Investigating Types of Errors: Checks and Balances

This activity uses the interactive Excel file to explore a variety of types of errors and how they influence data. Let's consider a scientist checking a set of standard masses on four different laboratory balances. We will assume each standard mass is the true value with no error. For each balance below, plot the balance mass as a function of the standard mass and perform a linear regression on each set.

Standards (grams)	Balance 1	Balance 2	Balance 3	Balance 4
2.00	2.00	4.00	1.70	1.81
5.00	5.00	7.00	4.25	5.57
10.00	10.00	12.00	8.50	9.57
15.00	15.00	17.00	12.75	15.67
20.00	20.00	22.00	17.00	20.89
25.00	25.00	27.00	21.25	24.12

Record the regression equation and value of r^2 in the table below.

	Balance 1	Balance 2	Balance 3	Balance 4
Linear regression equation				
r^2				
Describe the error				

For two variables that are in total agreement (no error) the y = x line fits the data. This is the case for Balance 1 above. There is no error in the mass determination on Balance 1. For Balance 1 the equation is $mass_{balance} = mass_{standard}$. This is a common comparison method we are going to use to investigate the possible errors that could occur. With this reference in mind, describe the errors in Balances 2, 3, and 4. You might want to plot these and compare them to the y = x line.

Now go to the interactive Excel spreadsheet. The errors worksheet allows you to explore a variety of errors. Residuals are the difference, Δy, of y-measured minus y-calculated from the regression equation. The residuals worksheet responds to the errors.

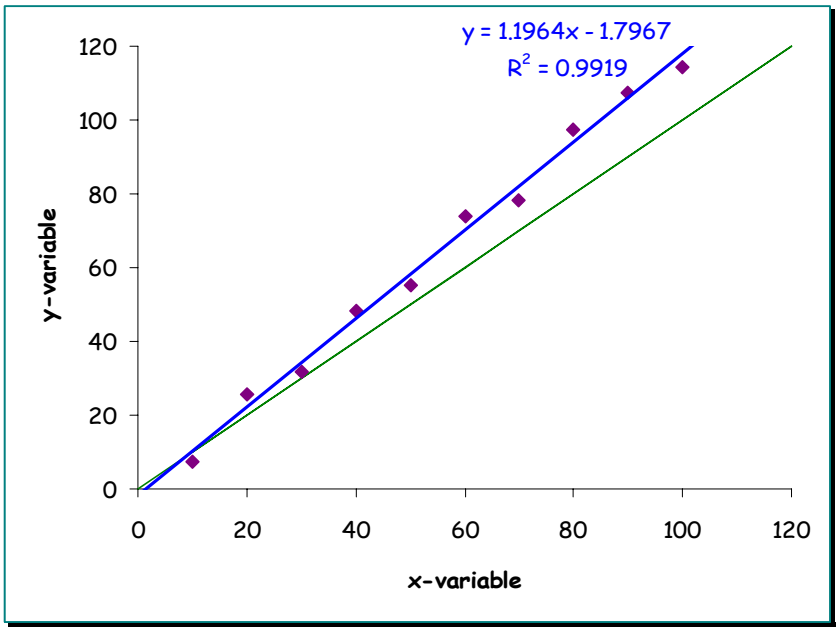
Fill in the information about the various errors. The data should start out lying on the y = x line. Adjust one parameter at a time while holding all the others to no error. For each type of error, consider how the slope, y-intercept, and r^2 respond. For the residuals plot, look for patterns or lack of patterns in the data.

Error	How does the graph respond compared to the y = x line?	How does r^2 change if at all?	View the residuals graph and explain how the points are distributed.
Random			
Proportional systematic			
Constant systematic			
Minimal curvature in data			
"Bad cal" or curvature at larger x values			

What type of error occurred for the balances?

	Balance 2	Balance 3	Balance 4
Type of error			

Now in the real world combinations of errors occur. What are the errors exhibited on the graph below?



Systematic errors are usually correctable after investigation. They may involve a miscalibrated measuring device or instrument. How would the graph above change if the systematic error was corrected?

Now place a small amount of curvature into the spreadsheet with all the other error set at none. Observe the residuals graph. What happens to the residuals in this case?

Could you discover the curvature from the r^2 value?

Curvature shows up in the residuals as a pattern in their distribution. The r^2 value changes (decreases) very little for a small amount of curvature. A good linear fit should show a random pattern in the residuals. Increase the random error in this data with a small amount of curvature. What happens to the residuals?

Random error can disguise the non-linear character of a data set. Random counting error in long-lived radioactive material could disguise the exponential nature of radioactive decay.