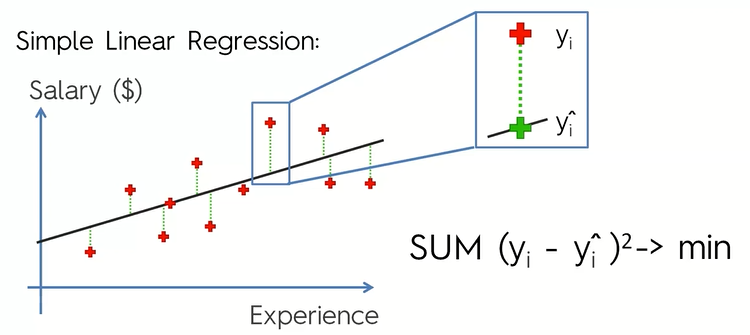
**R SQUARED INTUITION**

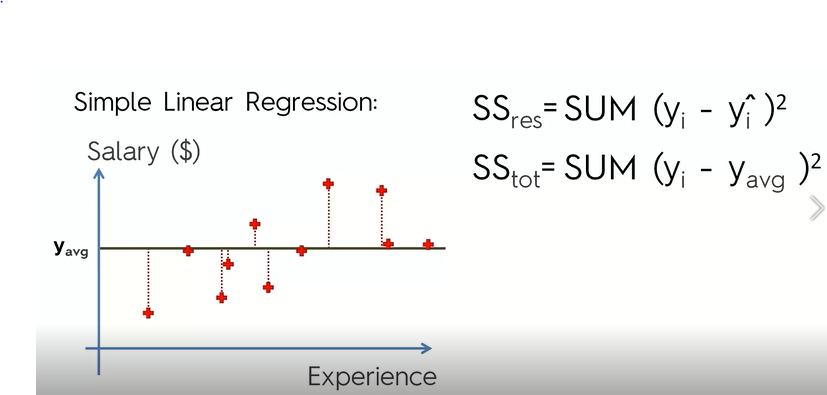
Simple linear regression is build by ordinary Least Square Method where we are minimizing the sum of squares of differences of yi and yicap, i.e,we are looking at the difference, we are squaring it for all the observations and then we are counting the sum. The line that has the smallest sum will be the best fitting line or will be the simple linear regression model.

ssres = SUM(yi - yicap)2



Now, instead of drawing the regression line, let us draw the average line and project our observations on this average line and then calculate the distances and square them.

sstotal = SUM(yi – yicap)2



R2  = 1 – ssres/sstot

R2 is telling us that how good is our line as compared to the average line.

In order to get the best fitting line you run a linear regression, so we minimize ssres

ssres decreases => ssres decreases => (1-ssres/sstot) increases

If ssres = 0

=> your tresd line that you are modeling goes through all the records and in that case R2 = 1

This is the ideal scenario.

The more R2closer to 1 the better

R2 can be negative( if ssres fits your data worse than your average line.

Normally it is between 0 and 1.