Model Building - Remedial Measures

- **Method of Weighted Least Squares.** This method is used when the common variance assumption is not satisfied.

  Idea: Minimize $Q = \sum_{i=1}^{n} w_i * (y_i - \hat{y}_i)^2$, where $w_i = 1/\sigma_i^2$

- **Using Robust Regression to dampen the effect of influential cases**

  ```
  # 1. Least Absolute Residuals (LAR) or Least Absolute Deviations (LAD) Regression
  # min(sum(abs(residuals)))

  data=read.csv("BodyFat.csv",header=T)
  attach(data)
  results.fat=lm(fat~triceps+thigh,data=data)

  library(quantreg)
  rq(fat~triceps+thigh,data=data)

  # 2. Least Median of Squares (LMS) Regression
  # min(median(residuals^2))
  ```

- **Nonparametric Regression.** Lowess (locally weighted regression scatter plot smoothing) Method

  1. The linear regression is weighted to give cases further from the middle X level in each neighborhood smaller weights

  2. To make the procedure robust to outlying observations, the linear regression fitting is repeated, with the weights revised so that cases that had large residuals in the fitting receive smaller weights in the second fitting

  3. To improve the robustness of the procedure further, step 2 is repeated one or more times by revising the weights according to the size of the residuals in the latest fitting

  ```
  lowess(triceps,fat)

  plot(lowess(triceps,fat),type='l')
  points(triceps,fat)

  par(mfrow=c(2,2))
  plot(lm(fat~triceps))
  ```