

Math 111  
Activity for Unit 2

For this project you are to collect and analyze data that has a linear trend. Linear data exhibits a constant increase or decrease over the domain of the independent variable. For a good statistical model you really should have 30 pairs of data points, but try to come up with 6 pairs or more.

The data you analyze may come from most any source. (You will need to cite your data source in the paper.) This includes information from internet resources, data you've collected yourself, data from other classes, or other sources. You cannot use a data set straight from the textbook; but if you are having trouble coming up with ideas, you might get some ideas from looking through the exercises.

Your paper should be around 2 pages long, and include graphs and tables. A major component of your assignment is to take a position in your paper and support it using your data.

The purpose of this assignment is to create a model for your data using the techniques of Section 2.4. Your paper should provide a good introduction where you explain what data you are analyzing and give arguments for the use of your model. The body of your paper should provide details to the reader. Specifically, for your linear model you should include interpretations of the slope and y-intercept. Do these seem reasonable? Why or why not? How good of a "fit" is your line? Discuss both the correlation coefficient and visual fit of the model to the data. This is not an exhaustive list of questions to consider, however.

Your conclusions should include how your "curve fitting" fits into the bigger picture around your data. What does the future look like? Should the domain of the model be restricted? Take this time to comment on the model itself and make arguments for or against its long-term effectiveness.

See below for an example of a rough draft.

### Activity 1 Report—Rough Draft

My data set looks at the cost of tuition at LBCC over time.

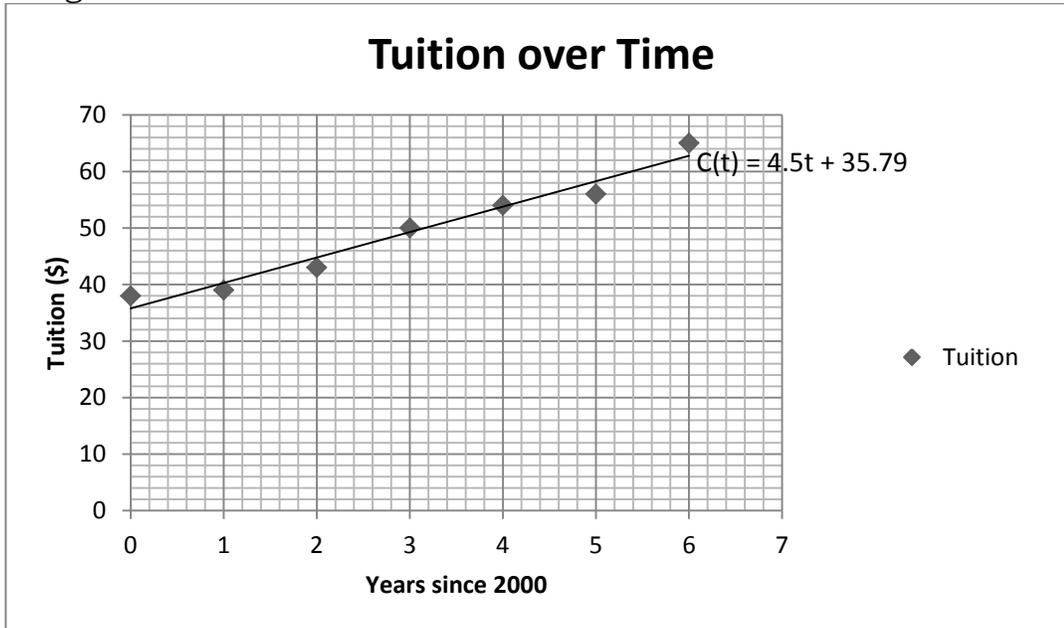
Tuition at Linn-Benton Community College from Fall 2000 to Fall 2006:

Year	2000	2001	2002	2003	2004	2005	2006
Cost per Credit (\$)	38	39	43	50	54	56	65

I will be using the year to predict the cost, so my independent variable will be  $t$  = the number of years since 2000, and my dependent variable will be  $C(t)$  = Cost (in dollars) of one credit of tuition at LBCC.

Using the Linear Regression capabilities of my TI83, the model is given by  $C(t) = 4.5t + 35.79$ , with an  $r$  value of .982. With an  $r$  value (correlation coefficient) so close to 1 this model looks like a good fit for the data. Visually, looking at the model with the original data the model appears to follow the data closely and the choice of a linear model seems reasonable. (See graph below.)

The slope of my model is 4.5 and represents the change in tuition per change in time. So my model indicates that tuition is increasing at an average rate of \$4.50 per year. The y-intercept of my model is 35.79, and represents the starting value of my model—the tuition in the base year of 2000. Of course the actual tuition in 2000 was \$38, but a regression model shows the overall trend and some variation is to be expected.



So why is this only a rough draft and not a finished report? It lacks an introductory paragraph (a sentence is not a paragraph) and a concluding paragraph. I have not cited the source of my data set. I should be discussing for what time period it would be reasonable to use this model. (Do I expect this trend to continue? Why?) I have not taken a position at all, is it reasonable for tuition to be increasing at \$4.50 per credit each year, what impact does this have on students? I could use my model to make a prediction; I could predict tuition for the year 2010 or I could predict when the tuition would reach \$100 a credit.

In short, the data is there, the graph is there, the model is defined and interpreted. I have done most of the math part, but it needs to be fleshed out.