Polynomial Regression

1. Linear Regression

* The model is linear in both the parameter and the features.

**Y = β₀ + β₁X₁ + β₂X₂ + ... + βₙXₙ + ε**

* Y is the dependent variable you're trying to predict.
* X₁ to Xₙ are the independent variables that influence Y.
* β₀ is the y-intercept, which represents the predicted value of Y when all the independent variables are zero.
* β₁ to βₙ are the regression coefficients, which represent the change in Y for a one-unit increase in the corresponding independent variable, holding all other variables constant.
* ε is the error term, which accounts for the difference between the predicted value and the actual value of Y.

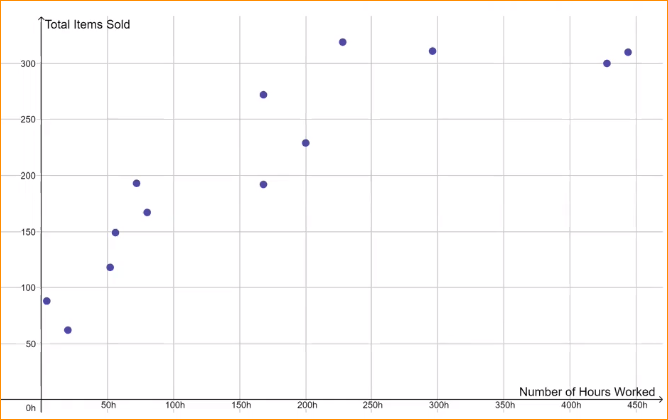
in here we assume the model will be linear, as well as the parameters are the linear.

**This is very strong assumption**

There can be non-linear variables between dependent and independent variable.

In here we are looking to behaviors that are non-linear features. (X and Y )

Example –



Here I have the data of- Relationship between **total items sold** by the salesperson and the **number of hours worked** in a shop.

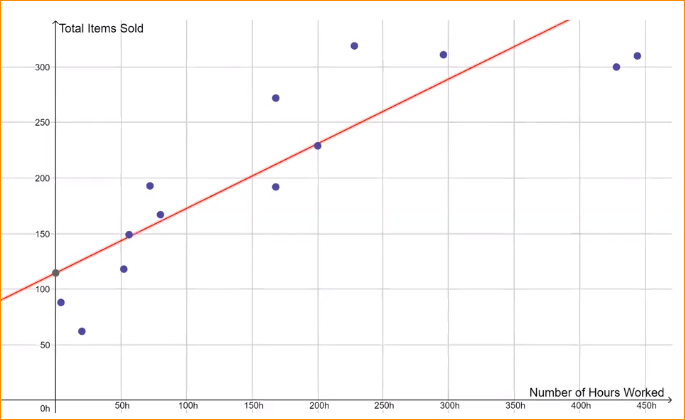
Here X axis – number of hours

Y axis – Total items Sold

Here there is a patterns.. if number of hours increases , number of item sold are increased.

But if the Employee Works more hours the efficiency goes down. But if we consider this as a linear. It will may be wrong.

If we try to plot that linear regression plot –



If we get coefficients,

Y = 0.5823X + 114.5

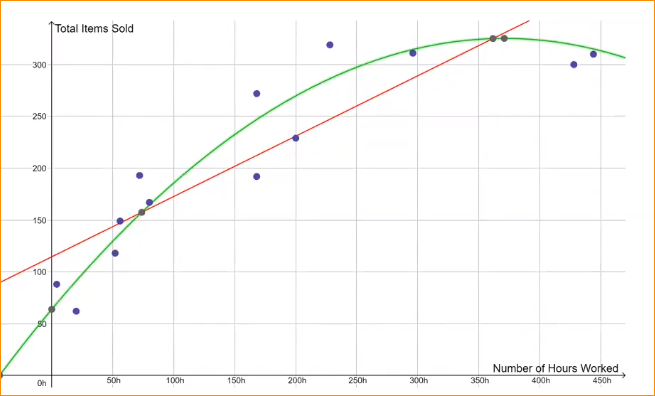
In here R square is = 0.8019 – in here it looks like close to 1

You feel this is a good model.

**But this is not a best Solution model**

Why?

* Nonlinear Regression

hee

here fitted this with non linear equation –



Consider of

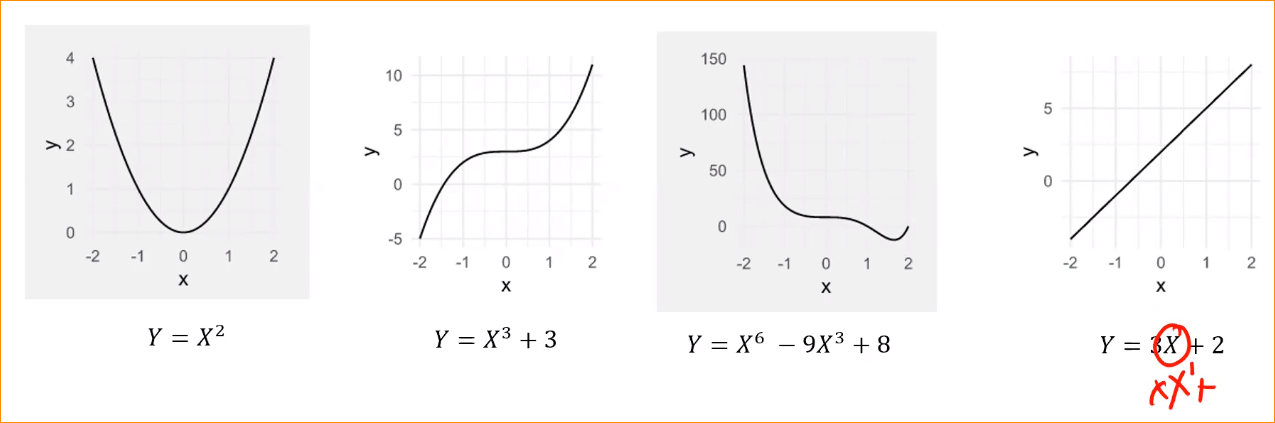


So this is closer to 1.

So the Green color line is the best.

What is the Polynomial Function?

Here we have a variable (X) – it will have addition, multiplication or subtraction and the power of variable, its may be negative or positive number.



Last one – you have X power of 1

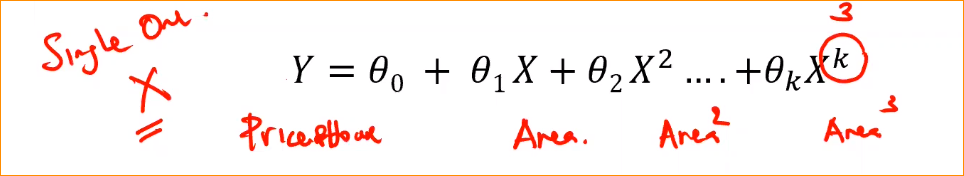
Has multiplication – 3 x X

Addition - + 2

In here we represent the data using a polynomial Function. Call **polynomial Regression**

**Polynomial Regression**

* Relationship between the independent variable x and the dependent variable y is modelled as a **k th degree polynomial in x.**
* A polynomial linear regression model of degree  has the form –



Example – here we get a dependent variable (Y) as a Price of house and just we want to find the **k** value.

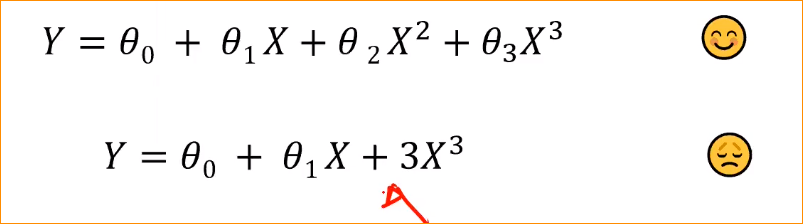
So if the k value equals to 3 equation is



Typically, statisticians aim for low order polynomial models (less than 5) – 5 or 6 too higher degree.

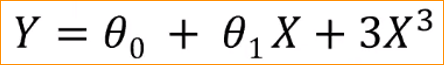
So the polynomial model of degree k should include all terms of lower order. Example 3 degree,

In the equation should include X^2, X



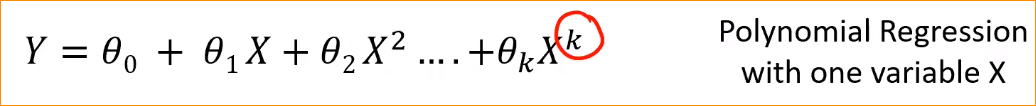
Some time model , will understand the theta 2 is not important that mean 0.

So we can put this equation



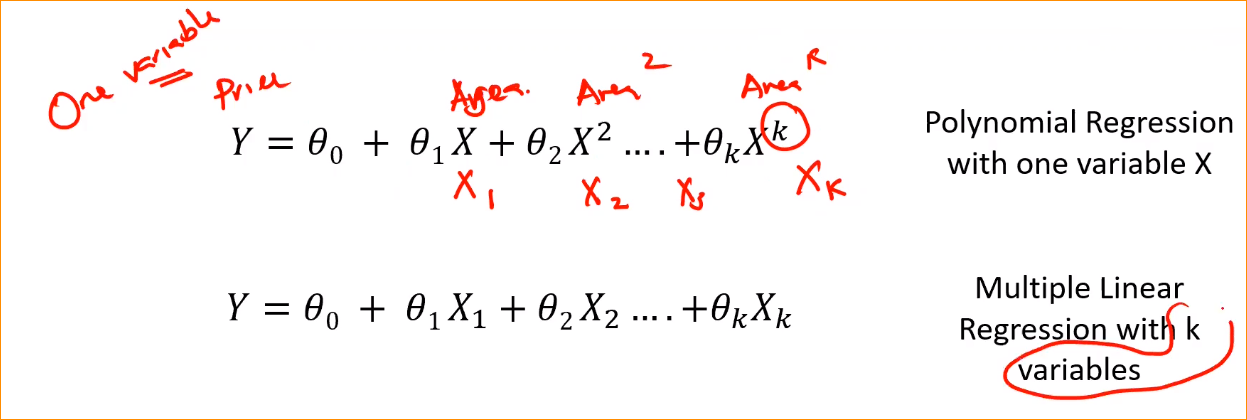
**Polynomial regression and Multiple Linear Regression**

Equation for one variable –



Multiple variable -

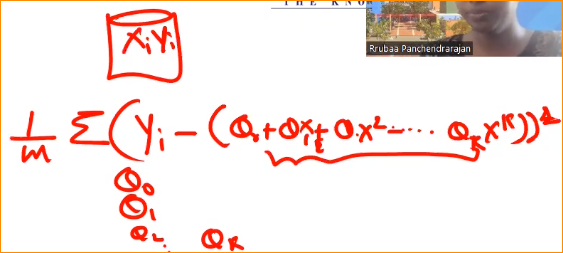
So in here we can consider this X^2, X^3 …. X^k equal to another(new- Example X^2 = X2) variables.. that mean Multiple Regression with k variables.



In here can recognize we have k number of variables. But in here the internal outcome is not linear.

**Estimation the parameters –**

Error function is still convex.



So get Coefficient value, should get each derivatives(theta1, theta2 .. theta k ) and should equal to 0.

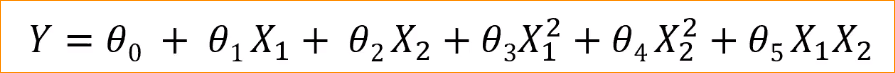
We can use the same Least Square Error method to estimate the parameters.

* Take the partial derivatives with respect to each parameter.
* Equate it to 0
* Solve the equation

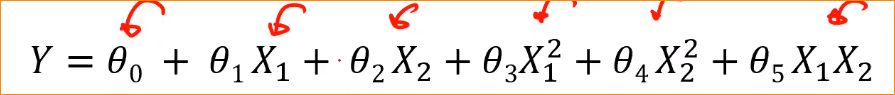
**More than One number of Variable**

When we have two variables(X1 and X2) –

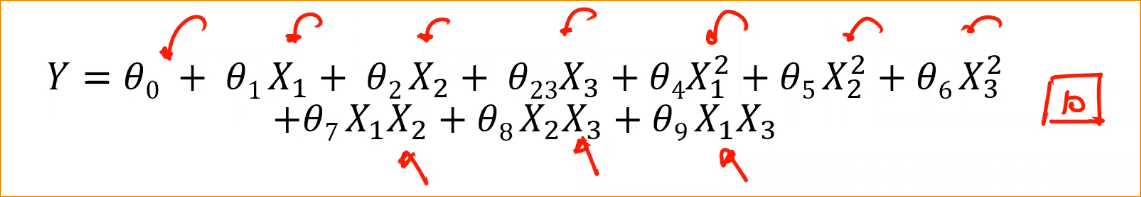
A second order polynomial is more used in practice and its model is specified by



In we have 6 parameters to learn



If we have 3 variables – here we have 10 parameters to learn



Here if we add one variable – the model parameters to learn increased by 4(Complex.. that model has to learn more parameters).

In here we got 2 degree, what happen is we have 3 degree.. its goes too complex.

**Usually best practice is this restrict to 2 degree , if you dealing with more variables in polynomial regression**