1. What does $\hat{y}$ mean/represent? **A prediction**

2. Suppose information was collected where x = # of hours you work per week and y = # of hours of sleep you get. If the LSRL (Least Squares Regression Line) equation is $\hat{y} = 63 - .375x$, what is your estimate of the number of hours sleep you get if you work 40 hours per week?

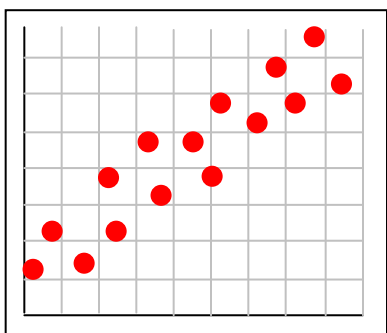   **$\hat{y}$ = 63 - .375(40)**
   = 48 hours

3. Give an example of an explanatory and response variable (make up your own situation!).

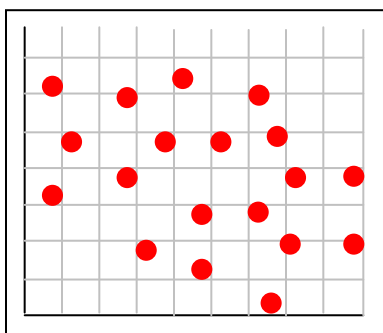   **Explanatory: # of close friends you have**
   Response: Happiness

4. Create a scatterplot with data points that would closely match the given correlation coefficient.
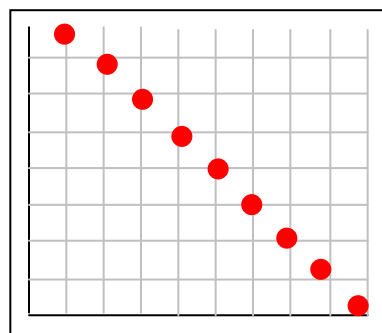   ****These are hard to make answers for in Word, but you get the idea!! ☺**
   a) r = 0.7   b) r = -0.3   c) r = -1.0

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 14.2 and the standard deviation is 1.7. The average income for an American adult is \$32,000 and the standard deviation is \$11,500. The correlation between education and income was found to be 0.8. Use this information and the formulas below to create a regression equation that **predicts** income when years of education is **given**.

$$\hat{y} = a + bx \quad where \quad b = r\frac{s_y}{s_x} \quad and \quad a = \bar{y} - b\bar{x}$$
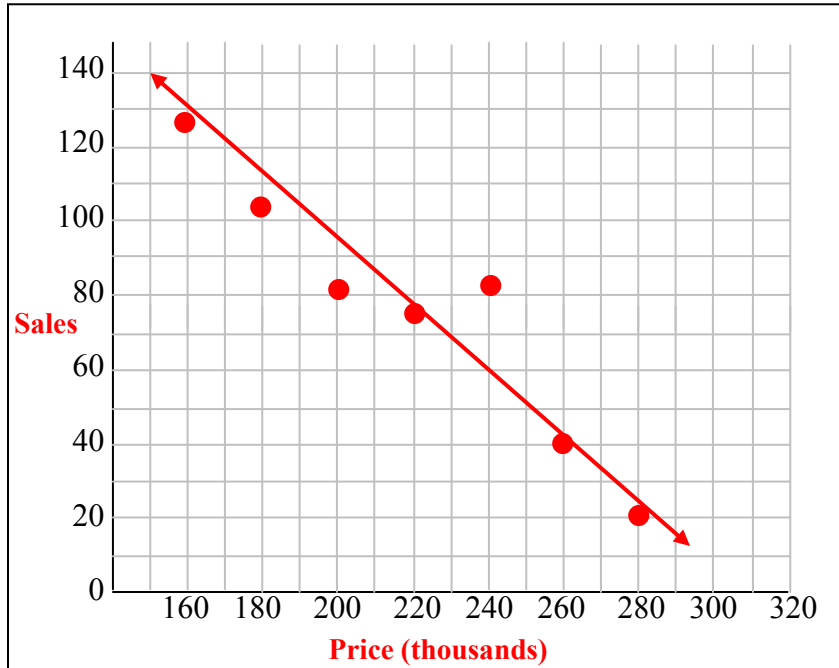
   b = .8(11500/1.7) = 5411.76

   a = 32000 – 5411.76(14.2) = -44847.06

   $\hat{y}$ = -44847.06 + 5411.76x

6. A sample of homes in Phoenix showed how many had sold at different price points in the last month.

Price (Thousands of $)	160	180	200	220	240	260	280
Sales of New Homes This Year	126	103	82	75	82	40	20

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

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

$$\hat{y} = 249.86 + \text{-}.79x$$

Slope: **For every extra $1,000 a house costs, the number of sales drops by .79**

y-intercept: **If a house costs $0, then 250 of them will sell (not really interpretable)**

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

d) Calculate the correlation coefficient. Is this weak or strong? **-.95; strong!**

e) What percentage of the variation in sales can be explained by this equation? **91% (r^2)**

f) Use the regression equation to predict the number of sales in a neighborhood in Norterra where the average price of a home is $170,000. What is the residual if the neighborhood's actual number of sales is 132?
*Use 170 in the calculation, not 170000.
$\hat{y} = 115.56$
$y - \hat{y} = 132 - 115.56 = 16.44$