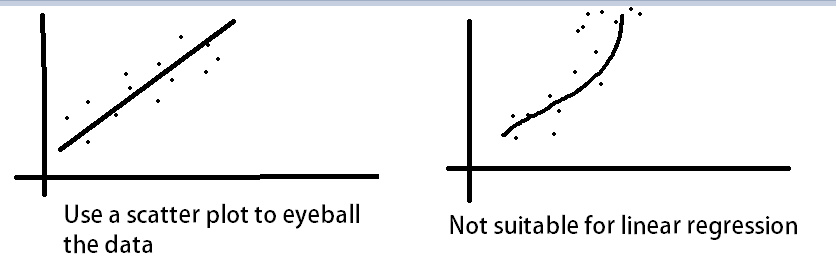
**Assumptions in linear regression:**

1. **The Quantitative Data Condition.**

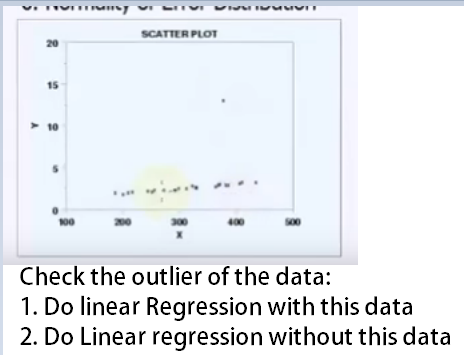
Should be numbers and not characters or alphabets.

One column has to be dependent and other has to be independent

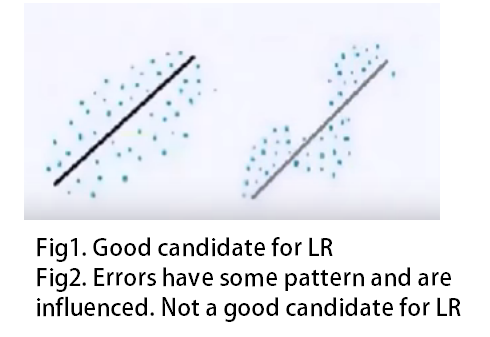
1. **The Straight Enough Condition (or “linearity”).**



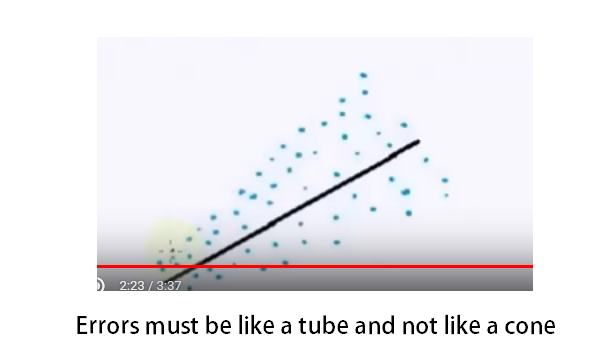
1. **The Outlier Condition.**



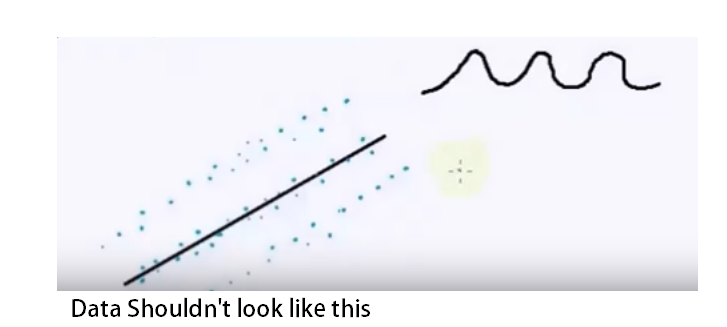
1. **Independence of Errors**



1. [**Homoscedasticity**](http://www.statisticshowto.com/homoscedasticity/)

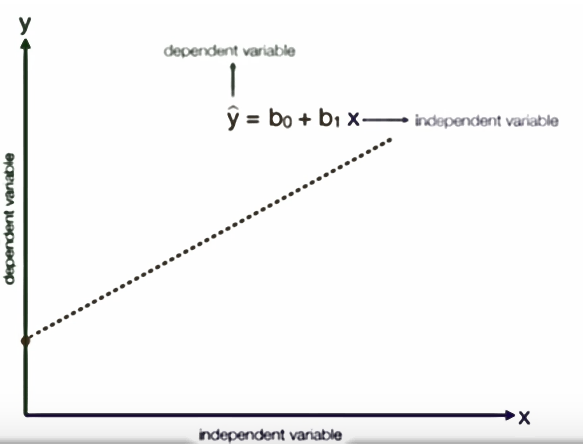


1. **Normality of Error Distribution.**



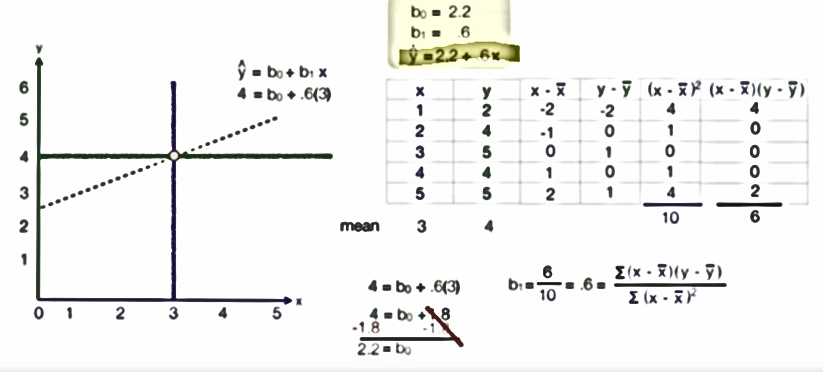


Things to find in Linear Regression:



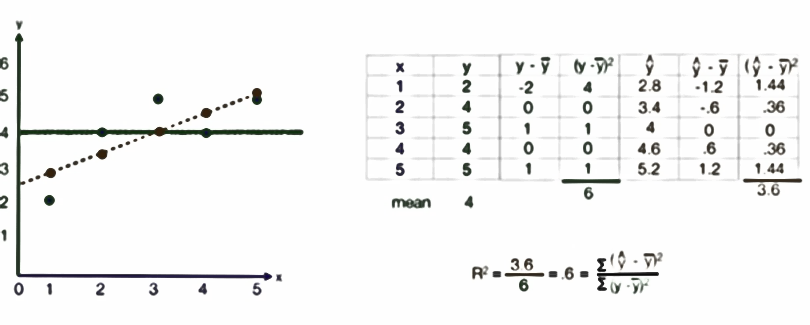
**B1 is also called Covarience**

Find the equation and plot the Regression line



**How to know if the Fit is good:**

1. **Finding R^2**

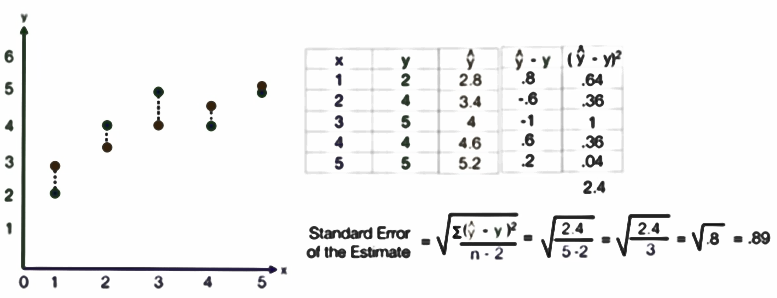


R^2 = 1 >> Best fit>> Less error>> variables are dependent

R^2 = 0>> No relationship at all

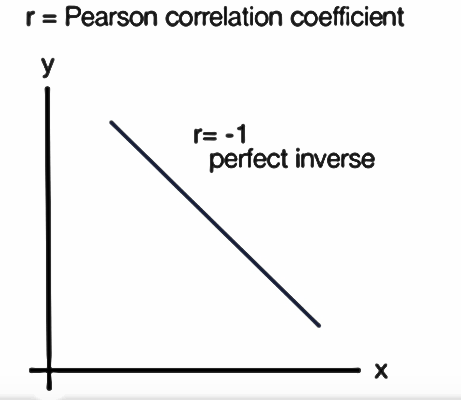
Lies between 0 and 1.

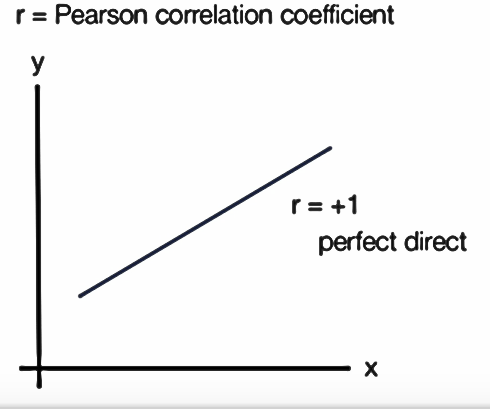
2. **Standard Error of the estimate OR Mean Square Error(MSE).**

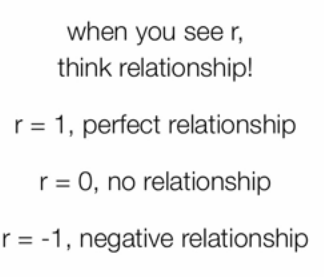


0.89 is pretty good. i.e the line will vary by +0.89 to -0.89

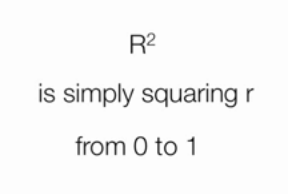
**Other things good to know:**



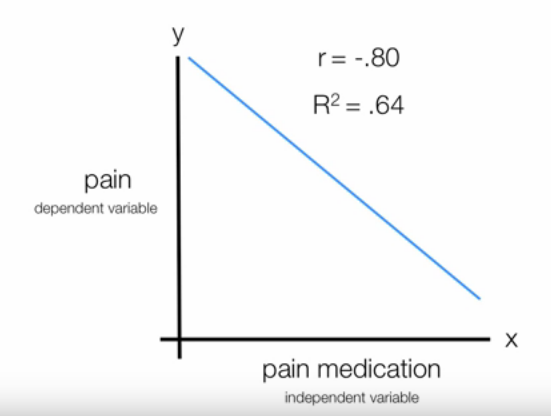




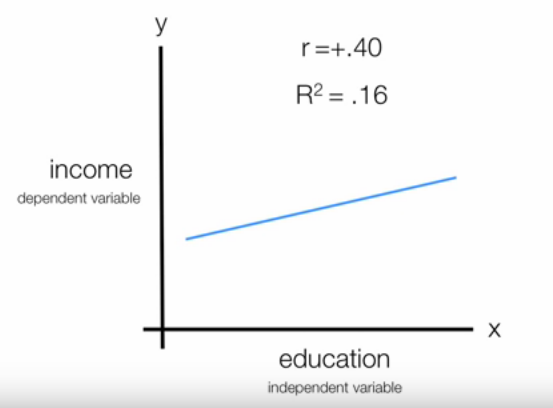
r measures relationship whereas R2 measures intensity.



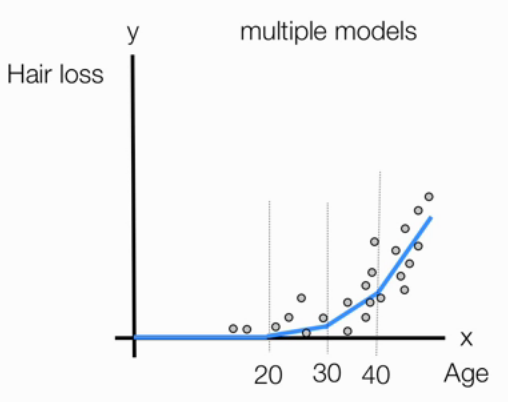
Eg:

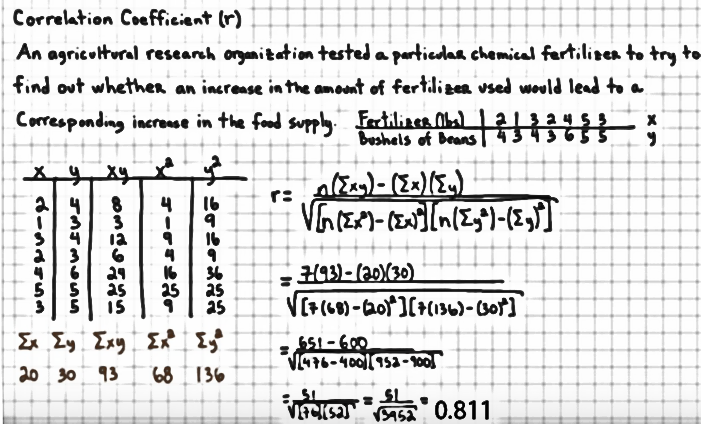


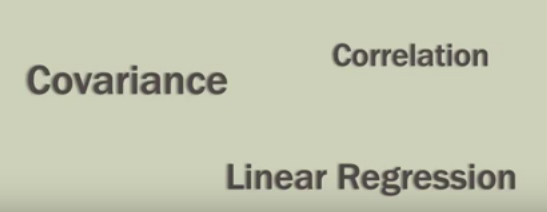
Change of pain medication can 64% chance to take care of the pain.



16% change in income, can be explained by education. r is just a relationship.







All 3 are closely related.

