The Meaning of Spatial Thinking

Michael F. Goodchild
University of California
Santa Barbara

Outline

- Spatial cognition
- Examples
- GIS representations
- Empirical laws of geographic information
- Basic concepts
- Implications for teaching

Spatial thinking

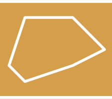
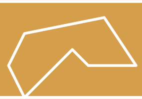
- Spatial cognition
  - how we think about the world around us
- Spatial reasoning
  - how we come to conclusions
- Knowledge discovery
  - space in support of science

Spatial cognition

- Fundamental ideas about how spatial skills develop
  - early childhood, Piaget
  - objects
    - object permanence
  - containment
    - Ontario is in Canada

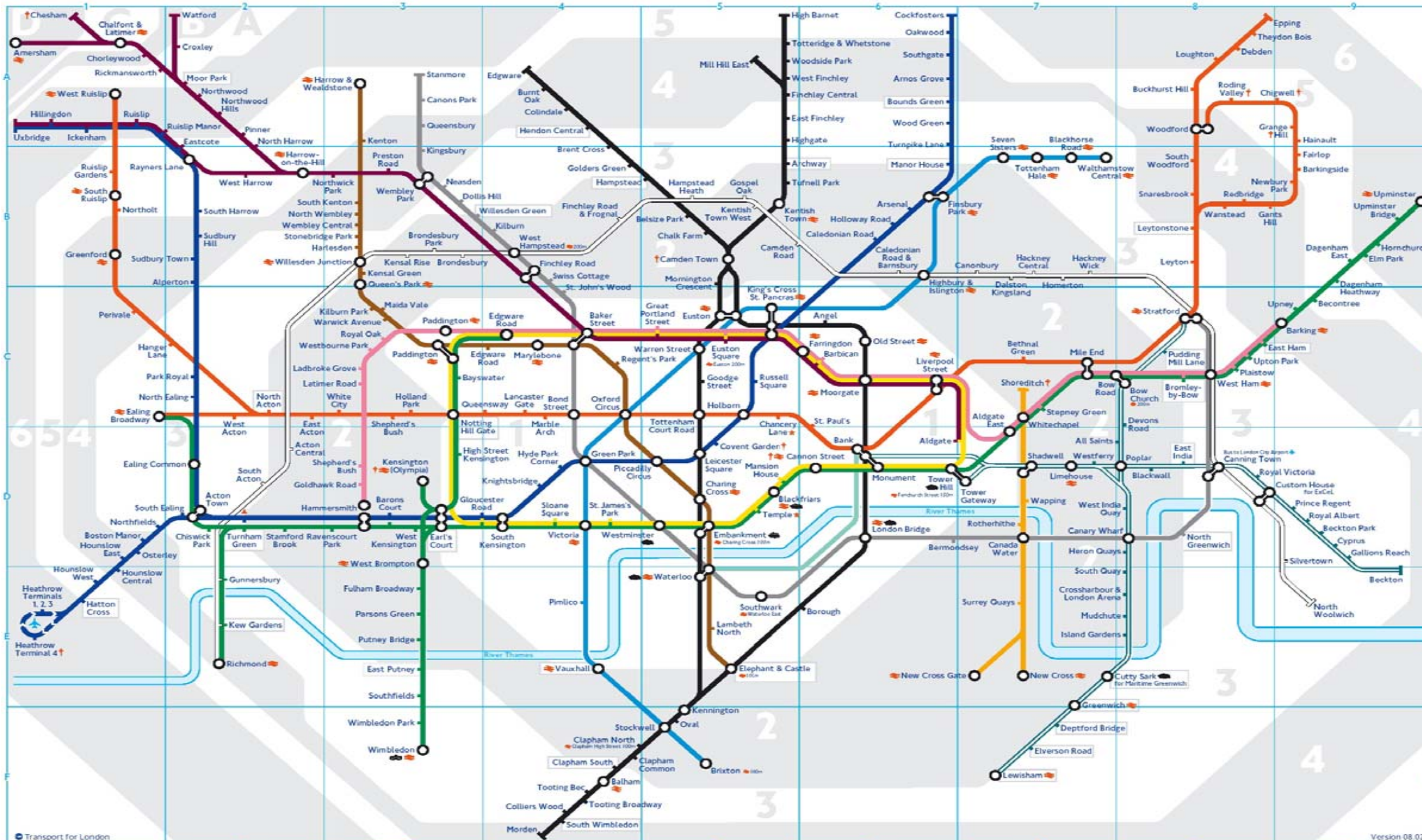
Wayfinding skills

- Landmark knowledge
  - a list of places
  - no spatial relationships
    - no adjacency
    - "if this is Tuesday it must be Belgium"
  - no spatial context
    - "how long is this flight?"
  - no intervening places

Tract	Pop	Location	Shape
1	3786	*x*,*y*	
2	2966	*x*,*y*	
3	5001	*x*,*y*	
4	4983	*x*,*y*	
5	4130	*x*,*y*	
6	3229	*x*,*y*	
7	4086	*x*,*y*	
8	3979	*x*,*y*	

Route knowledge

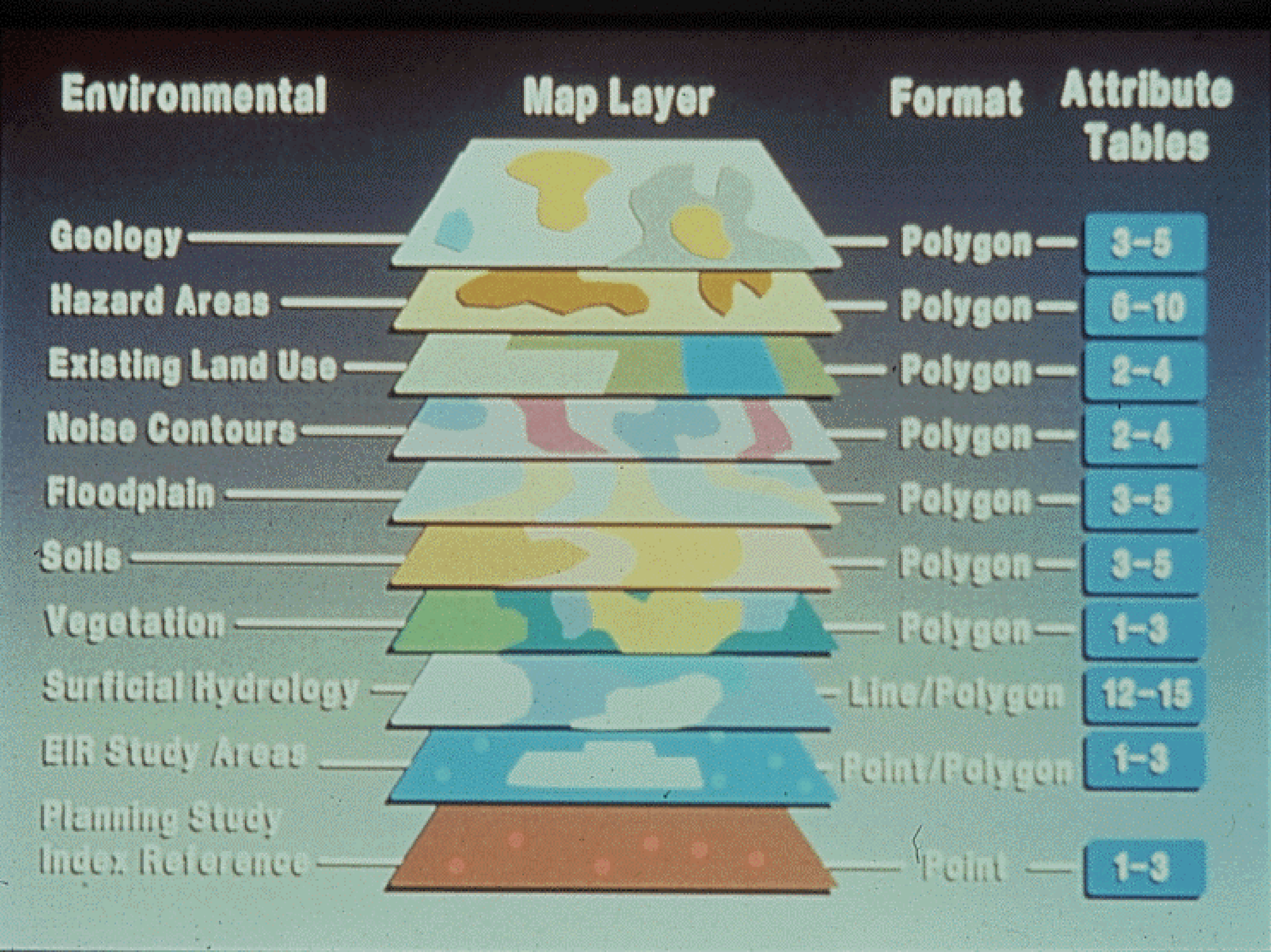
- Sequences of intervening places
  - no ability to short-cut
  - no directions, distances
  - context along the route but not off

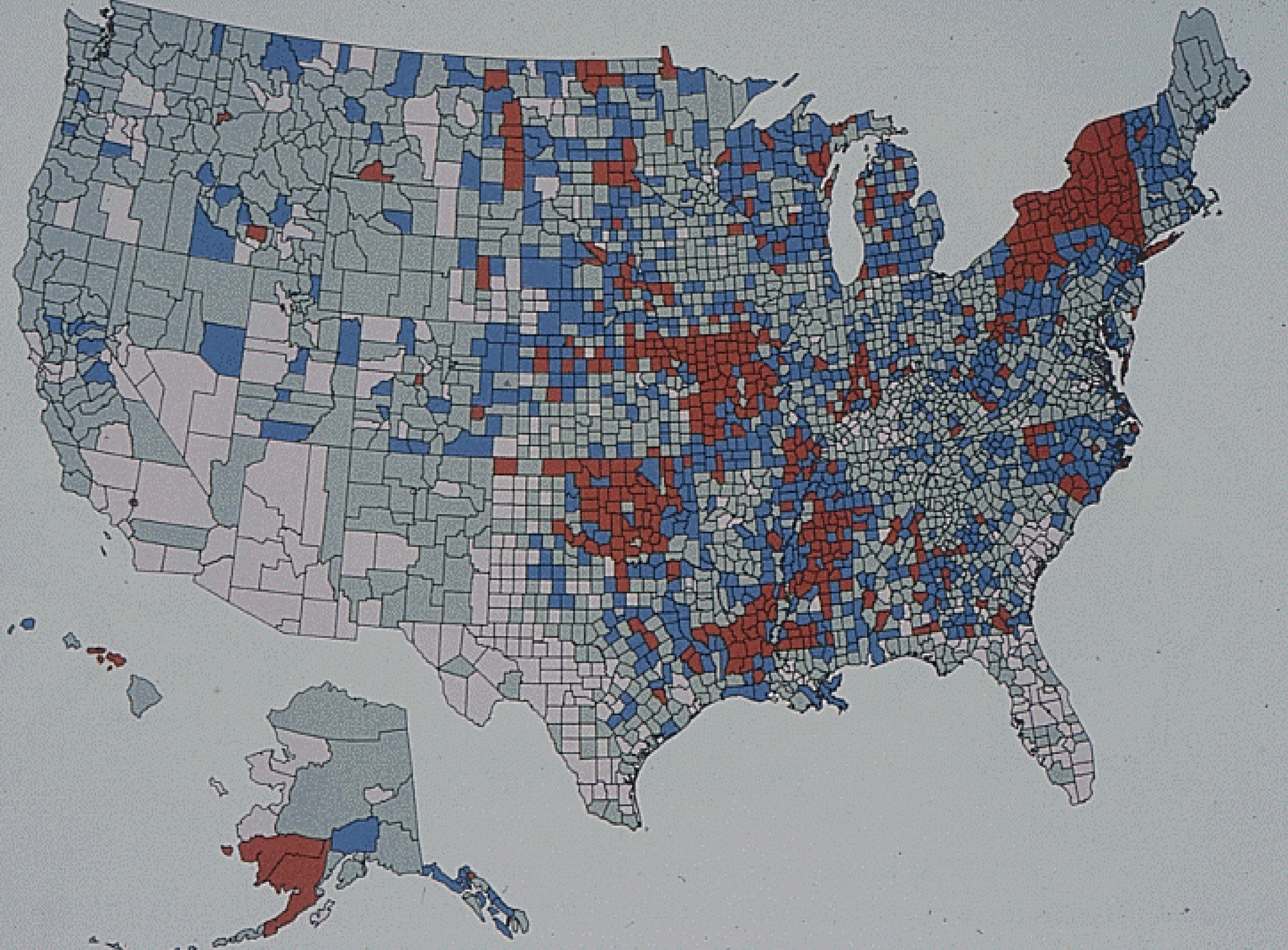


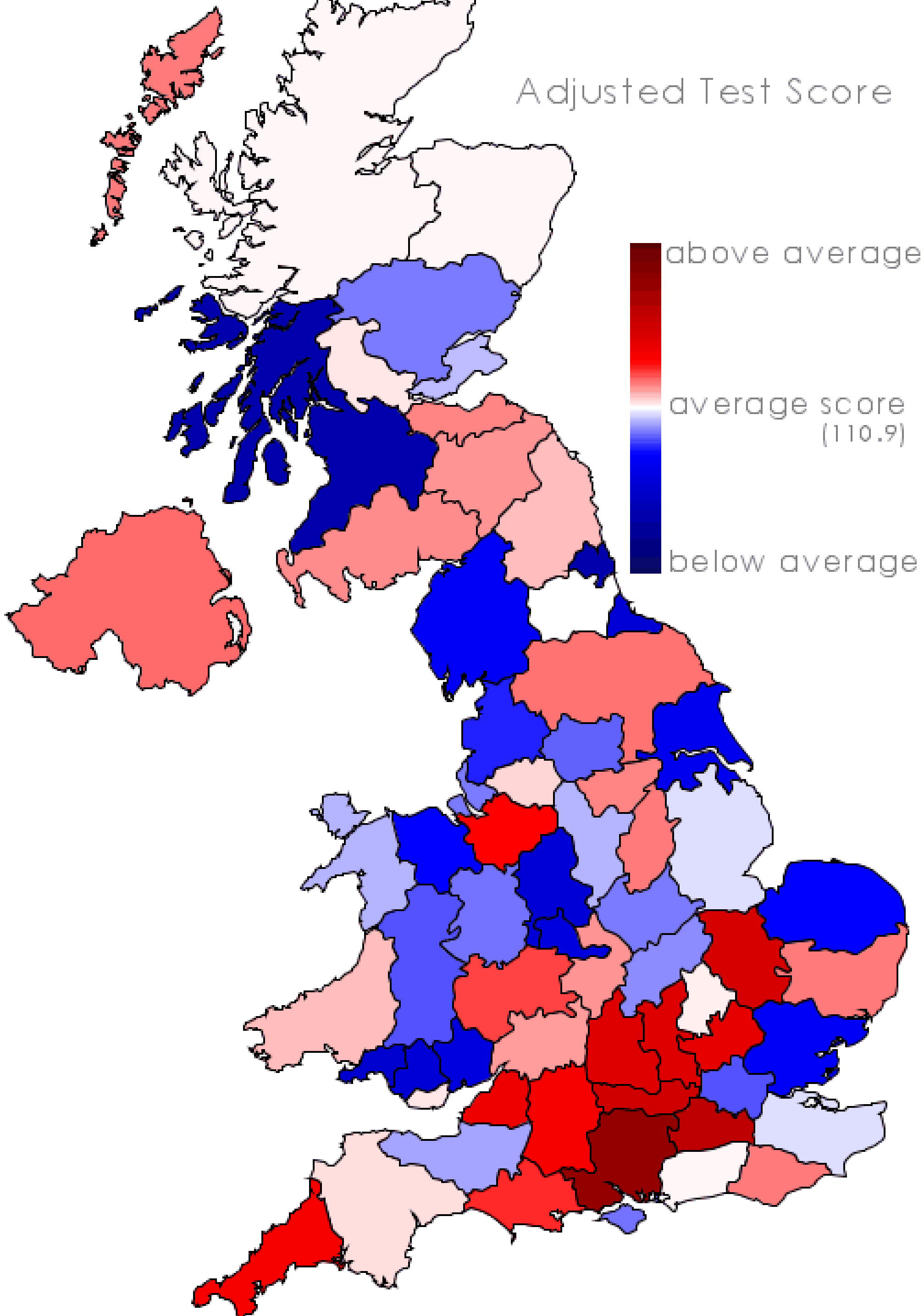
Chesham
Chalfont & Latimer
Amersham
Chorleywood
Rickmansworth
Watford
Croxley
Moor Park
Northwood
Northwood Hills
Pinner
North Harrow
Harrow-on-the-Hill
West Harrow
West Ruislip
Hillingdon
Ruislip
Ruislip Manor
Eastcote
Uxbridge
Ickenham
Ruislip Gardens
Rayners Lane
South Ruislip
Northolt
South Harrow
Sudbury Hill
Greenford
Sudbury Town
Alperton
Perivale
Hanger Lane
Park Royal
North Ealing
Ealing Broadway
West Acton
North Acton
East Acton
Ealing Common
South Acton
Acton Town
Acton Central
South Ealing
Northfields
Boston Manor
Osterley
Hounslow East
Hounslow Central
Hounslow West
Hatton Cross
Heathrow Terminals 1, 2, 3
Heathrow Terminal 4
Chiswick Park
Turnham Green
Stamford Brook
Ravenscourt Park
Gunnersbury
Kew Gardens
Richmond
Harrow & Wealdstone
Kenton
Preston Road
Northwick Park
South Kenton
North Wembley
Wembley Central
Stonebridge Park
Harlesden
Willesden Junction
Wembley Park
Stanmore
Canons Park
Queensbury
Kingsbury
Neasden
Dollis Hill
Willesden Green
Kilburn
West Hampstead
Brondesbury Park
Brondesbury
Kensal Rise
Kensal Green
Queen's Park
Kilburn Park
Maida Vale
Warwick Avenue
Royal Oak
Westbourne Park
Ladbroke Grove
Latimer Road
White City
Shepherd's Bush
Holland Park
Edgware
Burnt Oak
Colindale
Hendon Central
Brent Cross
Golders Green
Hampstead
Finchley Road & Frognal
Finchley Road
Swiss Cottage
St. John's Wood
Hampstead Heath
Gospel Oak
Belsize Park
Chalk Farm
Kentish Town West
Camden Town
Mill Hill East
High Barnet
Totteridge & Whetstone
Woodside Park
West Finchley
Finchley Central
East Finchley
Highgate
Archway
Tufnell Park
Kentish Town
Camden Road
Mornington Crescent
Cockfosters
Oakwood
Southgate
Arnos Grove
Bounds Green
Wood Green
Turnpike Lane
Manor House
Arsenal
Holloway Road
Caledonian Road
Caledonian Road & Barnsbury
Finsbury Park
Seven Sisters
Blackhorse Road
Tottenham Hale
Walthamstow Central
Highbury & Islington
Canonbury
Dalston Kingsland
Hackney Central
Hackney Wick
Homerton
Paddington
Edgware Road
Marylebone
Baker Street
Great Portland Street
Euston
Warren Street
Regent's Park
Euston Square
King's Cross St. Pancras
Angel
Farringdon
Barbican
Old Street
Moorgate
Liverpool Street
Goodge Street
Russell Square
Holborn
Bayswater
Queensway
Notting Hill Gate
Lancaster Gate
Marble Arch
Bond Street
Oxford Circus
Tottenham Court Road
Chancery Lane
St. Paul's
Bank
Aldgate
Aldgate East
Shoreditch
Bethnal Green
Mile End
Stepney Green
Whitechapel
High Street Kensington
Kensington (Olympia)
Hyde Park Corner
Green Park
Piccadilly Circus
Covent Garden
Leicester Square
Mansion House
Cannon Street
Monument
Tower Hill
Tower Gateway
Charing Cross
Blackfriars
Temple
Embankment
Knightsbridge
Gloucester Road
Sloane Square
St. James's Park
Westminster
Victoria
South Kensington
Barons Court
Hammersmith
Goldhawk Road
West Kensington
Earl's Court
West Brompton
Fulham Broadway
Parsons Green
Putney Bridge
East Putney
Southfields
Wimbledon Park
Wimbledon
Pimlico
Vauxhall
Waterloo
Southwark
Lambeth North
Elephant & Castle
Borough
London Bridge
Kennington
Oval
Stockwell
Brixton
Clapham North
Clapham Common
Clapham South
Balham
Tooting Bec
Tooting Broadway
Colliers Wood
South Wimbledon
Morden
Bermondsey
Canada Water
Rotherhithe
Wapping
Shadwell
Limehouse
Westferry
Surrey Quays
New Cross Gate
New Cross
West India Quay
Canary Wharf
Heron Quays
South Quay
Crossharbour & London Arena
Mudchute
Island Gardens
Cutty Sark
Greenwich
Deptford Bridge
Elverson Road
Lewisham
Epping
Theydon Bois
Debden
Loughton
Buckhurst Hill
Roding Valley
Chigwell
Grange Hill
Hainault
Fairlop
Barkingside
Newbury Park
Woodford
South Woodford
Snaresbrook
Redbridge
Wanstead
Gants Hill
Leytonstone
Leyton
Stratford
Pudding Mill Lane
Bow Road
Bow Church
Bromley-by-Bow
Devons Road
All Saints
Poplar
Blackwall
East India
Canning Town
Royal Victoria
Custom House for ExCeL
Prince Regent
Royal Albert
Beckton Park
Cyprus
Gallions Reach
Beckton
West Ham
Plaistow
Upton Park
East Ham
Barking
Upney
Becontree
Dagenham Heathway
Dagenham East
Elm Park
Hornchurch
Upminster Bridge
Upminster
North Greenwich
Silvertown
North Woolwich
River Thames
Transport for London
Version 08.02

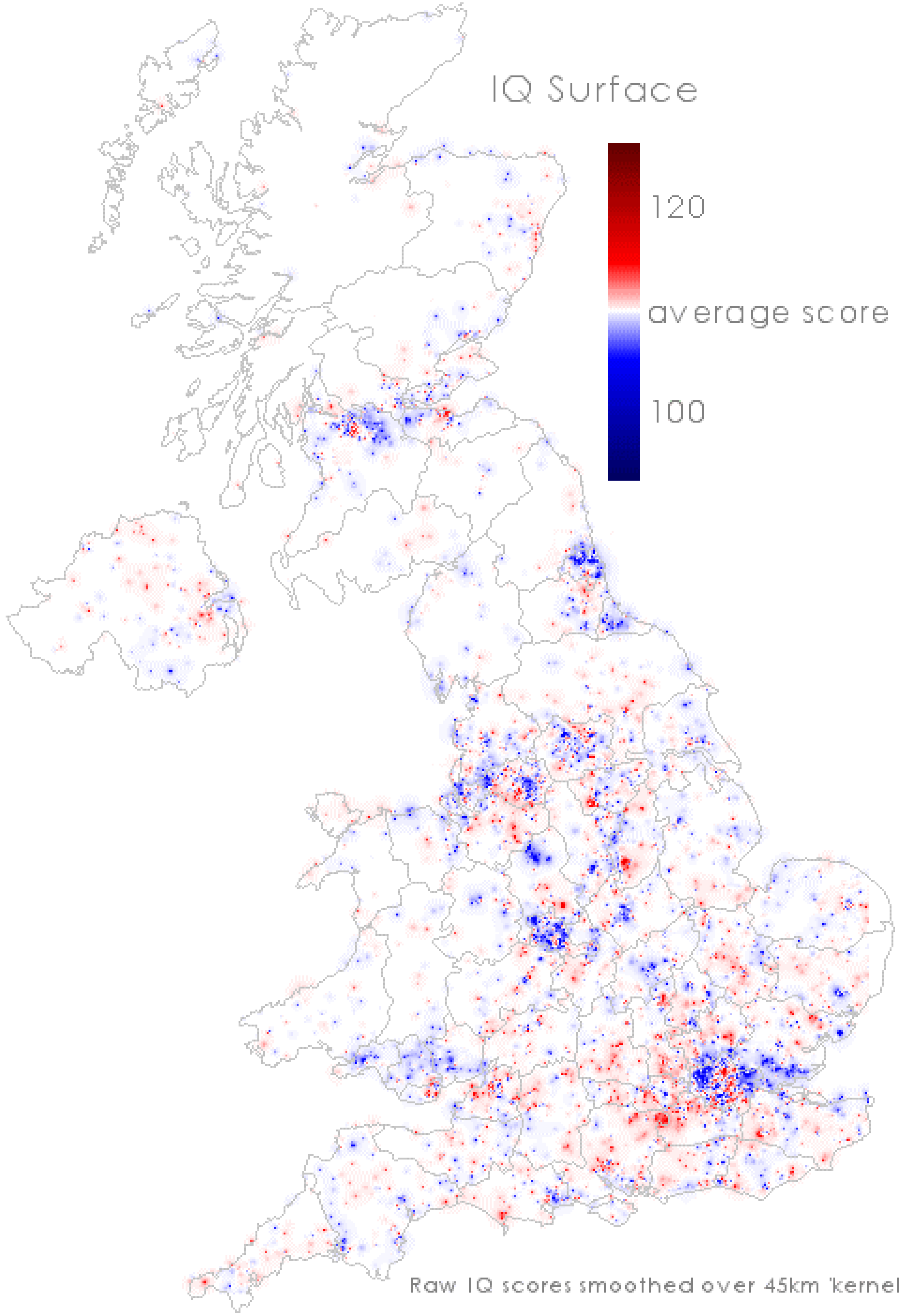
Survey knowledge

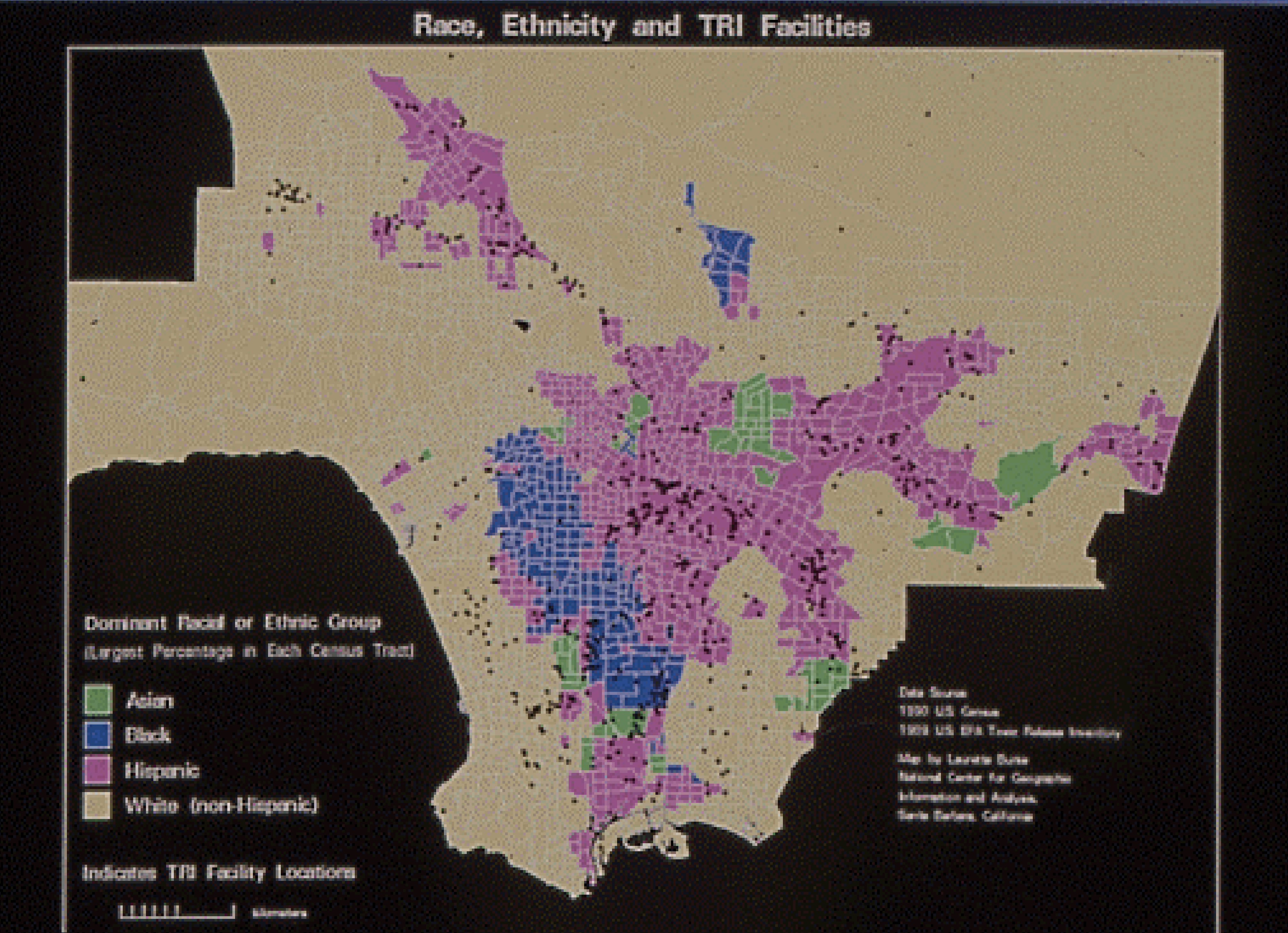
- Full two-dimensional representation
  - distances
  - orientations
  - shortcuts
  - context
    - vertical and horizontal

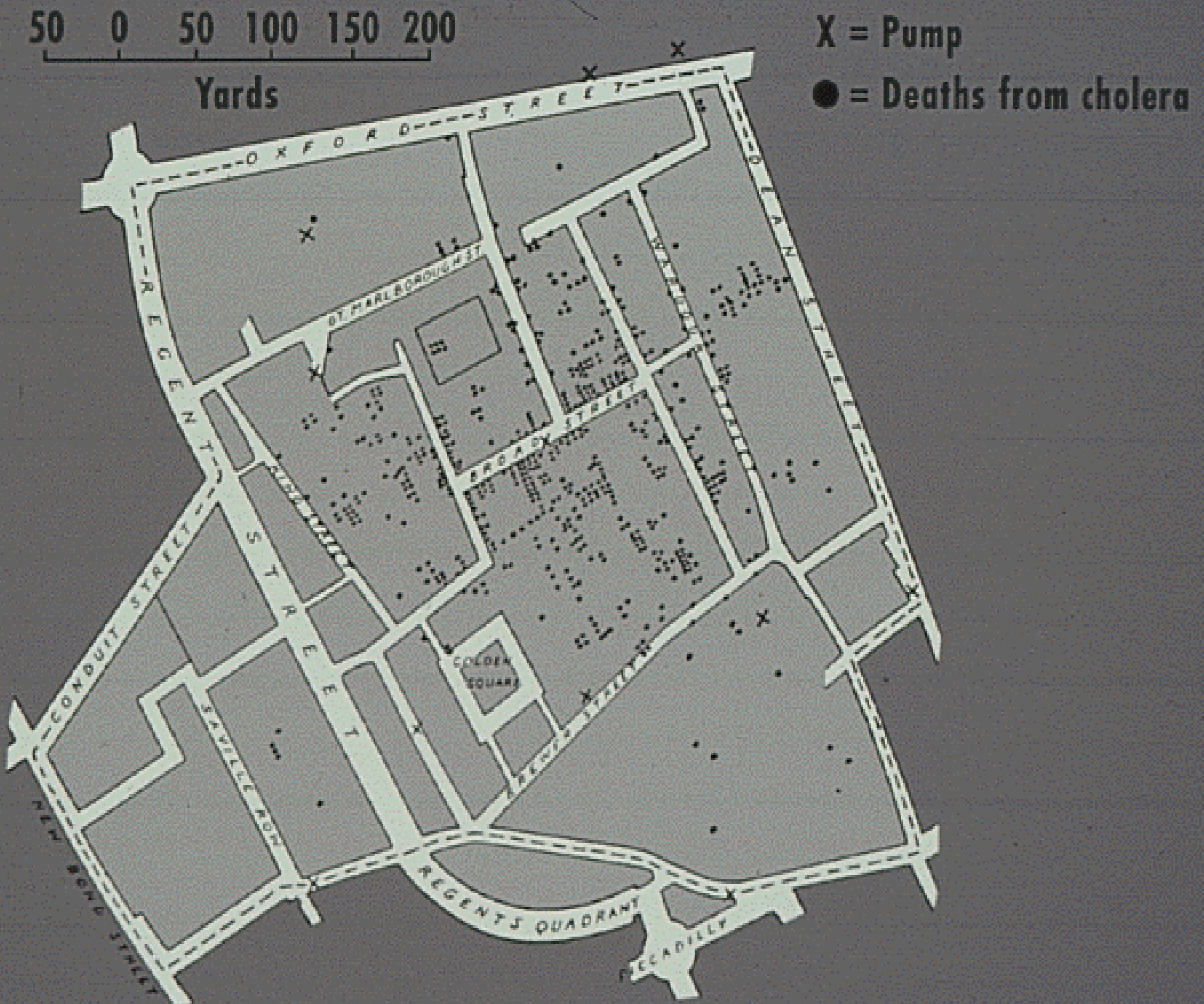

Environmental
Map Layer
Format
Attribute Tables
Geology
Polygon
3-5
Hazard Areas
Polygon
6-10
Existing Land Use
Polygon
2-4
Noise Contours
Polygon
2-4
Floodplain
Polygon
3-5
Soils
Polygon
3-5
Vegetation
Polygon
1-3
Surficial Hydrology
Line/Polygon
12-15
EIR Study Areas
Point/Polygon
1-3
Planning Study Index Reference
Point
1-3



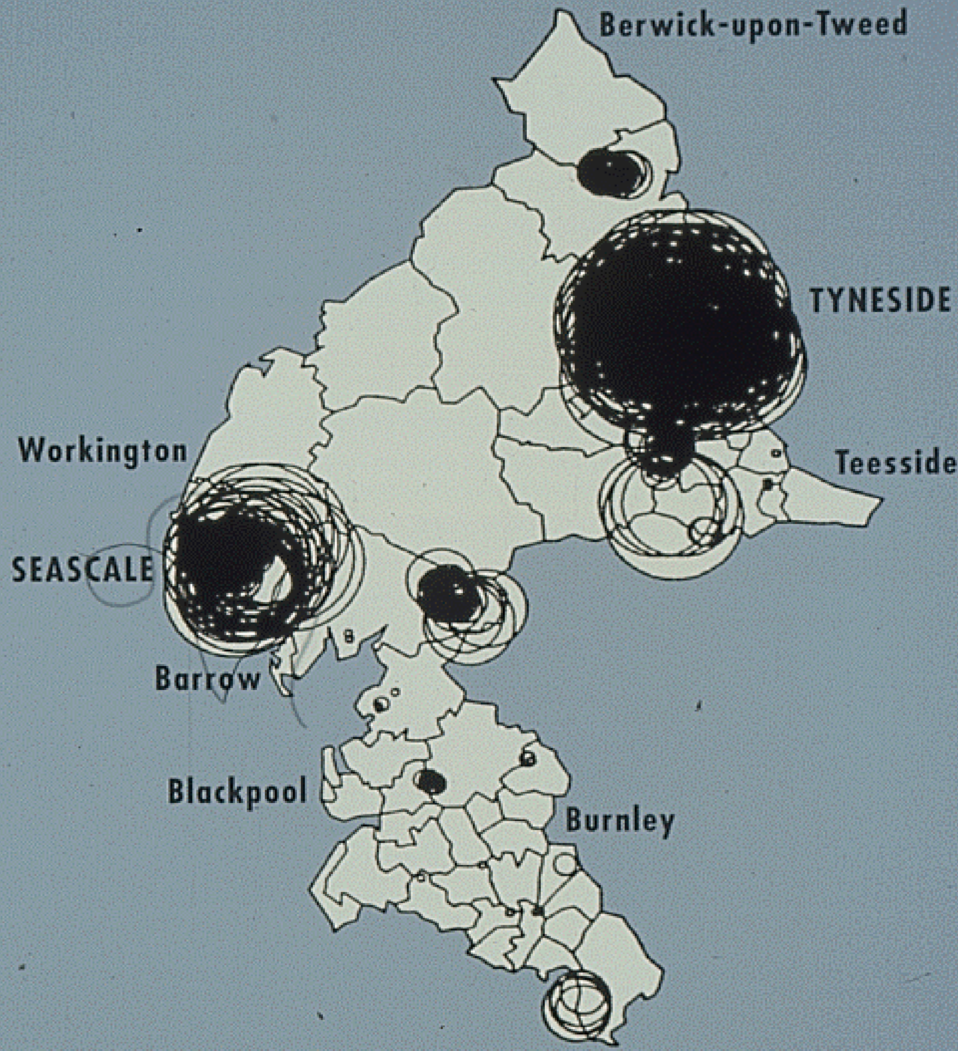

Adjusted Test Score
above average
average score
(110.9)
below average


IQ Surface
120
average score
100
Raw IQ scores smoothed over 45km 'kernel'


Race, Ethnicity and TRI Facilities
Dominant Racial or Ethnic Group
(Largest Percentage in Each Census Tract)
Asian
Black
Hispanic
White (non-Hispanic)
Indicates TRI Facility Locations
kilometers
Data Source
1990 U.S. Census
1989 U.S. EPA Toxic Release Inventory
National Center for Geographic
Information and Analysis,
Santa Barbara, California



The Snow Map of Cholera Incidence in the Area of Broad Street, London, in 1854.
The contaminated water pump is located at the center of the map, just to the right of the D in BROAD STREET.



Circles Indicating Significant Clusters of Cases of Acute Lymphoblastic Leukemia in the North of England, 1968-1985.

Naïve geography

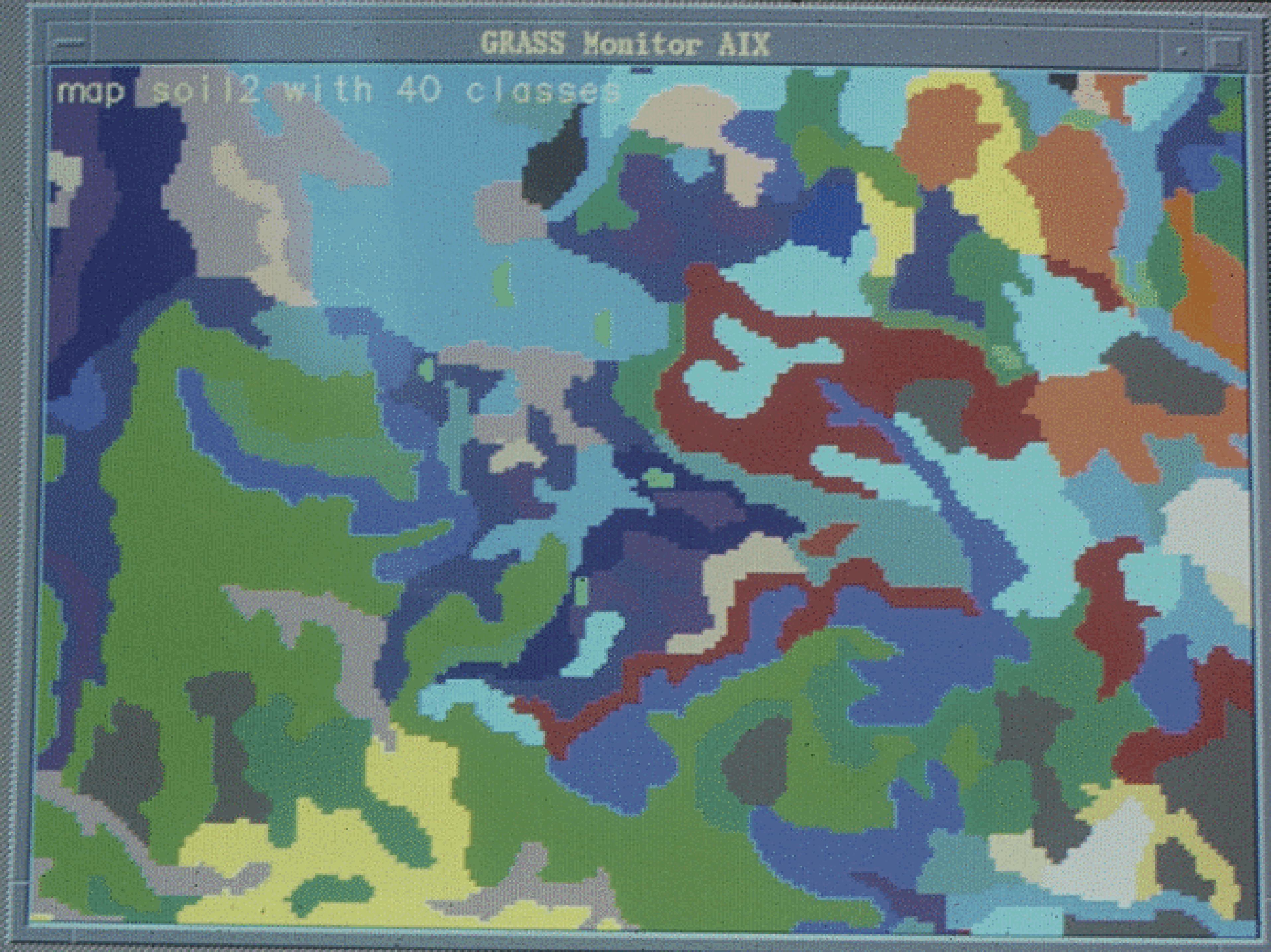
- Distortion of survey knowledge
  - to fit simple models
- The US as a rectangle
  - E and W coasts run N-S
    - Miami is S of NY
    - California is a rectangle
    - Santa Barbara is N of LA
    - Santa Barbara is W of Lake Tahoe
    - Las Vegas is E of San Diego
  - Canadian border is E-W
    - Detroit is S of Windsor

Geographic information

- Information linking a place **x** to a set of properties **z** (at some time *t*)
  - composed of atoms of the form <**x**,*t*,**z**>
  - mapping tradition is not sympathetic to *t*
  - atomic form only seen rarely
- Analysis and interpretation of geographic information
  - affected by changes in **x**

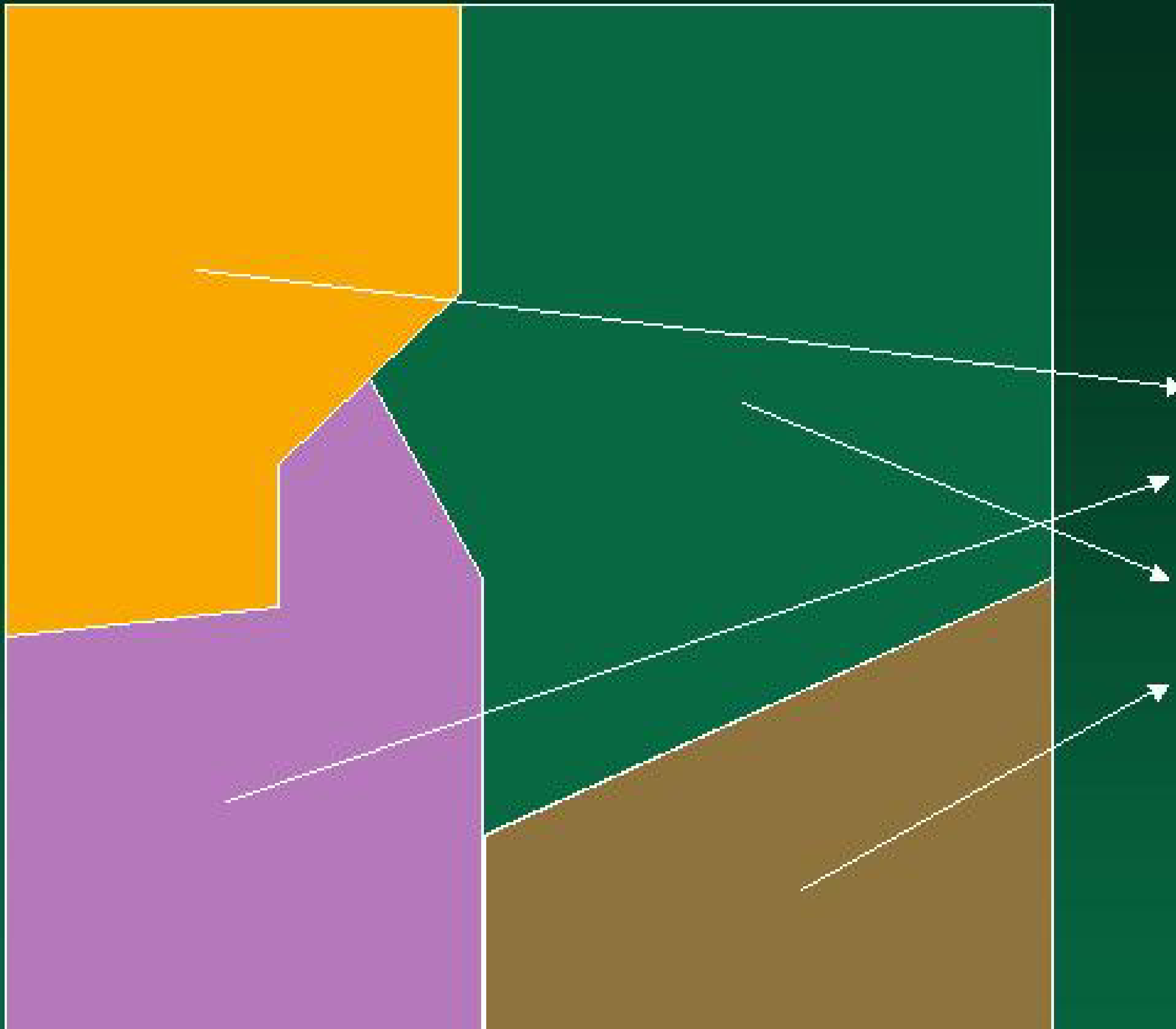
Standard coding schemes

- Music: MIDI, MP3
- Images: JPEG, TIFF, GIF
- FAX: CCITT
- Text: ASCII
- Planet Earth:
  - how to express knowledge about the planet's surface in 0s and 1s

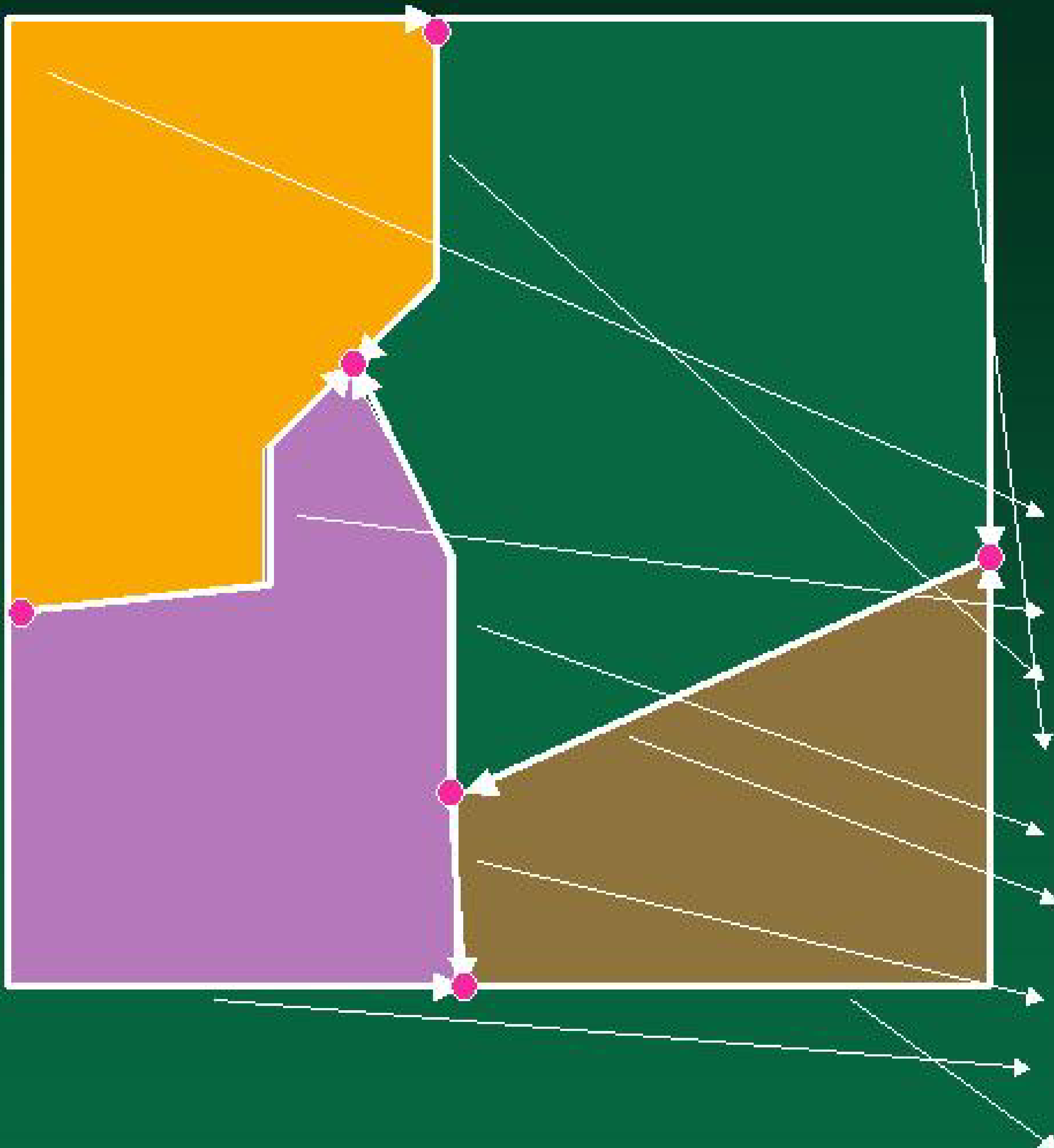

GRASS Monitor AIX
map soil2 with 40 classes

Coverage model

- Polygons
- Common boundaries
  - arcs
  - coded as polylines
  - pointers to leftpoly, rightpoly, fromnode, tonode
  - topology

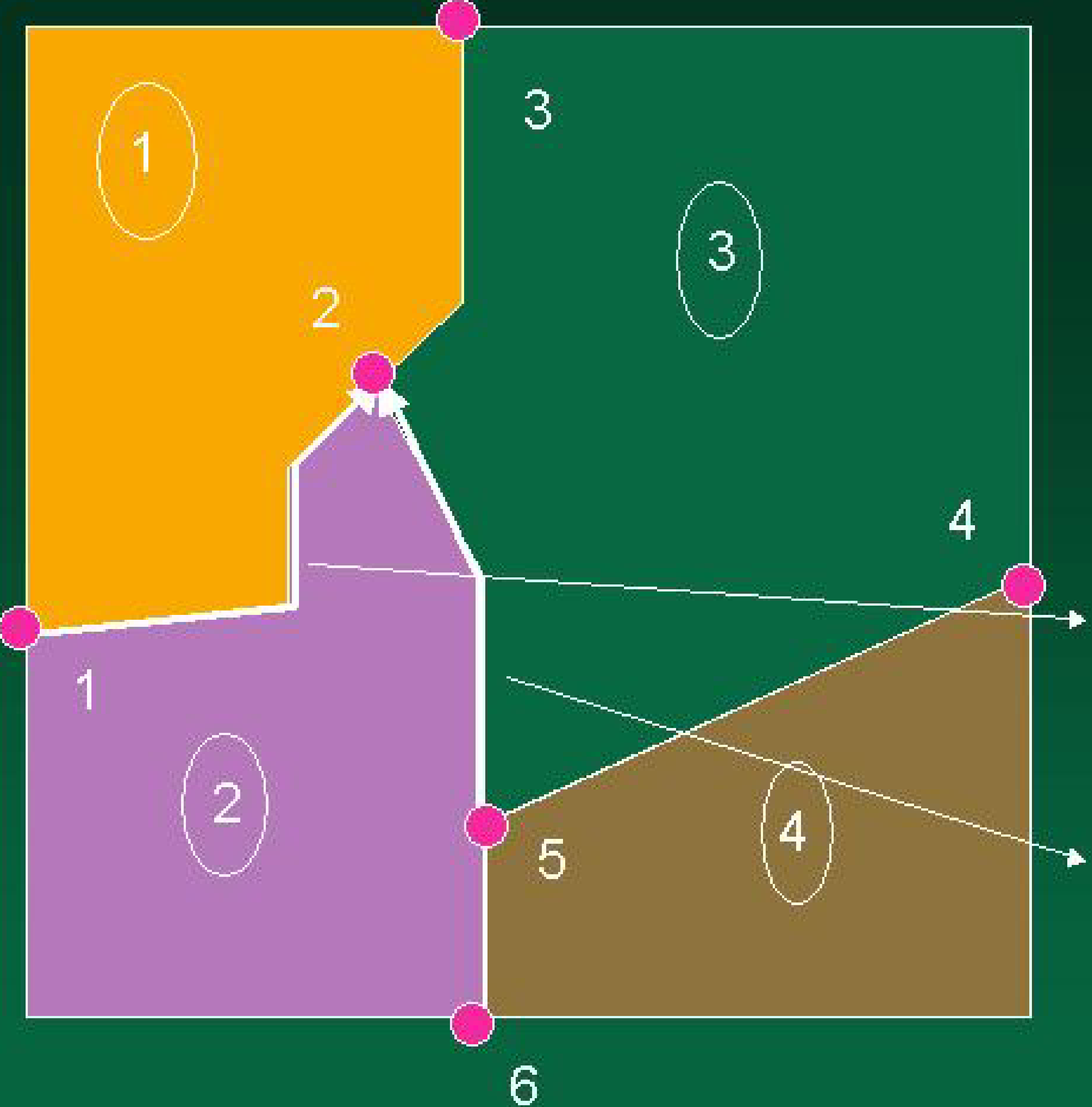


A	874	7	
B	95	8	
C	199	8	
D	767	5	

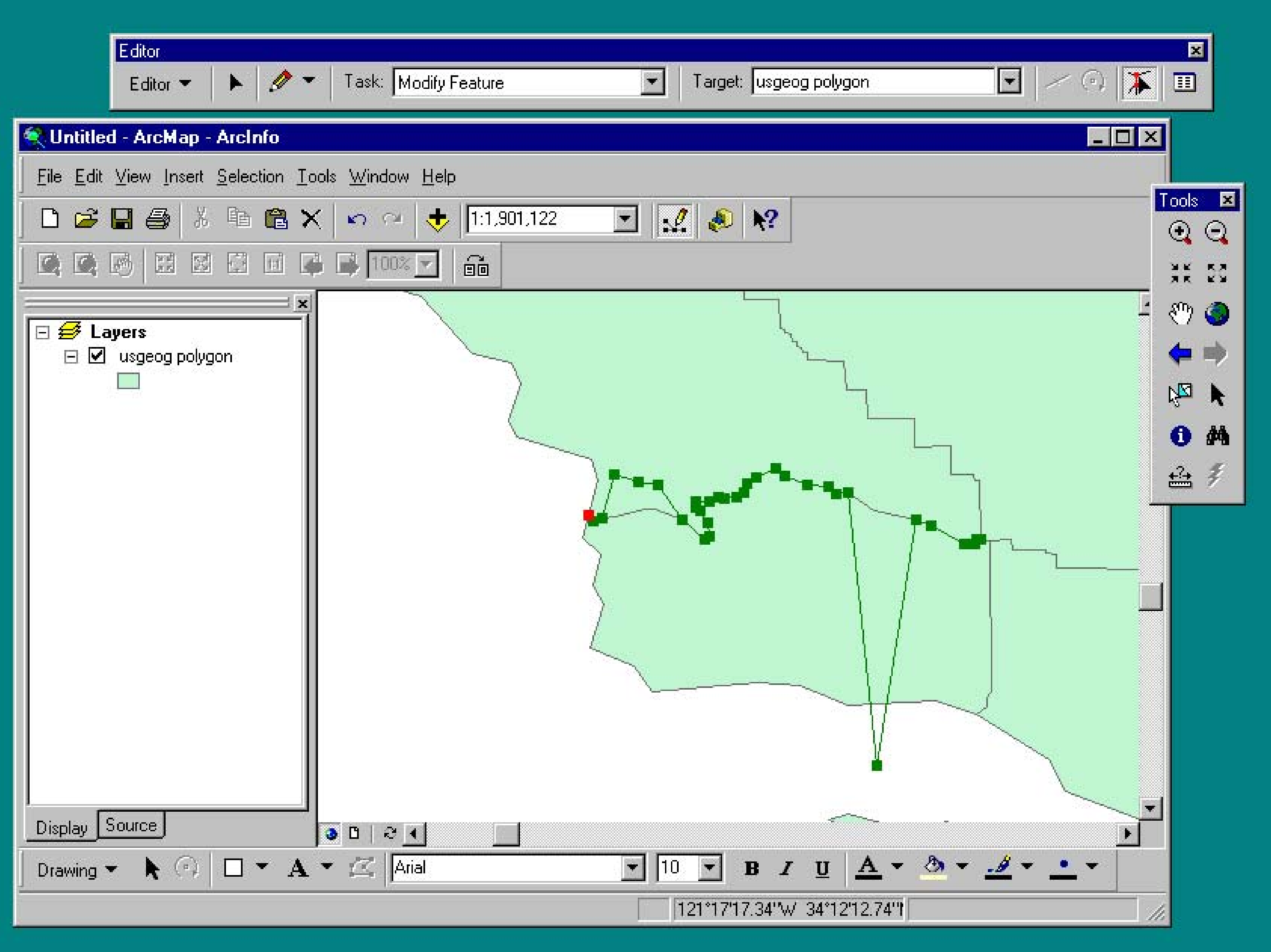


A	874		
B	95		
C	199		
D	767		

1						
2						
3						
4						
5						
6						
7						
8						
9						

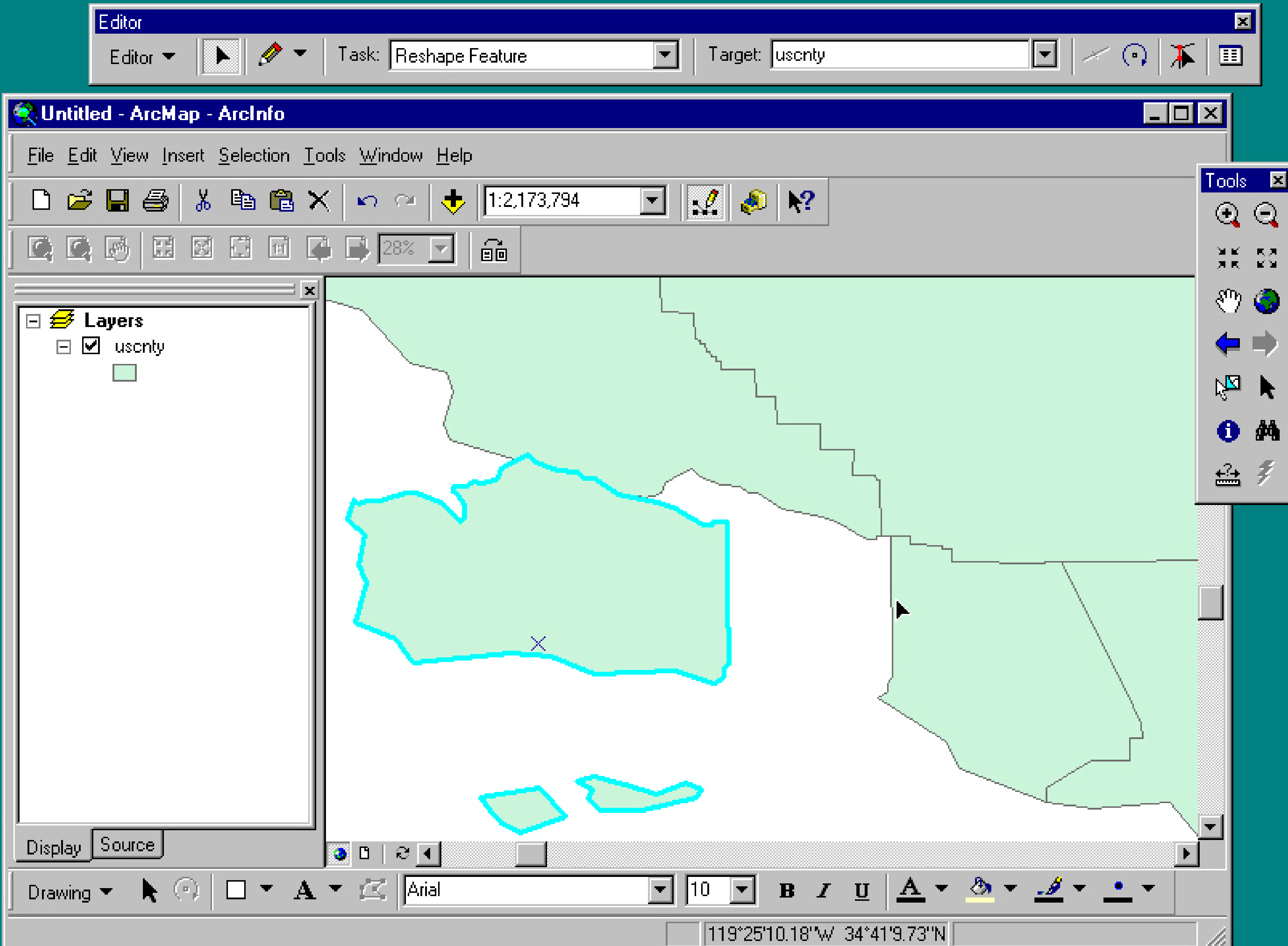


	FNODE	TNODE	LPOLY#	RPOLY#	
2	1	2	1	2	
5	5	2	2	3	


Editor
Editor
Task: Modify Feature
Target: usgeog polygon
Untitled - ArcMap - ArcInfo
File Edit View Insert Selection Tools Window Help
1:1,901,122
100%
Tools
Layers
usgeog polygon
Display
Source
Drawing
Arial
10
121°17'17.34"W 34°12'12.74"

Shapefiles

- A data set is a collection of all points, all lines, or all areas
  - lines as polylines
  - areas as polygons
- The features in a data set have associated attributes
  - stored in a table



Editor
Editor
Task: Reshape Feature
Target: uscnty
Untitled - ArcMap - ArcInfo
File Edit View Insert Selection Tools Window Help
1:2,173,794
28%
Tools
Layers
uscnty
Display
Source
Drawing
Arial
10
119°25'10.18"W 34°41'9.73"N

Real world

↓

GIS data model

↑

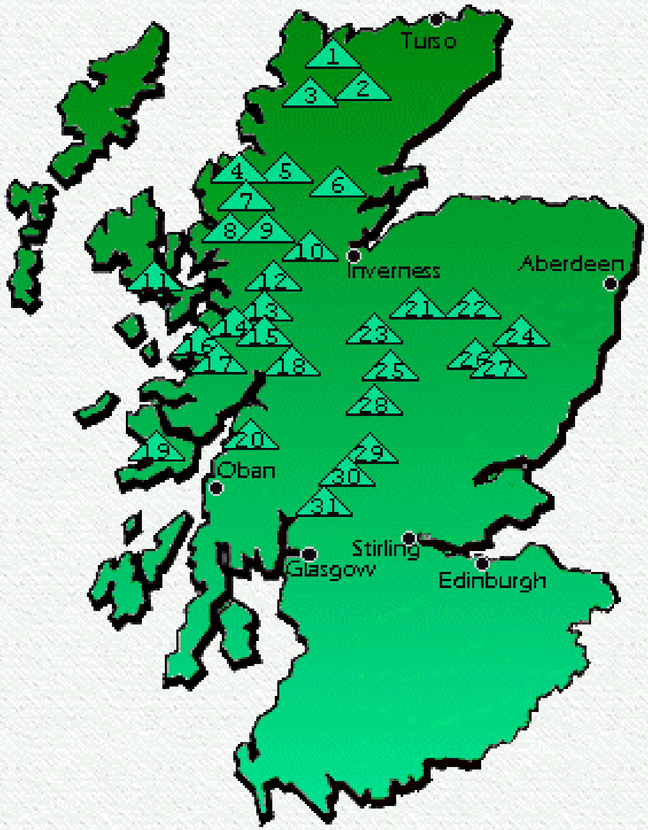
Industry-standard data model

The GIS data types

- Discrete geographic features
  - points, lines, areas
  - the contents of maps
  - with associated attributes
  - countable
  - conceived as tables with associated feature geometry
- ESRI shapefiles

Scottish Munros

1.. Ben Hope
2.. Ben Klibreck
3..Ben More Assynt
4..An Teallach
5..Seana Bhraigh
6..Ben Wyvis
7..Slioch
8..Sgorr Ruadh
9..Moruisg
10..Sgurr na Ruaidhe
11..Bia Bheinn
12..Sgurr na Lapaich
13..Ben Attow
14..The Saddle
15..Creag a' Mhaim
16..Ladhar Bheinn



17..Coireachan
18..Ben Nevis
19..Ben More
20..Ben Starav
21..Braeriach
22..Ben Avon
23..Meall Chuaich
24..Mt Keen
25..Beinn Dearg
26..Glas Maol
27..Driesh
28..Schiehallion
29..Ben Chonzie
30..Ben Lawers
30..Ben Challum
32..Ben Lomond


MIRAMAX CLASSICS
WIDESCREEN
HUGH GRANT
TARA FITZGERALD
AND COLM MEANEY
the Englishman
who went up a hill
but came down a mountain
"★★★★!"
—Sneak Previews
"TWO THUMBS UP!"
—Siskel & Ebert
DVD

Fields

- Geography as a collection of continuous variables
  - measured on nominal, ordinal, interval, ratio scales
  - vector fields of direction and magnitude
  - exactly one value per point
  - $z=f(\mathbf{x})$
  - population density, land ownership, zoning
- ESRI's coverage data model

What can we say about geographic information?

- Does it have general properties?
  - What's special about spatial?
- Design principles for GIS

Tobler's First Law

- "All things are related, but nearby things are more related than distant things"
  - W.R. Tobler, 1970. A computer movie simulating urban growth in the Detroit region. *Economic Geography* 46: 234-240
  - implies process as much as form
  - "nearby things are more similar than distant things"

Validity

- “Nearby things are less similar than distant things”
  - negative spatial autocorrelation
  - possible at certain scales
    - the checkerboard
    - retailing
  - but negative a/c at one scale requires positive a/c at other scales
  - smoothing processes dominate sharpening processes

Formalization

- Geostatistics
  - variogram, covariogram
  - measuring how similarity decreases with distance
  - parameters vary by phenomenon
    - does this make TFL less of a law?

Utility

- Representation
  - GI is reducible to statements of the form <**x**,**z**>
  - the atomic form of GI is unmanageable, encountered only in point samples
  - all other GI data models assume TFL
- Spatial interpolation
  - IDW and Kriging implement TFL

If TFL weren’t true

- GIS would be impossible
  - a point sample is useful only with interpolation
- Life would be impossible

Expanding the horizons

- Other spaces
  - are there spaces for which TFL is not true?
  - digits of π
  - genome
- Other laws of GIScience

Candidate laws

- All important places are at the corners of four map sheets
- Montello and Fabrikant, “The First Law of Cognitive Geography”
  - “People think closer things are more similar”

A second (first) law

- TFL describes a second-order effect
  - properties of places taken two at a time
  - a law of spatial dependence
  - is there a law of places taken one at a time?
- Spatial heterogeneity
  - non-stationarity
  - uncontrolled variance

Corollaries of the second law

- There is no average place on the Earth's surface
- Sampling is problematic
  - one must visit or map all of it to understand its full complexity
- Results depend explicitly on the bounds of the study
- The Noah effect
  - there is a finite probability of an event of any magnitude
  - to observe an event of a given magnitude it is simply necessary to wait long enough

A GIScientist's Noah effect

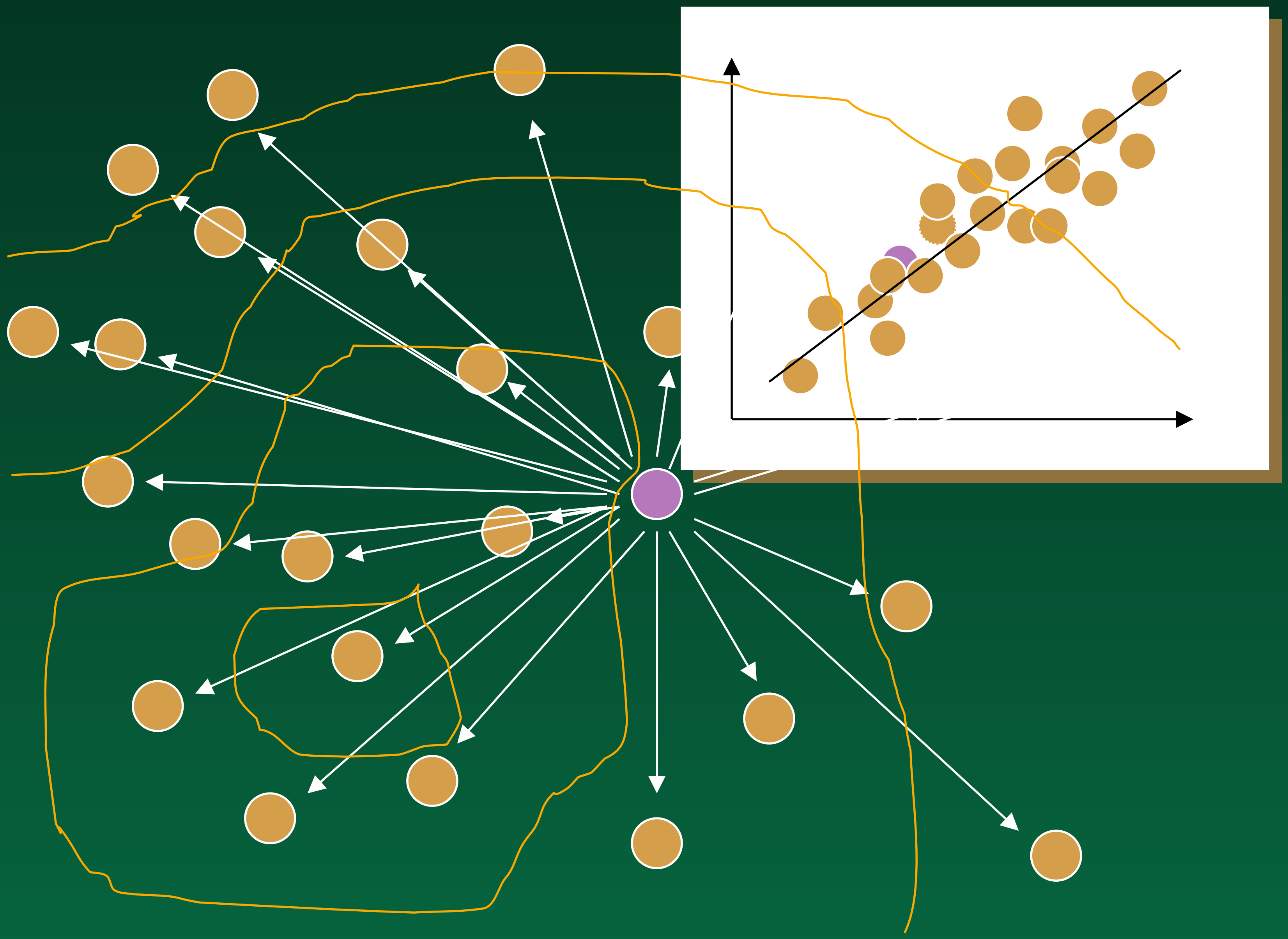
- The Eden effect
  - El Dorado
  - to find a feature of any magnitude it is sufficient to look far enough
    - but unlike time, the Earth's surface is limited
- The Pareto distribution or rank-size rule
  - plot log rank against log size
  - a model of the extreme upper tail of distributions
  - fits well to the world's largest:
    - cities by population
    - lakes by area
  - but not mountains by elevation

Practical implications of the second law

- A state is not a sample of the nation
  - a country is not a sample of the world
- Classification schemes will differ when devised by local jurisdictions
- Figures of the Earth will differ when devised by local surveying agencies
- Global standards will always compete with local standards

Implications for analysis

- Strong argument for place-based analysis, local statistics, geographically weighted regression
  - a middle ground in the nomothetic/idiographic debate



Possible corollary of the heterogeneity law

- For every conceivable pattern in two (three) dimensions there exists an instance on the Earth's surface
  - for every GIS algorithm/indexing scheme/data model there exists a data set for which that algorithm/indexing scheme/data model is optimal
  - "There are more things in Heaven and earth, Horatio, than are dealt with in your philosophy"

Scale

- A fundamental property of any geographic representation
  - defining what is included, what is left out
    - 1:24,000
    - 1:100,000
    - 1:2,000,000
- Two characteristics:
  - spatial resolution
  - spatial extent
- Conflict over "large" and "small"

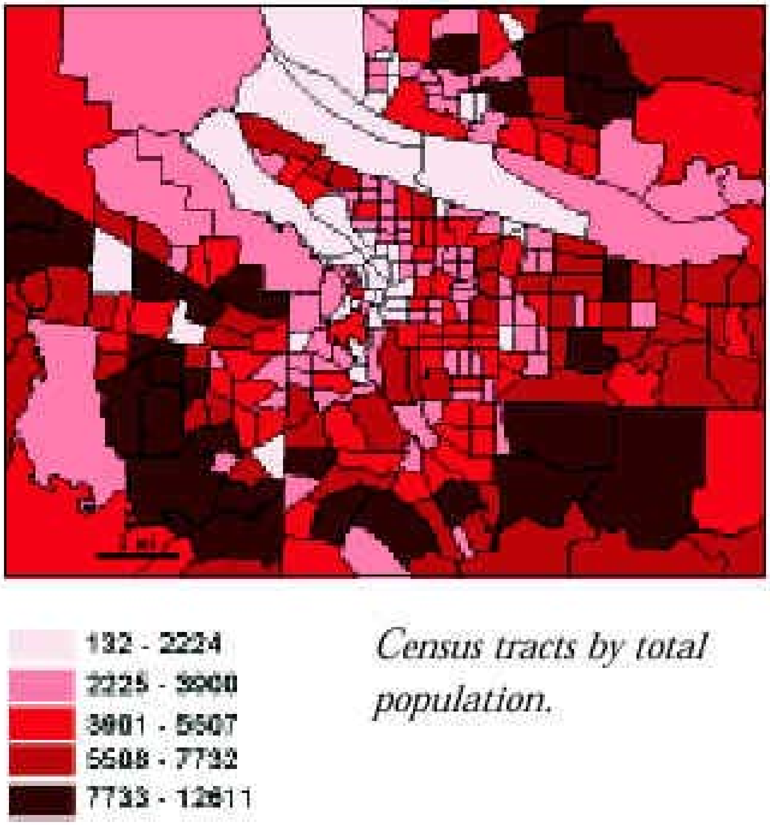
Scale and z

- Many properties are scale-dependent
  - slope
  - land cover class
  - soil class
  - travel behavior
    - daily
    - annual
    - lifetime
- What does scale mean for digital data?

Types of attributes

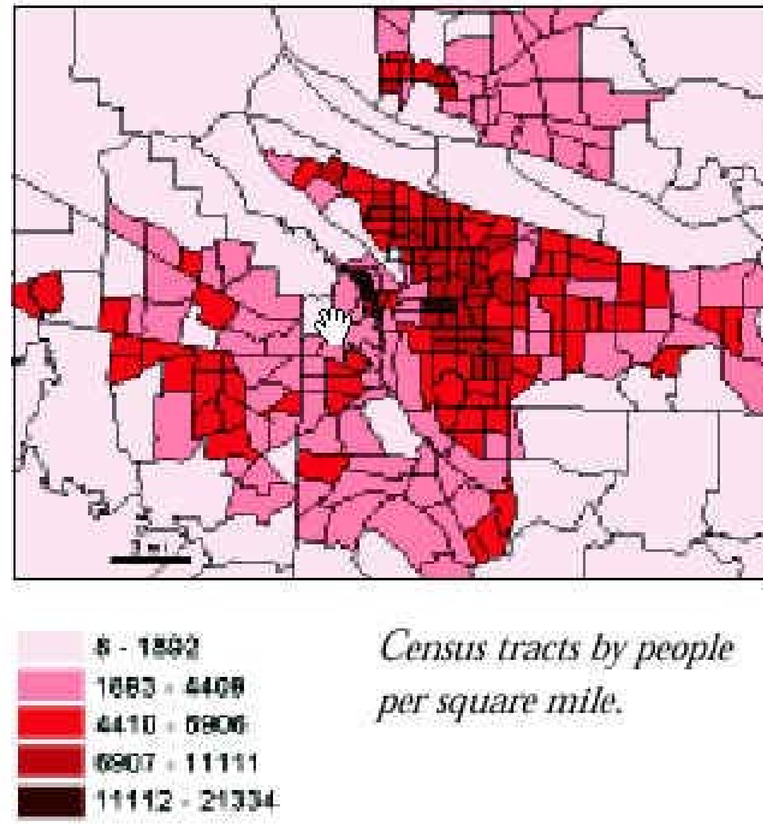
- Nominal
- Ordinal
- Interval, ratio
- Cyclic
- Vector
- Scaling properties
  - spatially intensive
  - spatially extensive

If you want to know approximately how many people each census tract has, map total population.



Census tracts by total population.

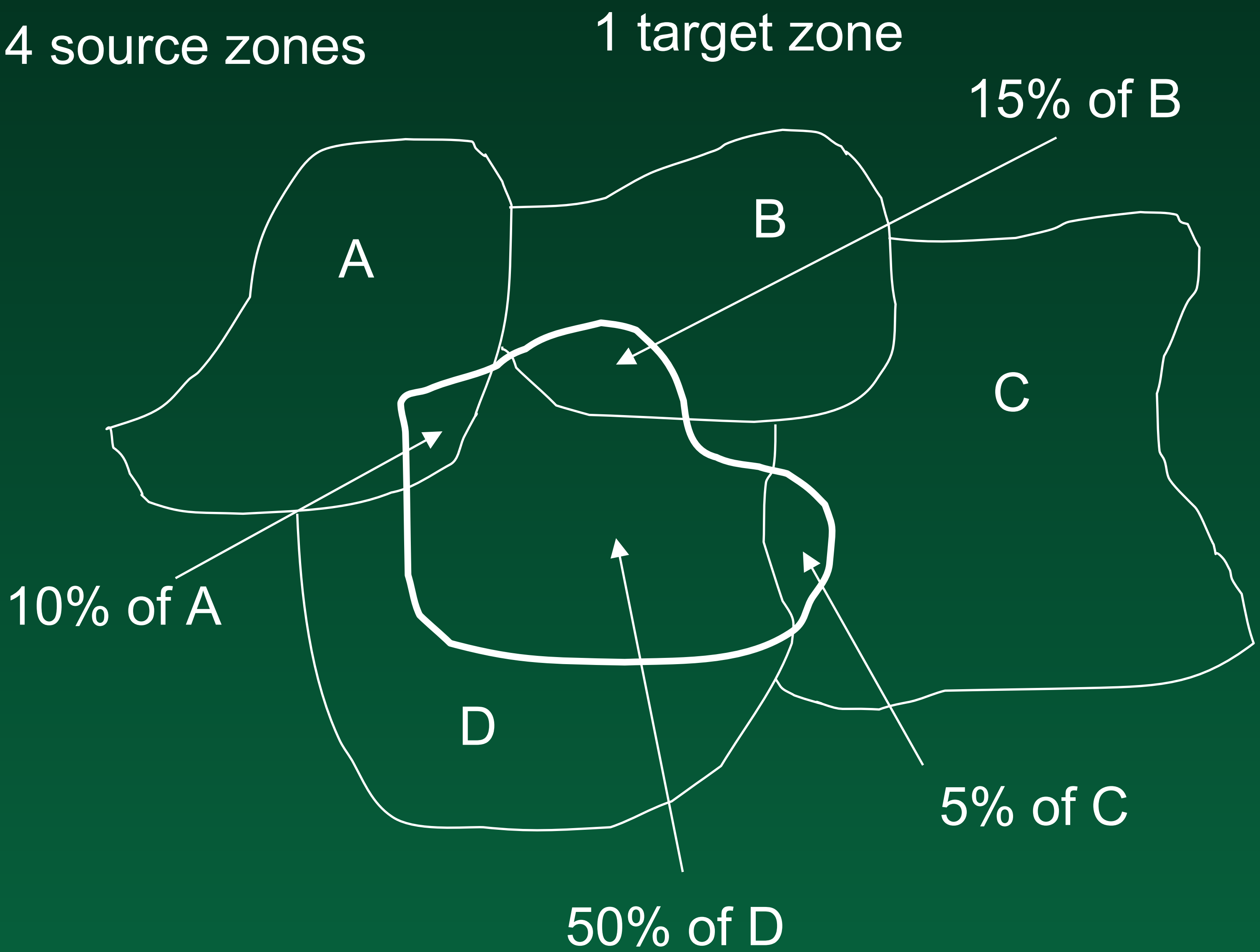
If you want to know where most of the people are concentrated, map population density.



Census tracts by people per square mile.

Spatial support

- The set of features used to characterize spatial distribution
  - sample points
  - reporting zones
- Spatial interpolation
  - guessing values at unmeasured/unobserved locations
  - areal interpolation



$$Pop_{TARGET} = 0.10\,Pop_A + 0.15\,Pop_B + 0.05\,Pop_C + 0.50\,Pop_D$$

Reasoning across support and scale

- The ecological fallacy
  - reasoning from aggregate to individual
  - Gary King
- The Modifiable Areal Unit Problem
  - reasoning across alternative support
  - an artifact of incomplete problem conceptualization

Summary

- Spatial thinking is powerful
  - providing context
  - support for knowledge discovery
- Geographic information is a unique information type
- Its representation in digital form involves many options
  - the GIS data models
- It has general properties
  - that are useful in system design

GIS education: past, present, future

- Thirty years of evolution
  - my own teaching
- Continuing issues
- A plan for the future

GIS education in 1974

- Fragmented
  - cartography, remote sensing, geographical analysis, planning
  - Tomlinson’s conferences (1970, 1972)
  - Harvard workshops (1967)
  - nobody talked about teaching
- Government systems
  - Canada Geographic Information System
  - Census Bureau and DIME

University labs in 1974

- Harvard Laboratory for Computer Graphics and Spatial Analysis
  - SYMAP, CALFORM, SYMVU
  - William Warntz and theoretical geography
- Experimental Cartography Unit in London

What to teach?

- No software
  - home-made
  - not integrated
  - Harvard lab
  - no COTS
- No principles
- No applications
- No text

Geography 300

- University of Western Ontario, 1975
  – Ross Newkirk
- Texts:
  – Bruce MacDougall, *Computer Programming for Spatial Problems* (1976)
  – Richard Baxter, *Computer and Statistical Techniques for Planners* (1976)
  – John Davis and Michael McCullagh, *Display and Analysis of Spatial Data* (1975)
- Software:
  – PlusX, Surface II, SYMAP, CALFORM

Principles circa 1975

- Algorithms
  - point in polygon
- Data structures
  - arcs and nodes
- Accuracy
- Analysis
  - overlay
  - location-allocation
  - political districting
  - spatial interpolation
  - corridor location

Significant milestones

- Research community
  - Topological data structures conference, Boston, 1977
- COTS software
  - 1982
- Peter Burrough's text *Principles of Geographical Information Systems*
  - 1985
- NCGIA Core Curriculum
  - 1989
  - http://www.ncgia.ucsb.edu, Education, Core Curricula

The three-course sequence (UCSB)

- 176A: Introduction to GIS
  - Keith Clarke, *Getting Started with GIS*
- 176B: Technical Issues in GIS
  - Mike Goodchild, *Geographic Information Systems and Science*
- 176C: Application Issues in GIS
  - Student projects

GIS education today

- Textbooks
- Courses and programs in thousands of institutions
- Distance learning
  - Virtual Campus, UNIGIS, Penn State, University of London
  - Intergraph's Online Education Training Program
    - http://imgs.intergraph.com/training/online.asp
- University Consortium for Geographic Information Science
- Alternative COTS solutions

Consensus on the principles

- Representation
- Data models
- Data structures
- Algorithms
- Visualization
- Analysis
- Uncertainty
- Metadata
- Data sharing
- Projections and coordinate systems
- Geodesy
- GIS design and implementation
- Data management
- Spatial decision support
- Dynamic modeling
- User interfaces

Teaching more of...

- Database design and management
  - relational
  - object-oriented
  - CASE tools
- Internet services
  - metadata
  - GIServices
  - location-based services, field GIS
- Social context
  - privacy, surveillance
  - costs and benefits

Teaching less of...

- Algorithms
- Analysis
- Programming
- Operating systems

The technology of teaching

- 1975: Blackboard, chalk, handouts
- 1996: WWW notes and overheads
- 1999: Online HTML
  - no failures in 6 years
- 2002: Online HTML and GIS demos
  - 2 failures in 3 years

Online

- http://www.geog.ucsb.edu, Faculty, Goodchild
- Linked lab exercises

Continuing issues

- Balance of pedagogic style
  - lecture or hands-on lab?
  - adaptation to individual learning style
- Balance of training and education
  - education in the persistent principles
  - training in today’s technology
- Balance of computer science and geography
  - focus on computers
  - focus on the real world and its digital representation

Continuing issues (2)

- GIS transparent or black box?
  - doing it by hand first
  - understanding what happens inside
  - educating the skeptical user
  - demanding better from the vendor
  - the rules of scientific reporting
- Education for the workplace or for research?

Where next?

- From geographic to spatial
  - geographic, spatial, geospatial
  - geographic data, spatial analysis
- The importance of spatial thinking
  - in all sciences
  - in life
  - a picture is worth a thousand words
  - spatial information is more easily comprehended

Generalizing to spatial

- Astronomy
  – information about the cosmos
- Bioinformatics
  – the human genome
  – "BLACKSBURG, VA -- The merits of linking two fields seemingly as disparate as geographic information systems (GIS) and bioinformatics might not seem obvious, but Virginia Tech's recent symposium linking the two and its roster of renowned participants from both fields raised expectations as well as eyebrows in national technology circles." (June, 2001)

From geography to STEM

- Traditional support for geographic research
  - expanded to resource management, geology, criminology, history, ecology, ...
- What about all of science, technology, engineering, math (STEM) education?

GIS as key to STEM education

- Student-centered learning
- Motivation

  a pathway to teaching
- Spatial thinking

  visualization and communication important in all sciences
- Principles of the scientific method
- Workplace skills
- Integration of technology in science and math education
- GIS for good citizenship

Vertical integration

- Upper division undergraduate
- Lower division, 2-year college
- K-12
- Non-traditional learners
- What to teach when?
  - tied development of conceptual and cognitive skills
- Current National Research Council study *Support for Thinking Spatially: GIScience in the K-12 Curriculum*

The proposed study will review knowledge about the teaching and learning of geographic information systems (GIS) and geographic information science (GIScience) in K-12 education. It will address two questions: (1) how can current versions of GIS and GIScience be incorporated into standards-based instruction in knowledge domains across the curriculum, and (2) how can cognitive developmental and educational theory be used to design age-appropriate versions of GIS and age-appropriate GIScience curricula; (3) what are the nature and character of spatial thinking: what is it, why do we need to know about it, and what do we need to know about it; and (4) how does the capacity for spatial thinking develop and how might it be fostered systematically by education and training? The review will develop application guidelines and research and development strategies. It will provide short-term guidance to incorporate GIS and GIScience into American schools and long-term research strategies to improve the design of GIS and reshape the teaching and learning of GIS and GIScience.

www.nas.edu, Earth Science, Board on Earth Science and Resources, Current Studies

A GIS for thinking spatially

- Capacity to:
  - Spatialize data sets
  - Visualize
  - Perform transformations and analysis

Design criteria

- Supportive of enquiry process
- Educationally appropriate
- Accessible to all learners
- Quick to learn and use
- Learning across disciplines
- Range of contexts
- Customizable
- Robust and realistic

Child of 10 or concert pianist?

- Conventions that complicate
  - scale
  - projections
  - language
  - TLAs
    - DEM, DOQ, DLG, DRG, DCW, DTM…

GIS adoption

- Approaching 100% in universities and colleges
- Early adopters in K-12
  - teachers, schools, districts
- Moving GIS beyond the early adopters
  - disadvantaged areas
  - learners with disabilities
  - gender issues

Infrastructure for GIS educators

- Physical meetings
  - national or regional
- Web resources
  - software, data, applications
  - online expertise
- P2P resource sharing
- Classroom internships for college GIS students
  - NSF’s GK-12 program
- GIS in credential programs
  - Schools of Education

Conclusion

- 30 years of progress
- A vision for the next level