



Premier League Exploratory Data Analysis

Fahad Alsaud



Introduction



Premier League

- Introduced in 1992
 - Dataset used was from 2006-2018
- Soccer League in the United Kingdom
- The league has 20 teams each year
 - The bottom three teams go to a lower league and are replaced with new teams

Dataset

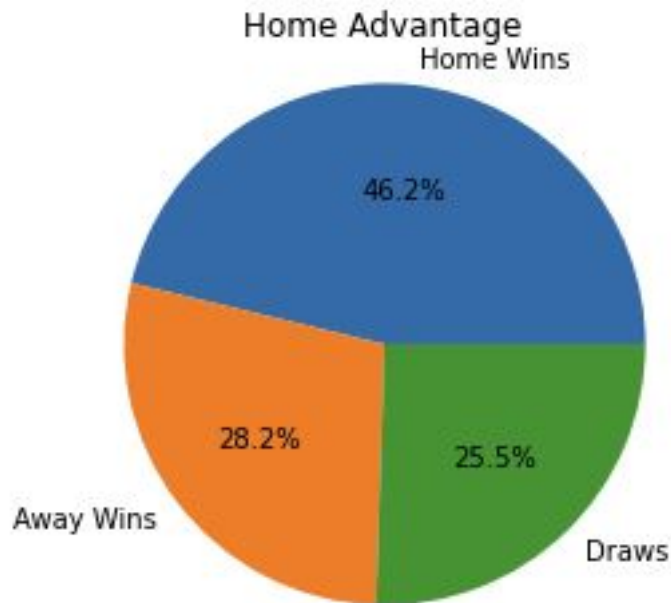


Premier League

- 2006-2018
- Measures numerous statistics
 - Wins, losses, touches, crosses, goals, etc.
- Also had statistics for each game
- Useful for exploratory analysis

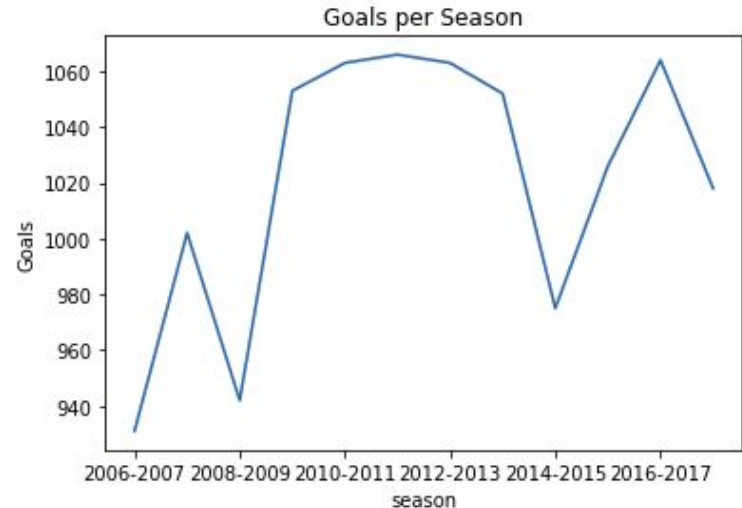
Home Advantage

- Home team won 46.2%
- Away team won 28.2%
- Draws were 25.5%
- Clear home advantage



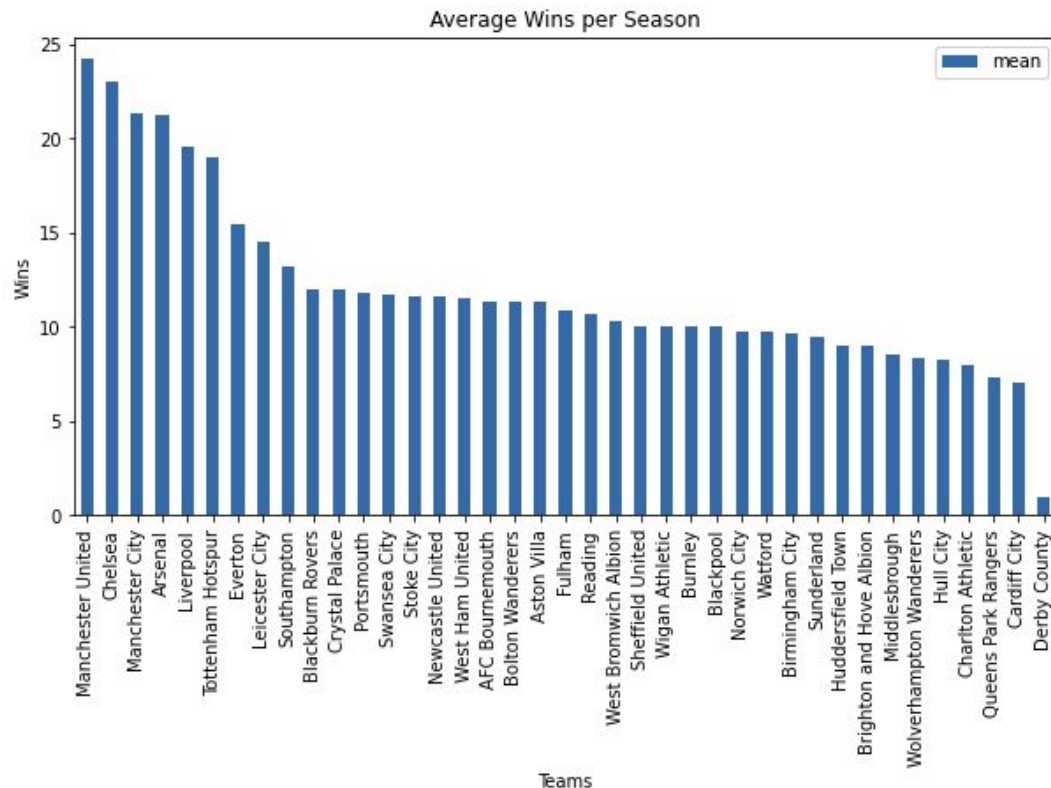
Goals per Season

- Goals per season has increased on average
- My hypothesis
 - More goals = attractive for fans
 - Thus, more goals= more fans
 - More fans = more revenue

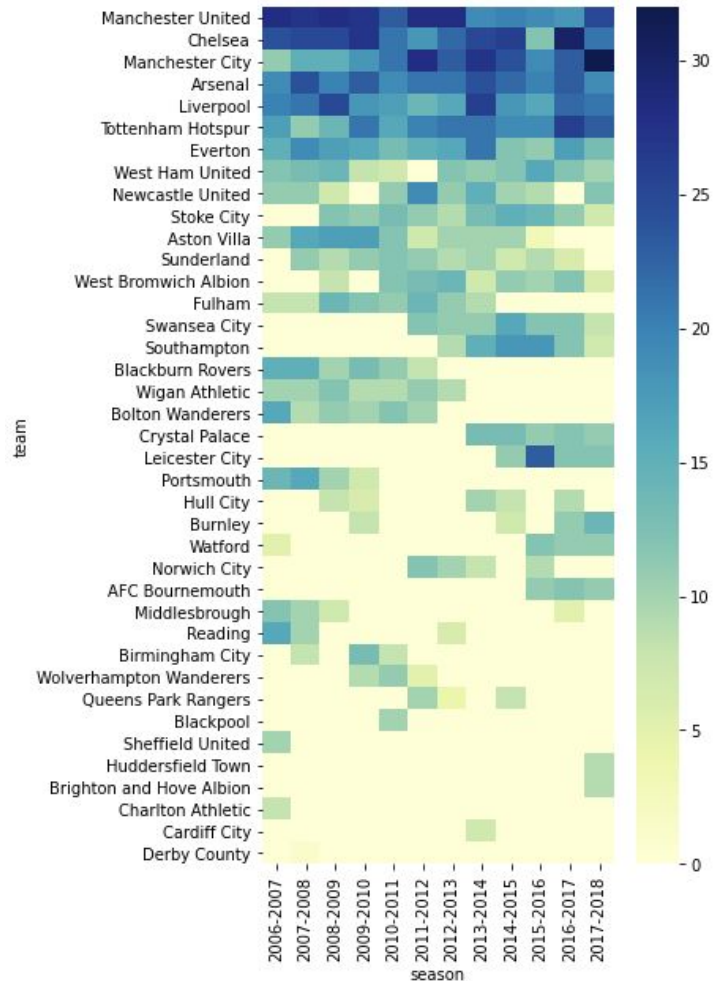


Most Successful Teams

- Manchester United are the most successful team in the era
- Derby County are the least successful team in the era
- Continued on next slide

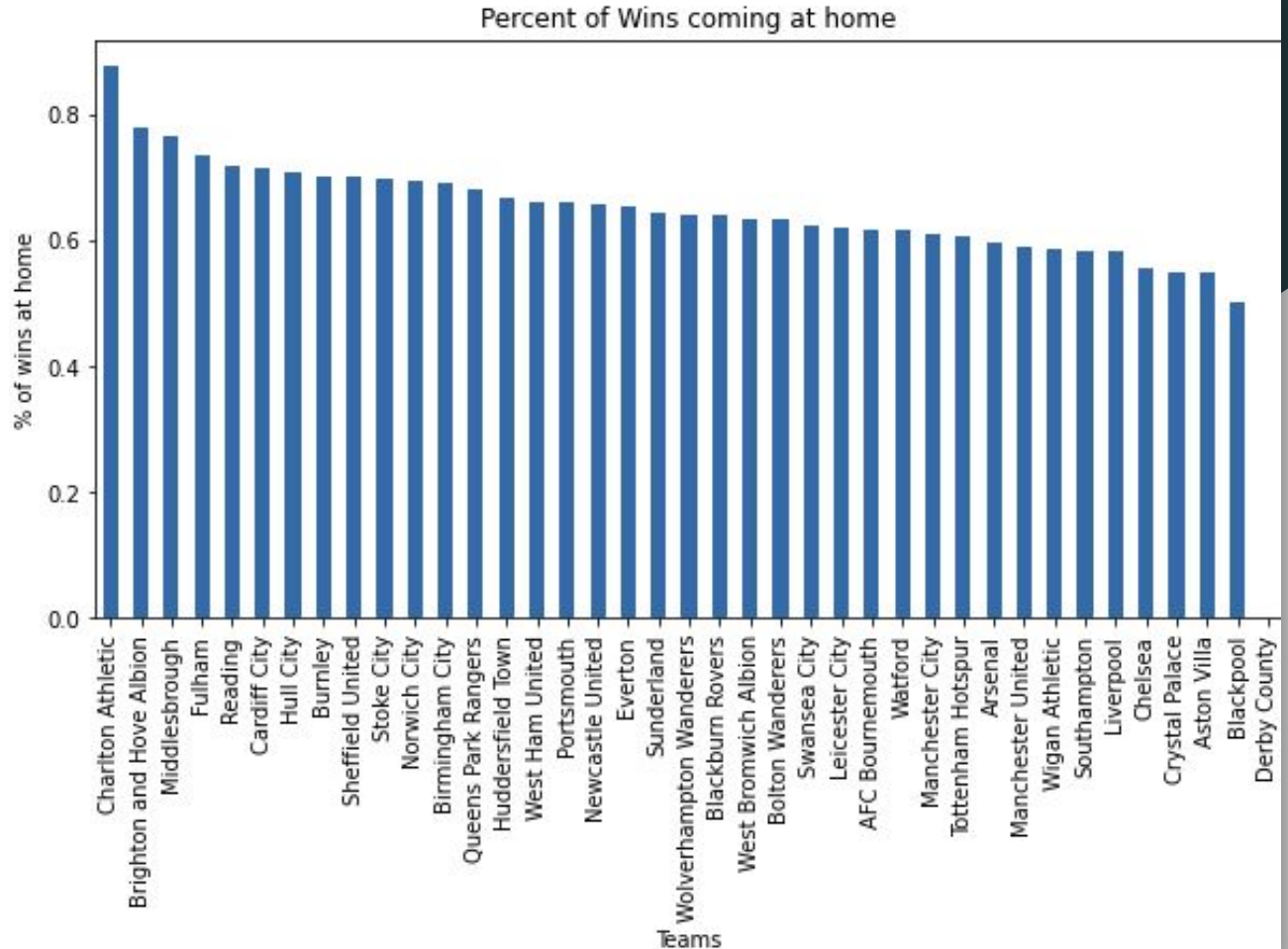


Most Successful Teams (per season)



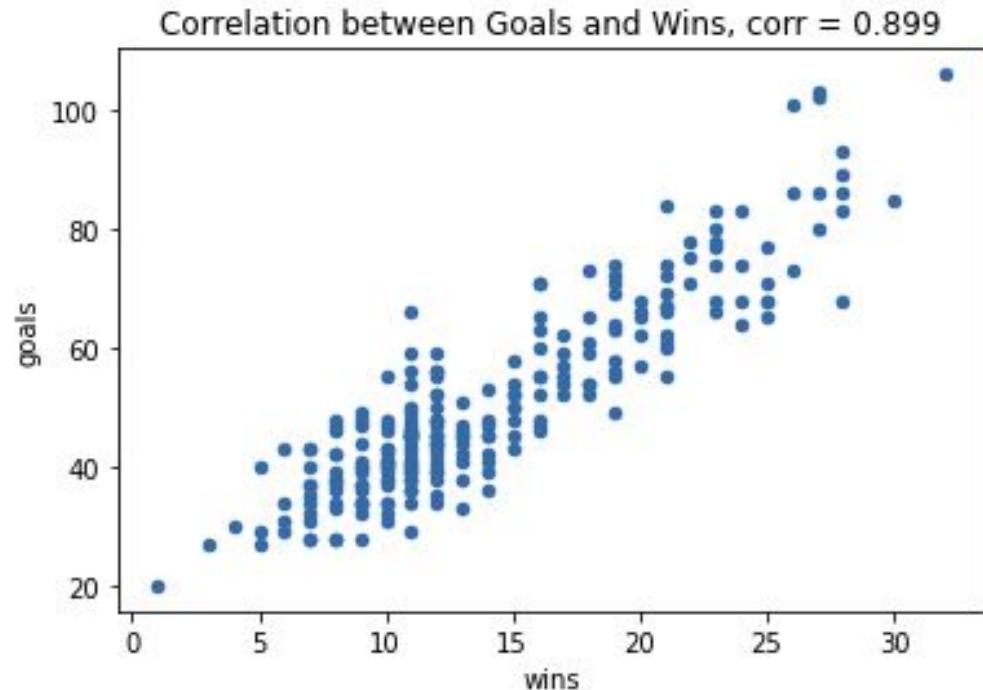
Reliance on Home Field Advantage

- Worse teams rely more on home field advantage
- Correlation = -0.5
- My Hypothesis
 - Worse teams require the extra boost in order to win games



Correlating Stats with Wins Goals

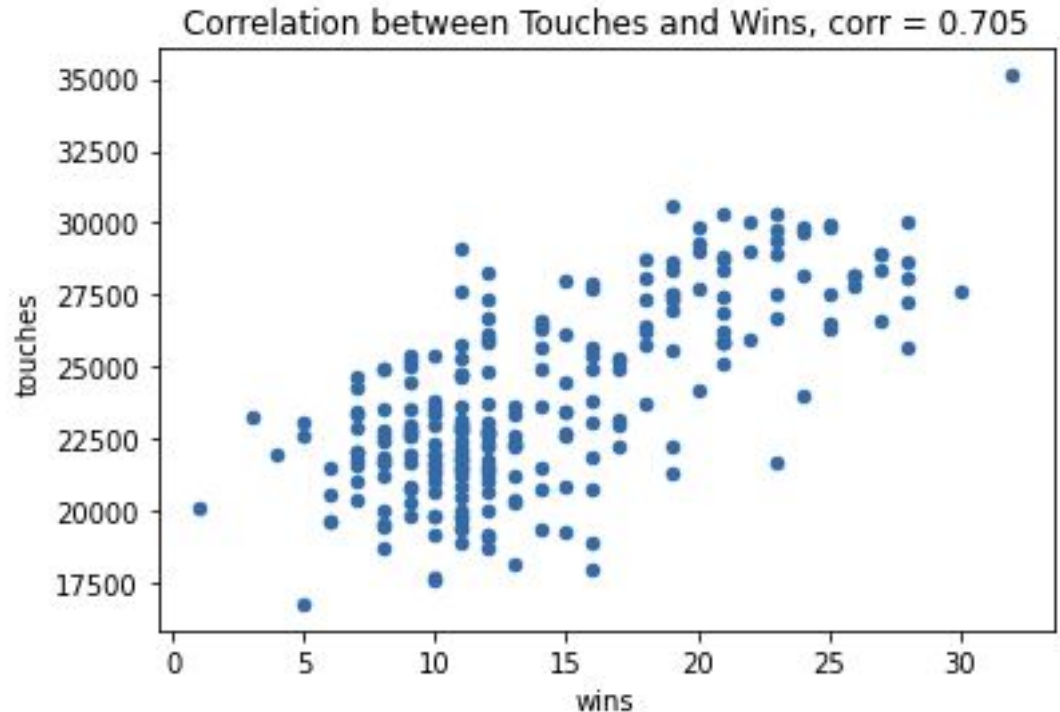
- Strong, Positive Correlation
- Teams with higher goals win more
- Attacking teams are more successful



Correlating Stats with Wins

Touches

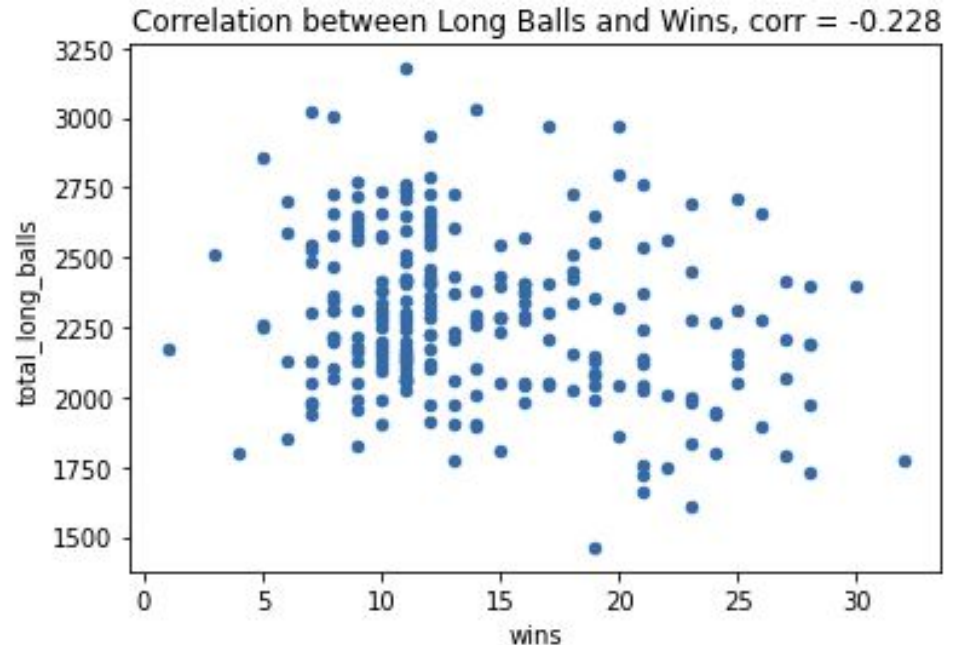
- Strong, Positive Correlation
- Teams that keep the ball more, win more
- Again, attacking teams (teams that have more touches) are more successful than defensive teams



Correlating Stats with Wins

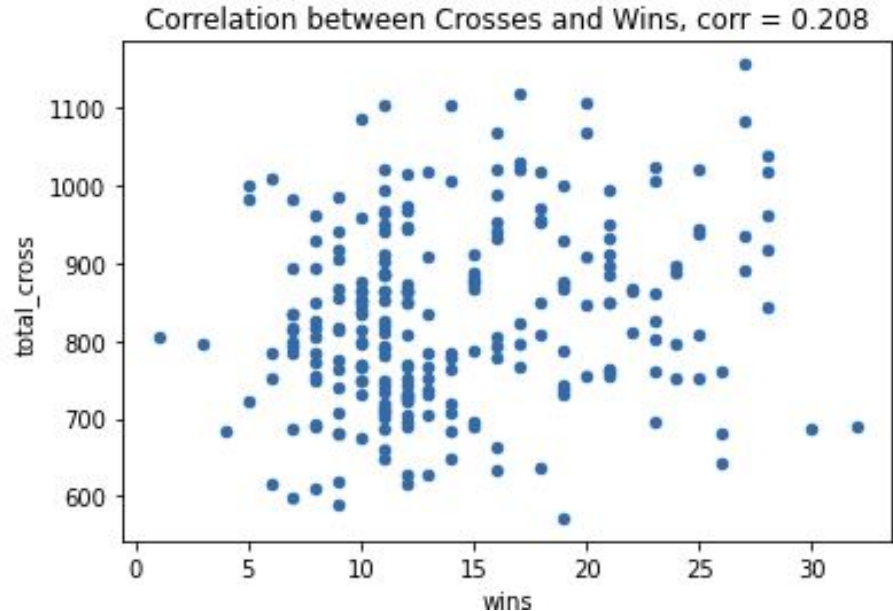
Long Balls

- Weak, Negative Correlation
- Teams that use long balls, win less
- Typically, long balls are used by defensive teams
- Thus, attacking teams (less long balls) are more successful than defensive teams



Correlating Stats with Wins Crosses

- Weak, Positive Correlation
- Teams that cross more, win more
- Crosses are produced in the attacking half
- Thus, attacking teams (more crosses) are more successful than defensive teams



Conclusion

- Attacking teams typically do better than defensive teams
 - More touches
 - More goals
 - Less long balls
 - More crosses
- More successful teams don't heavily rely on home field advantage

