CRP 4080: Introduction to Geographic Information Systems for planners

Lecture 4: Geoprocessing

Wenzheng Li, Ph.D.
City and Regional Planning
Fall 2024

Milestone

Intro: Getting to Know ArcGIS Pro and Basic Mapping

- What is GIS
- Operating ArcGIS Pro
- Thematic Mapping
- Map Projection

- **Now…Geo-spatial data management**

Final project

Proposals (1-2 paragraphs)

- Due: October 10 (submit via Canvas)
- Project idea (research question)
- Possible data sources (with website links if possible)

Project idea must require use of GIS (ie have some sort of spatial component)

- Analytical, not just descriptive: Use GIS to answer a question e.g. suitability study
- Exit (capstone) projects, client-based projects

Final Project examples will be posted on Thursday….

Recap

What is Spatial Analysis in GIS?

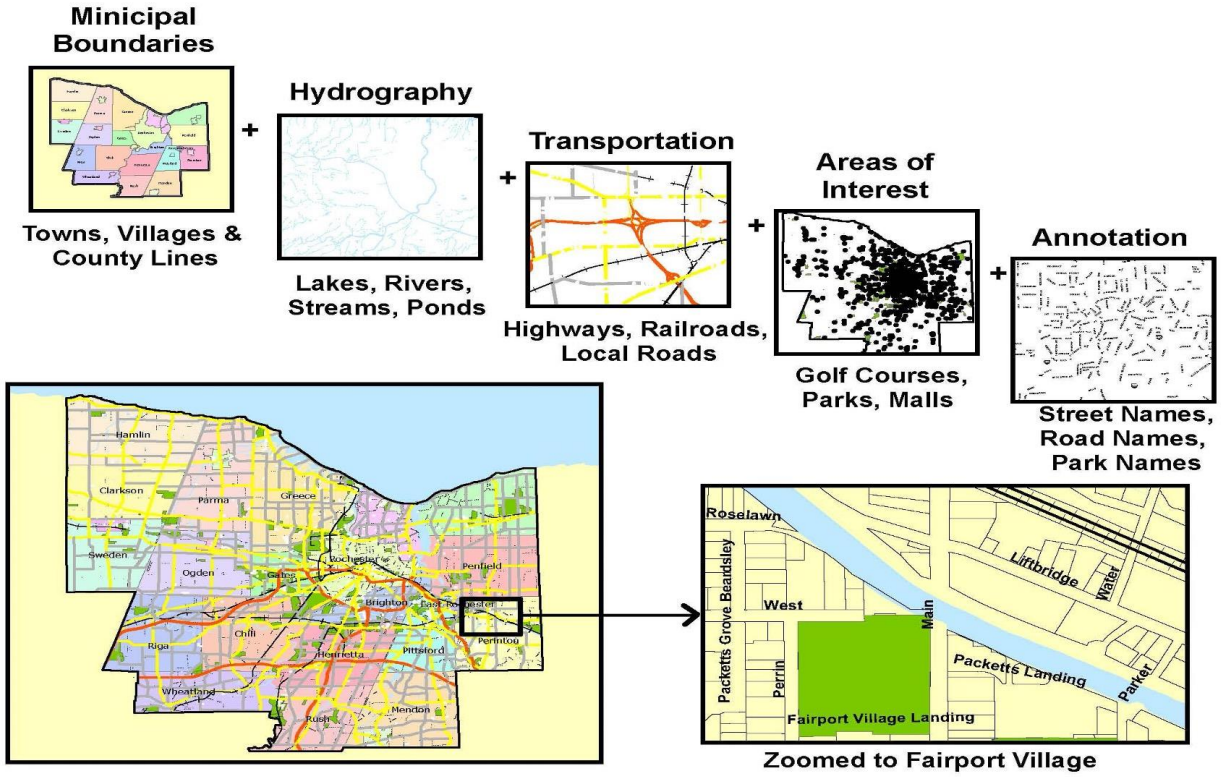


Image source

- First step: **Overlay** spatial features/data to tell a story, make an argument, solve or understand a problem, etc about **locations**.
- Second step: Share your spatial analysis in reports, maps, tables, and charts through visual communication.

Spatial Analysis

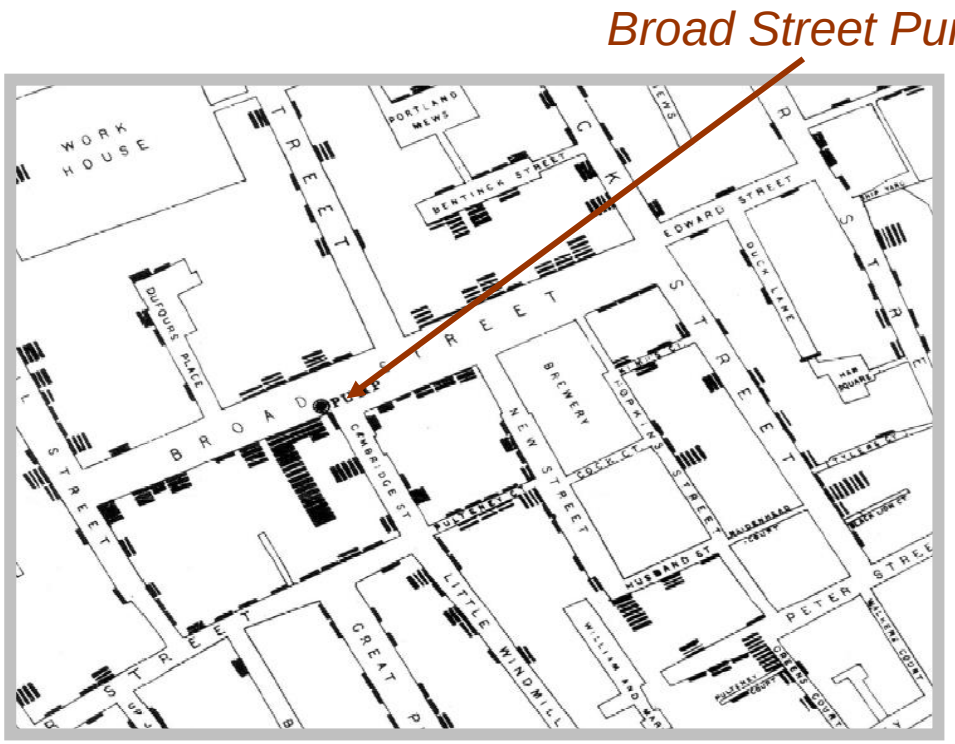
- "…to examine relationships between geographic features collectively and to use the relationships to describe the real-world phenomena that map features represent." (Clarke 2001, 182).

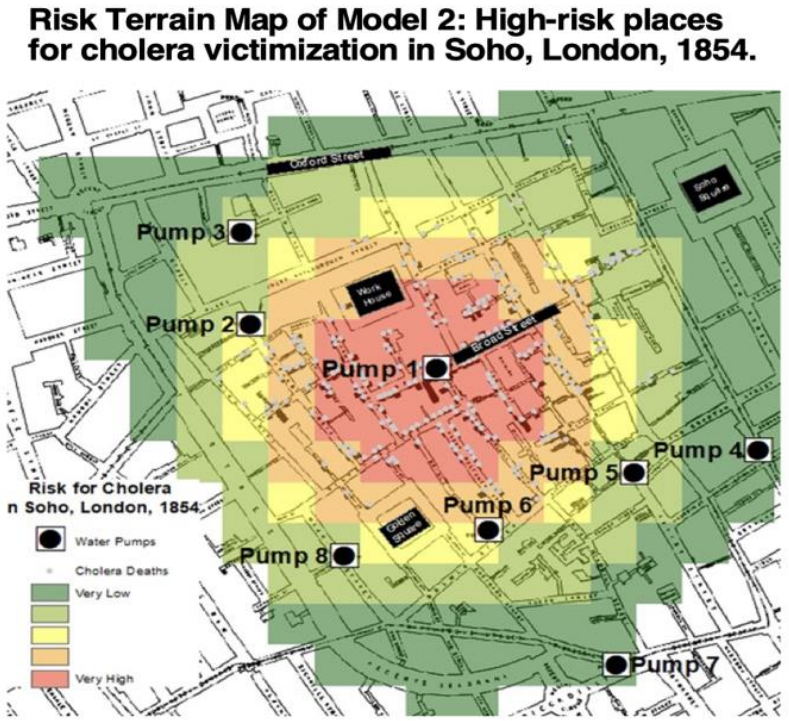
- A method of analysis is spatial if the results depend on the locations of the objects being analyzed
  - move the objects and the results change

- Spatial analysis requires both attributes and locations of objects
  - Spatial info to plot on map (e.g., lng/lat)
  - and the non-spatial data attributes (e.g., population, income, forests, animal habitats, etc.

Spatial Analysis in GIS

- The Snow Map of Cholera Incidence in the area of Broad Street, London, in 1854. The contaminated water pumps is located at the center of the map.





Source: Caplan JM, Kennedy LW, Neudecker CH (2020) **Cholera deaths in Soho, London, 1854: Risk Terrain Modeling for epidemiological investigations.** PLOS ONE 15(3): e0230725. https://doi.org/10.1371/journal.pone.0230725

Spatial analysis can be:

- **inductive**, to examine empirical evidence in the search for patterns that might support new theories or general principles.
- Four steps: **Observations, pattern discovery, analytics, and conclusion.**
  - e.g., the disease map
  - e.g., crime rates are higher in certain areas of a city

Spatial analysis can be – inductive

Why Some Immigrant Neighborhoods Are Safer than Others

FIGURE 1
Violent Crime and Immigrant Neighborhoods: Chicago

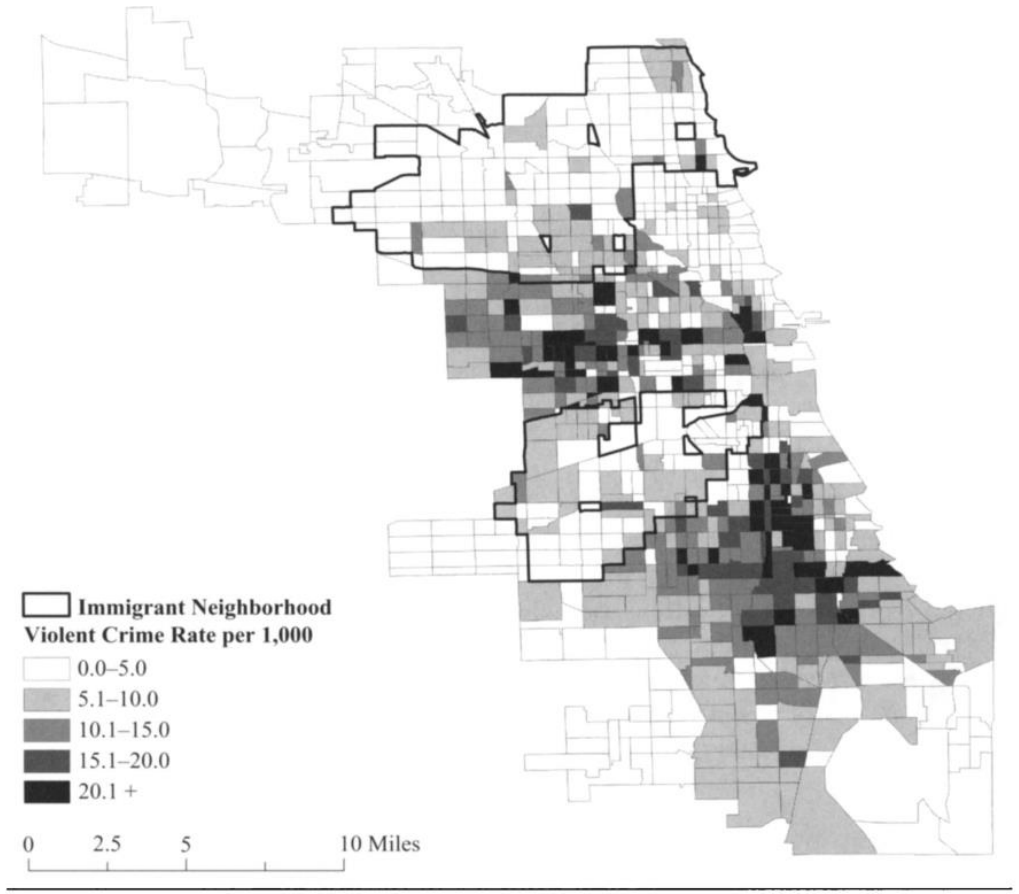
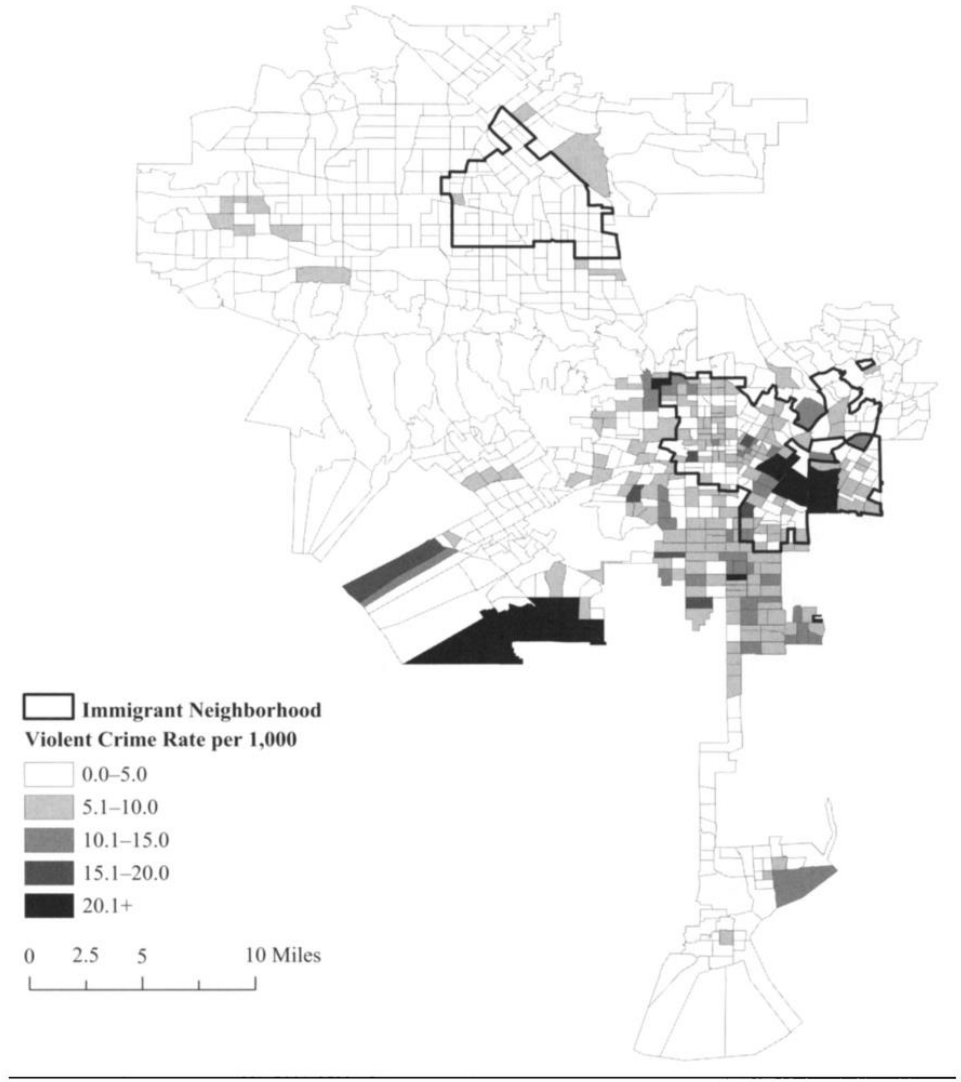


FIGURE 2
Violent Crime and Immigrant Neighborhoods: Los Angeles



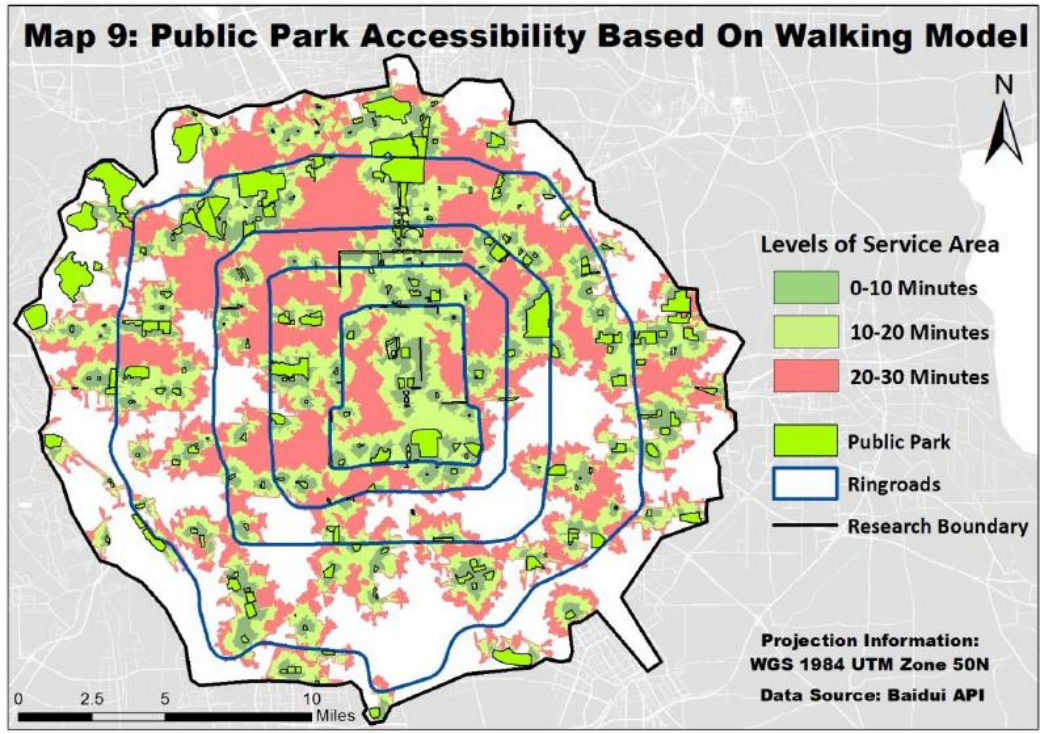
Source

Source: Wenzheng Li, CRP 5080, 2016

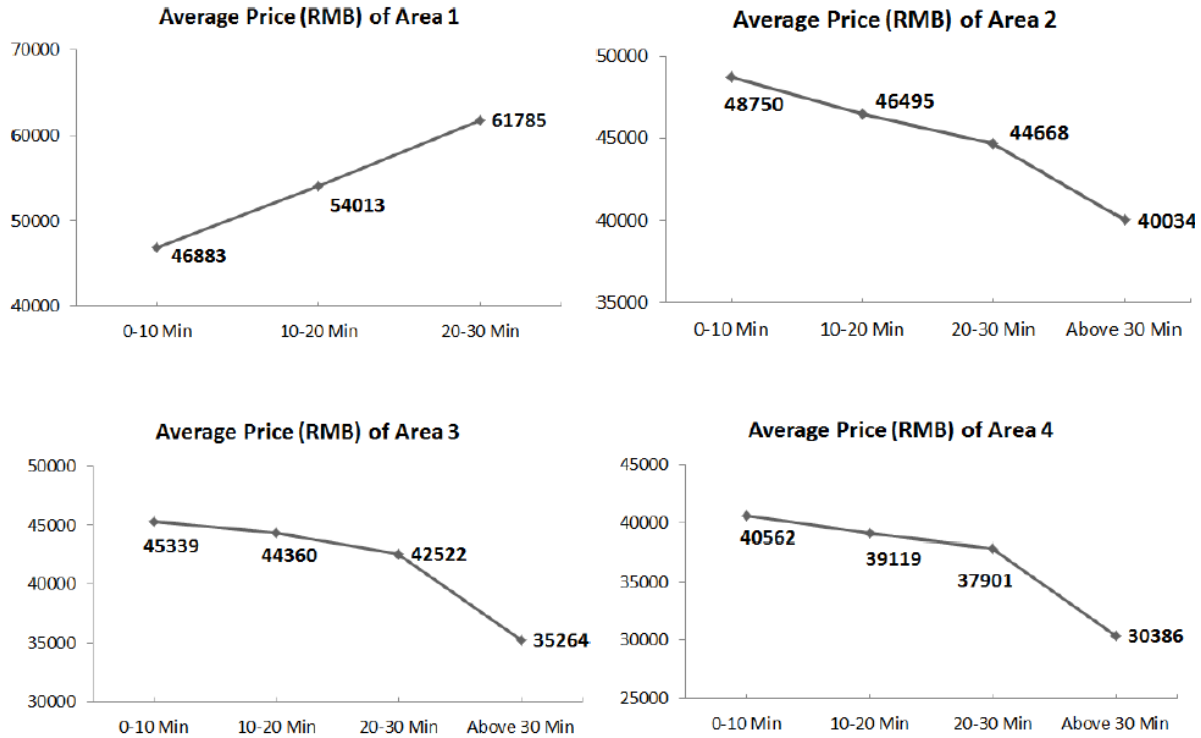
Spatial analysis can be:

- **deductive**, focusing on the testing of known theories or principles against data - starts with a hypothesis or theory, and then looks for evidence in the spatial data to support or refute it.

e.g., Hypothesis - Higher level of urban green spaces accessibility leads to higher property prices.



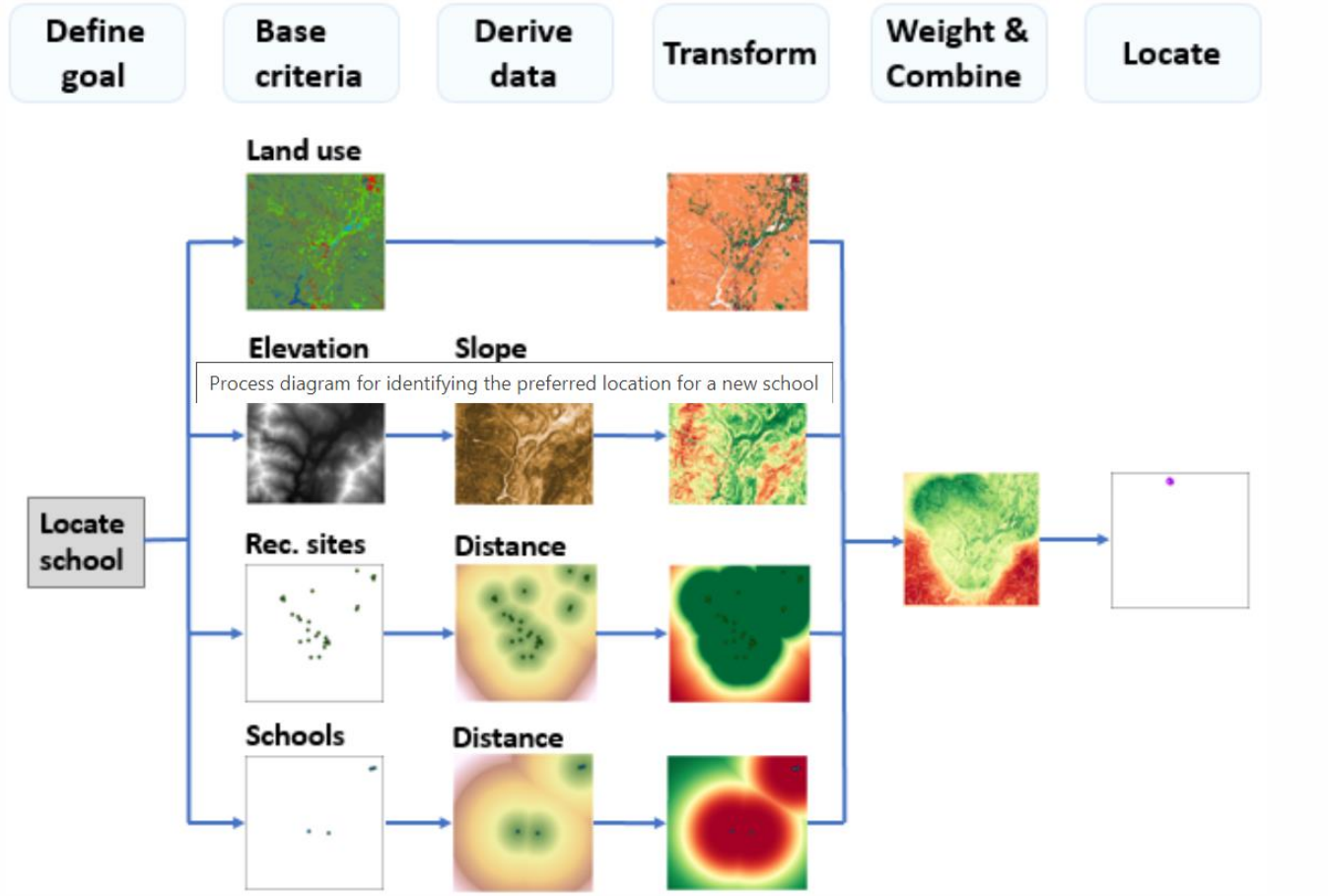
Public park accessibility in Beijing (Walking model)



The relationship between public park accessibility and housing price

Spatial analysis can be:

- **normative**, using spatial analysis to develop or prescribe new or better designs, patterns, etc.



Process diagram for identifying the preferred location for a new school.

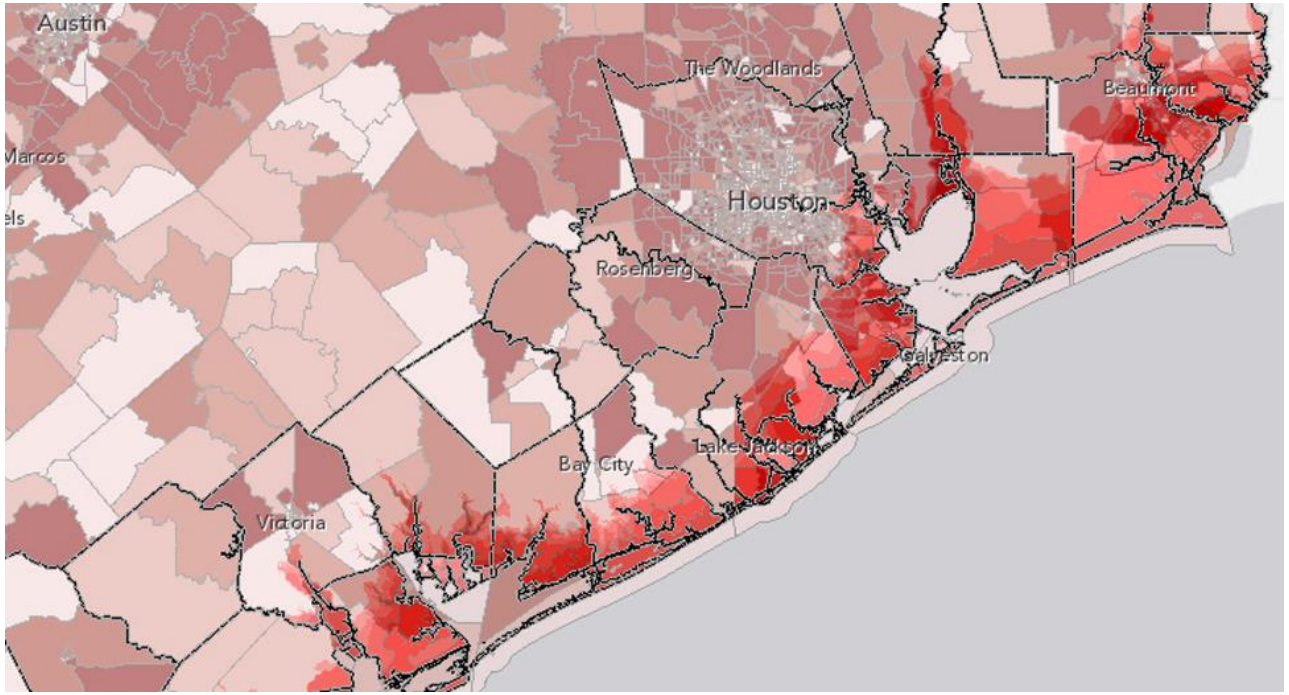
Public service Planning

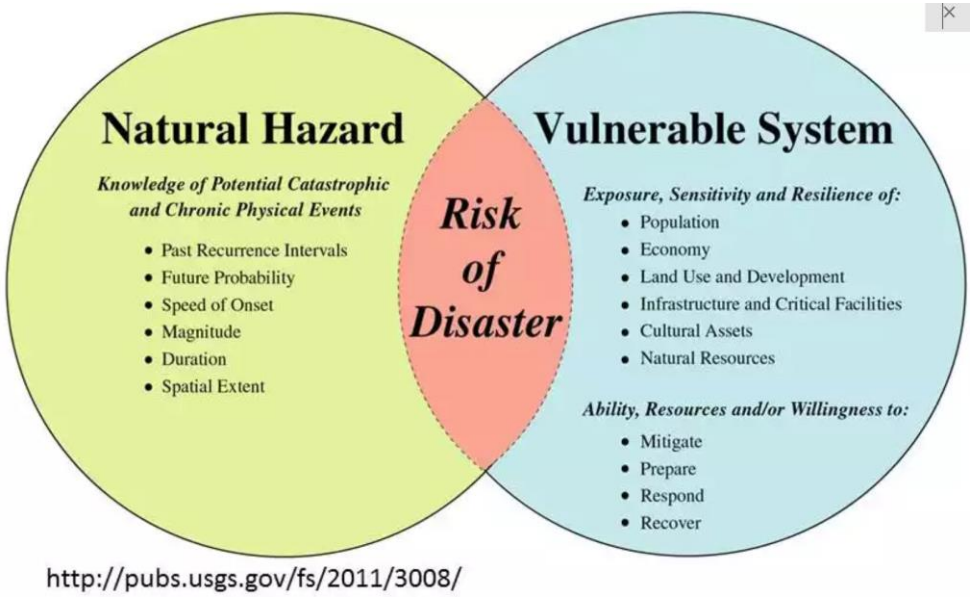
Land use suitability analysis

- e.g., Where should we locate an elementary school?
- Location-allocation modeling – know your demand (population?)
- Zoning/land use (permission to build?)
- Elevation (relatively flat)?
- Close to a recreation site

Spatial analysis can be - normative

Hazard Mitigation





http://pubs.usgs.gov/fs/2011/3008/

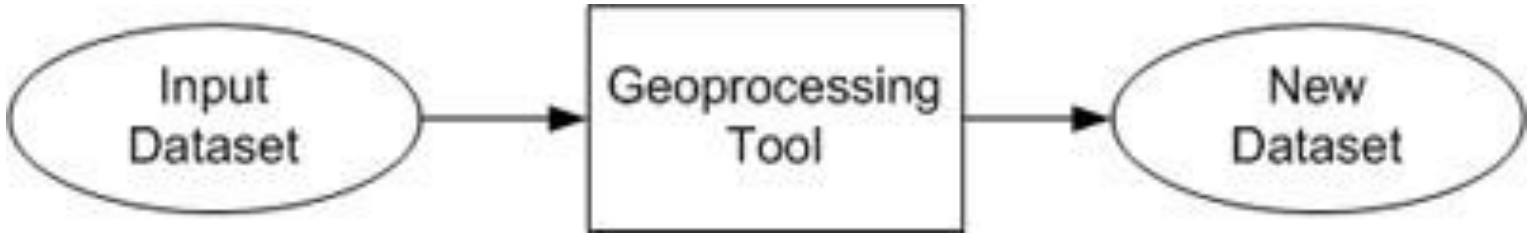
Example from Texas Sustainable and Resilient Planning Atlases: **Mapping the environmental hazards**

- Where are these hazards?
- Which neighborhoods face the most hazard risks?
- Where are mitigation mostly needed
- Where should future growth occur

Image source

Geoprocessing

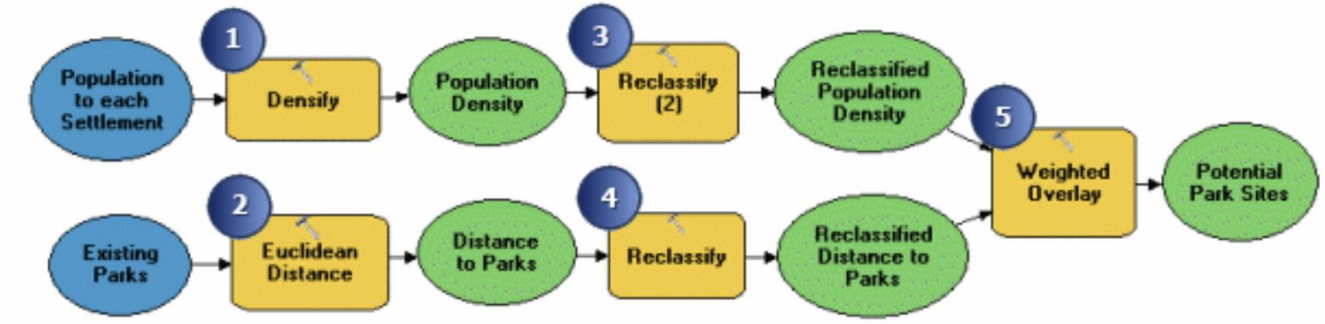
- Geoprocessing is the manipulation of geographic data.
- It provides a way to create new information by applying an operation to existing data.
- A typical geoprocessing operation takes an input dataset, performs an operation on that dataset, and returns the result of the operation as an output dataset or derived data

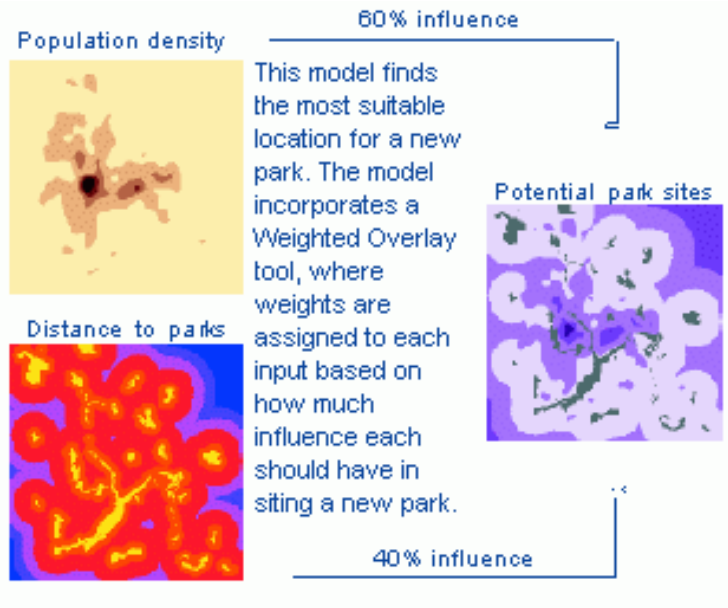


Geoprocessing

It can be done:

- 1) Interactively (through Arc toolbox)
- 2) Using a model builder (ex. Suitability Analysis) – clear to show the workflow of multiple geoprocessing.





- 3) Python – good for geoprocessing automation (batch analysis)

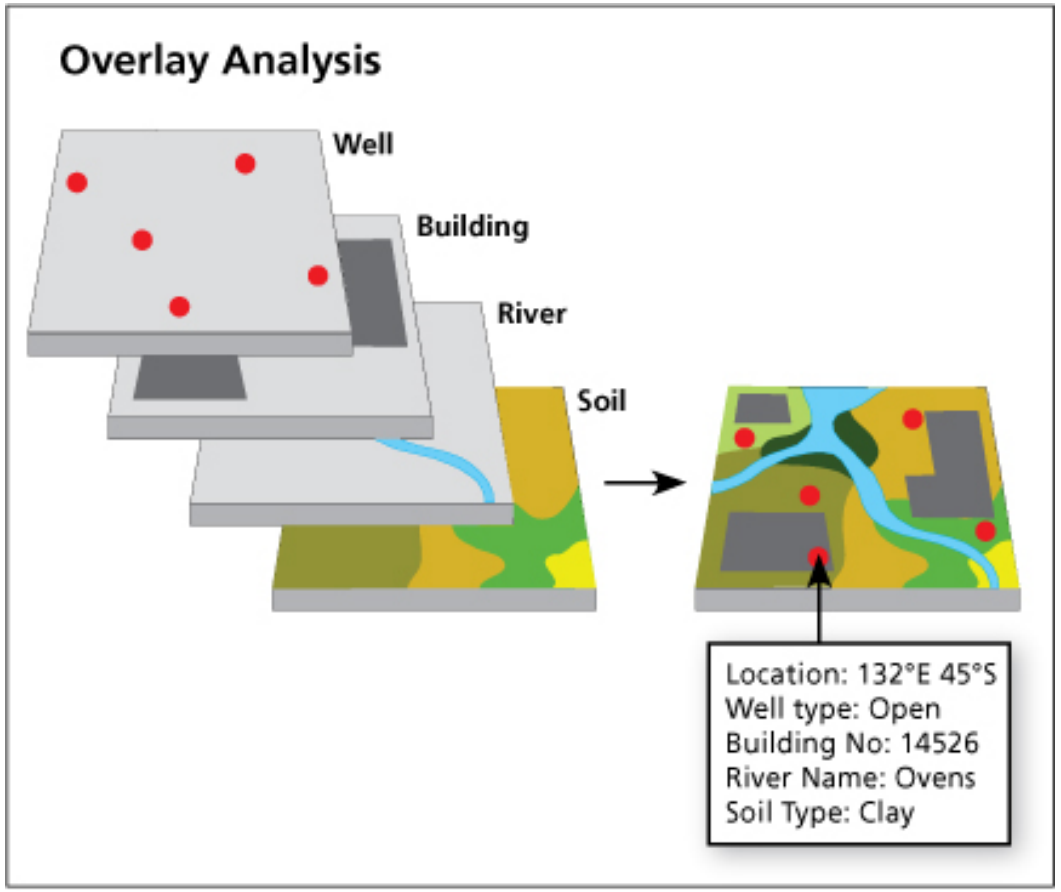
Geoprocessing Operations We learn today

- Overlay toolset
  - clip
  - Union
  - Intersect
  - Spatial join
- Dissolve (spatial data reclassification)
- Merge
- Proximity toolset
  - Buffer

Overlay analysis

Combining two layers to create a new output feature class containing *information from both of the inputs*.

- used for site selection and site suitability analysis
- Create new layer by combining features and attributes of two other layers
- Various methods - differ only in spatial extent of output
- Usually easier using raster data – more on this later



Overlay analysis

One of the most basic questions asked of a GIS is: "What's on top of what?"

- *Point-in-polygon:* What is the total employment of large employers (points) in Ithaca (polygon)? How many bus stops (points) are there within our study area (polygon)? - “Within” is just another way of saying “on top of.”

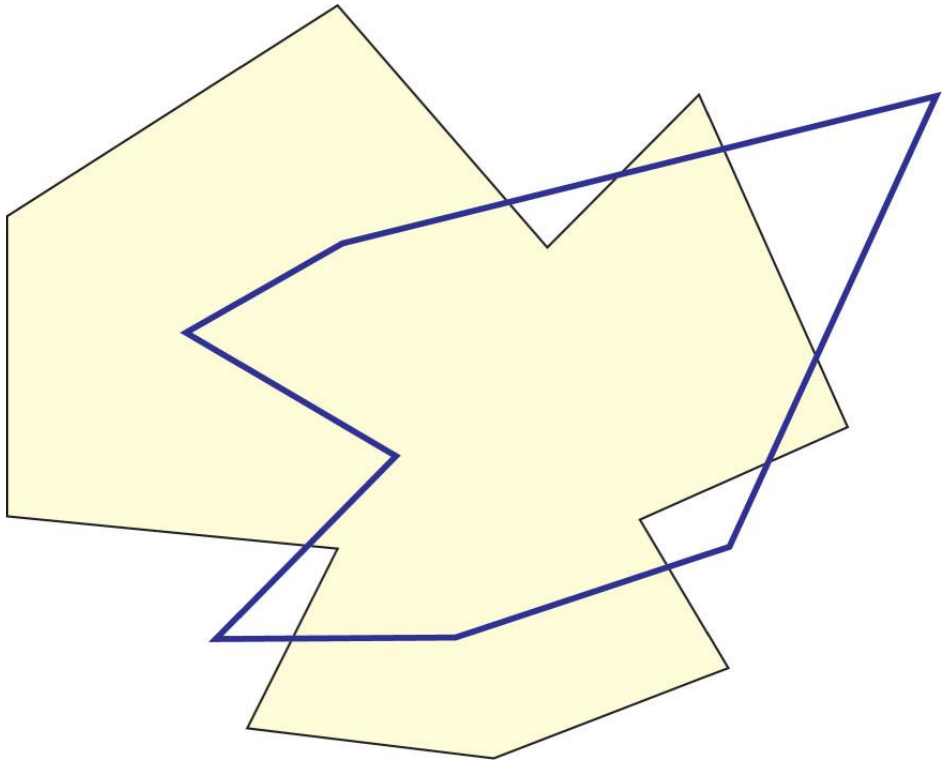
- *Line-in-polygon :* How many miles of bus routes (line) are within Tompkins County (polygon)?

- *Polygon-in-polygon:* Retrieve data for the census block groups (polygons) within the Central Business District (polygon).

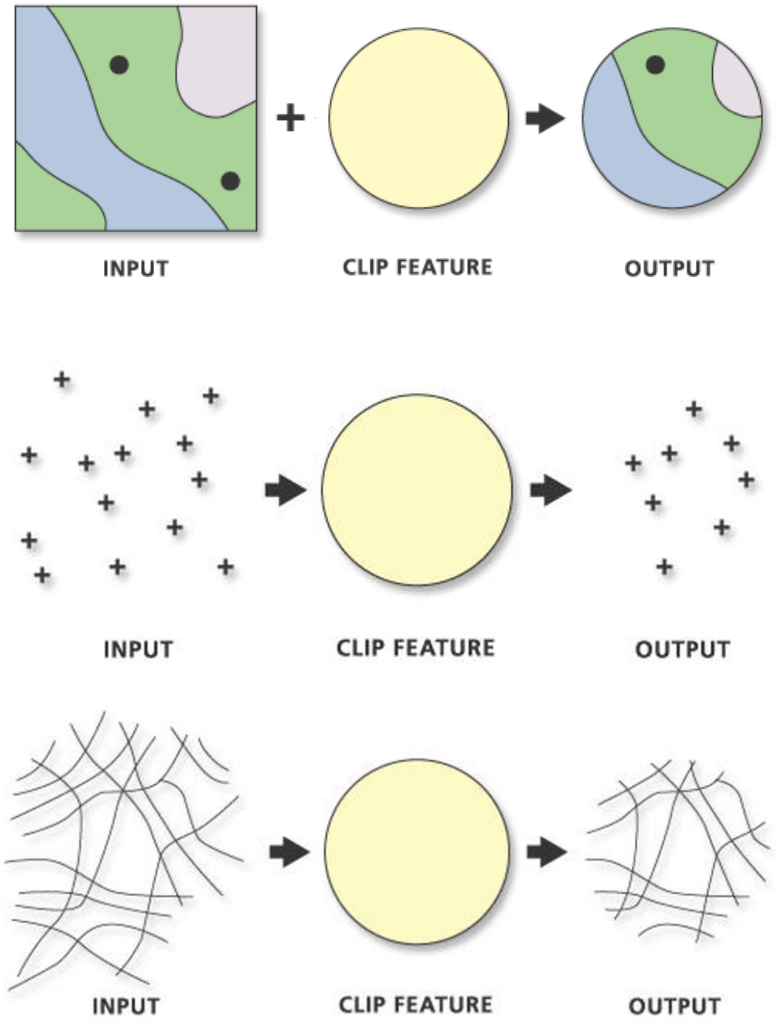
Polygon-in-polygon

Here the overlay of two polygons produces nine distinct polygons.

- One has the properties of both polygons,
- four have the properties of the yellow shaded polygon but not the blue (bounded) polygon.
- four are outside the yellow polygon but inside the blue polygon.



Overlay: Clip



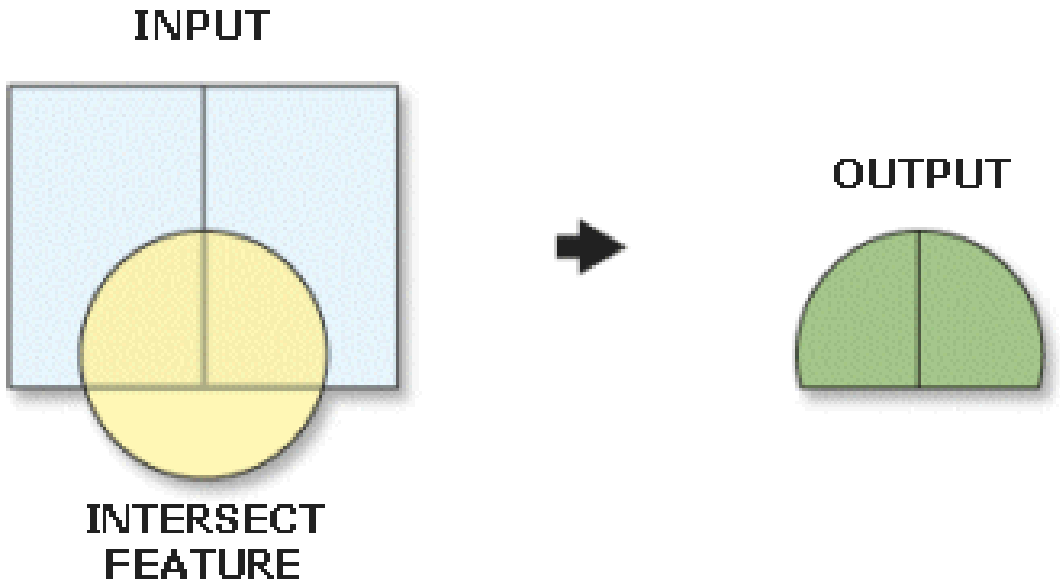
Extracts input features that overlay the clip features. Use this tool to cut out a piece of one dataset using one or more of the features in another dataset **as a cookie cutter**.

Spatial extent of 'clip' layer, attributes of original input layer

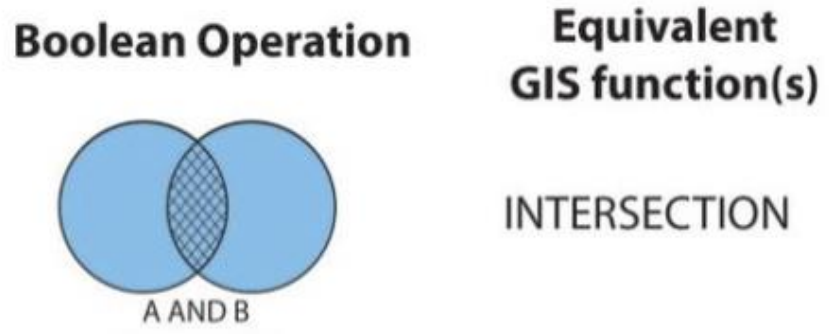
- Not combining attributes of inputs
- Input features retain their attributes.
- The attributes of boundary ("clip") layers are ignored.

Overlay: Intersect

Computes a geometric intersection of the input features. Features or portions of features that overlap in all layers or feature classes will be written to the output feature class.



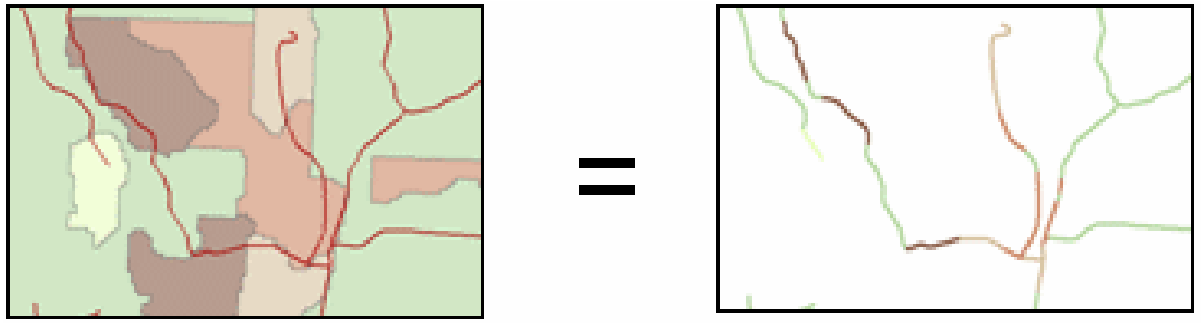
Contains attributes of both input layers, spatial extent only where they intersect



Environmental Analysis

Input:
Roads, vegetative cover

Output: Intersection of roads and vegetative cover, attributes of both



FID	Shape*	LOCALID	RD_TYPE
6	Polyline	222192	1500
7	Polyline	220893	1500

+

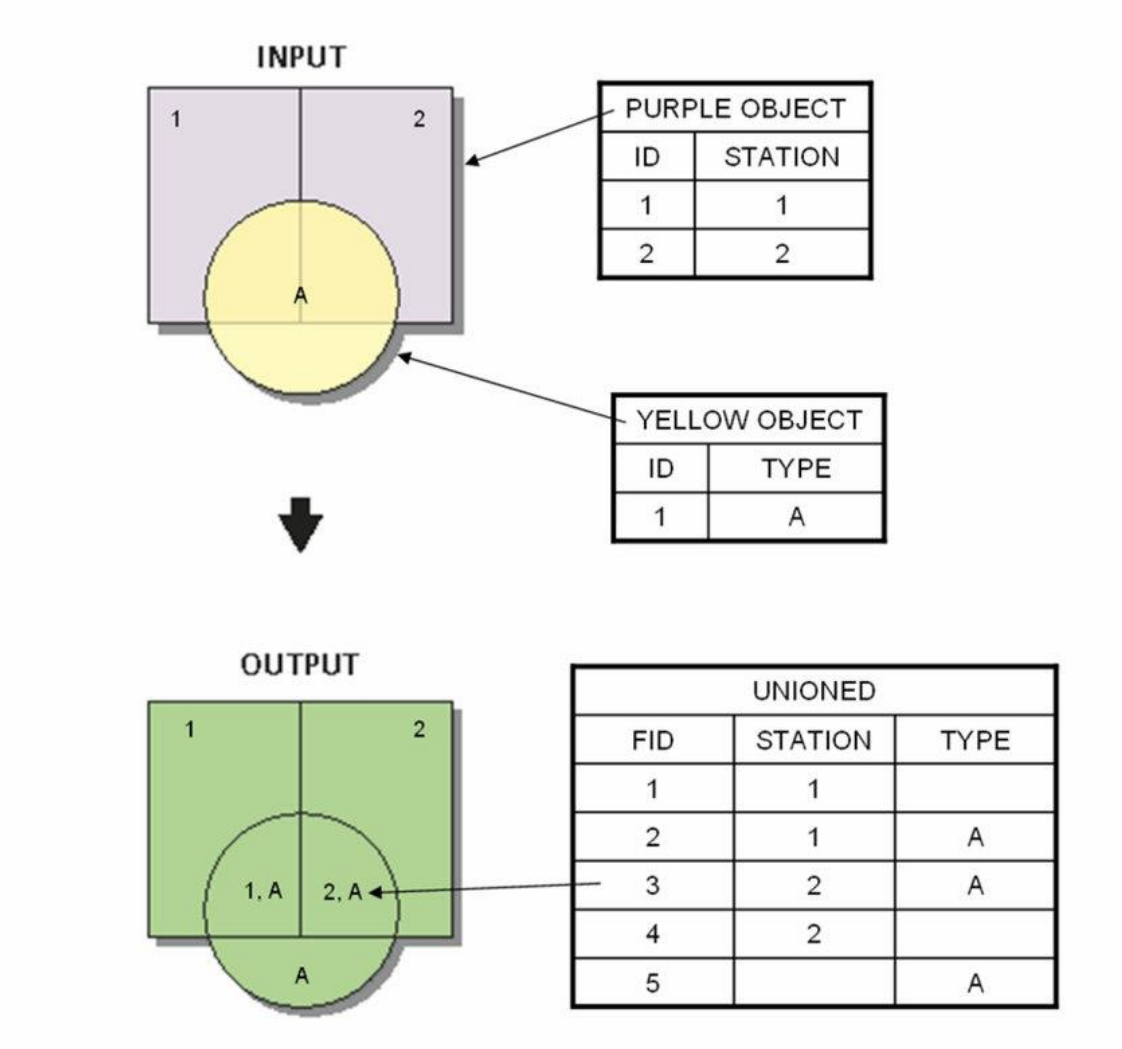
FID	Shape*	VEG_TYPE
6	Polygon	CL
7	Polygon	SD
8	Polygon	SS

FID	Shape*	LOCALID	RD_TYPE	VEG_TYPE
6	Polyline	219329	1500	FC
7	Polyline	219384	1500	FC
8	Polyline	219380	1500	FC
9	Polyline	219380	1500	SD
10	Polyline	224631	1500	FC

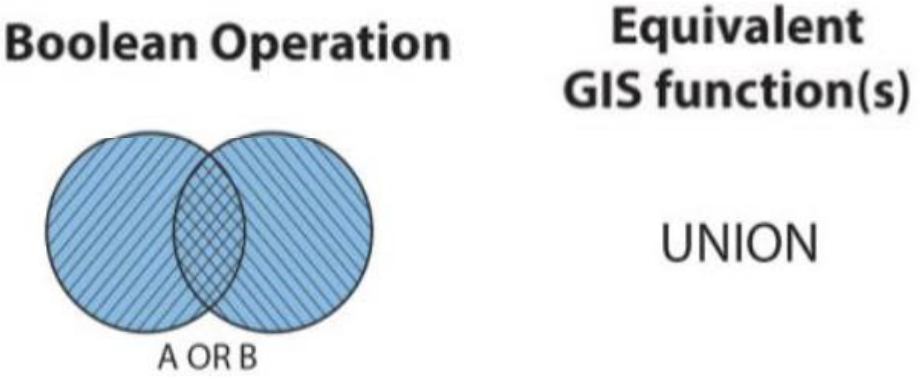
Intersect: Contains attributes of both input layers, but the spatial extent only where the input layers intersect

Intersect: logging roads (lines) and vegetation types (polygons) are overlaid to create a new line layer

Overlay: union



Contains attributes of both input layers, spatial extent of both input layers



Overlay: spatial join

- Joins attributes from one feature to another based on the spatial relationship.
- A spatial join matches rows from the **Join Features** values to the **Target Features** values based on their relative spatial locations.
- A new layer (output) is created

- **Target Layer:** The layer containing the features you want to enrich with additional attribute data
- **Join Layer:** The layer providing additional attribute data for the join.

Target Feature

Join Feature

Shape	TRI_ID	ZIP_Code
Point	1	10461
Point	2	10457
Point	3	10474
Point	4	10461
Point	5	10451
Point	6	10460
Point	7	10454

Overlay: spatial join

- Spatial Join Relation: **One to One** and One to Many
- **One-to-one operation**: If multiple **join features** are found that have the same spatial relationship with a single **target feature**, the attributes from the multiple join features will be aggregated using a field map merge rule (e.g., count, mean, first/last, sum, etc).

Overlay: spatial join

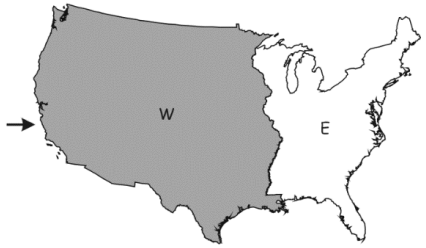
- Spatial Join Relation: One to One and **One to Many**
- One-to-many operation: If multiple **join features** are found that have the same spatial relationship with a single **target feature**, the output feature class will contain multiple copies (records) of the target feature.

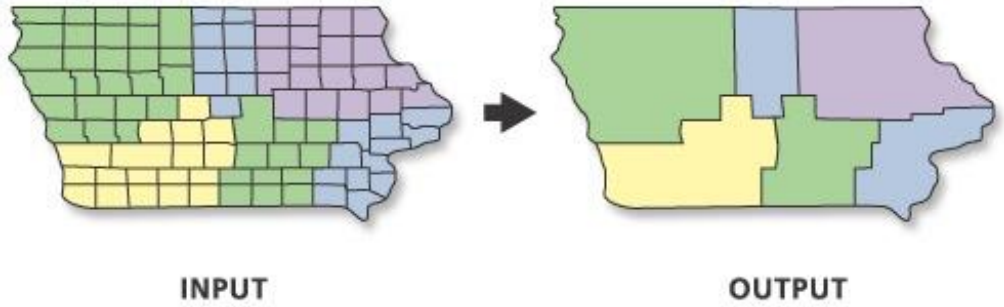
Dissolve

- Group features together based on attributes

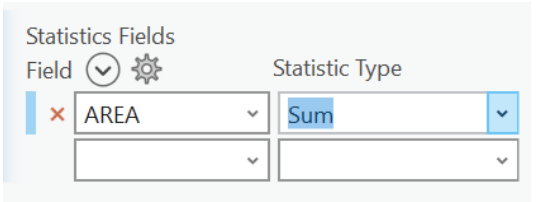
Dissolve table

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





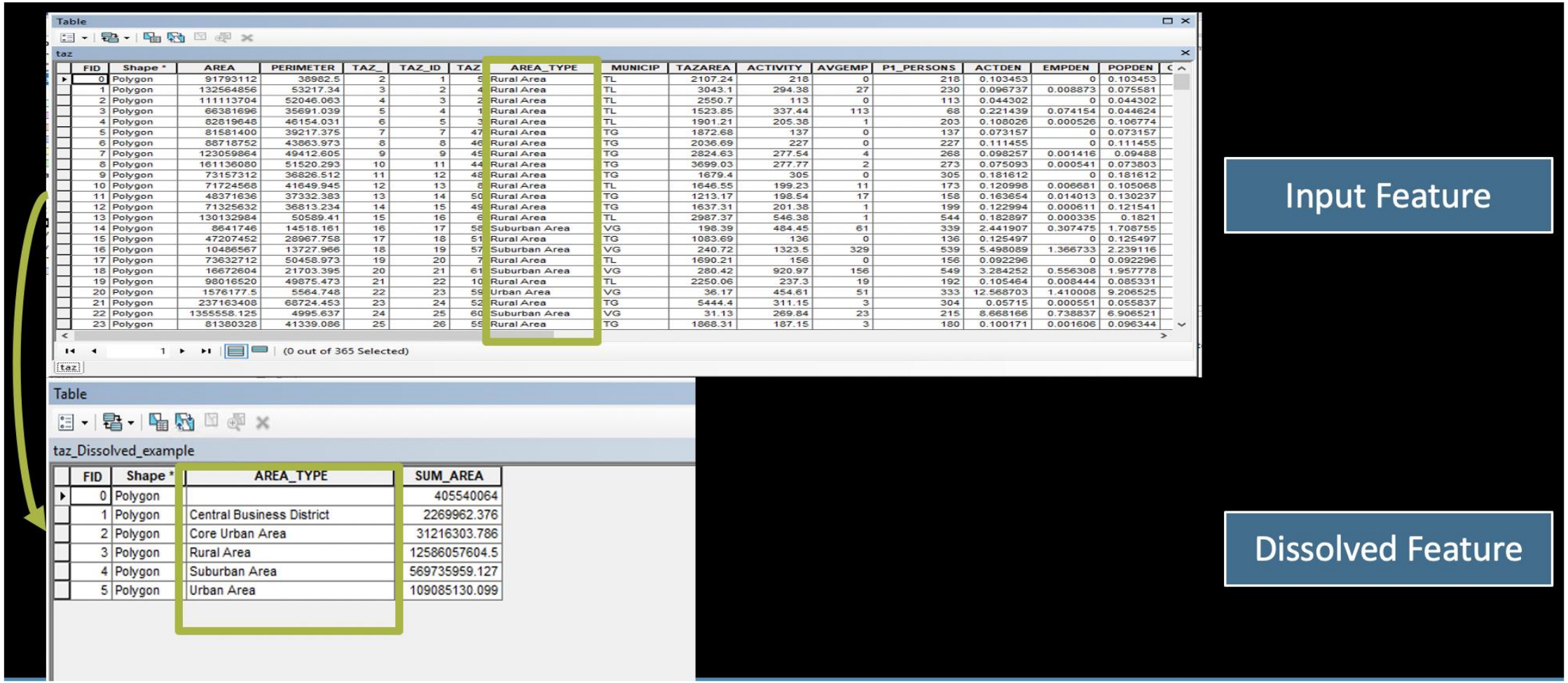
- The attribute table of the dissolved feature class
  - The field on which dissolve is based.
  - Additional fields to summarize information in the original features



Available statistics types are as follows:

- **Sum**—The values for the specified field will be added together.
- **Mean**—The average for the specified field will be calculated.
- **Minimum**—The smallest value for all records of the specified field will be identified.
- **Maximum**—The largest value for all records of the specified field will be identified.
- **Range**—The range of values (maximum minus minimum) for the specified field will be calculated.
- **Standard deviation**—The standard deviation of values for the specified field will be calculated.

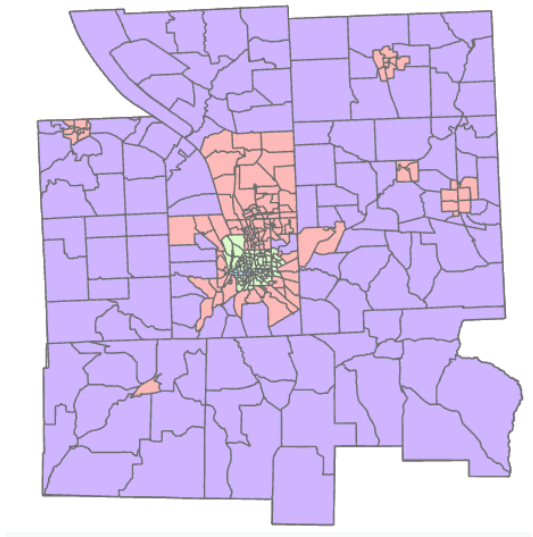
Dissolve – inside the attribute table

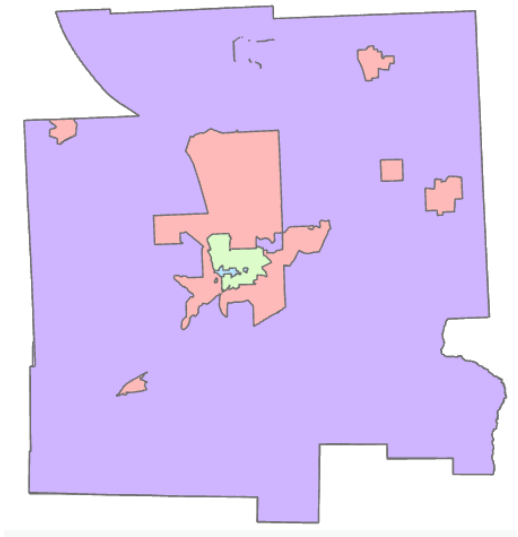


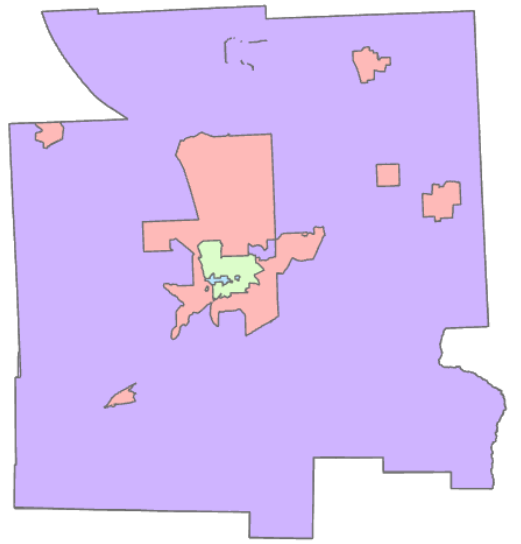
Dissolve

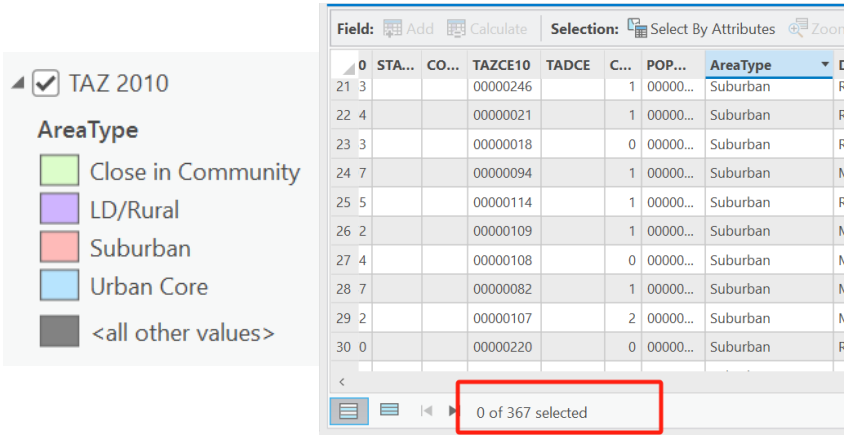
Specifies whether multipart features will be allowed in the output feature class.

- Checked—Multipart features will be allowed in the output feature class. This is the default.
- Unchecked—Multipart features will not be allowed in the output feature class. Individual features will be created for each part.







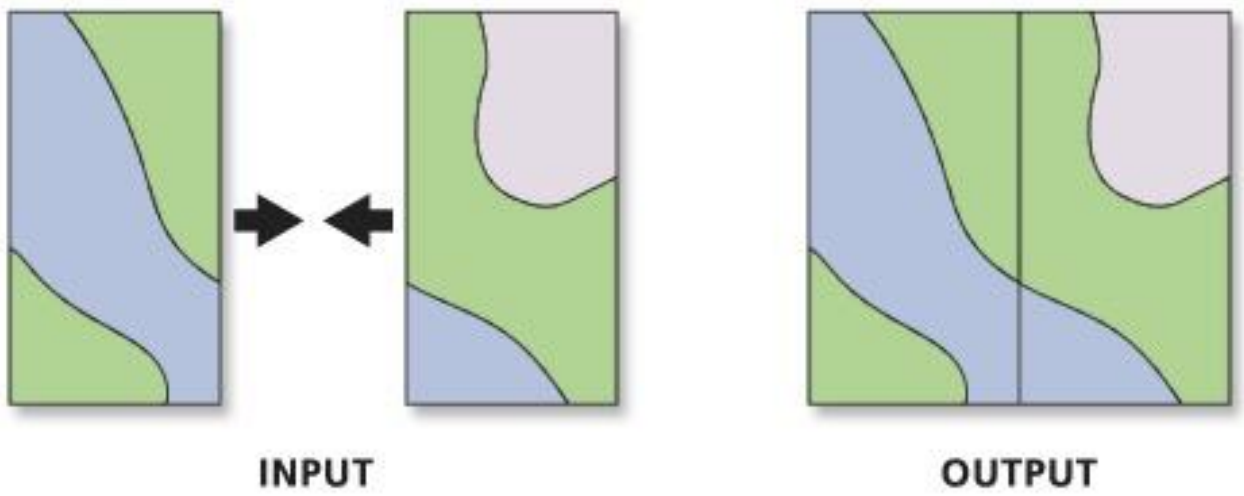


	FID	Shape *	AreaType	SUM_AREA
1	0	Polygon	Close in Community	131827615.82
2	1	Polygon	LD/Rural	12331441017.5
3	2	Polygon	Suburban	1209674247.11
4	3	Polygon	Urban Core	9925818.49
	Click to add new row.			

	FID *	Shape *	AreaType	Shape_Length	Shape_Area
4	4	Polygon	LD/Rural	999893.259787	12325372447.421152
5	5	Polygon	Suburban	28537.7278	18977190.881552
6	6	Polygon	Suburban	40403.731552	68570186.037054
7	7	Polygon	Suburban	22053.659887	30540406.38577
8	8	Polygon	Suburban	343883.550147	1012373011.282022
9	9	Polygon	Suburban	26395.186417	32564702.376664
10	10	Polygon	Suburban	36266.312779	48458406.550312
11	11	Polygon	Urban Core	4909.138892	977364.425255
12	12	Polygon	Urban Core	22516.589831	8963433.09672

Merge

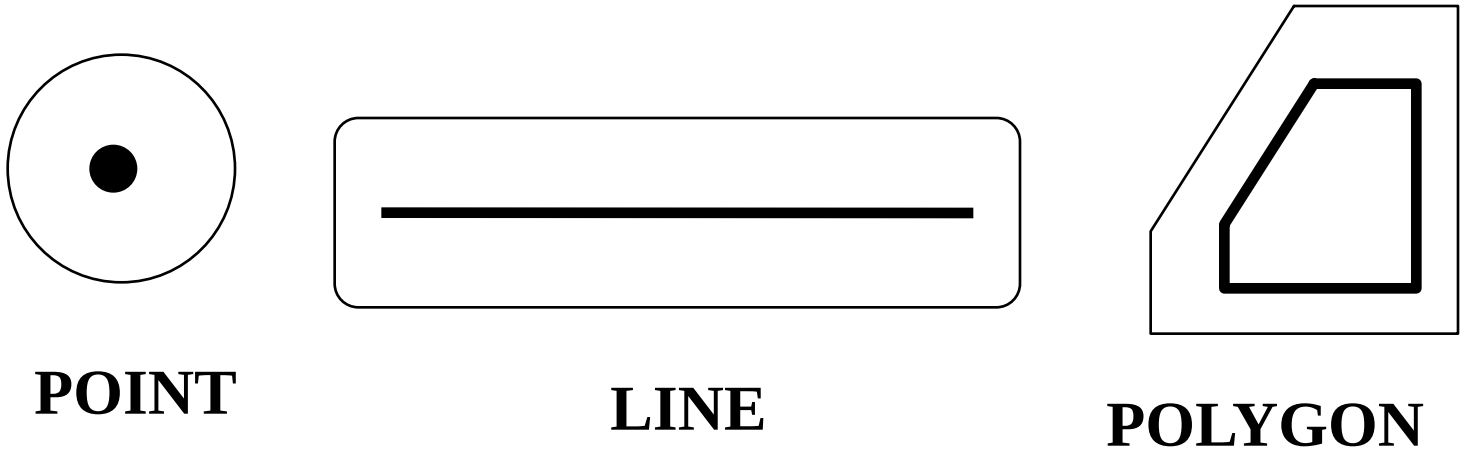
Merge is used when you have two or more geographically **adjacent** layers, each containing the same type of features, and you want one layer that contains all their features.



Proximity Analysis: Buffers

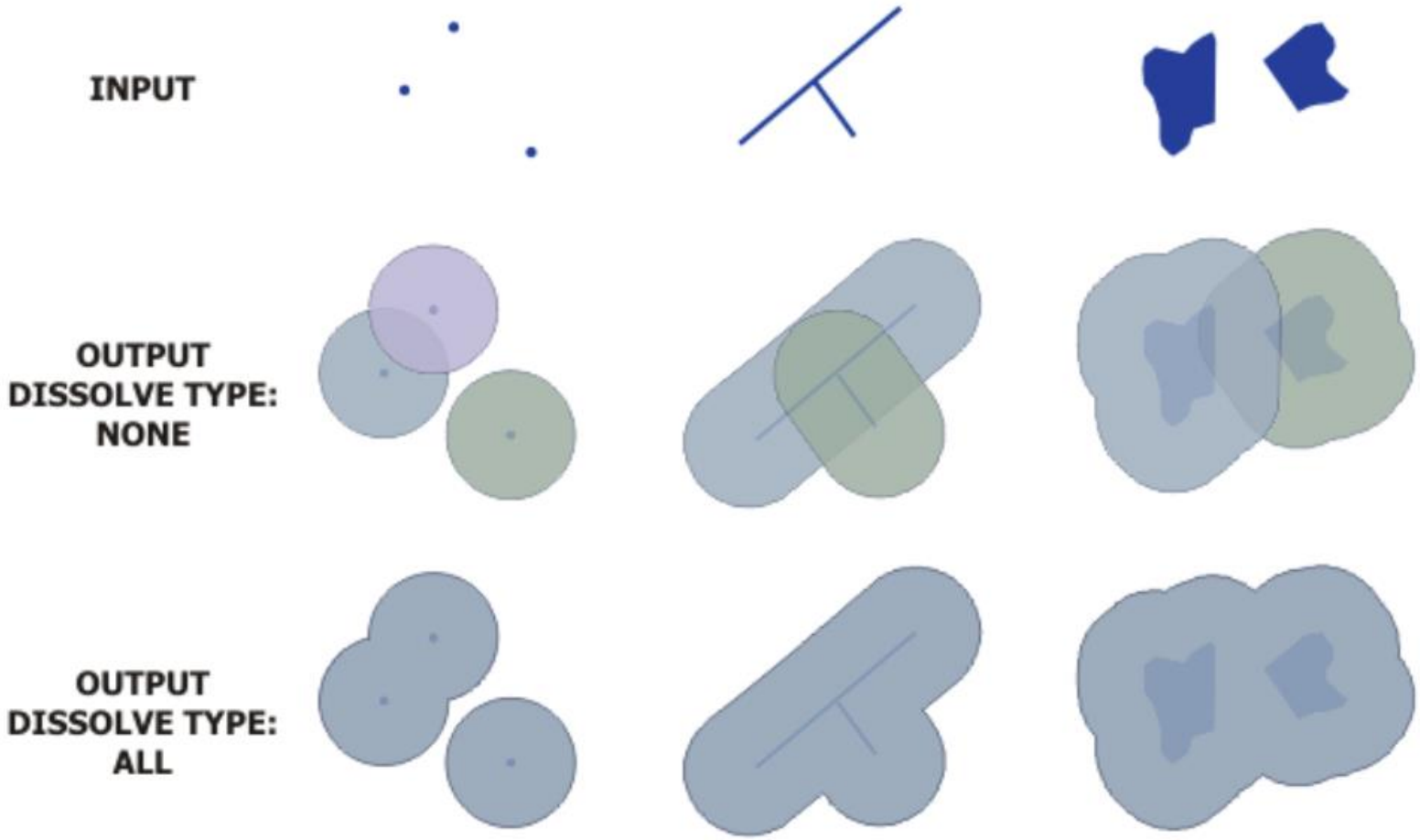
Buffering generates new polygons around existing points, lines, or polygons

- Identify all children within a 1 mile radius of a school (buffer around a point).
- Identify all large employers (points) within a 1/4 mile distance of a bus route (buffer around a line).
- Identify protected zones within 500 meters around a lake (buffer around lines)



Proximity Analysis: Buffers

Creates buffer polygons around input features to a specified distance.



References:

- Spatial analysis in ArcGIS Pro
- An overview of the Analysis toolbox
- Maantay, J., & Ziegler, J. (2006). GIS for the Urban Environment. Redlands, CA: Esri Press.