What's Special about Spatial?

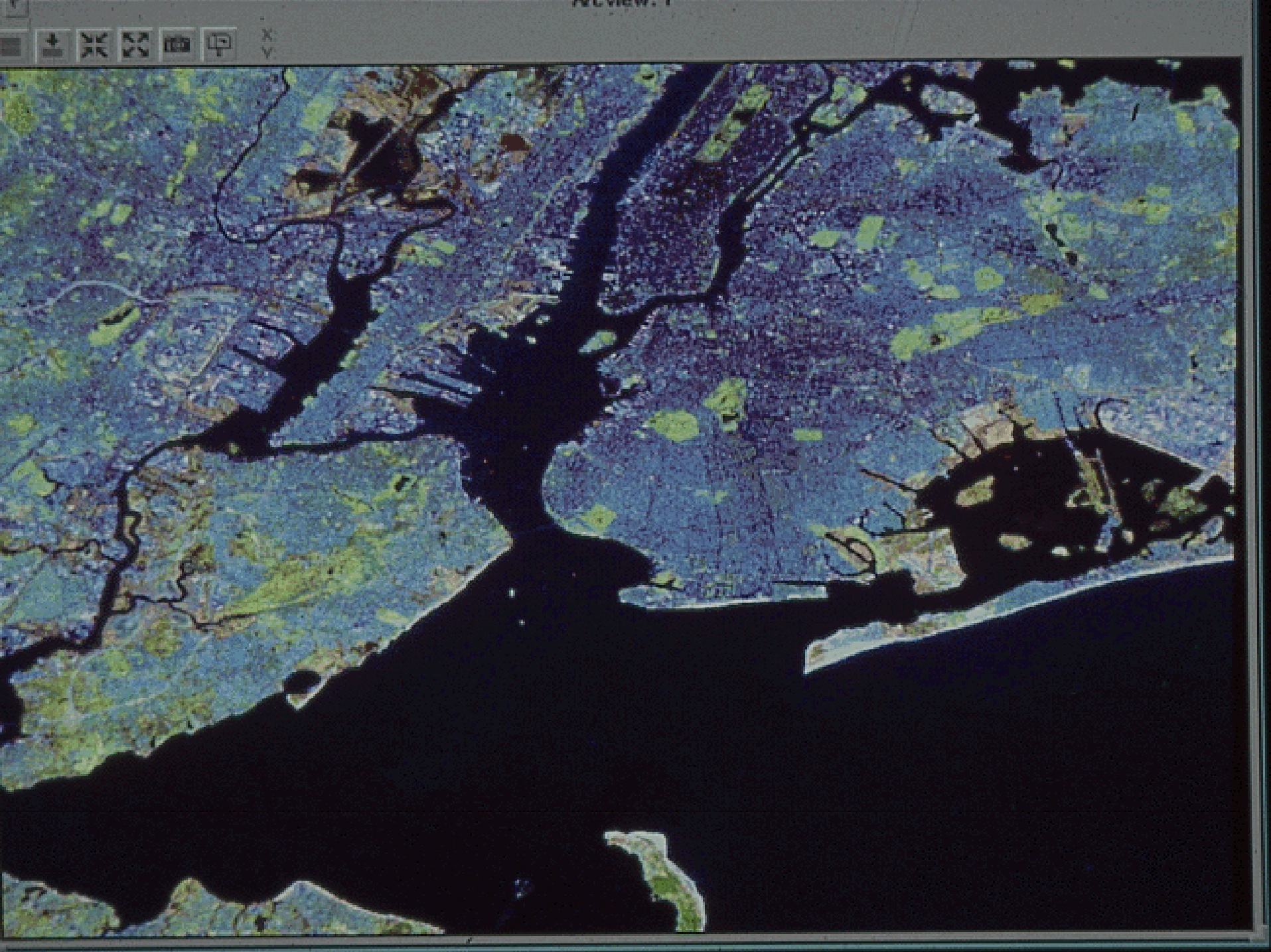
Michael F. Goodchild
University of California
Santa Barbara

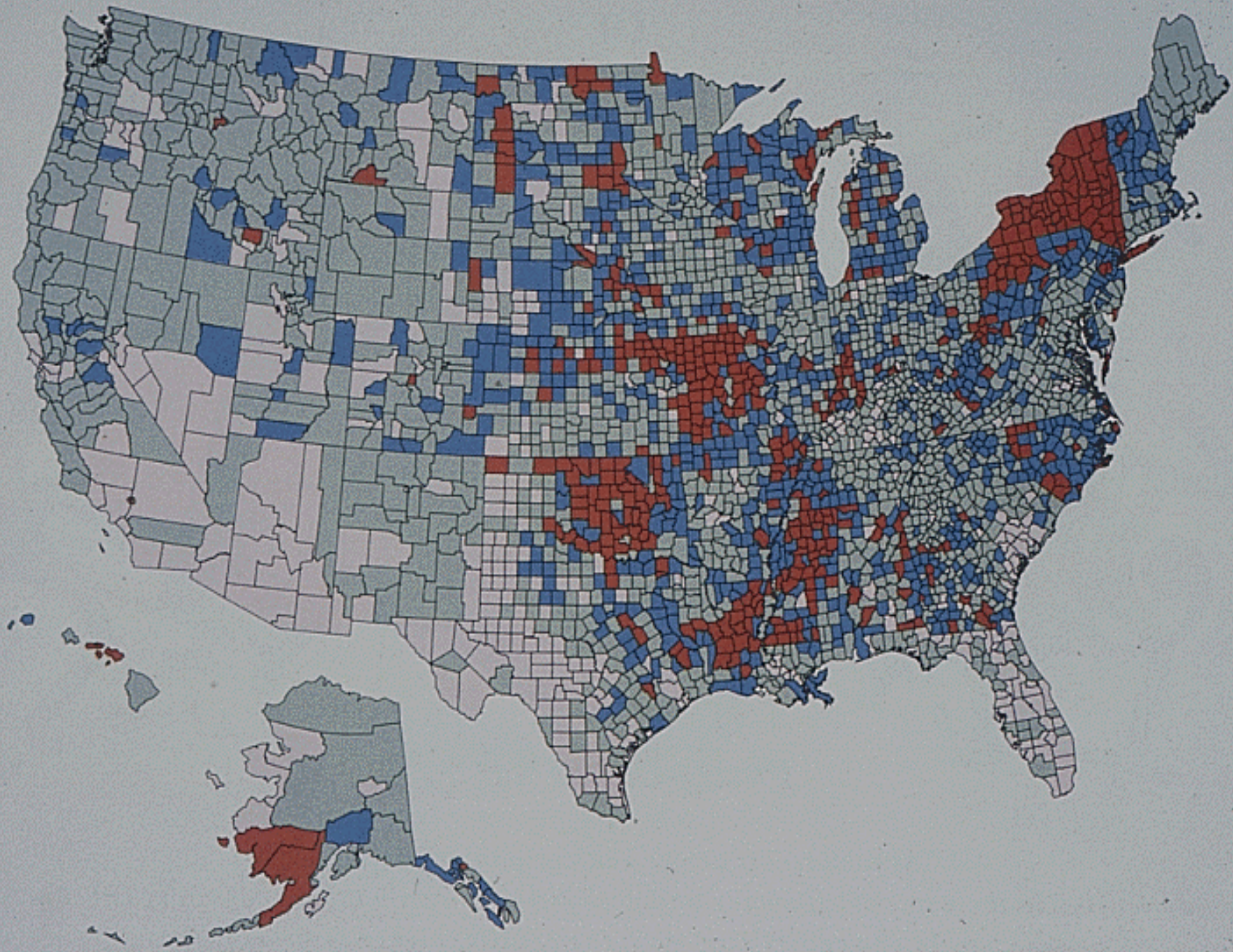
Outline

- GIS
- The Center for Spatially Integrated Social Science (CSISS)
- The ***W*** matrix

Definitions

- Spatial data
  - information about phenomena organized in a spatial frame
  - the geographic frame
- Methods applied to spatial data that
  - add value
  - reveal patterns and anomalies
  - support decisions





The role of the GIS

- The infrastructure for handling data types
  - to spatial data as Excel is to tables, as S-Plus is to statistical data, as Word is to text
  - spatial data or geographic data?
  - the housekeeper
  - the editor
- The visualization tool

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

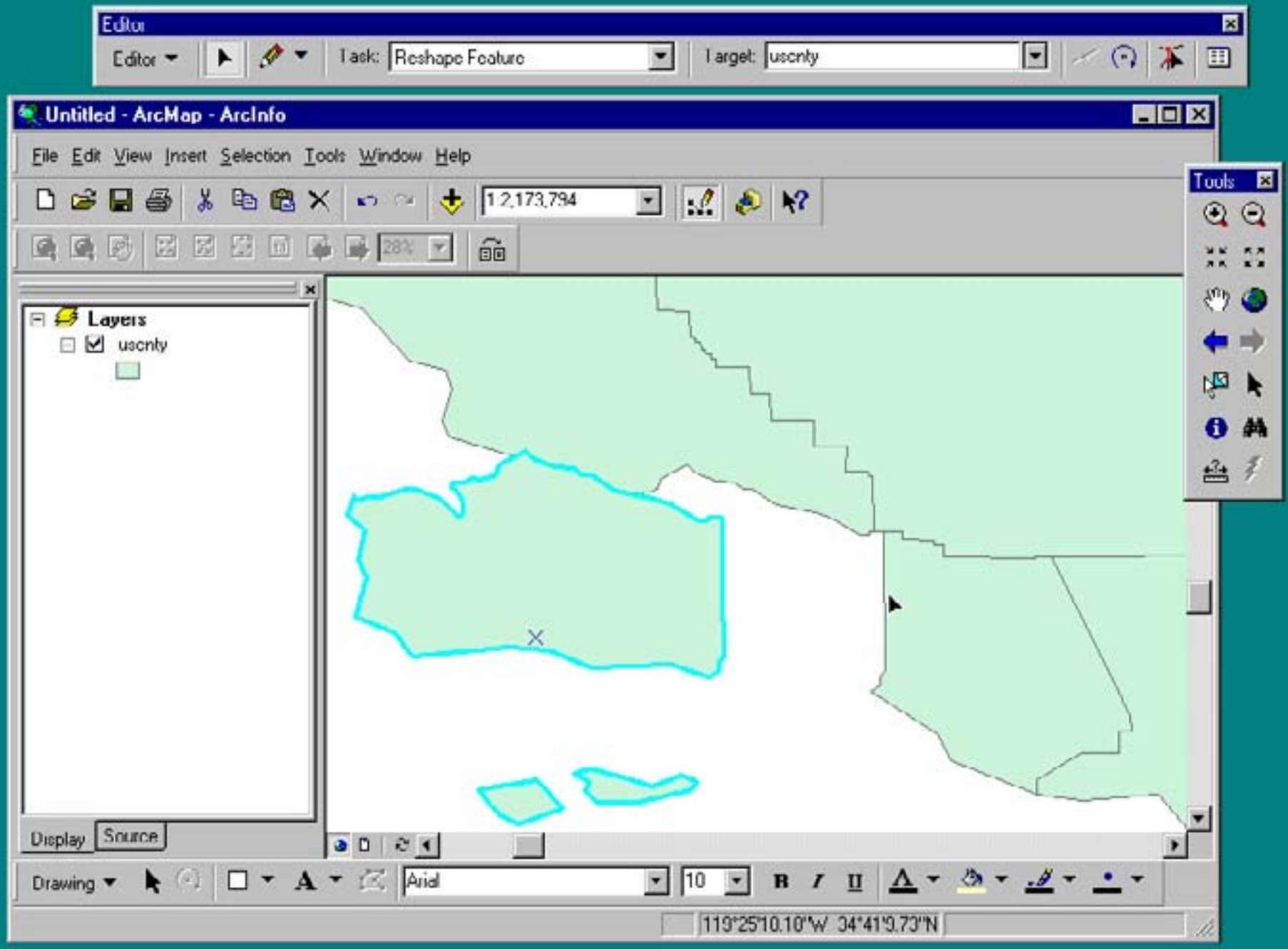
Scottish Munros

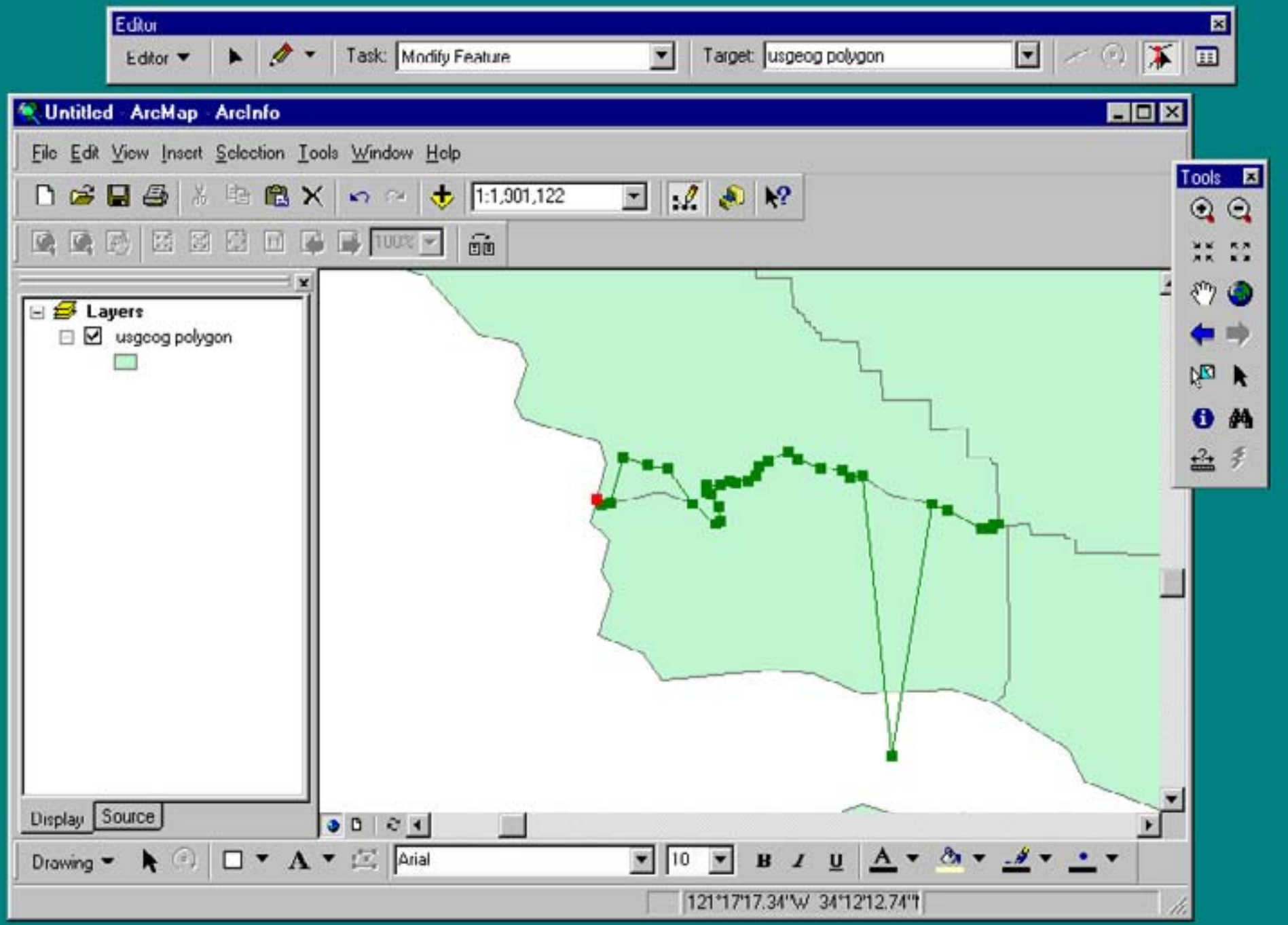
1.. Ben Hope
2.. Ben Klibreck
3..Ben More Assynt
4..An Teallach
5..Seana Bhraigh
6..Ben Wyvis
7..Slioch
8..Sgorr Ruadh
9..Moruisg
10..Sgurr na Ruaidhe
11..Bia Bheinn
12..Sgurr na Lapaich
13..Ben Attow
14..The Saddle
15..Creag a' Mhaim
16..Ladhar Bheinn

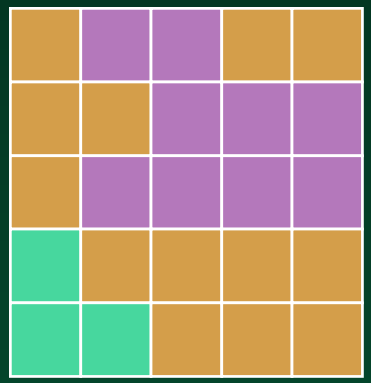
17..Coireachan
18..Ben Nevis
19..Ben More
20..Ben Starav
21..Braeriach
22..Ben Avon
23..Meall Chuaich
24..Mt Keen
25..Deinn Dearg
26..Glas Maol
27..Driesh
28..Schiehallion
29..Ben Chonzie
30..Ben Lawers
30..Ben Challum
32..Ben Lomond

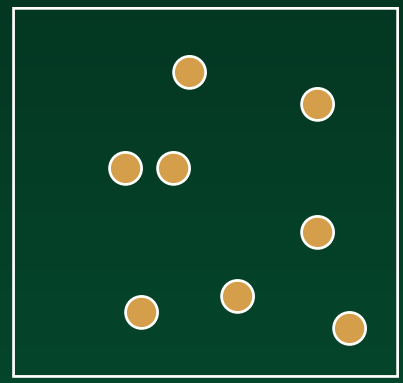
Fields

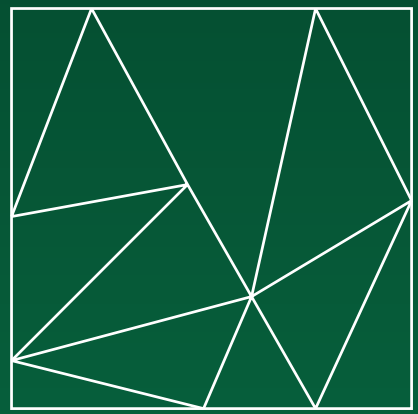
- Geography as a collection of continuous variables
  - measured on nominal, ordinal, interval, ratio scales
  - vector fields of direction and magnitude
  - exactly one value per point
  - $z=f(\mathbf{x})$
  - population density, land ownership, zoning


Editor
Editor
Task:
Reshape Feature
Target:
uscnty
Untitled - ArcMap - ArcInfo
File Edit View Insert Selection Tools Window Help
1:2,173,794
Tools
Layers
uscnty
Display
Source
Drawing
Arial
10
119°25'10.10"W 34°41'0.73"N

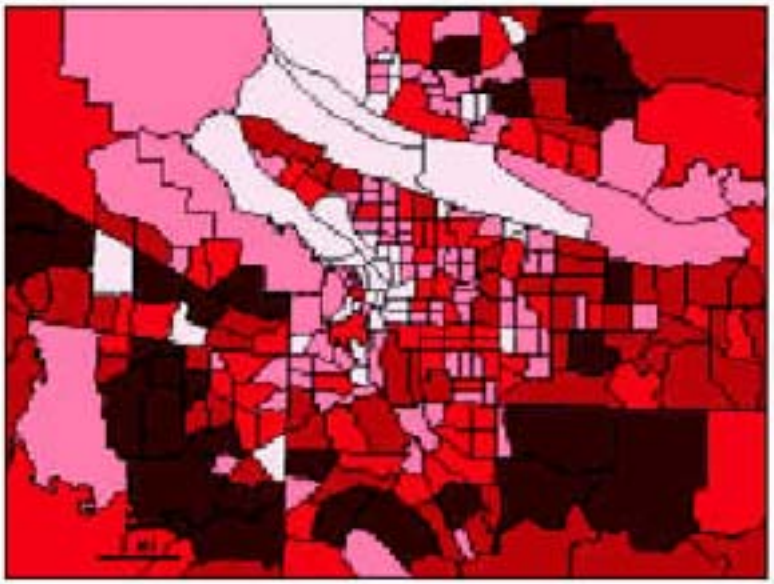

Editor
Editor
Task: Modify Feature
Target: usgeog polygon
Untitled - ArcMap - ArcInfo
File Edit View Insert Selection Tools Window Help
1:1,901,122
Tools
Layers
usgeog polygon
Display
Source
Drawing
Arial
10
121°17'17.34"W 34°12'12.74"N







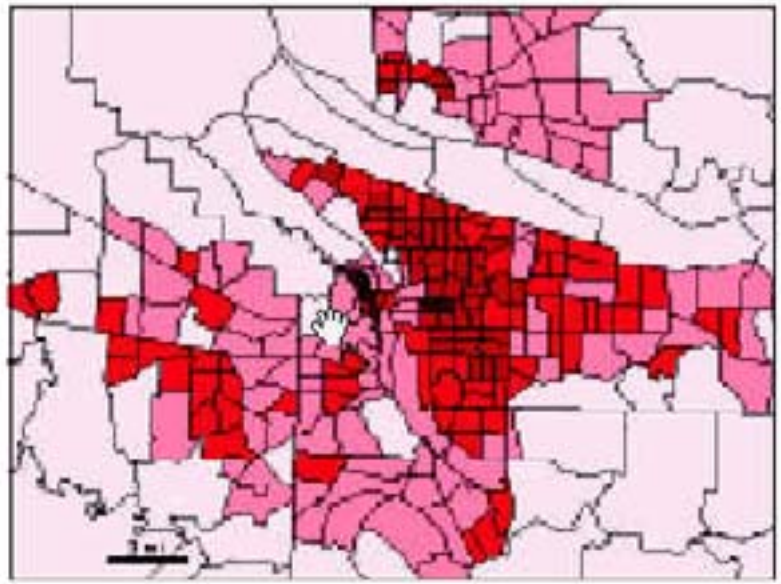
If you want to know approximately how many people each census tract has, map total population.



132 - 2224
2225 - 3900
3901 - 5507
5508 - 7732
7733 - 12611

Census tracts by total population.

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



8 - 1882
1883 - 4409
4410 - 6906
6907 - 11111
11112 - 21334

Census tracts by people per square mile.

Taxonomies of spatial analysis

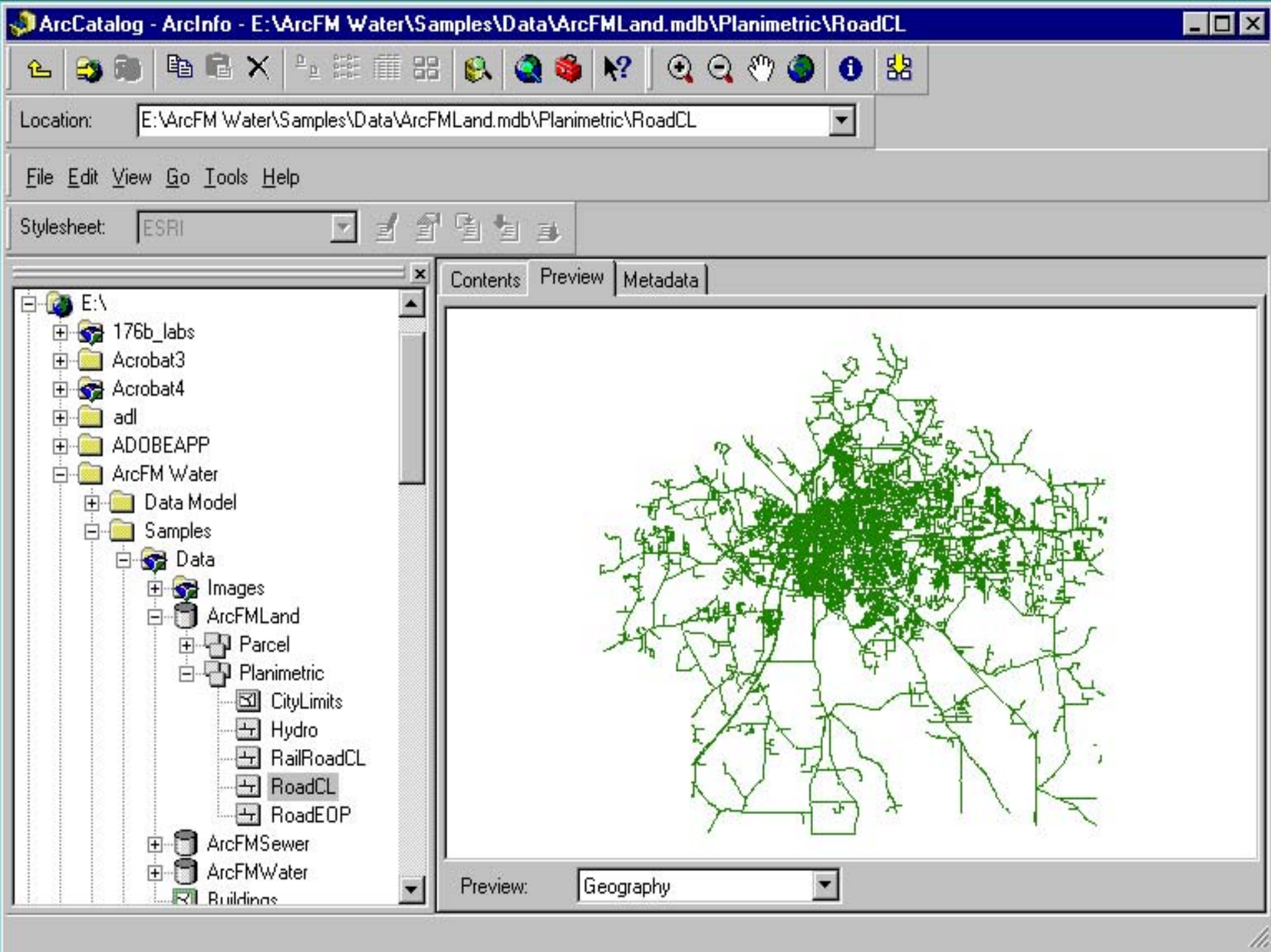
- Thousands of methods
  - every one a command, menu item, icon, …
- Based on data type
  - point pattern analysis
  - area (polygon) analysis
  - analysis of interactions
  - Bailey and Gatrell, Haining, Unwin

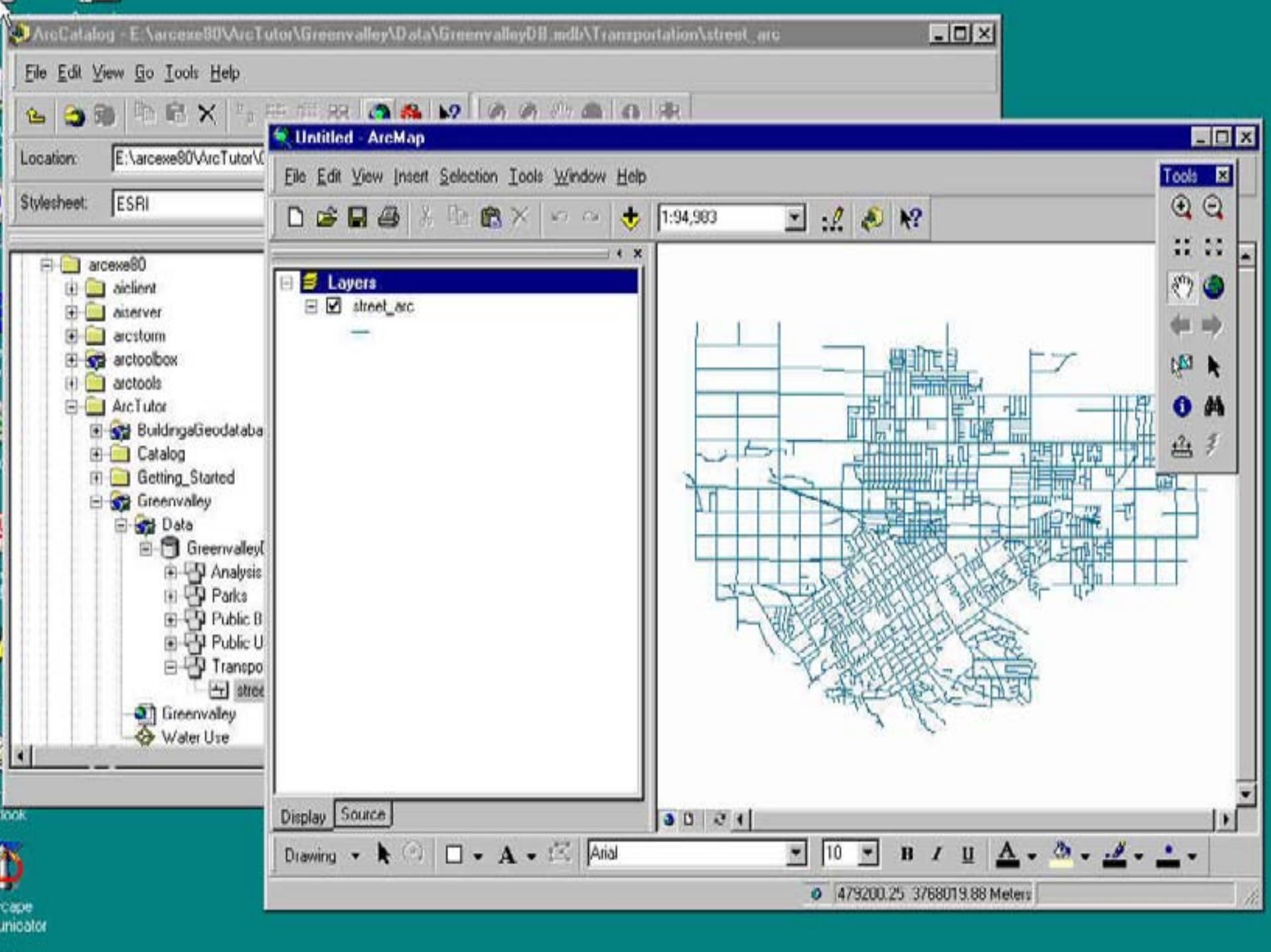
A six-way conceptual classification

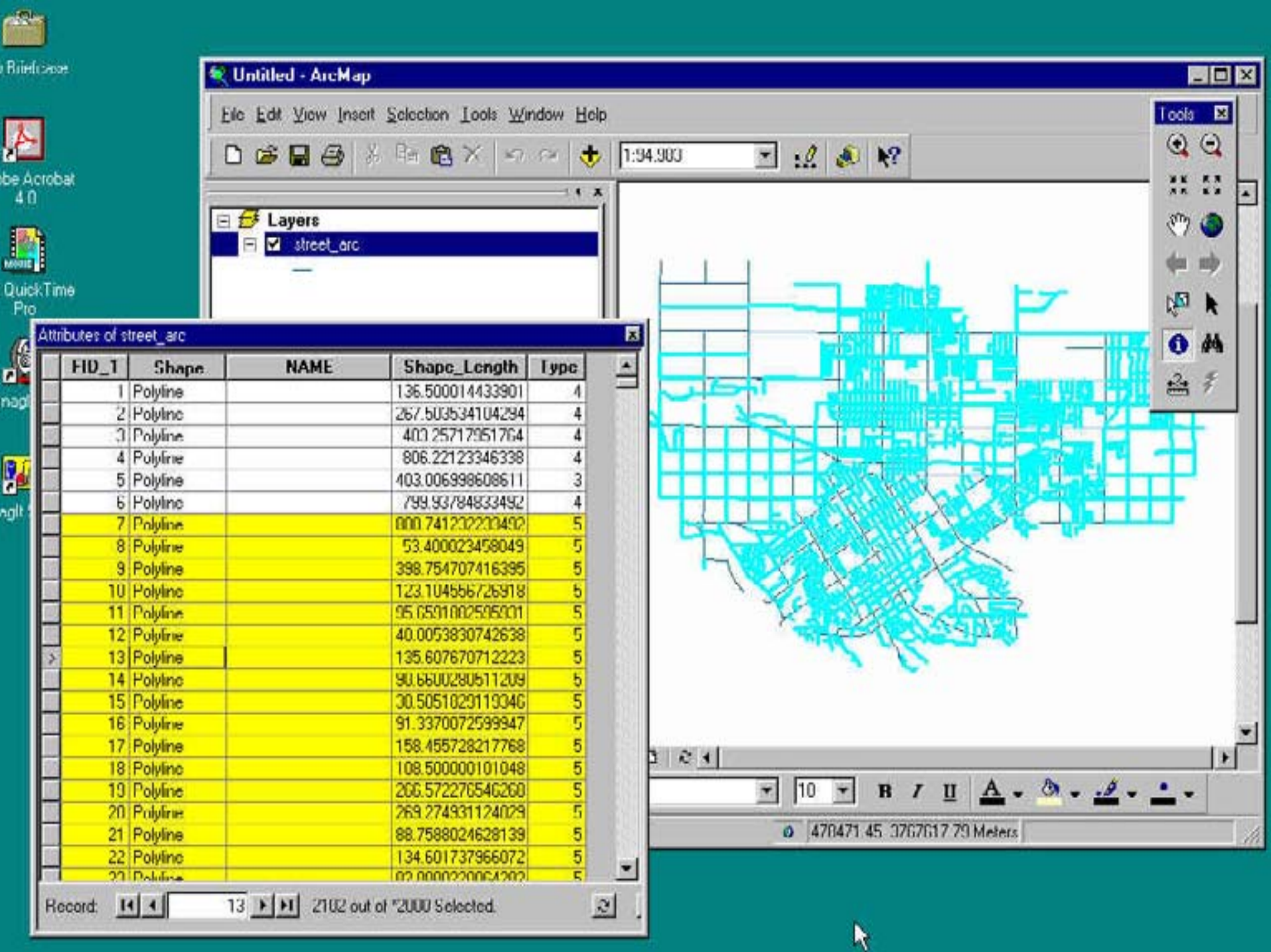
- Query and reasoning
- Measurement
- Transformation
- Descriptive summary
- Optimization
- Hypothesis testing

Queries and reasoning

- Real-time answers to geographic questions
  - Where is…?
  - What is this?
  - How do I get from here to here?
- Based on alternative views of a database

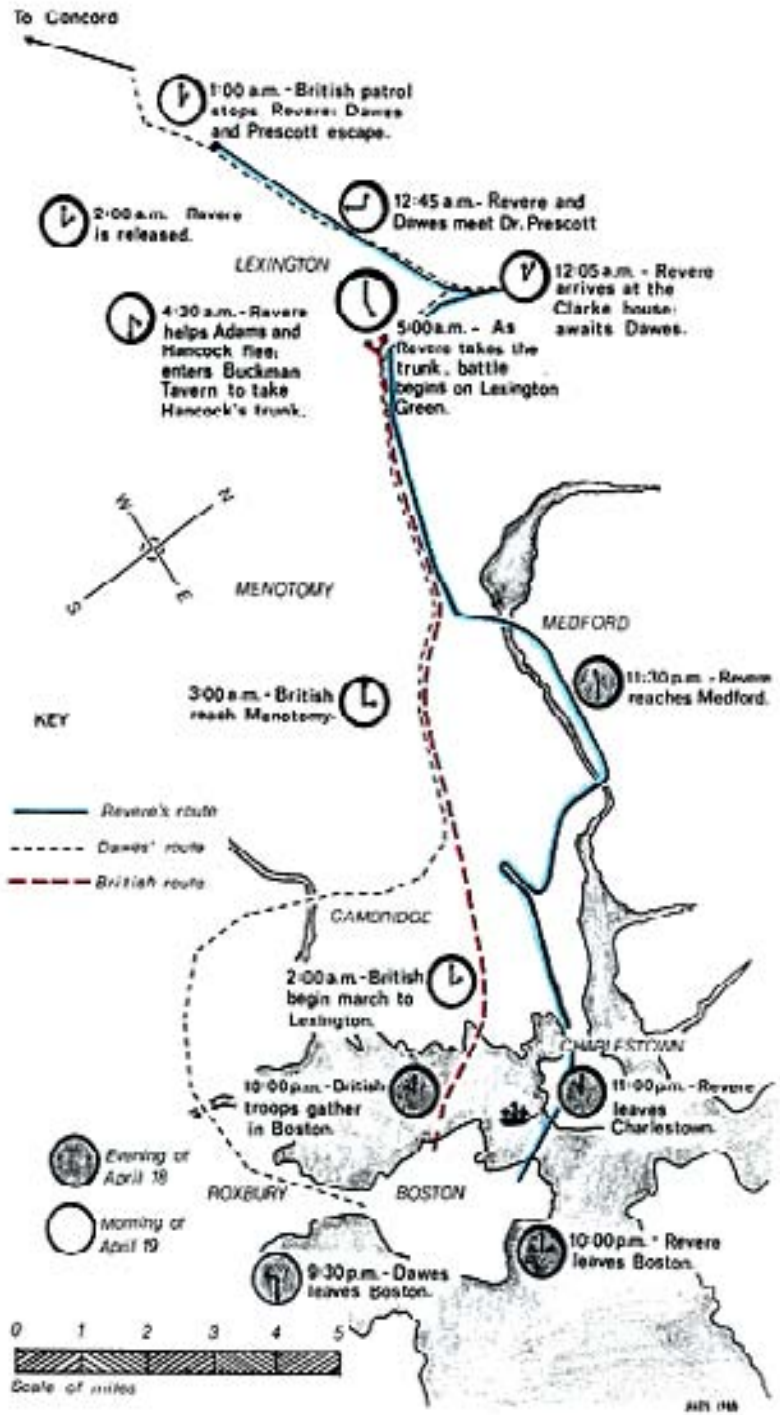

ArcCatalog - ArcInfo - E:\ArcFM Water\Samples\Data\ArcFMLand.mdb\Planimetric\RoadCL
Location:
E:\ArcFM Water\Samples\Data\ArcFMLand.mdb\Planimetric\RoadCL
File Edit View Go Tools Help
Stylesheet:
ESRI
E:\
176b_labs
Acrobat3
Acrobat4
adl
ADOBEAPP
ArcFM Water
Data Model
Samples
Data
Images
ArcFMLand
Parcel
Planimetric
CityLimits
Hydro
RailRoadCL
RoadCL
RoadEOP
ArcFMSewer
ArcFMWater
Buildings
Contents
Preview
Metadata
Preview:
Geography


ArcCatalog - E:\arcexe80\ArcTutor\Greenvalley\Data\GreenvalleyDB.mdb\Transportation\street_arc
File Edit View Go Tools Help
Location: E:\arcexe80\ArcTutor\
Stylesheet: ESRI
arcexe80
aiclient
aiserver
arcstorm
arctoolbox
arctools
ArcTutor
BuildingsGeodataba
Catalog
Getting_Started
Greenvalley
Data
Analysis
Parks
Public B
Public U
Transpo
Greenvalley
Water Use
Untitled - ArcMap
File Edit View Insert Selection Tools Window Help
1:94,983
Layers
street_arc
Tools
Display
Source
Drawing
Arial
10
479200.25 3768019.88 Meters


Untitled - ArcMap
File Edit View Insert Selection Tools Window Help
Layers
street_arc
Tools
Attributes of street_arc
FID_1 Shape NAME Shape_Length Type
1 Polyline 136.500014433901 4
2 Polyline 267.503534104294 4
3 Polyline 400.25717951764 4
4 Polyline 806.22123346338 4
5 Polyline 403.006998608611 3
6 Polyline 799.937848334S2 4
7 Polyline 000.7412022004O2 5
8 Polyline 53.4000234S8049 5
9 Polyline 398.754707416395 5
10 Polyline 123.1045S6726918 5
11 Polyline 95.0501002S05001 5
12 Polyline 40.0053830742638 5
13 Polyline 135.607670712223 5
14 Polyline 90.66002S0511209 5
15 Polyline 30.5051020113046 5
16 Polyline 91.3370072599947 5
17 Polyline 158.455728217768 5
18 Polyline 108.500000101048 5
19 Polyline 266.572276546260 5
20 Polyline 269.274931124029 5
21 Polyline 88.7588024628139 5
22 Polyline 134.601737966072 5
Record: 13 2102 out of *2000 Selected.
470471.45 3767617.79 Meters

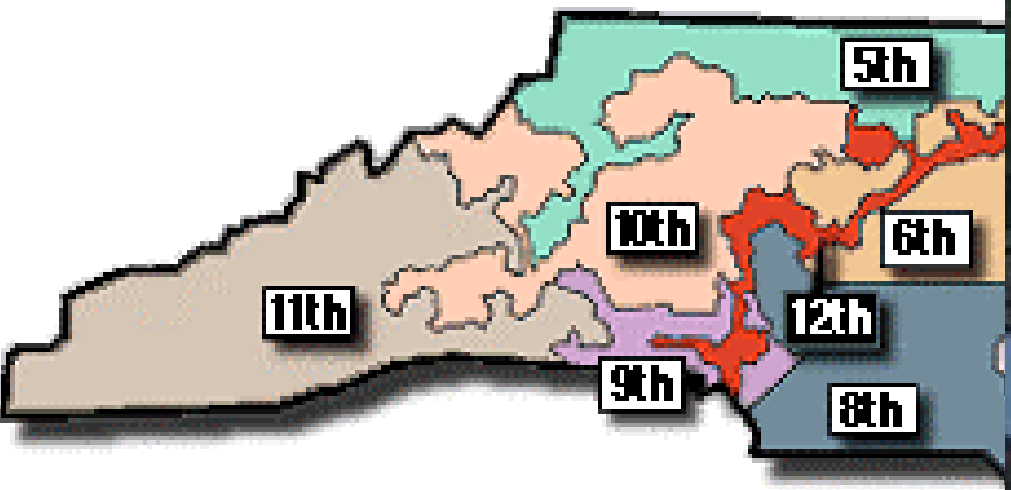
Directions	Distance
1:Start out going East on HENLEY ST towards WARREN ST.	0.1 miles (0.1 km)
2:Turn RIGHT onto WARREN ST.	0.0 miles (0.1 km)
3:Turn RIGHT onto CHELSEA ST.	0.0 miles (0.1 km)
4:CHELSEA ST becomes CHELSEA ST/CITY SQ.	0.1 miles (0.1 km)
5:Turn RIGHT onto CITY SQ/NEW RUTHERFORD AVE/SR-99 N.	0.0 miles (0.1 km)
6:Stay straight to go onto NEW RUTHERFORD AVE/SR-99 N.	0.2 miles (0.3 km)
7:Turn SLIGHT LEFT onto SR-99 N.	0.4 miles (0.6 km)
8:Turn SLIGHT LEFT onto SR-99 N/RUTHERFORD AVE.	0.1 miles (0.1 km)
9:Turn SLIGHT LEFT onto SR-99 N.	0.3 miles (0.4 km)
10:Turn SLIGHT LEFT onto SULLIVAN SQUARE OPAS.	0.4 miles (0.7 km)
11:Turn SLIGHT LEFT onto MYSTIC AVE.	0.7 miles (1.1 km)
12:MYSTIC AVE becomes MYSTIC AVE/SR-38 N.	1.2 miles (2.0 km)
13:Turn LEFT onto HARVARD ST.	0.6 miles (1.0 km)
14:HARVARD ST becomes WARNER ST.	0.2 miles (0.3 km)
15:Turn RIGHT onto POWDER HOUSE SQ.	0.1 miles (0.1 km)
16:Turn RIGHT onto BROADWAY.	1.0 miles (1.6 km)
17:Turn LEFT onto ALEWIFE BROOK PKWY/SR-16.	0.4 miles (0.7 km)
18:ALEWIFE BROOK PKWY/SR-16 becomes ALEWIFE BROOK PKWY/SR-16/US-3.	0.4 miles (0.7 km)
19:Take CONCORD TURNPIKE/SR-2 W.	4.7 miles (7.6 km)
20:Take the WALTHAM ST. exit, exit number 54B, towards LEXINGTON.	0.2 miles (0.3 km)
21:Merge onto WALTHAM ST.	1.9 miles (3.0 km)
22:Turn RIGHT onto MASSACHUSETTS AVE/MASS AVE/SR-225.	0.0 miles (0.0 km)
Total Distance:	12.9 miles (20.8 km)
Estimated Time:	24 minutes

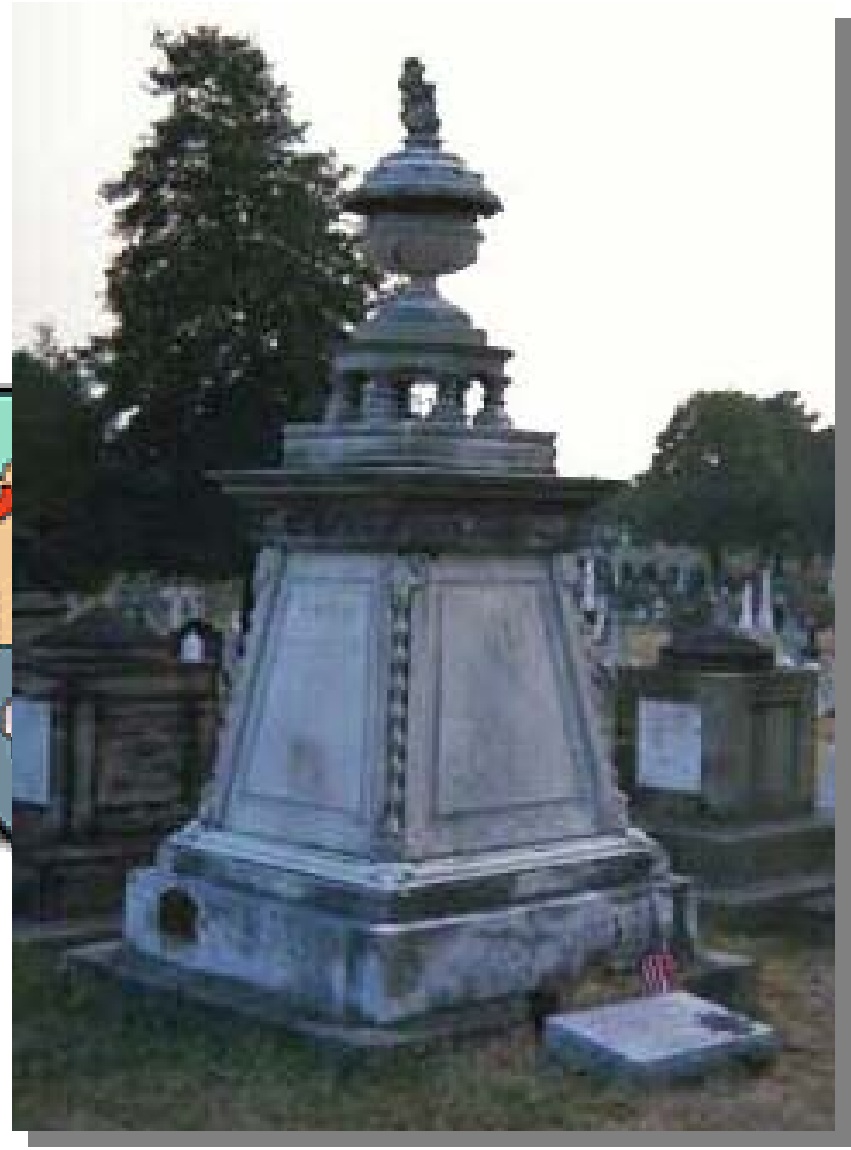
t



Measurements

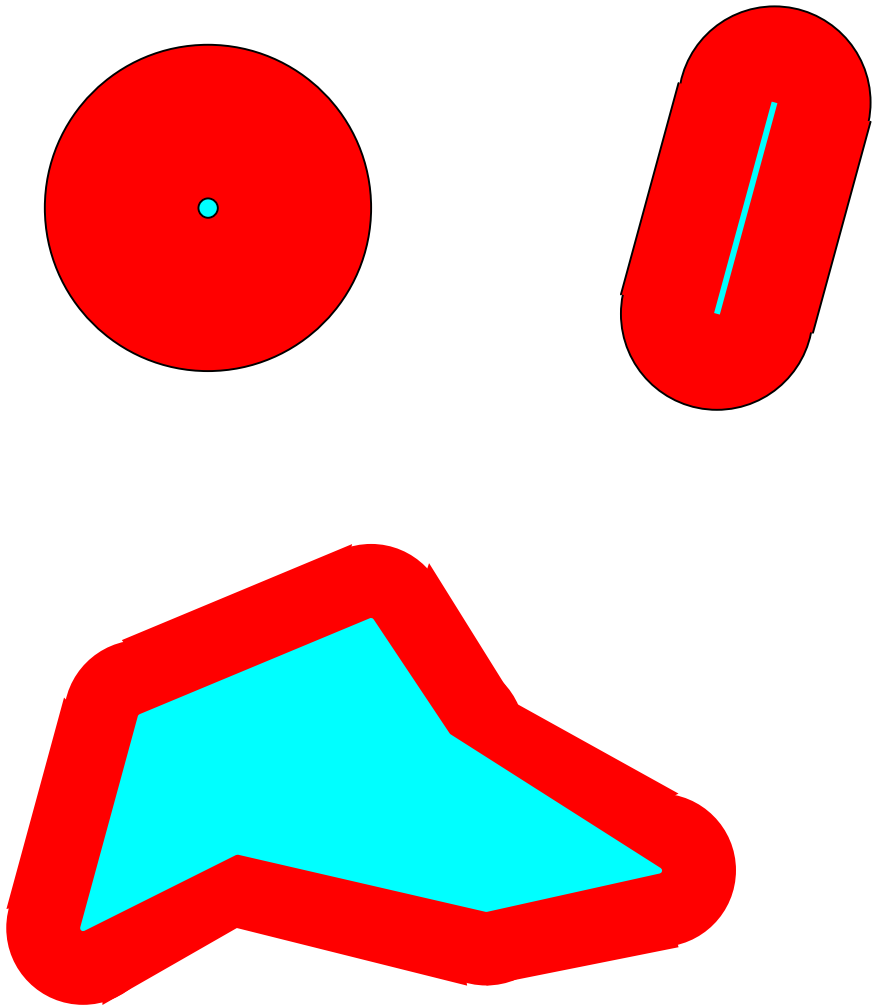
- Area
- Distance
- Length
- Perimeter
- Slope, aspect
- Shape

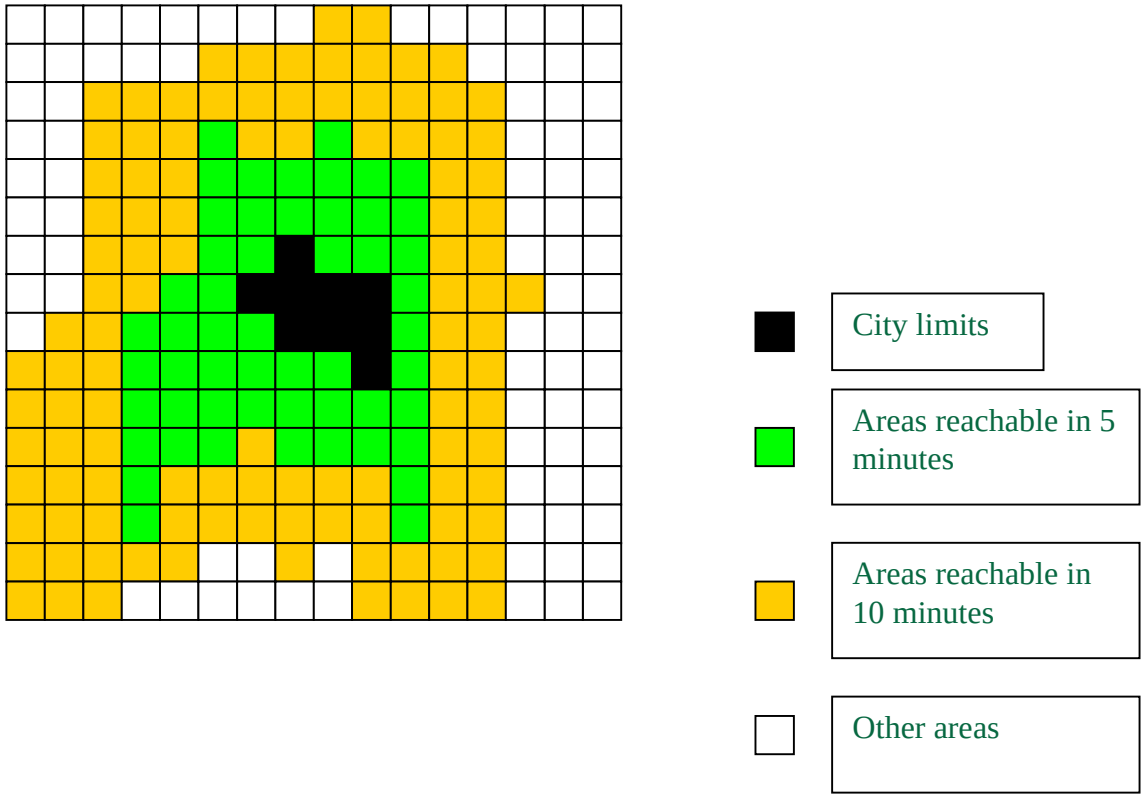

5th
10th
6th
11th
12th
9th
8th

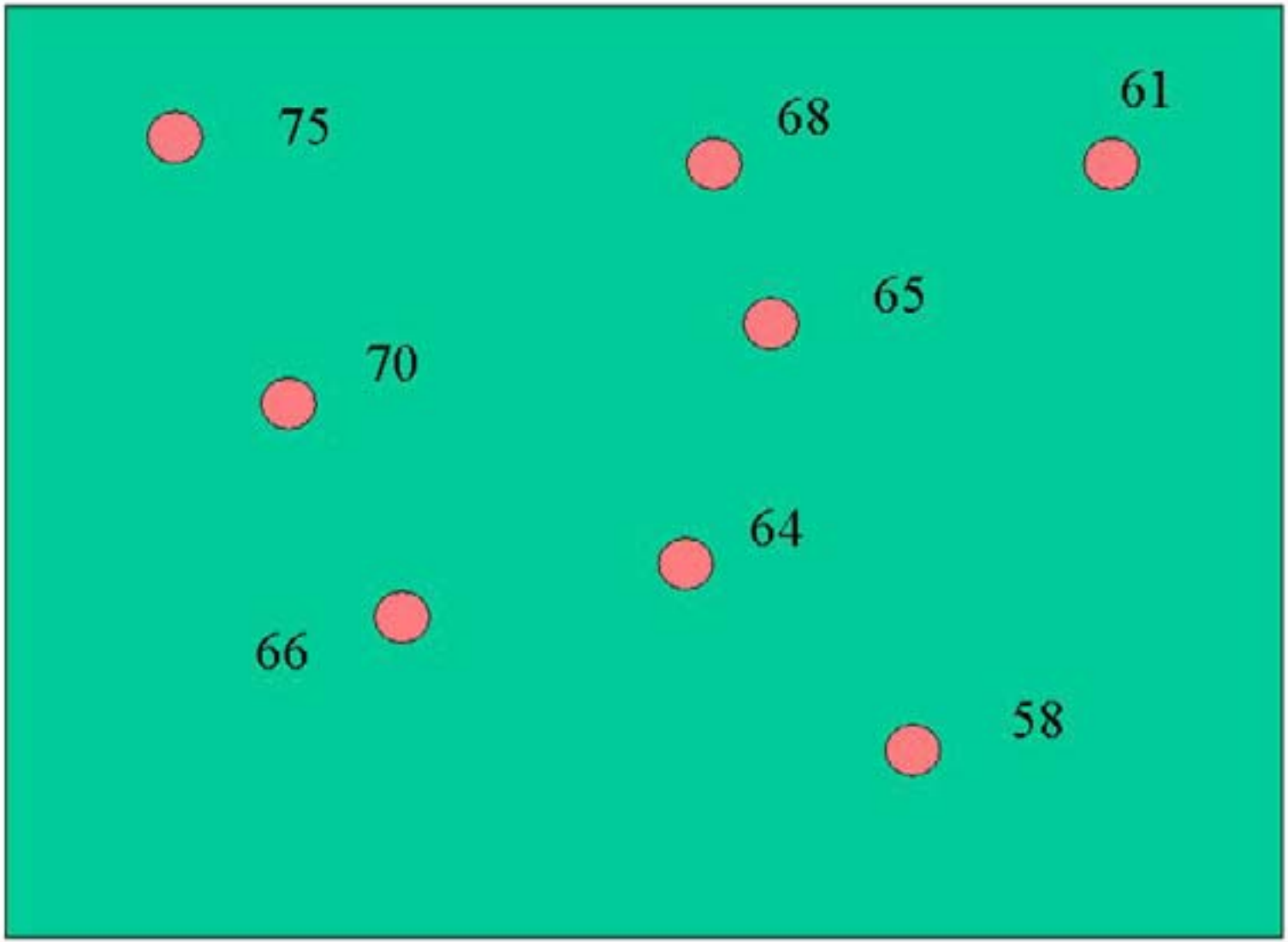


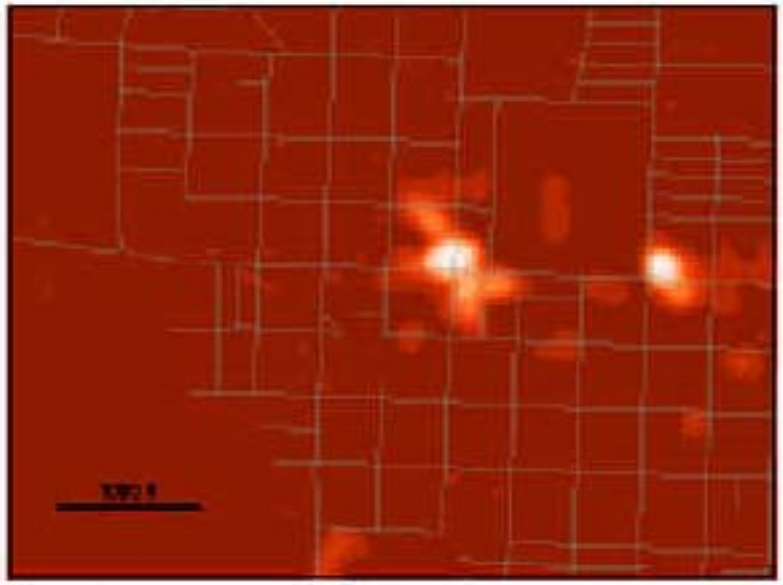
Transformations

- Buffering
- Points in polygons
- Polygon overlay
- Spatial interpolation
- Density estimation

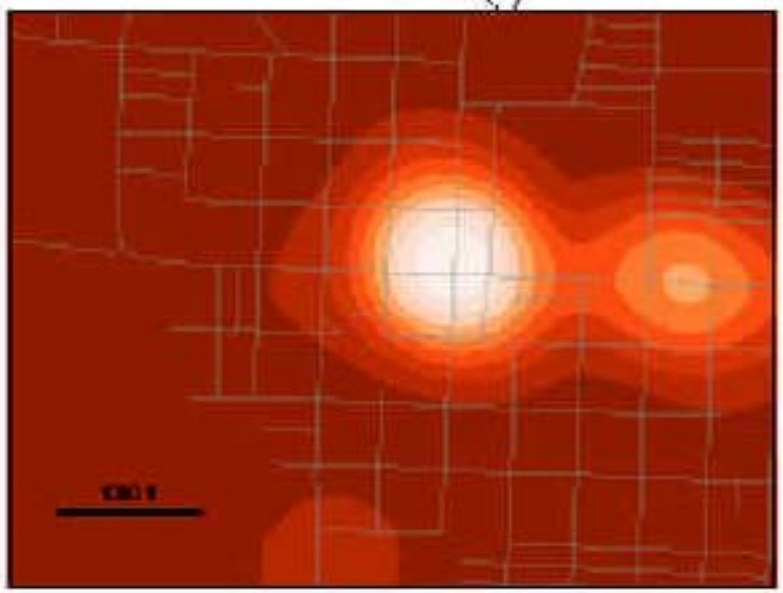



City limits
Areas reachable in 5 minutes
Areas reachable in 10 minutes
Other areas

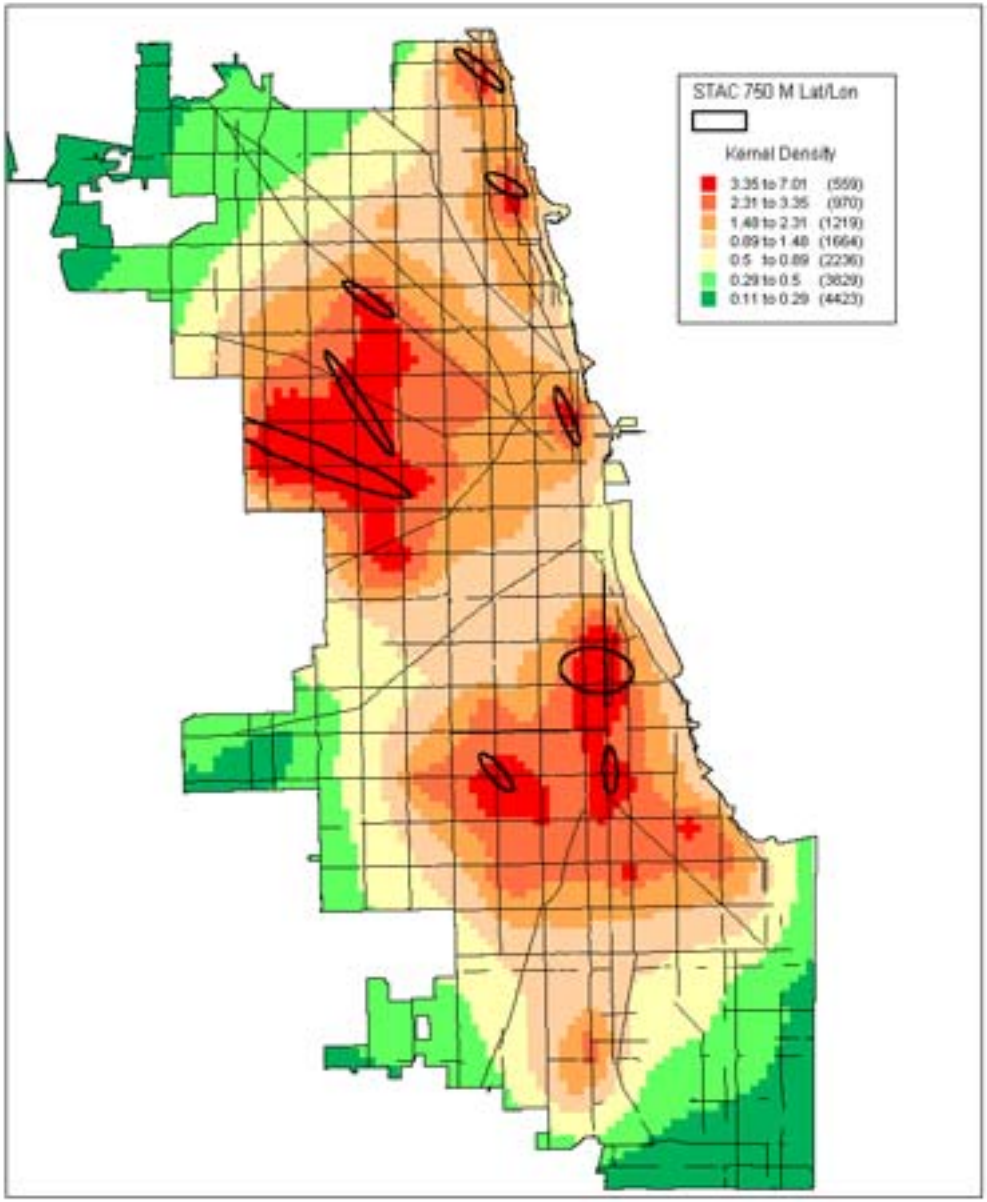

75
68
61
65
70
64
66
58



Search radius = 200 feet



Search radius = 1000 feet



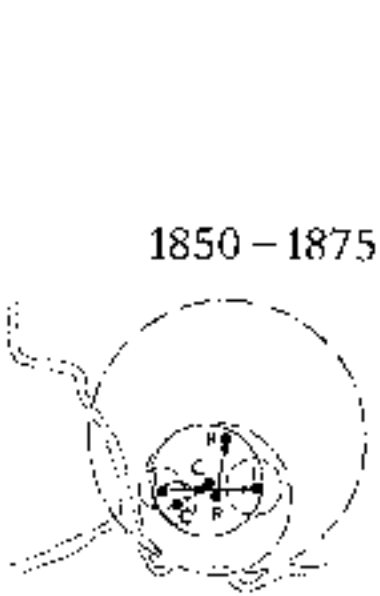
Courtesy of Dick Block

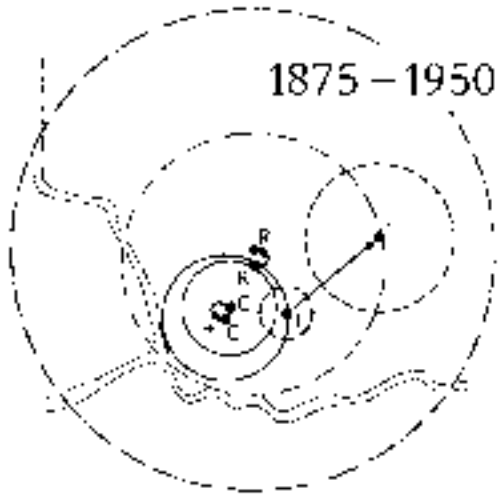
Descriptive summary

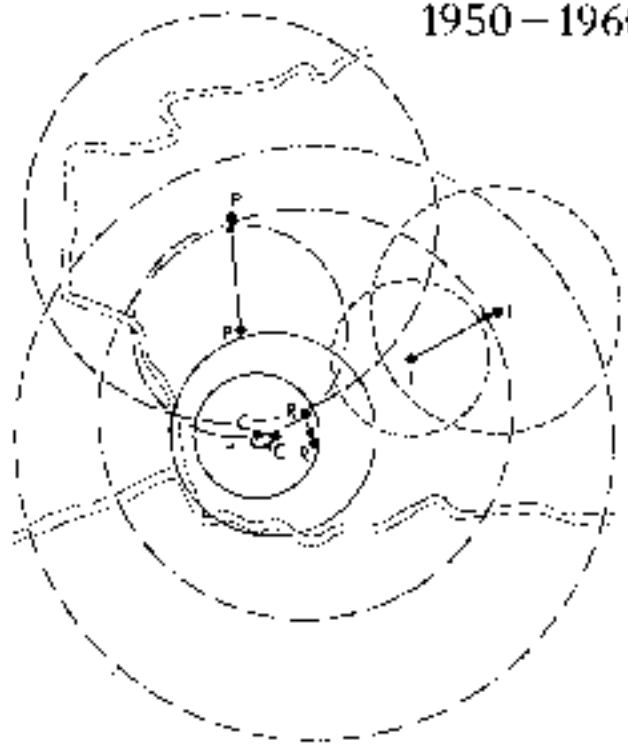
- Centers
- Measures of spatial dispersion
- Spatial dependence
- Fragmentation
- Fractional dimension

MASS MOVEMENTS OF LAND USE SURFACES
LONDON, ONTARIO
1850 – 1960

0 1
miles







N

Land Use Surfaces

— — · — · — — Residential (R)

– – – – – – Public – Institutional (P)

———··· Commercial (C)

- - - - - - - - - Industrial (I)

R C P Mean locations of land use surfaces

Direction of mass movements

Peak value intersection : 1960

Industrial land includes railroad and warehouse properties.

Public and institutional lands included with residential land for 1850 – 1875 and 1875 – 1950.

Optimization

- Design to achieve specific objectives
- Location of central point-like facilities to serve dispersed demand
- Location of linear facilities
- Design of boundaries for elections

PASS

File View Tools Window Help

List View

WorkCtr | Location | Visit | M067124

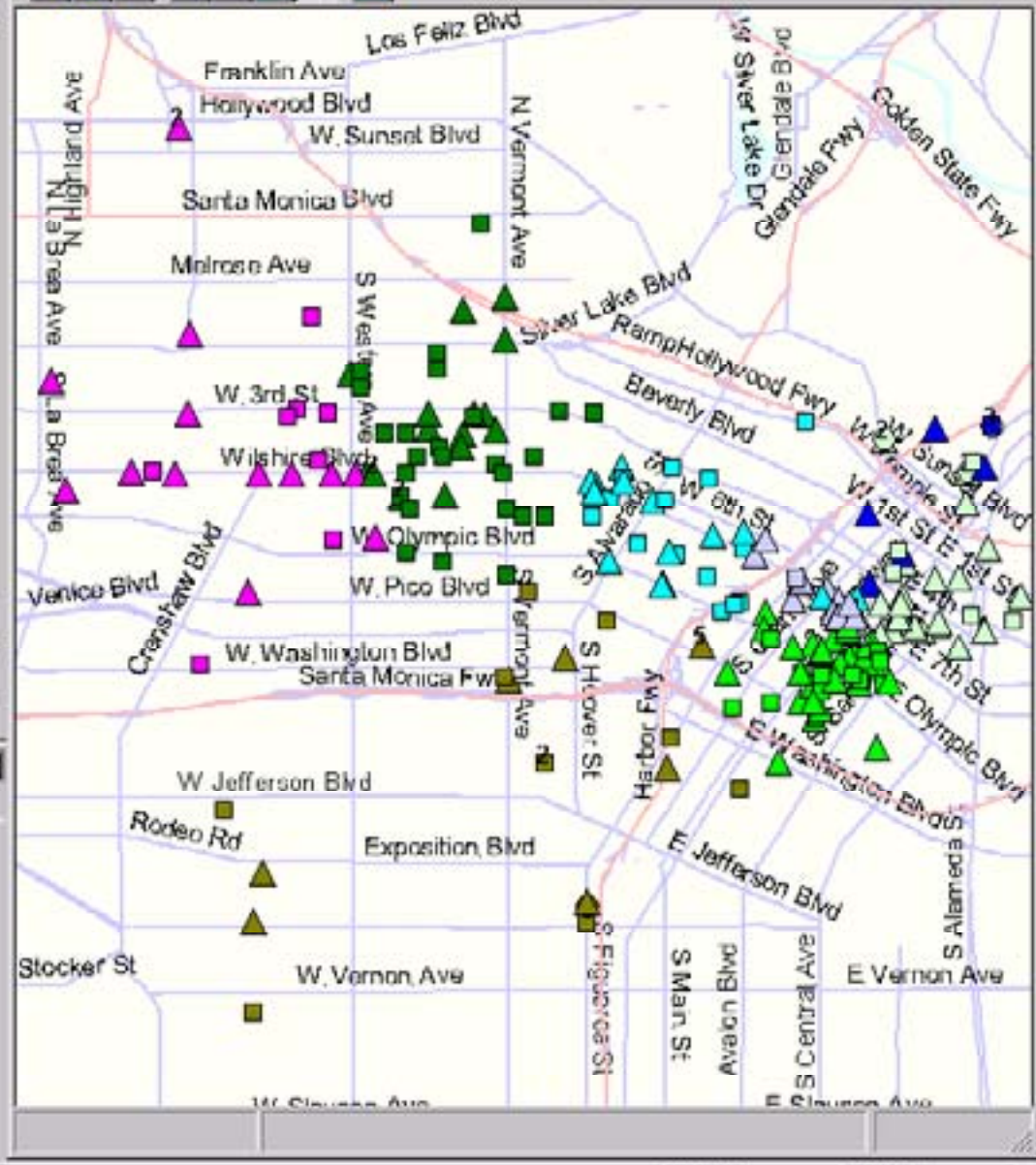
Name	Address	Units	Hrs/Mth	M	Type
1130 VERMONT	1130 VERMONT S	1	0.9	7	HY
625 NEW HAMPSHIRE	625 NEW HAMPSHII	1	0.9	2	HY
3611 WILSHIRE	3611 WILSHIRE BLV	1	0.6	6	HY
114 OXFORD	114 OXFORD S	1	0.8	3	HY
3099 OLYMPIC CAL KOREA B	3099 OLYMPIC BLVI	1	0.9	5	HY
542 MARIPOSA GORDON CHI	542 MARIPOSA S	1	0.8	5	HY
140 MARIPOSA AVE	140 MARIPOSA AVE	1	0.9	9	HY
209 MARIPOSA S	209 MARIPOSA S	1	0.8	5	HY
2930 FRANCIS	2930 FRANCIS	1	0.9	4	HY
248 OCCIDENTAL	248 OCCIDENTAL	1	0.9	3	HY
601 WESTMORELAND	601 WESTMORELAI	1	0.7	4	HY
303 COMMONWEALTH	303 COMMONWEAL	1	0.9	9	HY
2830 FRANCIS	2830 FRANCIS	1	0.9	7	HY
187 OXFORD OXFORD VILLA	187 OXFORD	1	0.9	8	HY
601 ARDMORE PUBLIC COUI	601 ARDMORE AVE	1	0.7	5	HY
3535 SIXTH ST.	3535 W SIXTH	1	0.7	5	HY
1052 MARIPUSA AVE	1052 S MARIPUSA	1	1.0	0	HY
128 MARIPOSA AV S	128 MARIPOSA AV S	1	0.8	4	HY
449 KINGSLEY BLDG	449 KINGSLEY DRIV	2	1.9	2	HY
445 HOBART	445 HOBART	2	1.8	6	HY
350 CATALINA BLDG	350 CATALINA STRE	1	0.9	6	HY
3500 EIGHTH STREET BLDG	3500 EIGHTH STRE	1	0.9	5	HY
3020 WILSHIRE BLDG	3020 WILSHIRE BLV	2	0.0	0	HY

Chart

Chart All | Dollars | Units | Hours | C1

Map

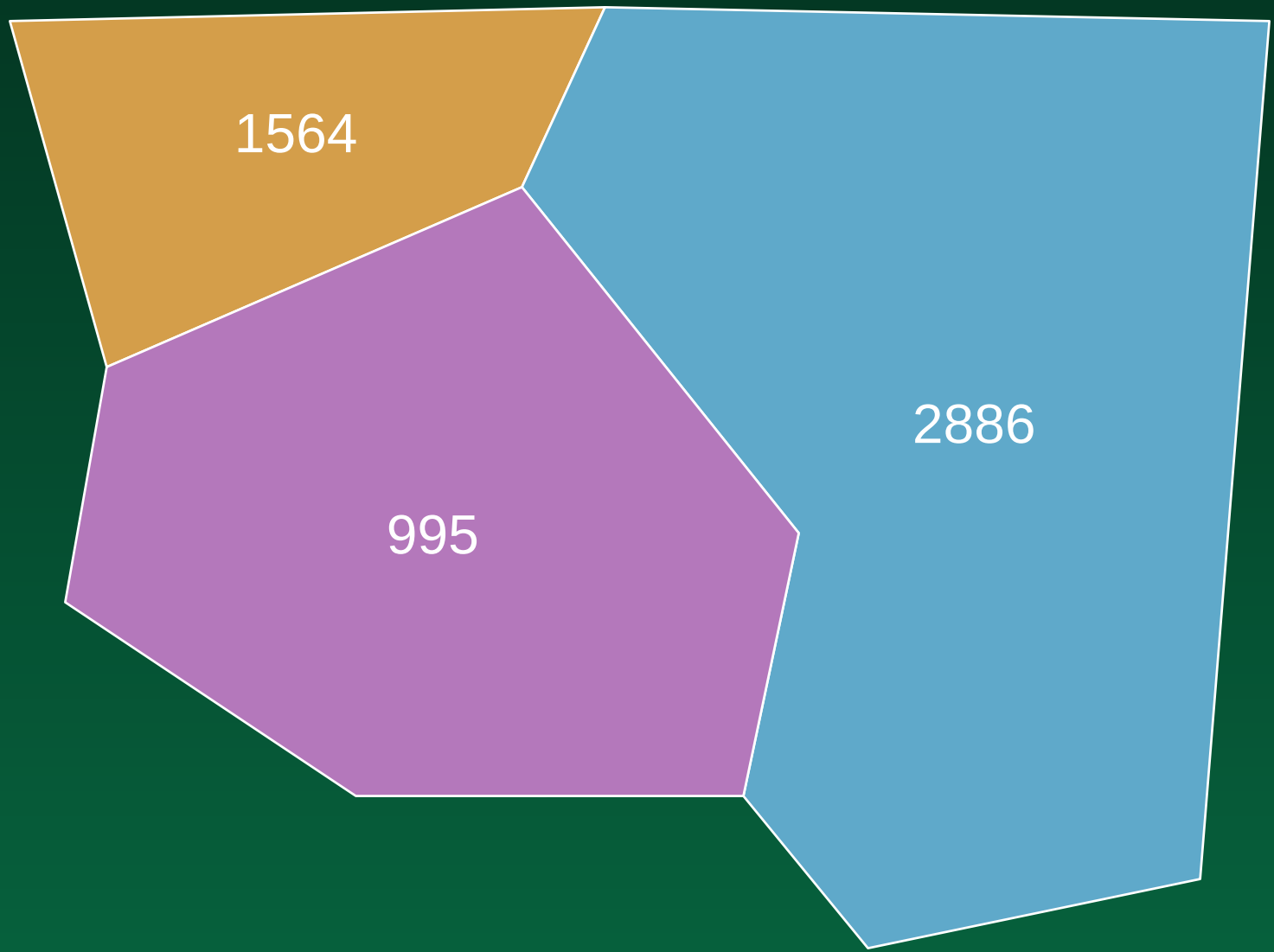
Show selected only



9/12/99 10:39 PM

Hypothesis testing

- Geographic objects as a sample from a population
  - what is the population?
- The independence assumption
  - the First Law of Geography
  - failure to find spatial dependence is always a Type II error
  - hell is a place with no spatial dependence

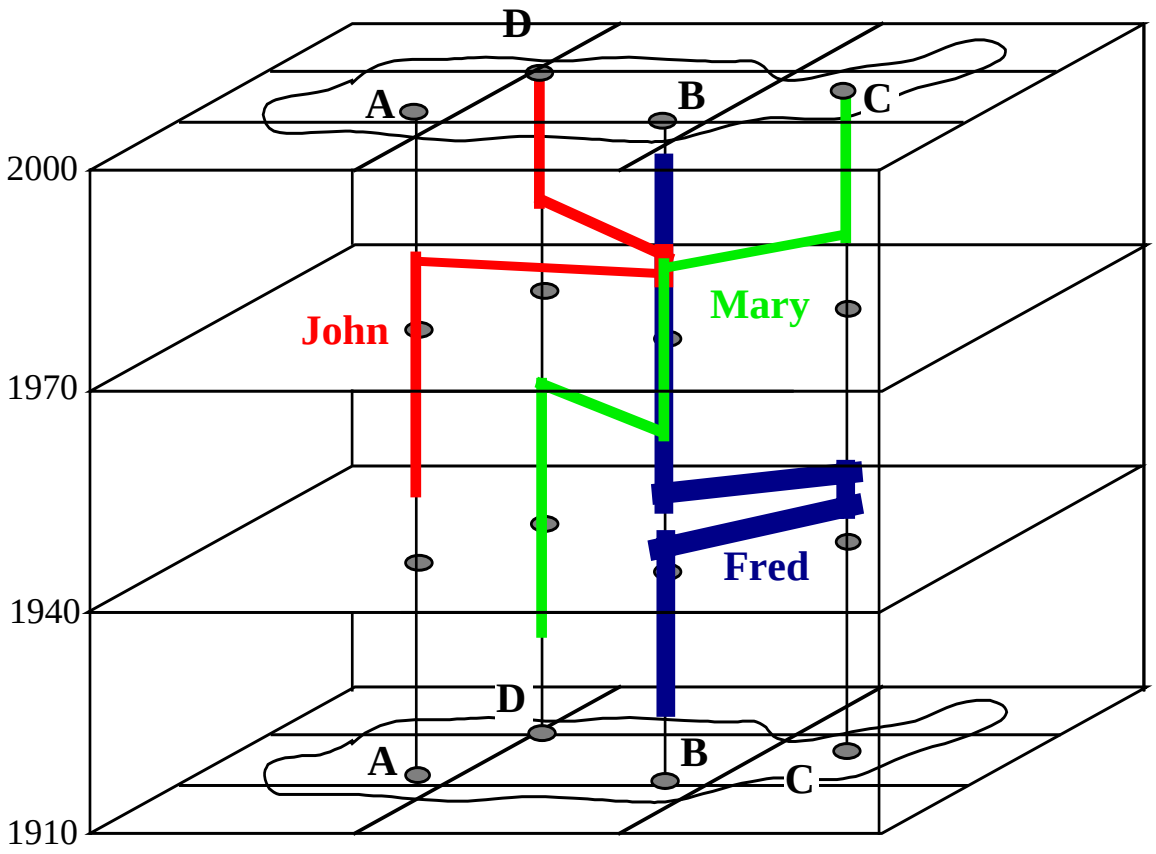

1564
2886
995
1990

Information lost to the representation

- All sub-polygon spatial variation
- All within-decade temporal variation
- All identities
  - instead of <xy, person> we have

<R, number>

<xy, xy, xy, xy, …, R>


D
A
B
C
2000
John
Mary
1970
Fred
1940
D
A
B
C
1910

Challenges of GIS

- How to characterize what is missing?
  - error, accuracy, uncertainty
- How to choose the best representation?
  - confounding influences
- How to support many data models in a single software package

Weaknesses of GIS

- There are too many possible data models
  - special-purpose GIS
  - lack of interoperability
- Difficult to add data models retroactively

Objectives, Structure, and History of CSISS

Research infrastructure

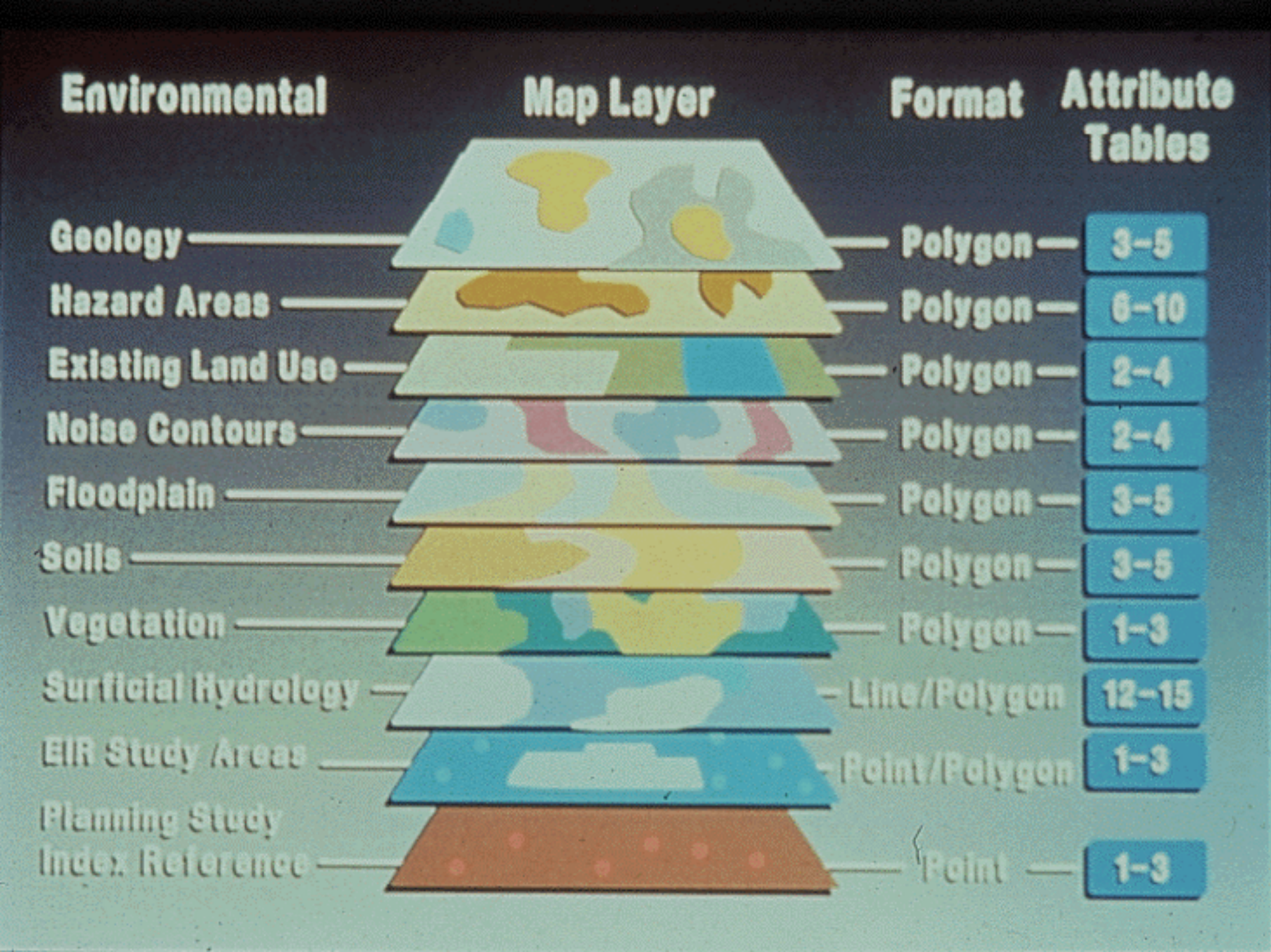
- Facilities that serve generic needs
  - economies of scale
- The Hubble Telescope
  - high fixed costs distributed over many users
- Infrastructure funding at NSF
  - South Pole
  - advanced computation

Generalizing the concept

- Shared computational facilities
- Shared data archiving and access
- Software tools
  - shared licenses
  - shared development
- Education and training
  - investments in skills
  - high leverage

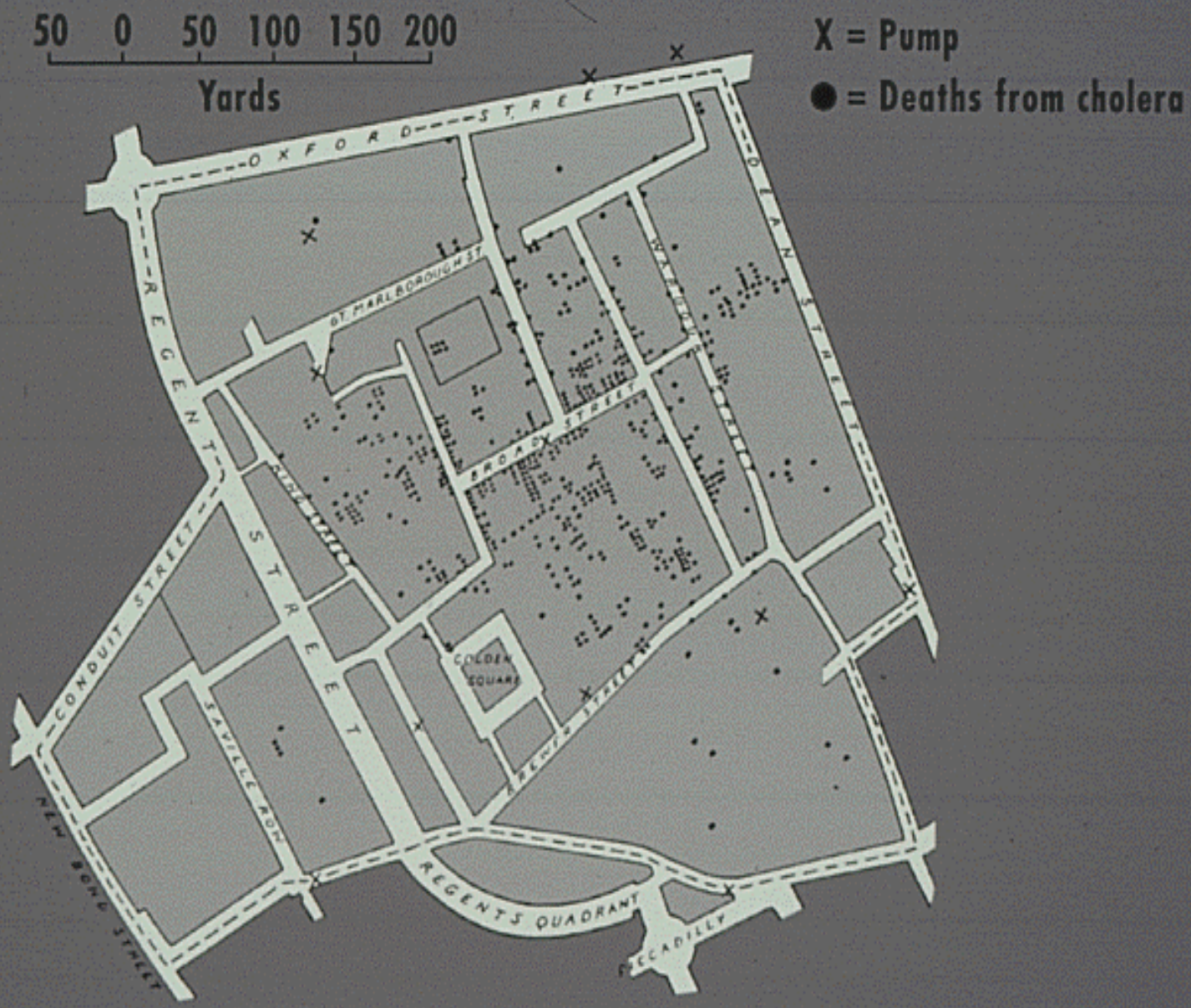
General principles: 1. Integration

- Linking data through common location
  - the layer cake
- Linking processes across disciplines
  - spatially explicit processes
  - e.g. economic and social processes interact at common locations


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

2. Spatial analysis

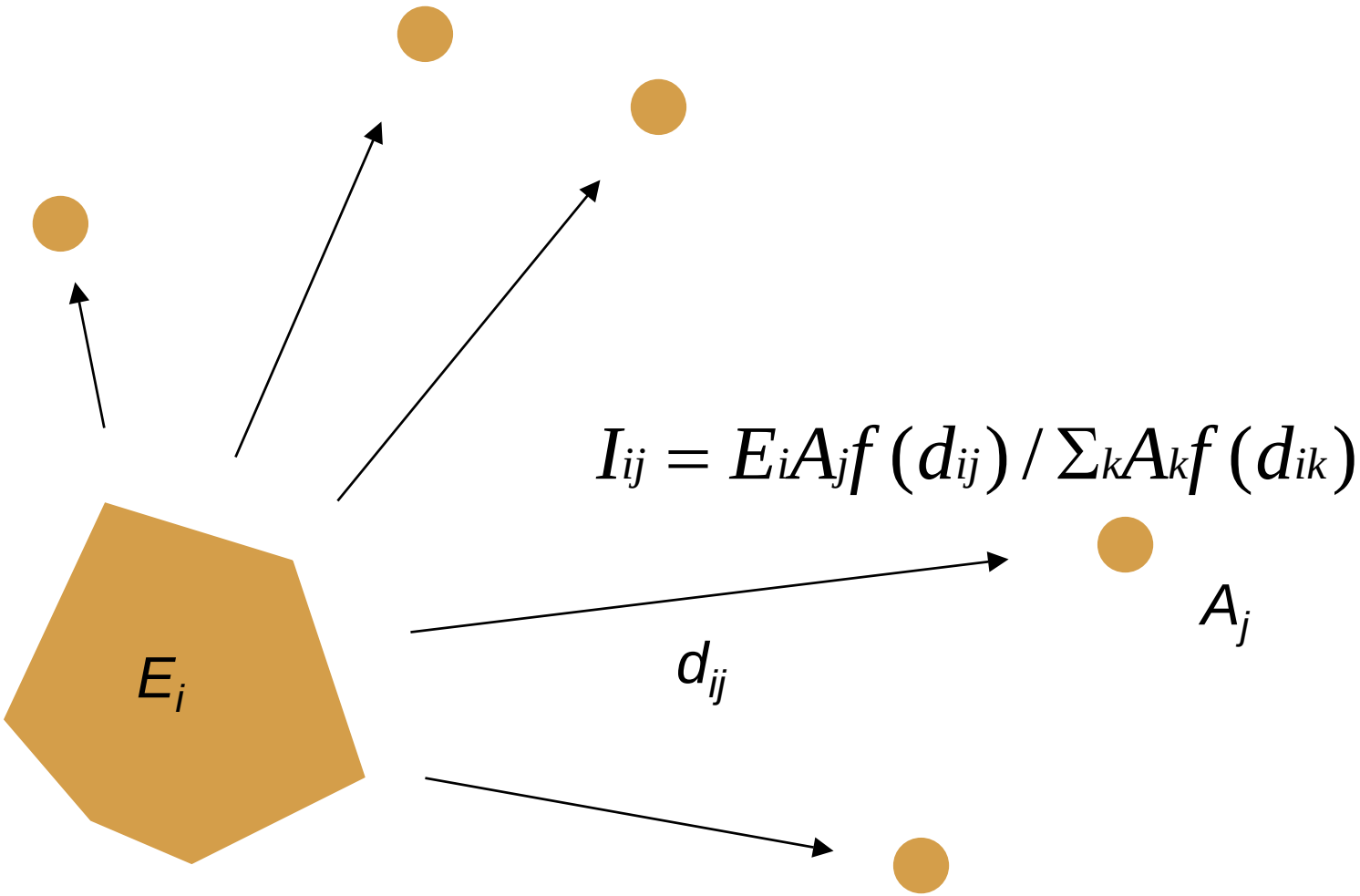
- Social data collected in cross-section
  - longitudinal data are difficult to construct
- Cross-sectional perspectives are rich in context
  - can never confirm process
  - though they can perhaps falsify
  - useful source of hypotheses, insights



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.

3. Spatially explicit theory

- Theory that is not invariant under relocation
- Spatial concepts (location, distance, adjacency) appear explicitly
- Can spatial concepts ever *explain*, or are they always surrogates for something else?


$I_{ij} = E_i A_j f(d_{ij}) / \Sigma_k A_k f(d_{ik})$
A_j
d_{ij}
E_i

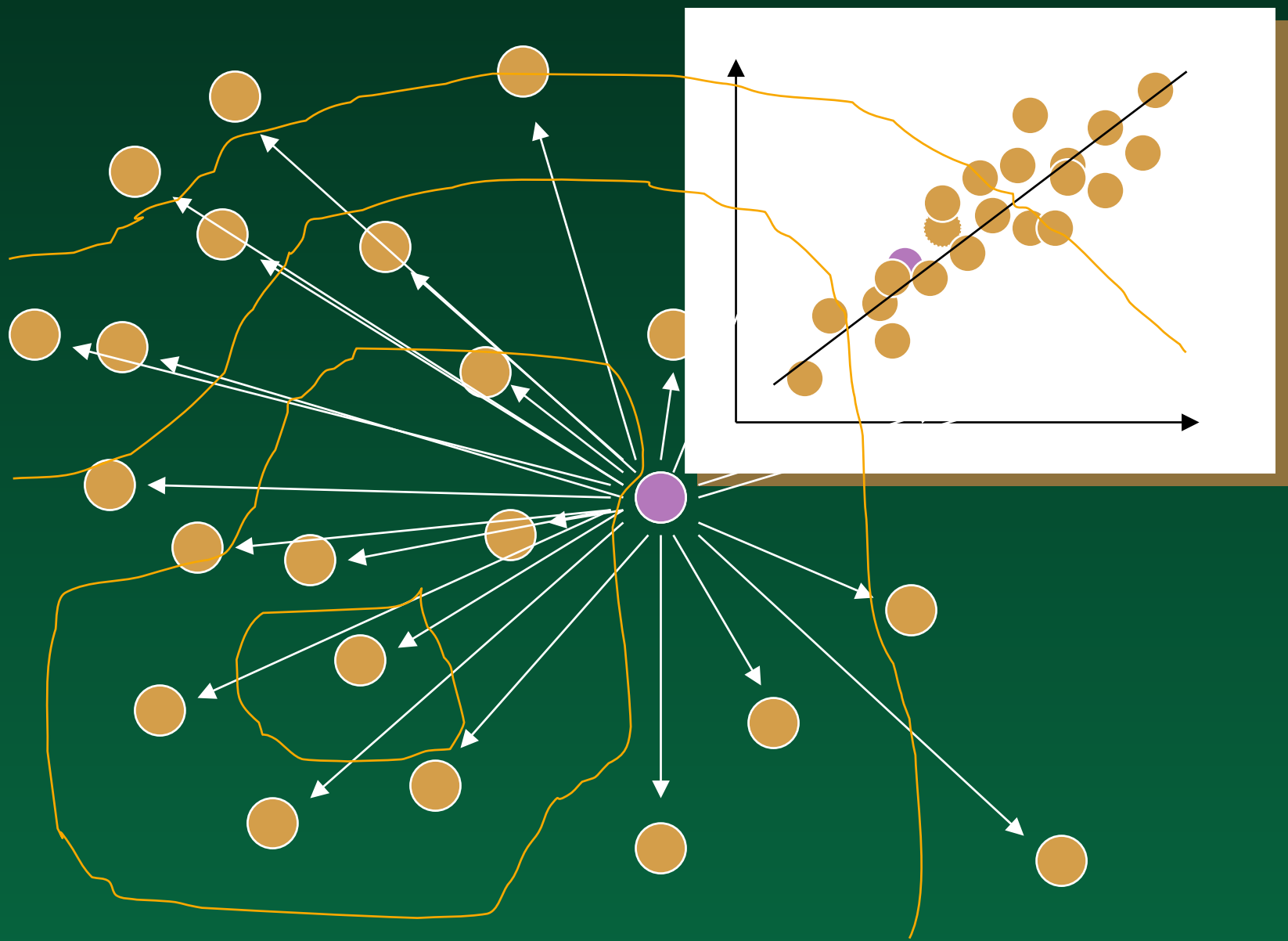
4. Place-based analysis

- Nomothetic - search for general principles
- Idiographic - description of unique properties of places
- An old debate in Geography

The Earth's surface

- Uncontrolled variance
- There is no average place
- Results depend explicitly on bounds
- Places as samples
- Consider the model:

  $y = a + bx$



5. Knowledge and policy

- Policy requires the projection of general knowledge in spatial context
  - the implications of this process in this location
  - alternative futures visualized under local circumstances
- GIS combines the general (processes, models, algorithms) with the specific (database of local details)

6. Place-based search

- Location as an organizing dimension to information
  - much information can be georeferenced
  - much more than maps and images
- The Geolibrary
  - what have you got about *there*?
  - impossible physically, feasible digitally

Prototype geolibraries

- National Geospatial Data Clearinghouse
  - *www.fgdc.gov*
- Microsoft's Terraserver
  - *terraserver.microsoft.com*
- Alexandria Digital Library
  - *alexandria.ucsb.edu*

EXANDRIA DIGITAL LIBRARY

CDL CALIFORNIA DIGITAL LIBRARY

ADL HOME · CDL HOME · HELP

NICK PLACENAME SEARCH

earch the entire world for...

Find

er "Rome" if you want Rome, Italy.
e information

NERAL SEARCH

Select collection to search

DL Catalog

ose collections

Set geographic region

the map to the right to set the geographic
ent of the search, or directly enter bounding
rdinates below.

N 34.43

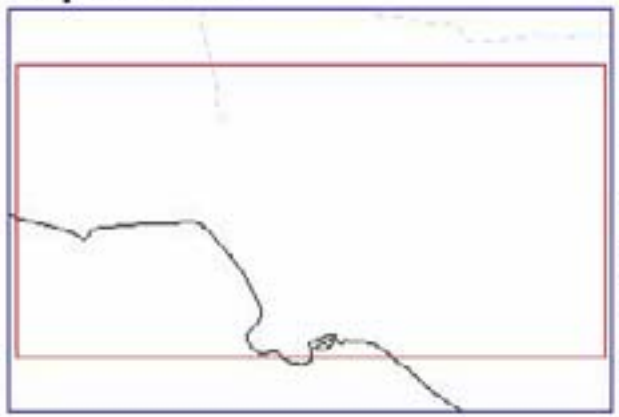
W -118.97 -117.57 E

33.73 S

Words to search for

Any of the above words
All of the above words
Exact phrase

Map Browser

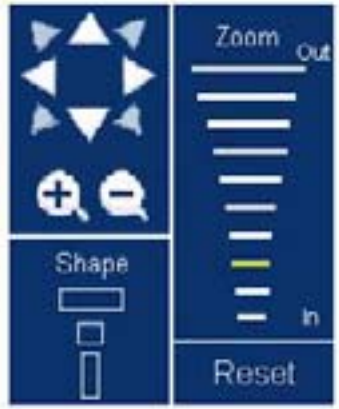


Click map to:

Recenter & Zoom in

Change location to:

--



ADL Search Results

The query that produced these results can be found at the bottom of this page.

1. **DRG o33117g6, Digital Raster Graphic of BLACK STAR CANYON, CA.**
   Type: maps. **Format:** TIFF. **Date:** 1988. **ADL identifier:** adl_catalog:800279.
   HIGHLIGHT IN MAP · COMPLETE DESCRIPTION · BROWSE GRAPHIC · ACCESS/DOWNLOAD

2. **DRG o33117g7, Digital Raster Graphic of ORANGE, CA.**
   Type: maps. **Format:** TIFF. **Date:** 1981. **ADL identifier:** adl_catalog:800280.
   HIGHLIGHT IN MAP · COMPLETE DESCRIPTION · BROWSE GRAPHIC · ACCESS/DOWNLOAD

3. **DRG o33117g8, Digital Raster Graphic of ANAHEIM, CA.**
   Type: maps. **Format:** TIFF. **Date:** 1981. **ADL identifier:** adl_catalog:800281.
   HIGHLIGHT IN MAP · COMPLETE DESCRIPTION · BROWSE GRAPHIC · ACCESS/DOWNLOAD

CSISS mission

- The CSISS mission recognizes the growing significance of space, spatiality, location, and place in social science research. It seeks to develop unrestricted access to tools and perspectives that will advance the spatial analytic capabilities of researchers throughout the social sciences.

Seven CSISS programs

- National Workshops
- Software Tools
- Virtual Community
- Best Practice Examples
- Place-Based Search
- Learning Resources
- Specialist Meetings

The W Matrix

Abstraction of geographic space

- Cartograms

- Invariance under rotation, displacement, reflection
- Reconstruction from a distance matrix
- Reconstruction from ranked distances
  - ordered metric data (Coombs)

Space as a matrix

- ***W*** where w_{ij} is some measure of interaction
  - adjacency
  - decreasing function of distance
  - invariant under rotation, displacement, reflection
  - readily obtained from GIS
  - reflects a discrete object conceptualization

The Modifiable Areal Unit Problem

- Openshaw and Taylor
  - 99 counties of Iowa
  - % Republican voters, % over 65
- 48 regions: -.548 to +.886
- 12 regions: -.936 to +.996
- Solutions:
  - manipulate to determine range
  - strengthen theoretical framework

Applications of the *W* matrix

- Spatial regression
  - add spatially lagged terms weighted by ***W***
  - Anselin's SPACESTAT
- Moran and Geary indices of spatial dependence

$$c = \frac{(n-1)\sum_i \sum_j w_{ij}(x_i - x_j)^2}{2\sum_i \sum_j w_{ij} \sum_i (x_i - a)^2}$$

Social versus spatial

- ***W*** estimated
- Well-defined discrete object conceptualization
- Direct measures of interaction

- ***W*** calculated
- Discrete objects as arbitrary regions, MAUP
- Surrogate measures of interaction

Cross-Product Statistics

- Let **C** be a matrix of similarities between objects
- Consider the cross-product

$$\Gamma = \sum_i \sum_j c_{ij} w_{ij}$$

Properties of Γ

- Generalizes Moran and Geary
- Measures the correlation between social and spatial
- Simple randomization tests based on permutation
- Hubert, Golledge, and Costanzo, *Geographical Analysis* 13(3): 224-233, 1981

Social versus Spatial

- The ability to explain
  - empirical estimates versus measurements
- Additional arguments supporting spatial
  - context
  - integration
  - implementation
- A fruitful basis for collaboration