**Linear Regression with Excel**

**Grade: 9-12**

Introduction:

Regression lines can be used as a way of visually depicting the relationship between the independent (x) and dependent (y) variables in the graph. A straight line depicts a linear trend in the data (i.e., the equation describing the line is of first order. For example, y = 3x + 4. There are no squared or cubed variables in this equation).

Note: A curved line represents a trend described by a higher order equation

(e.g., y = 2x2 + 5x - 8).

In addition to visually depicting the trend in the data with a regression line, you can also calculate the equation of the regression line. This equation can either be seen in a dialogue box and/or shown on your graph. How well this equation describes the data (the 'fit'), is expressed as a correlation coefficient, R2 (R-squared). The closer R2 is to 1.00, the better the fit. This too can be calculated and displayed in the graph.

**We will find the linear regression for the following set of data:**

|  |  |  |
| --- | --- | --- |
| **Year** | **Years since 2000** | **Tuition in $ per credit** |
| **2000** | **0** | **138** |
| **2001** | **1** | **139** |
| **2002** | **2** | **143** |
| **2003** | **3** | **150** |
| **2004** | **4** | **154** |
| **2005** | **5** | **156** |
| **2006** | **6** | **165** |

1. Enter the data into an excel spreadsheet. You can adjust and customize your input (change colors, fonts, number of decimal places, etc.) by going to the Home tab, in the Cells group, clicking Format, and then Format Cells. This is also where you can adjust the number of decimal places you wish to express on your data.



1. Once you have put in your data, highlight your data for the independent and dependent variable by clicking and dragging or clicking and holding down the shift key. Then, go to the insert tab and click on the arrow next to little picture of a scatterplot under the charts group.



1. Your data will show up as a scatterplot. You can enhance the graph you make by adding axis titles and a chart title. This option can be found under the chart tools design tab, under the group chart layouts, where you will select ‘add chart element.’





1. Make sure your labeled chart is selected and keep your attention on the chart tools design tab where we just labeled our graph. Under the ‘add chart element’ icon, you will click the ‘trendline’ option.

\*In exploring the options, try to click on ‘more trendline options. You will find that some of these options may not work in accordance to your data. Why might this be?



1. Now that we have our best fit line, you will scroll down on the ‘format trendline’ menu and select the ‘display equation on chart’ and ‘display R-squared value on chart’. Close out of the menu.



WHAT DOES IT MEAN??

4.5 means that on average the tuition is increasing at this rate each year ($)

135.79 (y-intercept) is the cost in the year 2000

 ^this is according to the LINE, not the actual cost!

0.9652 tells us how well the equation expresses the data (correlation coefficient)