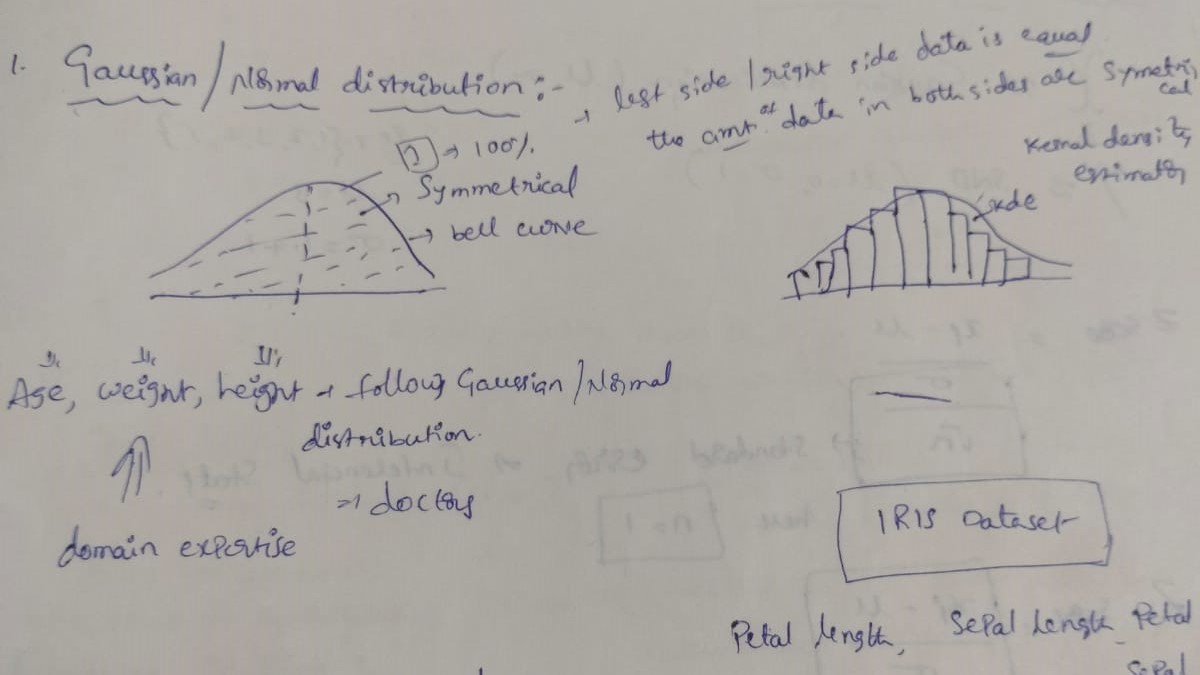
Today we will learn some information about types of distributions.

* Normal Distribution
* Standard Normal Distribution
* Normalization

**Normal / Gaussian distribution:**

The amounts of data in both sides are **symmetrical.** This distribution is called normal/Gaussian distribution.



**Ex:** Age, weight, height --- follows the Gaussian distribution.

**IRIS** dataset also follows the Gaussian distribution (Petal length, Petal width, Sepal Length and Sepal width)

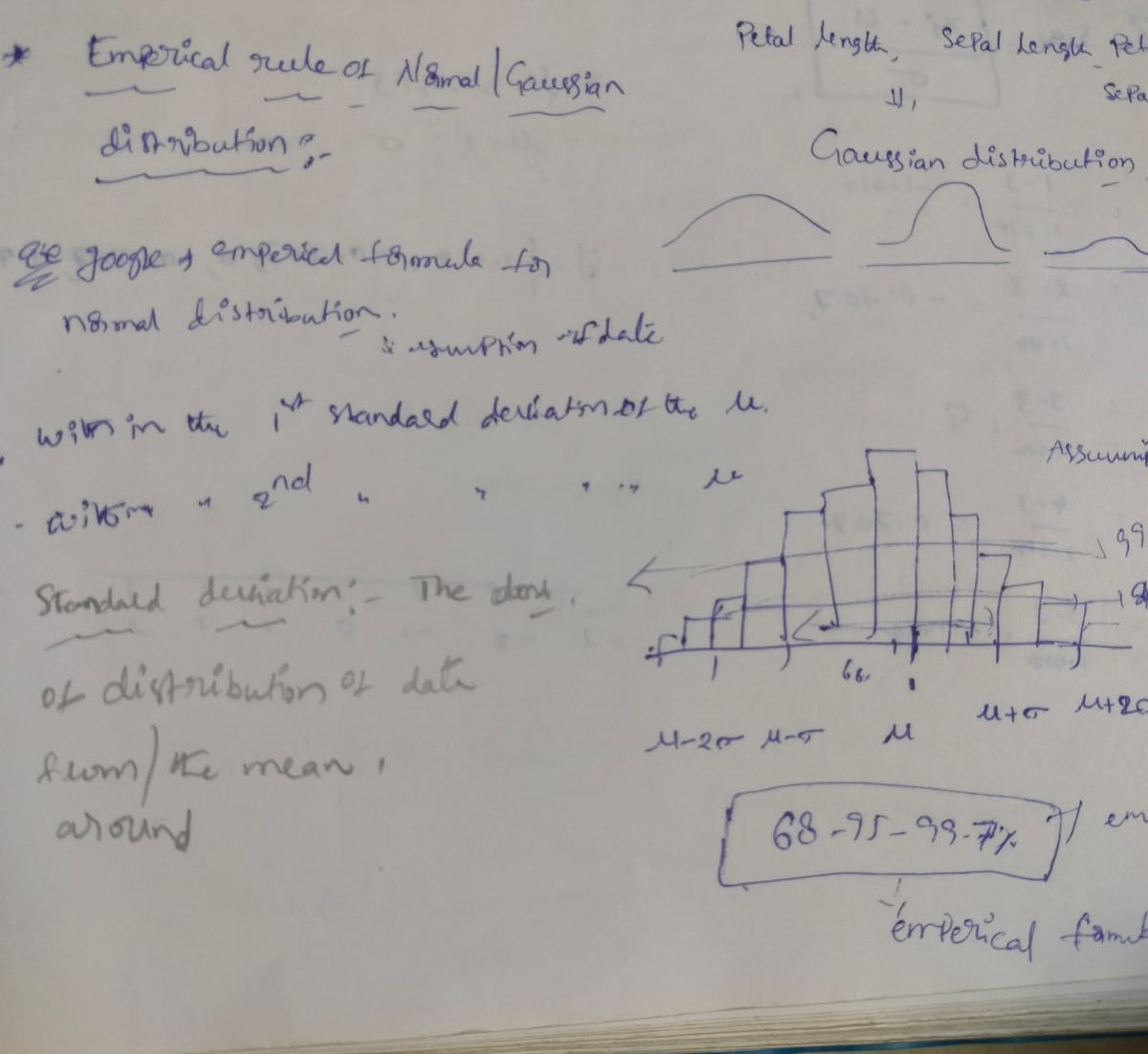
**Empirical rule of Normal distribution: 68 – 95 – 99.7%**

* Within the **1st standard deviation: 68%**

**(The amount of distribution of data from or around the mean – is called standard deviation.)**

* Within the **2nd standard deviation: 95%**
* Within the **3rd standard deviation: 99.7%**

**Q-Q plot:** To determine whether the distribution is Gaussian /Normal distribution

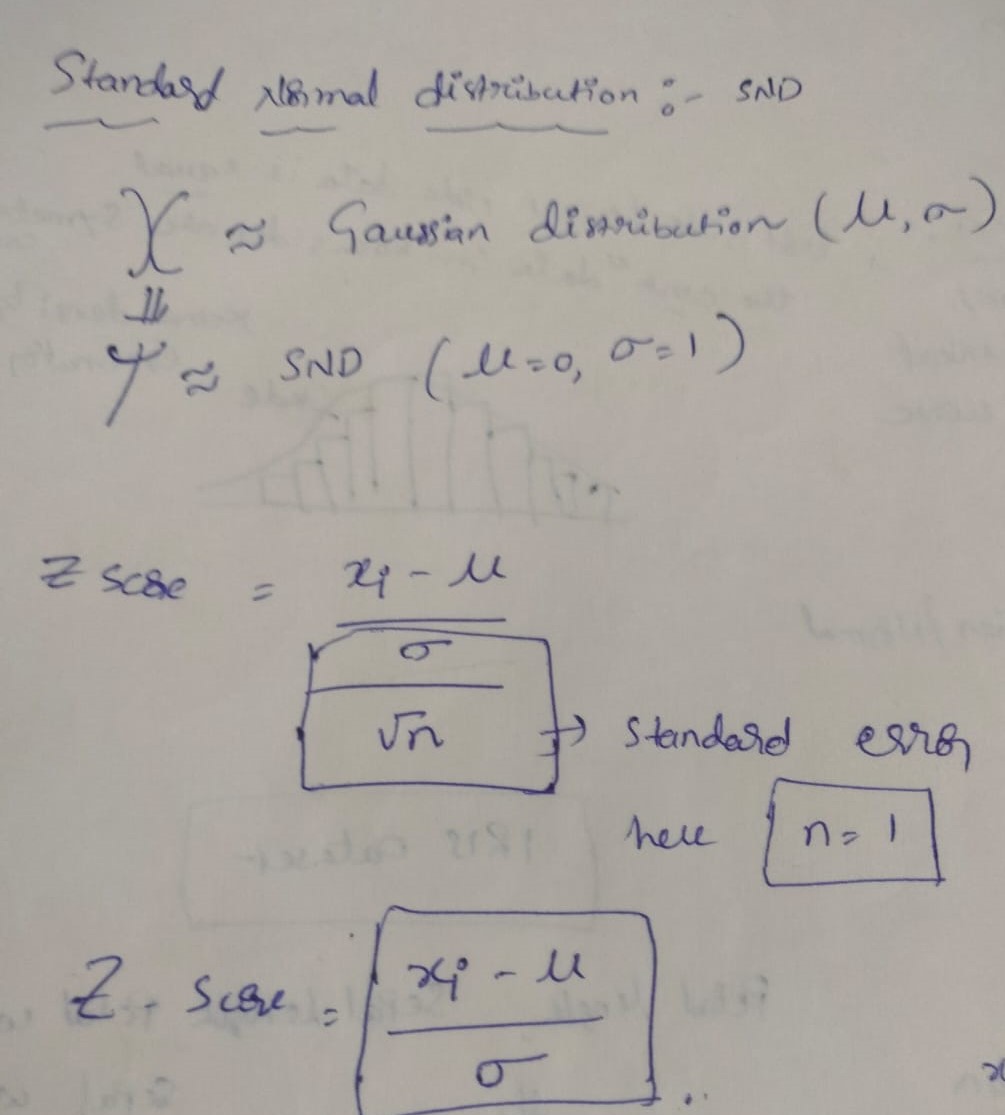
****

**Standard Normal Distribution (SND):**

**Why standardization required?**

* To bring the features in the same scale then the calculations will be easy. Where **µ=0, σ=1**

Using Z-Score, we can calculate standardization.



**Normalization:**

* Normalizing the data or values between min and max or **lower scalar -🡪 higher scalar**

**Min max scalar formula:**

**X scaled = (x-xmin)/(xmax – xmin)**