Met with Randy on 02-17-2017

Class date May 16 (3 hours; one person in ATRF and other travel down the road)

**Preliminary requirements:**   
R version (at least 3.3.2)   
RStudio Version (at least 1.0.136)  
Packages needed: Swirl, tidyverse,   
Familiarity with R. Novice/Beginner R users should install the Swirl package and complete “R **Programming:** The basics of programming in R” course before the class. <*probably will need to provide some instructions on this… It is pretty easy, so it shouldn’t be much of an issue*>

Dataset Tidying Data (Tentative schedule)

1. Introduction to R/RStudio **(20 minutes; dataset: BREAST Cancer Set)**
2. Importing data (**10 lecture + 25 minutes hands-on; two different data sets (Excel and Tab-delim) R script with hints**)
   1. Filetypes
   2. Reading function
      1. What to use or turn off in the default reading functions
      2. Memory, types of column data types
      3. Hands-on using two different data sets
   3. Importing big datasets <I assume this will need an additional 5-10 minutes after the 35 minutes from 2a and 2b?>
      1. Memory requirements; comparison to SAS
3. Break (**5 minutes**)
4. Tidy Data (**1 hour and 30 minutes**) <include a break in here somewhere>
   1. Tibble vs DF
   2. Spreading/Gathering
   3. String Parsings
      1. Date function
   4. Separating & Uniting
      1. Reg ex (passing reference; more on next class)
   5. Hands on using WHO dataset
5. Exporting and Sharing data (**15 minutes**)
   1. RData
   2. Rmd
   3. Dput
   4. Small vs big data
   5. RData, csv
   6. Hands on
6. Outline the whole process of importing, tidying and exporting using a data set. (**15 minutes**)
   1. A data set will be given in the class.