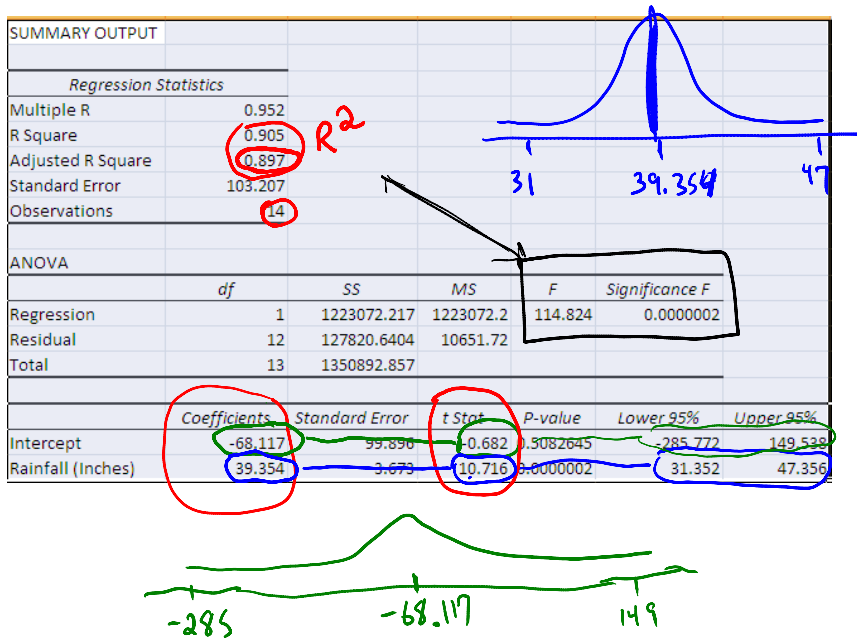


# MGMT 5113 Unit 11, Video 2



$Y = \text{constant term} + (\text{slope} \times \text{variable})$

Bales of hay = some number + (slope x inches of rain)  
*intercept*      *rainfall (inches)*

### Regression equation

Bales of hay = -68.117 + (39.354 x inches of rain) + error

Using the regression equation for forecasting:

Suppose you think there will be 40 inches of rain next year.  
 What is your best estimate for your hay harvest next year?

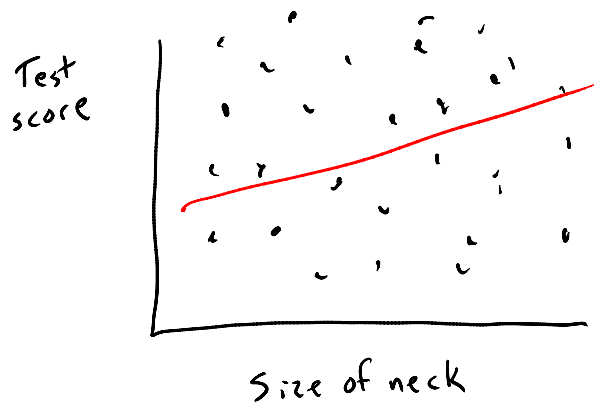
Bales of hay = -68.117 + (39.354 x 40) = 1,506.043 bales

$R^2$   
"R square"

The portion of the variation of the "y variable" (bales of hay) that is explained the variation in the "x variable: (inches of rain).

$R^2 = .90$  90% of the variation in bales of hay is explained by the variation in rainfall. (10% is explained by something not in the equation, e.g. temperature or bugs.

$R^2$  can range from 0 (nothing explained by the regression) to 1 (everything explained by the regression equation). Higher is better. In business, roughly speaking, an  $R^2$  of .50 or higher is pretty good.



Example of  $R^2 = 0$   
 Neck size does not cause test scores

t-statistics test the hypothesis that the values of the intercept and the coefficient are NOT zero. A higher t-statistic means we are more sure that the values are not zero.

Our t-statistic for the intercept = .682. This is too small to give us confidence that the value of -68.117 is correct, and in fact we cannot reject the possibility that the true value for the intercept is zero.

Our t-statistic for the slope = 10.716. This is very large (any value bigger than 2.0 is pretty good), so we can reject the possibility that the true slope is zero and we have confidence that the coefficient estimate of 39.354 is correct.

The f-statistic is another way to measure how much of the variation in the "y variable" (bales of hay) is caused by variation in the "x variable" (rainfall). A bigger F is better--it means that more of the variation is explained by your regression.

An easy way to figure out how good your regression is to use the "significance F" figure that is related to the F-statistic. A bigger F-statistic means a smaller significance f, and that is good.

For example, when significance F = .05 you are 95% confident that your regression is explaining some of the variation of the y variable.

For example, when significance F = .01 you are 99% confident that your regression is explaining some of the variation of the y variable.

**WARNING:** An MBA is not qualified to run regression analysis by her/himself. Many things can go wrong beyond the scope of this class. It is very important that a statistician or other trained person (economist) has a role in the regression.