Substance Abuse Among Youth

The majority of the project is taken from the course text, *Practicing Statistics*, chapter 6, research project, pages 247-250. The project will consist of doing an analysis of the data provided as specified in the research assignment. You should proceed through instructions 1-6 in your pairs.

R Hints

- The data is provided on the course website (it isn't on the CD).

- Once you've loaded in the data, try: `summary(IYSdata)` to make sure the data is loaded correctly and that you have a sense of the variables and their values.

- You will probably need to use the `ifelse` command in R to recode at least one of your variables to make it binary (see step #2).

- Come to office hours or go to the QSC R office hours on Wed from 8-10pm to get help with the R stuff.

Pair Write-up:

In your pairs, you should write-up the analysis you performed in parts 1-6. You don't need to provide every detail, but you do need to communicate the analysis process. The final write up should contain:

- any preliminary exploratory data analysis that was relevant to the choice in models (e.g., plots or summary statistics of the explanatory variables)

- a discussion of the Wald statistics, the drop-in-deviance test(s), and any model building decisions made along the way

- reliability of the model using measures of association (see section 7.10 and class notes) and/or ROC curves

- a discussion of the final model. Which variable(s) are in? Which are out? What do you conclude about substance abuse? What can you say about causation? What can you say about generalizing to a larger population?

- please do not include pages of output - provide only the relevant information. Remember, your reader is trying to figure out the story the data tells, and it will be very hard to understand your analysis if you have pages and pages of output.

To keep in mind:

- The CD from the book contains a lot of great information on how to write up a research paper.

- Your peers (and professor) will assess your work using the assessment on page 11 of the .pdf titled *How to Write a Scientific Paper or Poster*. It is important to keep those things in mind when writing up your report.

To turn in:

- Due: Thursday, Oct 30. Each individual within a group should post (to Sakai drop box) the **same** R markdown file. Please make sure that all group members' names are contained in the R markdown file. Note: I should be able to run the R markdown file from my computer (with the appropriate path to the data). If I can't run it, your grade will be lowered.

- Due: Thursday, Oct 30. Each individual within a group should post (to Sakai drop box) a **different** confidential assessment of the work allocation. Address the following:

  - Given 100 points, my partner should receive _____ points, and I should receive _____ points.
  - Provide a few sentences justifying the point allocation above.

- Due: Tuesday, Nov 11. Each individual fills out the assessment form for an anonymous project of one of their peers.