Assignment 4: Playing with published data, Part 1

Due: **Thursday, September 26th**

In class, we have started to play around with data associated with a publication to better understand how to use R for analysis and graphics. This is a powerful way to not only learn how to code in R—it can also force you to think hard about how others have analyzed data and help you learn how to do it yourself, or to question how you might do things differently (or better).

This will be the first component of a multi-part assignment. The ultimate goal is to force you to learn more about a study or a technique that will help you with your own work.

The assignment:

Part 1: **Choose a paper** (preferably empirical paper) with *publicly accessible raw data* from which you can re-create an analysis or plot.

Part 2: **Write a short (one paragraph) summary** of the paper (in .docx or .txt).

Part 3: **Download the dataset** and *make sure you can read it into R*. This step may take some re-structuring of the data on your part (e.g., saving it as a .csv file)

Part 4: **Try plotting some of the data on conducting a part of the analysis.** This may or may not be challenging. You may try to re-create a plot in the paper, or try plotting something different. The point is to learn a bit more about the data. You will submit the R code for your attempt at plotting, not the plot.

Things to turn in to the google drive folder under "Assignment 4":

1. The pdf of the publication

2. Your one paragraph summary of the publication (.docx or .txt or similar format)

3. The dataset(s) associated with the publication

4. R script for importing and plotting some aspect of the data. Upload this even if you are not successful.

5. Make this folder an Rstudio Project folder.