

ETC5513-Assignment3

```
library(tidyverse)
```

```
-- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
v dplyr      1.1.4      v readr      2.1.5
v forcats    1.0.0      v stringr    1.5.1
v ggplot2    3.5.1      v tibble     3.2.1
v lubridate  1.9.4      v tidyr      1.3.1
v purrr      1.0.4
```

```
-- Conflicts ----- tidyverse_conflicts() --
```

```
x dplyr::filter() masks stats::filter()
```

```
x dplyr::lag()     masks stats::lag()
```

```
i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become
```

```
epl_raw <- read.csv("data/epl_final.csv")
```

```
epl_data <- epl_raw %>%
```

```
  select(
    Season,
    HomeTeam, AwayTeam,
    FullTimeHomeGoals, FullTimeAwayGoals,
    FullTimeResult,
    HomeShots, AwayShots
  )
```

Executive Summary

Introduction

Methodology

Results

Discussions

The analysis examined the relationship between the average number of shots made by a team and the win rate of the team across different seasons. The findings from our visual analysis reveals that there is a positive relationship between the average shots and win rate. But, we can also see that there are several outliers as well which shows that there are some teams which even though attempting more shots were unable to win the matches. This also indicates that we will require more data to give precise analysis on the win rate for a team as the quantity of the shots may not be enough to give accurate results. We can look into the factors which specify the quality of shots as well.

Conclusions

This analysis demonstrate that EPL teams that produce more shots per game are more likely to achieve higher win rates. While the relationship is not absolute, the patter is strong enough to highlight that the offensive strategies for the teams are a good way for a season long success.

Recommendations

- Further Analysis: We can include more variables like defensive errors, actual goals, quality of shots, etc. to improve our model.
- Top Performers: We can also compare the outliers and other team (Top performing teams) to see the patterns of abnormality.

References