Name: _____ Hour: _____ Date: _____

How do outliers affect the LSRL?

1. Use the ***Correlation and Regression*** applet at www.tinyurl.com/regressionapplet
   - Click on the graphing area to add 10 points in the lower-left corner so that the correlation is about r = 0.50.
   - Check the boxes to show the LSRL and the mean X and Y lines.
   - Sketch it below.

2. For each of the following situations add the point to the scatterplot and decide if the slope, *y*-intercept and correlation will increase or decrease.

   a. If a point is added on the far right side of the graph on the horizontal line for the mean of Y.

      Slope: y-intercept: Correlation:

   b. If a point is added on the far left side of the graph on the horizontal line for the mean of Y.

      Slope: y-intercept: Correlation:

   c. If a point is added below the LSRL on the vertical line for the mean of X.

      Slope: y-intercept: Correlation:

   d. If a point is added above the LSRL on the vertical line for the mean of X.

      Slope: y-intercept: Correlation:

3. Which outliers had the greatest impact on the LSRL, vertical or horizontal outliers?

STATS MEDIC

Outliers and the LSRL

Important Ideas:

Check Your Understanding:

You've probably heard the saying "Practice makes perfect!", but does practice also help you complete a task faster? A study was conducted to find out. A random sample of 15 high school students were taught how to solve a Rubik's cube. Then they were each randomly assigned a number of times to practice this new skill. After they completed their assigned number of practices they were timed solving the Rubik's cube. Here is a scatterplot of the results along with the least-squares regression line.

a. Describe the influence the student who was assigned to practice following the steps to solve a Rubik's cube 14 times has on the equation of the least-squares regression line.
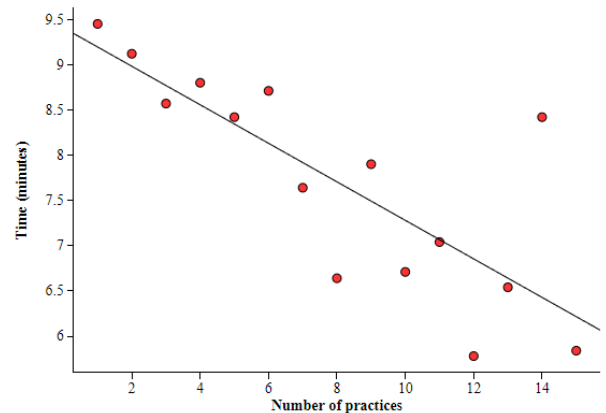
b. Describe the influence the student who was assigned to practice following the steps to solve a Rubik's cube 14 times has on the standard deviation of the residuals and r^2.

c. The mean and standard deviation of the number of practices are $\bar{x}$ = 8 practices and s_x = 4.47 practices. The mean and standard deviation of time are $\bar{y}$ = 7.71 minutes and s_y = 1.20 minutes. The correlation between number of practices and time to solve the Rubik's cube is r = –0.793. Find the equation of the least-squares regression line for predicting time to solve the Rubik's cube from the number of practices.

STATS MEDIC