

**Assignment D: Quantitative Thinking in the Life Sciences (Fall 2012)**  
(Assigned on November 7<sup>th</sup>. Due on November 28<sup>th</sup> – 100 points)

R code only! No write-up

- Create an array/matrix with simulated data for your system
  - Column one should be your y variable (aka dependent or response variable)
    - TSS remaining, phosphorus, Resilience to Climate Change Index, yield
  - The number of rows should be the number of expected samples
  - Columns 2-?: These are your x variables (aka independent or predictor variables)
    - year, site, pollen amount, harvest type
- Populate the array with dummy data using the distribution and distribution functions from your homework (you don't need to put in measurement error right now)
- Using the test agreed upon today, write the code and test your hypothesis
- Send well annotated code to me as a single document (plus data files if you choose to use your real data)

Hint: If you simulate data without any signal, you will likely not get a significant result. Use your knowledge of your relationships with the major components to simulate realistic data that includes patterns. For example, if you have three harvest types, and you expect different carbon storage capabilities with each harvest type, simulate data for each harvest type using three different distributions. Some of you may be thinking “Well isn't that cheating?” and in a sense it is cheating but you still need to incorporate all of the error you expect in each of the component variables and you are still limited by the number of samples that you expect to take.