NOVA IMS Information Management School

MGI

Master Degree Program in
Information Management

Unveiling the Relationship Between Fan Sentiment and Player Performance

Assessing Players, Key Factors, and Field Position Biases

Afonso de Sousa Gomes Serra Ramalho

Project Work

presented as partial requirement for obtaining the Master Degree Program in Information Management

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

NOVA Information Management School
Instituto Superior de Estatística e Gestão de Informação
Universidade Nova de Lisboa

UNVEILING THE DYNAMIC RELATIONSHIP BETWEEN FAN SENTIMENT AND FOOTBALL PLAYER PERFORMANCE

By

Afonso de Sousa Gomes Serra Ramalho

Project Work presented as partial requirement for obtaining the Master’s degree in Information Management, with a specialization in Business Intelligence and Knowledge Management.

Supervisor/Orientador(a): Professor Doutor Flávio Luís Portas Pinheiro

November 2023

STATEMENT OF INTEGRITY

I hereby declare having conducted this academic work with integrity. I confirm that I have not used plagiarism or any form of undue use of information or falsification of results along the process leading to its elaboration. I further declare that I have fully acknowledge the Rules of Conduct and Code of Honor from the NOVA Information Management School.

Lisboa, 8th November 2023

ACKNOWLEDGEMENTS

Two types of acknowledgments are due here.

The first, explicit, goes to my supervisor. Professor Flávio, thank you for believing in this work from the start, and for your assistance in shaping, reshaping, and refining my ideas, ultimately helping me make the most sense of them. Thank you for letting me forge my own path, for the support, and feedback.

The second, implicit, is to all the people that supported me in these past years, not only academically but also personally. To my friends, family, and girlfriend Matilde, my deepest gratitude. Thank you for continually showering me with your patience, wisdom, love, and support all the times.

Obrigado a todos,

Afonso

ABSTRACT

In today's dynamic sports landscape, characterized by an abundance of data and the influence of social media, we venture into a comprehensive exploration of the complex connection between player performance and fan sentiment. Leveraging the data generated by fan sentiment expressed on Twitter, we seek to determine patterns that reflect how player performance is perceived by fans, contributing with fresh perspectives to the field of sports analytics.

In this study, we developed and implemented a method for gathering Twitter data related to the 2022/2023 Premier League season, focusing on identifying the sentiments expressed in the opinions shared. Additionally, we created a player identification algorithm based on Levenshtein distance, enabling us to link specific tweets to individual players and subsequently calculate the sentiment ratio for each player in each game. Using this extensive dataset, we analyzed correlation studies between player performance and expert ratings from Fotmob and BBC. Furthermore, we employed multiple linear regression to identify the primary factors influencing fan sentiment, shedding light on what drives their assessments. Ultimately, we unveiled the relationship between players' on-field positions and the sentiments expressed by fans.

The outcomes of this study reveal that fan sentiment, following the Wisdom of Crowds principle, can serve as a meaningful indicator of player performance. Moreover, we have identified the key determinants influencing fan responses to player performances. This innovative approach is a valuable tool for future player performance assessments, emphasizing the significance of considering fan opinions in this domain.

KEYWORDS

Football; Sports Analytics; Text Mining; Sentiment Analysis; Performance Analysis;

INDEX

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

With the advancements in analytics and data-tracking technologies, performance analysis has become more data-driven than ever. This approach is now relevant not only for evaluating player performance but also for assessing the effectiveness of employees within a company. Nonetheless, the evaluation of individual performance remains susceptible to subjectivity. Indeed, people are not inhibited from expressing their opinions, often reflecting the momentary feelings related to specific events. Opinions that are often shared on social media platforms such as Facebook, Instagram, or Twitter. Football, the most popular sport worldwide, with approximately 3.5 billion fans and 265 million players, is central to this dynamic. For example, during the FIFA World Cup 2018, 115 billion tweets were people's reactions to what occurred during the competition, such as goals scoring, player injuries, or predicting the outcome of the next matches (Aloufi & El Saddik, 2018).

The regular non-expert viewer has become more critical when it comes to evaluating athletes (Su et al., 2020), and this newfound power has elevated fan voices to a level where they can influence public perceptions, player reputations, and even club decisions (Anstead & O'Loughlin, 2015).

Football and other sports now operate as a cycle of relationships between various entities. Experts, through their professional expertise and access to inside information, provide detailed analysis, match reports, and player interviews, shaping the narrative around the sports (Price et al., 2013). On the other hand, fans, fuelled by their passion and love for the game, actively engage in discussions, share opinions, and express emotions on various online platforms (Yu & Wang, 2015). This shift in the football landscape, from an expert-only opinion space to a public and decentralized one, has increased recognition of the importance of considering fans' opinions alongside experts' ratings in player performance evaluation.

Numerous studies have centered their attention on Sentiment Identification within Football-Specific Tweets (Aloufi & El Saddik, 2018), exploring the impact of game events on sentiment (Barnaghi et al., 2016; Jai-Andaloussi et al., 2015; Lanagan & Smeaton, 2011), and extracting pertinent topics from these tweets (Hidayatullah et al., 2018). Others evaluated the impact of social networks, particularly Twitter, on measuring in-game athlete performance (Gruettner et al., 2020) as a way to access the usage of social networks and the pre-match mood of professional athletes. Field (Li et al., 2022) concentrated on analyzing the correlation between statistical performance and sentiment in recent basketball research. Meanwhile, (Siemon & Wessels, 2022) employed fan sentiment to forecast performance.

Nevertheless, few works have tried to look at the relationship between the assessed performance of a football player, as evaluated by experts, and the quantification of fans' sentiments towards the same player. Can fans' opinions be utilized as a metric for evaluating player performance accurately? And if so, do they exhibit a greater inclination to assess a player's performance based on impactful events rather than relying on routine performance metrics? Or does a sentiment bias exist in fans' assessments of players that can be attributed to the players' field positions? Following the principle of Wisdom of the Crowds (Surowiecki, 2005) we should strive for improved quantification by considering the emotions expressed by populations of fans. Furthermore, if fan opinion is indeed a good proxy for performance, then we have the advantage of being able to monitor players in real time and in greater detail via social media sections. Consequently, it is crucial to determine whether sentiment expressed in tweets, analyzed using NLP Sentiment Analysis techniques, can serve as a valuable performance assessment tool for players.

To bridge the existing gap between traditional performance assessment approaches and the evaluations made by fans on social media platforms, we focus on the English Premier League's 2022/2023 season. Our dataset comprises 1.6 million tweets reacting to 135 matches, enabling us to uncover the underlying patterns, trends, and potential biases that shape public perceptions and evaluations.

This exploration seeks to understand the potential relationship between fan sentiment and the performance of football players, in line with Hypothesis 1. Moreover, we aim to ascertain the extent to which fans' evaluations of a player's performance are influenced by significant events, such as impactful events, and routine performance metrics, as suggested by Hypothesis 1b. We will also examine whether sentiment biases in fan assessments vary depending on the field position of the players, as proposed in Hypothesis 2.

This document will be structured into four additional chapters. Chapter 2 will review existing techniques, identify research gaps, and frame the topic. Chapter 3 will present approaches to address limitations and define the problem in more detail. Chapter 4 will analyse and explain the collected data within the research context. Finally, Chapter 5 will summarize key findings, suggest future steps based on limitations, and conclude the study.

2. LITERATURE REVIEW

2.1. Twitter

Twitter is a popular microblogging service (Kwak et al., 2010) launched in 2006. It allows users to share their thoughts and opinions on various topics within a character limit, initially set at 140 but later increased to 280. This change made it easier for people to express themselves and be more accessible on Twitter.

Due to its nature as a short-form communication platform, tweets tend to showcase individuals' opinions and emotions (Roberts et al., 2012). This is particularly evident when specific events act as triggers, prompting users to utilize Twitter to express their sentiments toward these events (Weng & Lee, 2011).

The real-time nature of Twitter allows emotions to be captured and shared instantly, creating a unique environment for studying user reactions (Wang et al., 2012). Researchers have utilized sentiment analysis techniques to understand users' opinions on a given subject (Barnaghi et al., 2016; Larsen et al., 2015; Naskar et al., 2020; Xue et al., 2020) and analyze the emotional tone of tweets (Errasti et al., 2017; Roberts et al., 2012; Thelwall et al., 2012), enabling a better understanding of the collective sentiment surrounding particular events or topics. This analysis helps identify patterns, trends, and shifts in emotional responses over time.

As events unfold, Twitter can serve as a real-time event detector, capturing the pulse of ongoing occurrences through user contributions and providing a live stream of updates. Past studies suggest that Twitter has even been used to detect earthquakes with a remarkable accuracy of 96% by analyzing tweets related to these natural disasters (Sakaki et al., 2010). This allows for faster alert systems compared to traditional methods. Other investigators provided frameworks for the real-time Twitter event analysis (Gaglio et al., 2016; M. Hasan et al., 2019). Evaluation results showed a significant improvement in recall and precision over five state-of-the-art baselines, highlighting the importance of event detection in social media and the need for more efficient and accurate real-time detection methos.

Overall, by studying the emotions expressed by Twitter users, it is possible to obtain insights into individuals’ online behavior (Acheampong et al., 2020). The next section will explore how these insights can be applied to the sports context by integrating event detection and sentiment analysis techniques. In the final chapter, we will discuss how this information can be used to evaluate athletes' performance on this social media platform.

2.2. Twitter in Sports

Twitter has emerged as a dynamic and indispensable component of the sports ecosystem, transforming how we engage with sports (Vale & Fernandes, 2018). From players and clubs leveraging the platform to interact with fans and promote their brands, to journalists and commentators sharing real-time news and analysis, Twitter has revolutionized the sports landscape (Hull & Lewis, 2014). Building upon this paradigm shift, this chapter investigates the multifaceted role of Twitter in sports, particularly its utilization in sentiment analysis, event detection, and harnessing the wisdom of the crowd principle.

As the influence of social media on sports continues to grow, researchers have explored fans' shifting sentiments and emotional responses during major tournaments like the World Cup and UEFA Champions League (Lucas et al., 2017). These studies have provided valuable insights into how fans react to emotionally charged events, their susceptibility to contextual factors, and the evolving nature of their emotions over time. For instance, research has shown that goals in sports competitions often generate a division of sentiment, with positive sentiments directed toward the scoring team and negative sentiments towards the opponent (Aloufi et al., 2018). Additionally, studies examining Japanese professional baseball games have revealed that Twitter captures immediate reactions to positive and negative events, leading to a burst of tweets. At the same time, subsequent retweets tend to be associated more with positive events (Takeichi et al., 2014). These findings highlight Twitter's role as a platform for fans to express excitement, support, and disappointment and engage in friendly banter with rival team supporters.

The potential of sentiment analysis for forecasting sports outcomes has garnered significant attention. Notably, (Schumaker et al., 2016) employed Twitter sentiment analysis to predict Premier League wins. Their analysis of 20 Premier League clubs demonstrated that tweet sentiment outperformed wagering on odds favorites, leading to higher payout returns but lower accuracy. This suggests that crowdsourced sentiment can serve as a more effective predictor of match outcomes than crowdsourced odds (Wunderlich & Memmert, 2021). In a different approach, (Miranda-Peña et al., 2021) employed a graph theory approach to analyze sentiment in 3,000 tweets during the last game week of the English Premier League 2019/2020 season, successfully predicting the probability of winning games by evaluating network centrality. These studies collectively indicate the potential of social media, such as Twitter, in predicting sports outcomes.

Twitter has also been extensively studied for its ability to identify and tag important events in live sports, even surpassing more complex audio-visual content analysis in accuracy and computational efficiency (Barnaghi et al., 2016; Jai-Andaloussi et al., 2015; Lanagan & Smeaton, 2011; Oorschot et al., 2012; Ruz et al., 2020). Remarkably, Twitter has proven capable of accurately detecting the most crucial events of a match, with the end of matches garnering higher tweet activity and exceptional events, such as remarkable shots or tackles, often prompting discussions during halftime. Leveraging publicly available tweets about sports events has the potential to aid in event detection and summary generation for associated media.

These phenomena collectively emphasize the power of crowdsourcing and are a clear example of the well-known phenomena of the wisdom of the crowds in sports-related contexts (Surowiecki, 2005). The collective knowledge and opinions of diverse groups often outperform even the most expert individuals in predicting game outcomes and evaluating players (Mannes et al., 2014). Research has increasingly focused on harnessing this power to predict match outcomes, with impressive accuracy achieved in predicting World Cup 2018 matches using Twitter data collected from fans (Talha et al., 2018). Furthermore, investigations into the market evaluation of players have demonstrated that crowdsourcing can provide insights into unmeasured, unreported, or uncertain outcomes of interest (Coates & Parshakov, 2022; Herm et al., 2014; T. Peeters, 2018). Similar studies conducted in other sports, such as the NFL, have shown that large volumes of tweets can equal or surpass the accuracy of more traditional statistical methods in assessing game statistics (Sinha et al., 2013).

Building upon the multifaceted role of Twitter in the sports ecosystem, the next chapter of this dissertation delves into the evaluation of players' performance and the utilization of Twitter data in this regard.

2.3. PERFORMANCE EVALUATION

In sports, "performance indicator" refers to an action variable defining a successful performance (Hughes & Bartlett, 2002). Performance can be successful for multiple reasons, and many factors influence match results based on individual performance. Players may exhibit lower performance in primary statistics, such as goals, shots, or tackles, yet their contributions to the team can be valuable when considering the contextual dynamics of each match.

A fair number of studies suggests that not only objective facts determine the quantifiable performance of an athlete (Carmichael et al., 2021; Lopes Dos Santos et al., 2020; Saw et al., 2016). The physiological (Reilly et al., 2008), psychological (Rice et al., 2016), and external or strategical (H. Liu et al., 2016) play a considerable part, indirectly, in the quantification of goals, assists or tackles.

However, while it is true that evaluating an athlete's performance goes beyond simple objective metrics, it is important to note that direct statistics still provide a valuable and precise overview of an athlete's in-game quality. These statistics capture measurable actions and outcomes, allowing for a quantitative assessment of an athlete's performance. As such, some literature has been produced to evaluate players' contribution to team success, as evaluated in Duch et al. (2010), where an alternative for quantifying individual performance in team success is presented. It stated that there is a correlation between the best individual performances with the best collective performances in the tournament. Reinforcing this point, Robertson et al. (2016), with the support of Castellano et al. (2012), not only validate the previous statement as they also show readers which features present more critical to performance outcomes, highlighting that winning, drawing and losing teams may be discriminated from one another based on variables such as ball possession and the effectiveness of their attacking play, resulting in goals and passes in behind being the best predictors in the model.

Moreover, performance considerations are influenced by an athlete's field position. Berber et al. (2020) delves into the psychological attributes of each playing position, providing detailed information about the critical factors for position-specific performance evaluation. This analysis delves into key moments of the game, such as "initiate build-up" and "connection," aligning with the framework presented by Decroos et al. (2019), which evaluates player actions based on their impact on the game outcome and the contextual factors surrounding each action.

Expanding on players' traditional performance evaluation methods, this dissertation investigates a novel approach by leveraging Twitter data to analyze and assess player performance from a fresh perspective. While traditional performance evaluation methods, such as statistics and game footage, are still widely utilized, using social media as a performance evaluator adds a new dimension to analyzing an athlete's ability. Although limited in number, studies have already emerged, employing this social network to assess and characterize athlete performance.

For instance, Gruettner et al. (2020) adopted a player-centric approach, evaluating the impact of Twitter usage on professional tennis athletes. Their findings revealed a significant negative effect of high social media usage on athletes' performance. Similarly, Dreyer et al. (2022) discuss how sentiment analysis can predict athletes' performance based on social media activity. Results show that social media-related features that reflect athlete mood, social media behavior, and sleep quality before games significantly contribute to predicting performance. Interestingly, the results also show that, in particular, the number of tweets athletes receive before matches contributes considerably to performance prediction.

In the realm of basketball, Li et al. (2022) explored a novel evaluation method for NBA players, employing Spearman rank correlation to compare statistical performance with sentiment analysis. Their research suggested that subjective sentiment holds greater value among the three types of sentiment: absolute, objective, and subjective. Furthermore, Siemon & Wessels (2022) presented a compelling approach to leveraging Twitter data in order to extract the personality traits of basketball players and predict their performance in the National Basketball Association (NBA). By scraping statistics and big five personality traits from 185 professional basketball players, they conducted correlation analysis and multiple linear regressions to forecast NBA performance based on college performance and personality traits. The results showcased the potential of automated personality mining of tweets to gather additional insights about basketball players and predict their future performance.

While limited in number, these emerging studies demonstrate the growing trend of utilizing Twitter data to evaluate and gain deeper insights into athlete performance, opening up for further exploration in this dissertation.

2.4. Hypothesis

Now that we have a good understanding of the context, the main approaches found in the literature, and their limitations, we can focus on defining the problem we will tackle and how to do so.

Building on the theoretical foundation of wisdom of crowds and recognizing the dynamic interplay between field positions, fan sentiment, and player performance, this study formulates three hypotheses to explore these relationships in the context of football.

Hypothesis 1 – There is a relationship between fan sentiment and performance.

Drawing on Surowiecki (2005) premise, the concept of the wisdom of crowds suggests that the collective opinion of a group of individuals can often be more accurate than the opinion of a single expert. This means that if many fans have a positive sentiment towards a particular player, it could indicate their good performance. That is because the collective opinion of the fans can provide valuable insights into the player's abilities and performance. Therefore, analyzing fan sentiment gives us reason to believe in a connection between fan sentiment and player performance.

Using sentiment analysis on Premier League 22/23 season tweets dataset, the study aims to reveal patterns and correlations that support the hypothesized connection between fan sentiment and player performance. The strength and nature of this connection will be quantitatively assessed using the Pearson and Spearman correlation coefficients, providing insight into the association between fan sentiment and player performance metrics.

Hypothesis 1b – Fans' assessment of a football player's performance is more affected by impactful events than routine performance metrics.

Building on Hypothesis 1, which posits a relationship between fan sentiment and performance, we extend our focus to explore the specific elements that shape fan sentiment. It is likely that impactful events, such as game-changing goals or critical assists, as outlined by Barnaghi et al. (2016) and Ruz et al. (2020), significantly influence how fans perceive a football player's performance.

As we anticipate that fan sentiment is intricately connected with player performance, this hypothesis investigates the extent to which these pivotal moments hold more weight in shaping fan assessments compared to routine performance metrics. By analyzing comprehensive individual player statistics and employing a multi-linear regression approach, we aim to discern the critical factors that drive variations in fan sentiment and identify the aspects that resonate most with fans.

Hypothesis 2 – There is a sentiment bias depending on the field position of the players.
In the context of Hypotheses 1 and 1b, which explore the relationship between fan sentiment, player performance, and the impact of pivotal events, it becomes relevant to investigate how the field positions of football players contribute to these dynamics.

Field positions within football not only dictate a player's role and responsibilities but also influence their visibility and capacity to engage in impactful events. This variation in roles may lead to differences in fan sentiment, as fans could exhibit more positive sentiments towards players in positions that are perceived to have a greater influence on the game. We propose that the distinct roles played by individuals across different positions contribute to sentiment bias among fans.

To substantiate this hypothesis, our study will thoroughly analyze sentiment ratios across various field positions, categorizing players into positions like strikers, midfielders, and defenders. By employing Chi-squared statistics and ANOVA, we will examine whether the number of mentions and sentiment ratio distributions significantly differ across these categorical groups.

3. DATA AND METHODS

This chapter will describe the study design, data collection procedures, and data analysis strategies utilized to compile the material and respond to the research questions. It will provide a roadmap for the reader, outlining the measures followed to ensure the study is carried out methodically. It also emphasizes the justification for the methodologies picked and the steps followed to guarantee the accuracy and quality of the findings. The chapter will be divided into:

- Data Collection
- Pre-processing
- Methods

3.1. DATA COLLECTION

During this phase, we retrieved data from multiple sources, which was made possible by our work being supported by a carefully designed architecture. This architecture is structured into two main data pipelines, each with its distinct focus and methodology.

The first pipeline is dedicated to the collection of Twitter data. We chose to store this data in a MongoDB local cluster due to its proficiency in managing unstructured data, which is often the case with social media information. The processing of this data is segmented into three critical tasks: spam removal, sentiment analysis, and player identification. These steps are instrumental in enhancing data quality, gauging public opinion, and making sense of our hypothesis, respectively.

Our second pipeline targets the collection of data from various football websites, via web scraping (T et al., 2019). Specifically, we gather player ratings from the BBC and Fotmob websites, which are crucial for our correlation analyses. Additionally, we source data on individual player statistics for each game from the FBRef website. Transfermarkt is another valuable resource for us, providing data on players' individual characteristics, such as market value, age, and field position.

Upon completing the processing within both pipelines, we then merge the data to create the final dataset. This dataset serves as the foundation for our analysis, allowing us to draw meaningful insights and conclusions from our research.

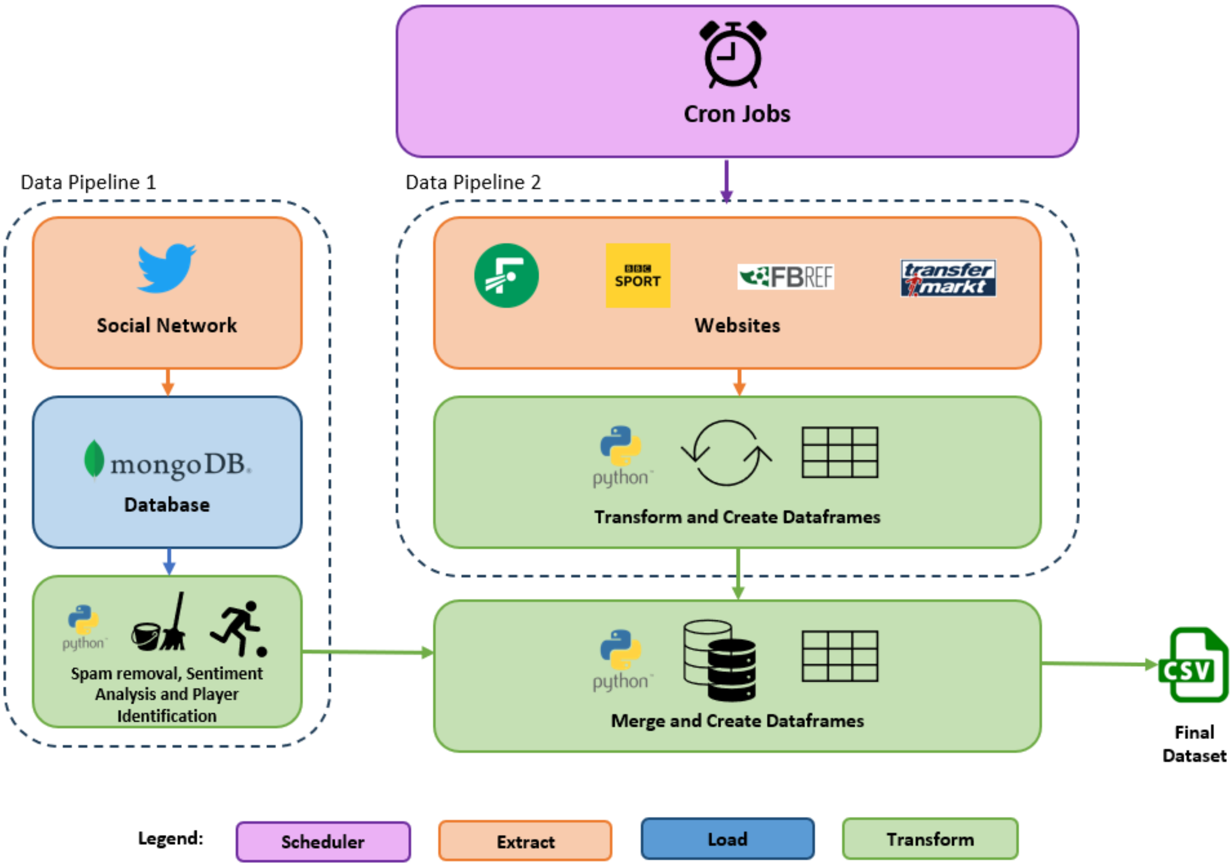


Figure 1 - Architecture of Data Collection

The selection criteria for choosing Twitter data were based on the following requirements:

- Large user base
- Public nature of tweets
- Real-time interactions
- Use of hashtags for games
- Data accessibility

Data collection started in December 2022 and finished in May 2023 for season 22/23 of the English Premier League with data based on Twitter. This league was chosen because it is the most competitive in the world and has the most significant number of fans, making obtaining large volumes of data easier. One of the positive aspects of the premier league is that there are hashtags for each game, so it is easy to filter information to obtain the desired data. As mentioned initially, the corpus of the data comprehends data about 135 games resulting in 1.6 million tweets collected.

This task runs every time a premier league game is live. The logic behind filtering out games is focused on producing results, as the flow for collecting tweets still needs to be fully automatized, as described in future research. For the Twitter data collection, some rules were applied.

Stream Rules:

- Only original tweets were collected;
- Only tweets in English were collected;
- Hashtags filter games (For example: Arsenal vs Liverpool #ARSLIV).

These rules are applied to obtain the most relevant sample of the extracted data. For instance, collecting only original tweets prevents duplicated information, and only tweets in English are more accessible for the model to interpret. In what concerns games, filtering by the game hashtag is the official way of communication for fans established by the Premier League.

For the collection, the Tweepi library is being used. It is an open-source Python library that helps programmers to interact with Twitter API more easily. This API provides streaming Twitter content, which is used for collecting fans' tweets during games.

In terms of aggregated data-points the dataset contains:

- 1,614,174 Tweets collected;
- 135 games;
- 2725 total players identified;
- 426 unique players identified;
- An average of 20 players identified by game.

Figure 1 illustrates the dynamics of the collected data. It presents a scatter plot that captures the correlation between the daily collected number of games and their associated mean tweet count per game.

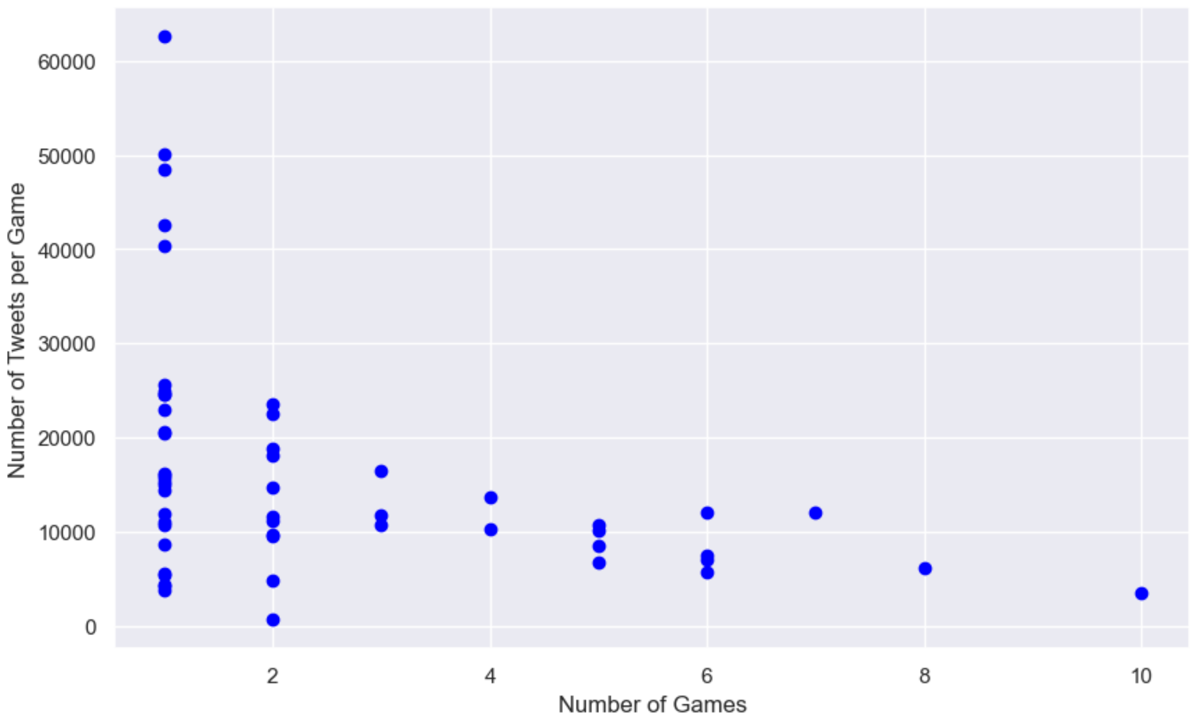


Figure 2 – Number of Games and Mean Number of Tweets by Date.

Next, we will showcase the disparity between games abundant in reviews and those with comparatively fewer. This description will contain essential metrics such as the number of reviews, player identifications within the game, and the count of distinct players identified. The focus will primarily be on the top three games that exemplify high review numbers and the bottom three games that exemplify comparatively lower review counts.

Table 1 - Top 3 and Bottom 3 Games Based on Number of Tweets with Summary Descriptions

Game	Result	Number of Tweets	Number of player identifications	Distinct players identified
#ARSMUN	3-2	62,576	14,289	22
#LIVMUN	7-0	50,086	11,217	22
#MCIARS	4-1	48,479	7,833	22
…	…	…	…	…
#SOUBRE	0-2	766	119	19
#FULWHU	0-1	727	205	18
#BHACRY	1-0	535	138	19

The next three tables offer readers an in-depth understanding of the features gathered from Twitter, Fotmob/BBC, along with Transfermarkt, thereby providing insight into all of data sources integrated into this work.

Table 2 -Twitter data collected

Feature	Description
Tweet	Full tweet text
Date/Time	Date and Time of tweet

Table 3 - Fotmob and BBC Sport data collected

Feature	Description
Name	Name of player
Shirt	Shirt Number
Club	Club player represents
Home Team	Team that plays home in that game
Away Team	Team that plays away in that game
Rating BBC	Rating [1 to 10] given by BBC Sports Website
Rating Fotmob	Rating [1 to 10] given by Fotmob Sports Website

Table 4 - Transfermarkt data collected

Feature	Description
Name	Name of player
Full Name	Full Name of player
Club	Club player represents
Shirt	Shirt Number
Position	Position in field
Age	Age of player
Market Value	Market value of player

3.2. Twitter data Pre-processing Steps

We initiated the process by eliminating emojis. Furthermore, we cleanse the text of special characters, which include numbers, punctuations, links, mentions, and new line characters. These elements, though often encountered in tweets, introduce noise into sentiment analysis. Additionally, we undertook the task of refining hashtags within the tweets. While hashtags can offer valuable context, they can also introduce bias into sentiment analysis. After that, our process involved filtering out special characters that occasionally appear within words, such as "&" or "$.". Recognizing the potential for spam tweets to skew sentiment insights, we adopted a strategy to exclude them by targeting specific keywords associated with spam. This proactive measure guarantees that our analysis remains insulated from the influence of irrelevant or misleading content. Lastly, we conducted tokenization where the text is split into words. The process is described in the below **Figure 2**.

Figure 3 - Sentiment Analysis Pre-processing steps.

In order to filter spam, we developed a specific strategy. By creating a word dictionary through manual inspection of tweets related to betting and streaming, we targeted the frequently occurring words found in such spam-type tweets. This approach enabled us to successfully filter our initial dataset of 1,614,174 tweets down to a refined set of 815,089 filtered tweets. To perform the player identification from the tweets, we use the Levenshtein distance (Yujian & Bo, 2007) to measure word similarity. For each word in a tweet, a comprehensive analysis of its similarity with player names from the obtained Transfermarkt dictionary is conducted.
Specifically, the similarity is calculated both for the first name, the last name, and the full name of each player. If either similarity score exceeds 85%, the word is considered a match. However, certain exceptions apply; the comparison word must be longer than three characters to exclude pronouns resembling player names. It is also essential to highlight that specific scenario, particularly when player names are identical, have been thoughtfully addressed in this analysis. For instance, in cases where names match exactly, a manual disambiguation process is employed to ensure precision and reliability in the calculated similarities.

$$\mathrm{lev}_{a,b}\,(i,j) = \begin{cases} \max(i,j) & \text{if } \min(i,j) = 0 \\ \min \begin{cases} \mathrm{lev}_{a,b}\,(i-1,j) + 1 \\ \mathrm{lev}_{a,b}\,(i,j-1) + 1 \\ \mathrm{lev}_{a,b}\,(i-1,j-1) + 1_{(a_i \neq b_j)} \end{cases} & \text{otherwise.} \end{cases} \tag{1}$$

3.3. Methods

In this study, we employ sentiment analysis to measure the emotional tone of Twitter data. The choice of methodology involves considering various algorithms for specific natural language processing (NLP) tasks, particularly those related to sentiment analysis.
A study by A. Hasan et al. (2018) demonstrates the utilization of Naive Bayes and Support Vector Machines (SVM) for sentiment analysis on Twitter data. Bayesian networks are employed for categorizing critical events in tweets, as highlighted in the work of Ruz et al. (2020). Additionally, neural networks (NN), particularly Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN), play a pivotal role in text classification. CNN, primarily designed for image analysis, is also effective at capturing features and relationships within larger text segments (Liao et al., 2017).

RNNs, on the other hand, are often utilized for text classification due to their recursive nature. Long Short-Term Memory (LSTM) models address the challenges posed by processing lengthy text sequences. Transformative models, exemplified by BERT (Devlin et al., 2019) and RoBERTa (Y. Liu et al., 2019), have revolutionized NLP tasks and paved the way for more sophisticated language representation models.

For this study, we adopted RoBERTa, a cutting-edge language model pre-trained on a vast corpus of approximately 58 million tweets (Barbieri et al., 2020). RoBERTa excels in classifying social media content due to its ability to comprehend contextual nuances and return the probabilities of a text's positive, negative, or neutral sentiment. To categorize a tweet, the sentiment category with the highest probability is assigned. For instance, if a text segment has probabilities of 0.7, 0.2, and 0.1 for positive, neutral, and negative sentiments respectively, it is classified as overall positive.

The result of the clean tweet and its sentiment is shown in the below **Figure 3**.

Figure 4 - Examples of Positive, Neutral, and Negative Tweets

4. RESULTS AND DISCUSSION

This chapter aims to discuss the results obtained and evaluate the hypotheses initially proposed. To achieve this, the chapter will be divided into three subchapters, providing a more detailed analysis for drawing conclusions. The information will be presented through a combination of complementary datasets and the player attribution method approach described earlier. The performance of our methods will be analysed, and the underlying reasons will be explored. The chapter will be structured as follows:

- Correlating Performance with Fan Sentiment
- Key determinants of fan's sentiment
- Sentiment across field positions

Before delving into the analysis of the hypotheses, in this chapter, we will start by presenting descriptive statistics of the dataset. This initial step offers an overview of the characteristics of the data, which in turn lays the groundwork for the exploration of results and the subsequent assessment of the hypotheses initially put forth.

To begin, we explore the characteristics of the described dataset. Firstly, out of approximately 815,089 tweets, we have 169,154 classifieds as positive, 416,795 as neutral, and 229,140 as negative. This distribution represents 21%, 51%, and 28% of the sample, respectively. Notably, positive reviews are underrepresented compared to negative ones, with neutral reviews comprising the largest proportion in the dataset.

Moving forward, the final dataset includes a total of 2,725 players, leading to the identification of 426 distinct players. Subsequently, we calculate a player sentiment ratio, which reflects the ratio of positive reviews to negative reviews for each player across different games. It's important to emphasize that multiple ratios are computed for the same player across various games. In the following figure, we present a bar chart that illustrates the distribution of these ratios and the corresponding number of players associated with each ratio.

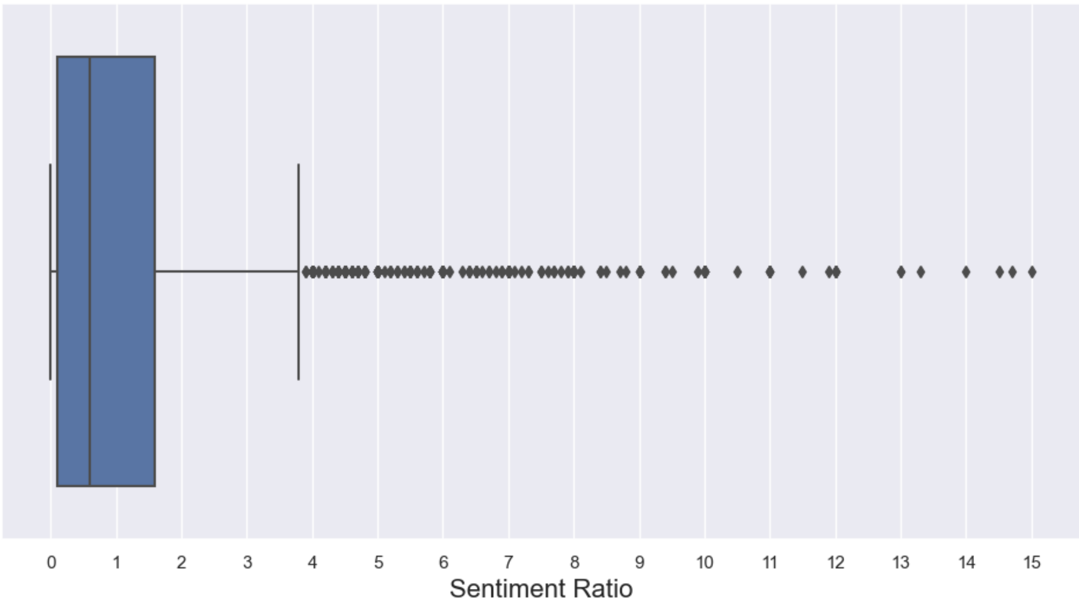


Figure 5 - Distribution of Players by Sentiment Ratio (number of positive tweets dividing per number of negative tweets)

In the majority of cases, ratios are concentrated between 0 and 2 units, with the lowest value at 0 units and the highest at 15 units. The mean ratio of player sentiment is approximately 1.3, with a significant influence from the largest ratios. To address outliers, they will be discussed in the subsequent subsection of this study.

Regarding quantiles, the first, second, and third quantiles correspond to values of 0.1 units, 0.6 units, and 1.6 units, respectively. This demonstrates a clear predominance of negative ratios, consistent with prior findings. This observation may be linked to the negativity bias phenomenon, which describes the tendency for individuals to give more attention and weight to negative information compared to positive information (G. Peeters & Czapinski, 1990).

In the upcoming figure, we will compare the ratios with the ratings for athletes in-game play from Fotmob and BBC.

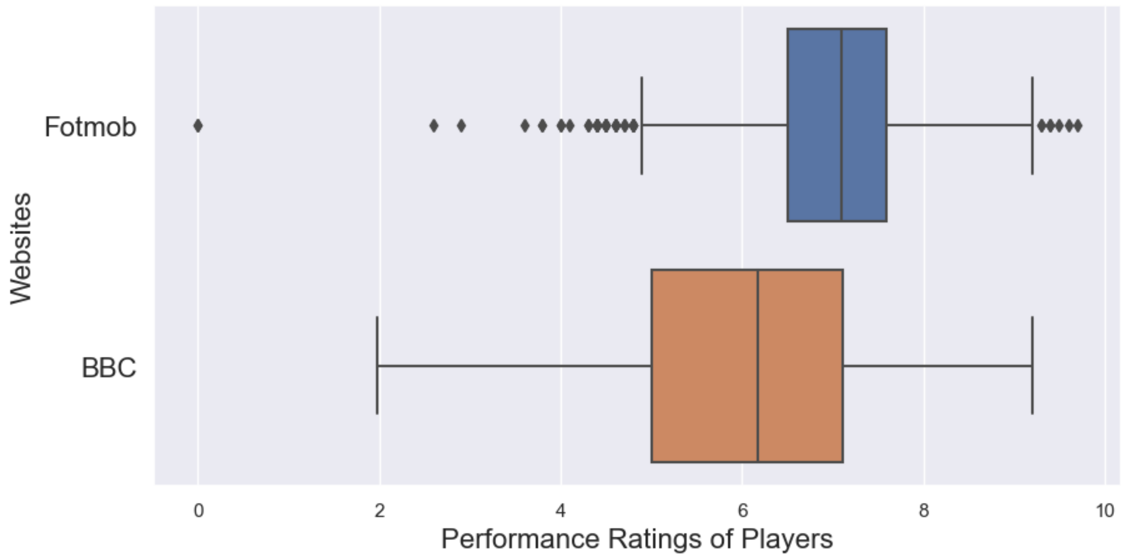


Figure 6 - Distribution of player’s Ratings sourced from fotmob and bbc

Fotmob, on average, assigns players a rating of approximately 7.02, illustrating a generally positive view of player performance. The dataset's relatively low standard deviation of approximately 0.87 indicates that Fotmob's ratings tend to cluster closely around this mean, suggesting a consistent approach to evaluating players. Notably, the dataset contains instances where player ratings reach a minimum of 0, signifying instances of exceptionally low ratings. At the 25th percentile, the rating stands at 6.5, indicating that a significant proportion of Fotmob's ratings fall above this value. The median rating, at 7.1, reflects a relatively symmetrical distribution of ratings. Additionally, the 75th percentile rating of 7.6 reveals a considerable number of high player ratings, with the highest observed rating reaching 9.7, suggesting occasional exceptionally positive evaluations.

BBC assigns players an average rating of approximately 6.03, slightly lower than Fotmob's mean rating, indicating a slightly less positive evaluation of player performance. This variance is further emphasized by a higher standard deviation of approximately 1.38, suggesting greater variability in BBC's player ratings. The presence of a minimum rating as low as 1.97 underscores instances where BBC assigns notably low ratings, potentially reflecting less favorable evaluations. At the 25th percentile, the rating stands at 5.0, implying that a substantial portion of BBC's player ratings falls below this value. The distribution is skewed, with a median rating of 6.18, indicating that ratings tend to cluster towards the lower end of the scale. Nevertheless, at the 75th percentile, a rating of 7.11 signifies a notable number of relatively high player ratings, although not reaching the same extremes as Fotmob. The highest rating observed for BBC is 9.21, indicating that while high player ratings are present, they do not reach the same extreme levels as Fotmob's evaluations.

Following this analysis of player ratings on both Fotmob and BBC, we proceeded to investigate the relationship between player sentiment ratios and the state of the game. To do so, we calculated the average sentiment ratio for all players at each state of the game. These findings are presented in Figure 6.

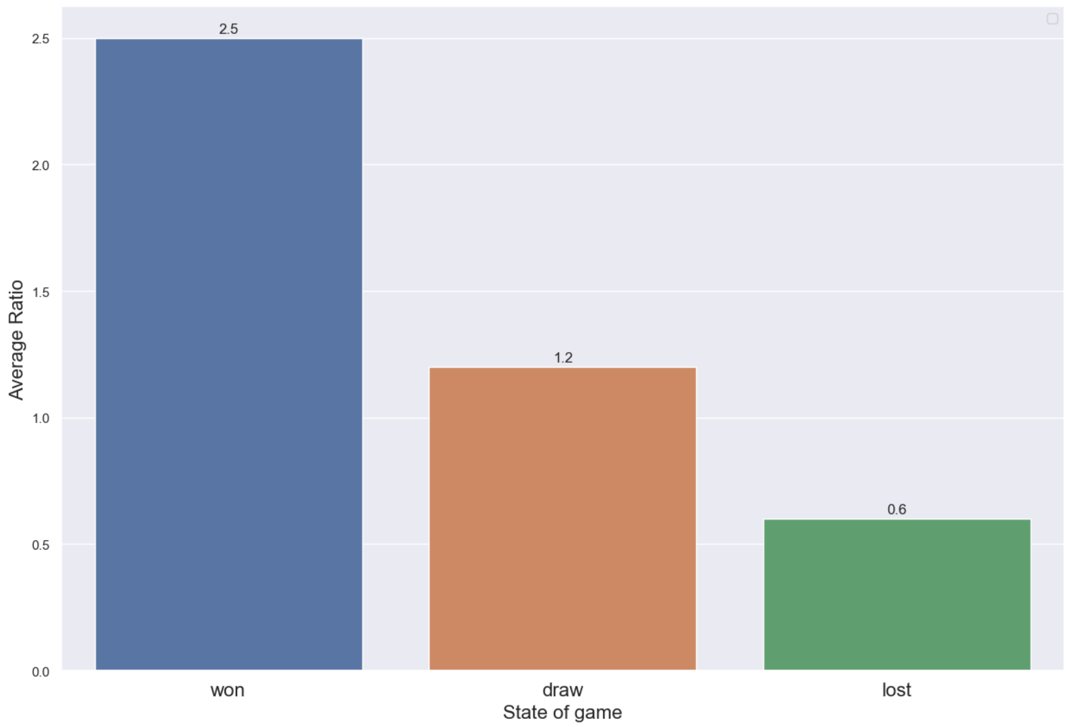


Figure 7 - Reviews per state of game

In the victory scenario, the average sentiment ratio stands at 2.5. This substantial ratio suggests that fans express significantly positive sentiments towards the players when their team wins. It signifies that triumph on the field resonates strongly with fan sentiment, leading to a surge in positive feedback and appreciation.

Contrasting the euphoria of victory, a game ending in a draw produces an average sentiment ratio of 1.2. Although this ratio is positive, it is notably lower than the 'won' state. The lower ratio suggests that while fans still maintain positivity, the outcome does not evoke as strong a sentiment as a clear victory. Draws may represent a mixed sentiment, where satisfaction coexists with a sense of missed opportunities.

The 'lost' state displays the lowest average sentiment ratio, at 0.6. This outcome triggers the least positive sentiment among fans. The decrease in the sentiment ratio during losses indicates fans' disappointment or frustration. It suggests that fans may be more critical or less forgiving when their team faces defeat, leading to a decrease in sentiment ratio.

We now turn our attention to how sentiment fluctuates throughout a game. For this analysis, we intentionally limited tweets generated exclusively within the two-hour timeframe of the game. Tweets collected minutes before the match started and those produced minutes after the game ended were deliberately excluded from our dataset for this time analysis. We will observe the temporal patterns of sentiment and present our findings through the following two figures.

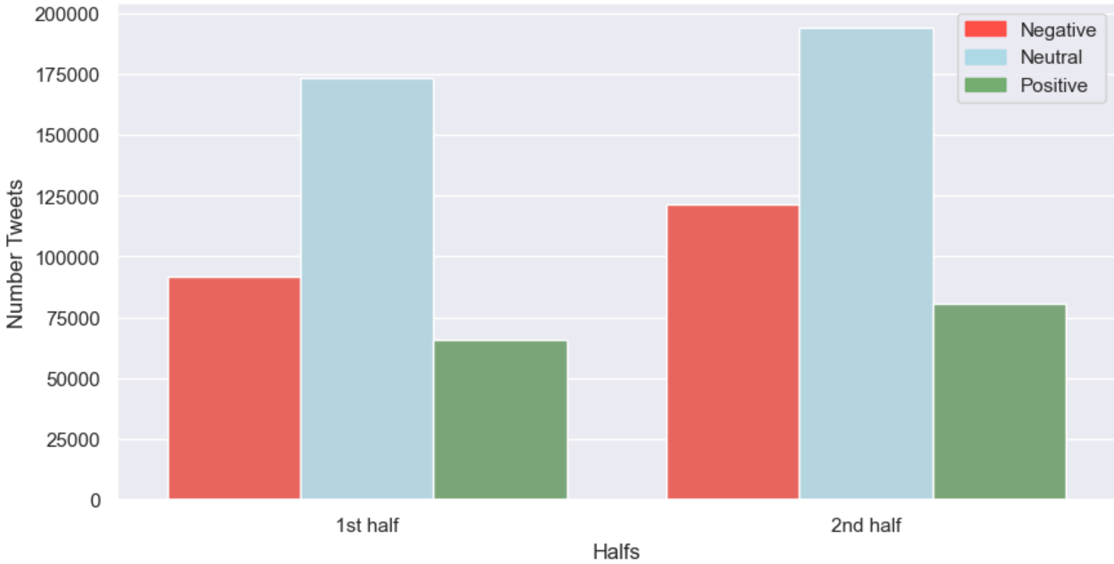


Figure 8 - Distribution of tweets by Sentiment across Game Halves

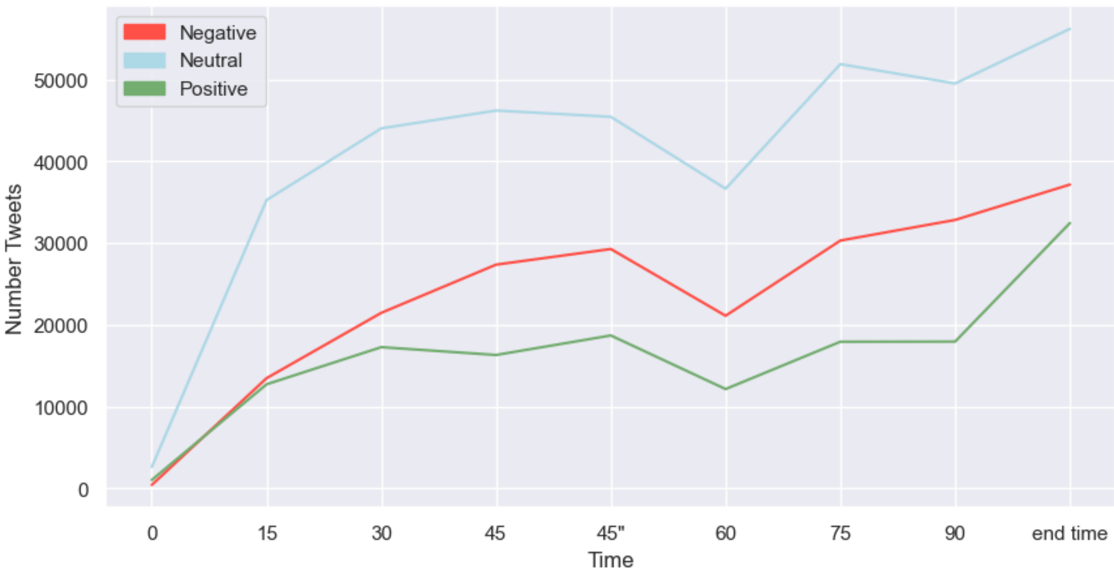


Figure 9 - Distribution of tweets by Sentiment over Time

In the initial half of the game, we observed a notable inclination toward positivity, with a substantial 66,058 tweets expressing 'Positive' sentiment. However, a distinct shift in the sentiment landscape is evident in the second half. During this phase, 'Negative' sentiment tweets surged to 121,411, indicating that fan sentiment tends to progressively become more negative as the game advances from the first half to the second. This shift suggests that changing game dynamics, outcomes, or events may have a pronounced impact on fan emotions as the game unfolds.

We further divided the game into quarters to gain a finer-grained understanding of sentiment progression. This division revealed the flow of sentiment throughout the game. In the initial quarter, 'Negative' sentiment started at 425 tweets, 'Neutral' at 2,623 tweets, and 'Positive' at 1,034 tweets. As the game unfolded, the sentiment landscape exhibited a notable shift, with 'Negative' sentiment steadily increasing to 29,277 tweets in the final quarter. 'Neutral' sentiment showed consistent growth, reaching 45,442 tweets, while 'Positive' sentiment gradually increased, peaking at 18,702 tweets.

Further delving into sentiment trends during specific time intervals, we noted intriguing patterns. Over a 15-minute interval, 'Negative' sentiment began at 13,476 tweets and reached 45,442 tweets, while 'Positive' sentiment gradually increased from 12,732 tweets to 18,702 tweets. Over a 30-minute interval, 'Negative' sentiment began at 21,471 tweets and progressively increased, eventually reaching 29,277 tweets. Meanwhile, 'Positive' sentiment exhibited a gradual upward trend. Over a 60-minute interval, 'Negative' sentiment initiated 21,116 tweets and experienced an increase, ultimately peaking at 30,309 tweets in the final 75-minute interval. 'Neutral' sentiment exhibited consistent growth, and 'Positive' sentiment fluctuated.

As the game approached its conclusion, sentiment appeared to stabilize, indicating the overall sentiment state as the game neared its end. At this point, 'Negative' sentiment amounted to 37,153 tweets, 'Neutral' sentiment to 56,198 tweets, and 'Positive' sentiment to 32,448 tweets. This stabilization provides a valuable reference point for understanding how fan sentiment culminates in response to the game's outcome and closing moments.

Our analysis also examined sentiment across various football clubs, shedding light on the sentiments most associated with each team. Notably, we observed a diverse range of sentiment distributions among these clubs, reflecting the multifaceted emotions expressed by fans in response to their respective teams' performances, as shown in the figure below.

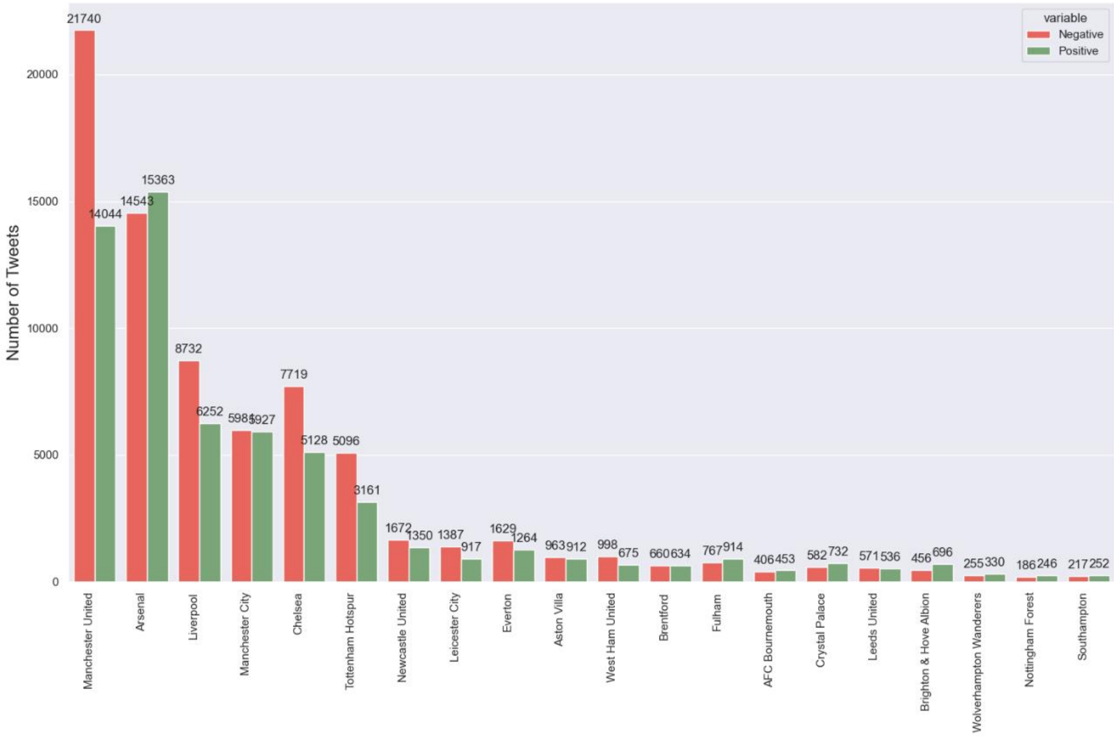


Figure 10 - Sentiment by Clubs where players were identified and Number of Games with Tweets

Among the clubs, 'Manchester United' garnered a substantial number of 'Positive' sentiment tweets, totalling 14,044, complementing a notable count of 'Negative' sentiment tweets at 21,740. This indicates significant discussion and emotion surrounding the club on social media platforms. Manchester United is also one of the most followed football clubs on Twitter, boasting an impressive 36 million followers.

'Arsenal' also stood out in our analysis, amassing 15,363 'Positive' sentiment tweets, indicating strong positive sentiment among fans. However, it was balanced by 14,543 'Negative' sentiment tweets, reflecting a mix of emotions associated with the club. Arsenal has a strong presence on Twitter, with 22 million followers.

'Liverpool' showed a similar trend with 6,252 'Positive' sentiment tweets and 8,732 'Negative' sentiment tweets, also suggesting a dynamic sentiment landscape. Liverpool has a large following on Twitter, with 24 million followers.

Additionally, 'Manchester City,' 'Chelsea,' and 'Tottenham Hotspur' exhibited various sentiment patterns, reflecting each club's unique fanbases and experiences. Manchester City has a considerable following on Twitter with 16 million followers, Chelsea has a strong presence with 25 million followers, and Tottenham Hotspur has fewer followers than the other clubs mentioned.

As we delve into the sentiment distribution per club, it's essential to recognize that the volume of sentiment discussions is often influenced by the clubs' social media presence, particularly their Twitter followers. Clubs with larger Twitter followings tend to attract more attention and generate more sentiment-related conversations. This trend is evident in clubs like 'Manchester United and 'Arsenal,' with extensive global fan bases, leading to many mentions and diverse sentiment expressions. Moreover, our analysis revealed intriguing patterns for other clubs as well. For instance, 'Fulham' exhibited a balance between 'Positive' and 'Negative' sentiments, reflecting the diverse emotions experienced by fans. 'West Ham United', while generating fewer mentions, showcased a notable count of 'Negative' sentiment tweets, indicative of the passionate discussions among its supporters. On the other hand, 'Brighton & Hove Albion' experienced a substantial count of 'Positive' sentiment tweets, demonstrating the positive sentiments associated with the club's performances.

Next, we will analyse the most frequently mentioned players, and the ensuing table will provide the relevant information about them.

Table 5 – Top 10 players with most mentions per game

Player	Club	Mentions per game	Market Value
Marcus Rashford	Manchester United	678	55M
Erling Haaland	Manchester City	489	170M
Antony	Manchester United	479	75M
Mohamed Salah	Liverpool	395	80M
Bukayo Saka	Arsenal	392	100M
Bruno Fernandes	Manchester United	386	75M
Gabriel Martinelli	Arsenal	384	60M
Harry Kane	Tottenham Hotspur	357	90M
Martin Ødegaard	Arsenal	334	60M
Casemiro	Manchester United	331	50M

The table before provides a detailed insight into the football players who have garnered substantial attention and discussion Twitter. To identify these players, we divided the number of mentions by the number of games collected, in order to avoid bias in the interpretation. This data reflects the prominence and influence of these players within the digital footballing landscape.

Among the players highlighted, we observe names such as 'Marcus Rashford,' 'Erling Haaland,' 'Antony,' 'Mohamed Salah,' and more. These players have consistently attracted the attention of fans, media, and online communities. 'Marcus Rashford,' representing 'Manchester United,' leads the list with a remarkable 12,893 mentions. This can be attributed to the player's exceptional performance during a season with a struggling club like Manchester United, where he delivered one of the best performances of his career. Similarly, 'Erling Haaland' from 'Manchester City' and 'Antony' from 'Manchester United' have garnered substantial attention with 8,817 and 9,117 mentions, respectively. These numbers underscore the significant impact these players have on digital football discussions.

As we delve into the analysis of the most mentioned players, it's worth noting that their popularity often extends beyond the digital realm and is intertwined with their market value and broader recognition in the footballing world. As represented in the next figure.

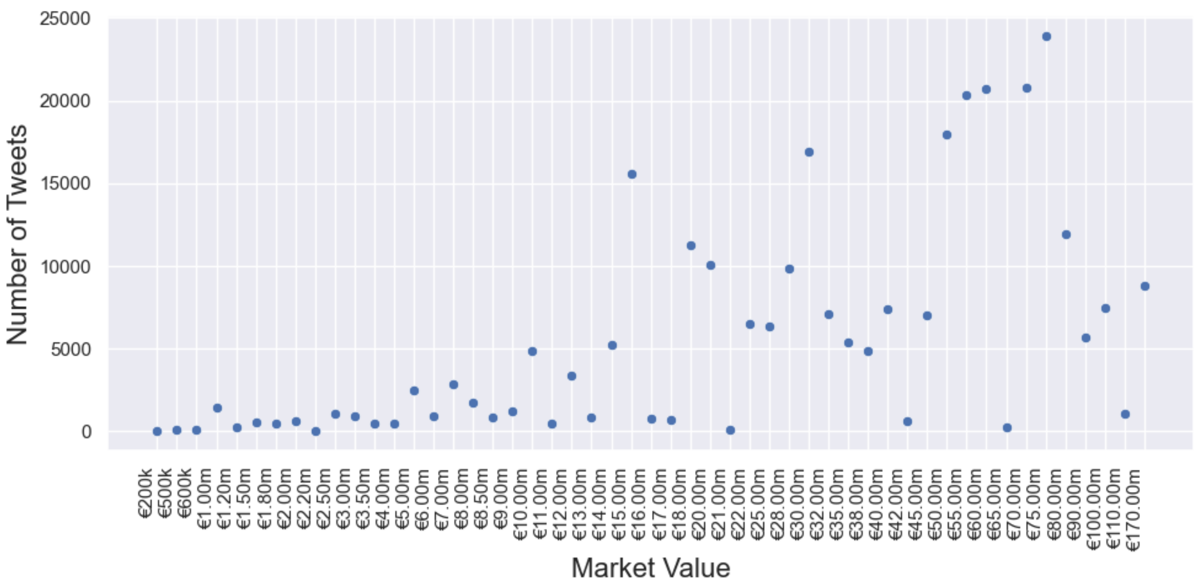


Figure 11 - Number of Tweets by Market Value of Players

The data presented in the table offers a nuanced understanding of the interplay between market value and sentiment counts for football players. As a general trend, it becomes evident that as a player's market value rises, so does the number of mentions for this player. This suggests that players with higher market values often command increased attention and evoke stronger opinions among football fans and the media alike.

A noteworthy aspect is the balance that emerges in sentiment as market value increases. In other words, high-value players tend to be a subject of both praise and criticism in equal measure. This equilibrium in sentiments underscores the multifaceted nature of public perception surrounding football players.

Furthermore, certain market value milestones, such as €10 million, €15 million, and €20 million, stand out as key points at which sentiment counts experience significant spikes. Players whose market values reach or exceed these milestones tend to attract a substantial volume of attention and generate a higher quantity of sentiments. These milestones signify the players' ascent into a realm of heightened recognition and discussion. Even players with moderate market values can become focal points of passionate discussions and sentiments, reflecting the dynamic nature of football fandom.

Furthermore, considering the average market value for all players in the dataset is approximately €12.51 million, it is evident that fans typically engage in more discussions about top players, even if their performance may not always align with their market value. For the top players identified, with an average market value of around €82.50 million, this phenomenon is particularly pronounced. It signifies that top players consistently garner more attention and spark discussions, irrespective of their current performance levels. This insight underscores the enduring fascination that football enthusiasts have with the game's elite players, emphasizing the complexity of sentiment dynamics within the footballing ecosystem.

With this comprehensive analysis clarifying the intricate relationships between player sentiment, market value, and various contextual factors, we have laid a solid foundation for exploring our hypotheses. In the upcoming sub-chapters, we will empirically test these hypotheses to uncover the intricate dynamics that shape fan sentiment in football.

4.1. CORRELATING PERFORMANCE WITH FAN SENTIMENT

As previously mentioned, our analysis is built on a dataset that includes data from all players identified. We've calculated a player sentiment ratio, which measures the ratio of positive to negative reviews for each player across different games within this dataset. It's worth emphasizing that, for this study, we excluded neutral comments, as they don't offer a subjective viewpoint but rather reflect factual observations. To ensure consistency with the principle of a substantial number of contributors and opinions, as outlined by Surowiecki, (2005), we established a threshold for the 'Number of Reviews' variable. Consequently, players with fewer than 100 mentions (combining positive and negative reviews) were excluded from the analysis, as their limited mentions bias the final results.

We use statistical methods such as Pearson's correlation coefficient to analyse the continuous quantitative data to determine the correlation between players' performance and fans' feelings.

Pearson's correlation coefficient (Hauke & Kossowski, 2011) is used when the variables are assumed to follow a linear relationship and their distributions are approximately normal. It measures both the strength and direction of linear relationships between variables.

$$r_{xy} = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^{n} (y_i - \bar{y})^2}} \quad (2)$$

Spearman's rank correlation coefficient (Hauke & Kossowski, 2011) is instead a nonparametric rank statistic proposed as a measure of the strength of an association between two variables.

$$\rho = 1 - \frac{6\sum d^2}{n(n^2 - 1)} \quad (3)$$

These two methods provide valuable insights into the strength and direction of the relationship between the variables in the investigation, and the coefficient can vary between -1 and 1, assuming that the value 0 represents the non-correlation between the variables and, consequently, the non-rejection of the null hypothesis.

The two figures below represent the correlation between the players' performance and the ratings obtained in the game for each website, Fotmob and BBC.

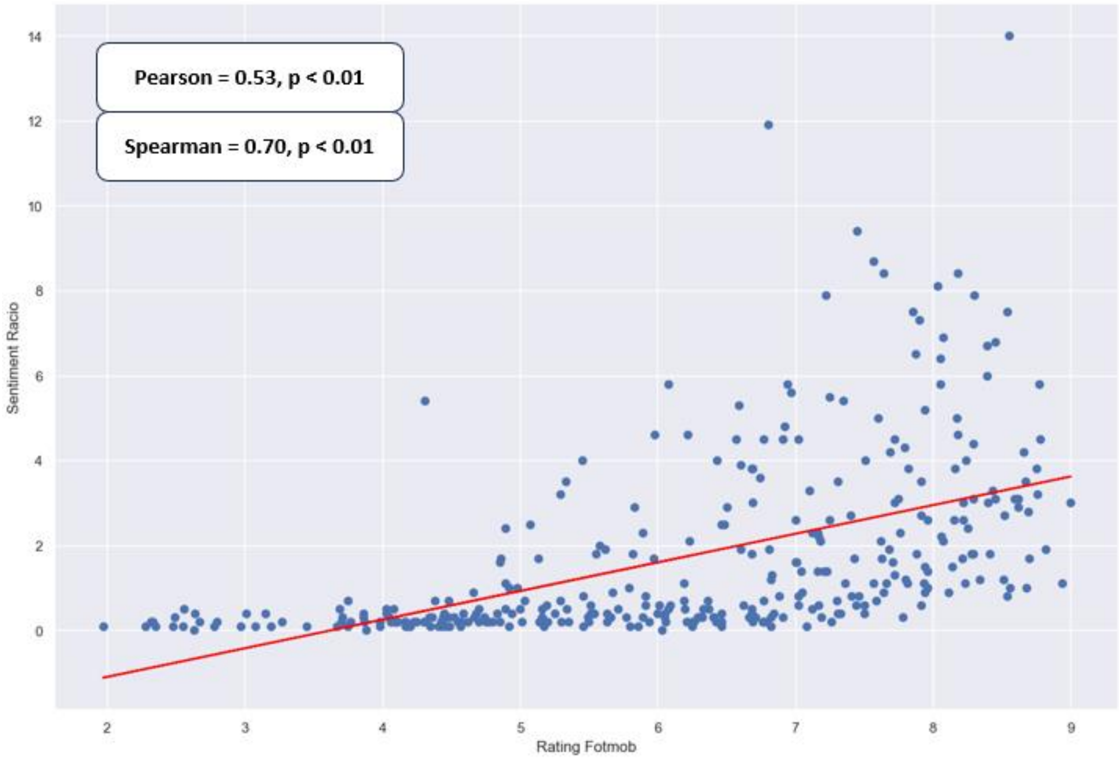


Figure 12 - Pearson Correlation for Fotmob Website between Ratings and Sentiment of Fans

In this analysis, the calculated correlation coefficient is approximately 0.53. This value indicates a positive correlation between "Rating Fotmob" and "Sentiment Ratio". A positive correlation signifies that as "Rating Fotmob" values increase, the "Sentiment Ratio" also tends to increase. However, it's important to note that while there is a clear positive relationship, it is not extremely strong, suggesting a moderate linear association between the two variables. The associated p-value for the correlation coefficient is less than 0.01, indicating that the observed correlation is statistically significant and unlikely to have occurred due to random chance. Consequently, the first hypothesis, proposing a correlation between rating and fan sentiment, is accepted, while the null hypothesis is rejected.

Additionally, when applying Spearman's rank correlation, we find a correlation coefficient of approximately 0.70, also with a significant p-value of < 0.01. This Spearman correlation suggests a strong monotonic relationship between the variables, supporting the notion of a robust association.

Within the scatter plot, we observe a notable cluster situated in the lower left quadrant. This cluster is characterized by players with 'Rating Fotmob' values typically ranging from 4 to 5, along with low 'Sentiment Ratio' values, which are close to 0. This cluster suggests that players in this group tend to have lower Fotmob ratings, indicating subpar performance, and receive predominantly negative sentiment on Twitter. This finding hints at a potential correlation between poor on-field performance and negative sentiment among fans on social media.

Moving to the upper left quadrant, we identify another cluster. Within this group, players exhibit 'Rating Fotmob' values ranging from 4 to 6, along with 'Sentiment Ratio' values that tend to be 1 to 5. This cluster suggests that players with moderate Fotmob ratings receive mixed sentiment on Twitter. This variance in sentiment may arise from inconsistent or unpredictable performance on the field, leading to diverse opinions among fans.

The upper right quadrant of the plot houses a distinct cluster comprising players with high 'Rating Fotmob' values, typically exceeding 7, and increased 'Sentiment Ratio' values, often surpassing 4. This cluster implies that players with outstanding Fotmob ratings receive overwhelmingly positive sentiment on Twitter. These players likely represent top-performing individuals who enjoy widespread acclaim and support from fans.

We also observe outliers in the scatter plot. Some of them have remarkably high 'Sentiment Ratio' values, indicating exceptionally positive sentiment, while others exhibit shallow 'Rating Fotmob' values, suggesting poor performance that leads to negative sentiment on Twitter.

Lastly, in the central region of the scatter plot, we notice a sparser distribution of data points. This suggests no strong correlation exists between 'Rating Fotmob' and 'Sentiment Ratio' in this area. Players here receive varying sentiments on Twitter, irrespective of their Fotmob ratings.

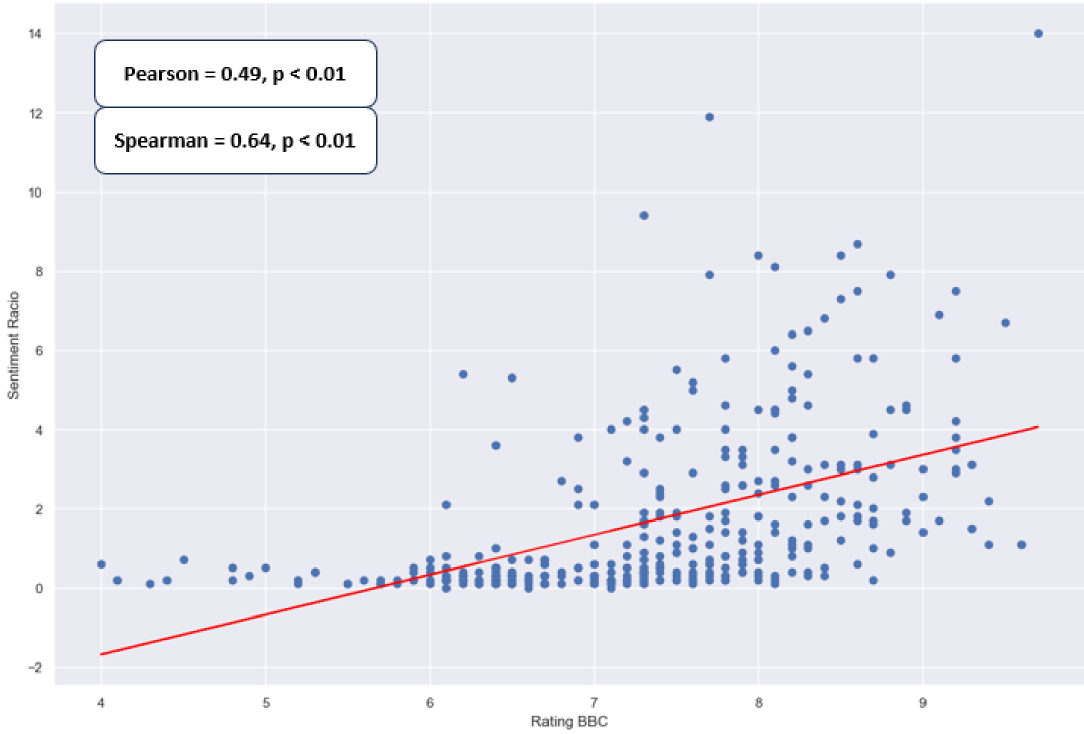


Figure 13 - Pearson Correlation for BBC Website between Ratings and Sentiment of Fans

Similarly, with BBC, the calculated correlation coefficient of approximately 0.49 indicates a positive moderate correlation between "Rating BBC" and "Sentiment Ratio" in the dataset. This positive correlation suggests that as "Rating BBC" values increase, the "Sentiment Ratio" also tends to increase.

The p-value, less than 0.01, confirms that the observed correlation between "Rating BBC" and "Sentiment Ratio" is statistically significant. Therefore, we accept the first hypothesis, which suggests a correlation between player ratings and fan sentiment, and simultaneously reject the null hypothesis for the BBC website dataset.

When examining the data with Spearman's rank correlation, we observe a correlation coefficient of approximately 0.64, also with a significant p-value of < 0.01. This Spearman correlation supports the presence of a strong monotonic relationship between the variables, reinforcing the idea of a robust association.

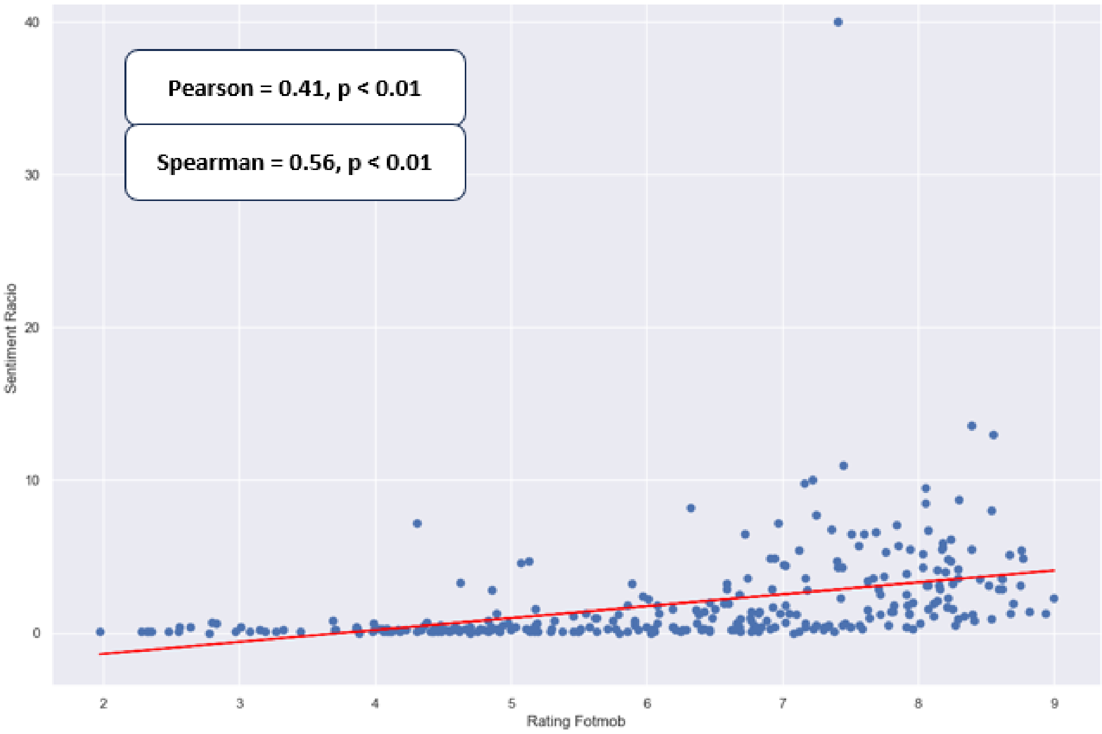
One such cluster in the lower left quadrant showcases players with low "Rating BBC" values, typically falling within the range of 4 to 5. These players also exhibit low "Sentiment Ratio" values, hovering around 0. This particular cluster indicates that individuals in this category tend to have lower BBC ratings, which signify below-par performance on the field. Moreover, they predominantly receive negative sentiment on Twitter, indicating a noteworthy correlation between weak on-field performance and a corresponding negative sentiment among fans on social media.

Moving to the upper right quadrant of the scatter plot contains a distinctive cluster of players with high "Rating BBC" values, often exceeding 8, and correspondingly high "Sentiment Ratio" values, frequently surpassing 4. This cluster strongly suggests that players with outstanding BBC ratings consistently receive overwhelmingly positive sentiment on Twitter. These players likely represent top-performing individuals who enjoy widespread acclaim and unwavering support from fans.

Additionally, specific clusters within the scatter plot also feature high and low "Rating BBC" values while maintaining relatively high "Sentiment Ratio" values. This observation suggests that fan sentiment can exhibit positivity even for players with moderate BBC ratings. This phenomenon may be attributed to specific circumstances, qualities, or narratives associated with those players.

Next, we will extend our analysis by applying these correlations to the same dataset, dividing it into halves corresponding to different game phases for each website.

1st Half



2nd Half

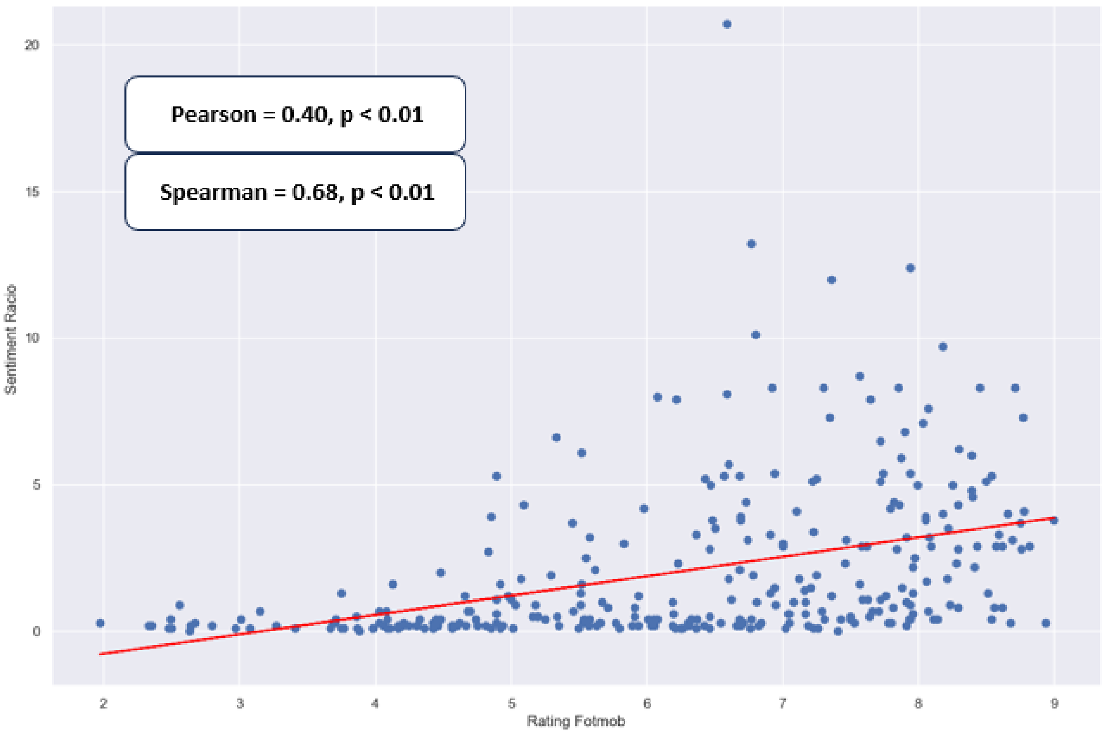
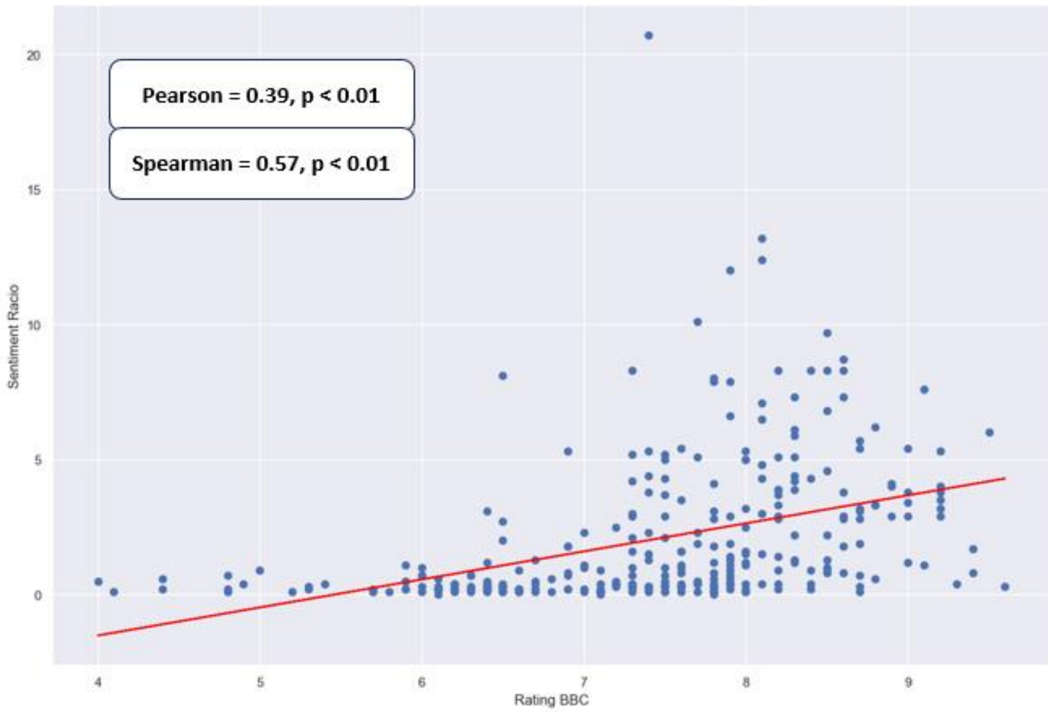


Figure 14 - Pearson Correlation for Fotmob Website based on halves of the game

1st Half



2nd Half

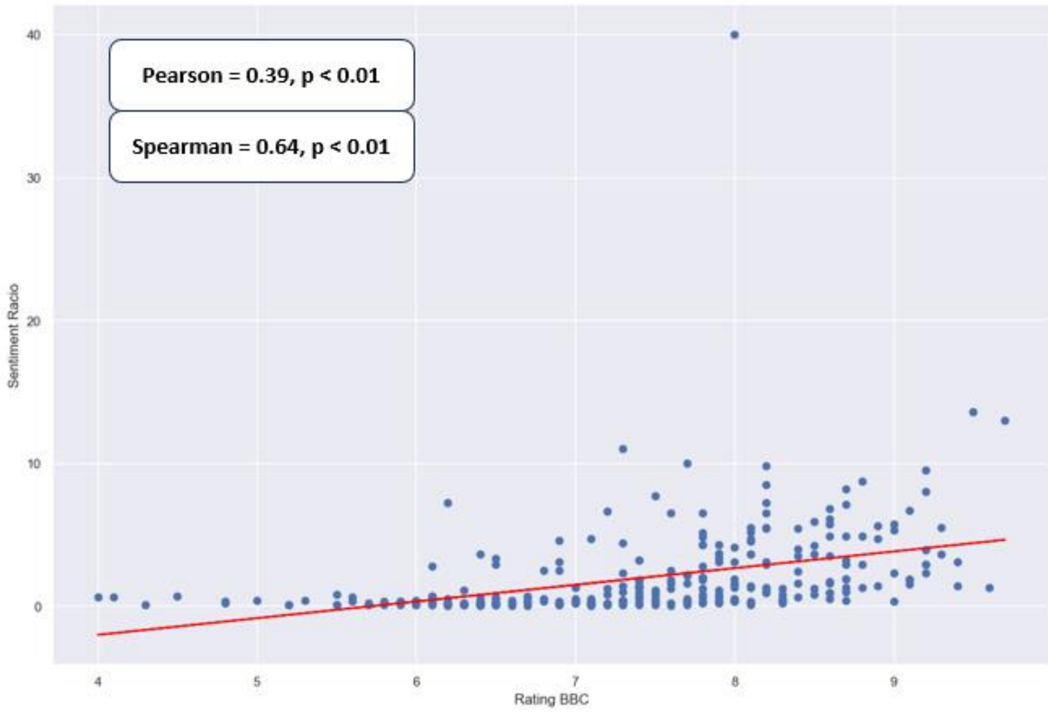


Figure 15 - Pearson Correlation for BBC Website based on halves of the game

In both the Fotmob and BBC datasets, we observed distinct player performance patterns during the 1st and 2nd halves. In the 1st half, a 'Lower Left Cluster' of players with low 'Rating Fotmob' and 'Rating_BBC' values, ranging from 4 to 5, underperformed, receiving predominantly negative sentiment on Twitter. In this phase, there was a statistically significant positive correlation between Fotmob ratings and Twitter sentiment with a Pearson coefficient of 0.41 (Fotmob) and 0.39 (BBC), both with p-values less than 0.01. Spearman correlations for the 1st half were also noteworthy, with coefficients of 0.56 (Fotmob) and 0.57 (BBC), both indicating a strong monotonic relationship (and p-values < 0.01).

Shifting to the 2nd half, both datasets displayed similar dynamics. An 'Upper Middle Cluster' with high 'Rating Fotmob' and 'Rating_BBC' values, often exceeding 7, excelled, garnering positive Twitter sentiment. In this phase, a moderate positive correlation between Fotmob ratings and Twitter sentiment was observed with Pearson coefficients of 0.40 (Fotmob) and 0.39 (BBC), both with p-values less than 0.01. Notably, Spearman correlations for the 2nd half were even stronger, with coefficients of 0.68 (Fotmob) and 0.64 (BBC), indicating a robust monotonic relationship (and p-values < 0.01). This noteworthy increase in Spearman correlations in the 2nd half may be attributed to the fact that fans have already witnessed the full performance of all players during the match, providing them with a strong foundation to assess and express their sentiments.

Overall, these findings highlight the consistent relation between player performance and fan sentiment, regardless of the dataset or game phase. The Pearson correlation coefficients, remarkably similar between the two datasets and halves, reinforce this relationship.

Lastly, we will investigate if player market values are correlated with fan sentiment, examining whether there is a consistent relationship between player market values and the sentiments expressed by fans on Twitter. The following figure will illustrate these potential correlations.

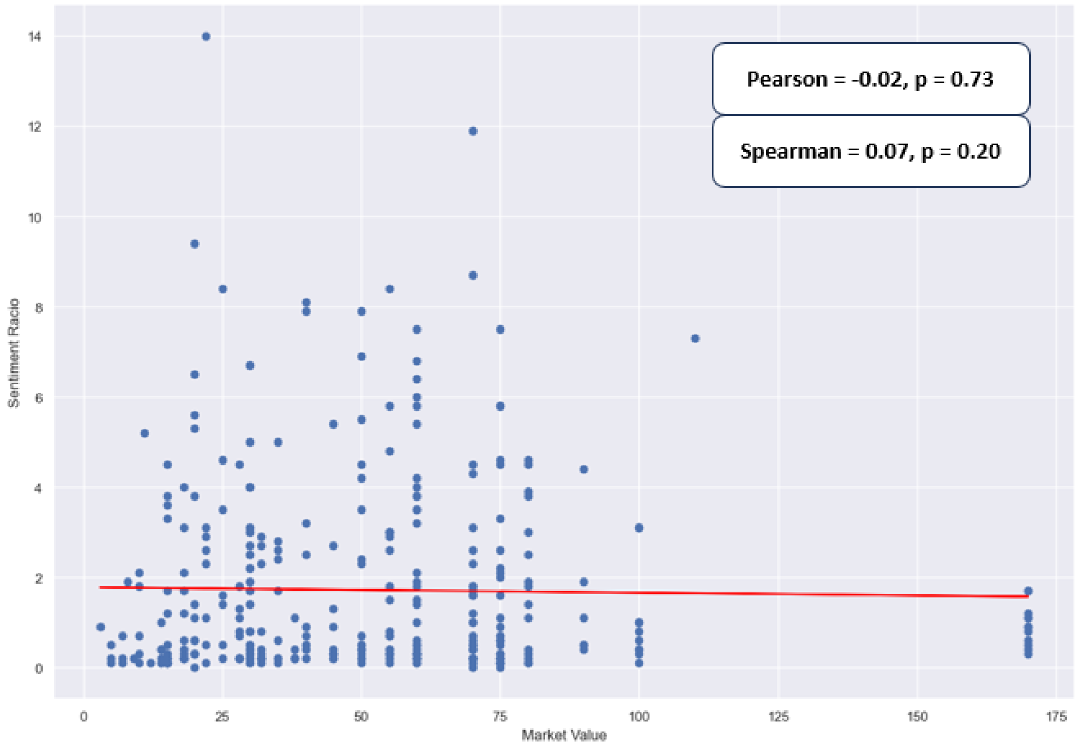


Figure 16 - Pearson Correlation between Market Value and Sentiment of Fans

Examining the scatter plot, it becomes evident that there is no discernible relationship between player market value and fan sentiment. The data points appear scattered without any noticeable pattern, corroborated by a low Pearson correlation coefficient of -0.02 and a significantly high p-value of 0.73. The p-value, notably exceeding the standard threshold for hypothesis rejection (0.1), reinforces the lack of statistical significance.

Contrary to our previous market analysis, a distinct trend emerges when considering player popularity measured by Twitter mentions. Notably, players with higher market values tend to garner more Twitter attention. However, this popularity does not necessarily translate into exclusively positive sentiments. Our findings suggest that fans can be particularly critical when evaluating high-profile players, often expecting exceptional performances. This highlights the nuanced nature of fan sentiment, which may be more demanding and scrutinizing toward renowned players.

4.2. KEY DETERMINANTS OF FANS SENTIMENT

This section aims to enhance the robustness of our descriptive findings by examining the connection between fan sentiment and athlete statistics. To achieve this, we've incorporated various player statistics into our dataset, categorizing them into summaries, defensive, attacking, and passing metrics.

To maintain a coherent and logical approach in building our model, we've decided to focus on players who have garnered at least 100 positive and negative reviews each, resulting in 340 athletes eligible for our multiple linear regression analysis. However, we've made an exception for goalkeepers, who are assessed using different criteria than outfield players. Consequently, we've excluded them from our analysis. While we briefly considered developing a separate model for goalkeepers, only around 19 goalkeepers in the sample made it statistically unviable to pursue such an approach.

By employing a multi-linear regression analysis, our aim is to unveil the intricate relationship between various statistics of the game for each player and the way fans perceive and form their opinions about them.

As a result, we have constructed three distinct models to discern the factors influencing fan performance perception. The first model encompasses individual player statistics, while the second focuses on collective team statistics. Lastly, the third model assesses the combined impact of these variables to validate our second hypothesis. We've compiled all statistical data from the FBREF website and previously collected variables such as club, age, position, and market value. The comprehensive dataset will be presented in **Table 6**, complete with detailed descriptions of these statistics. This data will be used to estimate a linear relationship between the Sentiment of fans (ratio_PN) and the remaining 32 variables.

The linear model to be estimated is given by:

$$
\begin{aligned}
ratio\_PN \; = \; & \beta 0 \; + \; \beta 1 * club \; + \; \beta 2 * age + \; \beta 3 * position + \; \beta 4 * market\_value + \; \beta 5 \\
& * result + \; \beta 6 * Performance\_Gls + \; \beta 7 * Performance\_Ast + \; \beta 8 \\
& * Performance\_PK + \; \beta 9 * Performance\_PKatt + \; \beta 10 \\
& * Performance\_Sh + \; \beta 11 * Performance\_SoT + \; \beta 12 \\
& * Performance\_CrdY + \; \beta 13 * Performance\_CrdR + \; \beta 14 \\
& * Performance\_Touches + \; \beta 15 * Performance\_Tkl + \; \beta 16 \\
& * Performance\_Int + \; \beta 17 * Performance\_Blocks + \; \beta 18 * SCA\_SCA \\
& + \; \beta 19 * SCA\_GCA + \; \beta 20 * Passes\_Cmp\% + \; \beta 21 * Passes\_PrgP + \; \beta 22 \\
& * Carries\_Carries + \; \beta 23 * Carries\_PrgC + \; \beta 24 * Take - Ons\_Att \\
& + \; \beta 25 * Take - Ons\_Succ + \; \beta 26 * Tkl + Int + \; \beta 27 * Clr + \; \beta 28 * Err \\
& + \; \beta 29 * Defense\_TklW + \; \beta 30 * Defense\_Tkl\_dribblers + \; \beta 31 \\
& * Defense\_Att + \; \beta 32 * Defense\_Lost
\end{aligned} \tag{3}
$$

, where β0 represents the intercept or constant term, which is the value of the predicted target variable when all independent variables are equal to zero.

Table 6 - Set of variables describing player performance

Variables	Description
club	Club of player
age	Age of player
position	Field position of player
market_value	Market value of player
result	Game result
Performance_Gls	Goals scored or allowed
Performance_Ast	Assists
Performance_PK	Penalty kicks made
Performance_PKatt	Penalty kicks attempted
Performance_Sh	Total Shots

Performance_SoT	Shots on target
Performance_CrdY	Yellow Cards
Performance_CrdR	Red Cards
Performance_Touches	Number of times a player touched the ball
Performance_Tkl	Number of players tackled
Performance_Int	Interceptions
Performance_Blocks	Number of times blocking the ball by standing in its path
SCA_SCA	Shot-Creating Actions
SCA_GCA	Goal-Creating Actions
Passes_Cmp%	Pass Completion Percentage
Passes_PrgP	Progressive Passes
Carries_Carries	Number of times the player controlled the ball with their feet
Carries_PrgC	Progressive Carries
Take-Ons_Att	Number of attempts to take on defenders while dribbling
Take-Ons_Succ	Number of defenders taken on successfully
Tkl+Int	Number of players tackled plus number of interceptions
Clr	Clearances
Err	Mistakes leading to an opponent's shot
Defense_TklW	Tackles Won
Defense_Tkl_dribblers	Number of dribblers tackled
Defense_Att	Number of unsuccessful challenges plus number of dribblers tackled
Defense_Lost	Challenges Lost
ratio_PN	Sentiment Ratio

We selected these variables because they encompass the most crucial individual statistics contributing to a player's success. Our approach involved initially identifying key general variables labelled with the prefix 'Performance.' Subsequently, we delved into position-specific attributes to assemble a diverse and comprehensive set of variables to accommodate all players except for goalkeepers, as previously explained.

We followed a series of systematic steps to facilitate the progression of the models. Initially, our exploration centred on examining the relationships between variables. This entailed conducting correlation analyses to discern which variables might introduce bias into our model. To this end, we established a correlation threshold of 0.7, which guided our process of eliminating highly correlated variables. The outcome of this analysis revealed several pairs of variables with notably high correlations: ('Performance_PKatt', 'Performance_PK'), ('SCA_GCA', 'Performance_Ast'), ('Carries_Carries', 'Performance_Touches'), ('Take-Ons_Succ', 'Take-Ons_Att'), ('Tkl+Int', 'Performance_Tkl'), ('Defense_TklW', 'Performance_Tkl'), ('Defense_TklW', 'Tkl+Int'), ('Defense_Tkl_dribblers', 'Performance_Tkl'), ('Defense_Att', 'Defense_Tkl_dribblers'), and ('Defense_Lost', 'Defense_Att').

Subsequently, guided by domain knowledge and the context of our study, we made informed decisions to retain or eliminate specific variables.

Consequently, our ultimate dataset excludes the following variables: 'Performance_PKatt', 'SCA_GCA', 'Carries_Carries', 'Tkl+Int', 'Defense_TklW', 'Defense_Tkl_dribblers', 'Defense_Lost', 'Defense_Att', and 'Take-Ons_Att.' This carefully considered variable selection process ensures that our models are based on a more focused and pertinent set of features.

Next, we normalized our integer variables, except those that already functioned as flags, such as 'Performance_CrY' or 'Performance_CrR.' Following this, we created dummy variables for our categorical data, encompassing the following variables: 'club,' 'position,' and 'result'.

To facilitate a more organized and interpretable presentation of our models, we have undertaken an aggregation process for the 'position' variable. For instance, players with roles encompassing both centre-backs and full-backs now fall into the 'Defenders' category, while similar aggregation applies to midfielders and attackers. This approach significantly enhances data clarity and mitigates potential bias. With this enhanced structure, we are now ready to present the models.

Table 7 – Regression results

Variables	Player Model (1)	Team Model (2)	Integrated Model (3)
Carries_PrgC	0.07		-0.14
	(0.86)		(0.87)
Clr	0.86		1.20*
	(0.64)		(0.68)
Err	-0.87		-0.18
	(0.78)		(0.83)
Passes_Cmp%	1.36*		1.43*
	(0.74)		(0.77)
Passes_PrgP	2.48***		2.00**
	(0.95)		(0.97)
Performance_Ast	3.65***		3.10***
	(0.68)		(0.66)
Performance_Blocks	-0.17		-0.01
	(0.49)		(0.48)
Performance_CrdR	-0.51		-1.16
	(0.84)		(0.82)
Performance_CrdY	-0.59*		-0.33
	(0.30)		(0.30)
Performance_Gls	2.76***		2.19***
	(0.48)		(0.47)
Performance_Int	1.98***		1.77***
	(0.66)		(0.66)
Performance_PK	-0.34		-0.69
	(0.83)		(0.79)
Performance_Sh	-0.95		-1.30*
	(0.76)		(0.75)
Performance_SoT	1.05		0.88
	(0.93)		(0.88)
Performance_Tkl	0.19		-0.40
	(0.65)		(0.65)
Performance_Touches	-2.48**		-1.27
	(1.24)		(1.36)
SCA_SCA	1.83**		0.90
	(0.78)		(0.78)
Take-Ons_Succ	0.03		0.26
	(0.48)		(0.47)
age	-0.94*		-0.28
	(0.56)		(0.59)
market_value	-1.40**		-0.00
	(0.61)		(0.71)
position_Attack		0.32**	0.17
		(0.15)	(0.22)
position_Defender		0.12	-0.32
		(0.21)	(0.32)
position_Midfielder		0.75***	0.29

		(0.18)	(0.24)
result_draw		-0.10	-0.34
		(0.19)	(0.23)
result_lost		-0.48**	-0.51**
		(0.19)	(0.22)
result_won		1.78***	1.00***
		(0.17)	(0.22)
Constant	0.49	1.20***	0.14
	(0.69)	(0.18)	(0.44)
Fixed Effects	No	Yes	Yes
Observations	319	319	319
R-squared	0.34	0.28	0.46
Adjusted R-squared	0.30	0.22	0.38
F-Statistics	7.74	5.38	5.80

Standard errors in parentheses.
* p<.1, ** p<.05, ***p<.01
“Yes” in the Fixed Effects row indicates that the model considers fixed effects for Clubs, while “No” indicates it does not.

Delving into the first model, we see that the coefficients for passing and attacking metrics provide insights into their impact on fan sentiment. A positive coefficient for "Passes_Cmp%" (pass completion percentage) suggests that fans react positively when their team completes a higher rate of passes. This is indicative of reasonable ball control and playmaking abilities. Similarly, "Passes_PrgP" (progressive passes) and "Performance_Ast" (assists) both exhibits significantly positive coefficients. This implies that fans value advanced passing and assists, recognizing their contributions to attacking plays and scoring opportunities. Additionally, the solid positive coefficient for "Performance_Gls" (goals scored) reaffirms that fans react positively when their team or players score.

On the other hand, the negative and highly significant coefficient for "market_value" suggests an interesting dynamic. As a player's market value increases, fan sentiment tends to decrease. This finding might be attributed to higher expectations or increased pressure associated with high-value players. Fans may become more critical in their evaluations of such players.

Regarding defensive metrics and player characteristics, "Performance_Int" (interceptions) shows a positive and highly significant coefficient. This implies that fans react positively to defensive actions like interceptions, disrupting the opponent's play and contributing to team success. Conversely, "Performance_Touches" has a negative and highly significant coefficient, indicating that fans might not positively perceive an excessive number of ball touches. This could lead to a decrease in sentiment. "SCA_SCA" (likely related to shot-creating actions) also exhibits a positive and highly significant coefficient, suggesting that fans appreciate actions like crucial passes and dribbles leading to shots. Lastly, "age" has a small yet substantial negative coefficient, indicating that fans may prefer younger players or that the performance of older players might negatively impact sentiment.

Analysing the second model, which focuses on team characteristics concerning fan sentiment, we can discern that midfielders and attackers are relevant features. This aligns with our findings from the first model, suggesting that goals, assists, and progressive passes are the most essential individual statistics when fans evaluate players individually.

As expected, wins and losses are associated with apparent effects: victories are linked to improved ratings, while losses have the opposite effect. However, drawing a game is the most minor definitive outcome regarding fan sentiment. This aligns with the idea that drawn games are typically less emotionally charged, as they lack a clear result for both teams.

Finally, looking at all the variables together, we realize that we are looking at the best model, with an R-squared of 0.46, meaning that these variables explain 46% of the total variation of fan sentiment. This suggests that our model provides a reasonably robust explanation of fan sentiment based on player and team characteristics. However, it's important to note that the significance and impact of specific variables may vary based on the context and preferences of football fans.

Firstly, in the provided data, we observe that variables representing impactful events such as goals ("Performance_Gls") and assists ("Performance_Ast") do indeed have statistically significant positive coefficients. This suggests that these noteworthy events positively influence fans' assessments. Furthermore, these coefficients are relatively large, indicating that goals and assists substantially affect fans' evaluations of players.

On the other hand, variables related to routine performance metrics like "Passes_PrgP" (progressive passes) and "Performance_Int" (interceptions) also exhibit statistically significant coefficients. This implies that fans consider routine metrics when assessing players, albeit to a lesser degree than impactful events.

Based on the provided data and analysis, it is reasonable to say that Hypothesis 2 is partially supported. Fans' assessment of a football player's performance is influenced by impactful events and routine performance metrics, with the former having a slightly stronger effect. However, the relative importance of these factors can vary based on the specific context and individual preferences of fans. It's essential to acknowledge that football is a dynamic and subjective sport, and fans' assessments reflect a blend of impactful moments and consistent contributions.

4.3. Sentiment across Field Positions

Continuing our study of football fan sentiment, we shift our focus from athlete statistics to sentiment classification across field positions. This change allows us to examine how fans perceive players in more detail.

As such, we'll discuss our methodology and findings in sentiment classification. We aim to provide a deeper understanding of how fans view players in various positions on the soccer field. Again, we limited our analysis to athletes who had received a minimum of 100 positive and negative reviews each to ensure the consistency of our work.

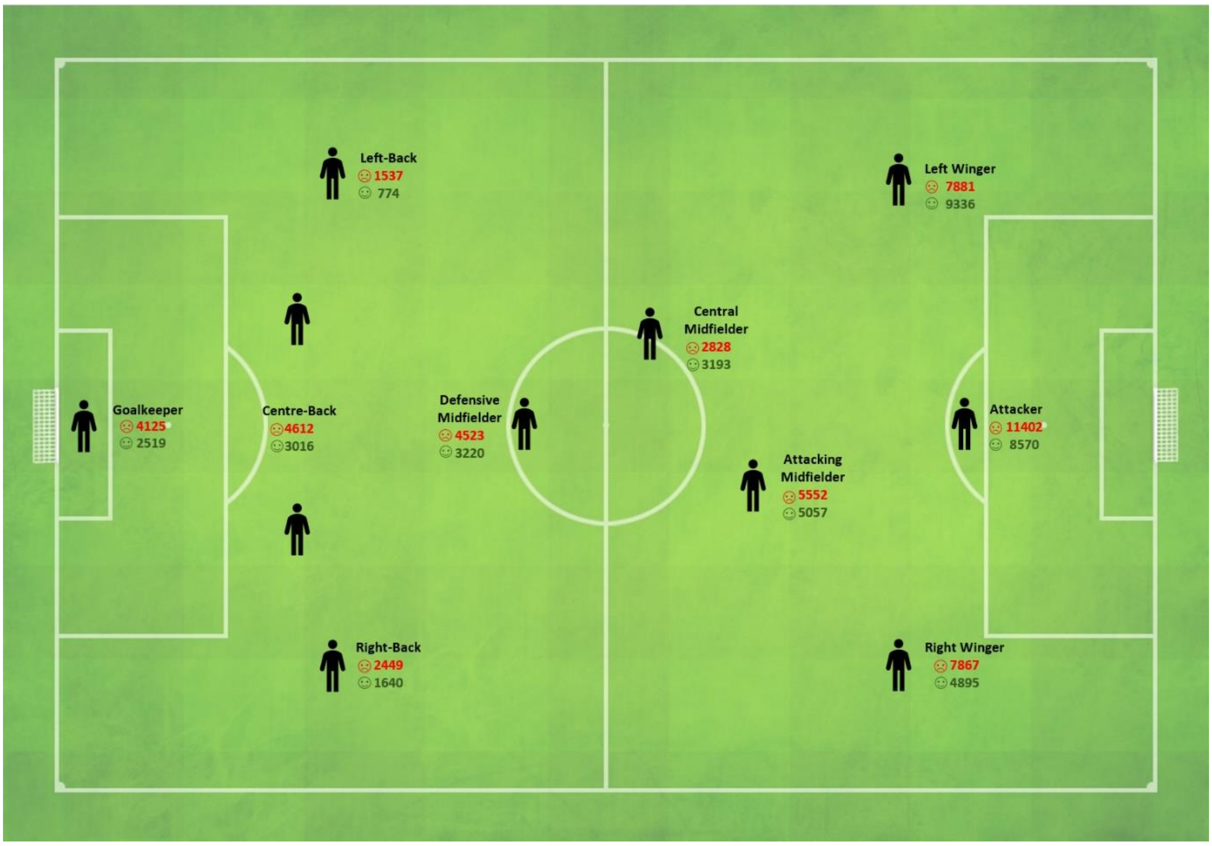


Figure 17 - Sentiment for Player Positions

The upcoming figure provides a visual representation of the distribution of negative and positive tweets, specifically excluding neutral comments, categorized by field positions.

As illustrated, attacking players stand out as the primary recipients of fan mentions, securing the top three spots regarding fan mentions. In corroboration, attacking midfielders secure the fourth position in fan reviews. Meanwhile, left and right backs receive comparatively fewer mentions in Twitter comments by fans. This discrepancy can be attributed to their limited involvement in pivotal game events such as goals, tackles, and saves, as we elaborated in the previous sub-chapter.

With a dataset comprising, including neutral tweets, 100,909 observations for attackers, 45,932 for midfielders, 27,247 for defenders, and 11,707 for goalkeepers, we employ the chi-square goodness of fit test (Magnello, 2005) to address our second pre-established hypothesis. In this analysis, the observed values align with those previously mentioned, and we designate the mean of reviews across the three positions as the expected frequencies. The computed chi-square value of 87,203 underscores a notable disparity between the observed and expected frequencies, affirming a substantial difference between the actual data and the anticipated distribution. Furthermore, the p-value < 0.05 leads us to reject the null hypothesis, conclusively confirming the presence of bias linked to the field position of players.

To build upon our earlier findings and considering our established evidence regarding the influence of player field positions on the number of mentions, we will now employ ANOVA to investigate whether significant variations exist in sentiment ratios among these player groups. The figure below offers a visual representation of the sentiment ratios among fans, in the form of violin plots.

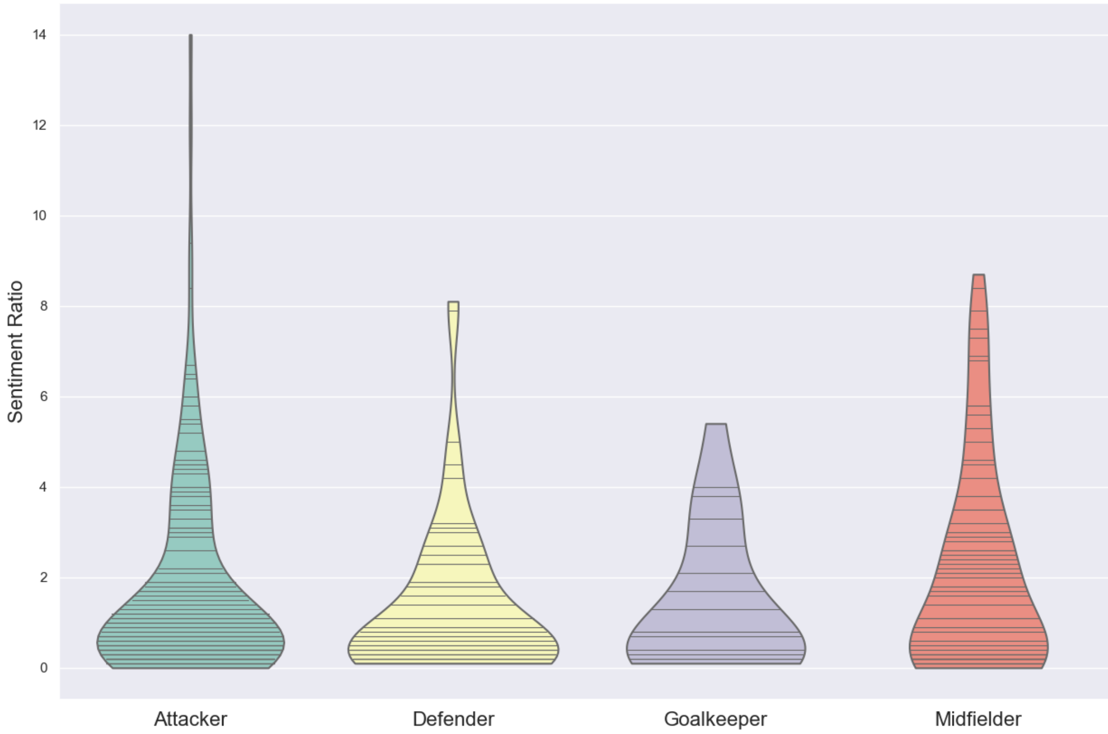


Figure 18 - Violin plots for positions

After conducting ANOVA on the position groups and visualizing the results using violin plots, we observe that the means of the groups appear to be relatively close to each other. This observation is supported by the ANOVA results, where we did not found a statistically-significant difference in average sentiment ratio according to field position ($F(3)=1.85$, $p > 0.1$). In conclusion, we reject the null hypothesis concerning the frequency of tweet mentions. However, we do not have sufficient evidence to reject the null hypothesis when assessing the mean variation of sentiment. This suggests that fans pay more attention to players in higher-pitch positions, mainly attacking players. Nevertheless, when evaluating performance, they tend to maintain a balanced and reasonable approach in their assessments.

5. CONCLUSIONS

This section will examine the topics and methodologies addressed in this study. We will summarize our findings, reevaluate our hypotheses, list our contributions, and propose potential directions for future research that align with the insights gained throughout this thesis.

In the initial part of the analysis, the correlation between player performance and fan sentiment was thoroughly examined—the utilization of Pearson's and Spearman's correlation coefficients allowed for a precise quantification of this relationship. The results indicated a positive correlation, suggesting that as player ratings on Fotmob and BBC websites increased, fan sentiment on Twitter also tended to become more positive. This statistical relationship was evident and highly significant, as the associated p-values were well below conventional thresholds, firmly supporting the hypothesis that player performance is correlated with fan sentiment.

However, it's important to acknowledge that the study's reliance on Twitter as the sole social network for analyzing fan sentiment may have limitations. Testing this hypothesis on other social networks, such as Reddit or Instagram, could provide a more comprehensive view of fan sentiment, as different platforms may exhibit varying patterns of sentiment expression.

Additionally, while the study focused on player ratings from Fotmob and BBC websites, it's crucial to recognize that there are numerous other online sources that fans may consult for player performance information. Expanding the sources of player ratings would help ensure a more comprehensive assessment of player performance and its relationship with fan sentiment.

Delving deeper into the scatter plots, distinct clusters of players emerged, offering valuable insights into the nuanced nature of the correlation. For instance, one cluster in the lower left quadrant signified players with subpar ratings, typically between 4 and 5, who received predominantly negative or neutral sentiment on Twitter. This observation showed a compelling connection between poor on-field performance and corresponding negative fan sentiment on social media.

Moving to the upper right quadrant, we encountered a distinct cluster of players with exceptional ratings, often exceeding 7, and markedly high sentiment ratios. These players enjoyed widespread acclaim and unwavering support from fans, reinforcing that top-performing individuals garner overwhelmingly positive sentiment.

Additionally, outliers in the scatter plot exhibited intriguing patterns. Some players with remarkably high sentiment ratios experienced exceptionally positive sentiment, while those with low ratings often faced negative sentiment on Twitter. These outliers added complexity to the analysis, suggesting that unique factors or narratives could drive fan sentiment in specific cases.

Transitioning to the exploration of factors influencing fan sentiment, multiple linear regression models were employed. These models revealed that impactful events such as goals and assists substantially shaped largely how fans perceived players. Fans valued these events, as indicated by the statistically significant and sizable coefficients associated with these variables. However, performance metrics also exhibit statistically significant coefficients, suggesting that fans are capable of forming a well-rounded perspective of the game. Interestingly, the market value of players harmed fan sentiment, with higher-valued players tending to receive more critical evaluations, possibly due to heightened expectations.

While its very relevant to address all the variables described in the model, it's important to acknowledge that these factors are not exhaustive. Other factors, such as team performance or off-field behaviour, could also impact fan sentiment. The model's explanatory power should be discussed, and any unexplained variance should be addressed.

Lastly, investigating sentiment across different field positions provided a comprehensive picture. Attacking players, were found to be the primary focus of fan comments, receiving the highest number of mentions. However, despite differences in the frequency of mentions, the sentiment ratios across player groups did not exhibit significant variations. This suggests that fans maintain a balanced and reasonable approach when evaluating players in different positions, ensuring fairness and objectivity in their assessments.

While the analysis suggests that fans maintain a balanced approach when evaluating players in different positions, it's worth noting that this finding aligns with the insights from the multiple linear regression model. The model highlighted the importance of various factors, including goals, assists, and performance metrics, in shaping fan sentiment. However, it appears that fans value characteristics beyond the main events of the match, often related to more attacking positions, in their assessments of players. This broader perspective contributes to fairer evaluations across all positions.

Yet, it's important to acknowledge the limitations of the analysis. The number of observations in the dataset may have constrained the depth of insights that can be drawn. With a larger dataset, it could be possible to conduct a more granular examination of how different field positions and their specific attributes influence fan sentiment. Additionally, the study could benefit from further research exploring how regional or cultural differences might impact fan sentiment related to player positions.

In conclusion, quantifying an athlete's performance through the lens of fan sentiment represents an innovative approach to performance evaluation. This methodology offers a unique perspective, revealing the intimate connection between athletes' on-field achievements and the emotional responses they evoke from their fan base. While this study has addressed the constraints imposed by Twitter's API limitations and certain algorithmic challenges, it also illuminates exciting possibilities for future research. These opportunities include analyzing full-season tweets, exploring sentiment variations across different sports and regions, streamlining fan sentiment data collection processes, and delving into performance assessments tailored to diverse countries and fan demographics. The framework presented here not only contributes to our understanding of the intricate relationship between athletes, fans, and sentiment but also paves the way for a more comprehensive and dynamic approach to performance assessment in the ever-evolving landscape of sports analytics.

6. BIBLIOGRAPHY

Acheampong, F. A., Wenyu, C., & Nunoo-Mensah, H. (2020). Text-based emotion detection: Advances, challenges, and opportunities. *Engineering Reports*, *2*(7), e12189. https://doi.org/10.1002/eng2.12189

Aloufi, S., Alzamzami, F., Hoda, M., & El Saddik, A. (2018). Soccer Fans Sentiment through the Eye of Big Data: The UEFA Champions League as a Case Study. *2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*, 244–250. https://doi.org/10.1109/MIPR.2018.00058

Aloufi, S., & El Saddik, A. (2018). Sentiment Identification in Football-Specific Tweets. *IEEE Access*, *PP*, 1–1. https://doi.org/10.1109/ACCESS.2018.2885117

Anstead, N., & O'Loughlin, B. (2015). Social Media Analysis and Public Opinion: The 2010 Uk General Election. *Journal of Computer-Mediated Communication*, *20*(2), 204–220. https://doi.org/10.1111/jcc4.12102

Barbieri, F., Camacho-Collados, J., Neves, L., & Espinosa-Anke, L. (2020). *TweetEval: Unified Benchmark and Comparative Evaluation for Tweet Classification* (arXiv:2010.12421). arXiv. http://arxiv.org/abs/2010.12421

Barnaghi, P., Ghaffari, P., & Breslin, J. G. (2016). Opinion Mining and Sentiment Polarity on Twitter and Correlation between Events and Sentiment. *2016 IEEE Second International Conference on Big Data Computing Service and Applications (BigDataService)*, 52–57. https://doi.org/10.1109/BigDataService.2016.36

Berber, E., McLean, S., Beanland, V., Read, G. J. M., & Salmon, P. M. (2020). Defining the attributes for specific playing positions in football match-play: A complex systems approach. *Journal of Sports Sciences*, *38*(11–12), 1248–1258. https://doi.org/10.1080/02640414.2020.1768636

Carmichael, M. A., Thomson, R. L., Moran, L. J., & Wycherley, T. P. (2021). The Impact of Menstrual Cycle Phase on Athletes' Performance: A Narrative Review. *International Journal of Environmental Research and Public Health*, *18*(4), Article 4. https://doi.org/10.3390/ijerph18041667

Castellano, J., Casamichana, D., & Lago, C. (2012). The Use of Match Statistics that Discriminate Between Successful and Unsuccessful Soccer Teams. *Journal of Human Kinetics*, *31*(2012), 137–147. https://doi.org/10.2478/v10078-012-0015-7

Coates, D., & Parshakov, P. (2022). The wisdom of crowds and transfer market values. *European Journal of Operational Research*, *301*(2), 523–534. https://doi.org/10.1016/j.ejor.2021.10.046

Decroos, T., Bransen, L., Van Haaren, J., & Davis, J. (2019). Actions Speak Louder than Goals: Valuing Player Actions in Soccer. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, 1851–1861. https://doi.org/10.1145/3292500.3330758

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). *BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding* (arXiv:1810.04805). arXiv. https://doi.org/10.48550/arXiv.1810.04805

Dreyer, F., Greif, J., Günther, K., Spiliopoulou, M., & Niemann, U. (2022). Data-Driven Prediction of Athletes' Performance Based on Their Social Media Presence. In P. Pascal & D. Ienco (Eds.), *Discovery Science* (pp. 197–211). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-18840-4_15

Duch, J., Waitzman, J. S., & Amaral, L. A. N. (2010). Quantifying the Performance of Individual Players in a Team Activity. *PLOS ONE*, *5*(6), e10937. https://doi.org/10.1371/journal.pone.0010937

Errasti, J., Amigo, I., & Villadangos, M. (2017). Emotional Uses of Facebook and Twitter: Its Relation With Empathy, Narcissism, and Self-Esteem in Adolescence. *Psychological Reports*, *120*(6), 997–1018. https://doi.org/10.1177/0033294117713496

Gaglio, S., Lo Re, G., & Morana, M. (2016). A framework for real-time Twitter data analysis. *Computer Communications*, *73*, 236–242. https://doi.org/10.1016/j.comcom.2015.09.021

Gruettner, A., Vitisvorakarn, M., Wambsganss, T., Rietsche, R., & Back, A. (2020). *The New Window to Athletes' Soul – What Social Media Tells Us About Athletes' Performances*. http://hdl.handle.net/10125/64045

Hasan, A., Moin, S., Karim, A., & Shamshirband, S. (2018). Machine Learning-Based Sentiment Analysis for Twitter Accounts. *Mathematical and Computational Applications*, *23*(1), Article 1. https://doi.org/10.3390/mca23010011

Hasan, M., Orgun, M. A., & Schwitter, R. (2019). Real-time event detection from the Twitter data stream using the TwitterNews+ Framework. *Information Processing & Management*, *56*(3), 1146–1165. https://doi.org/10.1016/j.ipm.2018.03.001

Hauke, J., & Kossowski, T. (2011). Comparison of Values of Pearson's and Spearman's Correlation Coefficients on the Same Sets of Data. *QUAGEO*, *30*(2), 87–93. https://doi.org/10.2478/v10117-011-0021-1

Herm, S., Callsen-Bracker, H.-M., & Kreis, H. (2014). When the crowd evaluates soccer players' market values: Accuracy and evaluation attributes of an online community. *Sport Management Review*, *17*(4), 484–492. https://doi.org/10.1016/j.smr.2013.12.006

Hidayatullah, A. F., Pembrani, E. C., Kurniawan, W., Akbar, G., & Pranata, R. (2018). Twitter Topic Modeling on Football News. *2018 3rd International Conference on Computer and Communication Systems (ICCCS)*, 467–471. https://doi.org/10.1109/CCOMS.2018.8463231

Hughes, M. D., & Bartlett, R. M. (2002). The use of performance indicators in performance analysis. *Journal of Sports Sciences*, *20*(10), 739–754. https://doi.org/10.1080/026404102320675602

Hull, K., & Lewis, N. P. (2014). Why Twitter Displaces Broadcast Sports Media: A Model. *International Journal of Sport Communication*, *7*(1), 16–33. https://doi.org/10.1123/IJSC.2013-0093

Jai-Andaloussi, S., Mourabit, I. E., Madrane, N., Chaouni, S. B., & Sekkaki, A. (2015). Soccer Events Summarization by Using Sentiment Analysis. *2015 International Conference on Computational Science and Computational Intelligence (CSCI)*, 398–403. https://doi.org/10.1109/CSCI.2015.59

Kwak, H., Lee, C., Park, H., & Moon, S. (2010). What is Twitter, a social network or a news media? *Proceedings of the 19th International Conference on World Wide Web*, 591–600. https://doi.org/10.1145/1772690.1772751

Lanagan, J., & Smeaton, A. (2011). Using Twitter to Detect and Tag Important Events in Sports Media. *Proceedings of the International AAAI Conference on Web and Social Media*, *5*(1), Article 1. https://doi.org/10.1609/icwsm.v5i1.14170

Larsen, M. E., Boonstra, T. W., Batterham, P. J., O'Dea, B., Paris, C., & Christensen, H. (2015). We Feel: Mapping Emotion on Twitter. *IEEE Journal of Biomedical and Health Informatics*, *19*(4), 1246–1252. https://doi.org/10.1109/JBHI.2015.2403839

Li, Q., Zhang, J., Guo, J., Li, J., & Kang, C. (2022). Evaluating Performance of NBA Players with Sentiment Analysis on Twitter Messages. *2021 2nd European Symposium on Software Engineering*, 150–155. https://doi.org/10.1145/3501774.3501796

Liao, S., Wang, J., Yu, R., Sato, K., & Cheng, Z. (2017). CNN for situations understanding based on sentiment analysis of twitter data. *Procedia Computer Science*, *111*, 376–381. https://doi.org/10.1016/j.procs.2017.06.037

Liu, H., Gómez, M.-A., Gonçalves, B., & Sampaio, J. (2016). Technical performance and match-to-match variation in elite football teams. *Journal of Sports Sciences*, *34*(6), 509–518. https://doi.org/10.1080/02640414.2015.1117121

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). *RoBERTa: A Robustly Optimized BERT Pretraining Approach* (arXiv:1907.11692). arXiv. https://doi.org/10.48550/arXiv.1907.11692

Lopes Dos Santos, M., Uftring, M., Stahl, C. A., Lockie, R. G., Alvar, B., Mann, J. B., & Dawes, J. J. (2020). Stress in Academic and Athletic Performance in Collegiate Athletes: A Narrative Review of Sources and Monitoring Strategies. *Frontiers in Sports and Active Living*, *2*. https://www.frontiersin.org/articles/10.3389/fspor.2020.00042

Lucas, G. M., Gratch, J., Malandrakis, N., Szablowski, E., Fessler, E., & Nichols, J. (2017). GOAALLL!: Using sentiment in the world cup to explore theories of emotion. *Image and Vision Computing*, *65*, 58–65. https://doi.org/10.1016/j.imavis.2017.01.006

Magnello, M. E. (2005). Chapter 56—Karl Pearson, paper on the chi square goodness of fit test (1900). In I. Grattan-Guinness, R. Cooke, L. Corry, P. Crépel, & N. Guicciardini (Eds.), *Landmark Writings in Western Mathematics 1640-1940* (pp. 724–731). Elsevier Science. https://doi.org/10.1016/B978-044450871-3/50137-6

Mannes, A. E., Soll, J. B., & Larrick, R. P. (2014). The wisdom of select crowds. *Journal of Personality and Social Psychology*, *107*, 276–299. https://doi.org/10.1037/a0036677

Miranda-Peña, C., Ceballos, H. G., Hervert-Escobar, L., & Gonzalez-Mendoza, M. (2021). Predicting Soccer Results Through Sentiment Analysis: A Graph Theory Approach. In M. Paszynski, D. Kranzlmüller, V. V. Krzhizhanovskaya, J. J. Dongarra, & P. M. A. Sloot (Eds.), *Computational Science – ICCS 2021* (pp. 422–435). Springer International Publishing. https://doi.org/10.1007/978-3-030-77980-1_32

Naskar, D., Singh, S. R., Kumar, D., Nandi, S., & Rivaherrera, E. O. de la. (2020). Emotion Dynamics of Public Opinions on Twitter. *ACM Transactions on Information Systems*, *38*(2), 18:1-18:24. https://doi.org/10.1145/3379340

Oorschot, G., Erp, M., & Dijkshoorn, C. (2012). Automatic Extraction of Soccer Game Events from Twitter. *CEUR Workshop Proceedings*, *902*.

Peeters, G., & Czapinski, J. (1990). Positive-Negative Asymmetry in Evaluations: The Distinction Between Affective and Informational Negativity Effects. *European Review of Social Psychology*, *1*(1), 33–60. https://doi.org/10.1080/14792779108401856

Peeters, T. (2018). Testing the Wisdom of Crowds in the field: Transfermarkt valuations and international soccer results. *International Journal of Forecasting*, *34*(1), 17–29. https://doi.org/10.1016/j.ijforecast.2017.08.002

Price, J., Farrington, N., & Hall, L. (2013). Changing the game? The impact of Twitter on relationships between football clubs, supporters and the sports media. *Soccer & Society*, *14*(4), 446–461. https://doi.org/10.1080/14660970.2013.810431

Reilly, T., Drust, B., & Clarke, N. (2008). Muscle Fatigue during Football Match-Play. *Sports Medicine*, *38*(5), 357–367. https://doi.org/10.2165/00007256-200838050-00001

Rice, S. M., Purcell, R., De Silva, S., Mawren, D., McGorry, P. D., & Parker, A. G. (2016). The Mental Health of Elite Athletes: A Narrative Systematic Review. *Sports Medicine*, *46*(9), 1333–1353. https://doi.org/10.1007/s40279-016-0492-2

Roberts, K., Roach, M. A., Johnson, J., Guthrie, J., & Harabagiu, S. M. (2012). EmpaTweet: Annotating and Detecting Emotions on Twitter. *Proceedings of the Eighth International*

Conference on Language Resources and Evaluation (LREC'12), 3806–3813. http://www.lrec-conf.org/proceedings/lrec2012/pdf/201_Paper.pdf

Robertson, S., Gupta, R., & McIntosh, S. (2016). A method to assess the influence of individual player performance distribution on match outcome in team sports. *Journal of Sports Sciences*, *34*(19), 1893–1900. https://doi.org/10.1080/02640414.2016.1142106

Ruz, G. A., Henríquez, P. A., & Mascareño, A. (2020). Sentiment analysis of Twitter data during critical events through Bayesian networks classifiers. *Future Generation Computer Systems*, *106*, 92–104. https://doi.org/10.1016/j.future.2020.01.005

Sakaki, T., Okazaki, M., & Matsuo, Y. (2010). Earthquake shakes Twitter users: Real-time event detection by social sensors. *Proceedings of the 19th International Conference on World Wide Web*, 851–860. https://doi.org/10.1145/1772690.1772777

Saw, A. E., Main, L. C., & Gastin, P. B. (2016). Monitoring the athlete training response: Subjective self-reported measures trump commonly used objective measures: a systematic review. *British Journal of Sports Medicine*, *50*(5), 281–291. https://doi.org/10.1136/bjsports-2015-094758

Schumaker, R. P., Jarmoszko, A. T., & Labedz, C. S. (2016). Predicting wins and spread in the Premier League using a sentiment analysis of twitter. *Decision Support Systems*, *88*, 76–84. https://doi.org/10.1016/j.dss.2016.05.010

Siemon, D., & Wessels, J. (2022). Performance prediction of basketball players using automated personality mining with twitter data. *Sport, Business and Management: An International Journal*, *ahead-of-print*(ahead-of-print). https://doi.org/10.1108/SBM-10-2021-0119

Sinha, S., Dyer, C., Gimpel, K., & Smith, N. (2013). Predicting the NFL using Twitter. *Proc. ECML/PKDD Workshop on Machine Learning and Data Mining for Sports Analytics*.

Su, Y., Baker, B., Doyle, J. P., & Kunkel, T. (2020). *Rise of an athlete brand: Factors influencing the social media following of athletes*. https://doi.org/10.34944/dspace/7092

Surowiecki, J. (2005). *The Wisdom of Crowds*. Knopf Doubleday Publishing Group.

T, K., Sekaran, K., D, R., V, V. K., & M, B. J. (2019). Personalized Content Extraction and Text Classification Using Effective Web Scraping Techniques. *International Journal of Web Portals (IJWP)*, *11*(2), 41–52. https://doi.org/10.4018/IJWP.2019070103

Takeichi, Y., Sasahara, K., Suzuki, R., & Arita, T. (2014). *Twitter as Social Sensor: Dynamics and Structure in Major Sporting Events*. 778–784. https://doi.org/10.1162/978-0-262-32621-6-ch126

Talha, A., Simsek, M., & Belenli, I. (2018). The Wisdom of the Silent Crowd: Predicting the Match Results of World Cup 2018 through Twitter. *International Journal of Computer Applications*, *182*(27), 40–45. https://doi.org/10.5120/ijca2018918125

Thelwall, M., Buckley, K., & Paltoglou, G. (2012). Sentiment strength detection for the social web. *Journal of the American Society for Information Science and Technology*, *63*(1), 163–173. https://doi.org/10.1002/asi.21662

Vale, L., & Fernandes, T. (2018). Social media and sports: Driving fan engagement with football clubs on Facebook. *Journal of Strategic Marketing*, *26*(1), 37–55. https://doi.org/10.1080/0965254X.2017.1359655

Wang, W., Chen, L., Thirunarayan, K., & Sheth, A. P. (2012). Harnessing Twitter 'Big Data' for Automatic Emotion Identification. *2012 International Conference on Privacy, Security, Risk and Trust and 2012 International Confernece on Social Computing*, 587–592. https://doi.org/10.1109/SocialCom-PASSAT.2012.119

Weng, J., & Lee, B.-S. (2011). Event Detection in Twitter. *Proceedings of the International AAAI Conference on Web and Social Media*, *5*(1), Article 1. https://doi.org/10.1609/icwsm.v5i1.14102

Wunderlich, F., & Memmert, D. (2021). A big data analysis of Twitter data during premier league matches: Do tweets contain information valuable for in-play forecasting of goals in football? *Social Network Analysis and Mining*, *12*(1), 23. https://doi.org/10.1007/s13278-021-00842-z

Xue, J., Chen, J., Hu, R., Chen, C., Zheng, C., Su, Y., & Zhu, T. (2020). Twitter Discussions and Emotions About the COVID-19 Pandemic: Machine Learning Approach. *Journal of Medical Internet Research*, *22*(11), e20550. https://doi.org/10.2196/20550

Yu, Y., & Wang, X. (2015). World Cup 2014 in the Twitter World: A big data analysis of sentiments in U.S. sports fans' tweets. *Computers in Human Behavior*, *48*, 392–400. https://doi.org/10.1016/j.chb.2015.01.075

Yujian, L., & Bo, L. (2007). A Normalized Levenshtein Distance Metric. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *29*(6), 1091–1095. https://doi.org/10.1109/TPAMI.2007.1078