**FRAMINGHAM DIDACTIC DATA SET**

**Resource 01 - Exercises   
*8-3-2015***

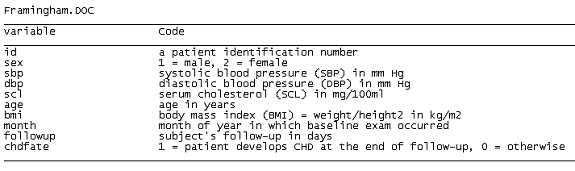
**The dataset represents data from the Framingham Heart Study, Levy (1999) National Heart Lung and Blood Institute, Center for Bio-Medical Communication.**

**Introduction to Exercises**

Cardiovascular disease (CVD) is the leading cause of death and serious illness in the United States. In 1948, the Framingham Heart Study - under the direction of the National Heart Institute (now known as the National Heart, Lung, and Blood Institute or NHLBI) - embarked on an ambitious project in health research. At the time, little was known about the general causes of heart disease and stroke, but the death rates for CVD had been increasing steadily since the beginning of the century and had become an American epidemic. The Framingham Heart Study became a joint project of the National Heart, Lung and Blood Institute and Boston University.

The objective of the Framingham Heart Study was to identify the common factors or characteristics that contribute to CVD by following its development over a long period of time in a large group of participants who had not yet developed overt symptoms of CVD or suffered a heart attack or stroke. The researchers recruited 5,209 men and women between the ages of 30 and 62 from the town of Framingham, Massachusetts, and began the first round of extensive physical examinations and lifestyle interviews that they would later analyze for common patterns related to CVD development. Since 1948, the subjects have continued to return to the study every two years for a detailed medical history, physical examination, and laboratory tests, and in 1971, the study enrolled a second generation - 5,124 of the original participants' adult children and their spouses - to participate in similar examinations. In April 2002 the Study entered a new phase: the enrollment of a third generation of participants, the grandchildren of the original cohort. This step is of vital importance to increase our understanding of heart disease and stroke and how these conditions affect families. Over the years, careful monitoring of the Framingham Study population has led to the identification of the major CVD risk factors - high blood pressure, high blood cholesterol, smoking, obesity, diabetes, and physical inactivity - as well as a great deal of valuable information on the effects of related factors such as blood triglyceride and HDL cholesterol levels, age, gender, and psychosocial issues. With the help of another generation of participants, the Study may close in on the root causes of cardiovascular disease and help in the development of new and better ways to prevent, diagnose and treat cardiovascular disease.

Here we use a subset of the 40-year data which consists of data on 4699 patients who were free of coronary heart disease at their baseline exam. At this exam, the following variables were recorded on each patient.



**Exercises:**

1. Briefly describe the study, including the type (observational, experimental), outcome variable(s), exposure(s) or treatment(s) of interest and any obvious advantages/disadvantages of the design.

2. Briefly describe the characteristics of the dataset, including sample size and the type of variables (e.g. categorical, nominal, ordinal, numeric, discrete, and continuous).

3. To capture the seasonal effect, create a new variable season where 1 = baseline exam occurred in January – March, 2 = April – June, 3 = July – September, 4 = October – December. Be prepared to show how you did this (screenshot of formula is fine) and the resulting distribution of your new variable in your presentation.

4. The researchers were interested in comparing those who develop CHD with those who do not with respect to potential risk factors. Create a “Table 1” demonstrating this comparison with respect to age, gender, season, systolic blood pressure, serum cholesterol and body mass index. Be sure to use appropriate descriptive statistics and be ready to defend your selections.

5. There is specific interest in congestive heart disease (CHD) and the obese. An adult who has a BMI of 30 or higher is considered obese. For the subset of obese participants, create a “Table 2” comparing those who develop CHD with those who do not with respect to age, gender and sbp.

6. Use graphs to explore the relationship between gender and the following: systolic blood pressure, body mass index and season. Interpret the findings.

7. Claims have been made that CHD is more common in males than females. Use graphs to investigate this claim. What do you conclude?

8. Use graphs to explore the relationship between serum cholesterol and body mass index by gender. Does the association look the same for males and females? Interpret the findings.

9. Describe the population to which these results can be generalized.