Performing Transformations on Data

Once you have considered the residual plot and seen a pattern, you know that a linear relationship does NOT exist.

Great .... now what ??

Now we try and 'straighten' the relationship between the variables by performing a transformation or changing them in one of 3 different ways.

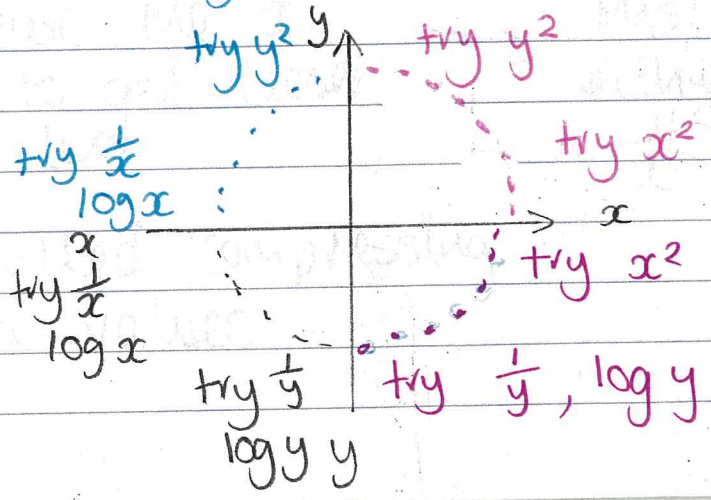
squaring x^2 or y^2

reciprocal $\frac{1}{x}$ or $\frac{1}{y}$

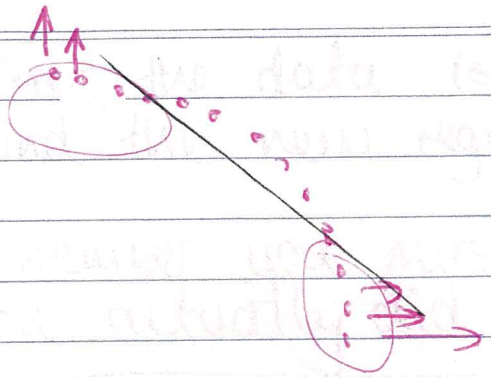
logarithm $\log x$ or $\log y$ *

* always use log base 10 * your CAS should do this automatically.

So how do you decide which one to do?



What shape does the original data curve take?



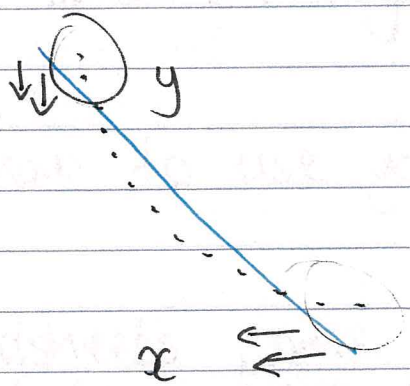
Take the curve in the first quadrant

we want to increase the x values to push them towards the line → done by x^2

or we want to increase the y values → done by y^2

* We don't know which one will work best
* You just have to try and see what you get.

Take the third quadrant



we want to decrease the y values to get closer to the line → $\frac{1}{y}$ or $\log y$

or decrease the x value to get closer to the line. → This is done by either $\frac{1}{x}$ or $\log x$

This is called compressing the data values.

Once the data is transformed, we find the new regression equation (or model)

Remember you must use the transformation you actually did in the new equation.

It is NO LONGER $y = a + bx$

it could be $y = a + bx^2$ or $\frac{1}{y} = a + b$

For each transformation you perform, determine the equation using your CAS

Also determine r^2 for each (coeff of determinatio

The transformation with HIGHEST r^2 is the most accurate model with the highest PREDICTIVE POWER.

Make sure you know how to use your CAS to help you.

* When in Lists & Spreadsheets, once you have entered your data, MAKE the CAS DO THE WORK.

* Label your columns wisely NOT x and y use xval, yval if you have to but better to use actual variable name

* Then use the next columns to calculate x^2, $\frac{1}{x}$, $\log x$ etc

* All you need to do now is select the appropriate columns when doing (menu) (6) (1) (4)

* Remember to label the columns wisely !! $xval^2$ or $\frac{1}{months}$ or $\log(age)$

etc.

* In the row below the column title use " = " then type in what transformation you are performing

Your CAS will then fill in the column automatically for you

If your CAS does NOT accept what you enter as a label or command, change your original headings and try again !!