**NORMALIZATION**

Indeed, it doesn´t exist strict rules on how normalize attributes. The simplest methodology to do it consists in the transformation of data to center them, eliminating the average value in each attribute, thereafter I scaled them, dividing the attributes by their standard deviation. Now, for an attribute:

This procedure is often the first stage, when I apply traditional statistical models. It makes sure that the observations have an average of 0 and a standard deviation of 1. Not always, like it happens to many things in Deep Learning, will give good results all the time, therefore, there are some alternatives: Here, I will show two popular options:

Where is the sum of the squares for , and and are the mean and the standard deviation of . Another different methodology that I can consider, is to scale the attributes so that they are between a given minimum and maximum value, frequently between zero and one, so the maximum absolute value in each attribute is scaled to the size of the unit:

Which one should be used? Well, in my opinion, the key of success in data science is to execute experimentations.