

Article

Leveraging Sports Analytics and Association Rule Mining to Uncover Recovery and Economic Impacts in NBA Basketball

Vangelis Sarlis , George Papageorgiou  and Christos Tjortjis * 

School of Science and Technology, International Hellenic University, 57001 Thessaloniki, Greece; e.sarlis@ihu.edu.gr (V.S.); gpapageorgiou2@ihu.edu.gr (G.P.)

* Correspondence: c.tjortjis@ihu.edu.gr

Abstract: This study examines the multifaceted field of injuries and their impacts on performance in the National Basketball Association (NBA), leveraging a blend of Data Science, Data Mining, and Sports Analytics. Our research is driven by three pivotal questions: Firstly, we explore how Association Rule Mining can elucidate the complex interplay between players' salaries, physical attributes, and health conditions and their influence on team performance, including team losses and recovery times. Secondly, we investigate the relationship between players' recovery times and their teams' financial performance, probing interdependencies with players' salaries and career trajectories. Lastly, we examine how insights gleaned from Data Mining and Sports Analytics on player recovery times and financial influence can inform strategic financial management and salary negotiations in basketball. Harnessing extensive datasets detailing player demographics, injuries, and contracts, we employ advanced analytic techniques to categorize injuries and transform contract data into a format conducive to deep analytical scrutiny. Our anomaly detection methodologies, an ensemble combination of DBSCAN, isolation forest, and Z-score algorithms, spotlight patterns and outliers in recovery times, unveiling the intricate dance between player health, performance, and financial outcomes. This nuanced understanding emphasizes the economic stakes of sports injuries. The findings of this study provide a rich, data-driven foundation for teams and stakeholders, advocating for more effective injury management and strategic planning. By addressing these research questions, our work not only contributes to the academic discourse in Sports Analytics but also offers practical frameworks for enhancing player welfare and team financial health, thereby shaping the future of strategic decisions in professional sports.

Keywords: association rule mining; basketball analytics; data mining; data science; sports analytics



Citation: Sarlis, V.; Papageorgiou, G.; Tjortjis, C. Leveraging Sports Analytics and Association Rule Mining to Uncover Recovery and Economic Impacts in NBA Basketball. *Data* **2024**, *9*, 83. <https://doi.org/10.3390/data9070083>

Academic Editor: Daniel Rojas-Valverde

Received: 6 June 2024

Revised: 18 June 2024

Accepted: 21 June 2024

Published: 24 June 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction and Background

The concept of Sports Analytics has revolutionized the way teams and organizations approach game strategy, player health, and financial management. In the NBA, where player contracts amount to millions and game outcomes significantly impact revenue streams, understanding the nuances of recovery and financial stability is paramount. Data Science and Machine Learning offer potent tools to delve into these aspects, providing insights that traditional methods might overlook [1–3].

Data Science for Sports Analytics is an emerging field that combines rigorous statistical and computational methods to interpret complex data generated in sports contexts. In NBA basketball, Sports Analytics can be applied to various domains, including player performance, injury prevention, and financial decision-making. Machine Learning and Data Mining techniques further enhance this analysis, enabling predictive models that guide strategic planning and recovery protocols [4–6].

Recovery in professional basketball encompasses physical, mental, and strategic dimensions. Data-driven approaches help identify optimal recovery times and methods,

reducing the risk of injuries and prolonging player careers. Financial aspects are equally critical; unplanned absences due to injuries can lead to significant financial losses for NBA franchises. This study explores how Data Science and Sports Analytics can mitigate these risks by providing deeper insights into player health and strategic game management [7–13].

In a previous study, various factors influencing player injuries and their impact on team performance were explored. For example, generalized joint hypermobility (GJH) was linked to an increased risk of ACL injury and inferior outcomes after ACL reconstruction. Research also indicates that injuries to NBA players occur more frequently during games, particularly at critical moments, such as landings from jumps or pivoting, and tend to be more common during the competitive season, due to higher game participation and overuse. Another study evaluated the return-to-sport and performance of NBA players after musculoskeletal reconstruction. The study found that NBA players were able to return to play after reconstruction, but their performance was not always the same as pre-injury [14–17].

Some studies have suggested that players who are older and have a history of previous knee injuries may be at a higher risk of sustaining an ACL injury [18–20]. Other studies have suggested that players who have a history of knee instability, or who have certain muscle imbalances or weaknesses, may also be at a higher risk of sustaining an ACL injury. It is worth mentioning that there are many factors that contribute to ACL injuries, which are not solely determined by the characteristics of the player. Factors such as the intensity and the frequency of the game, the surface of the court, and the type of shoes worn by the players can also play a role [21–23].

Knee tears have attracted researchers' attention in recent decades due to their high epidemiological rates in different sports (contact and non-contact), types of exposure (competition and practice), genders (male and female), and age groups (pediatric, adolescent, and adult populations) [24,25]. Moreover, there has been research into the economic and performance impact of injury on the player and team [26], post-injury or post-surgery consequences, and the increasing risk of injury to the contralateral or the same graft [24].

Work in [27] presents the use of deep learning to monitor on-field player workload exposure and knee injury risk in athletes. The authors aim to use their findings to identify injury risk factors and improve injury prevention in team sports. The study involved collecting data on player workload and injury occurrence, which was then analyzed using deep learning algorithms. The results showed that the deep learning approach can accurately predict knee injury risk and provide valuable insights into factors that contribute to injury.

Another study evaluated the effect of hip-focused injury prevention training on reducing the incidence of injuries in basketball players. A sample of basketball players was divided into two groups: one group received hip-focused injury prevention training, while the other group did not. Conducted over a 12-year period, the study found that the incidence of knee injuries was significantly lower in the group that received hip-focused injury prevention training compared to the control group. These results suggest that hip-focused injury prevention training is an effective intervention for reducing the risk of injury in basketball players. The study emphasizes the importance of incorporating injury prevention training into athletic programs to mitigate the risk of injuries [28].

Moreover, studies have shown that the COVID-19 pandemic affected injury patterns in the NBA, with changes in the types and rates of injuries. Interventions such as neuromuscular training programs have been identified as effective in reducing injury risks, emphasizing the importance of combining injury prevention with performance-enhancing strategies [29,30].

However, there are several studies on injuries in the NBA and beyond, showing that injuries tend to occur more frequently during games rather than during practice or other team activities. Additionally, injuries are more likely to occur when landing from a jump, cutting, or pivoting [31,32].

The research in [33] investigated the incidence of overuse injuries in young athletes participating in basketball and floorball. The study involved collecting data from the medical records of youth basketball and floorball players in Finland, and the data were analyzed to determine the types and incidence of overuse injuries in the participants. The study's results showed that overuse injuries are common among basketball and floorball players and highlighted the need for prevention programs to be implemented in these sports to reduce the risk of injury. The researchers compared hip biomechanics between responders and non-responders to the program and found significant differences between the two groups. The results suggest that biomechanics play a role in injury risk and that the effectiveness of injury prevention programs may be influenced by individual differences in mechanics [34].

A relevant unsupervised study uses entropy measurements to examine how soccer teams balance organized and disorganized play by analyzing ball passes. It correlates higher entropy with better team performance, demonstrating the significance of adaptable strategies in team sports [34].

Another research study explored the mechanisms, prediction, and prevention of injuries in athletes. The authors present three sharpened and validated tools to reduce the risk of injuries. The study was based on an analysis of the existing literature and aimed to help healthcare professionals and athletes prevent injuries through early detection and targeted interventions. The results of the study highlight the importance of considering knee joint kinematics, player movements, and other factors [35].

The causes of non-contact injury are a multi-factors phenomenon. According to studies with a prospective design, risk factors include the parameters of the knee joint but also proximal anatomical areas, such as the ankle and hip joints. They are categorized as anatomical (e.g., narrow femoral notch width, increased hip anteversion), anthropometric (e.g., increased body weight and Body Mass Index—BMI), and musculoskeletal (such as generalized joint laxity, greater amounts of navicular drop, increased knee joint laxity, decreased range of motion in ankle dorsiflexion, hip internal rotation, impaired hip abduction, and external rotation strength) factors. Most injuries occurred during games, rather than during practices or other team activities [36–40].

Another topic is the evaluation of explosive strength imbalance in professional basketball players. There was a significant difference in explosive strength between the dominant and non-dominant sides of the body, with the dominant side being stronger. The magnitude of the explosive strength imbalance varied among different player positions. The study highlights the importance of considering explosive strength imbalances in the assessment and training of professional basketball players and suggests that a focus on developing symmetry in explosive strength may reduce the risk of injury [21].

1.1. Aims and Objectives

This research aims to elucidate the relationship between professional basketball players' recovery times and the financial impact on their teams, while also considering the influence of player contracts and career duration. The study is designed to measure the extent of this relationship, identify significant trends and outliers, and apply these findings to improve strategic financial planning and contract negotiation practices in sports management. Ultimately, it seeks to provide actionable insights for mitigating the financial risks associated with player injuries and to suggest directions for future research in Sports Analytics and team management. The culmination of these objectives is intended to provide sports teams, managers, and stakeholders with actionable insights into the financial management of player health and performance, thereby enhancing the efficacy of their operations and strategic planning.

Research Questions

The health of athletes is a cornerstone of success in professional basketball, yet injuries and their subsequent recovery periods can have profound implications for a team's

financial health. This study aims to examine the relationship between player recovery times and the financial performance of basketball teams, scrutinizing how these dynamics are influenced by contract values and career lengths through Text Mining, Data Mining, and Association Rule Mining techniques. Furthermore, it explores the potential of these findings to refine financial management and contract negotiation strategies in the sport. Addressing these critical questions, the research provides insights that could help teams mitigate the economic risks associated with player injuries.

1. How can Association Rule Mining be utilized to uncover and understand the relationships between NBA players' salaries, physical attributes, and health conditions and their impact on team performance outcomes, including team losses and recovery times?
2. What is the relationship between basketball players' recovery times and the financial performance of their teams, and how do these factors interact with the players' salaries and career lengths?
3. How can Data Mining and Sports Analytics on player recovery times and financial losses inform the development of strategic financial management and contract negotiation practices in professional basketball teams?

These questions are addressed through the quantitative analysis of player recovery times, financial data, and the examination of outliers, which are then discussed in the context of team management and financial planning in professional sports. The conclusions drawn provide recommendations for future research and practical implications for the management of player health and contracts.

2. Materials and Methods

Our study harnessed a variety of data sources [41,42] to perform a thorough and detailed analysis. The retrieval and preprocessing of data posed multiple challenges, which we effectively managed by integrating diverse data into a supervised data model, prioritizing data quality. Our research methodologically unfolded in three key areas: Data Collection, Data Engineering, and Data Mining. In the Data Collection stage, we engaged in extensive data gathering [43] and web scraping, assembling an exhaustive dataset of NBA player performance, injury records, and salary details from the 2000-01 to 2022-23 seasons. The Data Engineering phase was dedicated to cleaning, organizing, and enhancing these datasets, including natural language processing (NLP) [44], to make the data uniform and analysis ready. During the Data Mining phase, we employed a combination of advanced techniques like anomaly detection using DBSCAN [45], isolation forest [46], Z-score analysis [47], and Association Rule Mining with the Apriori algorithm [48]. These methods were crucial in uncovering key patterns, anomalies, and correlations in the data, leading to valuable insights regarding player injuries in the NBA.

2.1. Data Collection

The raw data we collected were disorganized and varied in format, necessitating a robust methodology to understand and derive meaningful insights from them. Data were gathered using Python scripts [43], and the dataset underwent preprocessing to address information joins based on players' unique IDs and names, and irrelevant information for the research's goals was removed. Additionally, to ensure data quality, an Extract, Transform, and Load (ETL) process standardized and unified the data.

The corresponding data files used for data analysis can be found on GitHub at the following link: <https://github.com/vsarlis/associasionrules> (accessed on 15 June 2024).

This paper presents an exhaustive analysis of sociodemographic, injury, and salary data for all NBA players from the 2000-01 to 2022-23 seasons. It details the data acquisition process, data types, and dataset structures used in our study. Player sociodemographic and injury data were extracted using the `nba_api` [41], a powerful tool that pulls information directly from the NBA's official website and database. These data span 22 seasons, from

2000-01 to 2022-23, encompassing 2296 players and representing the entire player pool over these seasons.

We conducted two separate data extractions from `nba_api`. The first focused on regular season per-game records, while the second targeted playoff records. These datasets, containing detailed statistics for all NBA players within our timeframe, were merged using common `player_id` and `game_date` fields, and stored in a local PostgreSQL database. They also included metadata with reasons for player absences, such as coaching decisions and injuries.

Additionally, we obtained another dataset from ESPN's [42] website through web scraping, which provided detailed information on players' contracts, covering the same 2000-01 to 2022-23 seasons for consistency. These datasets collectively offer a thorough view of NBA player demographics, absences, and contracts, enabling deep insights into player behavior and team performance. The dataset schemas are summarized in Table 1.

Table 1. Aggregated datasets (seasons 2000 to 2023) of sociodemographic, injury, and salary data.

Name (Type)	Dataset (Rows, Columns)
Player sociodemographic info	(781,406, 26)
Injury data (on and off game)	(58,151, 4)
Salary data (signed over seasons)	(7257, 6)

Our study utilizes an over-two-decades-long aggregated dataset to ensure the stability and reliability of our unsupervised learning analysis, capturing comprehensive league-wide trends rather than short-term or individual variations. Shorter time spans and individual case studies, while insightful for specific scenarios, might not provide the extensive data necessary for robust Association Rule Mining and anomaly detection. The broader timespan enriches our dataset with a diversity of incidents and economic factors essential for an accurate analysis of long-term trends in the NBA. This approach aligns with our objective to explore overarching patterns that influence the league.

2.2. Data Engineering

The subsequent vital stage after Data Collection was Data Engineering, which entailed cleansing, structuring, and improving the datasets. This stage was pivotal for preparing the injury and contract datasets for anomaly detection and Association Rule Mining, particularly crucial for datasets rich in text, like injury reports and contract details. Organizing these datasets streamlined their access for text mining and facilitated the conversion of contract details into a comprehensive salary dataset for advanced Data Mining applications.

Text Mining and Categorization in Injury Data: We applied text mining techniques [49] to the injury dataset to extract pertinent information from the textual injury descriptions. A specialized dictionary categorized injuries into predefined groups. For example, an entry like "placed on IL with bone spur in left heel" was categorized under "Heel" → "Foot". The dataset also contained duplicates due to repeated injury reports, which we addressed by retaining only the initial injury instance and its date, flagging subsequent duplicates as 'TRUE' if they occurred within a 15-day window with the same injury type for a player. This method effectively reduced the dataset size while preserving essential information, crucial for analyzing injury data in sports.

Transformation of Contract Data into Salary Data: The goal here was to convert contract data into a more analytically useful salary dataset. Using text mining, we extracted details like contract length and amounts from descriptions, for example, breaking down "signed as a free agent (from Lakers) to a 2-year, USD 22 M contract" into a 2-year length and USD 22 million amount. We also adjusted salary figures for inflation [50] to enable more accurate comparisons, considering the economic context of the year and place of contract signing. In our analysis, the transformation of multi-year NBA player contracts into annual and per-game salary figures involves standardizing the total contract value to an annual

salary by dividing the total amount by the number of years in the contract. For instance, a USD 30 million contract spanning three years is broken down into an annual salary of USD 10 million. To further refine the data, we calculated the salary per game by dividing this annual salary by the total number of games the team plays in a season. For example, with an annual salary of USD 10 million and a team schedule of 82 games, the per-game salary would be approximately USD 121,951. This detailed breakdown allows us to assess the economic implications of player contracts with greater accuracy, providing insights into how salaries impact financial planning and team operations on a game-by-game basis.

2.3. Anomaly Detection Methodology

Our study employed a blend of DBSCAN, isolation forest, and Z-score algorithms to identify anomalies in NBA players' recovery times. This combination enhanced the accuracy and reliability of our analysis.

DBSCAN Algorithm Application: We used the DBSCAN algorithm [51,52] a density-based clustering method, to pinpoint clusters and outliers in the data. Adjusting the radius of neighborhood (ϵ) and the minimum number of points (min_samples) allowed us to fine-tune the algorithm's sensitivity and selectivity. The optimal ϵ was mathematically determined using the nearest-neighbors technique, identifying the point of maximum curvature (knee) in a k-distance graph as the optimal value indicator. The knee point was determined by the point of greatest gradient change in the sorted k-distance plot.

$$\epsilon_{\text{optimal}} = \text{k-distance}(\text{knee point}) \quad (1)$$

Statistical Anomaly Detection via Z-Score: We also incorporated Z-score analysis [53] as a statistical method to detect outliers. A Z-score represents how far a data point deviates from the group's mean in terms of standard deviations. We established a criterion for anomaly detection at a Z-score exceeding 3, which, in a normal distribution, encompasses about 99.7% of the data. This approach identifies outliers in the upper 0.3% tail.

Ensemble Anomaly Detection Strategy: Our study adopted a consensus strategy for anomaly classification. An observation is deemed anomalous if it is identified as such by at least two out of the three methods used. This ensemble approach ensures a more reliable and accurate detection of anomalies.

$$\text{Anomaly Indicator} = 1 \text{ if } \sum_{j=1}^3 M_j(i) \geq 2, \text{ otherwise } 0 \quad (2)$$

In the equation above, 'Anomaly Indicator' denotes the status for the (i th) observation, and $M_j(i)$ represents the detection result from the (j th) method for that observation. This strategy effectively reduces the likelihood of false positives, thus ensuring a more reliable identification of true anomalies.

The algorithms were meticulously calibrated to the unique characteristics of the dataset and applied iteratively to different groupings based on 'Detailed Body Part'. By leveraging the strengths of each algorithm—DBSCAN for local density, isolation forest for broader anomaly separation, and Z-score for statistical deviation—this ensemble method provided a rigorously filtered dataset. These refined data were crucial for the integrity of further analyses and the derivation of association rules.

In this study, we employed a combination of Z-score, DBSCAN, and isolation forest algorithms for anomaly detection. The selection of these algorithms was driven by their complementary strengths and their suitability for the specific characteristics of our dataset. By combining these three algorithms, we leverage their complementary strengths to provide a comprehensive and robust anomaly detection framework. The Z-score method helps identify extreme values, DBSCAN effectively handles clustering and density variations, and isolation forest provides computational efficiency and robustness in high-dimensional spaces. This ensemble approach ensures that our anomaly detection is thorough and reliable.

2.4. Association Rule Mining

Association Rule Mining is a vital technique in Data Mining focused on uncovering intriguing relationships, patterns, and causal structures among item sets in transaction databases or similar data sources. Its primary goal is to identify rules that frequently appear in a dataset.

Theoretical Basis: We employed the Apriori algorithm for extracting frequent item sets and generating association rules. This algorithm operates on the principle that all nonempty subsets of a frequent item set must also be frequent.

The process begins by determining the support for each item in the dataset, identifying frequent single-item sets. It then progresses to creating (k) -item sets by combining frequent $(k - 1)$ -item sets. During pruning, any (k) -item sets containing infrequent $(k - 1)$ -subsets are discarded, adhering to the Apriori principle. The algorithm continues to assess and retain candidate (k) -item sets above a minimum support threshold. This cycle repeats until no further frequent item sets are discovered, leading to the generation of association rules from these item sets.

The foundation of association rules is grounded in support, confidence, and lift metrics [49,54], which evaluate the rules' strength and significance:

- Support measures an item set's frequency or prevalence in the dataset.
- Confidence gauges the probability that a transaction containing the antecedent will also contain the consequent.
- Lift evaluates the performance of a rule beyond the chance co-occurrence of the antecedent and consequent, signifying the rule's predictive power for the consequent.

Implementation Process: In the present study, a methodological framework was meticulously designed to implement association rules for the analysis of NBA player data. The initial phase involved the discretization of continuous variables into categorical groups, adhering to established criteria. Subsequently, the data underwent a transformation via one-hot encoding, resulting in a binary matrix representation. Within this matrix, each row signified an individual player, and each column denoted a specific category, with binary values reflecting the presence or absence of each category for the players. The Apriori algorithm was then rigorously applied to discern frequent item sets from the dataset. Following this, association rules were extrapolated, with an emphasis on calculating the confidence and lift metrics for each rule. The culmination of this process involved a detailed analysis of the rules to extract pertinent insights, particularly focusing on the ramifications of player injuries, recuperation timelines, and their financial repercussions on the respective teams.

Significance in Research: This research endeavor utilized the potent analytical methodology of association rules to decode the intricate patterns and interrelations within the vast datasets encompassing NBA player sociodemographic profiles, injury records, and financial details. The adoption of this quantitative methodological approach was instrumental in addressing core research inquiries, chiefly pertaining to the economic ramifications of player injuries and the duration of recovery on NBA franchises. Through a comprehensive and methodical application of association rules, this study significantly augments the existing literature by elucidating the intricate interplay between player health, sociodemographic attributes, and the economic dynamics of professional sports.

The decision to use Association Rules (AR) Mining in this study is based on their strength in identifying and interpreting complex relationships between variables within large datasets. Unlike other Data Mining approaches, AR algorithms such as Apriori are particularly adept at discovering frequent item sets and generating interpretable rules that provide actionable insights. This capability aligns with our research objectives, which focus on uncovering patterns and correlations between NBA players' attributes, health conditions, and team performance outcomes.

Other Data Mining approaches, such as supervised learning, typically require labeled data and are used for prediction tasks. In contrast, our study aims to explore and understand the underlying relationships within the data without prior labeling, making methods

like AR algorithms more suitable. Additionally, AR algorithms allow us to set specific thresholds for support, confidence, and lift, ensuring that the rules generated are both significant and reliable.

3. Results

In the domain of professional sports, particularly in the NBA, understanding the intricate relationships between player attributes, health, and team performance is crucial. The dataset was derived from a comprehensive collection of NBA player data, which underwent several preprocessing steps, including discretization and one-hot encoding, to facilitate the mining of association rules. The Apriori algorithm was then applied to extract frequent item sets, from which association rules were generated. The final dataset for analysis was filtered based on stringent criteria to ensure the focus was on the most substantial and insightful rules.

3.1. Anomaly Detection and Outliers Identification

This research conducted an extensive analysis of NBA player injuries and their respective recovery periods. Our study utilized a dataset comprising 18,184 entries after preprocessing, within which we identified 189 anomalies, constituting approximately 1.04% of the total cases. These anomalies were characterized by atypical recovery durations, evidently deviating from the normative patterns observed across the dataset. Notably, these anomalous cases mainly involved players who exited the NBA due to injuries and returned to the league after a prolonged absence, resulting in unusually long recovery times. Our analysis encompassed various player-related variables, including position, age, affected body parts, and other applicable metrics.

3.2. Association Rules Mining Results

This study employs Association Rule Mining to explore these dynamics, focusing on the impact of players' salaries, physical characteristics, and health conditions on team losses and recovery times. The dataset has been meticulously filtered to focus on rules with a lift greater than 1.1, confidence greater than 70%, and support greater than 1%, ensuring the analysis is centered around the most significant and reliable associations.

The filtered rules provide compelling insights into how certain player characteristics and conditions are associated with team performance outcomes, as presented in the following tables (Tables 2 and 3):

Table 2. Association rule-based analysis between salary per game (antecedent) and team losses (consequent).

Antecedents	Consequents	Support	Confidence	Lift
Salary_Per_Game_0-150k, WEIGHT_70-100, HEIGHT_200-225	TeamLosses_0-25M, Forward	6.38%	70.22%	1.85761
Salary_Per_Game_0-150k, WEIGHT_70-100, HEIGHT_200-225, RECOVERY_0-10	TeamLosses_0-25M, Forward	3.43%	70.47%	1.86418
Salary_Per_Game_0-150k, AGE_group_AGE_25-30, WEIGHT_70-100, HEIGHT_200-225	TeamLosses_0-25M, Forward	2.54%	72.74%	1.92427
Salary_Per_Game_0-150k, HEIGHT_175-200, WEIGHT_100-130	TeamLosses_0-25M, Forward	1.89%	72.50%	1.91788
Salary_Per_Game_0-150k, WEIGHT_70-100, HEIGHT_200-225, AGE_group_AGE_30-35	TeamLosses_0-25M, Forward	1.15%	70.67%	1.86939
Salary_Per_Game_0-150k, HEIGHT_175-200, WEIGHT_100-130, RECOVERY_0-10	TeamLosses_0-25M, Forward	1.13%	70.27%	1.85890

Table 3. Association rule-based analysis between salary per game (antecedent) and recovery (consequent).

Antecedents	Consequents	Support	Confidence	Lift
Salary_Per_Game_0-150k, General Illness	RECOVERY_0-10	4.13%	76.18%	1.37363
Salary_Per_Game_0-150k, General Illness	RECOVERY_0-10, TeamLosses_0-25M	4.13%	76.18%	1.37363
Salary_Per_Game_0-150k, General Illness, TeamLosses_0-25M	RECOVERY_0-10	4.13%	76.18%	1.37363
Salary_Per_Game_0-150k, General Illness, HEIGHT_200-225, TeamLosses_0-25M	RECOVERY_0-10	2.51%	76.49%	1.37921
Salary_Per_Game_0-150k, HEIGHT_200-225, General Illness	RECOVERY_0-10	2.51%	76.49%	1.37921
Salary_Per_Game_0-150k, HEIGHT_200-225, General Illness	RECOVERY_0-10, TeamLosses_0-25M	2.51%	76.49%	1.37921
Salary_Per_Game_0-150k, General Illness, WEIGHT_100-130, TeamLosses_0-25M	RECOVERY_0-10	2.08%	76.40%	1.37758
Salary_Per_Game_0-150k, WEIGHT_100-130, General Illness	RECOVERY_0-10	2.08%	76.40%	1.37758
Salary_Per_Game_0-150k, WEIGHT_100-130, General Illness	RECOVERY_0-10, TeamLosses_0-25M	2.08%	76.40%	1.37758
Salary_Per_Game_0-150k, General Illness, WEIGHT_70-100, TeamLosses_0-25M	RECOVERY_0-10	2.02%	76.18%	1.37363
Salary_Per_Game_0-150k, WEIGHT_70-100, General Illness	RECOVERY_0-10	2.02%	76.18%	1.37363
Salary_Per_Game_0-150k, WEIGHT_70-100, General Illness	RECOVERY_0-10, TeamLosses_0-25M	2.02%	76.18%	1.37363
Salary_Per_Game_0-150k, HEIGHT_200-225, WEIGHT_100-130, General Illness	RECOVERY_0-10	1.89%	75.11%	1.35429
Salary_Per_Game_0-150k, HEIGHT_200-225, WEIGHT_100-130, General Illness	RECOVERY_0-10, TeamLosses_0-25M	1.89%	75.11%	1.35429
Salary_Per_Game_0-150k, General Illness, Forward	RECOVERY_0-10	1.75%	79.31%	1.43006
Salary_Per_Game_0-150k, General Illness, Forward	RECOVERY_0-10, TeamLosses_0-25M	1.75%	79.31%	1.43006
Salary_Per_Game_0-150k, General Illness, TeamLosses_0-25M, Forward	RECOVERY_0-10	1.75%	79.31%	1.43006
Salary_Per_Game_0-150k, HEIGHT_175-200, General Illness	RECOVERY_0-10	1.59%	75.52%	1.36164
Salary_Per_Game_0-150k, HEIGHT_175-200, General Illness	RECOVERY_0-10, TeamLosses_0-25M	1.59%	75.52%	1.36164
Salary_Per_Game_0-150k, HEIGHT_175-200, General Illness, TeamLosses_0-25M	RECOVERY_0-10	1.59%	75.52%	1.36164
Salary_Per_Game_0-150k, General Illness, Guard	RECOVERY_0-10	1.45%	74.58%	1.34479
Salary_Per_Game_0-150k, General Illness, Guard	RECOVERY_0-10, TeamLosses_0-25M	1.45%	74.58%	1.34479
Salary_Per_Game_0-150k, General Illness, Guard, TeamLosses_0-25M	RECOVERY_0-10	1.45%	74.58%	1.34479
Salary_Per_Game_0-150k, General Illness, HEIGHT_200-225, Forward	RECOVERY_0-10	1.42%	79.27%	1.42930
Salary_Per_Game_0-150k, General Illness, HEIGHT_200-225, Forward	RECOVERY_0-10, TeamLosses_0-25M	1.42%	79.27%	1.42930
Salary_Per_Game_0-150k, HEIGHT_175-200, WEIGHT_70-100, General Illness	RECOVERY_0-10	1.40%	73.71%	1.32916
Salary_Per_Game_0-150k, HEIGHT_175-200, WEIGHT_70-100, General Illness	RECOVERY_0-10, TeamLosses_0-25M	1.40%	73.71%	1.32916

Table 3. Cont.

Antecedents	Consequents	Support	Confidence	Lift
Salary_Per_Game_0-150k, General Illness, WEIGHT_70-100, Guard	RECOVERY_0-10	1.39%	74.20%	1.33797
Salary_Per_Game_0-150k, General Illness, WEIGHT_70-100, Guard	RECOVERY_0-10, TeamLosses_0-25M	1.39%	74.20%	1.33797
Salary_Per_Game_0-150k, AGE_group_AGE_20-25, General Illness	RECOVERY_0-10	1.31%	86.64%	1.56227
Salary_Per_Game_0-150k, AGE_group_AGE_20-25, General Illness	RECOVERY_0-10, TeamLosses_0-25M	1.31%	86.64%	1.56227
Salary_Per_Game_0-150k, AGE_group_AGE_20-25, General Illness, TeamLosses_0-25M	RECOVERY_0-10	1.31%	86.64%	1.56227
Salary_Per_Game_0-150k, HEIGHT_175-200, General Illness, Guard	RECOVERY_0-10	1.26%	74.52%	1.34362
Salary_Per_Game_0-150k, HEIGHT_175-200, General Illness, Guard	RECOVERY_0-10, TeamLosses_0-25M	1.26%	74.52%	1.34362
Salary_Per_Game_0-150k, HEIGHT_175-200, WEIGHT_70-100, Guard, General Illness	RECOVERY_0-10	1.21%	74.33%	1.34032
Salary_Per_Game_0-150k, HEIGHT_175-200, WEIGHT_70-100, Guard, General Illness	RECOVERY_0-10, TeamLosses_0-25M	1.21%	74.33%	1.34032
Salary_Per_Game_0-150k, General Illness, WEIGHT_100-130, Forward	RECOVERY_0-10	1.15%	78.73%	1.41962
Salary_Per_Game_0-150k, General Illness, WEIGHT_100-130, Forward	RECOVERY_0-10, TeamLosses_0-25M	1.15%	78.73%	1.41962
Salary_Per_Game_150k-300k, General Illness	RECOVERY_0-10	1.12%	77.65%	1.40015
Salary_Per_Game_150k-300k, General Illness	RECOVERY_0-10, TeamLosses_0-25M	1.12%	77.65%	1.40015

Player Salary and Team Losses: A recurring theme across the significant rules is the association between players with lower salaries (0–150 k per game) and team losses within the 0–25 M range. This suggests a nuanced relationship where the financial aspect of player contracts might relate to their roles and impact on team performance.

Physical Characteristics and Position: Several rules highlight specific combinations of players' height, weight, and playing position associated with team losses. For instance, forwards within certain height and weight categories appear frequently, indicating potential physical demands and risks associated with this position.

Age and Performance: The age group 25–30 is notably present in the rules, possibly reflecting a critical period in players' careers where their experience and physical condition profoundly impact their performance and, by extension, team outcomes.

Health and Recovery: General health conditions and recovery times are also prominent in the rules, underlining the significance of health management and recovery strategies in minimizing team losses.

The Results section of this study also presents a comprehensive analysis of basketball analytics, focusing on variables such as recovery time, team losses, seasons played by the athlete, and player cost per season in conjunction with inflation-adjusted salaries. Table A1 of Appendix A encapsulates the top 40 players, in descending order based on “Avg Cost per Season”, from an expansive dataset comprising 1365 players spanning the seasons from 2000 to 2023. In order to identify the most dominant features and conclude the top 40 volatile players, we applied filters to players who met the following criteria: an average recovery time of more than 20 days, an average inflated salary greater than USD 15 million, an average cost per season greater than USD 8 million, team losses totaling more than USD 30 million, and a total number of seasons equal to or greater than three.

Recovery Time: Analysis reveals a substantial range in average recovery times amongst players. A case in point is Karl-Anthony Towns, who exhibits an average recovery time

of 109.43 days, indicating potential variability in the duration of injuries or the efficacy of recovery processes across players.

Financial Impact (Team Losses): The data in the ‘Sum of team_losses’ column indicates considerable financial repercussions for teams. Noteworthy examples include John Wall and Kemba Walker, both associated with team losses surpassing the USD 100 million mark. Specifically, John Wall ranks at the pinnacle, with losses amounting to over USD 166 million throughout 10 seasons, which may be reflective of either an exorbitant contract value or substantial periods of inactivity due to injuries.

Seasons Played: The dataset shows variations in the number of seasons played by players, potentially signifying the length of their careers or the duration of their contracts. For instance, Mike Conley, with an extended number of seasons under his belt, may be perceived as a seasoned player whose value has been sustained over time.

Cost per Season: The ‘Avg Cost per Season’ metric sheds light on the economic load each player brings per season. An illustrative example is John Wall, who bears a markedly high average cost per season, suggesting a correlation with the magnitude of his contract or the financial toll of his absence on the team.

Inflation-Adjusted Salaries: The column titled ‘Average of salaries INF’ represents the inflation-adjusted salaries of players, providing an assessment of their earning capacity adjusted over time. This measure offers essential insights into the actualized value of player contracts for a robust financial and contractual analysis. To ensure the comparability of salary figures across different years within our decade-long analysis, all salaries were adjusted for inflation to reflect 2023 values, the most current year in our study period. This adjustment was made using Consumer Price Index (CPI) data, applying the formula $\text{Adjusted Salary} = \text{Reported Salary} \times (\text{CPI in 2023} / \text{CPI in Year of Salary})$. For example, a USD 5 million salary in 2013, with a CPI of 232.957 in 2013 and 288.012 in 2023, would adjust to approximately USD 6.18 million in 2023 dollars. This approach standardizes all financial data to 2023 economic conditions, allowing for an accurate analysis of salary trends and financial impacts over time without the distortions caused by inflation.

3.3. Descriptive and Correlation Analysis

Figure 1 illustrates the relationship between average cost per season and average recovery time for NBA players, categorized into four quadrants. The top left quadrant shows high-cost players with short recovery times, such as S. Marbury and R. Jackson, who provide significant financial value and quick return to play. The top right quadrant features high-cost players with long recovery times, like K.-A. Towns and K. Walker, indicating substantial financial investment but extended absences. The bottom left quadrant includes low-cost players with short recovery times, such as A. Iverson and L. James, who offer good value and high availability. Lastly, the bottom right quadrant highlights low-cost players with long recovery times, like Z. Randolph and D. Rose, who, despite lower costs, have prolonged recovery periods affecting their availability.

Figure 2 demonstrates the relationship between the sum of team losses (in millions of dollars) and the average recovery time for NBA players, categorized into four quadrants. The top left quadrant shows high-loss players with short recovery times, such as S. Marbury and R. Jackson, indicating substantial financial losses for their teams despite quick returns from injuries. The top right quadrant features high-loss players with long recovery times, like K.-A. Towns and K. Walker, representing significant financial burdens due to prolonged absences. The bottom left quadrant includes low-loss players with short recovery times, such as A. Iverson and L. James, who cause minimal financial impact and maintain high availability. Lastly, the bottom right quadrant highlights low-loss players with long recovery times, like Z. Randolph and D. Rose, who, despite lower financial losses, have extended recovery periods affecting their overall contribution to the team.

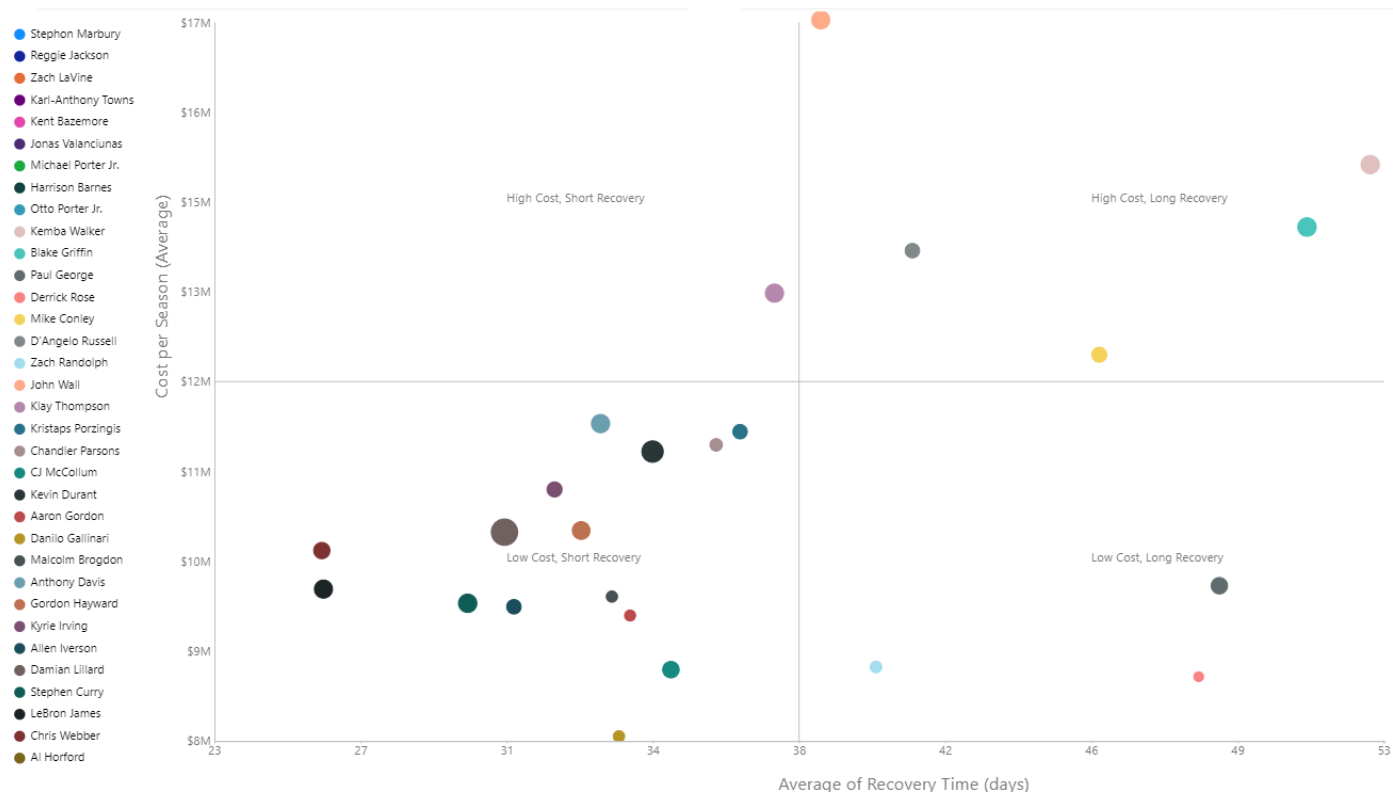


Figure 1. Cost per season vs. recovery time for NBA players.

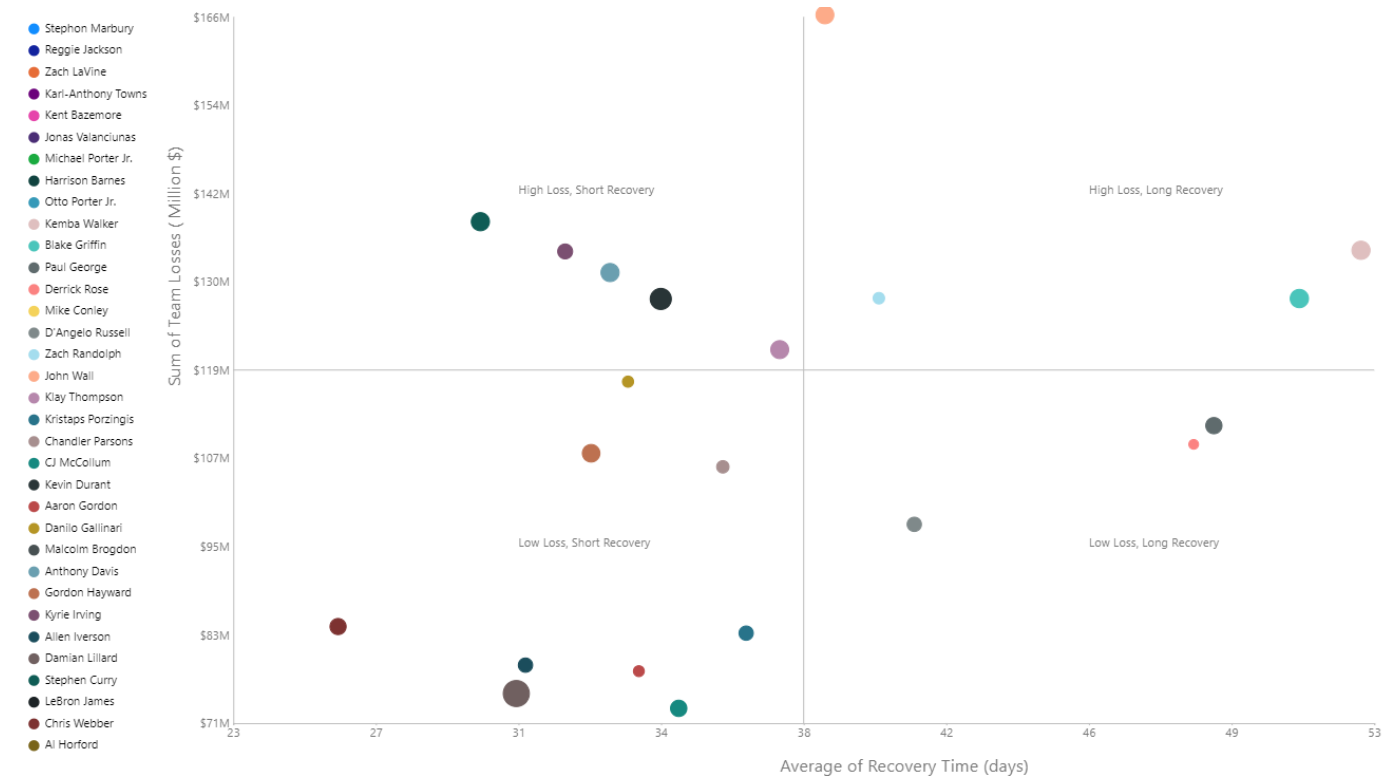


Figure 2. Team losses vs. recovery time for NBA players.

The findings show a pronounced positive correlation (Pearson correlation [55]) between the aggregate of team losses and the average cost per season ($r = 0.86$), inferring that as the financial losses incurred by a team due to a player escalate, the average cost

of that player per season also increases. Furthermore, a moderate correlation is observed between the longevity of players' careers, as denoted by the number of seasons played, and the sum of team losses ($r = 0.71$). This could imply that players with protracted careers may be linked to augmented team losses, conceivably due to elevated salaries or other contributory factors. Additionally, a subtle correlation between the average recovery time and the average cost per season ($r = 0.29$) alludes to a potential relationship where players with extended recovery times may likewise incur higher seasonal costs (Table 4).

Table 4. Pearson correlation coefficient analysis for team losses, recovery time, cost per season, and seasons played.

Variable	Sum of Team Losses	Average Cost per Season	Number of Seasons Played	Average Recovery Time
Sum of Team Losses	1	0.86	0.71	-
Average Cost per Season	0.86	1	-	0.29
Number of Seasons Played	0.71	-	1	-
Average Recovery Time	-	0.29	-	1

The statistical overview indicates that the average recovery time among players is approximately 33.47 days, with a standard deviation of 28.09 days. The mean sum of team losses stands at around USD 15.04 million, accompanied by a sizeable standard deviation of USD 23.43 million, signifying a broad dispersion within the data. On average, players have participated in 4.88 seasons, with the average cost per season amounting to approximately USD 4.62 million, and the average inflation-adjusted salary approximating USD 18.53 million.

The analysis also identifies several outliers, characterized by extreme values in recovery time and team losses. Notable among these are players such as John Wall, Kemba Walker, and Karl-Anthony Towns, who are associated with significantly higher team losses and seasonal costs in comparison to other players in the dataset.

4. Discussion

The findings from this study provide several insights and implications. The strong association between lower-salaried players and team losses suggests teams might need to reconsider how they invest in player contracts, perhaps focusing more on health and performance potential rather than just market value. The recurrent appearance of specific positions in association with team losses and physical attributes suggests that certain roles might be inherently riskier or more demanding. Teams might use this information to tailor training and health management programs for these positions. The prominence of the 25–30 age group in significant rules suggests this might be a pivotal age range for players.

4.1. Financial Repercussions of Player Recovery

The analysis underscores the significant financial repercussions of player recovery times on sports teams, with notable variability among individual cases. This variability can be attributed to factors such as the severity and nature of injuries, pivotal roles within teams, and overall marketability, which directly influences contract valuations. A discernible correlation exists between the tenure of players—evidenced by the number of seasons played—and the financial losses incurred by teams, suggesting that players with extensive careers, who typically command higher salaries, may exacerbate financial risks during periods of injury-induced absence.

4.2. Comparative Evaluation of Players

The dataset facilitates a comparative evaluation of players, incorporating a spectrum of financial and performance metrics. Players with prolonged average recovery times might be associated with escalated team losses, likely due to their integral role and the consequential impact of their absence on team performance and financial health. Disparities in the cost per season are attributable to determinants such as market value, on-field performance, and outcomes of contract negotiations. Adjustments for inflation in player salaries offer invaluable insights for longitudinal financial assessments, allowing for an equitable comparison across different timeframes, normalizing for the economic environment's temporal shifts.

4.3. Correlation and Outliers Analysis

The positive correlation between team losses and per-season costs indicates the higher monetary value ascribed to elite players who command larger salaries. The duration of a player's career appears to contribute to team losses, highlighting a potential financial liability for teams attributed to seasoned players with higher remuneration and a notable track record. A mild correlation between recovery time and seasonal costs suggests that extended absences of key players incur higher costs. An examination of outliers reveals cases that substantially diverge from the norm, suggesting exceptional circumstances such as long-term injuries, exceptionally high-value contracts, or other variables.

4.4. Economic Impact of Injuries

Financial losses are quantified as the economic impacts presented by NBA teams when players are unable to participate due to injuries. These losses are assessed in terms of decreased team performance and the resulting effects on various revenue streams such as ticket sales, merchandise sales, and broadcasting rights. While individual players may have insurance that covers personal income loss, our analysis focused on broader implications for the team's finances, including direct revenue losses and potential declines in the team's market value due to diminished fan engagement and competitive disadvantage. Our research highlights the significant economic stakes involved when key players are sidelined by injuries, setting the groundwork for future studies to explore the complex distribution of financial impacts among all stakeholders in professional sports leagues like the NBA.

4.5. Threats to Validity and Limitations

This study, while extensive in its quantitative analysis of the association between player injuries, sociodemographic data, and financial aspects within the NBA, acknowledges several limitations and threats to validity:

1. **Scope of Data:** The research primarily utilizes quantitative data, focusing on injuries, salary, and sociodemographic factors. While these provide valuable insights, they do not encompass the full spectrum of factors influencing player performance and team dynamics, such as psychological impact, player popularity, and personal narratives.
2. **Quantitative Focus:** The inherently quantitative nature of this study limits its ability to account for qualitative factors, such as team bonding, player morale, and the personal significance of player roles. These elements, while crucial to a comprehensive understanding of sports dynamics, require different methodological approaches and data sources that were outside the scope of this study.
3. **Injury Severity and Type:** The data include a broad range of injury types and severities, treated uniformly in the analysis. This approach does not differentiate between the varying impacts of different injury severities (e.g., a minor muscle stretch vs. a severe ACL tear) and their long-term consequences on player careers and team performance.
4. **Historical and Cultural Context:** The analysis does not account for the historical performance of teams or the presence of MVP players, both of which can significantly influence team strategies and player valuations. The NBA's star system and its influence on player contracts and recovery decisions are not explicitly modeled in this study.

5. **Data Availability and Reliability:** Reliance on publicly available data sources may introduce biases, especially if such data are incomplete or inconsistently reported. The accuracy and completeness of the data from external sources like ESPN and nba.com cannot be independently verified, which might affect the reliability of the findings.
6. **Generalizability:** While the findings provide significant insights into patterns over the last decade, they may not necessarily be generalizable to other sports or even to future seasons within the NBA, as player dynamics, team strategies, and economic conditions evolve.

By acknowledging these limitations, this study underscores the need for further research integrating both quantitative and qualitative data to paint a more complete picture of the complex interactions within professional sports. Future research should aim to incorporate mixed methods approaches to explore the nuanced effects of injuries, beyond contractual and immediate economic impacts.

Additionally, another limitation of this study is the exclusion of specific data on advancements in health technology and sports equipment (e.g., wearables or sensors data), which have significantly evolved from 2000 to 2023. These advancements likely influenced the rehabilitation times and overall recovery outcomes for injured players. Our reliance on publicly available data restricts our ability to incorporate detailed information on medical treatments and the technology used in sports gear, such as basketball shoes, which could impact injury recovery processes. This omission might limit the precision of our analysis concerning the speed and effectiveness of player recovery over the period analyzed. Future studies could benefit from collaboration with NBA teams or medical staff to access more detailed and proprietary data on these variables, potentially providing a more comprehensive understanding of the factors influencing recovery timelines and their subsequent economic impact on teams.

5. Conclusions

This study has illuminated several intriguing associations between players' salaries, physical attributes, health conditions, and team performance outcomes in the NBA. The insights derived underscore the importance of strategic financial planning, health and recovery management, and tailored approaches to player training and care. Specific patterns of injuries and recovery times significantly impact financial strategies in the NBA, revealing unexpected correlations between player positions, injury types, and financial outcomes. Tailored management strategies focusing on the dynamics of player injuries and salary structures can enhance both short-term recovery management and long-term financial planning. Applying these insights, the sports industry could optimize operational strategies, adding substantial value to the field of Sports Analytics.

The significance of recovery times in the rules implies that teams should invest more in medical and training facilities that specialize in faster and more effective recovery methods. Teams might focus on different strategies for player development, management, and rotation during this period.

The comprehensive analysis of NBA player data reveals that recovery times have a dual impact on both team performance and financial outcomes. This finding underscores the critical importance of strategic financial planning in sports management. Effective insurance policies, astute contract negotiations, and judicious salary cap management are essential to mitigate the substantial economic consequences associated with player injuries and recovery periods. Aligning player costs with their anticipated performance and availability is crucial for teams to balance performance demands with financial stability.

This study highlights the significant economic repercussions stemming from player availability and performance, emphasizing the necessity for comprehensive injury management and prevention strategies. By focusing on these strategies, teams can reduce financial risks and enhance overall team performance. The pronounced variability in the impact of individual players on their teams points to the need for nuanced management approaches that cater to both player health and contractual agreements. This individualized approach

ensures that teams can better manage the financial and performance-related aspects of player injuries.

While the observed correlations between various factors do not establish causality, they provide valuable insights for deeper exploration. Teams are encouraged to consider these correlations during contract negotiations and roster management, to alleviate financial risks. Additionally, the study calls for further research to investigate the underlying reasons for the observed outliers and discern whether strategic decisions by teams can mitigate the financial detriments associated with player recovery times. This approach emphasizes the sophisticated nature of Sports Analytics and the indispensable value of individualized player evaluations in financial and team management strategies.

Furthermore, the application of Data Mining techniques, specifically anomaly detection and Association Rule Mining, has proven effective in exploring the intricate relationships between NBA player injuries, sociodemographics, positions, and salaries. These findings do not only provide valuable insights for team management and financial planning within professional basketball, but also highlight the adaptability of the methodological framework to other datasets. This adaptability underscores the potential for broader application in the field of Sports Analytics, paving the way for future studies to build on these insights and further refine player management strategies.

Future Work

Future research could expand this work by integrating additional data types, exploring predictive models, and comparing these findings with other leagues and sports to gain broader insights into the dynamics of player performance and team management. It could encompass the examination of medical intervention efficacy, the deployment of injury prevention programs, and the application of analytical tools to enhance the predictability and management of player absences.

Enhancements to the study could be considered through work on the gamification of sports software to boost engagement and regarding the accuracy of wearable heart rate monitors. These additions could provide more precise data and make the recovery process more interactive, enriching the analysis and offering deeper insights into the management of sports injuries [56–60].

While the methodologies used in this study, particularly anomaly detection and Association Rule Mining, focus on the associations between NBA player injuries, sociodemographics, recovery times, positions, and salaries, extending these techniques to sports betting necessitates integrating bet-related data, such as odds and historical outcomes. Future research could explore this by incorporating betting data alongside player performance metrics to predict fluctuations in betting odds and outcomes. Such an approach would not only broaden the applicability of our current methodologies, but also contribute novel insights to the predictive analytics used in sports betting, provided ethical and legal considerations are carefully managed.

Acknowledging the critical role of mental qualities, team dynamics, and unique physical attributes in basketball, subsequent studies could integrate these aspects to provide a holistic view of player performance and valuation. Additionally, exploring the influence of management strategies and contractual negotiations could offer deeper insights into the economic and operational challenges faced by teams. By incorporating these complex and multifaceted elements, future work can enhance the understanding of what drives success and resilience in professional basketball, paving the way for comprehensive Sports Analytics research.

Additionally, performance potential refers to the anticipated capacity of a player to positively impact team outcomes, incorporating an evaluation of their health history and overall durability alongside current skill levels. This concept extends beyond immediate market value, which typically focuses on metrics like popularity and statistical output, to include a player's ability to sustain or enhance performance over time. We advocate for teams to integrate these health and performance assessments into their player valuation

strategies, aiming to balance short-term gains with long-term benefits. This approach encourages a comprehensive evaluation of a player’s worth, considering both immediate contributions and future reliability.

NBA teams are required to carry insurance for their highest-paid players, covering salaries in the event of long-term injuries. This coverage can mitigate the financial impacts when high-profile players are sidelined. Additionally, the presence of these insurance policies may influence recovery strategies, potentially extending recovery times, as there is a reduced financial urgency to rush players back to play. When analyzing financial losses, it should be clarified that a missed game does not equate directly to a financial loss based purely on the player’s salary. Furthermore, the role of reserve players is to be considered; their impact and contribution become crucial when starters are injured. The fact that injured players do not play all 48 min of a game, and the performance variability among substitutes, can significantly affect both team performance and financial resilience. Future research could benefit from a more detailed exploration of these dynamics, enhancing the understanding of team strategies and policies in managing player injuries.

Author Contributions: Conceptualization: V.S.; methodology: V.S. and G.P.; software: V.S. and G.P.; validation: C.T., V.S. and G.P.; formal analysis: V.S. and G.P.; investigation: V.S. and G.P.; resources: C.T., V.S. and G.P.; data curation: V.S. and G.P.; writing—original draft preparation: V.S.; writing—review and editing: C.T., V.S. and G.P.; visualization: V.S.; supervision: C.T.; project administration: C.T. and V.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Data are partly available within the manuscript.

Conflicts of Interest: The authors declare no potential conflicts of interest with respect to the research, authorship, and/or publication of this article. This manuscript complies with ethical standards.

Appendix A

Table A1. Descriptive and prescriptive analysis of top 40 players in descending order of average cost in dollars per season (2000 up to 2023).

Player Name	Average Recovery Time (Days)	Sum of Team Losses (USD)	Seasons Played	Avg Cost per Season (USD)	Average of Salaries INF (USD)
John Wall	38.56	USD 166,279,623.8	10	USD 16,627,962.4	USD 30,528,031.7
Kemba Walker	52.65	USD 134,609,844.2	9	USD 14,956,649.4	USD 31,063,692.4
Karl-Anthony Towns	109.43	USD 72,397,466.5	5	USD 14,479,493.3	USD 38,331,940.6
D’Angelo Russell	39.35	USD 98,661,341.6	7	USD 14,094,477.4	USD 24,687,237.2
Michael Porter Jr.	61.14	USD 55,012,949.9	4	USD 13,753,237.5	USD 28,583,903.4
Klay Thompson	37.37	USD 121,245,593.7	9	USD 13,471,732.6	USD 30,738,761.7
Blake Griffin	48.94	USD 132,430,957.2	10	USD 13,243,095.7	USD 30,916,497.2
Stephon Marbury	85.45	USD 78,512,384.0	6	USD 13,085,397.3	USD 23,979,243.4
Mike Conley	44.78	USD 178,982,363.7	14	USD 12,784,454.6	USD 25,075,153.9
Kristaps Porzingis	35.67	USD 84,042,695.4	7	USD 12,006,099.3	USD 24,750,948.3
Chandler Parsons	35.18	USD 105,500,608.0	9	USD 11,722,289.8	USD 20,717,144.6
Kevin Durant	34.24	USD 128,060,945.1	11	USD 11,641,904.1	USD 36,223,309.4
Otto Porter Jr.	50.28	USD 103,053,294.5	9	USD 11,450,366.1	USD 22,621,621.7

Table A1. Cont.

Player Name	Average Recovery Time (Days)	Sum of Team Losses (USD)	Seasons Played	Avg Cost per Season (USD)	Average of Salaries INF (USD)
Kyrie Irving	30.98	USD 136,717,270.5	12	USD 11,393,105.9	USD 25,404,397.4
Anthony Davis	30.48	USD 124,918,399.2	11	USD 11,356,218.1	USD 31,339,185.6
Damian Lillard	30.44	USD 74,977,418.3	7	USD 10,711,059.8	USD 44,654,569.9
Chris Webber	25.76	USD 83,987,646.2	8	USD 10,498,455.8	USD 27,522,807.0
Al Horford	24.67	USD 113,760,320.2	11	USD 10,341,847.3	USD 24,707,601.1
Malcolm Brogdon	32.47	USD 71,505,667.1	7	USD 10,215,095.3	USD 19,229,051.3
LeBron James	25.47	USD 181,389,033.4	18	USD 10,077,168.5	USD 30,900,149.5
Allen Iverson	28.82	USD 80,340,248.5	8	USD 10,042,531.1	USD 23,558,509.6
Aaron Gordon	32.31	USD 79,340,583.2	8	USD 9,917,572.9	USD 18,353,092.2
Gordon Hayward	30.97	USD 108,591,389.1	11	USD 9,871,944.5	USD 29,488,521.7
Zach LaVine	74.69	USD 77,476,670.0	8	USD 9,684,583.7	USD 18,262,790.9
Stephen Curry	29.32	USD 135,377,761.1	14	USD 9,669,840.1	USD 31,341,295.4
Harrison Barnes	58.55	USD 67,011,225.4	7	USD 9,573,032.2	USD 21,008,659.8
Paul George	40.46	USD 94,681,149.8	10	USD 9,468,115.0	USD 28,607,919.8
CJ McCollum	34.71	USD 72,983,930.1	8	USD 9,122,991.3	USD 27,937,269.9
Zach Randolph	42.02	USD 124,624,489.1	14	USD 8,901,749.2	USD 19,815,578.3
Jonas Valanciunas	64.27	USD 94,984,386.3	11	USD 8,634,944.2	USD 16,257,780.0
Danilo Gallinari	32.28	USD 117,793,165.6	14	USD 8,413,797.5	USD 18,863,024.5
Brandon Ingram	48.19	USD 50,359,606.3	6	USD 8,393,267.7	USD 21,900,605.9
Shaquille O'Neal	26.25	USD 92,232,810.8	11	USD 8,384,801.0	USD 28,909,107.6
Reggie Jackson	92	USD 66,340,792.2	8	USD 8,292,599.0	USD 17,127,992.0
Kent Bazemore	66.08	USD 57,472,889.9	7	USD 8,210,412.8	USD 17,355,423.5
Victor Oladipo	44.54	USD 82,093,002.1	10	USD 8,209,300.2	USD 16,626,329.9
Andrew Bynum	68.94	USD 57,153,545.2	7	USD 8,164,792.2	USD 17,455,602.4
Amar'e Stoudemire	25.94	USD 81,522,215.3	10	USD 8,152,221.5	USD 21,205,552.6
Derrick Rose	46.98	USD 112,776,529.1	14	USD 8,055,466.4	USD 16,124,754.6
Kevin Love	23.4	USD 120,642,532.7	15	USD 8,042,835.5	USD 23,875,334.4

References

1. Sarlis, V.; Tjortjis, C. Sports Analytics—Evaluation of Basketball Players and Team Performance. *Inf. Syst.* **2020**, *19*, 19. [[CrossRef](#)]
2. Van Haaren, J.; Zimmermann, A.; Renkens, J.; Broeck, G.V.D.; Beéck, T.O.D.; Meert, W.; Davis, J. *Machine Learning and Data Mining for Sports Analytics*; Springer: Berlin/Heidelberg, Germany, 2020; ISBN 9783030649111.
3. Sarlis, V.; Chatziilias, V.; Tjortjis, C.; Mandalidis, D. A Data Science Approach Analysing the Impact of Injuries on Basketball Player and Team Performance. *Inf. Syst.* **2021**, *99*, 101750. [[CrossRef](#)]
4. Sarlis, V.; Papageorgiou, G.; Tjortjis, C. Sports Analytics and Text Mining NBA Data to Assess Recovery from Injuries and Their Economic Impact. *Computers* **2023**, *12*, 261. [[CrossRef](#)]
5. Morgulev, E.; Azar, O.H.; Lidor, R. Sports Analytics and the Big-Data Era. *Int. J. Data Sci. Anal.* **2018**, *5*, 213–222. [[CrossRef](#)]
6. Kwartler, T. *Sports Analytics in Practice with R*; John Wiley & Sons: New York, NY, USA, 2022; ISBN 9781119598077.
7. Short, S.M.; MacDonald, C.W.; Strack, D. Hip and Groin Injury Prevention in Elite Athletes and Team Sport—Current Challenges and Opportunities. *Int. J. Sports Phys. Ther.* **2021**, *16*, 270–281. [[CrossRef](#)] [[PubMed](#)]
8. Kaplan, S.; Ramamoorthy, V.; Gupte, C.; Sagar, A.; Premkumar, D.; Wilbur, J.; Zilberman, D.; Chair, R. The Economic Impact of NBA Superstars: Evidence from Missed Games Using Ticket Microdata from a Secondary Marketplace. In Proceedings of the MIT Sloan Sports Analytics Conference, Boston, MA, USA, 1–2 March 2019; pp. 1–29.

9. Kim, S.W.; Shahin, S.; Ng, H.K.T.; Kim, J. Binary Segmentation Procedures Using the Bivariate Binomial Distribution for Detecting Streakiness in Sports Data. *Comput. Stat.* **2021**, *36*, 1821–1843. [\[CrossRef\]](#)
10. Kaplan, S. The Economic Value of Popularity: Evidence from Superstars in the National Basketball Association. *SSRN Electron. J.* **2020**, *50*, 3543686. [\[CrossRef\]](#)
11. Sarlis, V.; Tjortjis, C. Sports Analytics: Data Mining to Uncover NBA Player Position, Age, and Injuries Impact on Performance and Economics. *Information* **2024**, *15*, 242. [\[CrossRef\]](#)
12. Sarlis, V.; Papageorgiou, G.; Tjortjis, C. Injury Patterns and Impact on Performance in the NBA League Using Sports Analytics. *Computation* **2024**, *12*, 36. [\[CrossRef\]](#)
13. Papageorgiou, G.; Sarlis, V.; Tjortjis, C. Unsupervised Learning in NBA Injury Recovery: Advanced Data Mining to Decode Recovery Durations and Economic Impacts. *Information* **2024**, *15*, 61. [\[CrossRef\]](#)
14. Sundemo, D.; Hamrin Senorski, E.; Karlsson, L.; Horvath, A.; Juul-Kristensen, B.; Karlsson, J.; Ayeni, O.R.; Samuelsson, K. Generalised Joint Hypermobility Increases ACL Injury Risk and Is Associated with Inferior Outcome after ACL Reconstruction: A Systematic Review. *BMJ Open Sport. Exerc. Med.* **2019**, *5*, e000620. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Cole, B.; Arundale, A.J.H.; Bytowski, J.; Amendola, A. *Basketball Sports Medicine and Science*; Springer: Berlin/Heidelberg, Germany, 2020; ISBN 9783662610695.
16. Mateos Conde, J.; Cabero Morán, M.T.; Moreno Pascual, C. Prospective Epidemiological Study of Basketball Injuries during One Competitive Season in Professional and Amateur Spanish Basketball. *Physician Sportsmed.* **2022**, *50*, 349–358. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Harris, J.D.; Erickson, B.J.; Bach, B.R.; Abrams, G.D.; Cvetanovich, G.L.; Forsythe, B.; McCormick, F.M.; Gupta, A.K.; Cole, B.J. Return-to-Sport and Performance After Anterior Cruciate Ligament Reconstruction in National Basketball Association Players. *Sports Health* **2013**, *5*, 562–568. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Myer, G.D. Diagnostic Differences for Anterior Knee Pain between Sexes in Adolescent Basketball Players. *Bone* **2011**, *23*, 1–7. [\[CrossRef\]](#)
19. Trojan, T.H.; Cracco, A.; Hall, M.; Mascaro, M.; Aerni, G.; Ragle, R. Basketball Injuries: Caring for a Basketball Team. *Curr. Sports Med. Rep.* **2013**, *12*, 321–328. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Khan, M.; Ekhtiari, S.; Burrus, T.; Madden, K.; Rogowski, J.P.; Bedi, A. Impact of Knee Injuries on Post-Retirement Pain and Quality of Life: A Cross-Sectional Survey of Professional Basketball Players. *HSS J.* **2020**, *16*, 327–332. [\[CrossRef\]](#) [\[PubMed\]](#)
21. Schiltz, M.; Lehance, C.; Maquet, D.; Bury, T.; Crielaard, J.M.; Croisier, J.L. Explosive Strength Imbalances in Professional Basketball Players. *J. Athl. Train.* **2009**, *44*, 39–47. [\[CrossRef\]](#) [\[PubMed\]](#)
22. Jauhiainen, S.; Kauppi, J.P.; Krosshaug, T.; Bahr, R.; Bartsch, J.; Äyrämö, S. Predicting ACL Injury Using Machine Learning on Data from an Extensive Screening Test Battery of 880 Female Elite Athletes. *Am. J. Sports Med.* **2022**, *50*, 2917–2924. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Nakase, J.; Kitaoka, K.; Shima, Y.; Oshima, T.; Sakurai, G.; Tsuchiya, H. Risk Factors for Noncontact Anterior Cruciate Ligament Injury in Female High School Basketball and Handball Players: A Prospective 3-Year Cohort Study. *Asia Pac. J. Sports Med. Arthrosc. Rehabil. Technol.* **2020**, *22*, 34–38. [\[CrossRef\]](#)
24. Kaeding, C.C.; Léger-st-jean, B.; Magnussen, R.A. Epidemiology and Diagnosis of Anterior Cruciate Ligament Injuries. *Clin. Sports Med.* **2017**, *36*, 1–8. [\[CrossRef\]](#)
25. Shea, K.G.; Grimm, N.L.; Ewing, C.K.; Aoki, S.K. Youth Sports Anterior Cruciate Ligament and Knee Injury Epidemiology: Who Is Getting Injured? In What Sports? When? *Clin. Sports Med.* **2011**, *30*, 691–706. [\[CrossRef\]](#) [\[PubMed\]](#)
26. Vaudreuil, N.J.; van Eck, C.F.; Lombardo, S.J.; Kharrazi, F.D. Economic and Performance Impact of Anterior Cruciate Ligament Injury in National Basketball Association Players. *Orthop. J. Sports Med.* **2021**, *9*, 1–6. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Johnson, W.R.; Mian, A.; Lloyd, D.G.; Alderson, J.A. On-Field Player Workload Exposure and Knee Injury Risk Monitoring via Deep Learning. *J. Biomech.* **2019**, *93*, 185–193. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Omi, Y.; Sugimoto, D.; Kuriyama, S.; Kurihara, T.; Miyamoto, K.; Yun, S.; Kawashima, T.; Hirose, N. Effect of Hip-Focused Injury Prevention Training for Anterior Cruciate Ligament Injury Reduction in Female Basketball Players: A 12-Year Prospective Intervention Study. *Am. J. Sports Med.* **2018**, *46*, 852–861. [\[CrossRef\]](#)
29. Torres-Ronda, L.; Gámez, I.; Robertson, S.; Fernández, J. Epidemiology and Injury Trends in the National Basketball Association: Pre- and PerCOVID-19 (2017–2021). *PLoS ONE* **2022**, *17*, 0263354. [\[CrossRef\]](#) [\[PubMed\]](#)
30. Petushek, E.J.; Sugimoto, D.; Stoolmiller, M.; Smith, G.; Myer, G.D. Evidence-Based Best-Practice Guidelines for Preventing Anterior Cruciate Ligament Injuries in Young Female Athletes: A Systematic Review and Meta-Analysis. *Am. J. Sports Med.* **2019**, *47*, 1744–1753. [\[CrossRef\]](#) [\[PubMed\]](#)
31. McKeag, D.B. *Handbook of Sports Medicine and Science*; CRC Press: Boca Raton, FL, USA, 2003; ISBN 9781119130536.
32. Cumps, E.; Verhagen, E.; Meeusen, R. Prospective Epidemiological Study of Basketball Injuries during One Competitive Season: Ankle Sprains and Overuse Knee Injuries. *J. Sports Sci. Med.* **2007**, *6*, 204–211. [\[PubMed\]](#)
33. Leppänen, M.; Pasanen, K.; Kujala, U.M.; Parkkari, J. Overuse Injuries in Youth Basketball and Floorball. *Open Access J. Sports Med.* **2015**, *6*, 173–179. [\[CrossRef\]](#) [\[PubMed\]](#)
34. Neuman, Y.; Israeli, N.; Vilenchik, D.; Cohen, Y. The Adaptive Behavior of a Soccer Team: An Entropy-Based Analysis. *Entropy* **2018**, *20*, 758. [\[CrossRef\]](#)
35. Hewett, T.E.; Myer, G.D.; Ford, K.R.; Paterno, M.V.; Quatman, C.E. Mechanisms, Prediction, and Prevention of ACL Injuries: Cut Risk with Three Sharpened and Validated Tools. *J. Orthop. Res.* **2016**, *34*, 1843–1855. [\[CrossRef\]](#)

36. Amraee, D.; Alizadeh, M.H.; Minoonejhad, H.; Razi, M.; Amraee, G.H. Predictor Factors for Lower Extremity Malalignment and Non-Contact Anterior Cruciate Ligament Injuries in Male Athletes. *Knee Surg. Sports Traumatol. Arthrosc.* **2017**, *25*, 1625–1631. [\[CrossRef\]](#)
37. Khayambashi, K.; Ghoddosi, N.; Straub, R.K.; Powers, C.M. Hip Muscle Strength Predicts Noncontact Anterior Cruciate Ligament Injury in Male and Female Athletes: A Prospective Study. *Am. J. Sports Med.* **2016**, *44*, 355–361. [\[CrossRef\]](#)
38. Evans, K.N.; Kilcoyne, K.G.; Dickens, J.F.; Rue, J.P.; Giuliani, J.; Gwinn, D.; Wilckens, J.H. Predisposing Risk Factors for Non-Contact ACL Injuries in Military Subjects. *Knee Surg. Sports Traumatol. Arthrosc.* **2012**, *20*, 1554–1559. [\[CrossRef\]](#)
39. Uhorchak, J.M.; Scoville, C.R.; Williams, G.N.; Arciero, R.A.; St. Pierre, P.; Taylor, D.C. Risk Factors Associated with Noncontact Injury of the Anterior Cruciate Ligament. A Prospective Four-Year Evaluation of 859 West Point Cadets. *Am. J. Sports Med.* **2003**, *31*, 831–842. [\[CrossRef\]](#)
40. Woodford-Rogers, B.; Cyphert, L.; Denegar, C.R. Risk Factors for Anterior Cruciate Ligament Injury in High School and College Athletes. *J. Athl. Train.* **1994**, *29*, 343–346.
41. NBA.com. Available online: <https://stats.nba.com> (accessed on 1 November 2023).
42. ESPN NBA Stats. Available online: <https://www.espn.com/nba/stats> (accessed on 10 January 2023).
43. Mitchell, R. *Web Scraping with Python: Collecting Data from the Modern Web*; O'Reilly Media: Sebastopol, CA, USA, 2015; ISBN 9781491985564.
44. Ratner, B. *Statistical and Machine-Learning Data Mining: Techniques for Better Predictive Modeling and Analysis of Big Data*, 3rd ed.; CRC Press: Boca Raton, FL, USA, 2017; ISBN 9781498797610.
45. Witten, I.; Frank, E.; Hall, M. *Data Mining Practical Machine Learning Tools and Techniques*, 3rd ed.; Elsevier Inc.: Amsterdam, The Netherlands, 2011; ISBN 9780123748560.
46. Géron, A. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*; O'Reilly Media: Sebastopol, CA, USA, 2019; ISBN 9781492032649.
47. Winston, W. *Mathletics: How Gamblers, Managers, and Sports Enthusiasts Use Mathematics in Baseball, Basketball, and Football*; Princeton University Press: Princeton, NJ, USA, 2009; ISBN 9780691139135.
48. Brefeld, U.; Davis, J.; Goebel, R.; Van Haaren, J. *Machine Learning and Data Mining in Sports Analytics*; Springer: Cham, Switzerland, 2018; ISBN 9783030172732.
49. Dunham, M. *Data Mining—Introductory and Advanced Topics*; Prentice Hall: Upper Saddle River, NJ, USA, 2003; ISBN 0130888923.
50. Coinnews Media Group. US Inflation Calculator. Available online: <https://www.usinflationcalculator.com/inflation/current-inflation-rates/> (accessed on 1 September 2023).
51. Burkov, A. *The Hundred-Page Machine Learning Book*; Andriy Burkov: Quebec City, QC, Canada, 2019.
52. Ullmann, T.; Beer, A.; Hünemörder, M.; Seidl, T.; Boulesteix, A.L. Over-Optimistic Evaluation and Reporting of Novel Cluster Algorithms: An Illustrative Study. *Adv. Data Anal. Classif.* **2022**, *17*, 211–238. [\[CrossRef\]](#)
53. Berthold, M.R. *Journeys to Data Mining*; Gaber, M.M.G.M., Ed.; Springer: New York, NY, USA, 2012; ISBN 978-3-642-28046-7.
54. Tan, P.; Steinbach, M. *Introduction to Data Mining*; Pearson Addison-Wesley: Boston, MA, USA, 2006; ISBN 0321321367.
55. Lockie, R.G.; Beljic, A.; Ducheny, S.C.; Kammerer, J.D.; Dawes, J.J. Relationships between Playing Time and Selected Nba Combine Test Performance in Division i Mid-Major Basketball Players. *Int. J. Exerc. Sci.* **2020**, *13*, 583–596.
56. Ge, Z.; Prasad, P.W.C.; Costadopoulos, N.; Alsadoon, A.; Singh, A.K.; Elchouemi, A. Evaluating the Accuracy of Wearable Heart Rate Monitors. In Proceedings of the Proceedings—2016 International Conference on Advances in Computing, Communication and Automation (Fall), ICACCA 2016, Bareilly, India, 30 September–1 October 2016.
57. Alsadoon, A.; Al-Naymat, G.; Jerew, O.D. An Architectural Framework of Elderly Healthcare Monitoring and Tracking through Wearable Sensor Technologies. *Multimed. Tools Appl.* **2024**, *83*, 45–62. [\[CrossRef\]](#)
58. Worsey, M.T.O.; Espinosa, H.G.; Shepherd, J.B.; Thiel, D.V. Inertial Sensors for Performance Analysis in Combat Sports: A Systematic Review. *Sports* **2019**, *7*, 28. [\[CrossRef\]](#)
59. Worsey, M.T.O.; Jones, B.S.; Cervantes, A.; Chauvet, S.P.; Thiel, D.V.; Espinosa, H.G. Assessment of Head Impacts and Muscle Activity in Soccer Using a T3 Inertial Sensor and a Portable Electromyography (EMG) System: A Preliminary Study. *Electronics* **2020**, *9*, 834. [\[CrossRef\]](#)
60. Giannakis, K.; Chorianopoulos, K.; Jaccheri, L. User Requirements for Gamifying Sports Software. In Proceedings of the 2013 3rd International Workshop on Games and Software Engineering: Engineering Computer Games to Enable Positive, Progressive Change, GAS 2013, San Francisco, CA, USA, 18 May 2013; pp. 22–26.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.