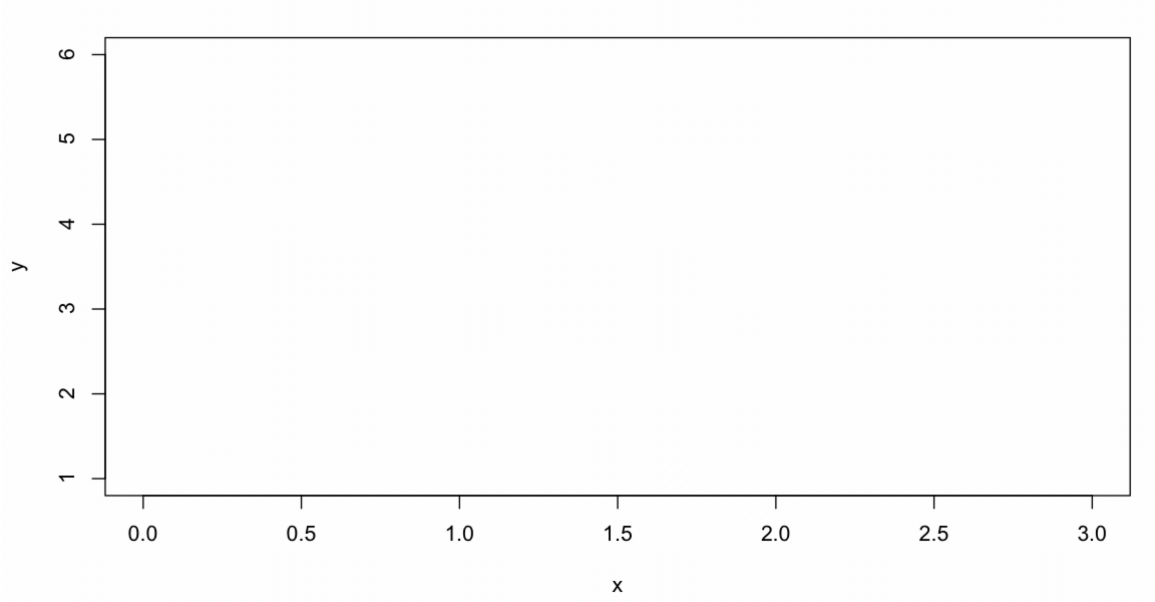
Stat 61: Simple Regression Worksheet

I believe it is worthwhile to go through the steps of fitting a linear regression and constructing an ANOVA table with a simple data set. I chose these numbers to make it easy enough to do by hand.

i	x_i	$x_i - \bar{x}$	y_i	$y_i - \bar{y}$	$(x_i - \bar{x})(y_i - \bar{y})$	$\hat{y}_i$	$y_i - \hat{y}_i$
1	1		1				
2	1		3				
3	2		1				
4	2		3				
5	3		4				
6	3		6				

a) Plot the x, y pairs on the graph below. Compute $\bar{y}$ and draw in a horizontal line at $y = \bar{y}$.



b) Fill in the blanks in the table and below. Add the least squares line to the graph.

$$s_x^2 = \_\_\_\_\_ \qquad s_y^2 = \_\_\_\_\_ \qquad r = \frac{1}{n-1}\frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{s_x s_y} = \_\_\_\_\_$$

$$\hat{\beta}_1 = r\sqrt{\frac{s_y^2}{s_x^2}} = \_\_\_\_\_ \qquad \hat{\beta}_0 = \bar{y} - \beta_1 \bar{x} = \_\_\_\_\_ \qquad (\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i)$$

c) Fill in the ANOVA table for regression and compute R^2 and R^2_{adj}. Verify that SSM+SSE = SST. Compare the sum of squares error for the least squares fit and for the line $\hat{y} = \bar{y}$.

source	df	SS	MS
Model			
Error			
Total			

d) Find the standard error for $\hat{\beta}_1$: $SE_{(\hat{\beta}_1)} = \frac{\text{rmse}}{\sqrt{\sum(x_i - \bar{x})^2}}$. Evaluate the t statistic (and degrees of freedom) for testing $H_o : \beta_1 = 0$ vs. $H_1 : \beta_1 \neq 0$. The F-Ratio is the Mean Square Model divided by the Mean Square Error. Show this equals the square of your t statistic (for simple regression, the "whole model" F test is equivalent to this t test).

e) Repeat for the binary x_i given in the table below. Show that the t statistic (and degrees of freedom) for testing $H_o : \beta_1 = 0$ vs. $H_1 : \beta_1 \neq 0$ match those of the pooled 2-sample t statistic for testing $H_o : E(Y_i|x_i = 0) = E(Y_i|x_i = 1)$.

i	x_i	$x_i - \bar{x}$	y_i	$y_i - \bar{y}$	$(x_i - \bar{x})(y_i - \bar{y})$	$\hat{y}_i$	$y_i - \hat{y}_i$
1	0		1				
2	1		3				
3	0		1				
4	1		3				
5	0		4				
6	1		6				

source	df	SS	MS
Model			
Error			
Total			