GEO401 Quantitative Project Analysis Instructions **FALL 2015**

Modeling Rare Plant Distributions in Michigan Today and Into the Future

The goal of the project is to learn how very simple species distribution models are made, using the "Climate Envelope" concept we discussed in class. Each student will choose 3 species of rare plants of Michigan, import data on their whereabouts into Excel then into R, and will use these locations, along with precipitation, temperature, and elevation data to estimate where the species might be found today, and where they might be found in 2100.

The following are very detailed instructions on how to do this. We will also step through this in class and spend 2 class days working on this, so don't freak out!

Note: You will be making three species distribution models. I highly recommend that you go through all the steps once, then start again, rather than trying to batch process all three.

(1) First, create a folder on your computer to store the data for this project. There should be **NO spaces in the entire directory path** (e.g. my data is stored in "C:/MSU_GEO401/data_project").
(2) Go to http://mnfi.anr.msu.edu/data/specialplants.cfm skim thorough the list of rare plants of Michigan and pick a plant that sounds interesting. Click on the "Explorer" link in the far right column to learn more.
(3) If you read the description and still like the plant, and *it has three or more counties listed in the "Occurrences" table with 'Year Last Observed' years after 1900*, proceed. If not, go back to #2 and try again.
(4) Highlight everything in the Occurrences table (but nothing else on the webpage), and copy this data (right click and select copy, or hit Ctrl+C or Cmd+C)
(5) Open a blank Microsoft Excel or other spreadsheet and paste this data into the table, so it looks just like the one on the webpage. It should look like this:

(6)   Reformat so it's just a simple table (remove fill, make all the text black), and rename the columns so there are no spaces but the column names are still understandable, using the names I use below (remember R is case-sensitive, so match the case, too!). Now it should look like this:

(7)   Sort by the "Year_Obs" column by highlighting everything in the table (easiest way to do that is just to click the upper left corner of the table), then go to Data > Sort and select "Year_Obs" from the "Sort by" dropdown menu. We're only interested in semi-recent observations, which, for this class, we'll define as observations from 1900 or later, so delete any rows with years before 1900 (highlight the entire row, right click and select delete).

(8)   Save this file, first as an Excel file, then as a .csv in the folder you created in #1. When you save as a csv Excel might give you an error message, but just say 'yes' and it should be ok. Now the data are ready to be read in to R! You can check this by looking at it in a text editor – should just be some number of lines of text, with each thing (County, Occurrences, Year_Obs) separated by a comma, with no commas at the end of the lines. Like this:

(9)   Before we do that, go to d2l.msu.edu and in the GEO 401 folder called 'Quant Project' > 'Analysis' and download the .R file. Then go to the 'Climate Data' folder and download the nine .tif files.

ALL OF THESE FILES SHOULD BE SAVED IN YOUR DESIGNATED FOLDER!

About these data: All of these files are outputs from a climate model – specifically the US National Center for Atmospheric Research's Community Earth System Model (CESM), run with biogeochemistry (carbon and nitrogen cycles) turned on (hence, CESM-BGC). Even the 1980-2000

data are from the model – running a model over a historic time period allows researchers to compare what the model says to our actual historical data. For our class, though, we're just treating the 1980-2000 data as 'truth'. This is just one of >30 earth system models that was used in the Coupled Model Intercomparison Project round 5 (CMIP5) and then 'downscaled' (as discussed in class) for the continental U.S., using two different Representative Concentration Pathways (RCPs). Monthly data were downloaded from here: https://cds.nccs.nasa.gov/nex/ and then processed to be maximum and minimum temperatures and average annual precipitation. The research papers in the 'readings' folder in the 'Quant Project' folder on D2L give more info about these data, which might be useful for your paper. If you are feeling super adventurous, I'm happy to share details of how I got processed these data, so you could do the same project again using output from a different climate model.

ON TO THE ANALYSIS!

(10) Assuming you've already installed R and RStudio and have confirmed that they function properly, open RStudio then open the .R file you downloaded from D2L.
(11) Proceed with the lab in R following the instructions in the code.