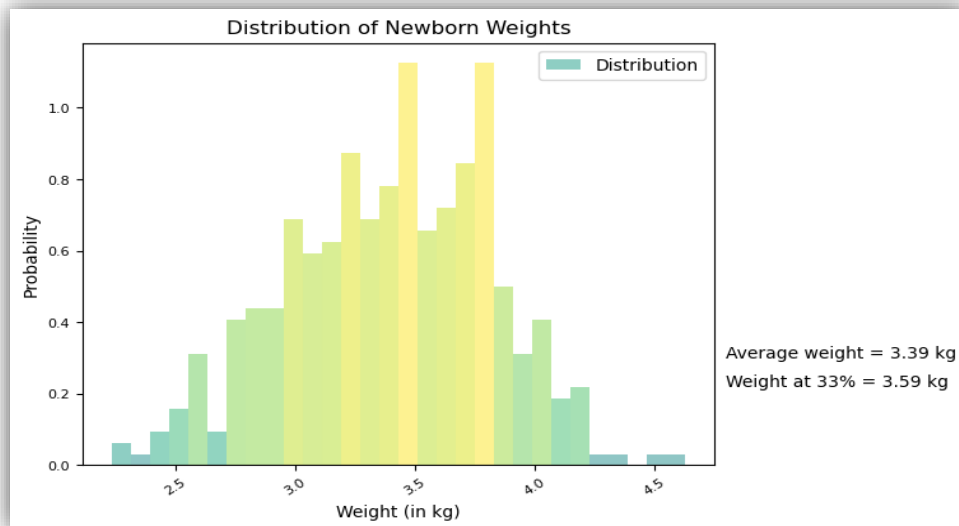# Fundamentals of Data Science - Coding Project

by Mukesh Avudaiappan

The given data is a list of weights of newborn babies in kgs. We have used this data to create a histogram to visualize the distribution of weights.

The distribution we get from the histogram appears to be roughly normal, with a mean weight around 3.39 kg. The majority of the weights seem to be concentrated around this mean, with fewer weights towards the tails of the distribution.

To calculate the mean weight, we use the formula:
mean = (sum of weights) / (number of weights)

where the sum of weights is calculated by sum(weights) and the number of weights is calculated by len(weights). Using this formula on the given data, we get a mean weight of 3.39 kg.

To calculate the required value X such that 33% of newborns are born above X, we first calculate the cumulative distribution function (CDF) of the weights using the formula:
CDF(x) = (number of weights ≤ x) / (total number of weights)

We then find the value of X such that the CDF of X is equal to 0.67 (33% of newborns born above X). This can be done using the searchsorted() function in numpy. We get a value of X as 3.59 kg.

In conclusion, we have analyzed the given data on weights of newborn babies and created a histogram to visualize its distribution. We have calculated the mean weight to be 3.39 kg and the weight X such that 33% of newborns are born above X to be 3.59 kg.