

Stat 1 Spring 2011 HW 7 (due Tuesday, Nov. 29.)

- (2 points) For a SRS of 1000 students at a large university, the correlation between the students' SAT scores and their GPA in the first year of college is $r = 0.4$. The mean SAT score is $\bar{x} = 1100$ (out of 1600 possible) with standard deviation $s_x = 150$. The mean first-year GPA is $\bar{y} = 2.5$ (out of a maximum of 4.0) with standard deviation $s_y = 0.6$.
 - Using the summary statistics, find the least squares fit for predicting GPA from SAT. Explain what the fitted slope represents in this situation.
 - What GPA would you predict for someone with an SAT score of $x = 1420$? What GPA would you predict for someone with an SAT score of $x = 975$?
 - Find the least squares fit for predicting SAT scores from first-year GPA. What percent of the variability in SAT scores is explained by the linear fit? (hint: this summary is easy to compute).
 - What SAT score would you guess for someone with a GPA of 3.0? What SAT score would you guess for someone with a GPA of 2.3?
- (2 pts) Problem 14.13 in SCC gives data on David Moore's swimming times and pulse rates. These data are loaded in the 2-variable calculator Applet at the SCC web page. Go to my Stat 1 page (which is linked to my math & stat department page) scroll to the bottom and click on the link for **SCC Web Page**. Once there, click on the link for **Statistical Applets** and then select **Two Variable Calculator**. Select **Data Sets** in the menu, and click the bubble by "Swim Time and Pulse Rate" (the first option). If you then click **Data** you will see the table of data values.
 - Click **Scatterplot** to see the graph of pulse rate against swim time. You can click "Show least squares line" and "Show mean X and mean Y" to display these summaries on the graph. Is the association between these variables positive or negative or neither? Explain why you expect the relationship to have this direction, and whether or not you think the correlation coefficient would be an appropriate summary for this association.
 - Click **Correlation and Regression** to get the regression summary statistics. Report the correlation coefficient and the least squares fit. If times were recorded in seconds instead of minutes, would the correlation change? Would the least squares fit change?
 - If the next day's time were 34 minutes, what would you predict for the professor's pulse rate?
 - Click **Swap X & Y** to get the regression of swim time on pulse rate. What time would you predict if the professor's pulse rate were 150 beats per minute?
- (1 point) Often intersections with the most accidents in a year are targeted for additional safety provisions. A study in one large metro area added rumble strips to the five intersections with the most accidents, and observed a lower average number of accidents in the following year. It is tempting to conclude that the rumble strips helped. Explain why the regression to the mean effect implies that we would have expected fewer accidents in the second year, even if rumble strips had not been added.

Additional Problems from SCC text book:

14.6, 14.20, 14.22, 15.6, 15.22 - be sure to explain how you know the sign of r .

Additional Practice Problems (not to turn in): 14.15, 14.19, 14.23, 14.29, 15.5, 15.19, 15.27, 15.35.