This project is based on the historical data downloaded from <http://www.football-data.co.uk/downloadm.php>.

Each excel spreadsheet contains all results for a given season divided into sheets containing different leagues. For instance, ‘E0’ – Premier League, ‘I1’ – Serie A, and so on.

The most important columns are:

FT(A/H)G – full time (away/home) goals

FTR – full time result (home/draw/away)

BBAv(H/D/A) – average bookmakers odds for (home/draw/away)

The stats\_tools.py file provides us with all necessary statistical tools we will be using to return calculations outlined in the pdf file.

The data\_grab.py file collects the data from the excel spreadsheet and applies the statistical tools defined above.

The model\_test.py file runs through a series of seasons to create a comparison between different types of models mentioned in the pdf file.