Types of Distributions - Normal

We'll cover the following

4. Normal Distribution (Gaussian)

A normal distribution, the bell curve or Gaussian Distribution, is a distribution that represents the behavior in most situations. For example, exams scores are typically a bell curve where most students get the average score, C, a small number of students scores a B or a D, and an even smaller number scores an F or an A. This results in a distribution that looks like a bell:

The bell curve is symmetrical, half of the data will fall to the left of the mean value and half will fall to the right of it.

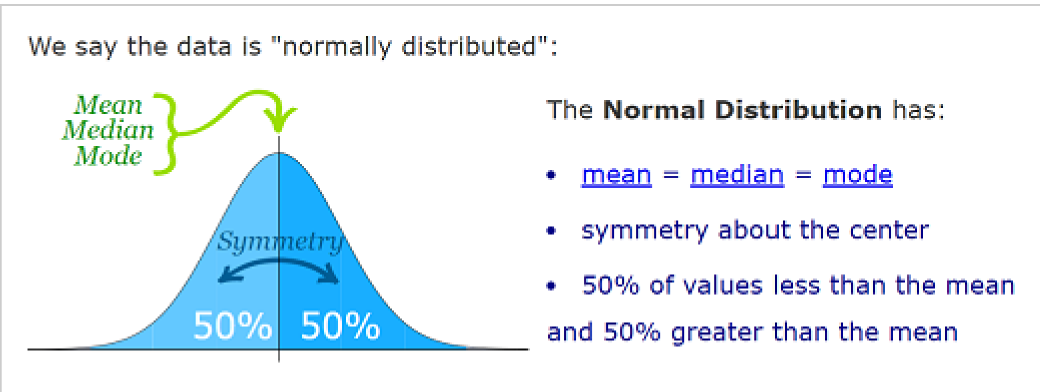

Image Credits: https://www.mathsisfun.com

Many things follow this type of spread, this is why this is used most often:

Many things follow this type of spread, this is why this is used most often,
you'd seen anywhere from businesses to academia to government. Here are

some examples to give an idea of the variety covered by normal distributions:

- Heights of people
- Blood pressure
- IQ scores
- Salaries
- Size of objects produced by machines

Some facts to remember about what percentage of our data falls within a certain number of standard deviations from the mean:

- **68%** of values are within 1 standard deviation of the mean
- **95%** of values are within 2 standard deviations of the mean
- **99.7%** of values are within 3 standard deviations of the mean

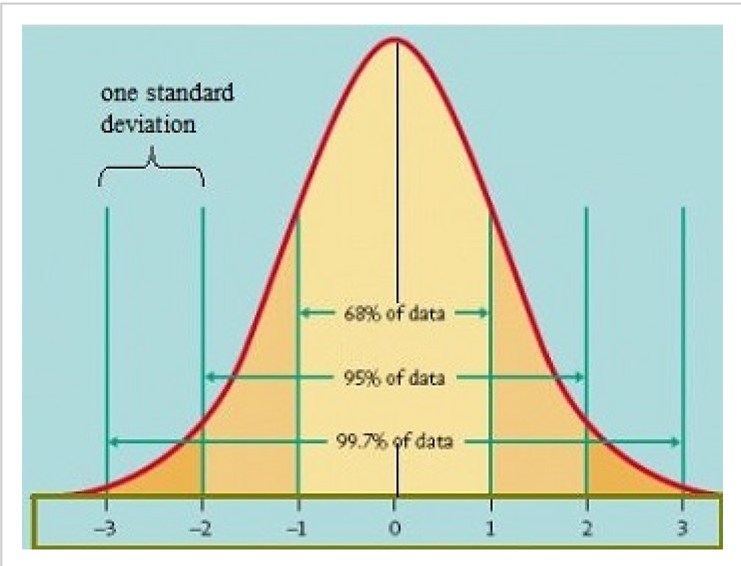

Image Credits: University of Virginia.

Why is it good to know standard deviations from the mean?

Because we can say that any value is:

- likely to be within 1 standard deviation (68 out of 100 should be)
- very likely to be within 2 standard deviations (95 out of 100 should be)
- almost certainly within 3 standard deviations (97 out of 100 should be)

The number of standard deviations from the mean is also called the **standard score** or **z-score**. Z-scores are a way to compare results from a test to a

score or Z score. Z scores are a way to compare results from a test to a "normal" population.

Say we are looking at a survey about heights. A z-score can tell us where a person's height is compared to the average population's mean height. A score of zero tells us that the value exactly matches the average, while a score of +3 tells us that the value is much higher than average.

The mean and variance of a random variable, *X*, which is said to be normally distributed is given by:

$$Mean = E(X) = \mu$$
$$Variance = V(X) = \sigma^2$$

where μ is the mean value and σ the standard deviation.

The z-score is then given by:

$$z = \frac{x - \mu}{\sigma}$$

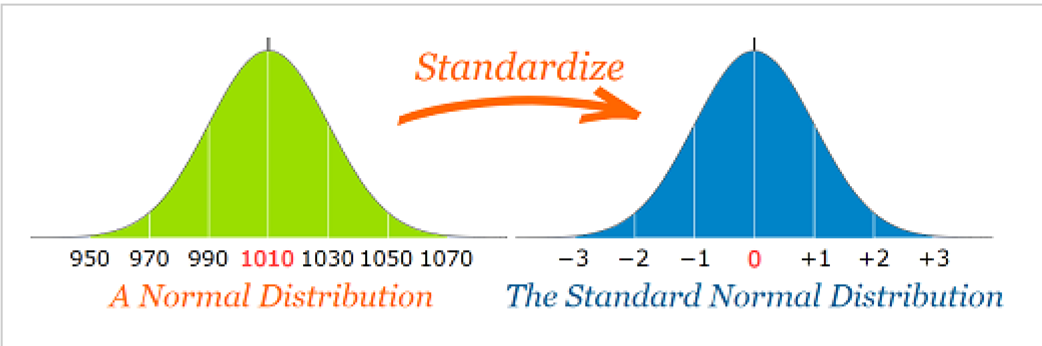

Image Credits: https://www.mathsisfun.com/data/standard-normal-distribution.html

All normal distributions do not necessarily have the same means and standard deviations. A normal distribution with a mean of 0 and a standard deviation of 1 is called a **standard normal distribution**. If all the values in a distribution are transformed to Z scores, then the distribution will have a mean of 0 and a standard deviation of 1. This process of transforming a distribution to one with a mean of 0 and a standard deviation of 1 is called **standardizing the distribution**. Standardizing can help us make decisions about our data more easily.

> ***Note***: *This is the most important continuous random distribution. So, make sure you understand it well before moving on to the next lesson.*