# Summary

This algorithm uses Poisson Distribution to determine the possible outcome of a future football match by using the number of goals scored by the home team and the number of goals scored by the away team. This is achieved by treating the two variables as independent set of events. The outcome of the game is predicted using three factors, the percentage probability of home team to win, the percentage probability of the game being a draw and the percentage probability of the away team winning.

We first start by importing the python modules to use for the project. Pandas is used to read, manipulate and visualize data, matplotlib is used to visualize data, numpy is used to manipulate mathematical functions on a dataset, seaborn is also used for data visualization and scipy is for scientific and technical computing.

We read the data using the Pandas read\_csv to read the dataset from the file system. We then use the pandas function head () to print out the column names in the dataset and the first few rows. To remove columns with null values, we use the method dropna ().

The first step is to find the average number of goals scored by the home teams, to achieve this, we first select the FTHG column using the following line Data[‘FTHG’]. We then use the pandas method mean () to calculate the average, we do the same to find the average away team goals.

To visualize the data, we use the pandas method plot() that takes the parameter kind which we set to hist for an histogram and line for a line graph. Because the data rows and a lot, displaying all of them in an histogram or a line graph will cause an overlap, we limit the data to the first 20 rows using the method head() that takes an integer n that represents the number of rows to retrieve from the dataset.

The next step is developing the prediction model using Poisson distribution. This is done by the assumption that the number of goals scored by the home team and the number of goals score by the away team are independent Poisson distribution. The probability of the home team winning is the probability of the home team scoring more goals than the away team and the probability of a draw occurring is when the two teams score the same number of goals or no goals at all.