Lab 1 Assignment, STAT 310
Due: Thursday, February 18

Directions: Your answers should be formatted using a word processor (e.g., Microsoft Word, Google Docs). For each exercise, copy and paste all the code, output, and graphs from RStudio to your document. Then convert your final document to PDF format. **Your submission to Blackboard must be in PDF format.**

Exercises in this lab assignment will use the `nhanes` data set. To load this data set into R run the following command:

```
nhanes <- readRDS(url("https://ericwfox.github.io/data/nhanes.rds"))
```

Exercise 1. Compute some summary statistics and make a histogram for the variable `Weight`. Describe the shape of the distribution in the histogram.

Exercise 2. Make a frequency table and bar plot for the variable `Smoke100`.

Exercise 3. Make a scatterplot with `BPSysAve` on the y-axis, and `BPDiaAve` on the x-axis. Label the y-axis "Systolic Blood Pressure", and label the x-axis "Diastolic Blood Pressure".

Exercise 4. Make side-by-side box plots of `HHIncome` across the categories of `Education`. Do you notice any trend or association in the plot? [Hint: your graph should contain 5 box plots, one for each category of `Education`]

Exercise 5. Make a contingency table between `Smoke100` and `HealthGen`. Use `addmargins()` to add the row and column totals. Then answer the following questions:

(a) What proportion of participants have smoked at least 100 cigarettes?

(b) What proportion of participants have smoked at least 100 cigarettes and reported being in very good health?

(c) Of the participants who reported being in very good health, what proportion have smoked at least 100 cigarettes?

(d) Of the participants who reported being in fair health, what proportion have smoked at least 100 cigarettes?