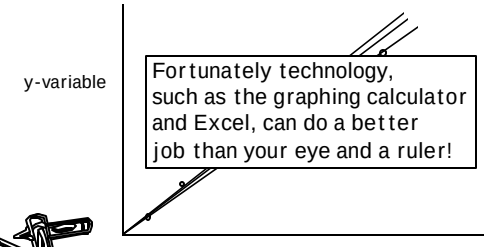
The Best-Fit Line

Linear Regression

How do you determine the best-fit line through data points?



PGCC CHM 103 Sinex

The Equation of a Straight Line

$$y = mx + b$$

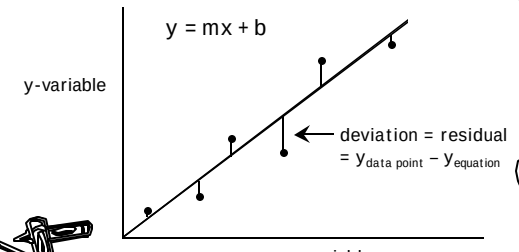
where m is the slope or $\Delta y/\Delta x$ and b is the y-intercept

In some physical settings, b = 0 so the equation simplifies to:

$$y = mx$$

PGCC CHM 103 Sinex

Linear regression minimizes the sum of the squared deviations



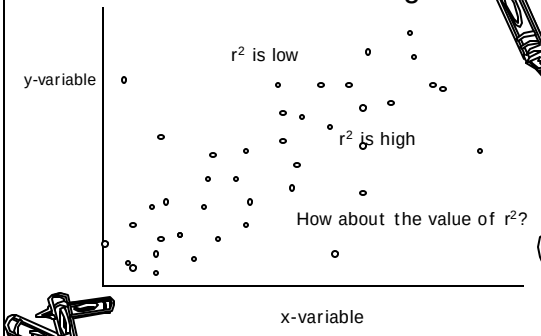
PGCC CHM 103 Sinex

Linear Regression

- Minimizes the sum of the square of the deviations for all the points and the best-fit line
- Judge the goodness of fit with r^2
- r^2 x100 tells you the percent of the variation of the y-variable that is explained by the variation of the x-variable (a perfect fit has $r^2 = 1$)

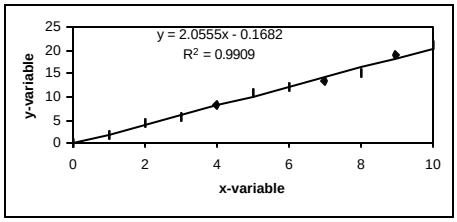
PGCC CHM 103 Sinex

Goodness of Fit: Using r^2



PGCC CHM 103 Sinex

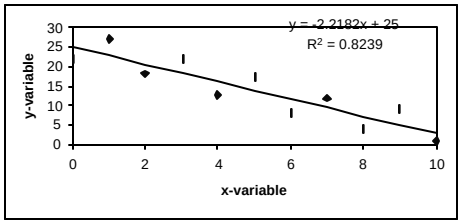
Strong direct relationship



99.1% of the y-variation is due to the variation of the x-variable

PGCC CHM 103 Sinex

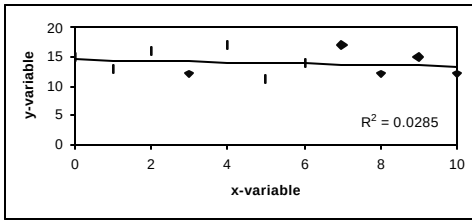
Noisy indirect relationship



Only 82% of the y-variation is due to the variation of the x-variable - what is the other 18% caused by?

PGCC CHM 103 Sinex

When there is no trend!



No relationship!

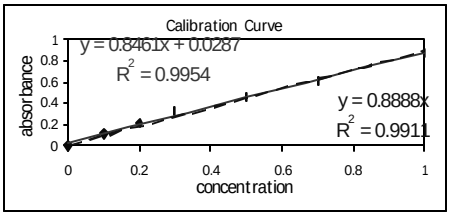
PGCC CHM 103 Sinex

In Excel

- When the chart is active, go to chart, and select Add Trendline, choose the type and on option select display equation and display r^2
- For calibration curves, select the set intercept = 0 option Does this make physical sense?

PGCC CHM 103 Sinex

Does the set intercept = 0 option make a difference?



Using the set intercept = 0 option lowers the r^2 value by a small amount and changes the slope slightly

PGCC CHM 103 Sinex

The equation becomes

$$A = mc$$

or

$$A = 0.89c$$

99.1% of the variation of the absorbance is due to the variation of the concentration.

PGCC CHM 103 Sinex