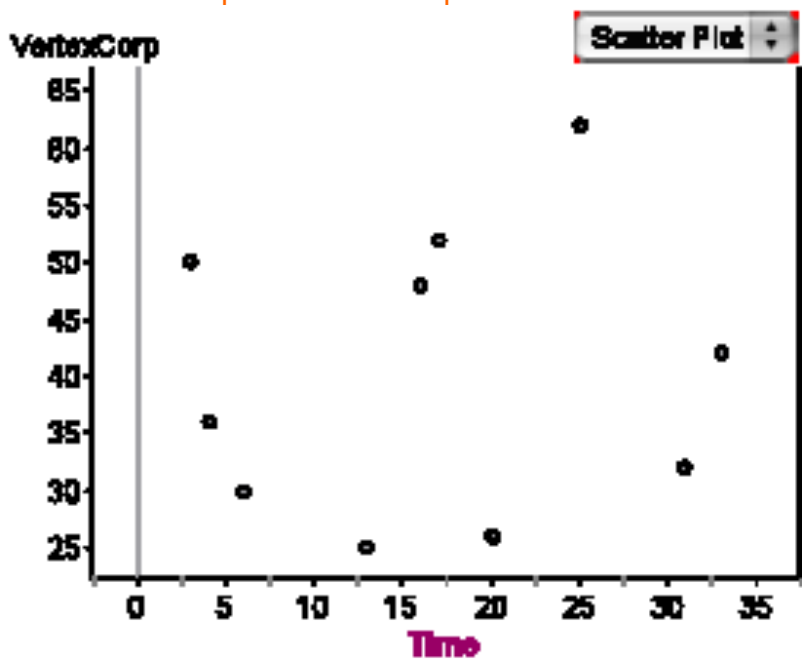
After watching the video, ***Regression***, complete the following problems. A graphing calculator or other similar technology will be helpful for #1 and #2. An online tool can be found at http://www.shodor.org/interactivate/activities/Regression/.

1. Vertexcorp is a large company that hires hundreds of people. Ten of the employees are picked at random for a survey. The results of the survey are given below:

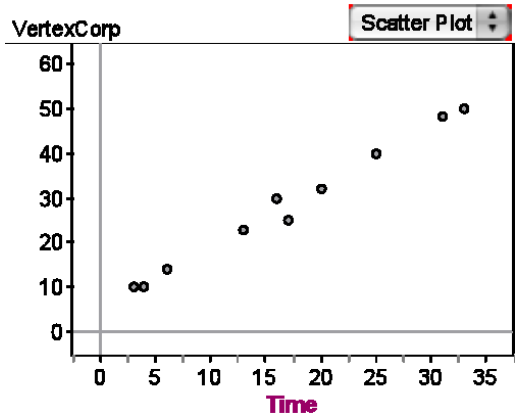
Interviewee number	Time with company (in months)	Salary (in thousands of dollars per year)	Job Satisfaction (scale from 0 - 50)
1	13	25	23
2	33	42	50
3	6	30	14
4	16	48	30
5	31	32	48
6	25	62	40
7	4	36	10
8	17	52	25
9	20	26	32
10	3	50	10

a. Make a scatter plot of salary vs. time with company.
Here is an example of a scatter plot.



b. Recall from the video that the data has a "linear correlation" if the data looks close to being linear. For this sample, is time with company correlated with salary?
Since the data does not have a linear trend, the data is not correlated. Using linear regression would not result in a reliable prediction.

c. Make a scatter plot of time with company vs. job satisfaction.
Here is a sample of a scatter plot.

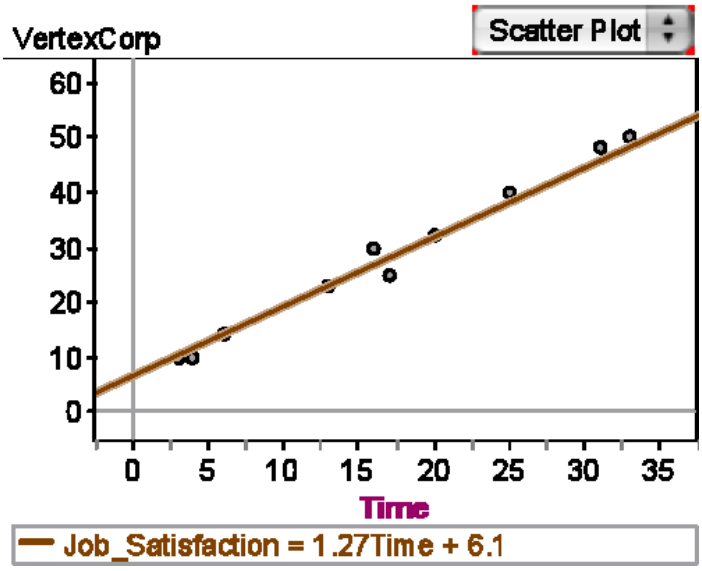


d. Is time with company correlated with salary?
It appears that the data has a high positive correlation because the points are in an upward linear trend.

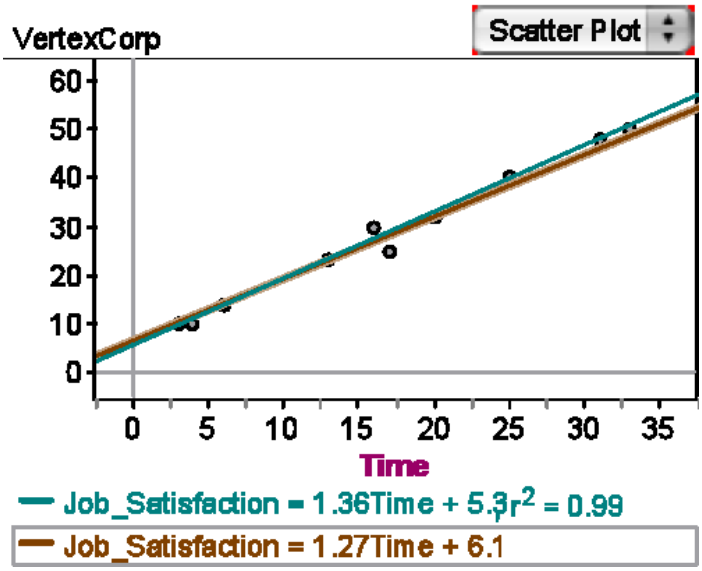
e. Can we conclude that working at the company for a long time causes an increase in job satisfaction? Why or why not?
We might be able to make a prediction about job satisfaction given time on the job but this does NOT mean that one causes the other. Correlation does not imply causation. Further testing and experimenting would need to occur in order to determine causation

2. Use your scatter plot from (1c): time with company vs. job satisfaction.
   a. Draw your best estimate of the least-squares regression line on your plot above.

Here is a sample of a scatter plot with a moveable line using Fathom.



b. Graph the following line on your grid above: $y = 1.36x + 5.3$.



c. The line $y = 1.36x + 5.3$ is the least-squares regression line. Was your estimate close to this line?

Answers will vary, but the lines should be fairly close.

3. Josh Helmcrease has been with Vertexcorp for two years.
   a. Use the least-squares regression line from problem 2 to predict Josh's Job Satisfaction, on our scale from 0 – 50.
   Using the graph, it looks like Josh's Job Satisfaction would be about 38. The value of job satisfaction for 25 months was 40 and since 2 years (24 months) is slightly less, then the job satisfaction would also be slightly less. If you use the regression equation, the value is 37.98, which rounds to 38.

   b. Josh fills out the job satisfaction survey, and it turns out his Job Satisfaction is 35. This isn't exactly what our line predicted. How would you compute the residual?
   The residual is found by finding the difference between the predicted value and the actual value. Since 35 – 38 = -3, the residual -3.

4. Which of the following quantities for United States adults have a positive correlation?
   a. Height and shoe size?
   Positive correlation—generally, the taller a person is, the larger the shoe size.

   b. Height and weight?
   Positive correlation—even though there are short, heavy people, and tall, light people, overall there is a trend that the taller a person is, the heavier the person.
   c. Shoe size and GPA?
   No correlation

   d. Years of education and annual salary?
   Positive correlation—although there are exceptions, there is an overall trend that the more years of education, the higher the salary.

   e. Hours of cell-phone use and incidence of cancer
   Current studies have shown there is no correlation.

   f. Annual income and happiness
   The research on this question is interesting. For low incomes (where being able to afford food and shelter is an issue), there is a positive correlation. For other income groups, there is no correlation. Someone's income is not a good predictor of their happiness. Within a community, however, there is a positive correlation. The richest family in the neighborhood tends to be happier than the poorest family in the neighborhood, regardless of the average income of the neighborhood.