Let’s start by defining a few things

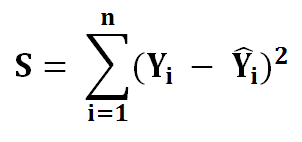
1) Given n inputs and outputs.



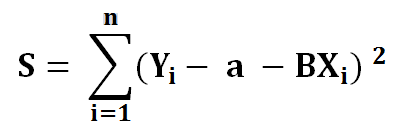
2) We define the line of best fit as:



3) Now we need to minimize the error function we named S



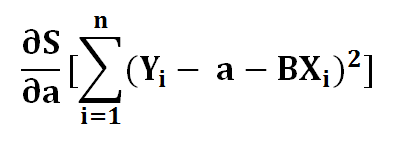
4) Put the value of equation 2 into equation 3.



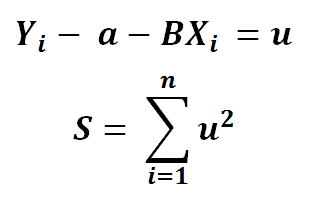
To minimize our error function, S, we must find where the first derivative of S is equal to 0 concerning a and b. The closer a and b are to 0, the less total error for each point is. Let’s find the partial derivative of a first.

**Finding a :**

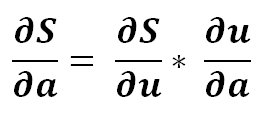
1 ) Find the derivative of S concerning a.



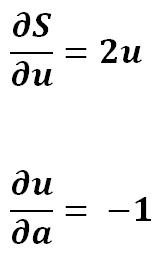
2 ) Using the chain rule, let’s say



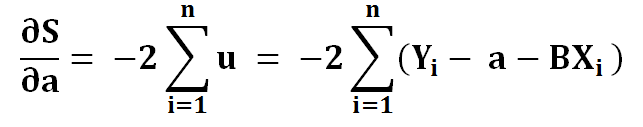
3) Using partial derivative



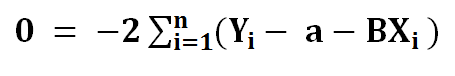
4) Expanding



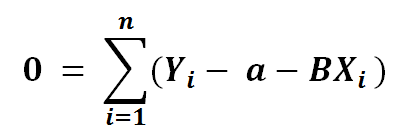
5) Simplifying



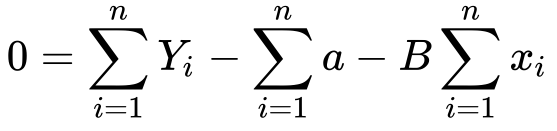
6) To find extreme values, we put it to zero



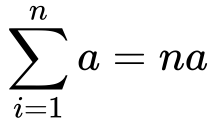
7) Dividing the left side with -2



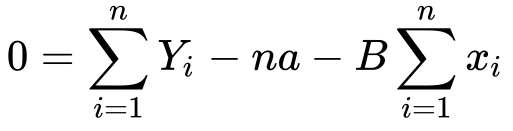
8) Now let’s break the summation in 3 parts



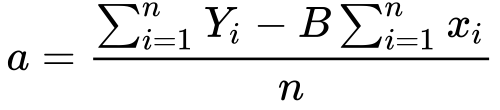
9) Now the summation of a is



10) Substituting it back in the equation



11) Now we need to solve for a



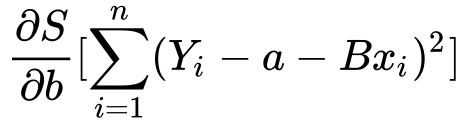
12) The summation of Y and x divided by n, is simply it’s mean



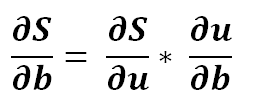
We’ve minimized the cost function concerning x. Now let’s find the last part which S concerning b.

**Finding B :**

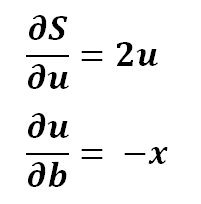
1 ) Same as we have done with a



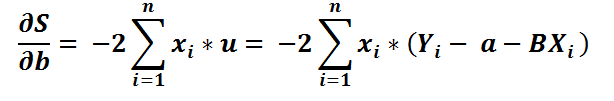
2) Finding the partial derivative



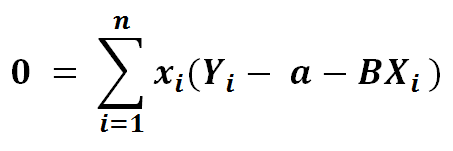
3) Expanding it a bit



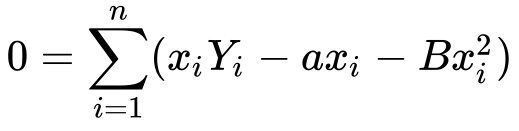
4) Putting it back in the equation



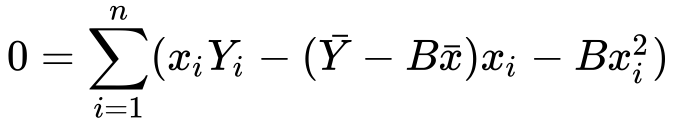
5) Let’s divide by -2 both sides



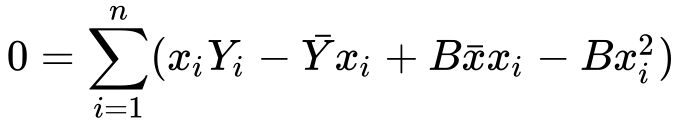
6) Let’s distribute x for ease of viewing



7) Substituting the value of a

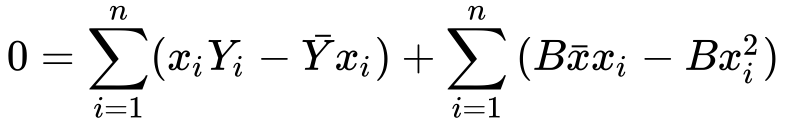


8) Let’s distribute the minus sign and x

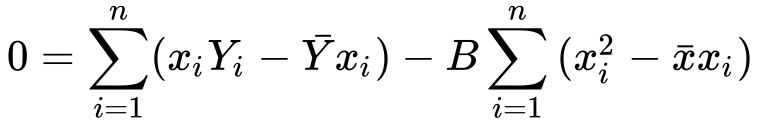


Well, you don’t like it? Let’s split up the sum into two sums

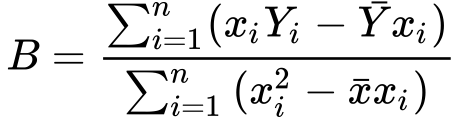
9) Splitting the sum



10) Simplifying



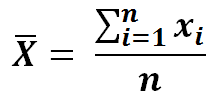
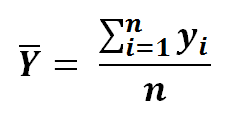
11) Finding B from it



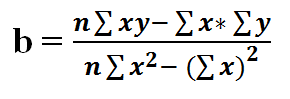
Great. We did it. We have isolated a and b in the form of x and y. It wasn’t that hard, was it?

I still have some energy and want to explore it a bit!

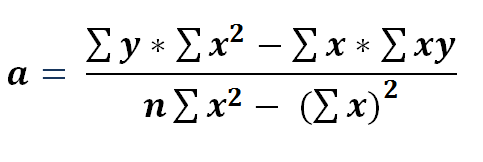
12 ) Simplifying the formula



13) Multiplying numerator and denominator by n in equation 11

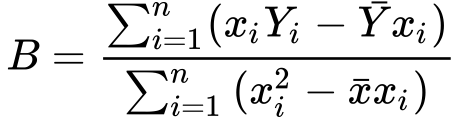


14) Now if we simplify the value of a using equation 13 we get



#### Summary 🙂

If you have a dataset with one independent variable, then you can find the line that best fits by calculating B.



Then substituting B into a



And finally substituting B and a into the line of best fit



**Moving Onwards,**

In the next article, we’ll see how we can implement simple [linear regression](https://towardsai.net/p/machine-learning/calculating-simple-linear-regression-and-linear-best-fit-an-in-depth-tutorial-with-math-and-python-804a0cb23660) from scratch (without sklearn) in Python.

And please let me know whether you liked this article or not! I bet you liked it.