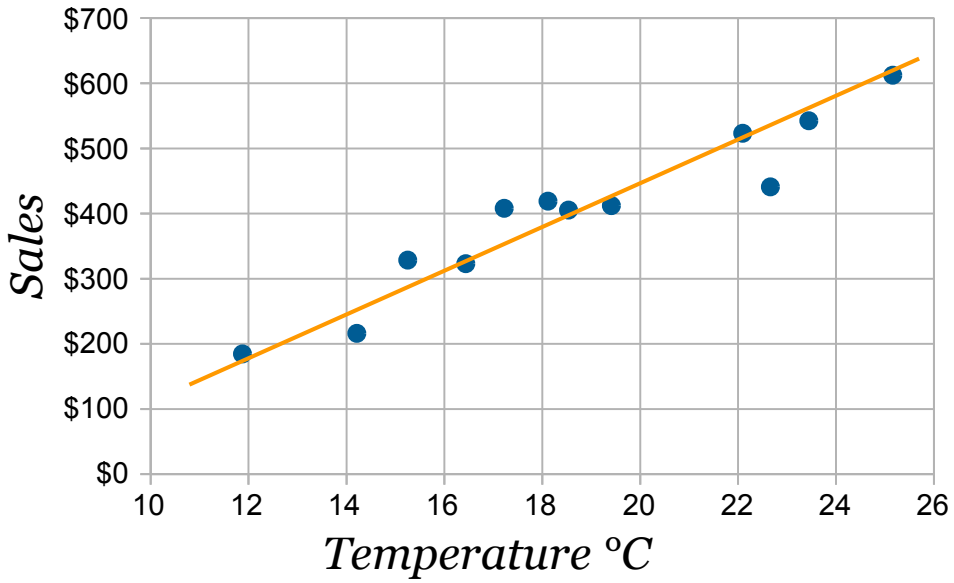
Least Squares Regression

https://www.mathsisfun.com/data/least-squares-regression.html

Line of Best Fit

Imagine you have some data values, and want to have a **line** that best fits them like this:



We can place the line "by eye": try to have the line as close as possible to all points, and a similar number of points above and below the line.

But for better accuracy, we can calculate the line using **Least Squares Regression**.

The Line

Our aim is to calculate the values **m** (slope) and **b** (y-intercept) in the equation of a line :

$$y = mx + b$$

Where

- **y** = the y variable
- **x** = the x variable
- **m** = Slope or Gradient (how steep the line is)
- **b** = the Y Intercept (where the line crosses the Y axis)

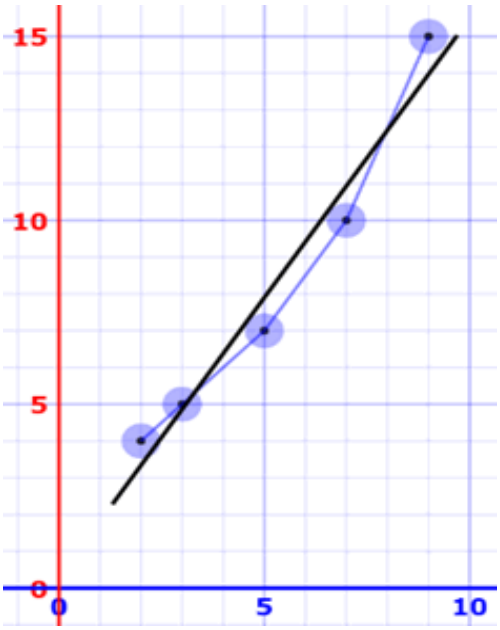
Example

Sam found how many **hours of sunshine** vs how many **ice creams** were sold at the shop from Monday to Friday.

"x" Hours of Sunshine	"y" Ice Creams Sold
2	4
3	5
5	7
7	10
9	15

Sam finds the best **m** (slope) and **b** (y-intercept) that fits these data.

$$y = mx + b$$



Nice fit!

Sam hears the weather forecast which says "we expect 8 hours of sun tomorrow," so he uses the above equation to estimate that he will sell

$$y = 1.518 * 8 + 0.305 = \mathbf{12.45\ Ice\ Creams}$$

Sam makes fresh waffle cone mixture for 14 ice creams just in case. Yum.