Image from NASA:
https://commons.wikimedia.org/wiki/File:Earth_Western_Hemisphere_transparent_background.png#filelinks

GIScience

Vector Analysis

WV VIEW

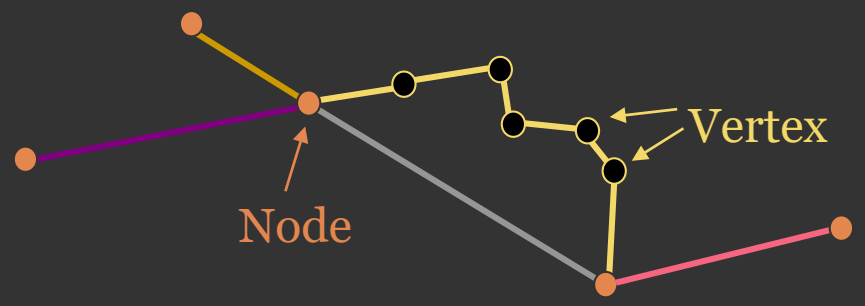
AmericaView℠
Empowering Earth Observation Education
AmericaView.org - Est. 2003

General Vector-Based Spatial Analysis

Spatial Analysis in ArcGIS

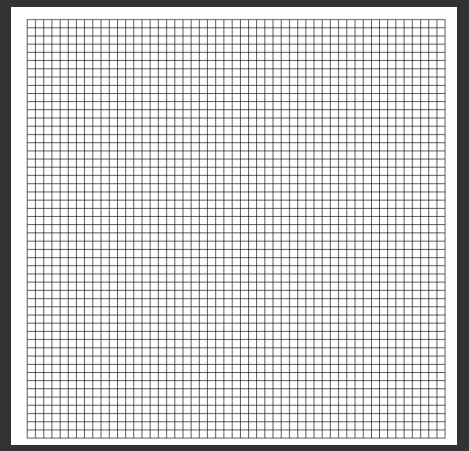
Analysis Toolbox for Vector Spatial Analysis

- Analysis Tools
  - Extract
  - Overlay
  - Pairwise Overlay
  - Proximity
  - Statistics



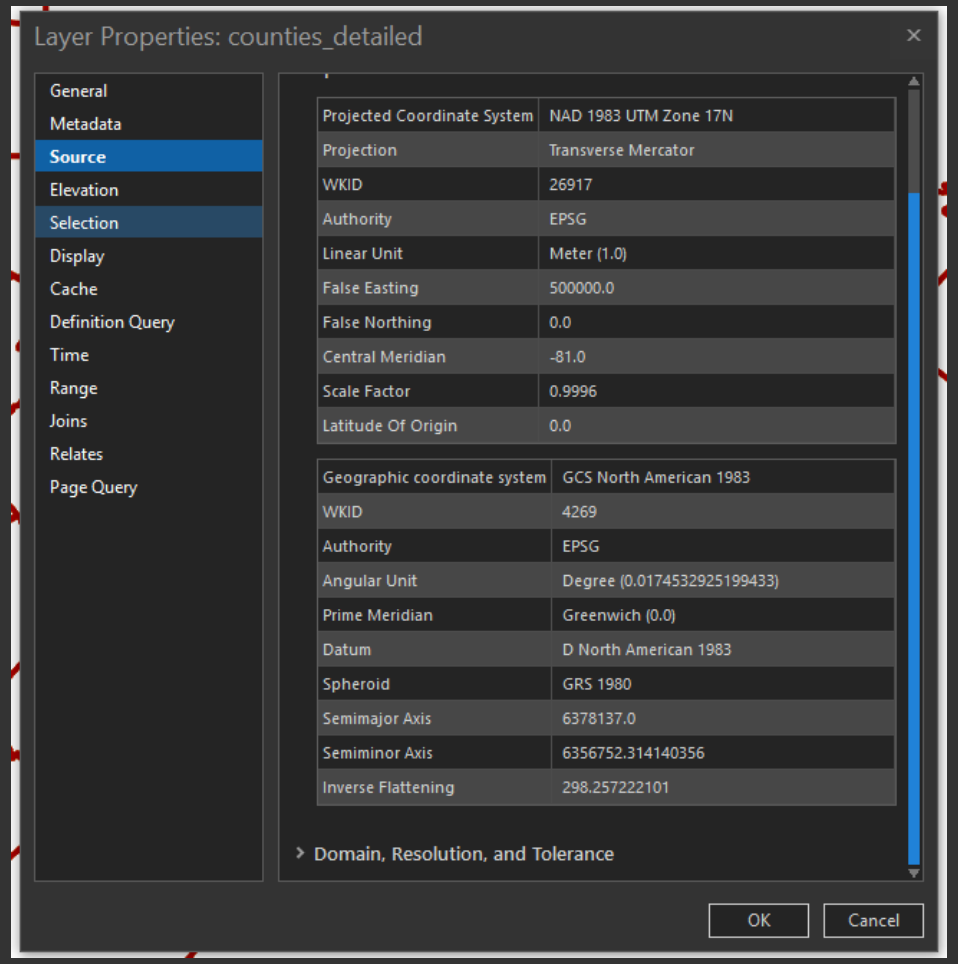
Spatial Analyst Toolbox for Raster Spatial Analysis

- Spatial Analyst Tools
  - Conditional
  - Density
  - Distance
  - Extraction
  - Generalization
  - Groundwater
  - Hydrology
  - Interpolation
  - Local
  - Map Algebra
  - Math
  - Multivariate
  - Neighborhood
  - Overlay
  - Raster Creation
  - Reclass
  - Segmentation and Classification
  - Solar Radiation
  - Surface
  - Zonal



A Note on Spatial Reference and Projection

- When undertaking spatial analysis, it is best if your data are all in the same spatial reference or projection
- This allows for overlay



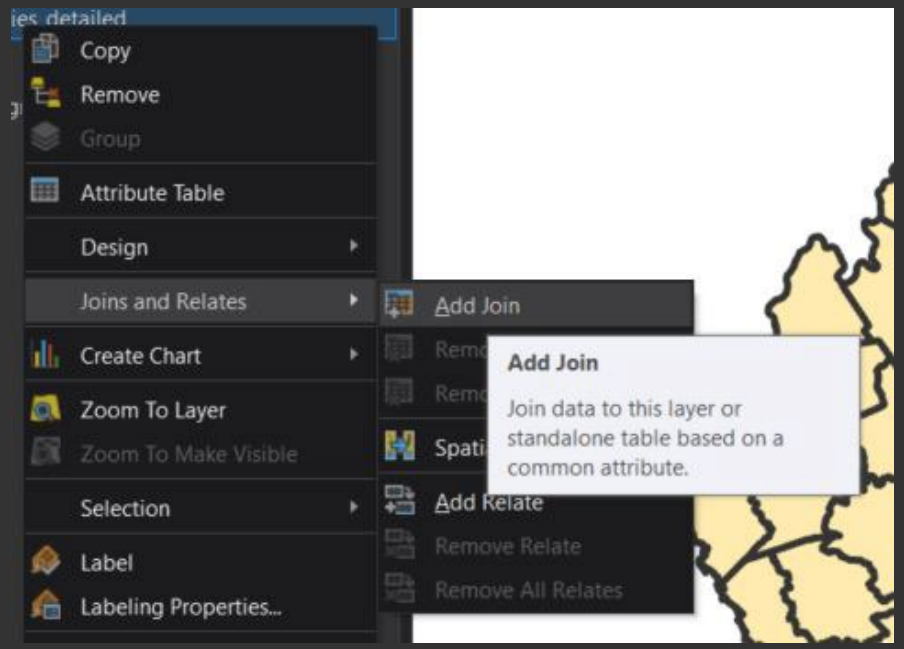
Joining non-spatial data and spatial data

Join

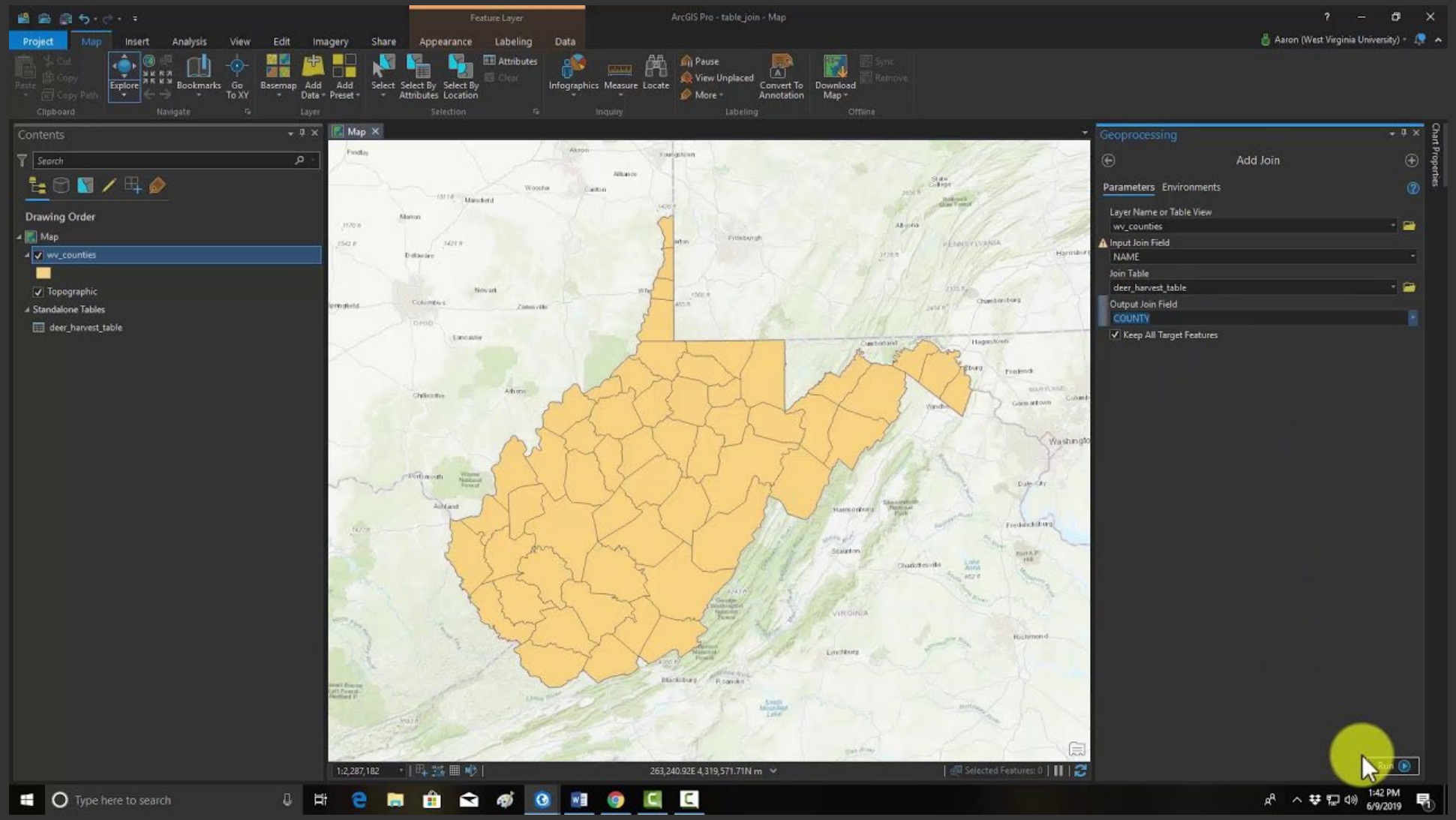
- Non-spatial data can be joined to spatial data as long as there is a common field between them
- Later, the joined fields can be used for analysis or to symbolize the features

Relate

- New fields are not added to table. However, a relationship is established

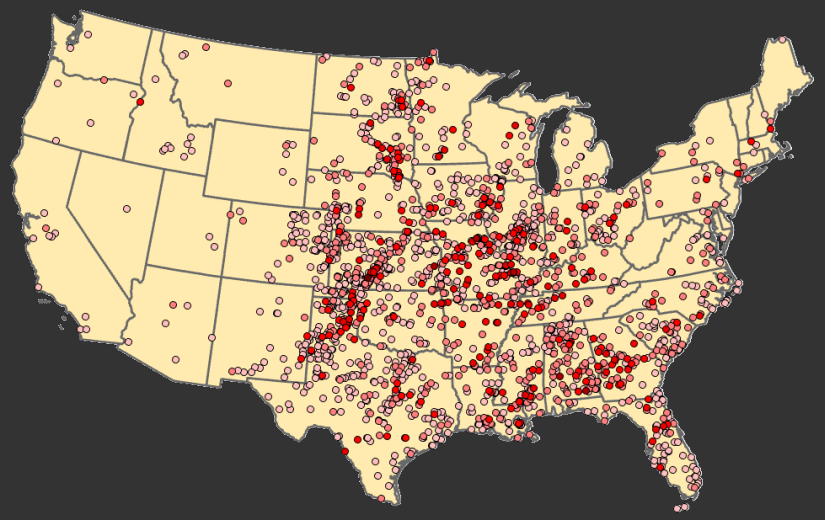


Video: Table Join

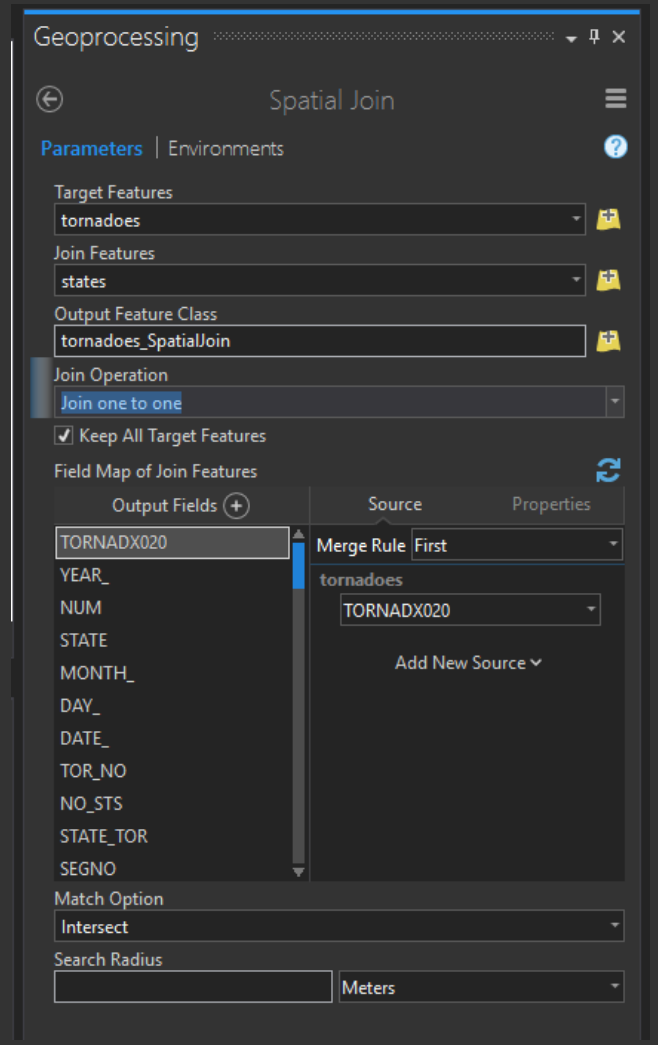


Spatial Joins

- Joins attributes based on spatial relationships



DAMAGE	F_SCALE	FID_	AREA	STATE_NAME	STATE_FIPS	SUB_REGION	STATE_ABBR
0	0	0	264435.873	Texas	48	W S Cen	TX
0	0	0	264435.873	Texas	48	W S Cen	TX
0	2	0	264435.873	Texas	48	W S Cen	TX
0	0	0	264435.873	Texas	48	W S Cen	TX
0	0	0	264435.873	Texas	48	W S Cen	TX
1	0	0	264435.873	Texas	48	W S Cen	TX
0	0	0	264435.873	Texas	48	W S Cen	TX
0	0	0	264435.873	Texas	48	W S Cen	TX
0	0	0	264435.873	Texas	48	W S Cen	TX
0	0	0	264435.873	Texas	48	W S Cen	TX
0	0	0	264435.873	Texas	48	W S Cen	TX
0	0	0	264435.873	Texas	48	W S Cen	TX
0	1	0	84520.49	Minnesota	27	W N Cen	MN
0	0	0	84520.49	Minnesota	27	W N Cen	MN



Video: Spatial Join

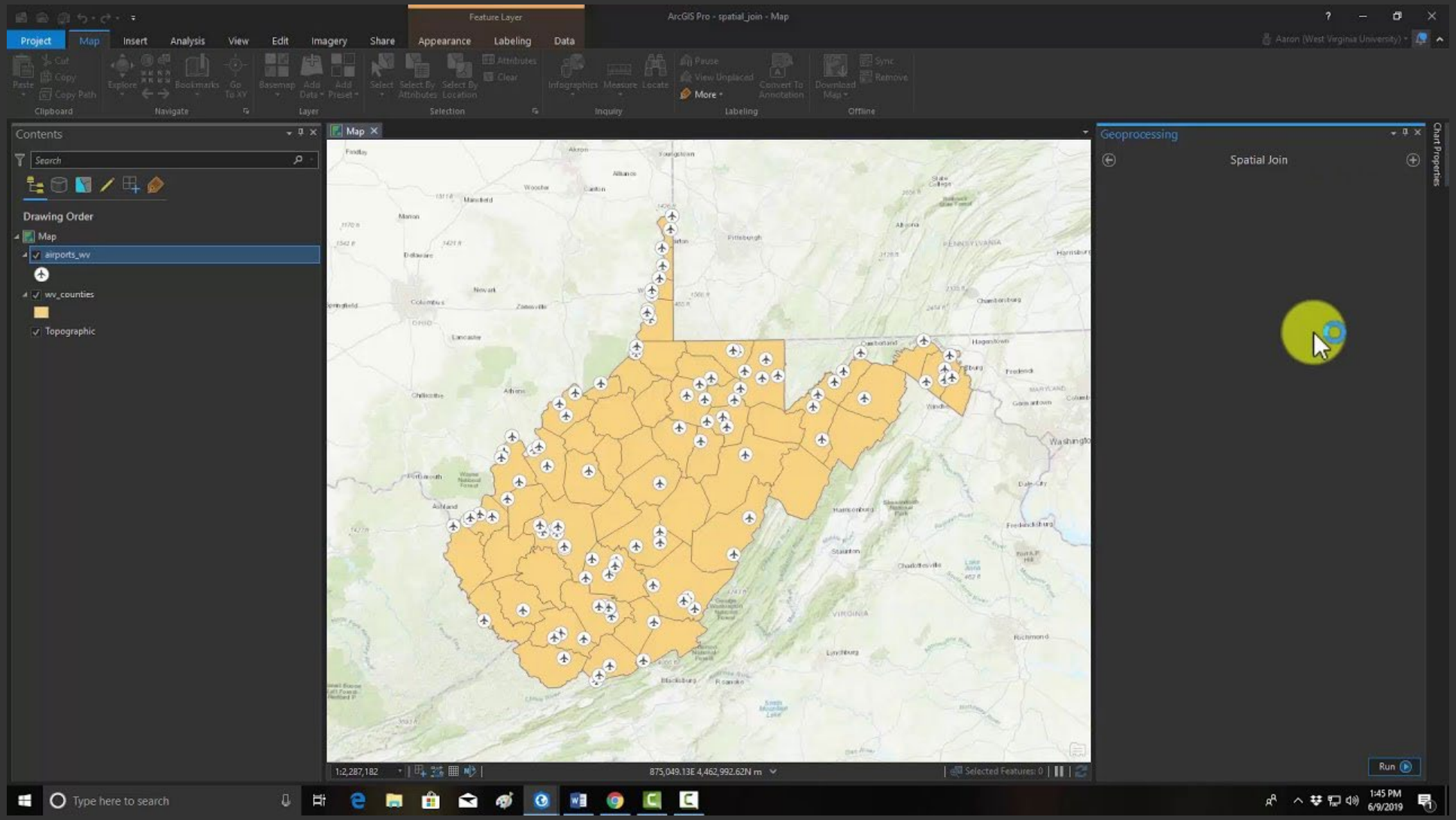


Table or Aspatial Queries

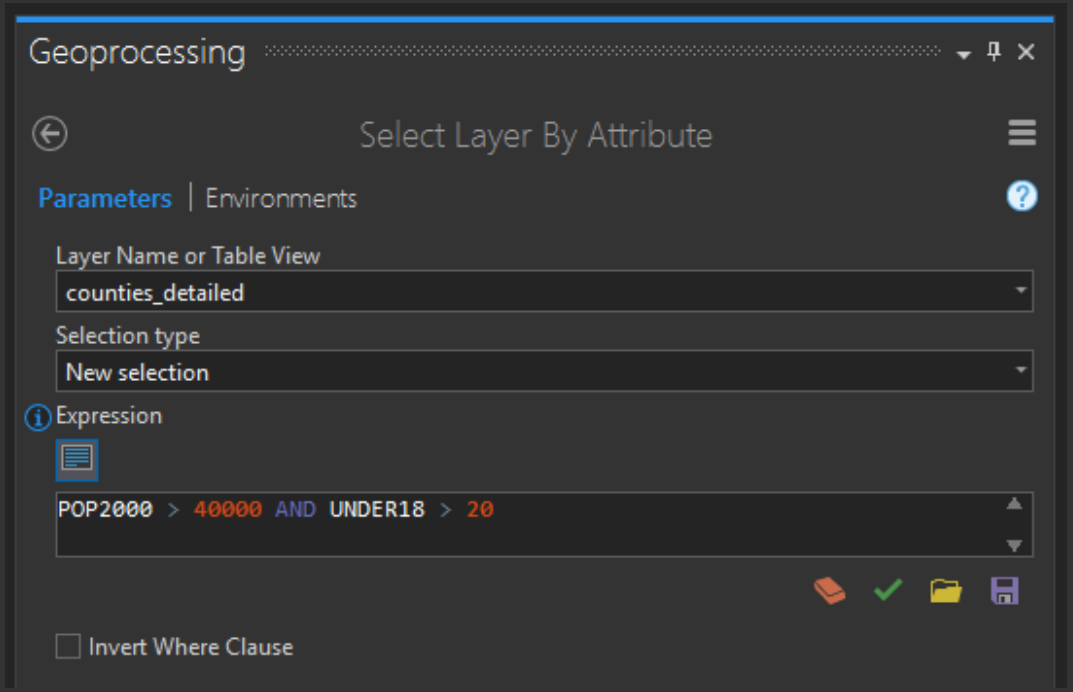
Logical Operators

- Equal: =
- Not Equal: <>
- Greater Than: >
- Greater Than or Equal To: >=
- Less Than: <
- Less Than or Equal To: <=

Boolean Connectors

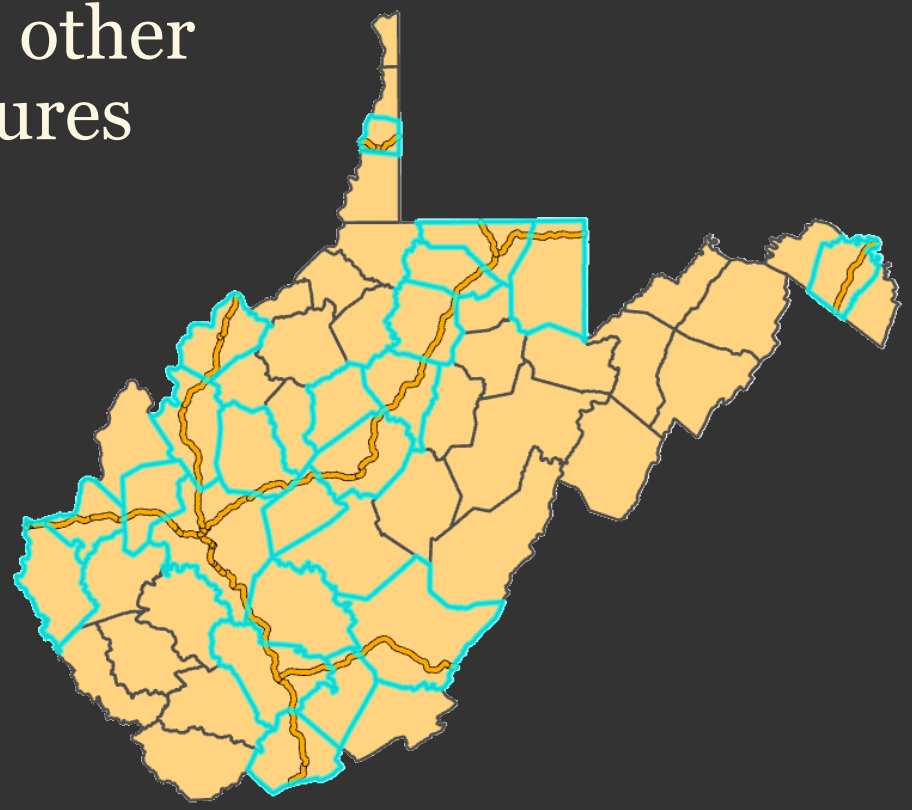
- AND (A and B)
- OR (A or B, including A and B)
- NOT (A but not B)
- XOR (A or B, but not A and B)

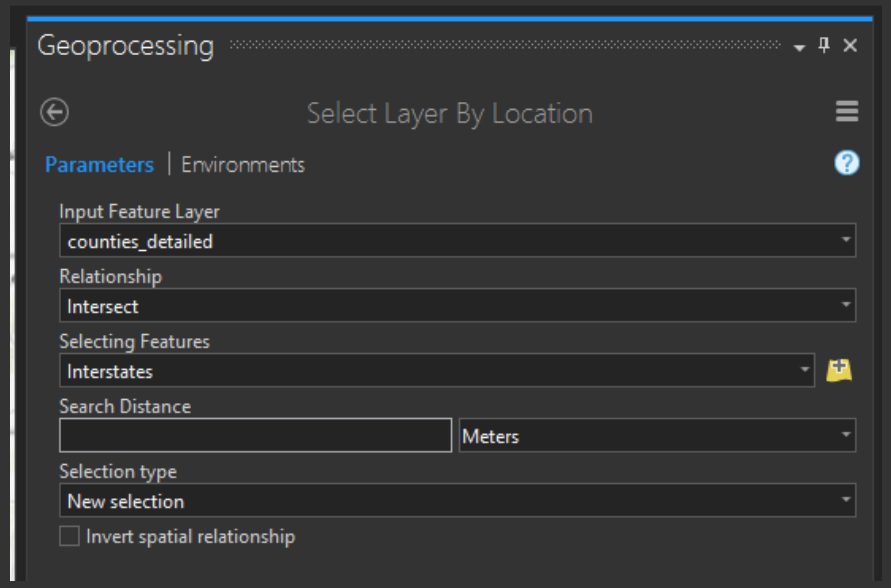
Example: Population >= 20,000 AND Region = Midwest



Select by Location

- Select based on spatial relationships to other geographic features





Select by Location

Intersect = select input features that intersect the selecting features. This is the default.

Within a Distance = select input features that are within the specified Euclidean distance of the selecting features.

Contains = select any input feature that contains, even partially, a selecting feature.

Completely Contains = select any input feature that completely contains a selecting feature.

Within = select any input feature that is within, even partially, a selecting feature.

Completely Within = select any input feature that is completely within a selecting feature.

Select by Location

Are Identical To = select any input feature that has identical geometry to a selecting feature.

Boundary Touch = select any input feature that has a boundary that touches a selecting feature.

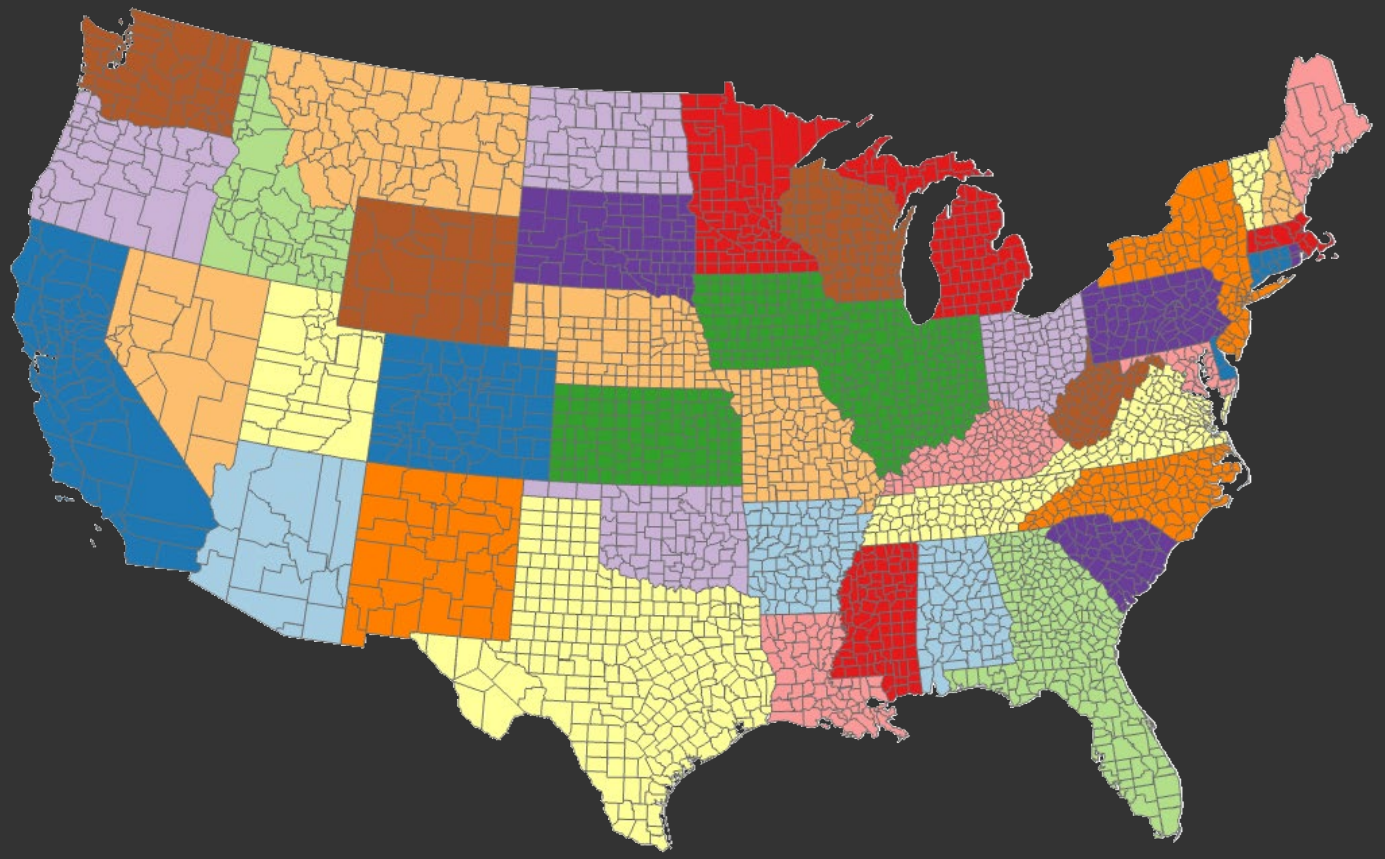
Share A Line Segment With = select any input feature that shares a line segment with a selecting feature.

Crossed By Outline Of = select any input feature that is crossed by the outline of a selecting feature.

Have Their Center In = select any input feature that has its center in the selecting features. The center for lines is defined as the midpoint and the center for polygons is defined as the centroid.

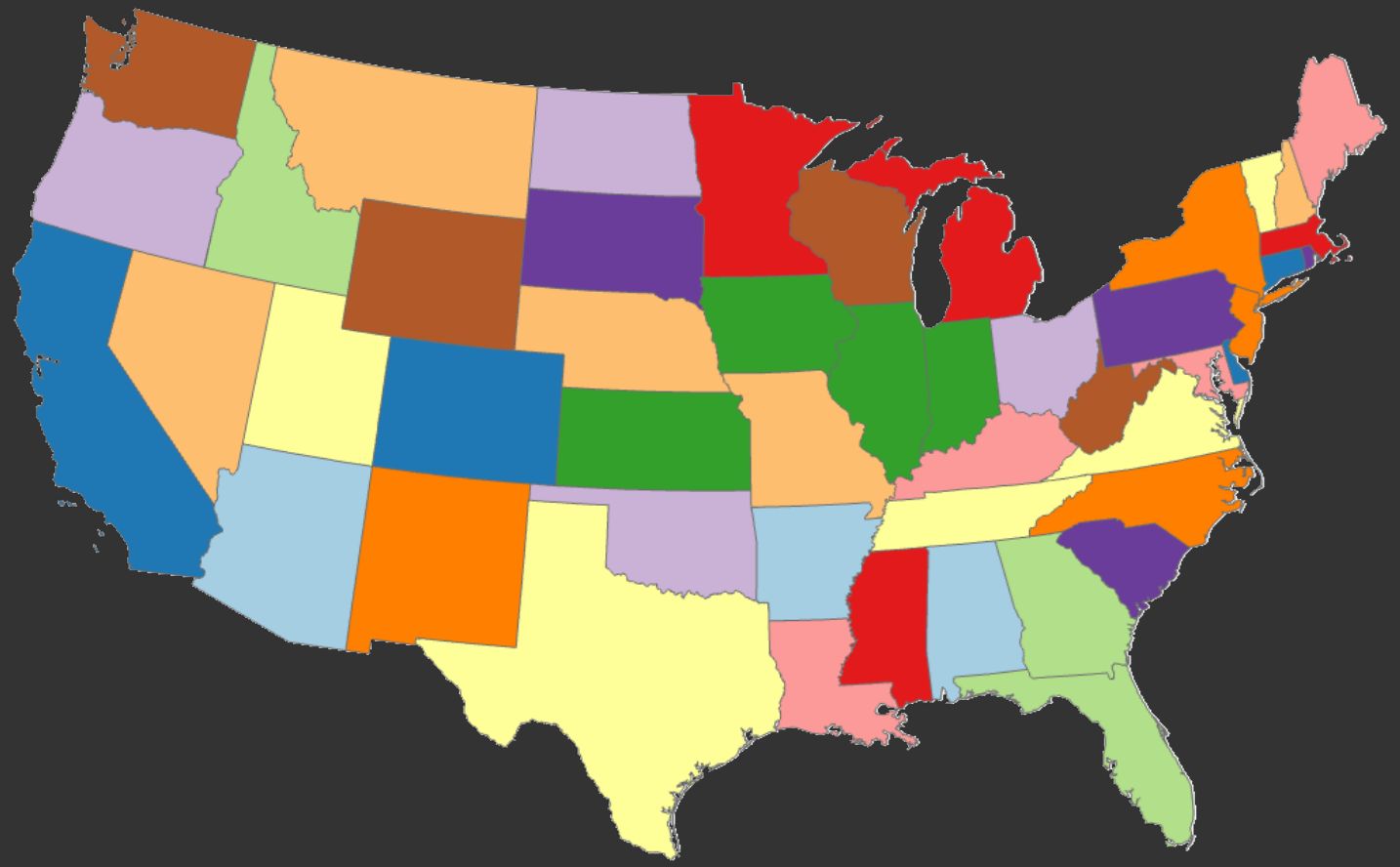
Generalization: Dissolve

- Generalize boundaries
- Combine all features to one feature or use a common attribute



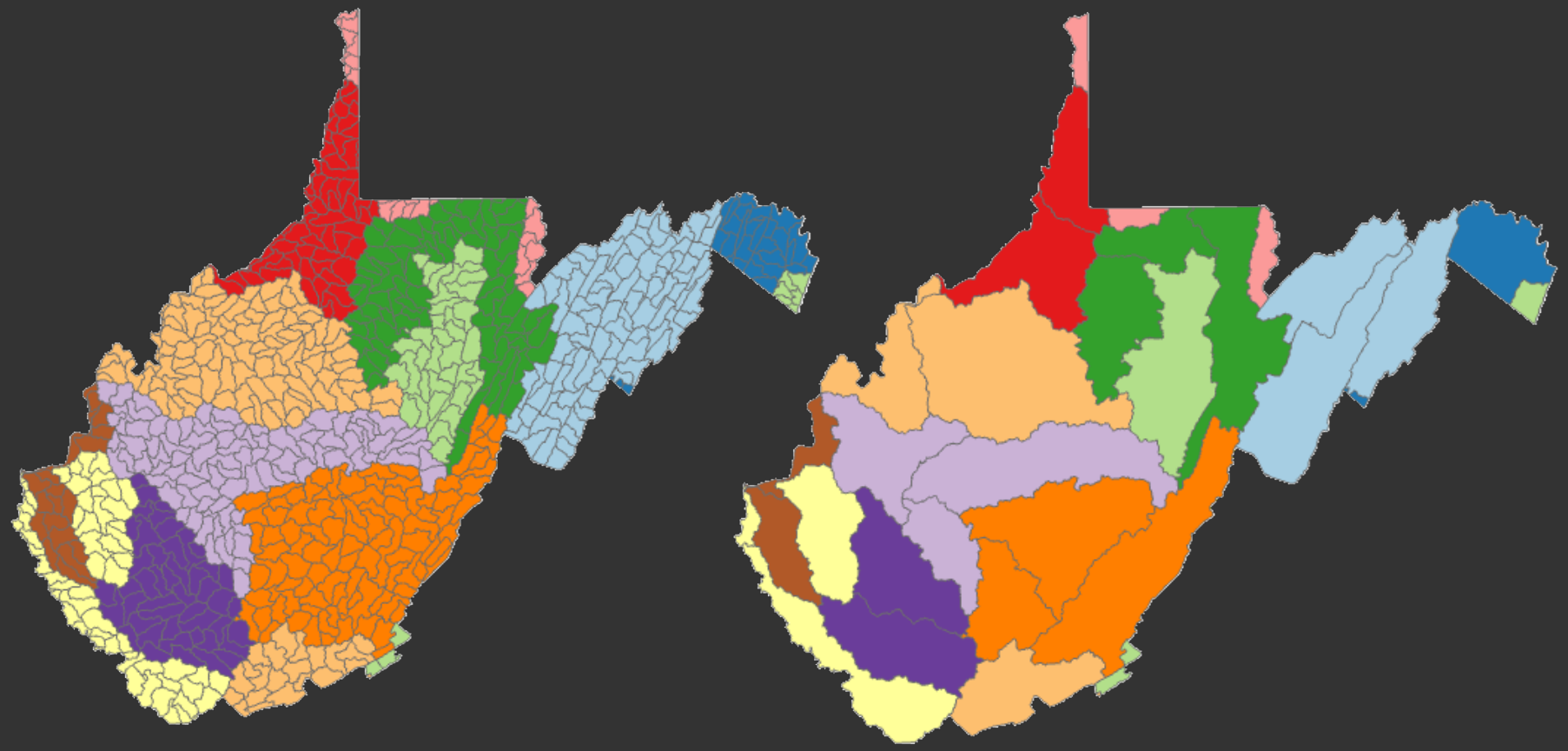
Generalization: Dissolve

- Generalize boundaries
- Combine all features to one feature or use a common attribute

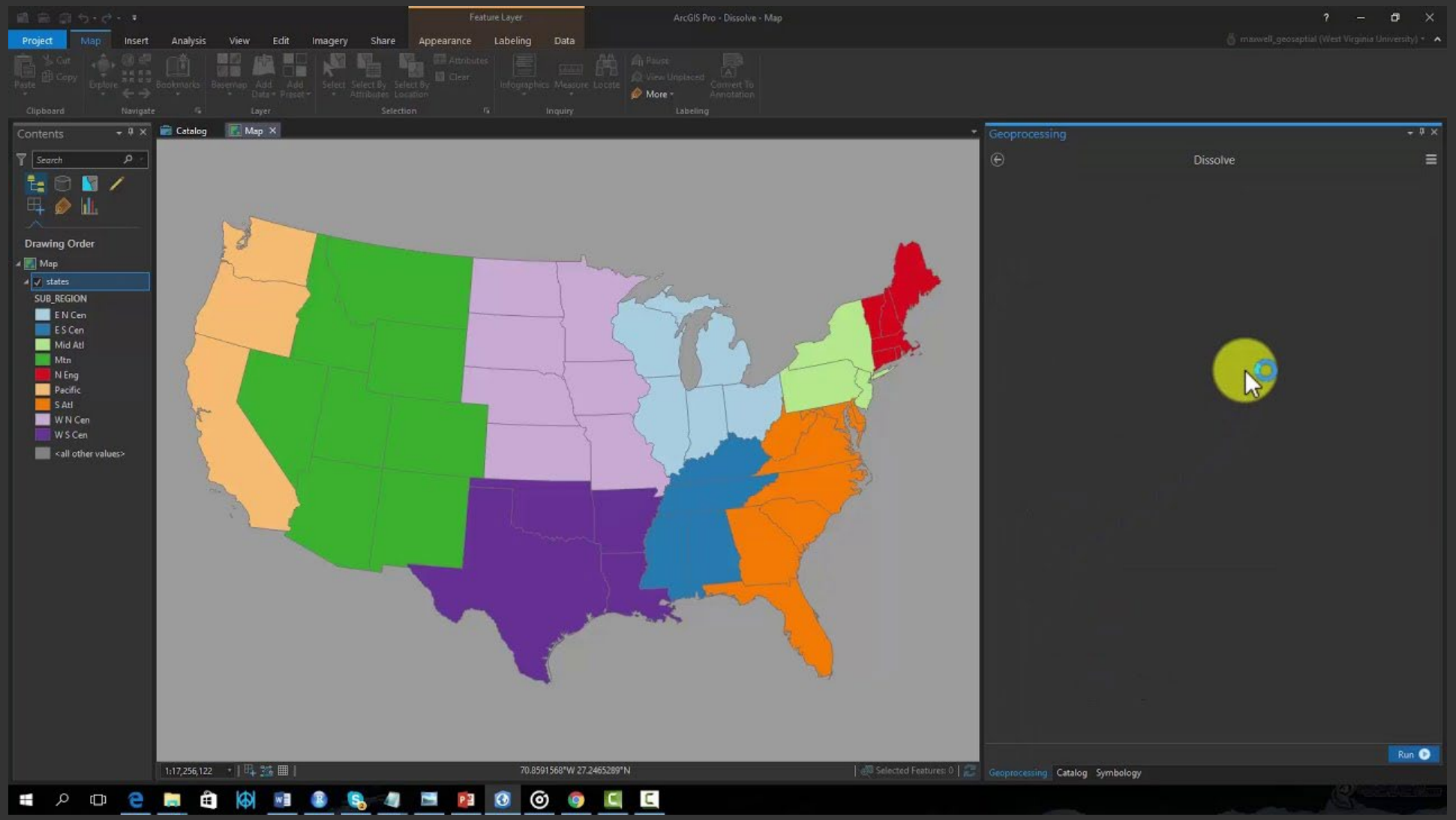


Generalization: Dissolve

- Generalize boundaries
- Combine all features to one feature or use a common attribute

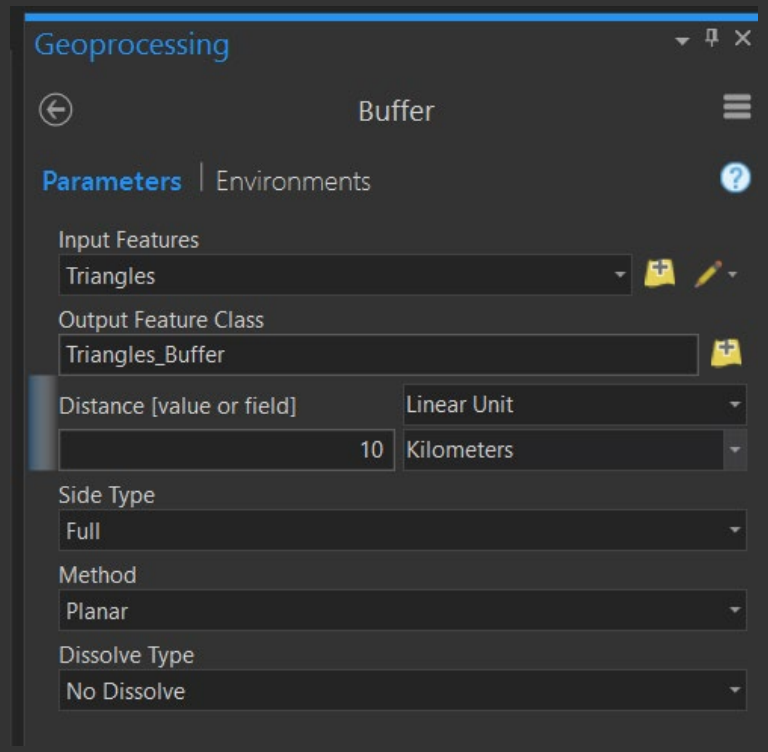


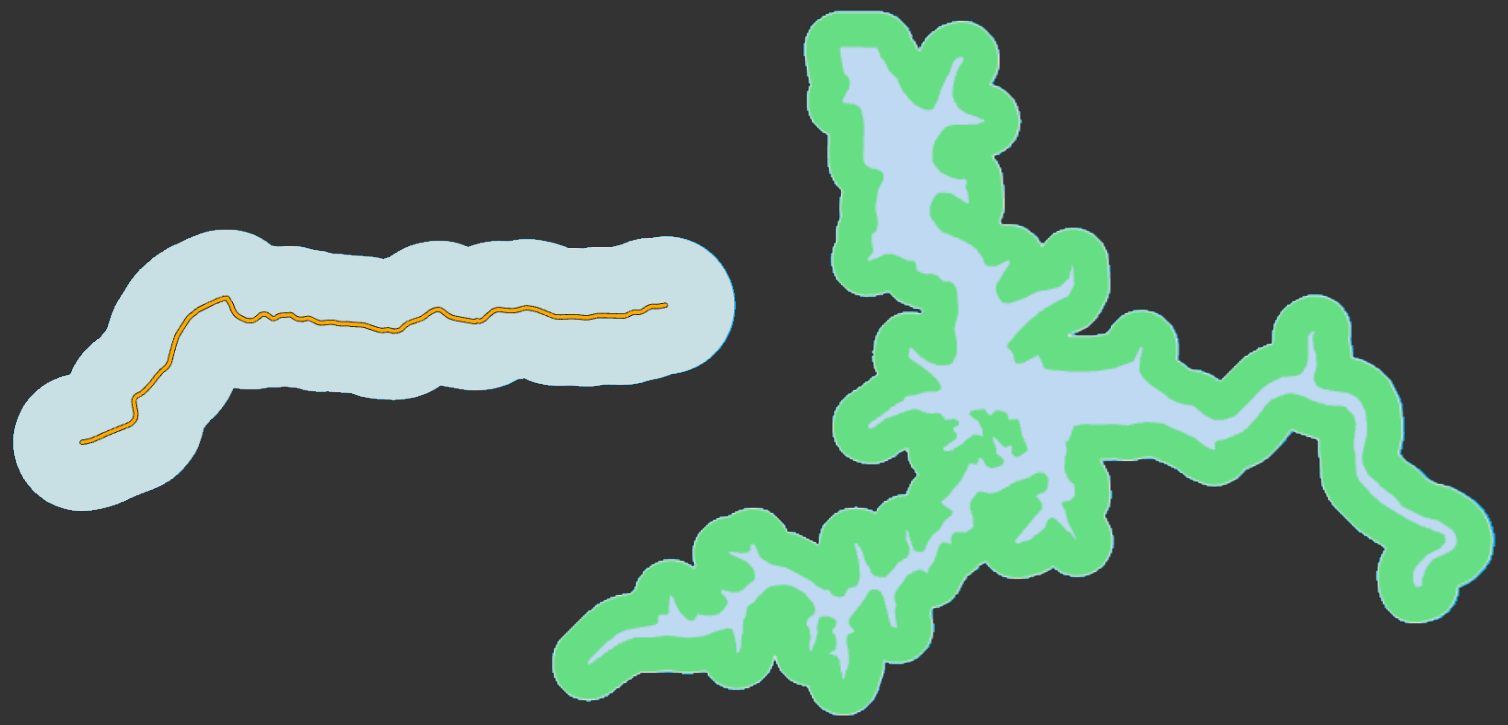
Video: Dissolve



Proximity Analysis: Buffer

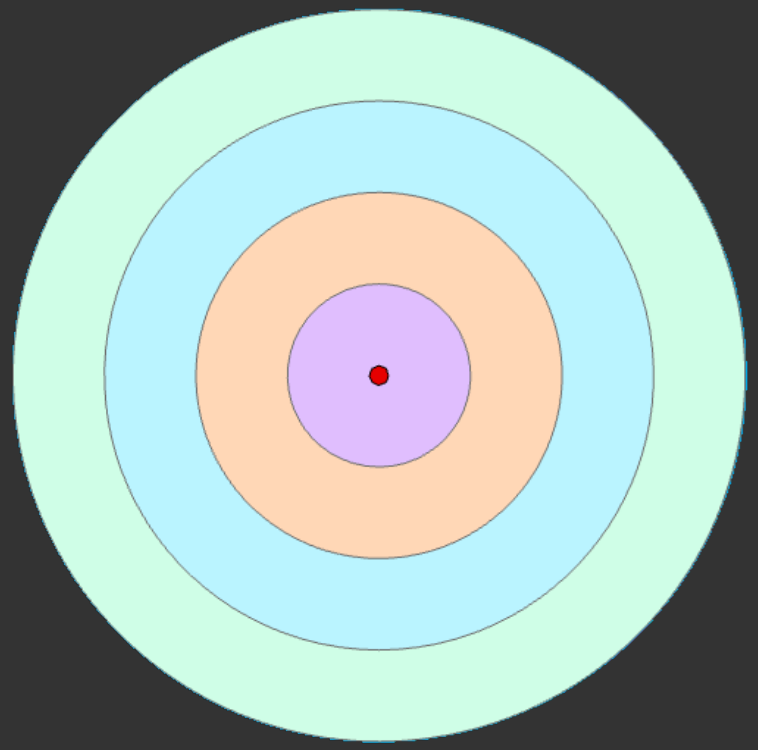
- Distance can be the same for all features or based on a field

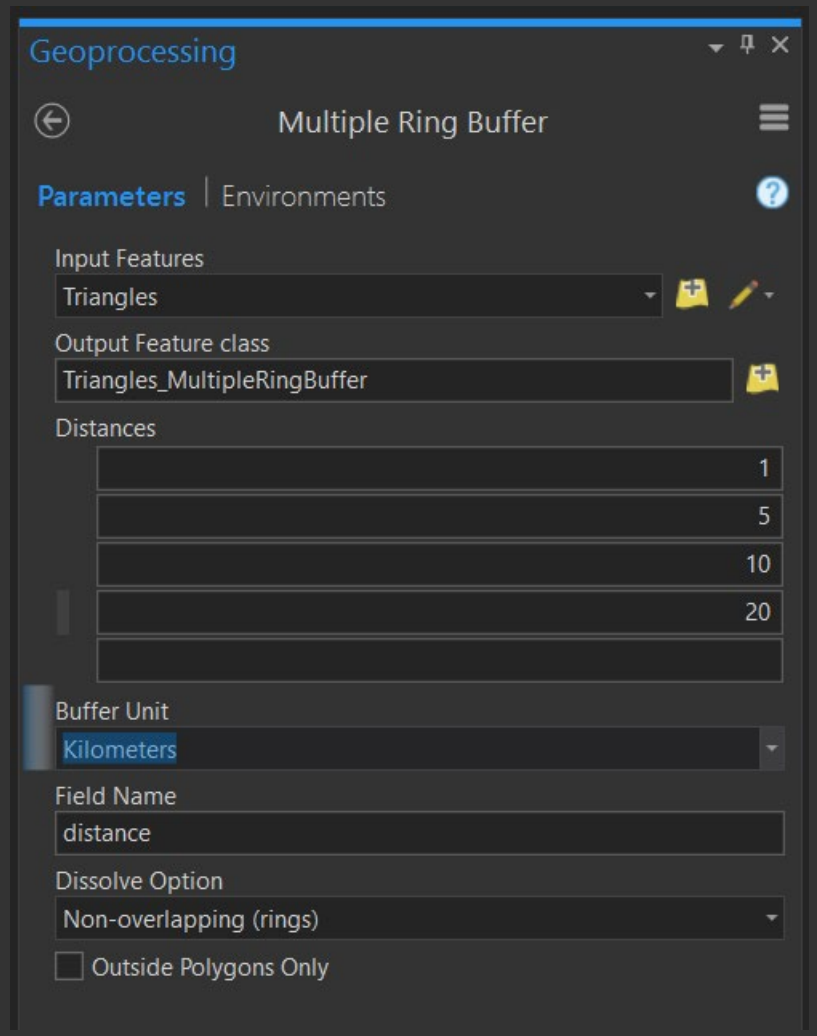




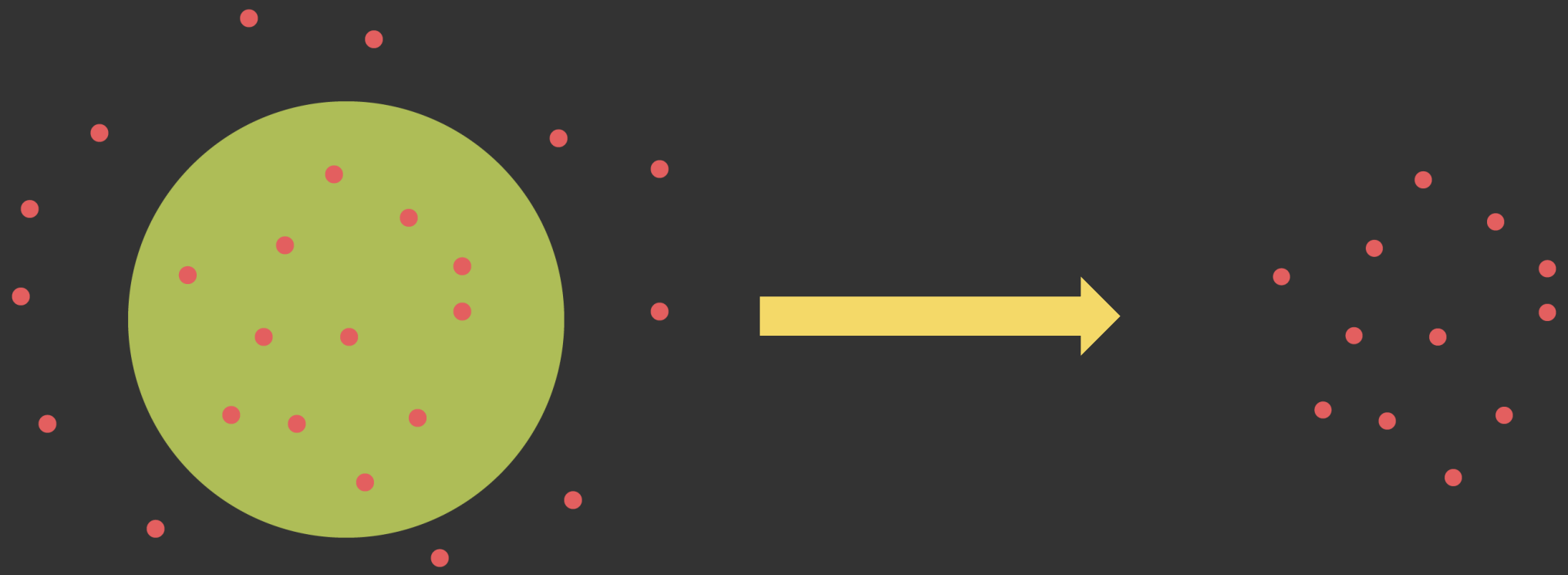
Proximity Analysis: Multiple Ring Buffer

- Buffer with multiple distances





Clip



Clip

Input Feature Attributes

Type
Point
Point
Point
Point

Clip Feature Attributes

Type
Circle
Circle
Circle
Circle

Output Feature Attributes

Type
Point
Point
Point
Point

Clip feature attributes not added to the output

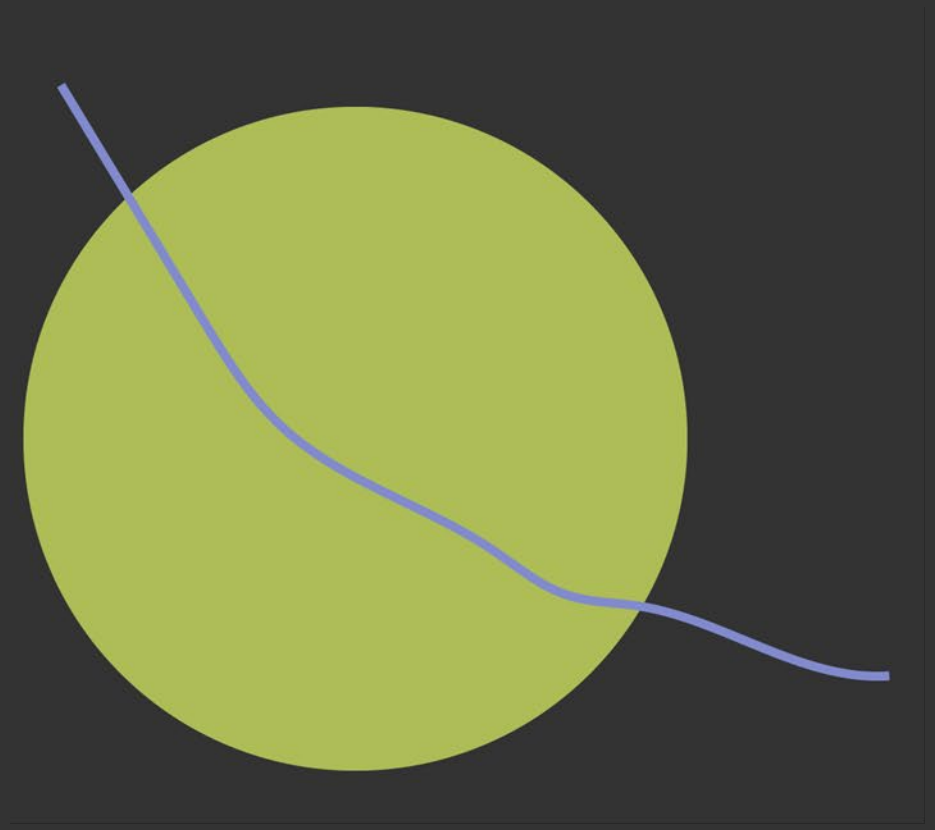
Clip

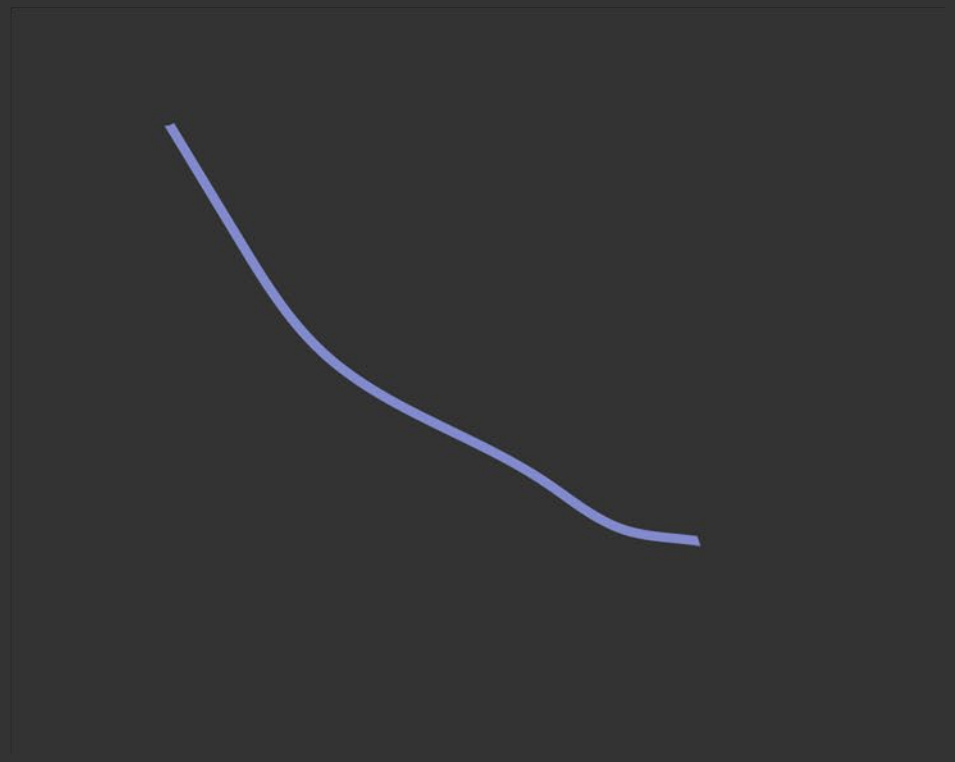
Input = Triangle
Clip Feature = Circle

Output = Portion of
Triangle in Circle

Clip

For shapefiles, geometric fields will not be updated in attribute table!!!



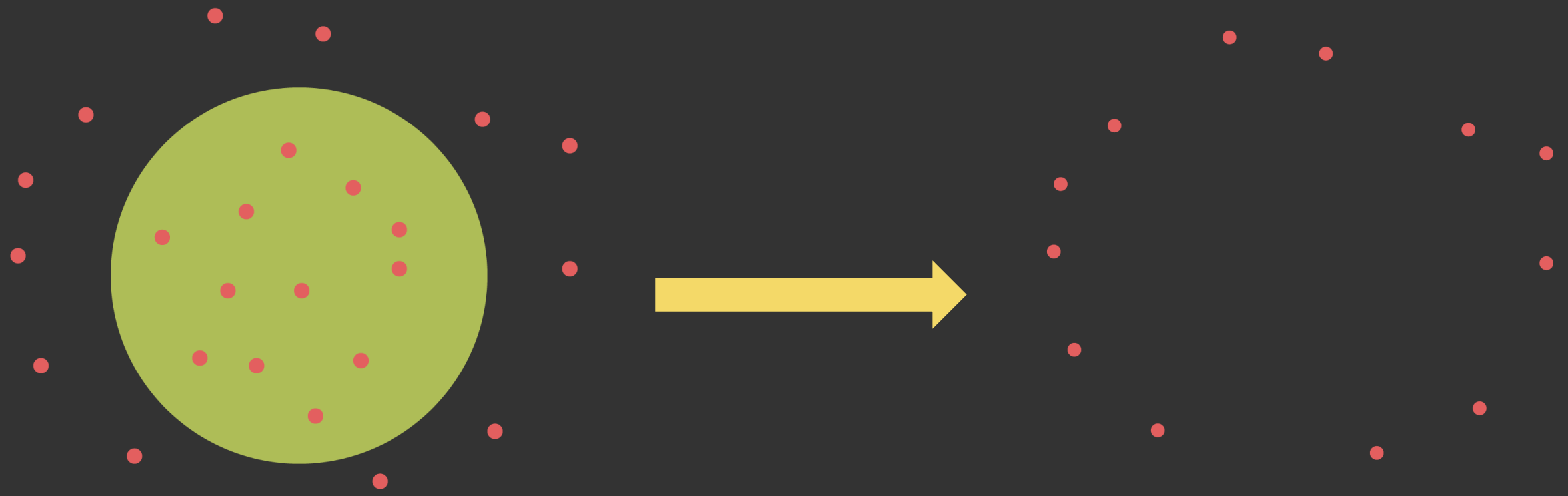


Input = Line
Clip Feature = Circle

Output = Portion of Line in Circle

Erase (NOT)

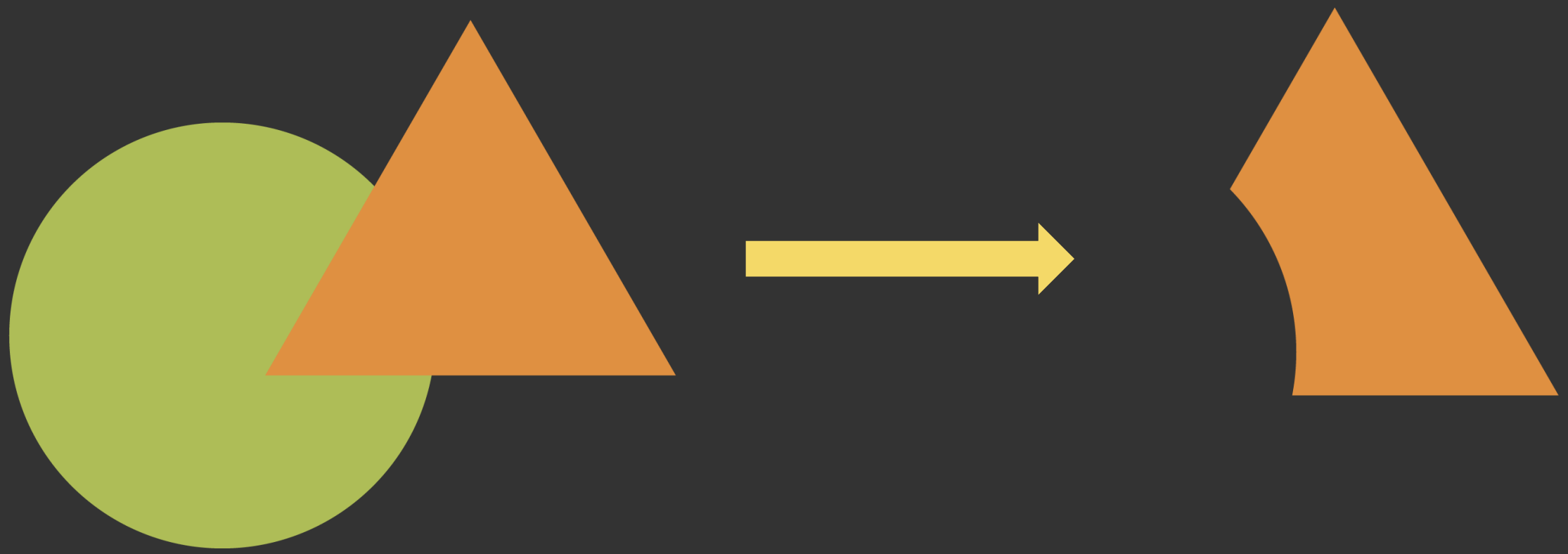
- The opposite of clip



Erase feature attributes not added to the output

Erase (XOR)

- The opposite of clip



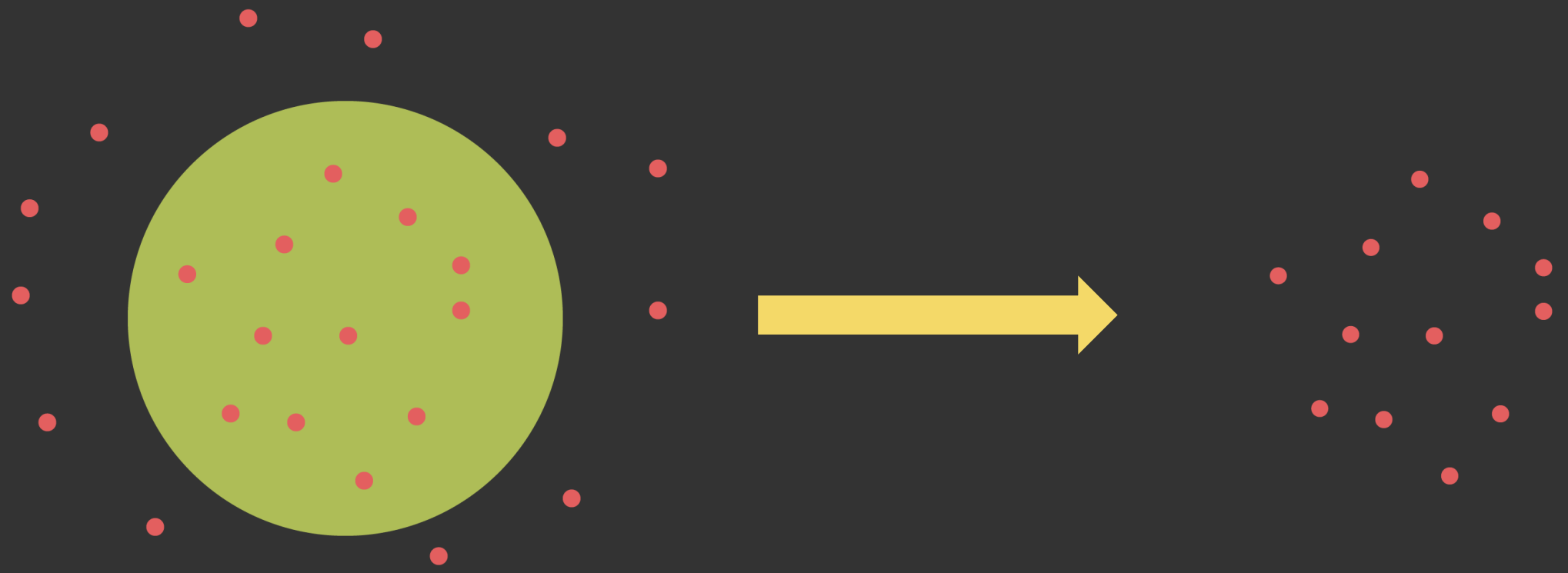
Erase feature attributes not added to the output

Intersect (AND)

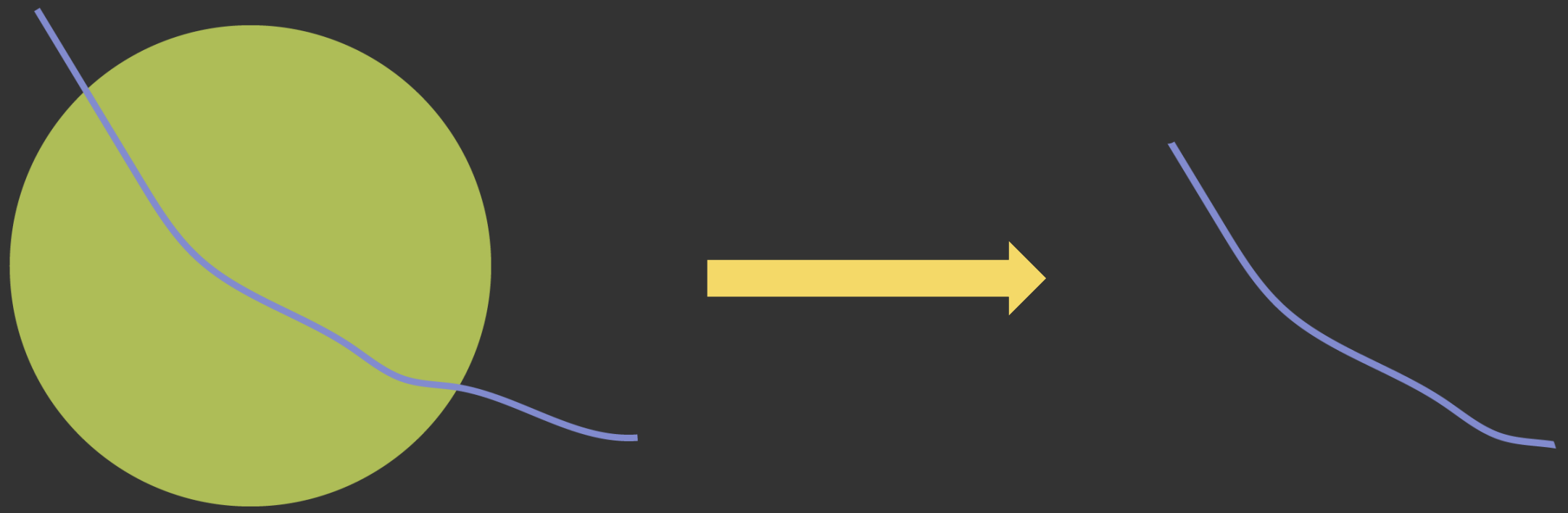
- Keep common area and boundaries in that common area
- Attributes of all intersected features in the resulting table

Intersect (AND)

- Although the geometry is the same as Clip, the attribute table will have all of the attributes

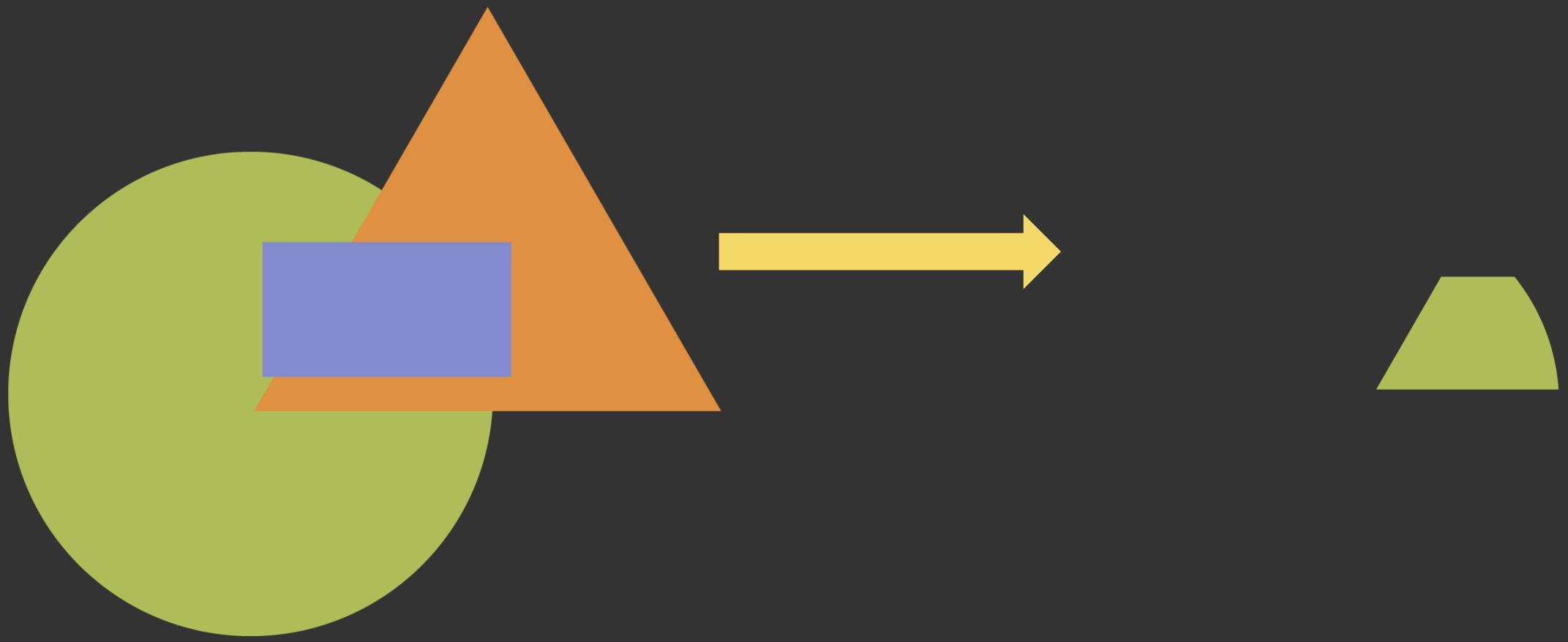


Overlay: Intersect



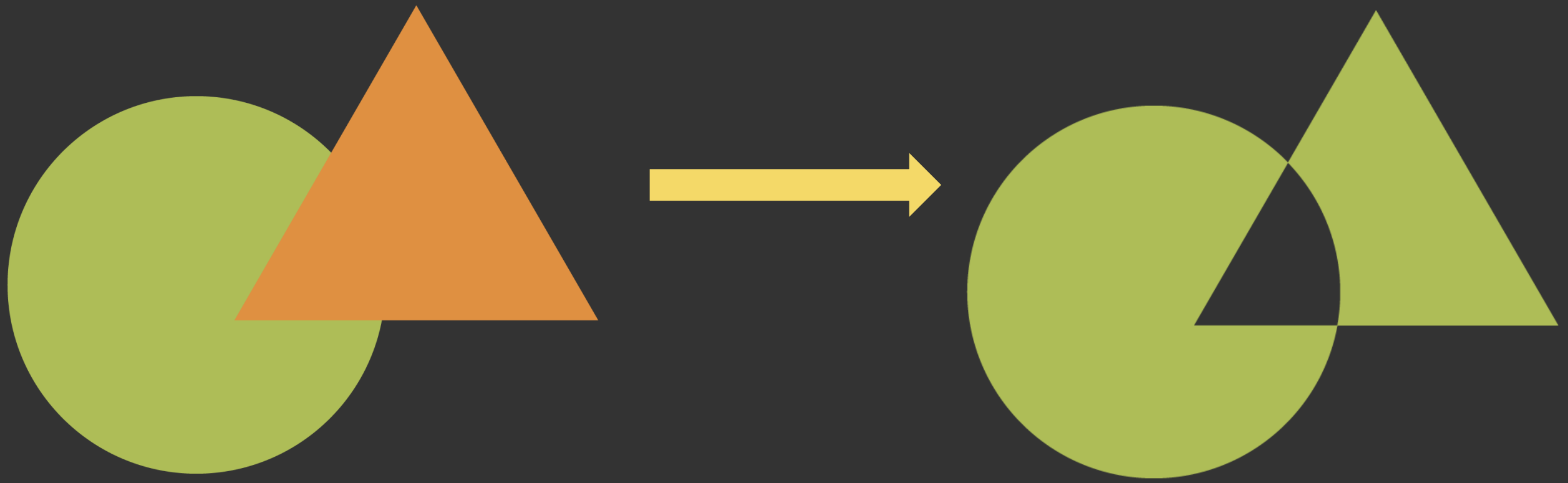
- Although the geometry is the same as Clip, the attribute table will have all of the attributes

Intersect (AND)



You can intersect more than two layers

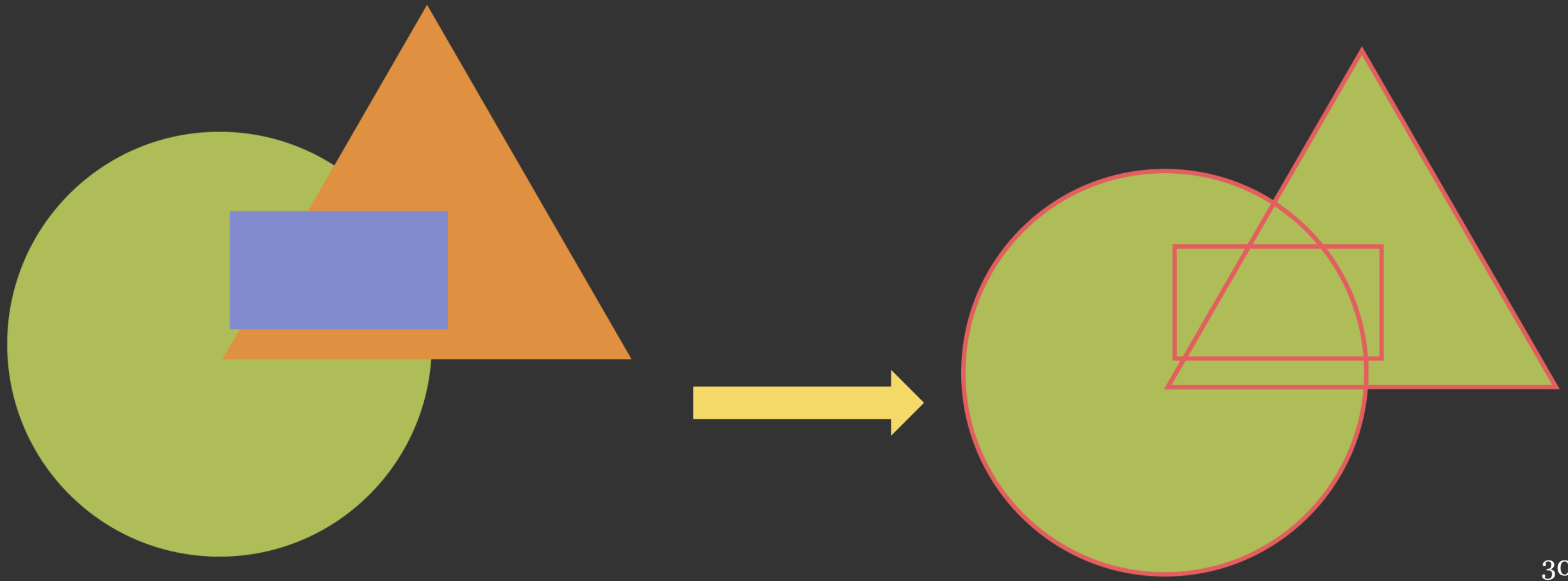
Symmetrical Difference (XOR)



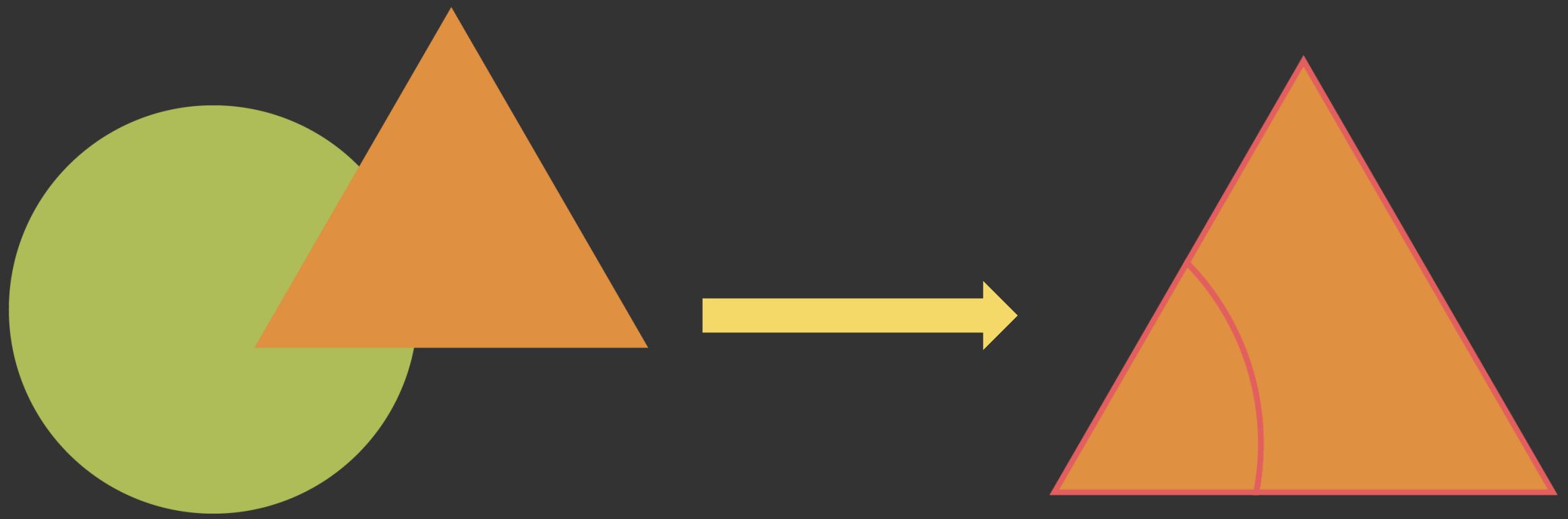
- Opposite of Intersect = only maintains areas that are not shared
- All input attributes are maintained

Union (OR)

- Keep all boundaries and attributes
- Used for polygon data only

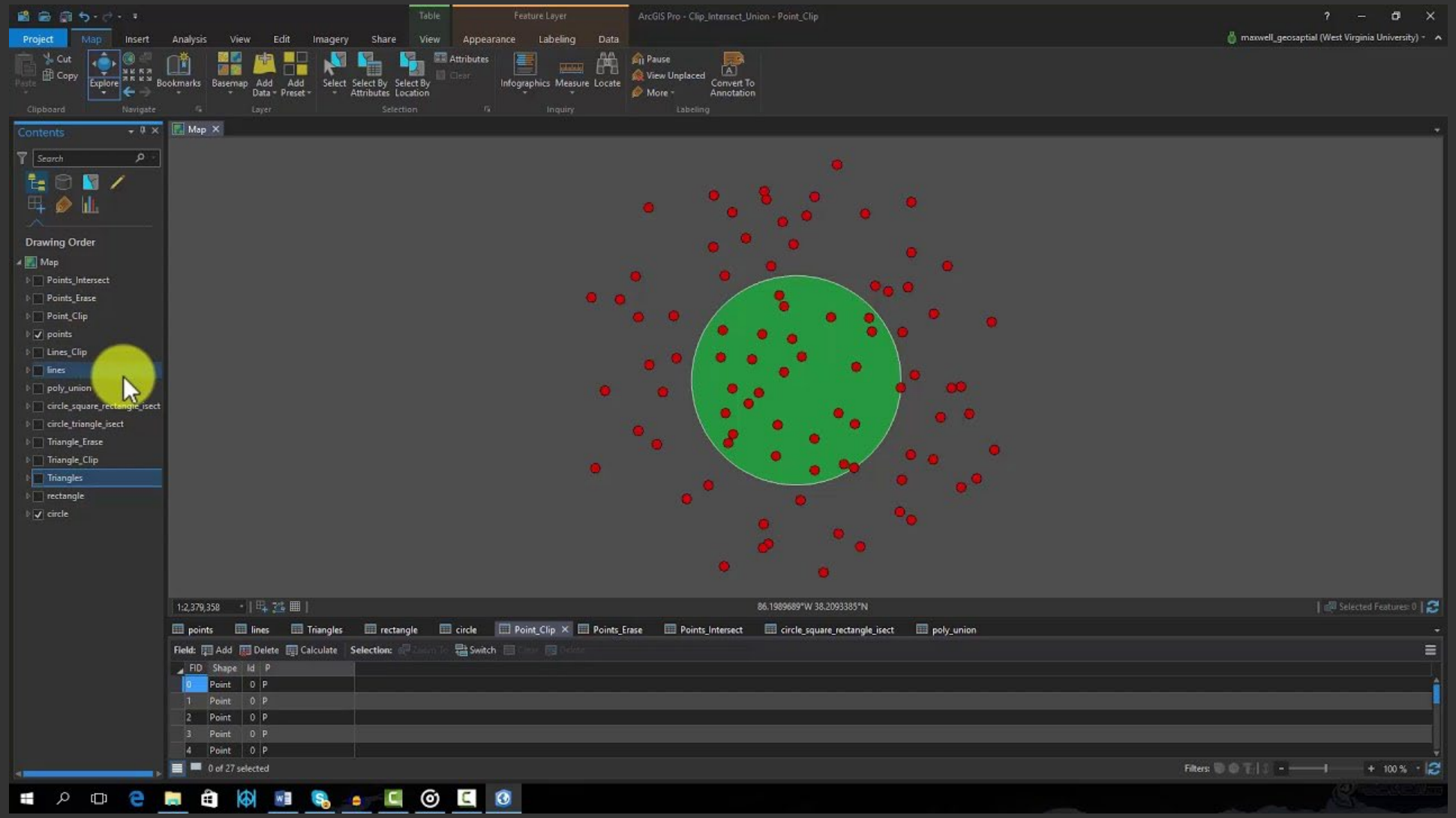


Identity



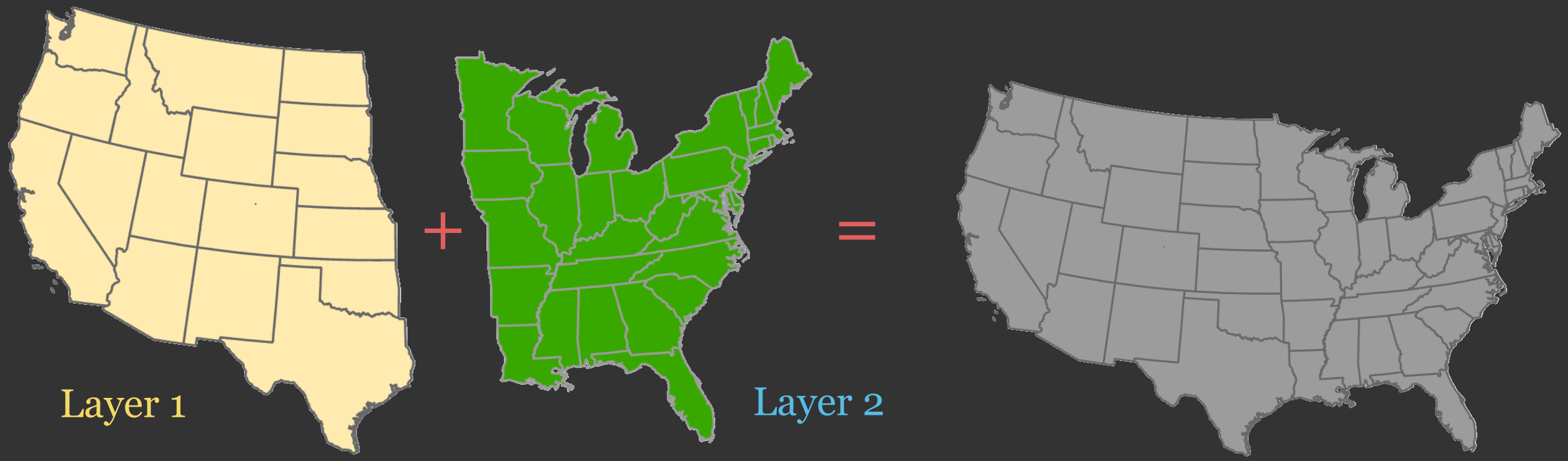
- Slices up an input layer relative to an identity layer
- Maintains all attributes

Video: Vector Overlay



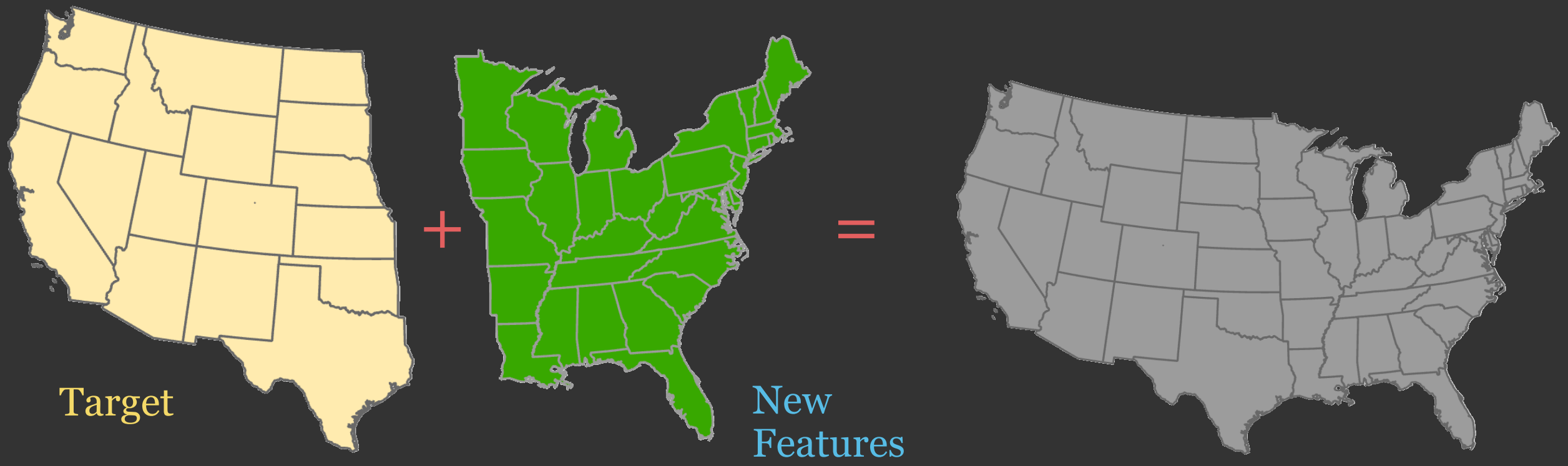
Merge

- Combine datasets of the same type: points, lines, or polygons



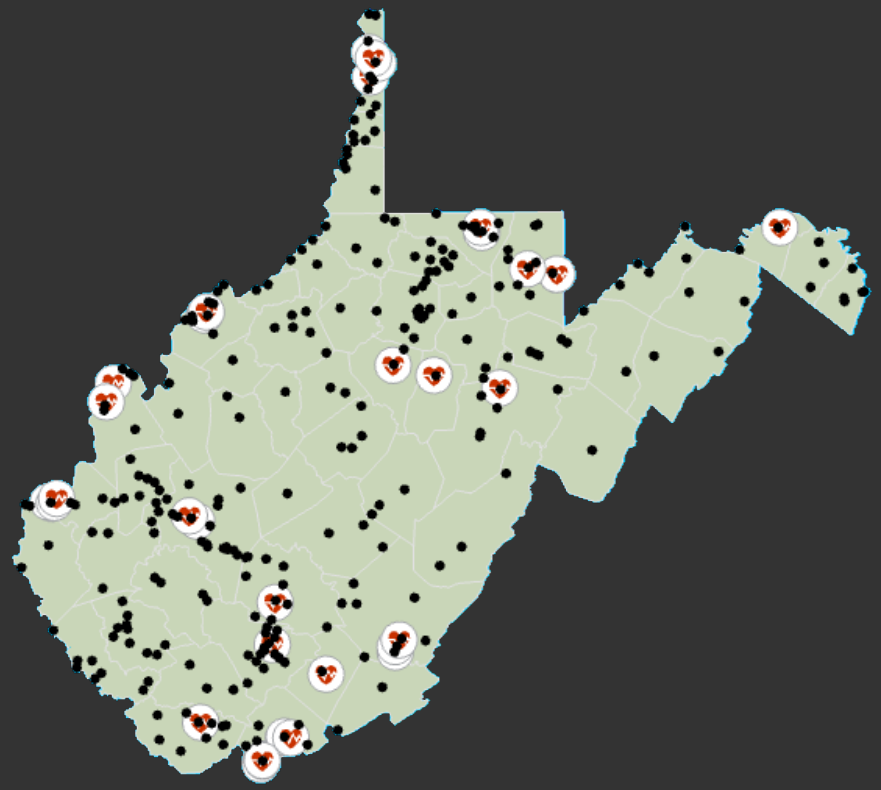
Append

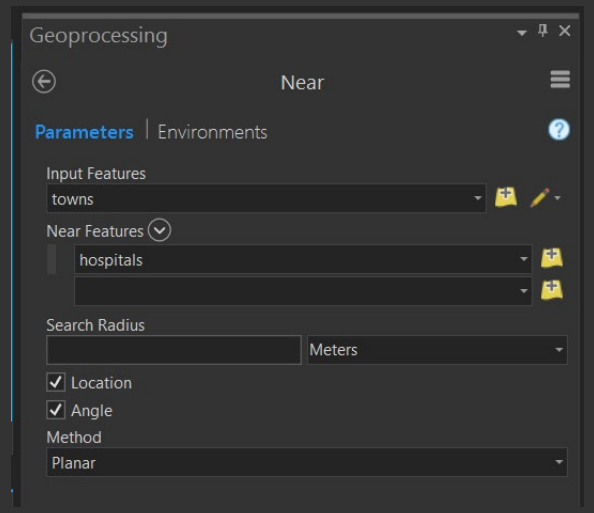
- Append new features to a target dataset



Near

- Calculates distance and additional proximity information between the input features and the closest feature in another layer or feature class.





NEAR_DIST	NEAR_X	NEAR_Y	NEAR_ANGLE
58894.721581	566123.56083	4316108.98758	75.571204
5302.150386	612877.109511	4369211.6301	-142.646654
15611.139892	546782.444547	4178069.65597	9.190721
19553.604681	443977.616489	4245995.17619	25.941755
42651.425334	449685.102582	4142924.79279	-65.304017
15679.871214	449685.102582	4142924.79279	135.684787
29021.050001	545740.876617	4321173.56963	-122.113975
18672.283133	486840.807273	4202724.97106	-103.076304
7620.342891	494728.242096	4135225.01702	-120.499255

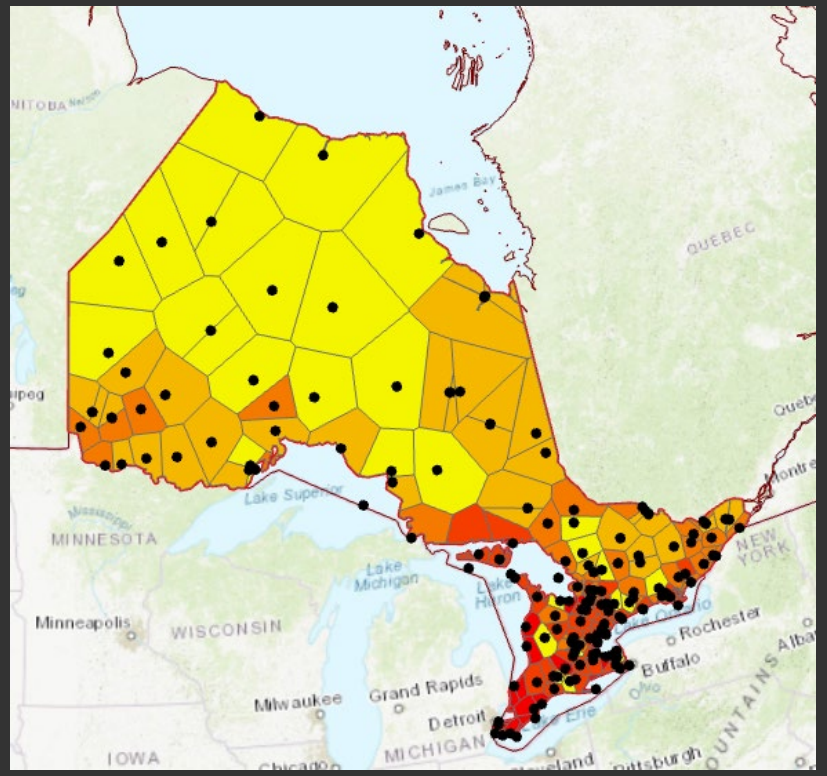
Calculate distance from towns to nearest hospital

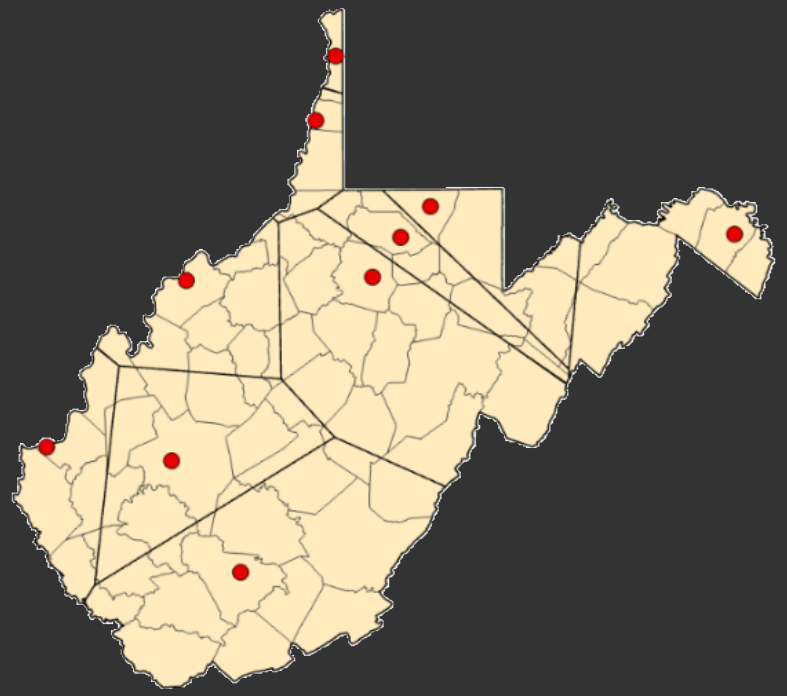
Feature to Point

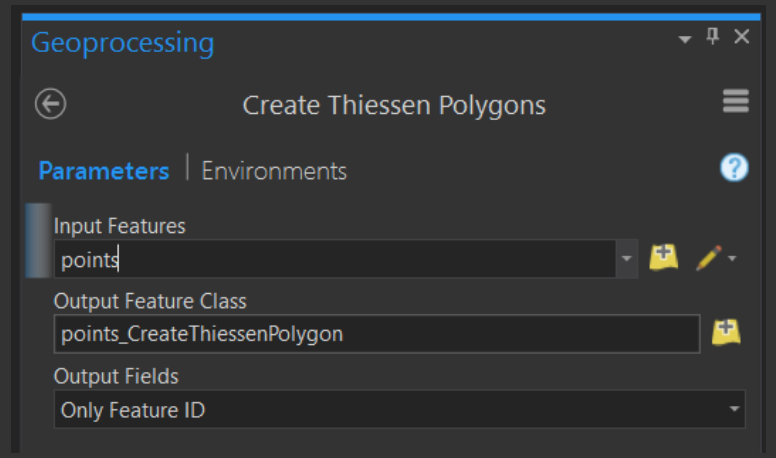
- Multipoint feature = the output point will be located at the average x and y coordinates of all the points in the multipoint feature
- Input line feature = the output point will be located at the weighted average x and y coordinates of the midpoints of all line segments in the line feature
- Polygon feature = the output point will be located at the center of gravity (centroid) of the polygon.
- Can force result to be inside of polygons

Thiessen (Voronoi) Polygons

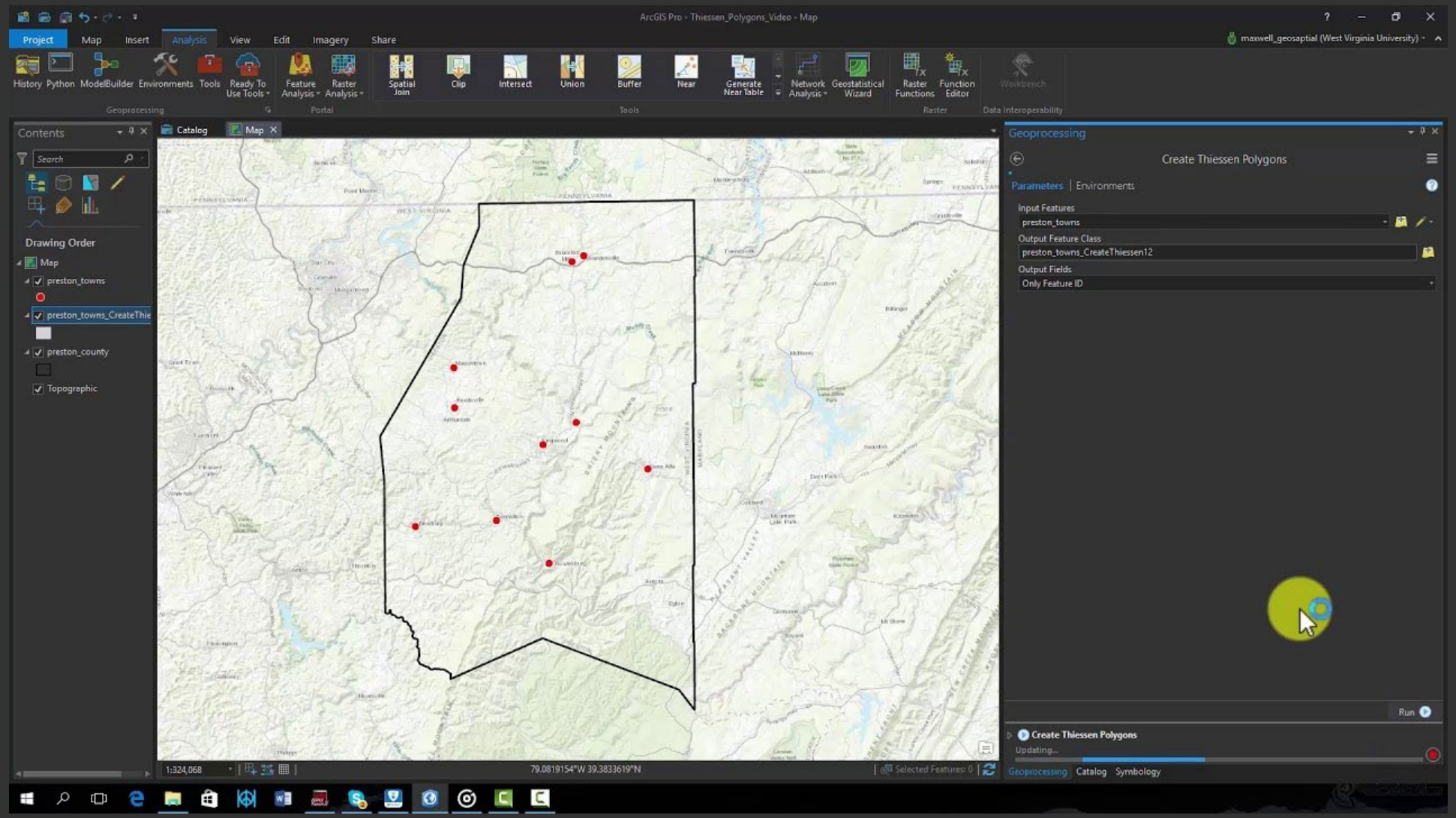
- Defines the area that is closer to one point than any other points in a dataset





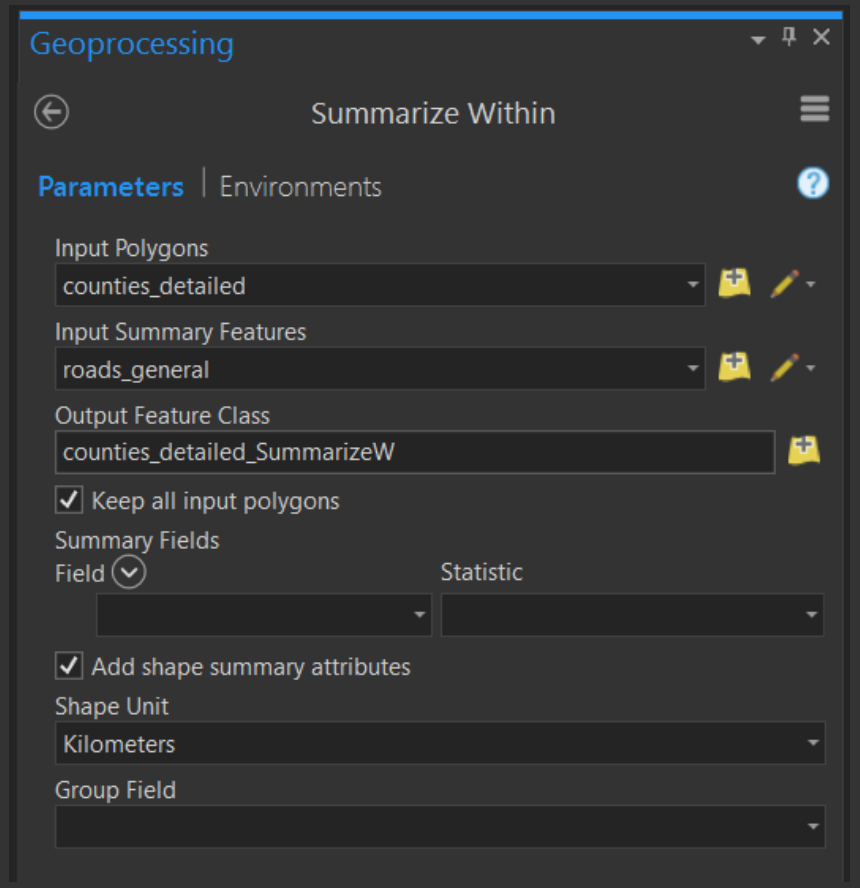


Video: Thiessen Polygons



Summarize Within

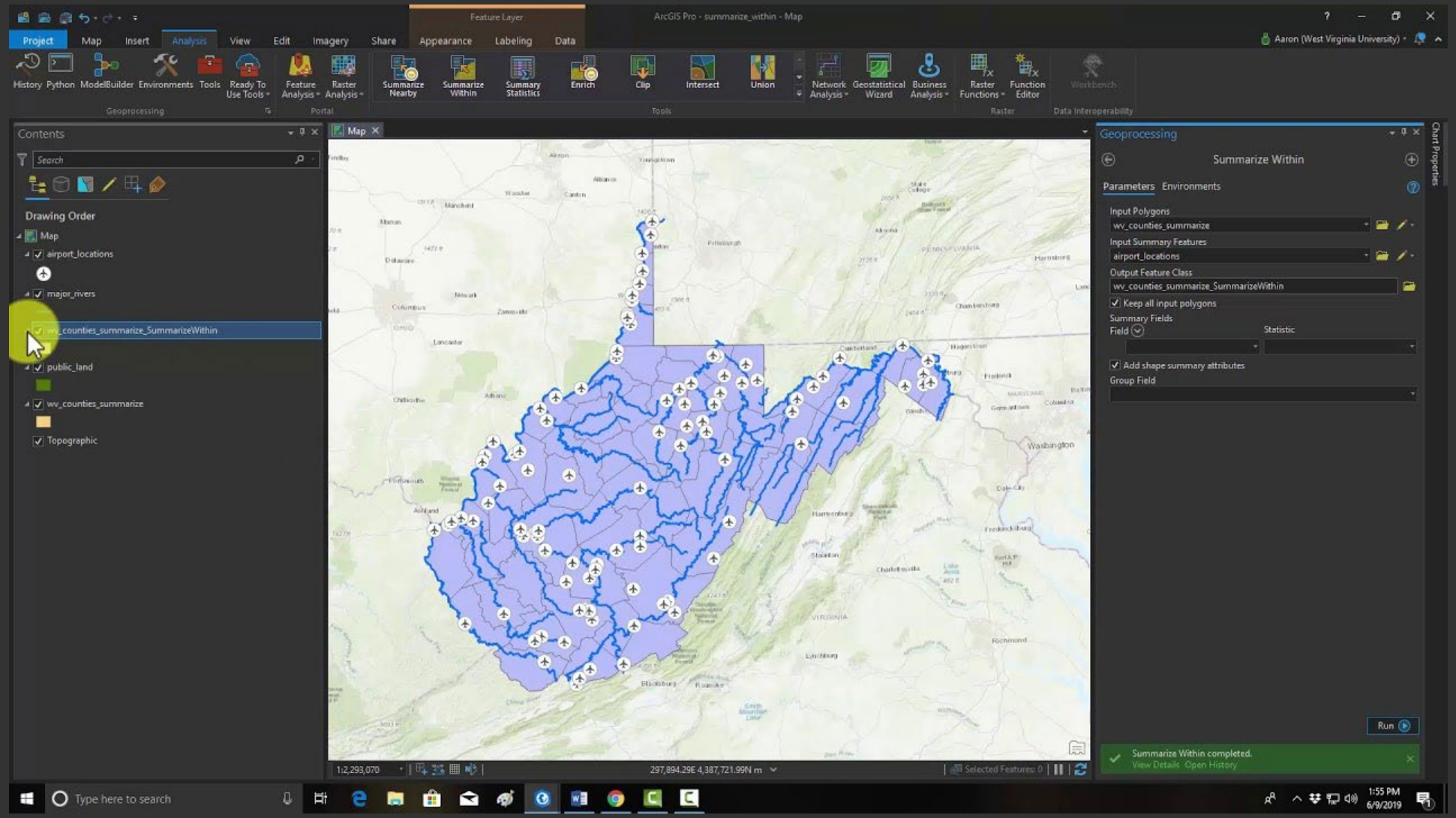
- Overlays a polygon layer with another layer to summarize the number of points, length of the lines, or area of the polygons within each polygon, and calculate attribute field statistics about those features within the polygons.



Summarized length in KILOMETERS
698.311215
1289.563736
2726.859736
830.083822
1848.318289
1188.3092
1193.580602
1365.397175
1519.167536

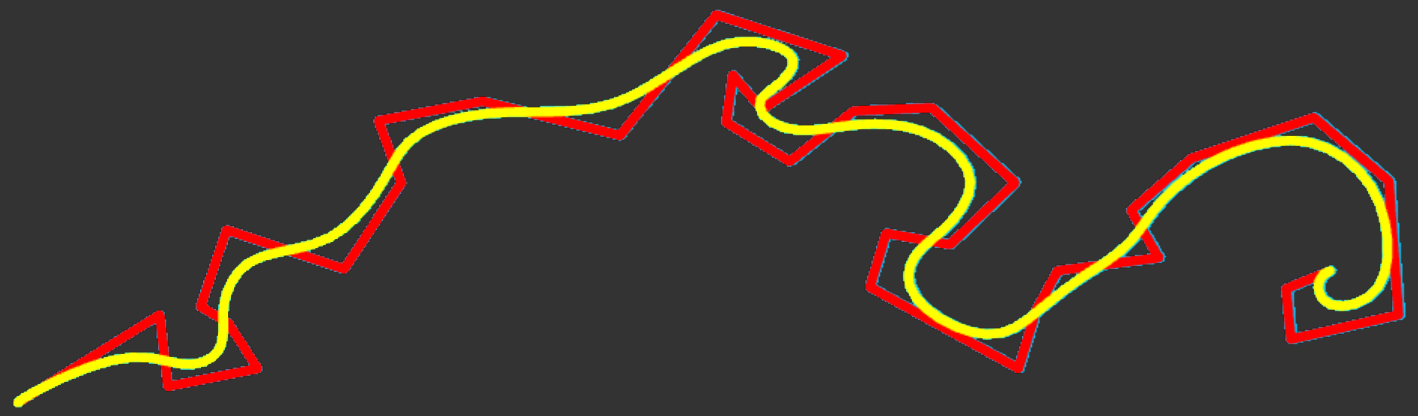
Example: summarize length of roads in each county

Video: Summarize Within



Line Generalization

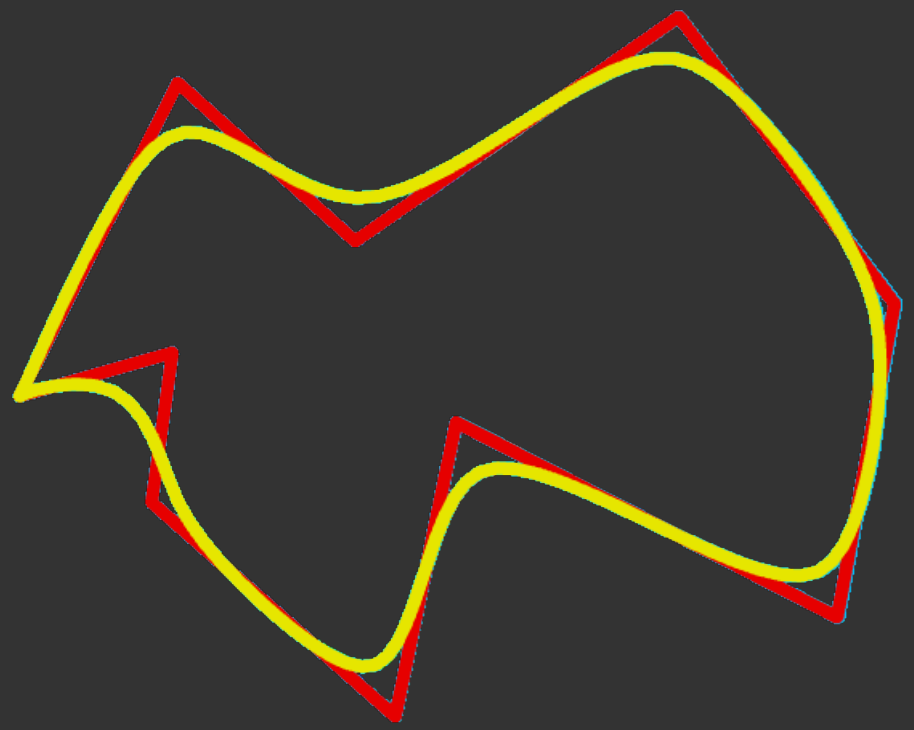
- Remove unnecessary vertices
- Smooth curves
- For cartographic generalization
- Simplify Line Tool
- Smooth Line Tool



Available under Generalization Tools in the Cartography Toolbox

Polygon Generalization

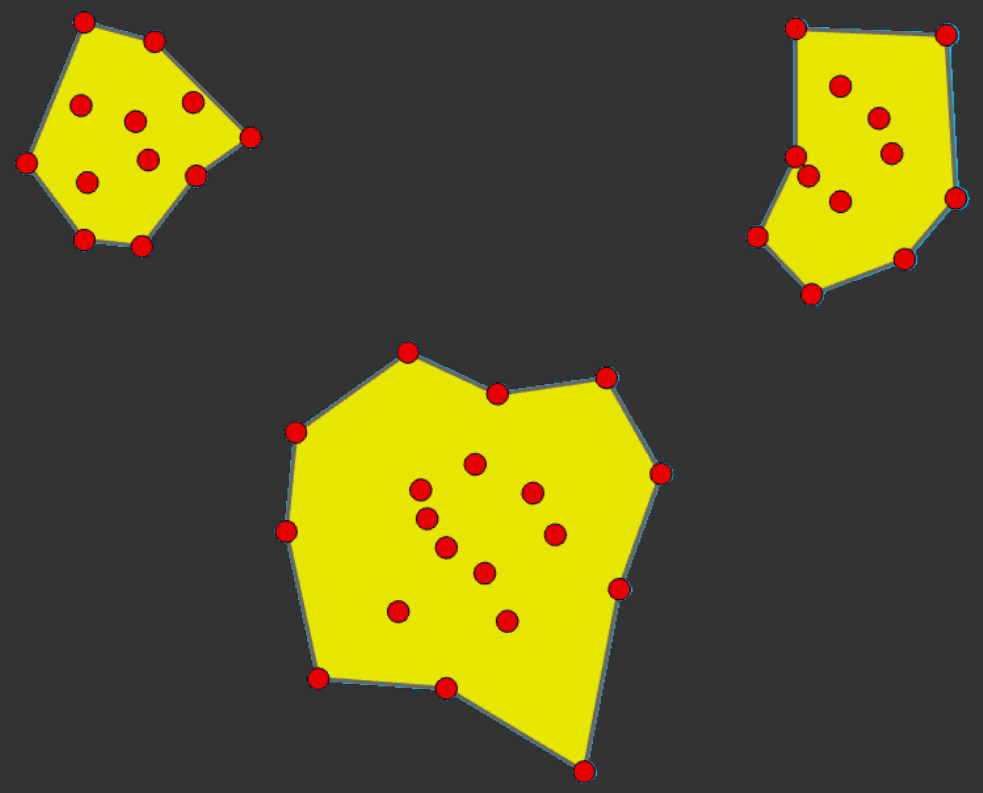
- Remove unnecessary vertices
- Smooth curves
- For cartographic generalization
- Simplify Polygon Tool
- Smooth Polygon Tool
- Aggregate Polygons Tool



Available under Generalization Tools in the Cartography Toolbox

Point Aggregation

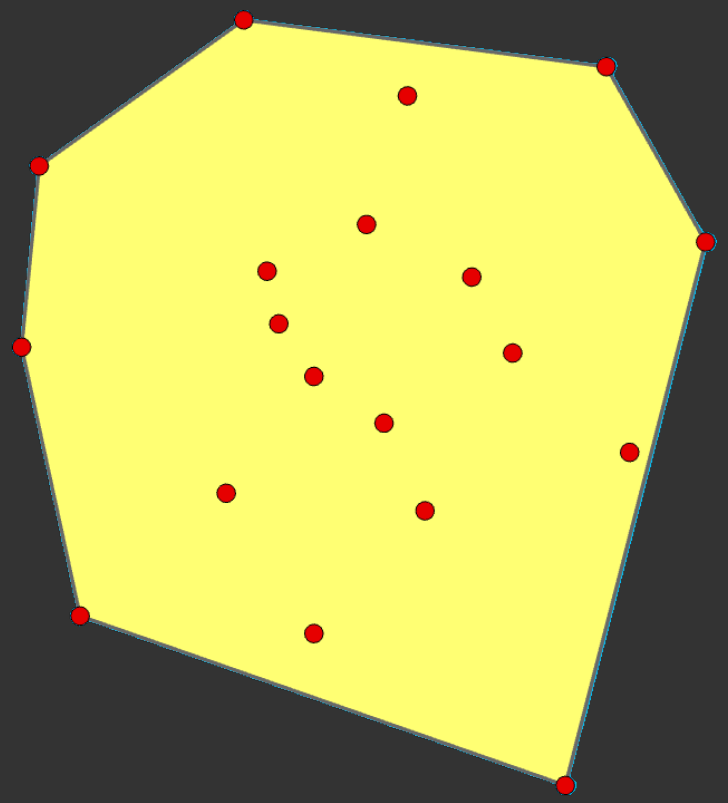
- Aggregate points to polygon
- Can be used for cartographic generalization
- Point Aggregation Tool



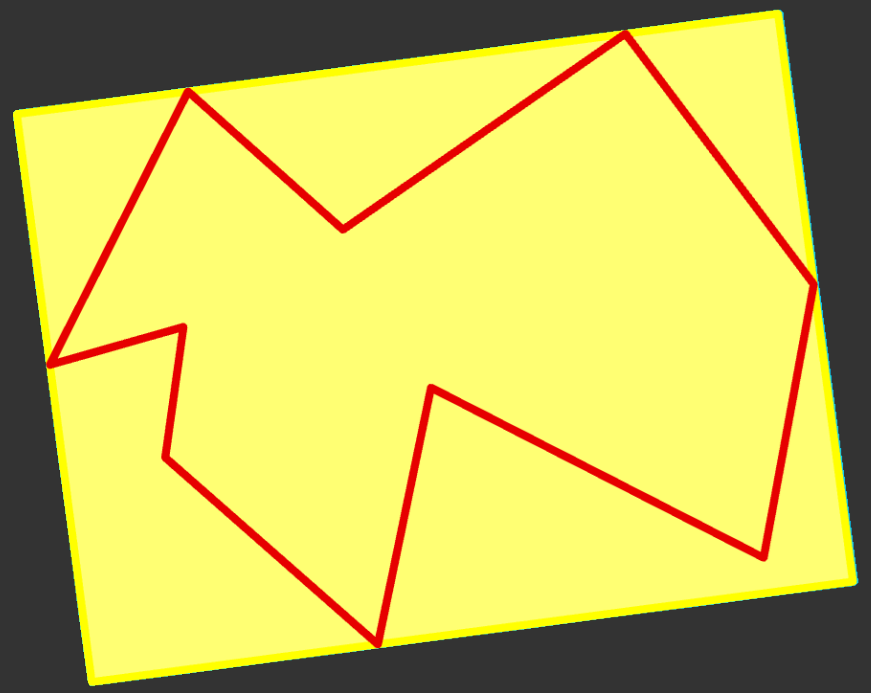
Available under Generalization Tools in the Cartography Toolbox

Minimum Bounding Polygon

- Creates a feature class containing polygons which represent a specified minimum bounding geometry enclosing each input feature or each group of input features



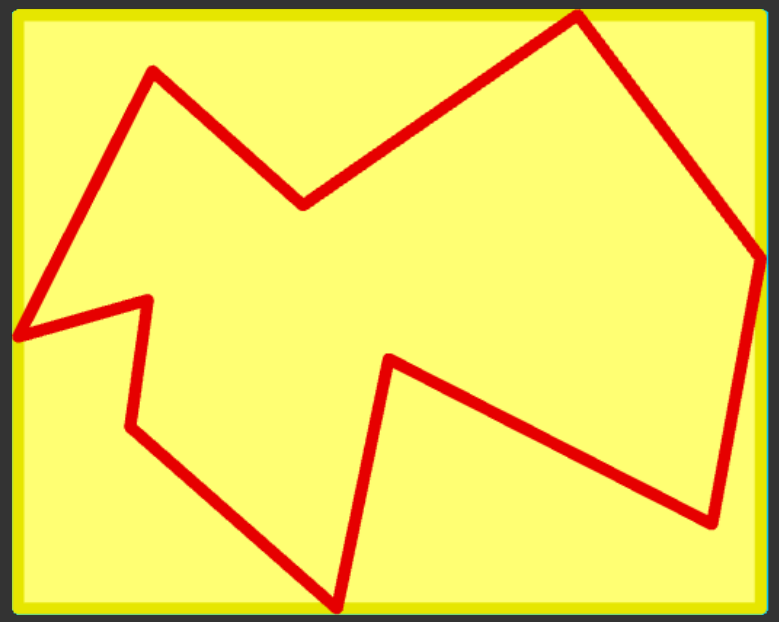
Convex Hull



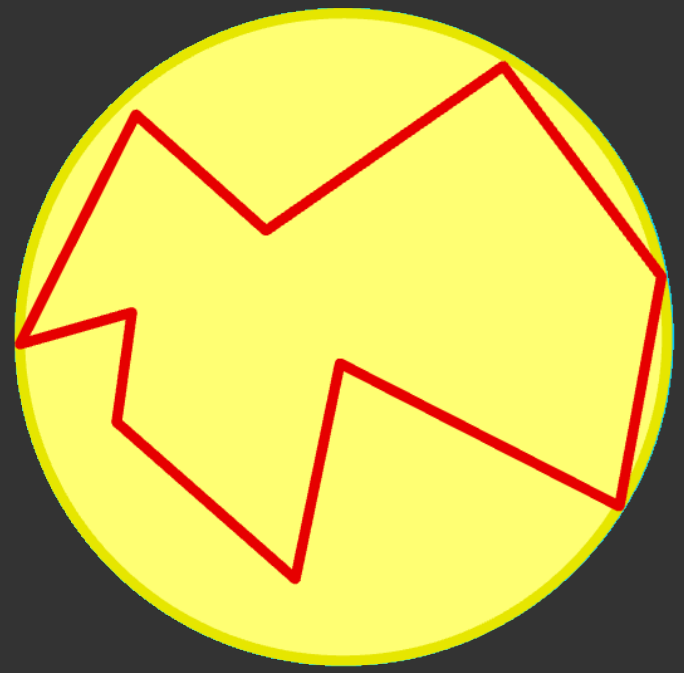
Rectangle by Area or Width

Minimum Bounding Polygon

- Creates a feature class containing polygons which represent a specified minimum bounding geometry enclosing each input feature or each group of input features



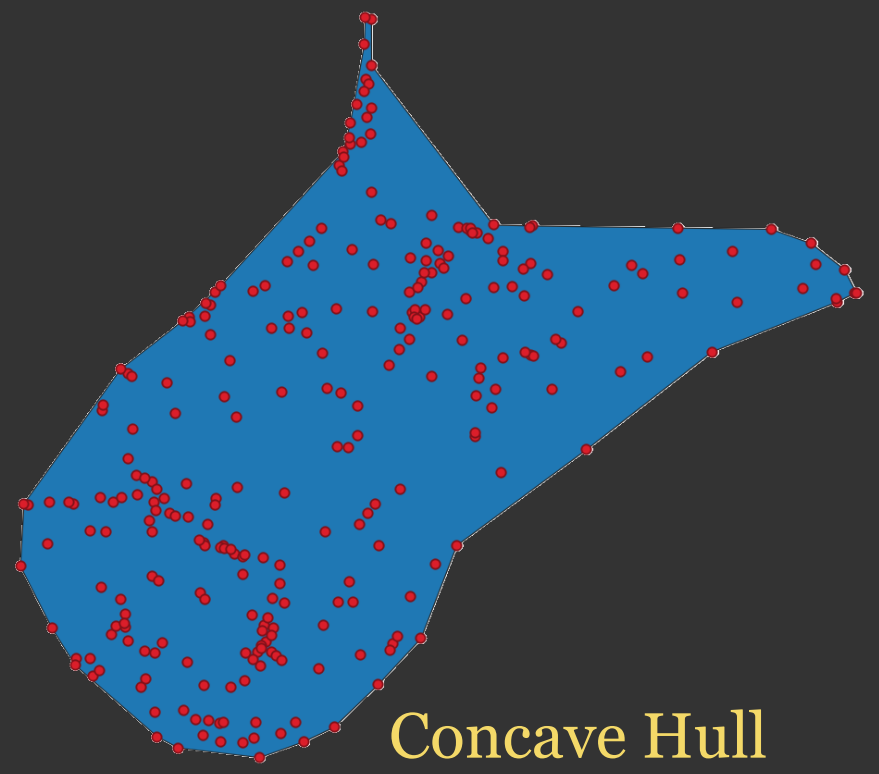
Envelope

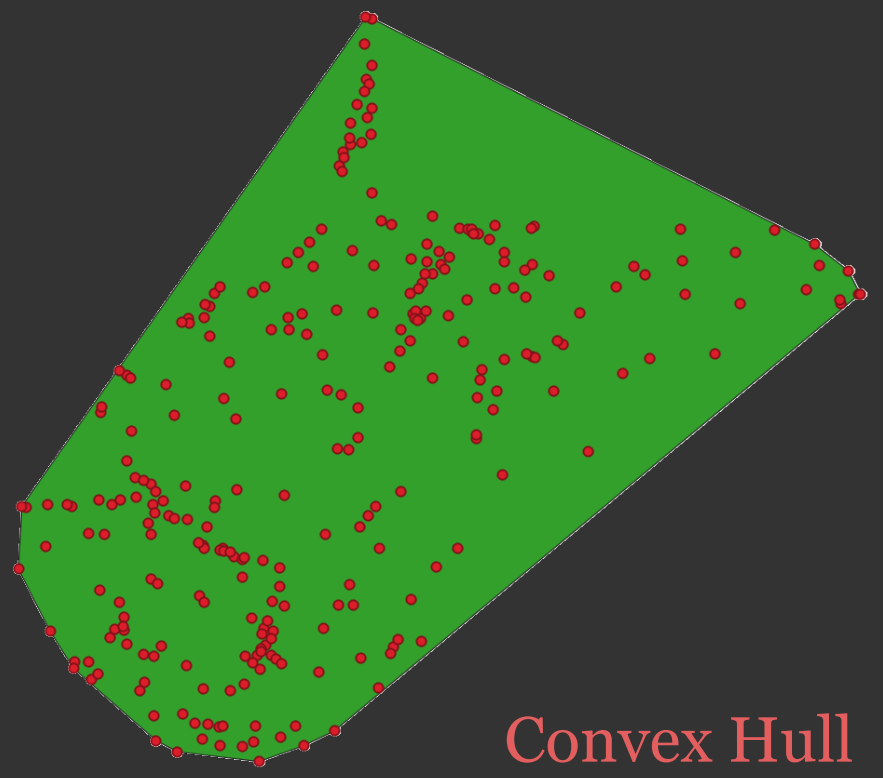


Circle

Convex Hull vs. Concave Hull

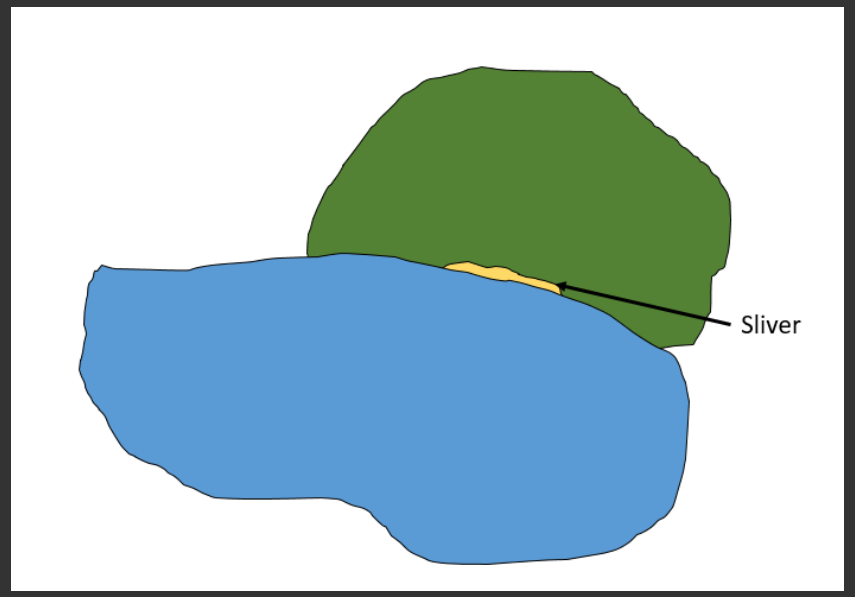
- Concave Hull = smaller than convex hull, can contain concave shapes
- Convex Hull = larger, cannot contain concave shapes





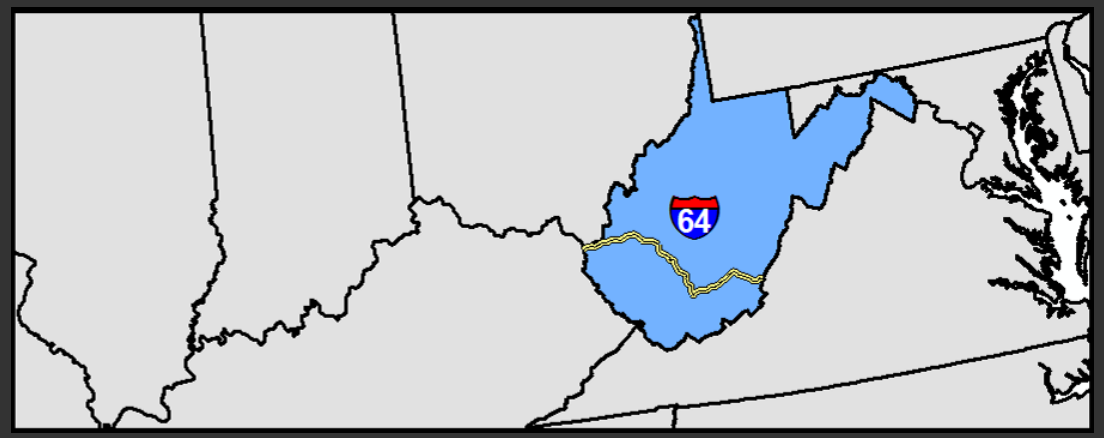
A Note on Slivers

- When there are issues with overlay
- Creates small residual areas of non-overlap
- Can remove features smaller than a certain size
- Can merge with adjacent polygons
- Consider minimal mapping unit (MMU)



A Note on Geometry

- Remember, that geometric fields are not updated with geoprocessing. Make sure geometry is correct before you use it.
- For example, remember to update length, area, perimeter, etc. after performing a clip.



Example from Mike and Jackie Strager

QGIS

Vector general
- Assign projection
- Convert layer to spatial bookmarks
- Convert spatial bookmarks to layer
- Create attribute index
- Create spatial index
- Define Shapefile projection
- Delete duplicate geometries
- Delete duplicates by attribute
- Detect dataset changes
- Drop geometries
- Execute SQL
- Extract selected features
- Extract Shapefile encoding
- Find projection
- Flatten relationship
- Join attributes by field value
- Join attributes by location
- Join attributes by location (summary)
- Join attributes by nearest
- Merge vector layers
- Order by expression
- Repair Shapefile
- Reproject layer
- Save vector features to file
- Set layer encoding
- Split features by character
- Split vector layer
- Truncate table

Vector geometry
- Add geometry attributes
- Affine transform
- Aggregate
- Boundary
- Bounding boxes
- Buffer
- Centroids
- Check validity
- Collect geometries
- Concave hull (alpha shapes)
- Concave hull (k-nearest neighbor)
- Convert geometry type
- Convert to curved geometries
- Convex hull
- Create layer from extent
- Create layer from point
- Create wedge buffers
- Delaunay triangulation
- Delete holes
- Densify by count
- Densify by interval
- Dissolve
- Drape (set Z value from raster)
- Drop M/Z values
- Eliminate selected polygons
- Explode lines
- Extend lines
- Extract M values
- Extract specific vertices
- Extract vertices
- Extract Z values
- Filter vertices by M value
- Filter vertices by Z value
- Fix geometries
- Force right-hand-rule
- Geodesic line split at antimeridian
- Geometry by expression
- Interpolate point on line
- Keep N biggest parts
- Line substring
- Lines to polygons
- Merge lines
- Minimum bounding geometry
- Minimum enclosing circles
- Multipart to singleparts
- Multi-ring buffer (constant distance)
- Offset lines
- Oriented minimum bounding box

Vector creation
- Array of offset (parallel) lines
- Array of translated features
- Create grid
- Create points layer from table
- Generate points (pixel centroids) along line
- Generate points (pixel centroids) inside polygons
- Import geotagged photos
- Points to path
- Random points along line
- Random points in extent
- Random points in layer bounds
- Random points in polygons
- Random points inside polygons
- Random points on lines
- Raster pixels to points
- Raster pixels to polygons
- Regular points

Vector overlay
- Clip
- Difference
- Extract/clip by extent
- Intersection
- Line intersections
- Split with lines
- Symmetrical difference
- Union

Vector analysis
- Basic statistics for fields
- Climb along line
- Count points in polygon
- DBSCAN clustering
- Distance matrix
- Distance to nearest hub (line to hub)
- Distance to nearest hub (points)
- Join by lines (hub lines)
- K-means clustering
- List unique values
- Mean coordinate(s)
- Nearest neighbour analysis
- Overlap analysis
- Statistics by categories
- Sum line lengths

Vector selection
- Extract by attribute
- Extract by expression
- Extract by location
- Random extract
- Random extract within subsets
- Random selection
- Random selection within subsets
- Select by attribute
- Select by expression
- Select by location

Network Analysis

Types of Networks

- Roads
- Streams
- Railroads
- Pipelines
- Electrical distribution

Geocoding

- Process of converting addresses in a table to coordinates on a map
- What is needed:
- Table of addresses
- Reference database
  - Street names
  - Address ranges
  - ZIP codes

Parsing = breaking an address up into its component parts

Address Standardization = setting up the components of an address in a regular format

Example: AVE vs. AV vs. Av vs Ave

98 Beechurst Avenue,	Morgantown,	WV	26506

Linear Referencing

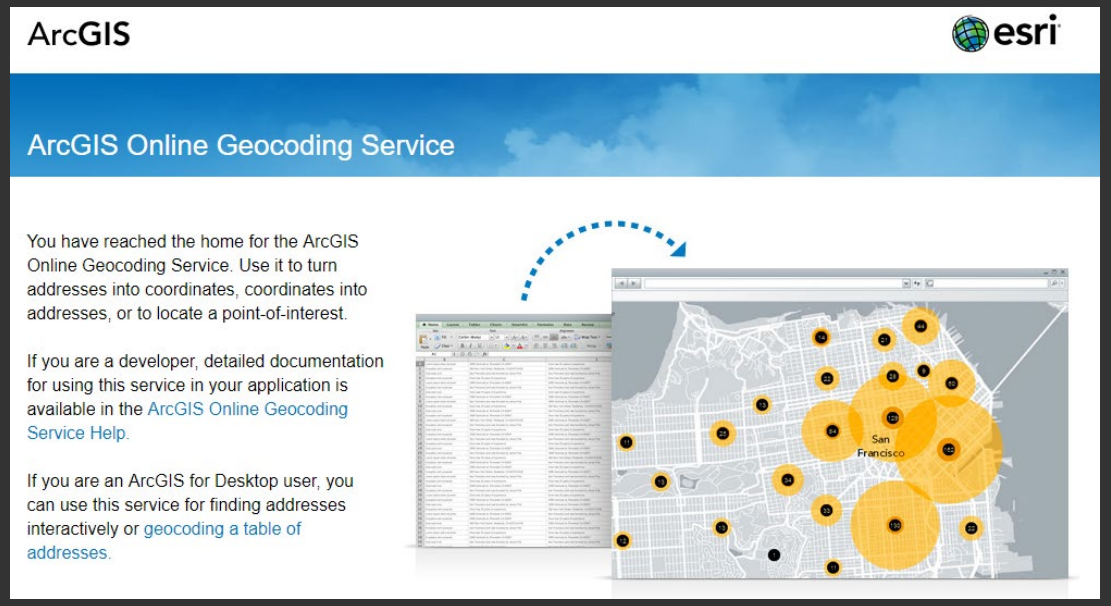
- Placing the address along a line segment based on where the address fits within the range of available addresses

- Example: If the address is 620 and the address range associated with the line segment is 600 to 700, the address will be located one-fifth of the street segment from 600 (Chang, 2012)

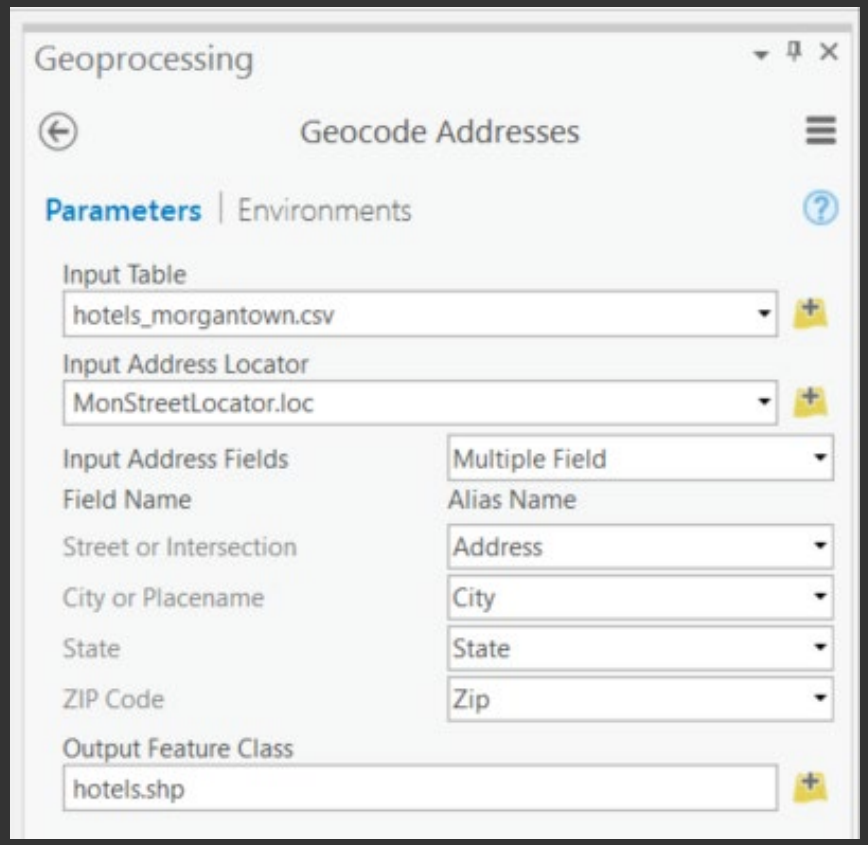
- You can also place addresses on the appropriate side of the line if you have the required information to do so

Address Locator

- Algorithm used to geocode addresses
- ESRI offers geocoding services



https://geocode.arcgis.com/arcgis/



West Virginia Geocoding

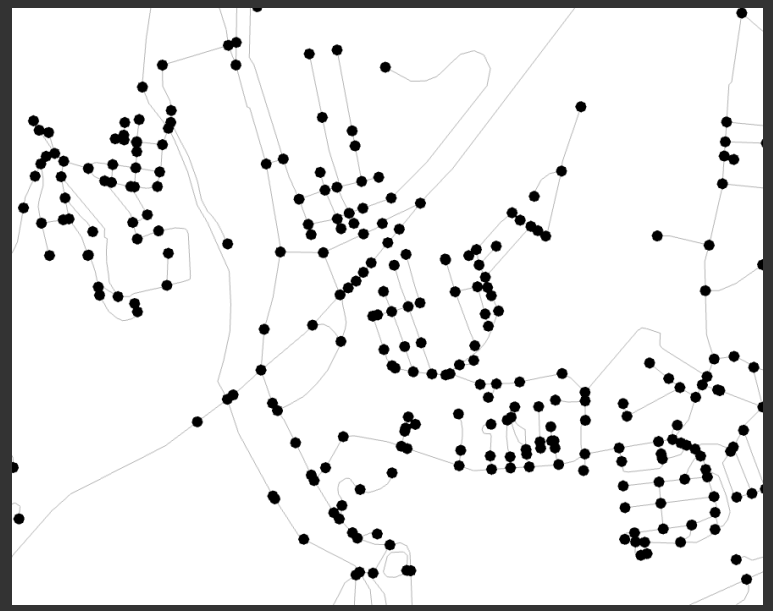
Address Geocoding

Name	Description	Projection	Date	Information	REST Services
Composite Geocoder	WV Street & Site Geocoder from Statewide Addressing & Mapping System (SAMS)	UTM 17N	2017	Link	Link
Site Geocoder	WV SAMS Site Locator Service	UTM 17N	2017	Link	Link
Street Geocoder	WV SAMS Street Locator Service	UTM 17	2017	Link	Link

https://www.mapwv.gov/gis_services.html

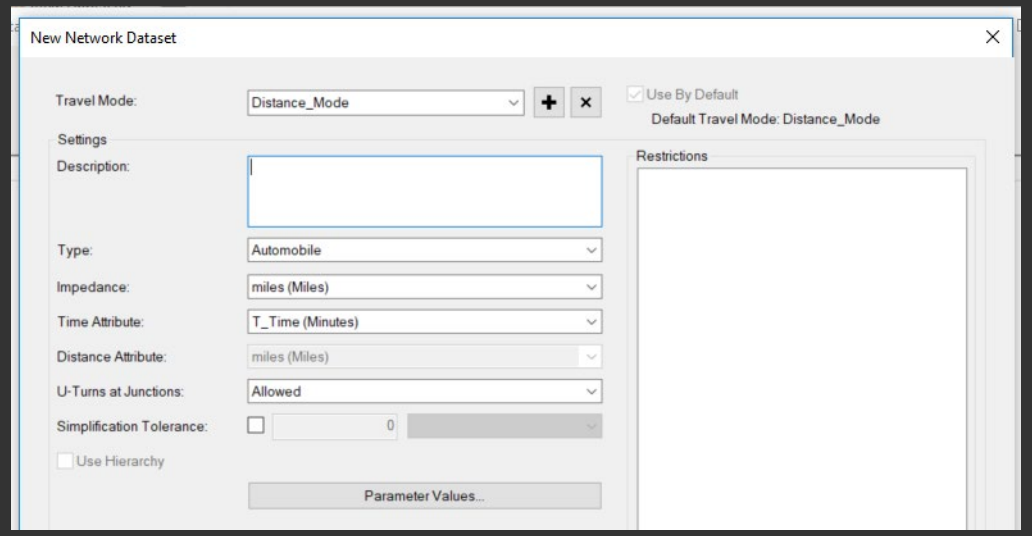
Components of a Network Dataset

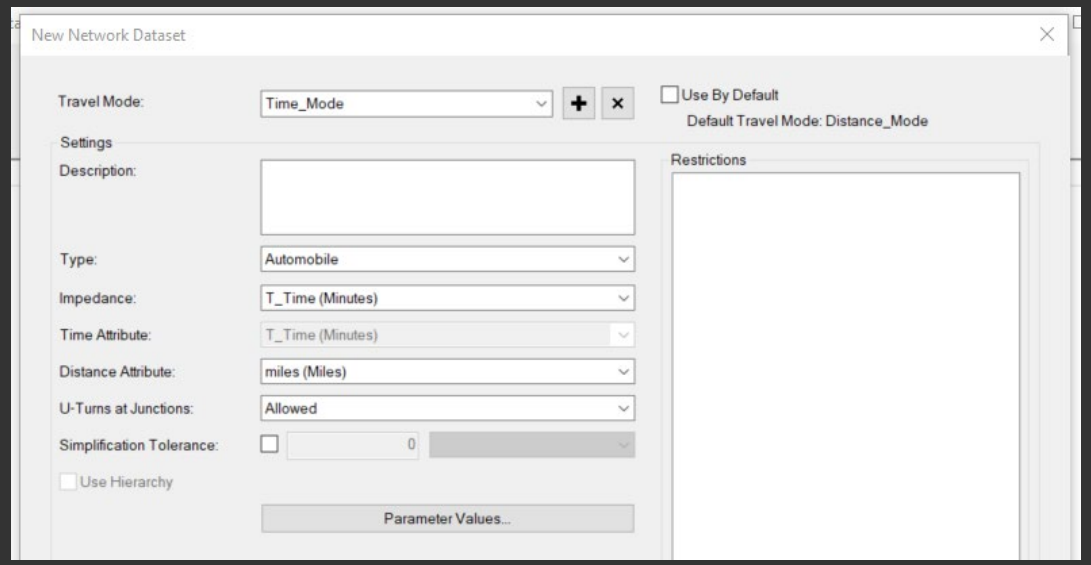
- Network = system of linear features that has the appropriate attributes for the flow of objects
- Topology = lines meet at intersections, lines cannot have gaps, lines have directions
- Junctions = where lines meet
- Edges = lines that connect the junctions



Impedance

- Cost of traveling between junctions on edges
- Can be defined based on travel distance or travel time
- If you now speed limit and distance, you can estimate travel time



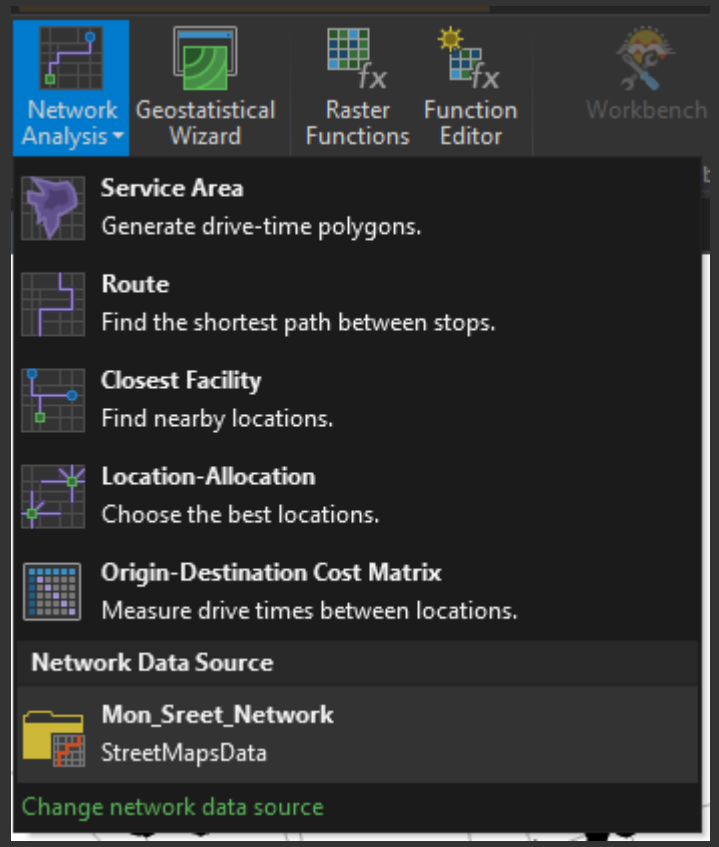


Complexities of Networks

- Incomplete data
- Outdated data
- Detours and road closures
- Changing traffic patterns
- Overpass and underpasses
- One-way streets
- Turning rules
- Barriers to movement

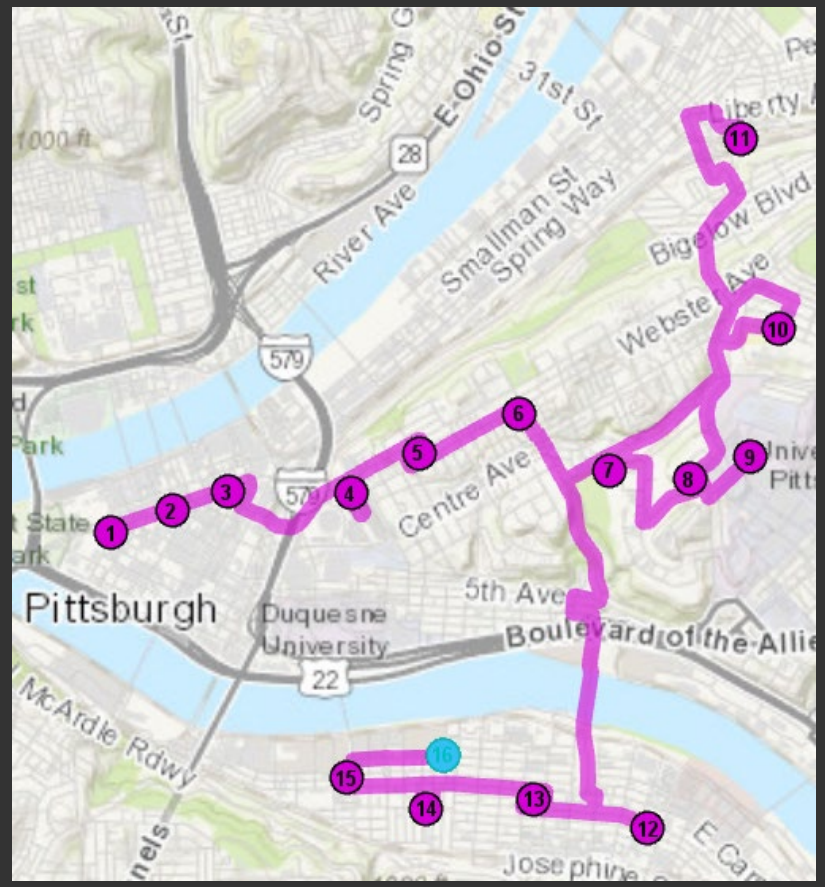
Types of Problems

- Shortest route between stops based on time
- Shortest route between stops based on distance
- Defining service areas based on travel time
- Defining service areas based on travel distance
- Find the closest facility to other locations
- Finding the best location at which to site a facility
- Travel costs between origins and destinations



Routes

- Must input stops
- Find the best (shortest) route based on time or distance
- Take into account the network rules and the impedance
- Can take into account: order of stops, type of impedance, and barriers

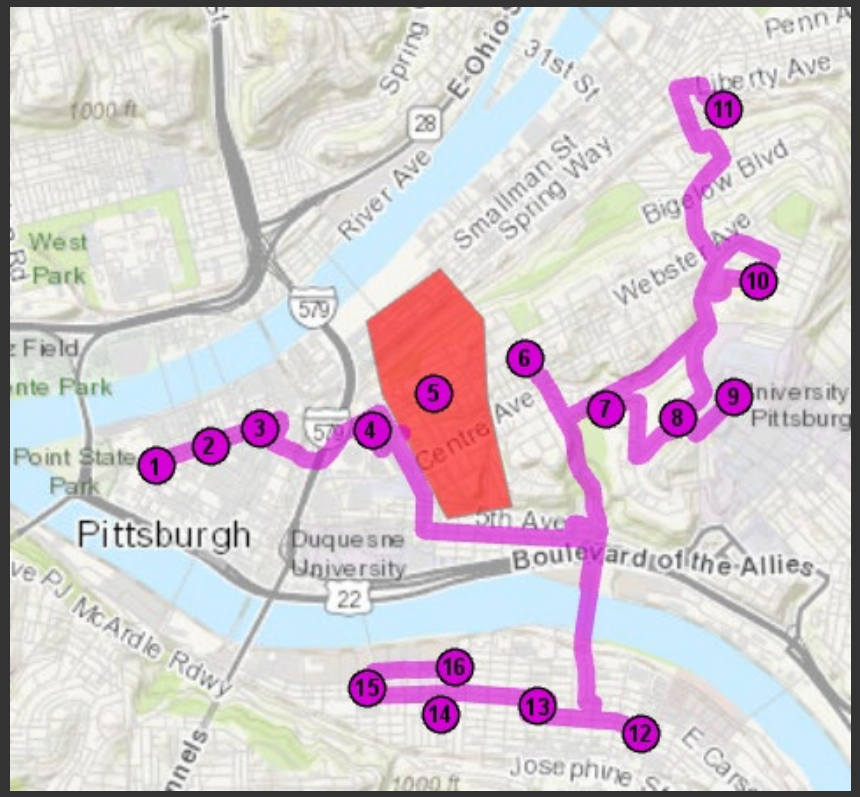


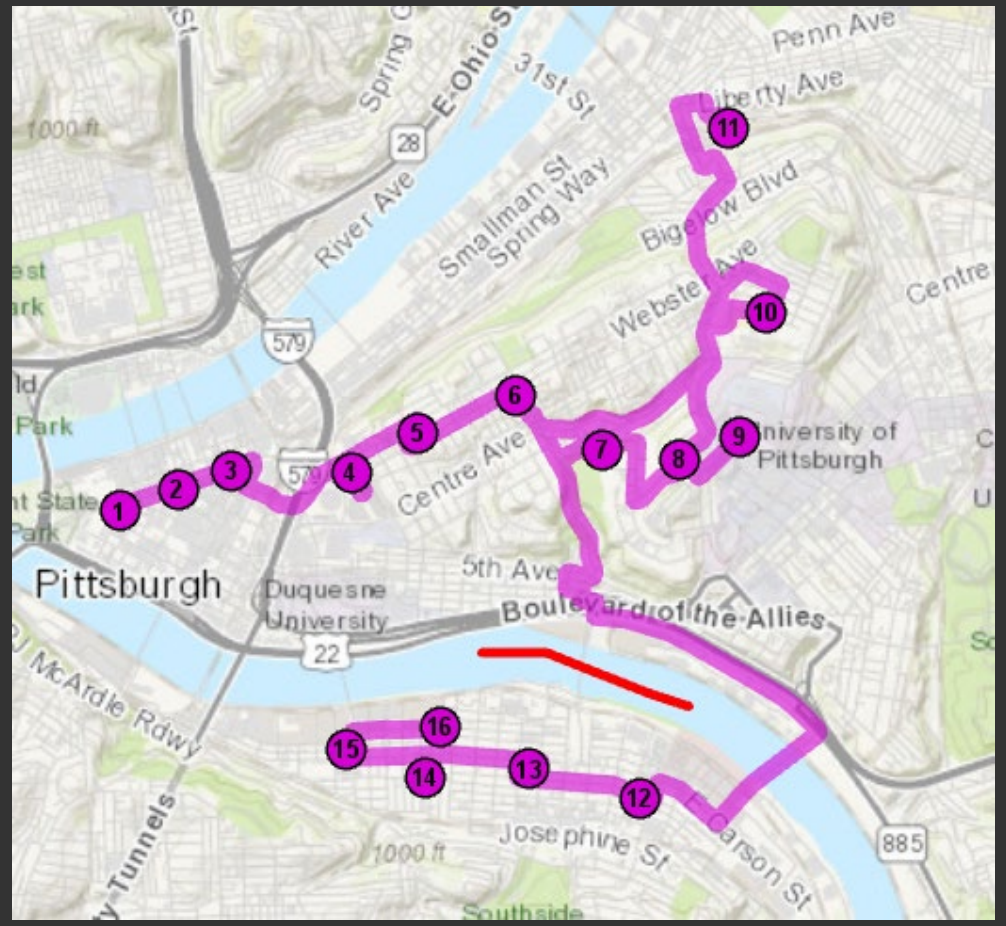
Route: Location 1 - Location 16

- Start at Location 1
- Go northeast on Stanwix St toward Forbes Ave
  117 ft ~ < 1 min
- Turn right on Liberty Ave
  0.2 mi ~ 2 min
- Arrive at Location 2, on the right
- Depart Location 2
- Continue east on Liberty Ave (PNC Plz)
  0.2 mi ~ 1 min
- At fork keep right on 7th Ave
  200 ft ~ < 1 min
- Turn left on Smithfield St
  118 ft ~ < 1 min
- Arrive at Location 3, on the left
- Depart Location 3
- Continue northeast on Smithfield St
  76 ft ~ < 1 min
- Turn right on Liberty Ave
  376 ft ~ < 1 min
- Turn right on William Penn Pl
  491 ft ~ < 1 min
- Turn left on 7th Ave
  502 ft ~ < 1 min

Barriers

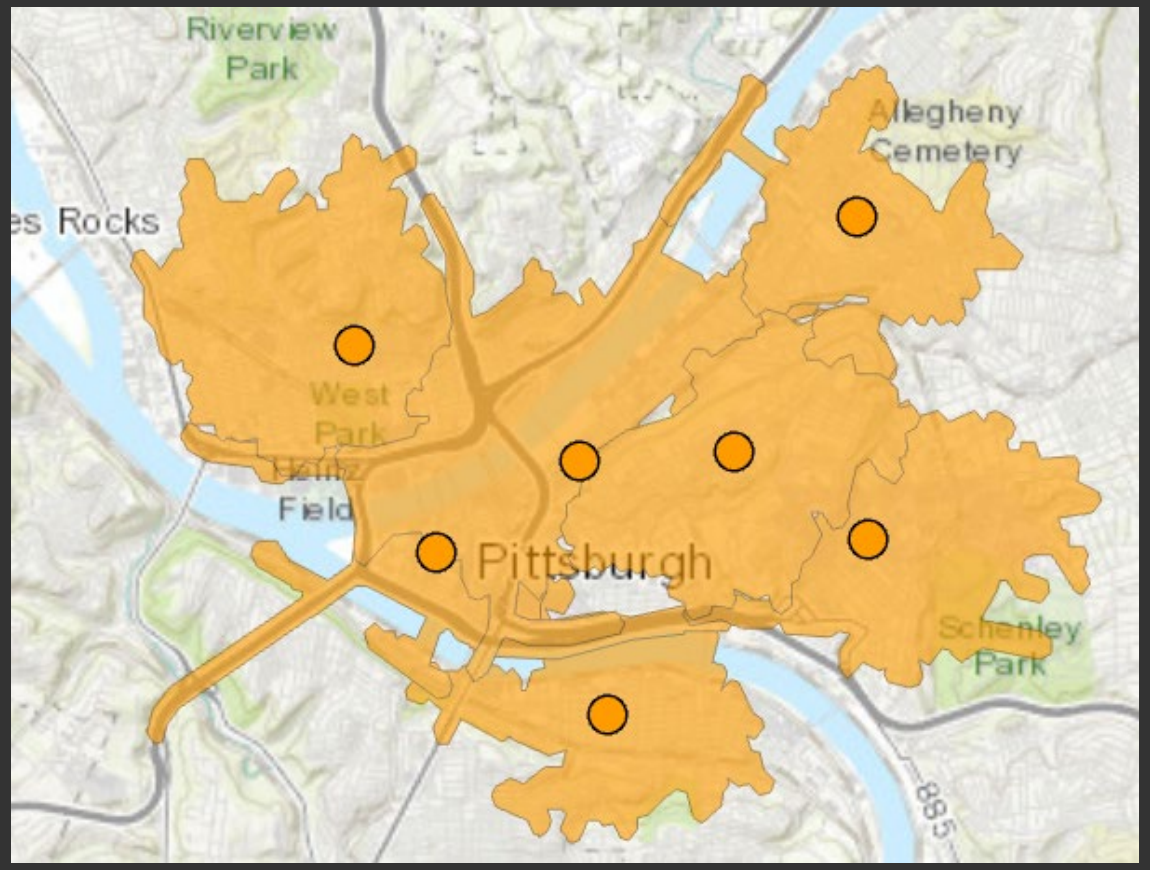
- Barriers can be defined as point, lines, or polygons





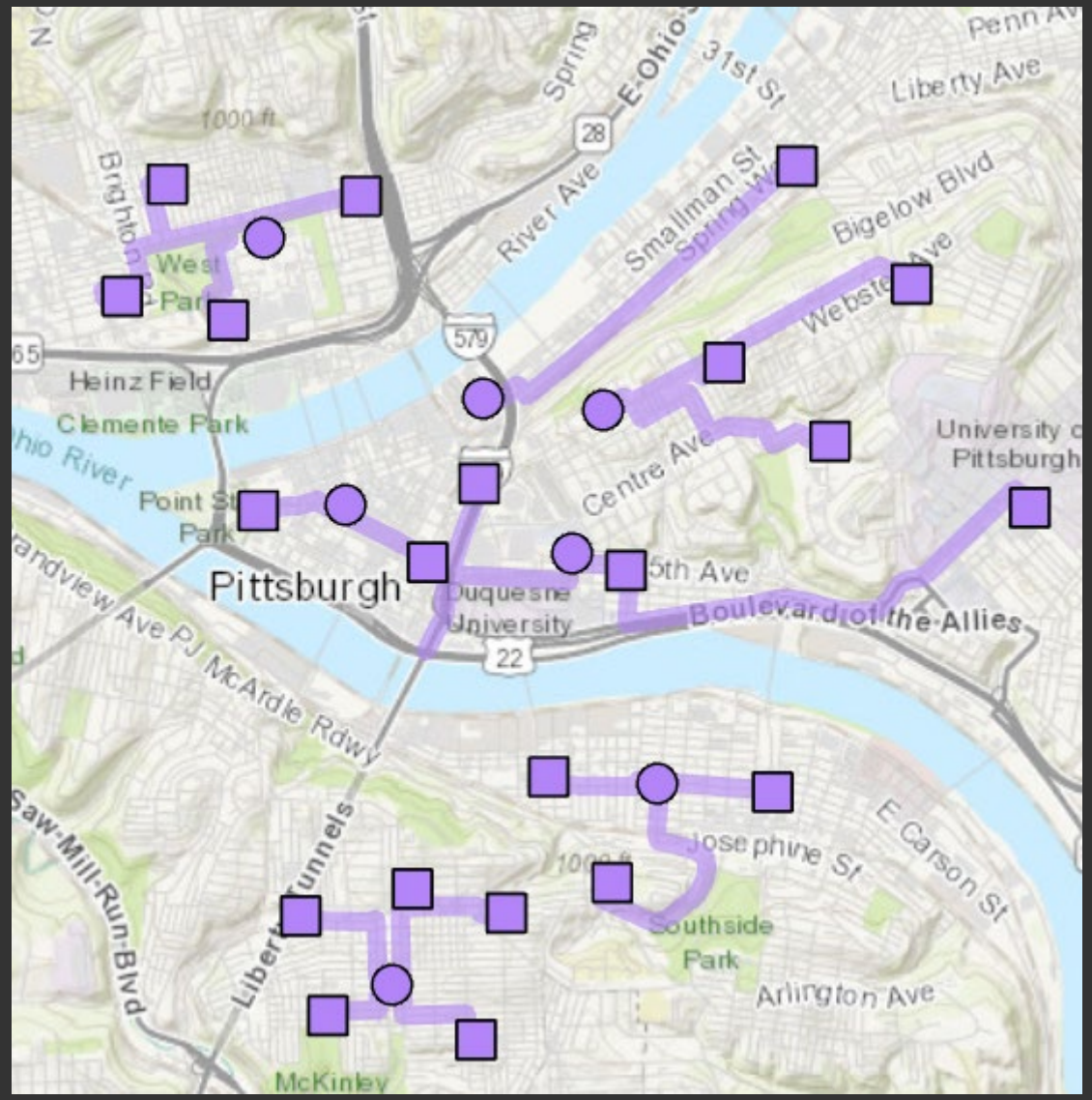
Service Areas

- Must input facilities
- Find areas that are within a certain distance or travel time from facilities
- Can specify direction of travel (to or from)
- Can solve for travel time or travel distance
- Can solve for multiple travel distances or times
- Can incorporate barriers



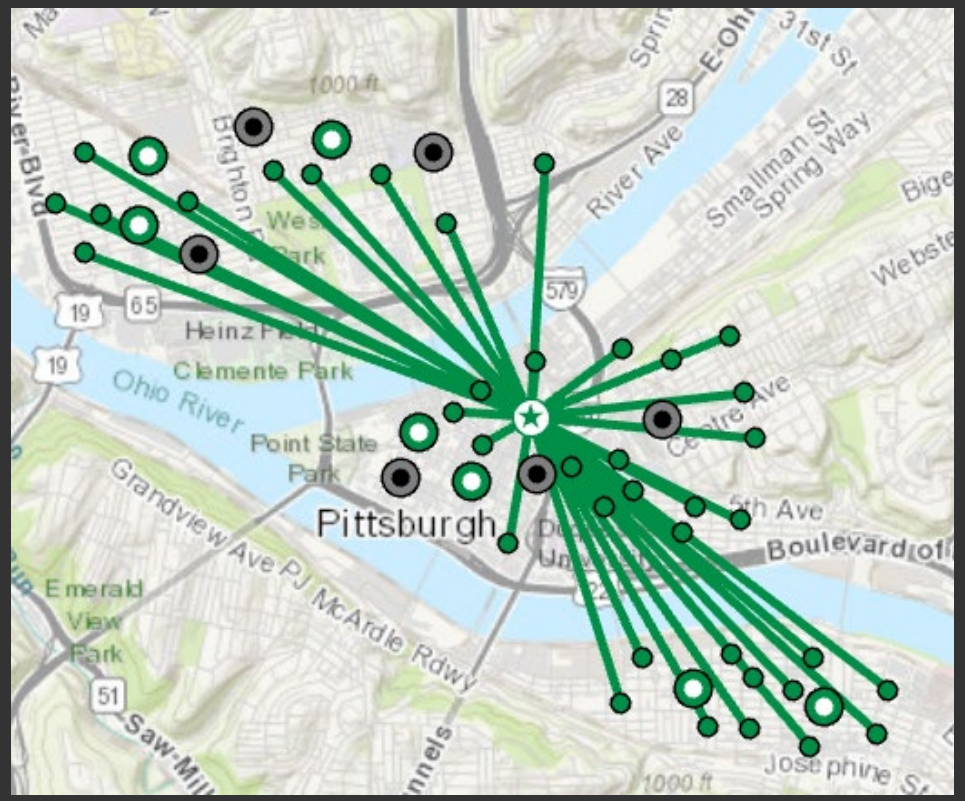
Closest Facility

- Must input facilities and incidents
- Find what facilities are closest to each incident based on travel time or travel distance
- Can incorporate barriers
- Can specify direction (to or from)



Location-Allocation

- Must provide potential facilities and demand points
- Find location that services the demand points most efficiently
- Can provide location of competitors
- Can solve for time or distance
- Can specify direction (to or from)
- Can incorporate barriers
- Many different methods to solve



Location-Allocation Problem Types

- Minimize Impedance = Best location is the one that provides the lowest cumulative impedance between demand locations
- Maximize Coverage = Facilities are located such that as many demand points as possible are allocated to solution facilities within the impedance cutoff
- Maximize Capacity Coverage = Facilities are located such that as many demand points as possible are allocated to solution facilities within the impedance cutoff; additionally, the weighted demand allocated to a facility can't exceed the facility's capacity
- Minimize Facilities = Facilities are located such that as many demand points as possible are allocated to solution facilities within the impedance cutoff; additionally, the number of facilities required to cover demand points is minimized

http://desktop.arcgis.com/en/arcmap/latest/extensions/network-analyst/location-allocation.htm#ESRI_SECTION3_F8EBCD830EDA464384FEC1E54DBA2707

Location-Allocation Problem Types

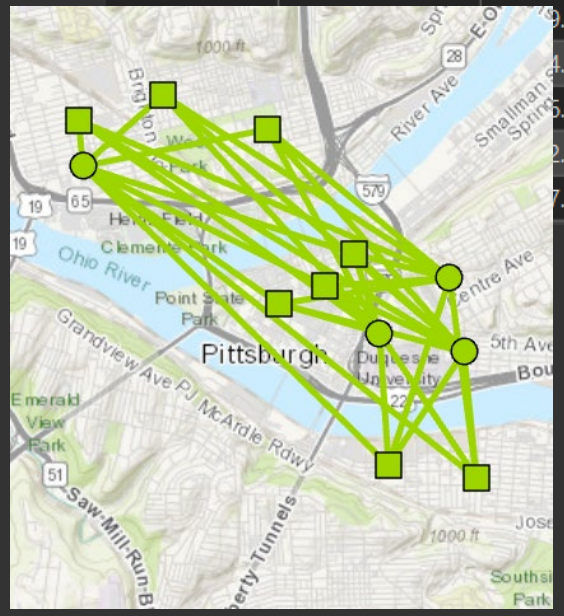
- Minimize Attendance = Facilities are chosen such that as much demand weight as possible is allocated to facilities while assuming the demand weight decreases in relation to the distance between the facility and the demand point
- Maximize Market Share = A specific number of facilities are chosen such that the allocated demand is maximized in the presence of competitors. The goal is to capture as much of the total market share as possible with a given number of facilities, which you specify. The total market share is the sum of all demand weight for valid demand points
- Target Market Share = Chooses the minimum number of facilities necessary to capture a specific percentage of the total market share in the presence of competitors. The total market share is the sum of all demand weight for valid demand points. You set the percent of the market share you want to reach and let the solver choose the fewest number of facilities necessary to meet that threshold

http://desktop.arcgis.com/en/arcmap/latest/extensions/network-analyst/location-allocation.htm#ESRI_SECTION3_F8EBCD830EDA464384FEC1E54DBA2707

Destination Cost Matrix

- Requires input destinations and origins
- Computes a matrix of cost of travel distance/time between all origins and destinations
- Can incorporate barriers

DestinationID	DestinationRank	Shape_Length	Total_Minutes	Total_TravelTime	Total_Miles	Total_Kilometers
8	1	1445.742157	<Null>	3.360779	<Null>	1.557115
1	2	1141.222499	<Null>	4.449583	<Null>	0.912656
2	3	786.808437	<Null>	4.624073	<Null>	0.957
7	4	1900.339831	<Null>	4.790759	<Null>	1.962175
3	5	903.573931	<Null>	4.840983	<Null>	1.109768
4	6	2532.706364	<Null>	8.929359	<Null>	2.322573
6	7	4009.961299	<Null>	9.403627	<Null>	4.214151
5	8	3505.266603	<Null>	10.930141	<Null>	4.723239
3	1	1073.110306	<Null>	4.751569	<Null>	1.426032
2	2	1361.803659	<Null>	5.290183	<Null>	1.466294
		9.165432	<Null>	6.658199	<Null>	1.713605
		4.830646	<Null>	7.126863	<Null>	2.711584
		5.981064	<Null>	8.556843	<Null>	3.116644
		2.683899	<Null>	9.141442	<Null>	2.733663
		7.324386	<Null>	11.623727	<Null>	4.944844



This is the end of this lecture module.

Please return to the West Virginia View Webpage for additional content.