

Boxuan Wang and Will Frimel
RMI 660 Final Project: LeBron James Triple-Double Simulations
May 10th, 2024

Data Source: [Lebron James Lakers' Statistics \(Pro Basketball Reference\)](#)

Code: [Triple Double Simulations](#)

Overview of Project: For the simulation project, our team aimed to investigate the impact of LeBron James achieving a triple-double on the Lakers' probability of winning a game in the regular season. A triple-double involves LeBron scoring double digits in three of five statistical categories: points, rebounds, and assists. The primary sources of randomness in this simulation will be LeBron's game-to-game performance fluctuations in these categories, along with the performance of both his team and the opposing team.

Data Collection: To construct this simulation, we utilized historical data of LeBron James' game-to-game performance while on the Lakers, along with the Lakers team performance (win vs. loss, margin of result, etc). We collected this information from Pro Basketball Reference, combining data from different seasons into one sheet for our dataset. Eventually, this data was used to fit distributions for LeBron's points, rebounds, and assists, as well as for the team's overall performance metrics. This comprehensive dataset allowed us to model various scenarios and run Monte Carlo simulations to estimate the impact of LeBron's triple-doubles on the Lakers' win probabilities in a realistic and statistically robust manner.

Data Analysis

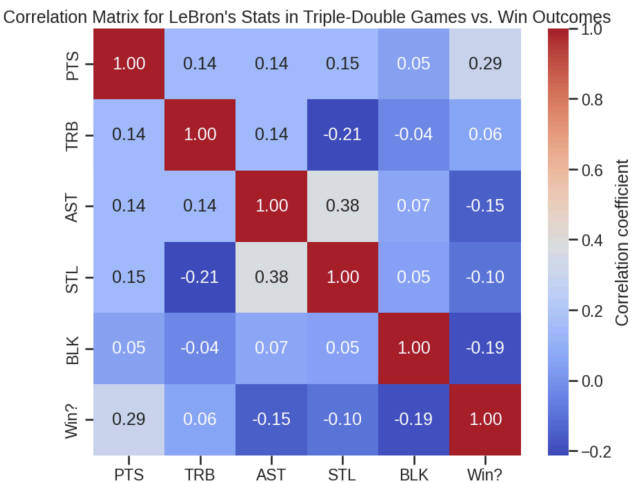
Preparing Data and First Analysis:

In our previous Python code, we performed several data cleaning steps aimed at ensuring the quality and appropriateness of the data for accurate statistical analysis. They include:

- 1) **Handling Missing Values:** We checked for any missing values in the dataset. Addressing missing data is crucial because it can lead to biased or incorrect analysis results if not handled properly.
- 2) **Converting Data Types:** We converted dates to the appropriate datetime format to ensure that operations involving dates were performed correctly. We also converted numeric columns like points (PTS), rebounds (TRB), and assists (AST) to float or integer types. This ensures that numerical operations, such as calculations and statistical analyses, are accurate and error-free.
- 3) **Adjusting Data Representations:** We replaced certain symbols and filled missing categorical values to standardize the data. For example, converting "@" to "Away" and filling unspecified entries with "Home" in the location column helps maintain consistency and clarity in our dataset's categorical information.

Analyzing the Probability of LeBron James Achieving a Triple-Double:

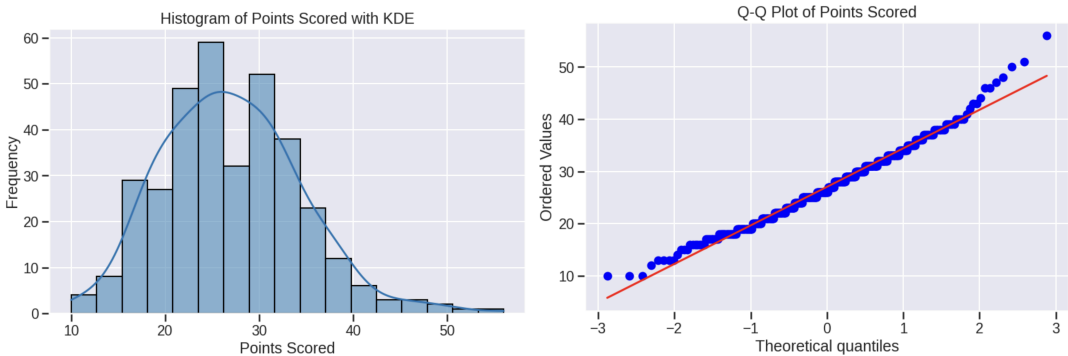
In our study, we treat the occurrences of points scored (Pts), assists (Ast), and total rebounds (Trb) by LeBron James as independent events. This assumption was substantiated through the use of a heatmap, which confirmed minimal correlation among these variables. The heatmap also highlighted that points scored are the strongest predictor of game outcomes, exhibiting the highest correlation with victories.



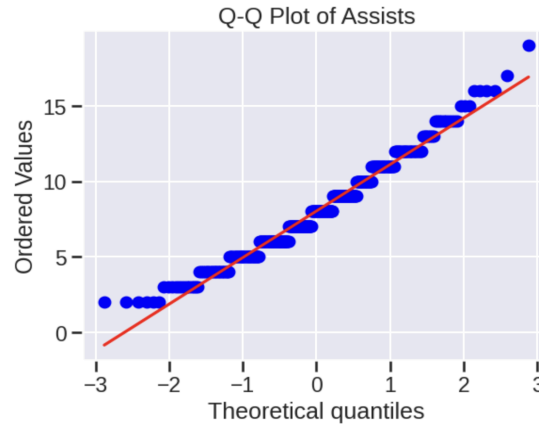
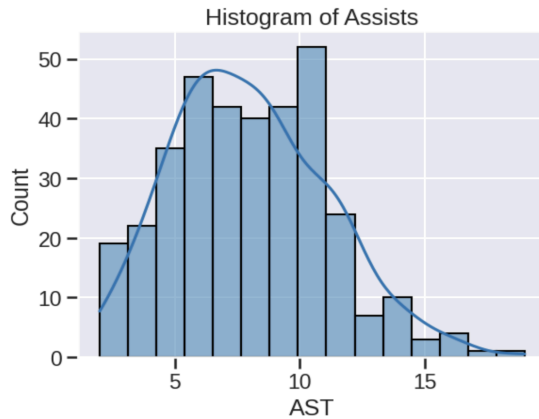
The heatmap indicates a minimal correlation between points (pts), assists (ast), and total rebounds (trb). It also shows that points are the most significant indicator of winning games due to their strong correlation with victories.

Statistical Distribution of Performance Metrics:

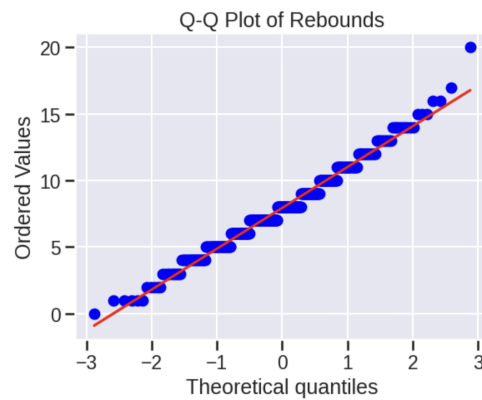
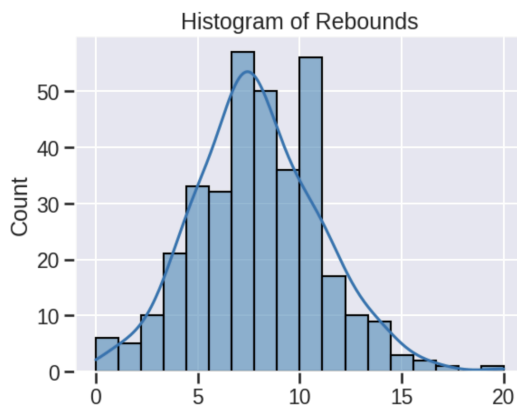
Points Scored (Pts): The distribution of points scored by LeBron follows a normal distribution, suggesting a symmetrical performance around the mean score:



Assists (Ast): The assist numbers are best described by a gamma distribution, indicating a skewed distribution that likely captures the variability in his playmaking performance.

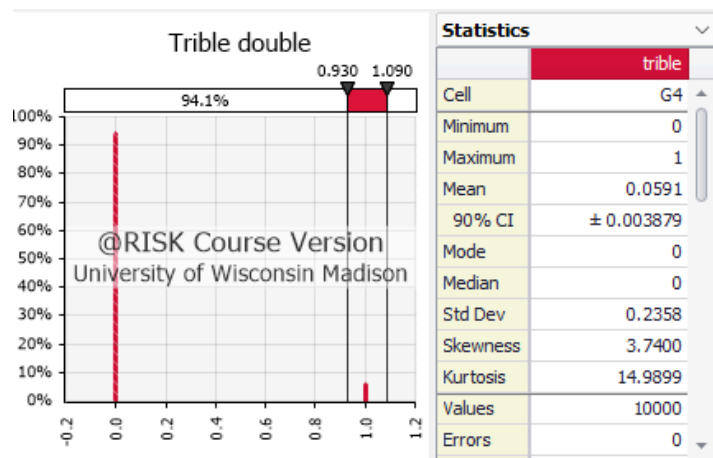


Total Rebounds (Trb): Rebound figures adhere to a Weibull distribution, a model often used for data that demonstrates a particular type of skewness or where the data is bounded on one side.



Risk Simulation of Triple-Doubles in Hypothetical Games:

To estimate the frequency of LeBron achieving a triple-double, we conducted a simulation across 10,000 hypothetical games using the statistical distributions described above. The results indicate that there is approximately a 5.6% chance of LeBron achieving a triple-double in any given game based on the independent performance metrics modeled.



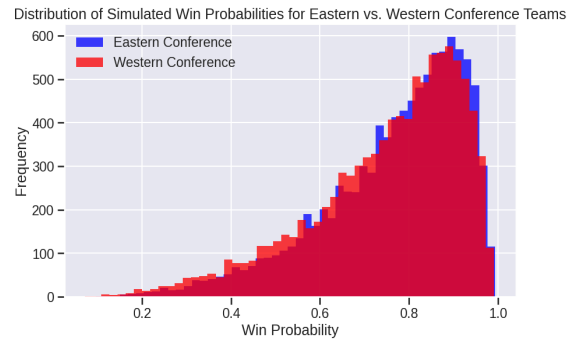
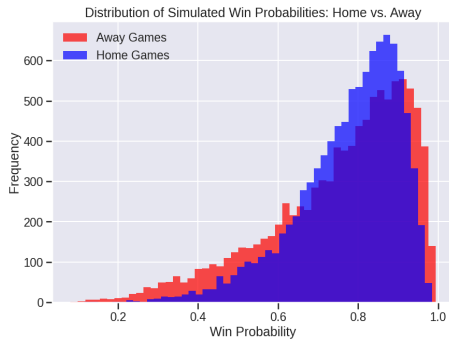
This analysis provides a quantitative basis for understanding the elements contributing to LeBron's triple-double performances and their impact on game outcomes. Further research could explore potential correlations between these performance metrics under different game conditions or opponent strategies to refine the predictive power of our model.

Logistic Regression: Before approaching other simulations, we utilized a logistic regression model to analyze the impact of LeBron James' performance metrics on the Lakers' probability of winning games where he achieves a triple-double. This statistical method allowed us to quantify how each of these performance indicators individually contributes to winning outcomes. While the results indicated that points showed a potential positive impact on win probability, the effects of rebounds and assists were not statistically significant on their own. This analysis helped us understand which aspects of LeBron's triple-doubles are most crucial for leading to wins before diving our simulations.

Python Simulations:

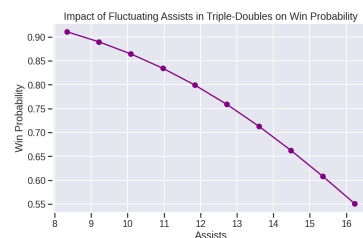
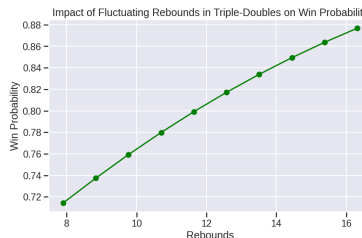
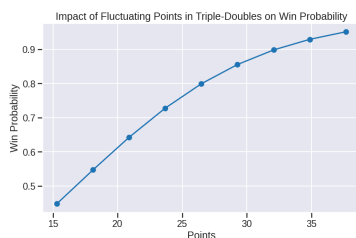
Triple-Doubles and Win Probability: First, we conducted a Monte Carlo simulation to understand how LeBron James' triple-doubles affect the Lakers' win probability - a major goal of our overall analysis. Using logistic regression coefficients derived from his historical performance data in the step before, we simulated 10,000 games, capturing the variance in his points, rebounds, and assists. This approach allowed us to generate a realistic estimate of win probabilities based on his performance metrics. The simulation results showed a significant estimated win probability of 76% when LeBron achieves a triple-double, highlighting how standout individual play can impact his team's success. This quantifies the value of his contributions in games with triple-doubles, demonstrating their importance in achieving positive outcomes. For a fun fact, if LeBron was to record a triple-double in every game of an NBA season (82 games), the Lakers would win a simulated 62 games, while losing 20 games.

Home Game vs. Away Game Impact: In our next simulation, we explored how the location of the game (home vs. away) impacts the Lakers' win probabilities when LeBron James achieves a triple-double. The results indicate a slightly higher win probability in home games (78%) compared to away games (75%), suggesting a home court advantage in scenarios where LeBron performs at a high level. By running 10,000 simulated games and averaging the outcomes, we've attempted to capture the effect of “game venues” on winning likelihood. In all, this analysis helps quantify the influence of home advantage on game outcomes when LeBron is at his best.



Eastern vs. Western Conference Opponents: In the following simulation, we assessed how LeBron James' performance in achieving a triple-double influences the Lakers' win probabilities against teams from the Eastern and Western Conferences. The simulation revealed very close win probabilities, with a 76% chance against Eastern Conference teams and a 75% against Western Conference teams. The close probabilities affirm LeBron's significant influence irrespective of the opponent when performing at a high level, highlighting his crucial role in games where he excels across multiple statistical categories.

Points, Rebounds, Assists: As a final simulation, our team wanted to dive deeper into how specific fluctuations in LeBron James' triple-double components—points, rebounds, and assists—affect the Lakers' win probabilities. We set up simulations to model variations in each of these metrics across a range, adjusted based on standard deviations from the mean performance data. This analysis allows us to understand not just the impact of achieving a triple-double, but how different levels within these achievements influence game outcomes. For each metric, we simulated win probabilities by incrementally adjusting LeBron's points, rebounds, and assists from their average values by up to two standard deviations. This approach reflects realistic game scenarios, where LeBron's performance might vary. For example, we saw that an increase in points generally contributes to a higher win probability, which aligns with basketball analytics that prioritize scoring impact - especially when LeBron is a primary scorer. The graphs from these simulations visually depict the relationship between each metric and win probability. Notably, points showed a strong positive correlation with winning, whereas assists displayed a more complex pattern - which is consistent with our earlier findings.



To build upon our data visualizations, we highlighted common statistical outcomes to see how winning probability was impacted (and fluctuated) when scoring a triple-double:

Points Scored (Win Probability): 15 (44%), 25 (76%), 35 (93%)

Total Rebounds (Win Probability): 10 (76%), 14 (84%), 16 (87%)

Assists Recorded (Win Probability): 10 (87%), 14 (69%), 16 (57%)

The data underscores the varying impact of LeBron James' performance metrics on game outcomes: higher points scored consistently enhance win probabilities, rebounds show a steady increase in winning chances with more rebounds, and assists exhibit a nonlinear relationship, with win probabilities decreasing as assists increase beyond ten. When considering this surprising result with assists, we can hypothesize that diminishing returns or different game dynamics are at play and contributing to this relationship.

Conclusion: When considering our simulations on the impact of LeBron James recording a triple-double on the Lakers' win probability, it's clear that it's extremely beneficial for the team. However, we also learned that external factors like "Home vs. Away Games" and "Western vs. Eastern Conference Opponent" do not significantly impact the winning probability when LeBron has a triple-double. While Points, Rebounds, and Assists all have different impacts on the win probabilities, one thing is certain - when LeBron plays well, the team wins at a higher rate.

Thank you for the great semester! - Will and Boxuan