**📊 The Math Behind Linear Regression**

**💡What Is Linear Regression?**

Linear regression is a way to **find a straight line** that best fits a set of data points.

Imagine you're trying to predict someone's **height based on their age**. You plot a few data points, and linear regression finds the best line that goes through those points to show the **trend**.

**📐 The Equation of a Line**

We use this equation:

y=mx+b

Where:

* **x** is the input (e.g., age)
* **y** is the output (e.g., height)
* **m** is the **slope** (how steep the line is)
* **b** is the **y-intercept** (where the line crosses the y-axis)

Let me show you visually 👇

A diagram of a graph

AI-generated content may be incorrect.

**🔍 How Do We Find the Best Line?**

We want the line that’s **closest to all the points**. That’s where the idea of **minimizing error** comes in.

We define **error** as the difference between the actual y-value and the predicted y-value:

Error=yactual−ypredicted

To measure overall error, we use something called **Mean Squared Error (MSE)**:

MSE=1n∑i=1n(yi−(mxi+b))2

We **adjust m and b** so this number becomes as **small as possible**.

This is where a method called **gradient descent** comes in, but I’ll save that for another post 😉

