Regression: Regression can be of two types

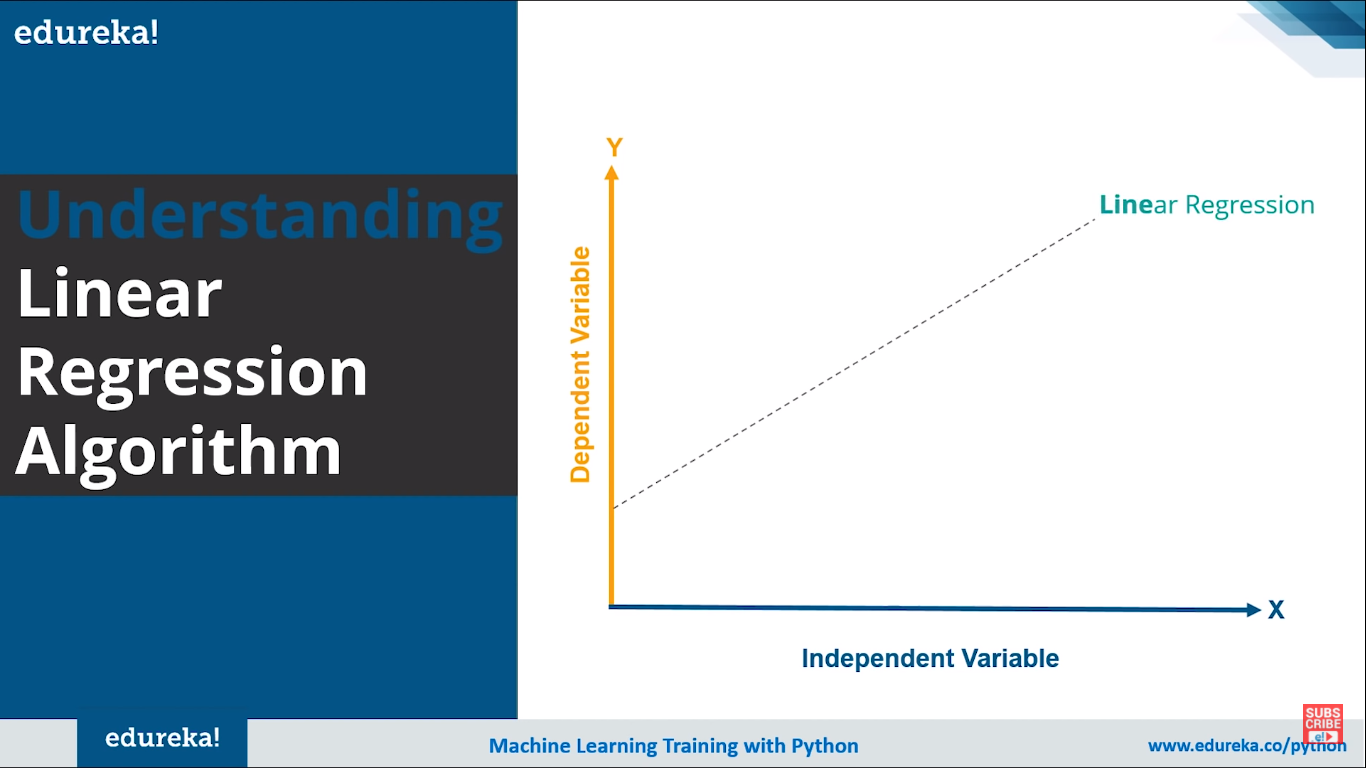
**Linear Regression:** It is used to predict the value of continous variable like sales and money spent on adds.

**Logistic Regression:** It is used to predict the cateogorical data…like prediction of True/False, Rain forecast[yes/No] for a day in week

Linear Regression:

Lets now know about the linear regression in detail. As said this regression deals with the prediction of the continous variable.

Here we have two variables dependent variable and independent variable and we plot the depenedt variable for the various values of the independent variable.



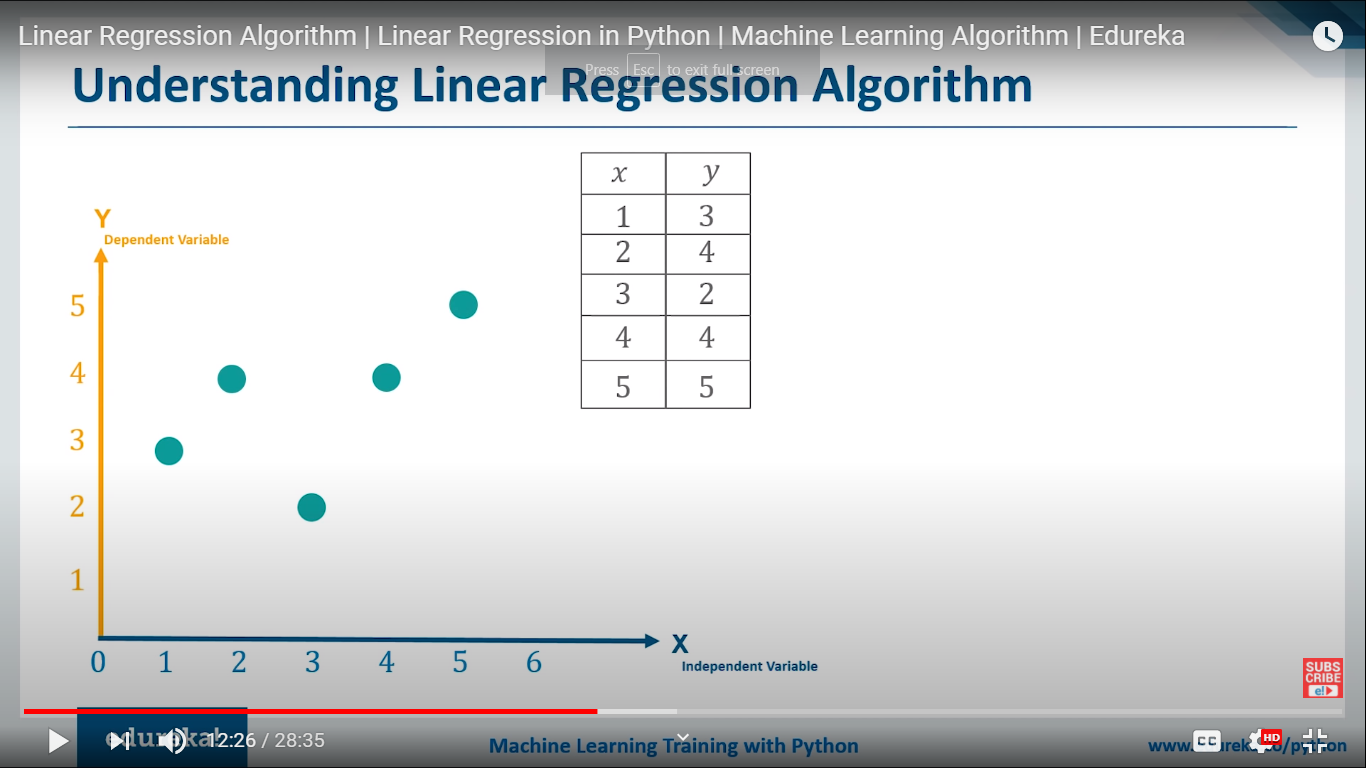
We predict the the dependent variable using the future or various values of independent variable in the Linear Regression.

If y increases as x increases its positive linear regression

If y decreases as x increases its negative linear regression

Lets us take the data and understand in clear:

Lets see the data for a LR using ordinary least squares:

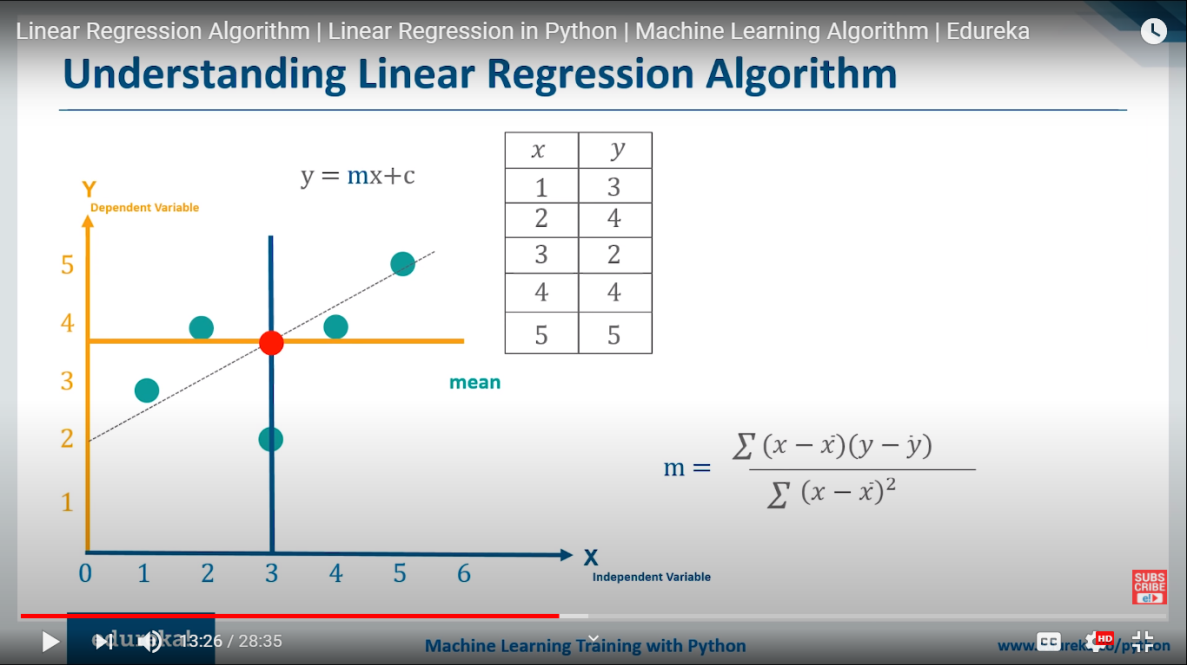


Here we plot the y values against the different x values

So we now find a line that fits the data for that we take the following actions:

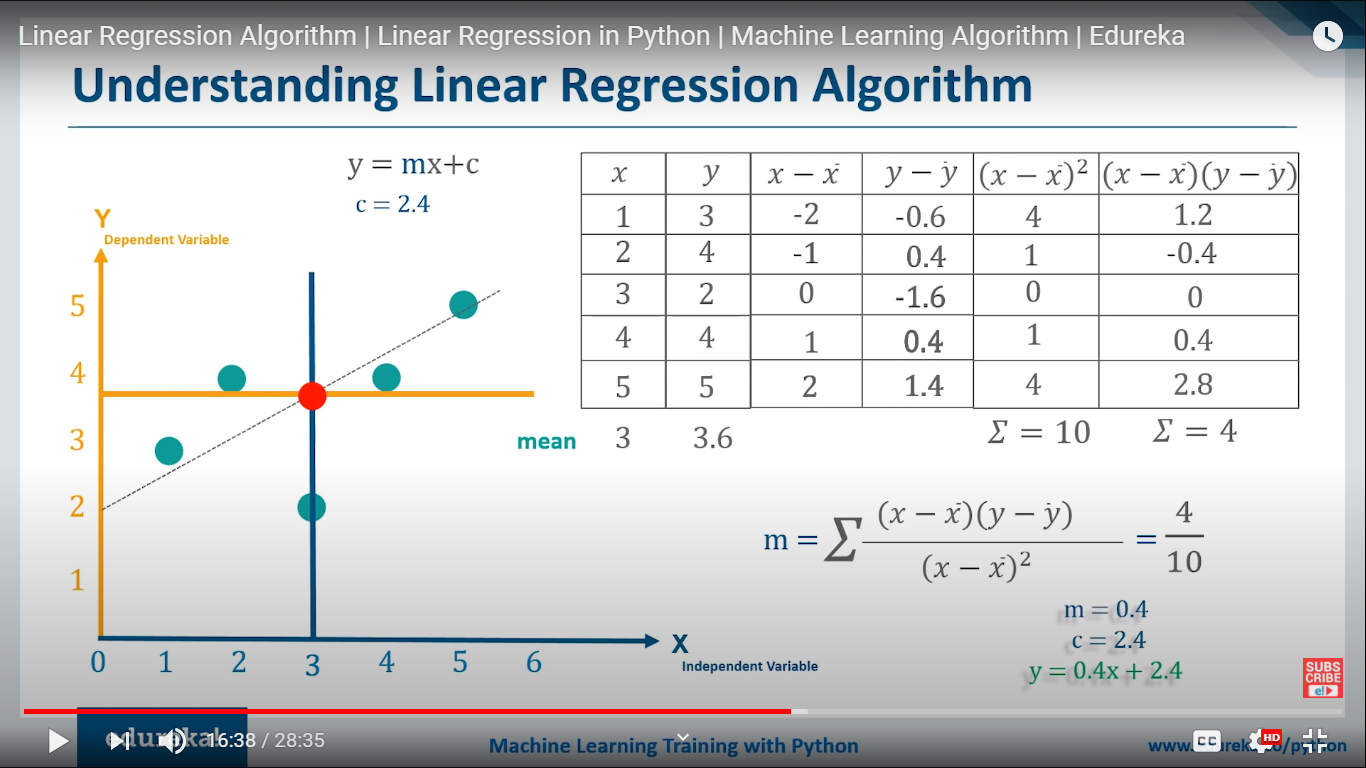
* Find the median of the x data
* Take that point as the measure of the our regression line
* As we know there can be infinite number of lines passing through that point our task is to find the best line that fits the data t mean we have find the slope and the costant that fits the data given

**How do we that?**



Finding the slope of the line passing through our center point.

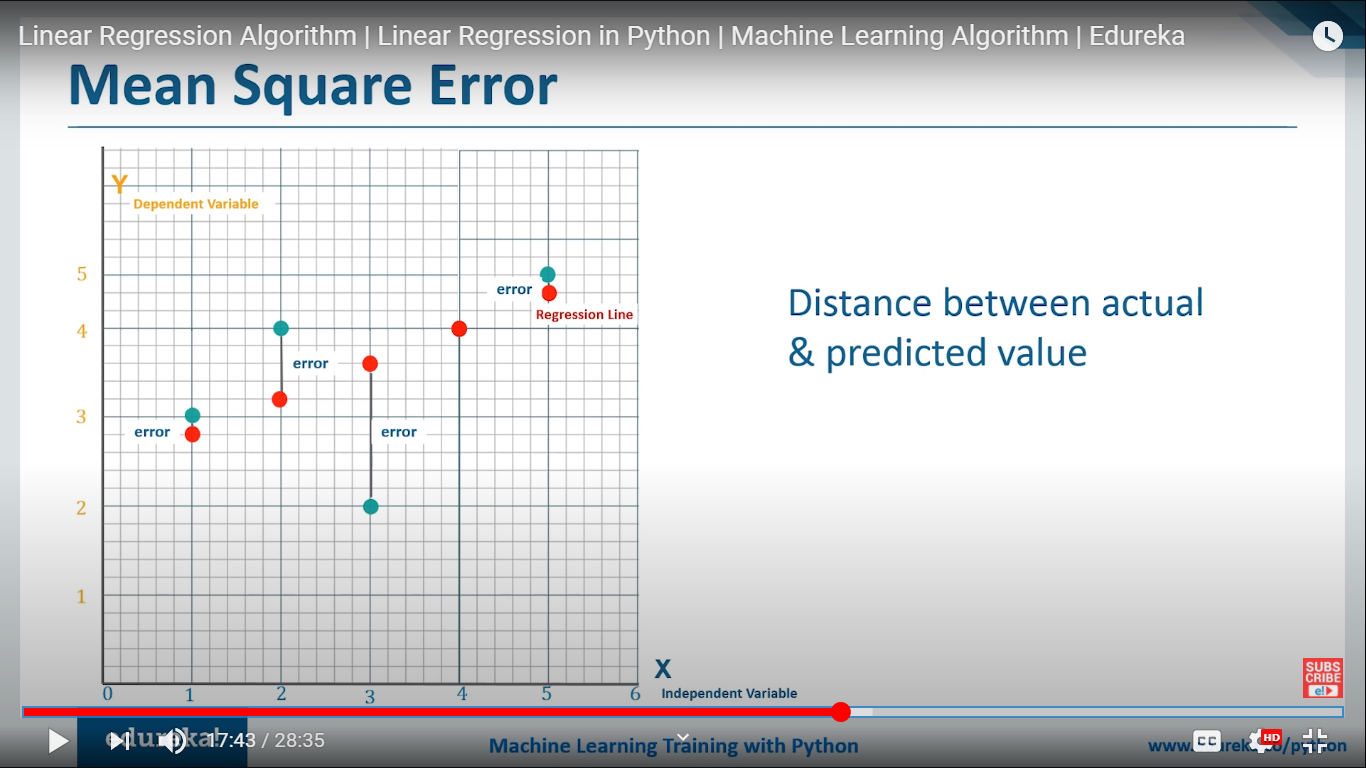
From the picture below the slpe of line from the point is calculated and find to best fit for the given data.



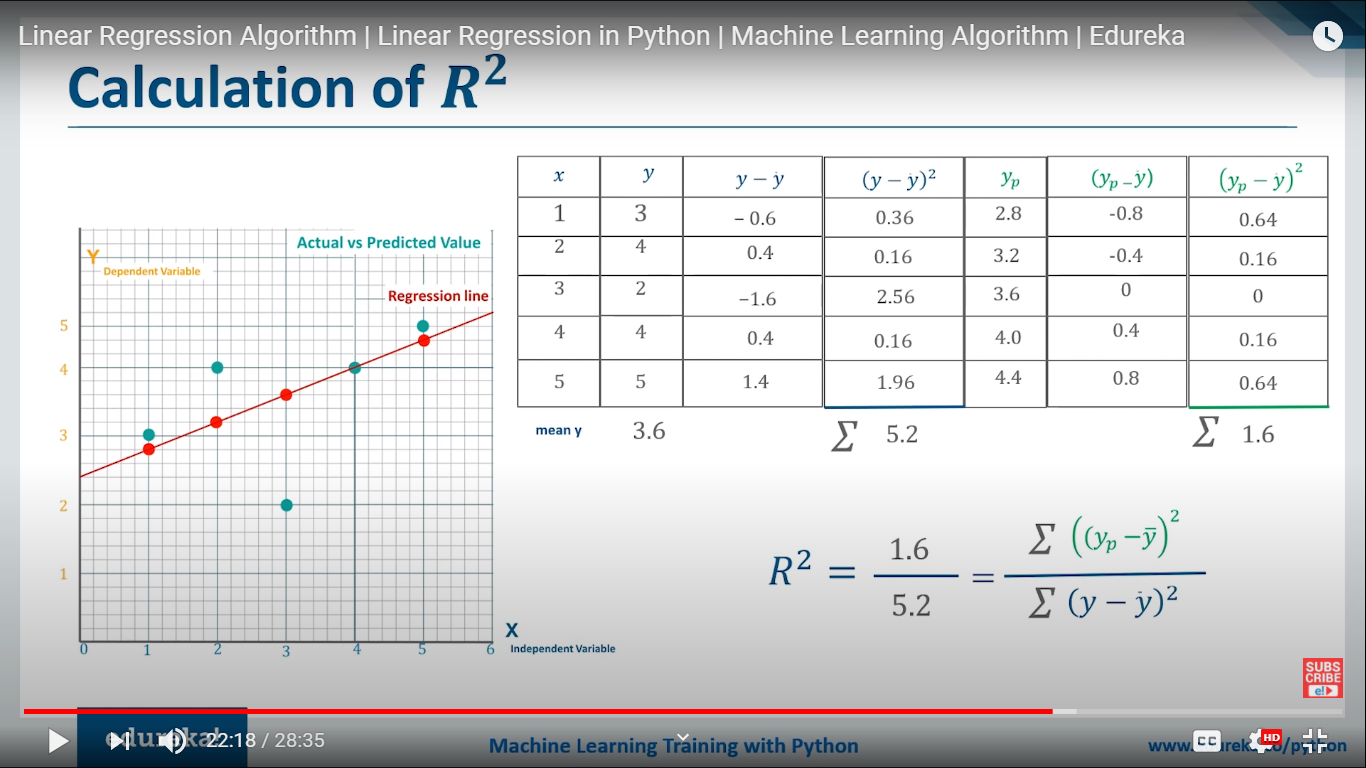
Now we have slope m and the constant c and our regression line is ready. Now we can find the other values of the y for the different values of the x.

**How can we measure the exactness of the line/How good is the line for the predictions?**

Mean Squre error is the measure to tell hoe good a line is suitable

\

Here we measure the error as the distance between the actual value of y at pont x to the value of y on the line at same point x for the all existing values of the x and is measured interms of Rsquare as a whole measure for the exactness of line

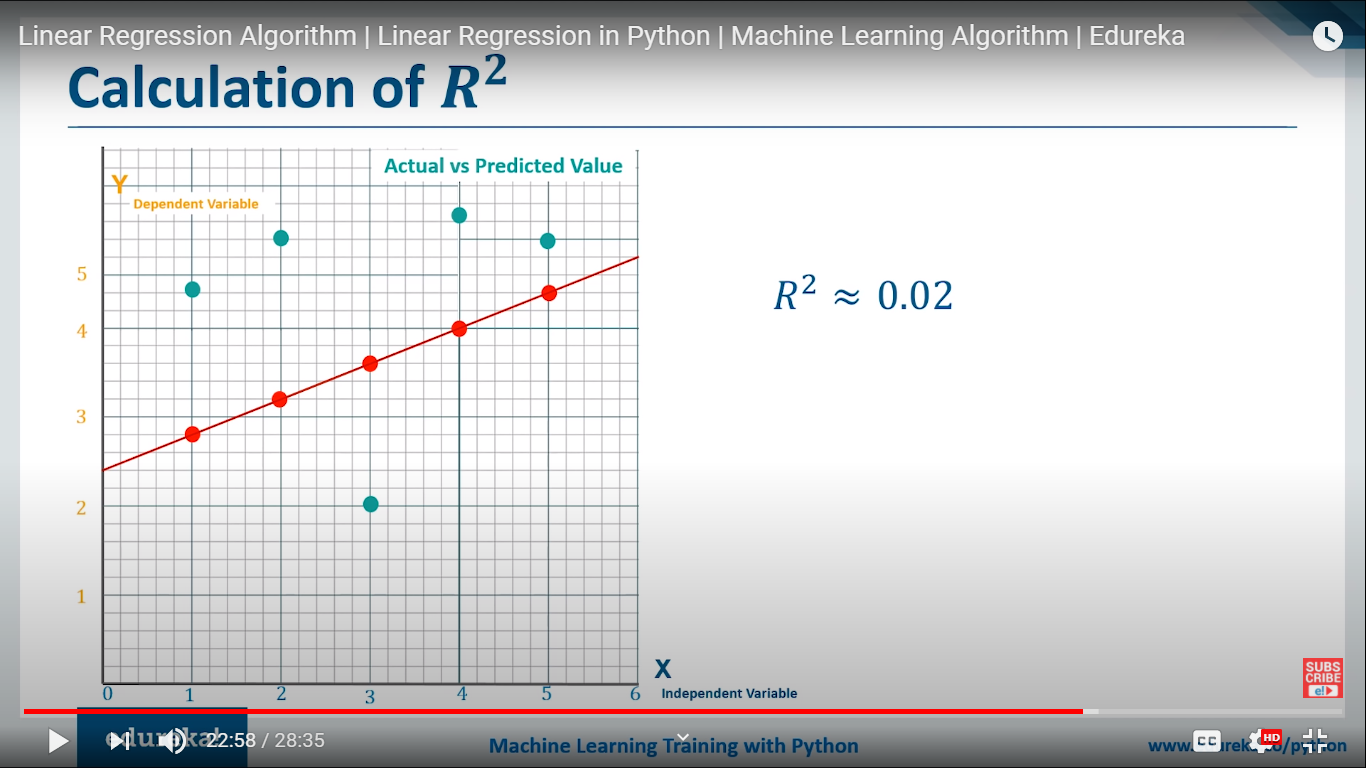


Value of R-square

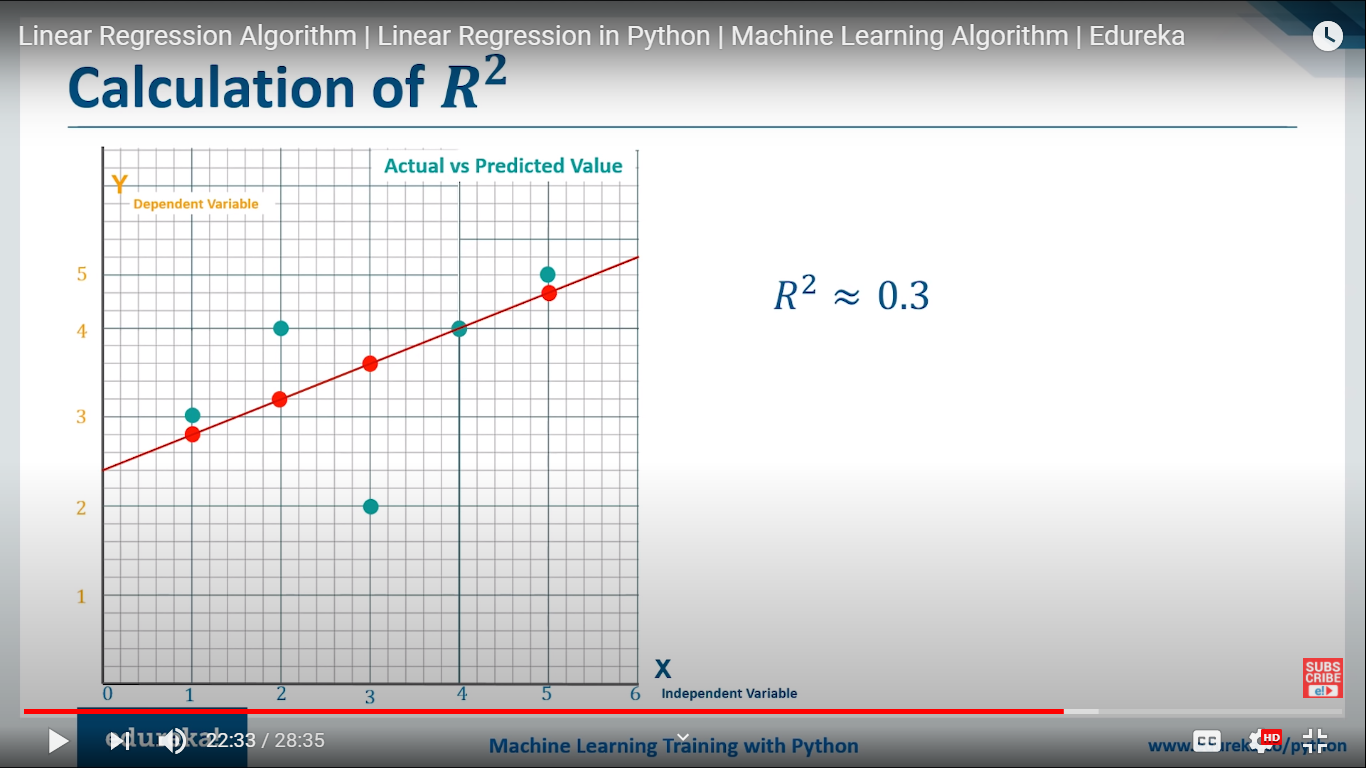
The value of rsquare lies in between 0 and 1. Ommonly the value close to 1 is treated as the best one

Let see some example graphs:

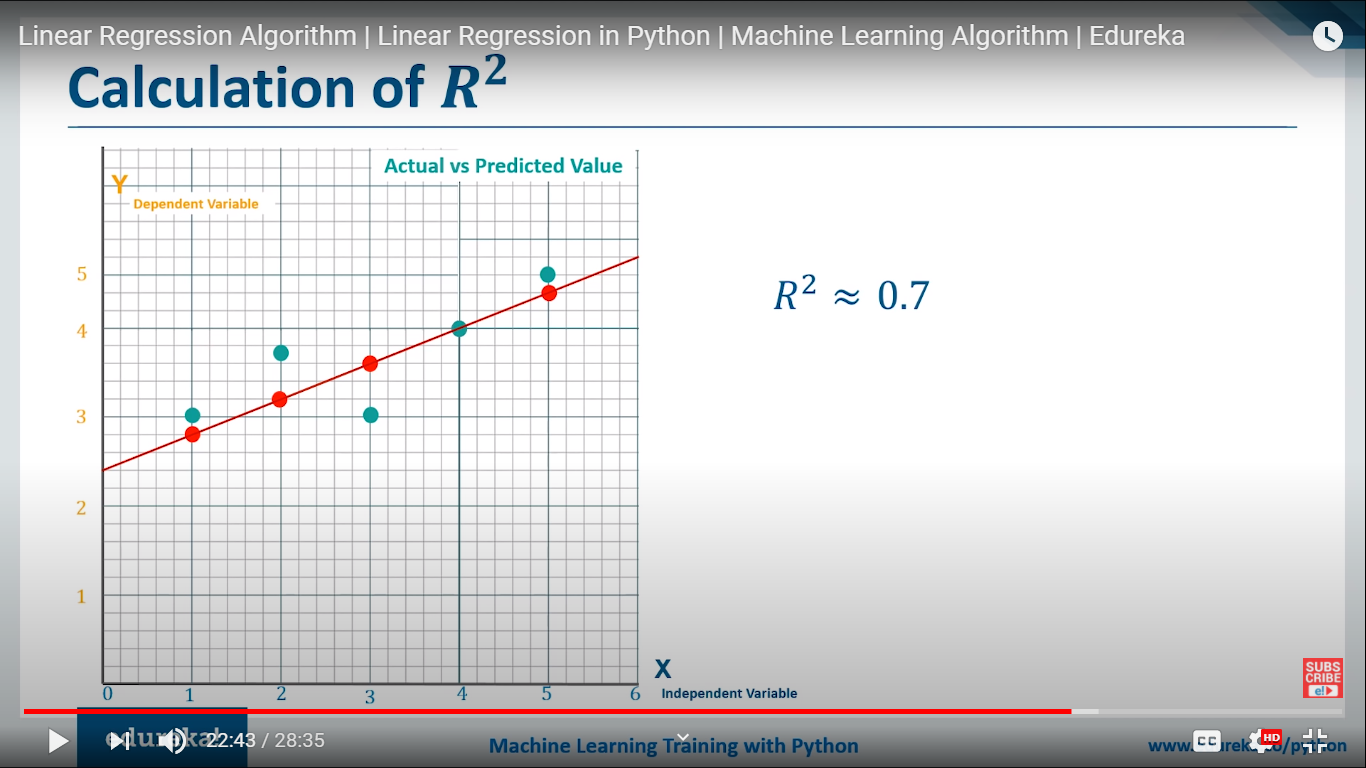
FOR R-SQUARE=0.02



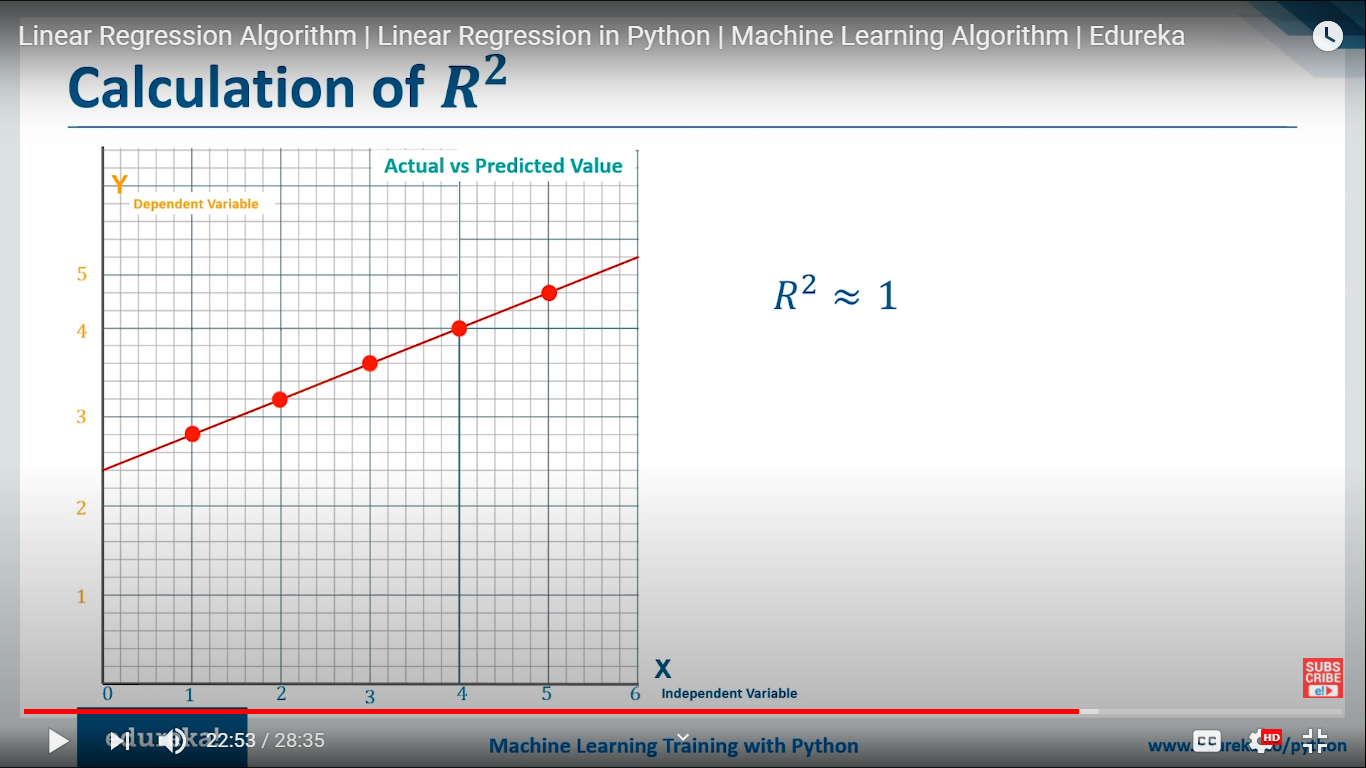
FOR R-SQUARE=0.3



FOR R-SQUARE=0.7



FOR R-SQUARE=1



So for the line for which the r squre value is max is taken

SUMMARY:

In short in linear regression what we do is we find a line for the data we have ,that line doesn’t include the all the points and the line which fits the max of the points in line is taken as best regression.So we take the center value of the x table and make it a base for the line , as many number of lines can be drawn through it,we see only one line is best and that line is calculatd by using some calculations an d slope and constant of that line is find.After finding the slope and conmstant we measure how much rsqaure fot the line predicte. If the vlue of r-square is find satisfied thenits okay or else we adopt some more other measures of finding the slope and line equation.