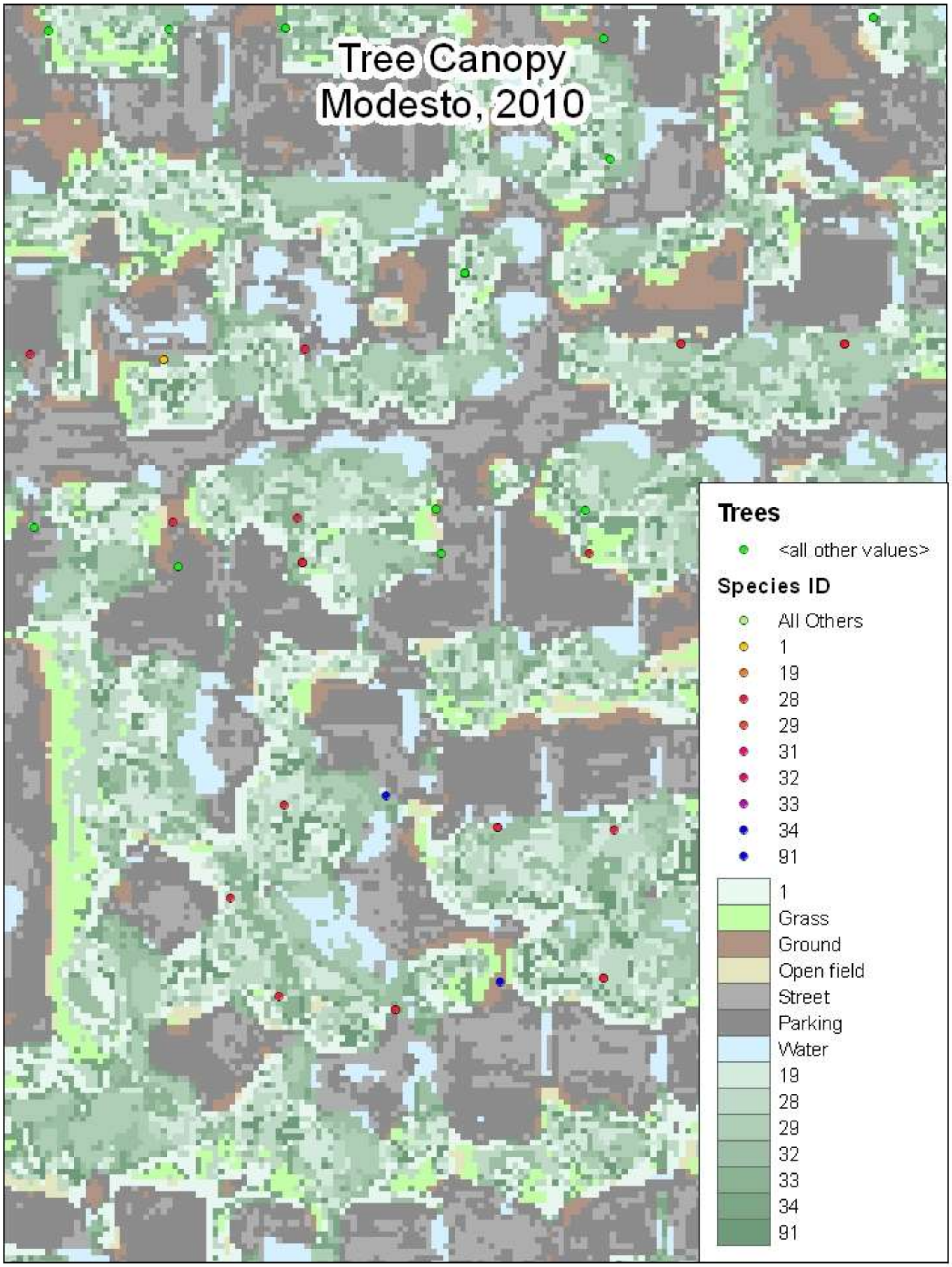
Project Summary

This report summarizes the efforts to attempt to classify the tree canopy for the City of Modesto. Tasks completed as part of this project included:

- Downloading Color Infrared images from the State of California Cal-Atlas Geospatial Clearinghouse
- Preparing a mosaic of the four images downloaded
- Extracting the area covered by the City boundaries
- Further extracting a subset of this data for testing the classification process and saving time
- Creating a Microsoft Access database of tree data provided by the City and querying the data to determine the number of different species of trees identified by the City
- Creating a file geodatabase and feature classes for the classification work
- Creating spectral signature polygons for a variety of feature classes in ArcGIS
- Creating spectral signatures from the polygons in ArcGIS
- Evaluating the spectral signatures using the Dendrogram routine in ArcGIS
- Running the Maximum Likelihood Classifier in ArcGIS
- Evaluating the results of the classification process

The work resulted in new image that showed the tree canopy but did not identify the tree species. A portion of the image is shown below. This portion of the overall image is enlarged to show the tree canopy at the neighborhood level for an area selected at random. Examination of the image will reveal that the tree canopy shows multiple tree species for individual trees, so this part of the classification did not work as desired.


Tree Canopy
Modesto, 2010
Trees
<all other values>
Species ID
All Others
1
19
28
29
31
32
33
34
91
1
Grass
Ground
Open field
Street
Parking
Water
19
28
29
32
33
34
91

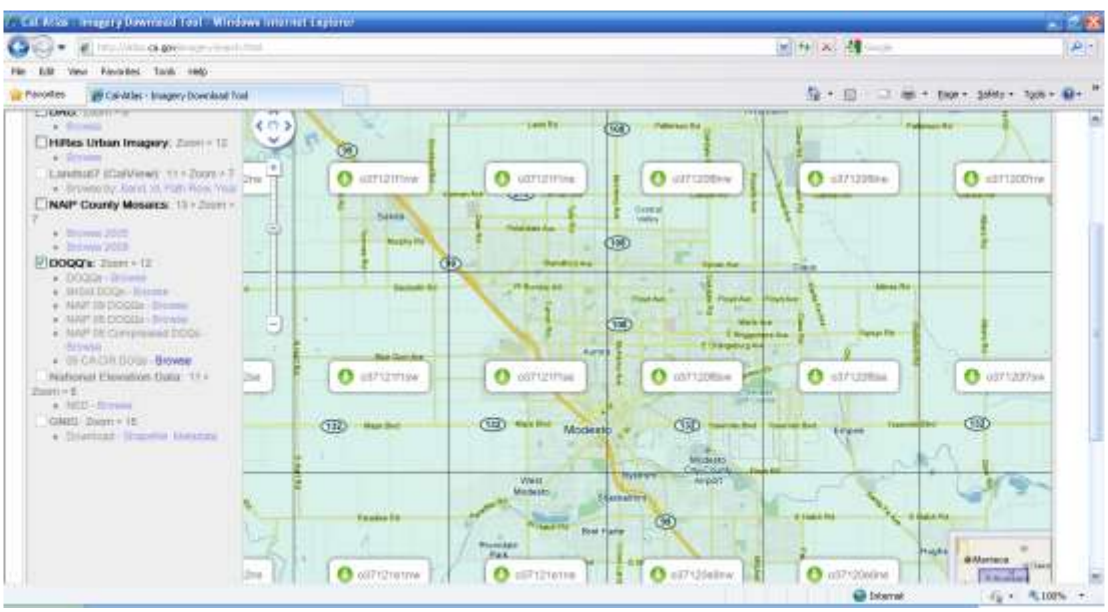
Purpose

My project was to prepare a tree canopy for Modesto. I was interested in doing this project because I have a number of datasets for Modesto. This includes a layer called trees, where the City provided me with point data showing the location of trees. In examining this data, I found that I have tree species identified for me which I will use for field reference information during the classification process and to compare against my image analysis project results.

Image processing tasks and methods used

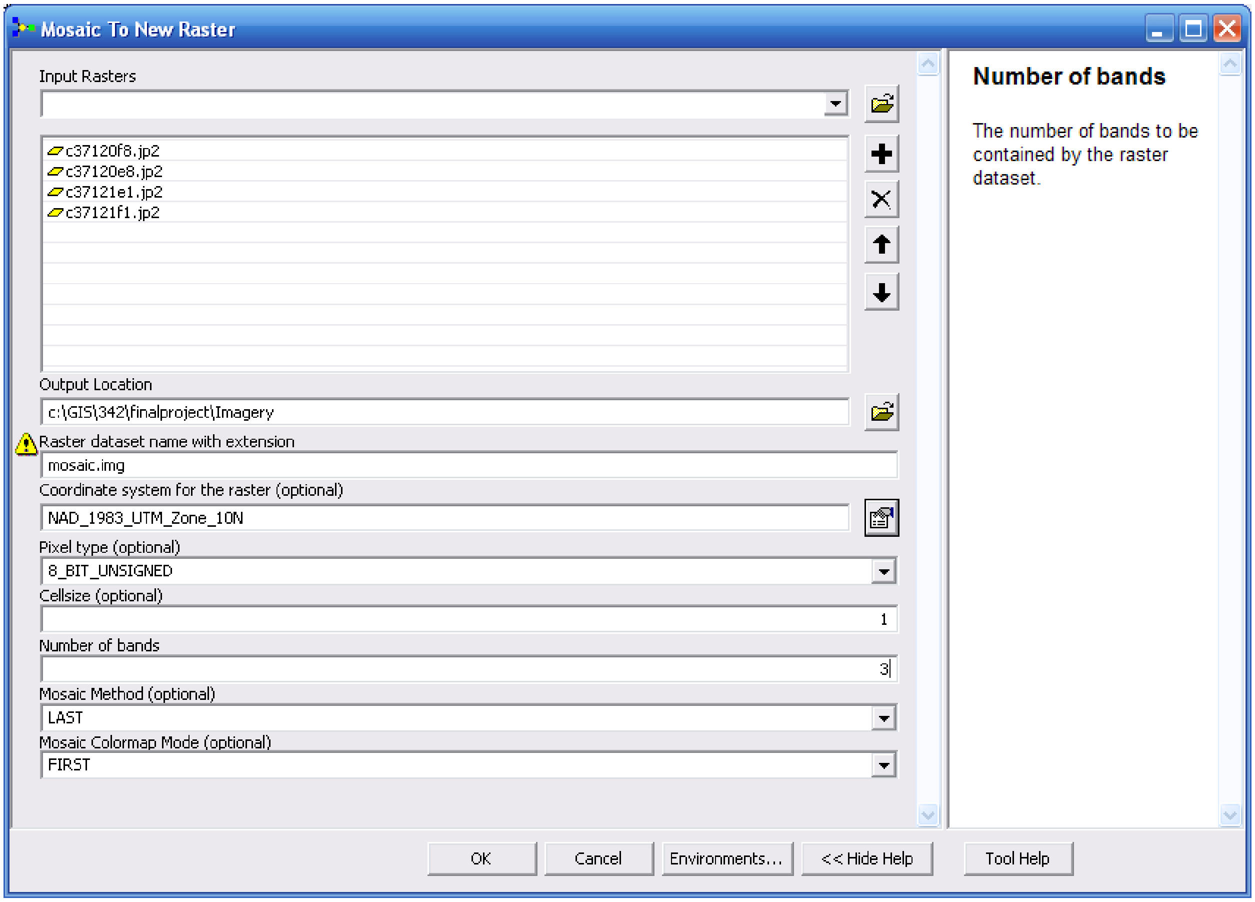
From the Cal-Atlas – Imagery web site I found that there were 9 possible images that I would need to get full coverage of the City of Modesto. These are listed below:

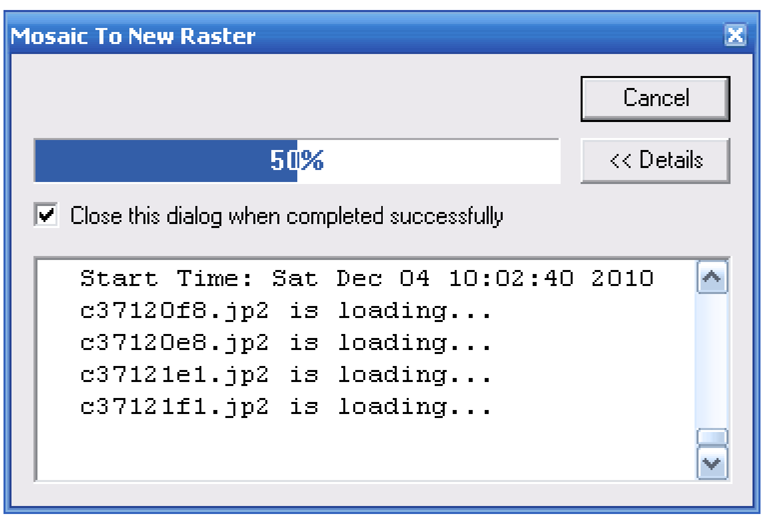
- o37121e1ne
- o37121f1se
- o37121f1ne
- o37120f8nw
- o37120f8ne
- o37120f8sw
- o37120f8se
- o37120e8nw
- o37120e8ne



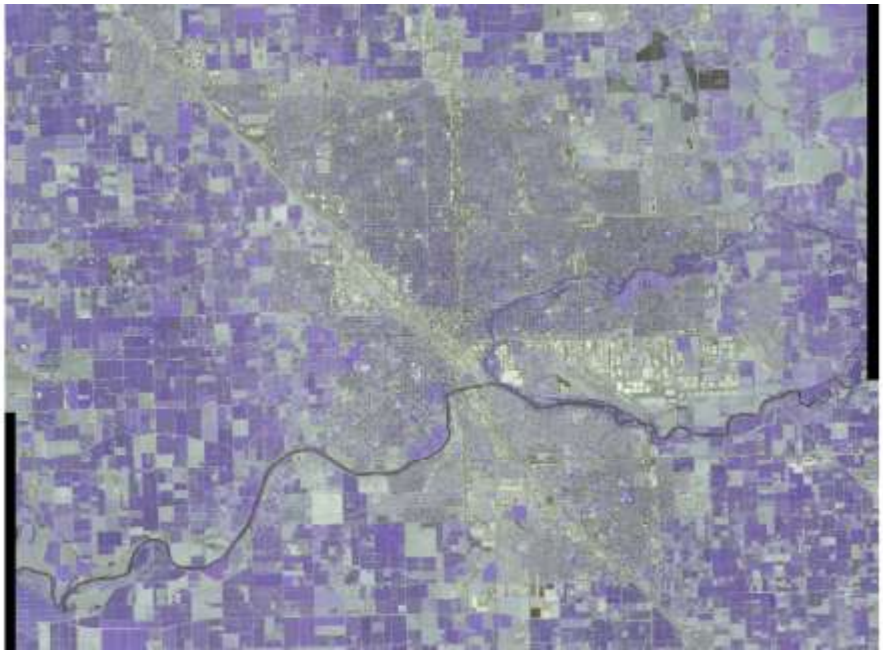
When I went to download the files, I found that four of the 2005 Color Infrared DOQs covered the entire area and more.

After downloading the four jp2 images, I had to mosaic them together since Modesto seems to be on the edge of all four images.

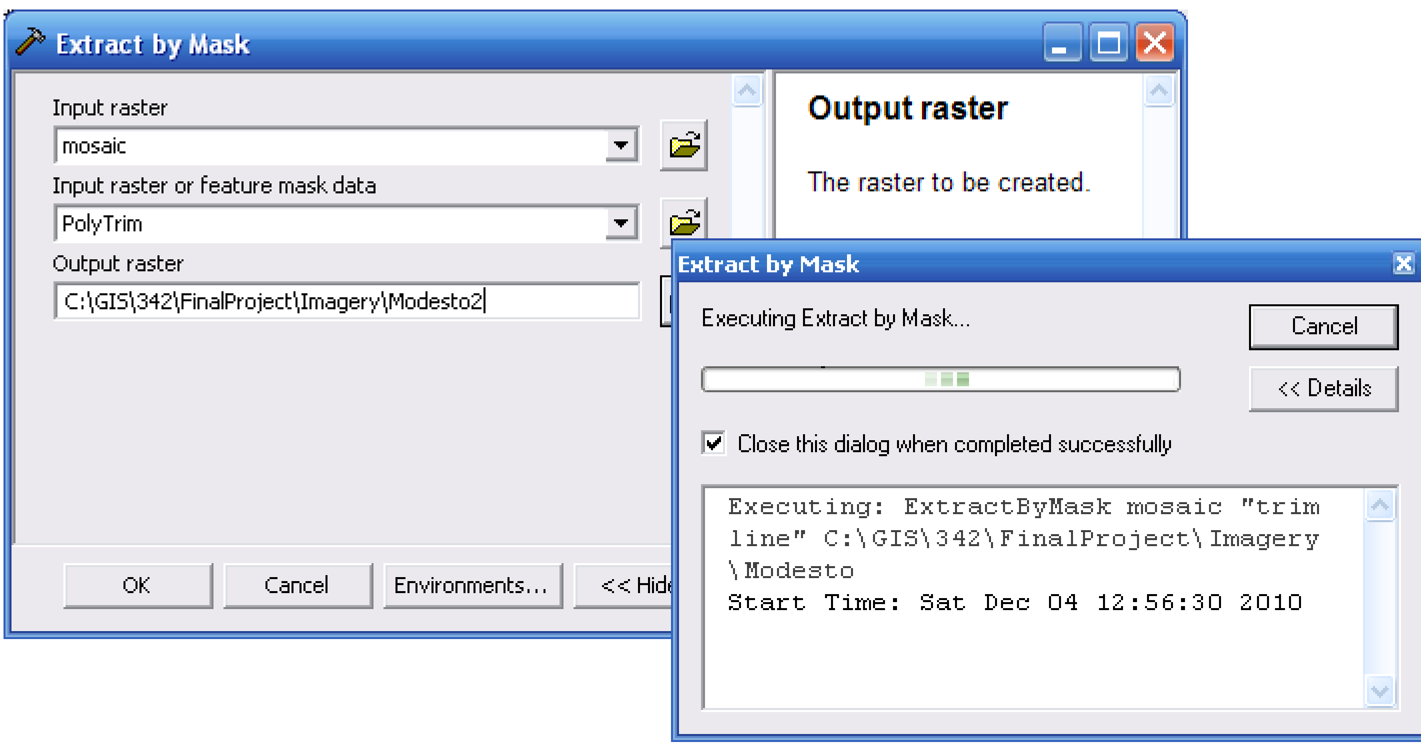




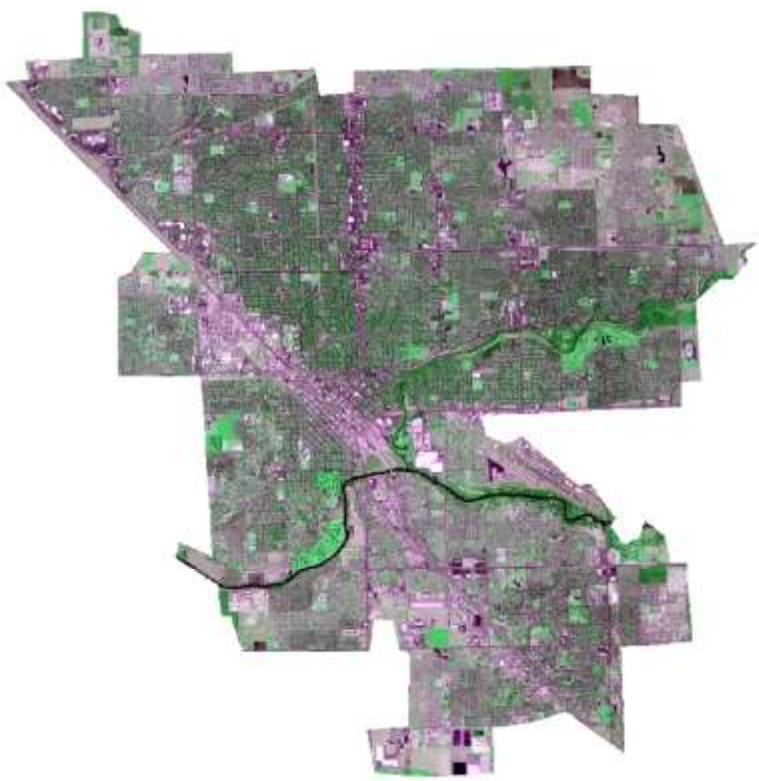
I tried to mosaic the images together during class work time using the data downloaded to my jump drive, but had to abort this operation after about 2-1/2 hours when the Ethan Center was closing. I performed this operation again at home on my PC with the data on the internal hard drive and it worked smoothly and took about an hour to process.



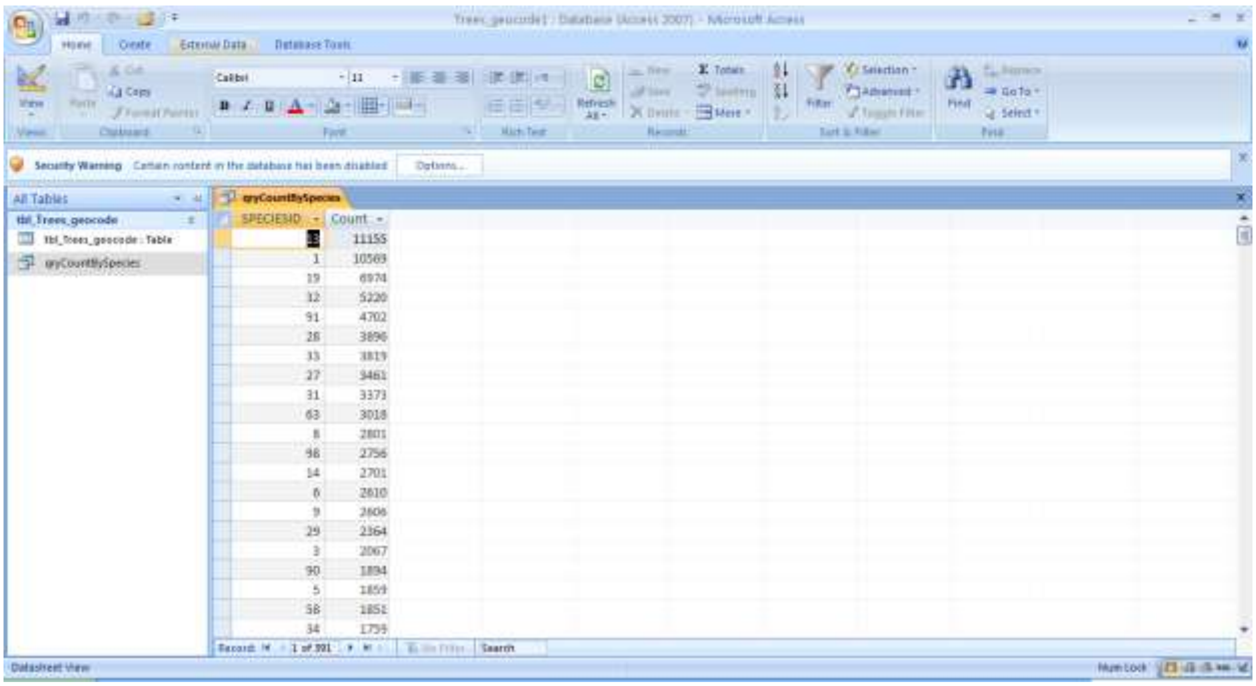
After mosiac-ing the images together, I extracted the Modesto area out using the Extract by Mask tool.



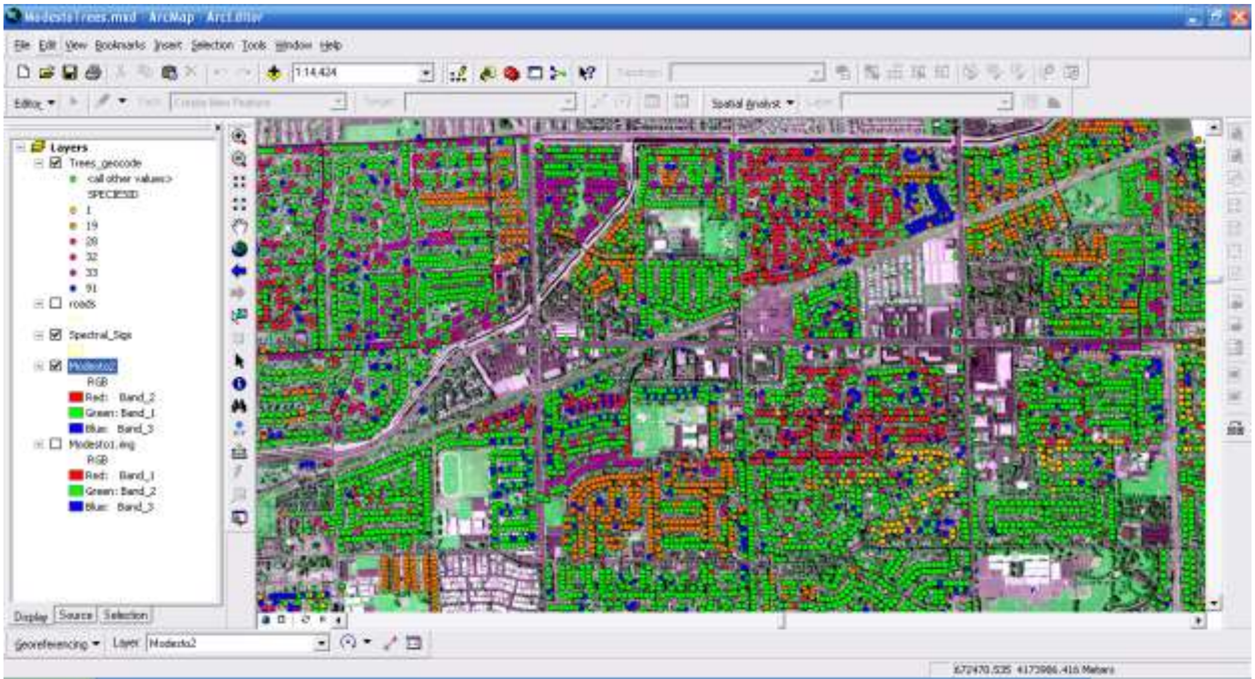
This operation took between a half hour and three quarters of and hour. When finished, the new image looked like the one shown below. An interesting thing that I learned during this step what that you need have a polygon feature class as the mask, or you will only get an extraction on the line if you try using a line feature class for the mask.

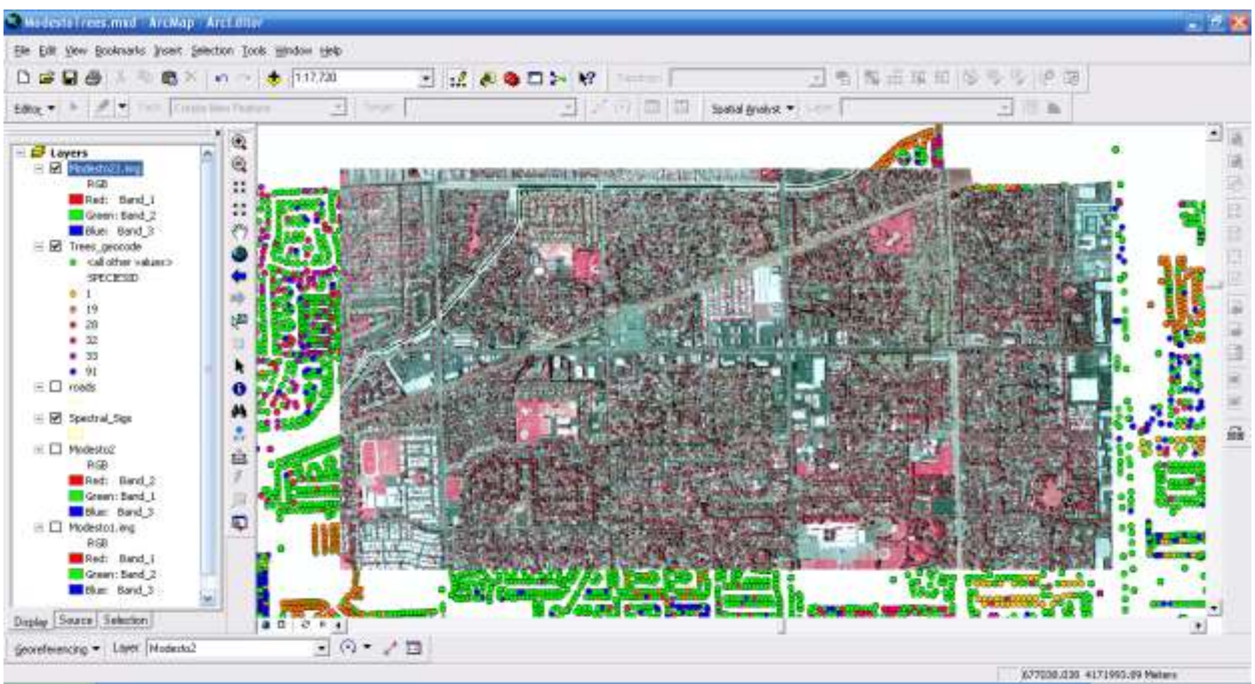


In the next step, I imported the tree information that I had received from the City and queried out the tree species by count so that I would know what species I was working with and how many trees of each species they had identified. This showed that the there were 391 different species which ranged from one tree of the species to over 10,000 trees in one of the species. There were 33 species with more than 500 trees and 24 species with more than 1000 identified trees. I knew that I did not have the time to try to classify this many tree species, so I limited my work to trying to classify the top 6 categories. Note: I requested a translation of the species ID and code to tree species names, but I have not received this information yet. After reviewing this data further, I think that the species with the most number of trees is the catch all "not identified" category.

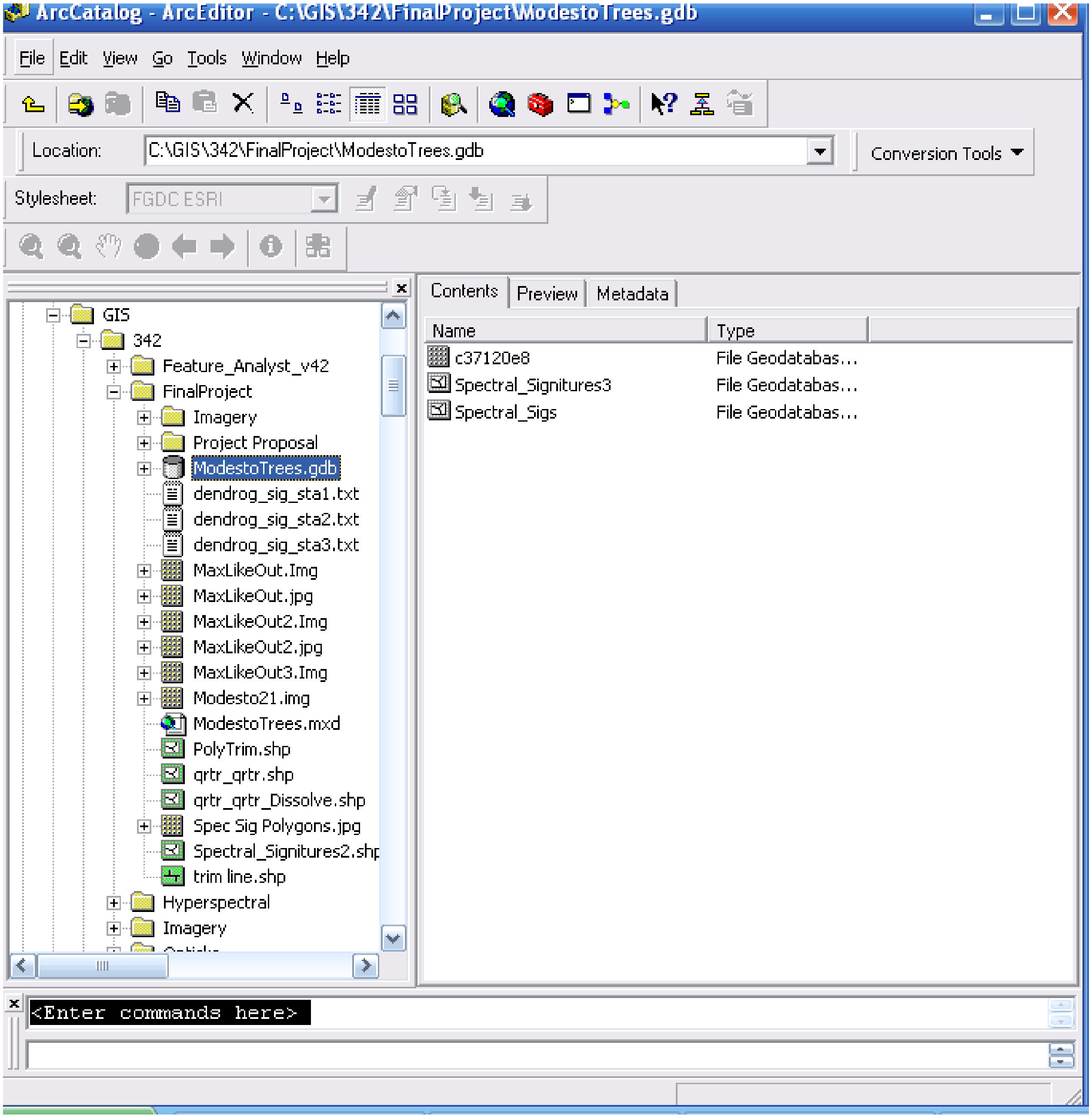


Following the suggestion of my instructor, Nathan Jennings, I then selected a smaller area to use for testing the classification process so that I did not get stuck with extremely long classification times. To select the test area, I symbolized the trees by species and displayed the species with highest number of trees. After doing this, I examined the image and selected an area which had a relatively large number of each of the species that I had selected. Shown below is an image showing the area selected with the different tree species shown and a second image showing the image subset.





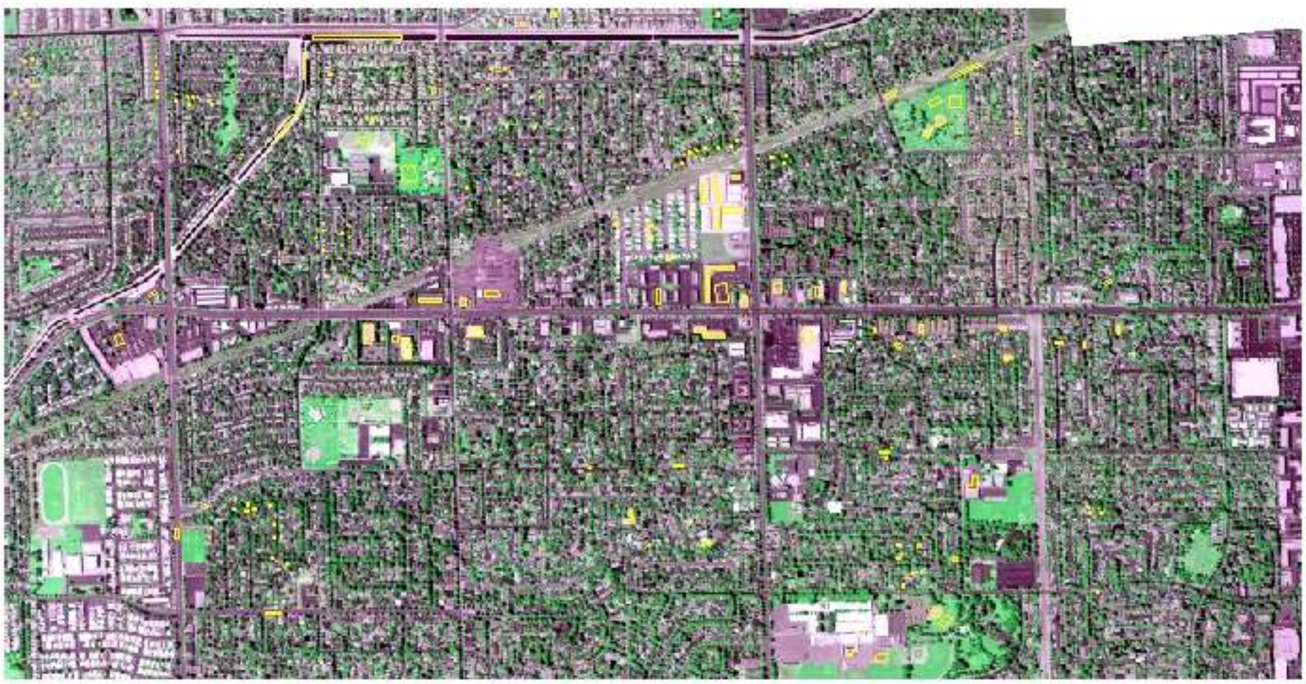
Using the image subset, I created and evaluated spectral signatures and then performed a supervised classification using ArcGIS. After creating a file geodatabase and new feature class called spectral_sigs, I created spectral signature polygons for a variety of feature classes. To try to get the software to learn the physical features from the trees, I included categories for streets, roofs (residential houses), industrial and commercial buildings (roofs), grass, bare ground, open fields, water (canals), and swimming pools.



Categories	**Signature Numbers**
Grass	500-509
Ground	530-533
Open field	560-561
Street	601-608
Parking lot	680-688
House (roof)	700-729
Ind or comm.	800-814
Water	900-904
Pool	910-919

Tree Species	**Signature Numbers**
1	101-108
19	1901-1908
28	2801-2811
29	2901-2909
32	3201-3210
33	3301-3309
34	3401-3407
91	9101-9109

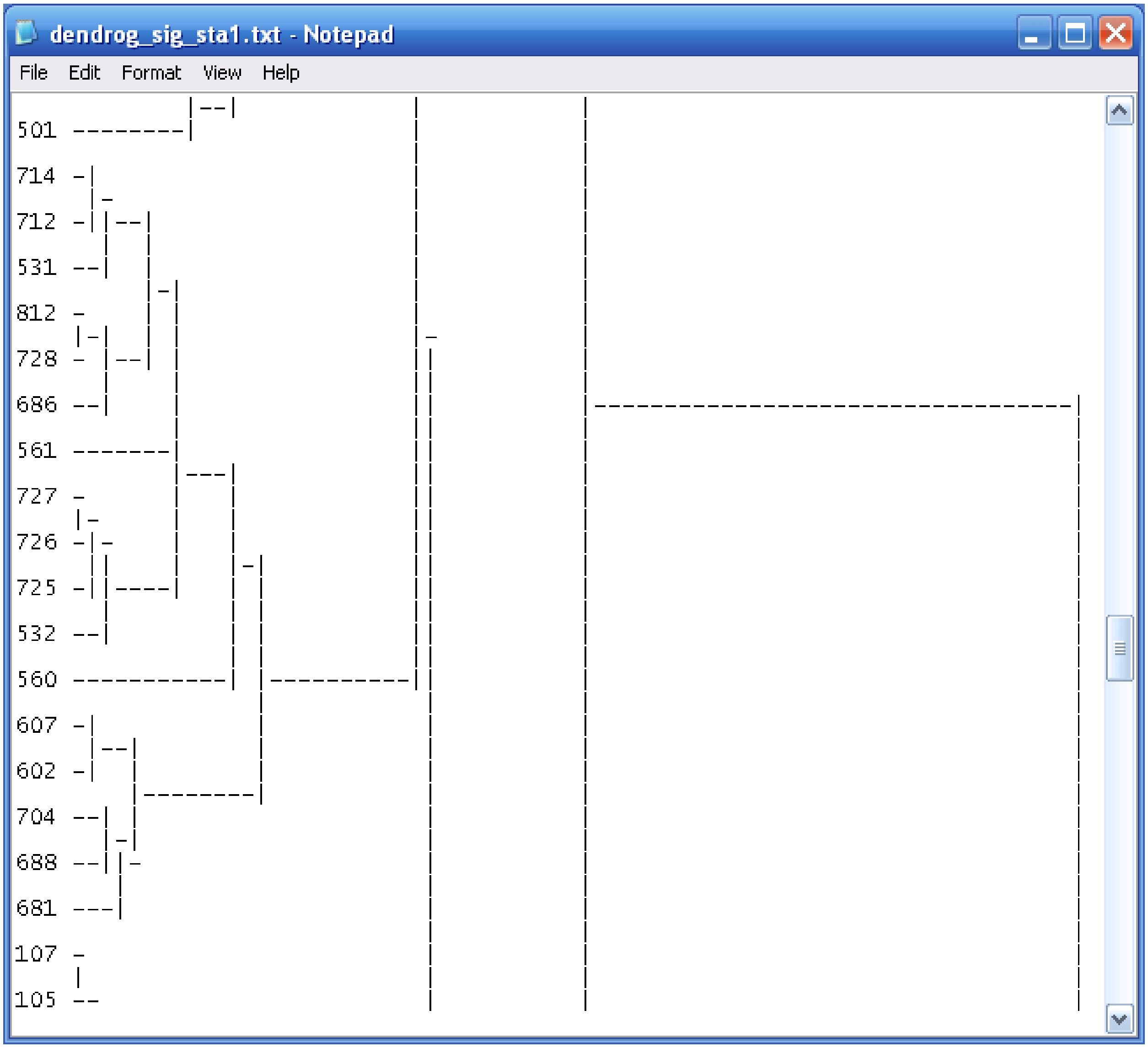
The image below shows the locations of the spectral signatures polygons (shown in yellow outline).



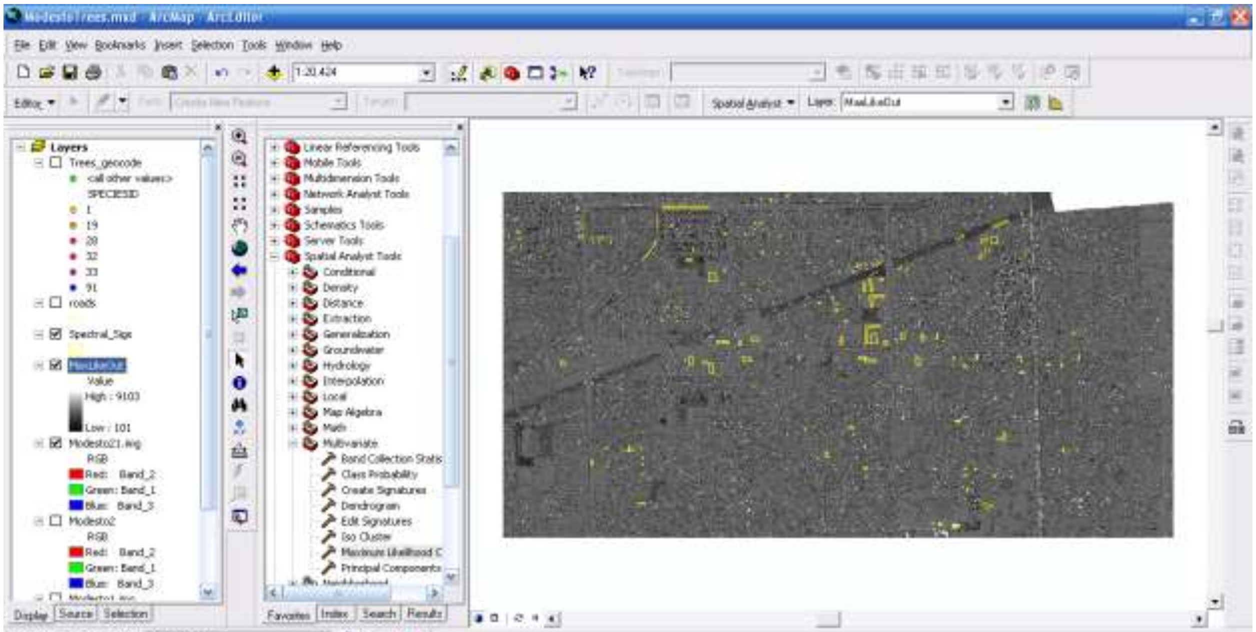
After creating the spectral signatures polygons, I created spectral signatures from the polygons using the Spatial Analyst Toolset—Multivariate Toolbox - Create Signatures Tool. I then performed a Signature Evaluation using the Dendrogram tool from the Spatial Analyst Toolset—Multivariate Toolbox.

To my surprise, the classes showed confusion with various spectral signatures polygons showing similar characteristics with spectral signatures polygons from dissimilar categories. With the large number of spectral signatures polygons, it was not easy to identify specific spectral signatures polygons that may be causing problems.

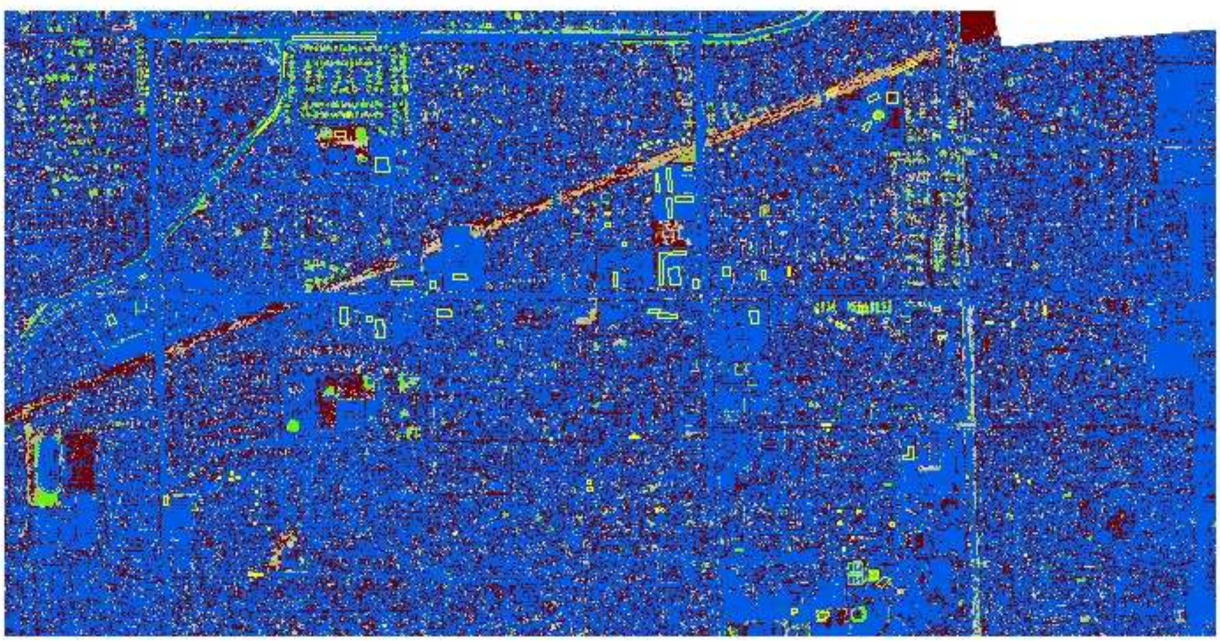
For example, as shown in the Dendrogram output below, while the different tree species were showing similar spectral characteristics, the bare ground (530s), open fields (560s), streets and parking lots (600s and 680s) were showing confusion with the House (700s) and Industrial and commercial buildings (800s).



I then ran the Maximum Likelihood Classifier tool from the Spatial Analyst—Multivariate toolbox to get a preliminary idea of the results.   This yielded the following image.

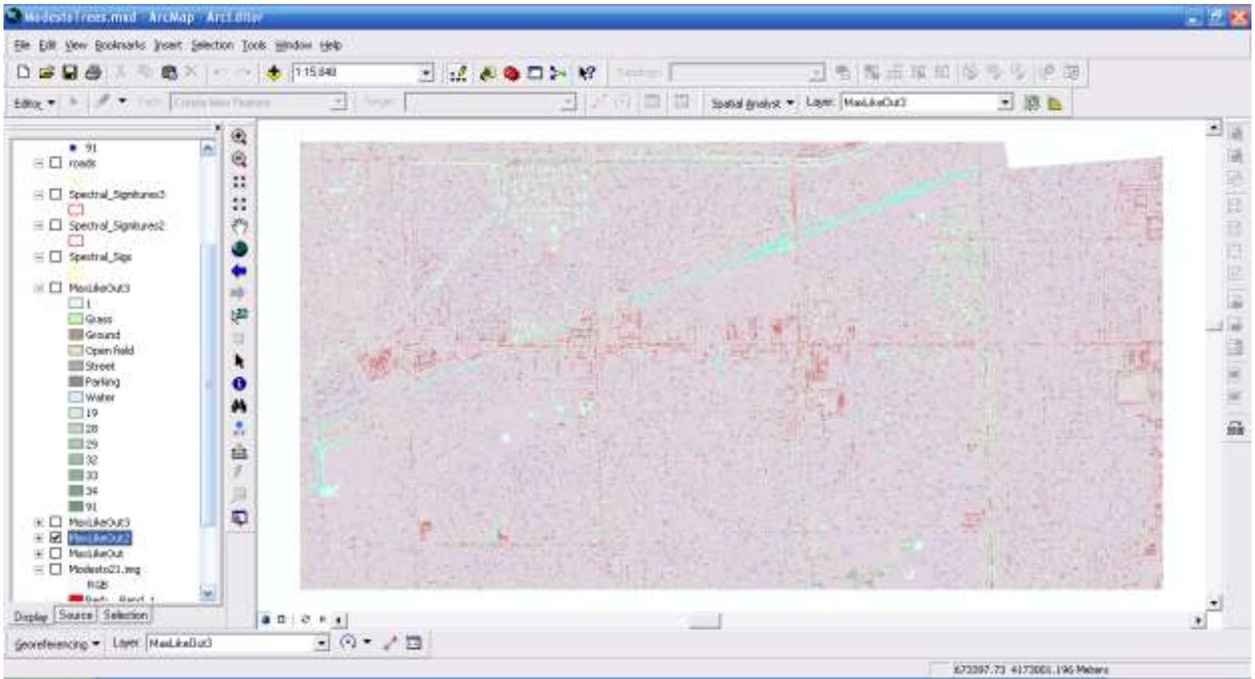


After symbolizing the categories, the image looked like this.

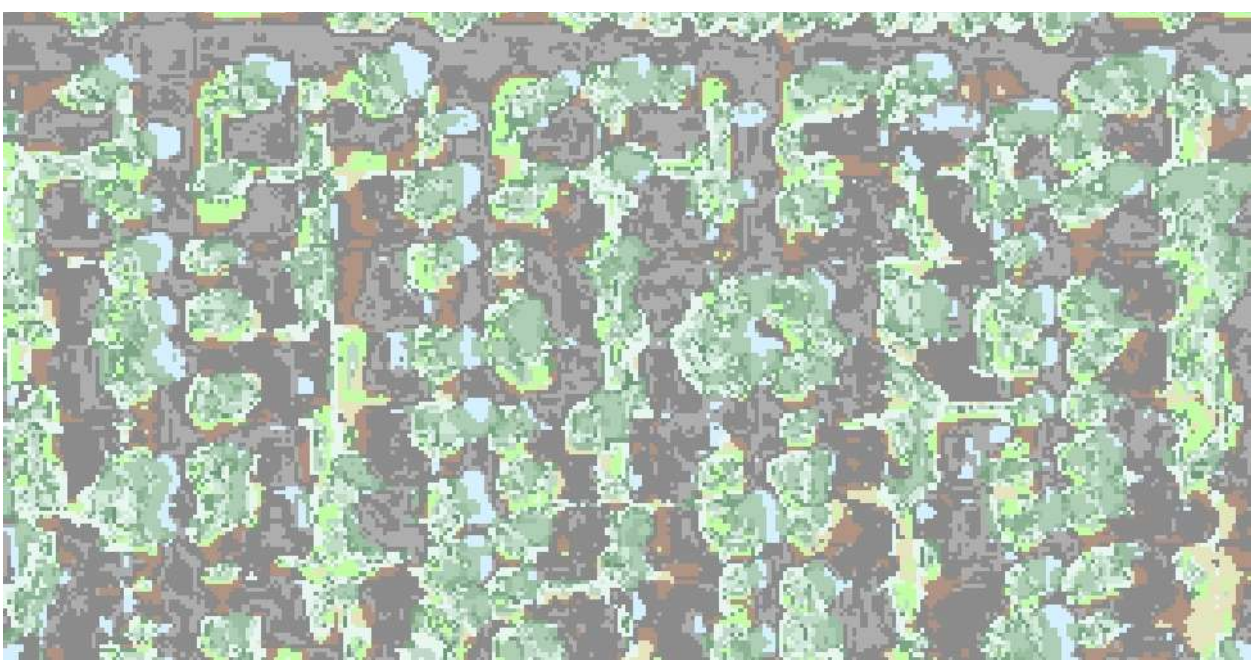


This image was highly biased toward the swimming pool signatures, which I had set the symbology to be a dark blue.

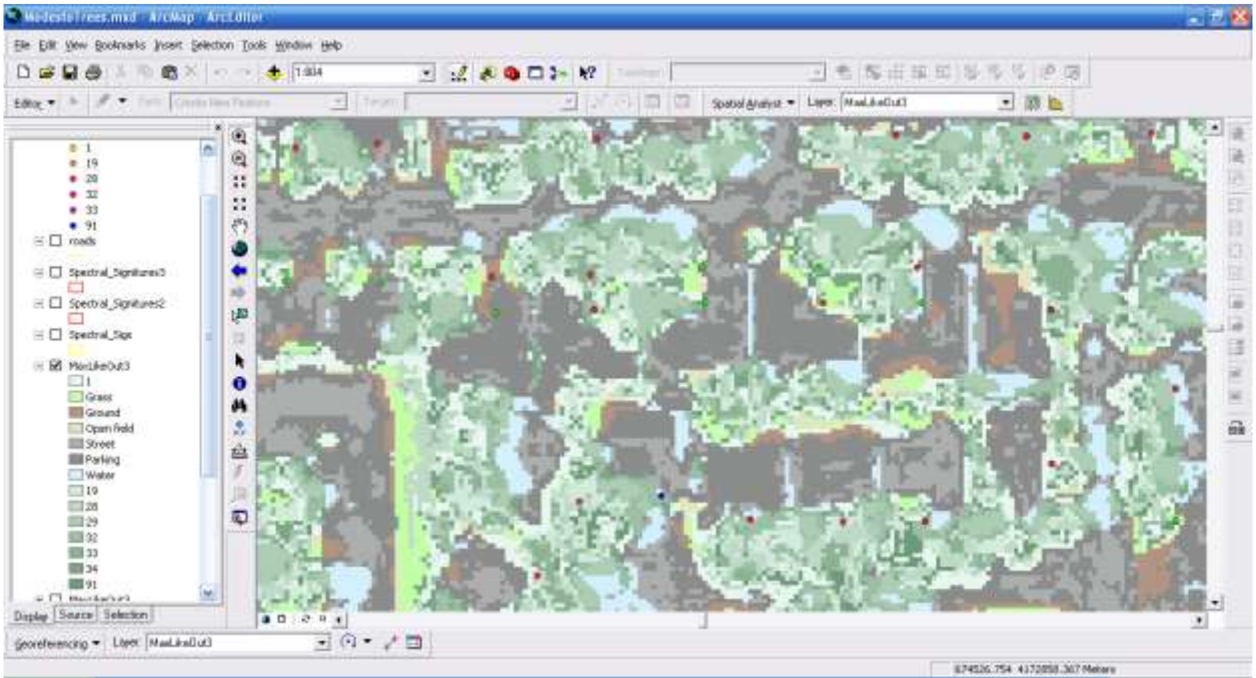
At this point I went back and modified the spectral signatures polygons to delete the swimming pool polygons, which seemed to have a lot of different colors when I selected them. I worked through the steps again and, as shown below, got another image that was highly biased toward one category.



Since the multiple categories seemed to be confusing the classification process, I tried deleting the roofs (residential houses) and industrial and commercial buildings (roofs) polygons and worked through the process again. This time it yielded results that seemed to show the tree canopy, which is what I had set out to do.



When I examine the results at the pixel level, however, as shown below I observe that the different tree species that I was hoping to differentiate are combined together within the classified trees. As shown below where I used a green color ramp to show the different tree species, there are multiple shades of green showing indicating that the tree species were not uniquely classified.



Had I gotten better results, my next step would have been to perform an accuracy assessment on the data. Since the results weren't strong enough, however, there was no point in going forward with this step.

Difficulties/issues encountered and how you resolved them

Most of the technical difficulties and issues encountered were discussed above. The primary technical difficulty was the issue with categories being apparently spectrally similar even though they appeared to be different when the spectral signatures polygons were prepared.

The other issue was time. Considerable time was spent preparing the mosaic of the 4 images downloaded from the Cal-Atlas site and then extracting the area of interest out of this mosaic. Considerable time was also spent preparing the spectral signatures polygons where I wanted to get about 10 polygons for each category that I selected. This required preparing 167 polygons which took the better part of a day.

Output, the interpretation(s) made on the output and/or intermediate image files.

The interpretation was that because of apparent spectral similarity, I was not able to perform a classification that would separate out the different tree species within the time I had available and with the tools available to me at the time of the classification. Nathan Jennings suggested when I shared my results with him that for the resolution I was using it would have been better suited to the Feature Analyst algorithms. Unfortunately, however, Feature Analyst is not available for student use.