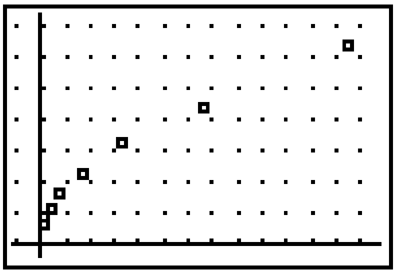
Plotting the RESIDUALS to Check Linearity.

If the data points don't seem to follow a straight line reasonably well, **STOP!** Your calculator will obey you if you tell it to perform a linear regression, but if the points don't actually fit a straight line then it's a case of "garbage in, garbage out."

For instance, consider this example from De Veaux, Velleman, and Bock, *Intro Stats* (Pearson Addison Wesley, 2009), page 179. This is a table of recommended f/stops for various shutter speeds for a digital camera:

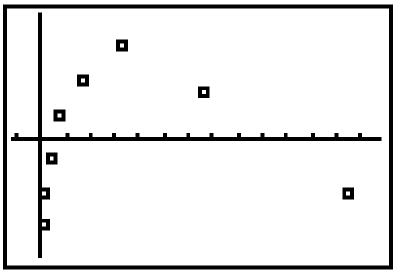
Shutter speed (x)	1/1000	1/500	1/250	1/125	1/60	1/30	1/15	1/8
f/stop (y)	2.8	4	5.6	8	11	16	22	32



If you try plotting these numbers yourself, enter the shutter speeds as fractions for accuracy: don't convert them to decimals yourself. The calculator will show you only a few decimal places, but it maintains much greater precision internally.

You can see from the plot at right that these data don't fit a straight line. There is a distinct bend near the left. When you have anything with a curve or bend, linear regression is wrong.

Suppose you continued anyway?



Refer to the scatter plot of f/stop against shutter speed. I said then that it was not a straight line, so you could not do a linear regression. If you missed the bend in the scatterplot and did a regression anyway, you'd get a correlation coefficient of r = 0.98, which would encourage you to rely on the bad regression. But plotting the residuals (at right) makes it crystal clear that linear regression is the wrong type for this data set.

This is a textbook case (which is why it was in a textbook): there's a clear curve with a bend, variation on both sides of the x axis is not consistent, and there's even a likely outlier.