1. Read pages 193, 194.

2. Big idea:
   (a) A good model makes the residuals small.
   (b) $R^2$ is another measure of how small the residuals are.
   (c) The book doesn’t give any computing formula for $R^2$ except to compute $r^2$ where $r$ is the correlation coefficient. From this however, we can conclude that $0 \leq R^2 \leq 1$, that 0 means no linear relationship, and that 1 means a perfect linear relationship.
   (d) Two ways of thinking about how $R^2$ is computed directly:
   $$R^2 = 1 - \frac{\sum (y - \hat{y})^2}{\sum (y - \bar{y})^2} = 1 - \left( \frac{sd(\text{residuals})}{sd(y)} \right)^2$$
   (e) In the middle fraction above: $\sum (y - \hat{y})^2$ is the sum of squares of residuals and is sometimes called sum of squares of error. (Later in the book this will be denoted $SS_E$). The expression $\sum (y - \bar{y})^2$ is called the total sum of squares. ($SS_T$).
   (f) We read $R^2$ as a percentage and say that it is the percentage of variation in $y$ that is accounted for (or explained by) the linear model.

3. Example: Ch08Body  We want to predict percentage of Body Fat from Weight in pounds. Weight accounts for 48.5% of the variation in Body Fat.

```r
> attach(Ch08Body)
> l = lm(Fat~Weight)
> summary(l)

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -27.37626   11.54743  -2.371 0.02912 *
Weight       0.24987    0.06065   4.120 0.00064 ***
---
Signif. codes:  0 *** 0.001 ** 0.01 * 0.05 . 0.1 1

Residual standard error: 7.049 on 18 degrees of freedom
Multiple R-squared: 0.4853,  Adjusted R-squared: 0.4567
F-statistic: 16.97 on 1 and 18 DF,  p-value: 0.0006434

> 1-(sd(residuals(l))/sd(Fat))^2
[1] 0.4852972
> anova(l)

Analysis of Variance Table

Response: Fat
            Df Sum Sq Mean Sq  F value Pr(>F)
Weight       1 843.33  843.33 16.9720 0.0006434 ***
Residuals    18 894.42   49.69

Test your understanding by referring to problem 9 of the Review problems at the end of Part II (page 263). The data is Rev02Manatees. Interpret $R^2$ in this example.