



Introduction to GIS

CSDE Workshop for EPI Scholars
June, 2017

Matt Dunbar
CSDE Assistant Director &
GIS Program Manager



Center for Studies in
Demography and Ecology



Դեմոգրաֆիկա և Եկոլոգիա



Lecture Overview

- Defining GIS
- What's so special about spatial?
- Public Health Research Examples
- Core GIS Concepts: storing, mapping, analyzing spatial data
- Detailed Concepts (future modules)



What is GIS?

One word at a time...



G Information S

Data is a fact or collection of facts

- A single interview
- The entire US Census

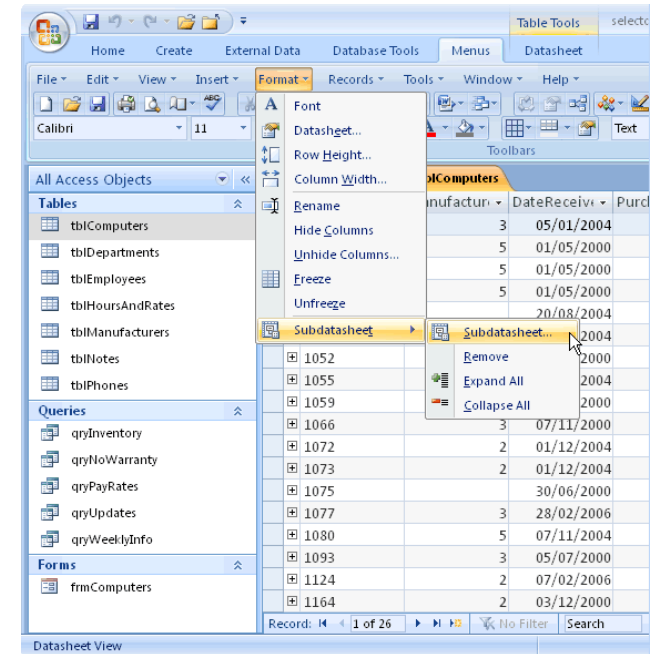
Data that is processed, organized, structured or presented in a given context to make them useful, are called Information



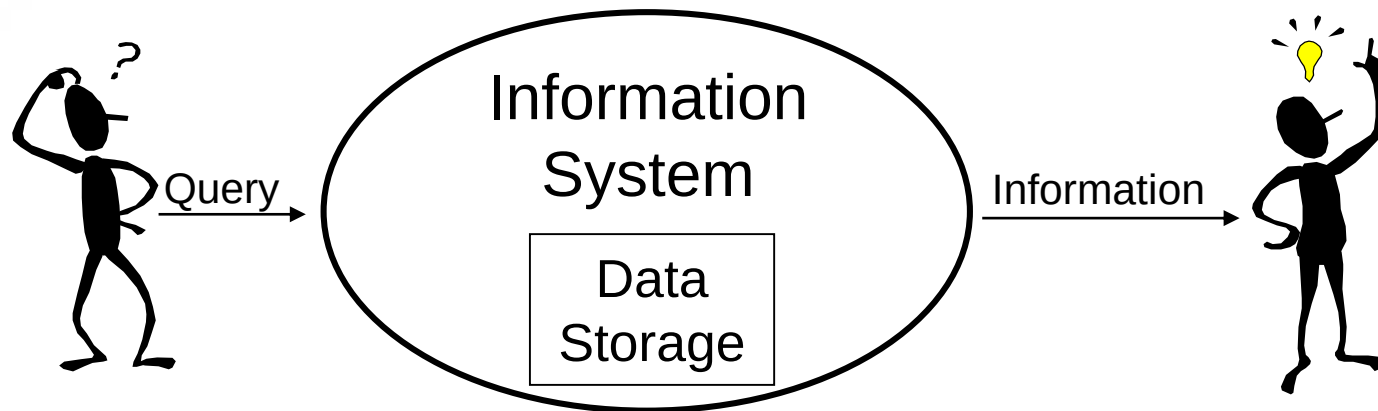
G Information System

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



Geographic Information System

Simple way to think about GIS

“A tool for deriving information from any data with a geographic/spatial component”

More technical definition

“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.”



What is GIS?

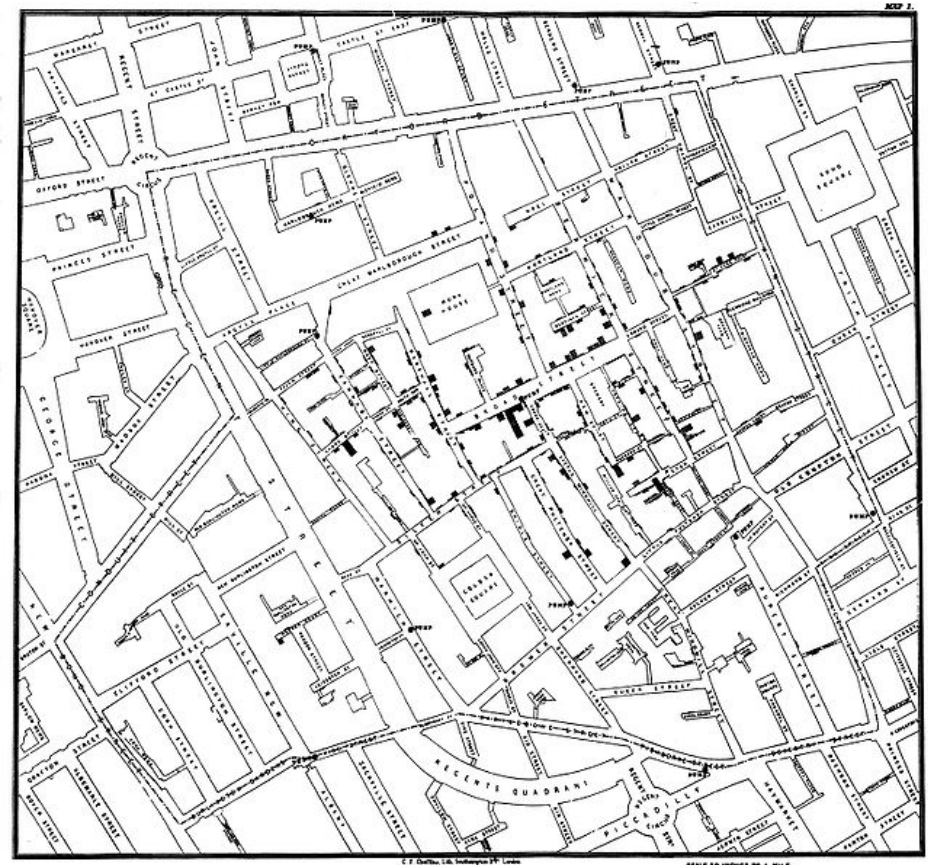
Spatial is special...



History: Dr. John Snow – Mapping Cholera and Health

1854 Cholera Outbreak in London

- First, identified WHERE people lived who were infected
- The clustered PATTERN suggested a particular water pump was the source
- Made MAPS to show the location of cholera cases in relation to water sources



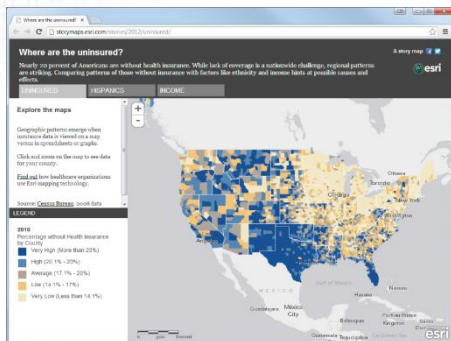


Spatial is Special

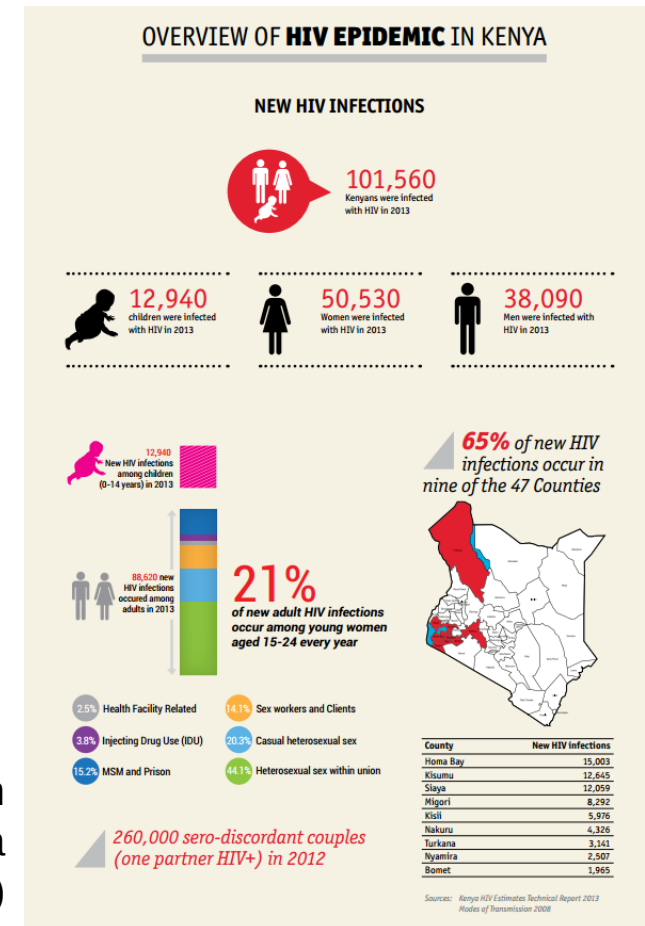
“Everything is related to everything else,
but near things are more related than
distant things”

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!).



Kenya HIV Prevention
Revolution Road Map (Kenya
MOH)



<http://storymaps.esri.com/stories/2012/uninsured/>



GIS and Health Research

GIS and spatial analysis methods are becoming a part of the health researcher's methodological arsenal; witness the emergence of

Textbooks

GIS and Public Health, 2nd Ed.
(Cromley and McLafferty 2011)

Medical Geography, 3rd Ed.
(Meade and Emch, 2010)

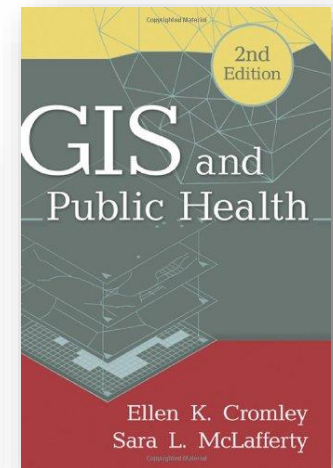
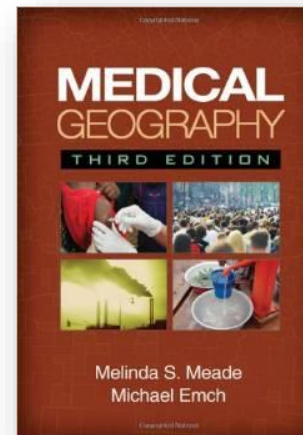
Journals

International Journal of Health Geographics
<http://www.ij-healthgeographics.com/>.

On-line resources and Geospatial data warehouses

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

The Health Resources and Service Administration (HRSA) hosts a geospatial data warehouse at <http://datawarehouse.hrsa.gov>



INTERNATIONAL JOURNAL
OF HEALTH GEOGRAPHICS

IMPACT
FACTOR
2.45



GIS/Spatial Analysis Myths

#1 A common views of GIS is that it is a technology that only does one thing: draw “**pretty maps.**”

#2 A common views of GIS is that it is an **expensive technology and the learning curve is steep.**

#3 A common views of GIS is that it should **only play a limited role in a project** (e.g. just mapping, or just data integration).



What is GIS?

Public Health Research
Examples...



Public Health Systems and Services Research

Local Health Departments: Exploring changes in funding, level of service, need, and health outcomes from 2000-2010.

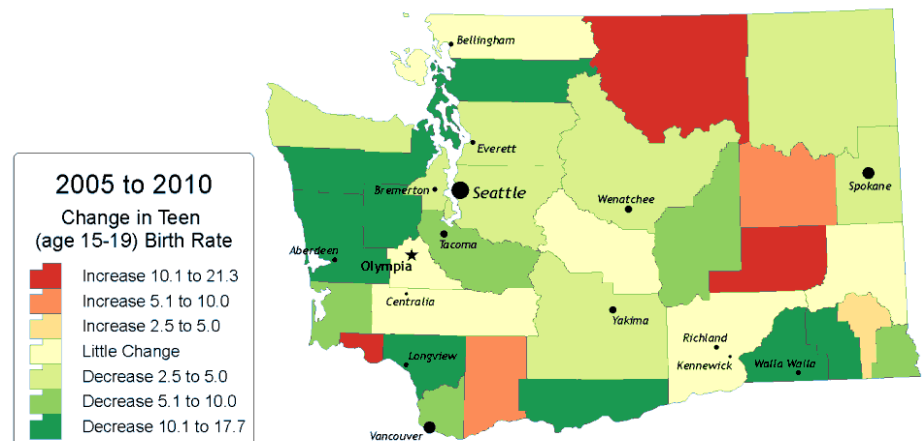
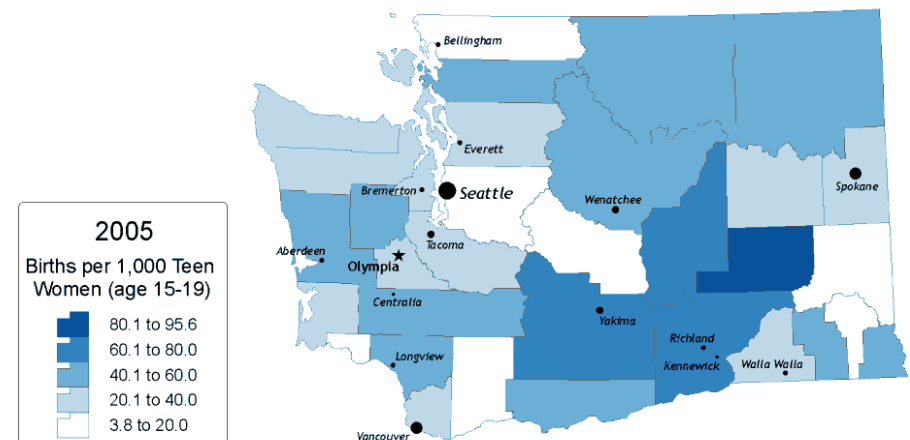
Local health departments and specific maternal and child health expenditures: relationships between spending and need.

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).

2005 to 2010 Change in Washington's
Teen Birth Rate (LHJ aggregation)



State and National Teen Birth Rates

WA	2005: 31.4	2010: 27.3	[decrease 4.1]
USA	2005: 39.7	2010: 34.3	[decrease 5.4]

Data Sources:
Washington Department of Health
(<http://www.doh.wa.gov/DataandStatisticalReports>)
National Center for Health Statistics
(<http://www.cdc.gov/nchs/data/databriefs/db69.htm>)

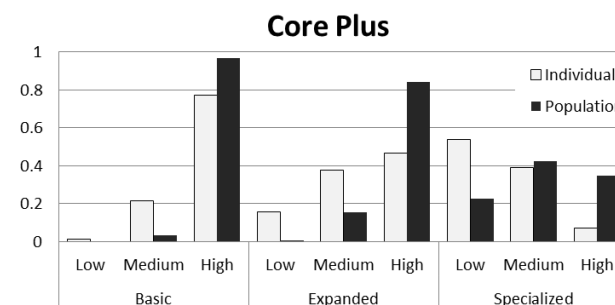
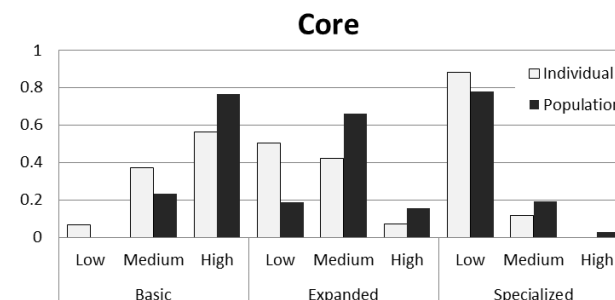
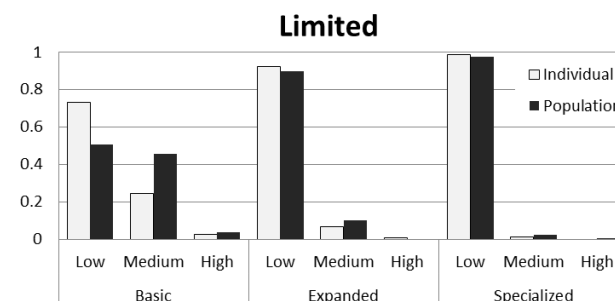
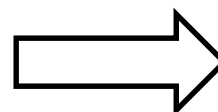


Understanding Local Health Department Service Philosophy:

A Latent Class Analysis of NACCHO Profile Data

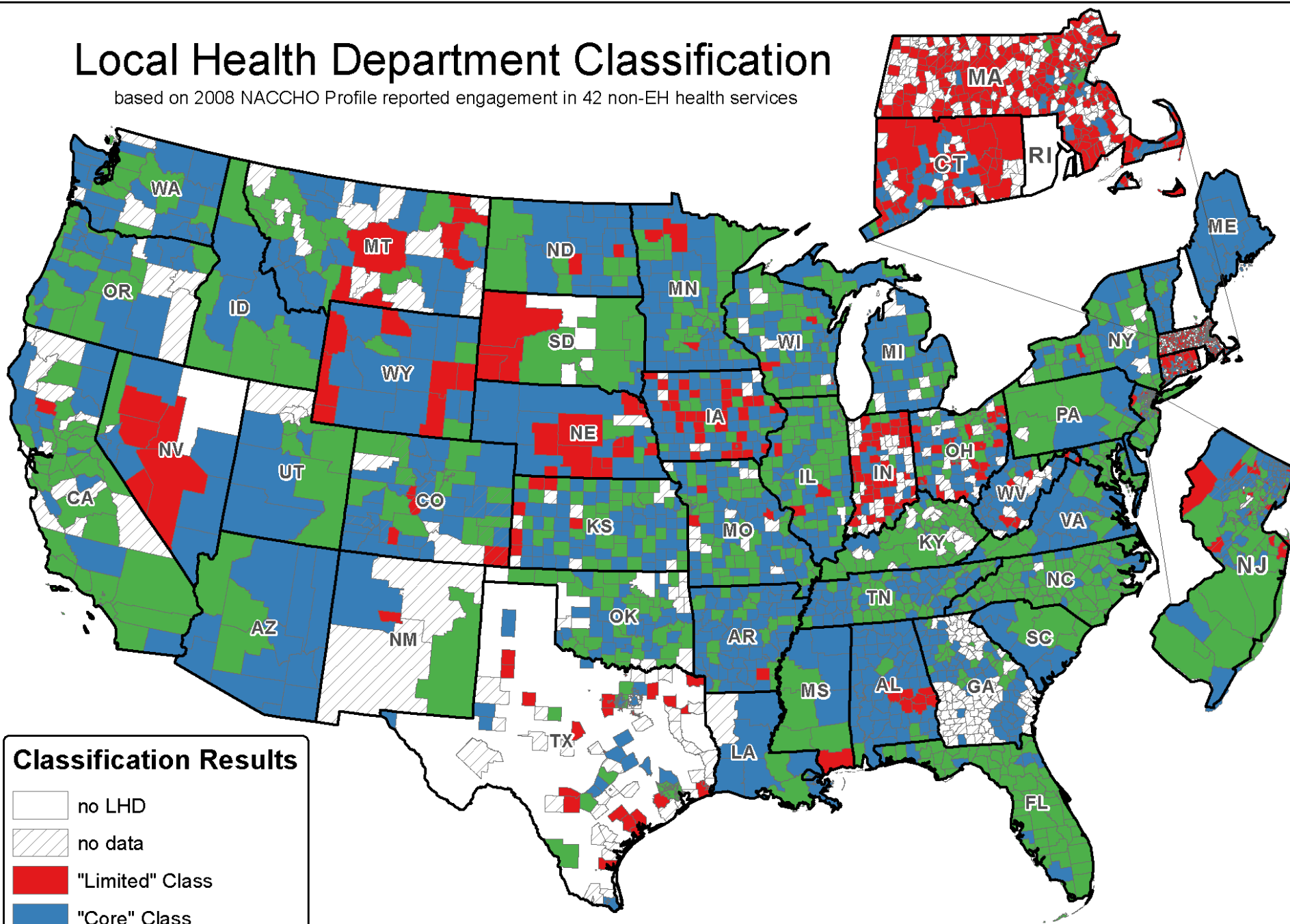
Betty Bekemeier, Athena Pantazis, Matthew Dunbar, Jerry Herting (in review)

	Basic	Expanded	Specialized
Individual Focused	Adult Immunizations (Imm) Child Immunizations (Imm) EPSDT (MCH) Family planning (MCH) MCH home visits (MCH) WIC (MCH)	Cancer (Scrng) Cardiovascular disease (Scrng) Diabetes (Scrng) High blood pressure (Scrng) Home health care (Other HS) Oral health (Other HS) Prenatal care (MCH) Primary care (Other HS) School health (Other) Well Child Clinic (MCH)	Behavioral/mental (Other HS) HIV/AIDS (Tx) Obstetrical care (MCH) School-based clinics (Other) Substance abuse (Other HS)
Population Focused	Blood Lead (Scrng) Communicable/ infectious disease (Epi) HIV/AIDS (Scrng) Nutrition (Pop) Other STDs (Scrng) Tuberculosis (Scrng) Tuberculosis (Tx) Tobacco (Pop) Unintended pregnancy (Pop)	Behavioral risk factors (Epi) Chronic disease programs (Pop) Maternal and child health (Epi) Physical activity (Pop) STDs (Tx)	Chronic disease (Epi) Injury (Epi) Injury (Pop) Mental illness (Pop) Substance abuse (Pop) Syndromic surveillance (Epi) Violence (Pop)



Local Health Department Classification

based on 2008 NACCHO Profile reported engagement in 42 non-EH health services



Classification Results

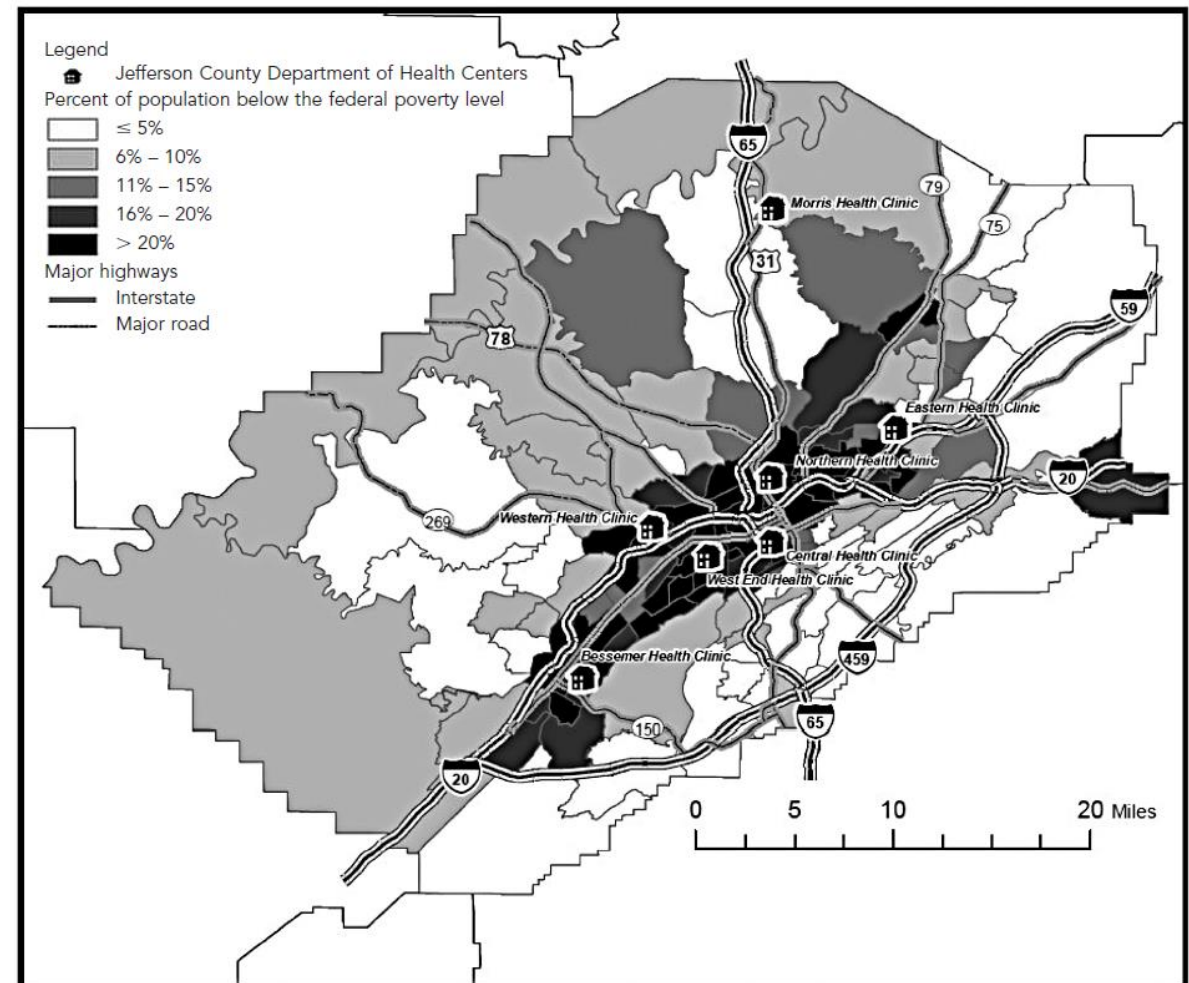
-  no LHD
-  no data
-  "Limited" Class
-  "Core" Class
-  "Core Plus" Class



Locating Health Facilities

Use of GIS for site placement of county-level public health primary care centers

Figure 1. Jefferson County Department of Health primary care center locations and percent of population below the federal poverty level: Jefferson County, Alabama, 2009^a



From:

Using GIS for administrative decision-making in a local public health setting.

Taylor DM, Yeager VA, Ouimet C, Menachemi N.
Public Health Rep. 2012 May-Jun;127(3):347-53



Locating Health Facilities

Figure 2. Primary care center locations and patient residences with six-mile catchment: Jefferson County, Alabama, 2009

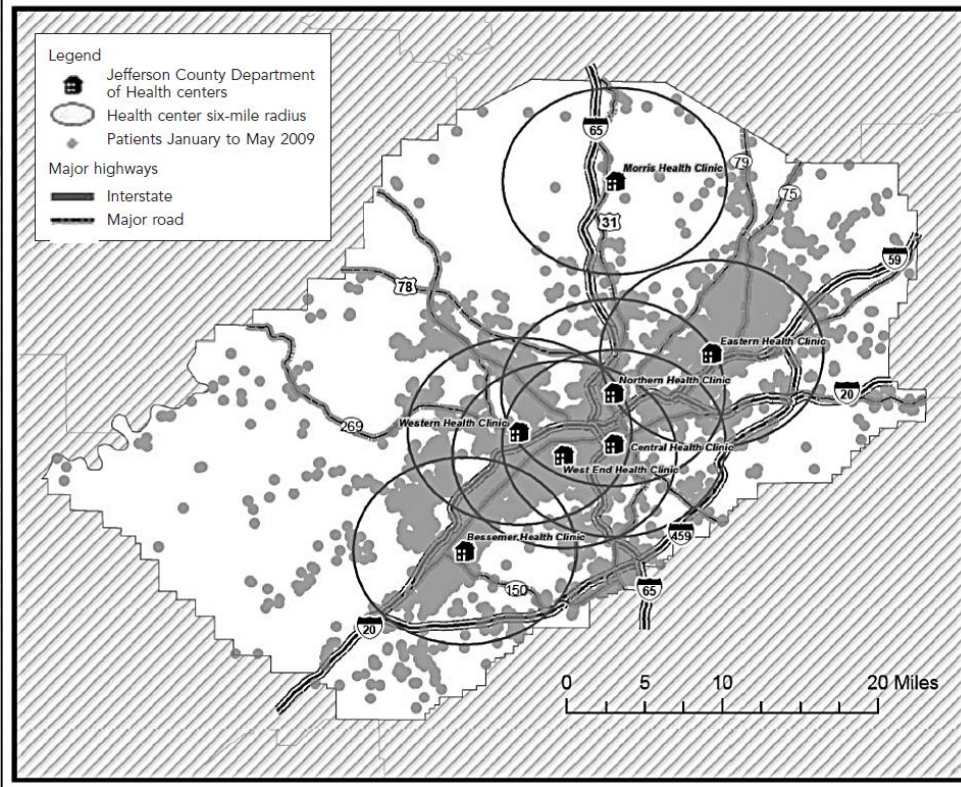
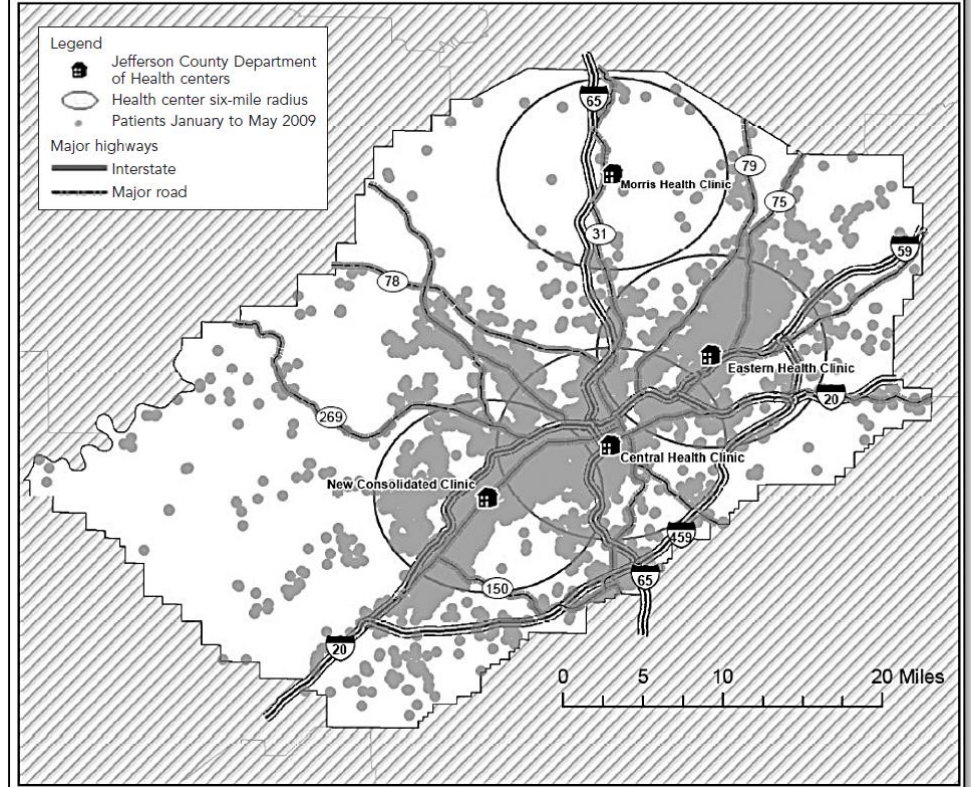


Figure 3. Revised primary care center locations and patient residences with six-mile catchment: Jefferson County, Alabama, 2009





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



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.

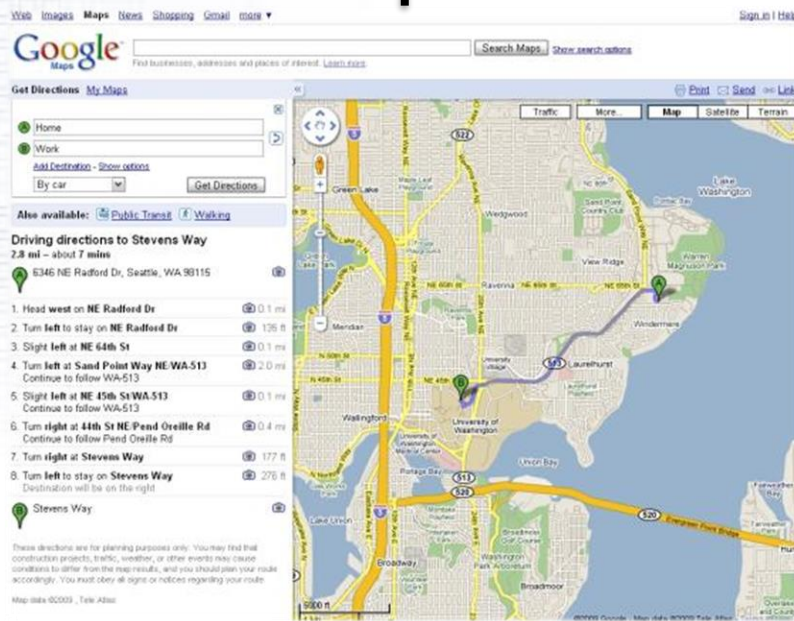
In an era of limited state and local budgets and legislators' reluctance to support higher fuel taxes or general tax increases, tolls on urban highways and bridges may be an attractive source of

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.

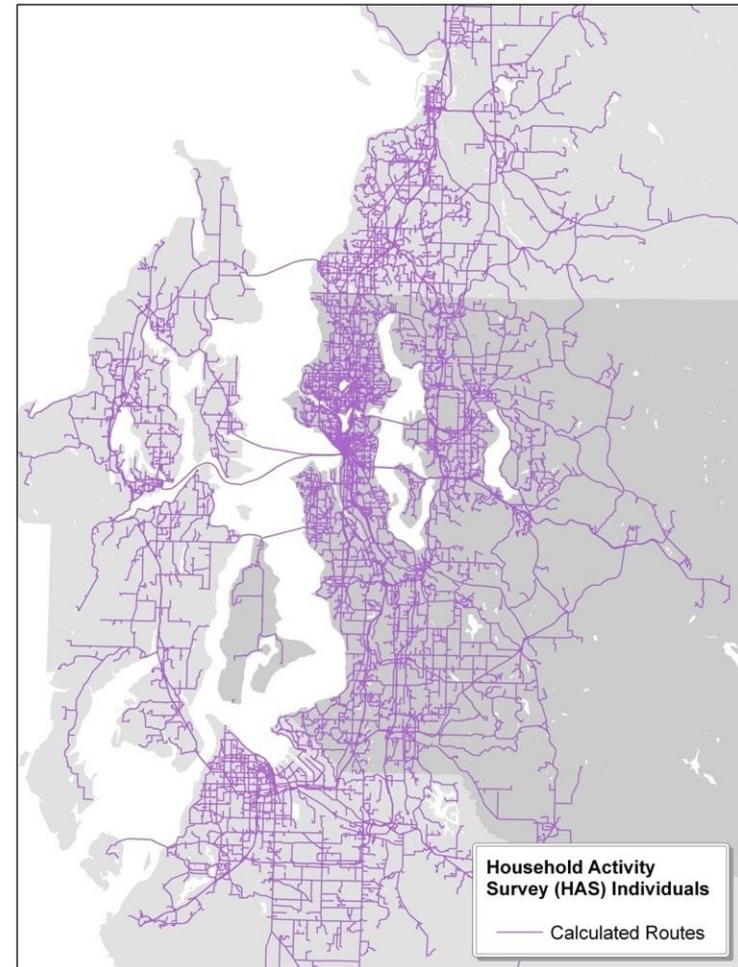
JOURNAL OF URBAN AFFAIRS, Volume 00, Number 0, pages 1–22.
Copyright © 2011 Urban Affairs Association
All rights of reproduction in any form reserved.
ISSN: 0735-2166.

DOI: 10.1111/j.1467-9906.2011.00551.x

Mapping Commuting Routes

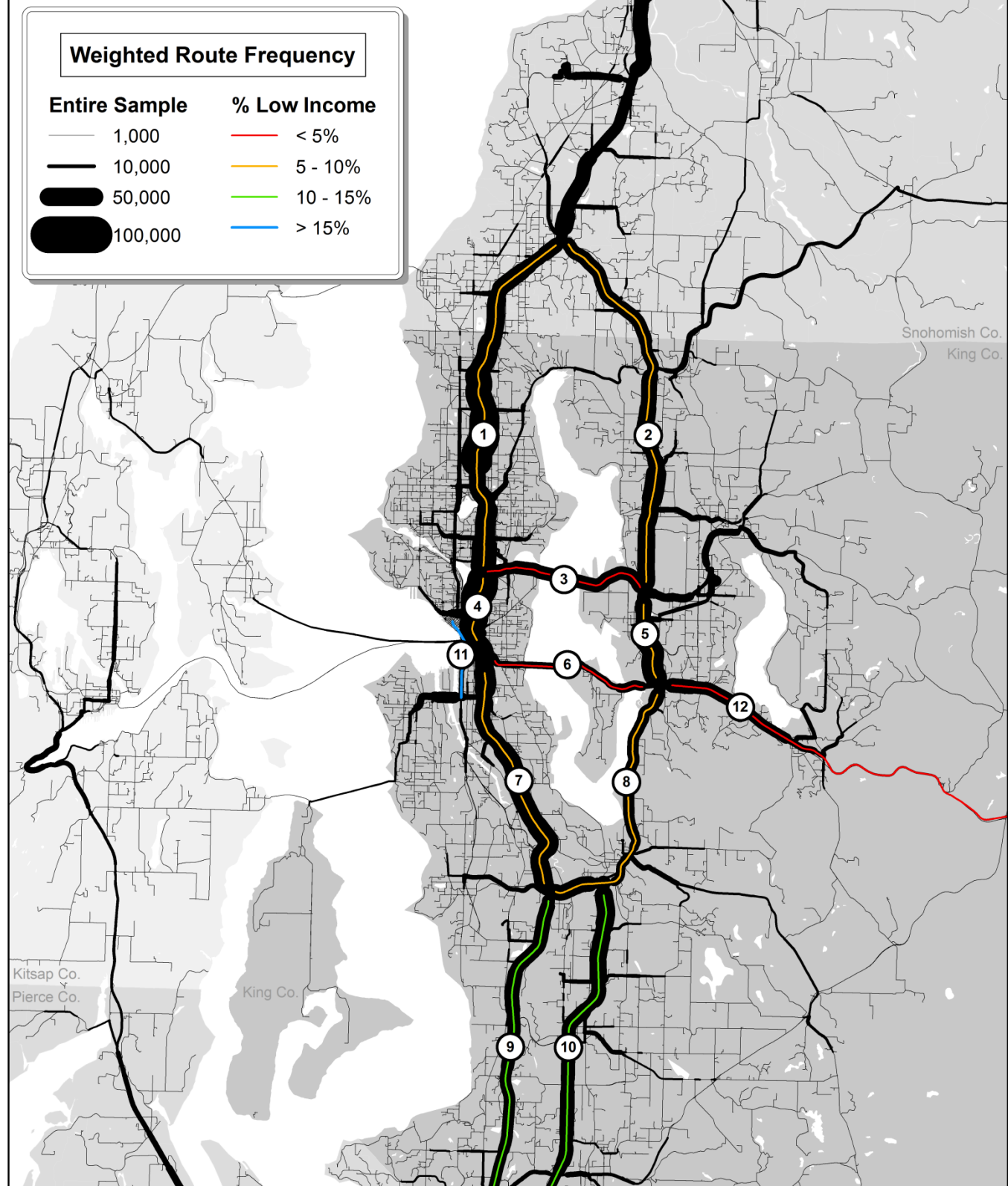


Home to Work Route
Single Individual (Google)



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

Commuting Routes





What is GIS?

Basics of Storing, Mapping, and
Analyzing Spatial Data...

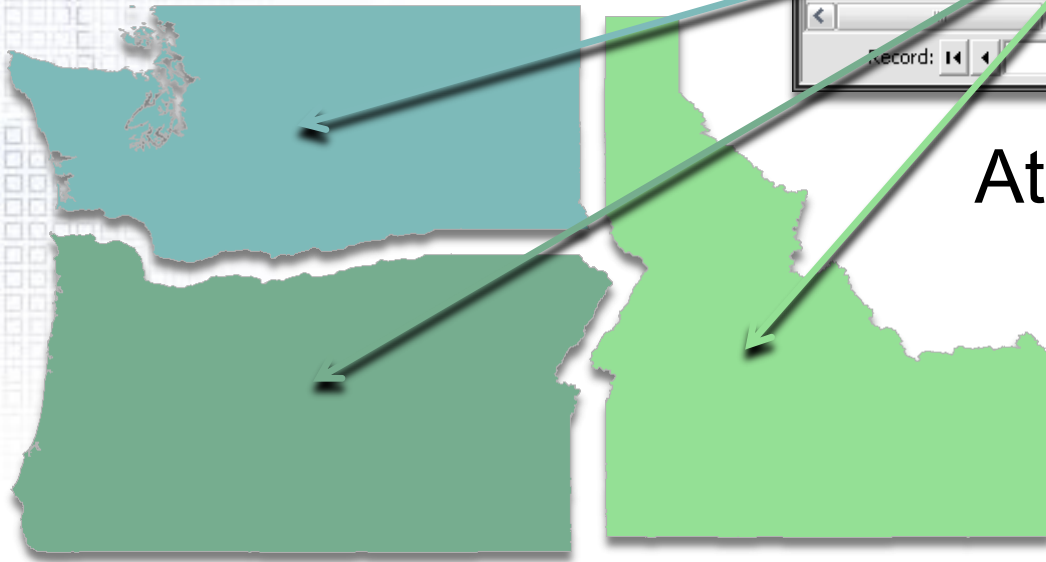


Storing Geographic Data

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

ObjectID *	Shape *	STATE_NAME	STATE_FIPS	SUB_REGION	STATE_ABBR	POP2000	POP2007	POP2008
48	Polygon	Washington	53	Pacific	WA	5894121	6516384	6880000
37	Polygon	Oregon	41	Pacific	OR	3421399	3752734	3831000
13	Polygon	Idaho	16	Mountain	ID	1293953	1513708	1567000

Record: 0 Show: All Selected Records (3 out of 52 Selected) Options



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

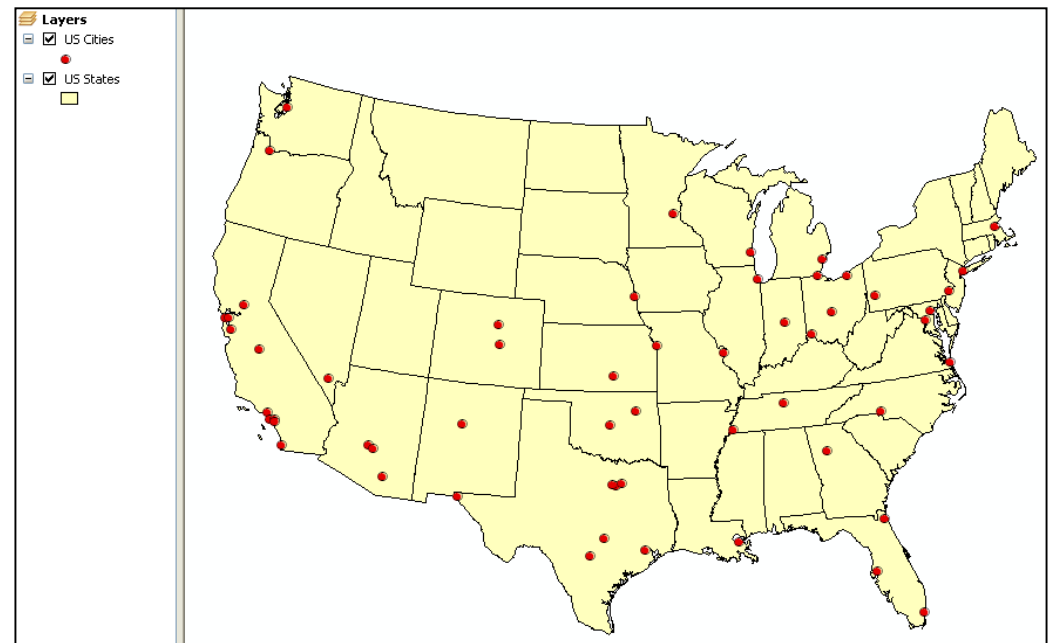
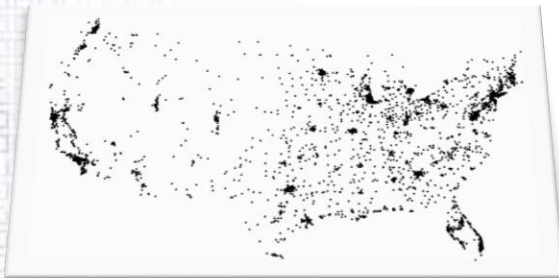
Areas (Polygons)

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)

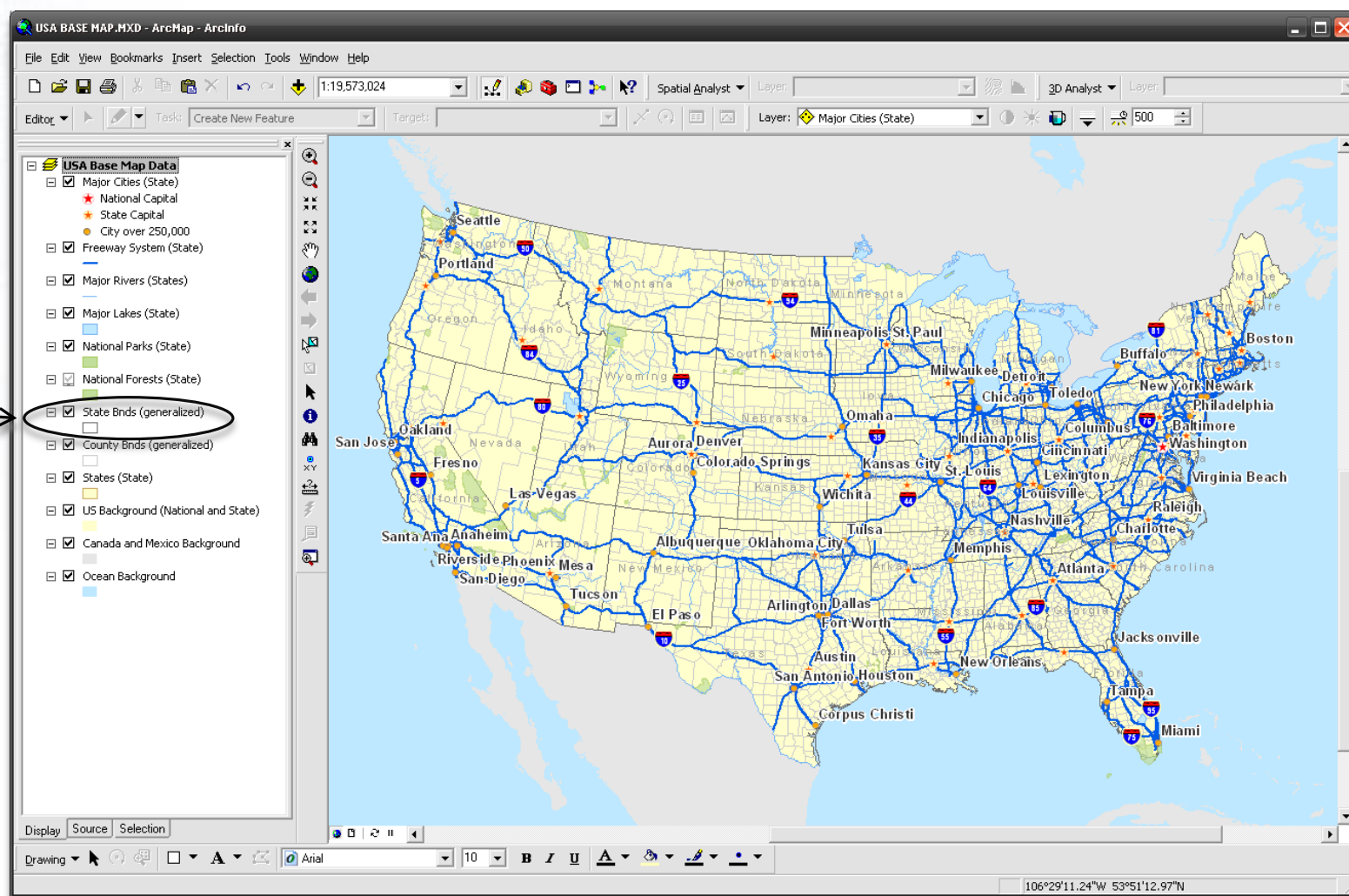




Mapping Geographic Data

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?

Select By Attributes

Layer: State Bnds (generalized)

☐ Only show selectable layers in this list

Method: Create a new selection

"SUB_REGION"
"STATE_ABBR"
"POP2000"
"POP2007"
"POP00_SQMI"

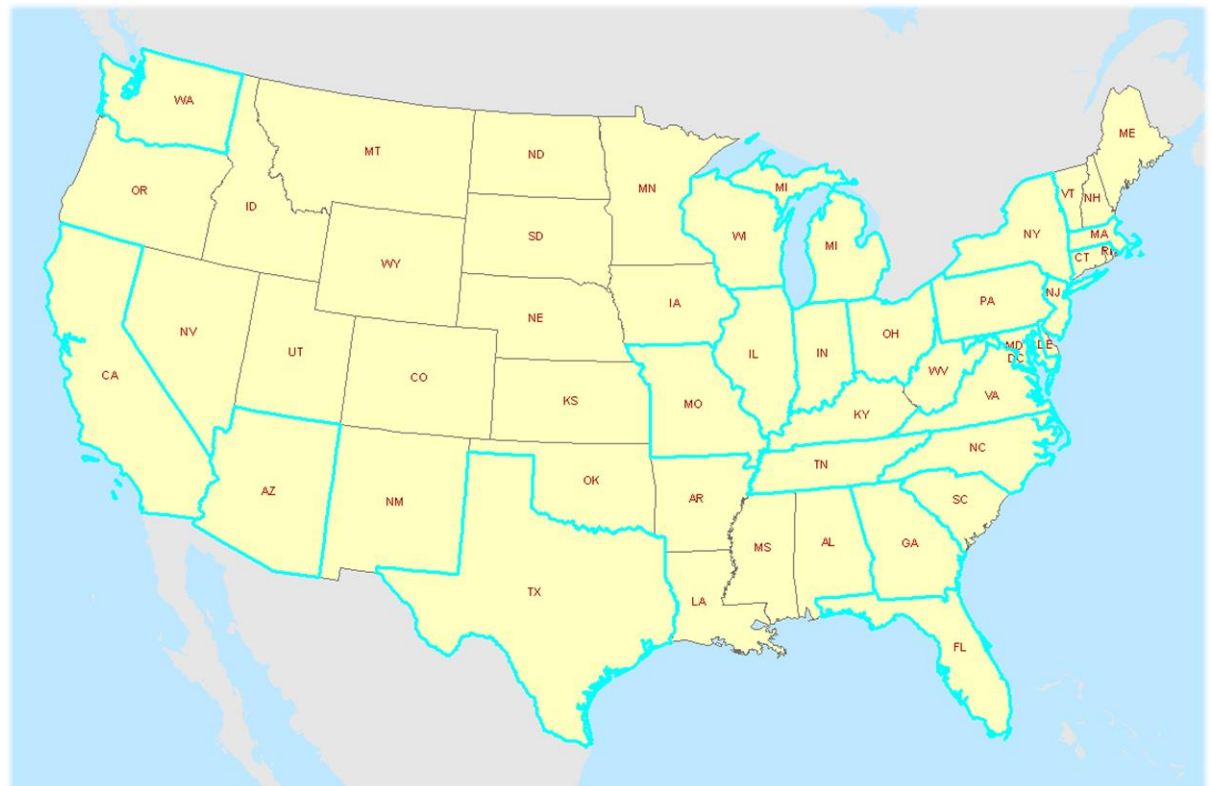
= < > Like
> >= And
< <= Or
_ % () Not

Is Get Unique Values Go To:

SELECT * FROM states WHERE:
"POP2000" >= 5000000

Clear Verify Help Load... Save...

OK Apply Close





Analyzing Geographic Data

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

2) Which states touch the border of Kansas?

Select By Location [?] [X]

Select features from one or more target layers based on their location in relation to the features in the source layer.

Selection method:
select features from [v]

Target layer(s):

- ☒ US States
- ☐ Canada and Mexico Background

☐ Only show selectable layers in this list

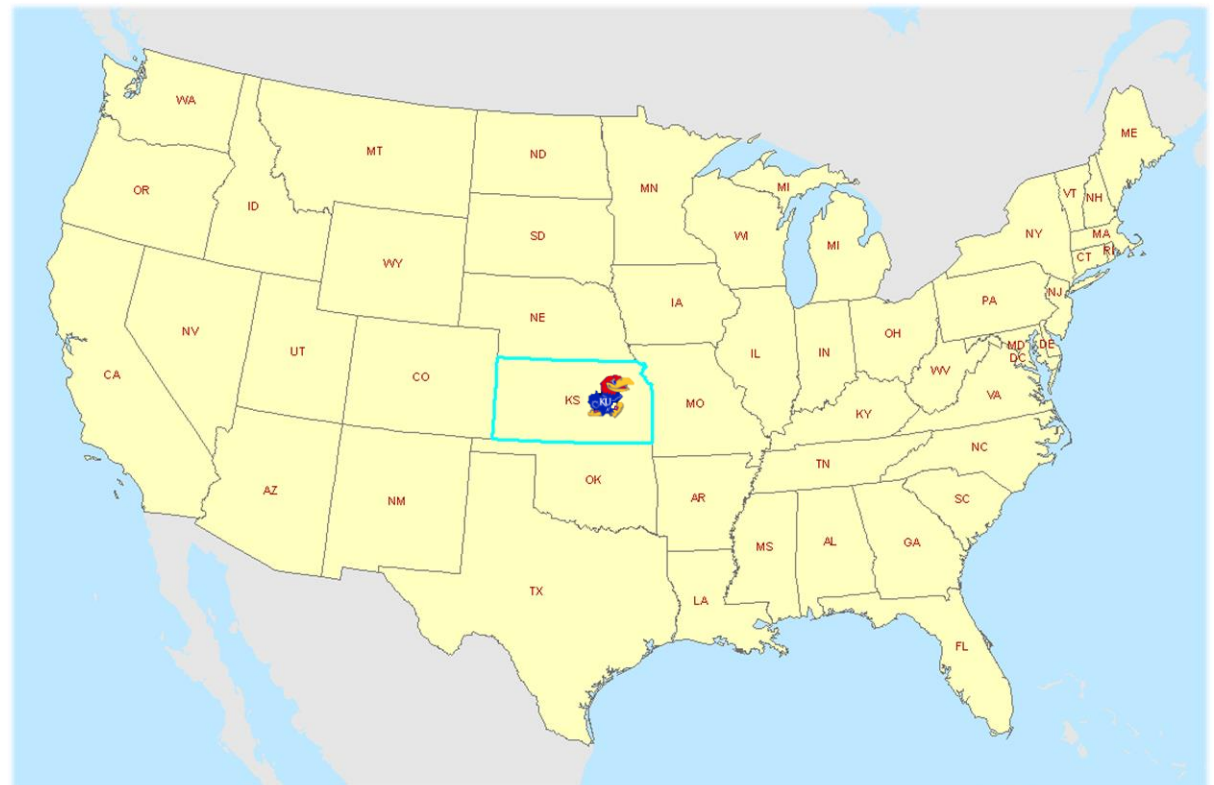
Source layer:
[US States] (1 features selected)

Use selected features

Spatial selection method:
[Target layer(s) features touch the boundary of the Source layer feature]

☐ Apply a search distance
[3.000000] [Decimal Degrees]

[Help] [OK] [Apply] [Close]





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

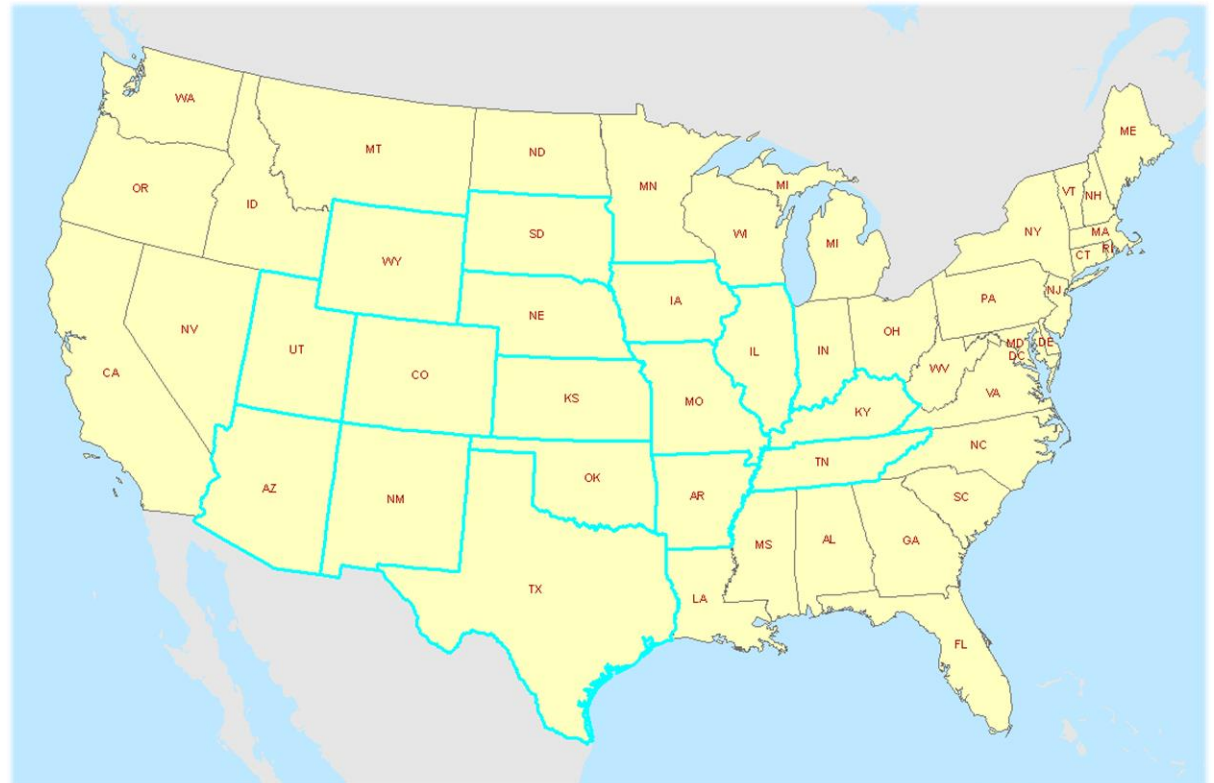
2) Which states touch the border of Kansas?

A map of the United States showing state boundaries and two-letter state abbreviations. A thick red line outlines a path starting from the Pacific Northwest, moving south through the West, then east through the central US, and finally south through the Southeast.



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

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



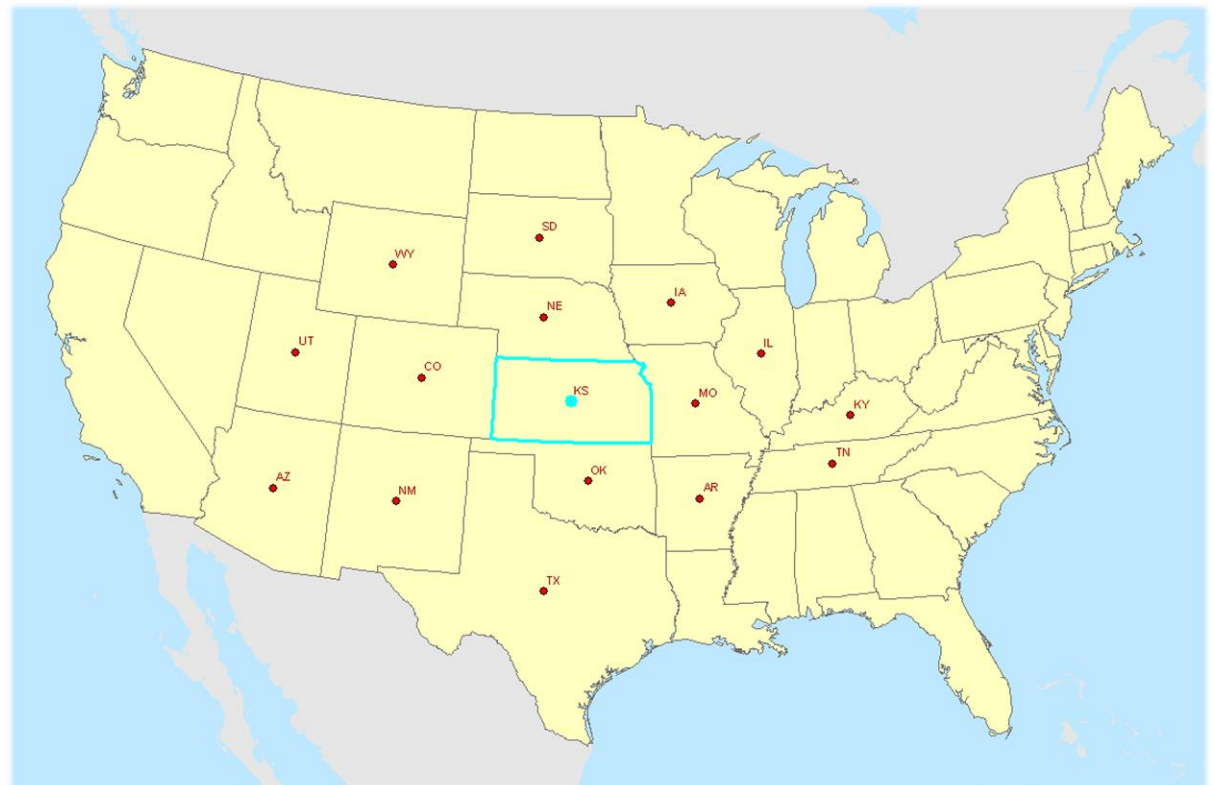


Analyzing Geographic Data

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

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

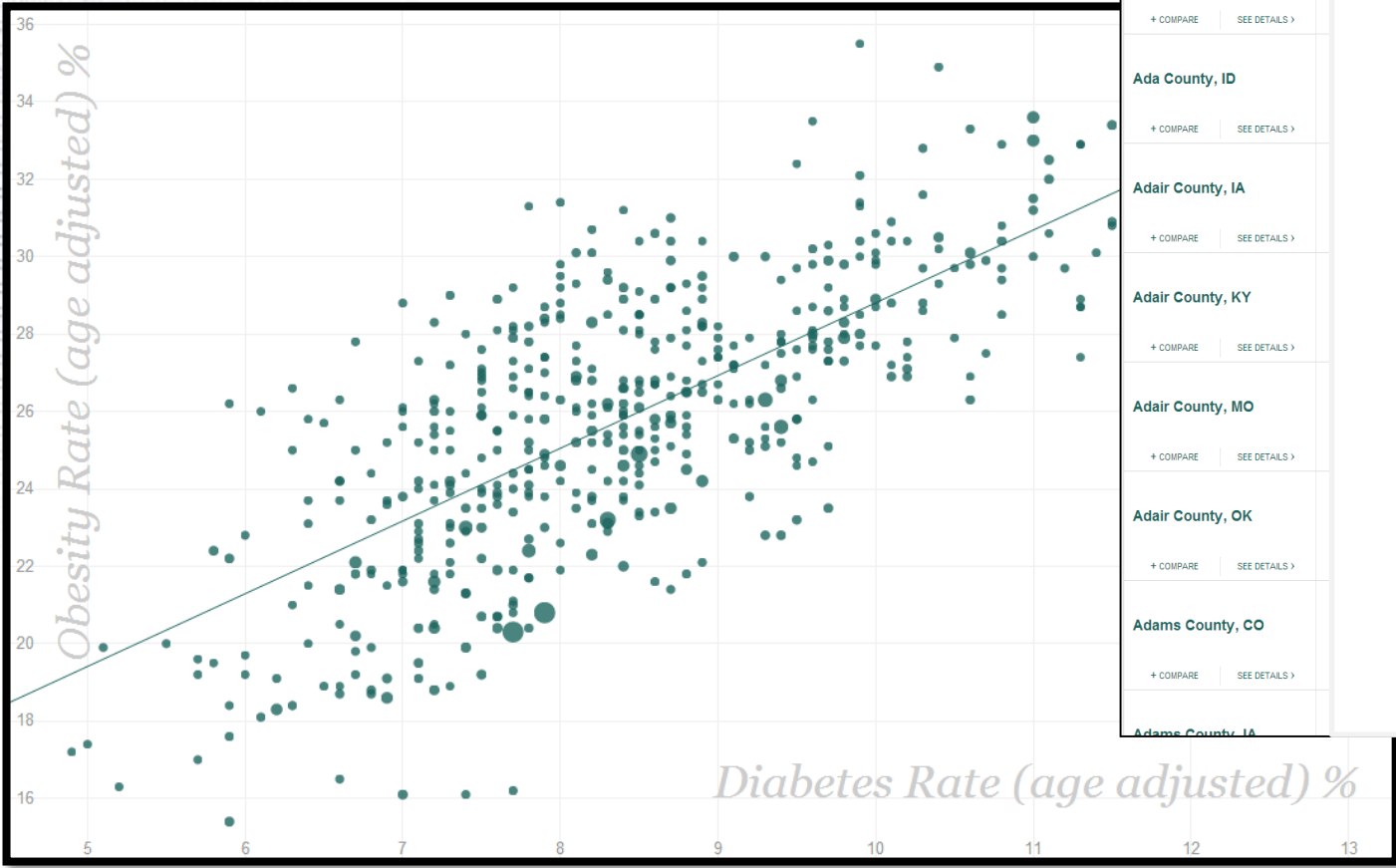
Table		
Export_Output		
STATE_NAME	DISTANCE	DISTANCE_MILES
Arizona	1272924.156589	790.95837
Arkansas	661910.092654	411.29187
Colorado	621990.494685	386.48697
Illinois	806745.278383	501.28827
Iowa	574132.587986	356.74945
Kentucky	1147940.661541	713.29724
Missouri	512794.420804	318.63568
Nebraska	358616.704365	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



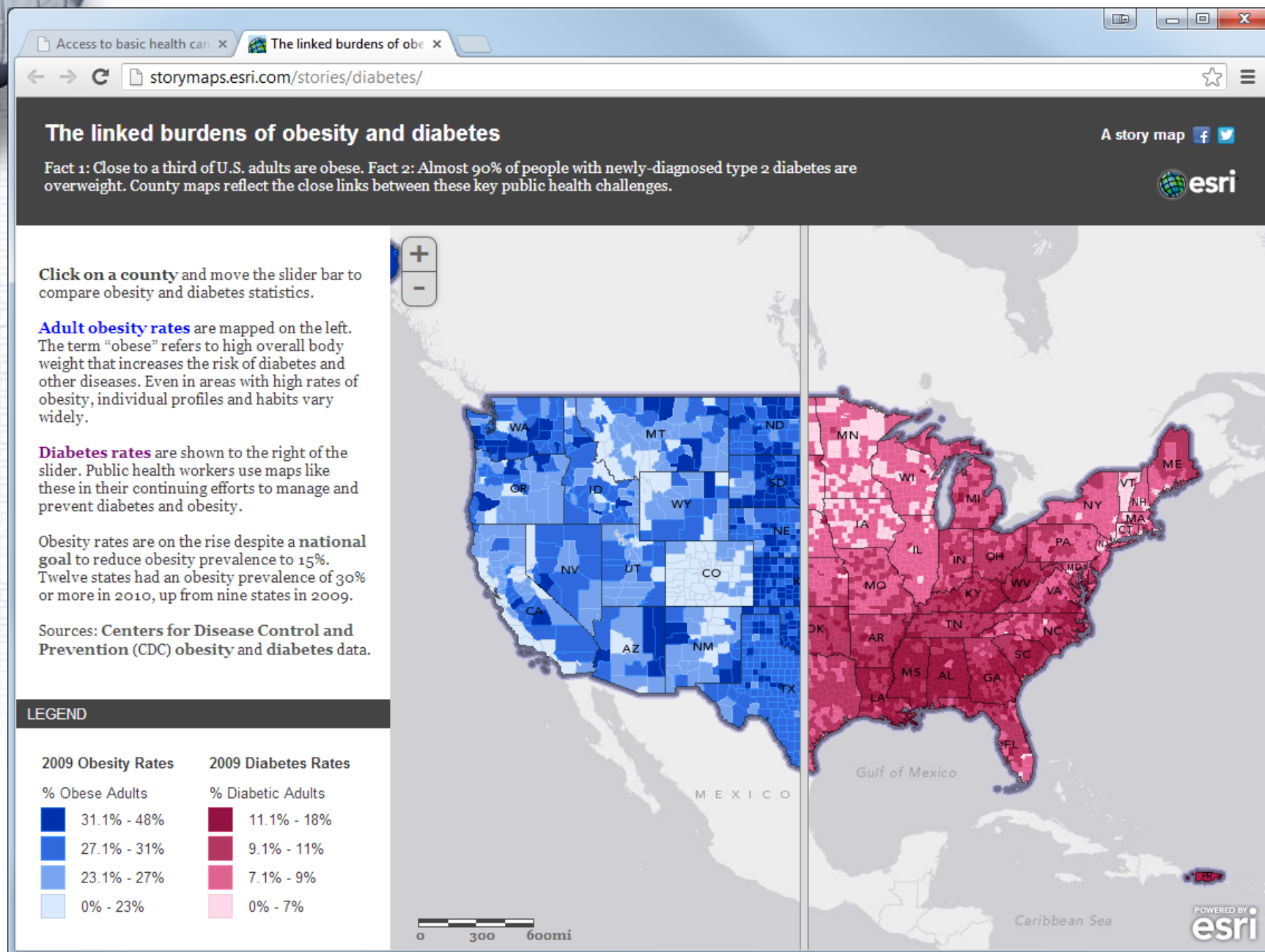


Storing, Mapping, Analyzing Example: County Obesity and Diabetes Rates

Traditional Vis
Scatterplot of y vs x



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
Adams County, IA	...	...	...	...	...	...	...

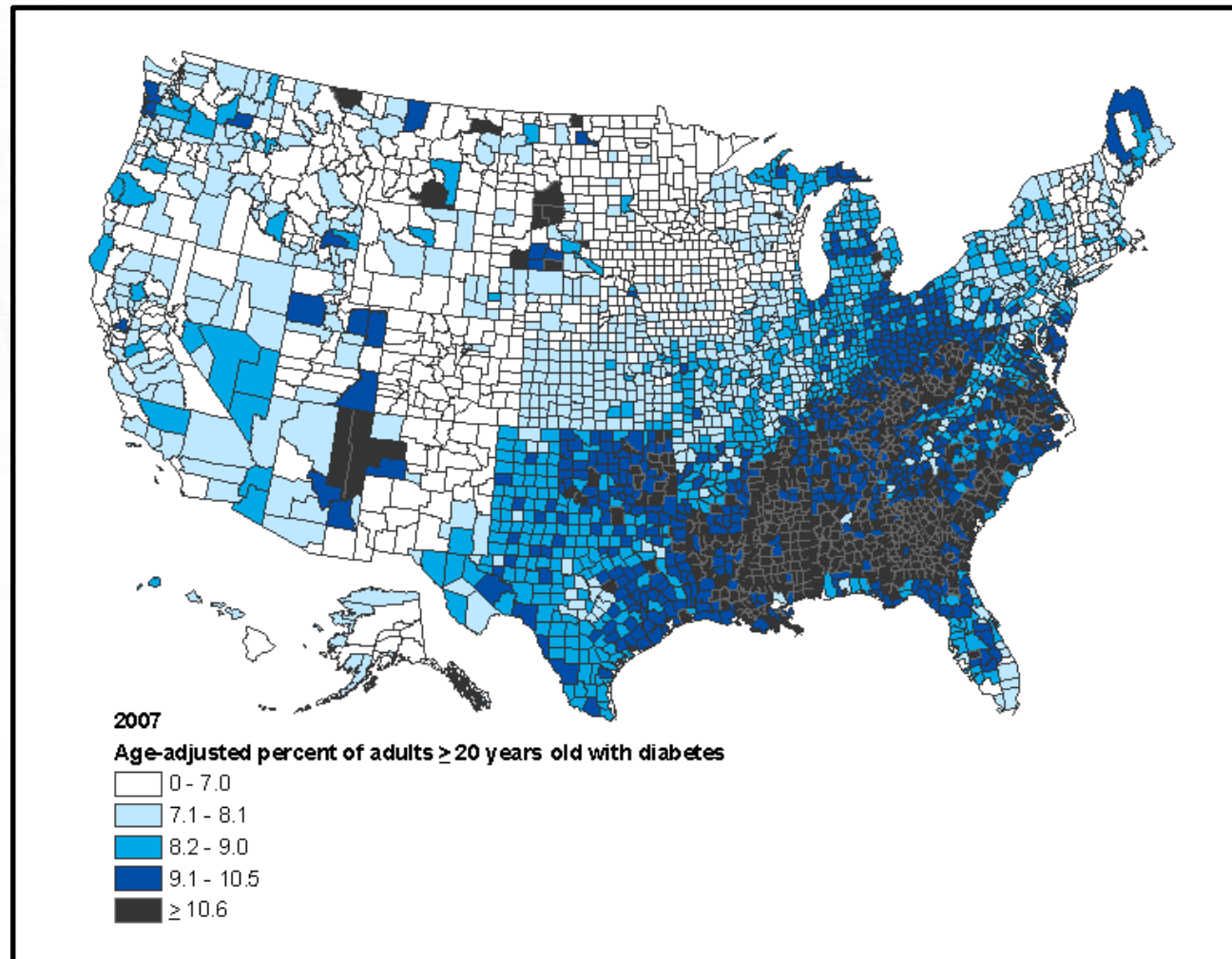


<http://storymaps.esri.com/stories/diabetes/>



Instruments for Reasoning about Quantitative Information

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



MMWR 2009;58:1259-1263



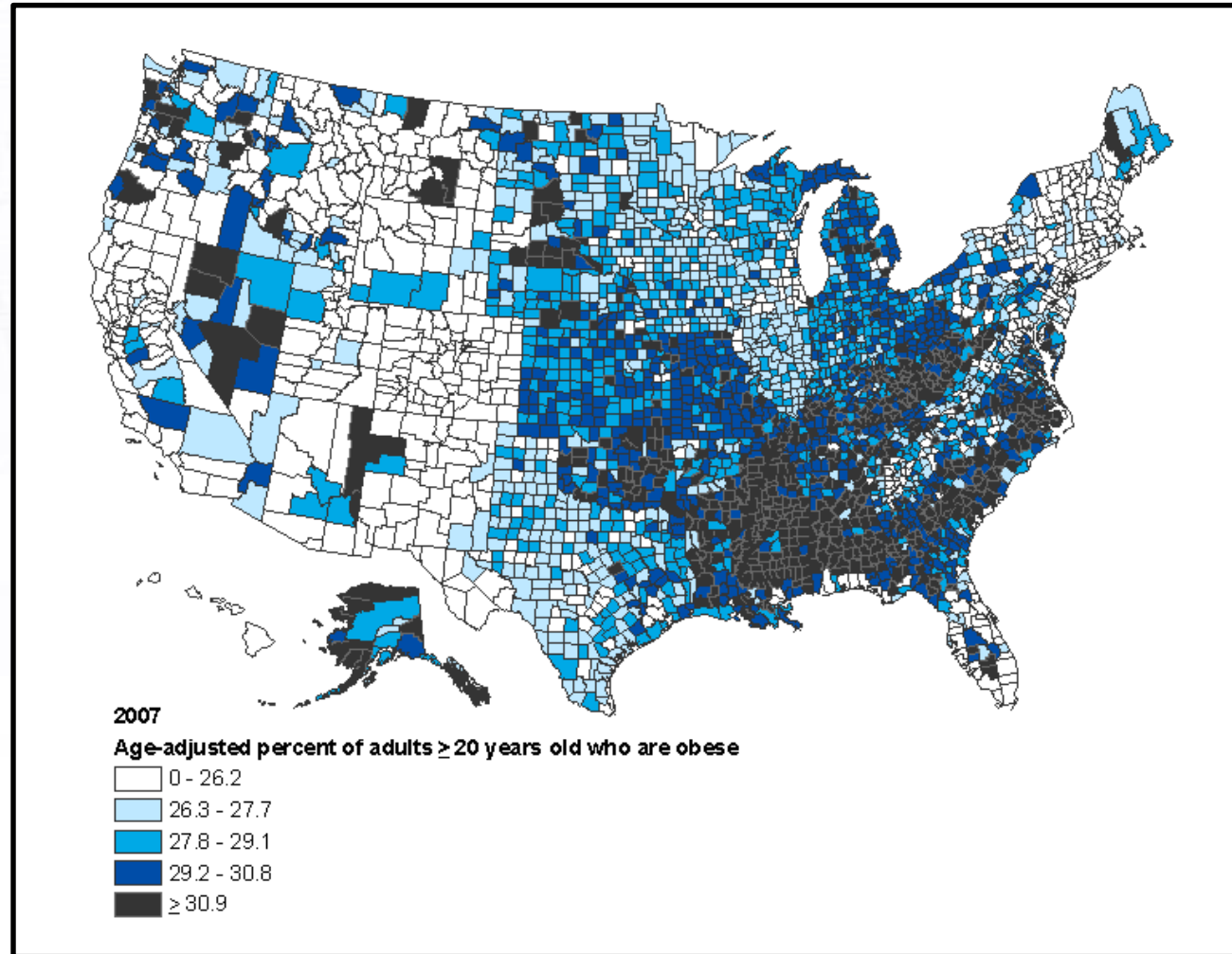
CDC's Division of Diabetes Translation. National Diabetes Surveillance System available at <http://www.cdc.gov/diabetes/statistics>





Instruments for Reasoning about Quantitative Information

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



MMWR 2009;58:1259-1263



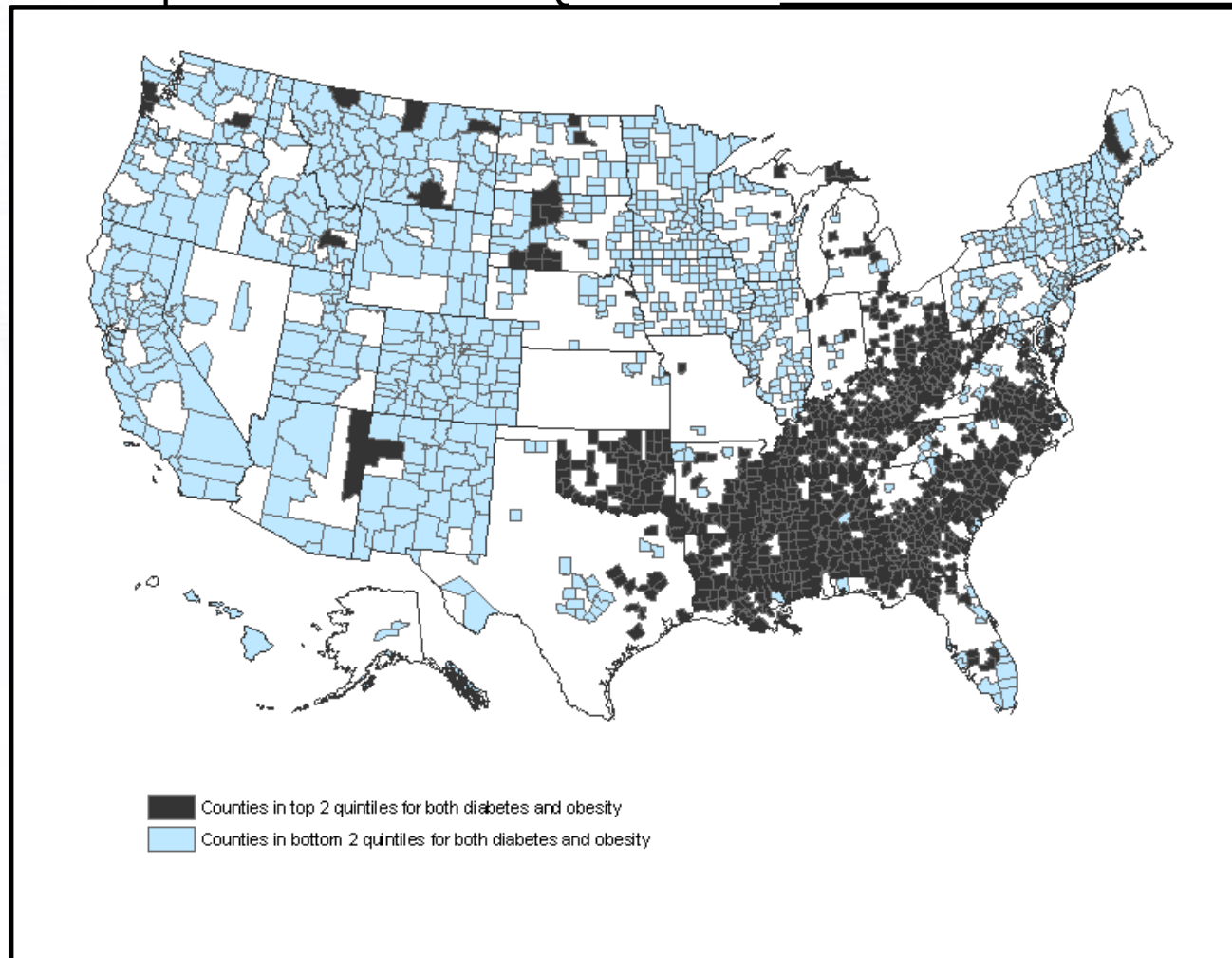
CDC's Division of Diabetes Translation. National Diabetes Surveillance System available at <http://www.cdc.gov/diabetes/statistics>





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>



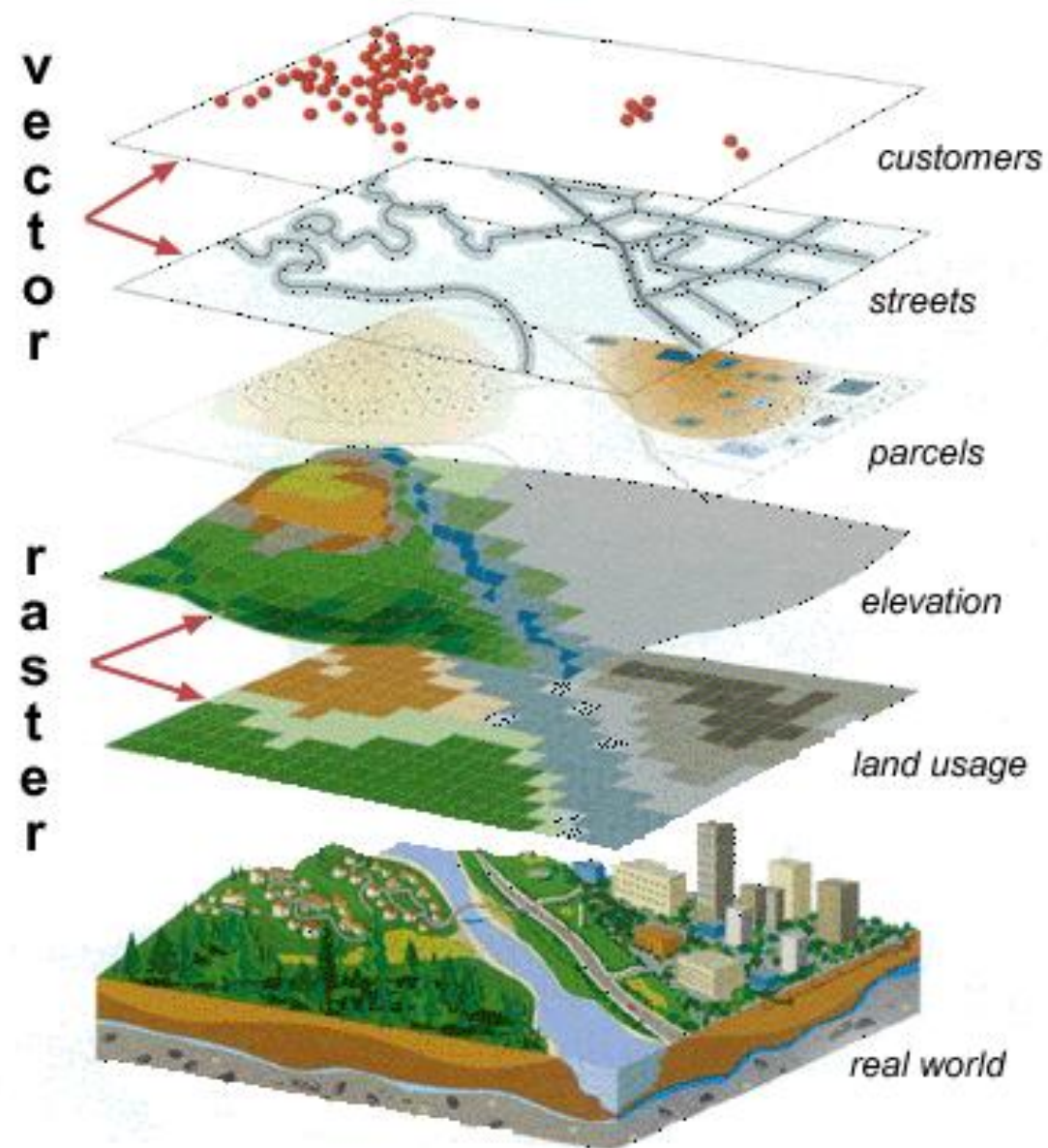


What is GIS?

In more detail...



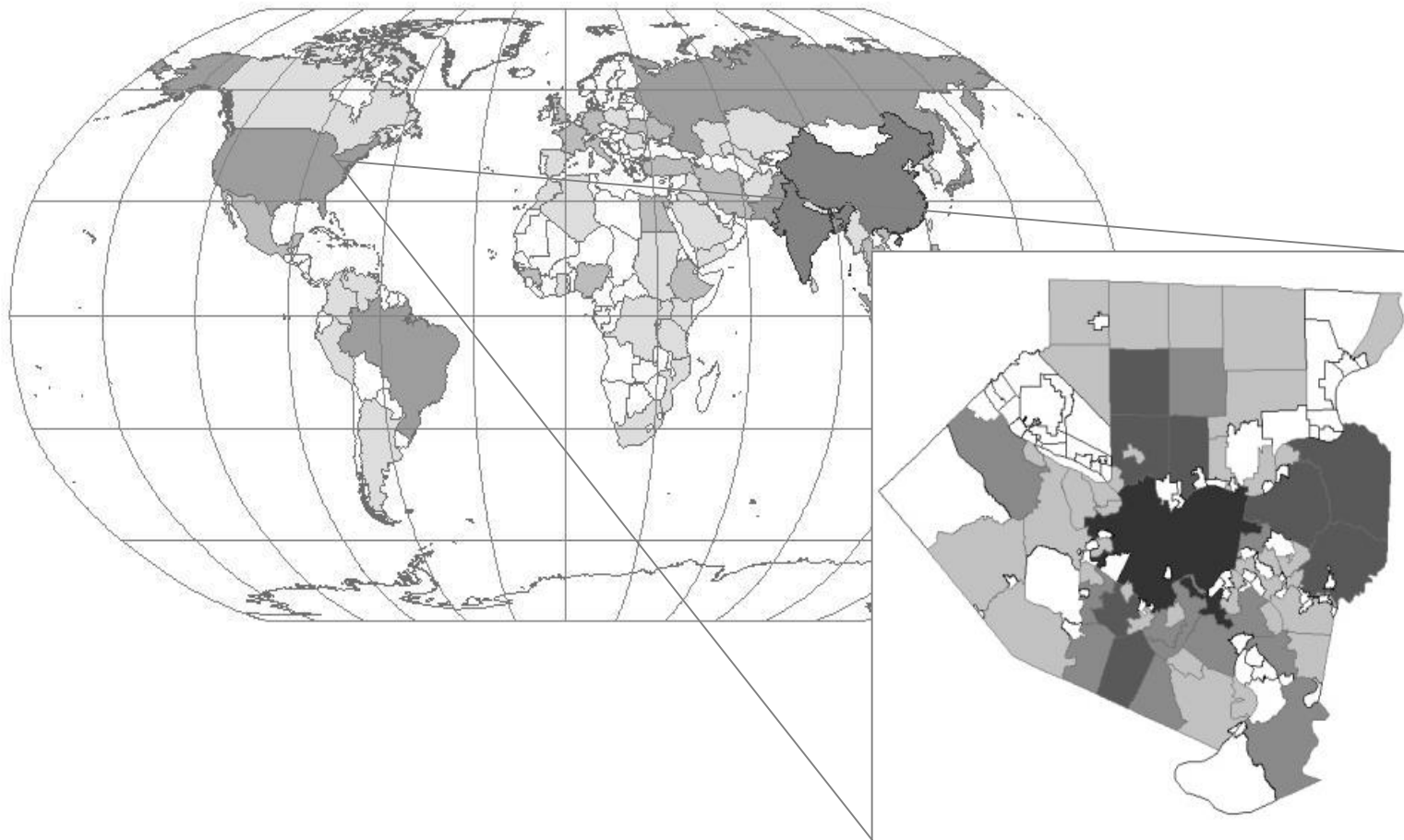
Representing Spatial Elements





Scale of GIS data

Global to local





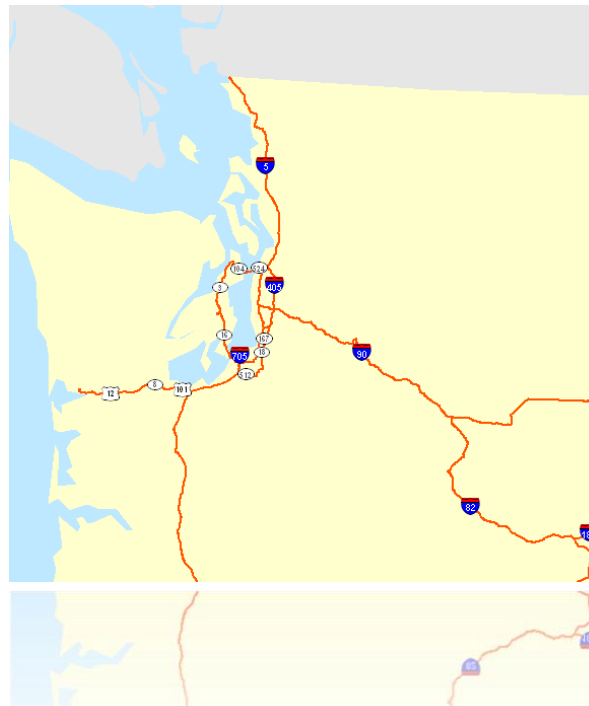
Vector (Data Type 1/3)

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

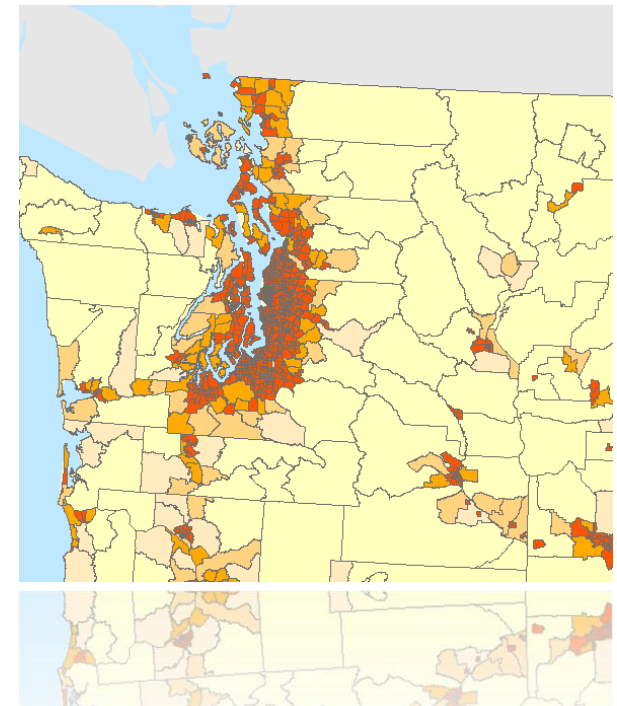
Points: People or Cities (center)



Lines: Roads or Other Linkages



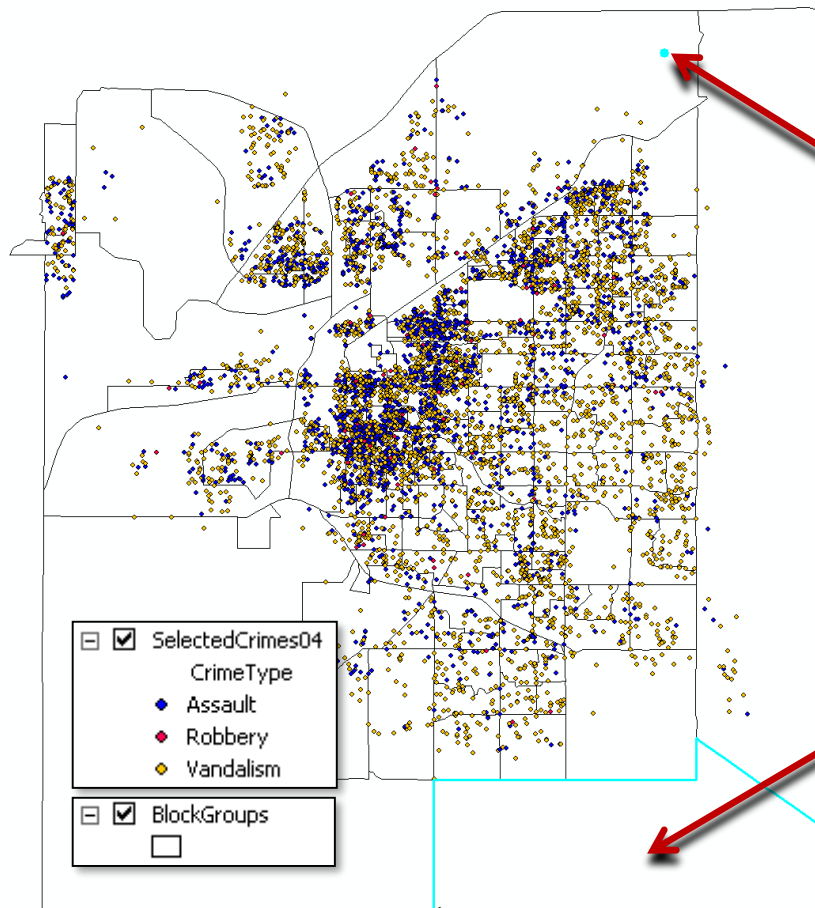
Polygons: Census Tract





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

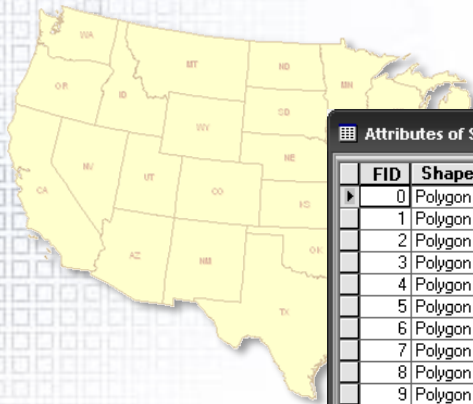
Record: 1 Show: All Selected records



Data Tables (Data Type 2/3)

No inherent geographic information.

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



Attributes of States

FID	Shape	STATE_NAME	STATE_ABBR
0	Polygon	Alaska	AK
1	Polygon	Alabama	AL
2	Polygon	Arkansas	AR
3	Polygon	Arizona	AZ
4	Polygon	California	CA
5	Polygon	Colorado	CO
6	Polygon	Connecticut	CT
7	Polygon	District of Columbia	DC
8	Polygon	Delaware	DE
9	Polygon	Florida	FL
10	Polygon	Georgia	GA
11	Polygon	Hawaii	HI
12	Polygon	Iowa	IA
13	Polygon	Idaho	ID
14	Polygon	Illinois	IL
15	Polygon	Indiana	IN
16	Polygon	Kansas	KS
17	Polygon	Kentucky	KY
18	Polygon	Louisiana	LA

Record: 1 Show: All Select

Vector Data Layer

Attributes of StatesPopStats

OID	STATE_NAME	STATE_FIPS	SUB_REGION	POP2000	POP2007	POP00_SQMI	POP07_SQMI	WHITE	BLAC
0	Alaska	02	Pacific	626932	679581	1.1	1.2	434534	217
1	Alabama	01	East South Central	4447100	4663715	86.1	90.3	3162808	11559
2	Arkansas	05	West South Central	2673400	2889091	50.3	54.3	2138598	4189
3	Arizona	04	Mountain	5130632	6363799	45	55.8	3873611	1588
4	California	06	Pacific	33871648	37483448	214.2	237	20170059	22638
5	Colorado	08	Mountain	4301261	4883413	41.3	46.9	3560005	1650
6	Connecticut	09	New England	3405565	3556875	684.5	714.9	2780355	3098
7	District of Columbia	11	South Atlantic	572059	591318	8412.6	8695.9	176101	3433
8	Delaware	10	South Atlantic	783600	880458	390.4	438.7	584773	1506
9	Florida	12	South Atlantic	15982378	18893813	282.3	333.7	12465029	23355
10	Georgia	13	South Atlantic	8186453	9654958	139.2	164.1	5327281	23495
11	Hawaii	15	Pacific	1211537	1299555	188.5	202.2	294102	220
12	Iowa	19	West North Central	2926324	3030140	52	53.8	2748640	618
13	Idaho	16	Mountain	1293953	1513708	15.5	18.1	1177304	54
14	Illinois	17	East North Central	12419293	13122246	220.4	232.9	9125471	18768
15	Indiana	18	East North Central	6080485	6413133	168	177.2	5320022	5100
16	Kansas	20	West North Central	2688418	2811114	32.7	34.2	2313944	1541
17	Kentucky	21	East South Central	4041769	4258898	100	105.4	3640889	2959
18	Louisiana	22	West South Central	4468976	4385281	95.7	93.9	2856161	14519

Record: 1 Show: All Selected Records (0 out of 52 Selected) Options

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

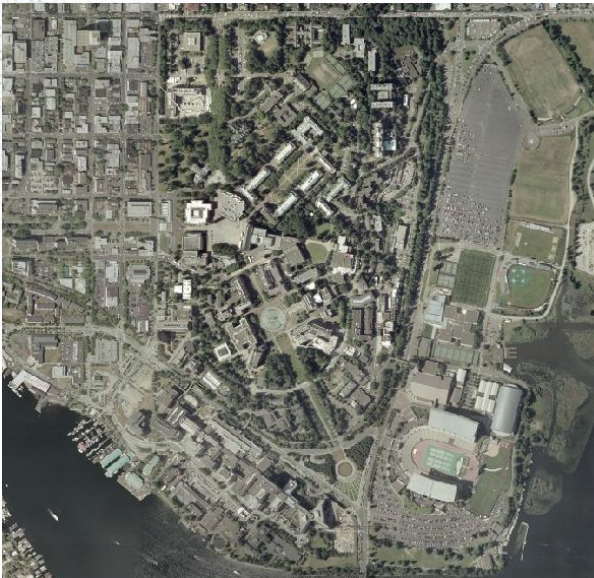


Raster (Data Type 3/3)

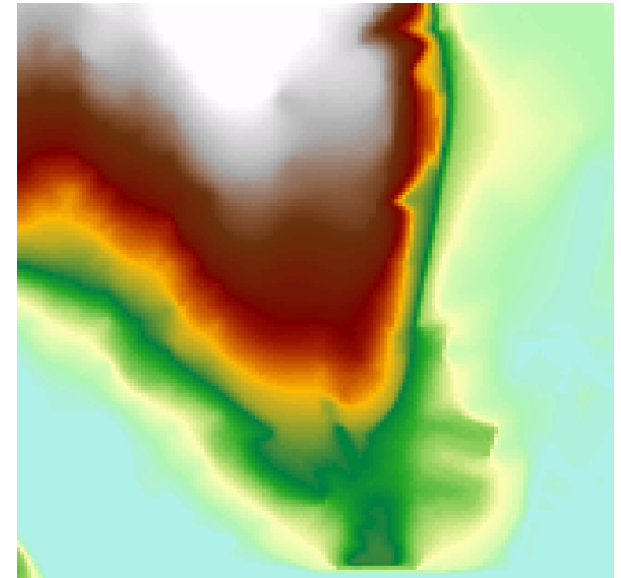
Raster data represents a continuous surface divided into a regular grid of cells.

Often used as background map layer.

Air Photo



Elevation Model

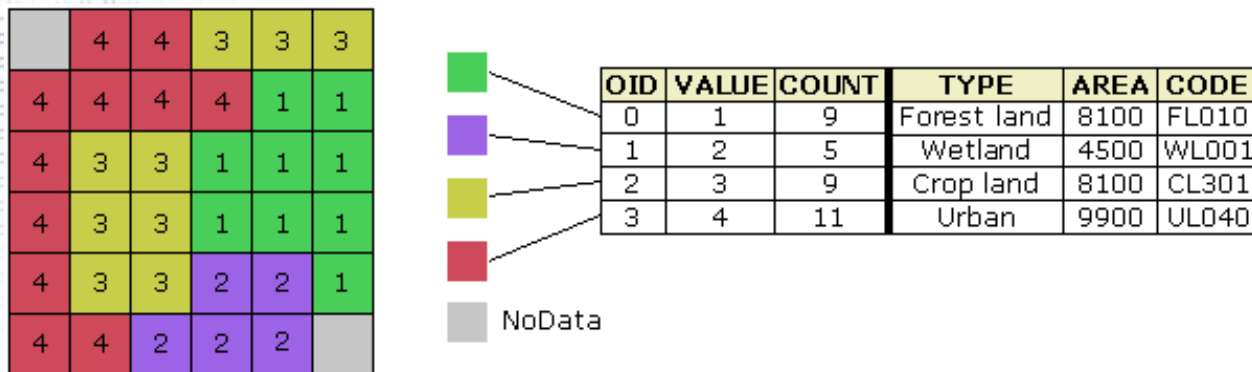




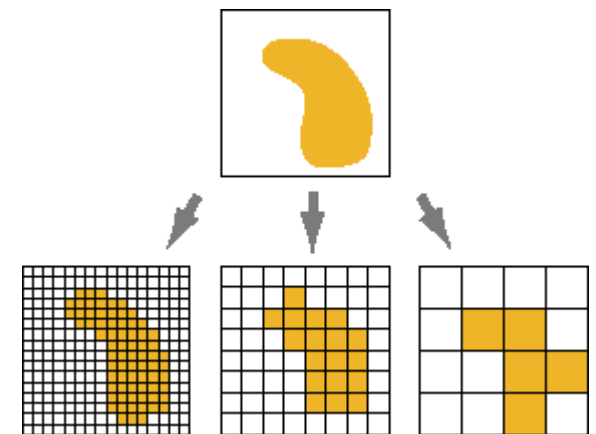
Raster Data Model

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

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



Raster Attribute Tables



Vector vs. Raster
(& raster resolution)

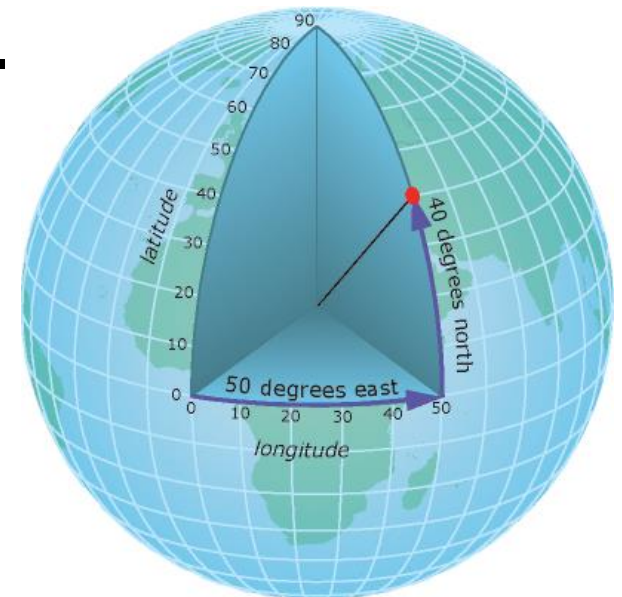


Coordinate Systems

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



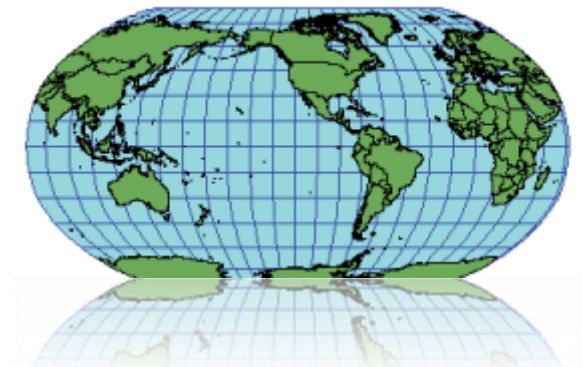
Lambert Conformal Conic



Albers Equal Area Conic



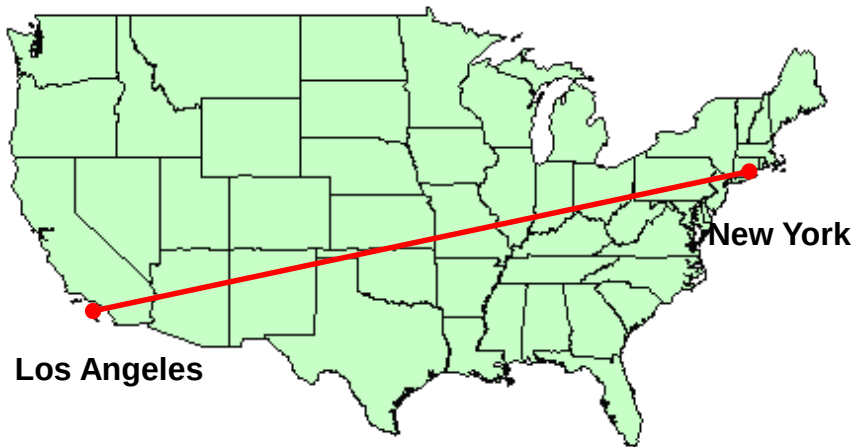
Robinson



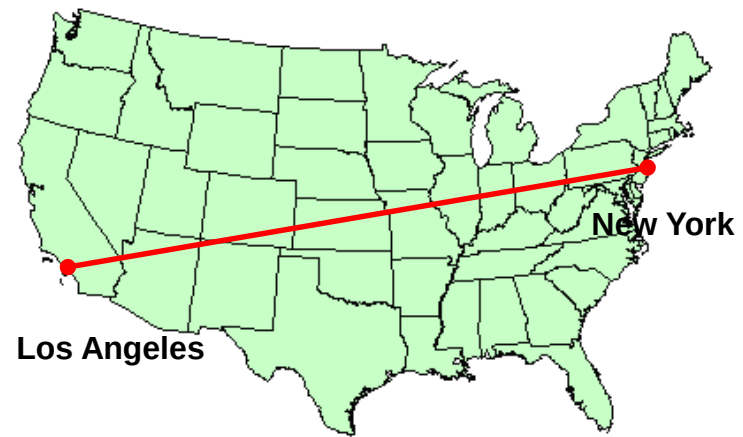


Map Projection is Important

- 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

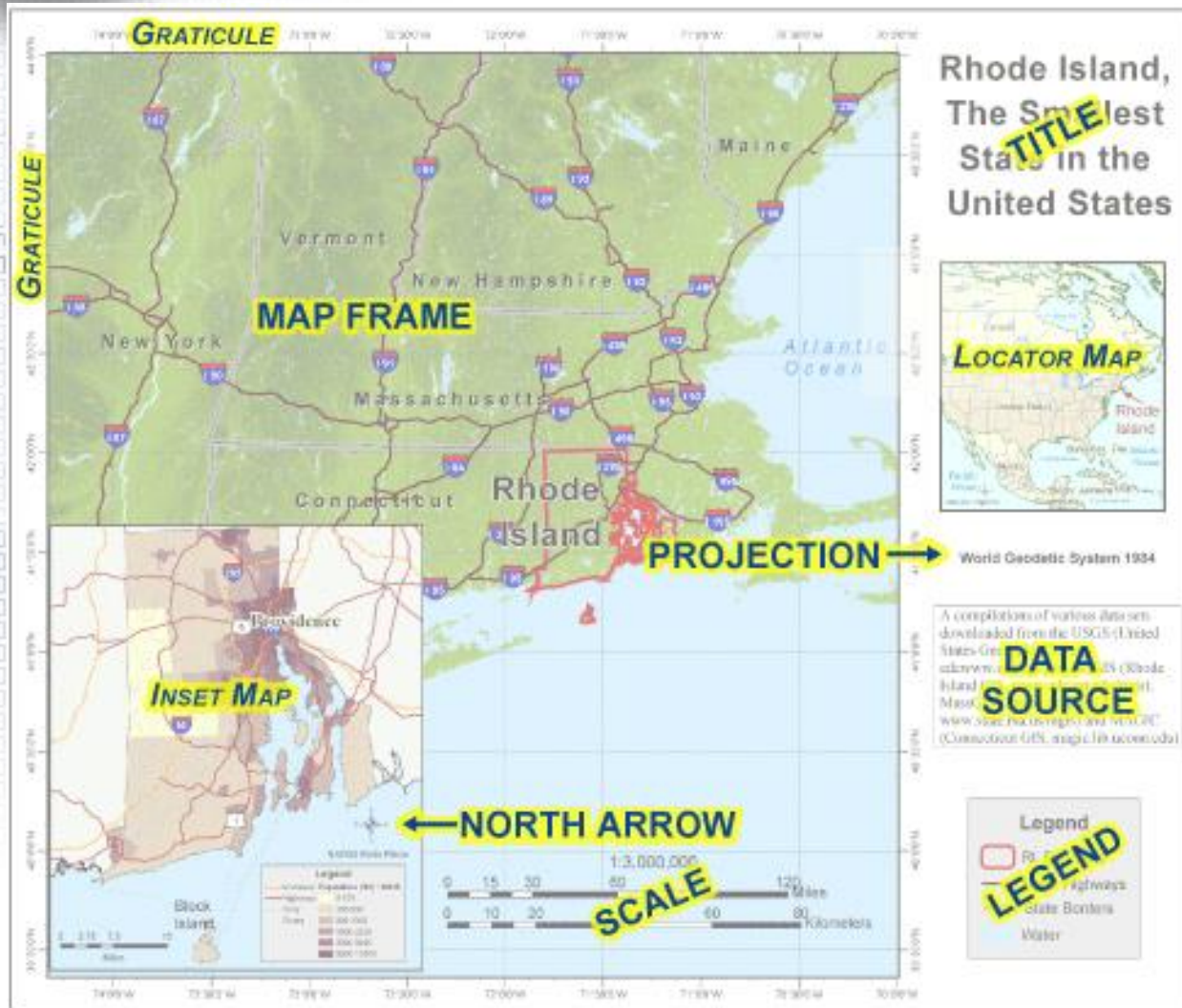


Projection: Albers Equal Area
Distance: 2,455.03 miles

Actual distance: 2,451 miles



Map Making: It's Complex



Many elements
& design theory

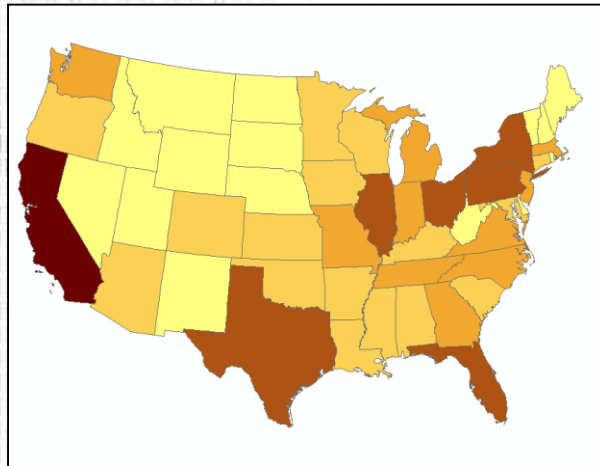
Sub-discipline
in Geography:
Cartography

Tip: Look at
Other Maps



Choosing Map Classification

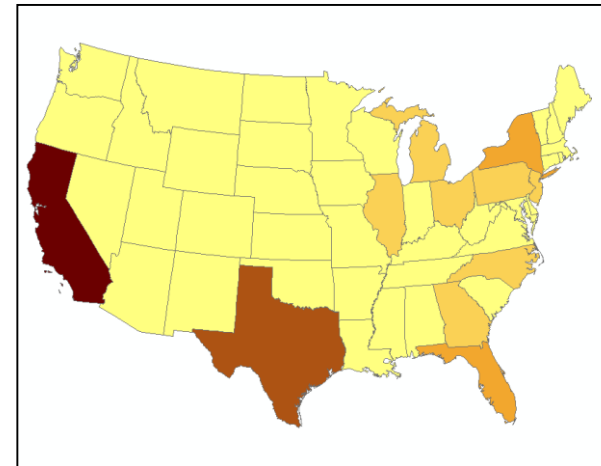
Natural breaks



POP2000

- 493,782 - 2,233,169
- 2,233,170 - 5,363,675
- 5,363,676 - 9,938,444
- 9,938,445 - 20,851,820
- 20,851,821 - 33,871,648

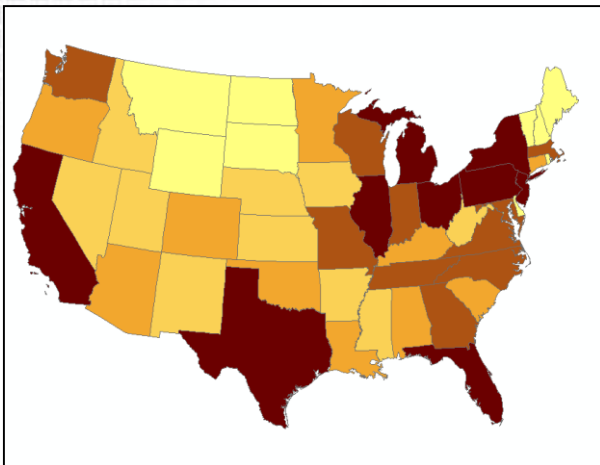
Equal Interval



POP2000

- 493,782 - 7,169,355
- 7,169,356 - 13,844,928
- 13,844,929 - 20,520,502
- 20,520,503 - 27,196,075
- 27,196,076 - 33,871,648

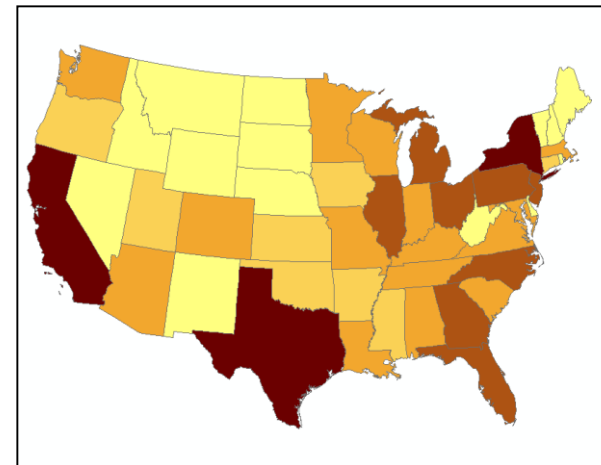
Quantiles



POP2000

- 493,782 - 1,274,923
- 1,274,924 - 2,926,324
- 2,926,325 - 5,130,632
- 5,130,633 - 8,186,453
- 8,186,454 - 33,871,648

Exponential



POP2000

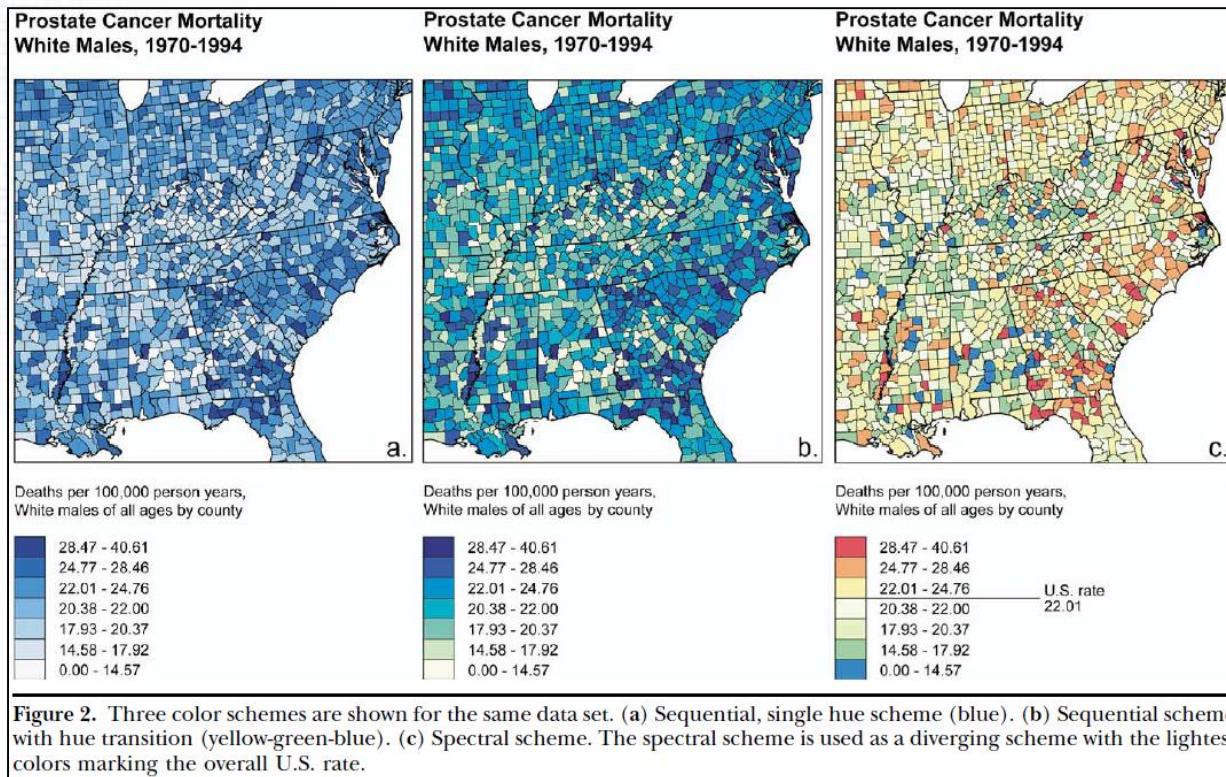
- 0 - 2,000,000
- 2,000,001 - 4,000,000
- 4,000,001 - 8,000,000
- 8,000,001 - 16,000,000
- 16,000,001 and greater



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

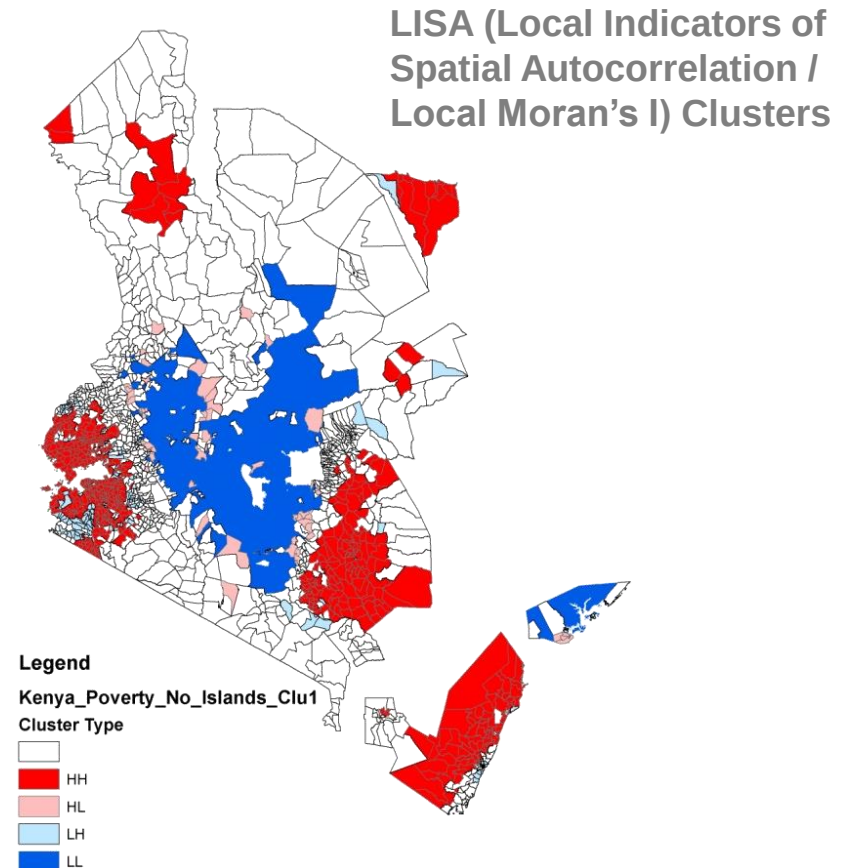
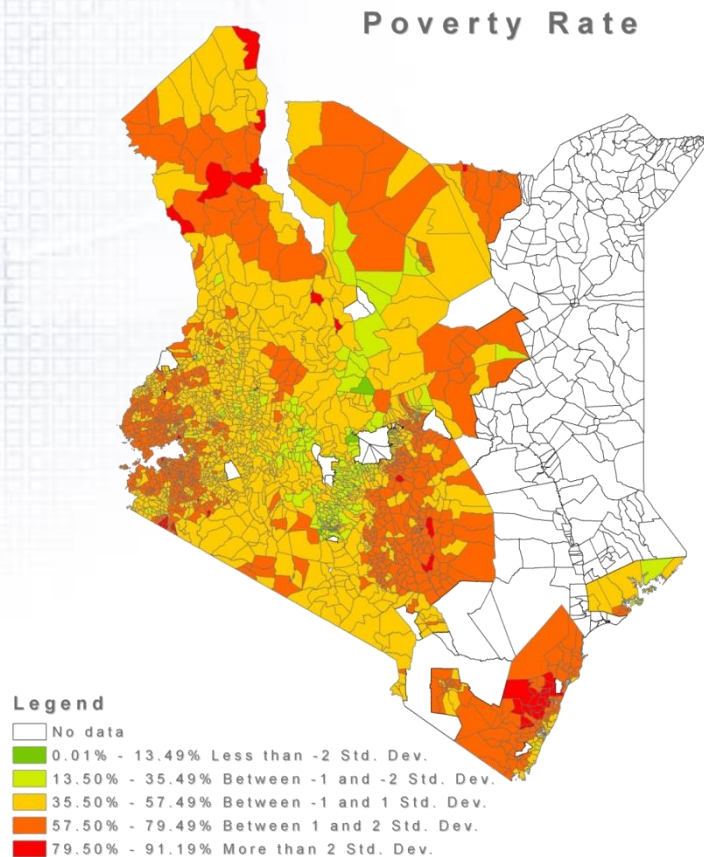


From:
Basic mapping principles for visualizing cancer data using Geographic Information Systems (GIS).
Brewer CA.
Am J Prev Med. 2006 Feb;30(2 Suppl):S25-36.



Analysis Within A Layer

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





Analysis Between Layers

- Comparing the spatial location of features in multiple data layers

