Linear Models and Regression with TI-89

To illustrate this topic, we use Problem 18 on page 115 as an example. Enrollment in public colleges (in thousands) in selected years is shown in the table below.

Year	1980	1985	1990	1995	2000	2003
Enrollment	9457	9479	10,845	11,092	11,750	11894

(a) Find a linear model for this data, with x = 0 corresponding to 1980.

1. We first rewrite our table.

Year	0	5	10	15	20	23
Enrollment	9457	9479	10,845	11,092	11,750	11894

2. Using your calculator, you want to make a scatter plot the data in the table. However, you must first enter the data into two lists, one for *x* and one for *y*.

   From the home page, hit APPS->6:Data/Matrix Editor->3:New. For Type, hit ⊙ ->3:List. For Folder, choose Main, and then for Variable, type in enroll alpha, followed by ENTER (see Figure 1). Now hit ENTER again. Then type in the numbers for the years in column c1, each followed by ENTER or ⊝. Then hit ⊙, followed by ◆ ⊙ to get to the top of the second column (see Figure 2). Then enter the six numbers for the enrollment similarly in column c2.
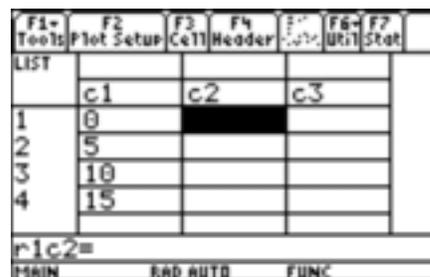

Figure 1

Figure 2

3. We plot our data.

   Hit F2 (Plot setup). Turn off any active plot that has a check in front of it by moving the cursor to the plot and hitting F4 (√), and then move the cursor to an undefined plot and hit F1 (Define). Choose Scatter for Type and Box for Mark. Type c1 for x (c alpha 1 should do it) and c2 (alpha c2) for y (See Figure 3). Then hit ENTER until you are back at the plot list. (Note that for regular function graphing, all plots should be off.)

   Now hit ◆[Y=], and turn off any function that has a check in front of it by moving the cursor to the function and hitting F4 (√). Then hit F2 (Zoom)-> 9:ZoomData to get the scatter plot. Now hit ◆[WINDOW], and make xmin=-2, xmax=25, xscl=5, yminl=9000, ymax=12500,

yscl=500, and xres=1. Then hit ◆[GRAPH] to get the scatter plot where each tic mark on the x-axis now represents a distance of 5 and each tic mark on the y-axis a distance of 500 (see Figure 4).
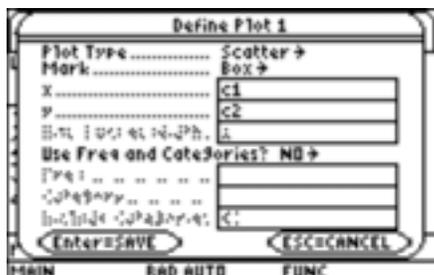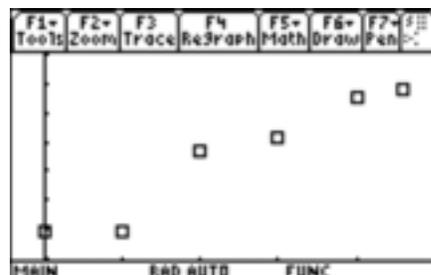

| Figure 3 | Figure 4 |

The points appear to be somewhat linear, so a **linear model** may be appropriate. But what line do we use? We know two points determine a line, so let's choose the first and last points (we could choose any pair of points). These points are (0, 9457) and (23,11894). The slope is 105.9565 to four decimal places, and we are given the y-intercept. The equation is y=105.9565 x + 9457.

4. We add the linear model of part 3 to the 6 data points on the graph in Figure 4.

Press ◆[Y=] and then enter the linear function from part 3 as y1 (for instance), pressing [ENTER] after it has been completely entered on the entry line. Then hit ◆[GRAPH]. On the TI-89, your data points should have remained and the linear graph will now be plotted with it. Notice that the graph goes through the center of the boxes for the points (0, 9457) and (23,11894).

5. Our line goes through the first and last points, but misses the others by varying distances. The ``best'' linear model uses all the data points instead of two arbitrary ones as we did in part 3 along with a statistical technique called regression. Linear regression finds the unique line that minimizes the sum of the squares of the vertical distances from each of the data points to the line. The technique for computing this line is learned in Calculus III. We now compute the linear regression formula for our data.

Press [APPS]->6:Data/Matrix Editor->1:Current to return to the Data/Matrix Editor. Hit [F5](Calc), for Calculation Type choose linear regression by hitting ⏵5:LinReg. Enter c1 for x and c2 for y as was done earlier. For Store RegEQ to, choose y2(x) or any other function that is clear or you want to replace (See Figure 5). Then hit [ENTER]. The formula is of the form y=a x + b. Using the a and b given, we have that the linear regression equation is y = 117.262687 x + 9326.137313.

Notice also that we are given **corr = .969132**. This is the correlation coefficient, and is often abbreviated **r**. It is a statistical measure of how close the data is to lying on a line. The correlation coefficient takes on values between negative and positive 1, inclusively. It has the same sign as the slope of the regression line. The closer the absolute value of the correlation
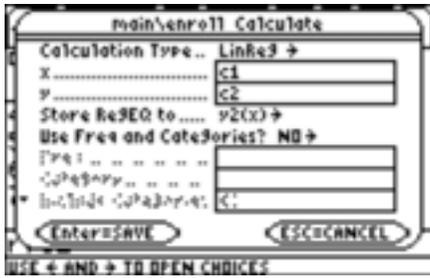
Figure 5

Figure 6

coefficient is to 1, the better the regression line fits the points. Perfect correlation (all the data points line on a single line – the regression line) has a correlation coefficient of ±1. Correlation coefficients near 0 indicate a poor fit.

6. Let's graph this **least-squares linear regression** model.

Press OK to clear the screen in Figure 6, then hit ⬥[Y=], move the cursor onto y1, and press F4(√) to deselect the first equation. Then hit ⬥[GRAPH] to get the regression line with the data points. You can see a better fit than previously.

(b) If the model remains accurate in the future, estimate public college enrollment in 2006 and 2009.

There are two ways to do this.

The first way to do this is to use the Numeric Solver. Press APPS->9:Numeric Solver, hit CLEAR to clear any previous equation. Assuming y2 is the equation used, enter y2(x) as the equation and press ⊖ or ENTER (See Figure 7 – exp= and x may have different values). Then press ⊖ or ENTER until you get to x, type in your value for x (26 or 29), press ⊕ to get to exp (which the display tells you equals y2(x)), and then press F2 (Solve) to get the corresponding y (see Figure 7). When finished, press APPS->6:Data/Matrix Editor->1:Current to return to the Data/Matrix Editor.
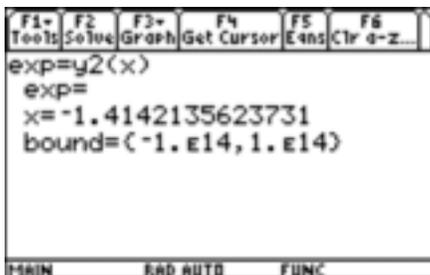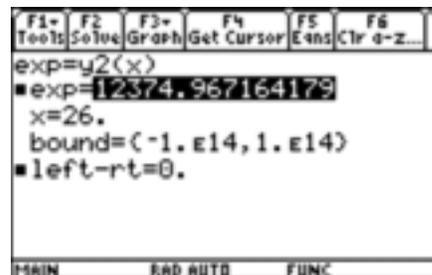

Figure 7

Figure 8

The second method is to hit HOME and assuming y2 is the equation used, enter y2(26) on the entry line for 2006, and press ENTER. The corresponding y value then appears. Then hit

⊳◁←←29 to get y2(29) on the entry line for 2009, and press ENTER. The corresponding y value then appears.

(c) According to this model, when will public college enrollment reach 13,000,000.

Go back to the Numeric Solver by pressing APPS ->9:Numeric Solver. Assuming y2 is the equation used, refer back to Figure 7. Then press ⊙ or ⊙ until you get to exp=, type in 13000, press ⊙ to get to x, and then press F2 (Solve) to get the corresponding x, which turns out to be 31.33. Thus an enrollment of 13,000,000 will be attained around May 1, 2111.