Università
Ca'Foscari
Venezia

Master's Degree

in Economics, Finance and Sustainability

Quantitative Finance and Risk Management

Final Thesis

Innovative Financial Strategies in Football: A Quantitative Analysis of Performance, Investment and Risk

Graduand
Sebastiano Gasparini (896810)

Supervisor
Antonella Basso

Academic Year 2023-2024

TABLE OF CONTENTS

INTRODUCTION

The subject of this thesis is the analysis of the introduction of modern and innovative technologies in the world of football, analysing both the technical and economic-financial aspects.

The reason why I chose to develop this topic is the intertwining of my greatest passions: technology, football and finance. I also believe that the impact of modern technology in people's daily lives is now more than ever a hot topic.

This thesis is divided mainly into two parts. The first part, which includes the first two chapters, mainly deals with the economic and financial aspects of modern football. The business models that exist today in the management of a football club, its revenues, its expenses and how these influence performance on the pitch will be analysed. I will also propose my personal vision of the modern football landscape, where I will offer food for thought on the problems of the current system and possible solutions.

The second part focuses on the more technological aspect of the thesis. These two chapters (chapter 3 and chapter 4) are connected but very different from each other. In the third chapter I will describe the enormous impact that big data and artificial intelligence have had on the world of football, both on the pitch and in terms of club management. I will highlight both the positive aspects of this impact (also through real case studies) and the negative aspects and limitations. Even in this case I will offer some personal ideas to reflect on.

The fourth and final chapter of the thesis, instead, (before the final conclusions) is an original and completely personal work. In fact, I will present an econometric model developed and built by myself, with the help of the RSTUDIO software, for the prediction of the market valuation of a football player.

CHAPTER I

EVOLUTION OF BUSINESS MODELS IN FOOTBALL CLUBS

1.1 From Patronage to Entrepreneurship

In their most primitive forms, football clubs were often community-centred entities, supplemented by locals who were enthusiasts. Driven by a passion for the game and a sense of community pride, early supporters provided the necessary financial backing for clubs. The patronage era remained, therefore, community- and sport-driven rather than profit-driven. It was often local businessmen and wealthy individuals, driven by personal passion and community status, who played important roles in sustaining football clubs. In Italy, one of the last concrete examples of patronage is Berlusconi's AC Milan[1], an entrepreneur with a strong passion for the game of football who was the longest-serving president of one of the most important clubs in football history, remaining its owner from 20 February 1986 to 13 April 2017.

Over the late years of the 20th century, a number of factors began reshaping football club ownership. The post-war boom[2], new developments in the field of broadcasting, and rising globalization meant that the game was now presenting ever-growing commercial opportunities. It was an absolute godsend when television broadcasting was introduced into the mix. Its advent changed everything for football, moving it directly into a millions-strong fan base worldwide. This

[1] Borghini, A., & Baldini, A. (2011). On the logic of soccer patronage. Soccer & Society, 12(5), 569–585

[2] I find this reading very interesting about it: Bummer, Shaun, "History on the Pitch: The Social and Economic Impact of Professional Soccer in Postwar London" (2015)

additional exposure did a great deal more than just spread the fan base; it also attracted advertising money in very large amounts.

Commercialization of football therefore increased in the second half of the 20th century[3]. Clubs realized the huge revenue-making potentials of selling broadcast rights, sponsorship, and merchandising. With increasing financial stakes, the requirement for professional management and strategic planning has become of paramount importance. An owner's role became transformational, from being patrons to strategic entrepreneurs who run a football club as business ventures.

With the turn of the 21st century, there was an immense turn towards corporate ownership and international investors. High-profile acquisitions by foreign billionaires and investment groups are now becoming reality, especially in the top European leagues. This new breed of owners brought a distinctly corporate mentality, focused on maximizing revenue and global reach and introducing advanced management practices. Most of the clubs were altered due to the rush of foreign capital that made them invest heavily in purchasing players and building state-of-the-art facilities and running marketing campaigns all around the world.

This shift in entrepreneurship and corporate ownership simply changes everything. On the one hand, new finance can enable clubs to become more competitive, earn the finest talent from anywhere in the world, and reach a worldwide audience. On the other hand, it has brought about problems associated with disparities in financial muscle between clubs, pressure for immediate results, and risks of financial instability.

[3] Busse Ronald and Damiano Jean-Pierre, “The Role of Commercialisation of the European Football Business for the Emotional Bond between Fans and Clubs”

The new landscape of football club management today will thus be informed by a review of the historical evolution that defined the ownership models. This narrative, shifting from patronage to entrepreneurship, is about economic change and broader social trends. Thereupon, with football clubs still trying to navigate the complexities of modern economics, lessons learned from its historical development offer valuable insights that may set a future direction of the game.

1.2 Ownership Models and Revenue Sources

The current scene has therefore been shaped by the transformation from patronage to entrepreneurship, bringing major changes in the ownership of football clubs, from ownership structures to the sophisticated revenue strategies of football clubs. As clubs transformed from passion projects of wealthy patrons into key strategic assets in global corporate portfolios, so did their financial models and revenue streams. Within this section, modern ownership models of football facilities and the main sources of revenue that maintain their financial well-being will be considered.

1.2.1 Ownership Models

1. Corporate Model

As the commercialization of football became evident, corporate ownership came to the fore. Companies rushed to acquire a football team, as it promised immense profits through broadcast rights, merchandising and global brand development.

This ushered in a new, more corporate approach to club management, with a strong focus on financial performance and strategic growth.

These, in turn, bring professional management practices, significant financial resources and extensive marketing knowledge and experience. One of the advantages of such combinations is the ease with which large investments in players, world-class facilities and extensive global marketing can be financed. Companies can exploit marketing and branding synergies by integrating a football club into the corporate portfolio, improving the club's commercial attractiveness and profitability.

2. Investment Groups and Consortia

Along with corporate ownership, the role of investment groups and consortia in football has become increasingly important. By pooling resources through wealthy businessmen, private equity firms and institutional investors it is possible to invest in and manage clubs. These investment groups bring a good share of expertise and capital to clubs, thus helping them to implement large projects and pursue ambitious strategic objectives.

The involvement of investment groups often leads to a focus on maximizing the club's financial performance and market value. Strategies include expanding the club's global fan base, optimizing commercial partnerships and developing new revenue streams. Investment groups are also starting to use data-driven decision-making and modern management techniques to improve both on-the-ground performance and business operations.

3. Fan Ownership

While most clubs are profit-driven models, some have embraced fan-ownership structures where fans directly own the club's shares. This is strongly favoured by the provisions of the 50+1 rule which guarantees that more than 50% of the voting rights are held by club members to protect the interests of the setters, this is especially the case in countries such as Germany where majority control of a single entity (person or company) is not permitted by the Deutsche Fußball Liga; shares of very important teams such as Borussia Dortmund are in fact traded on the German stock market and are widely owned by fans.

Fan-owned clubs prioritize community engagement, democratic governance and long-term sustainability over immediate financial returns. This model fosters strong loyalty and a deep bond between the club and its supporters, which translates into strong fan support and stable revenues from ticket sales and membership fees.

1.2.2 Revenue Sources

The financial viability of football clubs depends on a number of revenue streams, each of which contributes to the overall economic health of the organisation. The main sources of revenue for football clubs include broadcasting rights, sponsorship, matchday income, merchandising and player transfers[4].

[4] Gravina G. (2011). Il bilancio d'esercizio e l'analisi della performance nelle società di calcio professionistiche. Milano: FrancoAngeli, 40.

Broadcasting Rights

Broadcasting rights therefore represent one of the most important sources of income for football clubs, particularly those in the top leagues. It is through the excitement they bring to millions of people around the world that television networks and digital platforms pay huge sums for the rights to broadcast live matches. These agreements, which often last several years, include both domestic and international rights and provide a steady and substantial revenue stream.

Increasing competition between broadcasters and the rise of streaming services have further increased the value of broadcast rights. Clubs in leagues with lucrative broadcast deals, such as the English Premier League, benefit immensely, allowing them to invest heavily in player acquisitions, facilities and other strategic initiatives.

In the Italian Serie A in the 2022/2023 season the total revenues from TV rights were around one billion euros[5], which is then divided among the 20 teams belonging to the league in a meritocratic way based on the final position in the standings, the audiences of the individual matches and the history of the club.

Sponsorship and Commercial Partnerships

Commercial revenues represent the most important source of income for football clubs, they come from sponsorships and commercial agreements. These are major companies that pay to have their brands associated with the club, usually through shirt sponsorships, stadium naming rights, advertising on in-stadium billboards and other types of commercial deals.

[5] https://onefootball.com/en/news/cf-how-much-milan-and-other-serie-a-clubs-earned-from-tv-rights-in-2022-23-37672831

In fact, nowadays it is almost impossible to see a football shirt without sponsor patches. These deals are much more lucrative for clubs with a huge global following and large fan base. In fact, in 2022, six teams exceeded 300 million euros in revenue from this revenue section[6].

Matchday Revenue

Matchday revenue encompasses income generated from ticket sales, hospitality packages, and related activities on match days. This includes not only the ticket sales for regular season matches but also for cup competitions, friendlies, and other events hosted at the club's stadium.

With large stadiums and high fan representation, many clubs can generate a significant portion of their income from matchday revenue. Further investments in the stadium, particularly in the VIP boxes and areas that were developed into premium seating, have significantly enhanced the earnings made by clubs from stadium revenues. Moreover, clubs can earn ancillary revenues through food, beverage, and merchandise sales on match days.

Merchandising

Merchandising is another significant revenue stream, involving the sale of branded products such as jerseys, scarves, hats, and other memorabilia. Successful merchandising strategies leverage the club's brand and identity, appealing to fans' loyalty and enthusiasm.

[6] https://www.footballbenchmark.com/library/commercial_deal_landscape_in_the_big_five_leagues

The globalization of football has expanded merchandising opportunities beyond local markets. Clubs now reach international fans through online stores and global distribution networks. Limited edition items, player-specific merchandise, and collaborative collections with well-known brands have become popular ways for clubs to boost their merchandising revenue.

Player Transfers

The transfer market is a unique and dynamic source of income for football clubs. Profits from player transfers can be significant, especially for clubs with a strong scouting network and robust youth academies. Buying young talent at a low cost, developing them and selling them at a higher value is a common strategy used by many clubs to generate revenue. In Italy, for example, Atalanta (a team from Bergamo) has generated very strong revenues in recent years thanks to talents recognized throughout the world produced by their youth sector[7].

Beyond that, clubs conduct player swap and loan deals, which could yield both short- and long-term returns for strategic advantages. Transfer market fluctuations and risks associated with player performance make this revenue stream highly variable and dependent on effective talent management and market conditions.

This section has discussed the ownership models and revenue sources underlying football clubs in order to introduce the main financial strategies and performance metrics relevant to modern football industries. These insights will be further

[7] Neri, L., Russo, A., Di Domizio, M., & Rossi, G. (2021). "Football players and asset manipulation: the management of football transfers in Italian Serie A." European Sport Management Quarterly, 23(4), 942–962.

developed in Chapter 2, where we will examine in more depth the interplay between financial investments, performance outcomes and overall club sustainability.

CHAPTER II

FINANCES, PERFORMANCE AND EXPENDITURE OF FOOTBALL CLUBS

2.1 The Functioning of the Transfer Market and Club Expenditures

The transfer market is a very dynamic market and is a fundamental part of modern football; it in fact shapes the composition of the teams and influences the financial health of the clubs. This section delves into the complex workings of the transfer market and examines how club spending on transfers impacts both immediate and long-term performance. I decided to delve deeper into this part as the transfer market represents a club's largest expense[8], as well as being a very interesting element as it is characterized by great volatility.

2.1.1 The Transfer Market: Structure and Operations

The transfer market operates within clearly defined periods, known as transfer windows, during which clubs can buy, sell or loan players. These windows, in Europe, are in Summer and Winter (the latter known as the repair market) and are regulated by football governing bodies such as FIFA[9]. The timing and duration of these windows are

[8] https://www.footballbenchmark.com/data_analytics/starter/club_finance

[9] «The Fédération Internationale de Football Association (FIFA) is the international self-regulatory governing body of association football, beach soccer, and futsal»; https://en.wikipedia.org/wiki/FIFA

strategically designed to align with the competitive calendar (National Championship and European Championship).

During these periods, clubs enter into negotiations to acquire or sell players. These negotiations are mainly based on two phases. The first phase is the agreement with the selling club and the payment of the player's transfer fees, i.e. a sum paid by the buying club to the selling club in exchange for the player's membership and therefore the "ownership" of the latter. Transfer fees vary widely and are influenced by several factors such as the player's age, presumed skill level, remaining contract length with the club he plays for and market demand. The second phase is direct negotiation with the player in question regarding aspects such as salary, bonuses (which are rewards usually based on future performance) and contract duration. Agents and intermediaries often facilitate these negotiations, representing the interests of the players and ensuring that the deals are beneficial to all parties involved.

A club's transfer spending is a strategic decision made to improve the team's performance in achieving certain objectives. Financially strong clubs, which typically have wealthy owners or high revenues, can afford to spend large sums on renowned players. Such clubs seek major signings not only to improve immediate on-field performance, but also for global branding purposes to attract sponsors and sell merchandise.

For example, Manchester City, Paris Saint-Germain and other world-famous clubs have made headlines in recent years due to their spending sprees on many star players. The reason for this investment was immediate competitive gains which elevated the club's status in national and international competitions. This approach, however, hides risks of financial tension and underperformance of the players purchased, generating capital losses over the years.

On the contrary, clubs with limited financial resources become more cautious and strategic. In fact, these clubs base their market on the attention to scouting undervalued talents and investing in young people. By spending less by investing in young players and potential talents at lower prices, these clubs can build competitive teams over time, while also often generating large capital gains from future resales. This model greatly reduces potential financial risks and allows clubs to enjoy greater financial stability over the years.

2.1.2 Financial Risks and Implications

As mentioned above, spending in the transfer market carries significant financial risks. Transfer fees and high player salaries can put a strain on a club's budget, especially if the players acquired do not meet expectations. Such financial commitments can lead to long-term indebtedness and instability, as seen in recent years with clubs facing financial difficulties due to excessive spending or in some cases even administrative sanctions from leagues football.

Furthermore, the value of players is subject to market fluctuations. Factors such as injuries, form and external market conditions can cause significant fluctuations in a player's market value. Clubs must therefore be cautious and carry out rigorous analysis before entering the transfer market, to avoid overpaying for players whose value could decline in the years to come.

Despite these risks, the benefits of successfully navigating the transfer market can be considerable. Effective player acquisitions can completely transform a club's coffers, leading to better results, higher league positions and increased revenue from prize money, broadcast rights and greater fan engagement. The challenge is to balance ambitious spending with financial prudence, ensuring investments are sustainable and aligned with the club's long-term strategy.

The functioning of the transfer market is also influenced by market trends and imposed regulatory frameworks. Economic conditions, such as exchange rates and macroeconomic stability, influence the spending power of clubs and the overall valuation of players. Furthermore, there are regulatory measures such as Financial Fair Play (FFP) imposed by governing bodies such as UEFA[10] which aim to ensure financial discipline and sustainability within the sport, seeking to mitigate the gap that can be created between teams richer and poorer.

FFP regulations require clubs to balance expenses with income, preventing them from taking on excessive debt. Clubs that breach these rules risk sanctions, including fines, transfer bans and exclusion from competitions. These regulations encourage clubs to adopt more sustainable financial practices, ensuring that transfer market activities are conducted within the bounds of financial prudence.

2.1.3 Data, Analytics and Technological Innovations

In recent years, data and analytics have revolutionized decision-making in the transfer market. Clubs are increasingly relying on advanced statistical models and performance

---

[10] *«The **Union of European Football Associations** is one of six continental bodies of governance in football. It governs football, futsal and beach football in Europe.»* https://en.wikipedia.org/wiki/UEFA

metrics to guide their transfer strategies. These analytical tools help identify players who offer the best value for money, predict future performance and reduce the risks associated with high-cost transfers.

Data-driven approaches extend beyond player acquisition to include scouting and talent development. Clubs use sophisticated software and analytical techniques to monitor player performance across various leagues and competitions. By analysing attributes such as speed, stamina, tactical awareness and injury history, clubs can make informed decisions about potential signings and youth prospects.

Digital platforms and online databases provide clubs with comprehensive information on player performances, contract details and market valuations. These tools facilitate efficient scouting, negotiation and decision-making processes.

Innovations like artificial intelligence and machine learning are also starting to play a role in predicting player potential and market trends. Clubs use these technologies to develop sophisticated financial models and algorithms, optimizing their transfer strategies and improving their competitive advantage.

However, we will cover all this in more depth in the next chapter and I will then provide a practical example in the last chapter.

The transfer market is a complex and vital aspect of managing a football club, with significant implications for both financial performance and success on the pitch. By understanding the structure and mechanisms of the transfer market and using strategic spending and data-driven decision making, clubs can navigate this landscape effectively. This chapter has highlighted the multifaceted nature of the transfer market and the critical importance of balanced financial management, laying the foundation for a more in-depth exploration of the correlation between financial investments and sporting performance in subsequent sections.

2.2 Correlation Between Financial Investment and Sports Performance

Investment in player acquisition is perhaps the most visible manifestation of a club's financial commitment to achieving competitive success. High-profile transfers often dominate the headlines, reflecting the belief that acquiring top talent is essential to winning games and trophies. However, this is not always what happens on the football pitch.

It's true that clubs that invest heavily in acquiring elite players generally perform better on the pitch as they bring superior skill, experience and tactical awareness, improving the overall quality of the team. However, success does not depend solely on financial expenses. Strategic planning, effective management and a consistent gaming philosophy are equally crucial. In fact, almost every year there are "outsider" clubs that achieve surprising results despite not starting as favourites; a clear example is Leicester City, who won the Premier League in the 2015-2016 season, the most difficult league in the world. This proves that astute management and cohesive team dynamics can sometimes overcome financial strength. Leicester's success has been built on a well-balanced squad, tactical acumen and efficient use of available resources, challenging the idea that financial investment is the main driver of success.

The intricate relationship between financial investments and sporting performance in football is often the subject of considerable interest and debate[11]. This section will be divided into two parts and will first examine how spending on player acquisition can be decisive in sporting results and then whether player salaries influence the competitive performance of a club in the same way. To support these two analyses, I have developed two simple correlation models that take into consideration the first five and last five teams classified in the 2023/2024 Italian Serie A edition.

2.2.1 Investment in Player Acquisition and Team Performance

The purpose of this study is to try to understand whether a heavy investment in the market is actually decisive in obtaining satisfactory sporting results. As already mentioned, the correlation between financial investments in the acquisition of players and team performance is a fundamental area of study in football economics.

For simplicity I have collected the data of a single year and a single league as by doing so we avoid problems of inflation, different internal financial rules for different leagues and cash injections by some companies which occur every X years (of course in the year taken into consideration there was no injection of money from the companies).

The following table (Table 2.1) shows the transfer market balance data of the teams between sales and acquisitions of players of the season considered (summer session and winter session) in correlation with their corresponding positions in the championship.

---

[11] An in-depth study on this subject can be found in this article:
Luca Di Simone and Davide Zanardi (2020), "On the relationship between sport and financial performances: an empirical investigation"

Table 2.1 - First five and last five classified in Serie A 2023/2024 edition balances.

Club	League position	Expenses (€M)	Incomes (€M)		Balance (€M)
Inter	1	70,75	129,82		59,07
Milan	2	121,00	70,70		-50,3
Juventus	3	95,45	66,73		-28,72
Atalanta	4	78,20	152,59		74,39
Bologna	5	65,50	37,00		-28,5
Cagliari	16	23,10	14,73		-8,37
Empoli	17	10,40	59,00		48,6
Frosinone	18	3,00	9,25		6,25
Sassuolo	19	60,71	108,34		47,63
Salernitana	20	28,45	4,76		-23,69

Source: Transfermarkt.it (2024)

From the data in Table 2.1 It is immediately clear that the teams that have generally invested the most in the transfer market are actually the teams that occupy a better position in the standings. The only exception is represented by Sassuolo, who despite substantial spending in the transfermarket occupies the penultimate position in the standings, thus relegated to the Serie B (the lower league).

To mathematically support this discussion I used the following formula to calculate the correlation coefficient between spending and ranking position:

Formula 2.1 - Correlation coefficient

$$Correl(X,Y) = \frac{\sum (x-\bar{x})(y-\bar{y})}{\sqrt{\sum (x-\bar{x})^2 \sum (y-\bar{y})^2}}$$

Where x is the matrix of rankings and $\overline{x}$ the sample mean of that matrix, y the matrix of expenses and $\overline{y}$ the sample mean of the latter. Using Formula 2.1 with the data in our possession, the correlation coefficient is -0.818710059.

As we can see, in fact, the correlation coefficient is quite high, which implies a notable correlation between the two variables. Negativity implies that the greater the expense, the lower the number representing the ranking position (1 is less than 20, however in terms of ranking the lower the number and the higher the positioning).

Another noteworthy fact is the balances. You can also see by eye how the balances of the teams are not at all correlated to the sporting result, this makes it clear that a good market strategy consists in balancing short-term expenses and long-term expenses, i.e. the acquisition of players already ready but at the same time the attention in following young talents who will be ready in the future and who will be able to generate capital gains, this is exactly what It is called "prudent financial management".

In fact, for years in Italy and around the world Atalanta has been one of the teams with the most balanced market and with great attention to youth academies, this makes it a healthy and at the same time competitive club, an achievement that any team in the world would like to achieve .

The same cannot be said about Inter as the positive balance of the year taken into consideration is a “rare bird”, the general balance of the team in fact carries with it a debt of over 400 million euros[12], making it the team with the most debts in Italy.

12 https://www.corrieredellosport.it/news/calcio/serie-a/inter/2024/05/23-127916849/inter_a_oaktree_la_previsione_sui_conti_debiti_e_40-50_milioni_di_perdite

2.2.2 Player Salaries and Team Performance

At this point we still have to analyse player salaries which, together with player acquisitions, constitutes the budget that a club decides to invest in the transfer market[13].

The factors that mainly affect a player's salary are skill level (which raises the figure) and age (generally if you are very "young" or very "old" the figure tends to decrease)[14]. When a team dedicates a budget for the transfer market, it must always take into consideration that part of that budget will be allocated to player salaries. Salaries vary greatly from league to league as each nation has its own taxation, in fact in Italy it will be very difficult to see top players as taxation would make their salaries unaffordable. Unlike other sports, unfortunately, there is not yet a league that has implemented a salary cap, i.e. a maximum salary cap that acts as a balancing factor. Therefore the gap between the teams has become very evident.

For simplicity, "player salaries" here refers exclusively to the net salaries paid to players from clubs, excluding other forms of compensation such as bonuses, sponsorships and agent commissions. Also in this case we take into consideration the data from the 2023/2024 year to be consistent with Table 2.1.

[13] There are many articles that talk about it, one that struck me and from which I drew inspiration is the following: Thadeu Miranda Gasparetto (2012), "Relationship between Wages and Sports Performance", Federal University of Juiz de Fora, Minas Gerais, Brazil

[14] Thibaud Trichard (2021), "What are the main determinants of professional soccer players wages? An Examination of the top 5 European Leagues", Adam Smith Business School

Table 2.2 - First five and last five classified in Serie A 2023/2024 edition player salaries.

Club	League position		Player Salaries (€M)
Inter	1		76,00
Milan	2		62,22
Juventus	3		76,32
Atalanta	4		31,42
Bologna	5		20,19
Cagliari	16		20,78
Empoli	17		15,02
Frosinone	18		13,95
Sassuolo	19		23,55
Salernitana	20		22,49

Source: Capology.com (2024)

Looking at Table 2.2, the correlation between players' salaries and league performance is also evident in this case. Big clubs like Milan, Inter and Juventus, thanks to great popularity and thanks to the support of their fans, manage to obtain important sums of money which are allocated for the salaries of the players, which is reflected in their high-level positions in the championship. This phenomenon highlights a fundamental principle of football economics: investing in high-quality talent often translates into superior performance on the pitch.

As done for the data in the previous table (Table 2.1) we use Formula 2.1 to calculate the correlation coefficient between the players' salaries and the ranking position. Also in this case, as could be expected, the correlation takes on a significantly high value (-0.774212984) which confirms from a mathematical point of view how a greater investment in players' salaries actually reflects a better position in the championship.

I would like to point out, however, that investing in wages is different from investing in the acquisition of players as the money spent on wages has no way of directly generating revenue unlike acquisitions where the purchased player can be resold in the future.

Therefore a large investment in wages carries greater financial risk. High payroll expenses require robust revenue streams to avoid financial instability. This scenario highlights the importance of a balanced approach where clubs must not only invest wisely in talent but also ensure sustainable financial practices.

The case of Atalanta and Bologna offers an intriguing contrast. These clubs have achieved commendable league positions with relatively modest player salary budgets. Their success demonstrates that strategic financial management, effective scouting and player development can offset lower payroll expenses. These clubs serve as a model of efficiency, demonstrating that it is possible to compete at a high level without spending big bucks. This approach is aligned with the principles of sustainable sports management, emphasizing long-term planning and prudent financial practices.

Looking at the bottom of the standings we find clubs facing the twin challenges of limited financial resources and lower league positions. As the correlation index shows us, their lower budgets for player salaries contributed to an unsatisfactory performance. This situation highlights a wider issue in football: the financial disparity between clubs. Smaller clubs must be more adept at adopting innovative strategies, such as investing in youth academies and leveraging data analytics, to maximize their limited resources and improve their competitive advantage. However, this does not always happen, in fact Cagliari, Sassuolo and Salernitana came very far from obtaining the results that Bologna achieved with the same salary budget.

Another important aspect is the impact of player salaries on team morale and cohesion. High salaries can attract the best talent, but they can also bring disparity within the team, potentially affecting team harmony and consequently on-field performances. Effective management must be able to balance financial rewards with the promotion of a cohesive team spirit, ensuring that investment in player salaries translates into collective success on the pitch.

Analysing this data shows that the correlation actually exists and is strong, but fortunately it is not an absolute rule. Clubs that spend judiciously and focus on taking care of all the other key aspects of achieving a goal often punch above their weight. This insight is crucial for financial planners and club managers who aim to optimize both performance and financial health and especially for all football fans who follow and support teams with less financial power, giving them the chance to dream.

In conclusion, the correlation between player salaries and sporting performance is a vital aspect of football economics, but it is not the only determinant of success. Football clubs need to skillfully balance their investment in player salaries with wider financial sustainability. Effective management of player salaries, combined with comprehensive financial and operational strategies, is essential for long-term success in the competitive arena of modern football. This holistic approach ensures that clubs remain competitive while maintaining financial health, thus securing their future in the sport.

2.3 The Virtuous Cycle of Success: From Sporting Achievements to Financial Gains

From the previous paragraph we learned how huge transfer expenses positively influence a club's performance on the pitch. In this section we will close the circle by analysing how on-field performances in turn cause a knock-on effect that substantially improves the club's revenue streams through prize money from various competitions, increased ticket sales and merchandise and very lucrative sponsorship deals. In fact, it can be said that the financial health of football clubs is not only determined by their expenses, but is also greatly influenced by their performances on the pitch. We will delve into how these dynamics work and provide a comprehensive understanding of the virtuous cycle in which sporting success leads to financial gain, which in turn can fuel further investment in club development which will then be instrumental in maintaining that sporting success.

2.3.1 Direct and Indirect Resources deriving from Sporting Successes

1. Direct Incomes

The only direct financial benefit of success on the field is the prize money awarded based on league positions and tournament results. In Serie A the prize money for the final ranking is based on the distribution of revenues deriving from television rights. For the 2023-2024 season, Serie A has allocated approximately 1 billion and 73 million euros[15]. The distribution of revenues deriving from television rights is very particular: 50% in equal parts, 22% based on social roots (live spectators and average audience of

[15] https://www.economiaesport.it/2024/07/03/serie-a-la-ripartizione-dei-diritti-tv-ai-club-oltre-1-miliardo/

live TV broadcasts) and 28% based on sports results (of which 14% based on the ranking and the other 14% based on the results of previous years)[16]. The other two entries are the minimum participation fixed prizes for the UEFA Champions League, the most important european cup (top 5 classified) and for the Italian Super Cup (top 4 classified), a national cup for the best four teams of the season.

Below is the table with all the entries of the teams taken into consideration so far (first 5 and last 5 classified) for the positions of the 2023/2024 season.

Table 2.3 - Earnings from standings

Club	League position	League Awards (€M)	UEFA Awards (€M)	National Cup Awards (€M)	Total Earnings (€M)
Inter	1	101,00	20,30	1,60	122,90
Milan	2	87,40	20,30	1,60	109,30
Juventus	3	86,70	20,30	1,60	108,60
Atalanta	4	60,40	20,30	1,60	82,30
Bologna	5	56,20	20,30	0	76,50
Cagliari	16	37,10	0	0	37,10
Empoli	17	30,90	0	0	30,90
Frosinone	18	31,30	0	0	31,30
Sassuolo	19	35,80	0	0	35,80
Salernitana	20	31,60	0	0	31,60

Sources: La Gazzetta dello Sport, Eurosport - August 16th, 2024

As we can clearly see from Table 2.3, the difference in direct earnings between the first classified and the last classified is sidereal. There is a difference of around 40 million euros net between the fifth-placed team (Bologna) and the sixteenth-placed team

[16] Marco Iaria (2024), La Gazzetta dello Sport

(Cagliari), a figure which will then prove fundamental in the following year's purchasing campaign.

2. Indirect Incomes

To analyse the total income deriving from a good sporting performance, we must also look at the indirect income.

In fact, sporting success also has a profound impact on match revenues and fan engagement. A winning team often manages to bring more people to the stadium than teams that are not performing at their best. We analysed the effect that this flow of people has on a Club's earnings in paragraph 2 of chapter 1, explaining how this type of earnings should not be underestimated at all.

Finally, successful teams are more attractive to potential sponsors and commercial partners. In fact, the most important companies prefer to associate their brand with the most famous teams, recognizing the media coverage that they manage to obtain thanks to their successes on the pitch. As a result, clubs can negotiate more lucrative sponsorship deals and commercial partnerships. I would also add that agreements between clubs and sponsors often present bonuses, i.e. additional payments in the event of winning certain titles.

2.3.2 The Unstoppable Virtuous Circle

Up to now we have analysed the relationship between expenses, sporting results and incomes deriving from the latter. From the analysis carried out we have learned that a greater predisposition to wise and substantial investments in squads and infrastructure consequently leads to a greater probability of experiencing successful performances on the pitch, thus obtaining direct and indirect rewards that are crucial for a club. But why do I say that these resources are "crucial" for a club?

These resources are the key to access a virtuous circle whereby investments generate successes which in turn generate substantial extra revenues which are then reinvested increasing a company's investment power, thus perpetuating the success cycle and ensuring a disarming overall competitive advantage. This cycle creates a reinforcing loop where success begets more success. This is precisely what is called a "virtuous circle".

However, if on the one hand some teams fall within this virtuous elite, on the other hand all the other teams enter a vortex from which it is very difficult to escape because they find themselves faced with an unbridgeable disparity between clubs. This creates a real barrier between those few clubs that are always at the top of the table and all the others that struggle not to go bankrupt. Because precisely, if a club wanted to try to invest a lot of money to try to fill this gap it would find itself facing serious financial as well as legal problems, on the other hand, however, if it does not invest from its own pocket it will not be able to assemble a team capable of compete with those elite clubs.

Over the past decade, the financial gap between Europe's top clubs and their less successful counterparts has widened significantly[17]. Clubs such as Real Madrid, FC

[17] https://europeanleagues.com/wp-content/uploads/REPORT-THE-FINANCIAL-LANDSCAPE-OF-EUROPEAN-FOOTBALL.pdf

Barcelona, Bayern Munich and Manchester City have established themselves as perennial powers, consistently dominating both domestic leagues and European competitions. This dominance is not simply the result of superior on-field talent, but also the significant financial clout wielded by these clubs. As a result, the competitive balance within leagues and across Europe is skewed, with the same clubs often reaching the later stages of prestigious tournaments year after year.

It is arguable that European football has reached a point of no return, where the financial and competitive gap between the elite and the rest is insurmountable in the current framework. The implications of this gap are profound and impact the overall health and sustainability of the sport.

For the top clubs, the virtuous cycle of investment and performance will likely continue to push them forward, ensuring their dominance. However, for smaller clubs, entering this elite circle becomes increasingly difficult. Despite occasional breakthroughs and Cinderella stories, the systemic advantages held by the top clubs are overwhelming.

2.3.3 Breaking the Cycle: the Salary Cap Proposal

One of the most debated solutions to break the virtuous cycle that perpetuates financial and competitive disparities in football is the introduction of a salary cap[18]. The salary cap is not an innovative concept in sport, in fact it is often used in America. In fact, the NBA, the American basketball league, as well as the most followed and loved in the world, introduced the salary cap 40 years ago, in the 1984/1985 season. This rule sets a limit on the total amount that a club can spend on player salaries.

[18] I found this article interesting: Jason Stockwood (2024), "Time for a salary cap to keep leagues competitive and reduce agents' influence", The Guardian

The NBA's salary cap system can offer valuable insights into how such a model might be applied in football. The NBA uses a system called "soft cap", a flexible salary cap that allows teams to exceed the spending limit imposed but incurring a luxury tax, which is a heavy financial penalty ranging from 1.5 to 4.25 dollars for each dollar spent over cap[19]. This system maintains a level of competitive balance while providing flexibility for teams to retain key players and strategically manage their rosters. Obviously, if the model were applied to football the financial sanctions would have to be severe and rigorously applied, unlike how financial fair play rules are now applied by UEFA. This would generate a healthy "fear" in the clubs, which would force them to strictly follow imposed rules.

The main benefit of implementing a salary cap in football is the potential for greater competitive balance. By limiting player salaries, smaller clubs with fewer financial resources would certainly have a better chance of competing with richer teams, as global talent would be more dispersed across different clubs. This could lead to a more unpredictable and exciting league, where success is not dictated solely by financial power but also by managerial acumen and player development. This is precisely what happens in the NBA, where every year there are different teams competing for the title.

Furthermore, a salary cap could facilitate financial sustainability among football clubs. In fact, with the spending limits imposed by the salary cap, clubs would not invest disproportionate sums of money in the market and this would greatly lighten the burden on the company's balance sheet. This consequence would certainly reduce the risks of financial crises and bankruptcies which, unfortunately, often happen in football today. The football ecosystem would thus become more stable and sustainable.

---

19 https://www.sportingnews.com/us/nba/news/what-nba-luxury-tax-explained-penalties-high-spending-teams/p1sspaedfsmsit20rqewn9d7

However, implementing a salary cap in football also presents several challenges and potential drawbacks. The global nature of football complicates the implementation of a uniform salary cap. Unlike the NBA, which operates within a single country, soccer clubs compete in various national leagues and international competitions, each with its own economic context and legislative rules. Establishing a universally accepted salary cap would require extensive cooperation and coordination between football associations, leagues and governing bodies around the world.

Finally, the potential resistance from the best clubs and players should not be underestimated. Wealthier clubs may argue that a salary cap undermines their ability to use their financial success to attract top talent. Star players, accustomed to huge contracts, may also resist wage restrictions by not accepting a reduction in their wages, potentially leading to conflicts with unions and players' associations.

In summary, while a salary cap could help address financial disparities and competitive imbalances in football, its implementation would require careful consideration of the unique characteristics of the sport. It is critical to balance the benefits of greater equality and financial stability with the challenges of global enforcement and potential resistance from stakeholders. By learning from models like the NBA and adapting the approach to the specific context of football, it may be possible to create a more equitable and sustainable future for the sport.

CHAPTER III

DATA-DRIVEN REVOLUTION IN MODERN FOOTBALL

With this third chapter we move on to the second part of this thesis. From here, the focus will be shifted from the general financial aspects of football clubs to the new and more specific approaches in the world of football which involve the use of big data and artificial intelligence, while still maintaining some mention of the financial aspects. What we will delve into is how this modernization in football has impacted the internal management of clubs.

The integration of big data and artificial intelligence into football represents a paradigm shift that is revolutionizing the way clubs operate, make decisions and compete at the highest levels.

3.1 Introduction to the Data-Driven Era in Football

The evolution of football, just like many other aspects of modern life, has been profoundly influenced by the advent of advanced data analytics and artificial intelligence. Recent decades have seen a seismic shift in the way football is played, managed and understood, with data becoming an indispensable tool for decision-making at every level. This chapter marks the beginning of the second half of this thesis, moving from the financial mechanisms that govern football clubs to the innovative, data-driven strategies that are redefining the sport.

Historically, football management and strategy were dominated by intuition, experience and tradition. Coaches and managers relied heavily on their understanding of the game, honed over years of involvement in the sport, to make crucial decisions and to manage the club. Scouting networks were built on the basis of personal relationships and word of mouth, and player performances were evaluated through the subjectivity of those who studied and observed matches and training. While this approach has undoubtedly produced many of the sport's most iconic moments and legendary figures, it has not been without its limitations.

The data-driven revolution in football has its roots in the introduction of sabermetrics in baseball in the United States, a system popularized by a particular approach called “Moneyball”[20] first used in the early 2000s by Oakland Athletics[21] general manager Billy Beane, on which a very famous film with Brad Pitt entitled "Moneyball" was also based. Sabermetrics, a term derived from the acronym SABR (Society for American Baseball Research), emphasized the use of empirical data and statistical analysis in sports to evaluate player performance and develop strategies that were not considered parallel to traditional sports thinking. The success of this methodology in baseball has served as a catalyst for other sports, including football, to explore the potential of data-driven insights.

In football, the integration of data analytics initially struggled to take hold, with the first attempts appearing in the late 2000s and early 2010s, but it has since exploded into a far-reaching transformation[22]. Initially, the goal was to collect and interpret basic statistics, such as goals scored, assists and ball possession percentages. However, as technology has advanced, so has the sophistication of data collection and analysis.

[20] The term Moneyball comes from the title of the book "Moneyball: The Art of Winning an Unfair Game" by Michael Lewis, published in 2003.
[21] The Oakland Athletics is an American professional baseball team based in Oakland, California.
[22] https://www.soccertake.com/performance/soccer-analytics-how-data-is-changing-the-game

Today, it is very easy for football clubs to have access to large amounts of data, as there are many companies that collect it and then make it available to the teams that pay for it. Obviously the data found on the market today is much more complex and ranges from player tracking systems that capture every movement on the pitch to complex algorithms capable of predicting future performance trends and injury risks.

This chapter will delve into the various ways in which data analytics and artificial intelligence have shaped football, reshaping not only how it is played and the various tactics adopted by team managers, but also how clubs are managed. Information derived from big data is now an integral part of a club's operational strategy, influencing decisions related to player acquisition, match preparation and long-term planning.

Furthermore, the rise of data-driven methods has also impacted many financial aspects of football. We will also see that clubs that make effective use of data analytics often enjoy a significant competitive advantage, not only on the pitch but also in the marketplace. This advantage can manifest itself in the form of more strategic relocation decisions, optimized salary structures and increased business opportunities.

This chapter serves as a gateway to understanding how data has become the new currency of success in football, setting the stage for a deeper exploration of specific technologies, methodologies and case studies in the sections that follow. By the end of this chapter, it will become clear that the integration of data and artificial intelligence is not just a trend, but a fundamental shift in the paradigm of soccer management, a shift that promises to shape the future of the sport in ways that are beginning to be understood only in recent years.

3.2 The Role of Data Analytics in Football Management

The first uses of artificial intelligence and big data in football were limited to simple statistical analyses. However, it quickly evolved into a sophisticated system capable of bringing improvements that humans would have been unlikely to make, such as predicting results, evaluating the value of players and even simulating potential strategies[23]. This progression is similar to developments seen in other industries, where artificial intelligence and big data have been leveraged to predict consumer behaviour, optimize supply chains and drive innovation. Even in football, as in the world of industry, these technologies are significantly influencing the possible success of a club, offering new tools that were previously only imaginary.

To effectively understand the impacts of these technologies it is important to explore the specific areas in which they have made the most significant steps. One of the most important areas is match analysis, where artificial intelligence has enabled a deeper understanding of game dynamics and player performance. Additionally, AI has transformed player recruitment and scouting, providing clubs with more accurate assessments of potential signings. Finally, the use of artificial intelligence in injury prevention and player health management has also become a crucial aspect, helping those in charge of players' athletic performance to ensure athletes are in perfect physical shape throughout the season, making them perform at their best.

[23] Jim Totime (2023), "The Data-Driven Football Club: Strategies for Success"

3.2.1 AI in Match Analysis

Match preparation and tactical decision-making by a club's coaching staff have been significantly transformed by the application of artificial intelligence in match analysis. Traditionally, match analysis relied heavily on human observation, intuition and basic statistics, often leading to subjective interpretations and limited insights. Now, however, clubs have access to sophisticated tools capable of analysing large amounts of data in real time that provide a more accurate and complete understanding of the game[24].

AI-driven match analysis platforms allow clubs to evaluate and predict various aspects of the game that were previously impossible to quantify. For example, AI systems can track the movement of all players on the pitch, analysing their positioning, decision-making and overall impact on the game. This allows for a complete understanding of both individual and team dynamics, providing insights that go far beyond traditional video analysis.

The ability to process data in real time is one of the main advantages of artificial intelligence in match analysis. During a match, AI systems can instantly analyse player movements, identify patterns and suggest tactical adjustments. This real-time feedback is invaluable to coaches and players, who can make informed decisions based on data rather than relying solely on instinct or experience. The ability to dynamically adapt tactics during a match gives teams a significant competitive advantage.

Post-match analysis is crucial for a team and here too AI has taken on a fundamental role in the collection and interpretation of data and parameters. This includes not only basic metrics such as goals scored or passes completed, but also more nuanced factors such as the effectiveness of pressing, the success rate of defensive maneuvers and the efficiency

[24] Pavitt J, Braines D, Tomsett R. "Cognitive analysis in sports: Supporting match analysis and scouting through artificial intelligence". *Applied AI Letters*. 2021

of transitions from defence to attack. By analysing these elements, AI can help identify teams' strengths and weaknesses, allowing them to refine their strategies for future matches.

Additionally, AI allows for a more personalized approach to player development. By analysing individual player data, AI can provide personalized feedback and training recommendations. This could involve suggesting specific drills to improve a player's speed, advising on better positioning during certain phases of the game, or highlighting tendencies that need to be corrected. Such personalized insights can accelerate a player's development and improve his overall contribution to the team.

A very positive side of this technology is its easy accessibility, i.e. any club can adopt the use of AI to support team development. What was once the domain of elite clubs with significant resources is now available to smaller teams and even amateur clubs. This easy accessibility of technology allows a wider range of teams to benefit from advanced analytics, potentially levelling the playing field in competitive football.

In summary, artificial intelligence has transformed match analysis from a predominantly subjective practice to a data-driven science. By offering real-time insights, detailed performance assessments and personalized feedback, AI helps clubs prepare more thoroughly, adapt more quickly and ultimately perform more effectively on the pitch. As AI technology continues to advance, its role in match analysis will only grow, further changing the way football is analysed and played[25].

[25] Araujo, Duarte & Couceiro, Micael & Seifert, Ludovic & Sarmento, Hugo & Davids, Keith. (2021). Artificial Intelligence in Sport Performance Analysis

3.2.2 AI in Scouting and Recruitment

Perhaps the most important use of artificial intelligence is in the scouting and recruitment processes of football clubs, radically altering the way talent is identified, assessed and ultimately acquired. Traditionally, scouting relied largely on the experience and intuition of scouts, specialists who travelled extensively around the world to observe players in action and make subjective judgments about their potential. Although this method has produced many successful purchases, it is limited by human capabilities and the intrinsic biases that arise from personal evaluation, as well as having a very high cost.

AI-based scouting systems have addressed these limitations by providing clubs with the ability to analyse large amounts of player data from leagues around the world[26]. These systems use complex algorithms to evaluate players based on a wide range of parameters, including physical attributes, technical ability, tactical awareness and even psychological factors. By processing this data, AI can identify players who meet a club's specific tactical requirements or who possess the potential to become top talent, often uncovering prospects who may have been overlooked by traditional methods.

The AI's ability to objectively compare players from different leagues and levels of competition is certainly one of the most significant advantages of this tool. This ability is especially important when evaluating players from lesser-known leagues or younger age groups, where the quality of opponents can vary significantly. AI can normalize these differences, allowing clubs to make more accurate assessments of a player's potential performance in a higher or different competitive context.

---

[26] Guida, Caniato, Moretto, Ronchi (2023), "Artificial intelligence for supplier scouting: an information processing theory approach"

AI also plays an important role in the recruitment process by providing a deeper understanding of a player's compatibility with the existing team, a factor that should not be underestimated when evaluating an interesting profile. Highly advanced AI technology can even simulate the way a player might behave within the club's tactical framework, considering factors that are difficult for humans to quantify, such as understanding with current players, adaptability to the playing style of the club game and even cultural adaptation. This predictive ability reduces the risk associated with transfers, helping clubs make more informed decisions.

The AI can also proceed to follow players for an extended period, monitoring players' progress and adjusting ratings from time to time as experience and maturity grow. The ability to continuously monitor allows clubs to decide at the right time whether to sign a player before his value becomes out of reach or not to pursue a player whose development has been interrupted. This dynamic approach to scouting ensures clubs are agile and can act when opportunities arise.

Another very innovative application of artificial intelligence in scouting is the discovery of "hidden gems": these are players who do not shine according to traditional parameters but possess special attributes that can be very important in some tactical settings. For example, artificial intelligence could identify an atypical player who does not excel in quality but who is very effective. These are the insights that allow clubs to find, and then sign, talent that other clubs may not notice, giving them a competitive advantage in the transfer market.

While in many ways AI has helped facilitate scouting and recruiting, the process is not without its challenges. One of the biggest problems is related to the quality and completeness of the data used by AI systems. In fact, incorrect data will result in incorrect evaluations; therefore, clubs need to ensure they have access to complete, quality datasets. Furthermore, integrating AI into scouting requires a cultural change within clubs, as traditional scouts and managers must learn to trust and interpret AI insights.

In conclusion, artificial intelligence has radically revolutionized scouting and recruitment in football, equipping clubs with the most powerful tools to identify and evaluate talent with unprecedented precision. Through data analysis, AI enables clubs to make more informed decisions, reduce the risks associated with transfers and uncover hidden talent. AI technology will play a very pronounced role in shaping the future of football recruitment as the technology evolves further, likely changing the way talent is identified and developed within the sport.

An emerging platform for example is AiScout[27], a London-based technology company. AiScout app allows aspiring soccer stars to enter virtual trials for professional clubs by uploading self-recorded footage of themselves completing a series of drills. It offers 75 exercises, designed to test a range of skills, with videos showing users how to complete them. Performances are automatically scored by artificial intelligence (AI) technology. The data can then be accessed by clubs, allowing their scouts to search for valid talents, refining the search with a series of filters, from age and gender to position on the pitch. The application was fully launched in September 2023. To date it has more than 100,000 players in the current database and more than 100 professional clubs as partners, as well as Major League Soccer[28].

[27] Jack Bantock (2024), "Top soccer clubs are using an AI-powered app to scout future stars", CNN
[28] Highest soccer league in the United States

According to a “Grand View Research” report[29], The global sports analytics market size is expected to reach USD 14.41 billion by 2030 and expand at 21.5% CAGR[30] from 2024 to 2030

3.3 Case Studies: Success Stories of Data-Driven Football Clubs

To better observe the impact of big data and predictive analysis on football it is necessary to take into consideration true stories of clubs that have actually embraced these technologies and methodologies. Nowadays any team makes use of advanced technologies and big data to study football and to make it as efficient as possible, but initially there were few teams that "converted" to this type of approach to football.

In this paragraph I will consider three teams that were, each in their own way, pioneers of modern football. These case studies highlight how data-driven approaches can transform the fortunes of clubs, leading to lasting competitive advantages and, in some cases, unprecedented on-field success.

3.3.1 FC Midtjylland: The Pioneers of Data-Driven Success

FC Midtjylland, a Danish football club, is often cited as one of the most successful examples of a data-driven approach to football. Founded in 1999, the club has risen from relative obscurity to become a dominant force in Danish football, winning multiple

[29] https://www.grandviewresearch.com/press-release/global-sports-analytics-market#:~:text=Sports%20Analytics%20Market%20Growth%20%26%20Trends,by%20Grand%20View%20Research%2C%20Inc.

[30] Compound annual growth rate (CAGR) is a business, economics and investing term representing the mean annualised growth rate for compounding values over a given time period.

league titles. The club's success is largely attributed to its early adoption of big data and analytics, particularly in player recruitment and match strategy.

Midtjylland's journey into data-driven football began after its acquisition by Matthew Benham, who introduced sophisticated models to assess player potential and match tactics through his London-based analytics firm, Smartodds.[31], which provided the club with sophisticated models to assess player potential and match tactics. This collaboration allowed Midtjylland to identify undervalued players in the transfer market, often signing them at a fraction of their potential market value. The club's scouting process became highly efficient, focusing not only on traditional scouting reports but also on advanced metrics that predicted a player's future performance[32].

Probably the most famous success of Midtjylland's data-driven approach was the signing of Brazilian midfielder Evander. His name had been put forward by data-driven predictive analytics as a player with huge potential to become an important player for the club. Months after this analytical investment, Evander worked to justify it, playing a key role in national and European campaigns.

But it wasn't just in terms of recruitment that Midtjylland used the data. There were other ways the club used data analytics to improve match tactics, including set-piece strategies. By analysing their opponents' weaknesses and adapting their set-piece routines accordingly, Midtjylland have significantly increased their efficiency in scoring goals from dead-ball situations. This data-driven tactical innovation was instrumental to the club's success and has since been adopted by other clubs across Europe.

---

[31] Smartodds is specialised in providing in-depth quantitative and qualitative research and analysis on numbers sporting events all over the world

[32] https://outsideoftheboot.com/2016/07/20/rise-of-data-analytics-in-football-2/

3.3.2 Brentford FC: Data-Driven Ascent to the Premier League

Another very interesting case study from this point of view is Brentford FC[33], an English club playing in the second division that uses data analysis as a means to fulfil the ambition of playing football in the Premier League (the English first division). Matthew Benham, in addition to owning FC Midtjylland, also owned an English club, Brentford, where he applied the same data-driven philosophy in 2012, focusing on smart player recruitment and evidence-based management.

Brentford's approach to player recruitment is largely based on a proprietary statistical model which evaluates players based on a number of performance metrics. This model, known as the "Brentford Model", allows the club to identify players who are undervalued on the market but who have the potential to thrive in their system, very similar therefore to the model studied in the previous sub-section, but the technology used is slightly more advanced in when it studies new data compared to those previously studied. The model looks beyond traditional metrics and incorporates data such as expected goals (xG)[34], pass completion rates and pressing efficiency to build a comprehensive profile of each player.

One of Brentford's most successful data-based signings was that of Ollie Watkins, bought from Exeter City in 2017. Watkins, identified through Brentford's analytical model as having significant potential, quickly developed into one of the most prolific in the championship and still plays in the Premier League. His performances caught the eyes of many prestigious clubs, so much so that he was bought for €34 million by Aston Villa, providing Brentford with both sporting success and financial gain.

[33] Liam Callaghan Doyle (2022), "Is the Moneyball approach a feasible approach for success in Football (Soccer)? An analysis of Brentford's adoption of a statistical approach."

[34] The expected goals metric is generally calculated by determining the likelihood of a shot being scored based on various factors, taken from the moment before the player shoots. Nowadays it is one of the most relevant data in the world of football.

Brentford's data-driven approach also extends to match preparation and game strategy. The club's analysts use real-time data to make tactical adjustments during matches, such as altering the intensity of pressing or shifting defensive shapes to counter specific threats from opponents. This adaptability, informed by data, has been a key factor in Brentford's rise through the English football pyramid.

3.3.3 Liverpool FC: Big Data at the Elite Level

While smaller clubs such as FC Midtjylland and Brentford have used data analytics to mitigate the gap with stronger clubs, some of Europe's elite clubs have also embraced big data to maintain their status at the top of the game. Liverpool FC, under the direction of Jürgen Klopp, a revolutionary German manager among the best in the world, has been at the forefront of this trend.

Liverpool's data-driven approach comes with the involvement of a team of sports scientists who, aided by data analysts, work closely with the team's technical staff in pursuit of optimising performance and tactical choices[35]. Of more relevance in this use by the club, as in previous cases, is the way in which it implements its player recruitment strategy. Former sporting director Michael Edwards was integral to building a data-driven scouting and recruitment system credited with the arrival of world-class players such as Mohamed Salah, Sadio Mané and Virgil van Dijk.

Liverpool's scouting relies heavily on predictive analytics in an attempt to find players who not only fit the club's playing style but who can also develop further over the years.

[35] Lichtenthaler, U. (2021), "Mixing data analytics with intuition: Liverpool Football Club scores with integrated intelligence", Journal of Business Strategy

For example, Mohamed Salah (one of the most decisive players in the world in the last five years), was identified by data analysis which highlighted how high his xG was and his ability to create scoring opportunities from wide areas, thus adapting perfectly to Klopp's high-pressure, frenetic-pace attacking system.

Aside from player recruitment, Liverpool are using big data for the first time in managing players' fitness levels to reduce the chances of injury. The club's sports science team implements various data points, such as GPS tracking, heart rate monitoring and biochemical analysis, to monitor players' workloads and recovery. This data is then used to tailor individual training programmes, ensuring players are in the best physical condition for key matches.

Liverpool's data-driven approach has been a key factor in its recent successes, including winning the UEFA Champions League in 2019, the world's most prestigious club competition where the best teams in Europe compete, and the Premier League in 2020. The club's ability to integrate data into every aspect of its operations, from recruitment to match preparation, has set a benchmark for how elite clubs can use big data to support success at the highest level.

The case studies of FC Midtjylland, Brentford FC, and Liverpool FC underscore the transformative impact of big data and predictive analytics in football. Whether it’s a small club seeking to disrupt the status quo or an elite team aiming to maintain its dominance, the strategic use of data is proving to be a game-changer. These clubs have demonstrated that by embracing data-driven decision-making, it’s possible to achieve sustained success on and off the pitch, paving the way for a new era in football where data and analytics are as integral to victory as the players themselves.

3.4 The Impact and Challenges of Big Data in Football

The integration of big data into football has radically transformed the way clubs operate, strategize and compete. This revolution has opened new avenues to improve performance, optimise financial decisions and gain a competitive advantage[36]. However, alongside these opportunities come significant challenges that must be addressed to unlock the full potential of data analytics.

Big data offers football clubs an unprecedented level of insight into every aspect of the game, from player performance to fan engagement. By analysing large amounts of data, clubs can make more informed decisions that drive success both on and off the field. The strategic application of data has the potential to improve a club's performance, leading to better results, increased revenues and sustained success.

However, the journey towards a data-driven approach is not without obstacles. Successful implementation of big data in football requires overcoming several key challenges, including ensuring data quality, addressing cultural resistance within clubs and addressing complex ethical issues. These challenges, if not managed properly, can undermine the effectiveness of data-driven strategies and limit the benefits clubs can derive from their investments in analytics.

---

[36] The following article is not about football, but it provides a broader theoretical framework that can be applied to football, explaining how big data helps in improving decision-making processes and organisational efficiency, both of which can be crucial for clubs: Sabharwal, R., Miah, S.J. (2021) “A new theoretical understanding of big data analytics capabilities in organisations: a thematic analysis”. J Big Data

3.4.1 Financial and Strategic Advantages

The integration of big data and advanced analytics into football has provided important financial and strategic benefits for clubs who have developed their strategies using these technologies. One of these critical dimensions concerns how transfer operations are optimised, particularly in player transfers and contract negotiations. Thanks to the detailed analysis and precise statistics provided by these technologies, clubs can more accurately assess the true market value of players, thus reducing the risk of overpaying. This approach allows for more effective financial management, as it minimises the risk of capital losses through the acquisition of players, i.e. it reduces the possibility of clubs paying excessive amounts for talents that may not justify the investment and then be sold for less than they were paid by the club.

An exemplary case is the methodology adopted by clubs such as Brentford FC. Indeed, in the club, advanced scouting models based on metrics have been developed that especially look for undervalued players to help build a competitive team with very modest budgets. The strategy demonstrated that data analytics can be a game-changer in the financial dynamics of any club, helping to achieve very significant sporting results without having to compete with the giants of European football in terms of spending.

In addition to improving efficiency in transfers, the data-driven strategies also help clubs maximise return on investment (ROI) in player development. Performance data analysis allows a club to spot young talents who can potentially become great players at an early stage in their careers[37]. Early identification enables the clubs to make an investment in the player before his market value grows significantly, hence securing future stars at a fraction of the cost they would otherwise have to pay later on. Furthermore, collected

[37] Lucas Oaigen (2024), "How Artificial Intelligence Can Revolutionize Talent Management in Youth Soccer", Linkedin

performance data can tailor training programs based on the individual needs of each player to support the development and finally the market value of the player.

The financial benefits don't stop at player acquisition and development. Data analytics also plays a very important role in ticket sales[38], merchandise sales and sponsorship strategies. By analysing fan behaviour and engagement data, clubs can build targeted marketing strategies that they can use to increase revenue streams from these areas. For example, data can be used by clubs to segment their fan base allowing them to create personalised experiences that increase the chances of achieving repeated purchases by increasing fan loyalty. Furthermore, the fact that consumer trends can be tracked and analysed allows a club to leverage such information when negotiating sponsorship deals, as it allows it to substantiate its reach and influence with real data.

However, it is essential to note that these financial and strategic benefits are not uniformly accessible to all clubs. Adopting data analytics requires significant investments in technology, infrastructure and skills. While top-tier clubs with vast resources can afford to integrate these systems seamlessly, smaller clubs may struggle to keep up, potentially widening the financial and competitive gap between the elite and the rest of the world of football. Therefore, while data analytics offers immense potential, it also presents challenges that need to be addressed to ensure a fairer and more balanced playing field in football.

3.4.2 Challenges: Data Quality, Cultural Resistance and Ethical Issues

While the benefits of big data and advanced analytics in football are substantial, the integration of these technologies is not without significant challenges. These challenges

[38] https://www.hypesportsinnovation.com/how-rbfa-scored-big-with-ai-driven-ticket-sales/

can be broadly categorised into three main areas: data quality, cultural resistance, and ethical concerns[39].

Data Quality

One of the key challenges in exploiting big data in football is ensuring the quality of the data used. High-quality data is critical for accurate analysis and reliable decision making. However, collecting and storing this data is a complex task. The data may be incomplete, inaccurate or outdated, leading to imperfect analyses that can lead to poor strategic decisions. For example, data collected from match performances could be influenced by external factors such as weather conditions, referee decisions or even player injuries, which could skew the results if not adequately taken into account.

Additionally, integrating data from multiple sources, such as wearables, video analytics and scouting reports, requires a standardised approach to ensure consistency and comparability. Without adequate standardisation, the resulting analyses could suffer from bias, reducing the effectiveness of data-driven strategies. Clubs must invest in robust data collection and management systems, as well as employ expert data scientists who can process and clean the data to maintain its integrity. However, this requires significant financial resources, which can be a barrier for smaller clubs.

---

[39] For many of the themes in this subsection I took inspiration from the following book: Bormida, M.D. (2021), "The Big Data World: Benefits, Threats and Ethical Challenges", Iphofen, R. and O'Mathúna, D. (Ed.) Ethical Issues in Covert, Security and Surveillance Research (Advances in Research Ethics and Integrity, Vol. 8), Emerald Publishing Limited

Cultural Resistance

The adoption of data-driven approaches in football is also met with cultural resistance, both within clubs and from the broader football community. Football has traditionally been a sport driven by intuition, experience, and human judgement. Many coaches, scouts, and managers who have built their careers on these traditional methods may be sceptical or even resistant to the idea of relying on data and algorithms for decision-making. This means that this scepticism could be one of the main factors slowing the implementation of data analytics, as key stakeholders may be hesitant about its use and the reliability of the information it provides.

Since this requires a change in mindset, and therefore adoption at a cultural level, it becomes quite difficult to achieve. It is more than just training and educating staff on the need to perform data analysis, but involves showing the real benefits of this approach through success stories and case studies. Clubs that successfully implement these data-driven approaches to their business usually achieve this by creating an innovative culture that supports the pursuit of progressive performance improvement, where data is seen as a valuable tool and not a threat for traditional skills.

Ethical Issues

The use of big data in football also raises several ethical concerns that must be carefully considered. One of the most pressing issues is the potential for data misuse, particularly in relation to player privacy. The collection of vast amounts of personal and performance data can lead to situations where players feel their privacy is being invaded. Wearable devices that track biometric[40] data, for example, can provide valuable insights into player

[40] Biometrics are body measurements and calculations related to human characteristics and features.

fitness and performance, but they also raise questions about consent, data ownership, and how this sensitive information is used and shared.

Furthermore, the reliance on data-driven decision-making could lead to ethical dilemmas regarding player treatment and development. For example, a purely data-driven approach might favour players who fit certain statistical profiles, potentially leading to biases in player selection and development that overlook individuals with unique talents or attributes that are not easily quantifiable. This could result in a homogenization of player types and a reduction in the diversity of playing styles, which are integral to the richness of football.

In conclusion, while the integration of big data and analytics offers significant advantages for football clubs, these benefits are accompanied by considerable challenges. Addressing issues related to data quality, overcoming cultural resistance, and navigating the ethical implications of data use are essential steps for clubs looking to harness the full potential of these technologies. Ensuring that these challenges are met effectively will be crucial for the sustainable and equitable growth of the sport in the data-driven era.

3.5 Future Directions for Data-Driven Football

As the data-driven revolution transforms contemporary football, the next phase addresses the growing gap between elite clubs with significant financial resources and small teams struggling to remain competitive. While big data, advanced analytics and artificial intelligence have improved decision making, performance improvement and long-term strategy, some are increasingly concerned that this is concentrating power and resources in a small, elite group of organisations. For the sake of competitive integrity in sport, this cycle must be broken so that the benefits of data-driven approaches are democratised and available to all clubs, regardless of their size.

The following section will discuss some future directions for data-driven football, focusing on how smaller clubs might best equalise disparities, the role of regulation in the use of data and artificial intelligence in sport, and how these efforts can contribute to creating a more sustainable and balanced football ecosystem.

3.5.1 Bridging the Gap: Encouraging Smaller Clubs to Embrace Data Analytics

To clubs of less financial power, competing with the giants in this data-driven era seems to go uphill. This is while elite clubs can spend money on state-of-the-art technology, employ top data scientists, and develop highly complex analytics departments. However, the future of football analytics must not stay privileged and exclusive to only a few elite clubs. Instead, strategies to help smaller clubs access the same data-driven tools will have to be implemented.

This could be through the establishment of shared data platforms. Centralised systems or collaboration whereby smaller clubs can share access might overcome some of the cost implications. These platforms would house significant data resources, such as player statistics, scouting reports, video analysis, and biometric information. Second, analytics software could be made available to suit the needs of clubs at the lower ends of financial resources to ensure fair access to advanced technologies. Those positioned as third-party vendors with subscription-oriented services or modular data tools can considerably make big data more affordable and accessible to a wide range of teams.

Additionally, football leagues and other governing bodies can develop training courses and seminars that will help employees and managers of smaller clubs understand how to successfully apply data analytics. Such initiatives would help overcome cultural resistance to analytics-based approaches and, in turn, allow smaller clubs to make intensive use of analytics in talent identification, strategy and game performance analysis. Equipping smaller clubs with the know-how and means necessary for data analysis allows football to develop a more competitive and creative environment.

With these potential solutions in mind, you should consider that smaller clubs will find it even more difficult to implement data-driven approaches. Issues related to lack of funding, staff limitations, and problems trying to recruit high-quality people into data science and analytics positions represent quite considerable obstacles. Tackling these issues will require collaboration across the football industry, with larger clubs and governing bodies potentially playing a more proactive role in sharing knowledge and resources.

3.5.2 The Need for Regulation in AI and Data Usage

As the sport of football increasingly depends on data analytics and artificial intelligence technologies, the demand for regulatory frameworks intensifies. Presently, the application of data within football is predominantly unregulated, especially regarding decisions made with the assistance of AI. Although a certain level of innovation is beneficial and essential, the potential for unregulated AI usage to result in unintended outcomes, both in-game and beyond, also warrants concern.

One of the major concerns is regarding data privacy and the rights of players. Because the AI tool monitors everything related to the player's performance, from heart rate to sleep patterns, the question arises of how much control the players would have over their own personal data. In the protection of player privacy, football's governing bodies could be tasked to institute policies that spell out clearly the ownership and usage rights of the data on performance. This process would ensure that the athletes retain consent over their information while consequently forbidding clubs from the pursuit of even slight advantages crossing lines of morality.

Furthermore, the increasing involvement of artificial intelligence in the processes of player scouting and recruitment prompts ethical concerns pertaining to equity and bias. The reliance on algorithmic decision-making may unintentionally prioritise specific player attributes, such as measurable strengths (for instance, speed or passing precision), while neglecting non-tangible characteristics like leadership, creativity, or tactical acumen. If this bias remains unaddressed, it could result in a restricted variety of player profiles being chosen, ultimately reducing the diversity of playing styles that contributes to the distinctiveness of football.

Because of these potential pitfalls, it would be prudent if football governing bodies like FIFA and UEFA initiate governance structures related to artificial intelligence. Such guidelines could provide ethics regarding the deployment of AI in talent scouting, athlete development, and the analysis of games so that the clubs maintain transparency and accountability in their data-driven decisions. Moreover, such a regime can contribute to reducing competitive imbalances that may arise from larger clubs having better access to AI-enabled applications; the rules could encompass expenditure limitations on AI resources or requirements on equal access to certain key data platforms.

Finally, there is the overarching question of whether artificial intelligence or data-driven approaches might eventually impact the competitive balance in the sport. Still, the greater the integration of AI into the operation of soccer, the wider the gaps that have traditionally been between well-heeled clubs and less well-financed competitors. The larger-pocketed clubs would be able to invest in more highly developed AI mechanisms, creating sharp edges in recruitment, scouting, and performance analysis. This underlines the need for regulating measures, which can take the form of financial fair play initiatives or restrictions in using certain technologies with a view to maintaining equal competition between all clubs.

In conclusion, although data analytics and AI are potent instruments in the development of football, ethical and fair deployment of these technologies needs to be ensured. In this way, through the promotion of data-centric approaches in smaller clubs, as well as the institution of deliberate policies regarding the use of artificial intelligence and data, the sport can evolve in such a manner that all teams will profit, with both fairness and competitive integrity preserved.

The present chapter has discussed the big change that football has undergone with the integration of data analytics, AI, and complex software. The current chapter, through the examination of what role these technologies have played in enhancing on-field performance, underpinning strategic decision-making, and optimising the financial functioning of football clubs, showed that clubs making use of such tools secured competitive advantages. The importance of the transition comes out in this examination of case studies regarding data-driven football clubs and shows examples of teams in competitive advantage through an applied data-oriented strategy in their respective fields.

Yet, it has also highlighted various challenges identified with the use of big data: issues with data quality, resistance to cultural change, and ethical concerns with the deployment of AI. As football becomes increasingly dependent on advanced analytics, undeniably, the future of the game will be determined by how well clubs of all sizes are in a position to adopt such tools and indeed the regulatory frameworks created to govern their use.

The discussion here thus sets the stage for the next chapter, which shall demonstrate how econometric models have practical applications in football. Continuing from the previous discussions on data-driven decision-making and optimization of performance, this chapter shall seek to discuss how quantitative financial models, exemplified by the "Moneyball" approach, can be applied to football to enable football clubs better to optimize the use of their resources and maximise outcomes.

CHAPTER IV

DEVELOPING AN ECONOMETRIC MODEL FOR PLAYER EVALUATION IN FOOTBALL

4.1 Introduction to Econometric Modelling in Football

The use of econometric modelling in football reflects an important change regarding the analytical and managerial aspect of the game. Football clubs increasingly move away from exclusive reliance on traditional methods, such as subjective scouting and intuition-based decision making, toward increasing use of data-driven methods to advance player performance and financial strategies[41]. This development is part of a larger trend in many industries: one where quantitative analysis, deep data collection, and AI fundamentally reorder the operating functions of organisations and the decision-making process.

Although these econometric models have shown efficiency for application in the fields of finance, healthcare, and policy analysis, they remain in their infancy when it comes to football.

The analytics frameworks also enable football clubs to value athletes, predict their future performance, and determine undervalued talents for a competitive advantage. There are certain singularities in the football ecosystem, however. The interaction of sporting, financial, and cultural factors in this environment makes it far more complex

[41] https://thepfsa.co.uk/the-power-of-analytics-dissecting-the-role-of-data-in-football-strategies-and-performance/

compared with other sectors. Thus, adapting econometric models to this particular environment requires special attention to a number of factors and changes.

The aim of this chapter is to present applications from econometric modelling in football; it focuses on some fundamental models applied to player market value assessment and latent talent discovery. In fact, it provides the basis on which clubs willing to enhance their decision-making mechanisms, while optimising return on investment, can remain competitive in a different data-driven environment.

4.1.1 Overview of Econometric Models in Sports

Econometric modelling has its roots in economics and statistics, where it has long been used to study relationships between variables and to make predictions based on historical data. In sports, its adoption started primarily in baseball, thanks to the popularisation of the "Moneyball" approach, which revolutionised how teams evaluated player performance. The use of these models allowed clubs to identify underappreciated players who offered high performance relative to their cost, effectively reshaping how teams managed their rosters.

The basic premise of econometric models in sports is the quantification of complex relationships, between player performance metrics, financial data, and external factors such as market dynamics, into a structured framework that can be analysed and interpreted. These models typically rely on vast datasets, incorporating both on-field statistics (such as goals, assists, or passes) and off-field variables (such as age, transfer fees, and wages)[42]. The goal is to uncover patterns that might not be visible through traditional observation.

While baseball has been at the forefront of this movement, football has more recently begun to embrace these models. The unique aspects of football, including its fluidity, team-based dynamics, and the greater number of subjective variables, make it more challenging to apply traditional econometric methods. Nevertheless, the growing availability of data and the development of specialised football-specific models are overcoming these barriers, paving the way for a more analytical approach to football management.

In the sections that follow, we will explore how econometric models are applied in football, focusing on their use in evaluating player market value and identifying undervalued talent, which are critical aspects for any football club looking to maintain a competitive edge in the modern era.

4.1.2 The Role of Econometric Models in Football Club Management

In modern football, the application of econometric models has moved from a specialist tool of analysis to a strategic management instrument. Football clubs, in particular those operating under financial constraints, have been increasingly looking for means to optimise their resources and gain a competitive advantage relative to other wealthier opponents. Econometric models provide a conceptual framework whereby it will be possible to objectively evaluate past performances and forecast future performances

[42] Stewart, Mitchell, Stavros (2007), "Moneyball Applied: Econometrics and the Identification and Recruitment of Elite Australian Footballers", International Journal of Sport Finance

while concurrently assessing financial implications related to possible transfers and negotiations.

The transformation, therefore, must resolve the sporting ambitions with financial sustainability. Football clubs operate in a highly competitive and stressful environment, where success on the pitch leads to financial rewards by way of prize money, commercial sponsorships, and increased fan engagement. However, increasing transfer fees and player wages have put added pressure on clubs to ensure that their financial investments yield predicted returns. It is in this regard that econometric models become imperative.

A key role of such models is player valuation. By including several measures of performance-data on goals, assists, tackles, and other advanced metrics-along with non-performance variables such as age, market conditions, and length of contract, econometric models can determine a more precise market value of a player than traditional scouting methods have thus far been able to provide. This enables the club to drive a harder bargain in the transfer market and not spend too much on any player simply because he has a good reputation or is in current form.

Another important application of econometric models involves identifying undervalued or underpriced talent[43]. In football, for example, market forces are often driven by star names and media-generated news, and many talented players remain hidden or underused simply because they come from smaller clubs or have low levels of media exposure. Yet econometric models can sift through large volumes of data in order to identify those who perform most highly relative to cost. This enables clubs, especially those on a tight budget, to form competitive teams by signing players who offer the most value relative to their cost.

[43] Poli, Besson, Ravenel (2022), "Econometric Approach to Assessing the Transfer Fees and Values of Professional Football Players.", https://doi.org/10.3390/economies10010004

Besides the valuation of players, econometric models also play a very important role in risk management and long-term strategic planning. Analysing historical performance and forecasting future trends enable clubs to make evidence-based decisions on player development, injury-related risks, and transfers. It reduces uncertainty through data-based analysis, enabling clubs to make more strategic investments in teams.

Moreover, the models help balance the needs of short-term performance against the sustainability of long-term financial stability. For instance, clubs can conduct econometric analysis on when to sell a player at their peak market value, or invest in a younger, promising talent who is likely to outshine their peers in a few years' time. In this way, a club develops a valuable ability to strategically plan over a number of years and develop a competitive position while finely balancing its books.

In sum, econometric models assist football clubs on two fronts: one, to optimally evaluate players, and two, to undertake long-term financial planning by the clubs. In making decisions based on data, football clubs are able to secure their sporting ambitions with sound financial strategies while ensuring continued success both in the football pitch and at the box office.

4.1.3 Key Principles of Econometric Modelling in Football

While econometric modelling in football may be considered relatively new, it is nonetheless based on solid statistical and economic principles. The quality and relevance of the variables selected, the quality of the data, and the type of econometric methodology applied are all important success factors for such models. In this section, we will discuss the basic underlying principles behind econometric models in football, focusing on specific sport challenges and methodologies involved to account for them.

Selection of Variables

The choice of variables is the backbone of any econometric model. Performance metrics in football are diverse and may differ significantly for a player according to his role, position, and style of play. An appropriate selection of variables will determine the accuracy and relevance of the model's output.

In estimating player market value, the usual KPIs (Key Performance Indicators) used include goals, assists, successful dribbles, interceptions, tackles, and pass completion rates. Aggregating these statistics with physical attributes like sprint speed and stamina provides the quantitative measure of a player's contribution during matches.

Other non-performance variables also feature in player valuation models. Other factors determining a player's market value are his age, the length of his contract, injury record, inflation in the transfer market, and international experience. A player, for example, who is still quite young and has excellent performance data with a higher growth margin will have a higher valuation than an older player performing at the same level. Also, factors such as a player's brand equity or marketability-normally with elite athletes-can greatly influence the transfer fee applicable, especially for those teams that are based on commercial factors.

Data Quality and Availability

It is well known that an accurate and reliable dataset is necessary to develop robust econometric models. In football, during the last decade, both the availability and the granularity of data improved substantially due to improvement in data collection technologies and companies specialised in football analytics. Such progress enables clubs to access all those detailed datasets, which track almost every movement on the pitch and provide very important inputs for econometric analysis.

Nevertheless, quality and comparability are still issues with most of the data. In a few leagues and clubs, access to high-quality performance data is quite unequal, since smaller leagues and clubs cannot have access to highly developed tracking technologies. In such a context, the respective econometric models may suffer from biassed or incomplete data, leading to misestimated results. As a consequence, analytics processes will need to be adapted for clubs operating on limited data and may even have to develop alternative measures or proxy measures of performance.

Model Specification

The choice of adequate econometric methodologies forms one of the bases in formulating football models. Simple forms of linear regression analysis are used in preliminary tests to investigate the correlation between a player's performance indicators and his market value. However, the complexity, non-linearity and external interferences that accompany football data such as media exposure and agent[44] influence would require the inclusion of higher-order analysis techniques.

Fixed-effects models, panel data analysis, and machine learning algorithms are some of the sophisticated econometric techniques that have become increasingly popular in football analytics. Such methodologies allow clubs to consider exogenous factors,

---

[44] A sports agent is a legal representative (hence agent) for professional sports figures such as athletes and coaches. They procure and negotiate employment and endorsement contracts for the athlete or coach whom they represent. https://en.wikipedia.org/wiki/Sports_agent

identify temporal trends, and ultimately yield more accurate predictions about future performance. Machine learning models, for example, can incorporate hundreds of variables to find complex patterns that a traditional econometric model cannot. These techniques also allow for updating in real-time; thus, clubs can easily update their models when new data flows in.

Dealing with Uncertainty and Forecasting

The game is, by nature, full of uncertainty: from short-term loss of form, to severe injuries, to changes in coaching. Such econometric models thus need to include risk estimates and forecasts by accounting for such unpredictable factors.

Injury risk models aim at estimating, from past data, physical fitness, and the intensity of the match, the probability that a player may get injured. It is normally embedded in a wider econometric framework. From the valuation of such risks, clubs can optimise their team management and avoid heavy investments in injury-prone players, which may extend for several years. Predictive models projecting a player's development potential consider age, position, and career profile, among other factors; these forecasts are vital inputs to decisions regarding player buying and selling.

Interpretation of Results and Decision-Making

The primary objective of econometric models is to furnish decision-makers with practical insights. Nevertheless, the findings generated by these models necessitate meticulous interpretation. Insights derived from data should not be regarded as definitive truths but rather as instruments that may assist in the decision-making process. Human judgement, especially in a subjective context such as football, continues to play a crucial role in the interpretation of model results.

It can, for example, indicate that some player is undervalued, according to his performance data; after all, the coach may decide not to pursue the player because of issues regarding team balance or tactical mismatch. Concerning this aspect, econometric models should be augmenting and not entirely replacing traditional scouting methods. It is the combination of analytics and the professionals which allows one to make decisions most effectively in the management of football.

In the end, variable selection, data integrity, model formulation, uncertainty management, and results interpretation form the very foundation of econometric modelling in football. These can ensure that the outputs are meaningful, useful, and may have an impact on club strategies, both on the pitch and in organisational decisions.

Football, as an international enterprise, has become increasingly complex; hence, econometric models have become very essential in the correct valuation and identification of undervalued talent. By applying advanced statistical techniques and sophisticated analytics of data, these models give a quantitative basis that becomes the backbone of key transfer market decisions, contract negotiations, and detailed squad development.

4.2 Econometric Model for Player Market Value Evaluation

This section develops an econometric framework for estimating the market value of football players based on an analysis of various performance indicators and exogenous factors. Because the football transfer market has become increasingly competitive, reliance on subjective evaluations or traditional scouting methods often leads to mispricing among players. The aim of this model is to establish a data-centric methodology that offers a more objective and quantifiable assessment of a player's value, thereby equipping clubs with an analytical instrument to inform their transfer strategies.

The model will be tasked with finding the relationship existing between a player's pitch performance and market value, considering key variables like goals, assists, age, league, and even defensive actions-such as tackles or aerial duels won. With this information, we seek to construct an all-encompassing model applicable at many leagues and levels of football. The model will also be used to identify any inefficiencies in the market by identifying players whose market valuation may not be an accurate reflection of their potential value, thus helping smaller clubs maximise their often-limited resources.

4.2.1 Objectives of the Model

The econometric model we propose has several key objectives, each grounded in the practical needs of modern football clubs and the evolving role of data-driven decision-making in player valuation.

The main purpose of the model is to develop an accurate procedure for determining the market value of football players, relying on objective and measurable performance indicators. Traditional systems for valuing players are often based on subjective factors-media presence, prestige, or negotiations between agents and clubs, for example-which might produce significant differences between the market value of a player and his or her actual worth in a game. By anchoring the analysis in performance measures, we attempt to eliminate much of the subjectivity and ensure a more transparent and reliable means of arriving at valuations.

Market value is an important basis for the clubs in making decisions on transfers, the renewal of contracts, and investments in youngsters. Therefore, our model has to give a value for each player that could combine both the present value and future potential of the player, so as to bring the financial investment closer to real performances.

The second goal is to watch and appreciate the substantially different positions and skill sets required by each different field position: attackers, defenders, midfielders, and goalkeepers-each has different roles in maximising a team's success and thus requires changes to the assessment model as a function of player type being evaluated. For instance, forward players can be mostly evaluated through goals scored, xG, assists, and shot conversion rates, while the defenders can be evaluated based on successful tackles, interceptions, and aerial duels.

Here, we start working on separate models for the attackers and the defenders since these two roles have a sharp difference on most of the performance indicators. This approach will allow a better capture of the subtleties involved with different playing styles and contributions than would a single, all-encompassing model. While this distinction adds to the complexity of the system, it does ensure that the model is contextually relevant and increases the accuracy of its predictions.

Another crucial goal is not just to provide a theoretical framework, but to demonstrate how this econometric model can be applied in real-world decision-making within football management. This is a model that will be useful for clubs, scouts, and analysts in order to make better decisions on player purchase, sales, and also contract negotiations. By offering a data-driven objective opinion of player value, the model identifies the risks of an overpriced player yet opens the door to indicating players with talent who may be overlooked by traditional scouting.

The practical value of this model lies beyond the mere ranking of players currently available in the market; it also serves as a tool for strategic long-term planning. Football clubs can utilise this system to monitor the development and improvement of young rising stars and analyse performance trends over time, thus enabling them to make better, more informed decisions when it comes to team development and resource allocation.

Ultimately, it bridges the gap between financial management and sporting performance analysis into one coherent framework. As football clubs currently have to balance books while remaining competitive, a sophisticated, evidence-based player valuing tool provides clubs with a valuable asset in terms of the ability to optimise returns on their investments in talent purchase. By aligning financial valuations with material performance measurements, clubs can determine that their spending more precisely mirrors on-field performance.

4.2.2 The Choice of Econometric Technique

Several different methodologies can be employed when developing an econometric model to estimate market value for football players based on the level of complexity required and the particular objectives of the analysis. A player valuation problem involves complexity due to the differences in player positions and responsibilities within each position. For example, a defender plays a decisive role in a team's overall success, so using one strategy for all players is difficult because his contribution differs a lot from an attacker.

There are three main approaches we could consider:

1. Separate models for each role: this first approach involves creating two models for different positions[45], particularly one model for attackers and one model for defenders. This approach allows variables to be tailored according to the different types of skills and contributions that may be specific to each role. Attackers are more often rated based on their offensive output, such as goals, expected goals, and assists, while defenders are more often assessed with defensive indicators such as tackles, interceptions, and successful aerial duels. By separating them, it ensures that each position's specific variables are better captured. This is the approach that I have chosen to use for my studies.

2. A Unified Model with Categorical Variables: another approach might be a global model that includes all the players, in which the player position becomes a categorical variable (attackers, defenders, midfielders), possibly using different weights or dummy variables for the performance measures depending on the player's position. This approach keeps just a single model but adjusts the parameters for each position. With this approach, the framework is simplified, but it risks concealing some of the subtler interrelations of position-specific metrics with market value, as not all variances in the different roles may be acknowledged.

3. Advanced Techniques (Machine Learning): Another route would be to apply machine learning techniques such as random forest[46], gradient boosting, or neural networks. These models could handle the complex, non-linear relationships between performance

[45] Clarke (2021), "The Moneyball Method: Using Data to Build a Football Dream Team (On a Budget)", https://www.graphext.com/post/the-moneyball-method-using-data-to-build-a-football-dream-team-on-a-budget

[46] Chunyang Huang and Shaoliang Zhang, "Explainable artificial intelligence model for identifying Market Value in Professional Soccer Players", https://ar5iv.labs.arxiv.org/html/2401.16795
Sulimov, "Performance Insights-based AI-driven Football Transfer Fee Prediction", https://ar5iv.labs.arxiv.org/html/2401.16795

metrics and market value. Machine learning algorithms can also accommodate interactions between variables more flexible than traditional econometric models. However, these models often suffer from reduced interpretability, making it difficult to understand which variables drive a player’s value. Additionally, access to large-scale, high-quality data is required to properly train such models. Given the objectives and scope of this study, we have opted for a more interpretable and manageable technique: linear regression.

Within the possible approaches, we chose a position-specific model because we believe it allows for the best compromise between accuracy and interpretability. A single model using categorical variables risks overly simplifying the importance of given metrics by assuming its effect is constant across all positions. On the other hand, much more powerful advanced machine learning techniques involve a trade-off in interpretability and data requirements, making them less suitable for our goal of creating an interpretable, easily understandable system for player market value assessment.

Having different models for attackers and defenders will allow the capture of unique metrics relevant to each, and their relative importance in determining market value will be suitably expressed. This choice is the least complicated yet effective way of reaching our goal: to correctly express the market value of players with respect to their positions regarding their respective contributions to the game.

4.3 Econometric Model for Attacker Market Value

The evaluation of a football player's market value is a complex task, especially for attackers whose performance is often quantified using various metrics related to their contributions in scoring and assisting goals. In this section, we present an econometric model specifically designed to estimate the market value of attackers. The model uses key performance indicators that capture the most relevant aspects of an attacker's role on the field.

4.3.1 Key Variables for Attacker Market Value Estimation

In constructing the econometric model for attackers, the selection of appropriate independent variables is crucial. Each variable chosen reflects a different aspect of an attacker's contribution to the team's offensive play and his overall market value. Below, we define and explain each variable used in the model, assuming no prior knowledge of football statistics (in our case, all the variables will take into account the data of a single sporting season).

1. **Goals Scored**: Goals are the most straightforward measure of an attacker's effectiveness. Each goal represents a successful offensive action that directly contributes to the team's success. Attackers who score frequently tend to have higher market values, as scoring is their primary responsibility on the field.
2. **Assists Provided**: An assist occurs when a player passes the ball to a teammate who then scores. This variable reflects the attacker's ability to create goal-scoring opportunities for others, demonstrating a well-rounded contribution to the team's offensive efforts.
3. **Expected Goals (xG)**: Expected goals (xG) is an advanced metric that quantifies the quality of chances a player generates. It considers factors such as shot

distance, angle, and type. While goals scored measure actual outcomes, xG provides insight into the likelihood that a player’s shots will result in goals, offering a more predictive measure of offensive performance.

4. **Total Shots**: This variable represents the number of times an attacker attempts to score. More shots often indicate that a player is actively involved in offensive play. However, the efficiency of these shots (i.e., converting shots into goals) is equally important in determining the player’s value.
5. **Successful Dribbles**: Dribbling refers to a player’s ability to bypass opponents while maintaining control of the ball. Attackers who excel in dribbling can disrupt defensive lines and create space for themselves or their teammates, which is a valuable skill.
6. **Key Passes**: A key pass is one that directly leads to a shot, even if the shot does not result in a goal. This metric captures the creativity and vision of an attacker, as it reflects his ability to generate scoring opportunities for others.
7. **Minutes Played**: The total number of minutes the player has spent on the field during the season is essential for contextualising performance data. A player who has played more minutes is likely to have had more opportunities to influence the game. At the moment we keep this data "raw" even if it would need a standardisation. We will perform the transformation later.
8. **Age**: Age is a crucial factor in determining a player’s market value. Younger players often have higher market values due to their potential for development and longer career prospects. Older players, on the other hand, may see their value decrease as they approach the latter stages of their career.
9. **Market Value** (Dependent Variable): The market value is the dependent variable. It is expressed in monetary terms (in millions of euros) and reflects the player’s perceived worth in the transfer market.

In summary, the selected variables encompass a wide range of performance indicators, from basic metrics like goals scored to more nuanced data like expected goals and key passes. These variables are intended to provide a comprehensive view of an attacker's contribution to the game and how that contribution translates into market value.

4.3.2 Formulation of the Attackers Regression Model

To construct the econometric model for estimating attackers market value, data from 50 attackers who played in the 2023-2024 Serie A season were collected. The data includes the variables described above in the previous subsection. The dependent variable in this model is the player's market value, which is influenced by both on-field performance and personal characteristics.

Table 4.1 - Data collected for attackers in Serie A 2023-2024

Player_Name	Goals	Assists	Expected_Goals_xG	Total_Shots	Successful_Dribbles	Key_Passes	Minutes_Played	Age	Contract_Lenght	Market_Value_(€M)
Lautaro	24	3	18,78	108	19	36	2668	27	5	110
Kvaratskhelia	11	6	13,02	128	101	63	2752	23	3	80
Leao	9	9	9,14	70	73	58	2524	25	4	90

Thuram	13	7	14,82	77	33	29	2707	27	4	65
Soule	11	3	11,16	83	91	69	3141	21	5	25
Dybala	13	9	11,58	59	28	50	1977	30	1	20
Gudmundsson	14	4	10,18	57	41	78	3024	27	3	30
Gonzalez	12	2	11,25	87	34	22	1914	26	5	35
Zirkzee	11	4	8,93	86	52	43	2772	23	5	50
De Ketelaere	10	8	7,52	45	40	49	2045	23	3	34
Lookman	11	7	10,51	67	39	50	1903	26	2	40
Zaccagni	6	1	5,52	42	53	25	1974	29	5	20
Zapata	12	4	13,02	99	36	27	2893	33	2	8
Osimhen	15	3	16,85	90	16	23	1990	25	2	100
Chiesa	9	2	7,72	75	35	54	2207	26	4	35
Scamacca	12	6	7,92	60	15	33	1459	25	3	35

Giroud	15	8	15,48	77	3	32	2374	37	3	3
Politano	8	7	7,53	75	31	71	2387	31	3	13
Ngonge	7	2	4,64	42	34	21	1438	24	4	12
Vlahovic	16	4	18,01	110	18	29	2318	24	2	65
Lukaku	13	3	12,38	65	11	29	2648	31	3	30
Pulisic	12	8	8,42	65	47	44	2621	25	3	40
Pinamonti	11	1	10,05	83	12	18	3100	25	3	15
Orsolini	10	2	8,21	59	23	23	1795	27	3	16
Lucca	8	4	8,86	61	12	20	2602	24	4	12
Immobile	7	1	10,67	49	8	17	1662	34	4	4
Retegui	7	2	7,09	63	14	11	2226	25	4	16
Krstovic	7	1	12,17	92	29	41	2391	24	3	6
Cheddira	7	1	9,84	60	12	16	2132	26	4	6

Beltran	6	2	3,71	31	12	25	1696	23	4	16
Jovic	6	1	5,42	22	7	11	872	26	1	7
Pasalic	6	6	7,85	40	3	10	2051	29	1	13
Okafor	6	2	4,73	22	14	8	872	24	4	20
Sanabria	5	0	9,85	54	14	15	2134	28	2	7
Laurentie	5	4	6,41	84	63	46	2929	25	3	10
Anderson	5	6	5,07	43	42	38	2784	31	3	8
Thauvin	5	3	6,26	48	35	38	1717	31	1	3
Raspadori	5	3	5,7	61	14	36	1574	24	4	25
Arnautovic	5	3	4,93	17	1	16	785	35	1	4
Djuric	5	1	4,79	25	2	21	1208	34	1	1,5
Milik	4	1	5,56	37	5	15	902	30	2	6
Mota	4	3	5,11	38	15	33	2060	26	5	5

Pavoletti	4	1	3,03	25	0	7	620	35	2	0,8
Colombo	4	1	3,97	37	12	8	1367	22	4	6
Castellanos	4	3	7,39	54	11	19	1676	25	4	15
Luvumbo	4	5	5,05	51	26	21	1947	22	3	7,5
Caputo	3	0	4,14	23	5	10	1276	37	1	1
Kaio Jorge	3	0	6,57	27	5	8	807	22	4	4
Nzola	3	3	4,88	32	14	21	1550	28	3	7
Vlasic	3	2	2,72	46	25	32	2615	26	3	10

Source: https://www.transfermarkt.it/ (2024); https://it.whoscored.com/ (2024)

Using the collected data, the regression model was estimated using ordinary least squares (OLS) with the software RSTUDIO. The regression output is the following:

Table 4.2 - Regression output, primitive attackers model

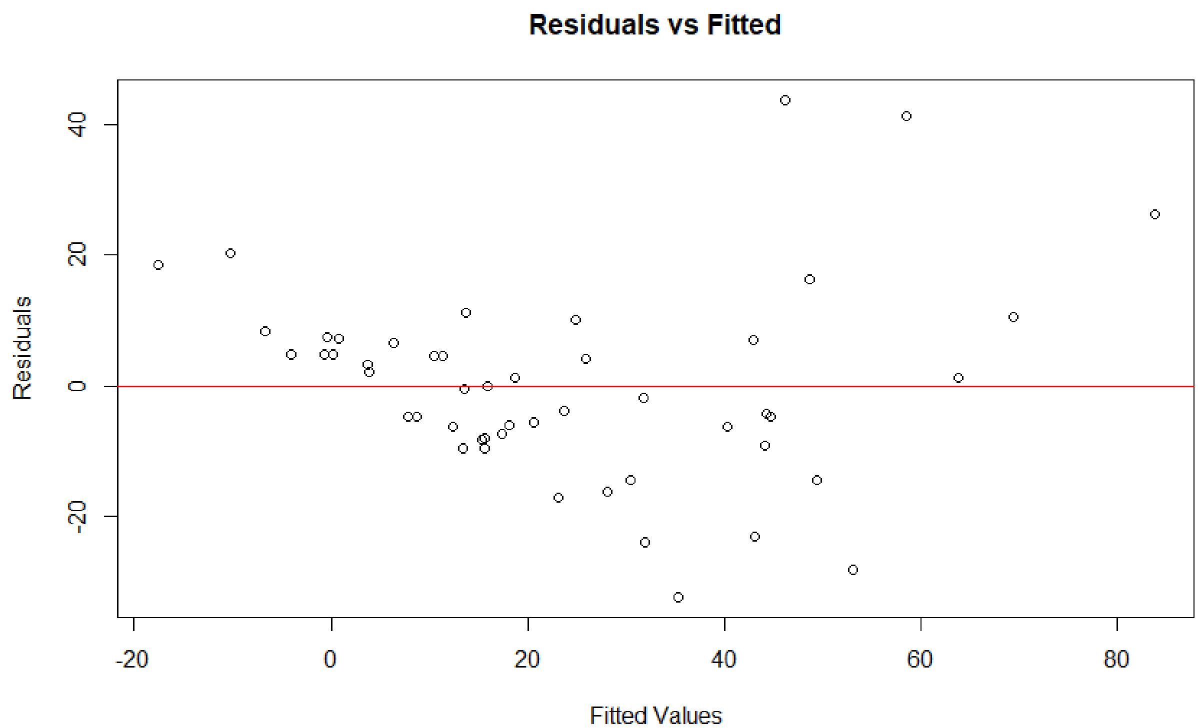
```
Residuals:
    Min      1Q  Median      3Q     Max 
-32.340  -7.916  -0.287   6.874  43.749 

Coefficients:
                     Estimate Std. Error t value Pr(>|t|)   
(Intercept)         28.759476  24.515410   1.173  0.24769   
Goals                3.432802   1.195073   2.872  0.00649 **
Assists              1.232729   1.250592   0.986  0.33020   
Expected_Goals_xG    1.088717   1.628887   0.668  0.50773   
Total_Shots          0.109540   0.226703   0.483  0.63160   
Successful_Dribbles  0.391383   0.189014   2.071  0.04489 * 
Key_Passes          -0.163498   0.236605  -0.691  0.49355   
Minutes_Played      -0.012166   0.005778  -2.106  0.04155 * 
Age                 -1.328620   0.713458  -1.862  0.06993 . 
Contract_Lenght      0.679417   2.423501   0.280  0.78066   
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 16.36 on 40 degrees of freedom
Multiple R-squared:  0.6853,	Adjusted R-squared:  0.6145 
F-statistic: 9.678 on 9 and 40 DF,  p-value: 1.163e-07
```

Now we will analyse our output step by step, starting from the residuals, then moving on to the coefficients, the model fit and possible improvements to the model.

Residuals:

- The range of residuals is still quite large, from -32.34 million euros to 43.75 million euros, indicating that the model has limitations in accurately predicting the market value of certain players.

Graph 4.1 - Residual vs Fitted, primitive attackers model



In a well-specified linear regression model, we expect the residuals to be randomly distributed around 0 and their variance to be constant (homoscedasticity). Here, however, there appears to be a slight increasing variance of the residuals as the fitted values increase, especially at the higher values. This may indicate heteroskedasticity, that is, that the variance of the residuals is not constant. At larger fitted values, the dispersion of the residuals appears to increase.

We expect the residuals to be randomly distributed with no obvious patterns. However, we can see a slight “U” shape in the plot (the residuals tend to go down and then up again), which suggests that there may be a non-linearity in the relationship between the explanatory variables and the dependent variable. This may indicate that the linear model may not be the best at capturing the relationship between the variables. It may be necessary to include non-linear terms or transformations (such as higher degree polynomials or logarithms) to improve the model.

Finally, while not perfect, the residuals appear to be roughly zero on average, which is a good thing for the model. There are no systematic deviations that would indicate obvious bias, which gives us hope that we can work on improving the model.

Coefficients Analysis:

Intercept: 28.76 (p-value = 0.24769)

The intercept is not statistically significant. This suggests that the baseline market value (when all independent variables are zero) is not significantly different from zero.

Goals: 3.43 (p-value = 0.00649)

Statistically significant at the 1% level ($p < 0.01$). Each additional goal increases the player's market value by 3.43 million euros, confirming that goals are a crucial driver of an attacker's market value.

Assists: 1.23 (p-value = 0.33020)

Assists are not statistically significant in this model ($p > 0.05$), which is somewhat surprising. This could indicate multicollinearity or that the relationship between assists and market value is less direct than expected.

Expected Goals (xG): 1.09 (p-value = 0.50773)

xG is not significant. While xG is an advanced metric that reflects a player's underlying performance, its statistical insignificance may stem from redundancy with the actual goals scored, which already captures the performance.

Total Shots: 0.11 (p-value = 0.63160)

Like xG, the total number of shots does not significantly affect market value in this model. This could indicate that raw shot volume is less important than metrics related to goal-scoring efficiency.

Successful Dribbles: 0.39 (p-value = 0.04489)

Statistically significant at the 5% level ($p < 0.05$). Dribbling success has a positive impact on market value, reflecting the importance of individual skill and offensive creativity for attackers.

Key Passes: -0.16 (p-value = 0.49355)

Not statistically significant and even has a negative sign, which is counterintuitive. This could suggest that key passes, at least in this model, do not play a substantial role in determining market value, or that there is some multicollinearity with other variables like assists.

Minutes Played: -0.012 (p-value = 0.04155)

Statistically significant at the 5% level, but with a negative sign. This is somewhat unexpected because more minutes usually imply higher value. It may indicate that playing more minutes could correlate with fatigue or injury, factors not directly captured by the model.

Age: -1.33 (p-value = 0.06993)

Borderline statistically significant ($p < 0.1$). Age negatively impacts market value, which is intuitive since older players generally have shorter career prospects.

Contract Length: 0.68 (p-value = 0.78066)

Contract length is not significant. This suggests that, in this model, the remaining length of a player's contract does not directly affect market value. However, this may be due to the relatively small sample size or because contract length effects could vary based on player profile.

Model Fit:

- **Multiple R-squared = 0.6853**: The model explains 68.5% of the variance in the market value of attackers. This indicates a decent fit, but there is room for improvement, especially considering the residuals' large range.
- **Adjusted R-squared = 0.6145**: This value, slightly lower than the R-squared, accounts for the number of predictors. It suggests that while the model is useful, there might be redundant or unnecessary variables.

- **F-statistic: 9.678 (p-value = 1.163e-07)**: The overall model is statistically significant, meaning that the variables together significantly predict market value.

Model Improvements

The insignificance of variables like assists, key passes, and contract length could suggest multicollinearity. Conducting a Variance Inflation Factor (VIF) analysis would help identify whether these variables are redundant. So i computed it and this is the results:

Table 4.3 - VIF analysis output

Goals	Assists	Expected_Goals_xG	Total_Shots	Successful_Dribbles
4.943363	1.789027	7.436557	6.160870	3.221999
Minutes_Played	Age	Contract_Lenght	Key_Passes	
2.739202	1.561128	1.590136	3.241111	

The key data that we are interested in analysing are the following:

Goals (VIF = 4.94):

A VIF below 5 is generally considered acceptable, but it is on the higher end, indicating some level of multicollinearity. However, since goals are a key variable, it is important to retain them in the model.

Expected Goals (xG) (VIF = 7.44):

This value exceeds 5, which indicates significant multicollinearity. Given that xG and goals are related metrics (xG is predictive of goals), this is expected. Including both in the model may reduce their individual significance, and we might need to consider removing or transforming xG.

Total Shots (VIF = 6.16):

This variable also exhibits a relatively high level of multicollinearity, which might stem from its correlation with goals or xG. Total shots may overlap with other goal-related variables, reducing its independent contribution.

Table 4.4 - Correlation matrix, primitive attackers model

	Goals	Assists	Expected_Goals_xG	Total_Shots	Successful_Dribbles	Key_Passes	Minutes_Played	Age	Contract_Lenght
Goals	1.0000	0.4184	0.8655	0.7138	0.2354	0.4070	0.5150	-0.0652	0.1404
Assists	0.4184	1.0000	0.3116	0.3245	0.4105	0.5821	0.4073	-0.0272	-0.0819
Expected_Goals_xG	0.8655	0.3116	1.0000	0.8040	0.1915	0.2902	0.5261	-0.0330	0.0735
Total_Shots	0.7138	0.3245	0.8040	1.0000	0.5311	0.5077	0.7272	-0.2676	0.2438
Successful_Dribbles	0.2354	0.4105	0.1915	0.5311	1.0000	0.7399	0.5868	-0.3768	0.3012
Key_Passes	0.4070	0.5821	0.2902	0.5077	0.7399	1.0000	0.6130	-0.1825	0.1472
Minutes_Played	0.5150	0.4073	0.5261	0.7272	0.5868	0.6130	1.0000	-0.2179	0.2983
Age	-0.0652	-0.0272	-0.0330	-0.2676	-0.3768	-0.1825	-0.2179	1.0000	-0.5147
Contract_Lenght	0.1404	-0.0819	0.0735	0.2438	0.3012	0.1472	0.2983	-0.5147	1.0000

The correlation matrix shows the linear relationships between the selected variables related to player performance. Correlation values range from -1 to 1, where a positive value indicates a direct correlation between two variables, while a negative value indicates an inverse relationship.

Some significant correlations emerge from the analysis of the matrix. For example, there is a very high correlation between Goals and Expected Goals (XG) (0.865), which confirms the direct relationship between the number of goals scored and the expected value of goals per player. The correlation between Total Shots and Expected Goals (0.804) also suggests that players who shoot more frequently are more likely to have a higher XG value.

Another significant correlation is between Key Passes and Successful Dribbles (0.739), which indicates that players who are good at dribbling tend to create more scoring opportunities. Furthermore, the correlation between Minutes Played and Total Shots (0.727) suggests that players with more time on the pitch have more opportunities to influence the game.

On the other hand, negative correlations are observed between Age and Contract Length (-0.515) and between Age and Successful Dribbles (-0.377), indicating that older players tend to have shorter contracts and fewer successful dribbles, consistent with a natural decline in physical performance.

This preliminary analysis suggests that some variables, such as Goals, Expected Goals and Minutes Played, are strongly correlated, which may require further evaluation to avoid multicollinearity issues in subsequent econometric analyses.

To increase the precision of the model I then decided to make some substantial changes. I tried to recalculate the model several times by removing some variables, modifying others and creating interactions between variables. I obtained the best result by eliminating from the model the non-significant or highly multicollinear variables such as Expected Goals, Total Shots, Contract Length, Key Passes and Assists, modifying through the variable Minutes Played in Minutes per 90 (minutes played per 90 minutes, the duration of a match) and transforming the variable Age into Age squared, so as to capture a possible non-linear relationship between age and the market value of a player.

The model is as follows:

Table 4.5 - Output regression, final attackers model

```
Residuals:
    Min      1Q  Median      3Q     Max
-31.700  -8.120  -1.128   6.446  47.663

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)         15.47676   11.61482   1.333  0.18940
Goals                4.62917    0.61200   7.564  1.5e-09 ***
Successful_Dribbles  0.36102    0.13306   2.713  0.00941 **
Minutes_per_90      -0.89401    0.42932  -2.082  0.04302 *
Age2                -0.02574    0.01034  -2.490  0.01654 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 15.89 on 45 degrees of freedom
Multiple R-squared:  0.6661,    Adjusted R-squared:  0.6364
F-statistic: 22.44 on 4 and 45 DF,  p-value: 3.055e-10
```

Source: RSTUDIO (2024)

The model has a **Multiple R-squared** value of **0.6661** and an **Adjusted R-squared** of **0.6364**, indicating that approximately **66.6%** of the variance in market value can be explained by the four independent variables. The **F-statistic** is 22.44 with a p-value of 3.055e-10, showing that the model is statistically significant overall.

Each variable contributes uniquely to the model:

Goals: This is the most influential variable with a coefficient of **4.62917**. For each additional goal scored by the attacker, their market value increases by approximately 4.63 million euros, assuming all other variables remain constant.

Successful Dribbles: A coefficient of **0.36102** indicates that dribbling success also positively impacts market value, although to a lesser extent than goals.

Minutes per 90: Interestingly, this variable has a **negative** impact, with a coefficient of **-0.89401**. This negative coefficient in the regression model suggests that, all else being equal, players who play fewer minutes per game tend to have higher market values. While this may seem counterintuitive at first glance, it can be explained by several contextual factors:

1. **Young, promising players often play fewer minutes**: Young and highly talented players, particularly those under the age of 21, may not always be regular starters in their teams. Clubs often manage the development of young prospects carefully, limiting their playing time to prevent burnout and allow them to adapt to the demands of professional football gradually. Despite playing fewer minutes, these players often have high market values due to their potential for growth and future success.

   For instance, an 18-year-old forward who shows flashes of brilliance in limited game time might be valued higher than an older player who regularly plays but has already reached his performance peak. This reflects the market's tendency to place a premium on young players with high potential, even if they haven't yet accumulated significant playing time.

2. **Interest in players who are not key to their current team**: Another reason for the negative coefficient could be that clubs are more likely to target players who are not integral parts of their current teams. Players who are not playing the majority of minutes may be viewed as more obtainable or transferable, making them attractive targets in the transfer market. These players may still demonstrate significant quality in the minutes they do play, which catches the attention of scouting departments. This is particularly relevant for large clubs looking to buy potential "undervalued" players from smaller teams where they aren't regular starters.

   For example, a player coming off the bench for a mid-table team might still be highly effective, and a bigger club might see the potential to give that player more game time in their squad. This would raise the player's market value despite their limited minutes on the pitch.

3. **Efficiency in limited minutes**: Another explanation for the negative impact of minutes played per 90 on market value is that players who are able to make a significant impact in limited playing time are often viewed as highly efficient and valuable assets. Forwards who can score goals or make key contributions even with fewer minutes demonstrate a level of quality and effectiveness that may increase their value.

   In contrast, players who require significant playing time to produce similar results might not be seen as being as "efficient" in the transfer market, particularly for clubs looking for players who can deliver results quickly in high-pressure situations. This could explain why players with fewer minutes but strong performance indicators (like goals per 90 minutes) may have higher market values.

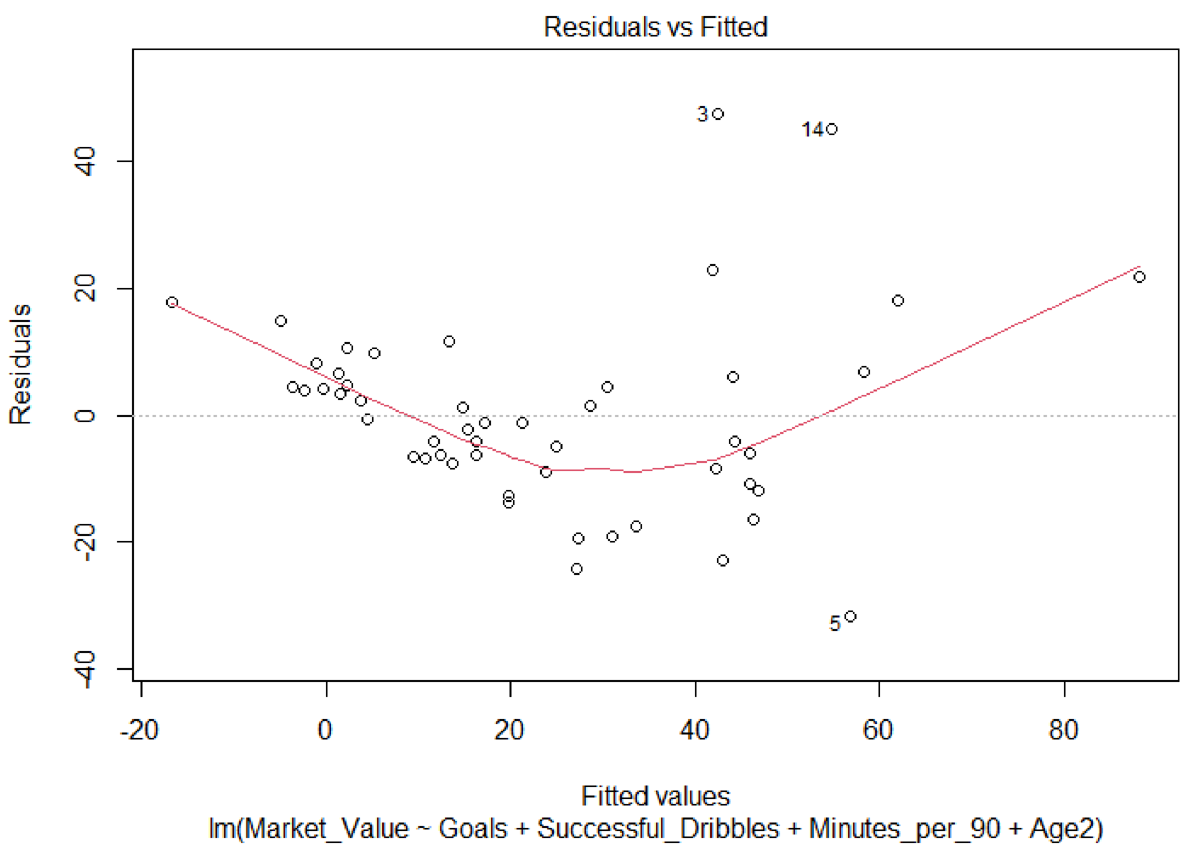
This result highlights that, while the minutes played is a relevant variable in the model, it must be interpreted in the broader context of player development, market demand, and tactical management within modern football.

Age²: The negative coefficient of **-0.02574** for the squared age term reflects the diminishing market value of players as they age, aligning with expectations that younger players, particularly those in their early 20s, tend to command higher market values.

To check the robustness of the model I performed some statistical tests on the residuals.

Residuals vs Fitted

Graph 4.2 - Residuals vs Fitted, final attackers model



The red curve (the smoothing line of the residuals) shows a "U" shape, which suggests that there could be a non-linear relationship between the independent variables and the market value (Market_Value). This type of pattern indicates that the linear model fails to fully capture the complexity of the relationship. In a good linear model, the residuals should be randomly distributed around zero, without any particular shape.

There are some points visibly far from the majority of the data, in particular the points labelled as "3" (Leao), "14" (Osimhen) and "5" (Soulé). These could be considered outliers, that is, observations that significantly influence the results of the model, the first two seem to be "overpriced" players, that is, the model predicts a lower market value than the actual observed one. I want to point out that the player Victor Osimhen spent most of the season injured, so it is normal that his performance in the year analysed does not reflect the market value. The third player, Matias Soulé, on the other hand, appears to be "underpriced", meaning that the model predicts a higher market value than observed, indicating that the market may not fully recognize its value based on the characteristics included in the model.

The "U" shape could also suggest possible heteroskedasticity, that is, that the variance of the residuals is not constant across all predicted values. This would violate one of the main assumptions of linear regression models.

Breush-Pagan test

The **Breusch-Pagan test** is used to check for heteroscedasticity in a regression model. The null hypothesis of the test is that the residuals have constant variance (homoscedasticity). If the p-value is below a certain significance threshold (typically

0.05), it indicates that heteroscedasticity is present and the assumption of constant variance is violated.

The results of the test, always carried out on RSTUDIO, are the following:

Test statistic (BP): 10,24

Degrees of freedom (df): 4

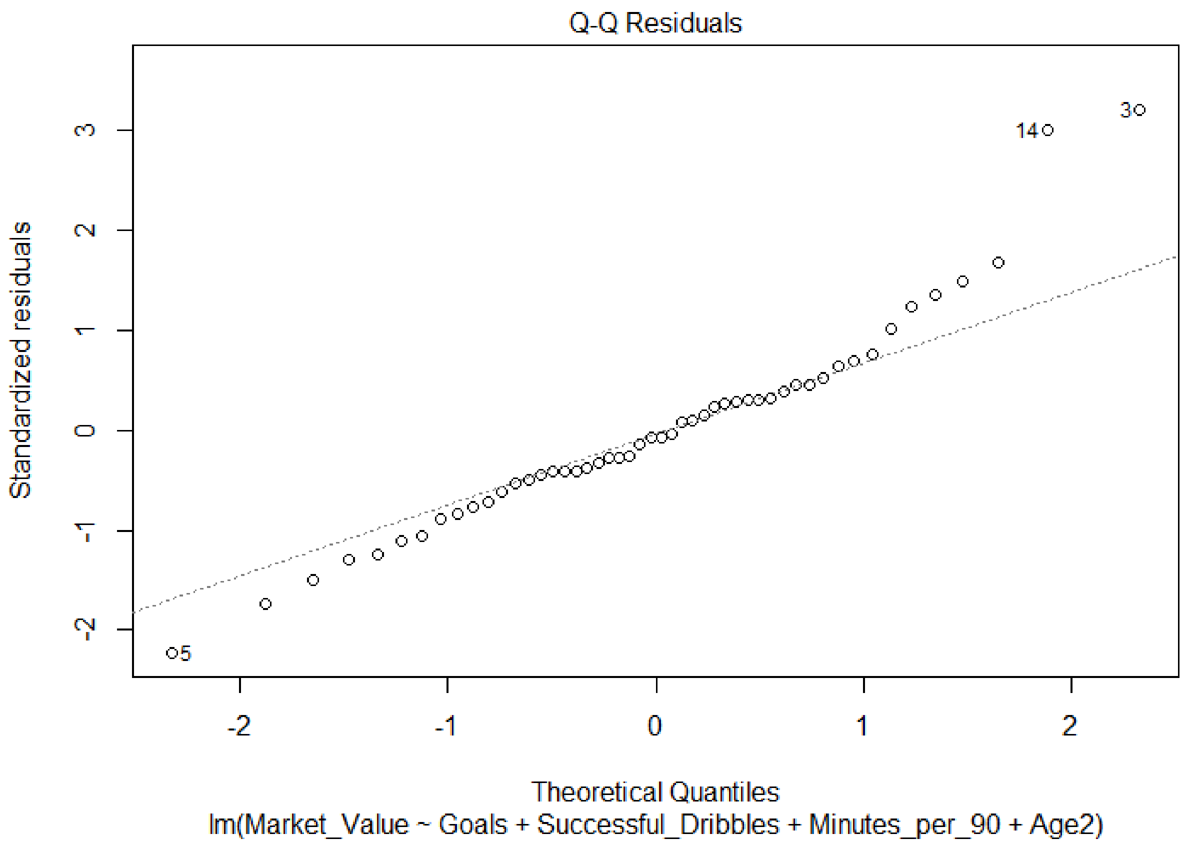
p-value: 0.03658

A p-value less than 0.05 suggests that we can reject the null hypothesis (that the residuals are homoscedastic, i.e., have constant variance). In this case, the p-value of 0.03658 is less than 0.05, indicating that there is significant evidence of heteroscedasticity in the model.

The Breusch-Pagan test confirms what was said above and suggests that there are signs of heteroscedasticity in the model. This means that the variance of the errors is not constant, and this could indicate that the model may need some correction, such as using a model that is robust to heteroscedasticity or transforming some variables.

Q-Q Plot

Graph 4.3 - Q-Q plot, final attackers model



Most of the points appear to follow the dashed line, which represents the theoretical normal distribution, quite well. This suggests that the residuals are nearly normally distributed, especially for the central part of the distribution.

We see significant deviations at the extremes, particularly for observations "3" and "14" in the right tail. Observation "5" in the left tail also deviates slightly, but not as dramatically. These are the same outliers we noted in Graph 4.2.

In summary, the residuals follow the normal distribution acceptably for most of the data, but there are some outlier observations that should be investigated further, especially if they significantly affect the model. This is consistent with what was observed in the Residuals vs Fitted plot, where points 3, 5, and 14 showed peculiar behaviours.

Shapiro-Wilk normality test

To support the Q-Q Plot I wanted to compute another test, the Shapiro-Wilk normality test. The Shapiro-Wilk normality test is a formal statistical test used to assess whether a sample (in this case, the residuals of the model) comes from a normally distributed population. It is particularly sensitive to deviations from normality, making it a helpful tool to supplement graphical analysis, like the Q-Q plot.

The results of the test, carried out on RSTUDIO, are the following:

W: 0.94508:

The W statistic is close to 1, which suggests that the data is approximately normally distributed.

p-value: 0.0215: The p-value is lower than the typical significance level of 0.05, meaning we reject the null hypothesis that the residuals are normally distributed. Since the p-value is quite low (0.0215), there is evidence to suggest that the residuals deviate from a normal distribution.

The Shapiro-Wilk test, combined with the results displayed in the Q-Q plot, suggests that there are deviations from normality, particularly at the extremes, as observed for observations 3, 14, and partly for 5. These deviations could affect the assumptions of the linear model, such as the normality of the residuals, which is essential for significance tests and for the validity of confidence intervals.

The model, designed to explain the market value of strikers based on variables such as goals scored, successful dribbles, minutes played per 90 minutes and age (Age2), has shown promising results, but is not without its problems.

The model revealed statistically significant relationships between some key variables and market value, confirming theoretical expectations. In particular, the inclusion of minutes played per 90 minutes and goals scored provided a robust explanation of the variations in market value, highlighting the importance of these metrics to evaluate the performance and relevance of a striker. The transformation of some variables such as Age2 allowed to capture non-linear effects, improving predictive accuracy.

On the other hand, the analysis of the residuals revealed some problems, especially in relation to normality and heteroskedasticity. The Shapiro-Wilk test rejected the assumption of normality, suggesting the presence of outliers or a skewed distribution of the residuals. Furthermore, the Breusch-Pagan test indicated a significant presence of heteroskedasticity, a sign that the variability of the residuals is not constant and could undermine the efficiency of the model estimates.

To improve the validity of the model, further developments are needed. It may be useful to explore alternative transformations of the variables or residuals to correct for heteroskedasticity and normality problems. Furthermore, it would be appropriate to include new variables that capture unconsidered aspects that can significantly influence the market value. In addition, robust regression techniques could be used to reduce the influence of outliers on the model.

In conclusion, the proposed model represents a good starting point to understand the determinants of strikers' market value, but it requires further optimizations to be considered fully reliable. The integration of advanced econometric techniques and a more in-depth analysis of the residuals can help improve the predictive accuracy of the model and provide a more robust framework for managerial decisions in the context of the football market.

4.4 Econometric Model for Defender Market Value

While attackers are primarily valued for their offensive contributions, the market value of defenders is shaped by their ability to prevent goals and support the team's defensive structure. Therefore, the econometric model for defenders includes variables that reflect their defensive capabilities, physical characteristics, and discipline on the field.

The key variables chosen for the model are tailored to evaluate defensive performance, focusing on factors such as tackles, interceptions, aerial duels won, and disciplinary issues. Additionally, age and physical fitness, measured via Body Mass Index (BMI), are considered, as these elements are crucial in determining a defender's longevity and effectiveness.

4.4.1 Key Variables for Defender Market Value Estimation

1. **BMI (Body Mass Index)**: Physical attributes are crucial for defenders, particularly in terms of strength and stamina. A balanced BMI can indicate a player's athletic condition, with extremes on either side (too high or too low) possibly leading to

lower market value due to concerns over fitness or agility. BMI is calculated by the following formula[47]:

$$BMI = weight_{kg} \,/\, height_{m}^{2}$$

I would add that in conjunction with the BMI it would be important to have data on the percentage of body fat for each player, but this data is not available online.

To penalise extreme data (underweight or overweight players), a further transformation to the formula is needed. I thought of creating an ideal value (given by the average of the most "expensive" defenders in the world) and subtracting it from the calculated BMI.

$$BMI\, penalty = |BMI - 23.8|$$

2. **Tackles**: Tackles represent one of the key defensive actions. A higher number of successful tackles indicates a defender's effectiveness in dispossessing opponents, which is highly valued.
3. **Interceptions**: Interceptions are another critical defensive action. A player who reads the game well and intercepts passes can prevent dangerous situations from developing, adding to their market value.
4. **Aerial Duels**: The ability to win aerial duels is vital, particularly for central defenders, as it helps in both defensive and offensive situations (e.g., set pieces). A high success rate in aerial duels can significantly increase a defender's value.
5. **Indiscipline**: This variable tracks the number of fouls committed and yellow and red cards received from a player. While defenders need to be aggressive, excessive fouls and suspensions due to indiscipline can detract from their market value. The weight attributed to fouls and cards is different, the indiscipline value

[47] https://en.wikipedia.org/wiki/Body_mass_index

is calculated using the following formula (the value of the weights was chosen by me, it is a personal and not objective evaluation):

$$Indiscipline = fouls + 5 * yellow\,cards + 10 * red\,cards$$

6. **Minutes Played**: As with attackers, the number of minutes played reflects a defender's importance to their team. Consistent playing time often correlates with higher value, as it shows the player's reliability and fitness.
7. **Age**: Like the attacker model, age plays a significant role in determining a player's market value. Younger defenders may have higher potential and value, while older defenders might face depreciation despite their experience.
8. **Market Value** (Dependent Variable): As in the model for attackers, the market value is the dependent variable. It is expressed in monetary terms (in millions of euros) and reflects the player's perceived worth in the transfer market.

4.4.2 Formulation of the Defenders Regression Model

Following the approach taken with attackers, we construct a multiple linear regression model using the selected variables. For this model, I collected data from 50 defenders playing in Serie A during the 2023/2024 season.

Table 4.6 - Data collected for defenders in Serie A 2023-2024

Player_Name	Minutes_Played	Age	BMI_Penalty	Tackles	Interceptions	Aerial_Duels	Indiscipline	Market_Value_(€M)
Luperto	3317	28	3,241	1	0,9	2,1	1,2	3,5
Baschirotto	3295	27	4,260	0,6	0,6	2,5	1,3	6
Pérez	3240	24	1,647	2	1,3	2,3	1,9	13
Bremer	3234	27	0,599	1,6	1,1	2,8	3,2	60
Pongracic	3220	27	0,981	1,2	1,4	1,7	2,6	7
Dossena	2979	25	0,035	1,6	1,3	3,7	2,1	5
Okoli	2915	23	1,144	1,6	0,9	2,3	2,3	8
Mancini	2874	28	2,470	0,9	1	1,3	3	25
Djimsiti	2832	31	0,808	1,1	1,5	3	1	10
S. Romagnoli	2716	34	1,183	1,1	0,9	3,2	1,8	0,8

Gatti	2641	26	0,531	1	0,6	1,7	2,3	18
Erlic	2627	26	0,175	1,2	0,9	1,9	2	3
Marì	2618	31	0,444	1,2	1,1	1,4	2,3	3,5
Rrahmani	2603	30	0,200	1,1	0,7	2,7	1,7	15
Buongiorno	2530	25	0,023	2,4	2,4	2,8	3,5	35
Milenkovic	2506	26	0,131	1,2	0,8	4	1,6	15
Beukema	2478	25	2,014	0,9	0,5	1,6	1,5	18
Scalvini	2466	20	2,544	1,6	1,8	2,9	2,2	45
Magnani	2442	28	0,254	1	1,4	1,8	2	3
A. Romagnoli	2410	29	2,193	0,9	1	3	2,8	15
Acerbi	2387	36	0,072	0,8	1,1	2,7	0,7	3,5
Danilo	2361	33	1,057	2,1	1,3	1,4	2,5	10

Vasquez	2315	25	3,055	1,4	1,2	2,5	2,2	10
Bani	2310	30	1,731	1	0,8	2	2,8	2,5
Martinez Quarta	2292	28	0,907	2,6	1	1,9	3,6	15
Bastoni	2283	25	3,024	1,6	0,8	1,5	2,2	70
Ferrari	2269	32	0,844	0,7	0,8	1,5	2	1,4
Calafiori	2258	22	1,098	1,6	1,7	2,4	2,4	45
Rodriguez	2234	32	0,554	1,4	0,8	1,5	1,3	3,5
Kolasinac	2183	31	1,581	1,5	0,7	1,4	2,2	10
Llorente	2180	31	1,943	1,1	0,8	1,9	2	7
Ndika	2171	25	1,556	1,3	0,6	1,7	2,1	25
Caldirola	2141	33	2,804	0,9	1	1,3	2	1,3
Tomori	2125	26	0,717	1,6	1	1,2	2,6	40

Lucumì	2124	26	1,495	1,7	0,6	1,2	0,9	17
Jesus	2118	33	0,743	1,8	1	2	2,7	2,5
Kristensen	2097	22	3,566	1,3	1,1	2,6	1,6	4,5
Bijol	2083	25	0,023	1,8	1	3,7	2,2	15
Ranieri	2068	25	2,638	0,5	1,1	1,7	4,4	8
De Winter	2066	22	2,419	1,6	0,8	2,5	1,9	12
Walukiewicz	2056	24	1,165	1,4	0,6	1,7	2,5	2,5
Dawidowicz	2054	29	0,249	1,8	0,8	2,7	2,4	3
Pirola	1993	22	2,763	1,4	0,9	2,6	2,5	5,5
Ismajli	1956	27	1,594	1	0,9	2,3	1,2	3
Gila	1724	24	0,425	1,5	1	1,6	2,6	13
Pavard	1679	28	0,387	1,5	1,4	1,9	2,7	50

Izzo	1655	32	0,387	1,2	1,2	1,1	4,2	2,5
Coppola	1636	20	0,614	1,5	1,1	4,5	4,3	6
Casale	1558	26	0,774	0,8	0,6	1,3	2,9	14
Thiaw	1531	23	0,152	1,7	0,6	2,1	3,7	25

Source: https://www.transfermarkt.it/ (2024); https://it.whoscored.com/ (2024)

Also in this case I carried out a multiple linear regression using the RSTUDIO software and the primitive output, without modifications, is shown in Table 4.6:

Table 4.7 - Regression output, primitive defenders model

```
Residuals:
    Min      1Q  Median      3Q     Max 
-15.026  -9.117  -3.591   5.975  49.815 

Coefficients:
                Estimate Std. Error t value Pr(>|t|)  
(Intercept)    36.199184  29.518335   1.226   0.2269  
Minutes_Played  0.002749   0.005421   0.507   0.6148  
Age            -1.423942   0.679785  -2.095   0.0423 *
BMI_Penalty    -0.088634   2.327797  -0.038   0.9698  
Tackles         6.938689   5.829453   1.190   0.2406  
Interceptions   8.683400   7.011435   1.238   0.2224  
Aerial_Duels   -4.259949   3.116827  -1.367   0.1790  
Indiscipline    0.872011   3.015107   0.289   0.7738  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 15.02 on 42 degrees of freedom
Multiple R-squared:  0.2432,	Adjusted R-squared:  0.1171 
F-statistic: 1.929 on 7 and 42 DF,  p-value: 0.08888
```

Coefficients Analysis:

Intercept: The intercept is 36.199, but it is not statistically significant (p=0.227). This suggests that the market value of a defender cannot be explained simply by a constant average without considering the independent variables.

Minutes Played: The coefficient for **Minutes_Played** is positive (0.0027), but very small and not statistically significant (p=0.615). This result suggests that, in this sample, the number of minutes a defender plays does not have a direct impact on their market value. This could be due to several reasons, such as defenders often having less visible roles compared to attackers or the fact that playing time alone does not fully capture a defender's contribution to the team.

Age: The coefficient for **Age** is negative (-1.42) and **statistically significant** (p=0.0423). This indicates that, for defenders, as age increases, the market value tends to decrease. This is consistent with general expectations in football, where younger players, especially those with potential for growth, tend to have higher market values.

BMI Penalty: The coefficient for **BMI_Penalty** is negative (-0.089) and not significant (p=0.97). BMI Penalty does not appear to be a relevant factor in explaining a defender’s market value. This may suggest that more accurate and advanced biometric data is needed to calculate a player's physical attributes. What we have is not enough to be significant in the model.

Tackles: The coefficient for **Tackles** is positive (6.94) but not significant (p=0.24). Although tackles are important for defenders, this result suggests that in this model, their direct contribution to market value may not be captured effectively or that there is variability in how tackles impact different defenders’ valuations.

Interceptions: Similarly, the coefficient for **Interceptions** is positive (8.68), but not significant (p=0.222). While interceptions are a key performance indicator for defenders, their influence on market value appears limited in this model, potentially due to the presence of other overlapping defensive metrics or due to sample size limitations.

Aerial Duels: The coefficient for **Aerial_Duels** is negative (-4.26) and not significant (p=0.179). This result suggests that winning aerial duels, although important in defensive play, may not have a clear positive impact on market value in this sample. This could be due to the fact that some defenders who excel in aerial duels might be less versatile in other aspects, affecting their overall market valuation.

Indiscipline: Finally, the coefficient for **Indiscipline** is positive (0.87) but not significant (p=0.773). This finding is somewhat counterintuitive, as higher indiscipline would typically be expected to negatively impact a player's value. The lack of significance here suggests that indiscipline may not be a decisive factor in the market valuation of defenders, though its impact could be more nuanced or indirect.

Model fit:

The model's Multiple R-squared is 0.2432, meaning that only about 24.3% of the variance in market value is explained by the model. The Adjusted R-squared is even lower, at 0.1171, indicating that the model does not fit the data particularly well. The overall F-statistic (1.929, p=0.088) is also not significant, meaning that the combined effect of the independent variables is not strong enough to reliably explain the market value of defenders. The model is not significant.

Model Improvements:

It is clear that the model is not satisfactory. I do not even need to see multicollinearity, there is a need for a radical change unlike the model for the attackers. I had to do a lot of studies and modify the model a lot in order for it to be acceptable. First I tried to transform all the variables to capture non-linear relationships. Then I proceeded by eliminating those variables that remained insignificant and adding two variables that I had not taken into account, the goals conceded by the team in which the player plays and the "rating"[48] performed by Whoscored.com, that is a value that includes the evaluation of all the matches played by the player based on defensive statistical data. I also found a good correspondence in transforming the dependent variable (market value) logarithmically.

[48] «WhoScored.com Ratings are based on a unique statistical algorithm, and calculated live during the match. There are over 200 raw statistics included in the calculation of a player or team's Rating, weighted proportionally to their influence in the game. Every relevant event is taken into account, with a positive or negative effect and with different weights on the Rating in relation to the area of the pitch and the success of the action.»

Table 4.8 - Regression output, final defenders model

```
Residuals:
     Min       1Q   Median       3Q      Max 
-1.40213 -0.36493  0.03923  0.39130  1.33892 

Coefficients:
                 Estimate Std. Error t value Pr(>|t|)    
(Intercept)    -3.5667189  4.5996885  -0.775 0.442056    
Age2           -0.0028335  0.0004583  -6.182 1.54e-07 ***
Goals_Conceded -0.0373798  0.0095316  -3.922 0.000291 ***
Rating          1.4274352  0.6272415   2.276 0.027565 *  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.6481 on 46 degrees of freedom
Multiple R-squared:  0.6545,	Adjusted R-squared:  0.6319 
F-statistic: 29.04 on 3 and 46 DF,  p-value: 1.091e-10
```

Residuals:

The residuals provide information about the distribution of model errors, which should be symmetric around zero to confirm that the model is well calibrated. The numbers in Table 4.8 show a slightly skewed distribution of residuals. However, the mean values very close to zero indicate that there is no excessive bias. Nevertheless, it may be useful to check the normality of the residuals with additional tests such as the Shapiro-Wilk test or to display a QQ plot (which we will do later).

Coefficient analysis:

Intercept: The intercept value is -3.5667, but it is not significant (p-value = 0.442), so it does not provide relevant information in the context of the model.

Age2: This variable, which represents age squared, has a coefficient of -0.002835, highly significant (p-value = 1.54e-07). The negative sign indicates that, all other things being equal, market value tends to decline with age at an accelerated rate. This result is consistent with the theory that very old players tend to lose market value, probably due to physical or athletic decline.

Goals Conceded: The coefficient is negative (-0.0374) and highly significant (p-value = 0.000291), indicating that as the number of goals conceded increases, the market value decreases. This result is intuitive, since defenders with a high number of goals conceded tend to be perceived as less effective.

Rating: The coefficient is 1.427, and it is also significant (p-value = 0.027565). This indicates that a higher performance rating (which likely aggregates various defensive metrics and overall impact on the game) is strongly associated with higher market value. It makes sense that defenders who consistently perform well will be valued more highly by clubs and transfer markets.

Model Fit:

- The **Residual Standard Error** is 0.6481, which represents the standard deviation of the residuals. The lower this value is, the closer the model is to the observed data. In this case, the value is moderately low, suggesting that the model has good predictive ability.
- The **Multiple R-squared** is 0.6545, indicating that 65.45% of the variance in market value is explained by the model. The **Adjusted R-squared** of 0.6319 takes

into account the number of variables in the model and shows a slight decrease compared to the R^2. This is a good sign, as the Adjusted R^2 does not decrease dramatically, suggesting that the inclusion of the additional variables actually improves the model without overfitting.

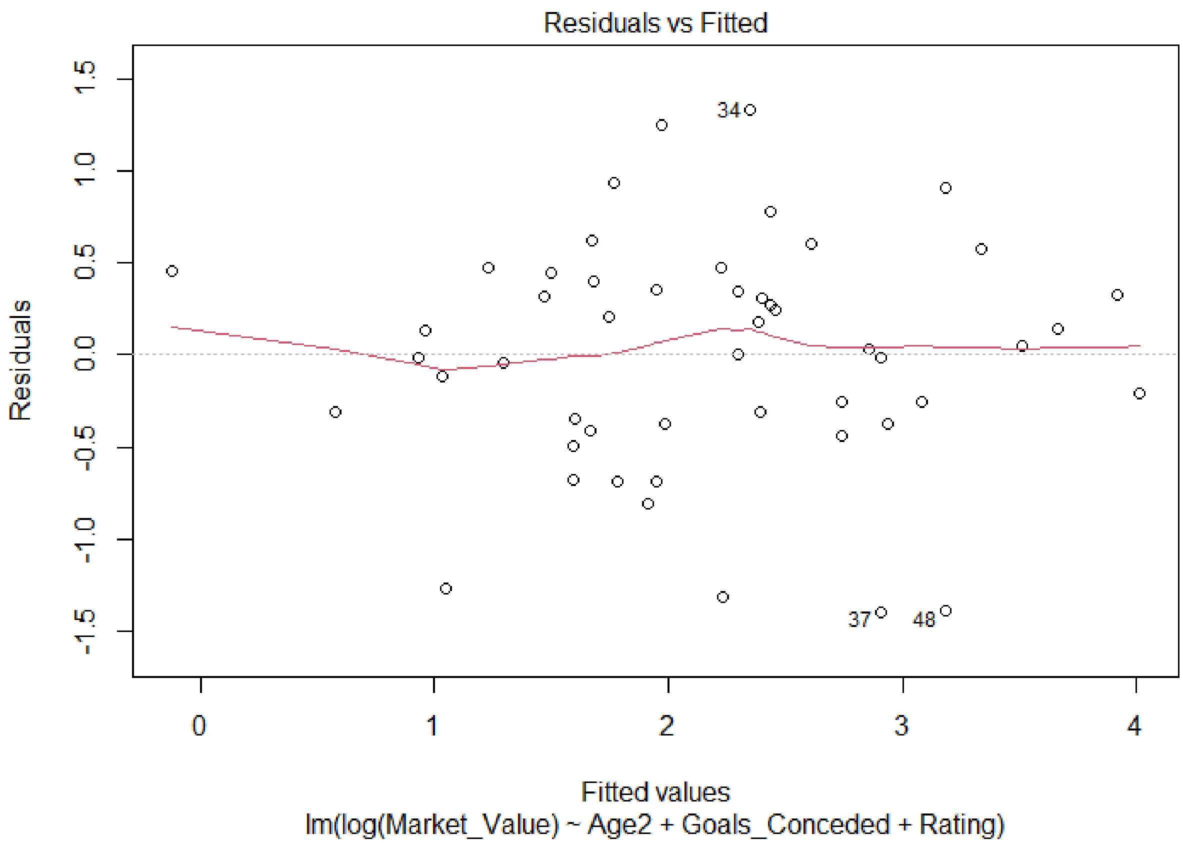
- The **F-statistic** is 29.04 with a very low p-value (1.091e-10), which suggests that the model as a whole is significant, that is, that at least one of the independent variables has a significant effect on the dependent variable (market value).

This improved model is a significant step up from the initial attempt, offering a much better explanation of a defender market value. The non-linear relationship between age and market value, the influence of team defensive performance, and the inclusion of overall rating provide a comprehensive, robust model for market valuation.

As we can see the model now seems to be solid. The changes made help a lot in predicting the market value of a defender. To confirm the robustness of the model, however, as was done for the attackers’ model, some tests need to be carried out.

1. **Residuals vs Fitted Plot**

Graph 4.4 - Residuals vs Fitted plot, final defenders model



As already mentioned, ideally, the residuals are expected to be randomly distributed around the horizontal zero line (the dashed line). This indicator tells us whether the model fits the data well without any noticeable patterns.

In the Graph 4.4, we see a slight curve in the central part, which could suggest the presence of non-linearity. Despite the logarithmic transformation applied to the dependent variable (log(Market_Value)), this curvature could indicate that there are factors that are not fully captured by the current linear model.

Some points distant from the line (observations 34, 37, 48) could represent outliers or influential observations that could distort the model results. In fact, I must add that, in the year taken into consideration, sample 34 (Fikayo Tomori) did not perform as he did in previous seasons, which indicates that according to the parameters observed in this model the player is "overpriced". Instead, samples 37 (Thomas Kristensen) and 48 (Diego Coppola) according to the model's parameters seem to be "underpriced", which suggests them as two good prospective players, as they are very young.

The variance of the residuals appears fairly constant along the predicted values, although there are some observations that deviate more significantly. This suggests that the homoscedasticity assumption may be partially satisfied, although it should be formally verified with tests such as the Breusch-Pagan test.

Breusch-Pagan test

Test statistic (BP): 0.6585

Degrees of freedom (df): 3

p-value: 0.8829

The test yielded a p-value of 0.8829, which is significantly higher than the commonly accepted significance threshold (0.05). Consequently, there is insufficient evidence to reject the null hypothesis of homoscedasticity.

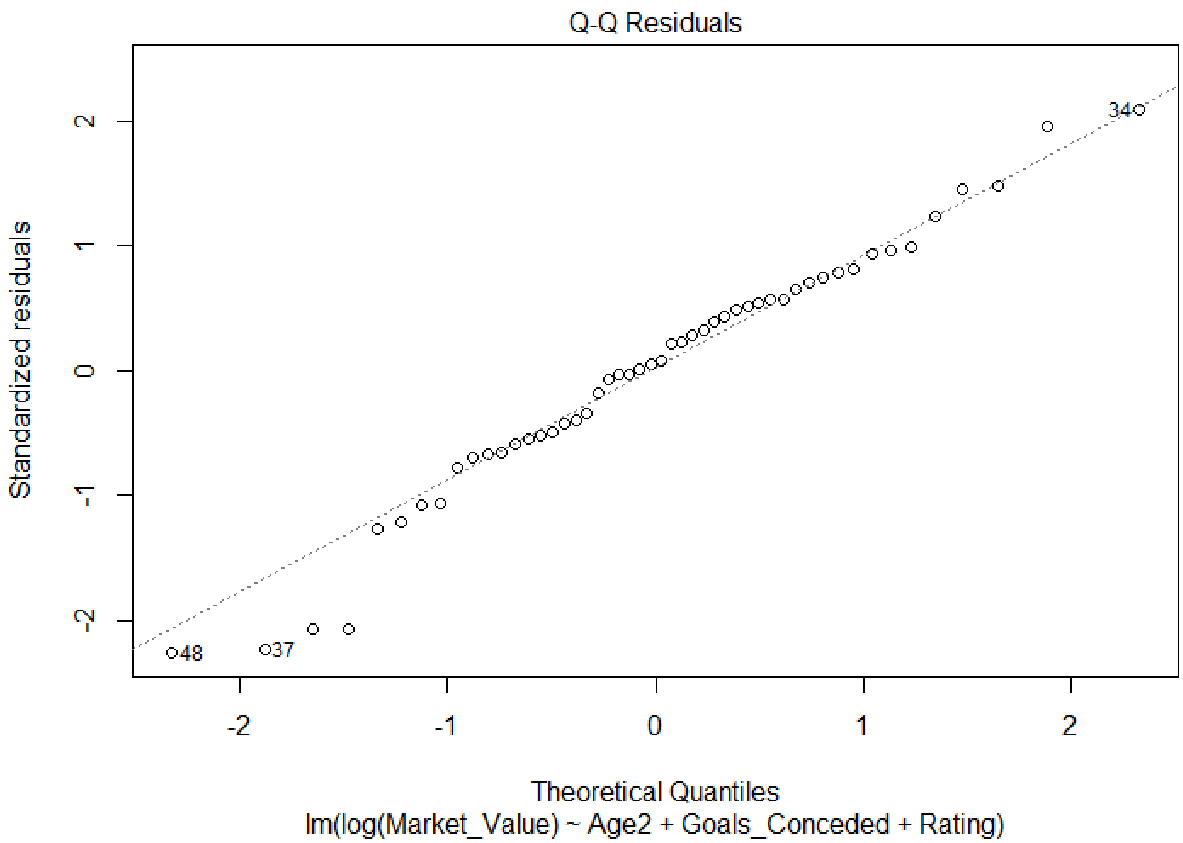
These results indicate that there is no heteroskedasticity in the model, which implies that the variance of the residuals is constant across the different values of the independent variables. This is important because the presence of heteroskedasticity can distort the results of linear regression, making the coefficient estimates inefficient and invalidating statistical inferences. However, in this case, the lack of heteroskedasticity confirms the robustness of the proposed model, which satisfies one of the main requirements of linear regression.

In practical terms, this means that the error distribution remains stable and does not vary significantly as a function of the predicted values. This aspect improves the reliability of the coefficient estimates, which can be used to make valid inferences about the relationships between the variables.

In conclusion, thanks to the positive outcome of the Breusch-Pagan test, it is not necessary to implement additional techniques to handle possible heteroskedasticity issues. Therefore, the model can be considered adequate and the coefficient estimates reliable, thus strengthening the interpretation of the relationships between the variables of interest.

Q-Q Plot

Graph 4.5 - Q-Q plot, final defenders model



Most of the residual points closely follow the dashed diagonal line, indicative of the expected normal distribution of the residuals. This suggests that the model residuals are approximately normally distributed, a key assumption in linear regression models.

However, as already seen in Graph 4.4 there are some labelled observations (such as observations 48, 37, and 34) that deviate slightly from the line, in particular at the extremes of the distribution. However, these outliers do not appear to be particularly influential, and the model overall shows a good adherence to the normality of the residuals.

The plot supports the assumption of normality of the residuals, since most of the points follow the expected normal distribution, despite the presence of some outliers. This strengthens the reliability of the inferences drawn from the regression model, indicating that the assumption of normality is not seriously violated.

In support of the Q-Q plot, however, I still want to perform the Shapiro-Wilk normality test.

Shapiro-Wilk normality test

W: 0.97388

A value of W close to 1 indicates that the data is close to a normal distribution. In this case, the value of W = 0.97388 is very close to 1, suggesting that the residuals are sufficiently normally distributed.

p-value: 0.3305

The p-value = 0.3305 is higher than the commonly used threshold of 0.05, which means that we do not have enough evidence to reject the null hypothesis of normality of the residuals. In other words, we can conclude that there are no significant violations of the normality assumption.

The Shapiro-Wilk test, along with the Q-Q plot shown above, confirms that the residuals of the improved model are normally distributed. This result further supports the validity of the model, since one of the main assumptions of linear regression is the normality of the residuals, which is necessary to ensure the accuracy of statistical inferences.

At this point it is possible to conclude that the model for the market valuation of defenders is valid, it does not have any important anomalies.

The econometric models for attackers and defenders had different responses to the residuals, however both provides valuable insights into the determinants of a player's market value. Although the key variables are different for different roles, we note that the Age variable is considered important regardless of the role under analysis, indicating that clubs are always looking for promising young players who can contribute to large market capital gains over the years.

For attackers, it was clear that performance-based metrics such as goals scored, successful dribbles, and age significantly influence the market value. The positive coefficient for goals suggests that a player's ability to score consistently is a critical driver of market value, as it directly correlates with on-field success and team performance. Conversely, the negative coefficient for age underscores how market values decrease as players grow older, due to shorter career horizons and decreasing physical prowess. Minutes played and age squared further capture the nuance of a player's experience and potential growth, with younger players who play fewer minutes are often seen as assets with untapped potential.

For defenders, the refined model demonstrated the importance of both defensive and overall performance metrics. Variables such as goals conceded and player rating were found to significantly impact market value. The negative coefficient for goals conceded implies that defenders who are part of stronger defensive setups, where fewer goals are conceded, command higher market values. Meanwhile, the overall rating of a defender, which incorporates a variety of performance indicators, also plays a crucial role in valuation.

Through the introduction of age squared, the models captured the non-linear relationship between a player's age and their market value, illustrating how value peaks at a certain point in a player's career before declining as they age further.

The models can certainly be improved through more in-depth and sophisticated analyses, but the aim of this study was to demonstrate how the market value of a player is potentially predictable with very high precision[49]. Unfortunately I do not have the same data (in terms of quality and quantity) and the same means that clubs have, but the result I obtained is satisfactory as I still see room for improvement.

4.5 Model Limitations and Future Directions

4.5.1 Data Quality and External Factors

The precision and dependability of any econometric model are basically related to the quality and completeness of data used. In this work, data used in the forecast of market values for football players is mainly from publicly available statistics concerning performance and physical data. As great a contribution these data sources have been, they maintain intrinsic deficiencies that limit the ability of the model to predict.

[49] As can be seen from the following study, with a more in-depth analysis and with less publicly accessible data, it is possible to achieve over 90% predictive accuracy through a preliminary regression. https://medium.com/@ofirmagdaci/football-insights-from-fifa-data-player-valuation-55b1b748e05d

One serious limitation is the absence of specific qualitative and contextual factors that could fundamentally change a player's market value but for one reason or another they are unavailable or hard to quantify. Examples of such qualities include leadership, mental toughness, and other intangible attributes both on and off the pitch crucial in modern football, but which cannot be easily measured by traditional statistical analysis. Similarly, injuries, tactical adaptability, and the versatility of a player often affect his value but are not reflected in basic performance metrics such as goals, tackles, or minutes played.

Also, the dataset used in this project has basically one-season performance indicators that limit the model to take into consideration long-term trends in player improvement or decline. An enhanced dataset with multi-season data or trajectories of performances can allow modelling continuity in player performance over time. This will provide a better way of managing the influence of temporary ups and downs in form or fitness that distort market value assessments when used on their own.

Besides the issues related to data availability, many external factors influence player market values in ways that are very hard to quantify using statistical models. Economic conditions, such as the financial health of the clubs and leagues, are key to setting transfer fees and salaries. For example, clubs playing in financially healthy leagues are likely to show an inclination towards spending more on transfers, thus raising the valuation of players above what could be expected through performance-based indicators. For example, the changes in the general economic environment-such as exchange rates or financial crises-make the purchasing power of clubs in various countries uneven, hence setting discrepancies in valuations between leagues.

Another external influence is the increasing use of social media and the commercial appeal of a player outside of the sport itself. In modern football, the value of a player's media presence, brand equity, and marketability has risen substantially.

Athletes with significant social media followings or major endorsement deals are valued on the market for more than their on-field contribution. These factors are not easily quantifiable through commonly accessible performance measures. The gap thus creates a divergence between the statistical models and real-life market dynamics.

In order to enhance the resilience of forthcoming models, a more comprehensive and intricate dataset should incorporate qualitative evaluations, including player ratings provided by managers and scouts, records of injuries, as well as indicators of commercial worth. The inclusion of these elements would yield a more complete understanding of a player's contributions and potential, thereby facilitating a closer alignment between the statistical model and the complex dimensions of player market valuations.

Despite these limitations, the model developed here provides a good starting point to begin the process of understanding the principal drivers of the market value of players, particularly in terms of on-field performance. Future studies should focus on enhancing data with the integration of exogenous factors, to further increase the accuracy and applicability of the model to a wide range of situations.

4.5.2 Model Expansion and Advanced Econometric Techniques

While the models developed in this study produce very useful results, significant scope exists for improvement by applying more advanced econometric techniques and incorporating a greater range of variables. The models so far have relied almost exclusively on linear regression, which, although capable of outlining basic relationships, perhaps oversimplifies what is in essence a more complex process of determining player market value. In fact, many of the factors driving player market value are likely to interact with one another in non-linear ways, therefore making the exploration of different modelling approaches worthwhile.

One possible direction for enhancement involves the implementation of nonlinear regression models or the utilisation of machine learning methodologies. Decision trees, random forests, and neural networks are some of the models that might find more complex relationships among variables and, therefore, can identify the interactions that might be difficult to spot in a linear modelling setting.

A nonlinear model could, for instance, better explain diminishing returns associated with particular performance measures. This idea is illustrated by the fact that a player scoring 20 goals, compared to one scoring 15 goals, does not necessarily face a commensurate increase in market value based on factors such as team composition or quality of the opponent team.

Another fruitful direction is the application of panel data models, which are based on both cross-sectional and time-series data. Panel data look at multiple seasons for each athlete, hence, in addition to capturing the fixed effects of individuals-such as a player's unique career path or innate ability-it is able to reflect temporal changes. This would eliminate the issue of short-term variations that affect single-season studies and provide more accuracy in predictive model market values based on performance trends.

In addition, integrating dynamic models that account for feedback loops could enhance the model's predictive accuracy. For instance, the market value of a player increases not necessarily as a result of his performance but because of the scrutiny given to elevated valuations by rich clubs to which the valuation itself becomes a feedback mechanism.

Accounting for such externalities would carry further depth on the mechanics of valuation that in cases of star players, market dynamics can be speculative and increase the transfer value.

The analysis of interaction effects is a direction that deserves more detailed attention. One could, for example, suspect that the age-performance relationship is not captured well by an additive structure: older players who continue performing well may see their value rise due to accumulated experience, leadership skills, and durability, while younger players who perform equally well are valued for their promise.

Similarly, interplays between various performance indicators, such as goals scored and assists, could provide a deeper insight into player value, particularly in hybrid positions where various skill sets are paramount.

Lastly, future studies could explore the role of Bayesian econometrics to introduce prior beliefs or industry-specific knowledge into the estimation process. This approach would allow the model to incorporate expert judgments in player potential or club-specific factors that determine market value over and above what information is contained in the available data. In addition, this would allow a policy of updating estimates of market value as new information becomes available, thus enhancing the model's flexibility and immediate relevance.

With these sophisticated methodologies in place, perhaps future versions of the model could further improve their predictive efficacy and shed deeper insights into the basic processes underlying player valuations, especially within the dynamic football market.

4.5.3 Expansion Across Leagues and Player Roles

The current models are tailored specifically to players in Italy’s Serie A during the 2023-2024 season, but extending the framework to other leagues and competitions is a natural next step. Market dynamics vary considerably across leagues, due to differences in financial power, style of play, and competitiveness. For instance, players in wealthier leagues such as the English Premier League may command higher market values than their counterparts in less affluent leagues, even with comparable performance metrics.

Each league's model would likely need changes to the set of explanatory variables, at the very least. Different leagues might focus on different aspects of the performance of the players; for example, in the Premier League, much more might be made of aerial duels and, indeed, physical contact, while in La Liga and Serie A, technical skills and tactical awareness are still considered more important. Market demand and transfer rules are also different across leagues, with some leagues having stricter financial fair play, while others may show a more frenzied approach to the transfer market depending on club ownership structures or sponsorship deals.

Furthermore, international tournaments, such as the UEFA Champions League or World Cup, have a significant influence on player market values, especially when a player performs well on a global stage. Incorporating data from such competitions could enhance the model’s ability to capture spikes in market value that arise from international exposure.

In addition to expanding geographically, the model can be improved by further differentiating between player positions and roles. The current models focus on attackers and defenders, but midfielders, goalkeepers, and even more specialised roles within these categories (e.g., defensive midfielders, wingbacks) have unique market dynamics that require tailored approaches. Each position has specific performance indicators that influence market value; for example, goalkeepers may see their value driven by save percentages, command of the penalty box, and distribution skills, while midfielders might be judged on passing accuracy, interceptions, and key passes.

Future research might create a truly comprehensive structure of world football player market valuations by further refinement of the model in order to catch the positional differences better and by applying it for various leagues. This would provide the stakeholders-clubs, agents, and analysts alike-with a more versatile and trustworthy tool for determining the value of the player in a wide array of situations.

4.5.4 Practical Applications of the Models

The models developed in this thesis are in an embryonic stage and are not yet suitable for practical application in the real world of football management, however they can be a good starting point for a more in-depth analysis. Their ability to quantify player market value based on objective performance metrics provides a structured approach to understanding how various aspects of a player's game contribute to their overall market valuation. These insights could be particularly useful in several key areas of the football ecosystem, though further development is necessary to fully realise their potential.

First and foremost, these models provide a framework that enhances decision-making greatly in terms of transfers and buying of players. Football clubs usually depend on subjective analysis, previous performances, and financial constraints regarding new signings. A data-driven model like the one below can help a club estimate how much a player may be worth versus how much he is expected to cost, in identifying undervalued talent or in refraining from overpaying for players whose market value has become overly inflated. Traditional scouting obviously continues to play an important role; models such as these supply another layer of consideration, enabling clubs to base decisions on quantifiable measures of performance.

Talent identification and player development are other very serious beneficial applications of these models. Through the analysis of performances and correlation with market value, clubs can take that information and identify younger or more obscure players that reflect many of the same characteristics as high-value players. It would be particularly useful to clubs which have to maximise returns against limited spending capacity because they could identify those talents bound to rise in value over time.

Other areas that could be most facilitated by such models include contract negotiations and the retention strategies of key players. By knowing the drivers of market value, clubs can more properly determine the fair value of contract extensions or wage negotiations, thus aligning player compensation with metrics reflecting performance. This may enhance cost efficiency for clubs, ensuring that they pay market-appropriate wages and decrease the risk of over- or underpaying their players relative to their contributions on the field.

The football valuation models might also serve in the overall valuation of football clubs. Players' value makes up a large part of a club's balance sheet, and an accurate determination of the player market value can directly affect the overall financial valuation of a club. In this respect, the models may enable more accurate financial valuations, especially during transfer-active periods when changes in the value of players may substantially alter the club's financial position.

Again, it is essential to note that even though these models present very encouraging prospects, they yet are not mature enough for wide application. The existing limitations of having a narrow range of variables and the peculiar data used mean that they should be considered as complementary tools and not as standalone solutions. The football market is inherently complex and determined by a variety of factors, most of which are not easily quantifiable. These include the chemistry a player brings into the team, his leadership qualities, or even market forces driven by external economic factors.

Nevertheless, with further refinement, broader data inputs, and ongoing validation against real-world scenarios, these econometric models have the potential to become integral to football management, offering valuable, data-driven insights that support more informed and efficient decision-making processes across the industry.

CONCLUSION

Having reached the end of this thesis, I hope to have led the reader to understand the current situation of the European football scene. Starting from how we got to this point, passing through various points of reflection and possible breaks in vicious circles, ending with a practical demonstration of how modern technologies can influence decisions by clubs.

Technology develops exponentially and it is impossible to predict how far technological progress will go in football. What is certain is that if you want to keep a sport as such, there is a need to implement regulations as soon as possible. I have taken football into consideration as it is the sport where there is the most money in circulation and consequently where the competition is highest. Personally, as a fan of this sport, I would never want the competition to move off the pitch, in a fight for who has the most advanced technologies and the most prepared research teams, but for it to remain a fight for who has the most talent and who puts in the most effort on the pitch.

Bibliography/Webliography

- Araujo, Duarte & Couceiro, Micael & Seifert, Ludovic & Sarmento, Hugo & Davids, Keith. (2021). Artificial Intelligence in Sport Performance Analysis

- Borghini, A., & Baldini, A. (2011). On the logic of soccer patronage. Soccer & Society

- Bormida, M.D. (2021), "The Big Data World: Benefits, Threats and Ethical Challenges", Iphofen, R. and O'Mathúna, D. (Ed.) Ethical Issues in Covert, Security and Surveillance Research (Advances in Research Ethics and Integrity, Vol. 8), Emerald Publishing Limited

- Bummer, Shaun, "History on the Pitch: The Social and Economic Impact of Professional Soccer in Postwar London" (2015)

- Busse Ronald and Damiano Jean-Pierre, “The Role of Commercialisation of the European Football Business for the Emotional Bond between Fans and Clubs”

- Capology - https://www.capology.com/it/serie-a/monte-ingaggi/2023-2024/c

- Chunyang Huang and Shaoliang Zhang, “Explainable artificial intelligence model for identifying Market Value in Professional Soccer Players”, https://ar5iv.labs.arxiv.org/html/2401.16795

- Ciavarella Francesco (2022), “Dalla Moneyball alla Football Analytics: la statistica scende in campo”, Ghigliottina

- Clarke (2021), “The Moneyball Method: Using Data to Build a Football Dream Team (On a Budget)”, https://www.graphext.com/post/the-moneyball-method-using-data-to-build-a-football-dream-team-on-a-budget

- Corriere dello sport - https://www.corrieredellosport.it/news/calcio/serie-a/inter/2024/05/23-127916849/inter_a_oaktree_la_previsione_sui_conti_debiti_e_40-50_milioni_di_perdite

- Datacamp - https://www.datacamp.com/blog/ai-in-sports-use-cases

- Economia e Sport - https://www.economiaesport.it/2024/07/03/serie-a-la-ripartizione-dei-diritti-tv-ai-club-oltre-1-miliardo/

- European Commission - https://ec.europa.eu/assets/eac/sport/library/documents/cons-study-transfers-final-rpt.pdf

- European Leagues - https://europeanleagues.com/wp-content/uploads/REPORT-THE-FINANCIAL-LANDSCAPE-OF-EUROPEAN-FOOTBALL.pdf

- Football Benchmark - https://www.footballbenchmark.com/data_analytics/starter/club_finance

- Football Benchmark - https://www.footballbenchmark.com/library/commercial_deal_landscape_in_the_big_five_leagues

- Grand View Research - https://www.grandviewresearch.com/press-release/global-sports-analytics-market#:~:text=Sports%20Analytics%20Market%20Growth%20%26%20Trends,by%20Grand%20View%20Research%2C%20Inc.

- Gravina G. (2011). Il bilancio d'esercizio e l'analisi della performance nelle società di calcio professionistiche. Milano: FrancoAngeli, 40.

- Guida, Caniato, Moretto, Ronchi (2023), "Artificial intelligence for supplier scouting: an information processing theory approach"

- HYPE - https://www.hypesportsinnovation.com/how-rbfa-scored-big-with-ai-driven-ticket-sales/

- Jack Bantock (2024), "Top soccer clubs are using an AI-powered app to scout future stars", CNN

- Jason Stockwood (2024), "Time for a salary cap to keep leagues competitive and reduce agents' influence", The Guardian

- Jim Totime (2023), "The Data-Driven Football Club: Strategies for Success"

- Liam Callaghan Doyle (2022), “Is the Moneyball approach a feasible approach for success in Football (Soccer)? An analysis of Brentford’s adoption of a statistical approach.”

- Lichtenthaler, U. (2021), "Mixing data analytics with intuition: Liverpool Football Club scores with integrated intelligence", Journal of Business Strategy

- Luca Di Simone and Davide Zanardi (2020), “On the relationship between sport and financial performances: an empirical investigation”

- Lucas Oaigen (2024), “How Artificial Intelligence Can Revolutionize Talent Management in Youth Soccer”, Linkedin

- Magdaci (2023), “Football Insights from FIFA Data: Player Valuation”, Medium.com

- Maraschi Andrea (2020), “INDICI STATISTICI PER LA MISURAZIONE DELLE PERFORMANCE NEL GIOCO DEL CALCIO: EXPECTED GOALS ED EXPECTED POINTS”

- Marco Iaria (2024), La Gazzetta dello Sport, https://www.gazzetta.it/Calcio/Serie-A/03-07-2024/diritti-tv-serie-a-inter-101-milioni-milan-e-juventus-87.shtml

- Michael Lewis (2003), "Moneyball: The Art of Winning an Unfair Game"

- Neri, L., Russo, A., Di Domizio, M., & Rossi, G. (2021). “Football players and asset manipulation: the management of football transfers in Italian Serie A.” European Sport Management Quarterly

- OneFootball - https://onefootball.com/en/news/cf-how-much-milan-and-other-serie-a-clubs-earned-from-tv-rights-in-2022-23-37672831

- OOTB - https://outsideoftheboot.com/2016/07/20/rise-of-data-analytics-in-football-2/

- Pavitt J, Braines D, Tomsett R. “Cognitive analysis in sports: Supporting match analysis and scouting through artificial intelligence”. Applied AI Letters. 2021

- Pilar Malagón-Selma (2023), “Machine Learning and Multivariate Statistical tools for Football Analytics”, Universitat Politecnica de Valencia

- Poli, Besson, Ravenel (2022), “Econometric Approach to Assessing the Transfer Fees and Values of Professional Football Players.”, https://doi.org/10.3390/economies10010004

- PSFA - https://thepfsa.co.uk/the-power-of-analytics-dissecting-the-role-of-data-in-football-strategies-and-performance/

- Sabharwal, R., Miah, S.J. (2021) “A new theoretical understanding of big data analytics capabilities in organisations: a thematic analysis”. J Big Data

- Soccer Take - https://www.soccertake.com/performance/soccer-analytics-how-data-is-changing-the-game

- Sporting News - https://www.sportingnews.com/us/nba/news/what-nba-luxury-tax-explained-penalties-high-spending-teams/p1sspaedfsmsit20rqewn9d7

- Stephen Dobson and John Goddard (2011), “The economics of football”, Cambridge University Press

- Stewart, Mitchell, Stavros (2007), “Moneyball Applied: Econometrics and the Identification and Recruitment of Elite Australian Footballers”, International Journal of Sport Finance

- Sulimov, “Performance Insights-based AI-driven Football Transfer Fee Prediction”, https://ar5iv.labs.arxiv.org/html/2401.16795

- Thadeu Miranda Gasparetto (2012), “Relationship between Wages and Sports Performance”, Federal University of Juiz de Fora, Minas Gerais, Brazil

- Thibaud Trichard (2021), “What are the main determinants of professional soccer players wages? An Examination of the top 5 European Leagues”, Adam Smith Business School

- Transfermarkt - https://www.transfermarkt.it

- Weimar, Daniel & Wicker, Pamela (2017), “Moneyball Revisited: Effort and Team Performance in Professional Soccer.”, Journal of Sports Economics

- WhoScored - https://it.whoscored.com/

- Wikipedia - https://en.wikipedia.org/wiki/Biometrics

- Wikipedia - https://en.wikipedia.org/wiki/FIFA
-
- Wikipedia - https://en.wikipedia.org/wiki/UEFA
-
- Zoppello Marco (2004), “Analisi e modelli di correlazione tra risultati sportivi ed economici: La serie B”