1. Always plot the data!

2. Better yet: always plot the residuals: $e_i = y_i - \hat{y}_i$.

3. Plot the standardized residuals: $e_i^* = e_i / s_{e_i}$.

4. Leverage: points far away from $\bar{x}$ pull the regression line toward them. A way to measure leverage is by the hat values

   $$h_i = \frac{1}{n} + \frac{(x_i - \bar{x})^2}{S_{xx}}.$$

5. Influence: the amount that a point affects the regression line.
   
   (a) One way to measure influence is to refit the line with the point removed. (R function \texttt{dfbeta()}).
   
   (b) A summary statistic is Cook’s distance:

   $$D_i = \frac{e_i^*}{2} \frac{h_i}{1 - h_i}.$$

---

**Homework - due Monday, May 5, 2008**

1. Read Section 8.4.

2. Do problems 8.2, 8.3, 8.5.

**Useful R**

```r
> l=lm(ohms~juice,data=fo)
> plot(residuals(l)^fitted(l))
> library(MASS)
> l=lm(brain~body,data=mammals)
> plot(rstandard(l)^fitted(l))
> dfbeta(l)
```