Overview

# Biostatistics course

Bill Perry

This course will serve as an introduction to reproducible data analysis using R. Specifically the goal is to introduce students to all facets of managing a research project with an emphasis on:

* Developing questions form observations, hypotheses, and predictions for testing statistically
* Designing data workflows with data entry, curation, QA/QC, and cleaning
* Using a controlled vocabulary and organized project structure and documenting the metadata for the project
* Importing data into R and doing calculations and transformations
* Visualizing data using ggplot
* Understanding how to decide on statistical tests that are appropriate - T-Tests - parametric and nonparametric

- Correlations  
  
- Linear models  
  
 - linear regression  
  
 - multiple linear regression  
  
 - ANOVA  
  
 - ANCOVA  
  
 - GLM and logistic regression  
  
- Analyzing Frequencies  
  
- Multivariate Statistics  
  
 - Principal component analysis  
  
 - NMDS  
  
 - Cluster analysis

In the main webpage I have provided links to all the information you will need:

* links to readings that should be read prior to class
* powerpoint lectures that should be reviewed prior to class
* in-class activities that will actively cover the materials in the powerpoints
* weekly homework to do the in class activities on your own using different data
* assignments that form the crux of the grade and there 4 of them