

Linear Regression Equation

Written by Professor Stephen T. Thornton, Department of Physics, University of Virginia, Charlottesville, VA 22903; adapted from earlier work of Daniel Mulligan, Great Bridge High School, Chesapeake, VA 23320.

In this example we will use real life data to produce a table, graph the data, and fit the data with a linear regression equation. We will use that fit to make predictions.

Robert Wadlow may have well been the world's tallest man. He lived in the United States and played basketball for his high school team. His height from age 8 until his death at age 22 is given below.

<u>Age</u>	<u>Height</u>	<u>Age</u>	<u>Height</u>	<u>Age</u>	<u>Height</u>
8	6.0	13	7.1458	18	8.2917
9	6.1875	14	7.4167	19	8.4583
10	6.4167	15	7.6667	20	8.5625
11	6.5833	16	7.8542	21	8.6875
12	6.8750	17	8.0417	22	8.9250

We want to make a plot of height (y-axis) versus age (x-axis) using a TI-83 graphing calculator.

STEPS

- 1) Clear the data in the registers L1 and L2: Press STAT and choose 4:ClrList. Then press L1, L2; press ENTER.
- 2) To enter the data in the table above, Press STAT and choose 1:Edit.
- 3) We will enter the age (x-axis) in L1 and the height (y-axis) in L2. Type all the ages into L1 before entering the heights in L2. Go through and check all your data again to make sure you entered it correctly.
- 4) Press STAT PLOT. Choose 1:Plot1. We then need to make some choices about the kind of plot we want to see.
- 5) Choose ON. Under Type: choose the kind of plot you want. The first or second one is a good choice. For Xlist, choose L1; for Ylist, choose L2. For Mark, choose the kind of data point you want. You can come back later and change your choices by pressing STAT PLOT.
- 6) Press ZOOM. Choose 9: ZoomStat. This will automatically fit your data plot to the size of the window. You can change it later by going to WINDOW and choosing your axes. The data plot should appear.
- 7) If at anytime you want to see the data plot, Press GRAPH.

- 8) Now let's try to fit the data. The data appear to be well approximated by a straight line, so let's attempt to fit the data with $y = ax + b$, a linear equation. Press STAT, scroll to the right and highlight CALC. Under CALC choose 4:LinReg(ax+b) and press ENTER. The best fit values of a and b are given. You might want to write those down. If you want to find out how good your fit is, press CATALOG, scroll down until you find DiagnosticOn and press ENTER to choose it, then press ENTER again to turn it on. DONE should be printed out on your screen. Then go back to STAT, choose CALC again and refit the data. Under the best fit values of a and b will be given a value of r (and r^2) which is the *Correlation Coefficient*. A value of r near 0 is a bad fit, and a poor choice of fitting curve was chosen. There are many other functions besides a linear line we might have chosen. Our value of r is near 1 indicating a linear fit was a good choice.
- 9) Now let's add our best fit line to the data plot. Press Y=. If there is an equation after Y1=, then press CLEAR to delete it. Press VARS, choose 5:Statistics, scroll to the right to highlight EQ and choose 1:RegEQ. Your best fit equation should then appear in Y1=.
- 10) Press GRAPH to have the best fit line appear with the data on the plot. It should approximate the data.

Exercises for this Example:

- 1) What is the average change (mean) of Robert's height from age 8 to 22? Decide how to find this.
- 2) Write down the linear regression equation that best represents the data above and gives Robert's height versus age?
- 3) Make some predictions about Robert's height at different ages assuming he lived until the age of 30? What would be Robert's height at age?
 - a) 3
 - b) 5
 - c) 25
 - d) 30

This may be a little perplexing, but there is more than one way to do this. First, let's change the plot so we can look at the predictions of Robert's height from ages 0 to 30. Press WINDOW, choose Xmin = 0, Xmax = 30. You will also have to change Ymin and Ymax to make the data fit on the plot. You do that. Then Press CALC, choose 1:Value. Type the value of X(age) for which you want to know the height, press ENTER, and Y (the height) will be displayed.