Data Analysis project: UK Soccer data analysis

Data source: [England Football Results Betting Odds | Premiership Results & Betting Odds (football-data.co.uk)](football-data.co.uk)

Metadata:

No. of fields in each file: 67

No. of CSV files analyzed: 30

Important data fields: Date of match

Home team/ Away team

Goals at half time/ full time for both home and away teams

Referee name

Half time and full-time result: D- Draw, A-Away team wins, H-Home team wins

Shots and fouls, yellows, reds, and corners by home and away teams

Betting data

## Questions answered in the analysis:

1. What is the best performing team overall (# of victories)?
2. What is the worst performing team overall?
3. Which team scored the most goals in each season?
4. Which team scored the least goals overall?
5. Which is the best performing team in the last 3 years? Is that the same team that scored most goals? From 2020,21,22 and 23 last 2 months
6. What are the odds of the best team in the last 3 years winning against the 2nd and 3rd best teams?
7. Plot the distribution of goals per match for every year (winner vs loser)
8. Plot the distribution of goals of the winner team for every match (winner vs loser)
9. Plot the correlation between high score and matches between 2 teams for the total goals per match
10. Plot the correlation between total goals of the winning team between 2 teams for the total goals per match
11. Plot the evolution of the average # of goals per match per week (including all seasons) compared to the least and max # of goals
12. Is there a team that is consistently improving the performance over the years?
13. Is there a team that is consistently losing the performance over the years?

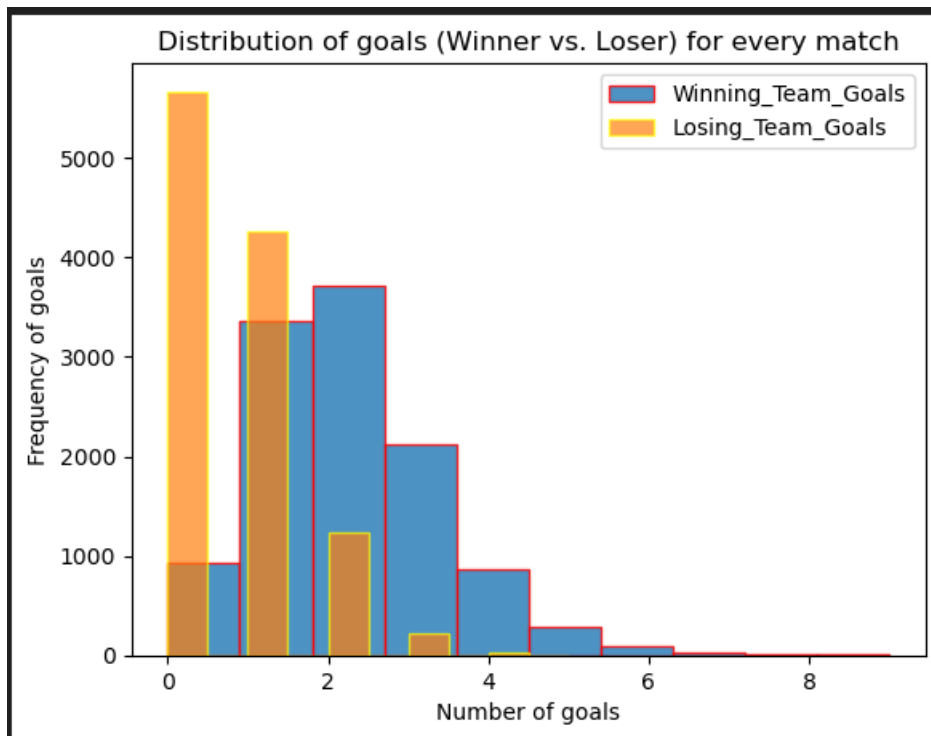## Answers and code file with Exploratory data analysis:
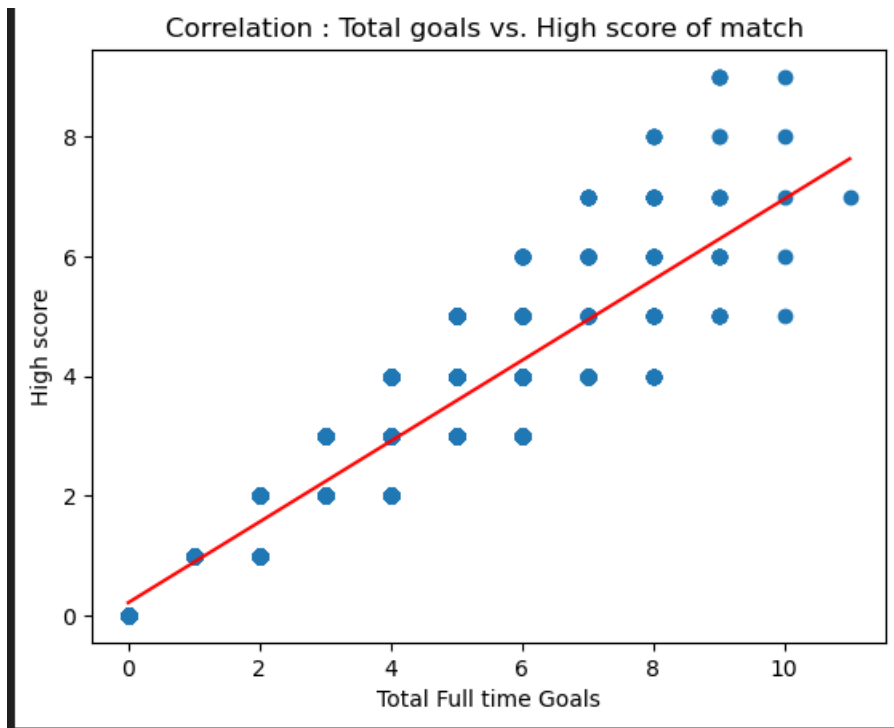
Exploratory data analysis examples:

Goals: Winner vs. loser
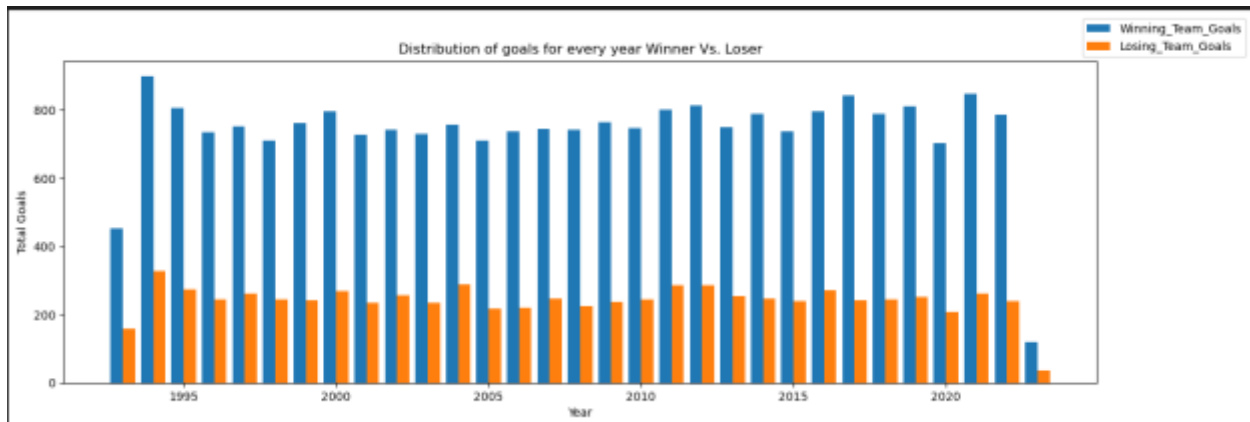
Histogram of Goals: Winner vs. loser

Correlation chart: Total goals in a match Vs. High score of the match

Correlation : Total goals vs. High score of match

Correlation between winning team's goals and the total goals

Correlation between winning and losing team's goals and the total goals
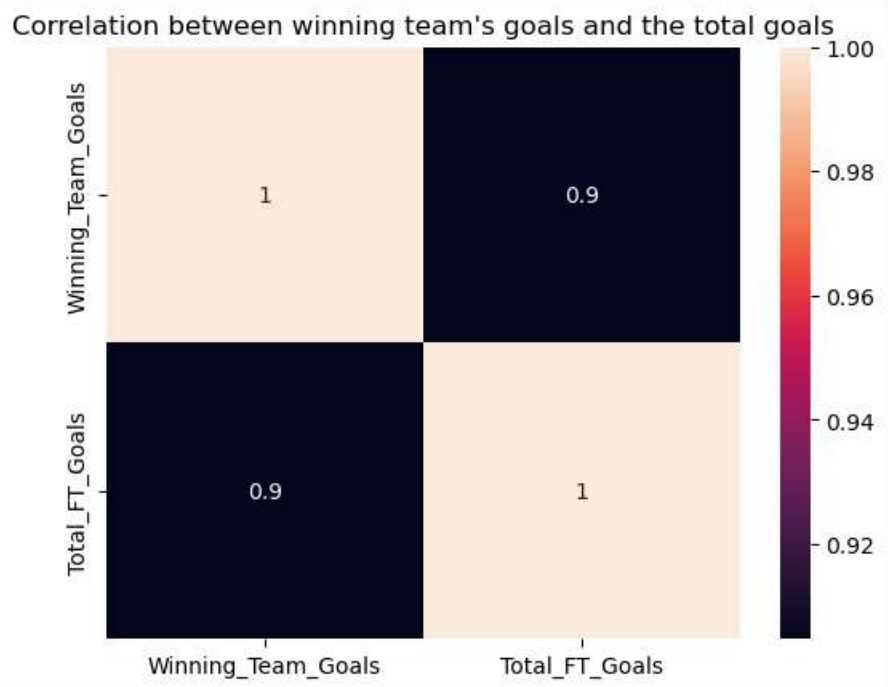
Complete goal analysis of all teams to find the team with consistent improvement: