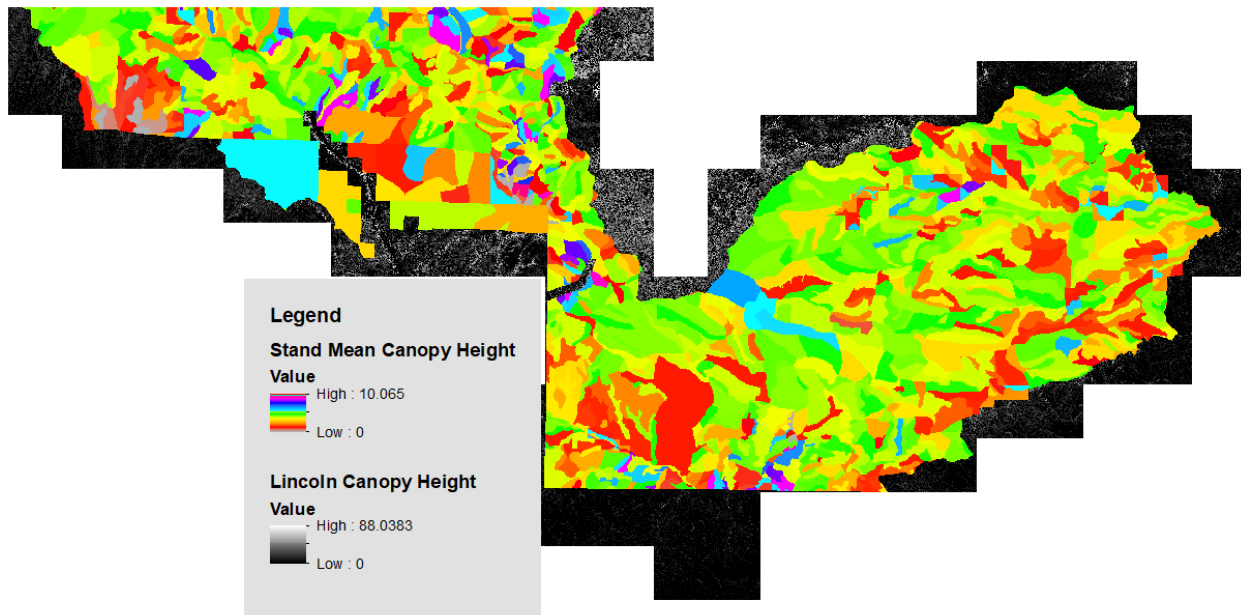# EXERCISE 3b
Zonal Statistics

Objectives

One tool that is very simple to use and very useful for spatial analyses is the Zonal Statistics tool. With the Zonal Statistics tool, a statistic is calculated for each zone (typically a polygon) based on values from a raster. For example, you may want to use a shapefile of delineated forest stands as the zone dataset in order to extract statistics (mean, range, standard deviation) from a lidar canopy height raster. Or perhaps you want to use watershed boundaries and a raster of fire severity to calculate the mean severity of the burn per watershed. The possibilities are endless. Basically, if you have a boundary and a raster with values of interest that you wish to summarize, you can do so using the Zonal Statistics tool.

Required Data

- **Lincoln_Stands.shp -** A zonal shapefile that contains data about the different forest stands in the study area. This shapefile does cover both the North and South of the Lincoln National Forest Study Area.

- **Lincolnnf_south_chm_1M.img –** Canopy height layer with 1-meter spatial resolution and vertical units in meters.

Prerequisites

- It is recommended that you are somewhat proficient using ArcMap however it is not necessary for this exercise to have an extensive knowledge of ArcMap.
- This exercise also assumes that you have the Spatial Analyst extension activated, and the toolbar displayed and docked in ArcMap.

Table of Contents

Part 1: Set up ArcMap

A. Start ArcMap

1. If it is not already open start ArcMap by clicking on the Start button and navigating to **All Programs, ArcGIS** and then open **ArcMap.**

2. Open a blank map.

B. Add The Data

1. Click the **Add Data** button.

2. Navigate to **South_Data\Lidar** and add **lincolnnf_south_chm_1M.img**

3. Click the **Add Data** button again, and navigate to **South_Data\Vector** and add **Lincoln_Stands.shp**

C. Change Appearance of Shapefile

1. Double click the **Lincoln_Stands.shp**, in the **Table of Contents** to open the **Layer Properties.**

2. Click the **Symbology** tab.

3. Double click the polygons fill color in the symbol section.

4. In the Symbol Selector window that opens, click **Fill Color:** and click **No Color.**

5. Click **Outline Color:** and select a **Red** color.

6. Adjust the **Outline Width:** to 1.00.

7. Click **OK** to close the symbol selector window and click **OK** again to close the Layer Properties window. Your display should look similar to the following graphic.
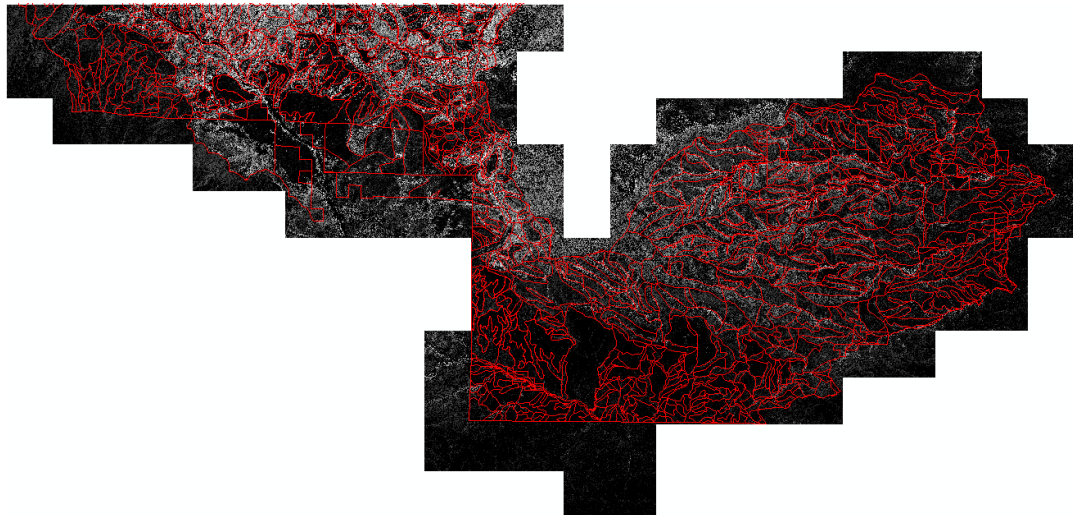
Part 2: Zonal Mean for Lidar Canopy Height Raster

A. Locate Zonal Statistics Tool

1. If it is not already open, open **ArcToolbox** now by clicking the **ArcToolbox** button.

2. From the **ArcToolbox** navigate to **Spatial Analyst Tools > Zonal > Zonal Statistics.**

3. Double click **Zonal Statistics** to open the dialog window.

B. Set Input Parameters, Run Process

1. Choose **Lincoln_Stands.shp** as your input for the Input raster or feature zone data.

2. Set the Zone field as **POLY_ID.**

3. Choose **lincolnnf_south_chm_1M.img** as the Input value raster.

4. For the Output raster choose **South_Data\Outputs** as your output location and name it **South_MeanCanopyHeight.tif.**

> ***Important Note:*** *ESRI Grid files have a character limit of 13. So, if you wish to leave your output in Grid format you will have to make sure the name length is 13 characters or less or change the file type to .img or .tif.*

5. Set the **Statistics type** to **MEAN** (Note the list of statistics you have the option of choosing from the drop-down menu).

6. Make sure the **Ignore NoData** in calculations box is checked.

7. Click **OK** to run the process.

   i. The output is the Mean canopy height calculated from the raster within the bounds of the vegetation zones.

   ii. Note: You could receive a warning about not all zones were rasterized, which is fine since we are working with stand polygons for the whole Lincoln National Forest area of interest and our CHM is for the southern portion.

C. Explore The Output

1. Change the **symbology** color ramp of your **South_MeanCanopyHeight.tif** layer so it is easier to see the differences among zones. Double Click the color ramp next to the layer.

2. In the color ramp dropdown menu, select a color ramp to better visualize the layer.

3. Click **OK**.

> ***Note:*** *if you did not save the output as a TIFF, you may need to change your Symbology to Classified and then choose a specific color ramp. Simply changing the color ramp for an Arc Grid output (default if no extension was specified for output) will not properly distinguish between the different stands.*

4. Use the **Zoom** and **Pan tools** to explore your output.

5. You can further explore the output by using the **Identify** tool.

   i. Be sure to change the **Identify from** field to your new output raster.

6. Single click on any of the output zones to identify the **mean canopy height** for that zone.
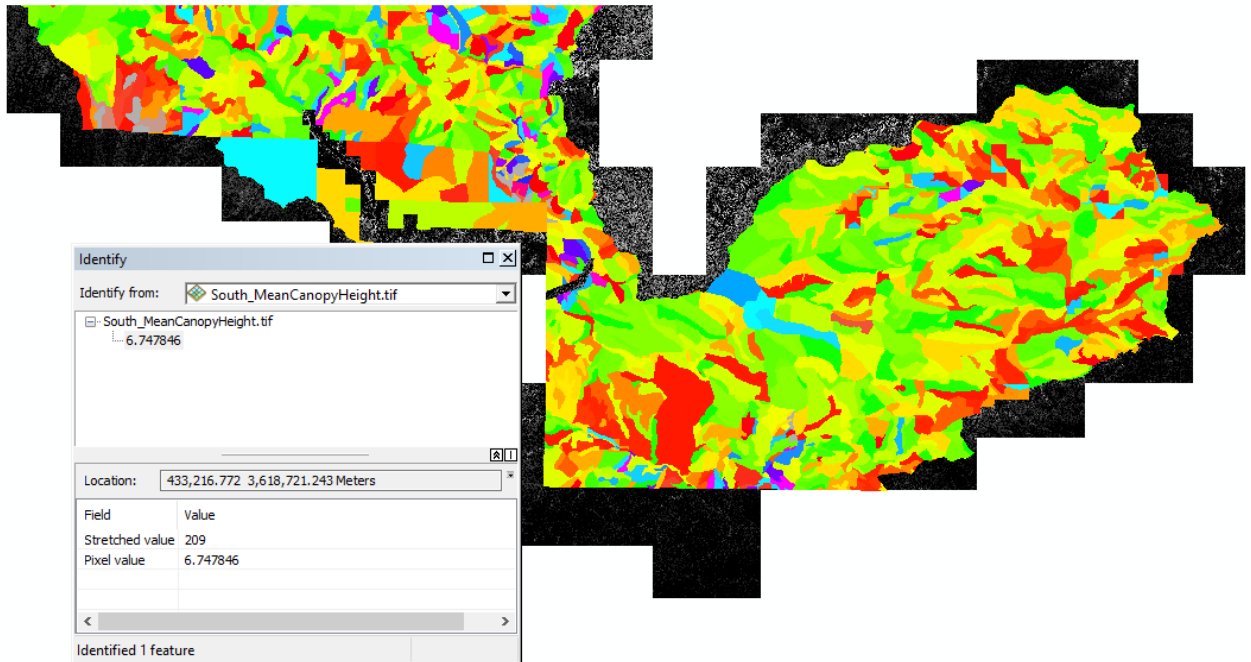
Part 3: Zonal Range for Lidar Canopy Height Raster

A. Locate Zonal Statistics Tool

1. From the **ArcToolbox** navigate to **Spatial Analyst Tools, Zonal, Zonal Statistics.**

2. Double click **Zonal Statistics** to open the dialog window.

B. Set Input Parameters, Run Process

1. Choose **Lincoln_Stands.shp** as your input for the **Input raster** or **feature zone data.**

2. Set the **Zone field** as **POLY_ID.**

3. Choose **lincolnnf_south_chm_1M.img** as the **Input value raster.**

4. For the **Output raster** choose **South_Data\Outputs** as your output location and name it **South_RangeCanopyHeight.tif.**

5. Set the **Statistics type** to **RANGE** (Note: The "range" in this case is a single value, representing the difference between the min and max value of that zone. The Min and Max values are not returned).

6. Make sure the **Ignore NoData** in calculations box is **checked.**

7. Click **OK** to run the process.

    i. The output is the Range of canopy heights calculated from the raster within the bounds of the vegetation stand zones. You may notice there are some really high values, this is because of a few outlier trees.

8. You may want to **adjust the color ramp** just as you did in the previous section to better visualize your outputs.
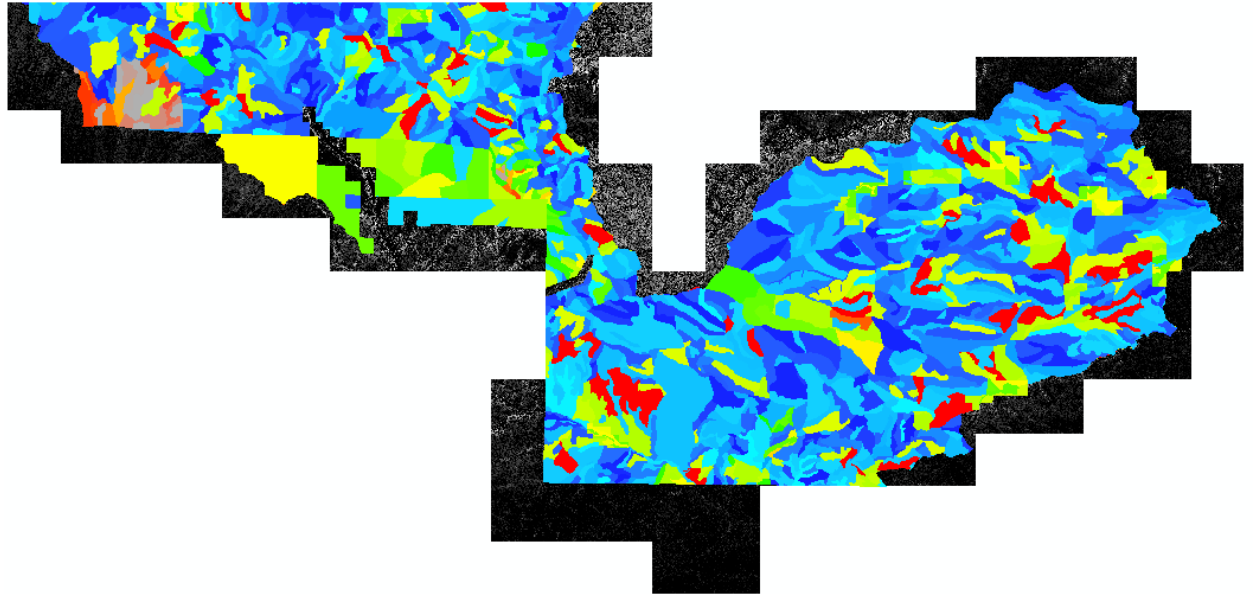
C. Explore The Output (See Above)

1. Change the symbology color ramp of your **South_RangeCanopyHeight.img** layer so it is easier to see the differences among zones.
2. Use the **Zoom** and **Pan tools** to explore your output.
3. You can further explore the output by using the **Identify** tool.
   i. Be sure to change the **Identify** from field to your new output raster.
4. Single click on any of the output zones to identify the **canopy height range** for that zone.

Congratulations! You can now effectively use a vector layer (e.g. shapefile) to extract statistical information from a raster of your choice. This is great tool to use when:

1. You need to understand how one layer relates to another.

2. You want to calculate the maximum tree height for each forest zone.

3. You want to calculate the mean elevation for each forest zone.

4. You need to know the average burn severity for each watershed.

These questions and their answers can be important when conducting more advanced types of analysis with your raster in a GIS. If you are interested in these zonal statistics but would like to incorporate a broader range of statistics all at once, explore the Zonal Statistics as Table tool.

This can create a suite of zonal statistics at once in tabular form. The output table can then be joined with your shapefile's attribute table for further spatial analysis.