HW 2 ~ Working with Spatial Data in R

Due Wednesday, Sept. 14, 2022 via GitHub

Assignment: HW#2 is designed to build your skills at working with, manipulating, and plotting spatial data in R. You are provided with a shapefile and must download a set of climate rasters and species occurrence data. You will then perform various tasks using these three datasets. You will be graded on your ability to produce clean, well commented R code that performs the tasks listed below. The hope is that I will be able to run your code without any errors. You can use R Markdown if you wish, but you are not required to do so. When you are done, push your code to GitHub, following the instructions provided in the document: `mees698C.submittingHW.pdf`.

Keep in mind this is a HW assignment, not an exam, so please do not hesitate to ask questions if you get stuck. Some of these tasks are challenging! Refer to the advanced spatial data tutorial we started in class, which provides examples of most of the steps required to complete this assignment.

Figure 1: A small grass tree in Western Australia

1. Use the `raster::getData` function to download the Worldclim climate data set at 2.5 arc-minute resolution. You want the bioclimatic variables (use `var="bio"`, see `?getData` in the `raster` package).
2. Make a raster stack of the `bio10`, `bio11`, `bio18`, and `bio19` bioclimatic variables only and clip this raster stack to the *outline of Australia (not the extent)* using the shapefile provided with this assignment. (Note that the shapefile also contains New Zealand, so you will have to do something about that before you perform the clipping operation, among other things. . .).
3. Use the `dismo::gbif` function to download records for the Austral grass tree (*Xanthorrhoea australis*). Clean up the resulting data frame by removing records without geographic coordinates & those that fall outside the Australian mainland. Convert the data to a `SpatialPointsDataFrame` with the correct CRS and containing only these attributes: `acceptedScientificName`, `institutionCode`, `lon`, `lat`, and `year`. Save your `SpatialPointsDataFrame` as a shapefile.
4. Make a simple map of the cleaned species occurrence records from GBIF, using a color ramp or

symbolization scheme to indicate the year the record was collected. Make sure to include the polygon of Australia and plot `bio10` as the background. **NOTE:** All data in this map should be in the original projection of the Australia & New Zealand shapefile (i.e., not WGS84). Save the transformed `bio10` raster as a GeoTiff.

5. Use the cleaned species occurrence data to extract the bioclimatic variables from the raster stack and compare the climate conditions where this species has been observed to the broader climate of Australia. A few hints: Have a look at the `raster::sampleRandom` function. To perform the comparison between climates where the species is present and Australia more broadly, you have a number of options. You might try scatter plots, box plots or histograms, but you do not need to do any statistical analyses (in other words, see what you can learn from simple plots alone - that's enough for this assignment). **Answer the question:** How does the climate where *X. australis* has been observed differ from that of Australian climates more generally?

6. Create a raster of the number of species observations in each grid cell. You might try using the `rasterize` function or perhaps by extracting the cell number for each observation and counting the number of times each cell number is duplicated (indicating the number of observations in that cell). This can be a tough one, so don't hesitate to check in if you get stuck.
