Linear Regression Using the TI-89 Calculator

- **Overview**

The following ordered pairs (t, N) of data values are obtained.

t	N
5	119.94
10	166.65
15	213.32
20	256.01
25	406.44
30	424.72
35	591.15
40	757.96
45	963.36
50	1226.58

Using these data we wish to estimate an empirical functional relationship $N(t)$ between t and N.

Regression is a statistical method that is used to estimate a functional relationship between variables when the underlying data are noisy. It assumes that for each fixed t, the observed value of N is a single realization from a distribution of N-values where the distribution is centered on the true functional relationship $N(t)$. The purpose of regression is to construct an empirical functional relationship that "best" explains the observed data. Because it is always possible to reproduce the data exactly by choosing a sufficiently complicated function, e.g., an n^{th} degree polynomial can be constructed that exactly passes through any set of $n + 1$ distinct points, we typically search for the simplest functional relationship that approximately reproduces the data.

The TI-89 can be used to fit various empirical models: linear (using least squares or median-median regression), polynomial (quadratic, cubic, and quartic), exponential, logarithmic, power, logistic, and sinusoidal. In what follows we fit linear and polynomial models to data and plot the results.

- **Enter the Data**

Press the **APPS** key and choose 6:Data/Matrix Editor and then 3:New… (Fig. 1). This brings up the NEW window (Fig. 2). Type should be set to Data. The default folder Main is fine, or you can create your own folder. (See the manual for how to do this if you are interested.) For Variable choose a name that reflects the contents of the data

Fig. 1 Applications Menu for Data Entry

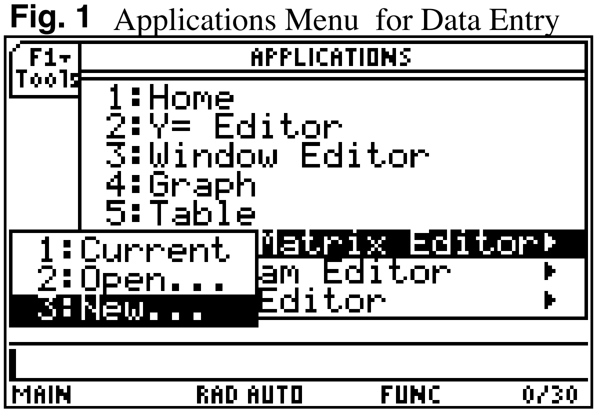
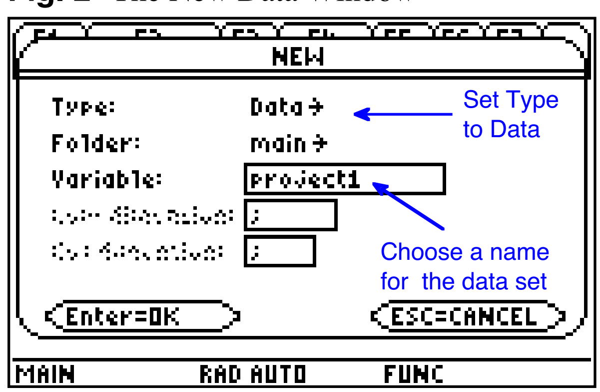


Fig. 2 The New Data Window



set, e.g., Project1. The name you choose is the one you will use to recall the data set. Press **ENTER** twice to bring up the Data/Matrix Editor window (Fig. 3).

Labels for the columns can be entered in the row above the column headings. These labels are for your reference only; the actual names for the variables are c1, c2, etc. and it is these names that are used in calculations. Enter the data values as columns each column corresponding to a variable. For our data I place the t-values in c1 and N-values in c2.

Fig. 3 The Data/Matrix Editor

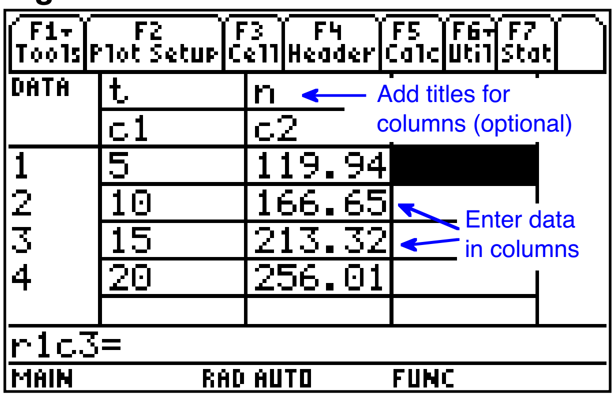
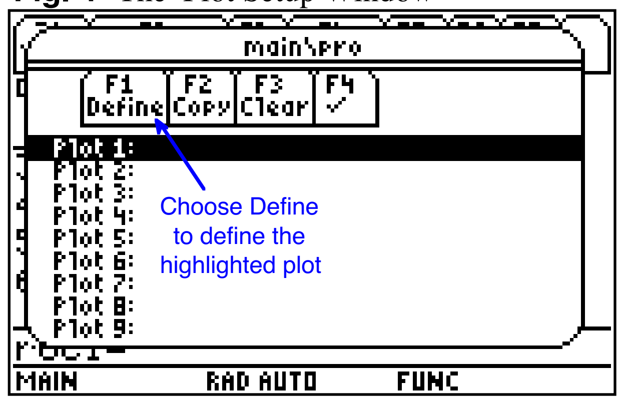


Fig. 4 The Plot Setup Window



- **Plot the Data**

To decide which empirical model to fit, create a scatter plot of the data. Begin by setting up the plot. From the Data/Matrix Editor window press **F2** to select Plot Setup. This brings up the display shown in Fig. 4. With the first plot selected, press **F1** to Define the settings for this plot (Fig. 5). The Plot Type should be Scatter. I choose the Box for Mark since this symbol is large and easy to see. Enter c1 for x and c2 for y. Press **ENTER** twice to save the settings.

Fig. 5 Plot Definition Settings

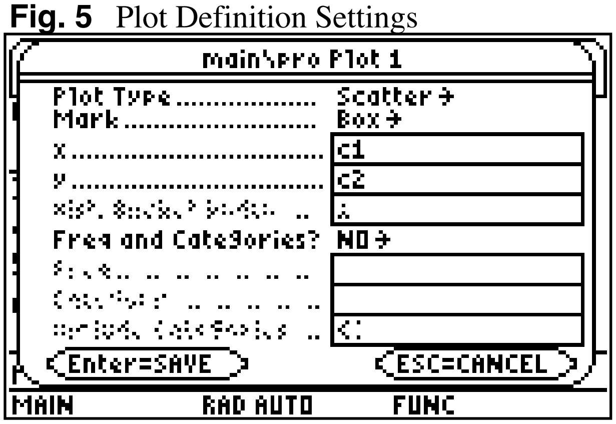
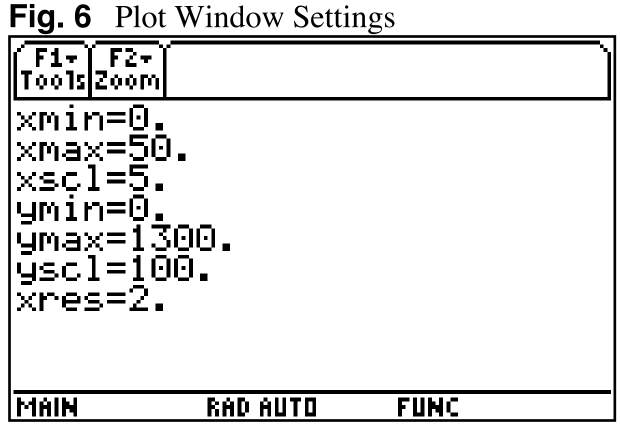


Fig. 6 Plot Window Settings



Define the data range for the plot. Press **WINDOW** (green key of **F2**) and enter the settings shown in Fig. 6. Next press **Y=** (green key of **F1**). If there are any functions displayed be sure they are deselected. A function is selected if it is preceded by a check mark, ✓ (Fig. 7). To deselect a function, use the arrow keys to highlight that function and press **F4**. **F4** is used to toggle function selection. Finally, press **GRAPH** (green key of **F3**) to produce the scatter plot (Fig. 8).

Fig. 7 Deselecting Functions

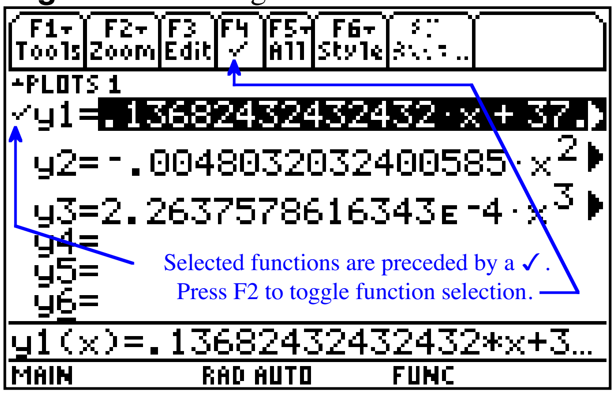
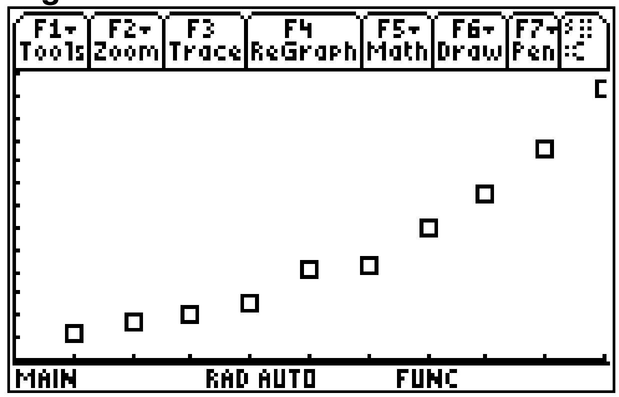


Fig. 8 Scatter Plot of Data

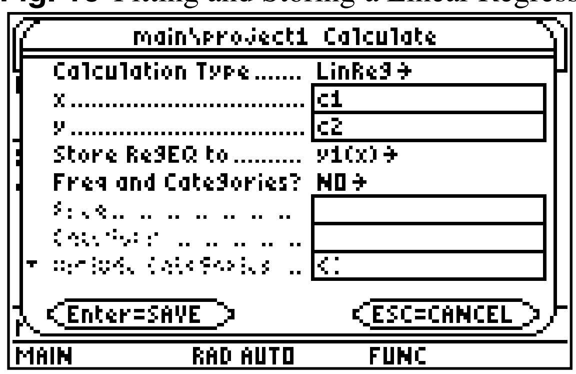


- **Fitting a Linear Function**

The scatter plot in Fig. 8 suggests that a straight line relationship is reasonable. Return to the Data Editor window by pressing the **APPS** key and choosing 6:Data/Matrix Editor and then 1:Current… (Fig. 9). To fit a regression, press the **F5** key to select the Calc menu option and bring up the Calculate window (Fig. 10).

Fig. 9 Reopening a Data Set

Fig. 10 Fitting and Storing a Linear Regression



From the menu selection, set Calculation Type to 5:LinReg. The *x*-variable is c1 and the *y*-variable is c2. In order to plot the results, I choose y1(x) as the function in which to Store RegEQ. This will overwrite the current contents of y1(x) with the regression function and automatically select y1(x) for plotting. Press **ENTER** to produce the regression results shown in Fig. 11. Finally, press **GRAPH** (green key of **F3**) to produce the scatter plot with the regression line superimposed (Fig. 12).

Fig. 11 Linear Regression Results

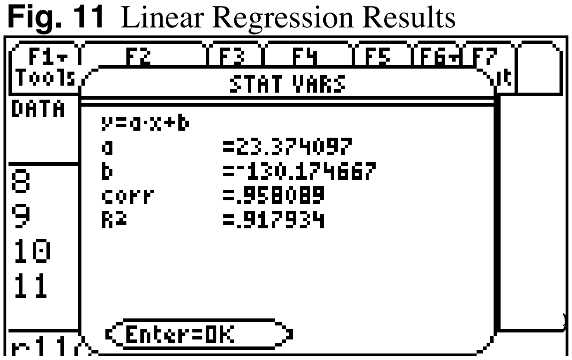
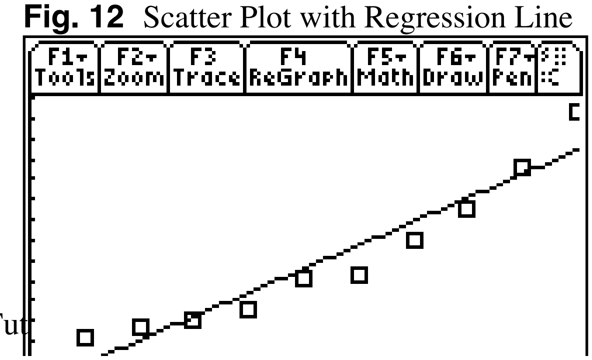


Fig. 12 Scatter Plot with Regression Line



The values of a and b are displayed on the screen along with model that was fit. Based on the output the fitted model is $N(t) = -130.17 + 23.374t$. Qualitatively it would appear from the graph in Fig. 12 that a linear function is a reasonable model. The standard quantitative measure of the usefulness of the regression model is R^2, the coefficient of determination. R^2 measures the fraction of the variability in y that is explained by its linear relationship to x and can take values between 0 and 1. The TI-89 prints R^2 as part of the standard regression output. Since $R^2 = 0.917934$ we conclude that approximately 91.8% of the variability in N is explained by its linear relationship to t. (If the fit were perfect R^2 would equal 1.)

- **Fitting a Quadratic Function**

The scatter plot in Fig. 8 reveals a slight curvilinear trend to the data suggesting a polynomial model might be appropriate. Return to the Data Editor window as explained in the "Fitting a Linear Function" section. Press **F5** to select the Calc menu option and bring up the Calculate window. Use the drop down men to set Calculation Type to 9:QuadReg. (Third and fourth degree polynomials can be fit by choosing 3:CubicReg and A:QuartReg respectively.) The x-variable is c1 and the y-variable is c2. In order to plot the results, choose y2(x) as the function in which to Store RegEQ (Fig. 13). Doing this will overwrite the current contents of y2(x) and automatically select y2(x) for plotting. Press **ENTER** to produce the regression results shown in Fig. 14.

Fig. 13 Fitting and Storing a Quadratic Regression

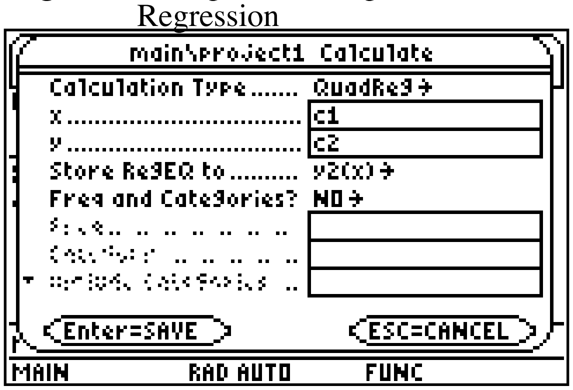
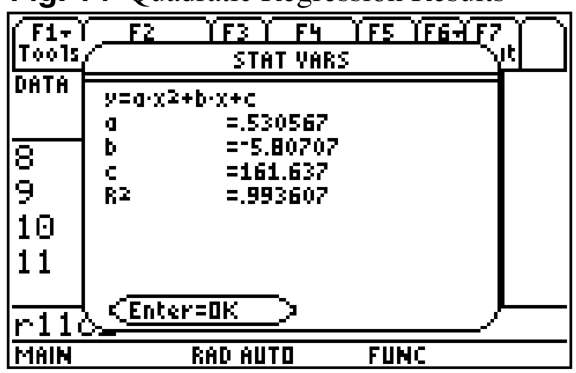


Fig. 14 Quadratic Regression Results



The values of a, b, and c are displayed on the screen along with model that was fit. Based on the output the fitted model is $N(t) = 0.531t^2 - 5.807t + 161.637$. From R^2 we conclude that approximately 99.4% of the variability in N is explained by its quadratic relationship to t. Finally, press **GRAPH** (green key of **F3**) to produce a plot with the quadratic regression function and linear regression function superimposed on a scatter plot of the data (Fig. 15).

Fig. 15 Quadratic and Linear Regression Plots