OPTIMIZING FANTASY TEAM SELECTION: AN INTEGER LINEAR PROGRAMMING APPROACH

Abhinav Venkatraman (av430)

Kent Business School

Word count: 8000 words

Acknowledgements

I extend my sincere appreciation to my Supervisor, Dr. Tri-Dung Nguyen, for his valuable guidance, consistent support, and constructive feedback that have greatly influenced the progress of this project. His expertise and mentorship have played a pivotal role in shaping the direction and outcome of this endeavor, and I am genuinely grateful for his contributions, which have been instrumental in achieving the successful completion of this undertaking.

Table of Contents

List of Figures

List of Tables

Abstract

A fantasy sport league is an online game where participants assume the role of team managers, thoughtfully selecting real players to form their virtual teams. The allocation of points to each team is based on the actual performance of these selected players in real-life sports events. The primary goal for every manager is to conclude the season with the highest points possible, aiming for victory. Among the largest leagues globally is the Fantasy Premier League (FPL), which boasts an impressive participation of approximately 7 million players from around the world.

The characteristics and rules of the FPL are thoroughly discussed, delving into the meticulous process of player choice, scoring systems, and budget constraints that heavily influence the strategic decisions made by team managers. Building upon these valuable insights, a mixed-integer linear programming model is introduced, specifically designed to optimize team selection each week with the goal of achieving the maximum points possible. The model's construction involves the use of retrospective data from a season, allowing for precise mathematical determination of the best scoring team for each week.

To demonstrate the model's practicality and effectiveness, data from the 2022/23 FPL season is utilized, and the model is executed on IBM ILOG CPLEX. The results obtained are thoroughly analysed and discussed, presenting a clear showcase of the model's potential in real-world scenarios.

1. Introduction

Fantasy sports, originating in the United States during the 1980s, have evolved into a captivating online phenomenon, revolutionizing sports engagement on a global scale. The concept of fantasy sports can be traced back to the ingenious idea of journalists Glen Waggoner and Daniel Okrent, who devised a game where a small group of participants would draft real players from a pool of active baseball athletes **(Davis & Duncan, 2006).** This innovative approach quickly gained popularity and laid the foundation for a new era of sports entertainment. As the Internet became more accessible and technology advanced, fantasy sports transitioned from a niche activity to a mainstream sensation. The integration of real-time statistics and digital platforms enabled millions of sports enthusiasts to actively take part in these virtual leagues. Today, fantasy sports have grown exponentially, captivating the hearts of tens of millions of people worldwide and encompassing a wide array of sports beyond baseball, including football, cricket, basketball, and more.

In these fantasy sports games, participants immerse themselves in the roles of team managers, meticulously selecting real-life athletes to form their dream teams. The athletes' on-field performances in real-life matches directly translate into points for their respective fantasy teams, making every game and player's performance crucial to the overall success of the fantasy managers. The thrill of competition, strategic decision-making, and the desire to outwit opponents fuel the fervour and excitement surrounding these virtual leagues. This profound connection with the sport has prompted many enthusiasts to derive their sports entertainment primarily through fantasy sports, influencing sports consumption behaviours significantly **(Nesbit & King, 2011).**

Among the multitude of fantasy sports leagues, Fantasy Premier League (FPL) stands out as the largest and most prominent, boasting an extensive following of approximately 7 million participants worldwide. This fantasy league operates in tandem with the top-tier of English football, the English Premier League. FPL commands unparalleled attention and fervent engagement from football enthusiasts, transcending geographical boundaries.

This dissertation seeks to conduct a comprehensive analysis and optimization of team selection in the context of fantasy football, with a specific emphasis on the Fantasy Premier League (FPL). The strategic decisions made by managers in assembling their fantasy teams are significantly shaped by various pivotal factors. These factors encompass guidelines for player selection, budgetary limitations that impose constraints on the total team cost, and the intricacies of scoring systems that attribute points based on players' performances, goals, assists, clean sheets, and other relevant criteria. The presence of such constraints poses intriguing challenges for FPL managers, compelling them to devise the most efficient and high-scoring teams on a week-to-week basis.

To achieve this goal, a mixed-integer linear programming model is presented that enables us to formulate an innovative approach for generating and analysing retrospective results for any given FPL season. Utilizing retrospective data from past seasons, the model allows us to tackle the following problems –

- What was the best-performing teams for each game week of the season?
- What was the maximum possible score that was theoretically possible at the end of the season?

Given the significant impact of fantasy sports on sports consumption and economic value, this study aims to shed light on the implications and challenges faced by fantasy team managers. By offering rigorous analysis and data-driven insights, the aim is to provide managerial recommendations to FPL managers that demonstrate the decisions they could have made to achieve the highest attainable score and enhance the overall FPL experience.

The subsequent chapters of this dissertation offer a thorough investigation of Fantasy Premier League (FPL) and the optimization approach employed. Chapter 2 conducts an extensive literature review, providing valuable insights into the existing body of work within this domain. In Chapter 3, the FPL characteristics are outlined, followed by an exploration of the data collection process. Additionally, a comprehensive account of the development of the mixed-integer programming model is presented. Moving forward, Chapter 4 unveils the findings and analysis of the optimal team selections, utilizing data from the 2022/23 FPL season. The results shed light on the efficacy of the optimization approach in team selection strategies. Lastly, Chapter 5 serves as the conclusion of the dissertation, summarizing the primary conclusions drawn from the research and emphasizing the relevance and significance of this study within the realm of fantasy sports and team selection optimization.

2. Literature Review

The research on team optimization in fantasy sports has gained significant attention in recent years. Various papers have explored mathematical models and mixed integer programming methods to enhance decision-making for team managers in different sports leagues. To provide a comprehensive understanding of the work done in this field, the review follows a funnel approach, starting with papers on the broader application of MILP models, gradually narrowing down to our case in hand.

2.1. Generic Models and Broader Applicability

Hunter, Vielma & Zaman (2016) analyse the general applicability of their model to various real-life situations. They analyse the problem of selecting an entry portfolio for competitions with top-heavy pay-out structures. The approach is generic in nature and can be applied to a plethora of problems faced in day-to-day business situations, such as venture capital firms identifying sound start-up companies or movie studios choosing movies to produce. However, to assess the model, the authors specifically focus on the daily fantasy sports contests, like hockey and baseball, wherein players enter line-ups of players into daily contests, and the top-scoring entry wins a significant amount of money. This presents a challenging optimization problem as multiple factors need to be considered when selecting a line-up, such as the performance history of individual players, the salary cap, and the need to diversify the line-up to reduce risk.

To address the problem, the authors employed a MIP model that finds the optimal set of decisions under the assumption that all data are available in advance. The method relies on a sequential greedy algorithm that creates entries for top-heavy contests with limited resources, where at each step, an integer programming problem is solved. The effectiveness of this approach is demonstrated by consistently good performance in daily fantasy sports contests across multiple sports, including hockey and baseball. The authors use the JuMP algebraic modelling language written in the Julia programming language to construct all formulations.

The paper also highlights that the application of integer programming in daily fantasy sports has gained significant success, which has brought about increased recognition and interest in analytics and optimization in this widely enjoyed pastime. It demonstrates the applicability of mathematical optimization techniques in solving real-world problems and how it can provide an edge to those who can effectively apply it.

Beliën, Goossens, & Van Reeth (2017), present a generic model with applicability restricted to various fantasy sports scenarios. Their research introduces an optimization model that analyses different fantasy sport games using mixed-integer programming (MIP). The approach assumes the availability of all relevant information about the sport before the team selection process.

They found that the optimal team from the previous season can be utilized as a promotional tool for the upcoming season, highlighting the discrepancy between the performance of the best participant and that of the optimal team. This comparison has the potential to attract more participants and increase revenue for organizers. Moreover, their model surpasses the performance of the best participant's strategy, offering enhanced team selection and player transfer capabilities.

To achieve this performance, they propose a relax-and-fix heuristic method, which finds application across a wide range of fantasy sports. To verify their approach, they test their model on Gigabike, a popular professional road cycling game, using IBM Ilog Cplex 12.6 to solve several editions of the game. The analysis yields multiple results, including valuable information for both organizers and participants, enabling the creation of new sub-competitions, strategic advice for participants, and the assessment of the impact of various game rules. They emphasize the efficacy of their model for both organizers and participants, as it generates informational and commercial value. They contend that their system can handle typical features of fantasy sports and can be employed in various game settings. Additionally, they provide a survey of fantasy sport games and enumerate shared characteristics related to the game regulations.

2.2. Fantasy Cricket Optimization Models

A series of three papers on the sport of cricket presents their findings in a chronological manner. **Sharp, Brettenny, Gonsalves, Lourens, and Stretch (2011)** developed an integer programming model based on binary response decision variables for team selection of an 11-person cricket team. Optimized through a linear function, the model is designed on predefined criteria and implemented using Excel Solver. To illustrate its efficacy, the researchers utilize data from the 2007 inaugural Twenty20 World Cup to select the ideal 11-person team for matches. Their proposed technique demonstrates greater versatility compared to conventional ranking methods, providing a means of forming a team that can be extended to a multi-stage team selection game. This approach holds relevance for fantasy league games, where teams need to be updated regularly as new information becomes available. In summary, this research article contributes valuable methods to the body of literature on cricket performance evaluations, combining mathematical optimization with cricket team selection.

Building upon their previous work, **Brettenny, Friskin, Gonsalves, J.W., and Sharp (2012)** explore a multi-phase integer programming technique for creating fantasy teams in Twenty20 cricket tournaments. Their aim is to select an 11-player team within a budget constraint of 1 million units, adhering to the formation restrictions of the game modified over four stages of the tournament to maximize the total number of points scored by the players. To address this problem, they employ a mathematical optimization procedure that utilizes performance measures to select and modify the team at each stage of the tournament. These performance measures are updated based on the career statistics of each cricketer as new data becomes available. Specifically, the model considers data from the most recent stage of the tournament, allowing optimization algorithms to select cricketers with exemplary performance. The authors' mathematical algorithm for sequential selection provides a solution to the fantasy team selection problem, operating in real-time and replicating the fantasy league environment. The effectiveness of this approach is demonstrated through computer simulations.

In summary, this paper presents a novel and effective approach to selecting fantasy cricket teams, accounting for the dynamic nature of the tournament and the evolving performances of players. The study constitutes a valuable contribution to sports analytics, showcasing the application of mathematical optimization in solving real-world problems in sports management.

Chand et al. (2018) explored a method for selecting the most optimal cricket teams in the Indian Premier League using a multi-objective approach with integer linear programming. Utilizing actual player data from the 2011 season, the authors consider several conflicting objectives, such as the batting average, bowling average, and total cost of the team. Furthermore, they discuss a method for identifying the most valuable player on the team to provide non-contractual incentives. The proposed approach guarantees optimality and produces results in a feasible amount of time. The paper demonstrates its scalability by presenting examples with five objectives. In conclusion, this approach serves as a valuable tool for team managers and owners to search for optimal teams based on their specific preferences.

2.3. Fantasy Football-Specific Optimization Models

The following papers pertain to the sport of American football, which shares similarities with football (soccer).

Becker, Xu, and Andy Sun (2016), focus on managing fantasy football teams through draft construction and weekly management. They introduce an analytical approach utilizing a Mixed-Integer Programming (MIP) model, which optimizes draft selection and weekly line-up management based on player performance predictions. The MIP formulation incorporates a comprehensive objective of winning the entire season, considering the uncertainty in opponent owners' drafting behaviour, modelled through robust optimization. They present a prediction methodology that leverages historical data and expert opinions to estimate player and team performance. Their model stands out due to its inclusion of the concept of innate abilities in different statistics of players or teams, forming the basis for making accurate predictions.

To evaluate the effectiveness of their proposed approach, they train the model using data from the 2004-2006 seasons and simulate the 2007-08 seasons. The results demonstrate that their model outperforms conventional strategies that rely on publicly available expert rankings. Notably, using the mixed-integer programming (MIP) drafting approach significantly increases the probability of winning money compared to the baseline approach, with a confidence level of 99.75%. They highlight that the unique structure of their model enables efficient computation, crucial for an online environment. Furthermore, they emphasize the broader implications of their approach for decision-making in various fields, such as finance and supply chain management, where optimization models can improve overall performance.

To conclude the funnel approach**, Edwards (2021)** delves into the use of Mixed-Integer Programming (MIP) models in football, specifically focusing on the AFL Supercoach league in Australia, which boasts a significant participant base of nearly 200,000 individuals. The study addresses three distinct problems encountered during the 2018 season of the competition.

The first problem centres around determining all the decisions a coach should have made to achieve the maximum possible score. The second problem involves identifying the minimum starting budget required to secure victory in the league. Lastly, the third problem seeks to determine the possibility of a team, created at the beginning of the competition, and subsequently neglected, still winning the competition. The primary goal of this paper is to showcase the discrepancy between the actual results achieved by real players in the AFL Supercoach competition and the maximum achievable score under theoretical optimal conditions. Due to the well-defined and straightforward objective function of the competition, it serves as an ideal candidate for MIP modelling and analysis. The study addresses three specific decision-making processes: maximizing the score, determining the minimum starting budget for victory, and investigating the potential success of neglected teams.

In their analysis, the researchers employ Gurobi 7.5.0 as the implementation tool for their models. Their study showcases a distinctive application of Mixed-Integer Programming (MIP) in the analysis of fantasy sports competitions, which have experienced a surge in popularity in recent years. These competitions present an exceptional opportunity to showcase the capabilities of MIP in addressing diverse questions of public interest. The author also proposes a natural extension of their work, with the aim of expanding the model to encompass the primary fantasy competition of the English Premier League, commonly referred to as the Fantasy Premier League (FPL).

In conclusion, the field of research on optimization models for fantasy sports, particularly in the context of football (soccer), remains limited with few existing studies. This work aims to contribute to this line of research by presenting a Mixed integer Linear Programming model specifically designed for the Fantasy Premier League (FPL) and extend the literature in this field.

3. Research methodology

This section is split into 3 sections. The first section details the problem description, highlighting the characteristics of the FPL. The second section explores the data collection and cleaning process. Lastly, the third section discusses the assumptions, simplifications and introduces the MILP model.

3.1. Problem Description

3.1.1. Player Budget and Team Composition

At the beginning of each season, every manager is given a fixed virtual budget of 100 million pounds to build their fantasy team. This budget serves as a financial constraint, making managers carefully select players for their squads.

Each player's price is determined beforehand, considering factors like past performance, popularity among managers, and expected contributions in the upcoming season. Players expected to perform exceptionally well are given a higher price. However, this price can change during the season based on player performance and the frequency of transfers made by managers. Premier League matches unfold every weekend, providing managers the opportunity to execute transfers during this period. Each of these weekends, encompassing a span of 38 consecutive weeks, is recognized as a distinct "game week". Each game week, if the ratio of transfer ins to transfer outs favours transfer ins for a player, his price will increase for the upcoming game week, and conversely, if the ratio tilts in favour of transfer outs, his price will decrease. As a result, managers often buy players at lower prices and sell them when their value rises, allowing them to improve their team budget for more acquisitions.

To add more strategy to team composition, the FPL restricts the number of players managers can select from a single Premier League club. This rule promotes diversity and prevents managers from having squads dominated by players from a particular high-performing team.

With these constraints in place, managers are required to assemble a team of 15 players as shown in image 1, comprising of:

- 2 goalkeepers
- 5 defenders
- 5 midfielders
- 3 Forwards

Figure 1: Squad Selection Process

3.1.2. Squad Management and Captain Selection

Managers are obliged to pick 11 players from the 15-member squad as the starting XI, whose performance will determine the points earned for the game week. The formation of the starting XI can be tailored to the manager's preference. However, it is mandatory to include a minimum of 1 goalkeeper, 3 defenders, and 1 forward in the selection.

Furthermore, on a weekly basis, a captain and a vice-captain must be chosen from the starting XI. In the typical practice, managers strategically designate a player whom they anticipate will deliver exceptional performance as the captain. This role holds added significance, as the captain's points are doubled, exerting a substantial influence on the team's cumulative score. If the chosen captain does not participate in the game week, the vice-captain assumes the role of captain, and their points are doubled instead. To make this selection, managers need to conduct a careful analysis of upcoming fixtures and the form of the players.

3.1.3. Transfers

Throughout the season, managers have the option to make one free transfer per game week to strengthen their teams. In case they choose not to utilize this transfer, it gets saved and can be carried over to the next game week, providing them with the opportunity to make two free transfers. However, only one saved transfer is allowed at a time.

For any additional transfers made beyond the allocated limit, managers will incur a deduction of 4 points for each additional transfer made. This rule forces managers to carefully consider the advantages of squad improvements against potential point losses, ensuring strategic decision-making throughout the season.

Moreover, if a player from the starting XI takes part in a game for a duration of less than one minute, an automatic substitution occurs. The player is replaced by the corresponding bench player of the same position who has garnered the highest points.

3.1.4. Chips

Managers in the Fantasy Premier League have access to special chips that can be strategically used to boost their team's performance. However, each chip can only be used once during the entire season, and only one chip can be utilized in any given game week. The available chips are as follows:

1. **Bench Boost:** The substitute players' points are included in the total points for that game week.
2. **Free Hit:** Unlimited transfers are allowed for a single game week and the new team formed in only valid for that game week.
3. **Triple Captain:** The captain's score is tripled instead of doubled.
4. **Wildcard:** All transfers that are made over a game week are free of point deductions and the team does not reset.

3.1.5. Scoring System

Points are awarded to players based on their real-life performance. The scoring system assigns points for various performance metrics, which contribute to a player's overall score. The image below shows the points awarded for different performance metrics –

Action	Points
For playing up to 60 minutes	1
For playing 60 minutes or more (excluding stoppage time)	2
For each goal scored by a goalkeeper or defender	6
For each goal scored by a midfielder	5
For each goal scored by a forward	4
For each goal assist	3
For a clean sheet by a goalkeeper or defender	4
For a clean sheet by a midfielder	1
For every 3 shot saves by a goalkeeper	1
For each penalty save	5
For each penalty miss	-2
Bonus points for the best players in a match	1-3
For every 2 goals conceded by a goalkeeper or defender	-1
For each yellow card	-1
For each red card	-3
For each own goal	-2

Figure 2: Scoring System

The final goal for each manager is to end the season with as many points as possible.

3.2. Data collection and Pre-processing

The FPL statistics are accessible to the public for playing the game and hence obtaining the necessary data for research purposes was straightforward. The data used in this study was obtained from two primary sources. Firstly, a CSV File containing relevant information was sourced from a GitHub user Anand, V. (2022) who had shared it on the platform. GitHub, a web-based platform widely utilized by developers for collaborative software projects, has significantly influenced the software development landscape. The projects and data shared on GitHub are open source, allowing free access, modification, and distribution to the public.

Additionally, several data points were collected from the official Fantasy Premier League site, which openly provides comprehensive data for public usage. Once the data was acquired the following steps were taken to process the data –

1. **Data Acquisition and Pre-processing:**
   The dataset, comprising 37 game weeks, was acquired from the site. As game week 7 had no data, the analysis was performed using data from 37 game weeks. In each game week's table, player names were accompanied by various metrics, including minutes played, price, points scored, red cards, goals scored, position and more. Subsequently, all 37 tables were uploaded into PostgreSQL for further processing. The following steps were undertaken to refine and achieve the desired data set.

2. **Handling Duplicate Entries:**
   To ensure a streamlined dataset, it was crucial to eliminate duplicate entries. Each game week's table was thoroughly examined, and duplicates were removed using a DELETE statement with a GROUP BY clause based on player name, position, and team.

3. **Data Consolidation:**
   To prepare the dataset for modeling, a LEFT JOIN operation was conducted, consolidating all players' information from the first game week to the last. Points, price, team, position were extracted for each game week.

4. **Null Value Handling:**
   A considerable number of cells in the "Points" column were found to be NULL, indicating that certain players did not participate in specific game weeks. To address this, NULL values were replaced with 0 to accurately represent zero points for non-participating players.

5. **Price Imputation:**
   Instances were observed where the "Price" column contained values of 0. To rectify this, the immediate previous game week's price was inputted to eliminate any zero values in the dataset.

6. **Price Format Standardization:**
   The "Price" column initially contained values that were multiplied by 10, for instance, 45 instead of the actual value 4.5. To ensure the correct representation, all price values were divided by 10 using pandas in Google Collab, standardizing the format to floating-point numbers. This was necessary to accurately reflect the total budget as 100 million, avoiding the erroneous representation of 1000 million that would have resulted from using the original integer values.

Following the completion of these data cleaning steps, the resultant dataset comprised 573 unique players, each associated with their respective points, price, team, position for all 37 game weeks.

Utilizing the dataset, three distinct sheets were generated in the CSV file, containing information on points, price, and position of each player. This was done for ease of data readability for the model.

3.3. MILP Model

Before the Model is presented, let us look at assumptions and simplifications made which significantly simplify the problem –

3.3.1 Assumptions

1. **Transfers:** It is assumed that managers are allowed only 1 transfer per week, and any unused transfers do not carry over to the next week, resulting in a total of 37 transfers over the season.

2. **Formation:** To streamline the team composition process, a 4-4-2 formation is adopted as the only formation considered in the model. This formation, widely used by managers, ensures an equal distribution of players among different positions (4 defenders, 4 midfielders, 2 forwards). This standardization allows for consistent comparisons between weekly teams.

3.3.2. Simplifications

1. **Chips:** The model does not incorporate chips like wildcards, free hit, bench boost, and triple captain. While these chips can significantly impact a team's performance in the actual FPL game, their exclusion in the mathematical model reduces unnecessary complexities.

2. **Automatic Substitution Exclusion:** The automatic substitution in the case of a starting player playing 0 minutes is omitted as an optimal approach is to select the best scoring players for each position. Instead, it focuses on selecting a 15-member squad and an 11-member starting lineup each week.

3. **Vice-Captaincy:** The automatic substitution process in the case of a starting player playing 0 minutes is excluded, as the optimal approach is to select the best-scoring players for each position. Instead, the focus is on assembling a 15-member squad and an 11-member starting lineup for each week.

3.3.3. Optimal Team Selection Model

The following notations, sets, and parameters are used in the model. Table 1.1 presents a description of all the sets, while Table 1.2 outlines the parameters along with their notations and values.

Indices	Sets	Details
p	*P*	Set of all the players
	P_{GK}	Set of players who are Goalkeepers.
	P_{DEF}	Set of players who are Defenders.
	P_{MID}	Set of players who are Midfielders.
	P_{FWD}	Set of players who are Forwards.
q	*Q*	Set of playable positions Q = {"GK", "DEF", "MID", "FWD"}
r	*R*	Set of all game weeks i.e., rounds

Table 1: Definition of sets

The index serves to access individual elements within their respective sets. Each of these sets were accessed from the excel file using the sheet read function.

Parameters	Details
$points_{p,r}$	points achieved by player $p \in P$ in game week $r \in R$
$price_{p,r}$	Price of player $p \in P$ in game week $r \in R$
B	Starting budget in millions B = 100.0
$Qmax_q$	Number of vacancies of each position $q \in Q$ $\mathrm{Q_{max_q}} = [2, 5, 5, 3]$
$Qpoints_q$	Number of players of each position $q \in Q$ allowed to score points. $\mathrm{Q_{points_q}} = [1, 4, 4, 2]$
X_r	Number of players whose points count towards team points in each game week $r \in R$. $\mathrm{X_r}$ = 11
T^{total}	Total transfers over complete season $\mathrm{T^{total}} = 37$
T_r	Transfers per game week $r \in R$ $\mathrm{T_r} = 1$

Table 2: Definition of parameters

The decision variables are defined as follows -

$x_{p,r}$ = 1 if player $p \in P$ is in the squad for game week $r \in R$; 0 otherwise

$y_{p,r}$ = 1 if points of player $p \in P$ is included for game week $r \in R$; 0 otherwise

$C_{p,r}$ = 1 if player $p \in P$ is the captain for game week $r \in R$; 0 otherwise

${t^{in}}_{p,r}$ = 1 If player $p \in P$ is transferred into the team for game week $r \in R$: ($r > 1$); 0 otherwise

${t^{out}}_{p,r}$ = 1 If player $p \in P$ is transferred out of the team for game week $r \in R$: ($r > 1$); 0 otherwise

b_r = Remaining budget at game week $r \in R$

Finally, we introduce the MILP model which can be used to find the most optimal team each week:

maximize:

$$\sum_{\boldsymbol{p} \in \boldsymbol{P}} \sum_{\boldsymbol{r} \in \boldsymbol{R}} \boldsymbol{points}_{\boldsymbol{p},\boldsymbol{r}} \left(\boldsymbol{y}_{\boldsymbol{p},\boldsymbol{r}} + \boldsymbol{C}_{\boldsymbol{p},\boldsymbol{r}} \right) \tag{1}$$

Subject to:

$$\sum_{p \in P} \sum_{\boldsymbol{r} \in \boldsymbol{R}: \boldsymbol{r} > \boldsymbol{1}} {\boldsymbol{t}^{\boldsymbol{in}}}_{\boldsymbol{p},\boldsymbol{r}} \leq \boldsymbol{T}^{\textbf{total}} \tag{2}$$

$$\sum_{\boldsymbol{p} \in \boldsymbol{P}} {\boldsymbol{t}^{\boldsymbol{in}}}_{\boldsymbol{p},\boldsymbol{r}} \leq \boldsymbol{T}_{\boldsymbol{r}} \qquad \forall \boldsymbol{r} \in \boldsymbol{R}: \boldsymbol{r} > \boldsymbol{1} \tag{3}$$

$$\sum_{\boldsymbol{p} \in \boldsymbol{P}} \boldsymbol{x}_{\boldsymbol{p},\boldsymbol{r}} - \boldsymbol{x}_{\boldsymbol{p},\boldsymbol{r}-\boldsymbol{1}} = {\boldsymbol{t}^{\boldsymbol{in}}}_{\boldsymbol{p},\boldsymbol{r}} - {\boldsymbol{t}^{\boldsymbol{out}}}_{\boldsymbol{p},\boldsymbol{r}} \qquad \forall \boldsymbol{r} \in \boldsymbol{R}: \boldsymbol{r} > \boldsymbol{1} \tag{4}$$

$$\sum_{\boldsymbol{p} \in \boldsymbol{P}} \boldsymbol{C}_{\boldsymbol{p},\boldsymbol{r}} = \boldsymbol{1} \qquad \forall \boldsymbol{r} \in \boldsymbol{R} \tag{5}$$

$$\sum_{\boldsymbol{p} \in \boldsymbol{P}} \boldsymbol{x}_{\boldsymbol{p},\boldsymbol{r}} \geq \boldsymbol{C}_{\boldsymbol{p},\boldsymbol{r}} \qquad \forall \boldsymbol{r} \in \boldsymbol{R} \tag{6}$$

$$\sum_{\boldsymbol{p} \in \boldsymbol{P}_{\boldsymbol{GK}}} \boldsymbol{x}_{\boldsymbol{p},\boldsymbol{r}} = \boldsymbol{Q}_{\boldsymbol{max}_{\boldsymbol{GK}}} \qquad \forall \boldsymbol{r} \in \boldsymbol{R} \tag{7a}$$

$$\sum_{\boldsymbol{p} \in \boldsymbol{P}_{\boldsymbol{DEF}}} \boldsymbol{x}_{\boldsymbol{p},\boldsymbol{r}} = \boldsymbol{Q}_{\boldsymbol{max}_{\boldsymbol{DEF}}} \qquad \forall \boldsymbol{r} \in \boldsymbol{R} \tag{7b}$$

$$\sum_{\boldsymbol{p} \in \boldsymbol{P}_{\boldsymbol{MID}}} \boldsymbol{x}_{\boldsymbol{p},\boldsymbol{r}} = \boldsymbol{Q}_{\boldsymbol{max}_{\boldsymbol{MID}}} \qquad \forall \boldsymbol{r} \in \boldsymbol{R} \tag{7c}$$

$$\sum_{p \in P_{GK}} x_{p,r} = Q_{max_{FWD}} \qquad \forall r \in R \qquad (7d)$$

$$y_{p,r} \leq x_{p,r} \qquad \forall p \in P; r \in R \qquad (8)$$

$$\sum_{p \in P_{GK}} y_{p,r} = Q_{points_{GK}} \qquad \forall r \in R \qquad (9a)$$

$$\sum_{p \in P_{DEF}} y_{p,r} = Q_{points_{DEF}} \qquad \forall r \in R \qquad (9b)$$

$$\sum_{p \in P_{MID}} y_{p,r} = Q_{points_{MID}} \qquad \forall r \in R \qquad (9c)$$

$$\sum_{p \in P_{FWD}} y_{p,r} = Q_{points_{FWD}} \qquad \forall r \in R \qquad (9d)$$

$$\sum_{p \in P} y_{p,r} = X_r \qquad \forall r \in R \qquad (10)$$

$$b_1 + \sum_{p \in P} price_{p,1} \times x_{p,1} \leq B \qquad (11)$$

$$b_r = b_{r-1} + \sum_{p \in P} price_{p,r} \times t^{out}{}_{p,r} - \sum_{p \in P} price_{p,r} \times t^{in}{}_{p,r} \qquad (12)$$

The objective function (1) aims to maximize the cumulative total points over the entire season. It accomplishes this by aggregating the points attained by each player in each game week, which are then multiplied by binary variables that indicate whether the player's score contributes to the team's overall score and whether the player is designated as the captain for that game week.

Constraint (2) guarantees that the combined incoming transfers made for all players during all game weeks, except the initial game week, do not surpass the predetermined maximum allowable number of transfers permitted for the entire season. On the other hand, Constraint (3) restricts the number of incoming transfers for each player within each game week, ensuring that it does not surpass the predefined maximum allowable transfers allowed for each individual game week.

Constraint (4) ensures that the difference in the number of times a player is in the team between two consecutive game weeks is equal to the difference between incoming and outgoing transfers for that player in the current game week.

- If player *p* is in the team for game week *r* ($x_{p,r}$ = 1), and they were also in the team for the previous game week ($x_{p,r-1}$ = 1), there should be no transfer made for this player between the two game weeks ($t^{in}{}_{p,r}$ = 0 and $t^{out}{}_{p,r}$ = 0).
- Conversely, if player *p* is in the team for game week *r* ($x_{p,r}$ = 1), but was not in the team for the previous game week ($x_{p,r-1}$ = 0), it implies that the player was transferred into the team for the current game week ($t^{in}{}_{p,r}$ = 1 and $t^{out}{}_{p,r}$ = 0).
- Similarly, if player *p* is not in the team for game week *r* ($x_{p,r}$ = 0) but was in the team for the previous game week ($x_{p,r-1}$ = 1), it indicates that the player was transferred out of the team for the current game week ($t^{in}{}_{p,r}$ = 0 and $t^{out}{}_{p,r}$ = 1).

This constraint emulates a scenario in which managers are allowed to either sell or purchase a player for the team during each game week.

Constraints (5) and (6) serve to guarantee the exclusivity of selecting a solitary captain for every game week while further ensuring that the chosen captain must be part of the designated squad.

Constraints (7a) and (7b) impose limitations on the upper bound of players permitted for each position within the composition of the 15-man squad.

Constraint (8) requires that the number of players whose points are considered for team points be less than the total number of players initially chosen for the squad.

Constraint (9) ensures that the overall count of players, whose scores are considered for the team's score, must not surpass the specified limitation set for each position. Constraint (10) guarantees that the total count of players whose scores contribute to the team's overall score is precisely 11.

Constraint (11) assures that the sum of the beginning team's value plus the remaining budget for the first game week does not exceed the entire budget. Constraint (12) computes the remaining budget for each game week by considering the initial budget and the prices of players traded in and out of the team. It considers both incoming and outgoing trades for each player to determine the available budget for subsequent game week.

4. Analysis and Results

The MILP model was implemented using IBM ILOG CPLEX Studio IDE version 22.1.0. The required data was extracted from an Excel sheet containing information from the 2022/23 FPL season. For reference, the code used in the implementation can be found in the appendix.

The average computational run time for the implementation was 45 seconds. The processor used was an Intel 11th gen Core i5-1135G7, 2.5Ghz with 8 GB RAM. The statistics display 83,956 constraints and 105,285 variables, with 105,248 being binary and 37 falling into other categories. Furthermore, a total of 334,665 non-zero coefficients are observed.

4.1. Optimal Solution

The optimal solution obtained from the model was 3,872 points. When comparing this final score to the actual winner of the FPL 2022/23 season (Rigga, S, 2023), who scored a total of 2,776 points, the model achieved a difference of 1,096 points higher than the actual winner's score.

The table below provides an overview of the scores achieved by the winners in the Fantasy Premier League over the past five years. Notably, this data highlights a considerable scope for managers to make substantial improvements in their strategies and performance.

Season	Winner's Score	Model's Score- 22/23
2022/23	2776	3872
2021/22	2844	
2020/21	2680	
2019/20	2557	
2018/19	2659	

Table 3: Past FPL Season Winners

Indeed, the data provided reveals a consistent trend where the model consistently surpasses the performance of the actual winners across various seasons. While it is crucial to recognize the year-to-year variability in statistics, he indicative benchmark for an average winning score is approximately situated at the 2600-point mark, information that is sourced from (Sharma A, 2023, Goal.com), a widely respected source for football news. This difference further emphasizes the model's significant effectiveness in achieving superior scores compared to the actual winners over various years. This implies that the model, when provided with data from any particular year, has the potential to exhibit a substantial performance increase compared to the actual season champions.

4.2. Player Selections

Let us now conduct a thorough examination of the player selections made by the model to gain a deeper understanding and substantiate its functioning.

The provided graphs illustrate the players who have been selected by the model at least 5 times, arranged in descending order from the highest selections to the lowest. Figure 4.1 displays the selections made in the squad, while Figure 4.2 showcases the selections made for the starting 11.

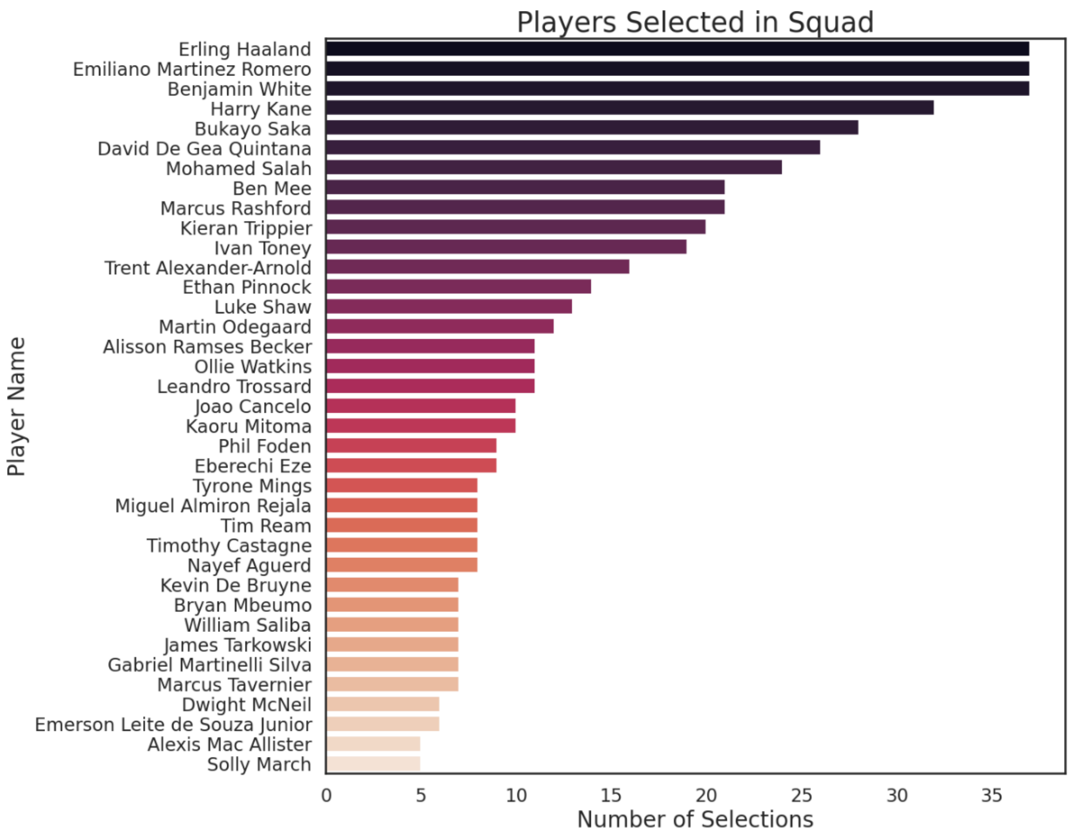

Figure 3: Top Players Selected in Squad

The data reveals that players such as Erling Haaland, Emiliano Martinez Romero, and Benjamin White have been consistently selected by the model for all game weeks, amounting to 37 selections. Additionally, players like Harry Kane, Bukayo Saka, and David De Gea Quintana were selected by the model more than 25 times.

Upon closer examination of the players selected in the starting 11, whose scores significantly contributed to the overall team score, a consistent pattern emerges. Specifically, Benjamin White was the predominant choice, selected an impressive 30 times. Following closely, Harry Kane's selection frequency reached 25 times, while Erling Haaland garnered 24 selections. Notably, Bukayo Saka secured 21 selections, while Kieran Trippier and Mohamed Salah each received 19 selections. The observed frequencies of these selections distinctly emphasize a noticeable inclination towards these players in the model's strategic preferences.

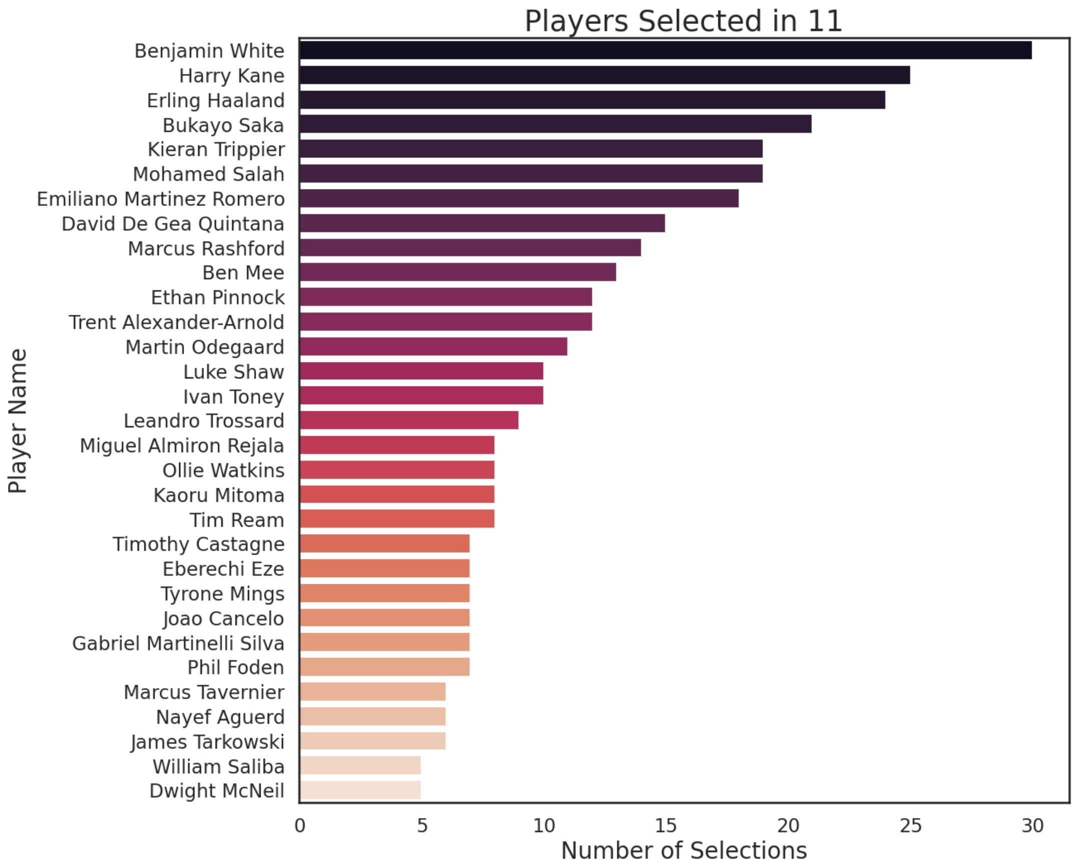

Figure 4: Top Players Selected In 11

These selections find justification in Kane and Haaland's top-performing status of the season with 272 and 263 points, respectively. Similarly, Kieran Trippier and Ben White's selections align with their attainment of the highest defender points at 198 and 156 points, respectively. Additionally, Mohamed Salah's selection coincides with his distinction as the leading midfielder, accumulating 239 points. Meanwhile, Bukayo Saka, securing the fourth-highest points, gathered a commendable total of 202 points.

However, the figures above showcase the top selected players from all playing positions. To gain a deeper insight into the model's player selections, let's examine the top selected players in the starting eleven from each position along with their total points and price during the 2022/23 season. This will provide a clearer understanding of the reasons behind the model's choices for each position.

4.2.1. Defenders

A total of 15 defenders were chosen for inclusion in both squads and starting elevens over the course of the 37 game weeks. Among these defenders, 8 players were favoured with selections exceeding 7 times. The following graph highlights these 8 defenders, their frequency of selection by the model, and the total points allocated to them during the 2022/23 season.

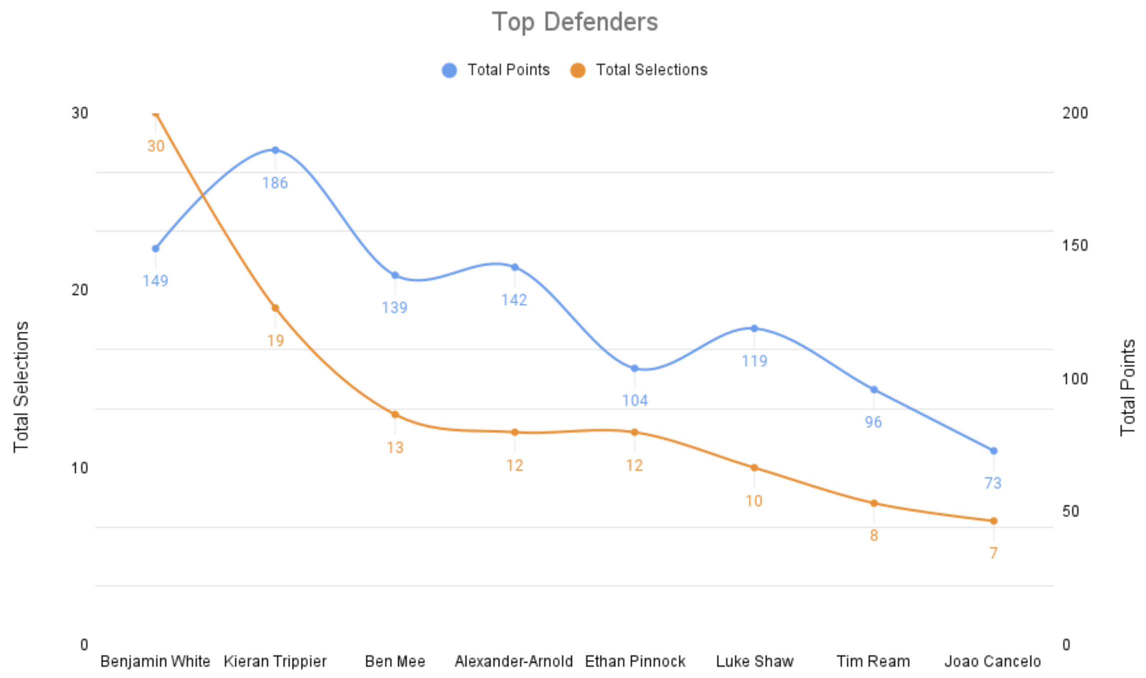

Figure 5: Top Selected Defenders

As observed, a clear correlation emerges between players' point totals and the model's selections. Players with lower points receive fewer selections. However, exceptions arise: Ben White, Ben Mee, and Ethan Pinnock received more selections than Keiran Trippier, Trent Alexander-Arnold, and Luke Shaw, despite the latter group accumulating more points.

This deviation finds its resolution in the context of value for money—an aspect that comes into play. The following table presents these players' overview, highlighting their average prices and average points attained per game week.

Name	Average Points	Average Price	Value = points/price
Benjamin White	4	4.7	0.85
Kieran Trippier	5	5.6	0.89
Ben Mee	3.8	4.6	0.83
Alexander-Arnold	3.8	7.4	0.51
Ethan Pinnock	2.8	4.3	0.65
Luke Shaw	3.2	4.9	0.65
Tim Ream	2.6	4.5	0.58
Joao Cancelo	2	7.2	0.28

Table 4: Points & Prices of Defenders

The points/price statistic accurately assesses a player's worth. The closer it is to 1, the more a player's points correspond proportionally to their cost. This pivotal measurement is appropriately termed "Value."

It is evident that Keiran Trippier possesses a higher value at 0.85 compared to Benjamin White. Nonetheless, White stands out as notably more budget-friendly, with a price difference of 0.9 million, yet still offering a closely comparable value with a mere 0.04 difference. This variance becomes even more pronounced when examining Ben Mee and Alexander Arnold. Ben Mee, for instance, provides equivalent average points at a price nearly 3 million lower. A similar trend emerges with Pinnock and Shaw, where Pinnock's slightly lower value is offset by its cost advantage of 0.6 million, thereby justifying its elevated selection rate.

In essence, the former trio embodies a convincing proposition of cost-effectiveness, given their considerable affordability in comparison to their counterparts. This divergence in the price-to-performance ratio substantiates the rationale behind their heightened selection rates, even in the presence of a slight decrease in point accumulation.

4.2.2. Midfielders

The model opted for a total of 22 defenders to be incorporated into squads and starting elevens. Once again, our attention is directed towards the most frequently chosen players, those who have been selected on more than 7 occasions. Within this category, a collective of 9 defenders garnered selections surpassing 7 selections.

The following graph displays these 9 defenders, indicating the model's selections and the total points they accumulated over the season.

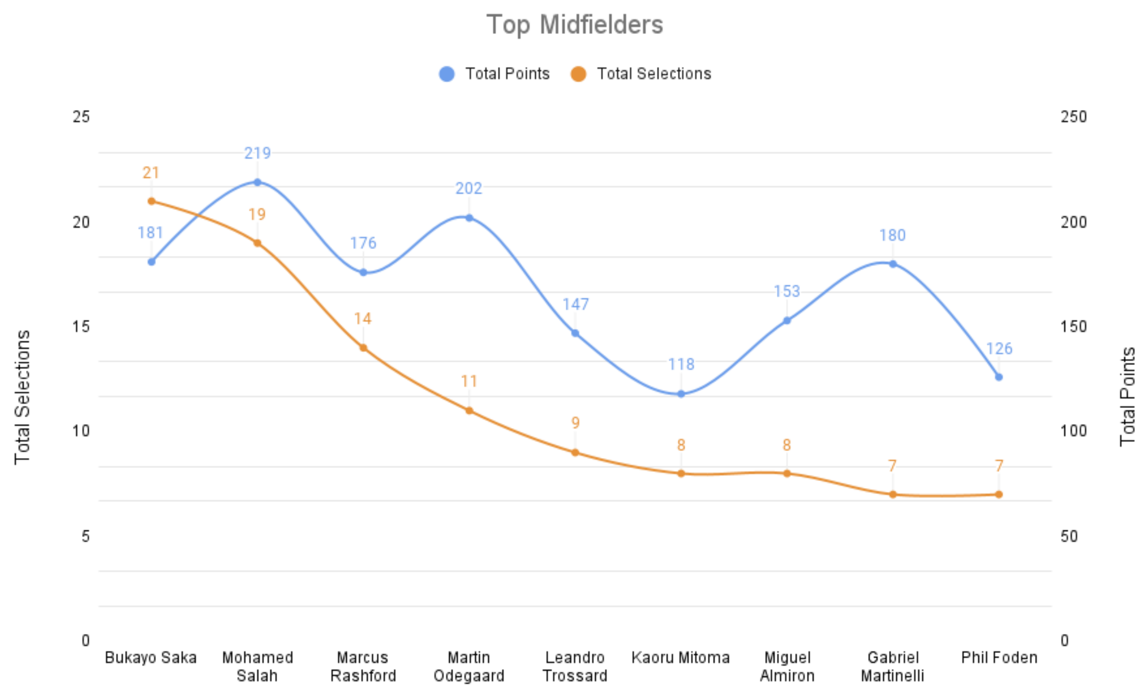

Figure 6: Top Selected Midfielders

The graph reflects a comparable overarching pattern, whereby selections tend to decrease as total points decline. However, the relationship is not strictly linear, with occasional fluctuations evident in the total points line. Notably, Martin Odegaard, chosen just 11 times, amasses 202 points—surpassing Marcus Rashford and Bukayo Saka, who accumulate fewer points at 181 and 176 respectively. Remarkably, Rashford receives 21 selections, and Saka, 14. Similarly, Gabriel Martinelli garners fewer selections than Leandro Trossard, Kaoru Mitoma, and Miguel Almiron, despite having higher points count. Let us look at the average price and average points table for these midfielders to get a better understanding.

Name	Average Points	Average Price	Value= points/price
Bukayo Saka	4.9	8.1	0.6
Mohamed Salah	5.9	12.9	0.46
Marcus Rashford	4.8	6.7	0.72
Martin Odegaard	5.5	6.6	0.83
Leandro Trossard	4	6.7	0.6
Kaoru Mitoma	3.2	5.1	0.63
Miguel Almiron	4.1	5.3	0.77
Gabriel Martinelli	4.9	6.6	0.74
Phil Foden	3.4	8.1	0.42

Table 5: Points & Prices of Midfielders

From this table, it is apparent why Kaoru Mitoma and Miguel Almiron have secured more selections than Gabriel Martinelli. Both players boast lower prices while delivering consistent performances, with values of 0.63 and 0.77 respectively—Almiron's value being higher than Martinelli's 0.74.

However, the model demonstrates a preference for selecting Bukayo Saka and Marcus Rashford more frequently than Martin Odegaard, despite the latter's lower price and higher value. This divergence might stem from the model's focus on transferring out other players during most game weeks. For instance, once Saka was chosen, the model discovered superior transfers in subsequent weeks, potentially yielding higher weekly points. Consequently, there was no need to substitute Saka for Odegaard. A similar scenario applies to Rashford. Regarding Salah, besides being the highest-scoring midfielder, he maintains the highest average points at 5.9 per week. Notably, he emerges as the season's best performing midfielder, proving his boosted selections.

4.2.3. Forwards

The model included a total of 7 forwards in its selections. Among these, 6 players were chosen more than 2 times, as depicted in the graph below, which also illustrates their total points.

Remarkably, a consistent downward trend aligns between total selections and total points, devoid of anomalies. Each player who received fewer selections compared to their predecessors also amassed fewer points. The dominance of Kane and Haaland in selections becomes evident, resulting in a limited inclusion of other forwards in the starting elevens.

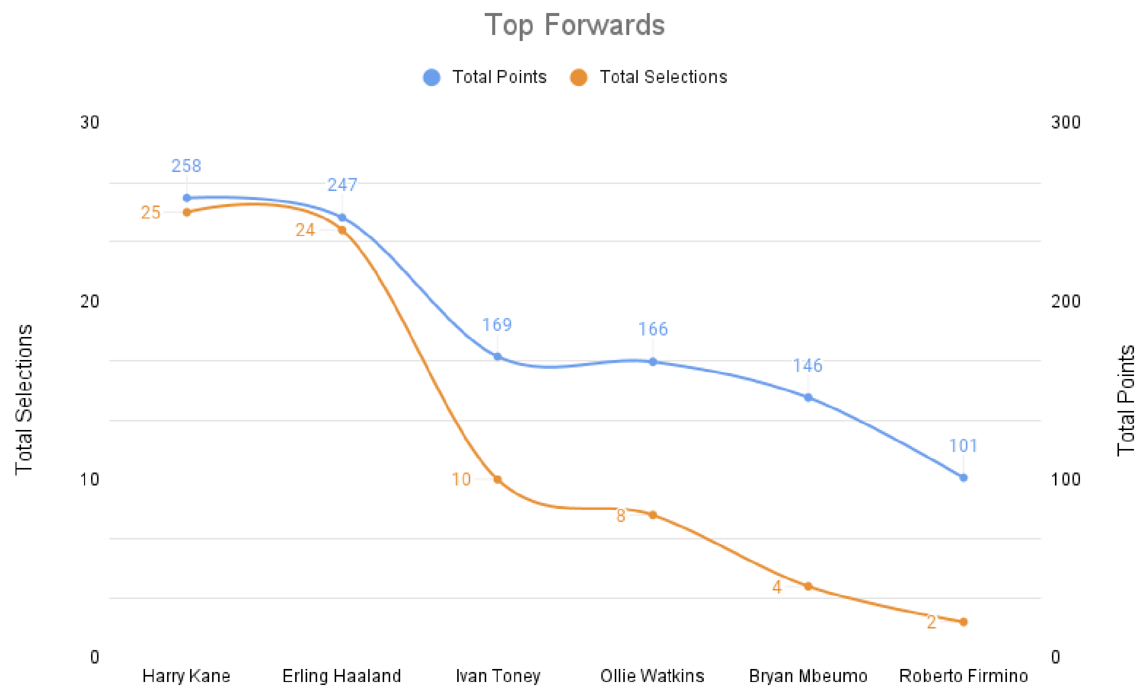

Figure 7: Top Selected Forwards

The average points and price table offers valuable insights into player performance and value. Ivan Toney, Ollie Watkins, and Bryan Mbuemo stand out as players who have delivered commendable value considering their prices. In contrast, Roberto Firmino received selections in Game weeks 4 and 5, capitalizing on a burst in points with a notable 22 and 9 respectively, before the model opted to transfer him out. Additionally, Gabriel Jesus, not depicted in this table, received a single selection during game week 2, amassing 19 points. However, his subsequent performances waned and was not selected again.

Name	Average Points	Average Price	Value= points/price
Harry Kane	7	11.6	0.6
Erling Haaland	6.7	12.1	0.55
Ivan Toney	4.6	7.2	0.64
Ollie Watkins	4.5	7.3	0.62
Bryan Mbeumo	3.9	5.7	0.68
Roberto Firmino	2.7	8	0.33

Table 6: Points & Prices of Forwards

4.2.4. Goalkeepers

The model's goalkeeper selections were limited to only three players, and their corresponding data is displayed in the table below:

Name	Total selections	Average Points	Average Price	Value= points/price
Emiliano Martinez Romero	18	3.6	4.9	0.73
David De Gea	15	3.9	4.8	0.81
Alisson Becker	4	3.8	5.4	0.70

Table 7: Points & Prices of Forwards

Alisson, with the lowest value, garnered the fewest selections. Interestingly, despite David De Gea's seemingly better value proposition, the model preferred Emiliano Martinez Romero with a higher selection count.

4.3. Captain & Squad Selections

The model designated a total of 25 different players as captains across the 37 game weeks. The following word cloud visually represents all the selected captains. The size of each player's name corresponds to the frequency of their selection as captain.

Figure 8: Captains

The analysis reveals that Erling Haaland emerged as the most frequently chosen captain, receiving this designation on 5 occasions. This outcome is in line with his impressive average weekly points of 6.7. Additionally, his status as the leading goal scorer of the previous season with a remarkable tally of 36 goals likely contributed to his popularity as a captain pick. Martin Odegaard secured the 4th highest number of captain selections, closely followed by Marcus Rashford in 3rd place. Bukayo Saka, Mohamed Salah, and Harry Kane each received 2 captain selections, while the remaining players were granted 1 captaincy selection each.

Let us now delve into a more detailed examination of the teams assembled by the model. To illustrate these selections, we have chosen teams from five game weeks spread across the 37 weeks, offering a snapshot of the model's decision-making process. The following teams of game weeks 1,9, 18, 27 and 35, complete with captain selections and the corresponding total points, are showcased below:

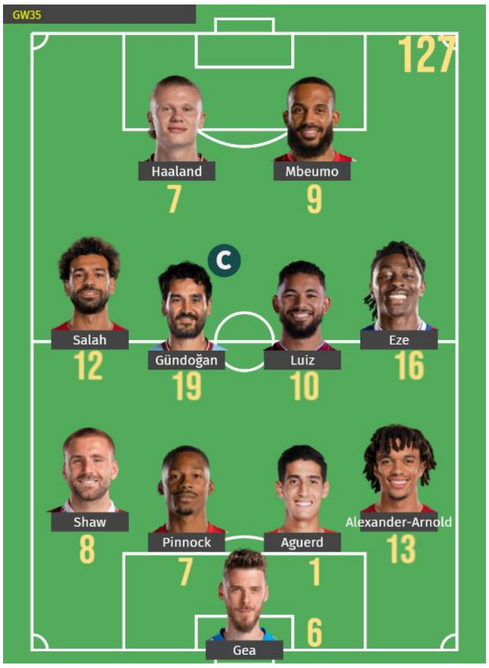

Figure 9: Team Selections

Each of the provided images portrays the team chosen for the specified game weeks. The formation adheres to the 4-4-2 setup that was previously established, comprising 4 defenders, 4 midfielders, and 2 forwards. The points accrued by each player are denoted by the yellow text below their names, while the cumulative points for each game week, incorporating the captain's doubled points, are presented in the top right corner. The captain for each respective week is marked by the "C" logo placed above their name, signifying their captaincy role for that specific game week. The teams exhibit a dynamic nature across the span of game weeks, with certain constants such as Haaland, Kane, and White. Remarkably, the midfield positions witness the most fluctuations, often featuring a varying set of players.

Examining the team composition for the initial game week, a budget of 95 million was utilized, adhering to the 100 million limit.

In the context of Fantasy Premier League, a basic reference point for managers is to achieve a score within the range of 50 to 59, representing the minimum goal. A score ranging from 60 to 75 is commonly considered as a praiseworthy performance, whereas the bracket of 76 to 90 is recognized as excellent. An exceptional achievement is represented by attaining a score of 100 or higher in a single game week. This information was also sourced from Jaganath, D. (2023).

Considering these aspects, it becomes evident that the model consistently attains impressive scores, often surpassing the 100-point mark. Notably, in game week 35, the model even achieved a remarkable score of 127. This trend highlights the model's ability to assemble high-scoring teams within the constraints provided. The average points per week for the teams formulated by the model stands at 104.

Additionally, it is evident that the player achieving the highest points in each game week is chosen as the captain, a strategic decision aimed at optimizing the overall point accumulation.

In summary, the findings affirm the model's intended functionality, effectively constructing optimal teams each week using retrospective data. Moreover, the model consistently surpasses the cumulative scores achieved by the actual winners, reaffirming its efficacy in the context of fantasy football team selection.

5. Conclusion

In conclusion, this study embarked on a journey to optimize Fantasy Premier League (FPL) team selection using a Mixed-Integer Linear Programming (MILP) model. By leveraging real-world FPL data from the 2022/23 season, the model demonstrated its remarkable ability to strategically compose a team that consistently outperformed actual season winners in terms of total points. The substantial difference of 1,096 points between the model's optimal score of 3,872 and the 2022/23 season winner's score of 2,776 stands as concrete evidence of the efficacy of the model's decision-making process. This achievement not only demonstrates the significance of employing advanced analytical techniques in complex decision-making scenarios but also indicates the potential prevalence of algorithmic approaches in the realm of fantasy sports optimization.

The exploration of the MILP model's player selections reveals an insight into the strategies that govern the model's performance. The consistent selection of key players such as Erling Haaland, Emiliano Martinez Romero, and Benjamin White showcases the model's emphasis on players with consistent performance over time. Furthermore, the model's repeated selections of players like Harry Kane, Bukayo Saka, and David De Gea Quintana highlight its recognition of players who consistently contribute to the team's overall points tally. These selections offer insights for enthusiasts seeking to refine their own selection methodologies.

It's essential to acknowledge certain limitations of the model. Notably, the model's predictions rely heavily on historical data, resulting in ex-post results. Furthermore, the model's inherent inability to promptly adapt to sudden shifts during the season, such as injuries, transfers, or tactical changes, suggests a potential avenue for future development—a more dynamic forecasting model. A forecasting model has the potential to offer real-time insights to managers, assisting them in navigating the complexities of a live football season. While the creation of such a forecasting model poses challenges, its success would be determined by its ability to consistently outperform weekly team selections.

Looking ahead, future work in this domain could delve into incorporating the chips which include the wildcard, free hit, and triple captain strategies. This would introduce additional complexities and dimensions to the optimization problem. Moreover, this expansion would directly address the practical challenges encountered by FPL managers during the season, thereby closing the gap between theoretical modelling and the dynamic reality of real-time decision making. Future endeavours could encompass the application of this model to other fantasy leagues, such as Ligue 1, Serie A, and the Bundesliga, among others. This extension would yield valuable insights into the model's adaptability and effectiveness across diverse contexts and leagues, contributing to the overall validation of its robustness.

To sum up, the intricate interplay between the world of sports and analytics is vividly demonstrated in the outcomes of this study. The systematic integration of data-driven methods with sports management yields promising results, exemplified by the consistent outperformance of the MILP model compared to real-world performance. This research underscores the potential advantages that arise from combining analytical methodologies with sporting activities, offering insights not only to fantasy sports enthusiasts but also to the broader field of data-driven decision-making across various domains. This symbiotic relationship between analytics and sports presents a compelling narrative that enhances our understanding of how quantitative approaches can enrich diverse aspects of human engagement and decision-making.

6. Appendix

CPLEX code:

Model file:

```
// SETS
{string} P = ...;
{string} Q = ...;
int Rounds = ...;
range R = 1..Rounds;

// Players in each position
{string} Q_GK = ...;
{string} Q_DEF =...;
{string} Q_MID =...;
{string} Q_FWD =...;

// PARAMETERS
float points[P][R]= ...;
float price[P][R]= ...;
float B = ...;
int Q_max[Q]= ...;
int Q_Points[Q]= ...;
int X = ...;
int T_total = ...;
int T = ...;

// DECISION VARIABLES
dvar boolean x[P][R];
dvar boolean y[P][R];
dvar boolean C[P][R];
dvar boolean t_in[P][R];
dvar boolean t_out[P][R];
dvar float+ b[R];

// 1.OBJECTIVE FUNCTION
maximize sum(p in P, r in R) (points[p][r] * (y[p][r] + C[p][r]));

// CONSTRAINTS
subject to {
  // 2.Transfer constraint per season
  forall (p in P, r in R: r > 1)
     t_in[p][r] <= T_total;

// 3.Limiting the number of transfers per GW
  forall (r in R)
     sum(p in P) t_in[p][r] <= T;

// 4.Update of transfers per GW
  forall (p in P, r in R: r > 1)
    x[p][r] - x[p][r-1] == t_in[p][r] - t_out[p][r];


// 5.Captain selection
  forall (r in R)
      sum(p in P) C[p][r] == 1;

// 6.Captain position selection
```

```
  forall (r in R, p in P)
     x[p][r] >= C[p][r];

// 7.Vacancies for each position
  forall (r in R) {
    sum(p in Q_GK) x[p][r] == Q_max["GK"];
    sum(p in Q_DEF) x[p][r] == Q_max["DEF"];
    sum(p in Q_MID) x[p][r] == Q_max["MID"];
    sum(p in Q_FWD) x[p][r] == Q_max["FWD"];
  }

// 8.Player must be in scoring position for score to be counted
  forall (p in P, r in R)
     y[p][r] <= x[p][r];

// 9.Max number of players from Each position allowed
  forall (r in R) {
         sum(p in Q_GK) y[p][r] == Q_Points["GK"];
         sum(p in Q_DEF) y[p][r] == Q_Points["DEF"];
         sum(p in Q_MID) y[p][r] == Q_Points["MID"];
         sum(p in Q_FWD) y[p][r] == Q_Points["FWD"];
       }

// 10. Players whose points count towards team points in each week
  forall (r in R)
     sum(p in P) y[p][r] == X;

// 11. Initial team budget lesser than total budget
  forall (r in R)
  b[1] + sum(p in P) price[p][1] * x[p][1] <= B;

// 12.Remaining budget calculation
  forall(r in R:r>1)
    b[r] == b[r-1] + sum(p in P) price[p][r] * t_out[p][r] - sum(p in P)
price[p][r] * t_in[p][r];

}
```

Data file:

```
//SETS
SheetConnection FplGwData("FINAL FPL DATA SET.xlsx");

// set of players p
P from SheetRead(FplGwData, "'Prices'!A2:A573");

// set of postions q
Q = {"GK", "DEF", "MID", "FWD"};

// Number of matchdays
Rounds = 37;


// Players in each postion
Q_GK from SheetRead(FplGwData, "'Positions'!A2:A57");
Q_DEF from SheetRead(FplGwData, "'Positions'!B2:B209");
Q_MID from SheetRead(FplGwData, "'Positions'!C2:C244");
Q_FWD from SheetRead(FplGwData, "'Positions'!D2:D61");
```

```
points from SheetRead(FplGwData, "'Points'!F2:AP573");

price from SheetRead(FplGwData, "'Prices'!C2:AM573");

X = 11;
B = 100.0;
Q_max = [2,5,5,3];
Q_Points = [1,4,4,2];
T = 1;
T_total = 37;

x to SheetWrite(FplGwData, "'x'!E2:AO573");
y to SheetWrite(FplGwData, "'y'!E2:AO573");
C to SheetWrite(FplGwData, "'C'!E2:AO573");
```

References

1. Davis NW, Duncan MC. 2006. *Sports knowledge is power: reinforcing masculine privilege though fantasy sport league participation*. J Sport Soc Issues. 30:244–264.
2. Nesbit TM, King KA. 2011. *Major League Baseball attendance and the role of fantasy baseball.* J Sports Econ. 13:494–514.
3. Hunter, D. S., Vielma, J. P., & Zaman, T. (2016). *Picking winners in daily fantasy sports using integer programming.* INFORMS Journal on Applied Analytics, 46(6), 742-757.
4. Beliën, J., Goossens, D. and Van Reeth, D., 2017. *Optimization modelling for analyzing fantasy sport games*. INFOR: Information Systems and Operational Research, 55(4), pp.275-294.
5. Becker, Adrian and Sun, Xu Andy. *"An analytical approach for fantasy football draft and lineup management"* Journal of Quantitative Analysis in Sports, vol. 12, no. 1, 2016, pp. 17-30. https://doi.org/10.1515/jqas-2013-0009
6. G D Sharp, W J Brettenny, J W Gonsalves, M Lourens & R A Stretch (2011) - *Integer optimisation for the selection of a Twenty20 cricket team*, Journal of the Operational Research Society, 62:9, 1688-1694,
7. Brettenny, W.J., Friskin, D.G., Gonsalves, J.W. and Sharp, G.D., 2012. *A multi-stage integer programming approach to fantasy team selection: A Twenty20 cricket study*. South African Journal for Research in Sport, Physical Education and Recreation, 34(1), pp.13-28.
8. Chand, S., Singh, H.K. and Ray, T., 2018, July. *Team selection using multi-/many-objective optimization with integer linear programming.* In 2018 IEEE Congress on Evolutionary Computation (CEC) (pp. 1-8). IEEE.
9. Edwards, S.J., 2021. *Analyzing Fantasy Sport Competitions with Mixed Integer Programming.* In Data and Decision Sciences in Action 2: Proceedings of the ASOR/DORS Conference 2018 (pp. 167-182). Springer International Publishing
10. Sharma A, 2023, *How many points do you need to win Premier League fantasy football?* Goal.com, Retrieved July 25, 2023, from https://www.goal.com/en-gb/news/how-many-points-do-you-need-to-win-premier-league-fantasy-football/blt53ca491b03f03b8c
11. Jaganath, D. (2023). *What Is Considered A GOOD Score In FPL? (3 Methods).* Football Handbook. Retrieved July 25, 2023, from https://footballhandbook.com/good-score-fpl/.
12. Premier League. (2023). *FPL champion: The one thing I changed to win title.* Premier League. Retrieved July 25, 2023, from https://www.premierleague.com/news/3495504

13. Rigga, S. (2023). *FPL 2022/23 winner: Who was the overall champion? Fantasy Football Scout.* Retrieved July 25, 2023, from https://www.fantasyfootballscout.co.uk/2023/05/28/fpl-2022-23-winner-who-was-the-overall-champion/
14. Anand, V. (2022). FPL Historical Dataset. Retrieved June 16, 2023, from [https://github.com/vaastav/Fantasy-Premier-League/].
15. Official FPL website. Fantasy Premier League Statistics. Retrieved from https://fantasy.premierleague.com/statistics