**Ready to find the line of best fit?**

Let’s start by defining a few things.

1. Given N inputs and outputs…

Shape

Description automatically generated with medium confidence

2. We define the line of best fit line as…

Shape

Description automatically generated with medium confidence

3. Such that the best fit line looks to minimize the cost function we named S…

Shape

Description automatically generated with medium confidence

For our reference, we will input the line of best fit into our cost function distributing the subtraction, resulting in…

Shape

Description automatically generated with medium confidence

To minimize our cost function, S, we must find where the first derivative of S is equal to 0 with respect to**a** and **B.** The closer **a** and **B** are to 0, the less the total error for each point is. Let’s start with the [partial derivative](https://mathinsight.org/partial_derivative_introduction) of **a** first.

**Finding a**

Shape

Description automatically generated with medium confidence

Use the chain rule by starting with the exponent and then the equation between the parentheses. Notice, taking the derivative of the equation between the parentheses simplifies it to **-1**.

Shape

Description automatically generated with medium confidence

Let’s pull out the -2 from the summation and divide both equations by -2.

Shape

Description automatically generated with medium confidence

Let’s do something semi clever. Let’s break the summation into 3 parts and pull the constant B outside the summation

Shape

Description automatically generated with medium confidence

We notice that summation of a to n is simply…

Shape

Description automatically generated with medium confidence

Substituting this back in and rearranging B, give us…

Shape

Description automatically generated with medium confidence

We’re almost there! The last thing we need to do is solve for **a,**so we add **na** to both sides and divide by **n**.

Shape

Description automatically generated with medium confidence

Check this out! The two summations of Y and x divided by the number of observations is simply the mean :). So after all that work, minimizing the cost function of **S**with respect to **a** is simply…

Shape

Description automatically generated with medium confidence

**Finding B**

We’ve minimized the cost function of S with respect to a. Let’s find the last part which is S with respect to b.

Shape

Description automatically generated with medium confidence

Shape

Description automatically generated with medium confidence

Let’s strip out the -2 and divide it by both sides.

Shape

Description automatically generated with medium confidence

Let’s distribute the x for ease of viewing.

Shape

Description automatically generated with medium confidence

Now, let’s do something creative. Remember we already figured out a? Let’s substitute **a**(derived formula below) into the partial derivative of S with respect to B above. We’re doing this so we have a function of a and B in terms of only x and Y.

Shape

Description automatically generated with medium confidence

Shape

Description automatically generated with medium confidence

Let’s distribute the minus sign and x

Shape

Description automatically generated with medium confidence

This looks messy but algebra kicks ass in this front. Let’s split up the sum into two sums.

Shape

Description automatically generated with medium confidence

Let’s take **-B** out of the summation on the right so we can isolate the variable and rearrange terms remaining terms in the summation.

Shape

Description automatically generated with medium confidence

Isolating B by subtracting the first summation and dividing by the second summation.

Shape

Description automatically generated with medium confidence

There we have it! We have now isolated **B** and **a** in terms of x and Y. You might think to yourself, wow….this looks like an awful formula! Don’t fret, there are simpler versions by manipulating the formula further — you can check those out [here](https://www.andrews.edu/~calkins/math/edrm611/edrm06.htm).

**Summing it up :)**

If you have a dataset with one independent variable, you can find the line of best fit by calculating **B**

Shape

Description automatically generated with medium confidence

Then substituting **B** into **a**

Shape

Description automatically generated with medium confidence

and finally, substituting **B** and **a** into the line of best fit!

Shape

Description automatically generated with medium confidence