Final Exam Review Sheet

**Everything from the first half of the class**

* What is a p-value in plain English?
* What are the null hypotheses for simple statistical tests we have covered?
* Understand how to use Bayes Theorem
* What are causes of the reproducibility crisis?
* What are the benefits to you in conducting your research in a reproducible manner?
* What are the advantages and disadavantages of using a program like R versus a program like Prism or Jump?
* What type of scientists and problems led to the development of almost all of the stats that you will every use.

**From the second half of the class**

* Be able to recognize the appropriate statistical test for a hypothesis and dataset and apply the simplest appropriate test.
  + Sometimes a T-test is all you need
* Munge data in R (subset, convert between data structures, etc.)
  + Select rows of a matrix or dataframe based on values in one column
  + Select elements of one vector based on values in another vector
* Run a linear model in R
  + Know how to describe a formula + vs \*
  + Know how to include a random effect
* Interpret the output of a linear model
  + What are the coefficients telling you in plain English
* Differentiate between statistical and biological significance
  + What should you evaluate beyond the p-value
* When do we use PCA or other dimensional reduction approaches?
* Make a publication quality plot
  + Use color and shape to your advantage
  + Follow the basic rules that we have discussed

I’m not going to give you a sample test but the datafiles you will be using on the final are available on the course website datafile 1 and datafile 2.

Good luck and thanks for being an awesome class.