

HW01 - Statistical Modeling

STUDENT NAME

```
knitr::opts_chunk$set(warning=FALSE, message=FALSE)
# load appropriate libraries and import data here

# For this assignment, to get you started, I have included code to directly import
# data sets from my website.
# After you create your own data management files, this approach should not be used.

fev <- read.delim("https://norcalbiostat.netlify.com/data/Lung_081217.txt",
                  sep="\t", header=TRUE)
hiv <- read.delim("https://norcalbiostat.netlify.com/data/PARHIV_081217.txt",
                  sep="\t", stringsAsFactors = FALSE, header=TRUE)
dep <- read.delim("https://norcalbiostat.netlify.com/data/depress_081217.txt",
                  sep="\t", header=TRUE)
```

Part I: Statistical Modeling

1. Fit a linear regression model

- a. Analyze the residuals
- b. Interpret each coefficient

2. Test gender as a moderator using a) using a stratified model

- b) using an interaction model.

3. Which of the two models in question 2 assumes that the affect of income on depression is constant (does not change) between males and females?

4. Determine whether the regression plane can be improved by also including weight. Use two measures of model fit to justify your answer to this question

5. Does weight *confound* the relationship between age or height and FEV1?

6. Fit a model of income using age, sex, educational level and religion as predictors.

- a. Use a general F test to determine whether religion has an effect on income.
- b. State the reference categories for both religion and educational level.
- c. Interpret the coefficient for each level of educational level

Part II: Variable Selection

1. PMA6 9.9
2. PMA6 9.11
3. PMA6 9.12
4. PMA6 9.13
4. PMA6 9.14