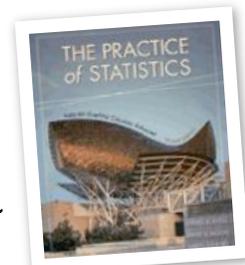


# AP STATS CHAPTER 4: MORE ON BIVARIATE DATA



"NUMBERS ARE LIKE PEOPLE...TORTURE THEM LONG ENOUGH AND THEY'LL TELL YOU ANYTHING..."

Tentative Lesson Guide					
Date	Stats	Lesson	Assignment	Done	
Mon	10/23	4.1	Transforming Relationships	Rd 192-205 <b>Do 3-5</b>	
Tues	10/24	4.1	Power and Exponential Models	Rd 214-219 <b>Do 13-16</b>	
Wed	10/25	4.1	Modeling Nonlinear Data	Nonlinear Modeling Practice	
Thu	10/26	<b>Rev</b>	<b>Review 4.1</b>	Rd 149-156 <b>Do 38-41</b>	
Fri	10/27	<b>Quiz</b>	<b>Quiz 4.1</b>	Rd 157-165 <b>Do 42-43, 45</b>	
Mon	10/30	4.2	Cautions about Correlation	Rd 225-230 <b>Do 27-31</b>	
Tues	10/31	4.2	The Question of Causation	Rd 231-237 <b>Do 33-37</b>	
Wed	11/1	4.2	Causation Practice	Rd 238 <b>Do 38-49</b>	
Thu	11/2	4.3	Categorical Relationships	Rd 241-245 <b>Do 53-55</b>	
Fri	11/3	4.3	Simpson's Paradox	Rd 246-250 <b>Do 59-61</b>	
Mon	11/6	<b>Rev</b>	<b>Review</b>	<b>Do 72-73, 81-83</b>	
Tues	11/7	<b>Exam</b>	<b>Exam Chapter 4</b>	<b>Online Quiz Due</b>	

### Note:

The purpose of this guide is to help you organize your studies for this chapter.

The schedule and assignments may change slightly.

Keep your homework organized and refer to this when you turn in your assignments at the end of the chapter.



### Class Website:

Be sure to log on to the class website for notes, worksheets, links to our text companion site, etc.

<http://web.mac.com/statsmonkey>

Don't forget to take your online quiz!. Be sure to enter my email address correctly!

<http://bcs.whfreeman.com/yates2e>

My email address is:

[jmmolesky@isd194.k12.mn.us](mailto:jmmolesky@isd194.k12.mn.us)

## Chapter 4 Objectives and Skills:

These are the expectations for this chapter. You should be able to answer these questions and perform these tasks accurately and thoroughly. Although this is not an exhaustive review sheet, it gives a good idea of the "big picture" skills that you should have after completing this chapter. The more thoroughly and accurately you can complete these tasks, the better your preparation.

### Transforming Data:

- If a scatterplot shows a curved pattern, it can perhaps be conveniently modeled by an exponential growth or decay function of the form  $y = ab^x$  or a power function of the form  $y = ax^b$ . In these situations, we can linearize the data by making use of logarithms.
- If a scatterplot suggests an exponential or power function, we should check to see which model is appropriate by transforming  $x$ ,  $y$ , or both variables and checking the linearity of the resulting scatterplot. If  $(x, \log y)$  appears linear, an exponential model may work best. If  $(\log x, \log y)$  appears linear, we may want to try a power model. Find the LSRL for the transformed data and then perform the reverse transformation to find the regression model. Know how to transform back to a power/exponential model

### CAUTIONS ABOUT CAUSATION:

- Correlation and regression need to be interpreted with caution. Two variables may be strongly associated, but this does not mean that one causes the other. Bottom line: High correlation does not imply causation. Among other things, lurking variables and common response should always be considered:
  - Extrapolation: Using LSRL to predict values outside the domain of the explanatory variable used to create the line. (It is generally best to avoid extrapolating.)
  - Lurking variable: A variable that effects the relationship of the variables in a study, but is not included among the variables studied.
  - Correlation based on averages are frequently higher than correlation based on all of the numbers that make up the averages.
- If there is a strong association between two variables  $x$  and  $y$ , any one of the following statements could be true.
  - $x$  causes  $y$
  - both  $x$  and  $y$  are responding to changes in some unobserved variable or variables.
  - the effect of  $x$  on  $y$  is hopelessly mixed up with the effects of other variables on  $y$ . Because of this, it is impossible to directly determine the effects of  $x$  on  $y$ .

### RELATIONSHIPS IN CATEGORICAL DATA:

- A two-way table of counts describes the relationship between two categorical variables... the row variable and the column variable. The row totals and column totals give the marginal distributions of the two variables separately, but do not give any information about the relationships between the variables. Probabilities, including conditional probabilities, can be calculated from two-way tables.
- In some cases, lurking variables can cause a reversal in the direction of a comparison or association when data from several groups are combined to form a single group. This phenomenon is called "Simpson's Paradox." Mmmmmmmmm...

