**BIOL 533: Remote Sensing Plants in R**

**Lab 5 Activities**

**Activity A:** Plotting maps of topography and plants and making predictions.

1. Make sure you are able to open the RemoteSensingPlantsinR project on RStudio Cloud
2. Follow along with the tutorial to load packages, run through the code, and ultimately generate the DTM, DSM, and NDVI maps.
3. Run the lines of code in order to import the data (dtm, dsm, ndvi) as well as the ones to calculate the values of vegetation height (DSM – DTM), slope, aspect, and northness. You will see those objects display in the “Environment” tab in the upper right as a way of determining if these steps were successful. IF they were not, then in the bottom left “console” you will see some form of an error, usually in a glaring red color.
4. Plot these four maps and consider the patterns you observe. Paste the map image here:
5. Based on these maps, make some predictions about the correlations you might find.
	1. Which topographic metrics (the calculated values of slope, aspect, vegetation height and northness) do you think would have a stronger association with NDVI? Draw a graph of your predicted relationship. As an example, here are some examples of correlation figures you may draw (nothing fancy!):



* 1. Which topographic metrics do you think have a stronger association with vegetation height? Draw a graph of your predicted relationship.
	2. Overall, does it seem like vegetation patterns are influenced by topography (natural physical features of the land)?
	3. What if anything else stands out among these maps?

**Activity B:** Plot the relationships between topographic variables and vegetation patterns.

This section prompts you to evaluate the p-value of data. If you are in need of a refresher regarding interpretation of a p-value, here’s a great [online video explanation](https://youtu.be/KS6KEWaoOOE). For correlation analysis, we conduct a statistical test to determine if there truly is a relationship between the variables, such that we can use that relationship as a predictive factor with other statistical tools (such as regression).

So for our data, if there is a significant difference (e.g. p-value is less than 0.05), then we can state there is a significant linear relationship between the two variables being tested (e.g. northness (north) and NDVI) and we would be able to use a regression analysis to model the relationship between the two.  Also recall that correlation coefficient ("cor") value is telling us a bit about the type and strength of the correlation.  Please note we are not performing a regression analysis in this particular lab, I just wanted you to be aware of what the hypothesis testing outcome would suggest.

1. Follow along with the code to convert these raster layers to a data table and take a 1% subset to speed along plotting and analysis.
2. Plot the correlations between vegetation variables (y-axes) and topographic variables (x-axes). Paste the figure here:
3. Look at these scatter plots and make some predictions about which will have statistical relationships.
	1. Do any of these plots look like there is a strong relationship?
	2. Is it harder or easier to see relationships with large data sets like this? Why?
4. Calculate the correlations and record them in the following table.

You will do this by running each of the “cor.test” functions (there are several because they are comparing two variables at a time). Highlight a line and hit the “Run” button. The results will output in the “Console” window at the bottom of your screen:



Within the output, the correlation coefficient indicated by "cor" in the console window (A)(the table in #4 indicates it is an "R" value), and the p-value (B).

 

Table to complete:

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  |  | DTM | Slope | Northness |
| NDVI | r (cor) |  |  |  |
| p-value |  |  |  |
| Veg Height | r (cor) |  |  |  |
| p-value |  |  |  |

1. How do these relationships compare to what you expected when you looked at the maps and the scatter plots?
2. Based on your results, which of the correlations resulted in a correlation coefficient that is significantly different from zero, or no relationship? Use an an apha value of 0.05 as the cutoff to determine significance. These are the variables that exhibit a relationship that could be used in an additional step of regression (wherein we would use one of the variables as a predictor to determine an unknown dependent variable). A reminder about interpretation of p-values:

**If the test concludes that the correlation coefficient is significantly different from zero, we say that the correlation coefficient is “significant.”**

Conclusion: There is sufficient evidence to conclude that there is a significant linear relationship between x and y because the correlation coefficient is significantly different from zero. What the conclusion means: There is a significant linear relationship between x and y. We can use the regression line to model the linear relationship between x and y in the population.

**If the test concludes that the correlation coefficient is not significantly different from zero (it is close to zero), we say that correlation coefficient is “not significant.”**

Conclusion: “There is insufficient evidence to conclude that there is a significant linear relationship between x and y because the correlation coefficient is not significantly different from zero.” What the conclusion means: There is not a significant linear relationship between x and y. Therefore, we CANNOT use the regression line to model a linear relationship between x and y in the population.

You can simply state the variable pairs that show a significant difference by writing below: (e.g. NDVI and DTM, etc)

**That’s all! We will use these techniques to compare other regions next week!**