

## MTH245 Unit 2 Module 2 Regression Lines

In the real world, few data sets are exactly linear. But many can be modeled by a linear function that shows the overall trend of the data.

Linear Regression will fit a straight line through a data set that minimizes the distance between each point and the linear regression line. This should be a review topic, and many of you have done this on a graphing calculator; today we will do it in Excel.

For our data set we will look at 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.

We are going to use Scatter Plot to first graph the data and then calculate and graph the equation of the linear regression line. If the scatter plot of the data does not appear to follow a linear trend then we should consider some of the other regression models available in Excel. Or it might be better to consider splitting a data set and creating a piecewise function for the data.

Be sure to watch the video, here is a summary of the steps:

Select the data (don't include titles!) and select the Insert tab, then Scatter. (Pick the option that does not connect the points, Scatter with only Markers.)

Next, under the Chart Tools ribbon, select the Layout tab, Select the Trend line tool and go to the bottom of the menu for More Trend Line Options, click on the box next to Display Equation on Chart and the Display R-squared value, and click Close. The graph, equation and correlation coefficient of the regression line will be displayed.

-  **None**  
Removes the selected Trendline or all Trendlines if none are selected
-  **Linear Trendline**  
Adds/sets a Linear Trendline for the selected chart series
-  **Exponential Trendline**  
Adds/sets an Exponential Trendline for the selected chart series
-  **Linear Forecast Trendline**  
Adds/sets a Linear Trendline with 2 period forecast for the selected chart series
-  **Two Period Moving Average**  
Adds/sets a 2 Period Moving Average Trendline for the selected chart series

[More Trendline Options...](#)

Go to the bottom of the Trendline menu and select More Trendline

**Format Trendline**

**Trendline Options**

Line Color


Line Style


Shadow


Glow and Soft Edges


**Trendline Options**


**Trend/Regression Type**


  Exponential

  **Linear**

  Logarithmic

  Polynomial    Order: 2

  Power

  Moving Average    Period: 2

**Trendline Name**

Automatic :    Linear (Series1)

Custom:   

**Forecast**

Forward: 0.0 periods

Backward: 0.0 periods

Set Intercept = 0.0

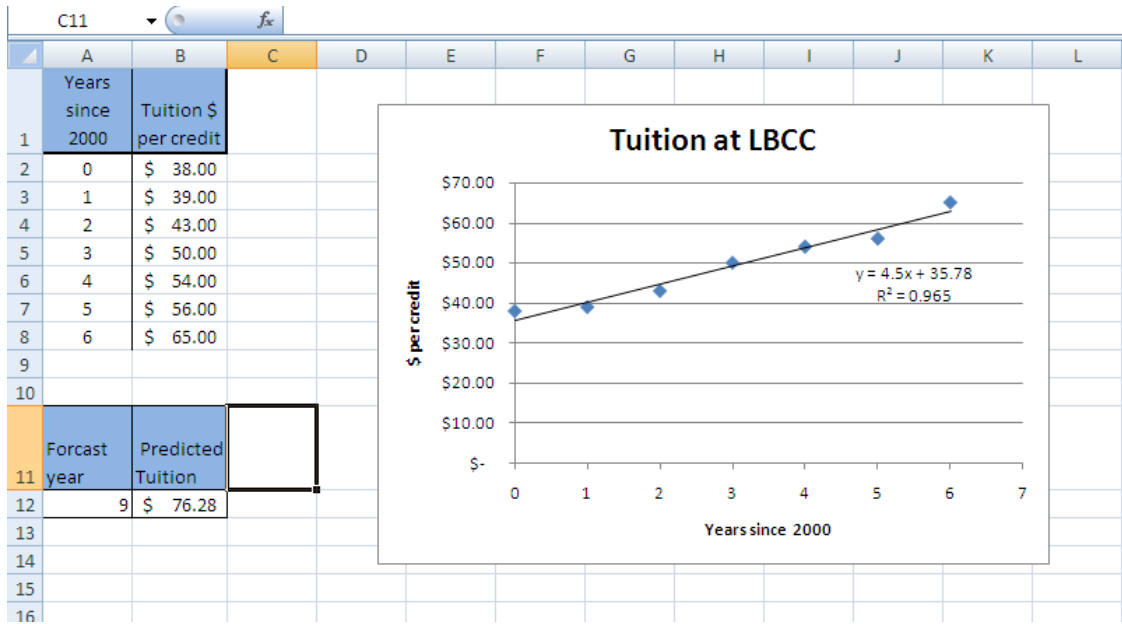
Display Equation on chart

Display R-squared value on chart

Linear regression is the default type.

Check both of these boxes.

Lastly, click on Close.



Trend line:  $(t) = 4.5t + 35.78$ ,  $t =$  years since the year 2000. To predict tuition for the year 2009 we would use  $t = 9$  and predict that tuition would be \$76.28 (it was actually \$74 that year). Could we use this model to predict tuition in 2016? No, that would be too much of an extrapolation.

The value of  $R^2$  is 0.963, what does that tell us? The closer  $R^2$  is to 1, the closer our data is to lining up in a straight line and the better our equation is for making predictions. A course in statistics will cover this topic much more thoroughly. For now, let's say that an  $R^2$  value 0.9 or higher is good and the equation can be used, if an  $R^2$  is less than 0.6 then the regression equation should probably not be used. For equations with an  $R^2$  value between 0.6 and 0.9, take a good look at the scatter plot and how well the equation appears to follow the data and use your judgment. There is a way to determine whether an  $R^2$  value is "good enough" but the details are beyond the scope of this class.