Home Advantage in Soccer: Myth or Reality?

Exploratory Data Analysis on 11,000 Match Results

What Are We Investigating?

1

Dataset

11,000 Soccer results from the english premier league 1993-2022

2

Goal

Determine whether playing at home offers a real performance advantage

3

Approach

Exploratory Data Analysis (EDA) & visualization

4

Final Output

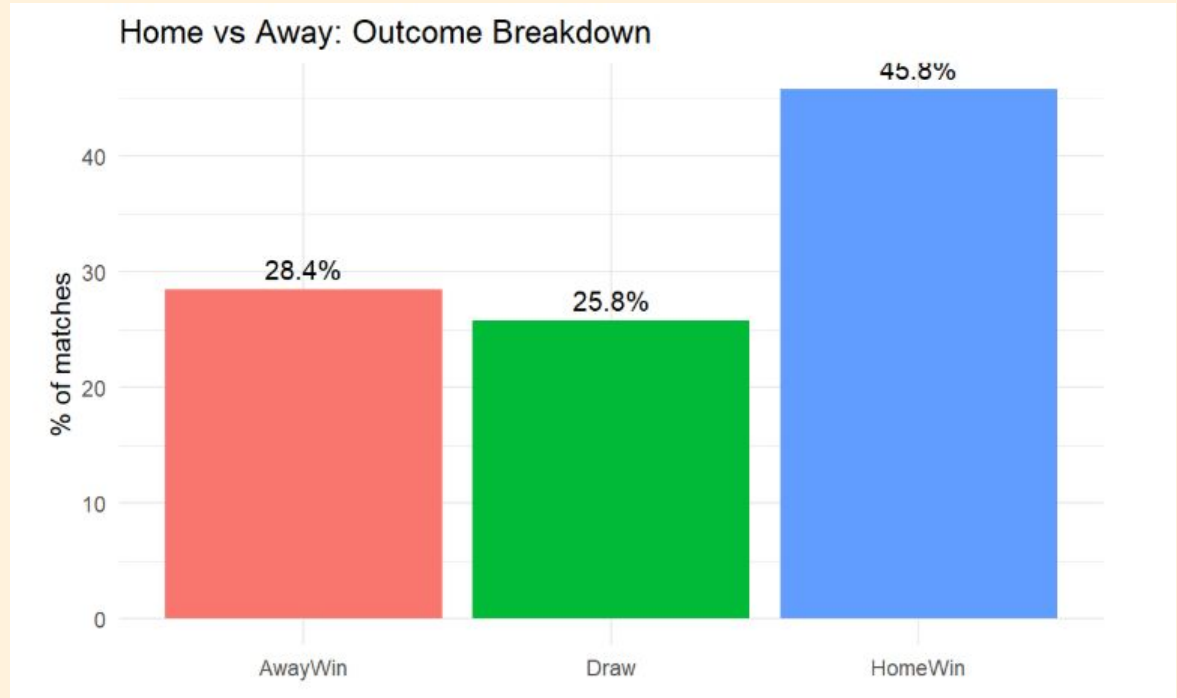Answering our SMART question

SMART question

Does playing at home significantly increase a Premier League team's chances of winning over the past seasons?

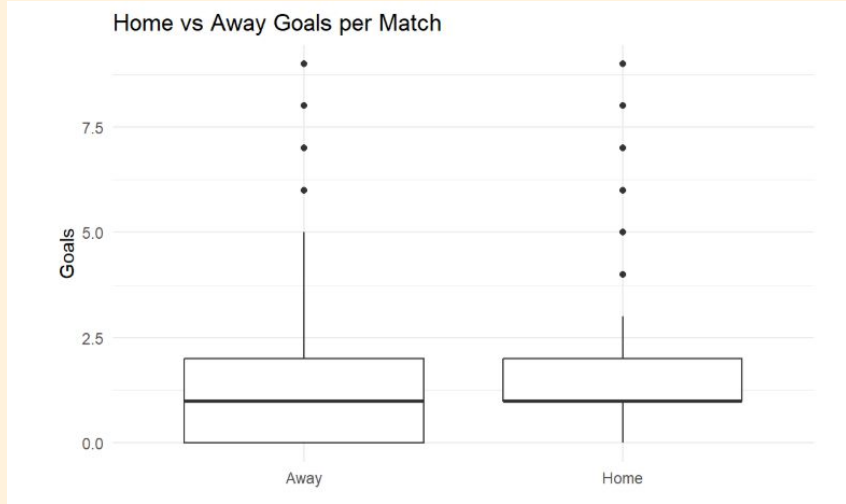# Understanding the Dataset

- Total Matches: 11,113
- Features:
  - Season   Match Season
  - DateTime   Match Date and Time (yyyy-mm-dd hh:mm:ss)
  - HomeTeam   Home Team
  - AwayTeam   Away Team
  - FTHG Full Time Home Team Goals
  - FTAG Full Time Away Team Goals
  - FTR   Full Time Result (H=Home Win, D=Draw, A=Away Win)
  - +
- Cleaned & processed for consistency
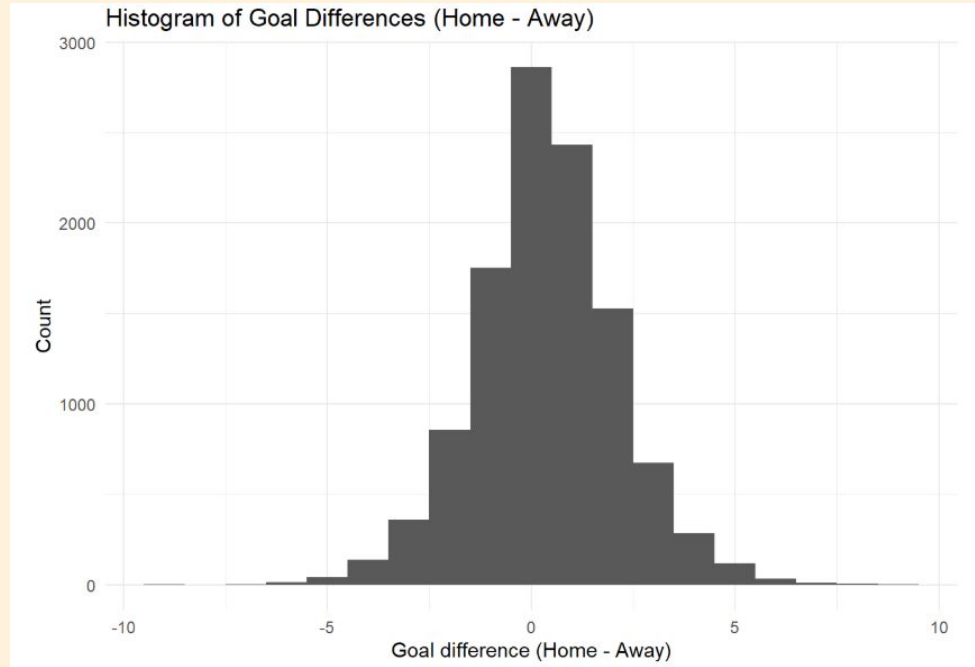
Home vs Away: Outcome Breakdown

- 45.8% of games won by Home Team
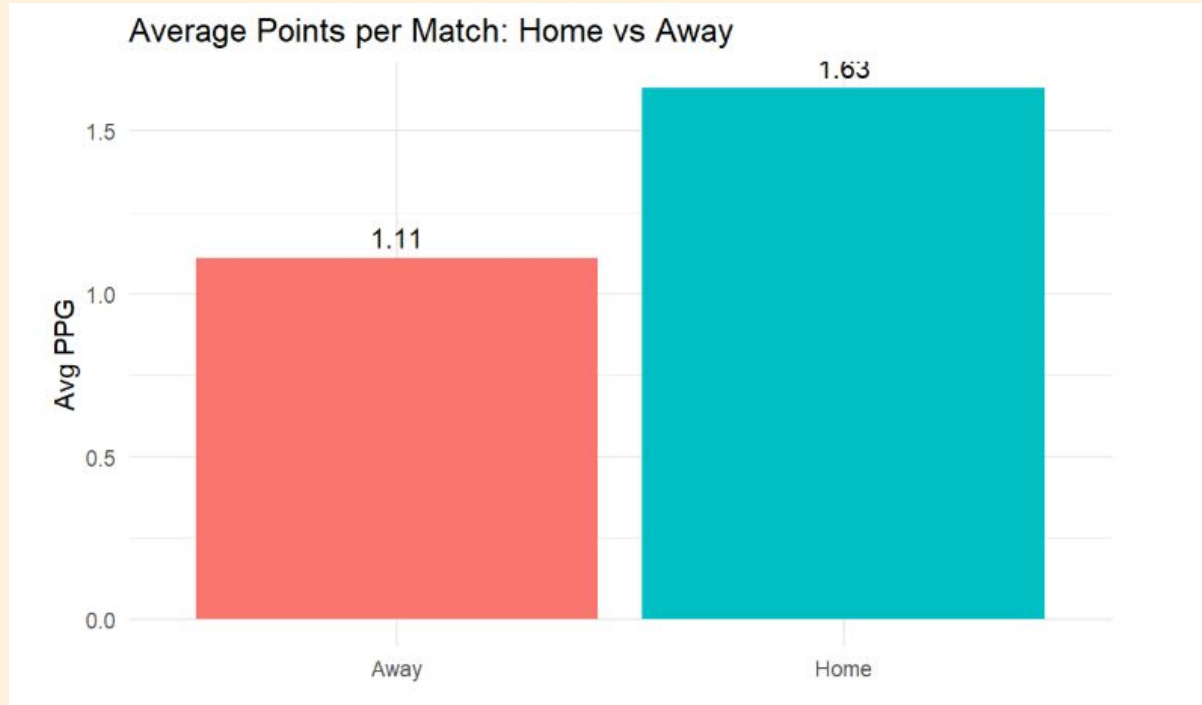- 28.4% of games won by Away Team
- 25.8% of games end in a draw

Home vs Away: Outcome Breakdown

Goals Scored: Home vs Away

Home vs Away Goals per Match

avg_home_goals	avg_away_goals
<dbl>	<dbl>
1.518222	1.141456

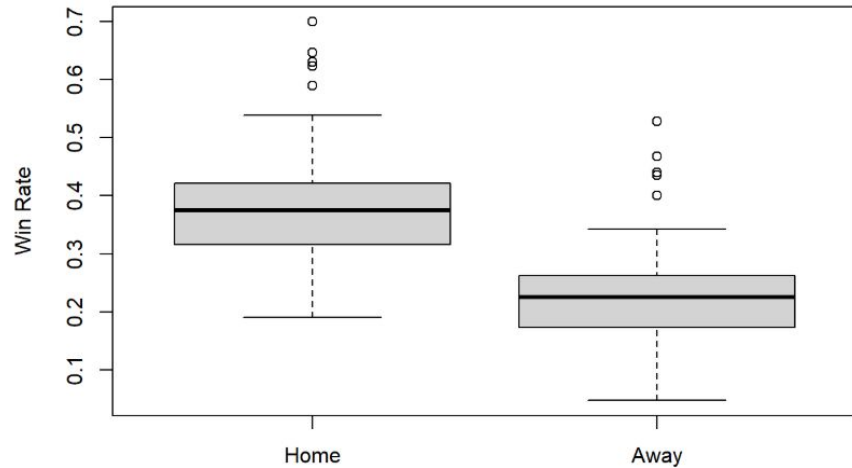Histogram of Goal Differences (Home - Away)

Points Per Game (PPG)
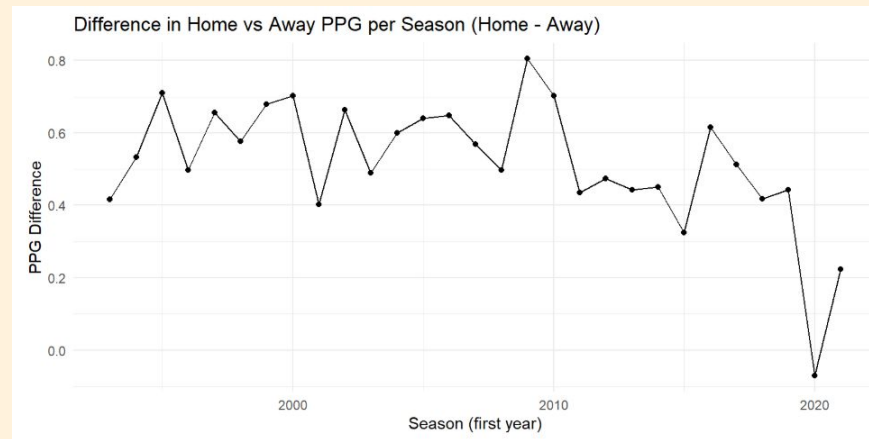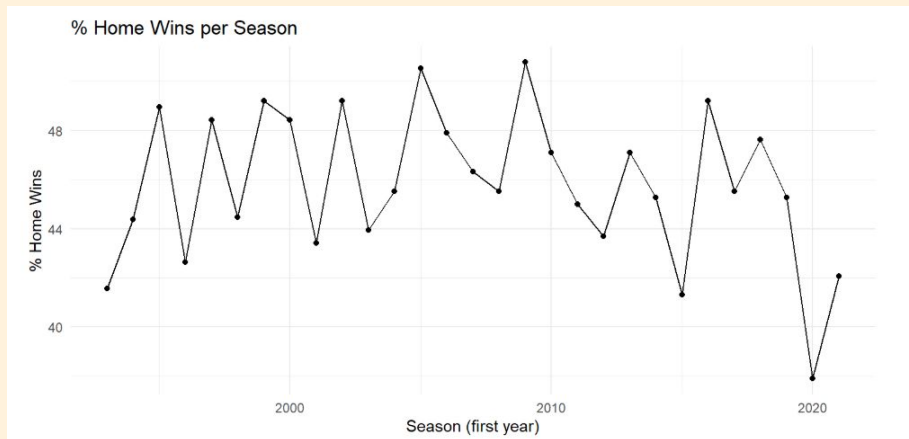
Home vs Away Win Rates

Premier League: Home vs Away Win Rates


```
        Paired t-test

data:  team_rates$home_winrate and team_rates$away_winrate
t = 19.784, df = 49, p-value < 2.2e-16
alternative hypothesis: true mean difference is greater than 0
95 percent confidence interval:
 0.1440735        Inf
sample estimates:
mean difference
     0.1574133
```

Has Home Advantage Changed Over Time?

Possible Explanations

- Crowd support
- Referee bias (unconscious)
- Travel fatigue for away teams
- Tactical adjustments at home and familiarity with pitch

Is There a Real Home Advantage?

- Data shows consistent advantage for home teams
- Higher win rate, more goals, more points
- Advantage is shrinking slightly over time
- Answer: **Yes**, but the impact is contextual

What Could Be Expanded?

- Include team strength or ranking to adjust for quality
- Consider special cases (derbies, weather, stadium size)
- Consider travel distance
- Look at other leagues for comparison
- Use modeling for predictive analysis

Thank you!