Name: **Answer Key** Period: **2, 3**

1. Construct a table of conditional proportions based on study time. Is there a difference between those who studied more than an hour and those who did not on the grade they received?

Study/Grade→	A or B	D or F	Total
1 or more hours	19 **.76**	6 **.24**	25
Less than 1 hour	5 **.19**	21 **.81**	26
Total	24	27	51

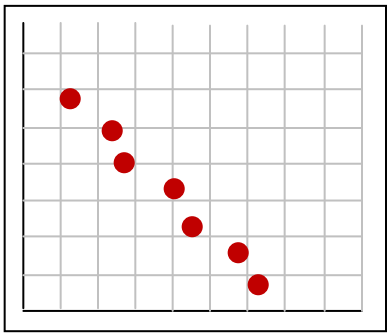
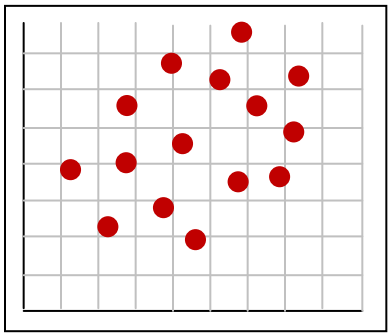
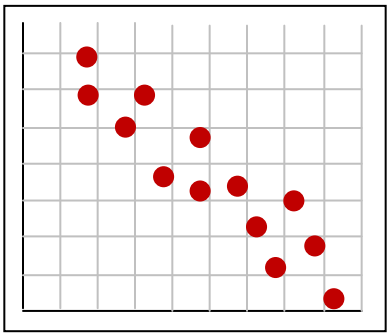
Yes! More A/B's for those who studied more than one hour.

2. Suppose information was collected where x = square footage of your home and y = # of bedrooms the home has. If the LSRL (Least Squares Regression Line) equation is $\hat{y} = -0.4 + 0.0023x$, what is your estimate of the number of bedrooms your house will have if you have a 2,000 square foot home?
Plug 2000 in for x; 4.2 bedrooms, so about 4 bedrooms in a house with 2000 square feet

3. In a study to determine whether surgery or chemotherapy results in higher survival rates for a certain type of cancer, whether or not the patient survived is one variable, and whether they received surgery or chemotherapy is the other. Which is the explanatory variable and which is the response variable?
Explanatory: Surgery or Chemo **Response: Whether the patient survived**

4. Create a scatterplot with data points that would closely match the given correlation coefficient.
****These are hard to make answers for in Word, but you get the idea!! ☺**

a) r = -0.9 b) r = 0.3 c) r = -0.7

5. The more years of schooling you receive, the more money you will make, in general. It was found that the average years of schooling for an American adult is 13.4 and the standard deviation is 1.5. The average income for an American adult is $31,000 and the standard deviation is $9,700. The correlation between education and income was found to be 0.7. Use this information and the formulas below to create a regression equation that **predicts** income when years of education is **given**. **(Income is y)**

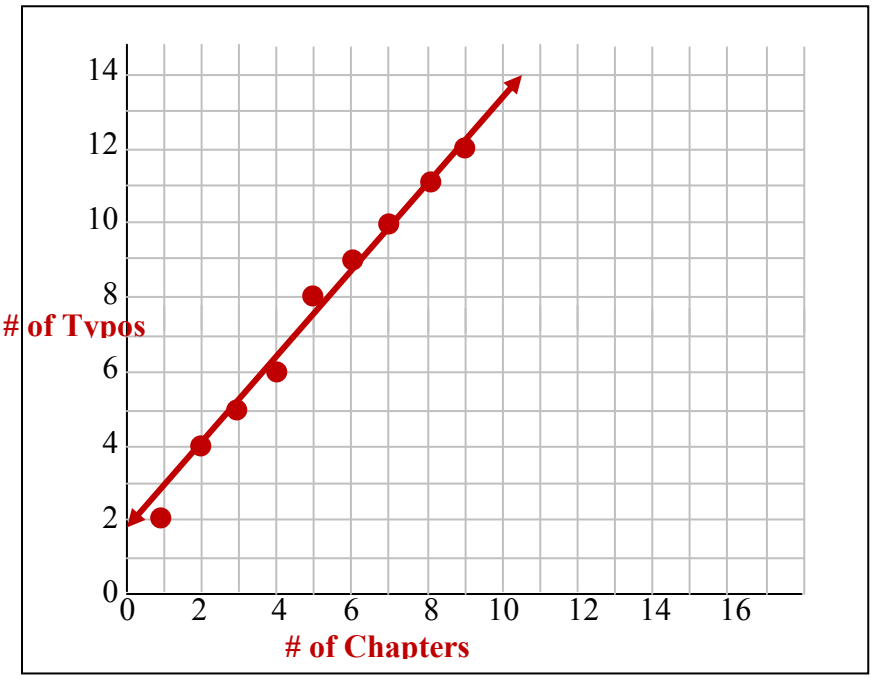
$$\hat{y} = a + bx \quad where \quad b = r\frac{s_y}{s_x} \quad and \quad a = \bar{y} - b\bar{x}$$

$\hat{y}$ **= -29657.38 + 4526.67 x**

6. A sample of novels showed how many chapters the book had and how many typos were found within the book.

Number of Chapters in a Book	1	2	3	4	5	6	7	8	9
Number of Typos Found in the Book	2	4	5	6	8	9	10	11	12

a) Complete an accurate scatter plot of the given data (label the axes).



b) Find the equation of the least squares regression line. Explain what the slope and y-intercept represent in reference to the given situation.

$\hat{y} = 1.28 + 1.23x$

Slope: **For every additional chapter, there are 1.23 more typos.**

y-intercept: **Even if the book had no chapters, there would still be 1.28 typos (doesn't really make any sense, but y-intercepts seldom do).**

c) Add the line of best fit to the scatter plot in part a. ✓

d) Calculate the correlation coefficient. Is this weak or strong? **.99; super strong!**

e) What percentage of the variation in typos can be explained by this equation? **98.96% (r^2)**

f) Use the regression equation to predict the number of typos in a book with 10 chapters. What is the residual if the book's actual number of typos is 16? **"y hat" is 13.58, so the residual is 2.42 (y – y hat)**

g) Would you use this same equation to predict the number of typos in a book with 30 chapters? Why or why not? **No – too far extrapolated! 38.18 typos! This same trend may not hold up!**