Biostatistics 212
Final Project instructions

50 points

The goal of the final is for you to use the material we learned in Biostat 212 to take raw data and turn it into a publishable product. This will require understanding the data, developing a research question, importing the data into Stata, generating derived variables that you will use for your analysis, labeling your variables, doing the analysis, saving and documenting your analysis, and producing a table and figure that readers can understand. In keeping with the spirit of TICR, we hope you will take advantage of this opportunity to analyze your own data and produce a table and figure that you can use in a manuscript.

- Create a publication quality <u>table</u> and <u>figure</u>:
    1. Write a short introductory paragraph (3-5 sentences) explaining what you are trying to show with your table and figures (the research question or the comparison you are trying to make). [2 points]
    2. Create a <u>table</u>, formatted as *suitable for submission to a journal*, complete with title, headings, data, and footnotes (including spelled-out versions of all abbreviations and other notes as needed to clarify meaning or methodology). You can create and export the table using any method covered in class, including `putdocx,` `putexcel,` or any of the user-generated commands. You can export directly into a word document or export to excel first, then move to word. [24 points]
        - Use commands that export your data directly into a table (avoid copying and pasting cells)
        - Additional formatting of the table may be done directly in word or excel
    3. Create a <u>figure</u> of quality *suitable for submission to a journal*. This should be produced within Stata. The figure should have a legend, title, and a note or caption. The reader should be able to easily understand the figure without any outside information. [24 points]
        - Use commands that export your figure directly into word (avoid copying and pasting)

- Turn in:
    - A docx or PDF with the table, figure, and short description
    - Your .do file, with the purpose of each block of code CLEARLY LABELED [**-1 point for each block of code not clearly labeled**]. For example, consider a header and subsection heading for table 1 code, table 1 export, etc. Comment on your code as necessary to explain decisions you made, etc.

- An example final project word doc (not the do file) is posted on the course website

- You will not be graded on the data-prep/data cleaning steps. It's your choice if you want to include them in the do-file you turn in; if you do, please clearly delineate where the code for your table and figure are located within your do-file.

A note on data sets:
- *If you do not have access to your own data, please ask for some from your mentor. If your mentor does not have any data you can look for publicly available datasets. Suggested sources are ACS data via IPUMS, CHIS, NHIS, or DHS data.*
- *Use a health-related data set (not auto.dta!)*
- *Do not use NHANES since we've used it often in class*