**Final Project Notes**

**Wednesday - Data Simulation**

* Types of data
  + Numeric/quantitative
    - Discrete
    - Continuous
  + Categorical/qualitative
    - Nominal
    - Ordinal
  + Character
  + Factor
  + Logical
* Rnorm() plots random continuous numbers from a normal distribution
* Runif() plots positive, continuous numbers within a range
* Rpois() plots random discrete integers from a poisson distribution
* Rep() and letter[] plot character variables repeatedly, as defined by a range
* Replicate() repeats the creation of a data set simulation n times

**Thursday – t-tests, ANOVA, factor analysis, structural equation modeling**

t-test

* Compare means of two samples relative to their CI
  + Unpaired= two samples from a different population
  + Paired= two samples from the same mean
* Compare means of one sample and the population (expected) relative to their CI
* Can run a one-tale or two-tale test
* Checks if the signal (difference of means) is significant or just noise
* More variance = more noise and a lower t-score
* Larger n = more signal
* Used for hypothesis testing
  + Null= no difference in means
  + Alt= difference in means
  + Reject null if your t-value is above your critical value
  + Fail to reject null if your t-value is below your critical value
    - Critical value is calculated from df and CI using a t-value table
* Assumptions
  + Normal distribution
  + Similar variance
  + N’s should be the same (or at least similar)
  + Sample size between 20-30ish
    - If higher, go for a z-test