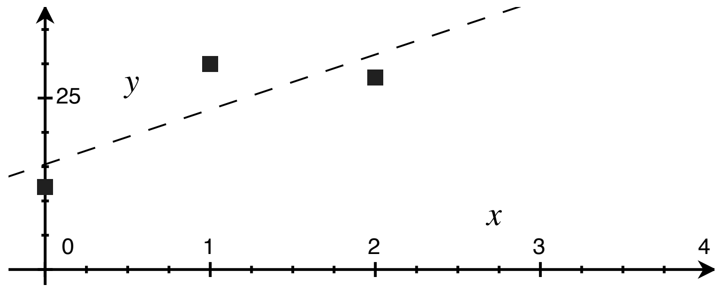
What is R^2? I mean, seriously...

In a previous class you found the line of best fit for the following data set

x_i	y_i
0	12
1	30
2	28



The equation of that line was:

$$\hat{y} = 15.33 + 8x$$

Let's explore what R^2 is all about:

(1) Compute the variability/variance in y. In other words, find: $\sum(y_i - \bar{y})^2$

(2) Next, compute the variability in the residuals. That is, find: $\sum(y_i - \hat{y}_i)^2$

(3) Divide the answer in (2) by (1).
This gives the percent of variability in the response variable that remains in the residuals – it is the variability in y_i that is NOT accounted for by the regression model.

(4) Compute 1− the answer in (3). That gives you R^2, the variability that IS accounted for by the regression model.

To do the above computations you will need the values of $\bar{y}$ and $\hat{y}_i$. They are:

$\bar{y} = 23.33, \quad \hat{y}_1 = 15.33,\ \hat{y}_2 = 23.33,\ \hat{y}_3 = 31.33$

If all goes well, the answer you should get is: $R^2 = 0.6575$.
If you fit the regression line in R [with `lm(y ~ x)`] you should get the exact same R^2.