Given some data X,Y you need to learn the parameters required to learn the relation between both variables. For example as shown in the equation below you need to learn Beta0 (Bais) and Beta1.

$$\hat{y}(x) := \hat{\beta}_0 + \hat{\beta}_1 x$$

Given the above equation you can then calculate **vector of** Betas ($\hat{\beta}$)using;

$$X^T X \hat{\beta} = X^T y$$

Last step to obtain Betas is to divide by (X transpose * X)or basically multiplying by (X transpose * X) inverse.

So in the end Beta can be calculated using;

Beta = (X^T . X)^(-1) .(X^T . Y)