Example:
n = 43 male college students
Y = weight (in pounds)
X = height (in inches)
Data available on class website (HtWt.txt) in list of data sets

Goals:
• Predict weight from height.
• Estimate average weight at any given height.

CHOOSE:
It looks like a linear model is appropriate.
Plot the data with regression line (commands given in discussion 2)

**FIT (output on next page):**
Regression line shown:
\[ \hat{Y} = -318 + 7X \]

Example:
\[ X = 75 \text{ inches} \]
\[ \hat{Y} = -318 + 7(75) \]
\[ = 207 \text{ pounds} \]

**ASSESS:**
Stem and leaf plot of standardized residuals looks good.

```
> stem(HtWt$StResids)
The decimal point is at the |

-2 | 0
-1 | 8
-1 | 331110
-0 | 97775
-0 | 4433322221100
0  | 0113
0  | 55789
1  | 022
1  | 569
2  | 0
2  | 5
```
• Some (partial) results from R; things you should already know in boxes
• Thing shown in bold red explained today on white board.

```r
> Mod<-lm(Weight~Height)
> summary(Mod)
Coefficients:
            Estimate Std. Error  t value  Pr(>|t|)    
(Intercept) -317.92      110.92  -2.866  0.00653 **
Height         6.996      1.581   4.425 6.98e-05

Residual standard error: 24 on 41 degrees of freedom
Multiple R-squared:  0.3232,  Adjusted R-squared:  0.3067
F-statistic: 19.58 on 1 and 41 DF,  p-value: 6.978e-05

> anova(Mod)
Analysis of Variance Table

Response: Weight
            Df Sum Sq Mean Sq F value Pr(>F)
Height       1 11277  11277 19.578 6.978e-05 ***
Residuals   41 23617     576
```
Before going to white board, create the same plot with a dotted line at the mean weight, which is 172.6 pound, for reasons that will become clear on white board.

R Commands:
Create plot, add the line for the model called “Mod”, then add horizontal line at the mean weight:
> plot(x=HtWt$Height, y = HtWt$Weight, xlab = "Height (inches)", ylab = "Weight (pounds)", xlim = c(64, 75), ylim = c(130, 240), type = "p")
> abline(Mod)
> abline(h=mean(HtWt$Weight), lty = 2)

In the abline command
• “Mod” is the name of the regression model used previously, and this tells R to add the regression line.
• “h” means to insert a horizontal line at the value given, using line type 2 (lty = 2), which is a dotted line.
If we didn’t know height, our best guess for any individual’s weight would be $\bar{Y} = 172.6$ pounds.