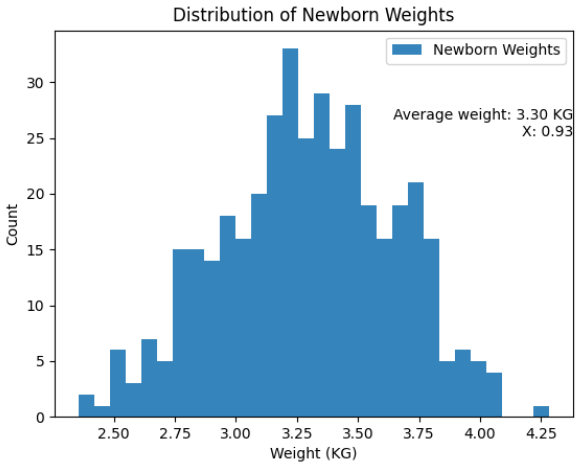
An Analysis of Newborn Weight Data in Europe Using data5.csv

Introduction & Data Description: In the assessment, a dataset was provided named **data5.csv** to analyze the weights of 400 newborns in a region of Europe. The reason to choose dataset5.csv is because my ID number ends with 5 so this particular dataset was required to be analyzed. The minimum weight in the dataset is 2.35406kg and the maximum weight is 4.28463kg, whereas the average weight is 3.29907kg. As per the analysis of the dataset, it is observed that the weight pattern is normally distributed. The aim is to use Python programming to analyze this dataset and learn more about how infant weights vary within the specified region. The Spyder environment in Anaconda is used for data analysis, where the necessary libraries i.e. CSV, Numpy, Panda, and Matplot are used to complete the given task. After reading the dataset, the histogram is plotted for the 400 entries. In the end, the mean of weight (W) and X is calculated as described below.

Description of Distribution: To visualize the distribution, we plot a histogram of the birth weights dataset. The histogram shows the frequency of the weights in the dataset having weights on X-axis and counts on Y axis. The histogram shows that babies' weight is normally distributed. The weights are evenly spread around the mean, and the peak of the histogram is about between 3.4 and 3.7 kilograms.

Mean Weight: the mean for weights can be calculated by adding all the baby's weights and dividing them by the total number of babies. The sum of whole entries is 1319.606, and the total number of babies (n) is 400, so by dividing the sum by 400 I got 3.299007. in simple words; mean weight = (sum of all weights) / (total entries of weight)

X Value: As my ID number ends with 5 so, I used data5.csv for which the value of X is a fraction of the 0.8W and 1.2W. where W is the average weight of babies i.e. 3.30kg. so by doing so the value of X I got is 0.93. As shown in the distribution plot below the blue bars represents the weights of newborn, average weight and X.



Name: Hassan Faraz Khan

Student ID: 21089475

Subject: Fundamentals of Data Science (Coding assignment)

Date: 3/May/2023

Data Source: data5

Repository Link: https://github.com/HassanFrazKhan/FDS-Assignment

ScreenShot