# Introduction to GIS

CSDE Workshop
October, 2017

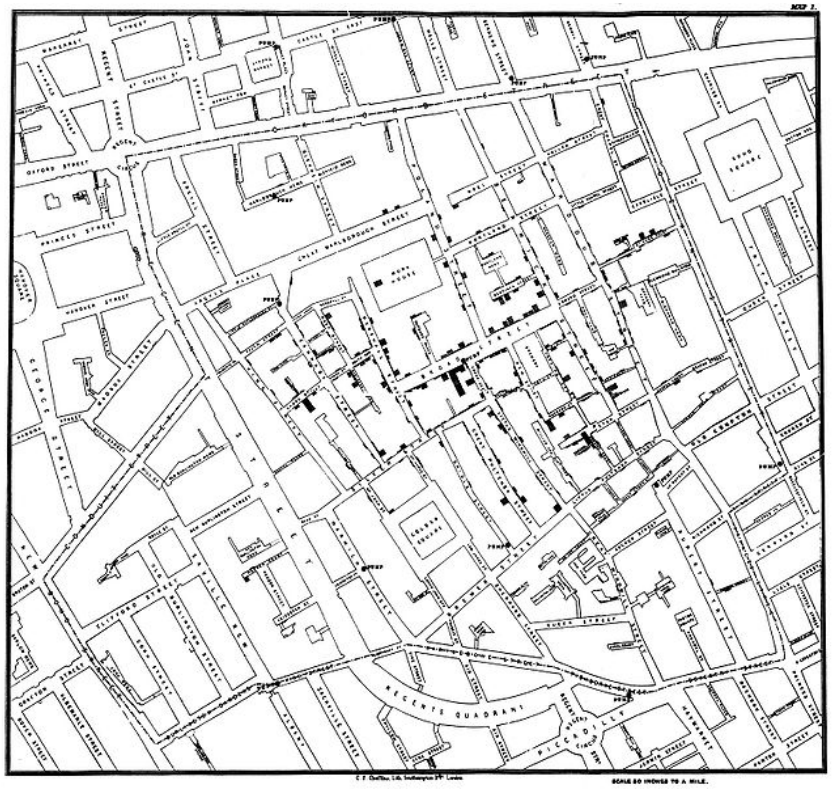
Matt Dunbar
Spatial Demography
Research Scientist

Center for Studies in
Demography and Ecology

# History: Mapping and Health

- First, identified **WHERE** people infected with this water-borne disease
  - *By making marks on a map for each sick person*

- The **pattern** suggested that water from a particular pump was infected and the cause of the disease.

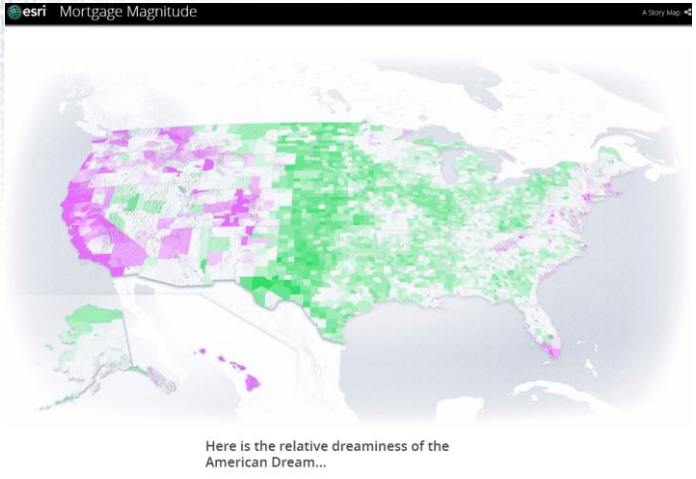


# Spatial is Special

"Everything is related to everything else, but near things are more related than distant things"

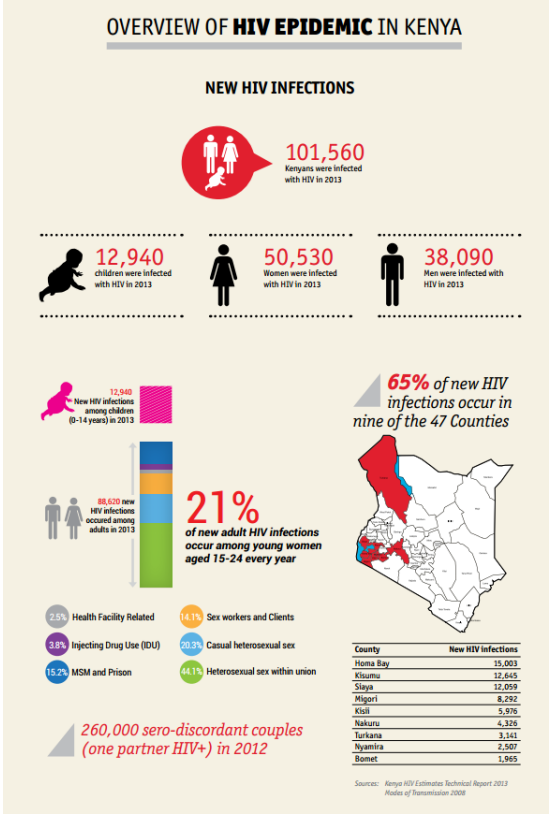
Tobler, W. 1970. A computer movie simulating urban growth in the Detroit region. *Economic Geography* **46**, 234–40.

Sometimes called the **First Law of Geography** (because it is generally true!).




Here is the relative dreaminess of the American Dream...

https://arcg.is/1fHaCy




Kenya HIV Prevention Revolution Road Map (Kenya MOH)

# GIS and Social Science

GIS and spatial analysis methods are becoming a part of the social scientist's methodological arsenal;

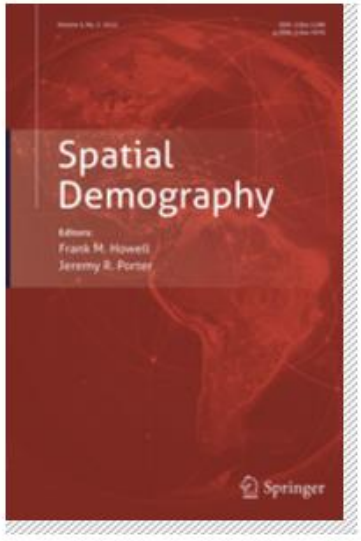
**Websites**

GISpopsci.org

http://gispopsci.org/

spatial@ucsb

http://spatial.ucsb.edu/




**Journals**

Spatial Demography

http://www.springer.com/social+sciences/population+studies/journal/40980

**On-line resources and Geospatial data warehouses**

OnTheMap at the US Census Bureau: http://onthemap.ces.census.gov/

GIS and Public Health at NCHS: http://www.cdc.gov/nchs/gis.htm

# What is GIS?

One word at a time…

# G Information S

Data is a fact or collection of facts

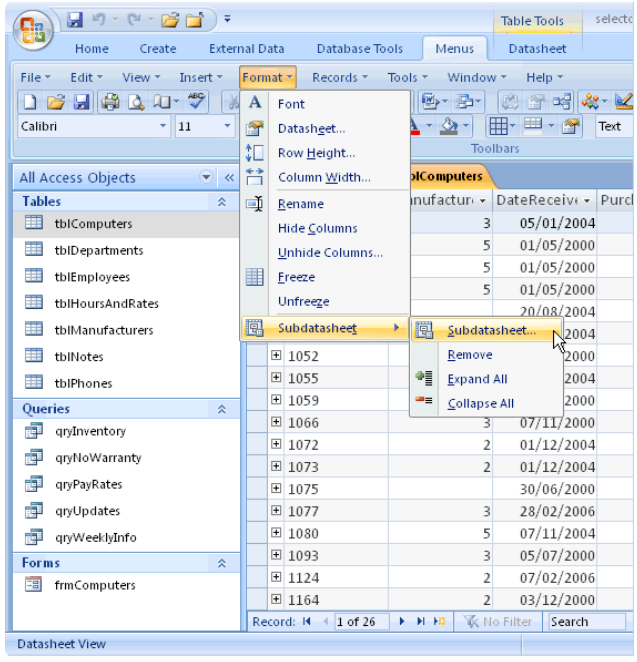
- A single interview: What generic word do you use to describe carbonated soft drinks? http://www.popvssoda.com/statistics/ALL.html
- The entire US Census

Data that is processed, organized, structured or presented in a given context to make them useful, are called Information

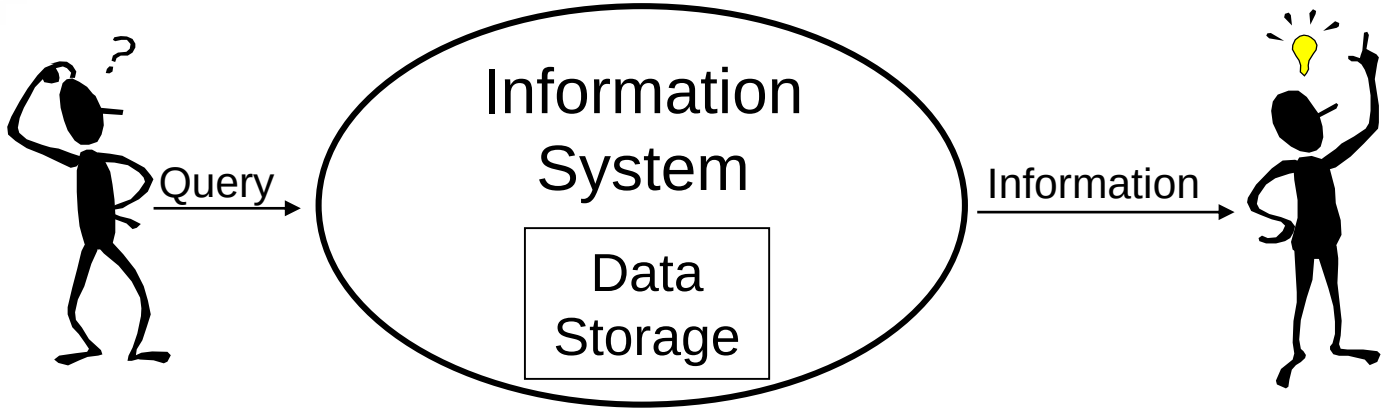
# G Information System

A set of components for:

*Storing*
*Displaying*
*Analyzing*
**DATA**




One example of an Information System:
Microsoft Access database



# Geographic Information System

## Information System

## + 

## Geographic Position

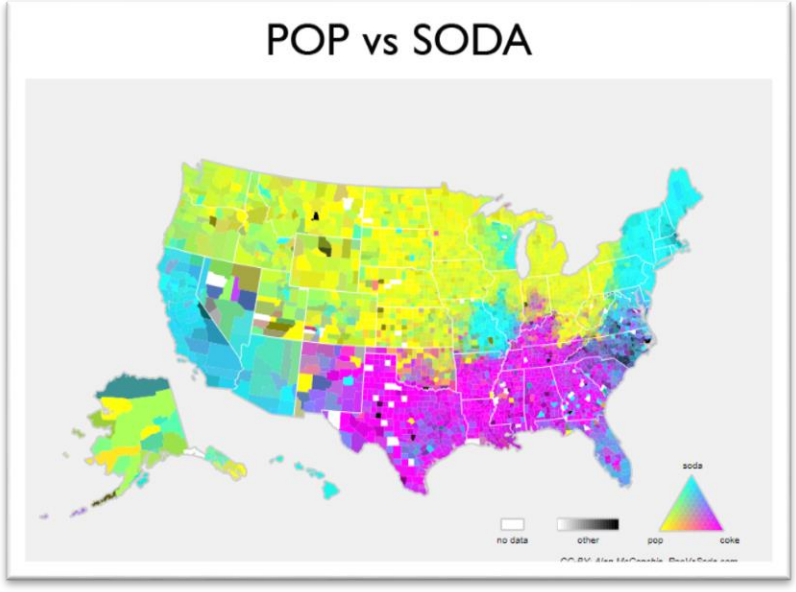
A means of:

*Storing*

*Mapping*

*Analyzing*

**Spatial Data**



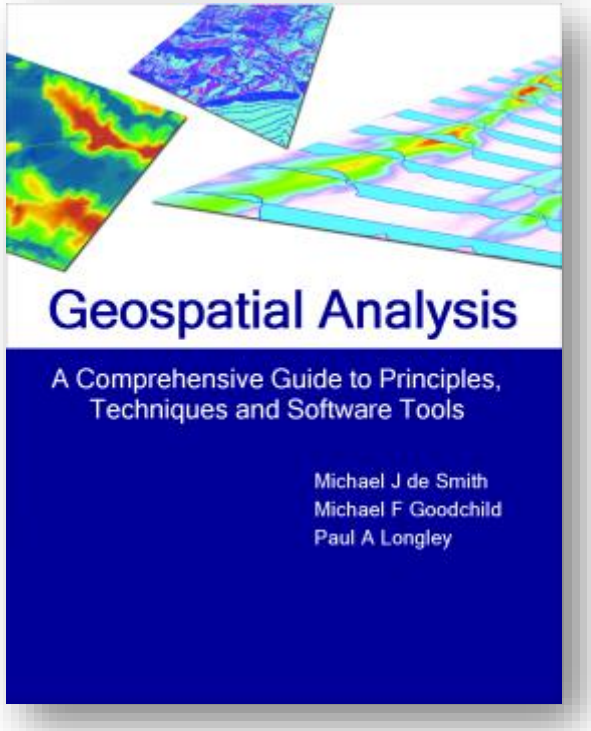

http://www.popvssoda.com/

# GIS Defined

A computer-based system for the manipulation and analysis of geospatial information in which there is an automated link between a data object and their spatial location.

**http://www.spatialanalysisonline.com/**

Free on-line textbook



# What is GIS?

Examples…

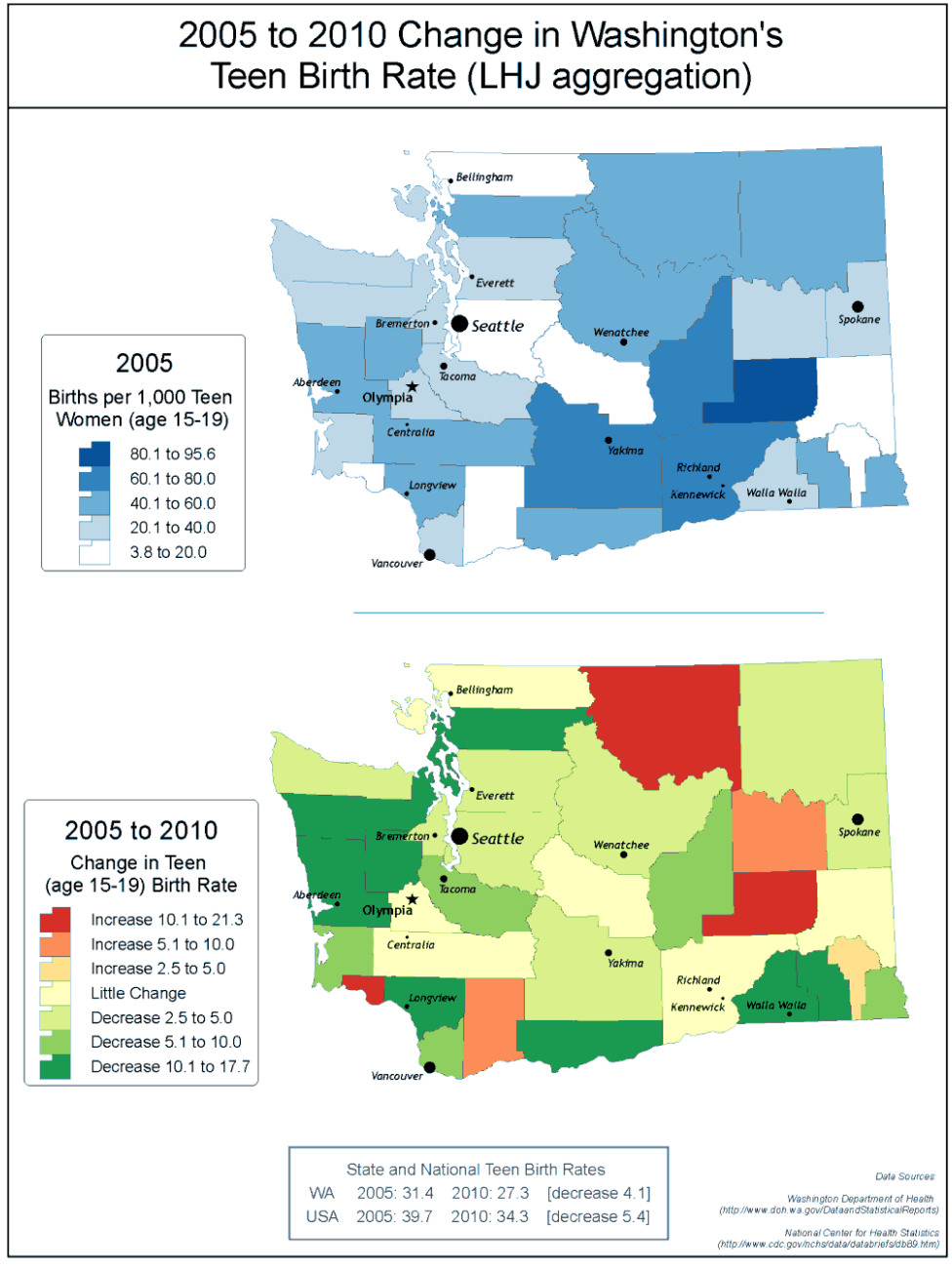
# Public Health Systems and Services Research

## Local Health Departments:

## Exploring changes in funding, level of service, need, and health outcomes from 2000-2010.

**Local health departments and specific maternal and child health expenditures: relationships between spending and need**. *Betty B Bekemeier, Matthew M Dunbar, Matthew M Bryan, Michael E ME Morris. J Public Health Manag Pract 18(6):615-22 (2012), PMID 23023288*

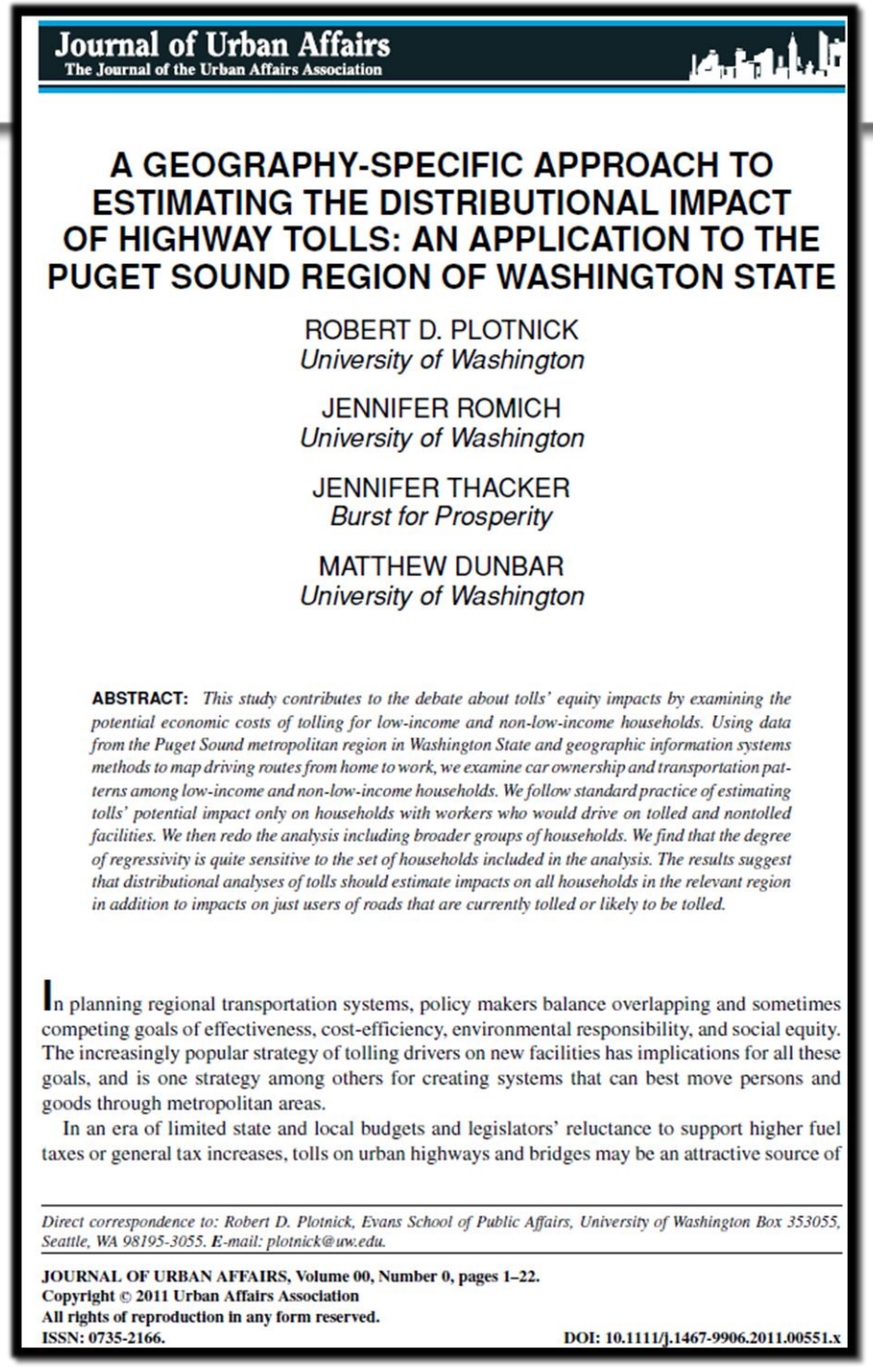
**Local Health Department Provision of WIC Services Relative to Local "Need"— Examining 3 States and 5 Years**
Bekemeier, B, Bryan M, Dunbar MD, Fowler, C. *Frontiers in Public Health Services and Systems Research 1(1)* (2012).



# Impact of Tolls

- Impact on Equity
  - Tolls are regressive: cost low-income households a higher % of their income than middle- or upper-income households
  - How regressive? Use GIS-based approach to provide empirical analysis

NOTE: compared to alternatives, tolls are least regressive option



Journal of Urban Affairs
The Journal of the Urban Affairs Association

**A GEOGRAPHY-SPECIFIC APPROACH TO ESTIMATING THE DISTRIBUTIONAL IMPACT OF HIGHWAY TOLLS: AN APPLICATION TO THE PUGET SOUND REGION OF WASHINGTON STATE**

ROBERT D. PLOTNICK
*University of Washington*

JENNIFER ROMICH
*University of Washington*

JENNIFER THACKER
*Burst for Prosperity*

MATTHEW DUNBAR
*University of Washington*

**ABSTRACT:** *This study contributes to the debate about tolls' equity impacts by examining the potential economic costs of tolling for low-income and non-low-income households. Using data from the Puget Sound metropolitan region in Washington State and geographic information systems methods to map driving routes from home to work, we examine car ownership and transportation patterns among low-income and non-low-income households. We follow standard practice of estimating tolls' potential impact only on households with workers who would drive on tolled and nontolled facilities. We then redo the analysis including broader groups of households. We find that the degree of regressivity is quite sensitive to the set of households included in the analysis. The results suggest that distributional analyses of tolls should estimate impacts on all households in the relevant region in addition to impacts on just users of roads that are currently tolled or likely to be tolled.*

In planning regional transportation systems, policy makers balance overlapping and sometimes competing goals of effectiveness, cost-efficiency, environmental responsibility, and social equity. The increasingly popular strategy of tolling drivers on new facilities has implications for all these goals, and is one strategy among others for creating systems that can best move persons and goods through metropolitan areas.

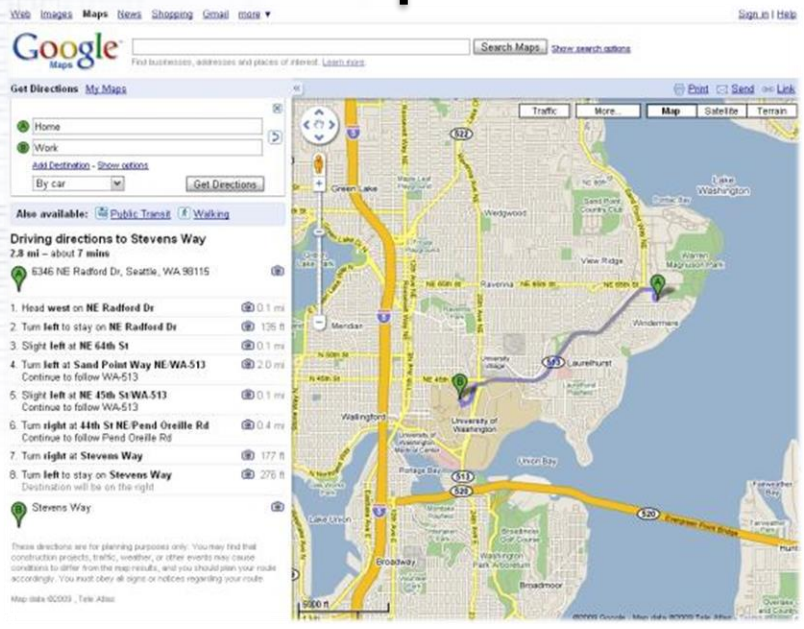
In an era of limited state and local budgets and legislators' reluctance to support higher fuel taxes or general tax increases, tolls on urban highways and bridges may be an attractive source of

*Direct correspondence to: Robert D. Plotnick, Evans School of Public Affairs, University of Washington Box 353055, Seattle, WA 98195-3055. E-mail: plotnick@uw.edu.*

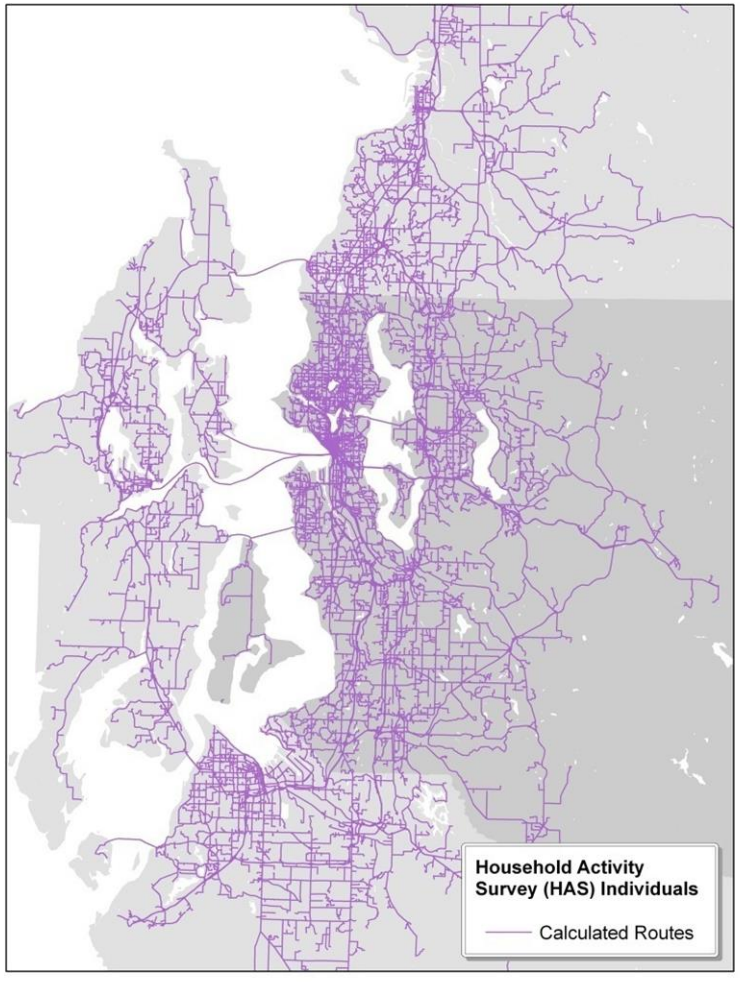
**JOURNAL OF URBAN AFFAIRS, Volume 00, Number 0, pages 1–22.**

**ISSN: 0735-2166.** **DOI: 10.1111/j.1467-9906.2011.00551.x**

# Mapping Commuting Routes

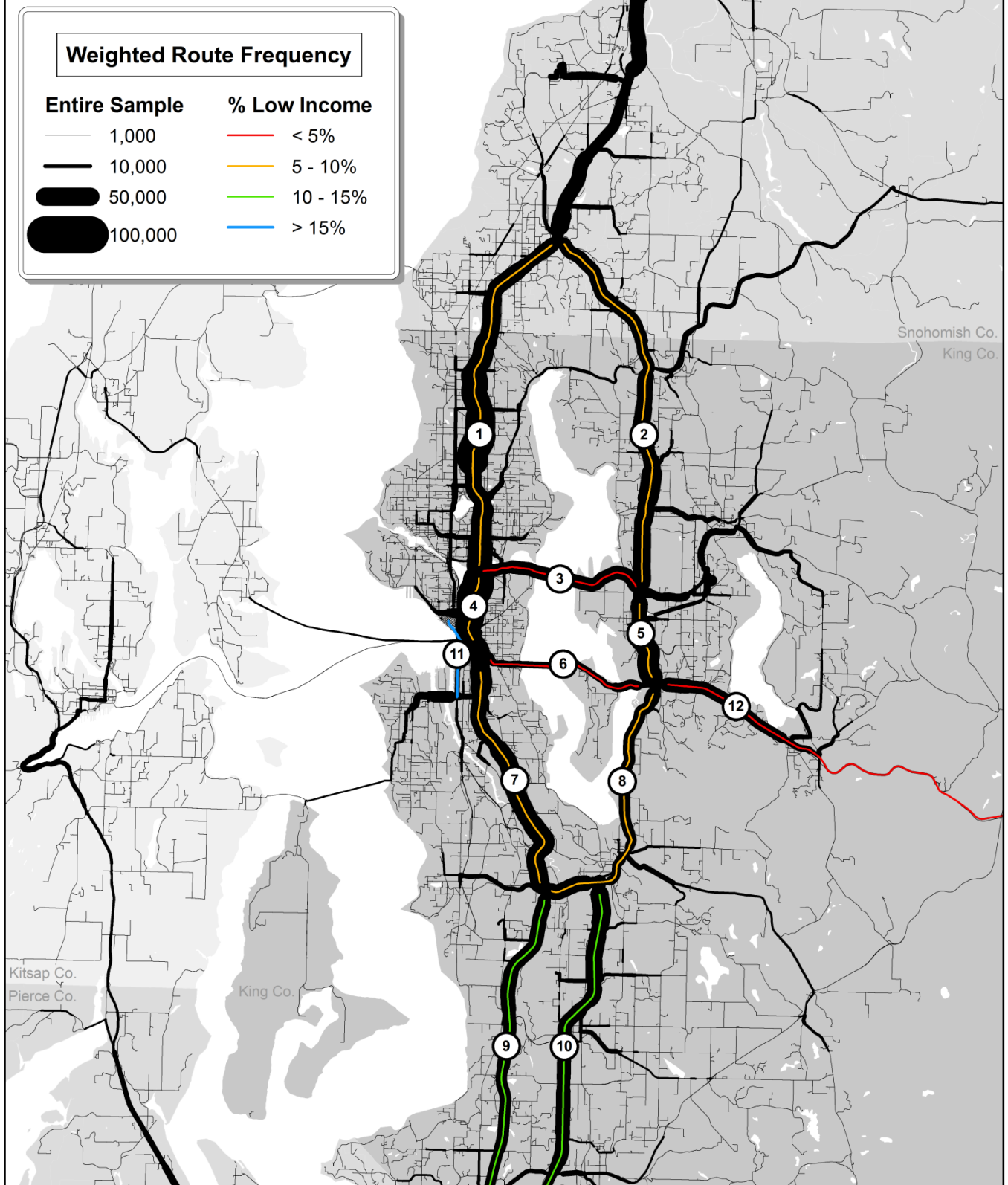



Home to Work Route
Single Individual (Google)

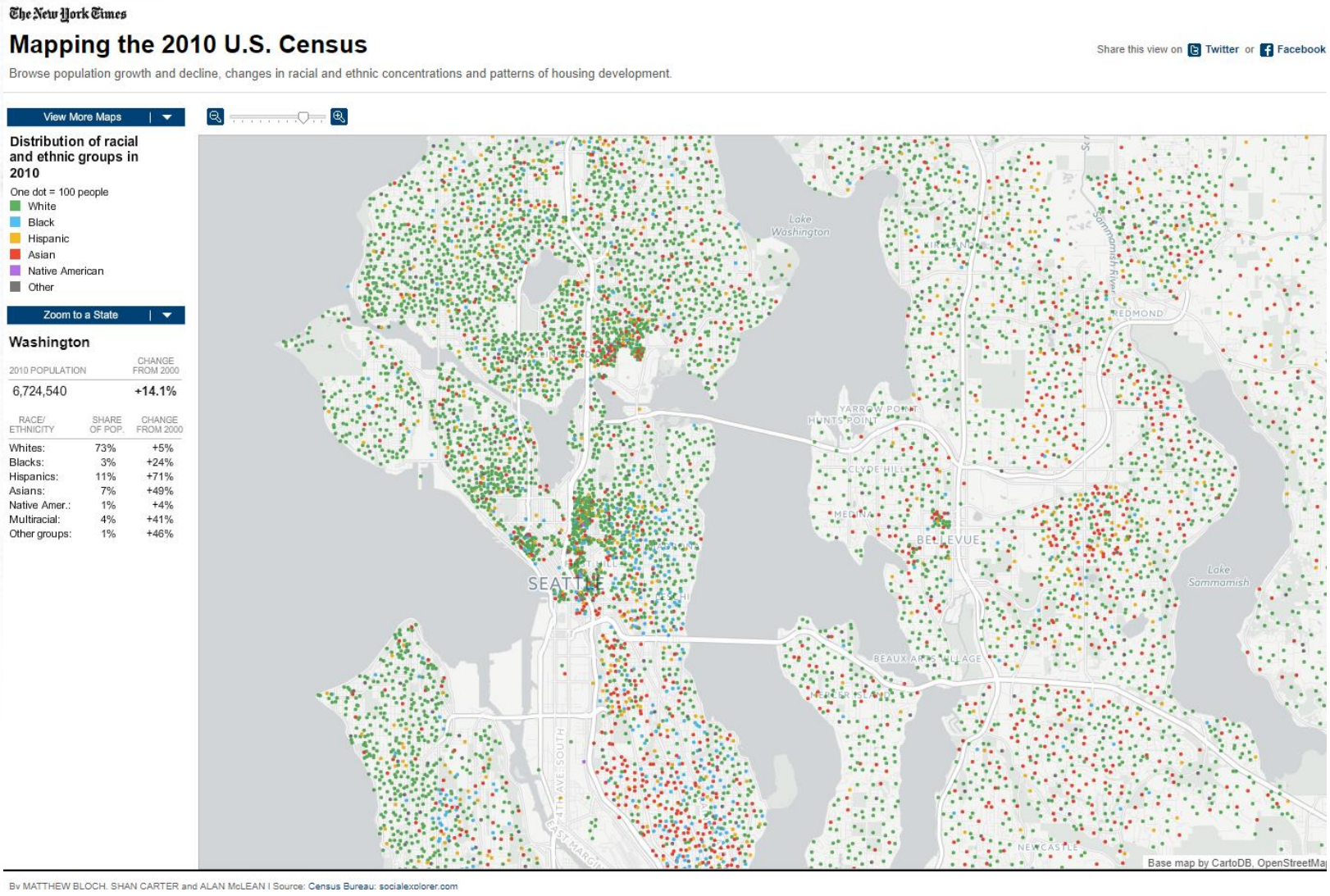



All 4,747 Respondents to Household Activity Survey (GIS + Tele Atlas Streets)

# Commuting Routes



# Mapping the 2010 U.S. Census




http://www.nytimes.com/projects/census/2010/map.html

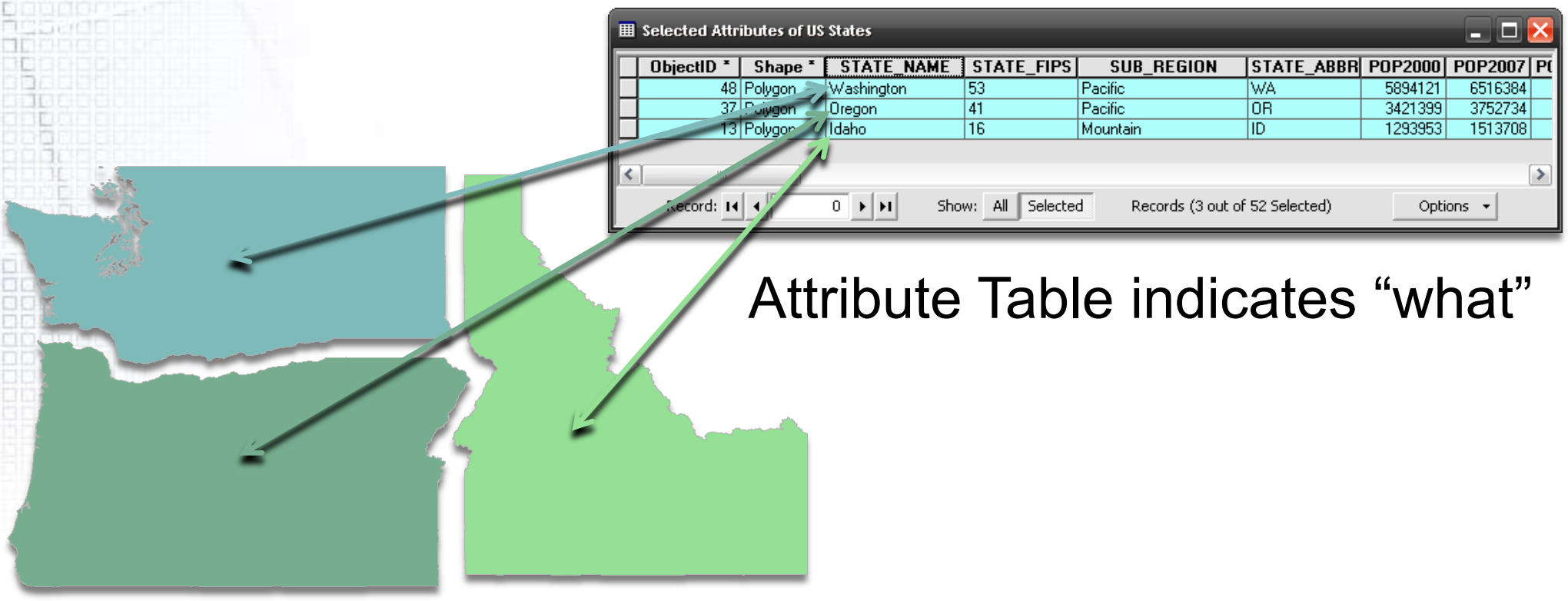
Tabular data from the US Census Bureau and pre-packaged web services from OpenStreetMap.org

# What is GIS?

Basics of Storing, Mapping, and Analyzing Spatial Data…

# Storing Geographic Data

One GIS data layer (US States) combines both Geographic Features and their Attributes




Attribute Table indicates “what”

Geographic Features indicate “where”

# Storing “Everyday” Geographical Objects

**Points**

home, day-care, health clinics, schools, retail and tobacco outlets, crimes & graffiti, bus stops, neighborhood anchor institutions, community assets, resources and risks

**Lines**

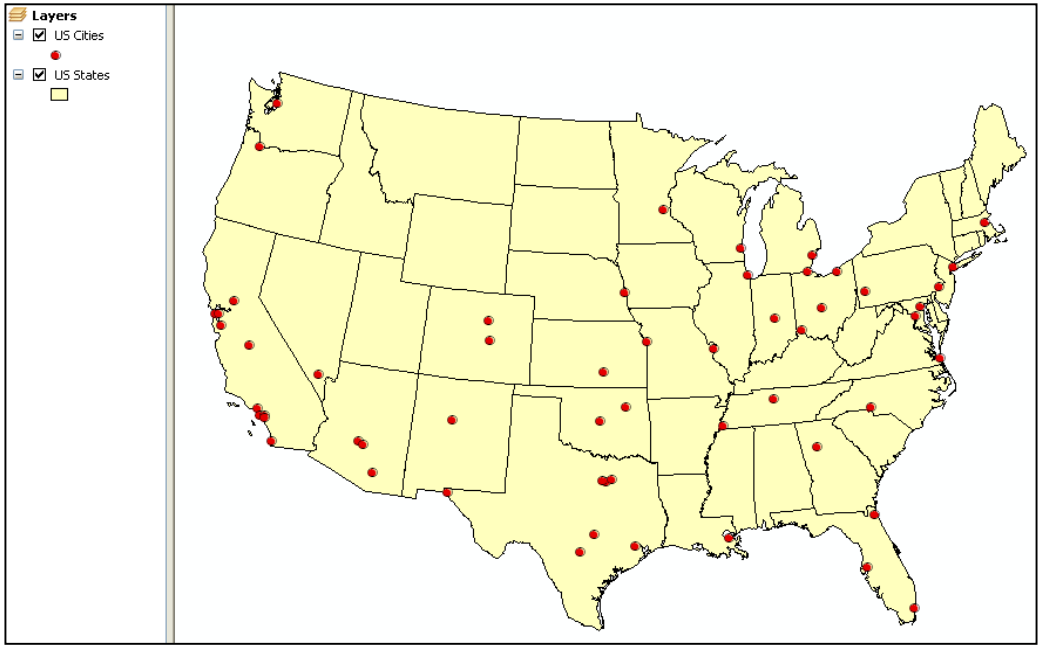
roads, railway, pathways, walking or bus routes, rivers

**Areas (Polygons)**

census unit, ZIP code, school district, police precinct, health service areas, counties, states, provinces, watersheds

# Mapping Geographic Data

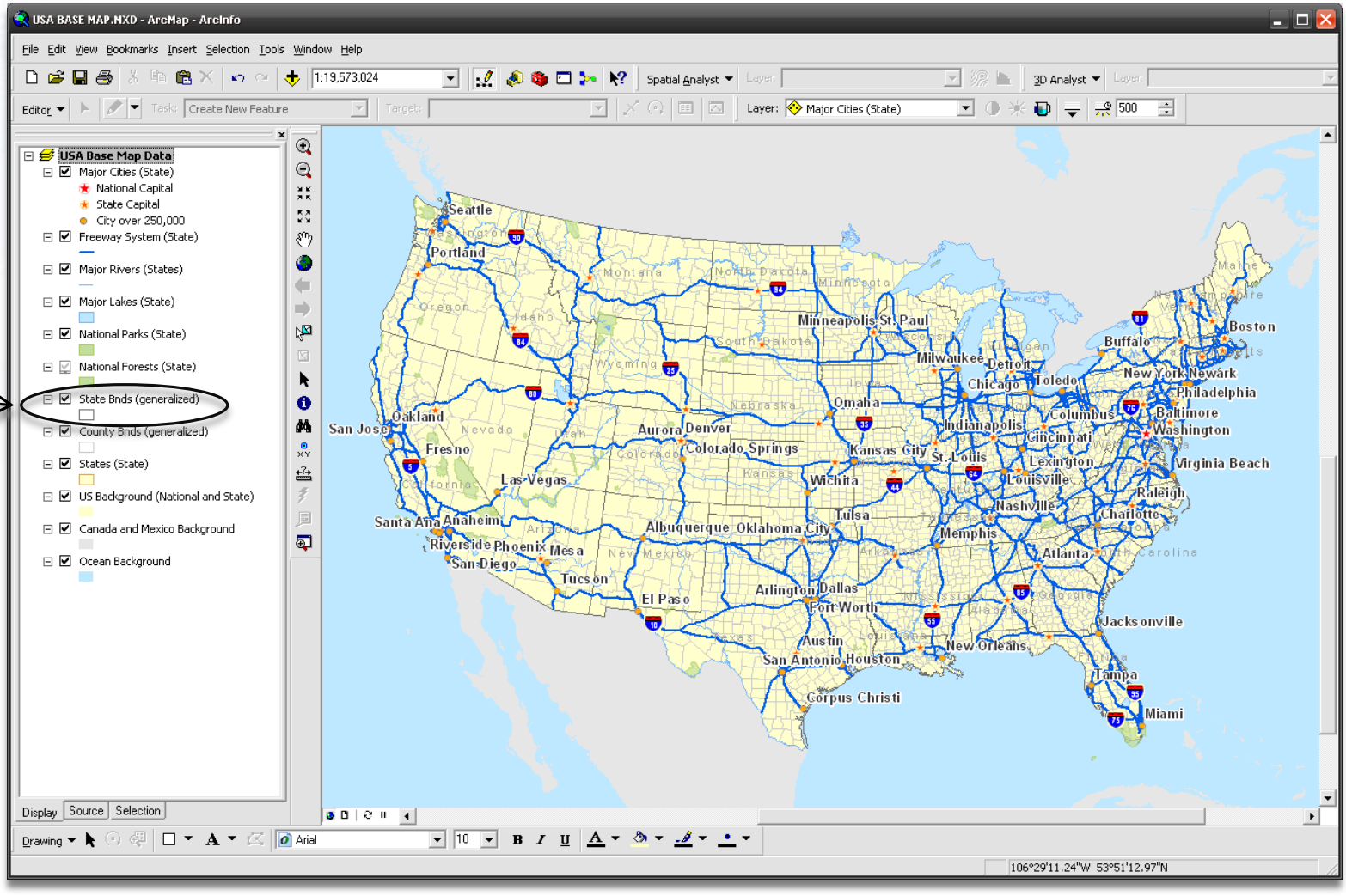
- USCities (red points, restricted to major cities)
- USStates (yellow polygons)



# Mapping Geographic Data

Select GIS data layers, how to symbolize, labels, map extent, etc.

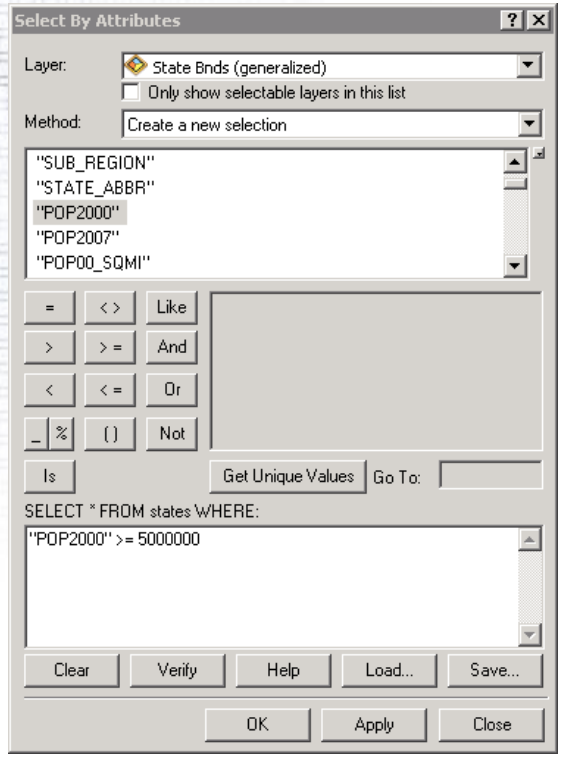
One GIS data layer

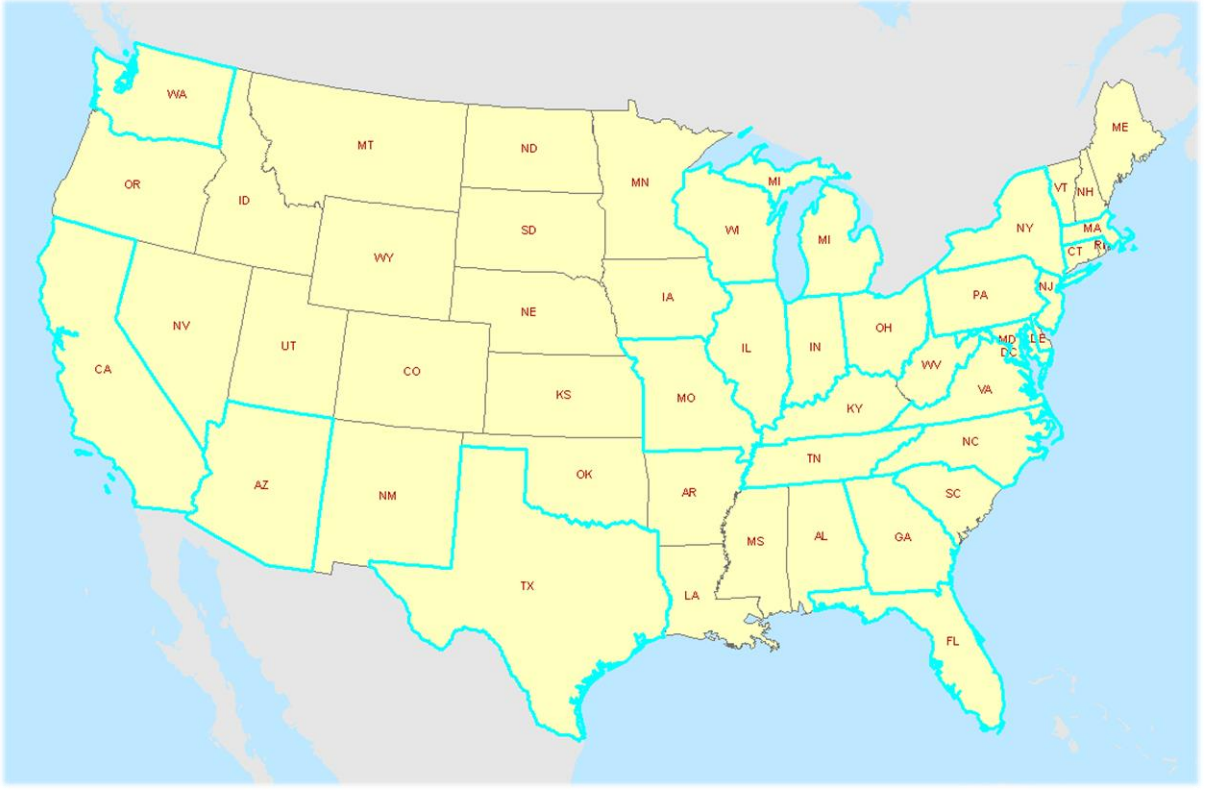


# Analyzing Geographic Data

Query GIS data layers based on attributes, geography, or both

1) Which states' population was more than 5 million in 2000?

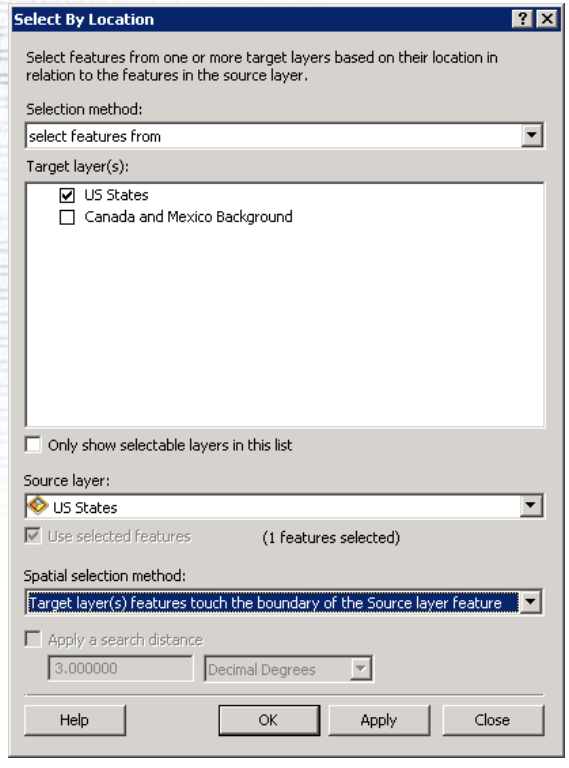


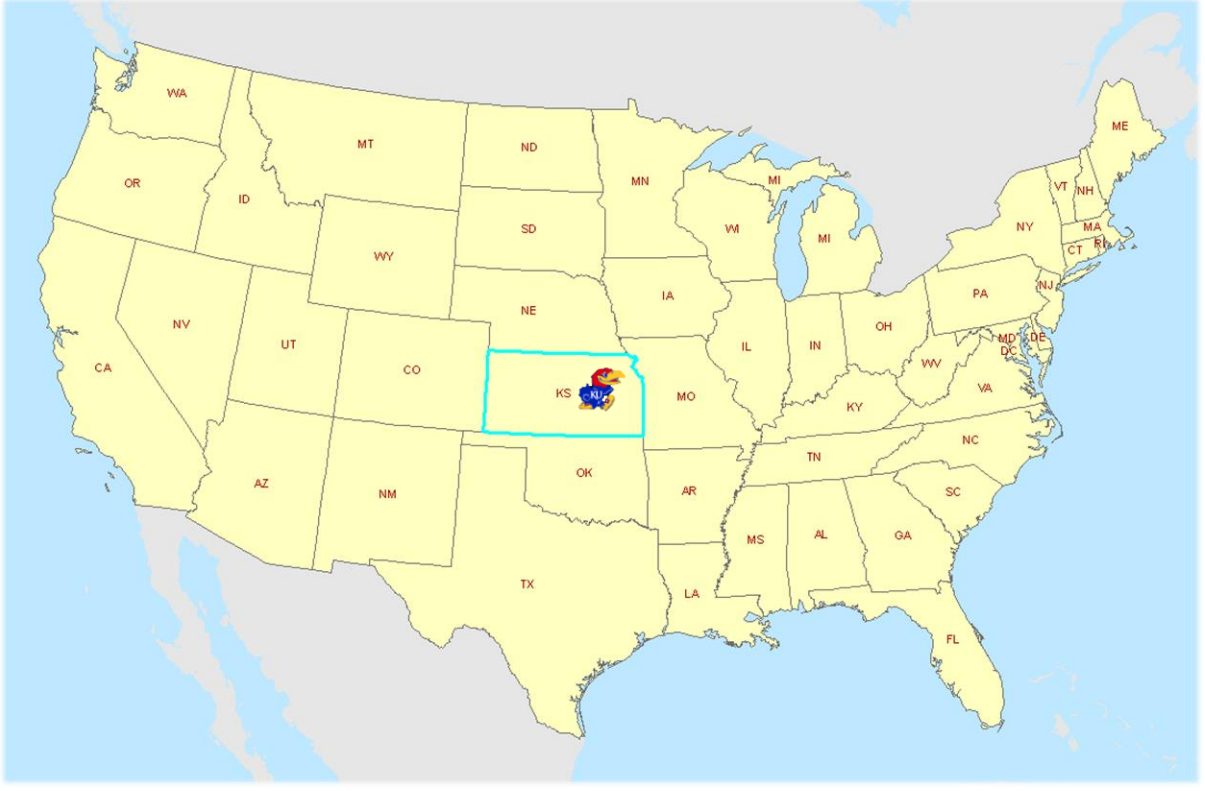



# Analyzing Geographic Data

Query GIS data layers based on attributes or geography, or both

2) Which states touch the border of Kansas?

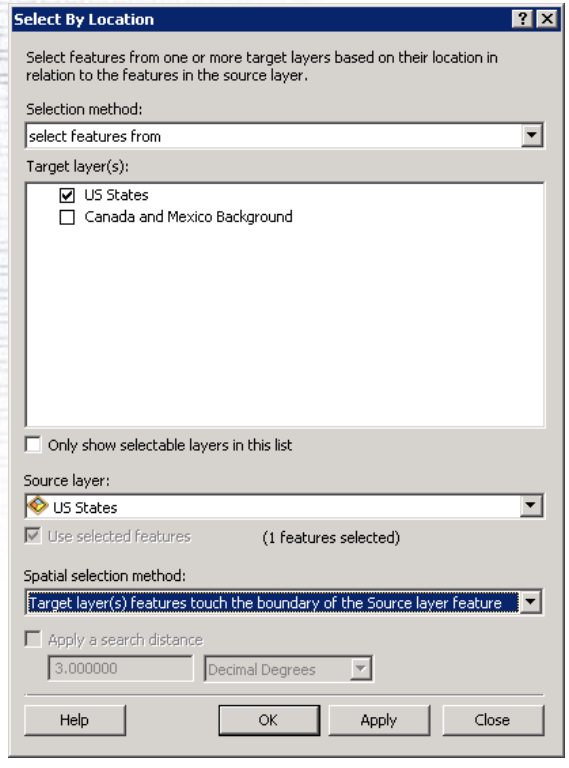


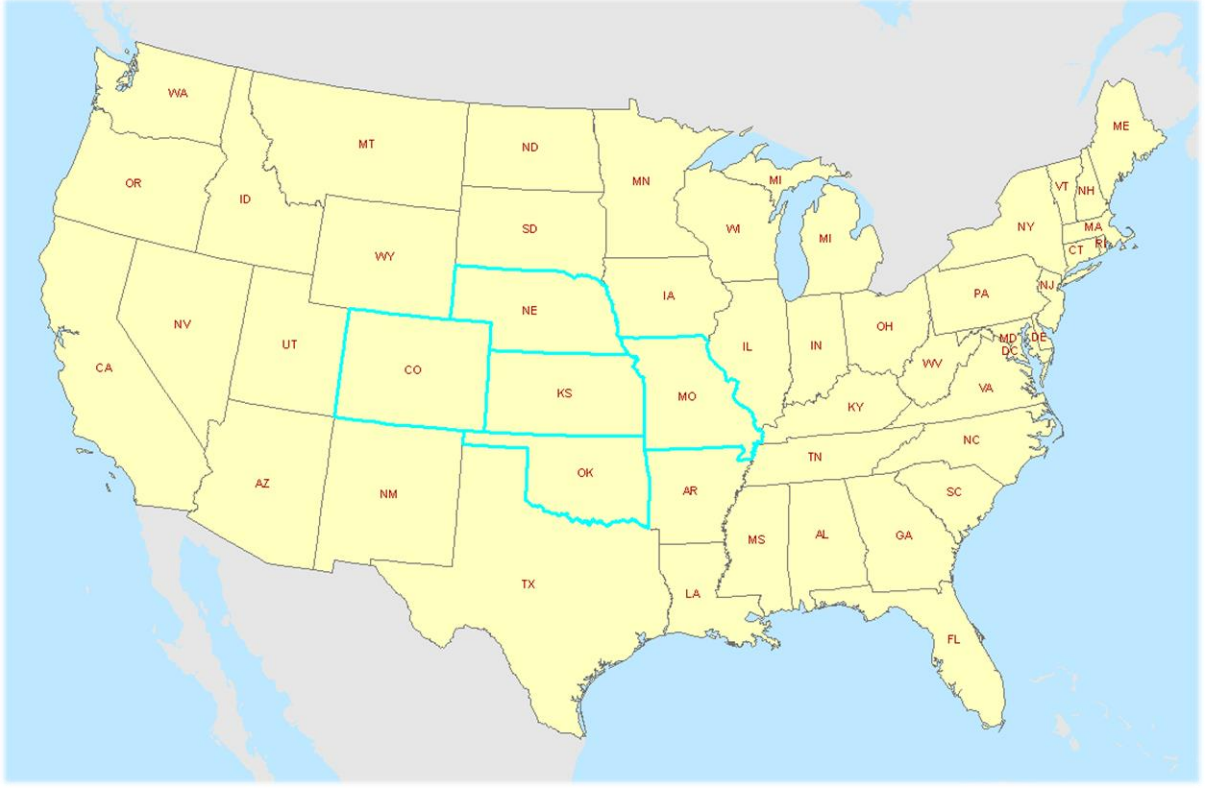



# Analyzing Geographic Data

Query GIS data layers based on attributes or geography, or both

2) Which states touch the border of Kansas?

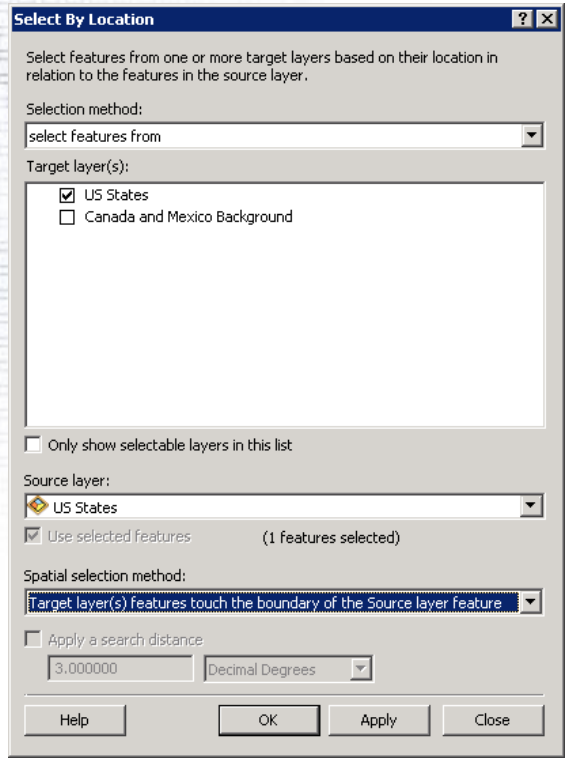


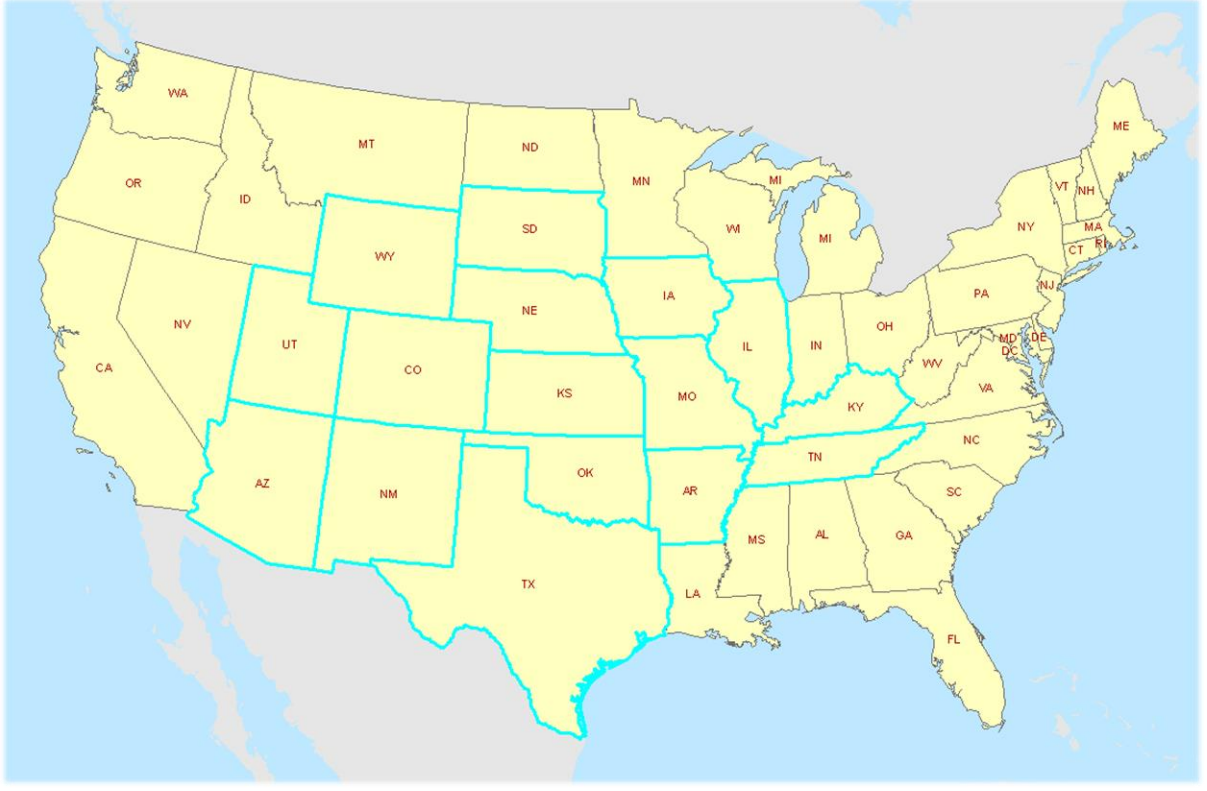



# Analyzing Geographic Data

Query GIS data layers based on attributes or geography, or both

3) Which states touch the border of Kansas border states?



# Analyzing Geographic Data

Query GIS data layers based on attributes or geography, or both

4) What is the distance from Kansas to 2nd order neighbors?

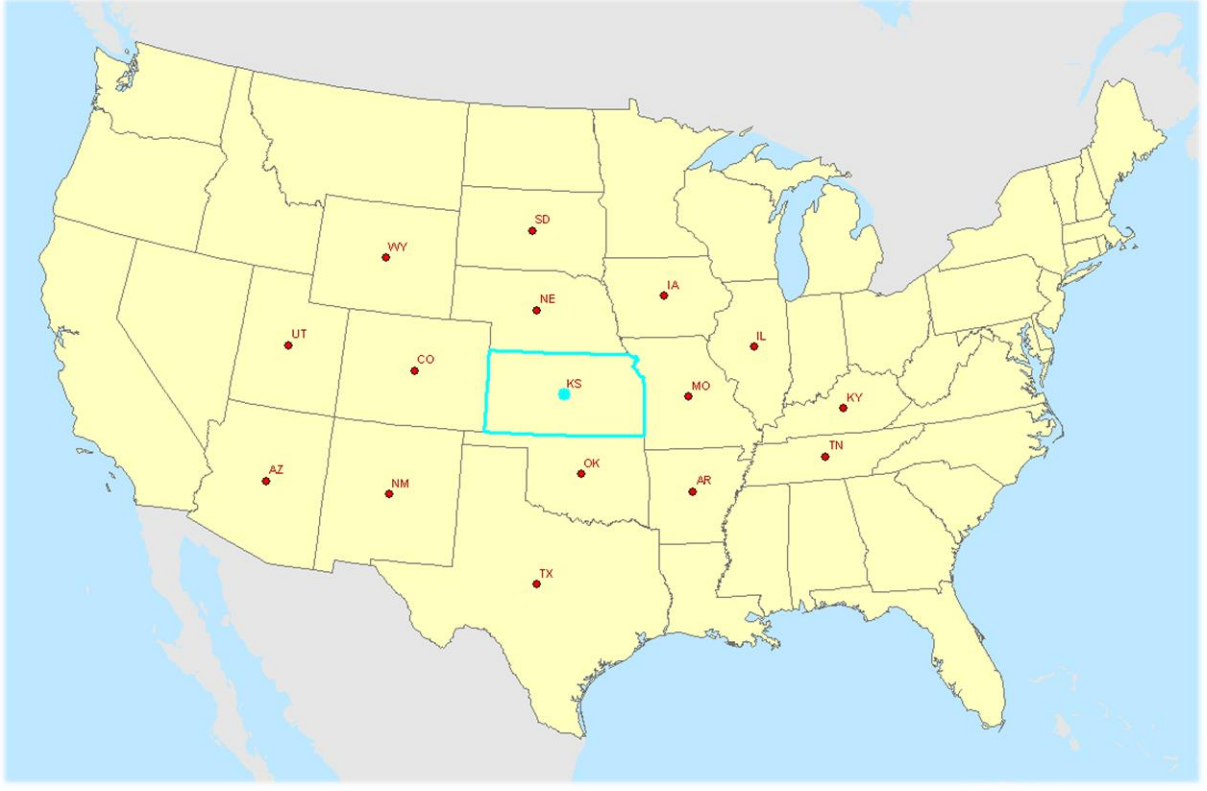
Table

Export_Output

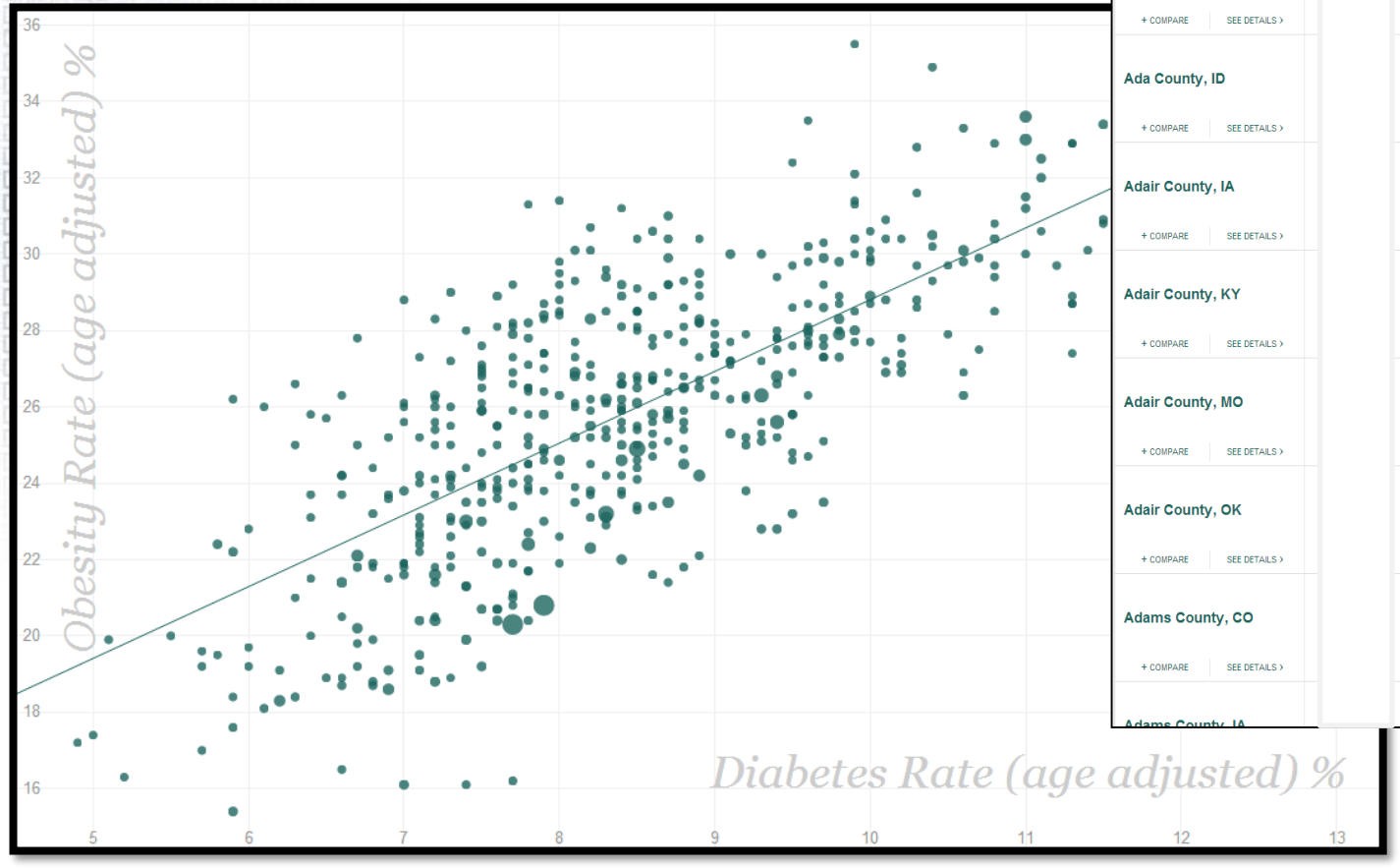
| STATE_NAME | DISTANCE | DISTANCE_MILES |
|---|---|---|
| Arizona | 1272924.156589 | 790.95837 |
| Arkansas | 661910.092654 | 411.29187 |
| Colorado | 621990.494685 | 386.48697 |
| Illinois | 806745.278383 | 501.28827 |
| Iowa | 574132.587986 | 356.74945 |
| Kentucky | 1147940.661541 | 713.29724 |
| Missouri | 512794.420804 | 318.63568 |
| Nebraska | 358616.704365 | 222.83409 |
| New Mexico | 822747.568369 | 511.23163 |
| Oklahoma | 330860.784034 | 205.58736 |
| South Dakota | 678572.428794 | 421.64536 |
| Tennessee | 1103525.966649 | 685.69922 |
| Texas | 780158.956915 | 484.76831 |
| Utah | 1149552.732445 | 714.29895 |
| Wyoming | 918758.804947 | 570.89026 |

0 (0 out of 15 Selected)

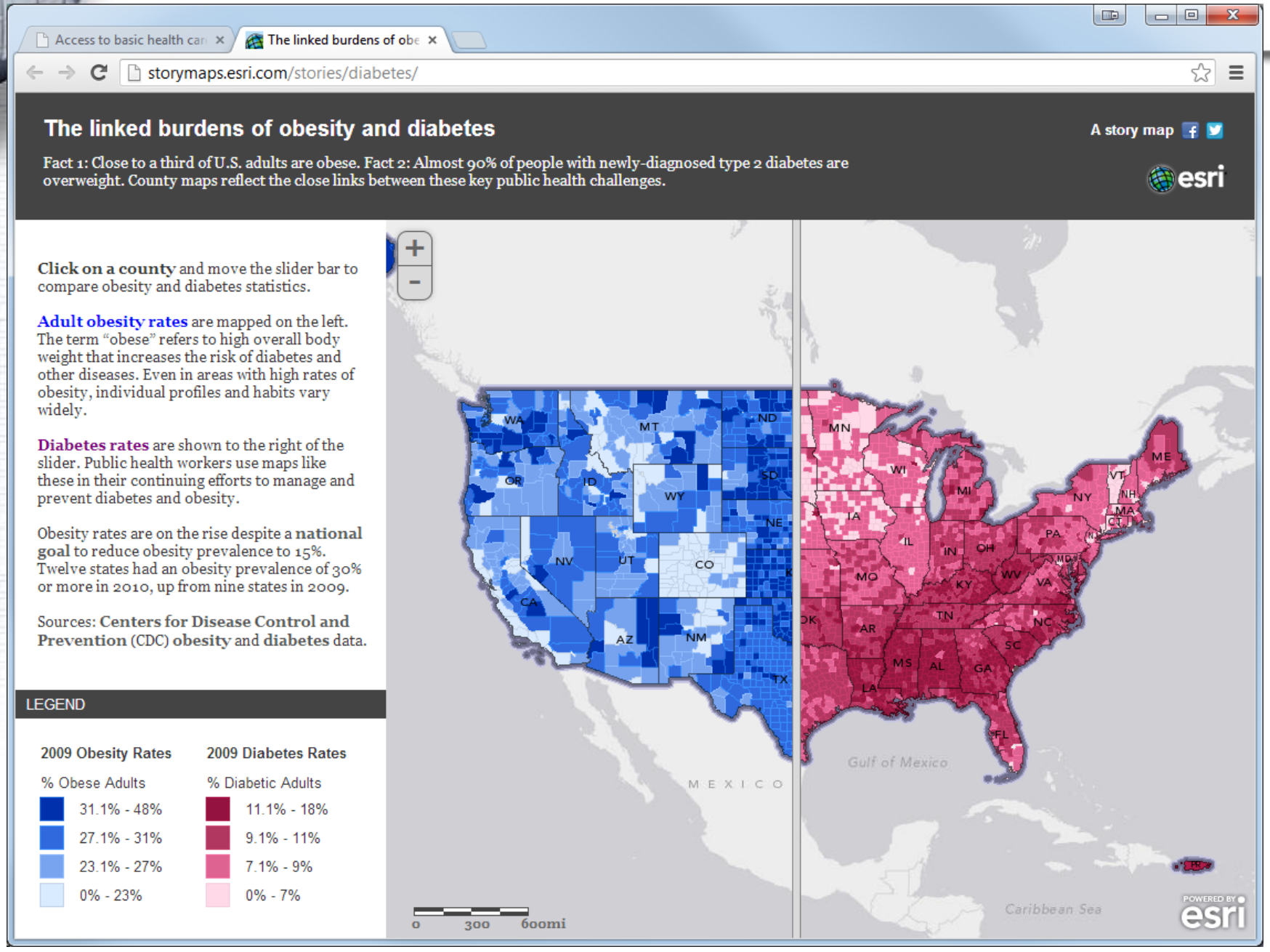
Export_Output



# Storing, Mapping, Analyzing Example: County Obesity and Diabetes Rates



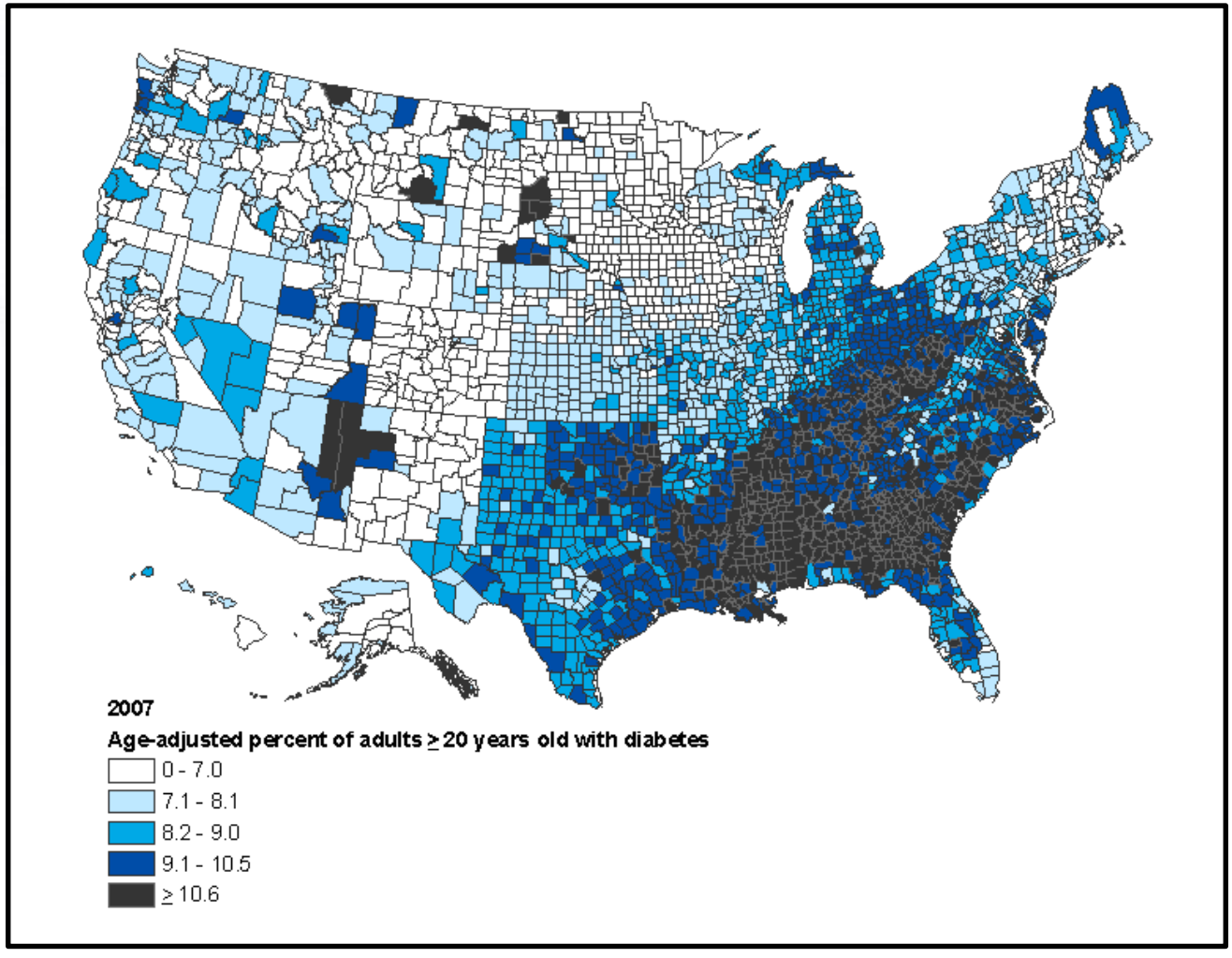

| COUNTY | DIABETES RATE % | OBESITY RATE % | PHYSICAL INACTIVITY RATE % | TOTAL DIABETES #PEOPLE | TOTAL OBESITY #PEOPLE | TOTAL PHYSICALLY INACTIVE #PEOPLE | YEAR |
|---|---|---|---|---|---|---|---|
| Abbeville County, SC | 13.4 | 30.3 | 32.5 | 2,521 | 5,693 | 6,114 | 2008 |
| Acadia Parish, LA | 10.3 | 31 | 28.5 | 4,318 | 12,960 | 11,940 | 2008 |
| Accomack County, VA | 11.8 | 31.4 | 29.8 | 3,393 | 9,019 | 8,550 | 2008 |
| Ada County, ID | 7.1 | 22 | 15.3 | 19,170 | 59,810 | 41,570 | 2008 |
| Adair County, IA | 9.3 | 27.7 | 27.1 | 542 | 1,620 | 1,587 | 2008 |
| Adair County, KY | 12.9 | 32.6 | 38.4 | 1,713 | 4,326 | 5,097 | 2008 |
| Adair County, MO | 8 | 28.8 | 28.6 | 1,446 | 5,193 | 5,150 | 2008 |
| Adair County, OK | 13.3 | 33.7 | 33.5 | 1,980 | 5,019 | 4,985 | 2008 |
| Adams County, CO | 6.5 | 24.3 | 20 | 19,230 | 72,220 | 59,500 | 2008 |
| … | … | … | … | … | … | … | … |




http://storymaps.esri.com/stories/diabetes/

# Instruments for Reasoning about Quantitative Information

Age-Adjusted Percentage of Adults Aged ≥20 Years with Diagnosed Diabetes, 2007

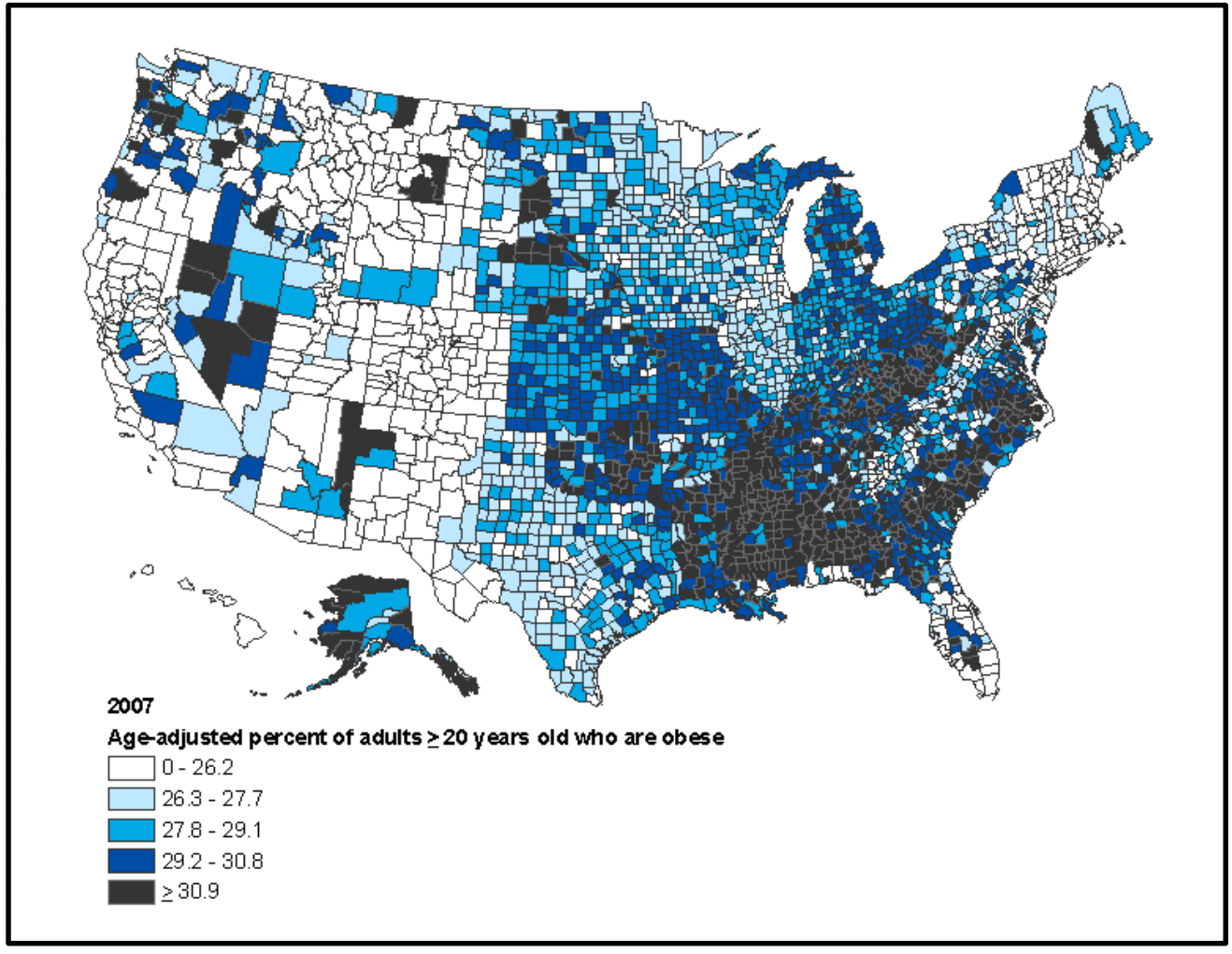



*MMWR* 2009;58:1259-1263

CDC's Division of Diabetes Translation. National Diabetes Surveillance System available at http://www.cdc.gov/diabetes/statistics

# Instruments for Reasoning about Quantitative Information

Age-Adjusted Percentage of Adults Aged ≥20 Years Who are Obese, 2007

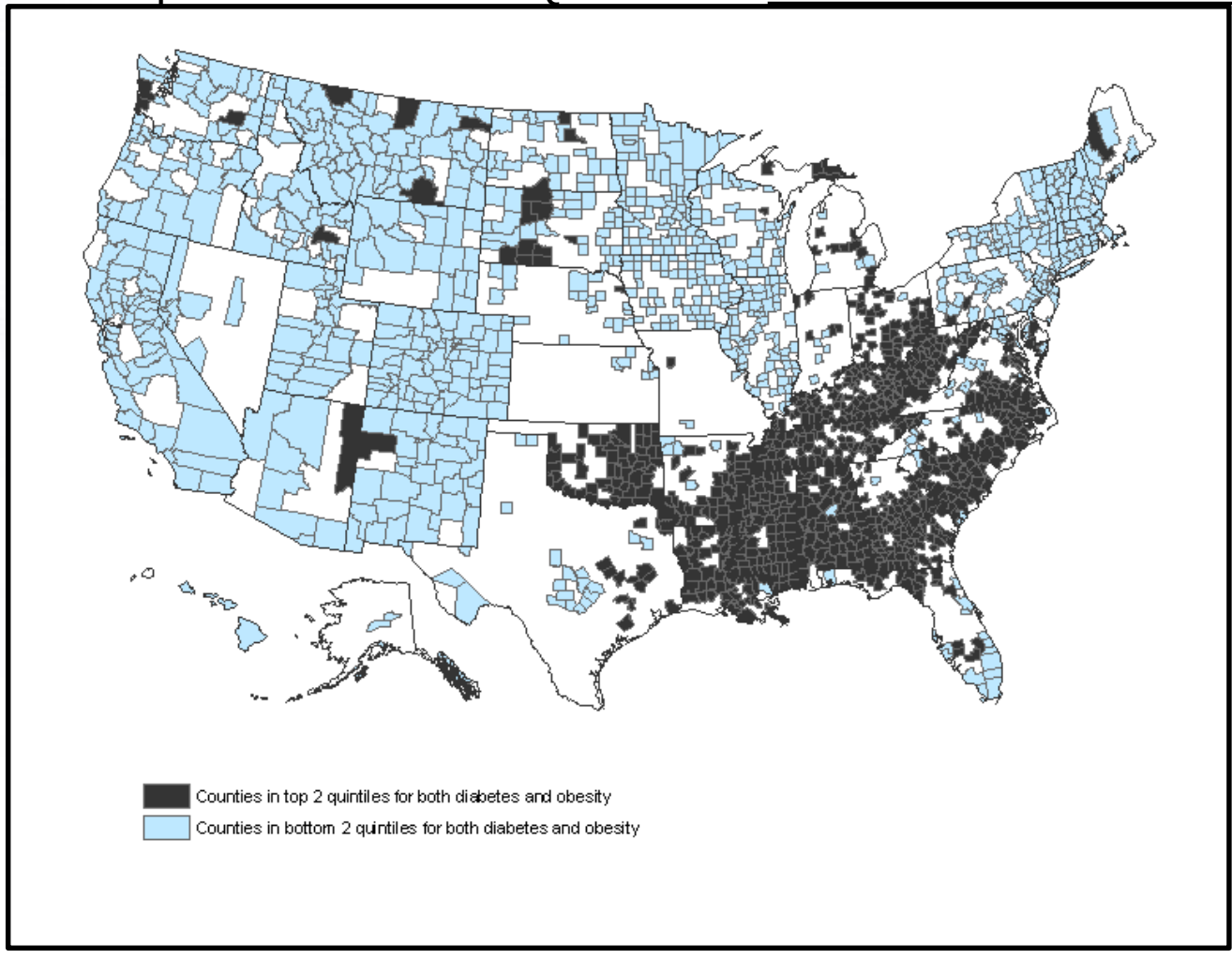



*MMWR* 2009;58:1259-1263

CDC's Division of Diabetes Translation. National Diabetes Surveillance System available at http://www.cdc.gov/diabetes/statistics

# Instruments for Reasoning about Quantitative Information

Counties in the Top and Bottom Two Quintiles of Both Diabetes and Obesity, 2007



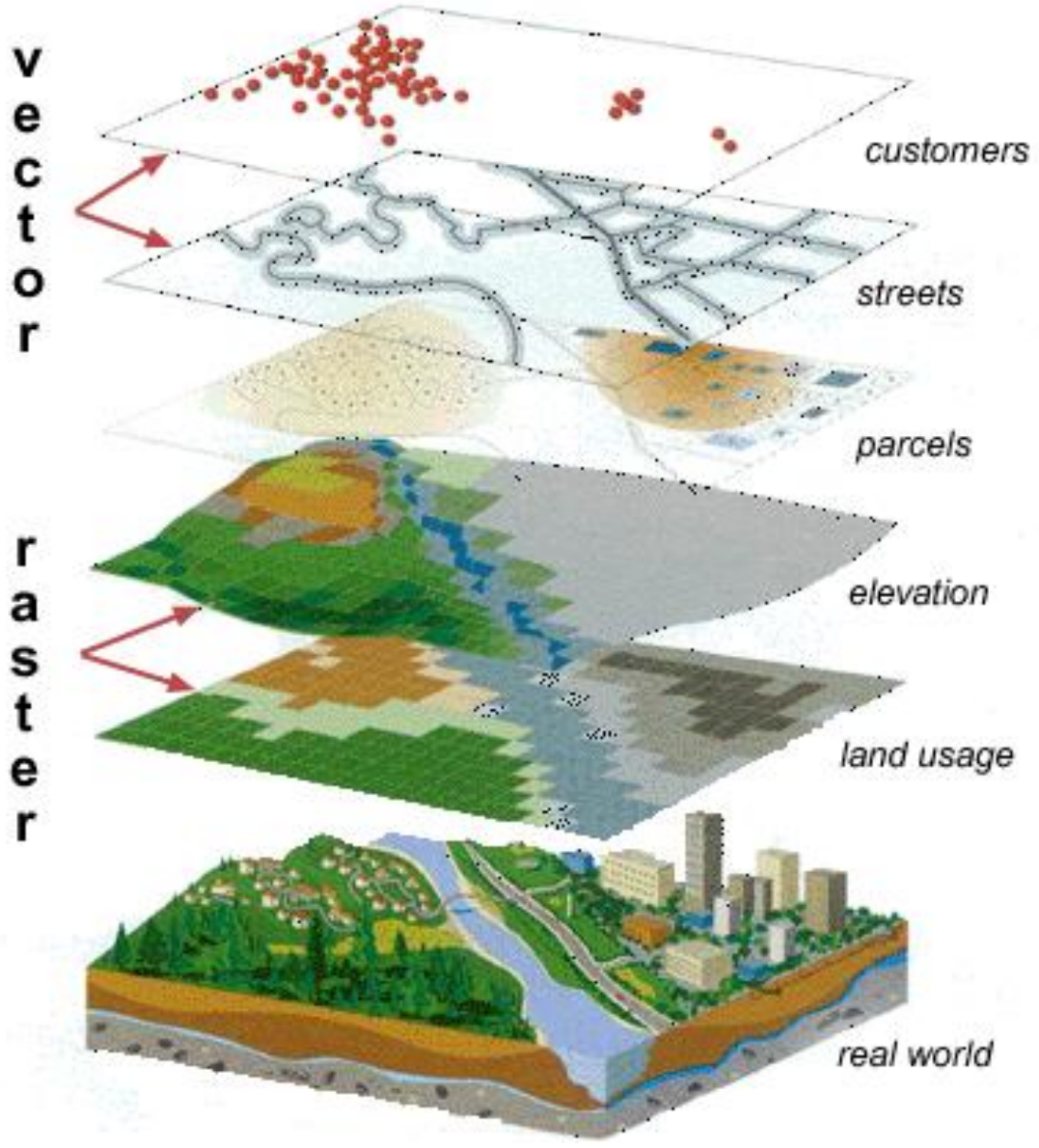

*MMWR* 2009;58:1259-1263

CDC's Division of Diabetes Translation. National Diabetes Surveillance System available at http://www.cdc.gov/diabetes/statistics

# What is GIS?

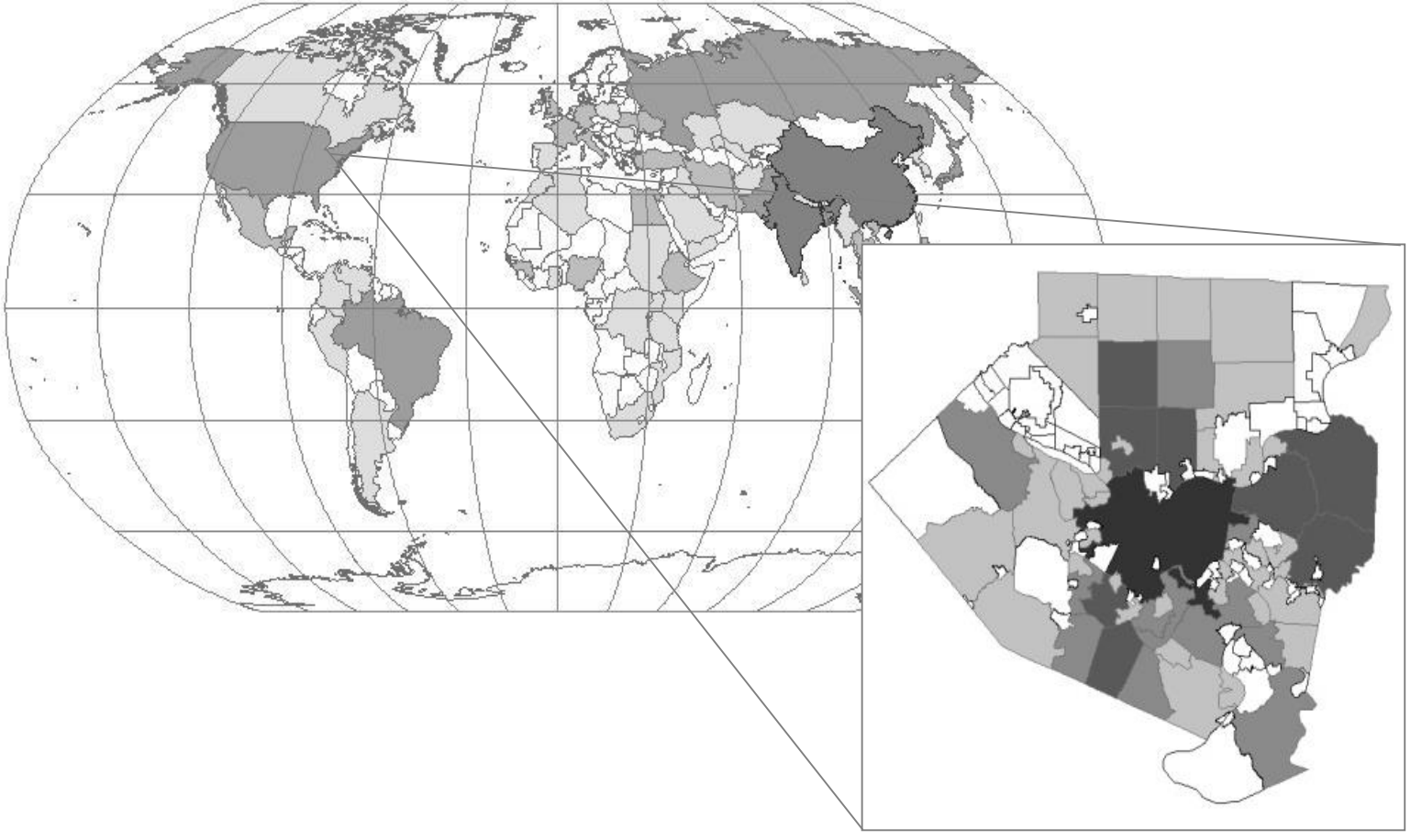
In more detail…

# Representing Spatial Elements



# Scale of GIS data

Global to local

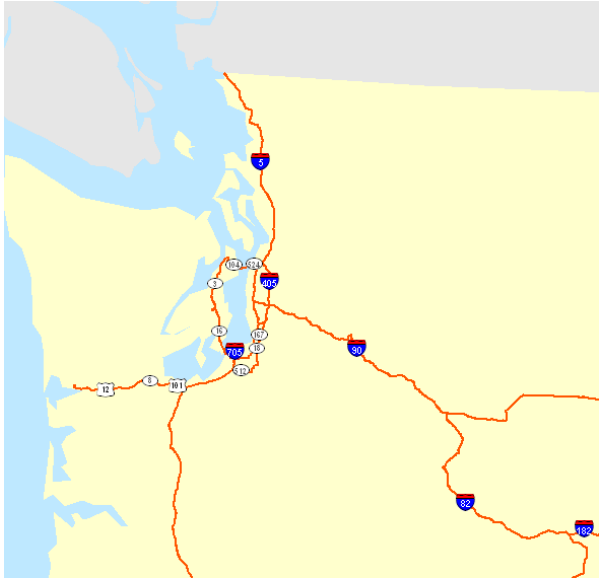


# Vector (Data Type 1/3)

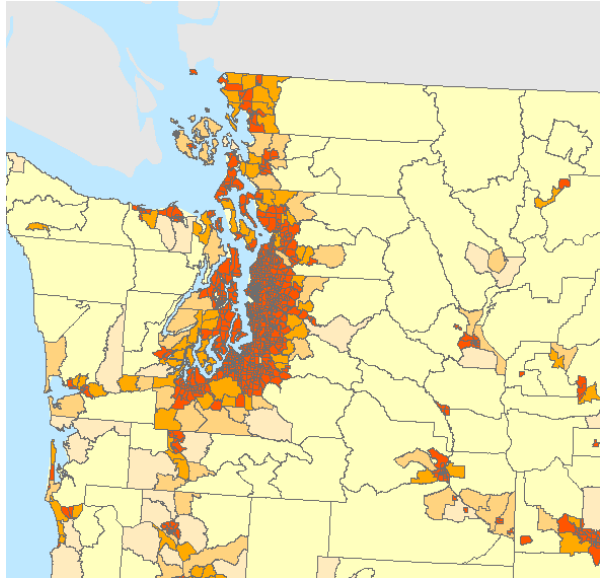
Vector data consists of discrete coordinates which can store the geographic position of points, lines and polygons

Points: People or Cities (center)


Lines: Roads or Other Linkages

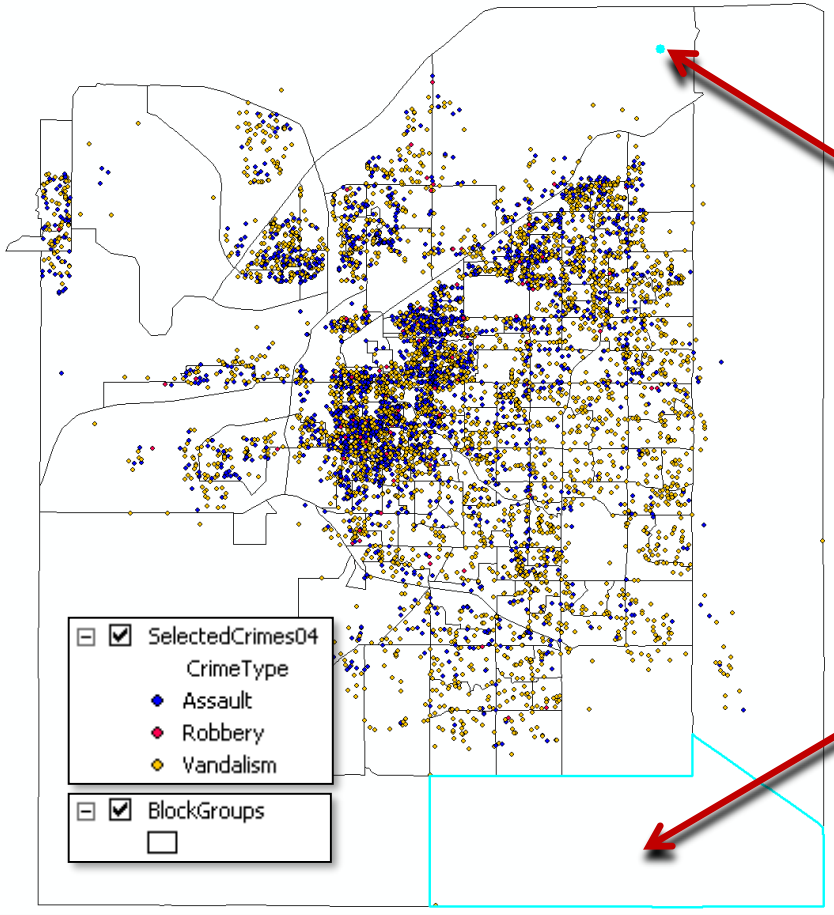



Polygons: Census Tract



# Vector Data Model

Geographic features are stored as X,Y coordinate pairs. Each vector layers has an attribute table, and each feature corresponds to a row in the table.




Attributes of SelectedCrimes04

| OBJECTID * | Shape * | CrimeType | CrimeTime | CrimeDate |
|---|---|---|---|---|
| 1 | Point | Assault | 0017 | 20040101 |
| 2 | Point | Robbery | 0112 | 20040101 |
| 3 | Point | Vandalism | 1036 | 20040126 |
| 4 | Point | Assault | 2347 | 20040214 |
| 5 | Point | Vandalism | 1215 | 20040220 |
| 6 | Point | Vandalism | 0050 | 20040225 |
| 7 | Point | Assault | 2352 | 20040226 |
| 8 | Point | Vandalism | 1815 | 20040415 |
| 9 | Point | Vandalism | 2339 | 20040421 |
| 10 | Point | Assault | 0703 | 20040602 |
| 11 | Point | Vandalism | 0800 | 20040716 |

Record: 6 Show: All Selected Records

Attributes of BlockGroups

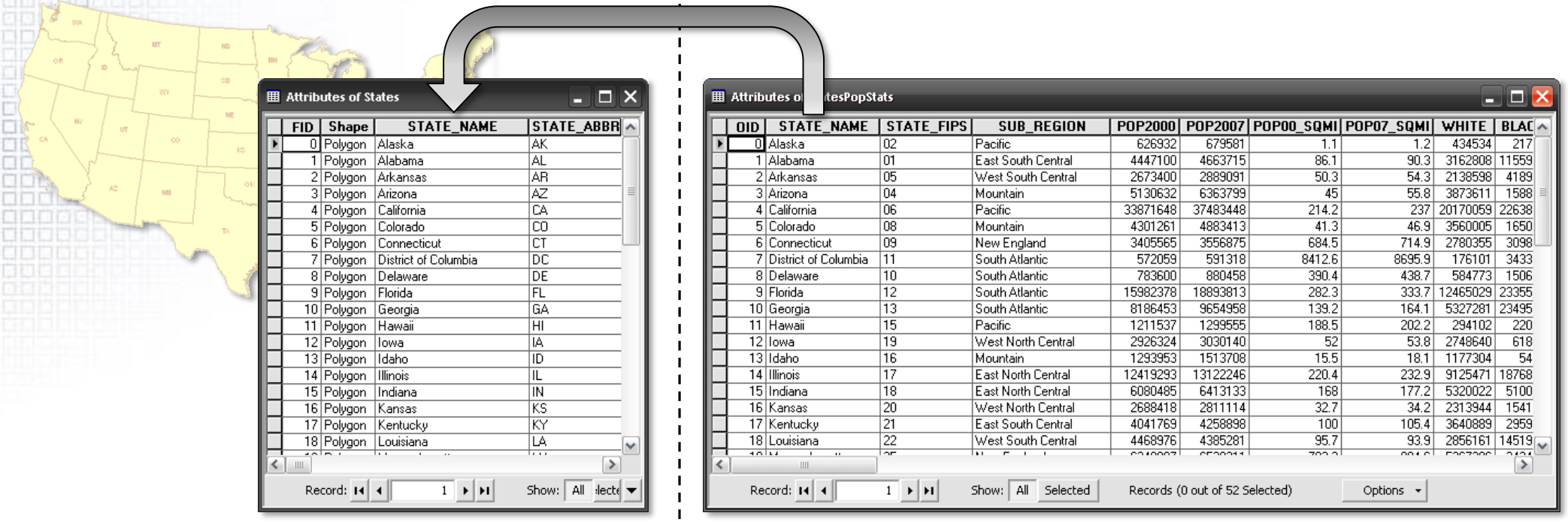
| OBJECTID * | Shape * | Shape_Length | Shape_Area | POP |
|---|---|---|---|---|
| 1 | Polygon | 19023.164382 | 16707150.211607 | 2843 |
| 2 | Polygon | 18344.586691 | 11844662.127896 | 1321 |
| 3 | Polygon | 12748.323138 | 8223385.482477 | 1834 |
| 4 | Polygon | 9613.753427 | 5794705.656414 | 1188 |
| 5 | Polygon | 12463.490459 | 7484094.115933 | 1378 |
| 6 | Polygon | 9896.231559 | 6016520.399411 | 1143 |
| 7 | Polygon | 20952.656815 | 27663080.254551 | 2741 |
| 8 | Polygon | 21192.478293 | 28175433.276053 | 1428 |
| 9 | Polygon | 42913.532763 | 111108699.017346 | 2143 |
| 10 | Polygon | 149870.527367 | 662027895.082628 | 3105 |
| 11 | Polygon | 86429.245137 | 331921312.368605 | 1361 |

Record: 1 Show: All Selected Records

# Data Tables (Data Type 2/3)

No inherent geographic information.

May be linked to a geographic data layer that shares a common attribute field.




Vector Data Layer

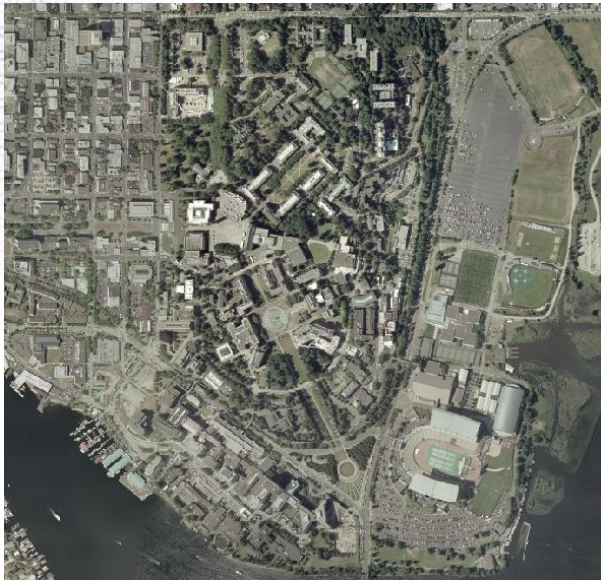
Data Table (.dbf, .xls, or .mdb)

# Raster (Data Type 3/3)

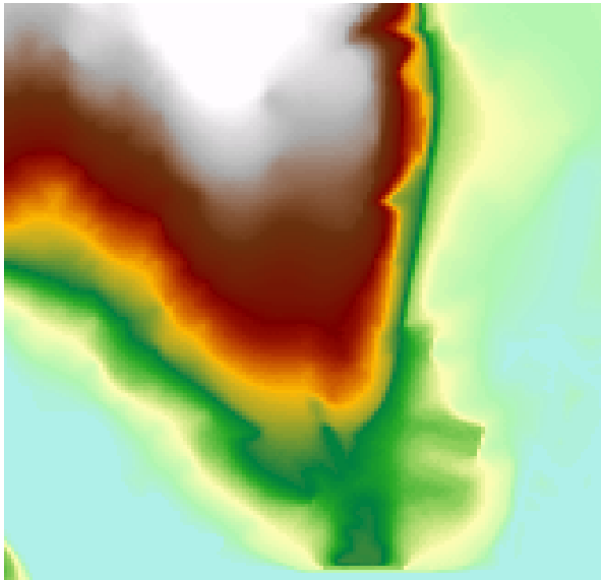
Raster data represents a continuous surface divided into a regular grid of cells.

Often used as background map layer.

Air Photo



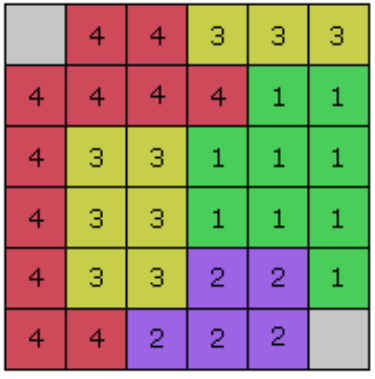
Elevation Model

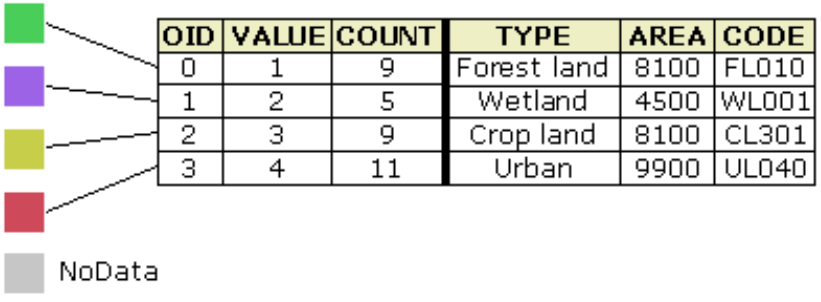


# Raster Data Model

Stores images as rows and columns of numbers, forming a regular grid structure.

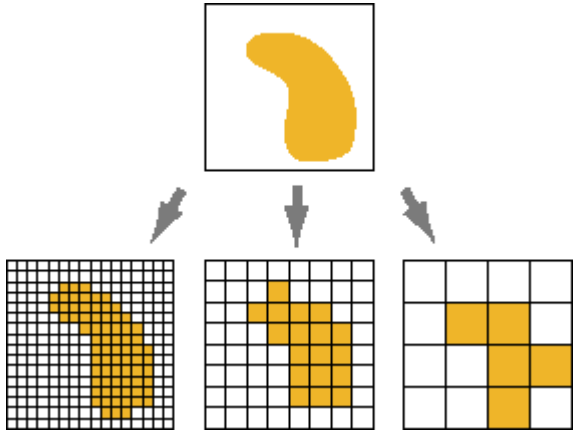
Great for computational analysis or modeling, bad for mapping precise locations.




| OID | VALUE | COUNT | TYPE | AREA | CODE |
|---|---|---|---|---|---|
| 0 | 1 | 9 | Forest land | 8100 | FL010 |
| 1 | 2 | 5 | Wetland | 4500 | WL001 |
| 2 | 3 | 9 | Crop land | 8100 | CL301 |
| 3 | 4 | 11 | Urban | 9900 | UL040 |

Raster Attribute Tables



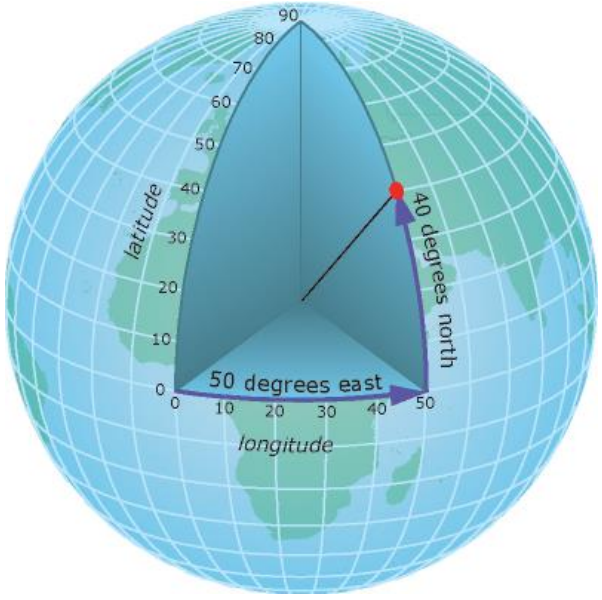
Vector vs. Raster
(& raster resolution)

# Coordinate Systems

Describing the correct location and shape of features requires a framework for defining real-world locations.

A geographic coordinate system is used to assign geographic locations to objects.

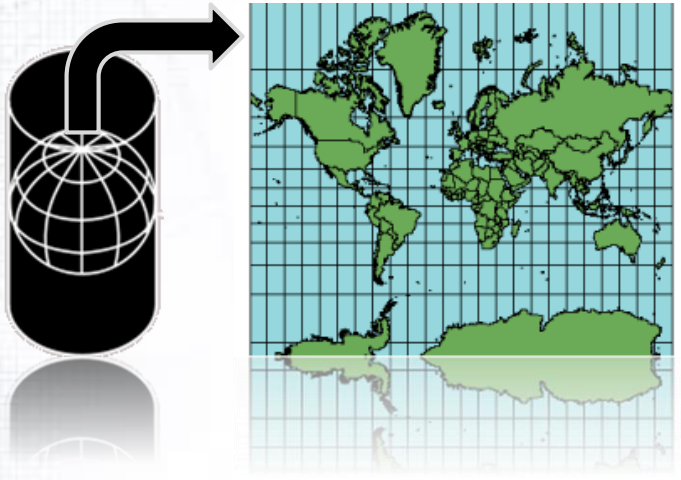
GIS data layers must have a coordinate system defined to integrate with other layers



# Map Projection

Transforming 3-dimensional space (Earth) onto a 2-dimensional map (GIS)

Mercator

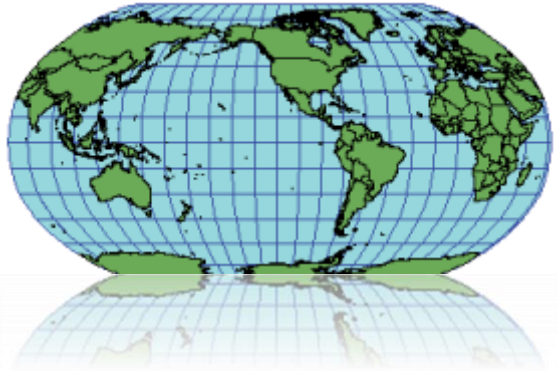


Azimuthal Equidistant

Lambert Conformal Conic

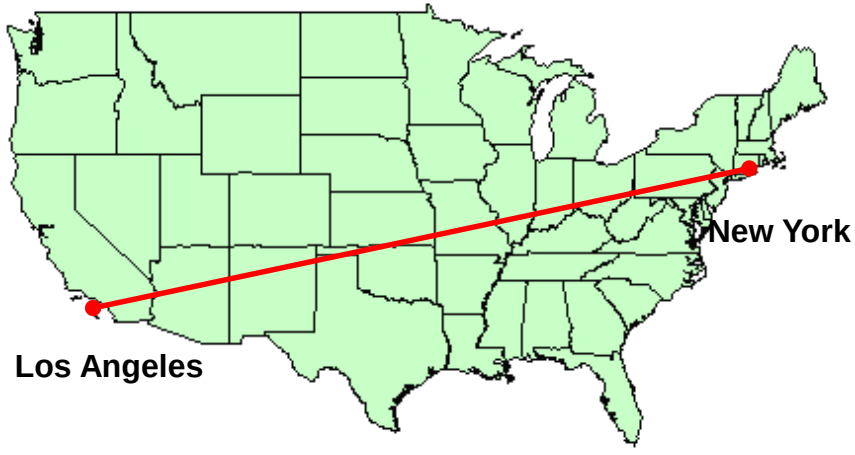
Albers Equal Area Conic

Robinson

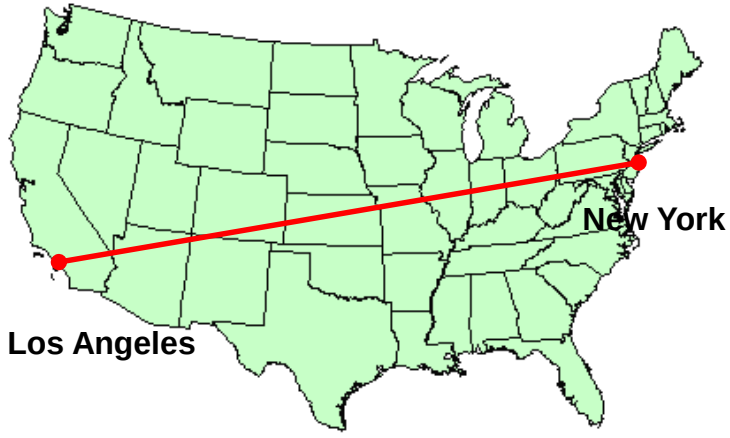


# Map Projection is Important

- Small-scale (large area) maps
  - Interested in Comparing shapes, areas, distances, or directions of map features?
  - Measurement errors can be quite substantial:



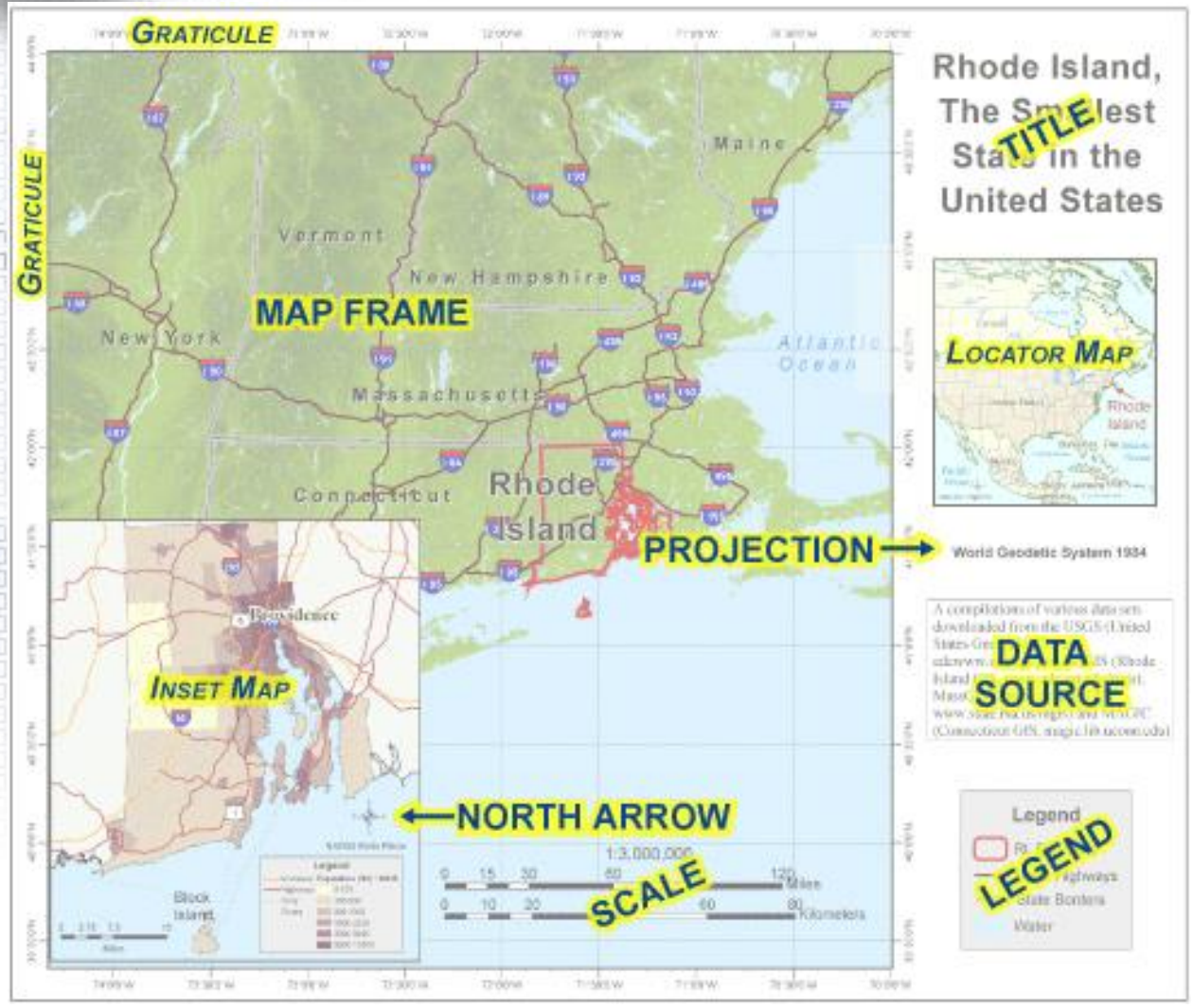

Projection: Mercator
Distance: 3,124.67 miles




Projection: Albers Equal Area
Distance: 2,455.03 miles

***Actual distance: 2,451 miles***

# Map Making: It's Complex



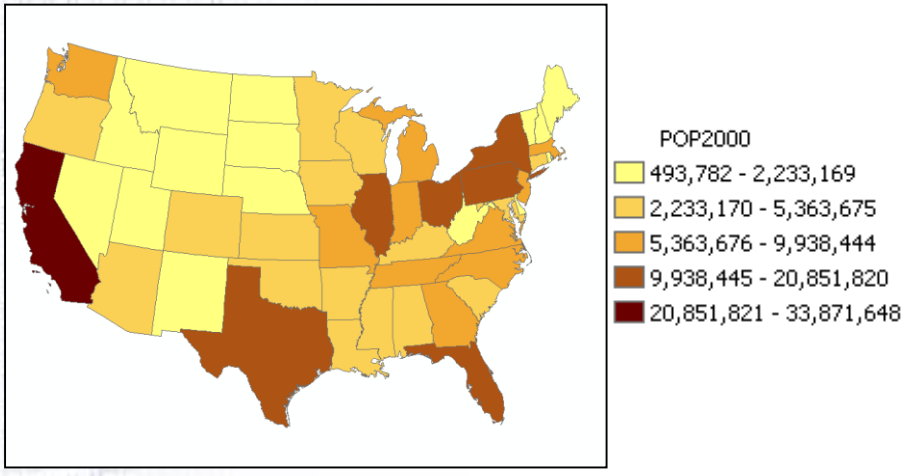

Many elements & design theory

Sub-discipline in Geography: Cartography

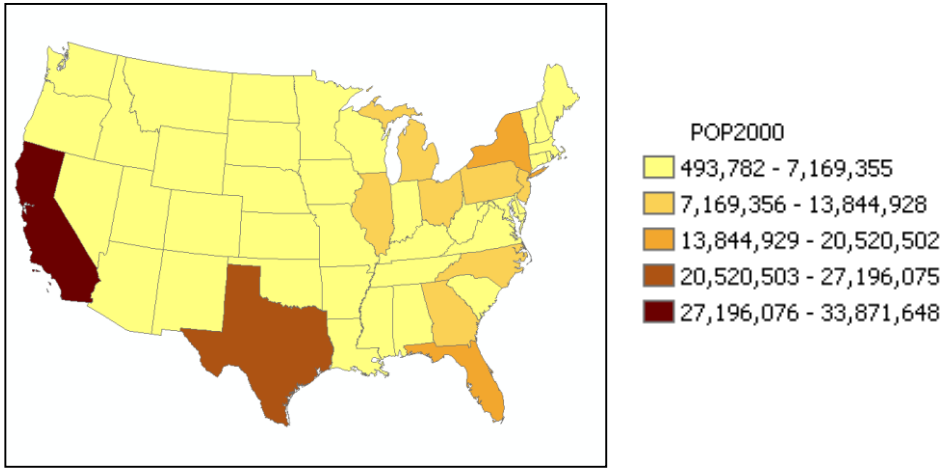
Tip: Look at Other Maps

# Choosing Map Classification

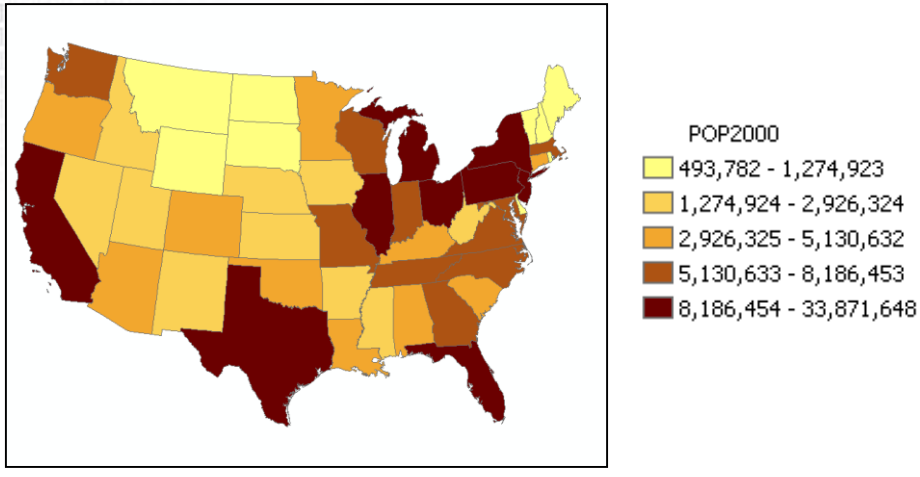
**Natural breaks**



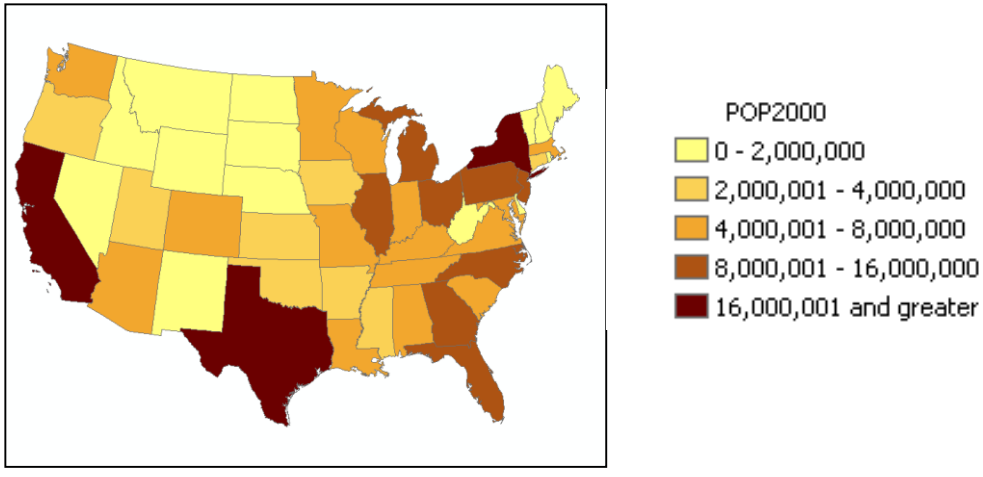

**Equal Interval**




**Quantiles**




**Exponential**



# Map Making: With Great Power…

## Factors influencing a Map's Message:

- Colors, tones and shading
- Font sizes and styles
- Map projection and distortion
- Classification schemes
- Symbology
- Geographic units

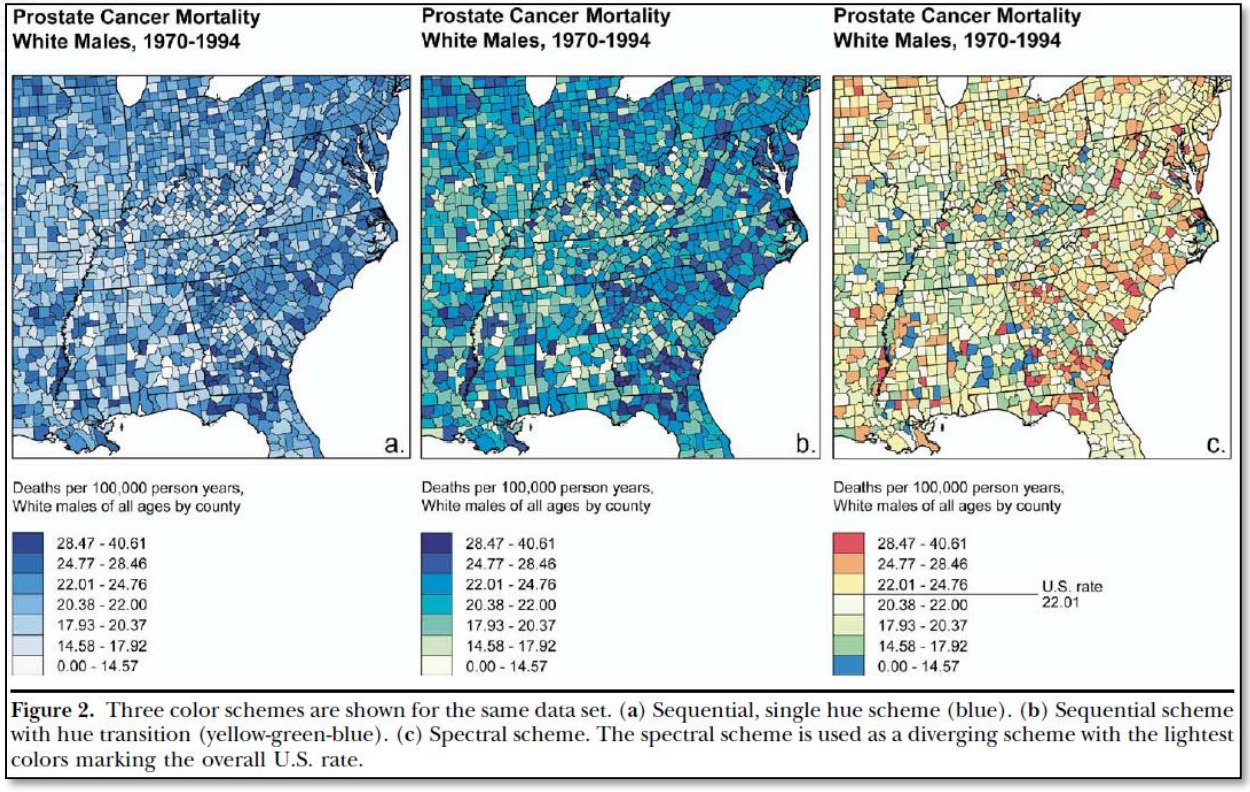
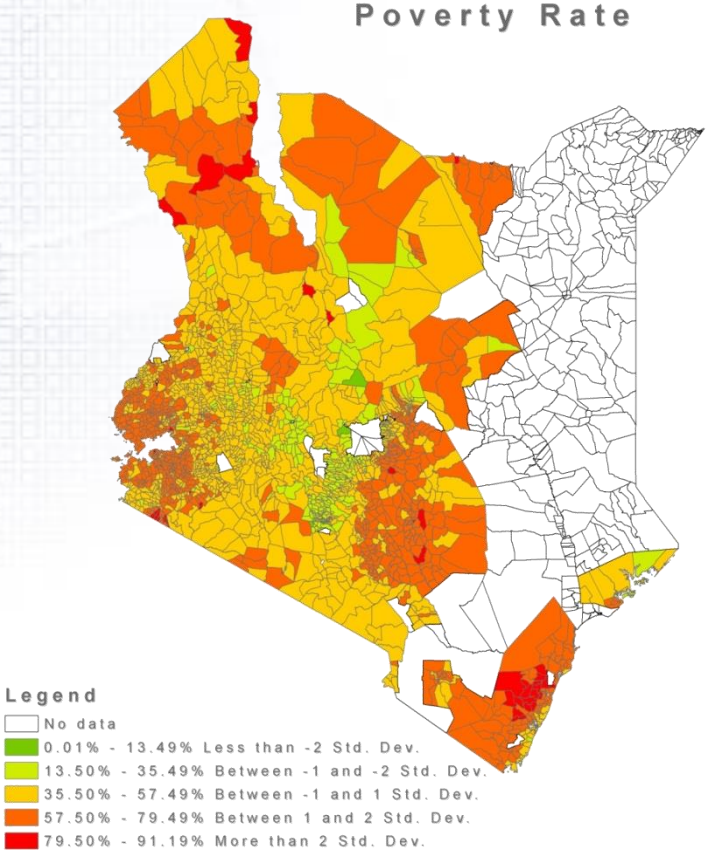



**Figure 2.** Three color schemes are shown for the same data set. (**a**) Sequential, single hue scheme (blue). (**b**) Sequential scheme with hue transition (yellow-green-blue). (**c**) Spectral scheme. The spectral scheme is used as a diverging scheme with the lightest colors marking the overall U.S. rate.

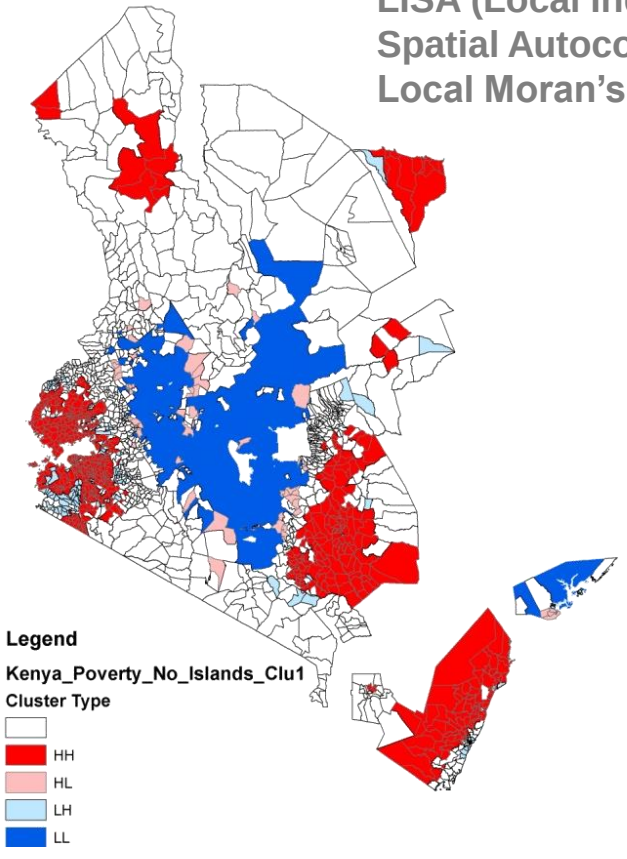
From:
**Basic mapping principles for visualizing cancer data using Geographic Information Systems (GIS).**
*Brewer CA.*
*Am J Prev Med. 2006 Feb;30(2 Suppl):S25-36.*

# Analysis Within A Layer

Comparing the spatial arrangement of features within a single GIS data layer




LISA (Local Indicators of Spatial Autocorrelation / Local Moran's I) Clusters



# Analysis Between Layers

- Comparing the spatial location of features in multiple data layers