Stat 242 – Assignment 5  
Due on 5/3/00

1) The provided data set contains estimates of the percentage of body fat determined by underwater weighing and various body circumference measurements for 252 men. Accurate measurement of body fat is inconvenient/costly (see data set bodyfat in the website of the course) and it is desirable to have easy methods of estimating body fat that are not inconvenient or costly.

A variety of popular health books suggest that the readers assess their health, at least in part, by estimating their percentage of body fat. In Bailey (1994), for instance, the reader can estimate body fat from tables using their age and various skin-fold measurements obtained by using a caliper. Other texts give predictive equations for body fat using body circumference measurements (e.g. abdominal circumference) and/or skin-fold measurements.

The file bodyfat contains introductory information that you should delete in order to be able to import the data in R. The variables provided in the data set, from left to right, are:

- Density determined from underwater weighing
- Percent body fat from Siri’s (1956) equation
- Age (years)
- Weight (lbs)
- Height (inches)
- Neck circumference (cm)
- Chest circumference (cm)
- Abdomen 2 circumference (cm)
- Hip circumference (cm)
- Thigh circumference (cm)
- Knee circumference (cm)
- Ankle circumference (cm)
- Biceps (extended) circumference (cm)
- Forearm circumference (cm)
- Wrist circumference (cm)

These data can be used to produce predictive equations for lean body weight. Analyze these data using the regression tree method and additive models. Does the regression tree model and the additive model use the same variables to make the prediction?