

## NOTES FOR THE WEEK OF OCT 16 TO OCT 23

The topic for this week is the relationship between two quantitative variables, such as SAT score and College GPA. (See Figure 5.12 on page 168 for this example.) Usually one of the variables will be designated as the explanatory variable, or X-variable, and the other as the response variable, or Y-variable. For instance, we would designate SAT score as the explanatory variable because we would be interested in using it to help explain future college GPA. When we find the *strength* of the relationship (called the *correlation coefficient*) it won't matter which variable is specified as explanatory and which as response. But when we *describe* the relationship with a *regression line* it will matter. We would get a different regression line if we switched the two roles.

We usually are interested in answering two basic types of questions about two explanatory variables:

(1) Can the response be at least partially predicted by the explanatory variable? For example, "Can college GPA be at least partially predicted from SAT scores?"

(2) Is there any relationship between the two variables? For example, "Is there any relationship between the amount of exercise people get and their resting pulse rates?"

You should be able to do the following:

1. Draw and interpret a scatter plot
2. Recognize approximate correlation based on the scatter plot
3. Understand how various types of outliers affect correlation and regression
4. Understand the interpretation of a regression line, including the slope and intercept
5. Be able to use a regression line to predict Y, given X
6. Understand some possible reasons for relationships, *other* than cause and effect
7. Understand and be able to find residuals
8. Understand why extrapolation is a bad idea for a regression situation
9. Understand how  $r^2$  relates to correlation, and how to interpret  $r^2$  on its own.

### Reading and Study Assignment for this week:

Book Chapter or CyberStats Unit	Focus on:
Chapter 5: Relationships between Quantitative Variables	All, but especially topics covered by assigned exercises, and the material above in these notes.
Unit A7: Scatter Plots	Self Assess Quiz and interactivities listed below.
Unit A8: Correlation, Describing Bivariate Data	Self Assess Quiz and interactivities
Unit D1: Overview of Linear Modeling	Warnings and interactivity below.

### Interactivities to play with:

#### Unit A7:

*Basics 1* - shows how to put points on a scatter plot

*Basics Practice 1* - practice creating a scatter plot; be sure to click "OK" when you plot a point

#### Unit A8:

*Basics Practice 3* (all) – matching scatter plots to their correlations; observing a scatter plot with specific correlation; given a "target" correlation, put points on the plot to achieve that correlation ([this one is also on the CD with the book – see page 178-179 of the book](#))

*Uses 4* – investigating the influence of outliers on correlation

#### Unit D1:

*Basics 2* – visualizing the line for given slope and intercept, allows you to vary them

### Exercises to hand in: Chapter 5: #5, **8abc\***, 17, 18, 21, 24, 29, 38, 49, **63\*\***, 69

\*You can use Excel for the scatterplot if you prefer, rather than doing it by hand. The data and an Excel Manual showing how to do this are both available on the CD with the book *and* on 1pass. (Click on the picture of the book, then on Chapter 5.)

\*\*The **correlation applet on the CD is also in Unit A8** – see note above, under Unit A8.