



Lecture 5

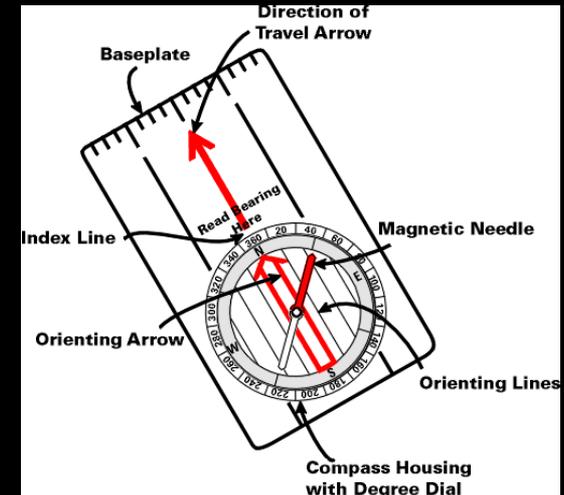
April 19, 2013



# GED 554 – IT & GIS

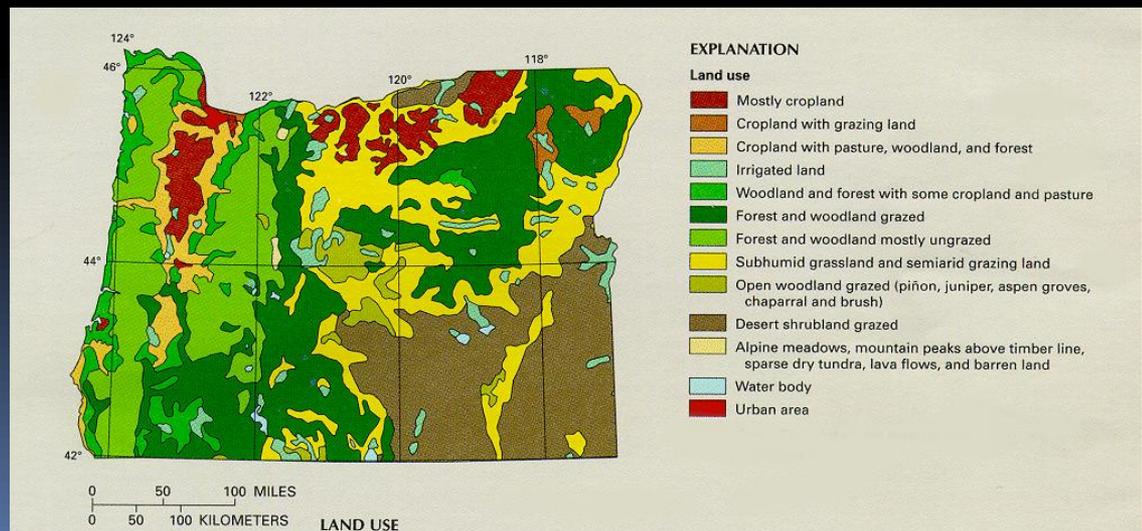
# Properties of Features

- size
- distribution
- pattern
- neighborhood
- shape
- scale
- orientation



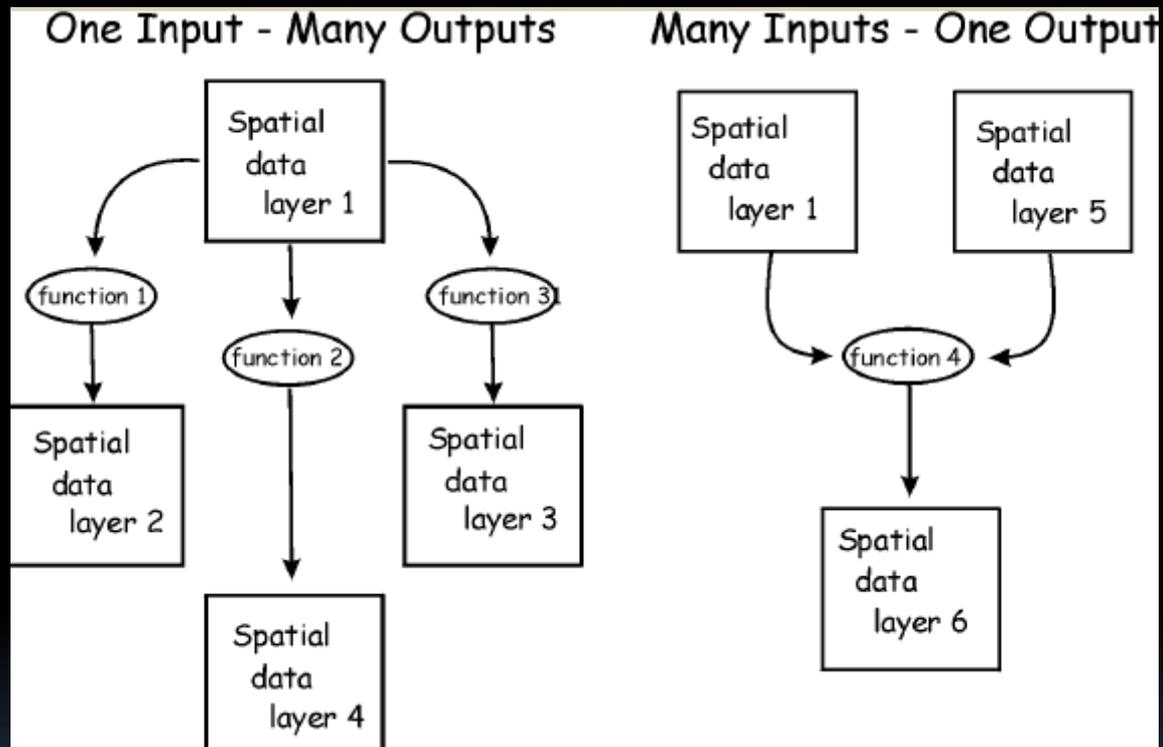
# GIS Analysis

- Much of GIS analysis consists of investigating the properties of geographic features and determining the relationships among them.



# Spatial Data Analysis

- Input -> *spatial operation* -> output



- Turns raw data into useful information
- Reveals patterns, trends, and anomalies that might otherwise be missed



# Spatial data analysis

- *Usually involves manipulations or calculation of coordinates or attribute variables with a various operators (tools), such as:*
    - Selection
    - Reclassification
    - Dissolving
    - Buffering
    - Overlay
    - Cartographic Modeling (a combination of the above)
- 

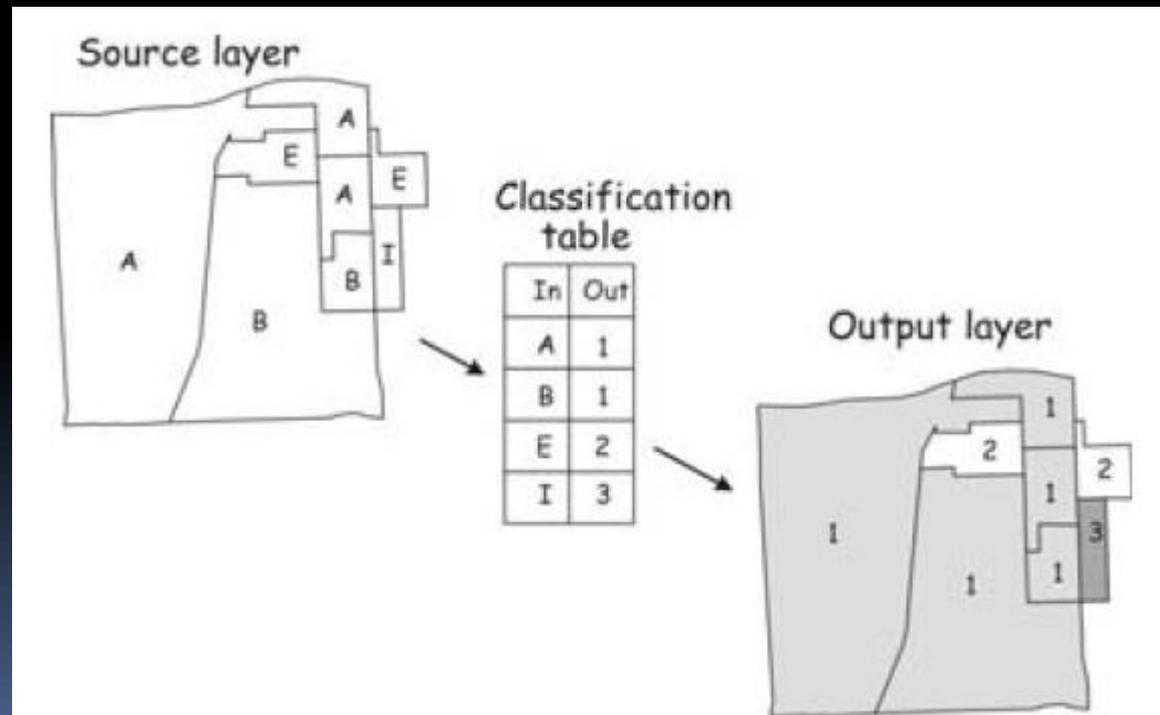
# Spatial data analysis: Spatial Selection

- Identifying features based on spatial criteria
  - Adjacency, connectivity, **containment**, arrangement



# Spatial data analysis: Reclassification

- An assignment of a class or value based on the attributes or geography of an object





## Spatial data analysis :dissolve

- A function whose primary purpose is to combine features within a data layer.
  - Dissolve removes the common boundary.
  - Dissolve is often used after reclassification.
- 

Dissolve: often used after reclassification

Dissolve operation



Boundaries are removed when they separate states with the same value. The dissolve attribute is `_west`.

Dissolve Table

state name	is_west	dissolve value
Alabama	0	E
Arizona	1	W
Arkansas	1	W
Colorado	1	W
Connecticut	0	E
....	...	...
Wyoming	1	W



States west of the main branch of the Mississippi River assigned 1, east of the River assigned 0

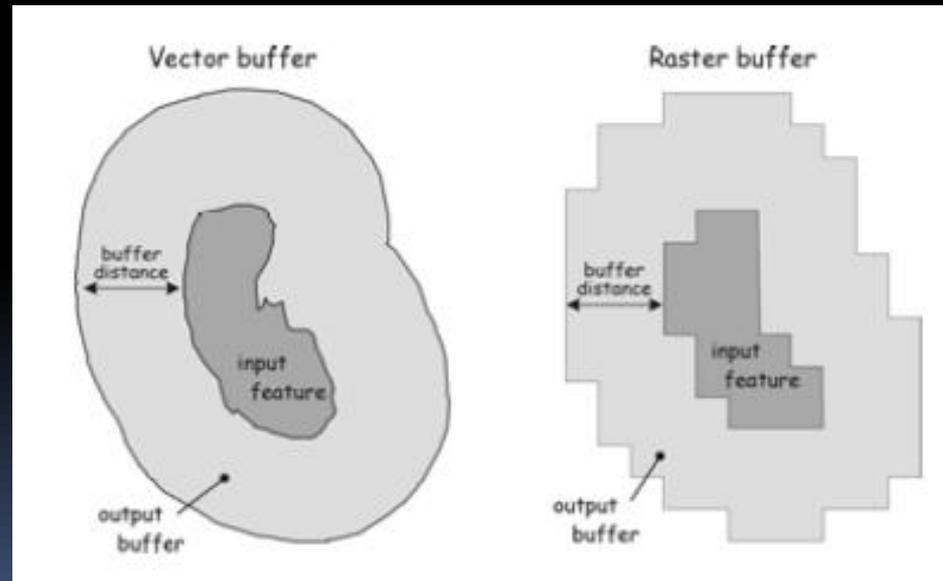
Classification table

state name	is_west
Alabama	0
Arizona	1
Arkansas	1
Colorado	1
Connecticut	0
....	...
Wyoming	1



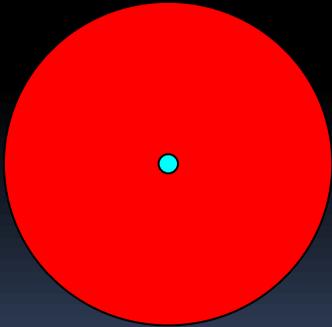
# Buffering (Proximity Tool)

- Buffering is the process of creating one or more zones around selected features, within a pre-specified distance from these features.

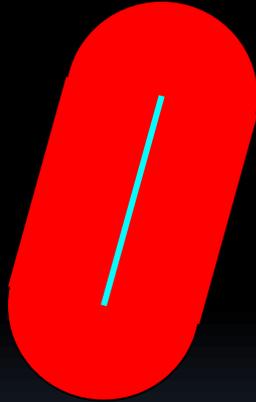


# Buffering

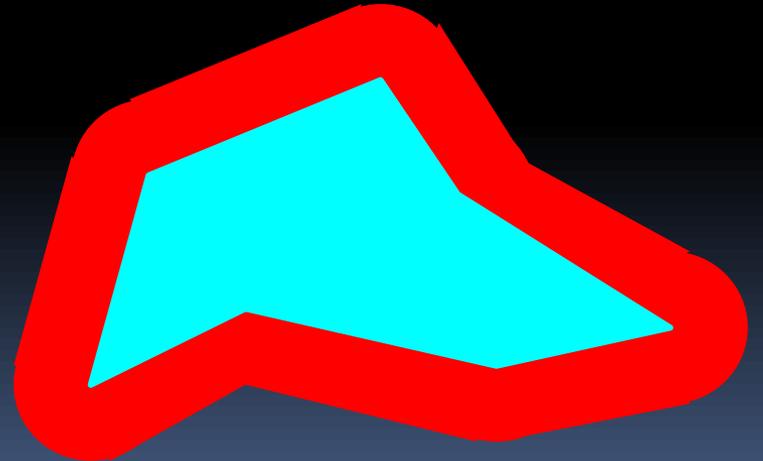
Point



Polyline



Polygon



# Applications of Buffers

- Site selection studies
- Environmental regulations

Number of people living at certain distances from planned bus stops...

Are there any other supermarkets NEAR the new housing development?



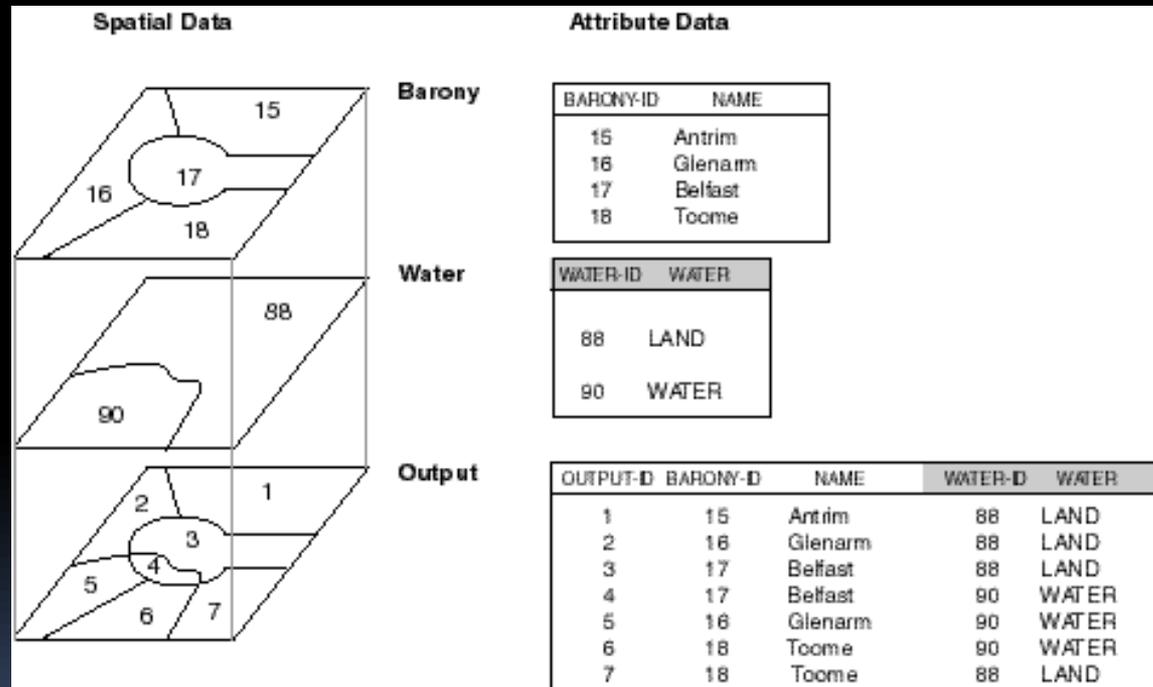
# Spatial Analysis: Overlay

- Combination of different data layers
  - Both spatial and attribute data is combined
  - Requires that data layers use a common coordinate system
  - A new data layer is created
- 

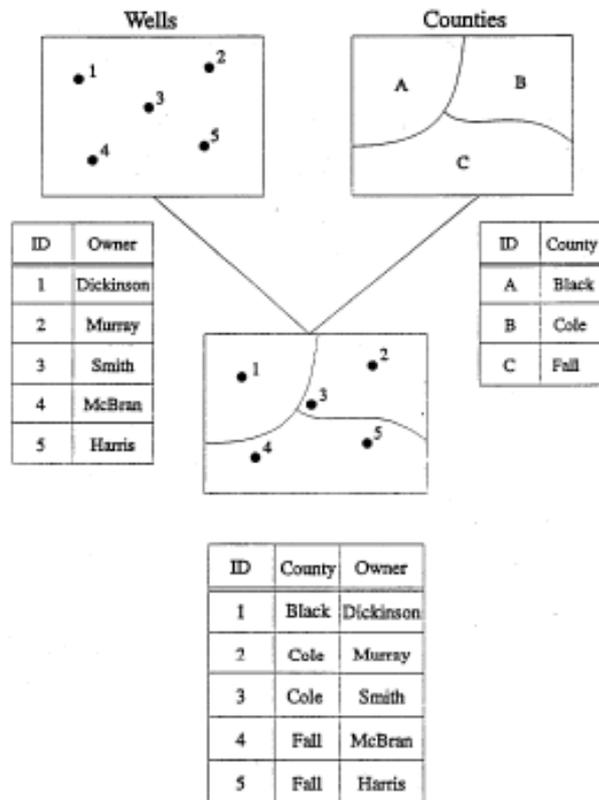
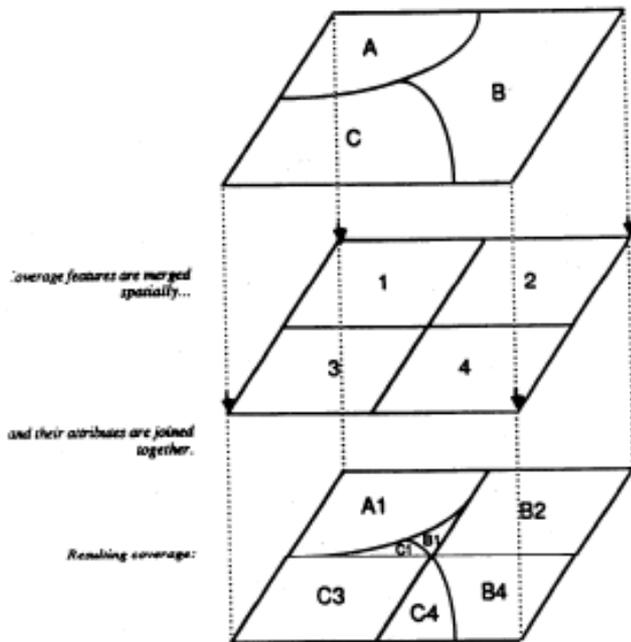
# Data integration: Overlay

Joins two layers to create a new layer

The output layer will contain both the spatial AND attribute data from both of the input layers



### Conceptual View of Map Overlay

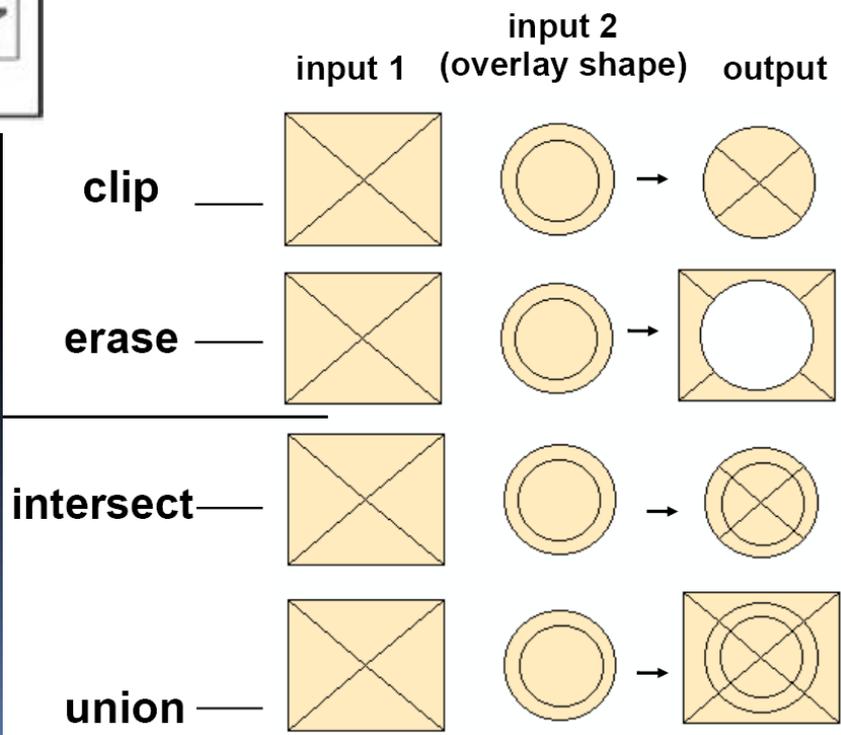
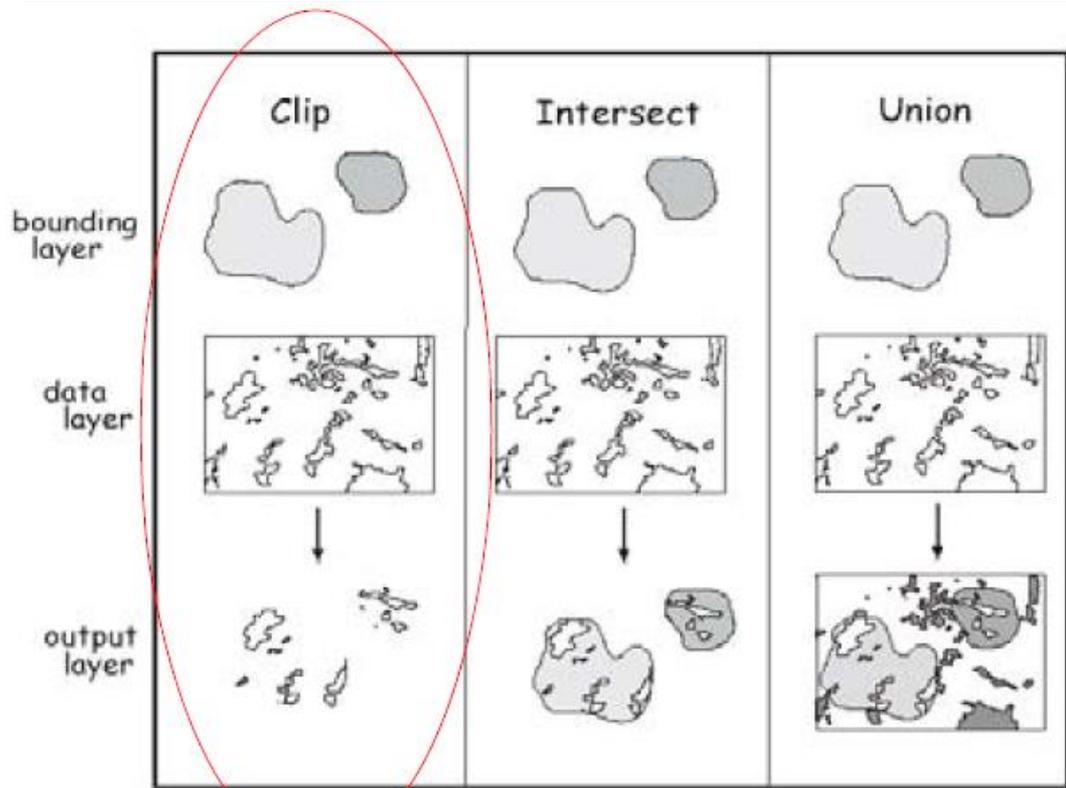


Overlay - Point in Polygon

# Overlay (common ways applied)

- CLIP
- INTERSECTION
- UNION
- ERASE







# CLIP

- Cookie cutter approach
- Bounding polygon defines the clipped second layer
- Neither the bounding polygon attributes nor geographic data (spatial data) are included in the output layer

# INTERSECTION

- Combines data from both layers but only for the bounding area
  - *(Bounding polygon also defines the output layer)*
  - *Data from both layers are combined*
  - *Data outside the bounding layer (1st layer) is discarded)*
- Order of intersection is important  
*(A to B or B to A)*



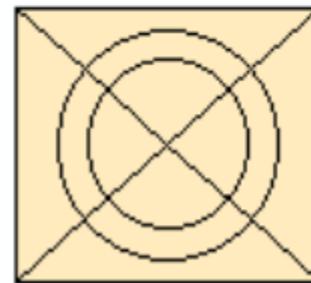
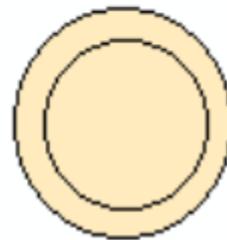
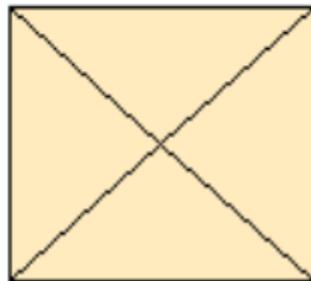
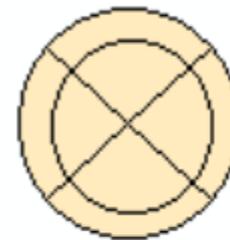
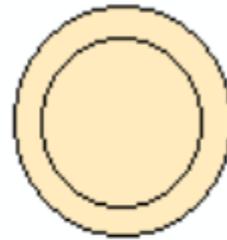
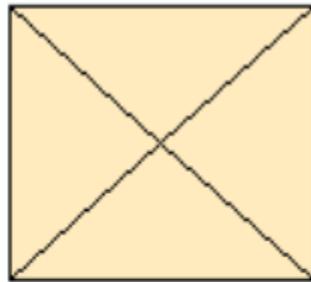
# UNION

- Includes all data from both the bounding and data layers
  - New polygons are formed by the combinations of the data from each layer
- 



QUESTIONS?

Which types of operations are shown here?



For which operation is the **order** of layers of importance?

Clip

Union

Intersect

All

Which of these operations affect the shape  
AND attribute table (graphics + DB)?

Clip

Erase

Intersect

All

# Tables in ArcGIS

- Tables contain attribute data ( “inside” a layer )
- Each line is a **record (per feature)**
- Each attribute or **field (column) has a type (numbers, words)**
- First attribute: **FID** (field ID) or **OID** (object ID) - don't alter!
- Dbase (.dbf) files, text files (.txt, .csv), Excel spreadsheet files can be imported/loaded into ArcGIS

# Types of tables

## Attribute table

- Stores attributes of map features
- Already associated with a spatial data layer (georef' ed.)

FID	Shape*	NAME	STATE_NAME
0	Polygon	Lake of the Woods	Minnesota
1	Polygon	Ferry	Washington
2	Polygon	Stevens	Washington
3	Polygon	Okanogan	Washington
4	Polygon	Pend Oreille	Washington

ArcCatalog: counties.shp

## Standalone table

- Stores any tabular data ( spreadsheet)
- **Not associated with** spatial data
- When imported into ArcGIS: **OID** instead of FID

OID	FIPS	POP1998	POP1997	POP1996
0	01001	42095	41284	40251
1	01003	132828	128820	124257
2	01005	26895	26791	26870
3	01007	18926	18595	18227
4	01009	46266	44930	43548

ArcCatalog: Tabo1.dbf

# Fields

- Fields have specific types (numbers, words, ...)
- Type must be defined before use
- Once defined, type cannot be changed
- Naming rules for field (attribute) names:
- No more than 13 characters
  - Use only letters and numbers (no space!)
  - Must start with a letter
  - You can define longer (better) names with an alias (in: Layer properties - Fields, "Pop Density")

# Field types (data formats)

<b>Short</b>	Small integers (stored as 4-byte binary numbers)	255 1201
<b>Long</b>	Large integers (stored as 10-byte binary numbers)	156000
<b>Float</b>	Floating point values with <b>8</b> significant digits in the mantissa	1.2893851E+12
<b>Double</b>	Double-precision floating point values with <b>16</b> significant digits in the mantissa	1.111111111111110E+1
<b>Text</b>	Alphanumeric strings (< 255 letters?)	'Maple St'
<b>Date</b>	Date format	07/12/92
<b>BLOB</b>	Binary large object; any complex binary data including images, documents, etc. (uncommon!)	

# QUESTION?

- You create a table column - what **type** for: Integer?
- **Population of countries in the world?** (Short or Long)
- 
- **Average price for 1 gallon of gas?**
- 
- **Number of counties in a state?** Text? (String)
- 
- **Highway names?** Floating point number?  
(Float or Double)
- 
- **Distances between cities, in meters**
-

# Field properties tab

The screenshot shows the 'Layer Properties' dialog box with the 'Fields' tab selected. The 'Primary Display Field' is set to 'RISDATA'. Below this, a table lists various fields with their properties. Annotations highlight specific features: 'Label with' points to the 'RISDATA' field in the dropdown; 'Show/Hide field' points to the checkboxes in the first column; and 'Field alias (long, descriptive)' points to the 'COVERLTYPE' alias for the 'COV\_TYPE' field.

Layer Properties

General | **Fields** | Symbology | Definition Query | Labels | Joins & Relates

Primary Display Field: RISDATA

Choose which fields will be visible. Click in the alias column to edit the alias for any field.

Name	Alias	Type	Length	Precision	Scale	Number Format
<input checked="" type="checkbox"/> Shape_1		Polygon				
<input checked="" type="checkbox"/> AREA	AREA	Double	8	0	0	Numeric
<input checked="" type="checkbox"/> PERIMETER	PERIMETER	Double	8	0	0	Numeric
<input checked="" type="checkbox"/> STANDS2#	STANDS2#	Long	4	0	0	Numeric
<input checked="" type="checkbox"/> STANDS2-ID	STANDS2-ID	Long	4	0	0	Numeric
<input checked="" type="checkbox"/> RISDATA	RISDATA	String	10	0	0	
<input checked="" type="checkbox"/> DATA	DATA	String	10	0	0	
<input checked="" type="checkbox"/> OWNER	OWNER	String	3	0	0	
<input checked="" type="checkbox"/> COV_TYPE	COVERLTYPE	String	3	0	0	
<input checked="" type="checkbox"/> SSTAGE96	SSTAGE96	String	4	0	0	

Select All Clear All

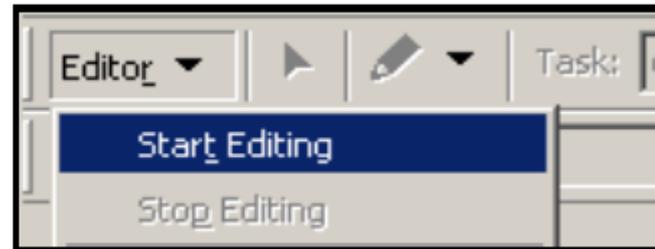
OK Cancel Apply

# Editing single records within fields (ArcMap)

Open Editor toolbar (Tools)



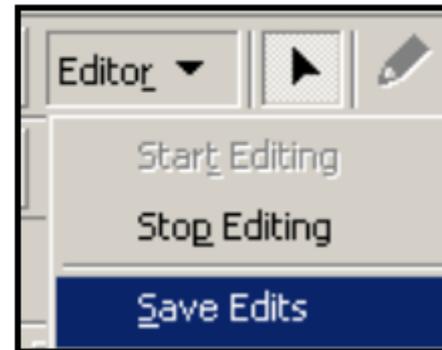
Editor Tool bar: Start editing



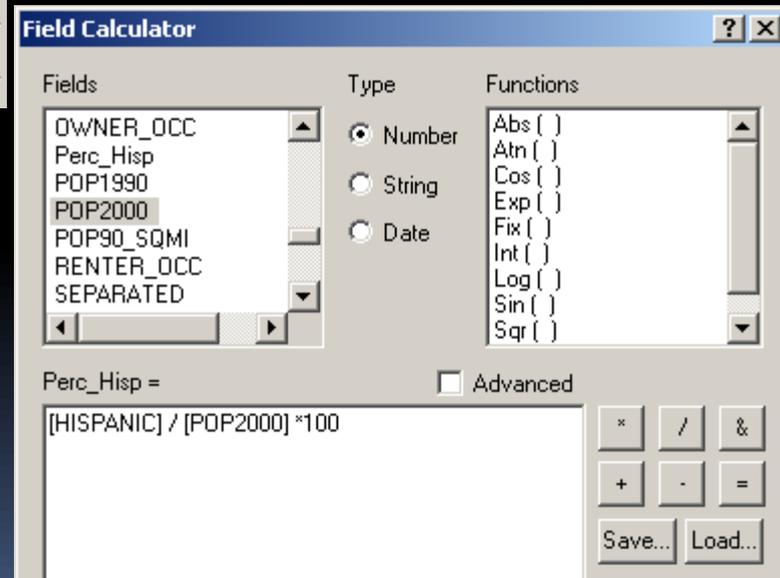
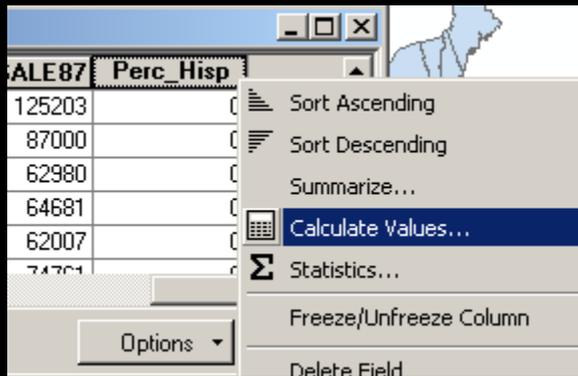
Type edits into Table

Ave_MAG	Ave_MMI	Risk
6.5042	3.7358	High
3.6143	6.1429	Low
5.0113	6.5625	High
2.135	5.9	Lo
6.35	7	
6.3167	8.5	

Save edits - stop editing



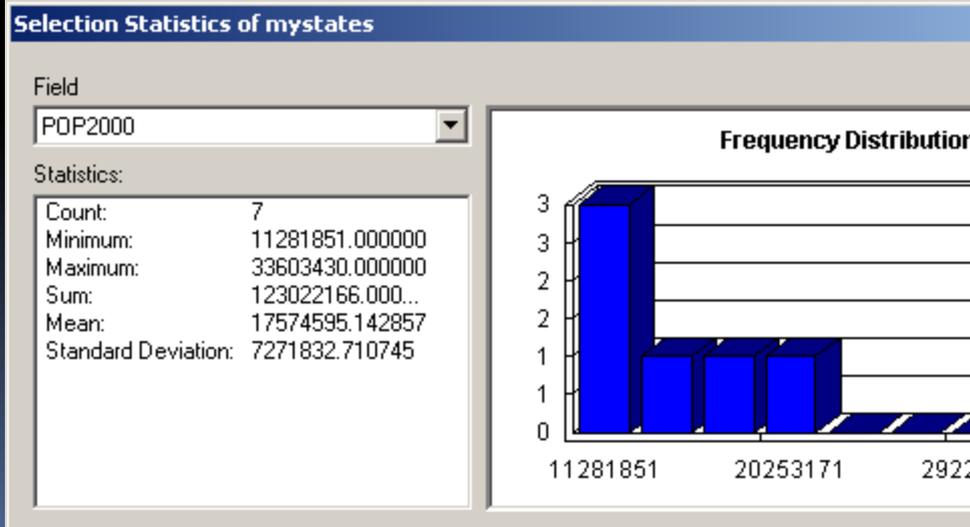
# Calculate field values



# Field statistics

POP2000	POP90	SAMI	HOUSEHOL
229	1184		
692	5835		
065	885		
928	1257		
800	631		
004	774		

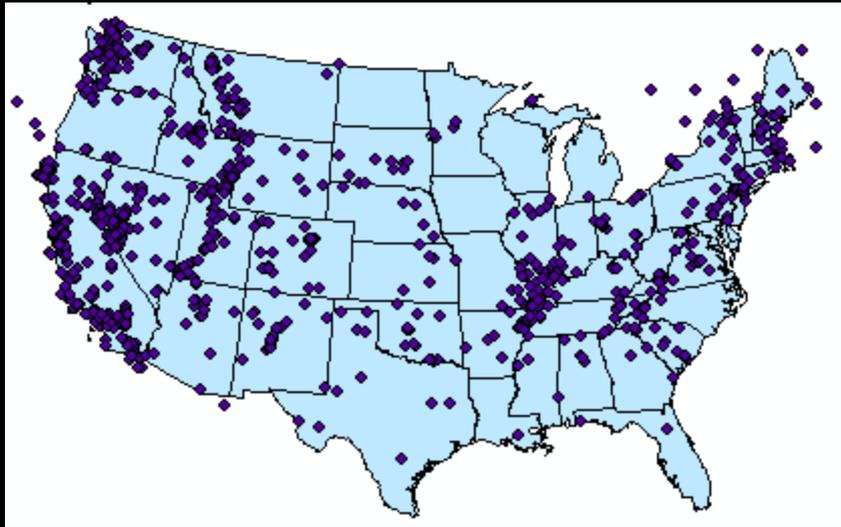
Sort Ascending  
Sort Descending  
Summarize...  
Calculate Values...  
**Statistics...**  
Freeze/Unfreeze Column  
Delete Field



## Summarize (different kind of statistics)

- Calculate statistics **for groups of features** in a table  
(Remember: “Statistics”: simply per attribute!)
- **Groups** are defined by unique values in the specified field
- Produces a **standalone table (not a layer!)** as output with groups and stats

# Example: Summarize earthquakes by state



- Group earthquakes **by state** field
- Determine total deaths, total damage, and average magnitude of earthquakes **per state** (in each state).

ID	STATE	DEPTH	DEATHS	DAMAGE	MAG	MMI	LOCATION
1	MO	0	7	0	7.88	12	New Madrid, Missouri
2	SN	0	51	0	7.36	12	Northern Sonora, Mexico
3	AK	0	0	0	8.15	11	Yakutat Bay, Alaska
4	AK	0	0	0	8.26	11	Southeast Alaska
5	AR	0	7	0	7.68	11	Northeast Arkansas
6	CA	20	3000	52400000	7.80	11	Near San Francisco, California
7	CA	16	12	6000000	7.48	11	South of Bakersfield, California
8	CA	8	65	50500000	6.62	11	North of San Fernando, California

# How to summarize

STATE	DEPTH	DEATHS	DAMAGE	MAG	MNI	LOCATION
MO	0	7	0	7.88	12	New Madrid, Missouri
SN	0	51	0	7.36	12	Northern Sonora, Mexico
AK	0	0	0	8.15	11	Yakutat Bay, Alaska
AK	0	0	0	8.26	11	Southeast Alaska
AR	7	7	0	7.68	11	Northeast Arkansas
CA	3000	52400000	7.80	11	Near San Francisco, Ca	
CA	12	6000000	7.48	11	South of Bakersfield, Ca	
CA	65	50500000	6.62	11	North of San Fernando,	

Choose grouping attribute: Right-click on STATE field - Summarize

### Summarize

Summarize creates a new table containing one record for each unique value of the selected field, along with statistics summarizing any of the other fields.

1. Select a field to summarize:  
STATE
2. Choose one or more summary statistics to be included in the output table:  
 FID  
 HOUR  
 LOCATION  
 MAG

Choose the type **stats** given for each attribute

- Minimum
- Maximum
- Average
- Sum
- Standard Deviation
- Variance

Specify output table:  
M:\GIS\mgisdata\temp\quakesum.dbf

Summarize on the selected records only

More about Summarize... OK Cancel

## ArcGIS Dialog Help

### About summarizing data

This dialog allows a table of summary information to be generated for the fields in the selected table. Unique values from a specified field are used as a basis for the summary information.

In this example, the table is summarized on field F1, and the sum of field F2 is added.

F1	F2
A	5
A	5
A	2
B	10
B	20

Input

→

F1	Sum_F2
A	12
B	30

Output

Summarize on the selected records only

# QUESTION?

USA county data (per county: NAME, POP and STATE)

- use query, "statistics" or "summarize" for:

- **Find all counties with > 20,000 people?** Summarize?
- Statistics?
- **Find total number of people in each state?**
- 
- 
- SQL query?
- **Total number of people in US?**
- 
-

# Joining tables (database joins)

- Temporary relationships between 2 tables to create a single table
- To Join tables must share a **common** field (key field)
- Keys do not have to have the same name, just the mean the **same concept**
- Joins are temporary and can be removed again
- To make permanent: export shapefile with joined tables

Destination



# Joining tables

Source



Destination table

Source table

Input

FID	Shape*	AREA	STATE_NAME	STATE_FIPS
0	Polygon	67230.061	Washington	53
1	Polygon	147244.653	Montana	30
2	Polygon	32161.925	Maine	23
3	Polygon	70812.056	North Dakota	38
4	Polygon	77195.055	South Dakota	46
5	Polygon	57803.139	Wyoming	56
6	Polygon	56088.178	Wisconsin	55

STATE_FIPS	POP1990	POP1999	POP90_SQMI	HOUSEHOLD
53	4866892	5773907	72	1872431
30	799065	884214	5	306163
23		1248908	38	465312
38				
46				
56				
55				
16				

Join tables on common field (STATE\_FIPS = key field)

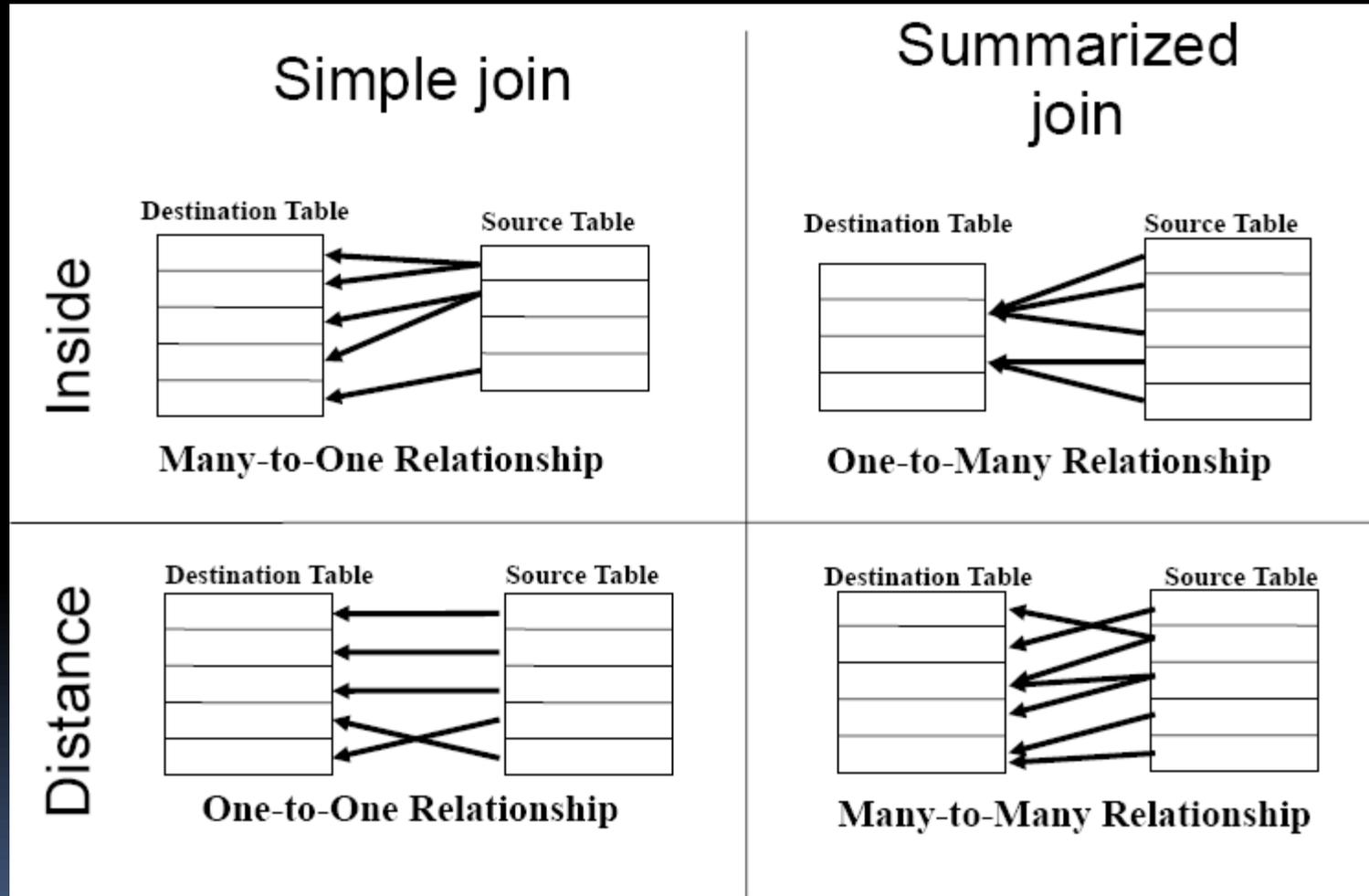
Input

Output

FID	Shape*	AREA	STATE_NAME	STATE_FIPS	POP1990	POP1999	POP90_SQMI
0	Polygon	67230.061	Washington	53	4866892	5773907	72
1	Polygon	147244.653	Montana	30	799065	884214	5
2	Polygon	32161.925	Maine	23	1227528	1248908	38
3	Polygon	70812.056	North Dakota	38	638900	637016	9
4	Polygon	77195.055	South Dakota	46	696004	739508	9
5	Polygon	57803.139	Wyoming	56	453588	482025	5
6	Polygon	56088.178	Wisconsin	55	4891769	5251093	57
7	Polygon	80343.643	Idaho	16	1006749	1250242	

Joined table

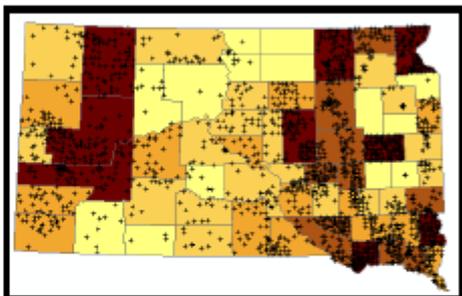
# Spatial joins: Four cases



## Simple

## Summarized

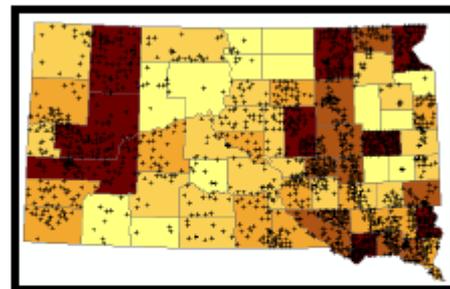
Inside



Hospitals  $\leftarrow$  Counties

Which county is each hospital in?

\_\_\_\_\_ to \_\_\_\_\_

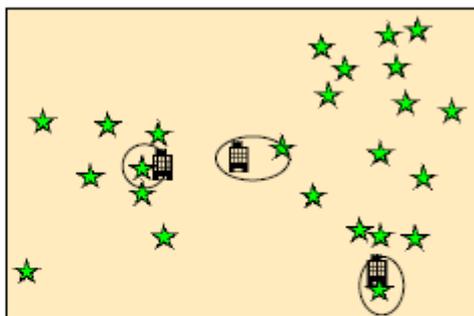


Counties  $\leftarrow$  Hospitals

How many hospitals in each of the

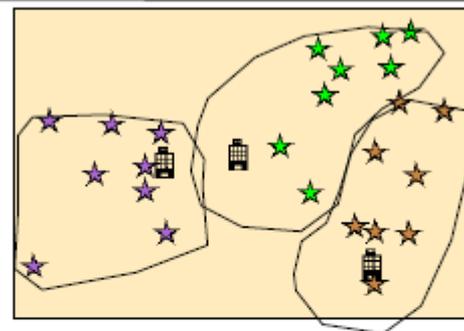
counties? \_\_\_\_\_ to \_\_\_\_\_

Distance



Hotels  $\leftarrow$  Attractions

Which attraction is **closest** to each hotel? How far is it? \_\_\_\_\_ to \_\_\_\_\_



Hotels  $\leftarrow$  Attractions

How many attractions are **closest** to each hotel? \_\_\_\_\_ to \_\_\_\_\_

# QUESTION?

What is the distance of each county to “its” closest hospital?

Type of join? \_\_\_\_\_

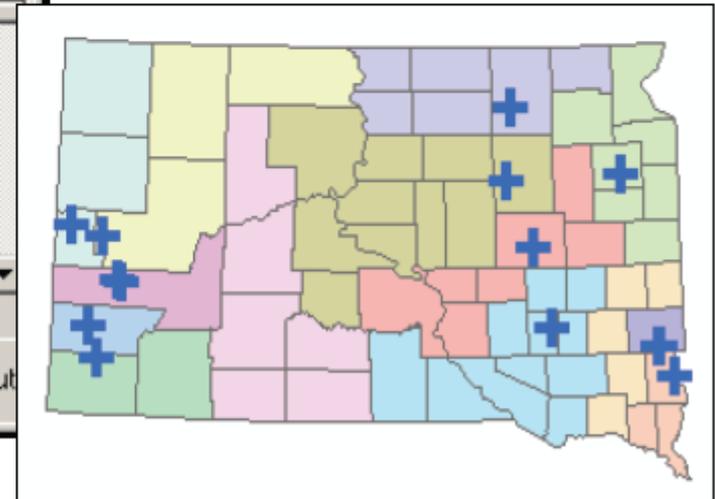
Dest.: \_\_\_\_\_ Source: \_\_\_\_\_

Each county features gets name of closest hospital and the distance (colored by distance)

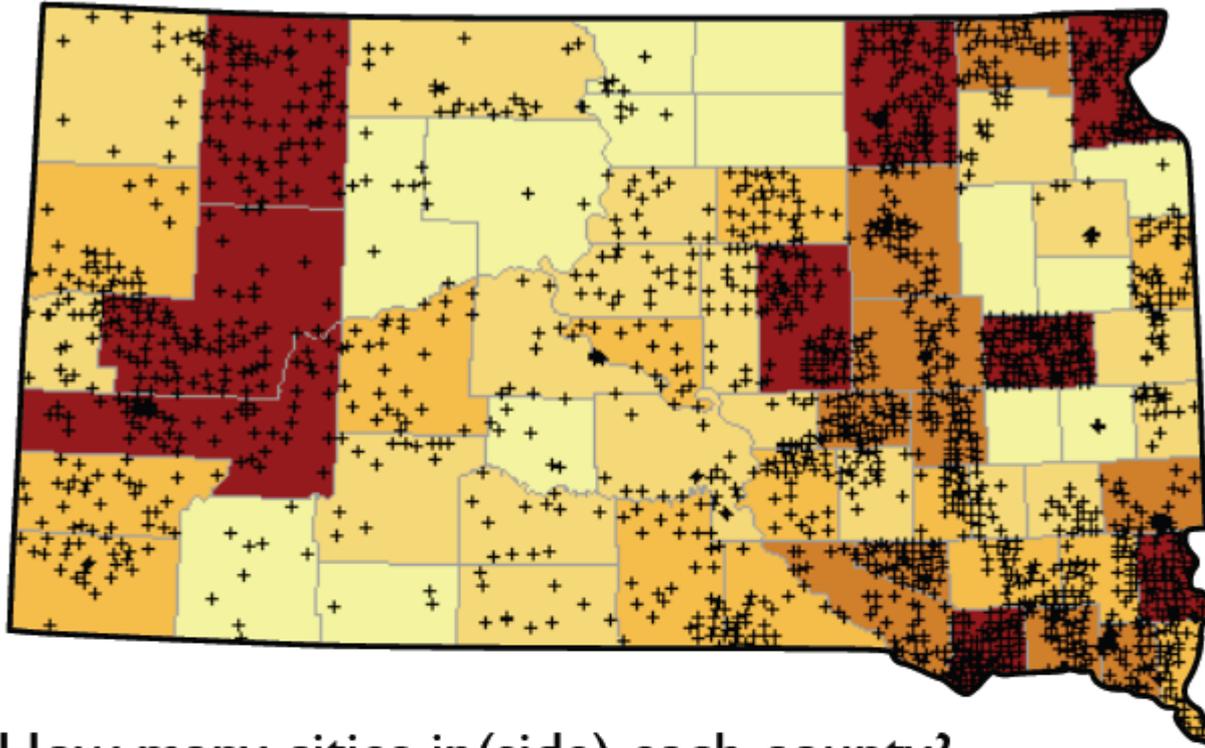
Attributes of Cty\_hosp

NAME	NAME_1	Distance
Shannon	Battle Mountain National Sanitarium	0.474633
Fall River	Battle Mountain National Sanitarium	0
Pennington	Bennett Clarkson Hospital	0
Lincoln	Canton-Inwood Hospital	0
Union	Canton-Inwood Hospital	0.216911
Clay	Canton-Inwood Hospital	0.336159
Custer	Custer State Hospital	0

Record: 0 Show: All Selected Records (0 out)



# QUESTION?



How many cities in(side) each county?  
(How many people live in each county's cities?)

Type of join? \_\_\_\_\_ Card.? \_\_\_ - \_\_\_

Dest.: \_\_\_\_\_ Source: \_\_\_\_\_

# ArcToolbox

is organized into toolsets that provide solutions for different types of tasks. The three main toolsets are:

- Data Management Tools
- Analysis Tools
- Conversion Tools

