Introduction to GIS

(GEOG/ES 470)

Raster Data Model

Table of Contents

PART II: Image Classification 2

I. Classifying Satellite Image Subset (Millcreek) 2

II. Supervised Classification 2

III. Unsupervised Classification 5

IV. Working with the Value Attribute Tables (VATs) 8

V. Working with Digital Elevation Moldels (DEMs) 10

VI. Your Turn 15

VII. Submit your Work 15

VIII. Grading Rubric 15

Raster Data Model

PART II: Image Classification

SUMMARY

In this exercise you are going to learn about different techniques in raster data classification, and DEM manipulation. At the end of this exercise you'll be replicating the same steps to create your own Township/area's classification and DEM representation. While you're working within your selected study area, make sure to grab screen shots of the steps that you carried out in order to create the final report of your methods and findings. The final report is going to be created in MS Word (or similar) and **saved as PDF**, you'll then upload it to the corresponding D2L assignment.

I. Classifying Satellite Image Subset (Millcreek)

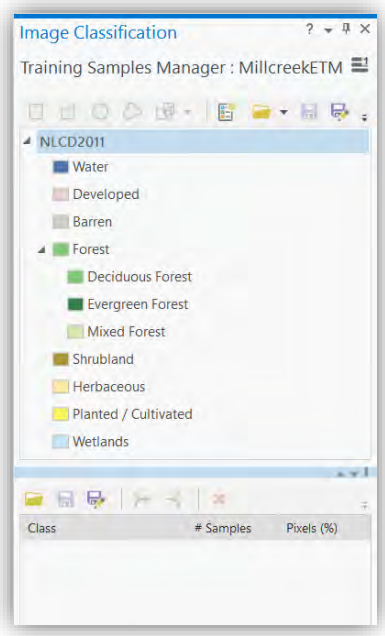
Satellite image classification will depend on feature reflectance in multi-band data. The class separation will be dependent on each feature's signature. There are two main methods for image classification: Unsupervised and Supervised. In the Unsupervised classification, a clustering method is applied to gather all pixels from the image that has similar behavior in each of the given bands. In the Supervised classification, on the other hand, the user will create sample training signature areas for each visually (or field-checked) feature. Those areas are then applied to the image as a whole, the routine is to find group of pixels that match each of the identified unique polygons. In the following text, you'll be performing both Supervised and Unsupervised classifications and visually compare them.

II. Supervised Classification

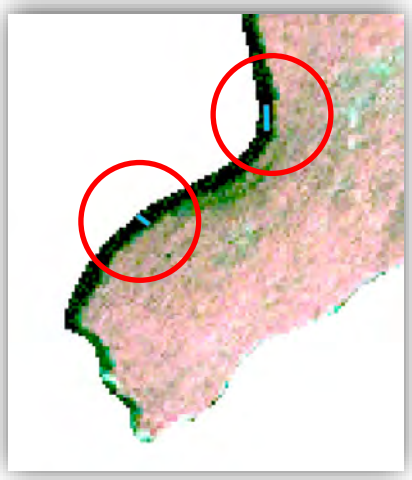
- In ArcGIS Pro, create a new Map; call it ***Image Classification***
- Add *MillcreekETM* to the *Contents*
- Change the band combination and histogram stretch to view the variations in land cover in a more enhanced way (refer to Part I for visual manipulation and contrast stretching).

A. Creating the Training Samples

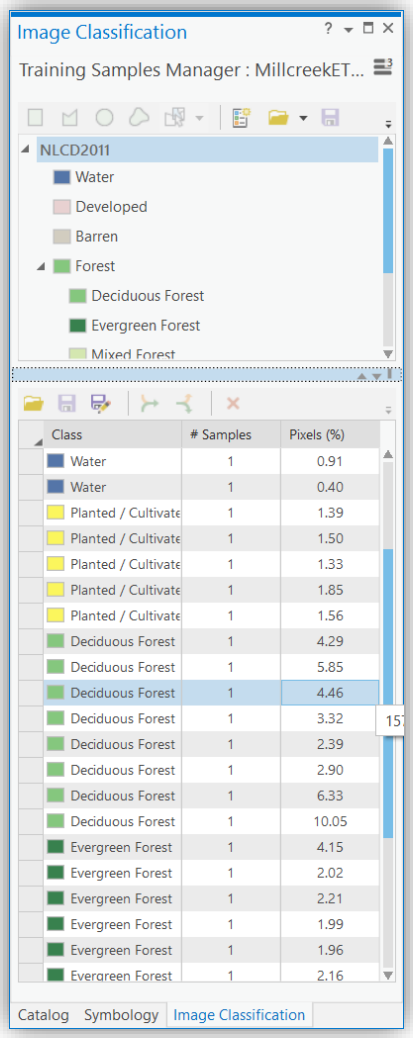
- With the *MillcreekETM* layer selected in the *Contents*, select *Imagery* tab >> *Image Classification* group >> *Classification Tools* >> *Training Samples Manager*
- The Image Classification Training Samples Manager pane opens and the default classification scheme (NLCD2011) is loaded.
- The classification scheme is what defines the classes expected to be available in the image. You can modify this list by creating custom schemes. For this exercise we'll work with the default NLCD2011 scheme.

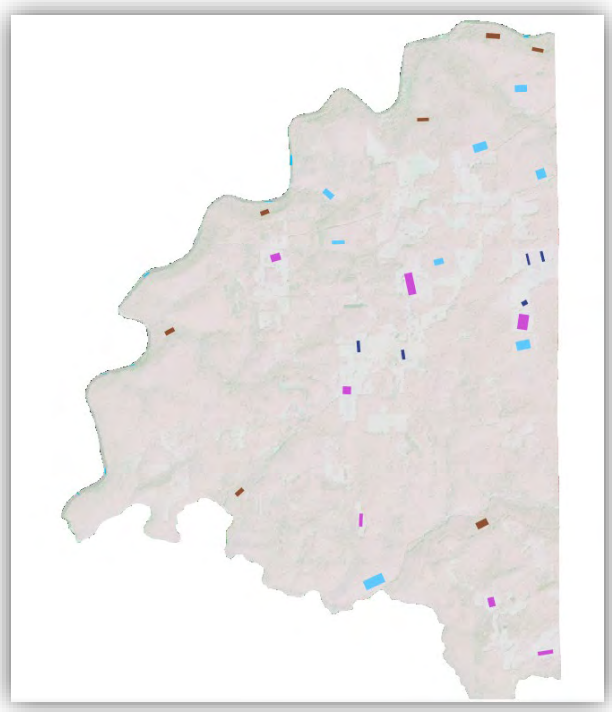


- Start picking training sites by clicking on the first class (*Water*) and draw small polygons or rectangles of homogenous areas with the same color, shade and texture.
- The following figure shows the expected size and nature of the polygons/rectangles that should be created for a couple of samples of water areas (lower left of Millcreek township)



- Add a few of those water polygons then switch to another class from the provided schema and digitize their corresponding polygons in a similar manner. Make sure that your polygons are small enough to include homogenous pixel values (they mostly look the same texture and color all over the training polygon).
- The Training Samples Manager window will include multiple classes (bottom part of the pane). **Save** the training samples inside the Project's database, call it ***Millcreek_TrainingSamples*** (use the save icon in the lower portion of the *Training Samples Manager* pane), the following map shows approximately the size of the training polygons. They should be small enough to represent the feature you're detecting, and large enough to contain enough pixels to help calculating the necessary statistics for the classification procedure (Signature file…etc.).



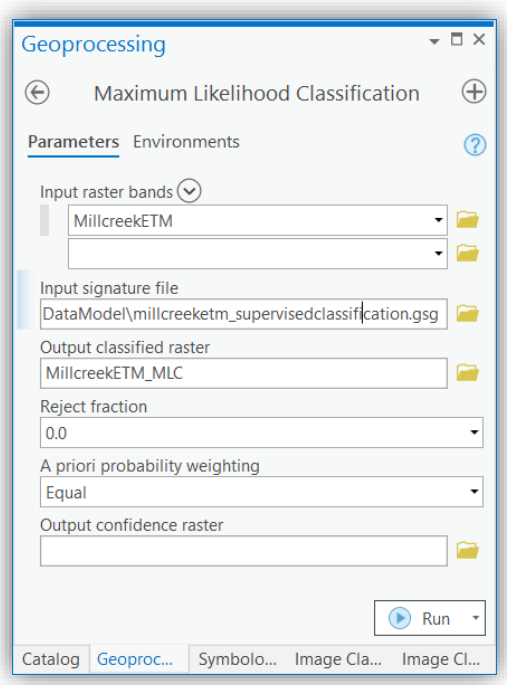


B. Creating the Signature File

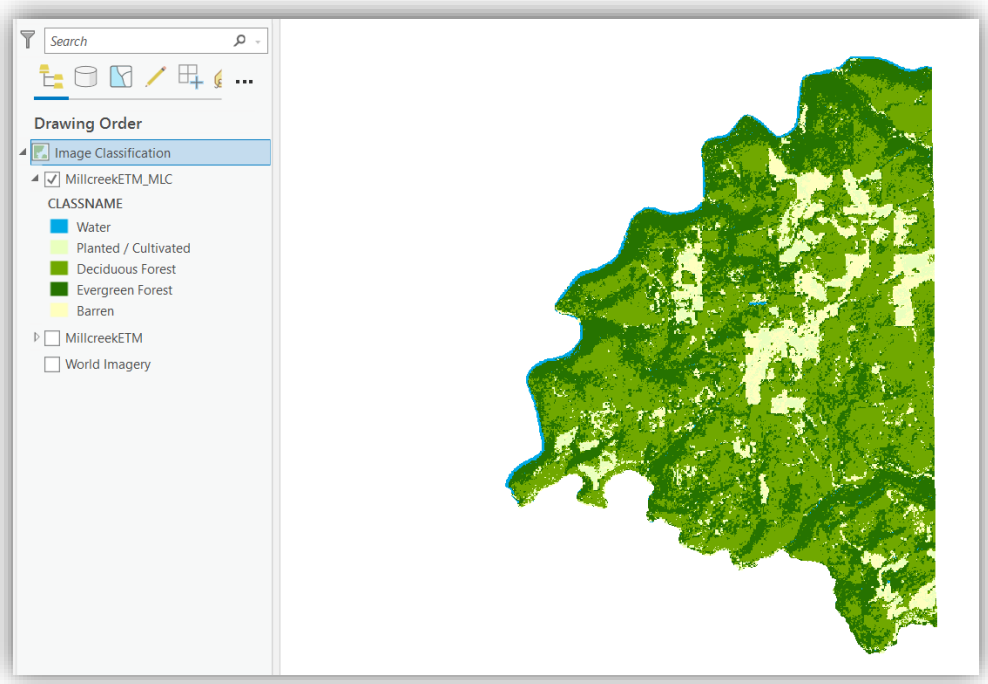
- From the *Analysis* tab >> *Geoprocessing* group, click on *Tools*
- In the *Geoprocessing* pane, search for "*signature*"
- Click on the second result: *Create Signatures* tool
- Add *MillcreekETM* as the *Input raster bands*, and the *MillcreekTrainingSamples* as the *Input raster or feature sample data*
- Keep the *Sample field* set to the default *Classname*
- Save the *Output signature file* to the Project's folder, name it ***Millcreek_SupervisedClassification***, a default *.gsg* file extension is going to be appended to the file name.
- Check the *Computer covariance matrices,* and click *Run*
- The signature file will be created and stored in the Project's folder.

C. Running the Supervised Classification

- From the *Analysis* tab >> *Geoprocessing* group, click on *Tools*
- In the *Geoprocessing* pane, search for "*maximum*"
- Click on the first result: *Maximum Likelihood Classification*
- Change the *Input raster bands* to *MillcreekETM*, the *Input signature file* to the previously created *Millcreek_SupervisedClassification.gsg* file, and the *Output classified raster* to inside the Project's database, name it ***MillcreekETM_MLC***
- Click *Run*
- The process will run in the background and the final result will be added to the table of contents.
- Change the color assignments to match the following:
    - Water: *Moorea Blue*
    - Planted / Cultivated: *Olivine Yellow*
    - Deciduous Forest: *Tarragon Green*
    - Evergreen Forest: *Fir Green*
    - Barren: *Yucca Yellow*



- The final result for the Supervised Classification should look similar to the following figure:



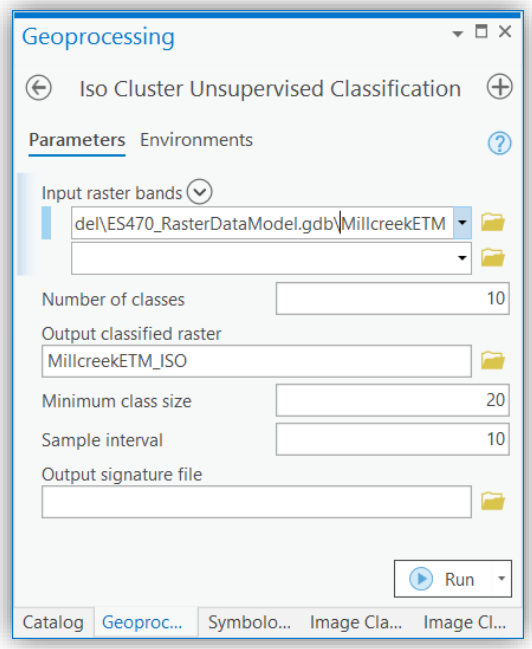
D. Inspect the classification

- From the *Map* tab >> *Layer* group, click on the *Basemap*, and select *Imagery*
- In the *Contents,* turn off all of the layer and keep the *MillcreekETM_MLC* classified image on.
- Select the *MillcreekETM_MLC* image from the *Contents*
- From the *Appearance* tab >> *Compare* group, click on *Swipe*
- Click, hold, and drag your mouse up/down or left/right anywhere on the image in the *Map*, now you can compare between the classified image and the most up-to-date satellite image of the township.
- You may also turn on/off the layer to compare it with the basemap and see how accurate the classification is.

III. Unsupervised Classification

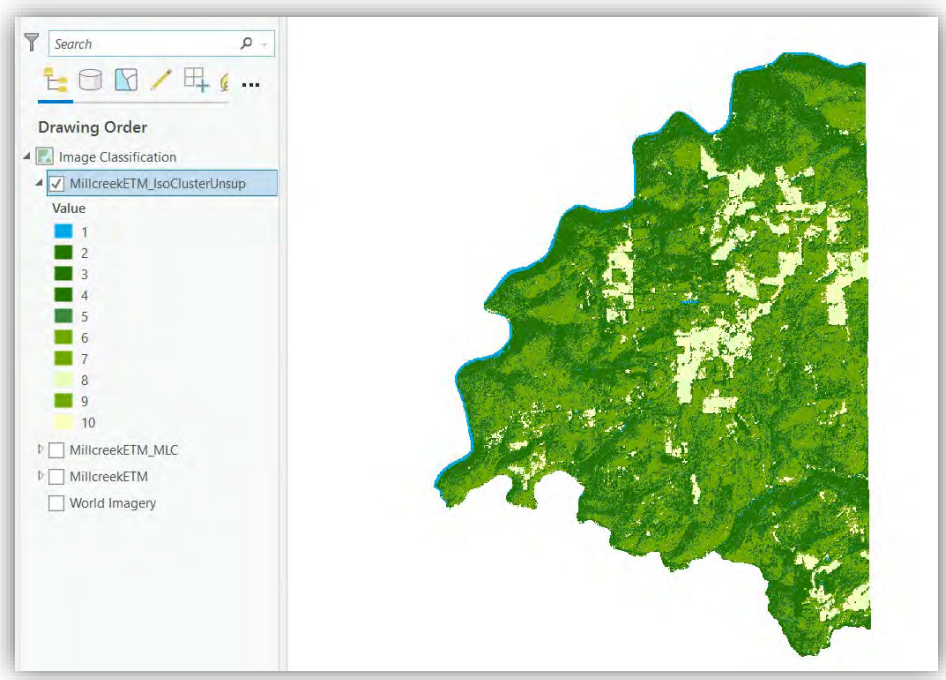
- In the unsupervised classification all what you need to specify is the number of classes and the minimum polygon size (number of pixels).
- From the *Analysis* tab >> *Geoprocessing* group, click on *Tools*
- In the *Geoprocessing* pane, search for "*iso cluster*"

- Click on the second result: *Iso Cluster Unsupervised Classification*
- For the *Input raster bands*, browse the Project's database and select the *MillcreekETM* image
- Change the *Number of classes* to **10**
- Save the *Output classified raster* to the Project's database, name it ***MillcreekETM_ISO***
- Keep all other default values (class size 20 and sample interval 10), and click *Run*
- The process will run in the background and the final result will be added to the *Contents*.

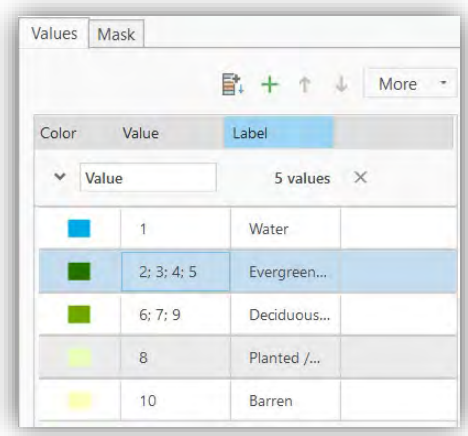


A. *Change the Layer's Symbology*

- Change the color assignments according to the produced classes; try to match them to the same color assignment of the Supervised Classification done earlier, although it is not necessary that this method would produce the same exact classes as the supervised classification. You'll notice that the 10 produced classes could be aggregated to match only 5 (e.g. 2 or more of the 10 classes could be identified as *Evergreen Forest*, in this case, give it the same color)
- The final result should look similar to the following figure:



- Notice that in this figure, classes 2, 3, 4, and 5 all belong to the *Evergreen Forest* class, and therefore they all were changed to the same color: *Fir Green*
- From the layer's symbology (*Appearance* tab >> *Rendering* group >> *Symbology*) group the similar classes together by holding down the Ctrl key and clicking on all the similar classes (e.g. 2, 3, 4, and 5), right-click and select *group values*
- Also, change the class labels to their corresponding class names instead of the given class codes (*values)*.



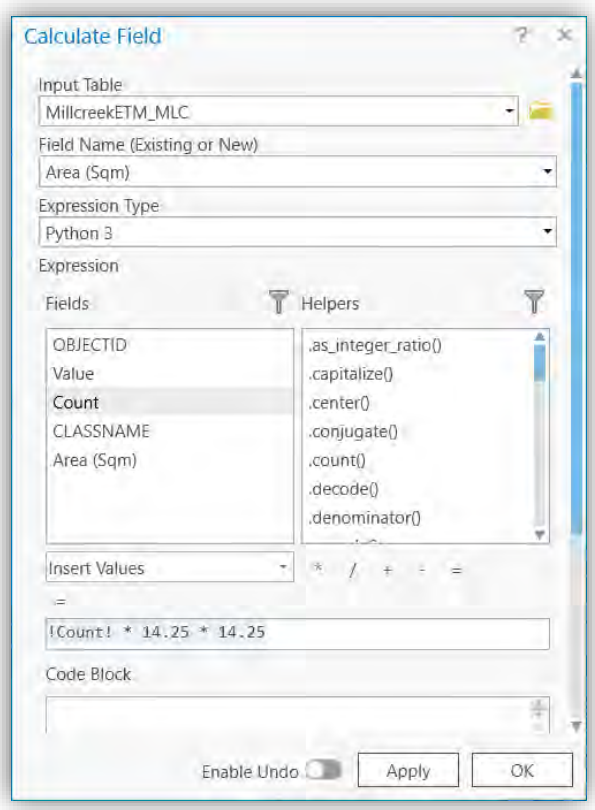
B. Inspect the classification

- If you didn't already do so, from the *Map* tab >> *Layer* group, click on the *Basemap*, and select *Imagery*
- In the *Contents,* turn off all of the layer and keep the *MillcreekETM_ISO* classified image on.
- Select the *MillcreekETM_ISO* image from the *Contents*
- From the *Appearance* tab >> *Compare* group, click on *Swipe*
- Click, hold, and drag your mouse up/down or left/right anywhere on the image in the *Map*, now you can compare between the classified image and the most up-to-date satellite image of the township.
- You may also turn on/off the layer to compare it with the basemap and/or the *MillcreekETM_MLC* supervised classification image, and see how accurate and different the classification is.

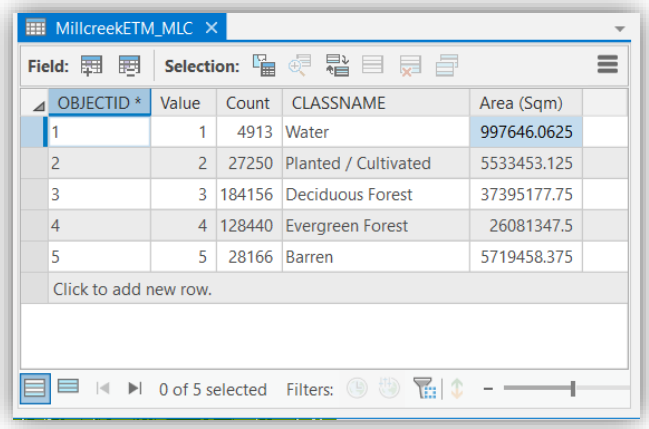
IV. Working with the Value Attribute Tables (VATs)

- Like vector data, rasters have attribute tables when they are classified (i.e. discreet number of values and not continuous raw data such as satellite image bands or elevation data)
- In the *Contents,* right-click on *MillcreekETM_MLC* and open its attribute table.
- A typical raster data attribute table will contain two essential fields:
    - Value: which is the assigned unique value stored in each pixel
    - Count: is the number of pixels assigned this unique value
- Another field, *CLASSNAME* is also present in this case, since it was a result of a supervised classification method.
- The raster Attribute Table is a summary of how many pixels in each class. You can calculate percentages and areas using those counts.
- It is therefore important to identify the pixel size in the classified image. From the *MillcreekETM_MLC* layer properties, go to *Source >> Raster Information.* Notice that the pixel size is 14.25 meters, so each pixel's area is 203.0625m^2 (14.25 x 14.25). If you multiplied this area by the count of pixels in each class (value), you'd get the total area occupied by this class in square meters.
- You can do this by adding a new field to the table. Click on the *Add Field* button from the attribute table view, for the *Field Name* type in *Area_SQM*, you may change the *Alias* to *Area (Sqm)* and change the *Data Type* to *Double*.
- From the *Fields* tab (main menu area) >> *Changes* group, click *Save*. This will save the changes made to the table structure (schema).
- Switch back to the *MillcreekETM_MLC* attribute table, and right-click on the newly created field (*Area (Sqm)*) and select *Calculate Field*
- Right-click Area_SQM and select Field Calculator.
- In the Field Calculator dialog, double click on the *Fields >> Count*, then click the * sign, type in 14.25, click on the * again, then type in 14.25. The following should show in the calculation field:

  !Count! * 14.25 * 14.25



- Click *Apply* and view the results in the attribute table. The areas in square meters are calculated for each class.
- Now if you identify any pixel (*Map* tab >> *Navigate* group >> *Explore)*, the value, the total count of this value in the whole raster layer and the Area in Square Meters for this value all over the layer are going to be reported.



MillcreekETM_MLC

OBJECTID *	Value	Count	CLASSNAME	Area (Sqm)
1	1	4913	Water	997646.0625
2	2	27250	Planted / Cultivated	5533453.125
3	3	184156	Deciduous Forest	37395177.75
4	4	128440	Evergreen Forest	26081347.5
5	5	28166	Barren	5719458.375

Click to add new row.

0 of 5 selected

V. Working with Digital Elevation Moldels (DEMs)

Digital Elevation Models are created as a raster representation of surfaces. For Pennsylvania, LiDAR data were acquired for the whole state at different stages. In 2006, Clarion County was covered by the LiDAR production. Data of all PA could be downloaded from the Pennsylvania Spatial Data Access (PASDA) portal (www.pasda.psu.edu). The DEM tiles are 10,000ft x 10,000ft named by their column and row numbers. Those tiles are available as shapefiles for the North and South Zones for PA in the *Downloads* folder that you extracted for Part I of this exercise.

- In a new Map, from the *Catalog* pane, browse, the Project's folder under *Folders*, add the *PA_SP_NAD83_north.shp* shapefile located under the *Downloads* located inside the Project's folder. Right click the color patch for the layer and select *No Color* in order to display just the borderlines of the tiles.
- Right click on the *PA_SP_NAD83_north* layer and select *Label.* This will turn on the labels for this layer and displays each tile ID that will help identifying the tiles corresponding to each municipality later on.
- Also add the *Municpalities.shp* shapefile (located under *Downloads\2000-Census*) to the *Contents.* You may label this layer in order to recognize the municipality names.

A. Downloading the DEMs

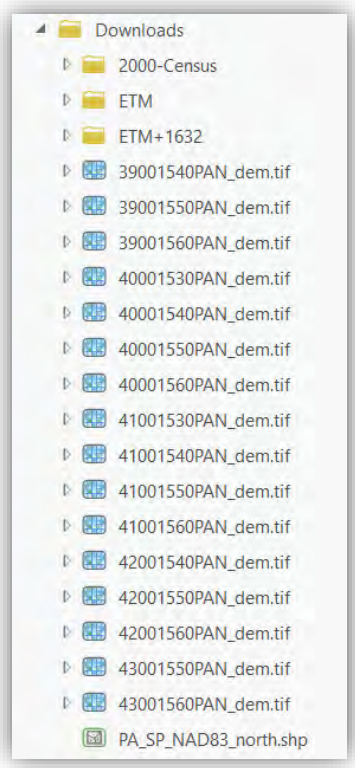
Displaying the tile (*PA_SP_NAD83_north*) and labeling it using the *TILEID* field would reveal the tile numbers necessary to cover your study area. For examples there are 16 DEM tiles that would cover Millcreek Township (see below)

- All downloaded data should be stored and extracted (unzipped) inside the ***DOWNLOADS*** folder (unzipping would need to be done from windows file explorer)
- Go to the PA Imagery Navigator website for all DEM data downloads (http://maps.psiee.psu.edu/ImageryNavigator/)
- From the *Display Tile Index,* show the *2006 >> Western PA Tile Index*
- Zoom in to Clarion County, and close to Millcreek township. Refer to the map you just created in ArcGIS Pro to identify the location of the tiles that needs to be downloaded.
- The designated tiles for download are:
    - 39001540, 39001550, 39001560;
    - 40001530, 40001540, 40001550, 40001560;

- 41001530, 41001540, 41001550, 41001560;
- 42001540, 42001550, 42001560; and,
- 43001550, 43001560.

You may need to zoom in more in order to display the tile labels.

- Start the tile data download by right-clicking on each of the tiles, the *PASDA* Download Links window will pop up.
- For the Lidar data, click on the "*Lidar, Topo*" tab from the PASDA *Download Links* window and download the DEM of this tile

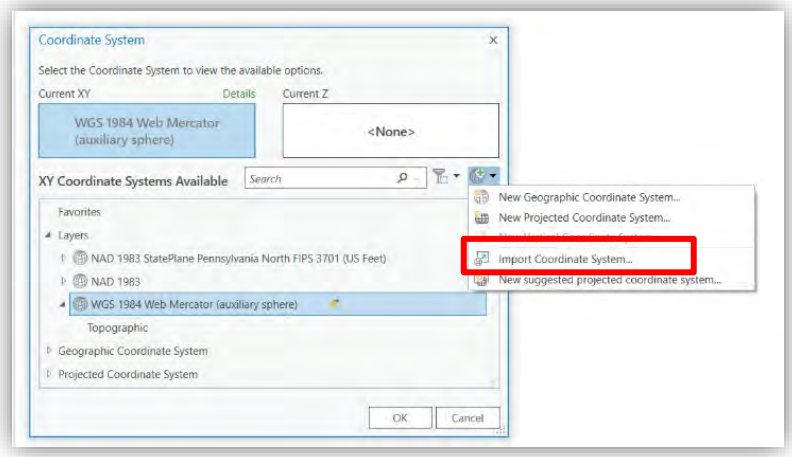


- repeat right-clicking on each of the designated tiles to Download their corresponding *DEM*. Note that when you right-click on a tile from the image, the tile name changes on the *PASDA Download Links* window. A total of 16 zip files for 16 tiles that cover Millcreek township should be downloaded.
- In windows explorer, unzip all 16 zip file inside the *Downloads* folder. After unzipping, you can delete the 16 zip files.
- In ArcGIS Pro, refresh the *Downloads* folder in the *Catalog* pane, the 16 dem tiles should show similar to the figure on the right.

B. Creating a Mosaic Dataset

- In the *Catalog* pane, right-click the Project's database *Databases >> ES470_RasterDataModel* and select *New → Mosaic Dataset*
- In the *Create Mosaic Dataset* dialogue, Change the *Mosaic Dataset Name* to ***Millcreek_DEM***.
- Change the *Coordinate System* to match the one of the downloaded DEM tiles:
    - Click on the browse icon (globe next to the *Coordinate System*)

- In the *Coordinate System* dialog, click on the arrow next to the globe on the right side of the window and select *Import Coordinate System…*
- Browse through the Project's folder, under *Downloads*, and pick any of the previously downloaded DEM tif files (e.g. *39001540PAN_dem.tif*)

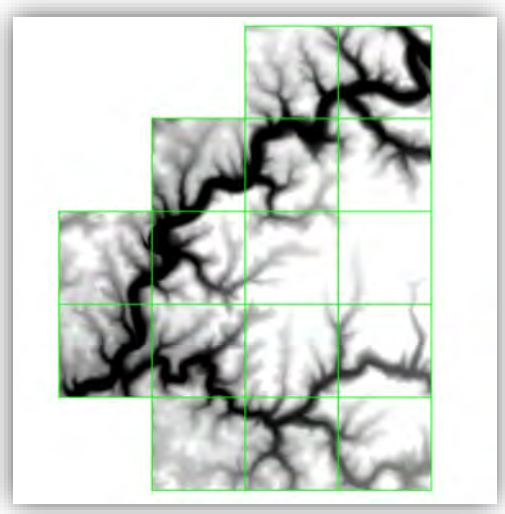


- This will set the coordinate system to *NAD 1983 StatePlane Pennsylvania North FIPS 3701 (US Feet)*.
- Click *OK*, then click *Run*

- The mosaic dataset will be added to the Project's geodatabase

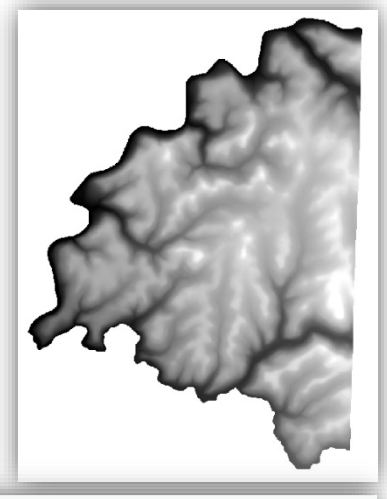
C. Adding tiles to the Mosaic Dataset

- Now, we're going to add the DEM raster tiles inside this mosaic dataset. Right-click the mosaic dataset (*Millcreek_DEM*) and select *Add Rasters…*
- In the *Add Rasters to Mosaic Dataset* dialogue, change the *Input Data* to *Dataset*, then browse the folder that contains all of the DEM tiles (*Downloads*) and select all the DEM tif files (by holding the shift key while clicking on the first then the last file in the list. Or, by selecting the necessary tiles by holding the Control Key on the keyboard to selectively pick the ones that form the current Township).
- Click *OK*.
- Under *Raster Post-processing* check the *Update Overviews*
- Under *Raster Processing*, Check the *Build Raster Pyramids* and the *Calculate Statistics* boxes.
- Click *Run*
- The resulting mosaic dataset will be added to the *Contents* and an overview of the DEM tiles will be displayed



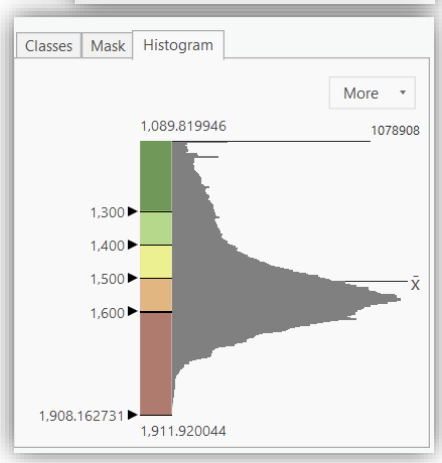
D. Clipping the DEM image

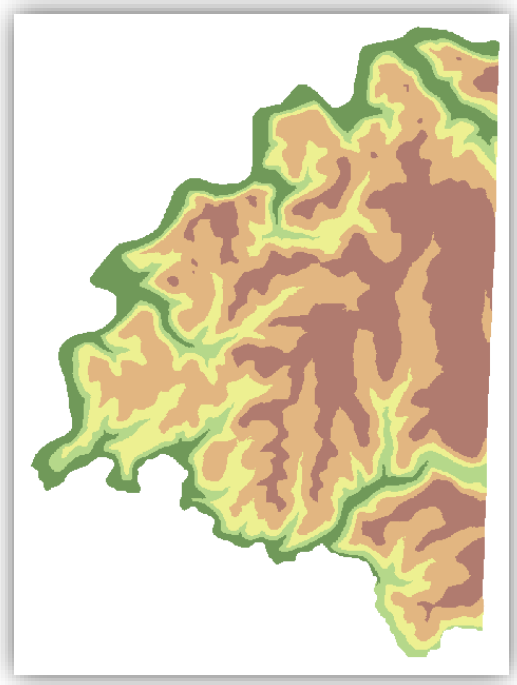
- Using the directions that you adopted earlier in **1 Image Enhancement_Understanding Histograms.docx** clip the *Millcreek_DEM Image* (from the *Contents)* to the shape of Millcreek Township
- Save the resulting Clipped DEM for Millcreek under the Project's geodatabase; name it ***Millcreek_DEMClipped***. The output should look similar to the following figure:



E. Reclassifying Symbolized Raster

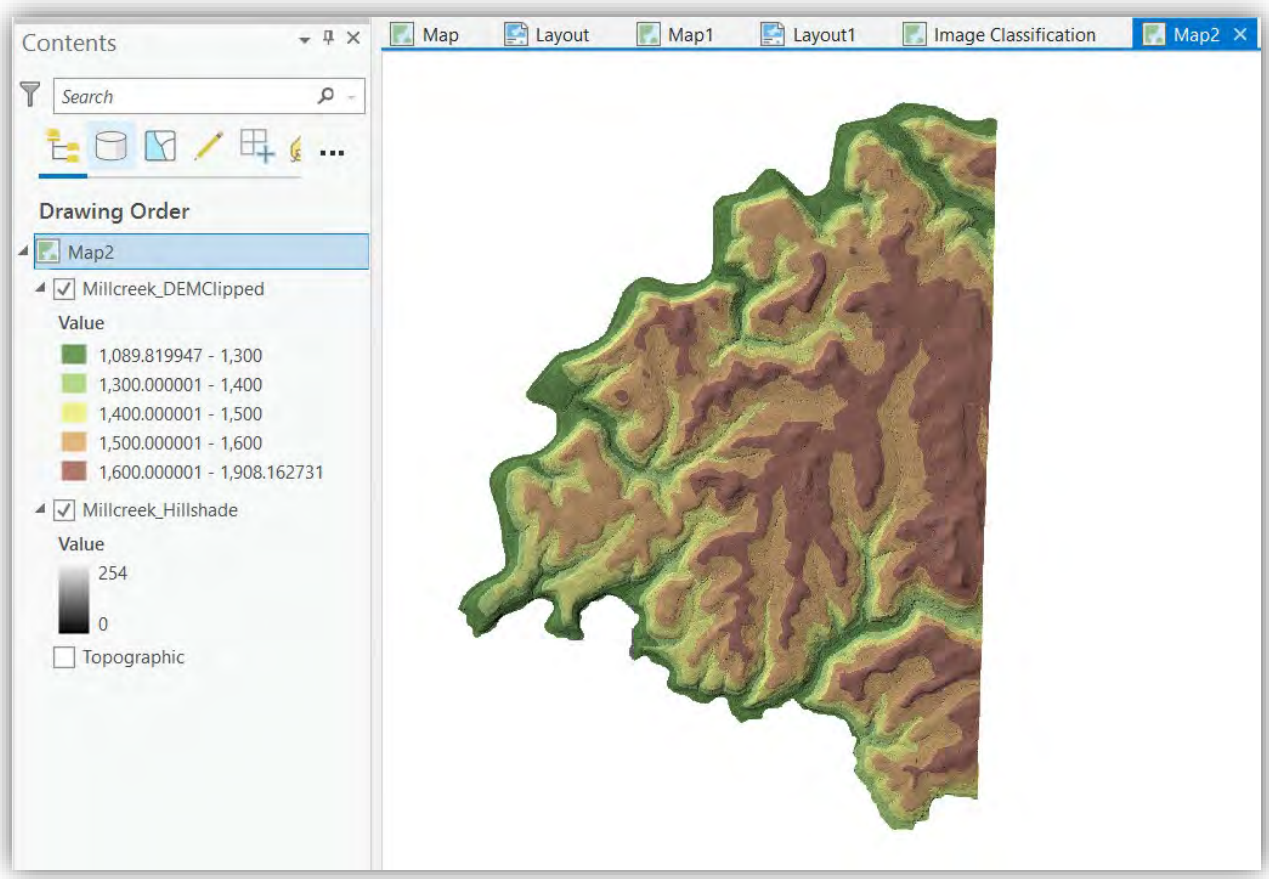
- When added to the table of contents, change the symbology for the *Millcreek_DEMClipped* raster to classified (*Appearance* tab >> *Symbology)*.
- Use 5 classes and the *Elevation #4* gradient for the *Color Scheme*.
- Refine the *Upper value* and their corresponding *Label* for every symbol to full elevation numbers (e.g. 1300, 1400, 1500, 1600, and 1912 (max) ... etc.), you may also use the *Histogram* to change those values, by either typing the values or dragging the arrows up/down.
- The final representation would look similar to the figure below





F. Adding Hillshade Effect

- From the *Analysis* tab >> *Geoprocessing* group, click on *Tools*
- Search for "*Hillshade*", and click on the first result: *HillShade (3D Analyst Tool).*
- Change the *Input Raster* to *Millcreek_DEMClipped* as your input (DEM of Millcreek), change the name of the *Output raster* to ***Millcreek_Hillshade***, make sure it is saved inside the Project's database. Click *Run.*
- The newly created *Millcreek_Hillshade* will be added to the *Contents*.
- In the *Contents* pane, move the *Millcreek_Hillshade* layer under the *Millcreek_DEMClipped*
- Click on the *Millcreek_DEMClipped* layer in the *Contents* to select it
- From the *Appearance* tab >> *Effects* group, change the *Layer Blend* to *Multiply*
- The result will be a blend between the elevation classes from the DEM layer and the effect of the hillshade from the bottom layer



VI. Your Turn

Pick a Township of your choice in Clarion County, make sure it is equivalent in size to the Millcreek township (Do NOT pick a small borough), clip the ETM image to its boundaries, run a supervised and an unsupervised classification, present the steps and results of the classification. Also, create a mosaic dataset for the DEM, load the proper DEMs for your selected Township, create a Hillshade and clip the result to the shape of the selected Township. Produce multiple screen shots of the steps, include them in a MS Word document and write a brief description of the methods used in creating each. Save the document as ***ES470_RasterDataModel<MunicipalityName>.pdf*** and upload it to the corresponding D2L Dropbox.

VII. Submit your Work

Save the produced PDF (***ES470_RasterDataModel<MunicipalityName>.pdf***) and upload it to the corresponding D2L Assignment

VIII. Grading Rubric

- Clipping Satellite Image to Study Area..... 10pts
- Supervised Classification..... 10pts
- Unsupervised Classification 10pts
- Creating DEM Mosaic..... 10pts
- Clipping DEM to Study Area..... 10pts
- Reclassification of DEM..... 10pts
- Hillshade Creation..... 10pts

R,G,B	Potential Information Content
4,3,2	The standard "false color" composite. Vegetation appears in shades of red, urban areas are cyan blue, and soils vary from dark to light browns. Ice, snow and clouds are white or light cyan. Coniferous trees will appear darker red than hardwoods. This is a very popular band combination and is useful for vegetation studies, monitoring drainage and soil patterns and various stages of crop growth. Generally, deep red hues indicate broad leaf and/or healthier vegetation while lighter reds signify grasslands or sparsely vegetated areas. Densely populated urban areas are shown in light blue. This TM band combination gives results similar to traditional color infrared aerial photography.
3,2,1	The "natural color" band combination. Because the visible bands are used in this combination, ground features appear in colors similar to their appearance to the human visual system, healthy vegetation is green, recently cleared fields are very light, unhealthy vegetation is brown and yellow, roads are gray, and shorelines are white. This band combination provides the most water penetration and superior sediment and bathymetric information. It is also used for urban studies. Cleared and sparsely vegetated areas are not as easily detected here as in the 4 5 1 or 4 3 2 combination. Clouds and snow appear white and are difficult to distinguish. Also note that vegetation types are not as easily distinguished as the 4 5 1 combination. The 3 2 1 combination does not distinguish shallow water from soil as well as the 7 5 3 combination does.
7,4,2	This combination provides a "natural-like" rendition, while also penetrating atmospheric particles and smoke. Healthy vegetation will be a bright green and can saturate in seasons of heavy growth, grasslands will appear green, pink areas represent barren soil, oranges and browns represent sparsely vegetated areas. Dry vegetation will be orange and water will be blue. Sands, soils and minerals are highlighted in a multitude of colors. This band combination provides striking imagery for desert regions. It is useful for geological, agricultural and wetland studies. If there were any fires in this image they would appear red. This combination is used in the fire management applications for post-fire analysis of burned and non burned forested areas. Urban areas appear in varying shades of magenta. Grasslands appear as light green. The light-green spots inside the city indicate grassy land cover - parks, cemeteries, golf courses. Olive-green to bright-green hues normally indicate forested areas with coniferous forest being darker green than deciduous.
4,5,1	Healthy vegetation appears in shades of reds, browns, oranges and yellows. Soils may be in greens and browns, urban features are white, cyan and gray, bright blue areas represent recently clearcut areas and reddish areas show new vegetation growth, probably sparse grasslands. Clear, deep water will be very dark in this combination, if the water is shallow or contains sediments it would appear as shades of lighter blue. For vegetation studies, the addition of the Mid-IR band increases sensitivity of detecting various stages of plant growth or stress; however care must be taken in interpretation if acquisition closely follows precipitation. Use of TM 4 and TM 5 shows high reflectance in healthy vegetated areas. It is helpful to compare flooded areas and red vegetated areas with the corresponding colors in the 3 2 1 combination to assure correct interpretation. This is not a good band combination for studying cultural features such as roads and runways.
4,5,3	This combination of near-IR (Band 4), mid-IR (Band 5) and red (Band 3) offers added definition of land-water boundaries and highlights subtle details not readily apparent in the visible bands alone. Inland lakes and streams can be located with greater precision when more infrared bands are used. With this band combination, vegetation type and condition show as variations of hues (browns, greens and oranges), as well as in tone. The 4,5,3 combination demonstrates moisture differences and is useful for analysis of soil and vegetation conditions. Generally, the wetter the soil, the darker it appears, because of the infrared absorption capabilities of water.
7,5,3	This band combination also provides a "natural-like" rendition while also penetrating atmospheric particles, smoke and haze. Vegetation appears in shades of dark and light green during the growing season, urban features are white, gray, cyan or purple, sands, soils and minerals appear in a variety of colors. The almost complete absorption of Mid-IR bands in water, ice and snow provides well defined coast lines and highlighted sources of water within the image. Snow and ice appear as dark blue, water is black or dark blue. Hot surfaces such as forest fires and volcano calderas saturate the Mid-IR bands and appear in shades of red or yellow. One particular application for this combination is monitoring forest fires. During seasons of little vegetation growth the 7 4 2 combination should be substituted. Flooded areas should look very dark blue or black, compared with the 3 2 1 combination in which shallow flooded regions appear gray and are difficult to distinguish.
5,4,3	Like the 4 5 1 combination, this combination provides the user with a great amount of information and color contrast. Healthy vegetation is bright green and soils are mauve. While the 7 4 2 combination includes TM 7, which has the geological information, the 5 4 3 combination uses TM 5 which has the most agricultural information. This combination is useful for vegetation studies, and is widely used in the areas of timber management and pest infestation.
5,4,1	This will look similar to the 7 4 2 combination in that healthy vegetation will be bright green, except the 5 4 1 combination is better for agricultural studies.
7,5,4	This combination involves no visible bands. It provides the best atmospheric penetration. Coast lines and shores are well defined. It may be used to find textural and moisture characteristics of soils. Vegetation appears blue. If the user prefers green vegetation, a 7 4 5 combination should be substituted. This band combination can be useful for geological studies.
5,3,1	This combination display topographic textures while 7 3 1 may display differences in rock types.

Source: http://web.pdx.edu/~nauna/resources/10_BandCombinations.htm