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

GIScience

Digitizing, Georeferencing, and Resampling

WV VIEW

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

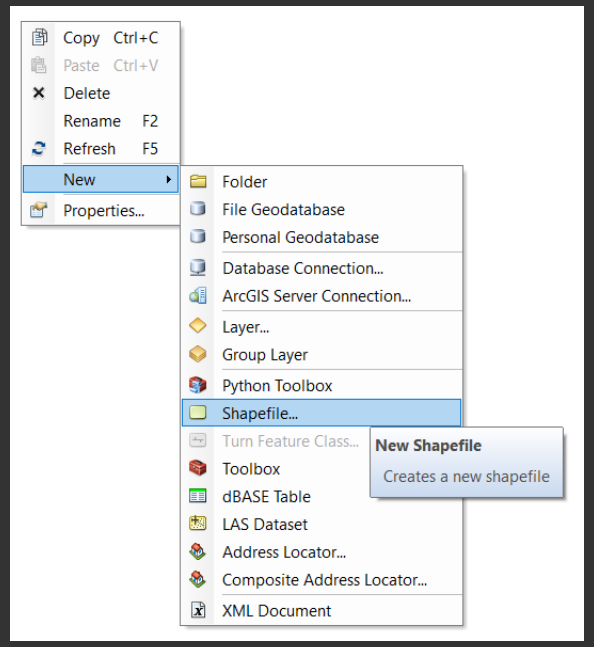
Ways to Create New Spatial Data

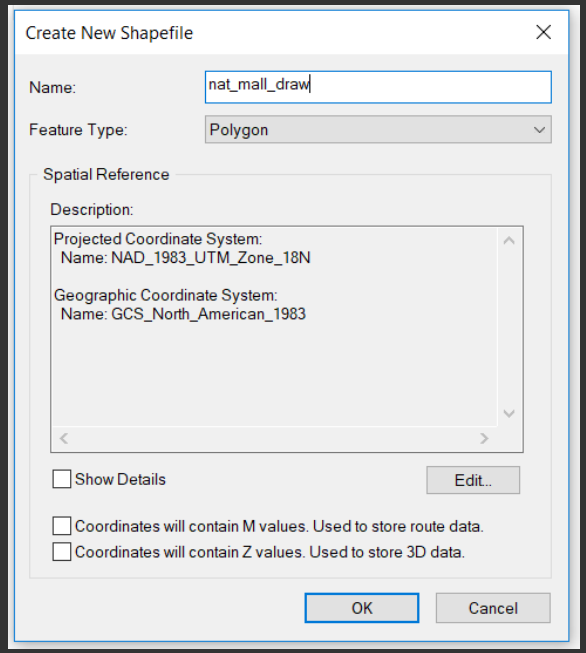
1. Field collection with GNSS
2. Collecting and georeferencing sensor-based data
   - Aerial imagery
   - Satellite imagery
   - Radar
   - LiDAR
3. Scanning and georeferencing paper maps
4. Heads-up digitizing
5. Automated extraction of features
   - Unsupervised classification
   - Supervised classification
   - Feature extraction

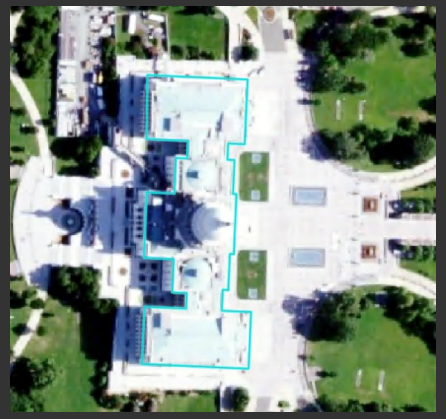
Digitizing

Digitizing Steps

1. Obtain input imagery or reference data
2. Create a new file to edit into
   - Define geometry (point, lines, or polygon)
   - Define projection
   - Create fields
3. Create attribute columns
4. Draw
5. Edit attribute information
6. Check and validate your results

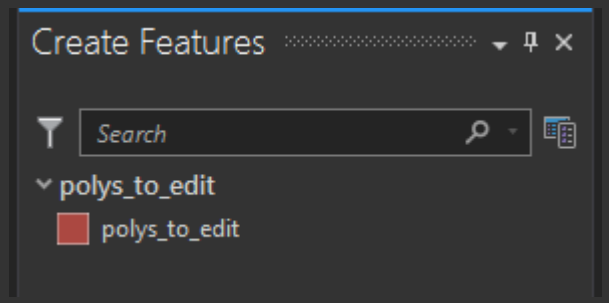


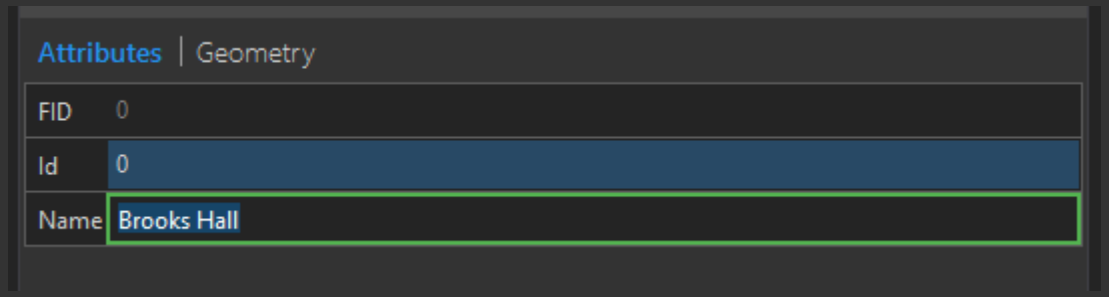


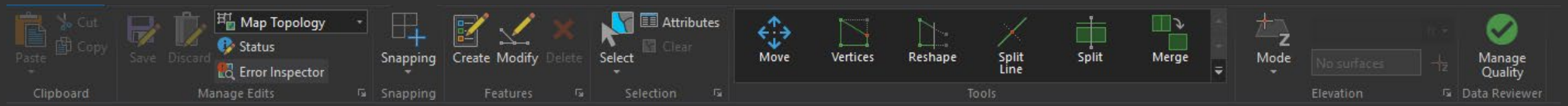


Digitizing in ArcGIS Pro

- Use Edit tab
- Remember to consider topology
- Remember to edit attributes
- Save often!!!!!!!







Adding Fields

- Make sure to define the correct type

Visible	Read Only	Field Name	Alias	Data Type	Allow NULL	Highlight	Number Format	Default	Precision	Scale	Length
✓	✓	FID	FID	Object ID			Numeric		0	0	
✓		Shape	Shape	Geometry					0	0	
✓		Id	Id	Long			Numeric		6	0	
✓		name		Text							255

Attributes | Geometry

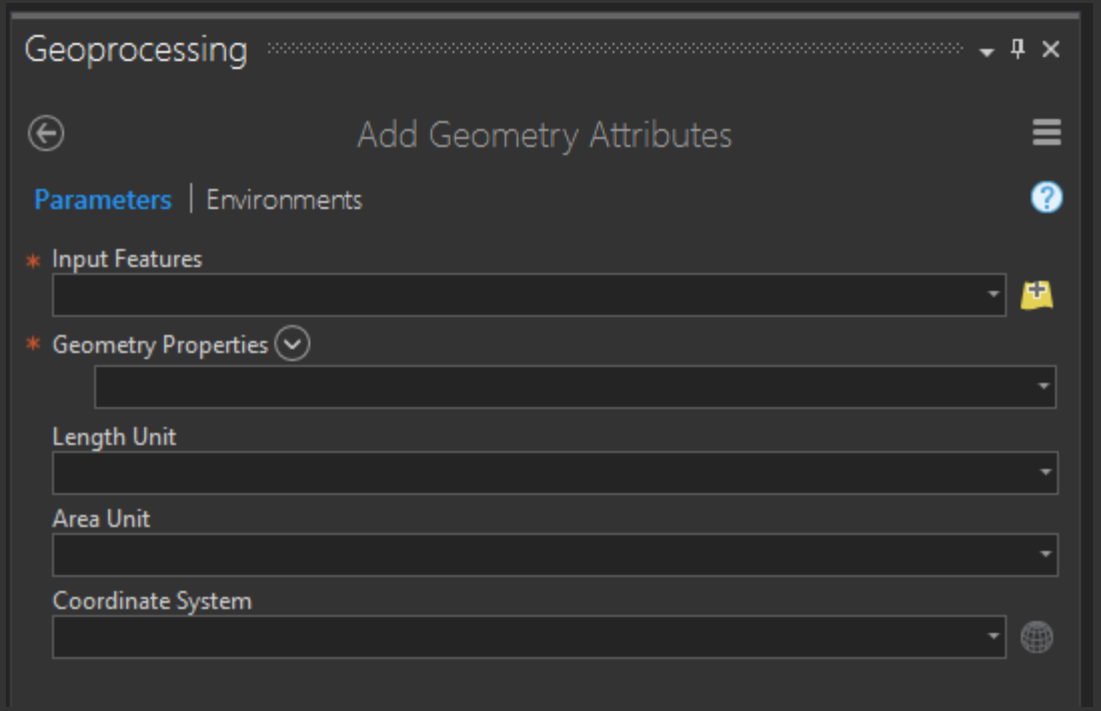
FID	0
Id	0
Name	Brooks Hall

Data Type	Description
Short Integer	Integer values between -32,768 and 32,767
Long Integer	Integer values between -2,147,483,648 and 2,147,483,647
Float	Decimal values with 1-6 decimal places
Double	Decimal values with more the 6 decimal places
Text	Text strings
Date	Data info (mm/dd/yyyy)

Add Geometric Fields

- Use Add Geometry Attributes Tool

Points	Lines	Polygons
X-coordinate	Start-X	Area
Y-coordinate	Start-Y	Perimeter
Z-coordinate (if z-enabled)	End-X	Centroid-X
M-coordinate (if m-enabled)	End-Y	Centroid-Y
	Mid-X	
	Mid-Y	
	Length	
	Bearing	

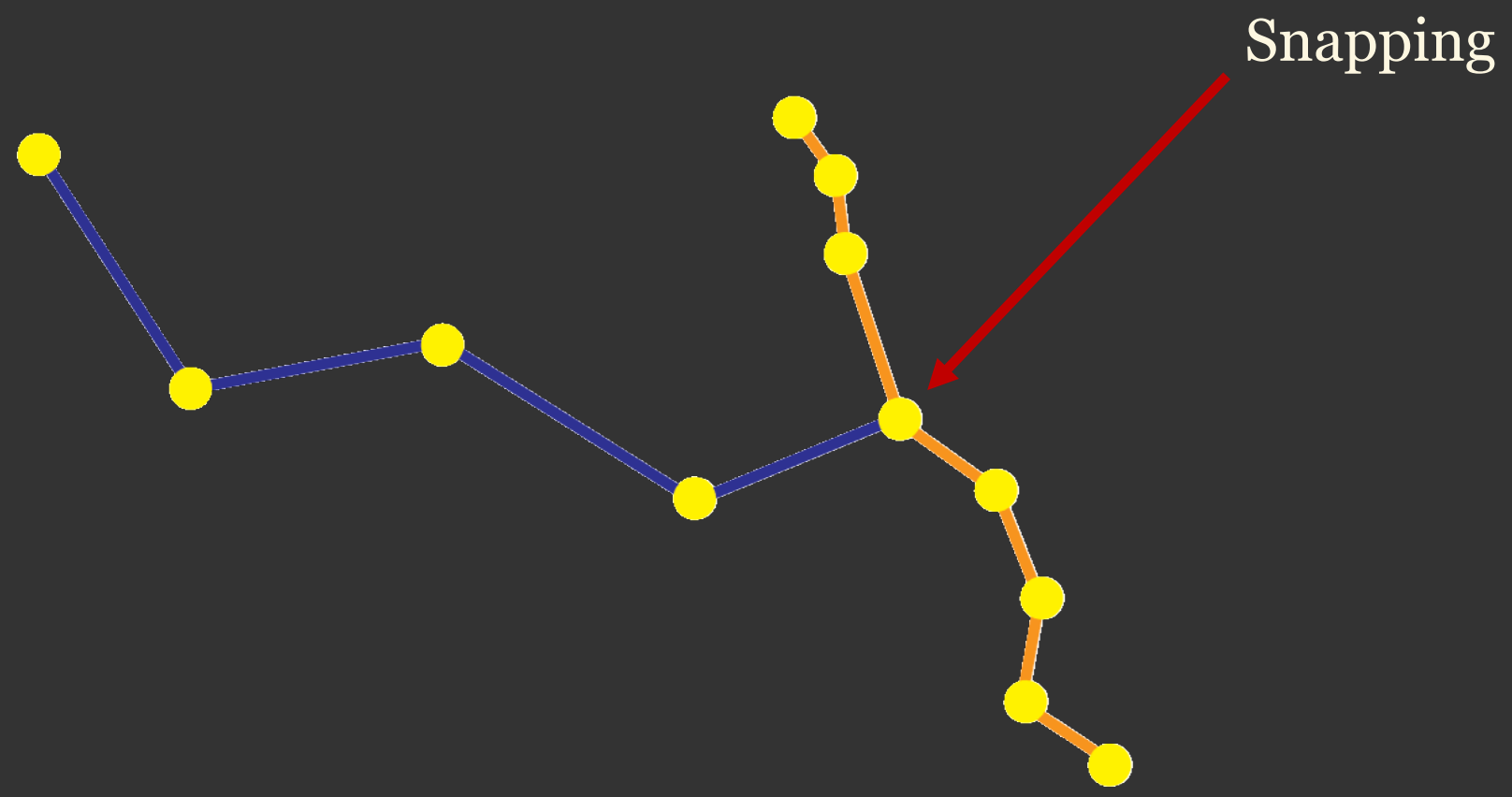


Impacts on Digitizing Accuracy/Precision

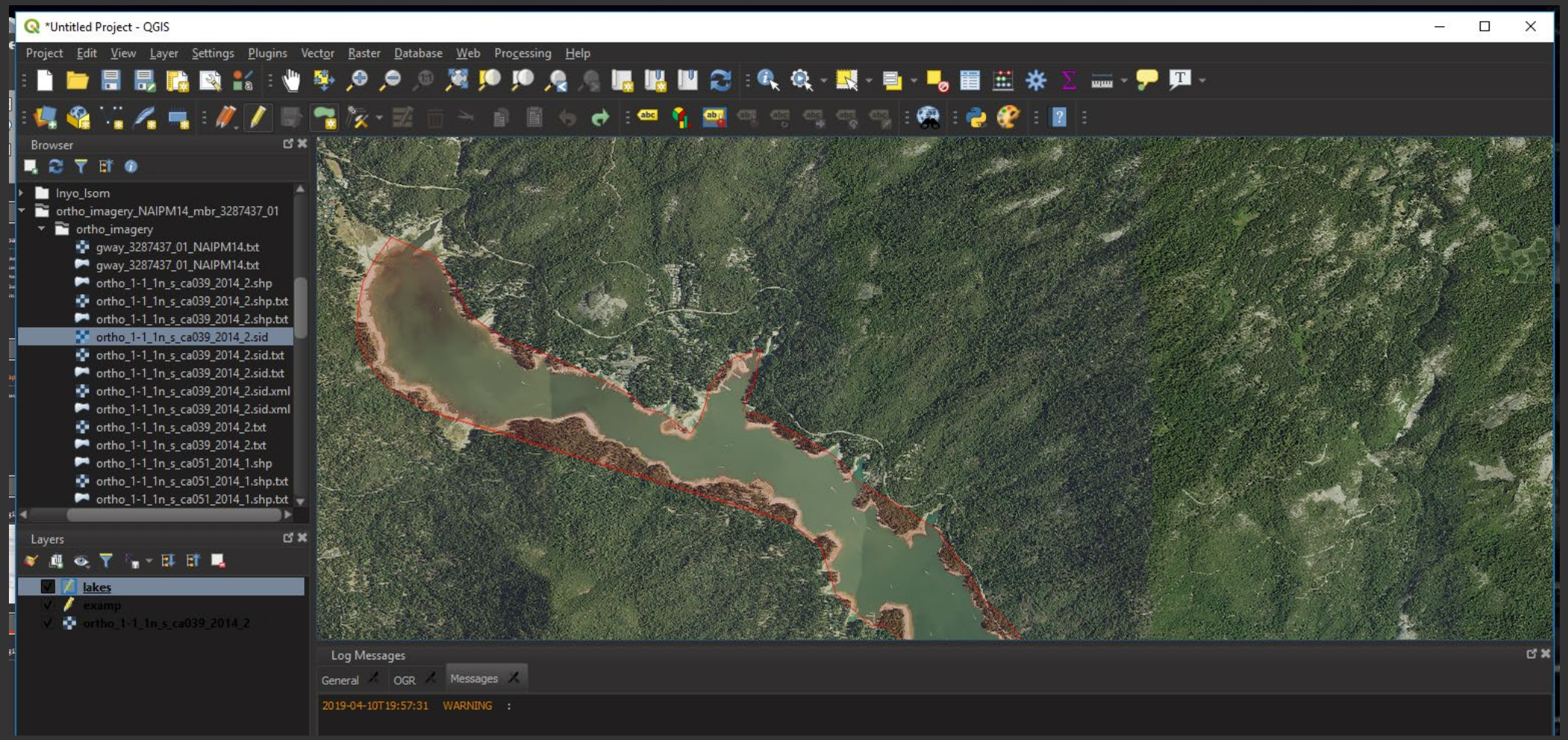
1. User skill, precision, and patience
2. Ease of human interpretation
3. Precision of equipment
4. Quality of reference layers
5. Etc.

Snapping

- Causes/forces vertices to match up

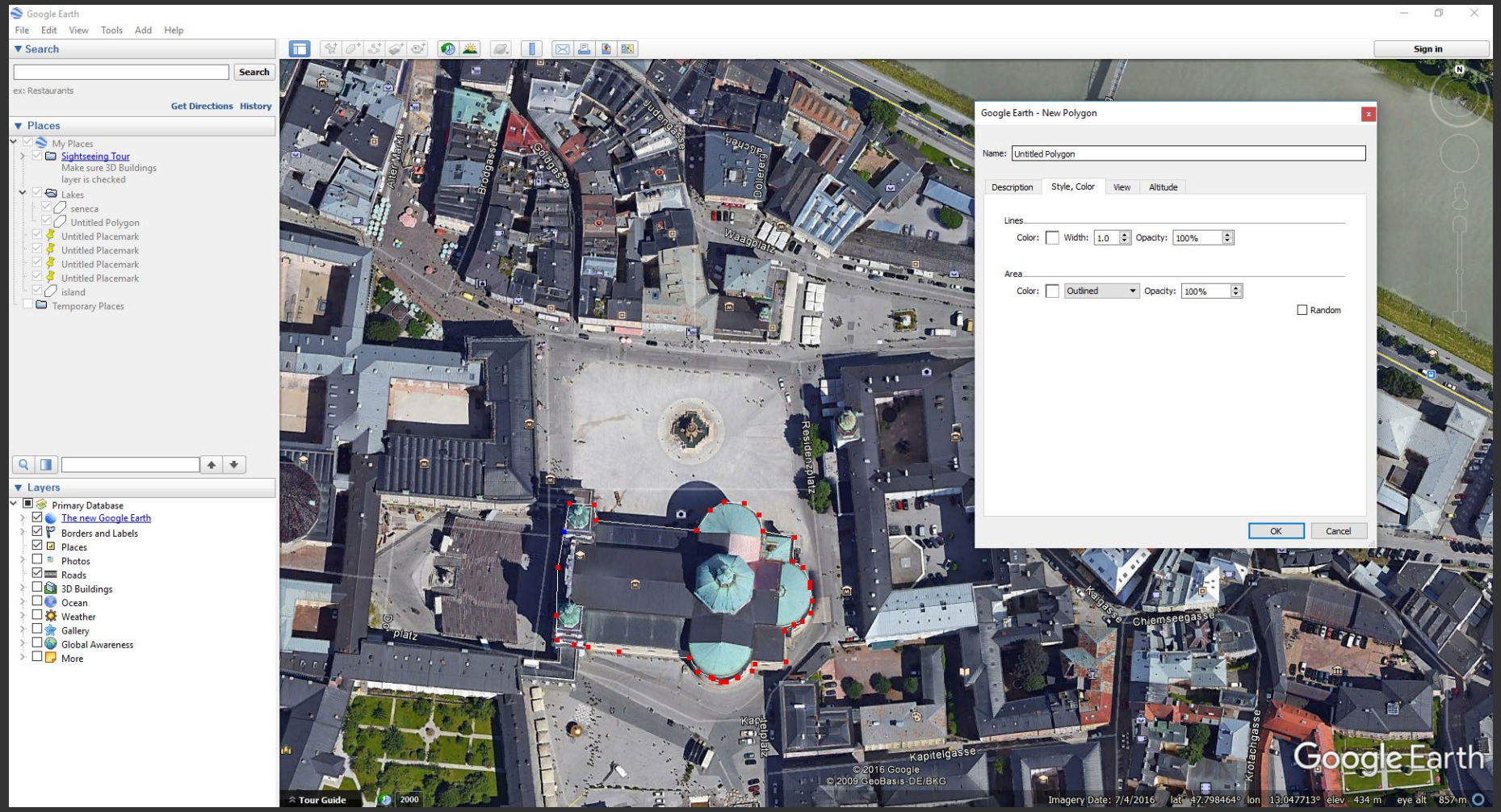


Digitizing Outside of ArcGIS (QGIS)



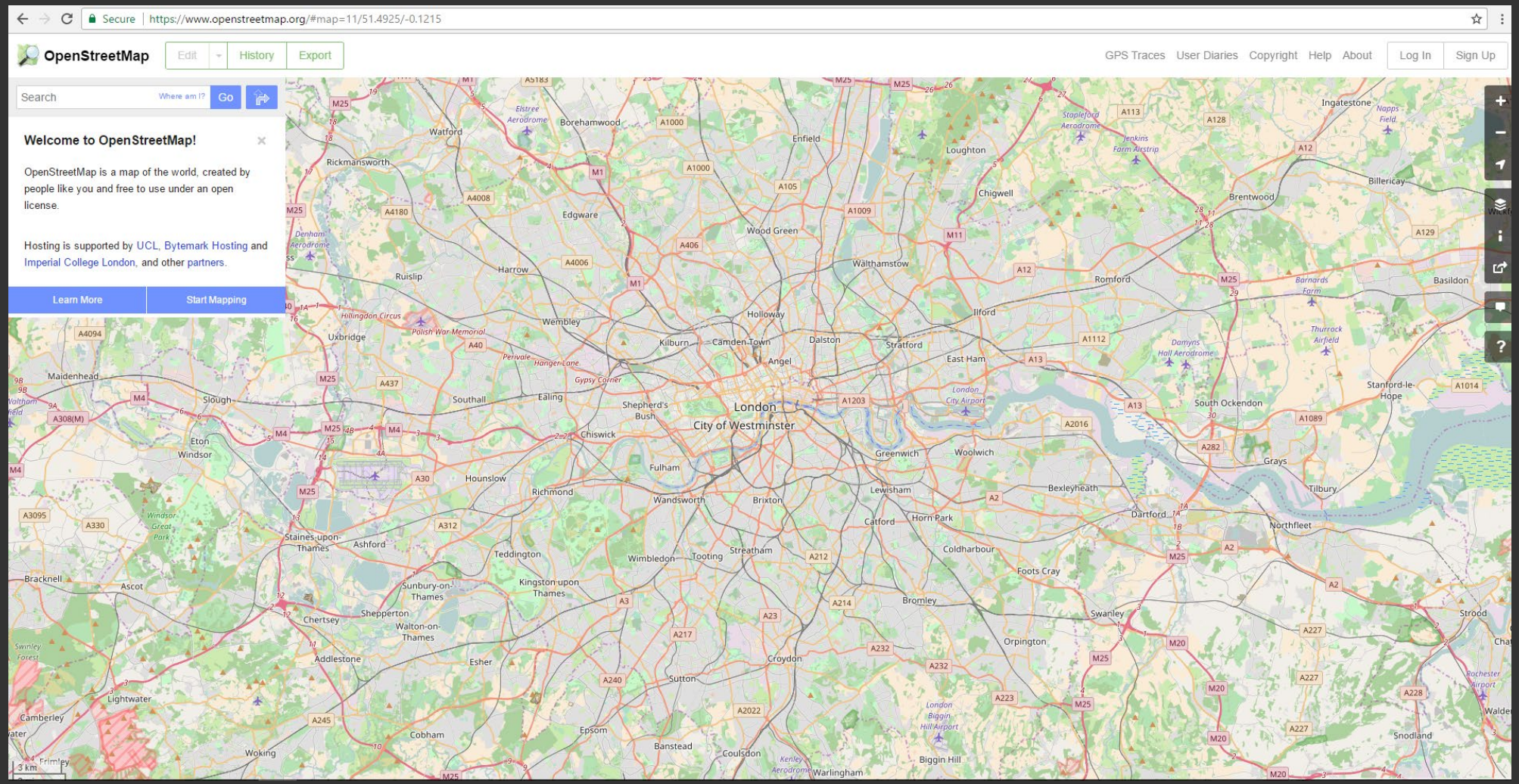
Screen capture from QGIS

Digitizing Outside of ArcGIS (Google Earth)



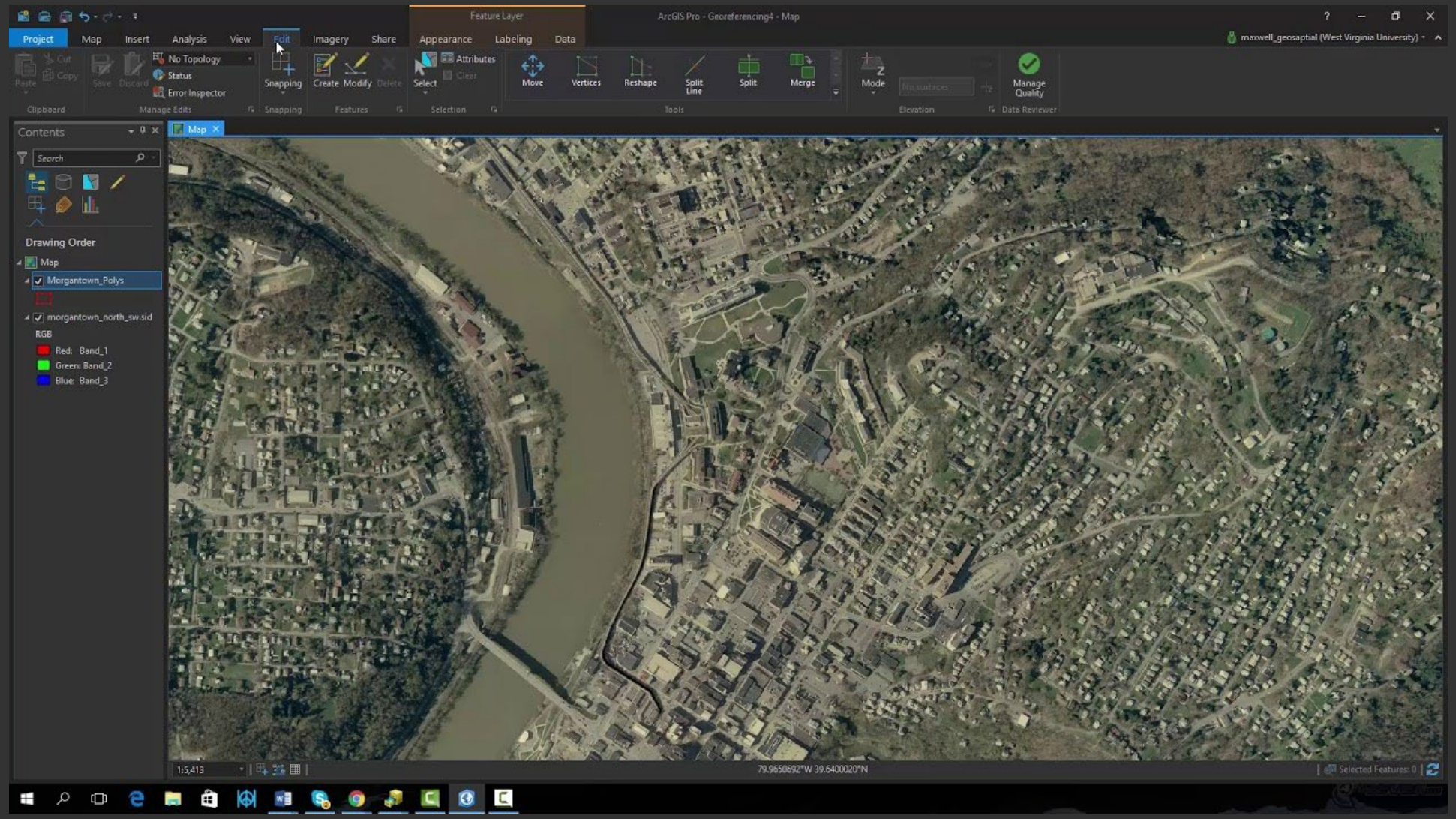
Screen capture from Google Earth Pro

Digitizing Outside of ArcGIS (OpenStreetMap)



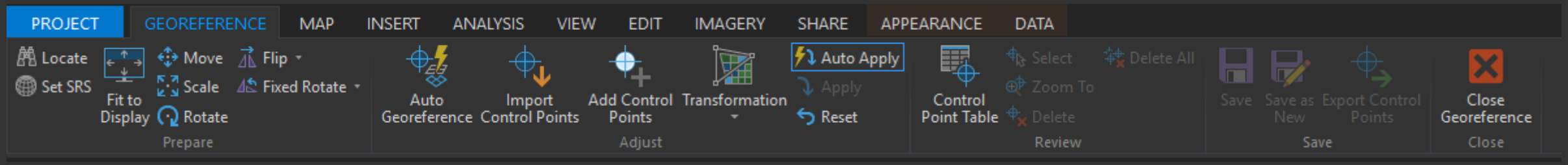
Screen capture from OpenStreetMap

Video: Digitizing in ArcGIS Pro



Georeferencing

Georeferencing



- Process of aligning an unreferenced dataset with one that has spatial reference information

Scanning

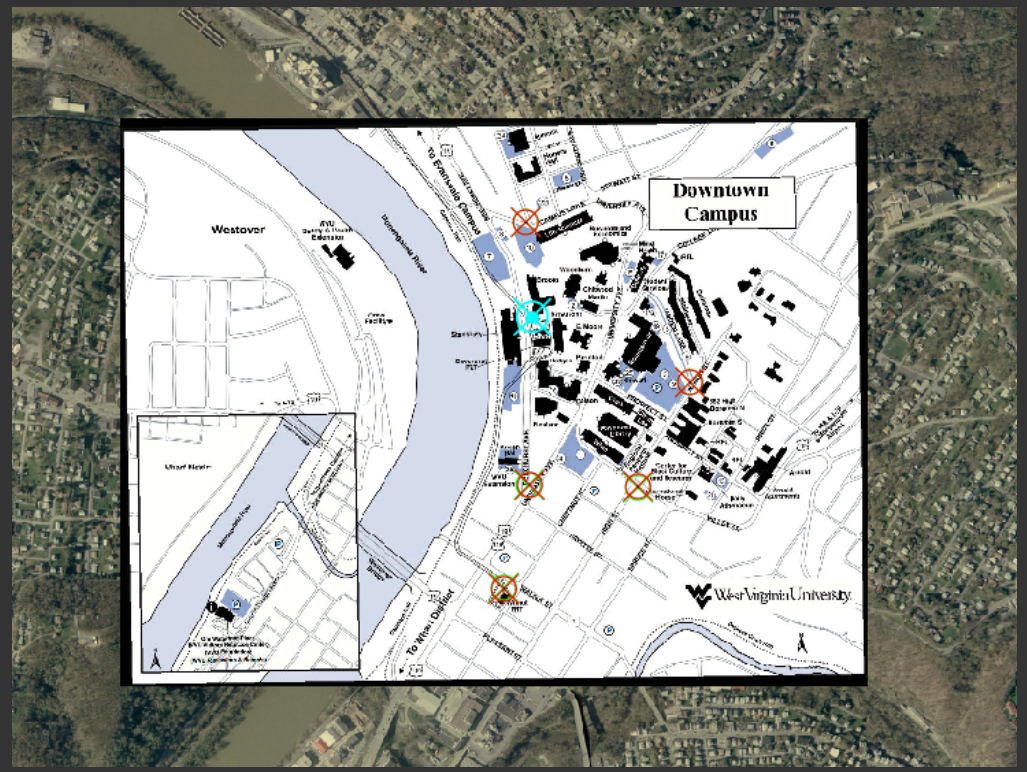
Convert analog map to digital raster data

Scanning

- Convert analog map to digital raster data
- Resolution and quality depends on dots per inch (dpi)
- Can be converted back to vector data via tracing

Control Points

- Used to register dataset to known location



WVU: DowntownCampusMap.jpg

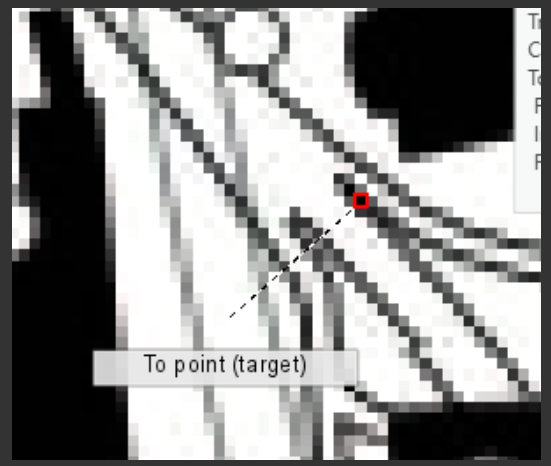
Zero Order Polynomial (Only Shift)

Link	Source X	Source Y	X Map	YMap	Residual X	Residual Y	Residual
2	642.095046	-304.056601	589,558.042166	4,387,806.527743	588,915.947120	4,388,110.584344	4,427,452.596381
3	591.781440	-727.549348	589,497.725870	4,387,231.508718	588,905.944430	4,387,959.058066	4,427,301.086062
4	889.943736	-409.666139	589,892.546175	4,387,669.548631	589,002.602439	4,388,079.214769	4,427,433.032895
5	634.244094	-565.724327	589,550.818827	4,387,450.899573	588,916.574733	4,388,016.623899	4,427,359.554363
6	636.229751	-155.838320	589,543.704963	4,388,008.172735	588,907.475212	4,388,164.011055	4,427,504.421486
7	803.933326	-570.431343	589,781.692548	4,387,445.060220	588,977.759223	4,388,015.491563	4,427,366.571118

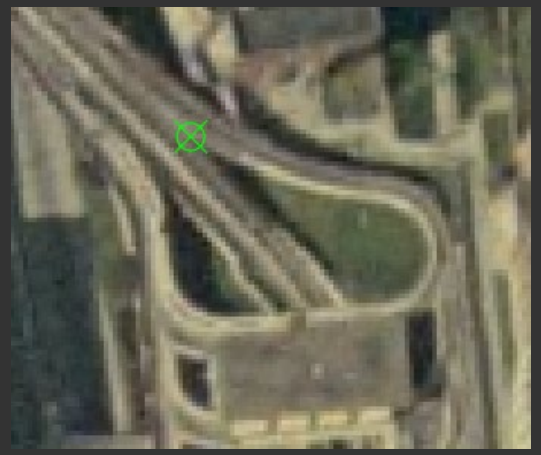
Control Points

Criteria for control points

1. Spread evenly around study area
2. Enough points to use the algorithm of choice
3. Use highest accuracy dataset possible
4. Use locations identifiable on both datasets (known & unknown)



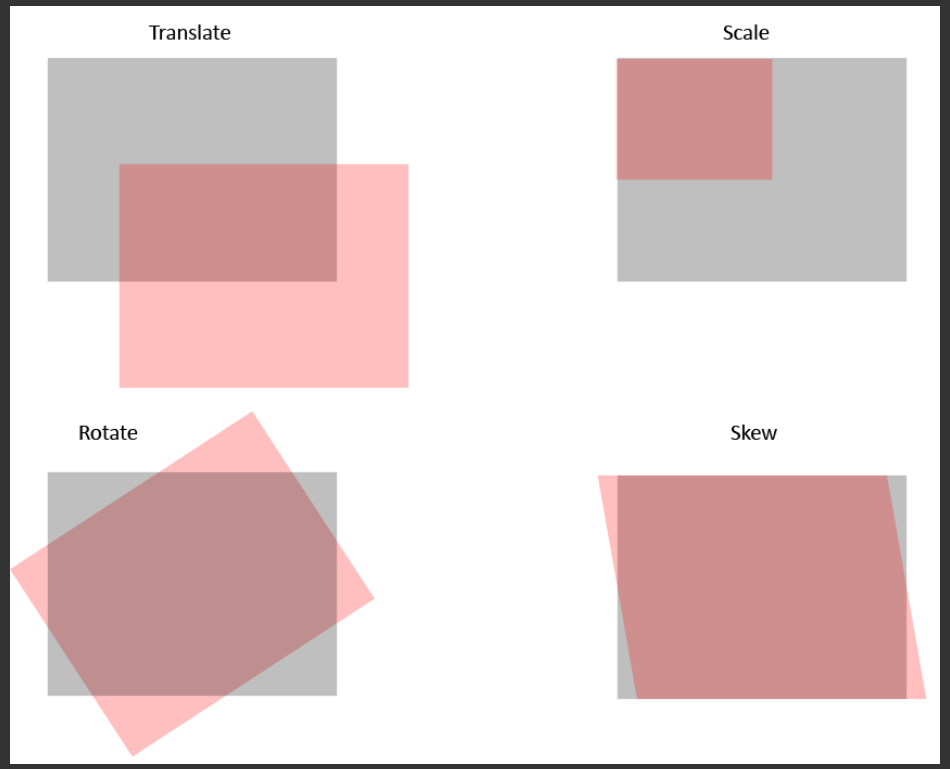
- From Point = location on the data being referenced
- To Point = same location on the reference data



Transforming the Data

In order to georeferenced data, it may need to be:

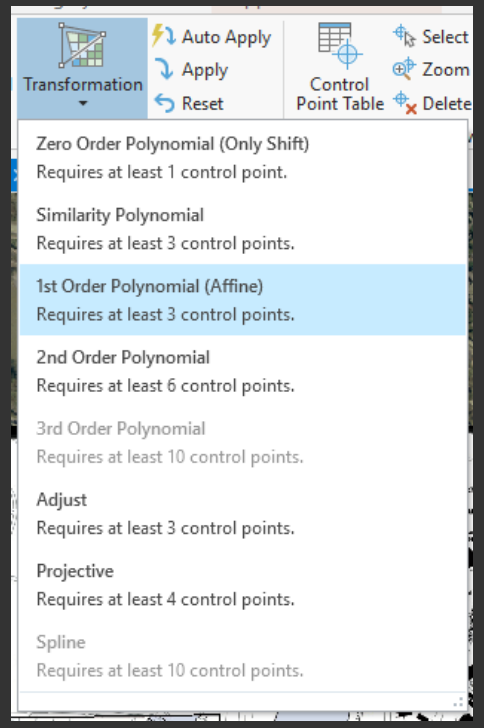
- Translate
- Scale
- Rotate
- Skew



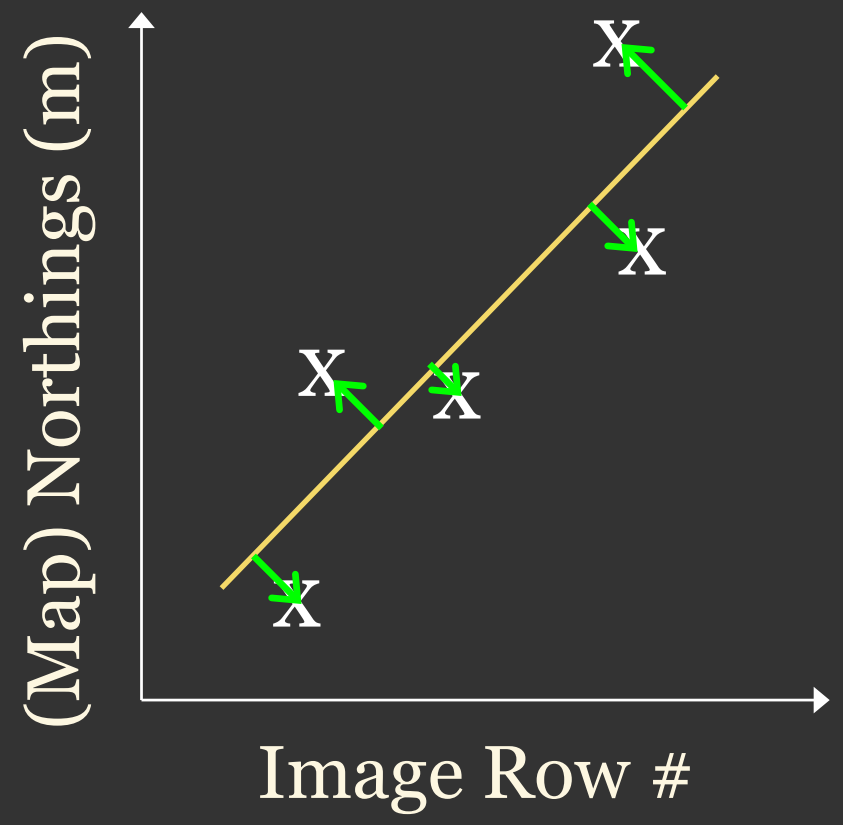
These transformations are applied mathematically.

Transformation Methods

- Zero-Order Polynomial = just translate the data (need 1 control point)
- First-Order Polynomial = translate, scale, and rotate (need 3 control points)
- Higher Order = more complex distortions can be corrected
  - Rarely use higher than third order
  - Need more control points
  - If raster dataset must be bent or curved, use a second- or third-order transformation



First-Order Polynomial (Affine)



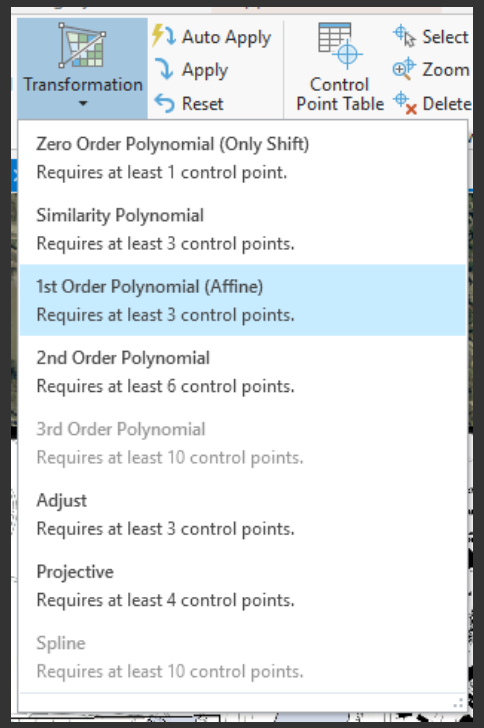
$$X = Ax + By + C$$
$$Y = Dx + Ey + F$$

X = new x coordinate
Y = new y coordinate
x = original x coordinate
y = original y coordinate

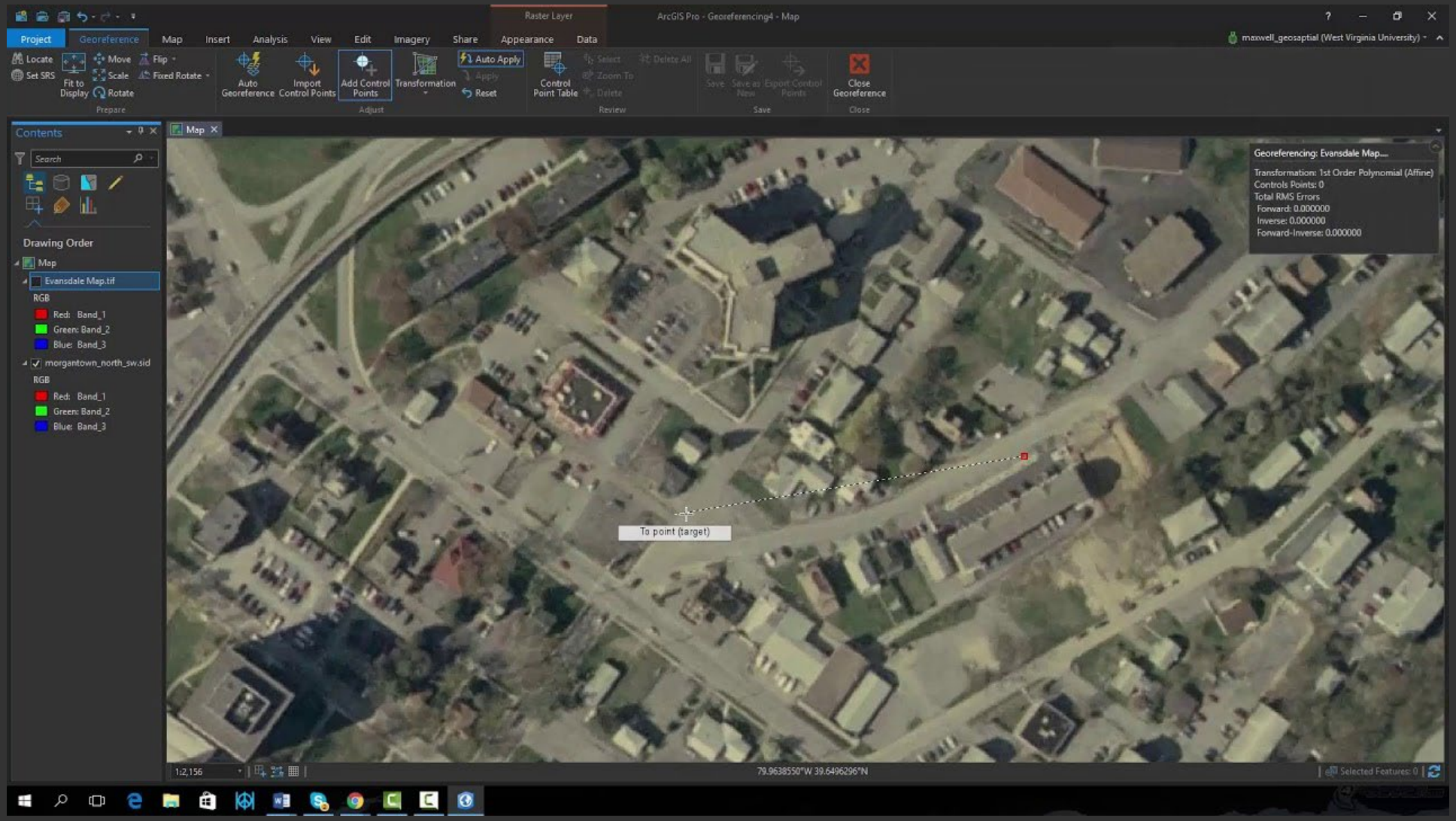
Example by Dr. Tim Warner

Transformation Methods

4. Adjust = combines polynomial and triangular irregular network (need 3 control point)
5. Similarity = first order transformation which tries to preserve the shape of the original raster (need 3 control points)
6. Projective = commonly used for image data (need 4 control points)
7. Spline = uses rubber sheeting (need 10 control points)



Video: Georeferencing in ArcGIS Pro



How do we assess the quality of our georeferenced result?

Root mean square error (RMSE)

$$RMSE_x = \sqrt{\Sigma(x_{predicted} - x_{actual})^2/n}$$

$$RMSE_y = \sqrt{\Sigma(y_{predicted} - y_{actual})^2/n}$$

$$RMSE = \sqrt{(RMSE_x^2 + RMSE_y^2)}$$

- Measure of how well the now-referenced data matches with the source
- Obtained by comparing the known location of the control point with the computed coordinates of the control point.

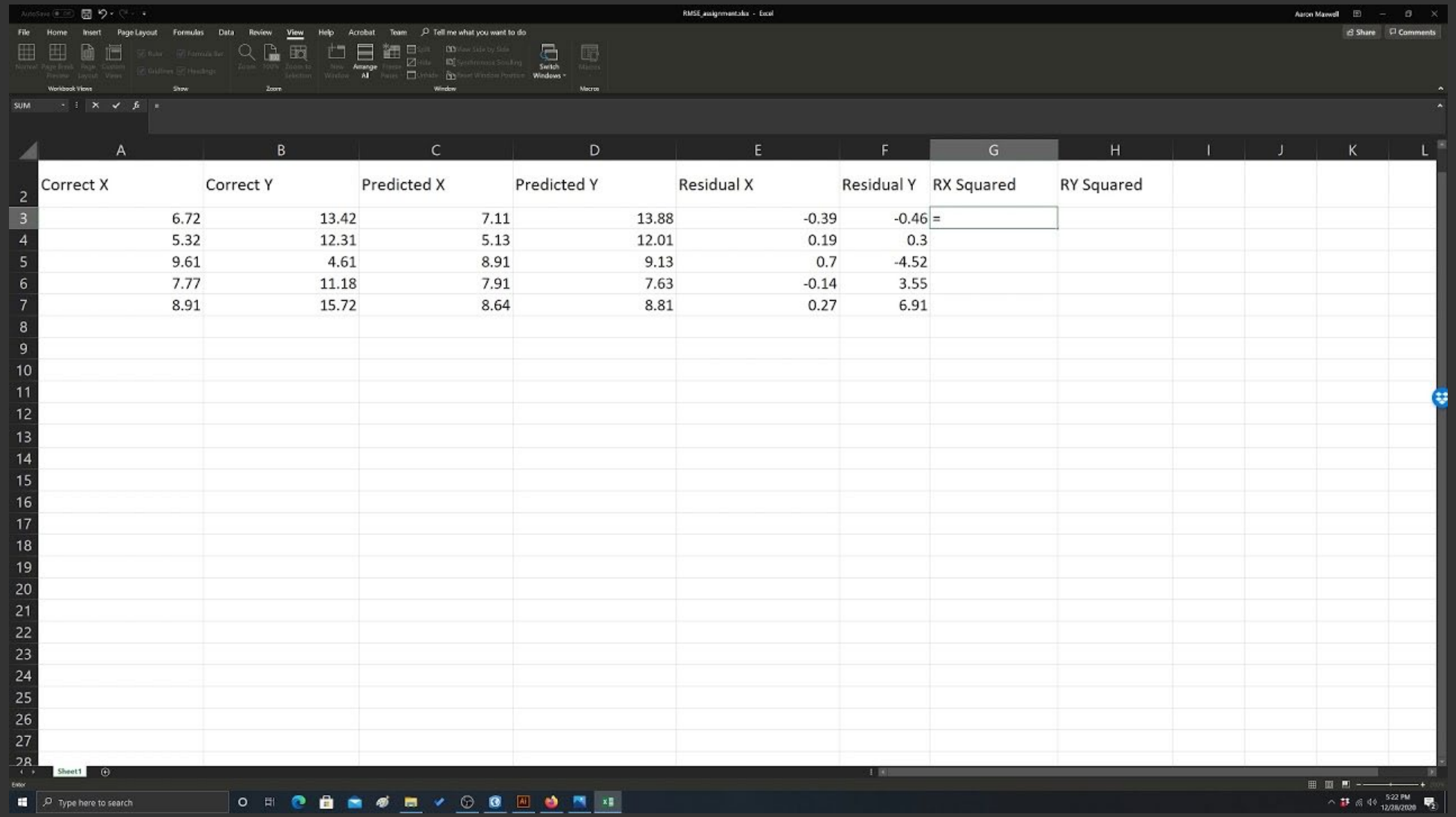
How do we assess the quality of our georeferenced result?

Correct X	Correct Y	Predicted X	Predicted Y	X residual	Y residual	X residual squared	Y residual squared
513,126.80	4,101,294.60	513,127.30	4,101,293.40	0.5000	-1.2000	0.2500	1.4400
513,123.20	4,101,303.40	513,118.10	4,101,302.10	-5.1000	-1.3000	26.0100	1.6900
513,203.80	4,101,209.10	513,203.10	4,101,207.20	-0.7000	-1.9000	0.4900	3.6100
513,182.60	4,101,377.80	513,181.60	4,101,379.90	-1.0000	2.1000	1.0000	4.4100
513,217.60	4,101,403.60	513,216.10	4,101,405.20	-1.5000	1.6000	2.2500	2.5600
						RMSE X	2.45
						RMSE Y	1.66
						RMSE	2.96

RMSE Interpretation

- The residuals is the difference between the predicted x and y coordinates and the actual coordinates, as specified with the control point
- Low RMSE can be misleading if the quality of the control points is poor
- Sometimes the algorithm overfits to the control points, so the RMSE can be misleading
- For imagery, we generally want the RMSE to be lower than the size of the pixel

Video: RMSE Calculation

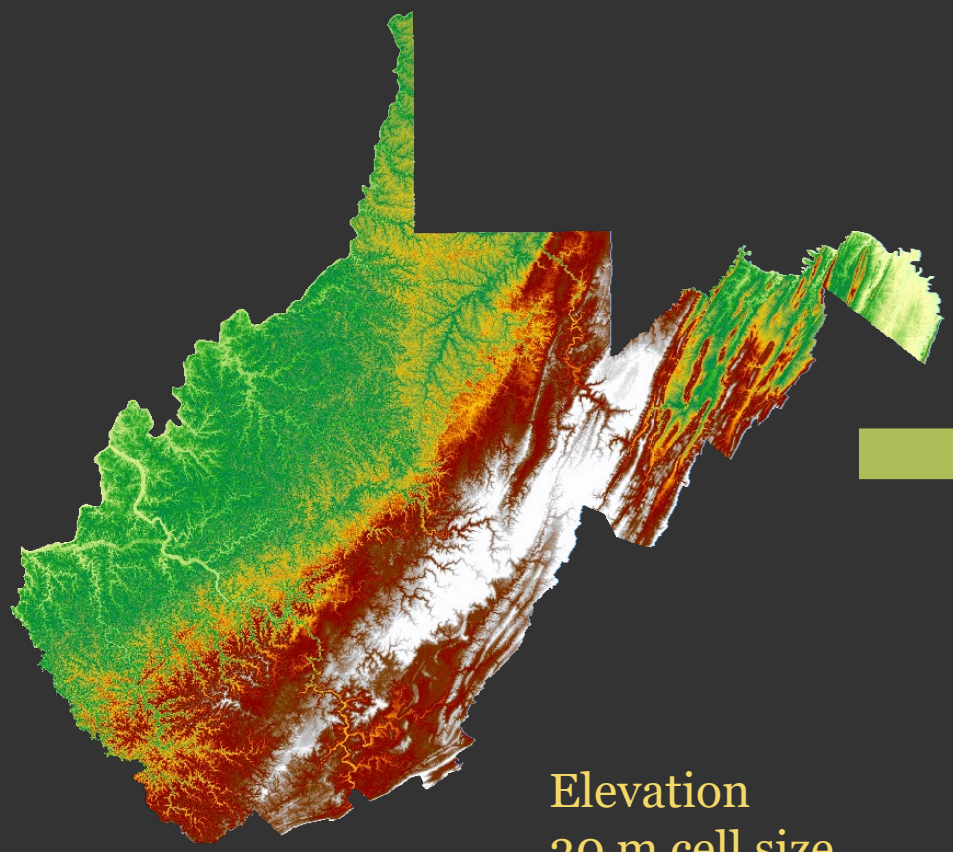


Correct X	Correct Y	Predicted X	Predicted Y	Residual X	Residual Y	RX Squared	RY Squared
6.72	13.42	7.11	13.88	-0.39	-0.46	=	
5.32	12.31	5.13	12.01	0.19	0.3		
9.61	4.61	8.91	9.13	0.7	-4.52		
7.77	11.18	7.91	7.63	-0.14	3.55		
8.91	15.72	8.64	8.81	0.27	6.91		

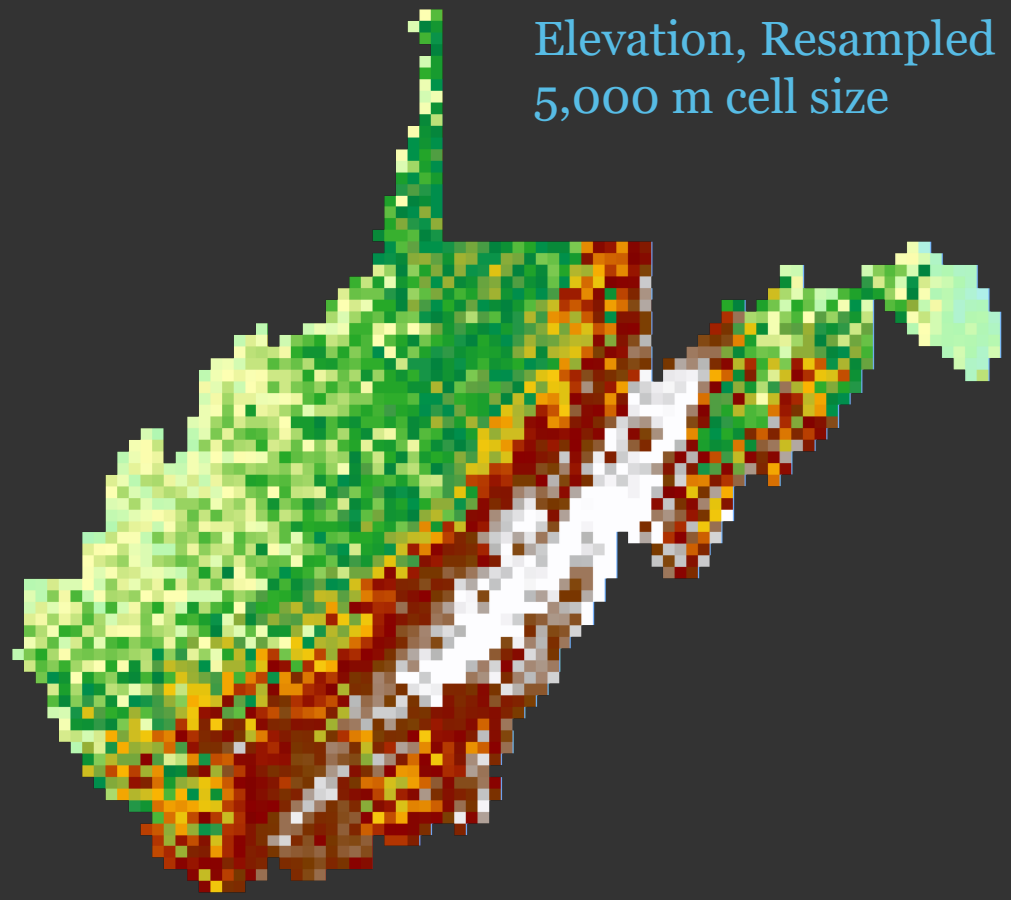
Resampling

Resampling

Change the cell size



Elevation
30 m cell size



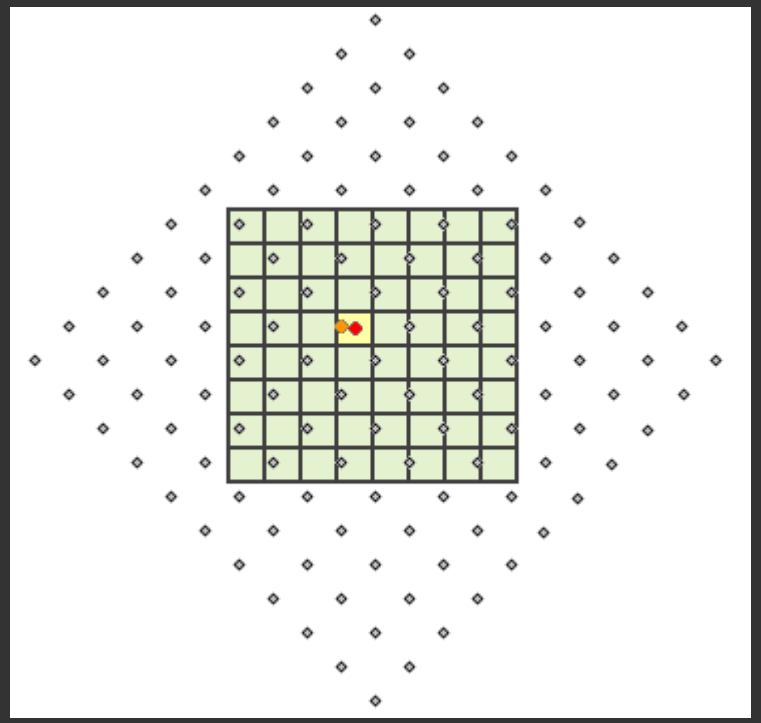
Elevation, Resampled
5,000 m cell size

Resampling Methods

- Majority Resampling
- Resamples using the most common cell value
- Good for categorical data

Resampling Methods

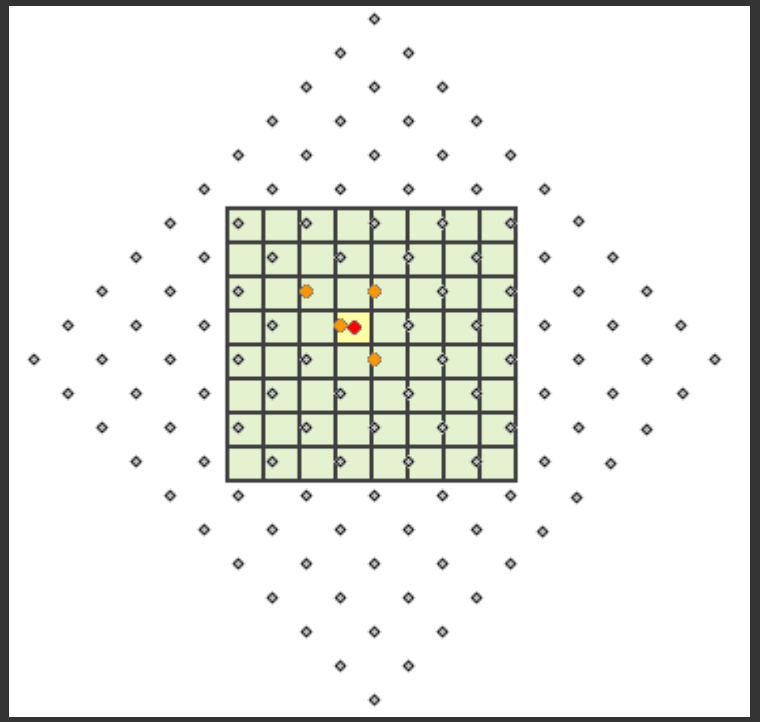
- Nearest Neighbor
- New value will be the value of the original cell closest to the center of the new (larger) cell
- Good for categorical data



http://desktop.arcgis.com/en/arcmap/latest/extensions/spatial-analyst/performing-analysis/cell-size-and-resampling-in-analysis.htm

Resampling Methods

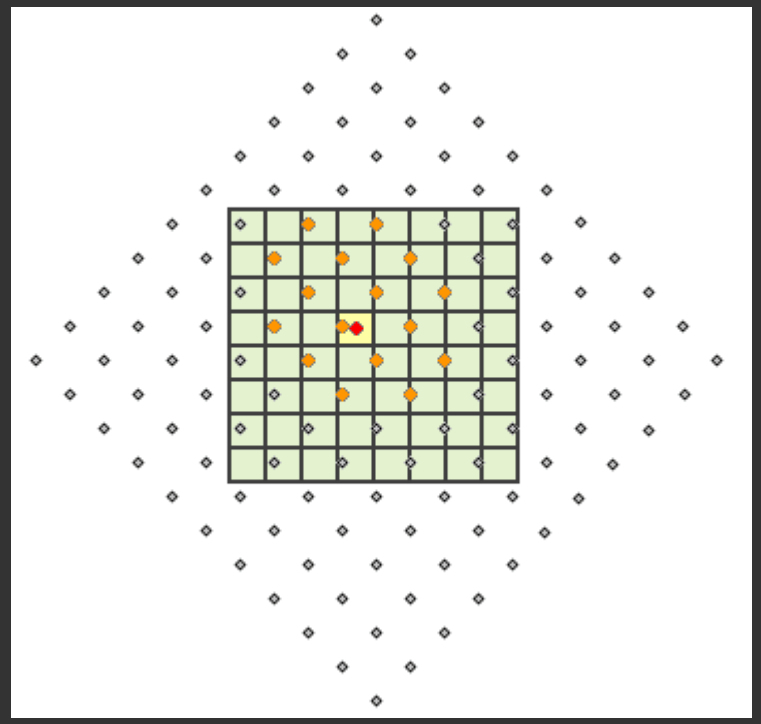
- Bilinear Interpolation
- Uses the inverse distance-weighted average of the four nearest pixel values to estimate a new pixel value
- The four cell centers from the input raster that are closest to the cell center for the output processing cell will be weighted based on inverse distance and then averaged
- Good for continuous data
- Creates a smoothed surface



http://desktop.arcgis.com/en/arcmap/latest/extensions/spatial-analyst/performing-analysis/cell-size-and-resampling-in-analysis.htm

Resampling Methods

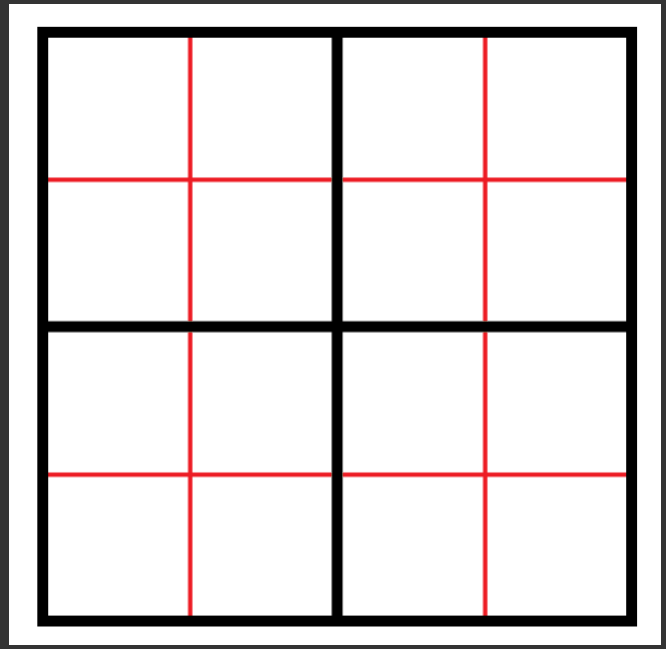
- Cubic Convolution
- Similar to bilinear interpolation, except the 16 nearest input cell centers and their values are used
- Tends to sharpen edges
- Good for continuous data

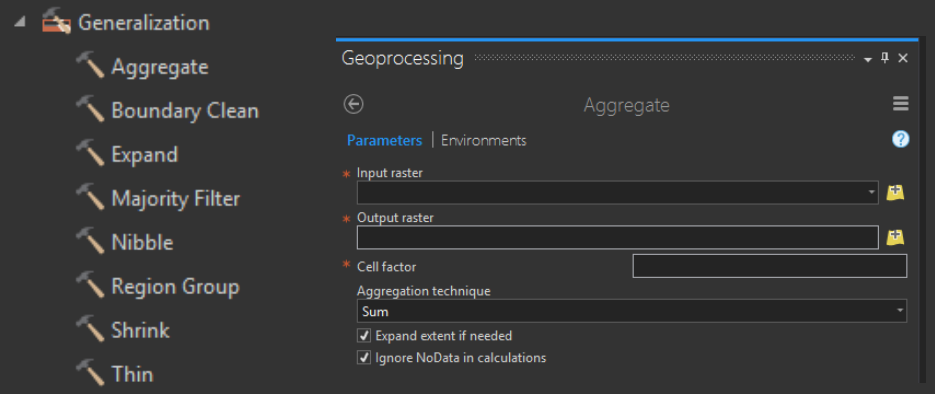


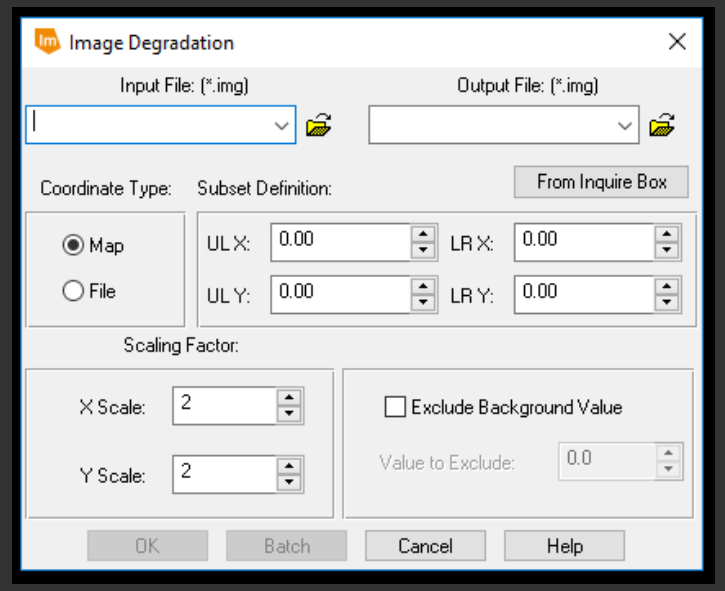
http://desktop.arcgis.com/en/arcmap/latest/extensions/spatial-analyst/performing-analysis/cell-size-and-resampling-in-analysis.htm

Pixel Aggregation

- Generates a reduced-resolution version of a raster grid
- Each output cell contains the Sum, Minimum, Maximum, Mean, or Median of the input cells that are encompassed by the extent of that cell



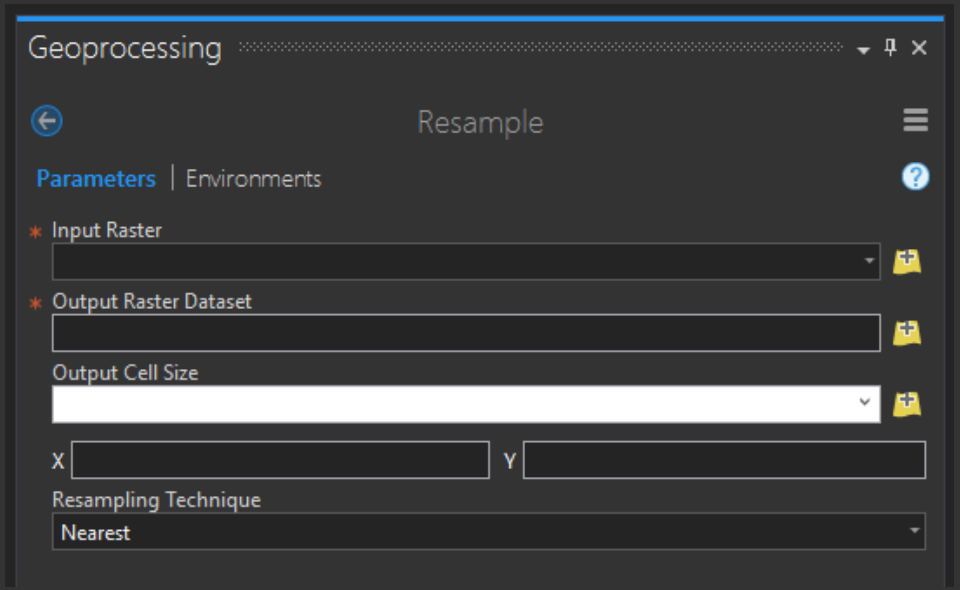


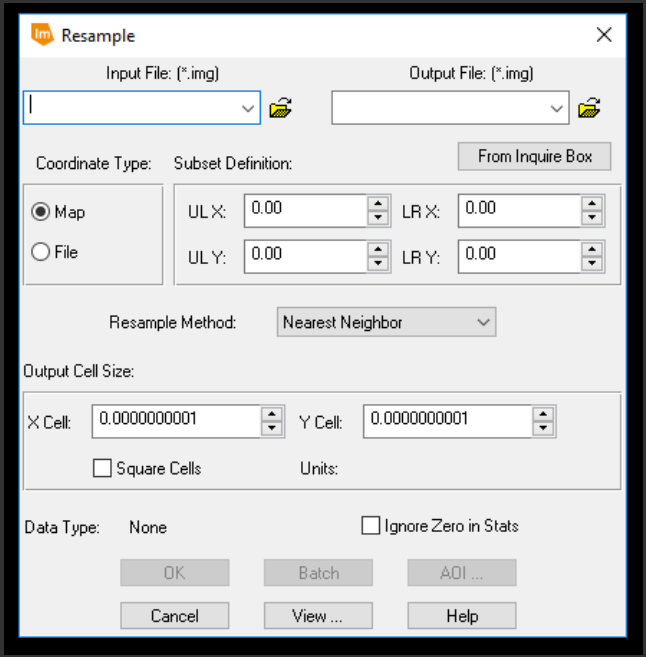


Resampling in ArcGIS Pro

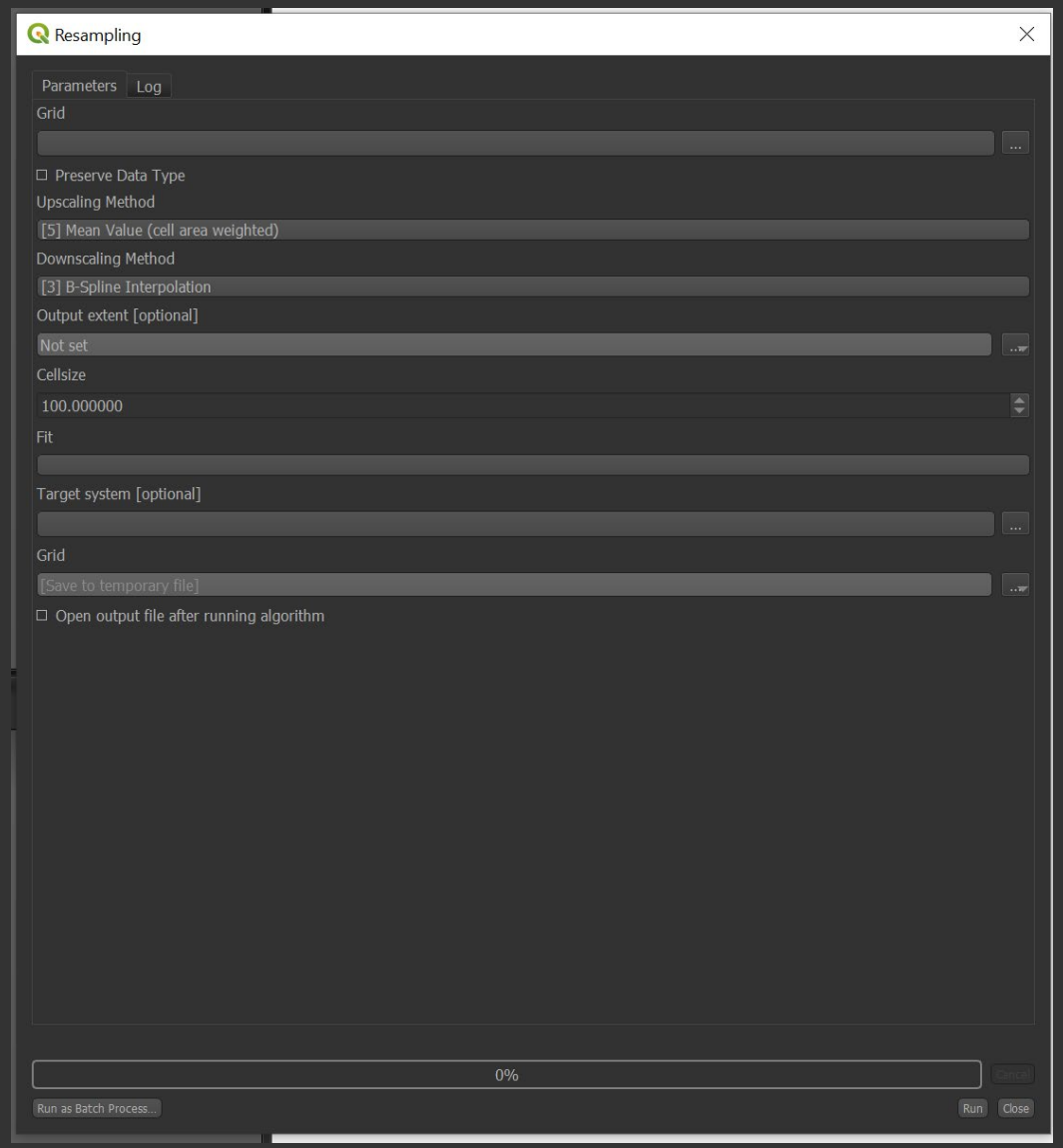
- A good resource:

http://desktop.arcgis.com/en/arcmap/latest/extensions/spatial-analyst/performing-analysis/cell-size-and-resampling-in-analysis.htm





Resampling in QGIS



- Resampling and pixel aggregation tools are available in QGIS, GDAL, GRASS, and SAGA

This is the end of this lecture module.

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