**Stat 400 Project 1 Fall 2019**

*Purpose:* To analyze large sets of “real” data in the form of a spread sheet and to use descriptive statistics and graphical methods to determine any association between heart disease and the risk factors age, sex, blood pressure, and cholesterol level.

*Introduction*: The Framingham Heart Study was initiated in 1948 in Framingham, Mass., to investigate the relationship between the incidence of coronary heart disease and certain risk factors such as age, sex, serum cholesterol, serum glucose, body-mass index, systolic blood pressure, and cigarette smoking.

*Procedure*: Access the Microsoft Excel file “2.20.Framingham” on ELMS under “Files” which represent a partial listing of the data. The data key is:

# sex : Sex (1 = male; 2 = female)

# sbp : Systolic Blood Pressure

# dbp : Diastolic [Blood Pressure](http://www.math.montana.edu/shancock/courses/stat539/r/Framingham-Feb7.R)

# scl : [Serum](http://www.math.montana.edu/shancock/courses/stat539/r/Framingham-Feb7.R) Cholesterol

# chdfate : Coronary [Heart Disease](http://www.math.montana.edu/shancock/courses/stat539/r/Framingham-Feb7.R)

# followup : Follow-up in Days

# age : Age in Years

# bmi : Body Mass Index (wt (kg) / h^2 (m)

# month : Study Month of Baseline Exam

# id : Subject ID

1. Investigate the relationship that sex may have on the incidence of coronary heart disease in the study. Use descriptive statistics and graphical methods (histograms, boxplots, etc.) for your analysis. What are your conclusions based on your findings?

2. Now investigate the relationship between systolic blood pressure and coronary heart disease by making a suitable histogram. What are your conclusions based on your histogram?

3. As in (2), do the same for the risk factor serum cholesterol.

4. Investigate the relationship that age may have on the incidence of coronary heart disease in the study. You may combine both men and women and/or consider them separately. Use descriptive statistics and graphical methods (histograms, boxplots, etc.) for your analysis. What are your conclusions based on your findings?

**NB** This is a well known study with countless research articles available on the web. It is acceptable to read such information; however, work submitted must be your own and in your own words. Your answers should be precise and concise, we are not refereeing your project for publication in *The Lancet*!