

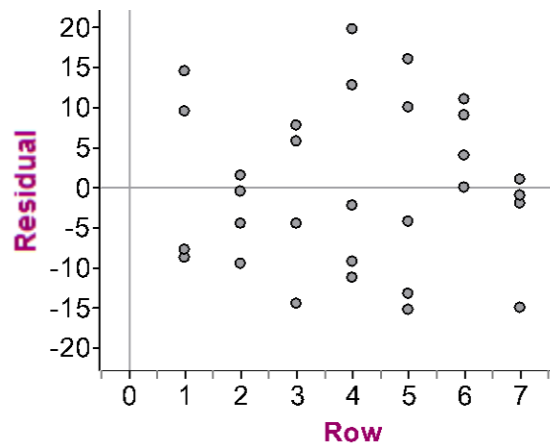
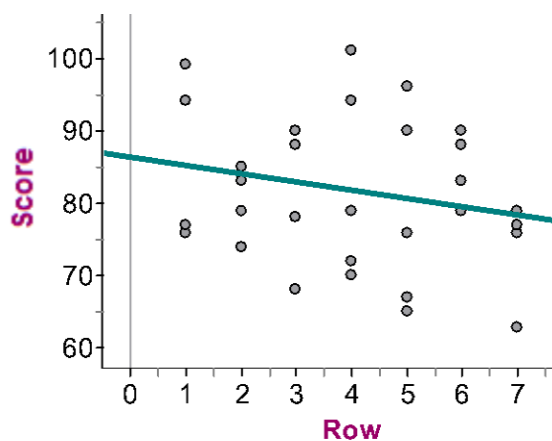
## 3.2 Interpreting Computer Output, Regression to the Mean

Read 181–182

Many statistical software produce computer output for linear regression. Such software include Minitab and JMP. We will learn how to read & use these computer outputs to write equations of the least-squared lines.

**Alternate Example:** *Does seat location affect grades?*

Many people believe that students learn better if they sit closer to the front of the classroom. Does sitting closer *cause* higher achievement, or do better students simply choose to sit in the front? To investigate, an AP<sup>®</sup> Statistics teacher randomly assigned students to seat locations in his classroom for a particular chapter. At the end of the chapter, he recorded the row number (row 1 is closest to the front) and test score for each student. Least-squares regression was performed on the data. A scatterplot with the regression line added, a residual plot, and some computer output from the regression are shown below.



Predictor	Coef	SE Coef	T	P
Constant	85.706	4.239	20.22	0.000
Row	-1.1171	0.9472	-1.18	0.248

S = 10.0673    R-Sq = 4.7%    R-Sq(adj) = 1.3%

(a) What is the equation of the least-squares regression line? Define any variables you use.

(b) Interpret the slope of the least-squares regression line.

(c) What is the correlation?

(d) Is a linear model appropriate for this data? Explain.

- (e) Would you be willing to use the linear model to predict scores when for students who sit in row 15? Explain.
- (f) Calculate and interpret the residual for the row #4 that had a score of about 95.
- (g) Interpret the values of  $r^2$  and  $s$ .

Read 182–185

How can you calculate the equation of the least-squares regression line using summary statistics?

What happens to the predicted value of  $y$  for each increase of 1 standard deviation in  $x$ ?

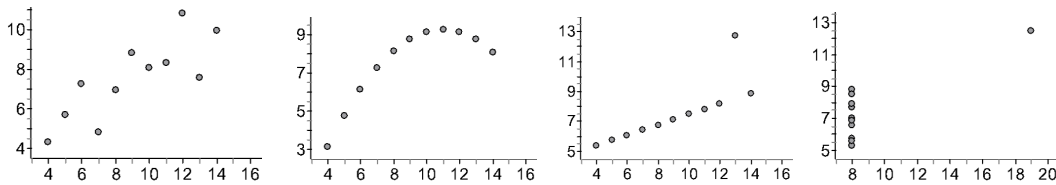
**HW page 195 (56, 59, 61, 63, 65)**

## 3.2 Putting it all Together: Regression and Correlation

Read 185–191

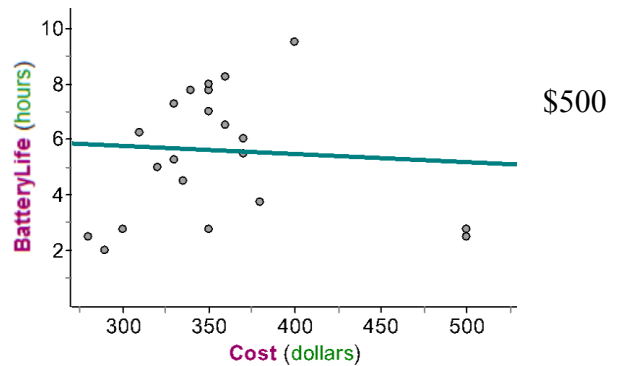
Does it matter which variable is  $x$  and which is  $y$ ?

Which of the following has the highest correlation?



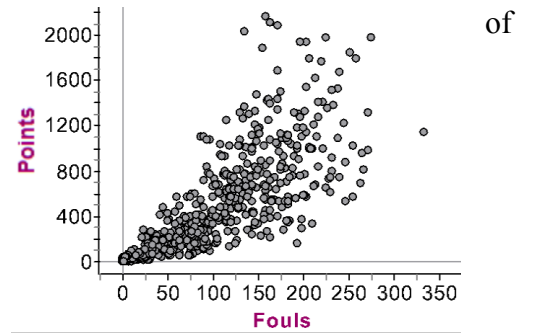
How do outliers affect the correlation, least-squares regression line, and standard deviation of the residuals?  
Are all outliers influential?

Here is a scatterplot showing the cost in dollars and the battery life in hours for a sample of netbooks (small laptop computers). What effect do the two netbooks that cost \$500 have on the equation of the least-squares regression line, correlation, standard deviation of the residuals, and  $r^2$ ? Explain.



Here is a scatterplot showing the relationship between the number of fouls and the number of points scored for NBA players in the 2010-2011 season.

a) Describe the association.



b) Should NBA players commit more fouls if they want to score more points? Explain.