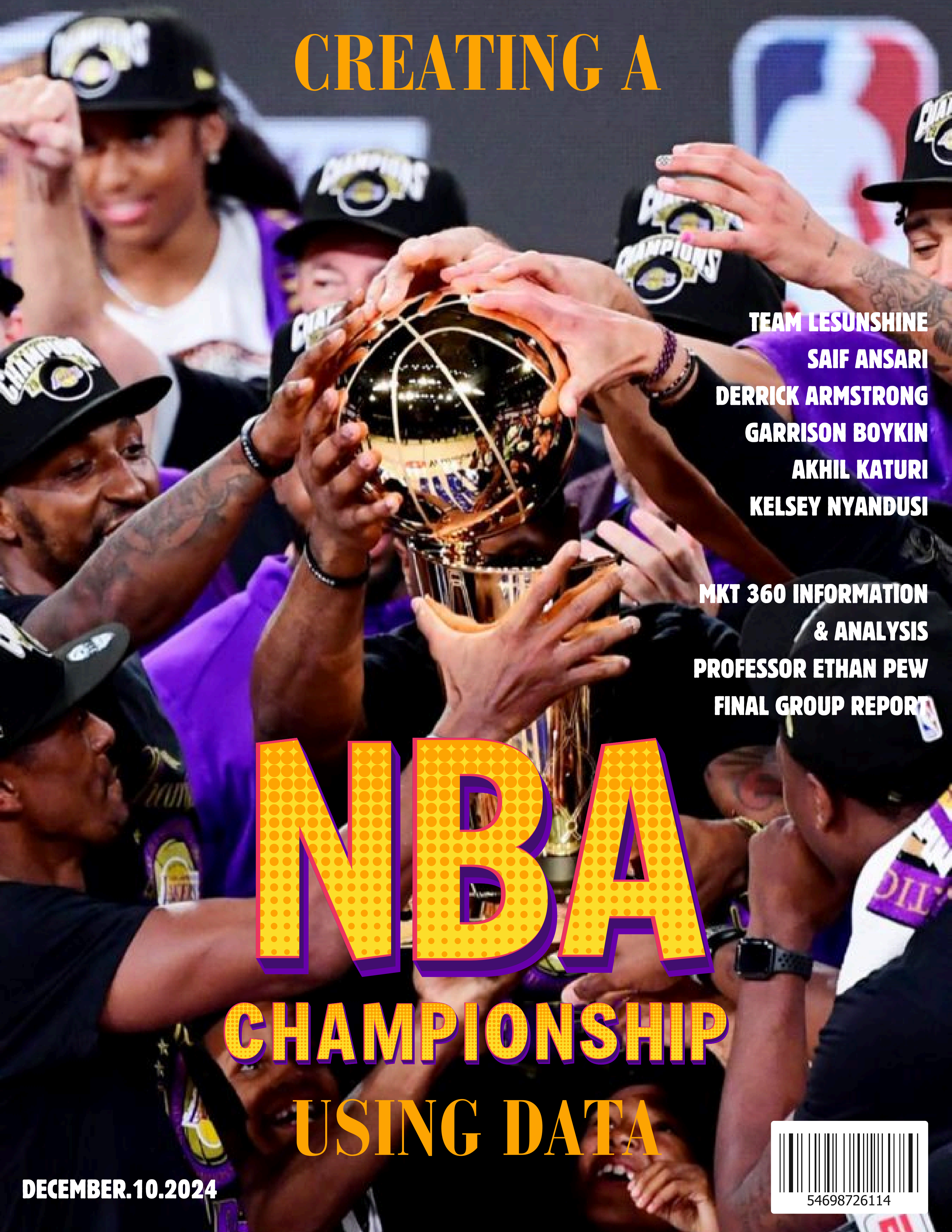


CREATING A NBA CHAMPIONSHIP USING DATA

TEAM LESUNSHINE
SAIF ANSARI
DERRICK ARMSTRONG
GARRISON BOYKIN
AKHIL KATURI
KELSEY NYANDUSI

MKT 360 INFORMATION & ANALYSIS
PROFESSOR ETHAN PEW
FINAL GROUP REPORT

DECEMBER.10.2024

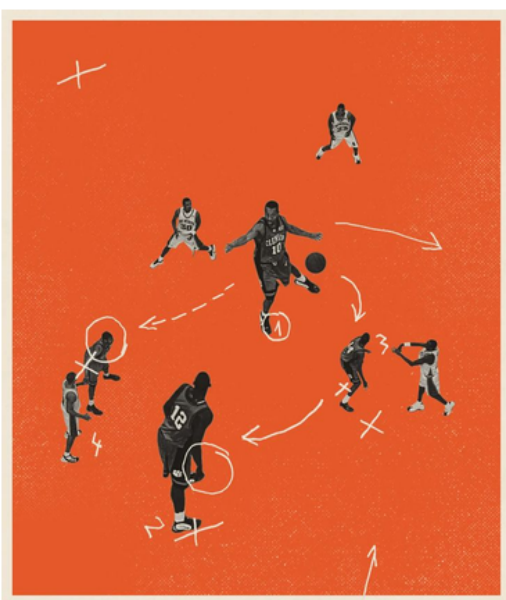
54698726114

TABLE OF CONTENTS

Background

When we first started brainstorming questions to explore, we quickly gravitated toward the topic of basketball. As sports fans, we believed the best way to gain meaningful insights from a data-driven project was to choose a subject we were passionate about.

Initially, while examining NBA datasets, we considered creating predictive models to help amateur sports bettors determine a winning team or even make unique bets, such as predicting who would play in the NBA All-Star Game. However, we quickly encountered data issues with the lack of high-quality data and the need to create highly advanced predictive models using machine learning techniques. It also didn't help that only 24 of the over 500 players in the NBA actually made it to the All-Stars. Our idea of creating a logarithmic-based predictive model was a bust.



Along the process, we came across an interesting observation. Many of the variables describing player performance like shooting accuracy, offensive and defensive rebounds, assists, steals, blocks, etc. were all highly correlated. In statistical terms, we faced multicollinearity issues, which made it difficult to determine how these variables independently affected our observations - specifically, whether a player made it to the All-Stars. We came to learn that the best way to tackle this issue was to combine the variables into a single variable that represents a player's overall efficiency - which we will now refer to as "EFF".

To expand the scope of our analysis beyond the 24 All-Star players, we decided to investigate the relationship between EFF and player salaries. By analyzing and segmenting the data, we aimed to generate valuable insights for an NBA team's general manager (GM) who could benefit from our findings.

Introduction

Meet Rob Pelinka, General Manager of the Los Angeles Lakers, a key decision-maker tasked with shaping a championship-winning roster. From signing and trading players to evaluating draft picks and discovering hidden talent, Pelinka's decisions are critical to the team's success. However, two fundamental questions consistently drive his choices: *how effective are my players? And how much should I pay them?*

Methodology

Dataset Setup

To assist in answering these questions, we utilized the 2023-2024 NBA regular seasons statistics sourced from Kaggle. In addition to our primary dataset, we utilized salary and team information from ESPN and the NBA's official website to add new columns to the data: salary, team name, team win percentage, and market size. With these enhancements, our preliminary dataset setup established a solid foundation for analysis.

Data Cleaning

Fortunately, the dataset used did not contain null values or invalid entries relative to the respective columns. However, the data required cleaning and filtering in order to fit the context of our business problem.

First, the dataset included players on two-way deals and therefore included their stats from the G-league. Since these stats would not accurately represent NBA performance, we removed these players from the dataset. Next, the dataset included players who were traded during the NBA season, resulting in multiple rows for the same player with stats recorded for each team they played for. To maintain one entry per player, we retained the row containing their total average stats in combination of all the teams they had played for. For the team column, we assigned the team for which the player played the most games. For example, OG Anunoby played 26 games for New York before he was traded to the Toronto Raptors where he played 32 games. In this case, we assigned the Raptors as his team.

It is worth noting that a change in team could lead to a change in the player’s role and, consequently, their efficiency score. While we could have retained separate records for the same player on different teams, treating them as distinct entries, we opted to normalize the data by consolidating duplicate records for the purpose of this study.

Data Filtering

Two main filters were used on the data in order to remove potential outliers when running analysis. First, players that played less than five games were removed. This is because a minimum of five games eliminates the anomaly of a player having one or two great games. Next, we removed players that did not average at least 10 minutes a game. Players falling below this threshold are not in the playing rotation regularly and therefore would not greatly contribute to our analysis, especially relative to the new metric (EFF per 36 minutes) that we next created.

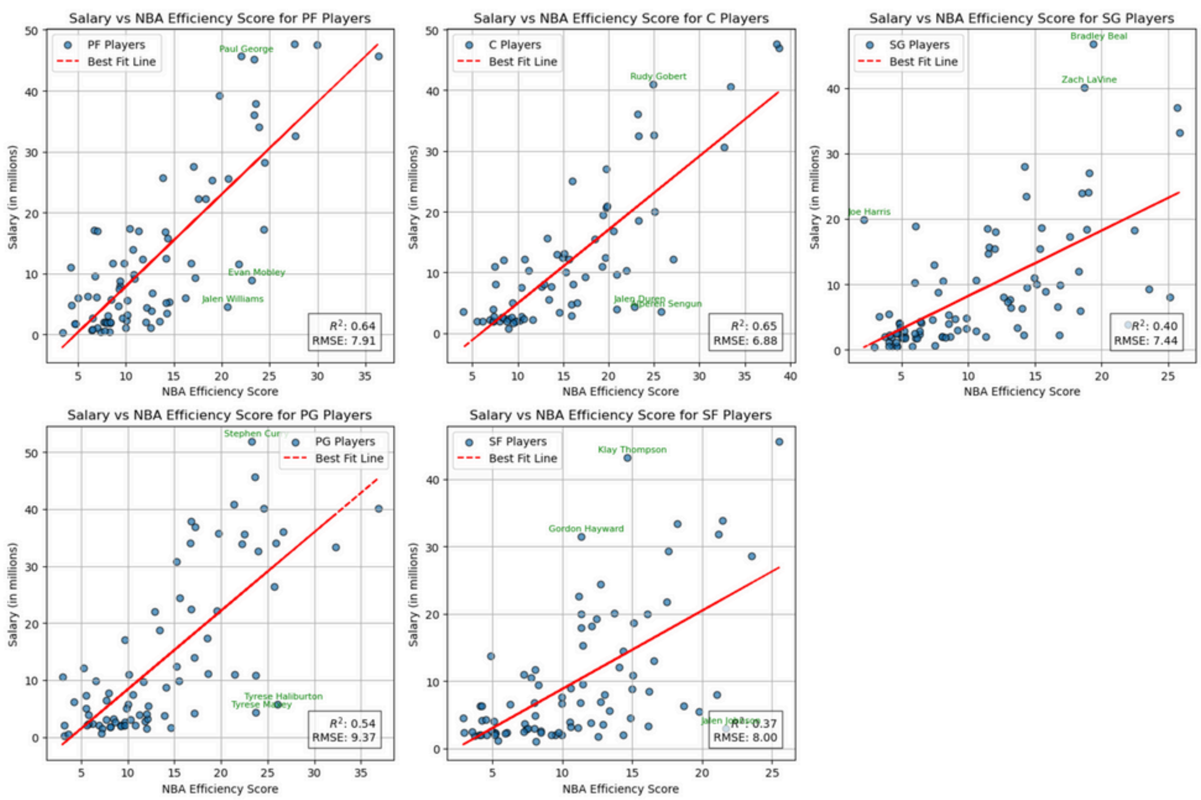
New Metrics

A few metrics were implemented and created in order to perform the analysis on our data. First, we added the efficiency metric that the NBA currently uses to track individual player efficiency. It takes a player’s points, assists, rebounds, blocks, and steals, then subtracts turnovers, missed field goal attempts, and missed free throw attempts. This step was important as our preliminary analysis showed multicollinearity between many of the player stats. This made sense as for example, a player that gets an offensive rebound is likely to score a point.

Next, we created a new metric to find a player's efficiency per 36 minutes of play, but at a regressive rate. This way, a player’s efficiency would not be exactly double per 2x minutes that they played, which helps account for fatigue and defensive adjustments to a player playing well. Each stat within this metric was also weighted with reference to how difficult it was to obtain and how much it contributes to the team overall. For example, a point is important, but a steal takes away a possession from a team and could also lead to a score. The last metric created was a value per million metric which took a player’s efficiency divided by their salary (in millions) to understand the efficiency they provided relative to their salary. This metric would be the cornerstone of our most important analysis and findings.

Through the five visualizations in Figure 1, each focusing on the relationship between salary and EFF for a specific position, we identified a positive correlation showing that more efficient players tend to earn higher salaries. This relationship provided the foundation for deeper exploration into salary efficiency dynamics, ultimately offering a GM like Pelinka actionable insights for building a stronger Lakers roster.

Figure 1



Insight #1

The NBA's market size directly influences player salaries, team performance, and how teams build their rosters. By examining win percentages, individual player efficiency (EFF), and star player contributions to total team efficiency across big, mid, and small markets, we can identify key factors driving success and areas where improvement is necessary.

Big-market teams lead the league in terms of average win percentage, with 56.09% compared to 49.57% for mid-market teams and 44.52% for small-market teams. This disparity reflects the financial advantages that big-market teams enjoy, including higher revenues from ticket sales, sponsorships, and media deals. These resources allow them to attract top talent, invest in better facilities, and build competitive teams.

Figure 2

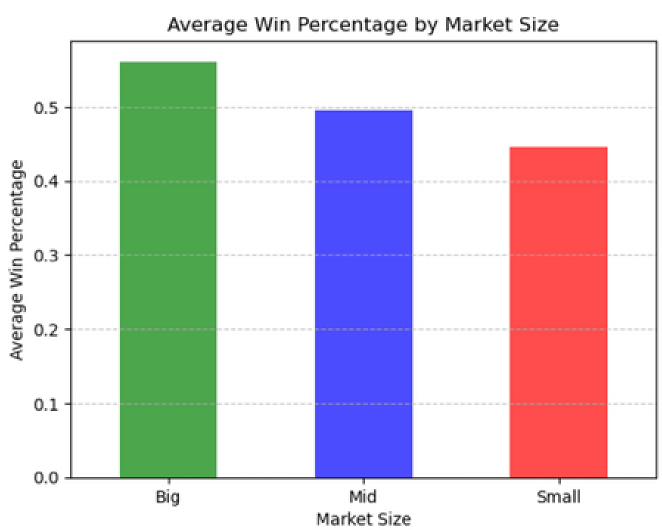
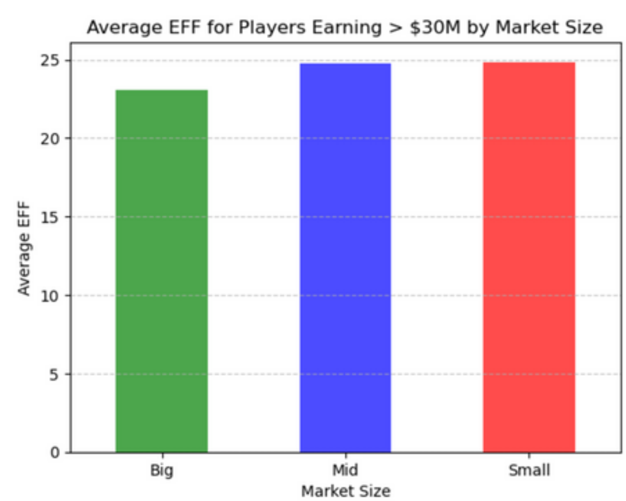


Figure 2 underscores the clear advantage that big-market teams hold, further emphasizing the challenges faced by mid- and small-market teams in competing at the same level.

Figure 3

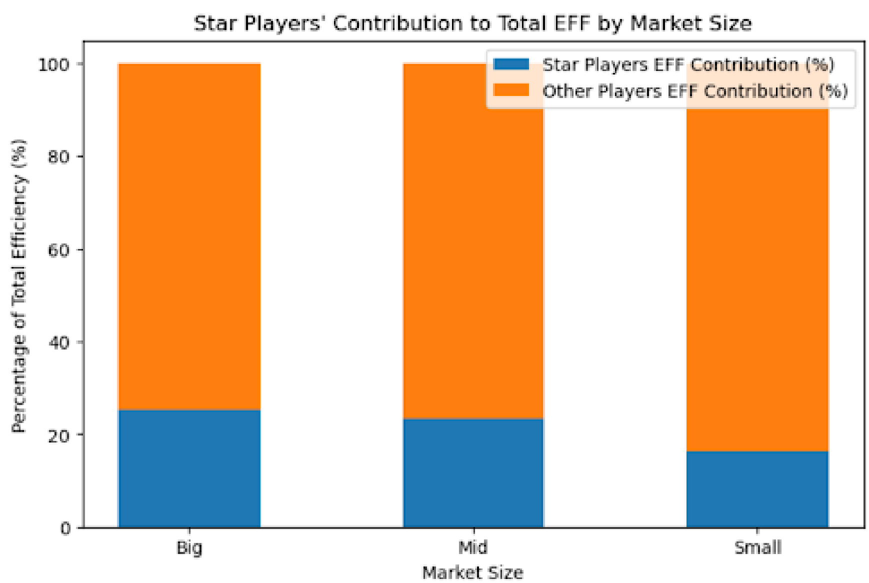


Interestingly, players earning over $30 million demonstrate similar efficiency levels (average EFF) across all market sizes. Figure 3 shows that high-salary stars deliver consistent individual contributions, regardless of the team's market size.

This consistency in performance suggests that star players alone cannot explain the difference in team success. Instead, it highlights the importance of other factors, such as team depth, role players, coaching quality, and overall team. Big-market teams likely excel in these areas, which allows them to maximize the value of their stars.

A deeper look into how star players contribute to their teams reveals that their impact differs significantly based on market size. Figure 4 shows that star players contribute a smaller percentage of total team efficiency (EFF) across all markets.

Figure 4

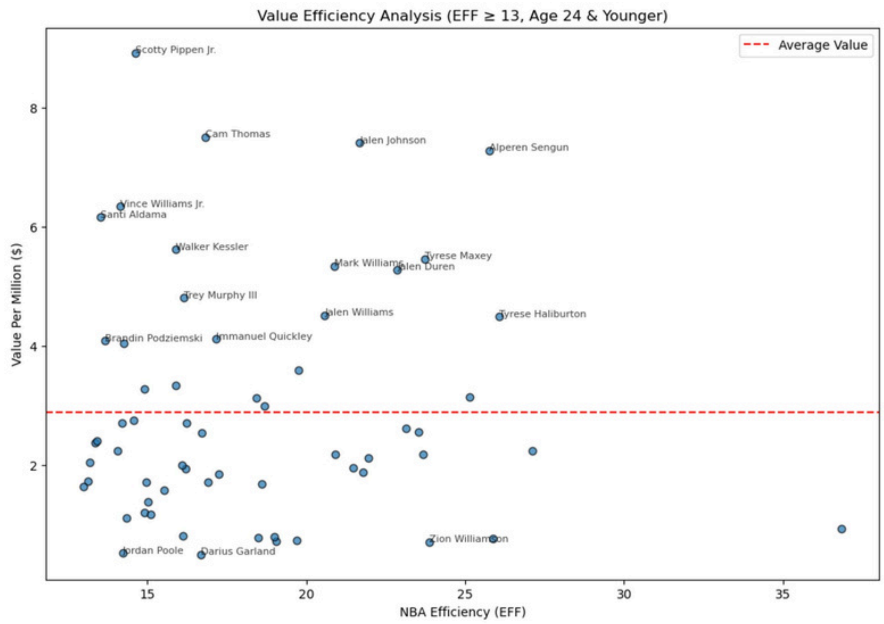


This finding highlights the importance of roster depth in all markets, where role players contribute significantly to overall team performance. Conversely, mid- and small-market teams rely disproportionately on their star players, which limits their ability to compete effectively when these stars underperform or are unavailable.

Big-market teams should maintain their competitive edge by continuing to invest in role players, ensuring balanced rosters that reduce reliance on individual stars. This depth has proven critical to their sustained success. Mid-market teams, while competitive, must look at the efficiency of their high-paid stars and strengthen their supporting role players through strategic investments in role players and infrastructure. Small-market teams, heavily reliant on star players, should prioritize building balanced rosters by increasing player development and coaching, reducing star underperformance. These adjustments can help teams across all markets close performance gaps and create a more competitive league.

Insight #2

Figure 5



To give a quick overview of the characteristics of Figure 5, we placed two filters on our players. We are only looking at players with an EFF >=13 because we believe that this is the lowest threshold for a player that a GM would need to be concerned about, especially when they are looking for trades or contract extensions. We want them to focus on players that have an impact on the game and who can be developed into the future faces of the franchise and even the NBA. Our second filter allows us to only observe players that are 24 or younger. This is to only look at players that are still within their rookie contract, as they are likely to be a lot more mobile than players who are well seasoned and settled into a franchise that works well with them. This allows a GM to really focus on players who are either at risk of being poached by other teams that are willing to pay their players more, or to poach players themselves if their team has the budget or resources to offer a more enticing compensation package.

A GM would interpret this by looking at players who have similar NBA EFF scores and study players in that vertical. Players who are higher on the y-axis, showing a higher Value per Million Dollars, can be considered as having a better “bang for the buck”, as it indicates that they are paid far less than the players along the mean VPM of $3M.

To put this in context, let's say you are Zach Klieman, the GM of the Memphis Grizzlies, and you're looking at some of your rookies to see how you can strategize and create your championship winning team. Looking at Figure 5, you see that your Shooting Guard, Scottie Pippen Jr., has a similar efficiency score to Jordan Poole, Shooting Guard from the Washington Wizards, but gets paid $1.6M compared to Jordan's $27.5M. This raises a concerning flag for you. Now you have a player that's got an efficiency score similar to players that get paid closer to $3M and even higher! You start to wonder whether Scottie Pippen Jr. needs a better bonus to retain him, or he might be bought by another team that seeks his talents but has the pockets to pay for him.

Figure 6

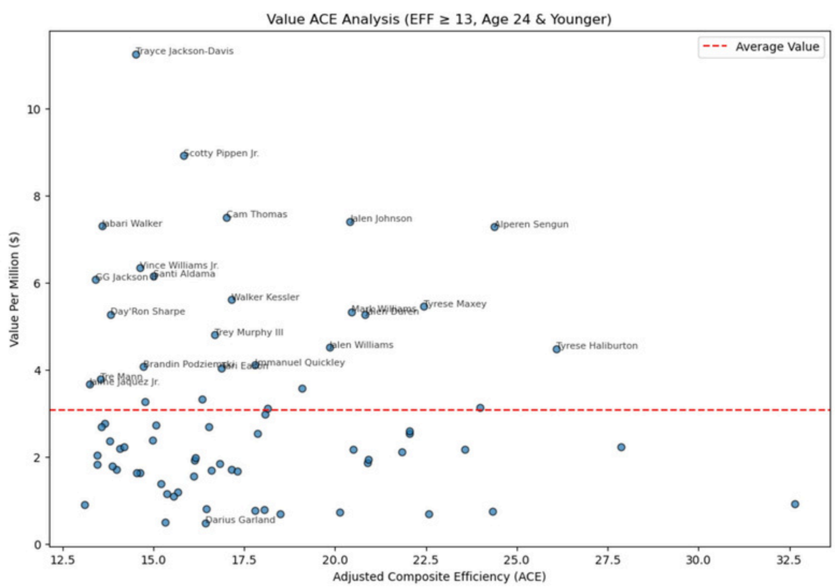
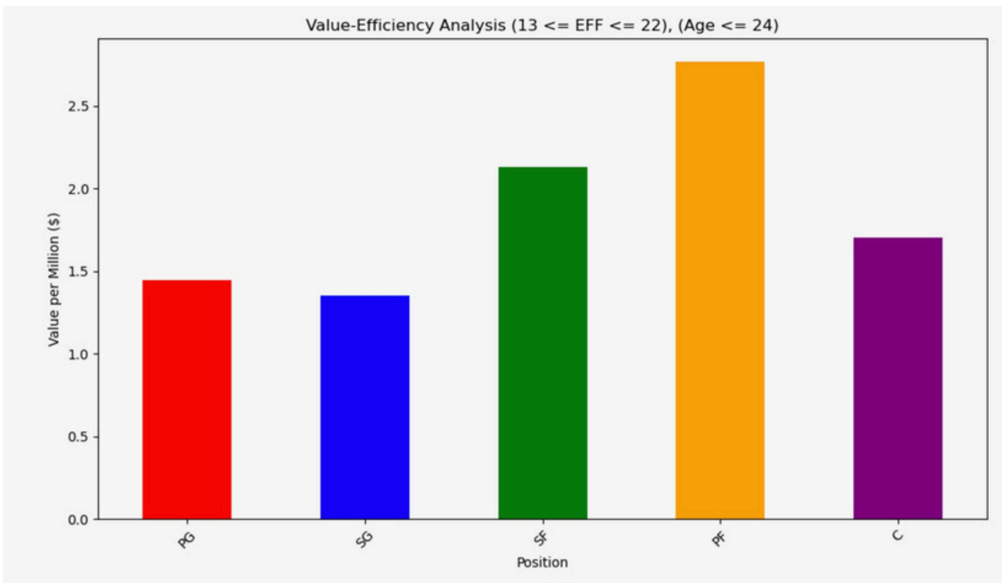


Figure 6 takes what we've observed and adds an extra layer of normalization to it. We're now looking at VPM according to our ACE Efficiency metric. ACE normalizes players according to total minutes played, putting them on an equal playing field, so we're now looking at player efficiency scores per 36 minutes played. The point of adding this is to see whether we notice any massive changes in the players' VPMs. From our observations, everything seems consistent except for a few players who now emerge to the top. We see players like Trayce Jackson-Davis, the Forward for the Golden State Warriors, who from what we can see is not paid nearly as much as other players in his similar efficiency bracket. It helps GMs identify further upcoming players that they might want to look out for for contract extensions or trades.

Insight #3

Lastly, we looked at efficiency per million at the positional level. Once again we looked at players with efficiency scores between 13 and 22 so that only players in a rotational role would apply, as well as so the diminishing returns of highly paid superstars would not skew the analysis. This analysis helps a general manager evaluate their current team. Further, it helps identify which positions offer the best value for the money and highlights potential inefficiencies in spending. In order to do this effectively, it was important that we broke this analysis into two age groups: younger players likely still on their rookie contracts, and veteran players who have somewhat established themselves in the NBA.

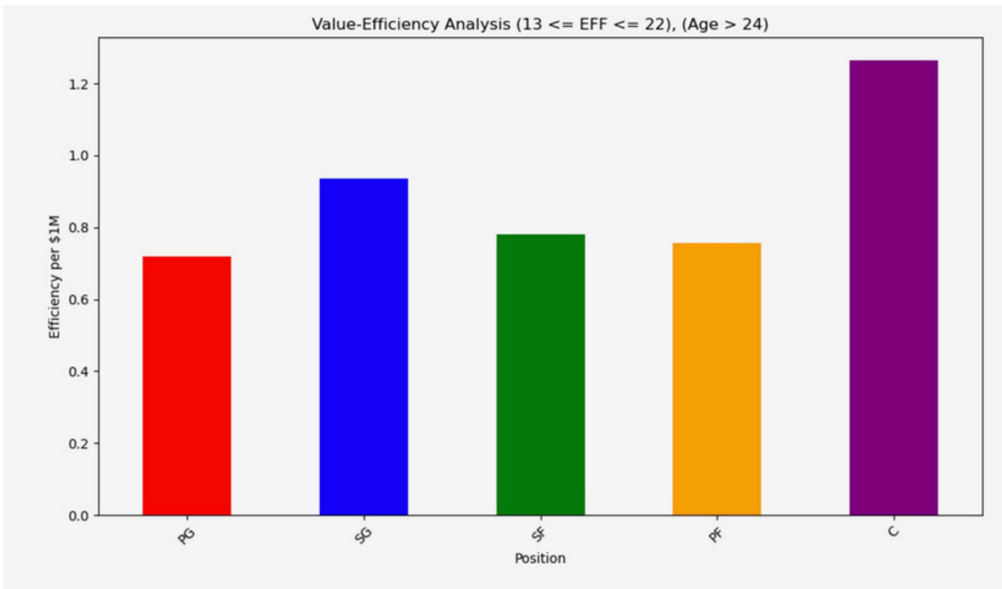
Figure 7



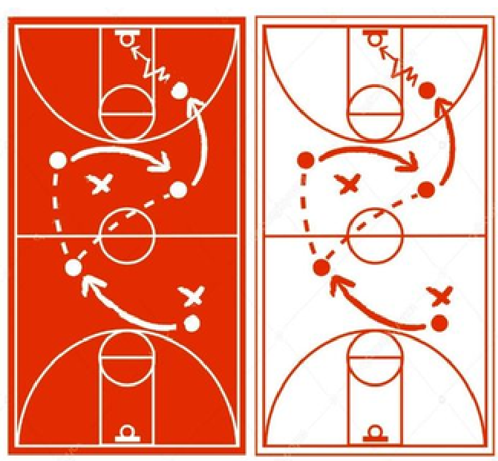
When examining players at the age of 24 and below in Figure 7, it shows that power forwards provide the best value, followed by small forwards. This indicates that teams are, on average, getting the most “bang for their buck” from rookies at these positions. Point guards, shooting guards, and centers still provide over a unit of one in terms of value per million, but are not as cost effective relative to the other positions.

Moving on to players over the age of 24 in Figure 8, or those likely past their rookie deals, the center position has the highest efficiency to salary ratio. Point Guards, power forwards, and small forwards have the lowest Efficiency-to-Salary ratio, meaning they tend to be the most expensive relative to the efficiency they contribute. These are simply average values, but it provides GMs with a great starting point to examine the value they're receiving from their players relative to other teams.

Figure 8



What exactly do these findings indicate? First, teams should look to invest in rookie power forwards and small forwards, as well as non-rookie centers. This is a cost effective way to improve team performance Second, GMs should reassess spending on non-rookie point guards, shooting guards, and small forwards to ensure they are receiving adequate contribution to justify the amount paid to these players. By implementing these measures, a GM can ensure that they are being as frugal as necessary to build their winning roster.



Conclusion

Step 1
Market Share
Evaluating Team Market

Step 2
Apply EFF
Positional Analysis

Step 3
Player Spread
Team Evaluation

Step 4
Adjust Budget
Team Improvement
(Add, Subtract, Develop)

To build a championship-caliber NBA team, a general manager must take a strategic, data-driven approach. Our analysis, grounded in market size, positional value, and player efficiency (EFF), provides a clear roadmap for achieving sustained success. The first step in this process is evaluating the market size, as it plays a significant role in shaping both a team's financial strategy and player acquisition approach. Teams in larger markets typically have greater financial flexibility, allowing them to attract and retain star players. Our data shows that bigger markets tend to outperform smaller ones in terms of win percentage. However, this financial advantage comes with the challenge of balancing high salaries with player efficiency. Teams in smaller or mid-sized markets, on the other hand, must adopt a more cost-conscious strategy. By focusing on undervalued players with high efficiency-to-salary ratios, these teams can remain competitive without overspending. Interestingly, teams with high-salary players (over $30 million) in smaller and mid-sized markets tend to achieve similar efficiency metrics to big-market teams, suggesting that smarter spending is more effective than simply acquiring high-salary players.

Next, an in-depth positional analysis reveals the importance of cost-effective spending across the roster. Centers provide the best value for money, offering superior efficiency-to-salary ratios compared to other positions, making them a priority for retention or acquisition. Conversely, point guards often see overinvestment, with teams frequently spending more than necessary unless the player is a top-tier contributor. Teams should reassess their spending in this position and explore opportunities for greater value in other areas. Forwards, especially those who are versatile and can contribute across multiple categories such as scoring, rebounding, and defense, offer significant value and should be prioritized in roster construction.

Evaluating the current roster is the next critical step, where GMs should focus on identifying strengths, weaknesses, and inefficiencies. High-value players, particularly those under rookie contracts with high ACE scores, should have their contracts extended to secure long-term value at a reasonable cost. On the other hand, high-salary players who are underperforming should be targeted for loss, through trade or release. This approach will help free up cap space, enabling the team to make more efficient roster moves.

Finally, improving the team requires a careful balance of player gains, losses, and development. GMs should target undervalued players in free agency or through trades, focusing on those who offer high efficiency relative to their salary. Losses are equally important; moving on from overpaid underperformers will free up cap space, allowing for strategic reinvestment in more efficient players. Additionally, providing more opportunities for young, high-value players, such as rookies with strong ACE scores, will accelerate their contributions and help build a sustainable team for the future.

In summary, creating a championship-winning NBA team involves making smart decisions at every level, from evaluating market constraints and team needs to strategically making gains and losses. By prioritizing player efficiency and maintaining a balanced, value-oriented roster, GMs can build a team that competes at the highest level while preserving financial flexibility for long-term success.

References

ESPN. (2024). NBA salaries by year: 2024. ESPN. Retrieved December 6, 2024, from https://www.espn.com/nba/salaries/_/year/2024

National Basketball Association (NBA). (2024). NBA standings and margins for the 2023-24 season. NBA. Retrieved December 6, 2024, from https://www.nba.com/standings?Section=margins&Season=2023-24

VivoVinco. (n.d.). 2023-2024 NBA player stats [Data set]. Kaggle. Retrieved December 6, 2024, from https://www.kaggle.com/datasets/vivovinco/2023-2024-nba-player-stats/data