**OEB 137 – Assignment 8**

**Due November 11, 2020 at 11:59 am (Eastern)  
  
Contingency and Multivariate Analyses**

1. Read in the “Bivalve Contingency.csv” dataframe on surveys of 3 different marine bivalve larvae (Family) in water from sites with high dissolved oxygen (DO) and other sites with low dissolved oxygen. Create a contingency table for the counts of individuals in each family that were observed in each DO condition. Now run a chi-square test on these data. Do you need to employ the Yates Continuity Correction based on your contingency table? If so, use it, if not, omit it. **(2 pts)**
2. Load the vegan package and use the data() function to pull in the varespec and varechem data sets, which provide data on lichen cover and the chemical properties of the sites surveyed by Vare et al. 1995. The varespec dataset has abundance data for 44 different lichen species (columns) at 24 sites (rows), which provides a robust assessment of the lichen community at each site. The varechem dataset has various nutrient/abiotic properties (columns) at those same 24 sites
   1. Run a Principle Component Analysis on the varespec data. Briefly explain what the PCA analysis does and what, in general terms, the eigenvalues of the PCA axes represent. **(2pts)**
   2. Calculate what percent of the total variation in the data is explained by the first two Principle Component Axes. **(1pts)**
   3. Create a biplot that shows PCA 1 on the x-axis, PCA 2 on the y-axis, displays the sites by their site number and displays the loadings for each species as eigenvectors (arrows). Based on this plot, what is dominant species driving the community at site 27? **(2pts)**
   4. Copy/Paste/Adjust your above code to run the same analyses/plots on the varechem dataset. Based on this plot, what site is most likely to be strongly co-limited by (having low values of) Calcium and Aluminum? (**1 pt**)
3. Read in the “Soil Microbes and Disturbance.csv” data frame. This contains data on the abundance of different microbial OTU’s (conceptually similar to species) for 40 sites – 10 sites in each of 4 different disturbance categories.
   1. Use metaMDS() to run an NMDS analysis on these data. How many dimensions did it take to get convergence? (1pt)
   2. Run an ANOSIM test on these data to determine if there is statistically significant separation by disturbance treatments in NMDS space (1pt)

**Bonus Question**: With the “Soil Microbes and Disturbance.csv” data, make an NMDS ordination plot where: There are 4 convex hulls for the 4 different disturbance treatments and each convex hull has a different color, the OTU’s are plotted in black in their correct locations in NMDS space, and the sites are plotted with the same/similar color scheme as the convex hulls that describe the disturbance treatments that each site belongs to (**1pt**).