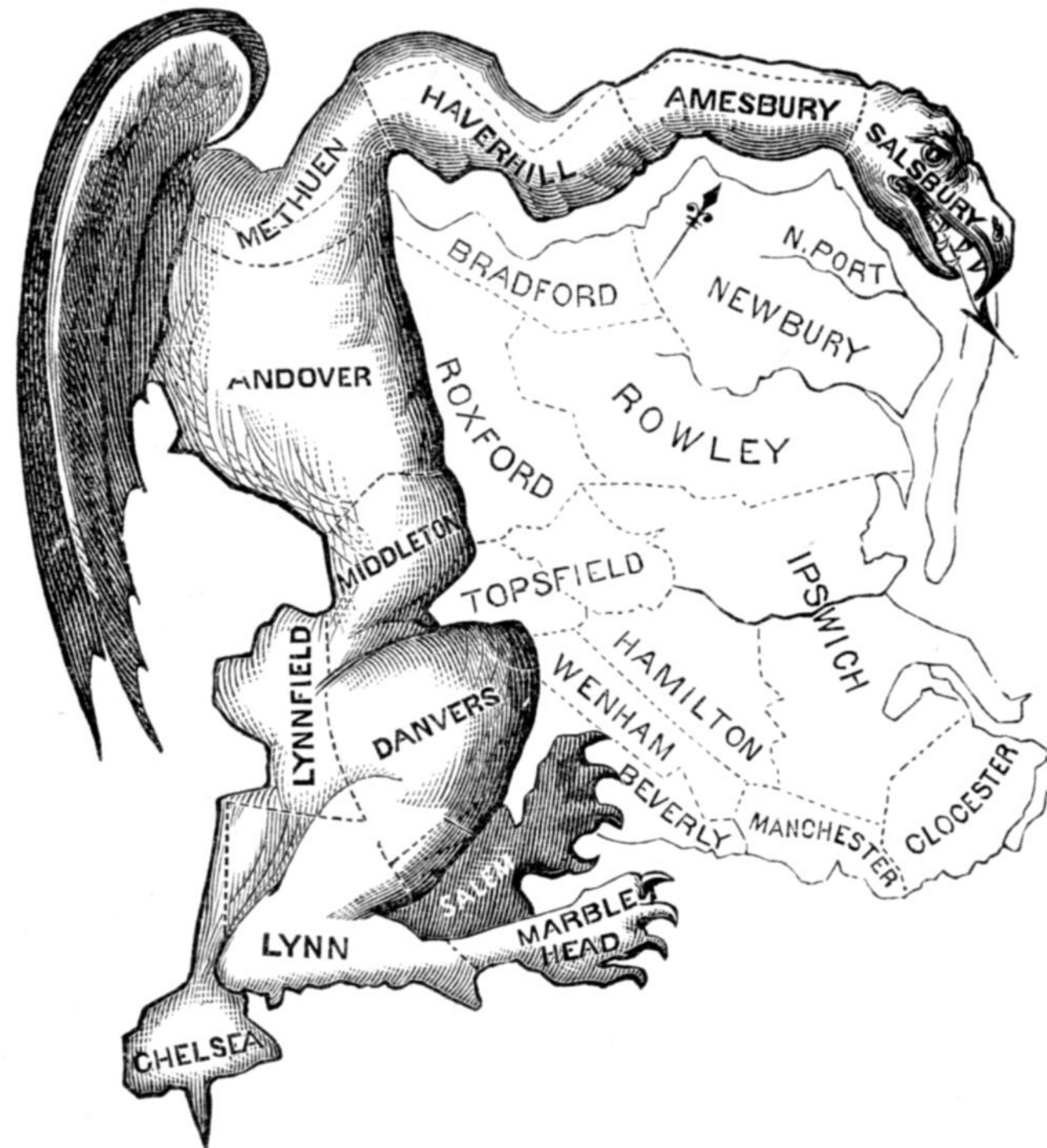


GEO 200CN

Session #3



Is spatial special?

- Complex data: geometry and attributes
- Coordinate reference systems and datum
- Spatial operations: distance, spatial intersection
- Nearby things are more related than distant things
- Distinct statistical methods to analyze data
(or different names for the same idea?)
- Size: lots and lots of it, multivariate, time series
- Specialized software: GIS and Remote sensing
- Special plots: maps

Types of spatial analysis

- **Query and reasoning**

Where is? How much is this here? How to get from A to B?

- **Measurement**

Area, Distance, Length, Slope

- **Transformation**

Buffering, overlay, interpolation

- **Exploration and description**

clusters, trends, spatial dependence, fragmentation

- **Optimization**

Site selection, re-districting, traveling salesman

- **Inference**

Samples from a population, problem of spatial autocorrelation

- **Modeling**

Climate change effects, impact of nuclear accident, dispersal

Representation of space

Object vs Field view

Object

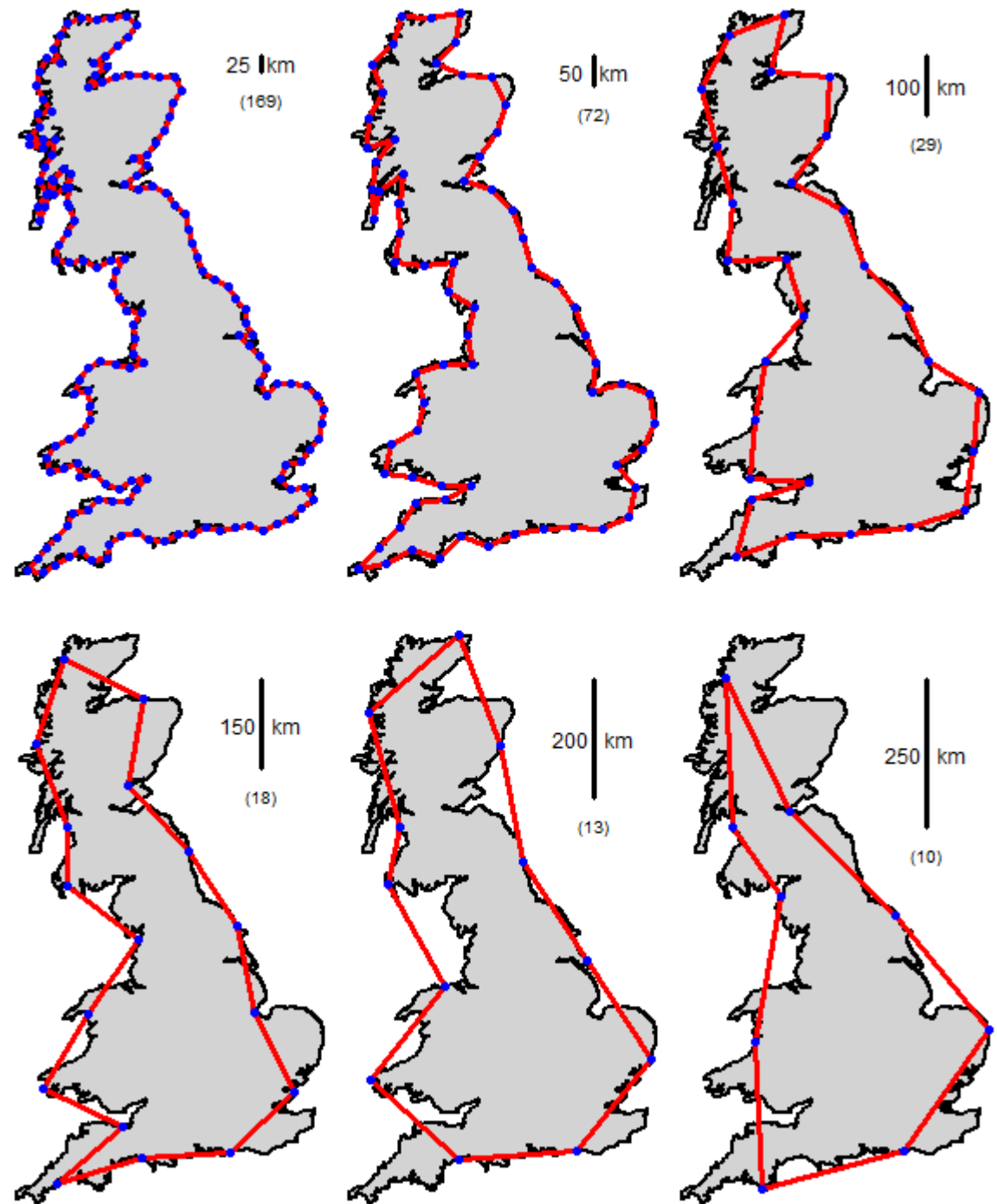
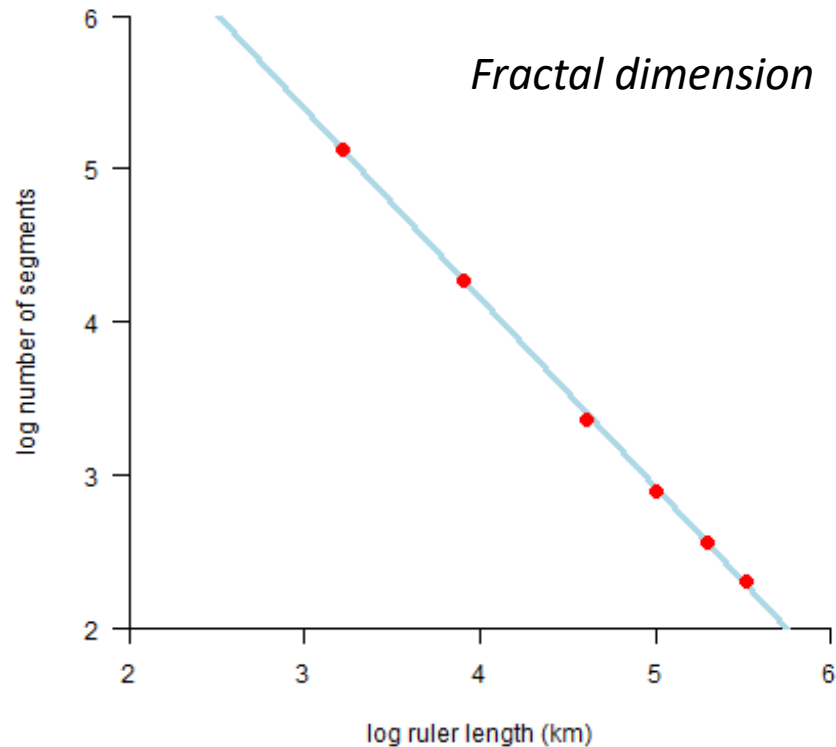
Discrete entities, defined by coordinates
points, lines, areas

Field

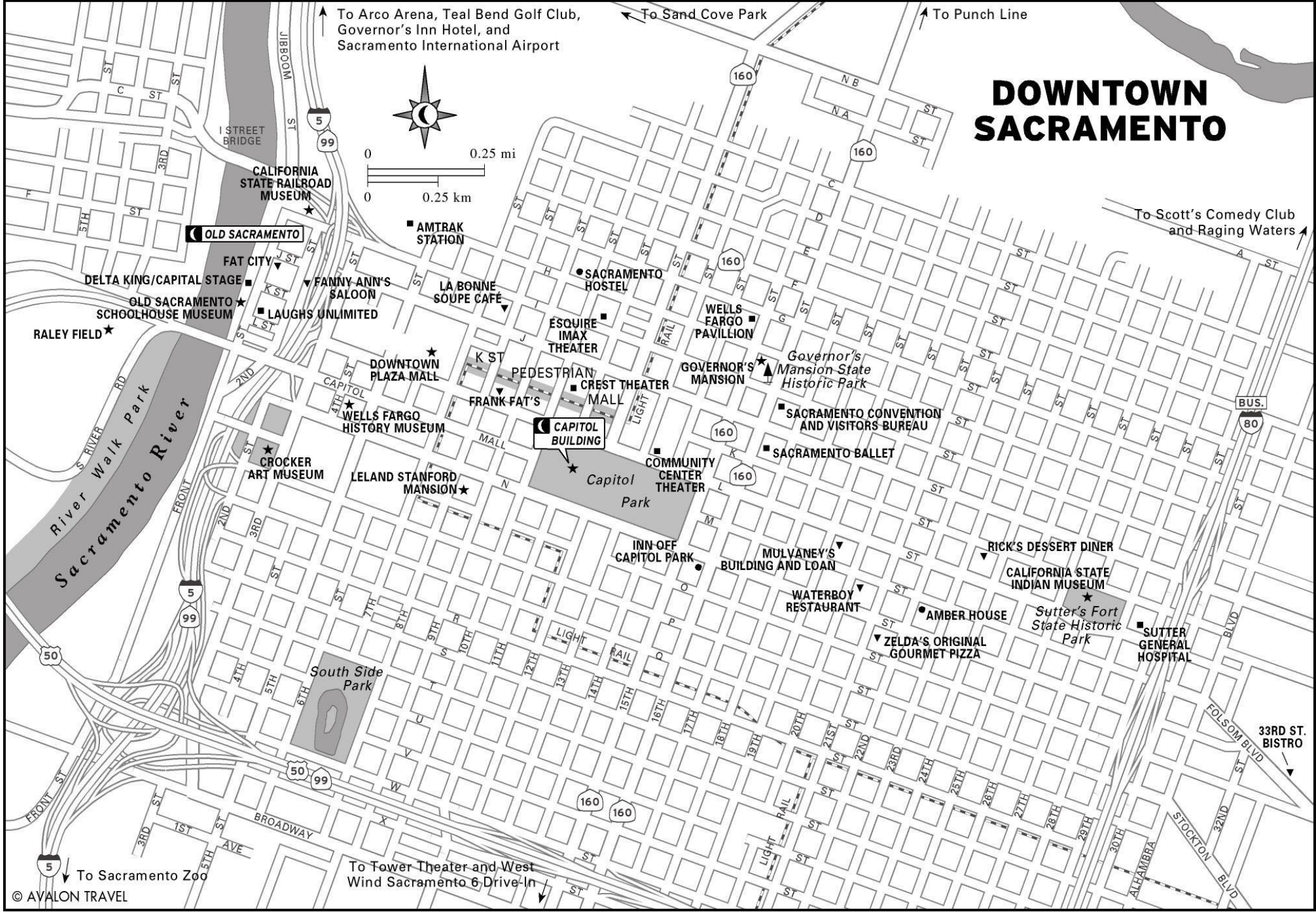
Continuously varying properties,
self defined area

The problem of scale

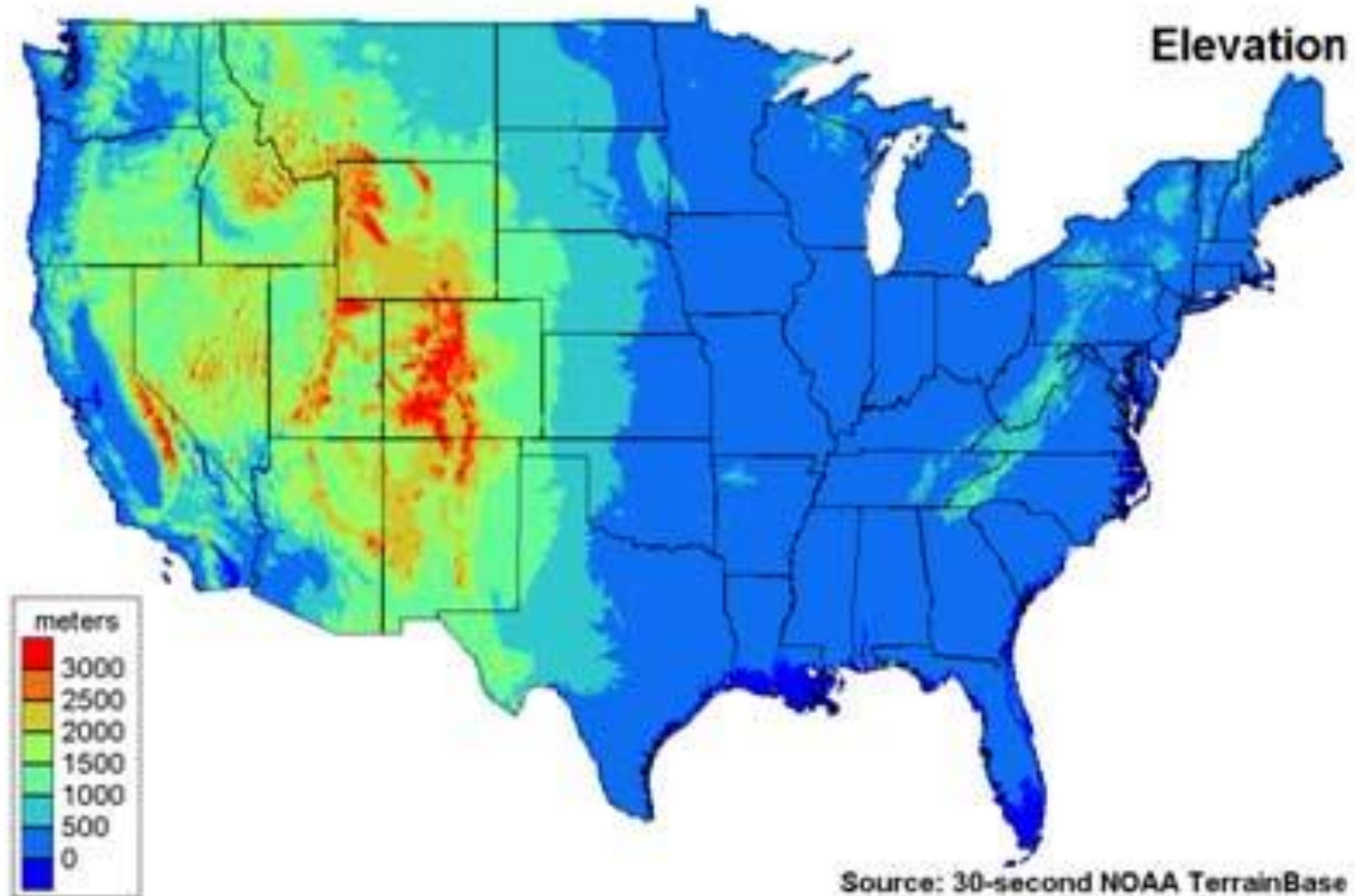
Representation and analysis
are scale dependent



Fuzziness

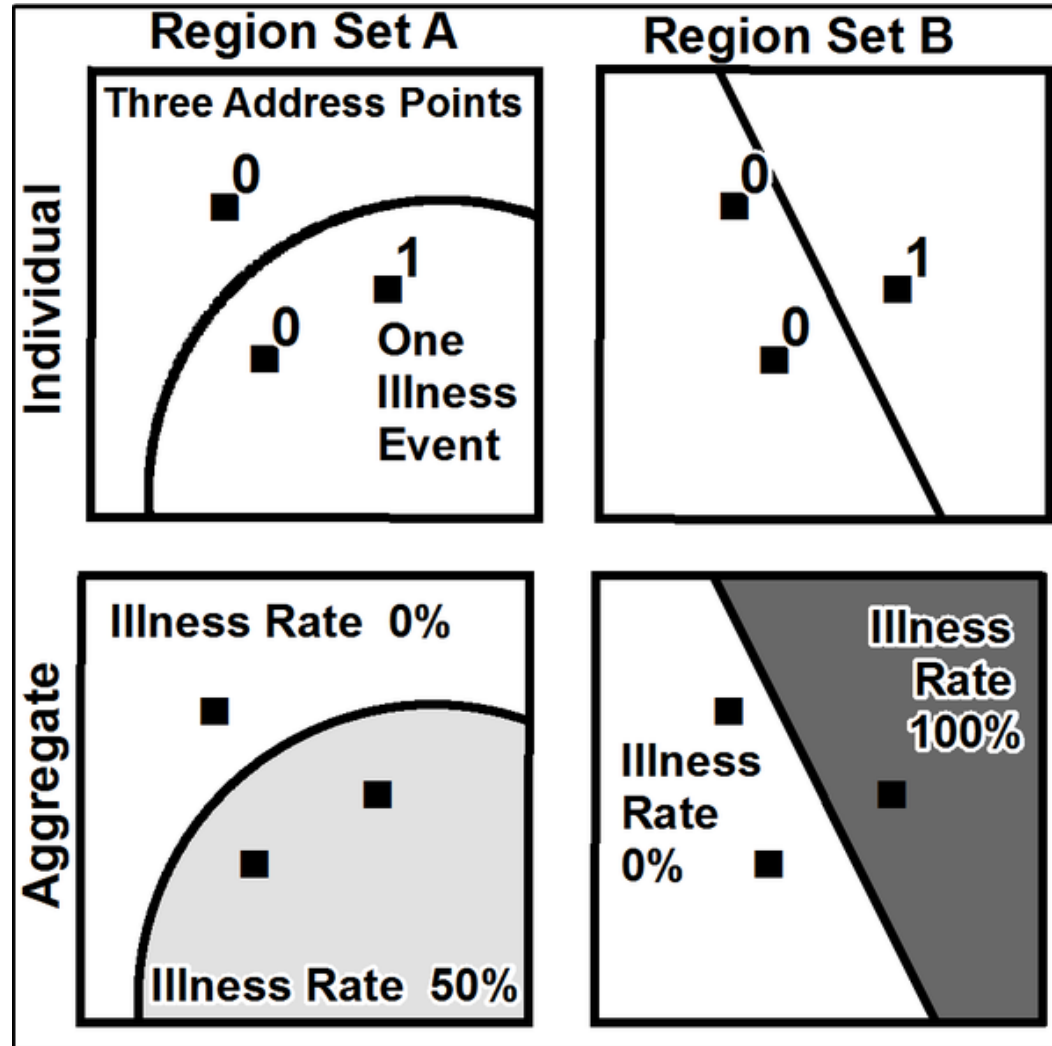


Spatial autocorrelation



Modifiable Areal Unit Problem

Zoning effect



Statistical relationships depend on the spatial grouping of the data

Do old people vote Republican?

Correlation between % people over 60
and % voting Republican
in 1968 Iowa election

Combination of counties	r
6 Congressional dist. proposed by Republicans	0.5
6 Congressional dist. proposed by Democrats	0.6
6 actual Congressional districts	0.3
6 districts based on urban vs. rural	0.9
99 Iowa counties	0.4

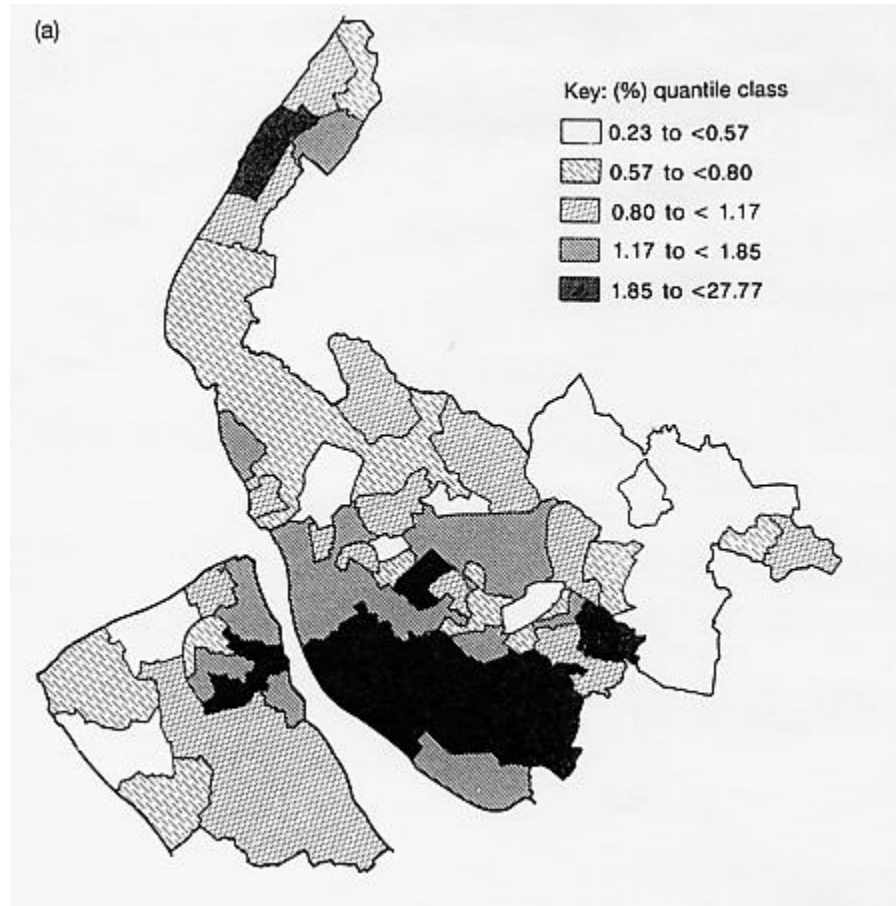
Openshaw, S., and P. Taylor. 1979.

A million or so correlation coefficients: three experiments on the modifiable areal unit problem.

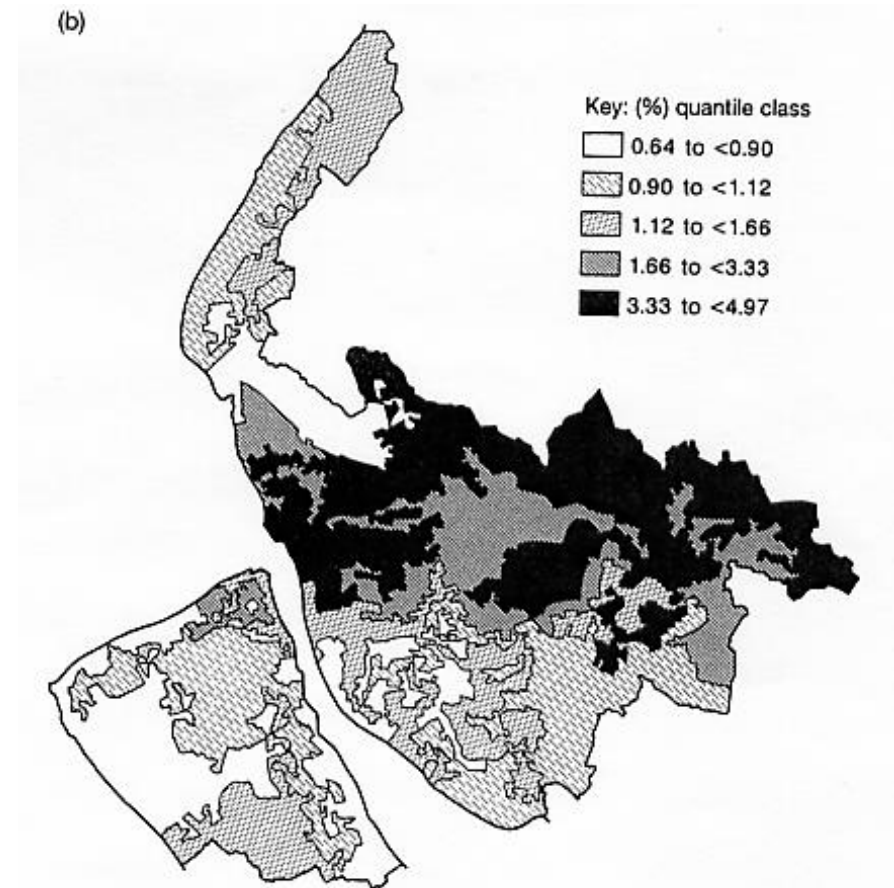
In N. Wrigley, (ed.), Statistical Applications in the Spatial Sciences. pp. 127-144.

Recent immigrant population in Merseyside, UK

119 wards



119 zones with equal population



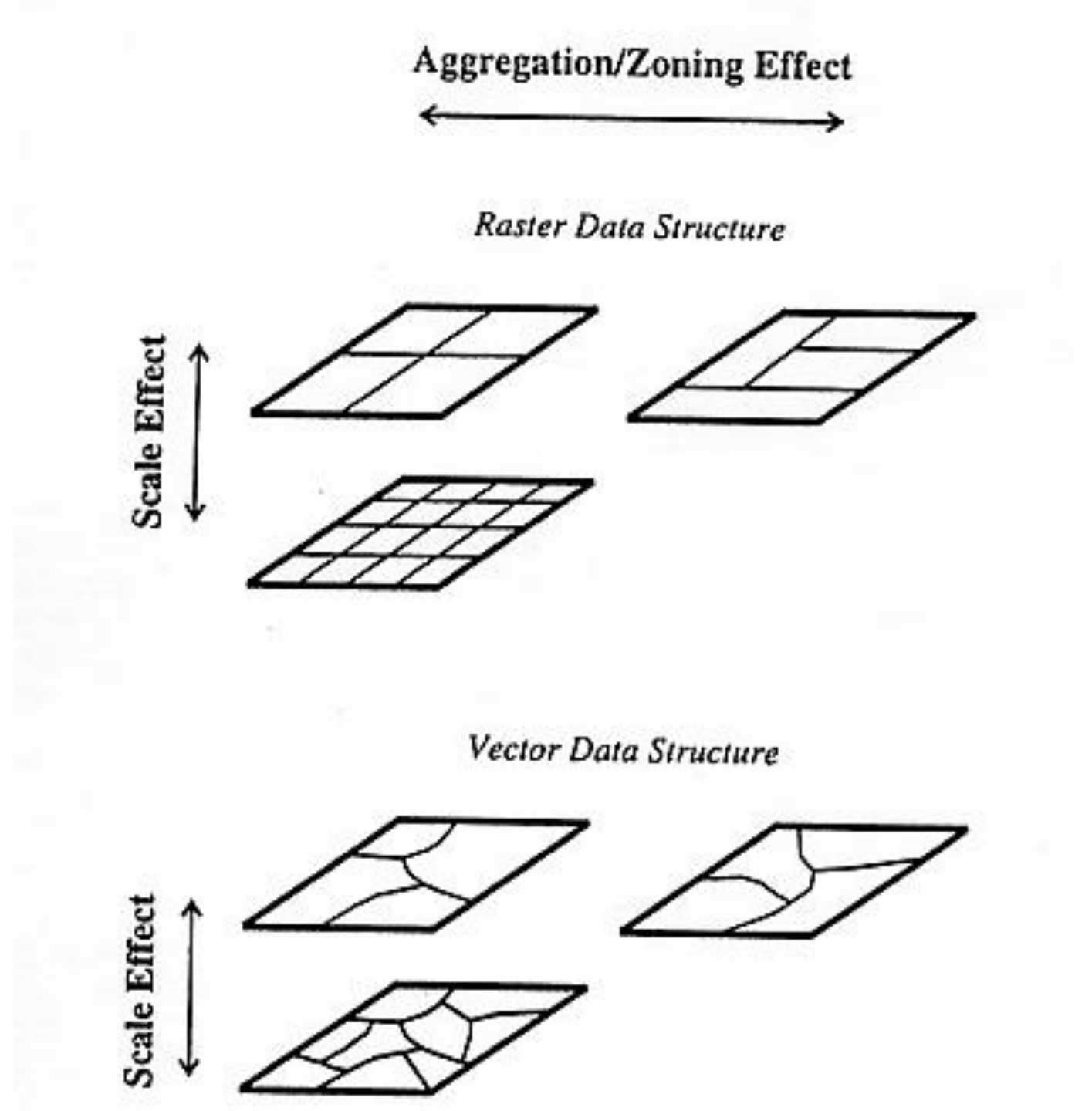
Aggregation effect

Correlation coefficients between
wheat yield & elevation and yield & NDVI

Cell size (m)	Yield & Elevation	Yield & NDVI
9 x 9	-0.25	0.50
27 x 27	-0.31	0.62
55 x 55	-0.36	0.63
110 x 110	-0.50	0.41
165 x 165	-0.63	0.00
329 x 329	-0.71	-0.65

Scale and zoning effects

(Wong, 1995)



Approaches to “solving” the MAUP (Wong, 1995)

- “Optimal” zoning systems
- Identification of meaningful zones
- Sensitivity analysis
- Abandonment of statistical analysis
- Focus on rates of change

Ecological inference / fallacy

Analyze grouped data (“ecological data”) to infer individual level relationships

For example:

What is the relationship between tree height and topsoil depth?

Ecological inference:

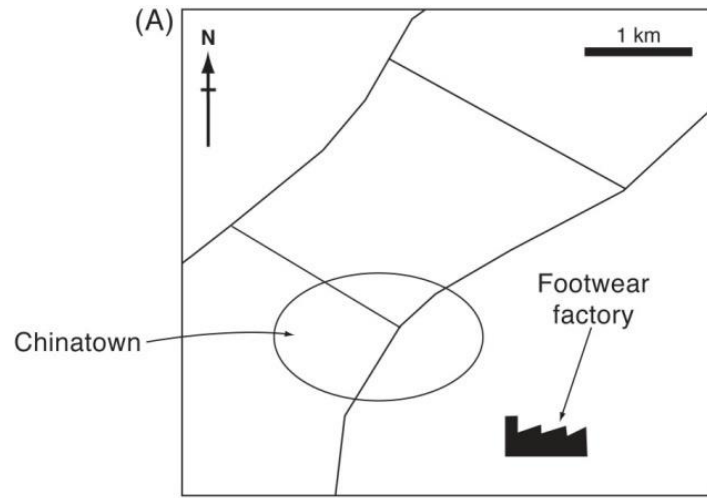
For each region, estimate average tree height (y) and average topsoil depth rather than individually measuring these parameters for each tree.

1930 census

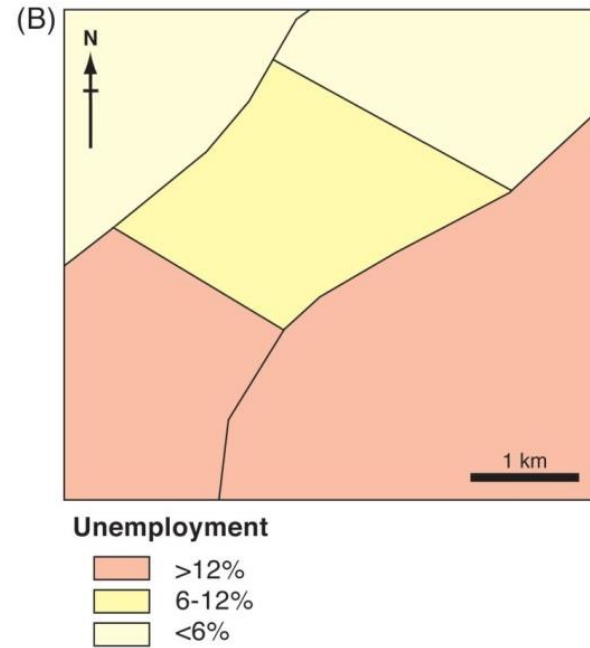
Relation between the illiteracy rate and the proportion of the population born outside the US

Aggregated by state: $r = -0.53$

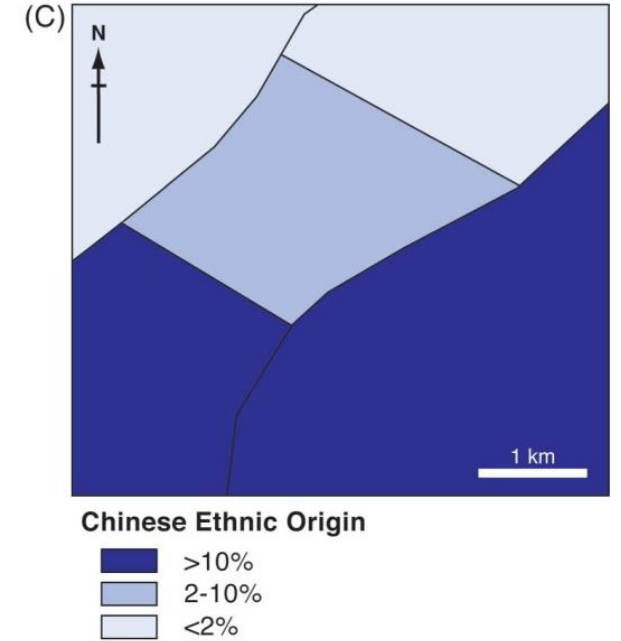
Individual level: $r = +0.12$



Before it closed down, the footwear factory drew its labor from its local neighborhood and a jurisdiction to the west



The closure caused high unemployment, but not amongst the workers of Chinatown

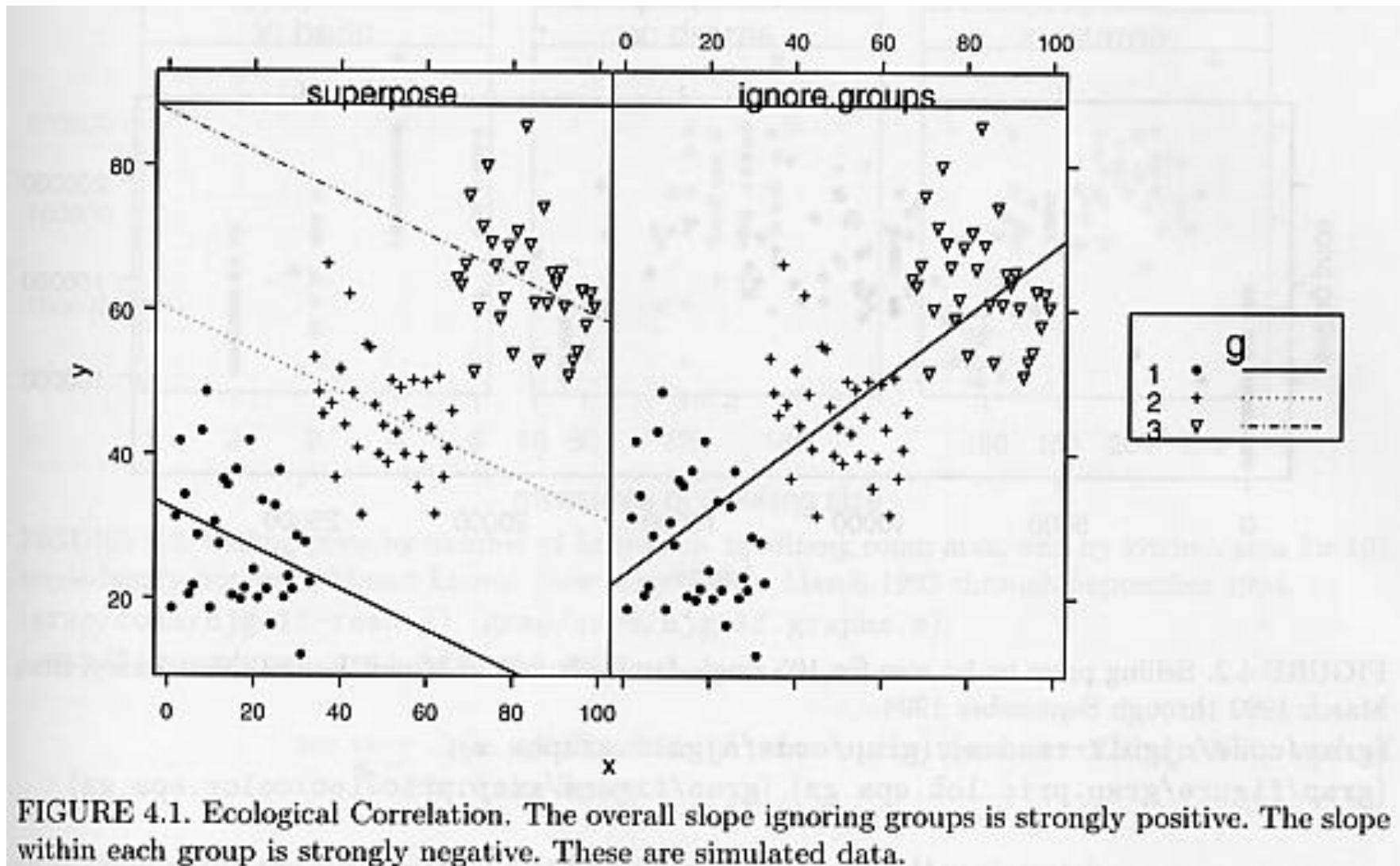


A spurious relationship between Chinese ethnicity and unemployment

In the 2004 presidential election, 62% of voters with annual incomes over \$200,000 voted for George W. Bush and 36% of voters with annual incomes of \$15,000 or less voted for Bush.

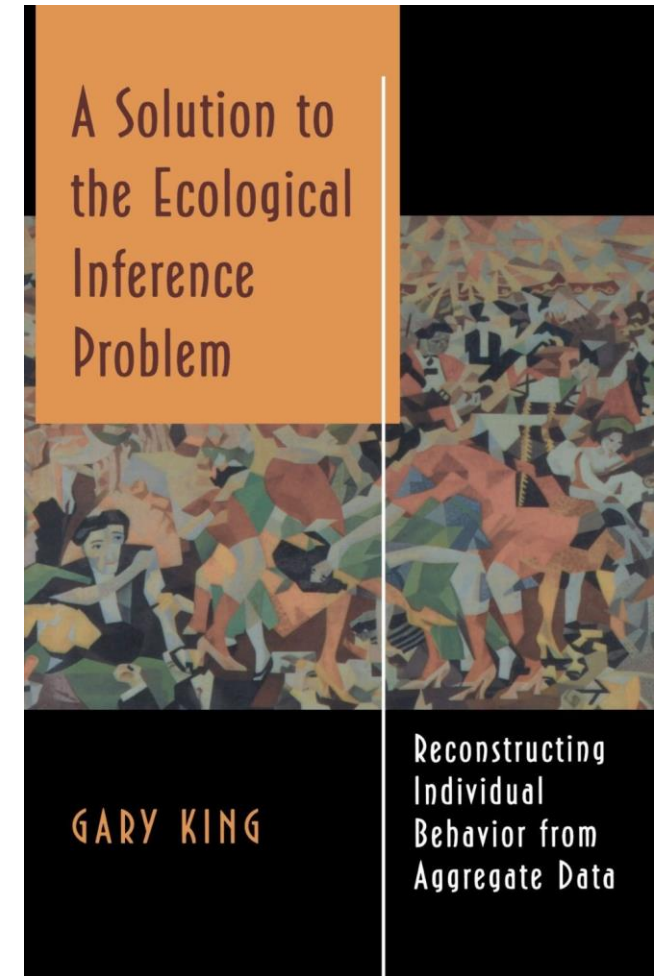
Yet, George W. Bush, won the fifteen poorest states, and John Kerry won 9 of the 11 wealthiest states.

Correlation and aggregation

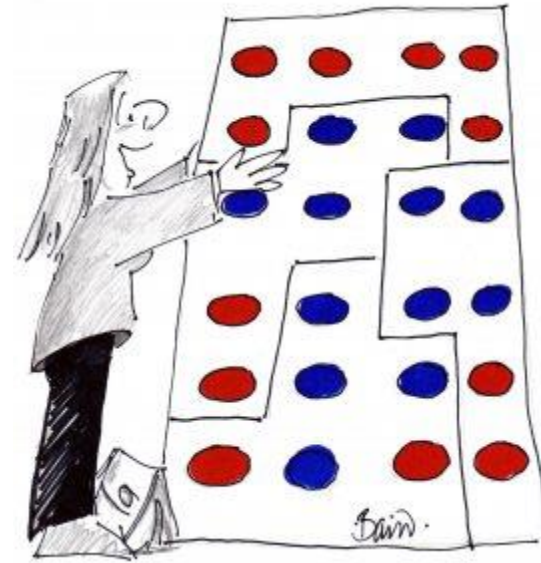
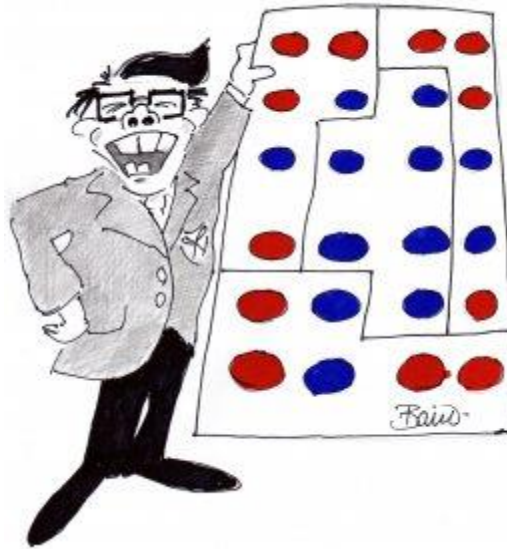
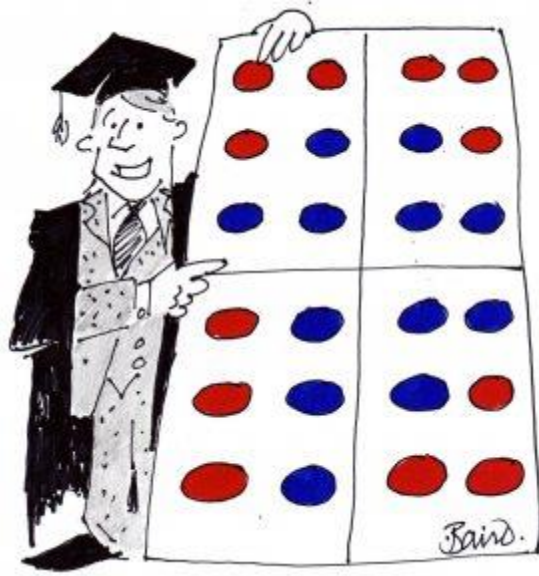


Ecological inference

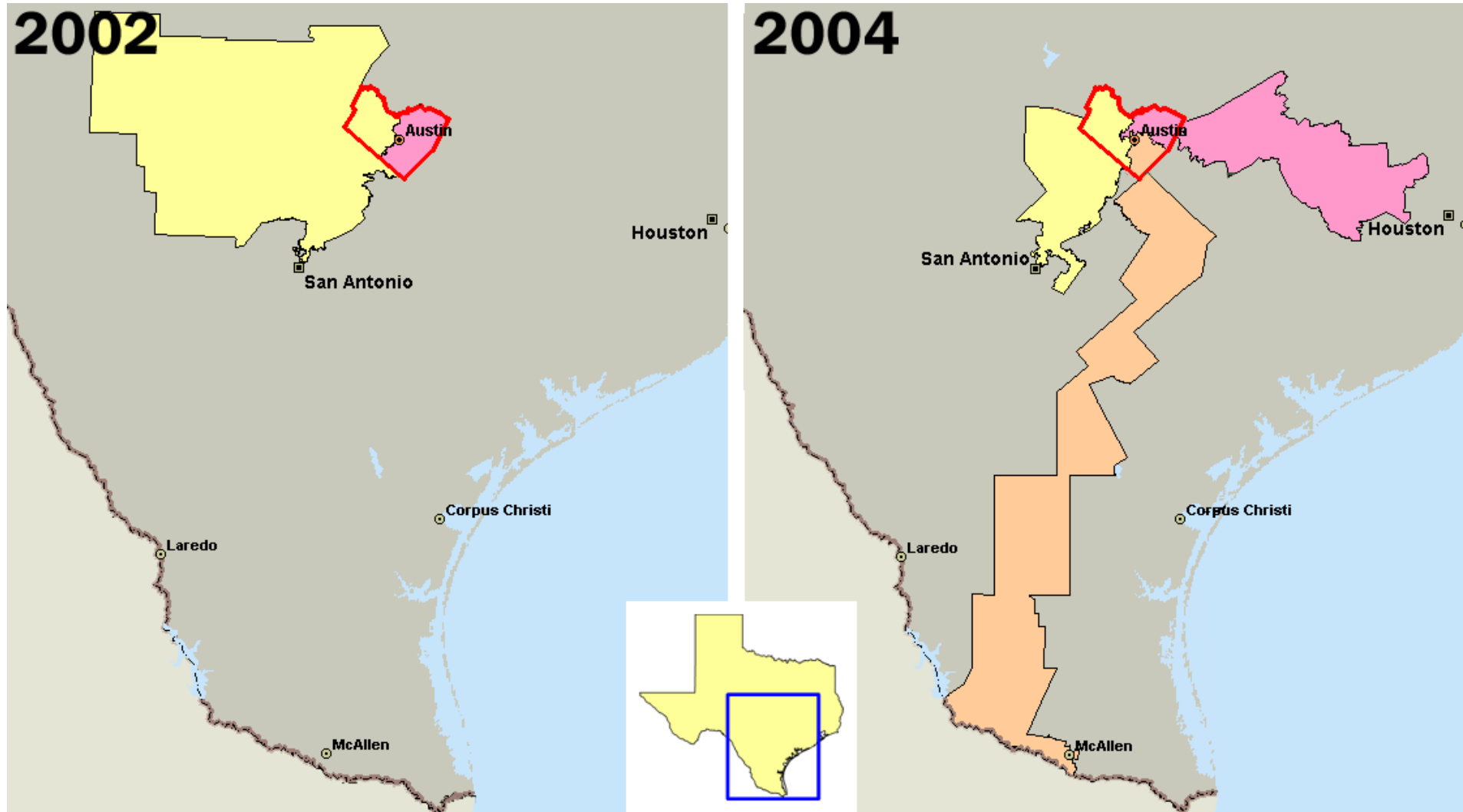
- Data is often only available on a regional basis.
- The data may be used in an attempt to draw conclusions about behavior or properties of individuals. These conclusions may not be valid.
- There is no general solution (although there is some work that aims to find it)



The MAUP *application*



Redistricting



Congressional District 38



Congressional District 23



California Citizens Redistricting Commission

California Proposition 11, the Voters First Act, 2008

Five Democrats, five Republicans, and four commissioners from neither major party.

Rank-ordered criteria that the Commission followed to create new districts:

Population Equality: Districts must comply with the U.S. Constitution's requirement of "one person, one vote"

Federal Voting Rights Act: Districts must ensure an equal opportunity for minorities to elect a candidate of their choice

Geographic Contiguity: All areas within a district must be connected to each other (except for islands)

Geographic Integrity: Districts shall minimize the division of cities, counties, local neighborhoods and communities of interests to the extent possible. A community of interest is a contiguous population which shares common social and economic interests that should be included within a single district for purposes of its effective and fair representation.

Geographic Compactness: Districts must not bypass nearby communities for more distant communities

Nesting: Each Senate district will be composed of two whole Assembly districts, Board of Equalization districts will be composed of 10 Senate districts.

CONGRESSIONAL DISTRICTS

113th Congress (January 2013–January 2015)

The Constitution prescribes Congressional apportionment based on decennial census population data. Each state has at least one Representative, no matter how small its population. Since 1941, distribution of Representatives has been based on total U.S. population, so that the average population per Representative has the least possible variation between one state and any other. Congress fixes the number of voting Representatives at each apportionment. States delineate the district boundaries. The first House of Representatives in 1789 had 65 members; currently there are 435. There are non-voting delegates from American Samoa, the District of Columbia, Guam, Puerto Rico, and the Virgin Islands.

1	15	29	43
2	16	30	44
3	17	31	45
4	18	32	46
5	19	33	47
6	20	34	48
7	21	35	49
8	22	36	50
9	23	37	51
10	24	38	52
11	25	39	53
12	26	40	
13	27	41	
14	28	42	

