Analyzing the Dynamics of Home Advantage: A Statistical Investigation into Goals, Corners, and Shots on Target in Premier League Football*

Jerry Lu (Yen-Chia Lu)

April 2, 2024

This study explores the phenomenon of home advantage in the Premier League for the 2022–2023 campaign by analysing the correlation between goals, corners, and shots on target for the home club. Advanced statistical studies are performed using a season-long dataset to find trends and insights that clarify the significance of playing at home. Significant correlations that have practical consequences for clubs and coaching techniques as well as contributing to the scholarly discourse on sports analytics are highlighted by the findings. In light of the most recent season, this research provides fresh perspectives on the tactical and psychological factors affecting football game results.

1 Introduction

Since its founding in 1992, the Premier League has grown to become one of the world's most competitive and watched football competitions. Each of the twenty teams that compete each year plays 38 games, half of which are played at their home stadium and the other half away. The league is renowned for its competitive matches, fervent fandoms, and the exceptional skill level of its clubs and players. This competitive atmosphere creates an ideal setting for investigating several aspects of a team's performance, especially the home field advantage.

In sports, the idea of home team advantage is widely accepted, and the Premier League is no different. Playing at home is frequently linked to a higher chance of success, which can be related to things like comfortable surroundings, the backing of local supporters, and the lack

*Code and data are available at: https://github.com/Jerryluuuu/-Premier-League.git

of travel fatigue. Metrics that are essential for assessing a team's attacking prowess at home games include Full Time Home Team Goals, Home Team Shots on Target (HST), and Home Team Corners (HC). Understanding how home advantage might affect a team's attacking dynamics and overall performance in a match depends on these factors.

Not to be overlooked in this situation is the referee's involvement. Although it is the referees' responsibility to maintain fair play, their judgements have a big influence on how the game goes and how it ends. There is interest in the idea of "home bias," which holds that choices may unintentionally benefit the home team, for example, by giving them extra time off from injuries or more lenient rulings for fouls. Understanding the psychological effects of home advantage and the impact of outside variables on game outcomes can be gained by examining the relationship between referee decisions and home team performance indicators like goals, shots on target, and corners.

Furthermore, a key factor in maximising home advantage is the home team's confidence. When combined with the encouragement of home fans, the psychological boost of playing in a familiar environment can improve team performance. Higher percentages of shots on target and in corners are signs of an aggressive offensive strategy that frequently reflects this confidence. We can measure the effect of home confidence on a team's performance by looking at these variables, which provides a greater understanding of the psychological and strategic components that contribute to the Premier League's home advantage occurrences.

2 Data

You can and should cross-reference sections and sub-sections. We use R Core Team (2023) and Wickham et al. (2019).

3 Model

3.1 Model set-up

$$y_i|\mu_i, \sigma \sim \text{Normal}(\mu_i, \sigma) \tag{1}$$
$$\mu_i = \alpha + \beta_i + \gamma_i \tag{2}$$
$$\alpha \sim \text{Normal}(0, 2.5) \tag{3}$$
$$\beta \sim \text{Normal}(0, 2.5) \tag{4}$$
$$\gamma \sim \text{Normal}(0, 2.5) \tag{5}$$
$$\sigma \sim \text{Exponential}(1) \tag{6}$$

We run the model in R (R Core Team 2023) using the `rstanarm` package of Goodrich et al. (2022). We use the default priors from `rstanarm`.

3.1.1 Model justification

We expect a positive relationship between the size of the wings and time spent aloft. In particular...

We can use maths by including latex between dollar signs, for instance θ.

4 Results

5 Discussion

5.1 First discussion point

If my paper were 10 pages, then should be be at least 2.5 pages. The discussion is a chance to show off what you know and what you learnt from all this.

5.2 Second discussion point

5.3 Third discussion point

5.4 Weaknesses and next steps

Weaknesses and next steps should also be included.

Appendix

A Additional data details

B Model details

B.1 Posterior predictive check

References

Goodrich, Ben, Jonah Gabry, Imad Ali, and Sam Brilleman. 2022. "Rstanarm: Bayesian Applied Regression Modeling via Stan." https://mc-stan.org/rstanarm/.

R Core Team. 2023. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.

Wickham, Hadley, Mara Averick, Jennifer Bryan, Winston Chang, Lucy D'Agostino McGowan, Romain François, Garrett Grolemund, et al. 2019. "Welcome to the tidyverse." *Journal of Open Source Software* 4 (43): 1686. https://doi.org/10.21105/joss.01686.