Stat 400 Project - Group 4

Megan Dunnahoo, Jasmine DeMeyer, Macey Dodd

12/6/2021

Article

We chose the article "Using Monte Carlo Simulation to Calculate Match Importance: The Case of English Premier League" by Jiri Lahvicka. This article describes the process of using Monte Carlo simulations to predict the outcome of a match given the results of previous matches. It then goes further and uses Monte Carlo simulations to predict the final ranking of the teams in the English Premier League at the end of a season.

Soccer Information

What is necessary information on soccer? There are 20 teams in the English Premier League. Each team plays 19 away games and 19 home games?

Data Manipulation

We wanted to use the specific variables FTR, FTAG, and FTHG, along with the identifier variables of Date, Away Team, and Home Team. There were no NA values in any of our selected variables.

Results

```
## # A tibble: 1,254 x 9
## # Groups:   AwayTeam [20]
##    Date     HomeTeam    AwayTeam  FTHG  FTAG FTR      Id
##    <chr>    <chr>       <chr>    <int> <int> <chr> <int>
##  1 13/08/11 Newcastle   Arsenal      0     0 D         4
##  2 28/08/11 Man United  Arsenal      8     2 H        26
##  3 17/09/11 Blackburn   Arsenal      4     3 H        41
##  4 02/10/11 Tottenham   Arsenal      2     1 H        69
##  5 29/10/11 Chelsea     Arsenal      3     5 A        90
##  6 19/11/11 Norwich     Arsenal      1     2 A       112
##  7 03/12/11 Wigan       Arsenal      0     4 A       136
##  8 18/12/11 Man City    Arsenal      1     0 H       157
##  9 21/12/11 Aston Villa Arsenal      1     2 A       162
## 10 02/01/12 Fulham      Arsenal      2     1 H       192
## # ... with 1,244 more rows

## # A tibble: 1,254 x 9
## # Groups:   HomeTeam [20]
##    Date     HomeTeam AwayTeam    FTHG  FTAG FTR      Id
```

Going Forward

In the next week we plan to continue to expand on our exploration of this data. We plan to use Monte Carlo simulations to predict the final rank of the two teams we have chosen which are Chelsea and Liverpool. We also hope to create a final contingency table that consists of the rank for these two teams as well as how many wins, losses, and draws they had in our simulations.