

GRADE 12 APPLIED
UNIT C – FUNCTIONS
REGRESSION – FITTING CURVES TO DATA



Name: _____

Date: _____

Often you collect some data (as a scientist or a social worker, etc) and you want to see if there is a trend or a function (formula) you can use that explains the data. Having a formula allows you to predict what happens given some variable. Eg: does a child's school marks 'depend' ('correlate' is a better word) on the family income? Does attendance correlate with a course mark? Good attendance means a good mark, statistically speaking? Let us see!

I have collected the following data on ten students over a 40 day course.

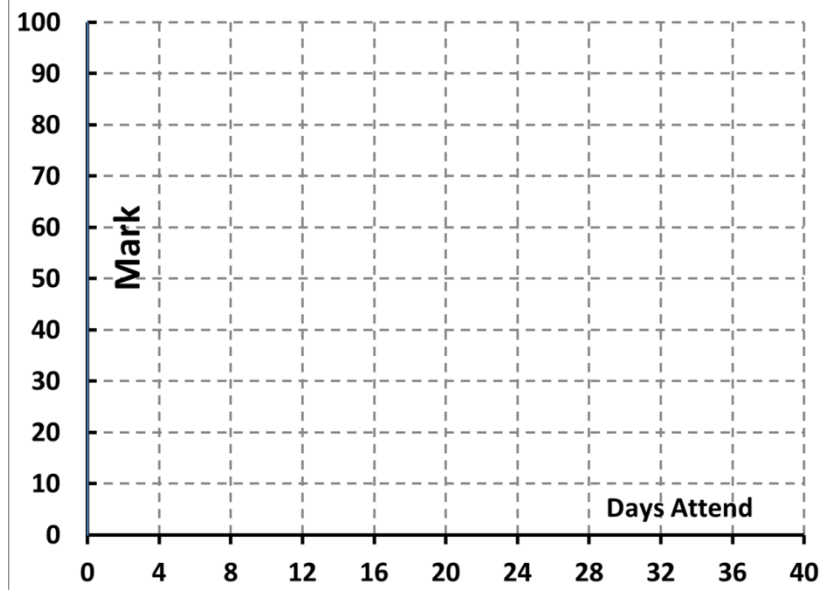
Days Attended	40	30	20	35	30	15	38	25	20	5
Mark	85	52	35	98	80	25	95	100	25	10

Plot the data as a 'scatter' plot

Draw a 'best fit' line

Identify the 'outlier'

Calculate the equation that best fits the data (see steps below)



Entering Data. Enter data points as follows. (You need at least 3 or 4 points)

a. Press STAT Select 1:EDIT

b. Clear Lists 1 and 2 (Cursor to the top of each column and press CLEAR ENTER)

c. Entering independent data in L1 . Enter dependent data in L2. (the data must be in increasing order by independent variable)

Plotting Data. Plotting data is similar to graphing data.

a. De-select all **Y=** formulas or clear them so they will not graph. The equals sign will be highlighted if they are to be graphed.

b. Select **STAT PLOT** by pressing **2nd Y=**. Select **Plot 1**. Turn

on Plot 1. Put plot into the *Scatter Plot* mode. Make sure the data is being taken from lists L_1 and L_2 . Select the largest *mark* possible.

Press **GRAPH**. You should have a plot of your data! You will likely need to use **ZOOM 9:ZOOMSTAT** to nicely fit the data.

Data Regression¹. Now that you have your data entered let the TI83 calculate the *coefficients* of the type of curve you want. This is a statistical operation, **fitting the best curve to the data**. Do it like this:

- a. Go to Catalogue [**2nd 0**] and select **DiagnosticOn**²
- b. Press **STAT** . Select **CALC**. Select desired regression (in this case: LinReg). Press **ENTER**
- c. The screen will show you the A, B, C and sometimes D for the equation that best matches the data. Make sure it makes sense!
- d. No graph the line function. Go to the Y= window. Type in the equation that was calculated OR Put the cursor in the first function (Y1=). Press **VARS** . Select 5:STATISTICS. Select **EQ**. Select **RegEQ** to paste the Regression Equation in (if you were too scared to type it in manually)
- e. Press **GRAPH**. Both your raw data and the best fit curve will appear.

The R^2 value tells you whether the equation is a good fit, how reliable and accurate it is. An R^2 of **1.00 is a perfect fit, the equation is reliable**.

An R^2 of 0.5 or 50% means you can only be 50% certain the regression is good (sort of, as an easy explanation)

¹ The **TI 83 Regression** operation is very powerful. There are additional 'parameters' that can be entered. We are just using the most basic and simple mode of **Regression** above.

² Turning on the diagnostic will enable the calculator to display the correlation factors!. Remember how if r^2 is close to 1 then something is a good fit? Once the calculator is set up like it will never have to be done again.