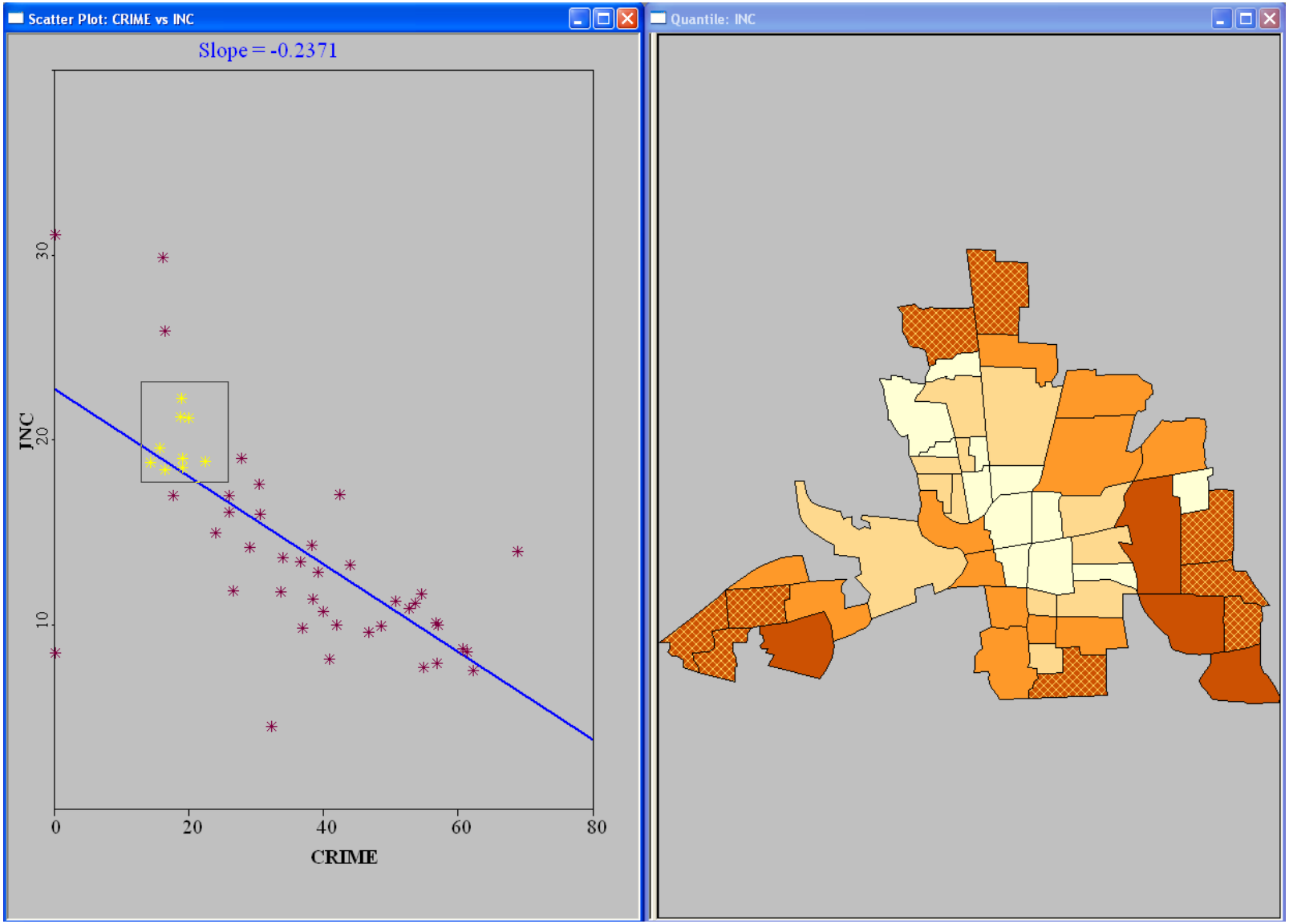

pcr
0.673 11
0.896 9
1.031 10
1.284 9
1.873 9


Spatial Lag pcr 1929
1.602
1.337
1.073
0.809
0.544
0.441
0.799
1.157
1.515
1.873
pcr 1929

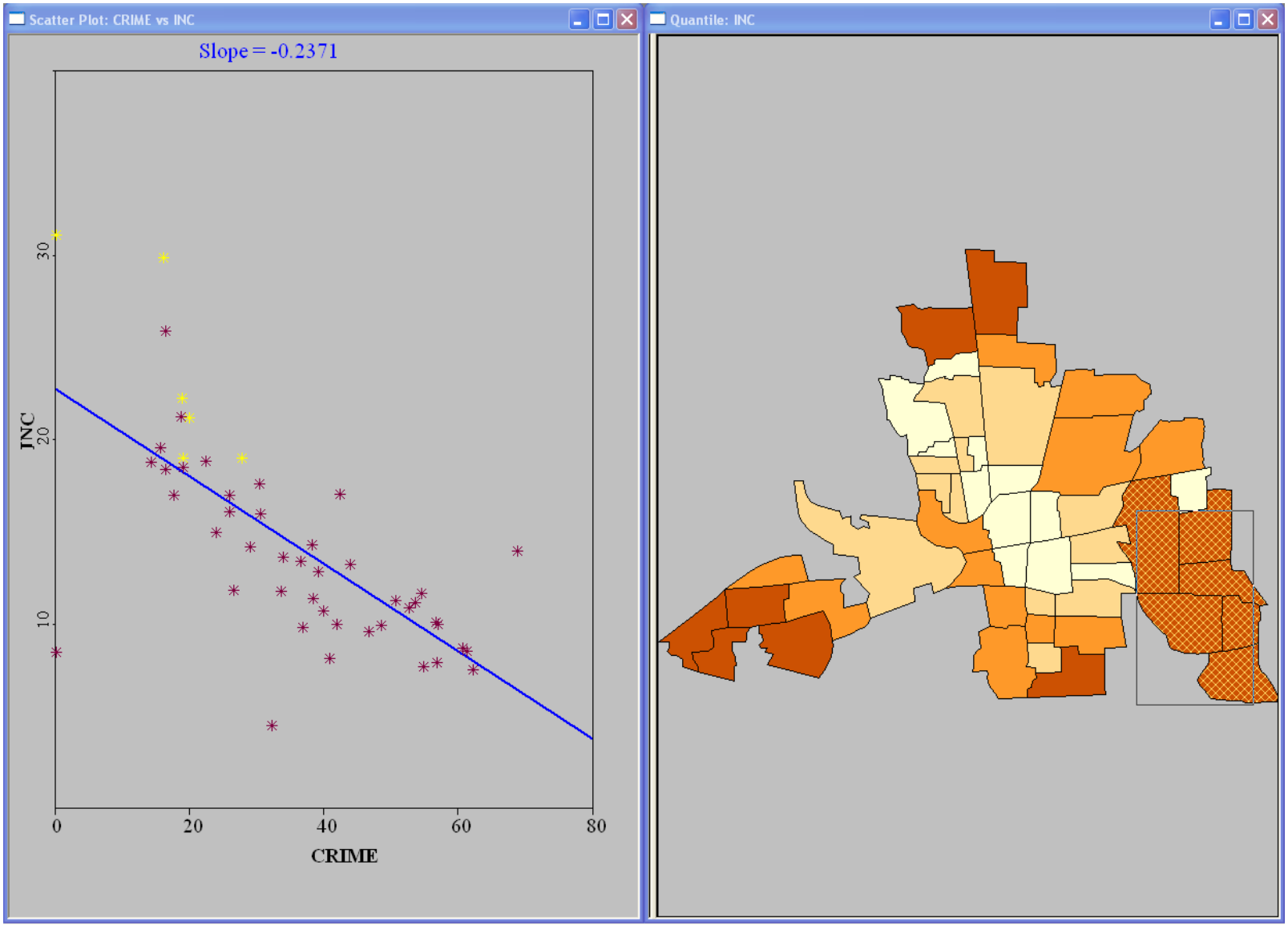

Moran's I
0.667
0.562
0.457
0.353
0.248
1929
1946
1963
1980
1997



Brushing a Scatter Plot



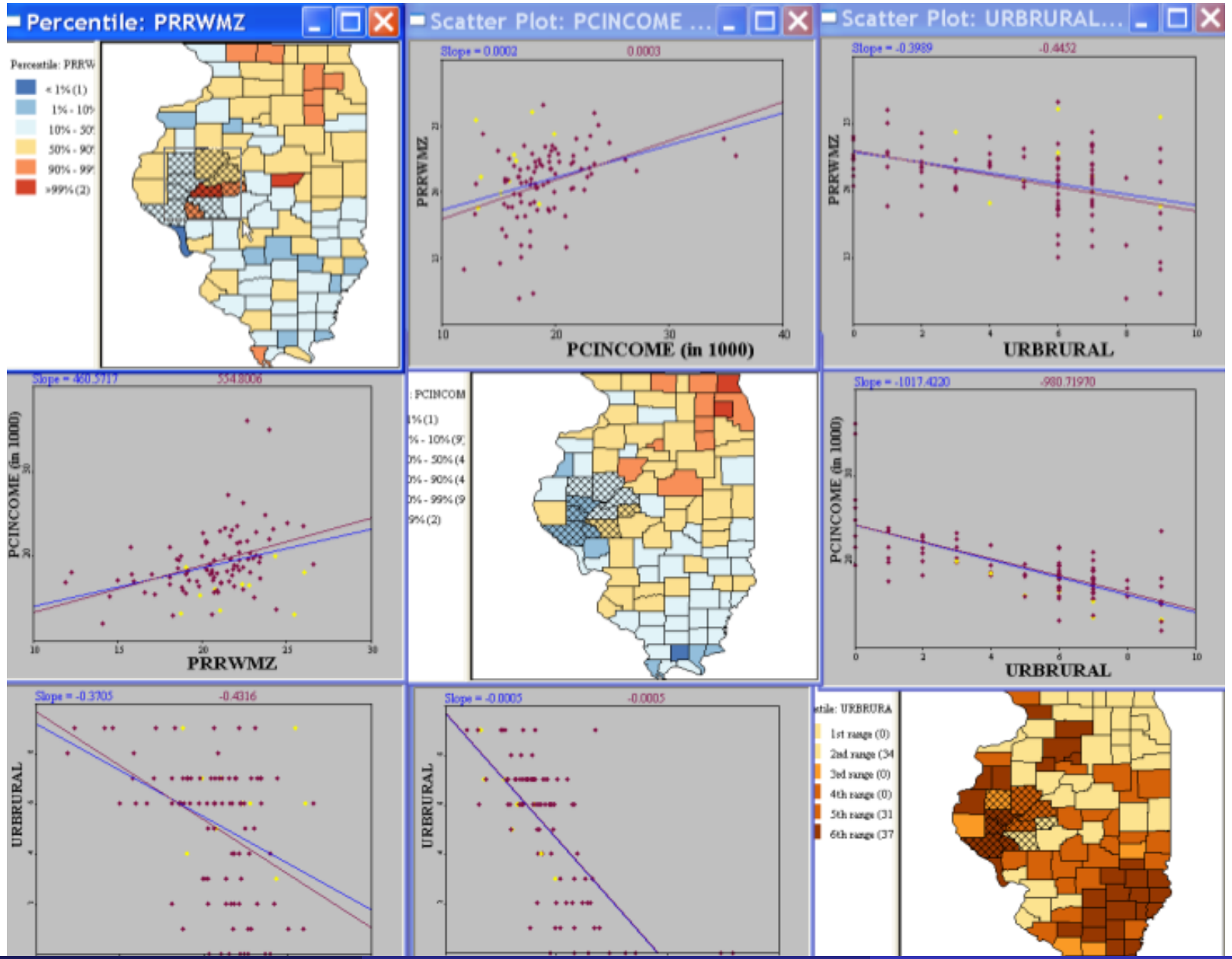
Brushing a Map



Multivariate EDA

- Scatter Plot Matrix
- Parallel Coordinate Plot
- 3-D Scatter Plot

Scatter Plot Matrix



Brushing a Parallel Coordinate Plot

Brushing in 3-D