Aalto-yliopisto
Aalto-universitetet
Aalto University

Bachelor's Programme in Information and Service Management

The Impact of Fan Attendance on Expected Goals and Finishing Efficiency

A Comparative Analysis of Premier League Matches With and Without Fans

---

Walter Kantola

Bachelor's thesis
2024

Aalto-yliopisto
Aalto-universitetet
Aalto University

Author Walter Kantola

Title of thesis The Impact of Fan Attendance on Expected goals and Finishing Efficiency

Programme Bachelor of Science in Economics and Business Administration

Major Information and Service Management

Thesis supervisor Senior Fellow Tomi Seppälä

Thesis advisor(s) Senior Fellow Tomi Seppälä

Date 31.12.2024 **Number of pages** 33 **Language** English

Abstract

This thesis investigates how fan attendance impacts expected goals (xG) and finishing efficiency (FE) in the Premier League by comparing the 2019/2020 season with fans to the 2020/2021 fanless season. The analysis shows that xG values significantly decreased in fanless matches, particularly at home, while finishing efficiency exhibited slight changes, with a minor decrease in home matches and an increase in away matches Fan absence also reduced home advantage (HA), leveling performance between home and away teams. These findings increase understanding of how external factors like fan presence shape football performance, contributing to football analytics.

Keywords expected goals, finishing efficiency, fan attendance, Premier League, home advantage, football analytics

Tekijä Walter Kantola

Työn nimi Yleisön vaikutus odotettuihin maaleihin ja viimeistelytehokkuuteen

Koulutusohjelma Kauppatieteiden kandidaatti.

Pääaine Tieto- ja palvelujohtaminen

Vastuuopettaja/valvoja Johtava tutkija Tomi Seppälä

Työn ohjaaja(t) Johtava tutkija Tomi Seppälä

Päivämäärä 31.12.2024 **Sivumäärä** 33 **Kieli** Englanti

Tiivistelmä

Tämä tutkielma selvittää, kuinka yleisön läsnäolo vaikuttaa expected goals (xG) -malliin ja viimeistelytehokkuuteen Valioliigassa vertaamalla yleisön kanssa pelattua kautta 2019/2020 ja pääosin ilman yleisöä pelattua kautta 2020/2021. Analyysin mukaan xG-arvot laskivat merkittävästi otteluissa ilman yleisöä, erityisesti kotijoukkueilla, kun taas viimeistelytehokkuus (maalit/maaliodottama) muuttui hieman: kotijoukkueiden tehokkuus laski hieman, kun taas vierasjoukkueiden tehokkuus kasvoi. Yleisön puuttuminen vähensi myös kotietua, mikä tasoitti kotijoukkueiden ja vierasjoukkueiden suorituksia. Nämä havainnot lisäävät ymmärrystä siitä, kuinka ulkoiset tekijät, kuten yleisön läsnäolo, vaikuttavat jalkapallon suorituskykyyn ja rikastuttavat jalkapalloanalytiikan tietämystä.

Avainsanat maaliodottama, viimeistelytehokkuus, yleisön läsnäolo, Valioliiga, kotietu.

Table of contents

1 Introduction

Football is a sport where performance is influenced not only by skill and tactics but also by external factors such as fan attendance (Pollard & Gomez, 2014). The presence of fans has long been recognized as a crucial element affecting the dynamics of a match, most notably through its impact on home advantage (HA) and player performance (Bilalić et al., 2023).
Fan support plays a crucial role in providing additional motivation and confidence to home teams (Seckin & Pollard, 2008) while putting extra pressure on away teams. It can also considerably affect refereeing decisions in favor of the home team (Nevill et al., 2002). This effect is known as the "12th man", referring to the impact of the home crowd on the course of the game (Bilalić et al., 2021).

The COVID-19 pandemic, which broke out at the beginning of 2020, brought about an unusual situation worldwide. This also meant a major change for professional football. During a large part of the 2020/2021 season and the end of the 2019/2020 season, competitive games were played without spectators in attendance due to pandemic-induced restrictions. This unprecedented situation provided a unique opportunity to study the effects of crowd absence on match outcomes and team performances. This presents an opportunity to examine expected goals (xG), a key statistical measure of football analytics, and investigate how the absence of fans might affect the predictive power of xG models and the efficiency of teams in finishing (Goals/xG).

Therefore, this study examines the relationship between fan attendance and xG by comparing the 2019/2020 Premier League matches, where fans were mostly present, with the 2020/2021 season, where most of the season was played in empty stadiums. The goal is to investigate whether the lack of fans of fans influenced xG as a predictive metric for match outcomes and teams' ability to finish their goal-scoring opportunities. Moreover, it seeks to understand the impact it had on HA, a phenomenon usually amplified by the motivation of home fans. By analyzing these factors, the research aims to contribute to the growing body of literature on the external effects of football performance and provide practical insights to performance evaluation.

A large body of research has been published on HA and crowd effects, but little on whether these dynamics interact with modern performance metrics such as xG and finishing efficiency (FE). Much prior work has focused on traditional match outcomes, like win rates and goal differences, without fully integrating analytical tools. Additionally, research has shown a positive relationship between fan attendance and home teams' performance. However, there is a limited amount of research on how the absence of fans is impacting predictive models and especially performance metrics like FE in football analytics. This study seeks to help understand these gaps by integrating advanced performance metrics with insights into the influences of fan presence.

To explore these objectives, the study employs a methodology that combines statistical analysis of match data with a comparative approach. Key performance metrics, including xG, goals scored, and FE, will be analyzed for home and away teams across the two seasons. The research will also assess whether the absence of fans led to changes in HA, particularly in terms of decreased performance for home

teams and improved performance for away teams. By focusing on these elements, the study contributes to the existing literature on how external factors influence football analytics and offers practical insights, especially for coaches and football analysts.

1.1 Impact of COVID-19 on Football Matches and Fan Attendance

On March 13th, 2020, all professional football games across Europe, including the English Premier League were suspended which wiped out the European football calendar for the first time since WWII (Bond et al., 2022). The English Premier League 2019/2020 season resumed behind closed doors on June 17th, 2020, initiating a period where matches were played without fans. This situation lasted mostly throughout the 2020-2021 season apart from a short period in limited areas in December 2020 and at all grounds in May 2021, when clubs were allowed a limited number of attendance. Fans finally fully returned for the 2021-2022 season on July 19, 2021.

1.2 Purpose of the Study

The primary aim of this study is to explore how fan attendance affected xG and FE in Premier League matches during the 2019-2020 and 2020-2021 seasons. Specifically, it investigates whether the absence of fans influenced teams' ability to convert xG into goals and examines the impact on HA by comparing xG and actual goal outcomes in matches played with and without crowds.

1.3 Research Questions

The core focus of this study is to analyze the impact of fan attendance on the predictive accuracy of xG and to compare performance between seasons with and without fans. The research is guided by the following questions:

1) How did the predictive accuracy of xG differ between the 2019-2020 Premier League season (with fans) and the 2020-2021 season (without fans)?

2) To what extent did teams overperform or underperform their xG during the fanless 2020-2021 season compared to the 2019-2020 season?

3) How did the absence of fans affect home and away xG performance, and what differences can be observed between these contexts across the two seasons?

1.4 Structure of the Research

This thesis is organized into six chapters. Chapter two provides a review of relevant literature on xG, FE, and the influence of fan attendance. Chapter three outlines the research methodology, including data collection and analysis methods. Chapter four presents the results of the comparative analysis between matches with and without fans. Chapter five concludes the study, discussing the findings, limitations, and recommendations for future research. References are provided in chapter six.

2 Literature Review

This section examines the relevant concepts and frameworks that will help in understanding the relationship between fan attendance, xG, and FE in football. It emphasizes defining xG as a metric for evaluating performance, discusses how fan attendance influences match dynamics and player behavior, and identifies the factors affecting FE based on previous studies.

2.1 Overview of Expected Goals

XG is a commonly used metric in football analytics designed to estimate the probability of a shot leading to a goal based on several key factors. These include shot location, player's angle to goal, and the quality and type of the pass received (e.g. through ball or cross). Defensive pressure, such as the number of defenders between the shooter and the goal, is also a crucial variable (Narayanan & Pifer, 2024). Other variables, such as the part of the body used to take the shot (e.g. foot or head) and the goalkeeper's position, affect the xG calculation. Comparing each shot to thousands of historical attempts with similar characteristics yields an xG value that ranges between 0 and 1, reflecting the probability that a given shot will result in a goal, with higher values representing better chances (Narayanan & Pifer, 2024). For example, in penalty kicks, the xG value is 0.79.
The total xG of a team in a match is calculated by summing all the individual xG values for the chances created during the game. This cumulative number serves as a measure of the overall quality of chances created by two teams during a match (Narayanan & Pifer, 2024).

Evaluating the performance of teams only by looking at the results of the matches can on many occasions be very misleading. Often the results of matches do not reflect the dynamics on the field. This is where the concept of xG becomes very useful as it provides a more insightful understanding of a team's performance rather than just looking at the scoreboard. By focusing on shot quality, xG helps analysts distinguish between teams or players who create many high-quality chances and those who create fewer but potentially convert chances more effectively (Hewitt & Karakus, 2023). For example, a team with a high xG but few actual goals might be considered unlucky. On the other hand, a team that scores more than the xG suggests that they are overperforming or benefiting from exceptional finishing. These performance differences are valuable information for analysts and coaches when evaluating individual and team performances (Mead et al., 2023).
Over time, xG has become a standard tool for both analysts and football clubs. It offers valuable insights for tactical setups, player acquisitions, and even contract negotiations. A player's xG performance now plays an important role in assessing their contribution to the team (Mead et al., 2023).

In addition, xG allows coaches to compare expected outcomes with actual results, helping them to understand whether their teams are creating enough quality chances or whether players' finishing needs improvement (Hewitt & Karakus, 2023). While xG provides an objective measure of scoring potential, its application

and accuracy can be influenced by external factors, such as fan attendance, which can significantly alter match dynamics and performance.

2.2 Importance of Fan Attendance in Football

The presence of fans is widely recognized as a crucial external factor affecting various aspects of football matches, especially HA, player performance, and match dynamics (Wang & Qin, 2023). This section examines the impact of fans based on a number of studies, several conducted during or after the COVID-19 pandemic.

Effects on Players

HA in football is a complex concept influenced by several elements beyond the footballing side of things. These mainly include the combination of logistical, environmental, and psychological factors. Most importantly, playing in their own beloved stadium gives home teams a distinct advantage. They are more familiar with the dimensions of the pitch, surface, and other localized factors, such as even better dressing rooms for the home team can provide an extra edge. This familiarity with their surroundings allows for better preparation and execution of tactics, which can be particularly beneficial in competitive matches (Ponzo & Scoppa, 2014). Additionally, home teams benefit from maintaining their usual pre-game routines, which are often disrupted for away teams due to the need to travel to another stadium and logistical challenges. This stability helps optimize both physical and mental performance (Pollard, 2008).
Another critical element of HA is crowd dynamics. Studies have shown that crowd size is not the most significant factor in HA. Instead, it's the crowd density and pitch proximity that have a bigger impact. Dense, engaged crowds, which often mean small, compact stadiums, create an intense atmosphere that can enhance the home team's performance while putting psychological pressure on away teams (Ponzo & Scoppa, 2014). This unique dynamic often gives home teams a psychological advantage, especially in high-stakes matches.

The importance of the presence of fans becomes very clear during matches played without fans. The presence of fans increases the home team's motivation, confidence, and energy levels, creating a sense of urgency and focus. In their absence, this advantage diminishes, often leading to poorer performance (Bryson et al., 2021). According to McCarrick et al. (2021), who analyzed matches across major European leagues during the pandemic, found a significant decrease in HA, which is traditionally measured by winning percentages. Before the pandemic, home teams gained an average of 0.39 more points per game at home compared to away games. However, this changed in the 2020/21 season, when matches were played without fans and home teams gained just 0.22 more points, reflecting a nearly 50% drop in points.
In addition, home teams' key performance indicators, such as corners and shots, dropped significantly in games without spectators. This decline highlights the psychological role of fans in motivating the home team's performance and intimidating the opposition.

In addition, the absence of fans appeared to create more even playing conditions, with away teams showing better performance metrics such as increased xG and dominance (which is a standardized factor of the three HA indicators: corners, shots, and shots on goal) in games without fans. These findings suggest that fan presence plays an important role in match dynamics and outcomes.

Influence on Referee Decisions

The crowd has a known dramatic effect on the decisions of the referees, which often stems from subconscious favoritism towards the home team. Throughout the 90 minutes of a game, fans of the home team typically voice their dissent over even the most insignificant calls against their team, and through this constant pressure, the referee is kept under stress. This occurs more in the United Kingdom, where fans are notorious for their vocal and occasionally aggressive responses to refereeing.
In the 2019/20 season, Bryson et al. (2021) analyzed 6,481 matches, of which 1,498 were played without spectators, and found some significant changes in refereeing behavior when fans were not present.
Visiting teams received 0.29 fewer yellow cards, while home teams increased marginally by 0.07 cards. The difference, therefore, decreased the gap in yellow cards given to visiting teams by 0.36 ($p < 0.01$). There was also a decrease of 0.22 yellow cards in matches played behind closed doors ($p < 0.01$).
These findings suggest that referees tend to make more neutral decisions without the pressure from the home crowd. Without fans, referees did not similarly favor home teams in disciplinary actions, creating a fair competitive environment.

McCarrick et al. (2021) further observed that the absence of spectators caused a more level playing field. Visiting teams adopted a more aggressive playing style during games, including a shift to more offensive and less defensive tactics compared to pre-pandemic conditions. This natural effect of the tactical change reduced the total number of yellow cards and fouls for the away team, as passive teams are usually characterized by fewer shots on goal and corners; they commit more fouls when defending in their own half against more dominant opponents. The average number of penalties awarded to away teams also significantly increased in the top five leagues (Destefanis et al., 2022). As such, measuring the overall impact of fan attendance on officiating decisions is quite challenging.

Impact on Match Dynamics

The influence of fan absence extended beyond individual players and referees to affect overall match dynamics. McCarrick et al. (2021) observed a significant reduction in HA during matches played without fans. Home teams were less effective in capitalizing on their familiar environments, scoring fewer goals and earning fewer points compared to matches with fans. This decline highlights the critical influence of crowd presence in boosting home teams' performance through motivation and psychological support.

Similarly, Wunderlich et al. (2021) reported a decline in attacking performance metrics for home teams in fanless games, including reductions in the number of shots taken and shots on target. These findings reinforce the idea that the energy

and encouragement provided by fans are integral to driving home teams' offensive intensity.

The absence of fans also appeared to impact away teams. Wunderlich et al. (2021), found that away teams experienced reduced psychological pressure in the absence of home crowds. This reduction in pressure may have allowed away players to perform with greater composure, contributing to a more evenly contested match environment.

2.3 Finishing Efficiency and Its Determinants

FE, defined as the ratio of actual goals to xG, is a measure of how well a team or player turns goal-scoring opportunities into goals (Rathke, 2017). This metric highlights the factors driving deviations from xG predictions, such as overperformance or underperformance. Here are the main factors influencing FE.

Technical Quality of Players

The technical quality of the players is an important component of FE. Some of the attributes, such as shot precision and effective positioning on the field, are integral parts of determining a player's ability to score goals based on xG. Mead et al. (2023) highlight that including player-specific attributes, such as individual market value, improves the predictive power of xG models. This suggests that players with a higher skill level, which is often associated with a higher value, have a greater chance of successfully converting their scoring chances, thus bringing into relief the importance of individual competence in FE.

Psychological Factors

Football being the most followed sport in the world, it is natural to create immense pressure on players, especially at the highest level. Psychological factors significantly influence players' performance in high-pressure situations where some players struggle to cope with the demands. The ability to deal with the pressure can have a direct impact on FE. The constructs of confidence, mental preparation, and anxiety control are some of the crucial elements that influence a player's ability to convert scoring chances into real goals. According to Abdullah et al. (2016), variables such as motivation, concentration, and the ability to manage anxiety greatly contribute to performance variations among elite athletes in football. Thus, emphasizing the importance of psychological preparation in achieving success.

Equally important are composure and emotional management, especially in dealing with high-pressure situations like penalties. A study by Jordet et al. (2009) showed that players who rushed in preparation for their penalty shootout performed worse compared to those who took more time before shooting. The study illustrates how stress and rushed decision-making result in poor performance. Even though the research focused on penalties, its findings give valuable insights into how players behave in open play, where maintaining composure can similarly enhance FE through better decision-making.

At the team level, psychological factors such as group cohesion are just as important. Setić (2020) emphasizes that a strong sense of unity within the team creates a supportive environment, helping to reduce individual players' psychological stress and improve overall team performance. Cohesive teams tend to communicate more

effectively and work better towards shared objectives. These are directly linked with the ability to convert goal-scoring opportunities into goals.

External Influences

External factors, including refereeing decisions influenced by crowd presence, also affect FE. Destefanis et al. (2022) found that the absence of fans led to more neutral refereeing, with fewer fouls and penalties awarded to home teams. This reduction in home-biased officiating helped further improve the away team's performance and FE. For instance, in the top five leagues in Europe, away penalties per game rose from 0.11 to 0.18 in games without fans, while home penalties increased only slightly (from 0.17 to 0.20).

Another critical external factor would be the crowd. As pointed out by Pollard (2006), home supporters give the home players a psychological boost while increasing the pressure on the away players, which may affect the decisions and performances of both teams. The presence of a crowd often bolsters the home team's performance through increased confidence for those playing for the home team, while it is harder for the visiting teams to maintain their poise in a hostile environment.

Travel fatigue is among the major variables that have great impacts on FE, especially for the visiting teams. As noted by Pollard (2006), the distractions of travel may lead to physical fatigue as well as reduced cognitive sharpness, which in turn, affect the abilities of the players to remain calm and focused while trying to score. This becomes even worse in leagues that involve much travel since the extra load from long-distance travel negatively affects both the physical and mental health of the players.

Furthermore, match congestion, which is currently a hot topic due to the increasing number of games being played, also affects FE. It refers to the situation in which teams are forced to play many games in a relatively short period. This leads to increased physical and mental fatigue. Dellal et al. (2010) observed that a high volume of games degrades the performance of players, especially in tasks requiring accurate execution and concentration. Fatigue due to fixture congestion may lead directly to reduced FE, as players are unlikely to perform optimally when they are mentally and physically fatigued.

Understanding external factors such as referees, crowd presence, travel fatigue, and match congestion provides critical insights into the wider impacts on FE. These factors highlight the need to consider external variables in football performance analysis to gain a more accurate understanding of team and player capabilities.

2.4 Gaps in Research

Despite significant advancements in football over the past decade, much remains unknown about the impact of fan attendance on team performance. Metrics like xG and FE still have many aspects of their relationship with fan presence that remain surprisingly underexplored. The absence of fans during the COVID-19 pandemic offered a great opportunity to isolate these influences, yet key aspects remain largely unanswered.

One major area that has not been explored enough is the effect of fan presence on HA and its relationship with xG predictability and FE. Furthermore, little attention has been given to focusing on individual seasons since most research is done on the top European leagues across various seasons.

This study addresses these research gaps by analyzing Premier League matches over two seasons to assess the impact of fan presence on xG performance, FE, and the dynamics between home and away teams. The goal is to provide a deeper understanding of the external factors that shape football performance

3 Methodology

This chapter discusses in detail the methodological approach taken to answer the research questions and introduces the methods of data collection, processing, and analysis. The main aim is to analyze the effect of fan presence on xG and FE using three statistical methods: descriptive statistics, regression analysis, and paired t-tests.
Descriptive statistics provide an overview of key metrics, allowing for general trends and fluctuations in performance over the seasons to be understood. Regression analysis explores how fan attendance, the match location, and their interaction affect performance, leading toward a more holistic understanding of dynamics around performance (Liu et al., 2015). This also uses paired t-tests for the two data sets analyzed here to identify any statistically significant differences between them—a common procedure in sports analytics (McIntosh & Robertson, 2023). Taken together, these methods form a coherent and methodical approach to address the research questions proficiently.

3.1 Data Collection

For this study, FBref was chosen as the data source because it thoroughly covers match statistics and uses the StatsBomb xG model, a well-known and respected source of football analytics. The StatsBomb xG model considers various factors of the shot, such as position, type of shot, and defensive pressure, ensuring a reliable estimate of the probability of scoring. The data for xG may differ a little from other data sources (e.g. Opta) due to different methods in xG calculations on different models.
This study focuses on two Premier League seasons:
Season 2019/2020: 380 matches, of which 288 were played with fans and 92 were played without fans.
Season 2020/2021: 380 matches, of which 348 were without fans and 32 with limited attendance (average crowd size 5,487 for matches with limited attendance).

Each game's data across the two Premier League seasons was exported to an Excel file, where each match included the following variables: date, teams, result, attendance, home xG, and away xG. The two seasons were organized into separate Excel pages for further processing.

To maintain a clear comparison matches with and without fans were excluded from matches with limited attendance, resulting in a sample of 636 (288 with fans and 348 without). This meant that the 92 games played without fans in the 2019/20 season were excluded, so only matches with full attendance were analyzed. Similarly, 32 limited-attendance games were removed from the 2020/21 season to focus solely on matches without fans. This filtering ensured that the analysis compared the two separate conditions we were looking to analyze: matches with the fan at full capacity versus those with no fans. To facilitate the analysis, additional variables were calculated:

Total xG: the sum of home xG and away xG for each match.
Total goals: the total number of goals scored by both teams in each match.
FE: goals scored divided by xG for home and away teams.

Once the data was cleaned and organized, the data set was ready for analysis, allowing a clear comparison between the two seasons and minimizing the confounding effects of limited fan presence.

3.2 Analytical Approach

For this study, we adopted a structured three-step analytical framework to investigate the impact of fan attendance on xG performance and FE in Premier League matches. The framework includes descriptive statistics, regression analysis, and statistical comparisons, which have been used before by established methodologies in football analytics research.

3.3 Descriptive Statistics

The analysis started with descriptive statistics to summarize key performance metrics in two different situations: matches with fans from the 2019/20 season (before COVID) and matches without fans from the 2020/21 season. This preliminary step provides insights into general trends and differences in performance between these two periods. The primary variables of interest included:

Table 1. Definitions and Calculations of Key Metrics

Variable	Calculated as
Home and Away xG	Raw data from the datasets
Total xG	Home xG + Away xG
FE	Goals / xG
Total Goals	Home Goals + Away Goals

The descriptive analysis summarizes xG and goal-scoring performance for matches with and without fans, comparing variables like total xG, total goals, and FE. Metrics were calculated for each match and grouped by condition (with or without fans), providing a clear initial comparison.

The descriptive statistics in Table 1 provide an overview of key performance metrics across matches with and without fans. To illustrate some of these trends, Figure 1 visualizes the daily average FE (total goals to total xG ratio) over time. The timeline compares matches with full fan attendance (2019/2020 season) to fanless matches (2020/2021 season), highlighting variations in performance metrics across the two periods. This visualization sets the stage for further analysis of how fan presence influenced FE.

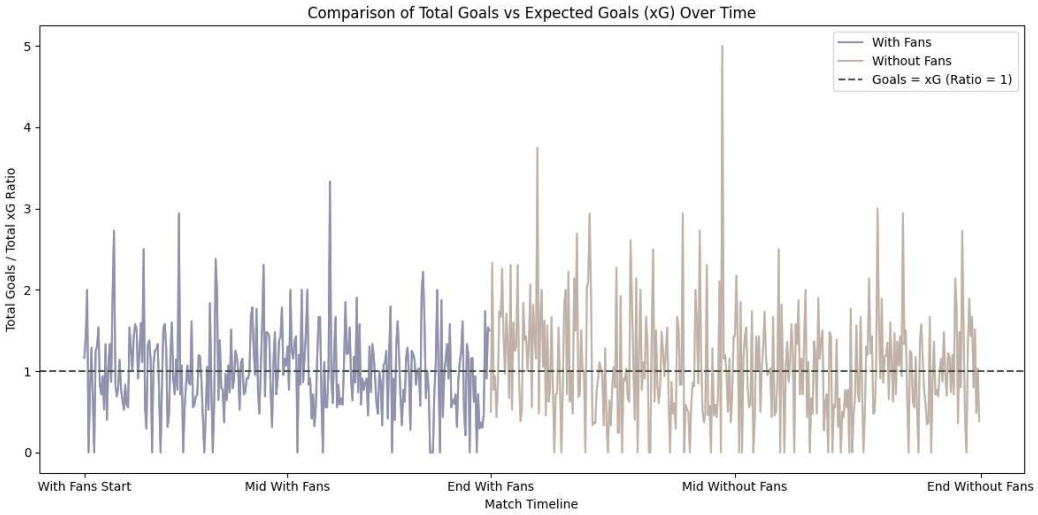

Figure 1. Trends in Total Goals and xG Over Time.

3.4 Regression Analysis

Two regression models were developed to understand the effects of fan presence on xG and FE. These models allow us to analyze the relationship between fan attendance, venue, and team performance. Regression analysis is a well-established approach in sports analytics for the solution and understanding of the relationships between complex variables.

The regression analysis focuses on FE and xG as the dependent variables—these are the outcomes under study, showing how they are affected by the coefficients. The coefficients used are: β0: basic performance when all other factors are 0 (FE and xG). β1: fan presence effect. β2 refers to the effect of the location of the match (home vs. away). β3 captures the interaction between the presence of fans and the match venue, exploring if there is an interaction effect in terms of the response of home and away teams to the presence or absence of fans. The model further includes an error term, ε, which captures the variation in FE and xG not explained by the variables mentioned above. This term includes the effects of different factors, such as the performance of the players, the tactics used, or any unexpected events in the game, thus ensuring that the analysis remains very realistic and detailed.

Each coefficient measures the extent to which the dependent variable, FE or xG, is influenced by changes in the relevant independent variable, ceteris paribus. For example, a positive β3 in the FE model (2) implies that the presence of fans affects HA by improving efficiency in converting opportunities into goals.

Regression models are designed to measure how the presence of fans and the location of matches affect team performance. Fan Presence (1 = fans present, 0 = no fans) captures the potential impact of spectators, while match location(1 = home, 0 = away) reflects the effect of playing home or away.

The Interaction Term (fan presence * match location) examines whether the effect of fans varies between home and away teams, such as amplifying HA or easing pressure on away teams in fanless matches. Coefficients (β) provide numerical estimates of these effects, offering insights into how external factors like fans and location shape team performance. The regression is run with all possible variations, such as when fans are present (Fan Presence = 1), and the match is played away (Match Location = 0). This allows us to determine how much the FE or xG changes relative to the baseline performance and assess if these results are statistically significant. Table 2 provides the variables used in the regression analysis.

Table 2. Regression Variables and Their Descriptions

Term	Description	Example values
Dependant Variable	FE and xG	xG= 1.1 , FE = 1.4
Fan Presence	Fans present or not	Binary (1 or 0)
Match Location	Team is playing Home or Away	Binary (1 or 0)
Interaction Term	Effect of fan presence on xG differs by location	Fan Presence x Match Location
Coefficients (β)	Impact of each variable on dependant variable	$\beta 1$=0.5, $\beta 2$=−0.9

The regression coefficients provide insights into how the independent variables impact team performance:

Let's say β1 (Fan Presence) is positive, this indicates that fan presence positively affects team performance overall, whereas a negative β1 suggests fan presence has a negative impact (e.g. reduced FE). A positive β2 (Match Location) reflects to HA, where teams perform better at home, and a negative β2 would imply a disadvantage when playing at home. A positive β3 (Interaction Term) would indicate that fan presence strengthens home performance. A negative β3 indicates that fan presence benefits away teams, possibly by reducing the pressure on away teams or dampening the HA.

The following regression models incorporate these coefficients to analyze the dependent variables. Model 1 (1) analyzes the impact of fan presence, match location, and their interaction on FE. On the other hand Model 2 (2) analyzes similarly the impact of fan presence, match location, and their interaction but this time on xG.

Model 1: FE

$$FE = \beta 0 + \beta 1(Fan\ Presence) + \beta 2(Match\ Location) + \beta 3(Fan\ Presence * Match\ Location) + \varepsilon \quad (1)$$

Model 2: xG

$$xG = \beta 0 + \beta 1(Fan\ Presence) + \beta 2(Match\ Location) + \beta 3(Fan\ Presence * Match\ Location) + \varepsilon \quad (2)$$

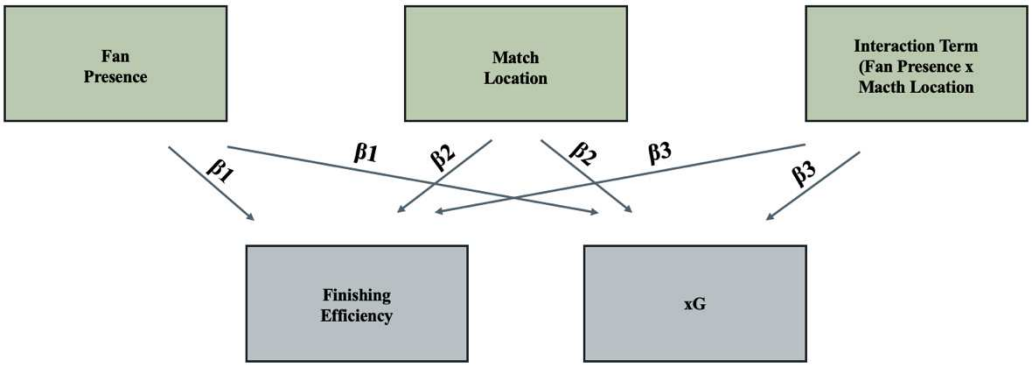

Figure 2. Regression Model Framework

3.5 Statistical Comparisons

This section uses statistical comparisons to validate the findings from the regression models and further analyze the effects of fan presence and match location on performance metrics. One key method, paired t-tests was selected. This method is also widely used in sports analytics and in many statistical studies to analyze the differences and interaction effects of different metrics. The study highlights key trends and simplifies general comparisons for clarity.
To ensure a fair comparison, the datasets with and without fans have been aligned so that only games are preserved in both conditions. In the 2019/2020 season, 288 games were played with fans, while in the 2020/2021 season, 348 games were played without fans. For comparability of the analysis, only the 288 matches from each data set that overlapped by venue are included. Equalization is achieved by using "Venue" as a common identifier, ensuring that matches played at the same venue were directly compared in both conditions.
An inner join was applied during the merging process to retain only games appearing in both datasets, creating a paired dataset suitable for statistical testing.

The paired t-test is used to compare FE and xG. This method is particularly suited to this analysis, as it evaluates whether the mean differences in these metrics are statistically significant, isolating the influence of fan absence.
The paired t-test is formulated as

$$t = (\frac{\bar{d}}{\frac{sd}{\sqrt{n}}}) \quad (3)$$

where $\bar{d}$ refers to the mean difference in the metric (Goals/xG or xG), sd refers to the standard deviation of these differences and n is the number of matches
The t-statistic in the paired t-test quantifies the difference between the means of matches with and without fans relative to the variability in the data. A larger absolute t-statistic indicates that the observed difference is substantial compared to the natural variability within the sample. Specifically:
A high t-statistic suggests a meaningful difference between the two conditions, indicating that fan presence or absence significantly influenced the metric.
A low t-statistic implies that the observed difference is small or not statistically distinguishable from random variation.
This value, when compared to a critical value from the t-distribution, determines statistical significance. A p-value below 0.05, derived from the t-statistic, indicates that the observed difference is unlikely to have occurred by random chance, confirming that fan presence meaningfully impacts performance. For this study, a p-value below 0.10 is also considered quite significant.

In summary, the paired t-test evaluates whether FE (Goals/xG) and xG differ significantly between matches with and without fans. It calculates the mean difference ($\bar{d}$) and a p-value, which indicates the likelihood that the observed difference is due to chance. A p-value below 0.05 signifies a statistically significant result, confirming that fan presence meaningfully impacts performance.

3.6 Summary and Plan for Analysis

This methodology merges descriptive statistics, regression analysis, and paired t-tests to evaluate the impact of fan attendance on xG and FE. The results will first be analyzed overall, comparing simple metrics between the two seasons. One with fans and one largely without. This will provide some sort of understanding of how fan presence influences team performance.

A thorough analysis will then compare home and away teams separately, focusing on whether teams show reduced FE without fans or if there were differences in xG generation in different match settings. Regression analysis helps quantify these effects using coefficients ($\beta 1, \beta 2, \beta 3$), which shows how fan presence and match location influence performance. For example, a significant positive or negative $\beta 3$ would indicate that fan presence has different impacts on home and away teams.

Finally, paired t-tests confirm whether the differences in FE and xG with and without fans are statistically significant. Together, these approaches ensure a thorough assessment that considers both general seasonal trends as well as home-away dynamics.

4 Results

This section presents the findings from the Premier League match analysis for the seasons 2019/2020 (with fans) and 2020/2021 (without fans). Key performance metrics, including xG, total goals, FE, and performance difference (Goals – xG) are analyzed to explore the impact of fan presence. Differences are assessed at the season level, for home and away performance, supported by descriptive statistics, regression analysis, and paired t-tests.

4.1 Descriptive Statistics

This analysis presents the differences in key metrics, including xG, total goals, FE, and performance difference between the 2019/2020 and 2020/2021 Premier League seasons. These metrics are examined at the season level, as well as by home-away performance, and at venue-level. To examine season-level differences, average xG per game, average goals per game, and metrics such as FE and performance difference were calculated across all matches in each season. Table 3 summarizes these results.

Table 3. Average Season-Level Metrics per Game

Season	Total xG	Total Goals	FE	Performance Difference
2019/2020	2,79	2,72	1,00	-0,07
2020/2021	2,58	2,68	1,06	0,10

During the 2020/2021 Premier League season, which was largely played without fans, teams created slightly fewer chances, averaging 2.58 xG per game compared to 2.79 in the 2019/2020 season, when fans were present.
This reduction in xG suggests that teams may have adopted more passive playing styles with less tempo in the absence of crowd influence.

Interestingly, despite the decrease in xG, the goals-per-game figures remained stable between the two seasons, with 2.68 goals per game in 2020/2021 compared to 2.72 in 2019/2020. This stability points to an improvement in FE, calculated as the ratio of goals to xG rose from 1.00 in the fan-attended season to 1.06 in the fanless one, indicating that players were slightly more effective at converting their chances into goals when playing in empty stadiums.

This trend is further supported by the shift in performance relative to xG. While teams underperformed slightly against their xG in 2019/2020 with a negative difference of -0.07 goals per game, they transitioned to overperforming in 2020/2021 with a positive difference of +0.10 goals per game.

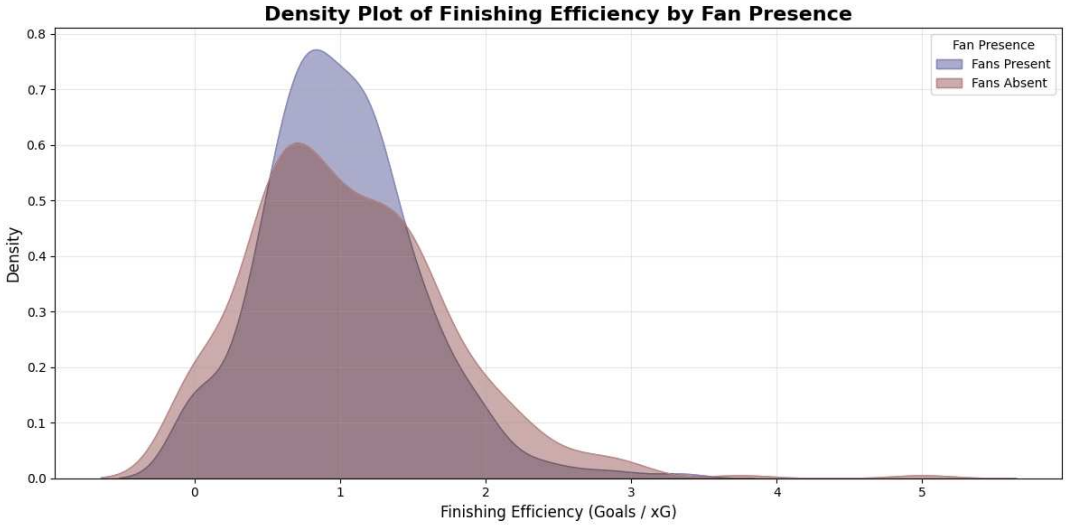

Figure 3. Distribution of FE by Fan Presence

The density plot illustrates overlapping distributions of FE for matches with and without fans. While the curves show slight shifts, with FE having greater variability in matches with fans, these differences are not statistically significant.

4.1.1 Home Team Analysis

Table 4 shows that in 2019/2020, home teams had an average xG of 1.51 and a FE of 1.09, indicating strong chance creation and conversion. In 2020/2021, home xG decreased notably to 1.34, while FE also declined slightly to 1.04, indicating a reduction in performance metrics compared to the previous season.

Table 4. Comparison of Home xG and FE Across Seasons

Season	Home xG	FE
2019/2020	1,51	1,09
2020/2021	1,34	1,04

To understand the impact of fan absence on home team performance, performance difference (average per match) was analyzed for each season. Figure 4 visualizes the distribution of home performance difference across the fan-attended and fanless season.

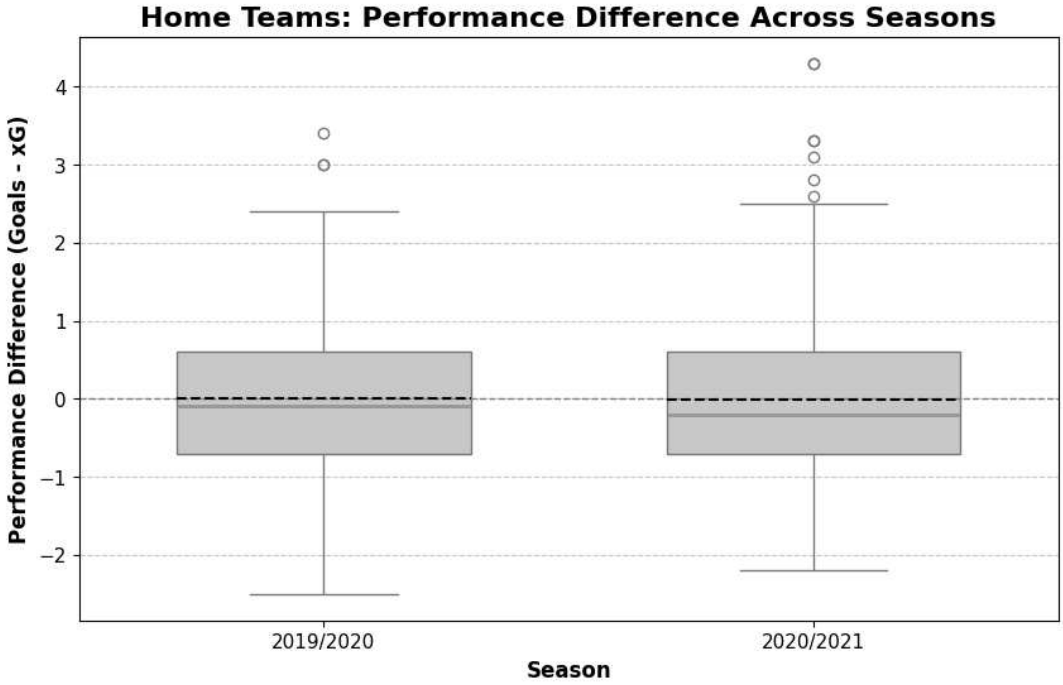

Figure 4. Boxplot of Home Team Performance Difference Across Seasons

In 2019/2020 (fan-attended), home teams displayed a relatively narrow distribution of performance difference, with most matches gathered near the value of 0 as seen from the dashed line (average). This indicates consistent finishing close to xG predictions, with only a small percentage of extreme overperformance (positive performance difference) or underperformance (negative performance difference). The variance for home teams during this season was 0.9771, indicating low variability.
In contrast, the 2020/2021 season (fanless) displayed a wider distribution of performance difference for home teams, with a little increase in variability. The median (solid line) shifted slightly lower, and several matches showed significant underperformance, while others indicated extreme overperformance. The variance for home teams increased to 1.1135, indicating greater inconsistency in FE without fan attendance. This suggests that the absence of fans led to more outcomes diverging considerably from the xG.

To further examine how fan absence influenced home team performance, venue-level performance differences (Goals - xG) were analyzed for the 2019/2020 and 2020/2021 seasons. Figure 5 presents a heatmap illustrating the performance difference across various Premier League stadiums. This venue-specific analysis clarifies how individual venues experienced changes in home team performance dynamics under fan-attended and fanless conditions.

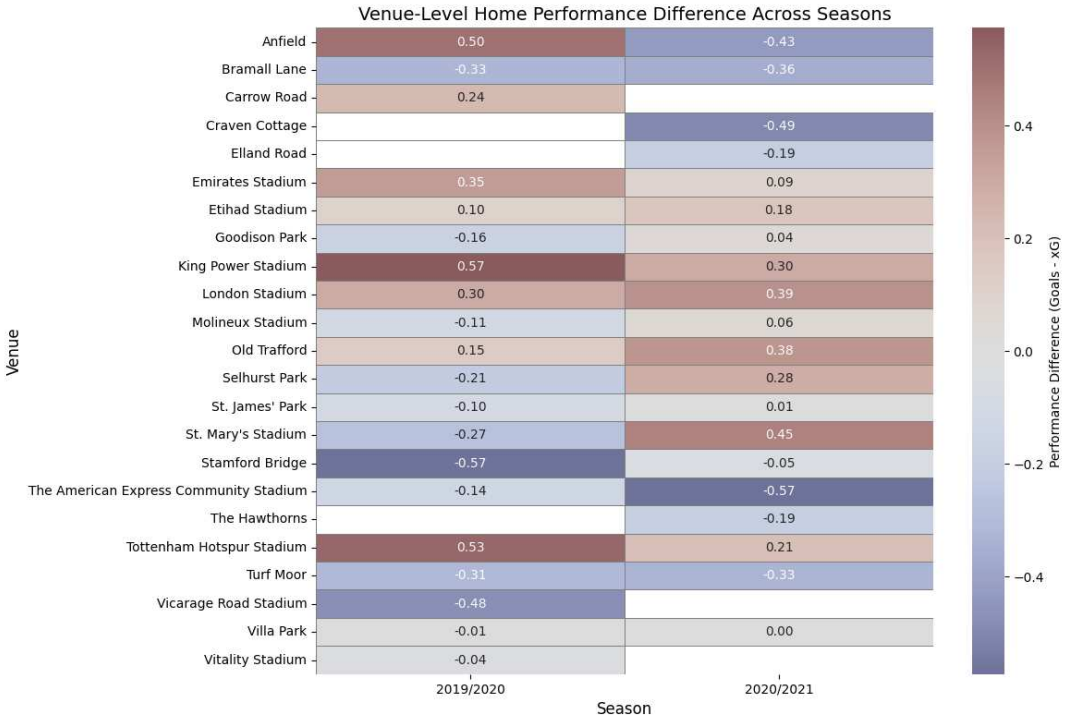

Figure 5. Venue-Level Home Performance Difference Across Seasons

Figure 5 highlights variations in performance difference across venues between the two seasons. In the fan-attended 2019/2020 season, stadiums like King Power Stadium (0.57) and Anfield (0.50) showed strong positive performance differences, indicating overperformance relative to xG, while Stamford Bridge (-0.57) and Bramall Lane (-0.33) showed notable underperformance. This shows the varied influence of HA under normal conditions.

In the fanless 2020/2021 season, significant shifts occurred at several venues. Anfield dropped from a positive 0.50 to a negative -0.43, demonstrating a big decrease in FE without fans. In contrast, St. Mary's Stadium improved from -0.27 in 2019/2020 to 0.45 in 2020/2021, suggesting some teams adapted better to fanless conditions.

These findings reveal the diverse effects of fan absence on home team performance. Further research could explore the factors behind these home team changes, such as previous years' data on performance difference, the impact of empty stadiums on team tactics, or environmental influences, to understand the dynamics of HA in football better.

4.1.2 Away Team Analysis

The table highlights that away teams recorded a slight decline in average xG from 1.28 in 2019/2020 to 1.24 in 2020/2021, indicating a reduction in the quality of chances created. However, FE increased from 1.09 in 2019/2020 to 1.18 in 2020/2021, suggesting improved conversion rates despite the lower xG values.

Table 5. Away teams: Average xG and FE Across Seasons

Season	Away xG	FE
2019/2020	1,28	1,09
2020/2021	1,24	1,18

To further understand the impact of fan absence on away team performance, performance difference (average per match) was analyzed for each season. Figure 6 visualizes the distribution of away performance difference across the fan-attended and fanless season.

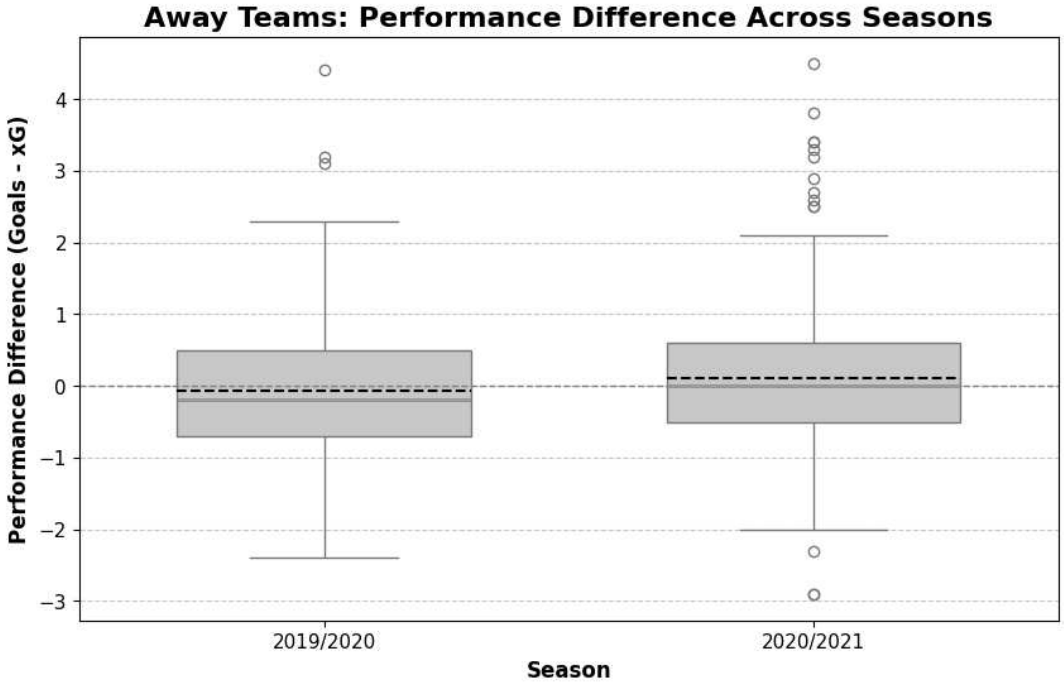

Figure 6. Boxplot of Away Team Performance Difference Across Seasons

Figure 6 illustrates the distribution of performance difference (Goals - xG) for away teams across both seasons. In 2019/2020 (fan-attended), away teams showed a narrow distribution around a median slightly below zero, reflecting underperformance relative to xG. The variance for away teams during this season was 0.9418, indicating relatively low variability in performance outcomes.

In contrast, 2020/2021 (fanless) displays a shift toward overperformance, with the median slightly above zero but also greater variability (variance of 1.0495), indicating more matches where teams significantly exceeded or fell short of xG predictions.

4.1.3 Home vs Away Analysis

xG overachievement reflects FE when teams score more goals than predicted by xG. HA, often influenced by fan presence, typically boosts home team performance. This section examines how xG overachievement differed between home and away teams

in 2019/2020 and 2020/2021 demonstrating fan absence's impact on performance dynamics. The bar chart in Figure 7 compares the number of teams overachieving xG at home and away across both seasons:

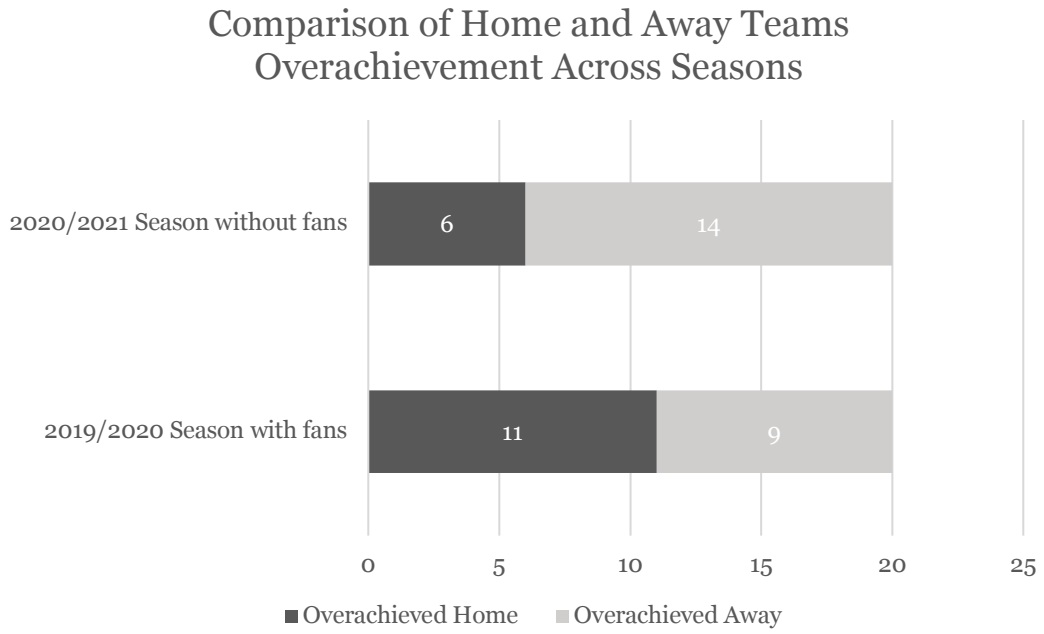

Figure 7. Comparison of Home and Away Teams' Overachievement Across Seasons

In the 2019/2020 Premier League season out of the total 20 teams, 11 teams more frequently overperformed their xG (goals/xG > 1) at home compared to 9 away, reflecting stronger FE at home.
In contrast, during the season without fans, only 6 teams overperformed at home, while 14 did so away. The absence of fans appears to have diminished the HA, allowing away teams to display improved FE.

Further analysis of xG overachievement percentages highlights this trend. In 2019/2020, 47.46% of matches saw home teams overachieve their xG compared to 41.43% for away teams. In 2020/2021, the trend reversed, with away teams overachieving in 50.23% of matches, surpassing the 42.89% recorded for home teams. These findings illustrate a reversal in performance dynamics, with away teams benefiting more from the absence of fans than home teams, disrupting the traditional HA.

Table 6. Percentage of xG Overachievement in Home and Away Matches Across Different Conditions

Match setting	% of xG Overachievement
Home (Fans)	47,46
Away (Fans)	41,43
Home (No Fans)	42,89
Away (No Fans)	50,23

4.2 Regression Analysis

To assess the impact of fan attendance and match location on team performance, two regression models were developed. Model 1 (1) focuses on FE (Goals/xG) and Model 2 (2) on xG. This approach allowed for a thorough evaluation of how fan presence influenced these metrics across different contexts. The results of these methods are summarized in Table 7 below.

Table 7. Regression Results for FE and xG Models

Variable	FE (Model 1)	xG (Model 2)
Intercept (β0)	1,143	2,575
Fan Presence (β1)	-0,035 (p = 0.682)	0,382 (p = 0.003***)
Match Location (β2)	-0,146 (p = 0.032**)	0.008 (p = 0.940)
Interaction Term (β1 **x** β2)	-0.006 (p = 0.958)	-0.249 (p = 0.112)
R-squared	0.015	0.019
Observations	636	636

*Significance levels: ***p < 0.01; **p < 0.05; *p < 0.10.

Fan Presence

The regression results indicated that fan presence had no statistically significant impact on FE (β = -0.035, p = 0.682). This suggests that teams converted their xG into actual goals at quite similar rates regardless of whether matches were played with or without fans. However, fan presence significantly increased total xG (β = 0.382, p = 0.003), indicating that teams created better or more frequent chances when fans were present. This finding highlights the motivational or tactical influence of fans in increasing offensive actions, particularly for home teams, even though it does not translate to a measurable improvement in FE.

Match Location (Home vs. Away)

The regression results for match location (β = -0.146, p = 0.032) revealed a significant effect on FE, with away teams performing better in converting xG into goals than home teams. This finding highlights the impact of fan presence in HA and particularly FE, suggesting that away teams may adapt relatively better to crowdless environments, as seen during the 2020/2021 season. Conversely, match location had no significant impact on a team's xG (β = 0.008, p = 0.940), indicating that the quality or quantity of chances created was not strongly affected by whether a match was played at home or away.

Interaction Between Fan Presence and Match Location

The interaction term (β = -0.0055, p = 0.958 for FE; β = -0.249, p = 0.112 for xG) failed to reach statistical significance in both models. This suggests that the combined effects of fan presence and match location, such as amplifying HA or decreasing pressure on away teams, did not have a measurable impact on either FE or xG in this dataset. These findings indicate that fan presence and match location more likely influence these metrics independently rather than in a combined or interactive manner.

Explanatory Power of the Models

Both models had low explanatory power, with R^2 values of 0.015 (FE) and 0.019 (xG). This indicates that only a small portion of the variation in these metrics is explained by fan presence, match location, and their interaction. These results highlight the need to consider additional factors, such as tactical strategies, player attributes, or team quality, which may substantially shape FE and xG generation.

4.3 Paired t-tests

The results of the paired t-tests are summarized in Table 8 below. These t-tests compare key performance metrics between the 2019/2020 season and the 2020/2021 season to assess the impact of fan presence on team performance. Significant differences were observed for total xG and home xG, indicating that fan presence influenced the quality or frequency of chances created. However, there were no significant differences in total goals or FE, suggesting that the overall chance creation and the ability of teams to finish these chances remained quite unaffected.

Table 8. Paired T-test Results Comparing Performance Metrics Between Fan-attended and Fanless Seasons (Δ = Mean Difference: With Fans - Without Fans).

Metric	t-Statistic	p-value	Mean difference (Δ)	Result
Total Goals	0.4463	0.6557	0.0627	No Significant Difference
Total xG	3.3676	0.0009***	0.2422	Significant Difference
Home FE	0.3107	0.7562	0.0309	No Significant Difference
Away FE	-0.2811	0.7788	-0.0352	No Significant Difference
Home xG	2.8508	0.0047***	0.1902	Significant Difference
Away xG	0.8268	0.4091	0.0519	No Significant Difference

*Significance levels: ***p < 0.01; **p < 0.05; p < 0.10.

The paired t-tests revealed clear insights into how fan presence influenced performance metrics. The most notable results were observed in total xG ($p = 0.0009$) and home xG ($p = 0.0047$), which were significantly higher in fan-attended matches. This suggests that teams, particularly home teams, created better quality or more frequent chances when fans were present.

The differences between the two conditions, with and without fans, are quantified using the delta symbol (Δ). Δ represents the mean difference, calculated as with fans - without fans. For example, $\Delta = 0.2422$ for total xG indicates that teams generated, on average, 0.242 more xG when fans were present compared to matches played without fans. Similarly, $\Delta = 0.1902$ for home xG shows a meaningful increase in chance creation for home teams in fan-attended matches.

However, total goals ($p = 0.6557$) and FE for both home and away teams ($p > 0.75$) showed no significant differences, indicating that the ability to convert chances into goals was not affected by fan presence. Similarly, away xG ($p = 0.4091$) remained stable, suggesting that away teams' chance creation did not change meaningfully between the two seasons.

5 Conclusions

The comparative analysis of the Premier League's fan-attended (2019/2020) and largely fanless (2020/2021) seasons provides critical insights into how fan attendance influences football performance. The main focus of the study is the role of xG as a predictive metric and the overall performance metrics such as FE in home and away contexts.

This study reveals several interconnected findings. First, while the predictive accuracy of xG remained quite consistent across both seasons, fan presence significantly shaped xG generation. Matches with fans showed higher xG values, particularly for home teams, suggesting that fan presence boosts home teams' performances. Conversely, fanless matches showed reduced xG values, pointing to a shift in team performances, likely toward more even matches played with home teams not dominating games as the season before or diminished motivation in the absence of the home crowds.

Second, the absence of fans disrupted traditional HA, creating a more balanced playing field. Home teams, which are traditionally boosted by fan support, experienced greater variability in performance during the fanless season. Some matches demonstrated significant overperformance relative to xG, while others displayed underperformance, highlighting an inconsistency in home team FE. In contrast, away teams converted their chances into goals more effectively during fanless matches. They also shifted from slight underperformance relative to xG in fan-attended matches to overperforming in over 50% of matches in the fanless season.

Third, statistical analysis confirmed that fan absence significantly impacted total xG and home xG, with the fan-attended season showing statistically considerably higher values for both variables compared to the fanless season. However, total goals and FE remained quite stable, indicating that while the quantity and quality of chances created varied, the ability to convert these chances into goals was somewhat unaffected by fan presence. Notably, away teams outperformed their previous season's results during the fanless season, while home teams performed worse. This finding still highlights the resilience of xG as a predictive model, even under unusual conditions, while also indicating the shifting dynamics of home and away performances in the absence of fans.

Lastly, venue-level analysis revealed that the impact of HA varied a lot, with certain stadiums (e.g. Anfield) experiencing a complete loss of advantage, while some teams adapted better to fanless conditions. This highlights the complexity of psychological pressures and environmental factors on football performance. Future research could delve deeper into this topic.

These results emphasize the role of fan attendance in shaping match dynamics, particularly by boosting HA and influencing team psychology. The absence of fans diminished these effects, leveling the playing field and enabling a more neutral

environment for away teams. Yet, even after these shifts, the xG model showed great consistency, reinforcing its value as a tool for understanding football performance.

In summary, this thesis addresses the research questions by uncovering the effects of fan attendance on xG generation, FE, and the balance between home and away performances. It advances the understanding of crowd dynamics in the English Premier League, providing theoretical insights for football analytics.

5.1 Limitations

Despite its contributions, this study has several limitations that must be acknowledged. First, the analysis is limited to two Premier League seasons, restricting its applicability to other leagues or competitions that may experience varying crowd effects. Second, this study relies on team-level performance metrics such as xG and FE, which do not account for individual players' performance. Players may react differently to the presence or absence of fans, but these differences have not been analyzed in this study. Additionally, a packed schedule, COVID-19 protocols, and other variables during this abnormal era likely affected performance in the 2020/2021 season. While the absence of fans is the focus, ignoring its effects from these other factors remains a challenge. Regarding the study itself, descriptive statistics, paired t-tests, and regression analyses provide valuable insights but at the same time, they offer a rather broad-level perspective that can miss complex interactions between variables such as HA, fan presence, and tactical adjustments.

5.2 Recommendations for Future Research

To address these limitations and deepen our understanding of fan attendance in football, several directions for future research are recommended. Incorporating data from leagues with different crowd cultures (e.g. Bundesliga, Serie A) and competitions of varying significance (e.g. international tournaments) could provide insights into how fan presence influences performance in other contexts. To emphasize individual performances, examining player data, such as shot accuracy and FE, could provide valuable insights. This approach would help reveal how the presence or absence of a crowd impacts individual players and identify those who handle pressure well. Additionally, individual stadiums could also be analyzed further to determine which teams benefit most from HA.
To delve even deeper into player motivation and pressure, interviews with players, coaches, and referees could complement these quantitative results. It would provide interesting subjective perspectives on how the presence of fans actually affected motivation and decision-making. Advanced methods, such as machine learning models or network analysis, could also be used to capture complex interactions between variables, offering a deeper understanding of the varying effects of fan presence.
By analyzing these areas, future research may reveal deeper insights into the role of fan presence in shaping football performance. Such findings would provide meaningful contributions to coaches, analysts, and the football world.

6 References

Abdullah, M. R., Musa, R. M., Maliki, A. B. H. M. B., Kosni, N. A., & Suppiah, P. K. (2016). Role of psychological factors on the performance of elite soccer players. Journal of Physical Education and Sport, 16(1).

Bilalić, M., Graf, M., & Vaci, N. (2023). The effect of COVID-19 on home advantage in high- and low-stake situations: Evidence from the European national football competitions. Psychology of Sport and Exercise, 69. https://doi.org/10.1016/j.psychsport.2023.102492

Bilalić, M., Gula, B., & Vaci, N. (2021). Home advantage mediated (HAM) by referee bias and team performance during covid. Scientific Reports, 11(1). https://doi.org/10.1038/s41598-021-00784-8

Bond, A. J., Cockayne, D., Ludvigsen, J. A. L., Maguire, K., Parnell, D., Plumley, D., Widdop, P., & Wilson, R. (2022). COVID-19: the return of football fans. In Managing Sport and Leisure (Vol. 27, Issues 1–2). https://doi.org/10.1080/23750472.2020.1841449

Bryson, A., Dolton, P., Reade, J. J., Schreyer, D., & Singleton, C. (2021). Causal effects of an absent crowd on performances and refereeing decisions during Covid-19. Economics Letters, 198. https://doi.org/10.1016/j.econlet.2020.109664

Dellal, A., Wong, D. P., Moalla, W., & Chamari, K. (2010). Physical and technical activity of soccer players in the French first league- with special reference to their playing position. International SportMed Journal, 11(2).

Destefanis, S., Addesa, F., & Rossi, G. (2022). The impact of COVID-19 on home advantage: a conditional order-m analysis of football clubs' efficiency in the top-5 European leagues. Applied Economics, 54(58). https://doi.org/10.1080/00036846.2022.2074361

Hewitt, J. H., & Karakuş, O. (2023). A machine learning approach for player and position adjusted expected goals in football (soccer). Franklin Open, 4. https://doi.org/10.1016/j.fraope.2023.100034

Jordet, G., Hartman, E., & Sigmundstad, E. (2009). Temporal links to performing under pressure in international soccer penalty shootouts. Psychology of Sport and Exercise, 10(6). https://doi.org/10.1016/j.psychsport.2009.03.004

Liu, H., Hopkins, W. G., & Gómez, M. A. (2016). Modelling relationships between match events and match outcome in elite football. European Journal of Sport Science, 16(5). https://doi.org/10.1080/17461391.2015.1042527

McCarrick, D., Bilalic, M., Neave, N., & Wolfson, S. (2021). Home advantage during the COVID-19 pandemic: Analyses of European football leagues. Psychology of Sport and Exercise, 56. https://doi.org/10.1016/j.psychsport.2021.102013

McIntosh, S., & Robertson, S. (2023). Relationships between contract status and player performance in the Australian Football League. Journal of Sports Sciences, 41(2). https://doi.org/10.1080/02640414.2023.2190564

Mead, J., O'Hare, A., & McMenemy, P. (2023). Expected goals in football: Improving model performance and demonstrating value. PLoS ONE, 18(4 April). https://doi.org/10.1371/journal.pone.0282295

Narayanan, S., & David Pifer, N. (2024). An xG of Their Own: Using Expected goals to Explore the Analytical Shortcomings of Misapplied Gender Schemas in Football. Journal of Sport Management, 38(2). https://doi.org/10.1123/jsm.2023-0022

Pollard, R. (2008). Home Advantage in Football: A Current Review of an Unsolved Puzzle. The Open Sports Sciences Journal, 1(1). https://doi.org/10.2174/1875399x00801010012

Pollard, R., & Gómez, M. A. (2014). Comparison of home advantage in men's and women's football leagues in Europe. European Journal of Sport Science, 14(SUPPL.1). https://doi.org/10.1080/17461391.2011.651490

Pollard, R., & Pollard, G. (2005). Home advantage in soccer. A review of its existence and causes. International Journal of Soccer and Science, 3(1).

Ponzo, M., & Scoppa, V. (2018). Does the Home Advantage Depend on Crowd Support? Evidence From Same-Stadium Derbies. Journal of Sports Economics, 19(4). https://doi.org/10.1177/1527002516665794
Quantifying_the_bullwhip_effec. (n.d.).

Rathke, A. (2017). An examination of expected goals and shot efficiency in soccer. Journal of Human Sport and Exercise, 12(Proc2). https://doi.org/10.14198/jhse.2017.12.proc2.05

Seçkin, A., & Pollard, R. (2008). Home advantage in turkish professional soccer. Perceptual and Motor Skills, 107(1). https://doi.org/10.2466/PMS.107.1.51-54
Šetić, R. (2020). CONTRIBUTION OF PSYCHOLOGICAL CHARACTERISTICS TO TEAM SUCCESS IN FOOTBALL. Homo Sporticus, 1.

Wang, S., & Qin, Y. (2023). The impact of crowd effects on home advantage of football matches during the COVID-19 pandemic—A systematic review. PLoS ONE, 18(11 November). https://doi.org/10.1371/journal.pone.0289899

Wunderlich, F., Weigelt, M., Rein, R., & Memmert, D. (2021). How does spectator presence affect football? Home advantage remains in European topclass football matches played without spectators during the COVID-19 pandemic. PLoS ONE, 16(3 March 2021). https://doi.org/10.1371/journal.pone.0248590