Winning Factors in Volleyball

Jasmine Dong, Jeremy Li, Allison Zhang, Jodie Zheng

2025-04-28

Introduction

Volleyball is a highly structured, data-rich sport, but unlike basketball or baseball, it hasn't received the same level of mainstream attention from data analysts. While rich with measurable statistics like kills, digs, and errors, the sport often sees match outcomes attributed broadly to "momentum" or "execution" without rigorous statistical validation. Our goal was to dig deeper: **What measurable aspects of in-game performance most predict whether a team wins or loses a men's volleyball match?**

We focused on match-level statistics to explore these questions:

- In particular, we wanted to understand: Which parts of the game — offense, defense, or error control — matter most?
- What statistical patterns consistently show up in winning performances?
- Can we build a model that predicts match outcomes using key metrics?

Using match-level data from the 2023-2024 NCAA Men's Division I season, we combine exploratory data analysis (EDA) with logistic mixed-effects modeling to explore relationships between team performance metrics and match outcomes. We focus on offense (kills, assists, aces), defense (digs, blocks), and different types of errors (serve, receive, attack) to understand how different phases of the game contribute to winning. By combining exploratory data analysis with predictive modeling, we identified the most important areas where teams excel — and where small mistakes make a big difference. Our findings reveal clear, actionable insights for players, coaches, and analysts looking to move beyond intuition and toward evidence-driven strategies for success.

Data

This project uses team-level match statistics from NCAA Men's Division I volleyball to explore the factors that contribute most to winning a match. Data was collected using the `ncaavolleyballr` R package, specifically the `team_match_stats()` function. In our dataset, each row represents one team's performance in a single match, including a wide range of match statistics.

The original dataset includes offensive metrics such as kills, assists, and aces; defensive metrics such as digs, block solos, and block assists; and error-related metrics such as attack errors, serve errors (SErr), receive errors (RErr), and block errors (BErr). Pre-calculated fields like hitting percentage and total attacks are also included.

To deepen our analysis, we created a new metric called block points (Blocks), calculated as:

$$\text{Block Points} = \text{Block Solos} + 0.5 \times \text{Block Assists}$$

For tactical clarity, we grouped key performance metrics into four major "buckets":

- Offense: Kills, Assists, Aces
- Defense: Digs, Blocks (Block Points)
- Offensive Errors: Errors, Serve Errors (SErr)
- Defensive Errors: Receive Errors (RErr), Block Errors (BErr)

Additionally, the following key derived metrics were used throughout our analysis:

- Hitting Percentage:

$$\text{Hit Percentage} = \frac{\text{Kills} - \text{Errors}}{\text{Total Attacks}}$$

- Total Attacks: Sum of kills, errors, and other non-scoring attack attempts.

Exploratory Data Analysis

Using a combination of EDA techniques, we first began looking for patterns and differences between winning and losing teams. We used correlation analysis, scatterplots, density plots, and faceted visualizations to uncover how teams play and how those patterns relate to match outcomes.

EDA 1: What separates winners and losers at a glance?

We first compared average match statistics for winning and losing teams to get a high-level view of which performance areas most distinguish successful teams. This baseline helps identify where winners typically outperform losers.

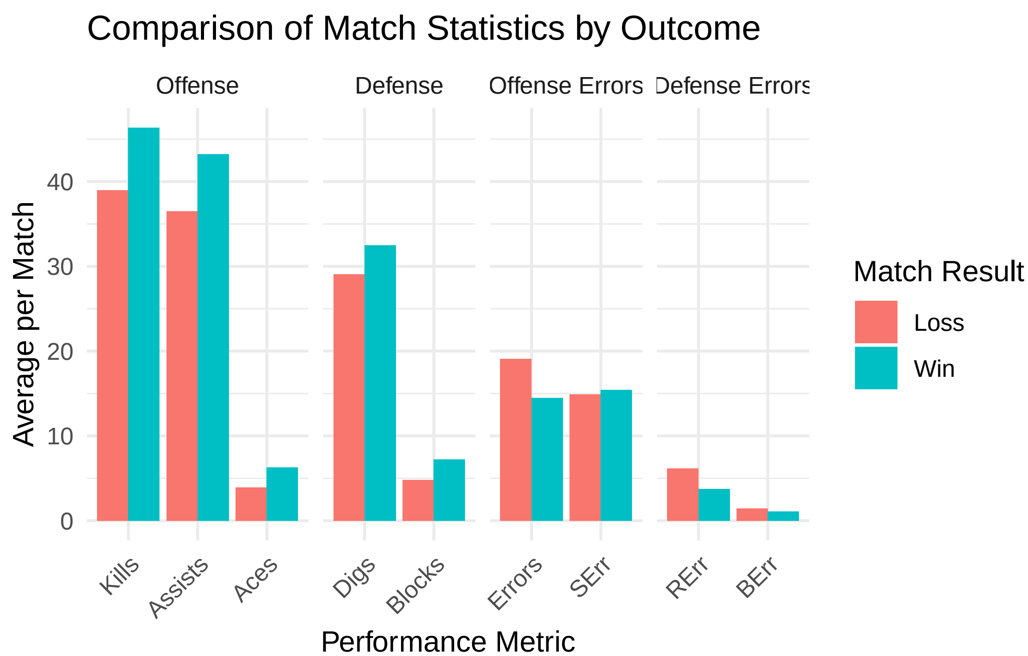

Figure 1: Average match statistics across winners and losers, grouped by play category.

After reviewing the averages, we observe that winning teams record more kills, assists, aces, digs, and blocks compared to losing teams. Conversely, losing teams commit more receive errors and general attack errors. Interestingly, winners show slightly higher serve errors, suggesting that taking more risks on serve may lead to rewards if managed well. This gives us an initial view that offense, defense, and mistake management all factor into match outcomes.

EDA 2: Does aggressive serving help or hurt?

We further investigated the observed serve error pattern, exploring whether higher serve errors correlate with more aces — and if this tradeoff benefits winning teams.

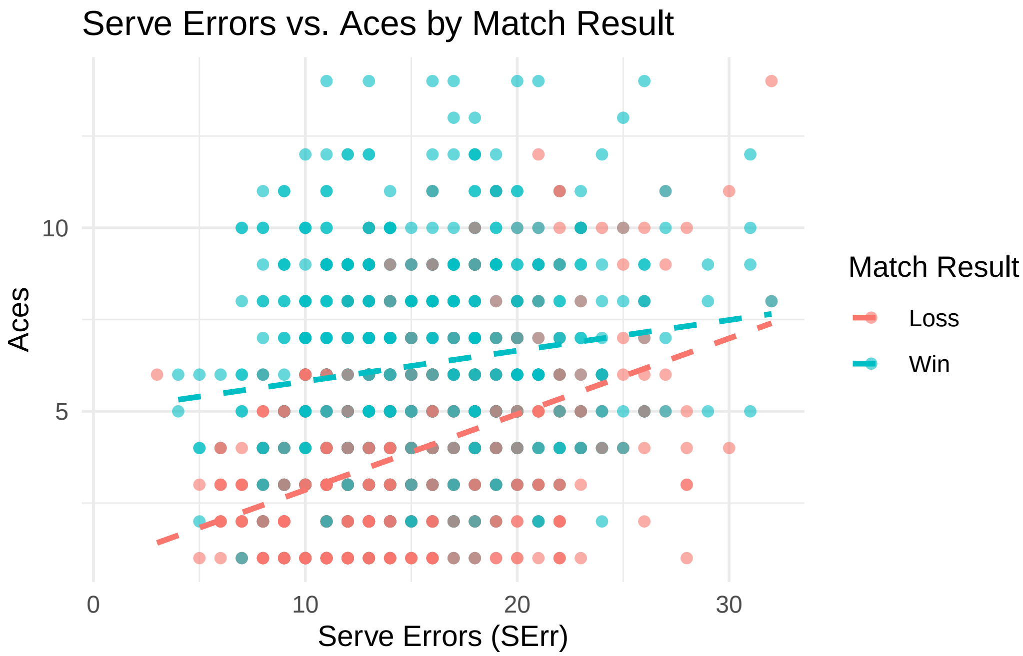

Figure 2: Relationship between serve errors and aces, split by match outcome.

This analysis shows that both winners and losers tend to earn more aces as their serve errors increase. However, winners consistently generate more aces at any given error rate compared to losers. This suggests that aggressive serving is beneficial for winning teams as long as it results in more scoring opportunities, supporting the idea that controlled risk is part of an effective serve strategy.

EDA 3: Which stats are most correlated with winning?

From EDA 1 and EDA 2, we explored the patterns of winning teams, such as what they tend to do more or less of on average, and how specific metrics like serve errors and aces might interact. In EDA 3, we decided to look at how all match stats relate to winning using correlation analysis. This will allow us to understand which stats actually have the strongest relationship with winning and which variables consistently matter.

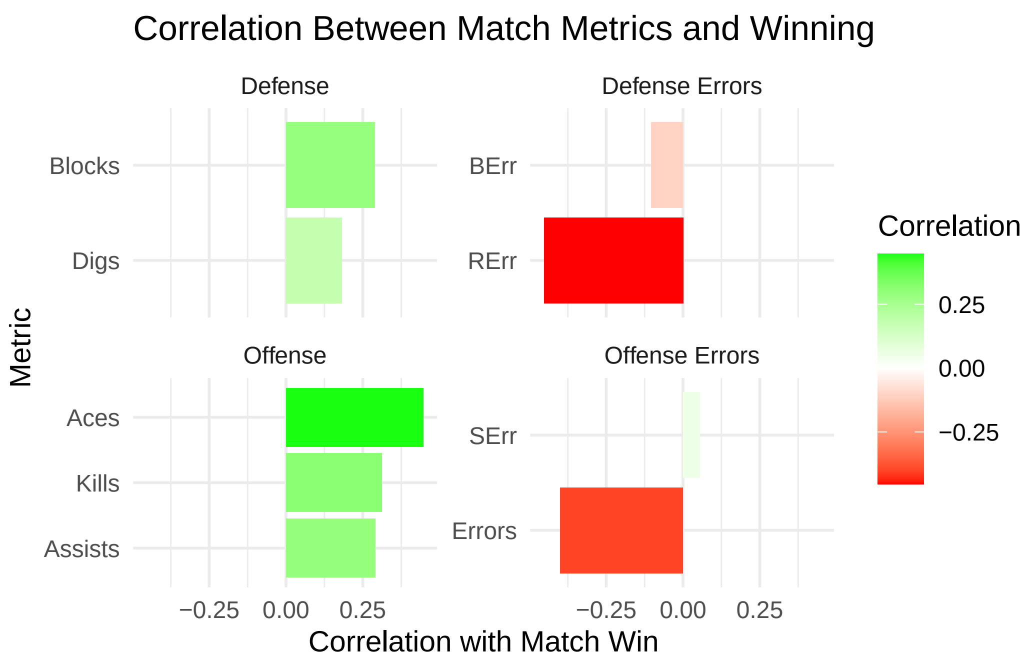

Figure 3: Correlation coefficients between key match statistics and winning outcome.

Positive correlations are strongest for aces, block points, and kills, meaning teams that excel in these areas are more likely to win. On the other hand, receive errors and attack errors show the strongest negative correlations, indicating that minimizing mistakes is crucial. Serve errors interestingly show a slight positive correlation, reinforcing that aggressive serving, if executed well, tends to help rather than hurt.

EDA 4: Does efficiency matter more than total output?

In the previous EDAs, we looked at absolute totals such as average values per metric, such as average number of kills, errors, and aces, to understand what winning teams do more or less of. However, we also realize that how efficient those attacks and defense volumes are might matter. For example, a team might attack a lot but still be inefficient, or attack less and make every attempt count. In this EDA, we used the hitting percentage, which is precalculated in the dataframe as $\frac{Kills - Errors}{Total\ Attacks}$ to offensive efficiency, and see how it differs between winners and losers.

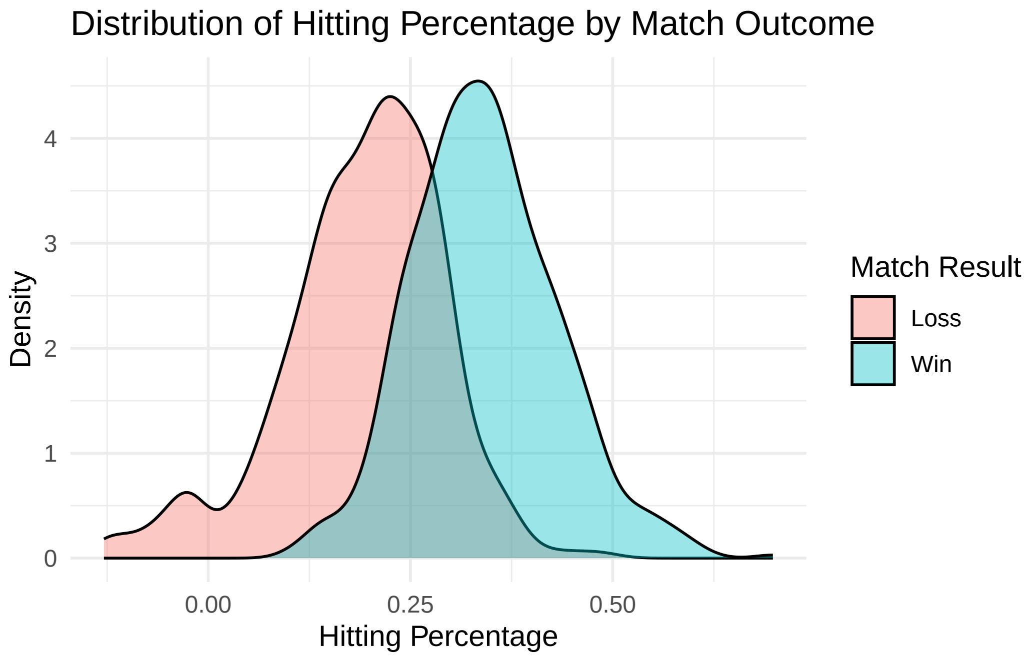

Figure 4: Distribution of hitting percentage for winning and losing teams.

Winning teams seem to have higher hitting percentages on average than losing teams. The distribution for winners is not only shifted to the right but also shows fewer teams with very low hitting efficiency. In fact, the mean of losing teams is actually on the lower tail of the winning distribution. This supports the idea that not just attacking more, but attacking well, is tied to better outcomes. Efficient execution, where there are fewer errors and more conversions, can be a key reason why teams win.

EDA 5: Does hitting efficiency vary across different attack volumes?

Finally, we investigated whether the importance of hitting efficiency changes depending on a team's attack workload.

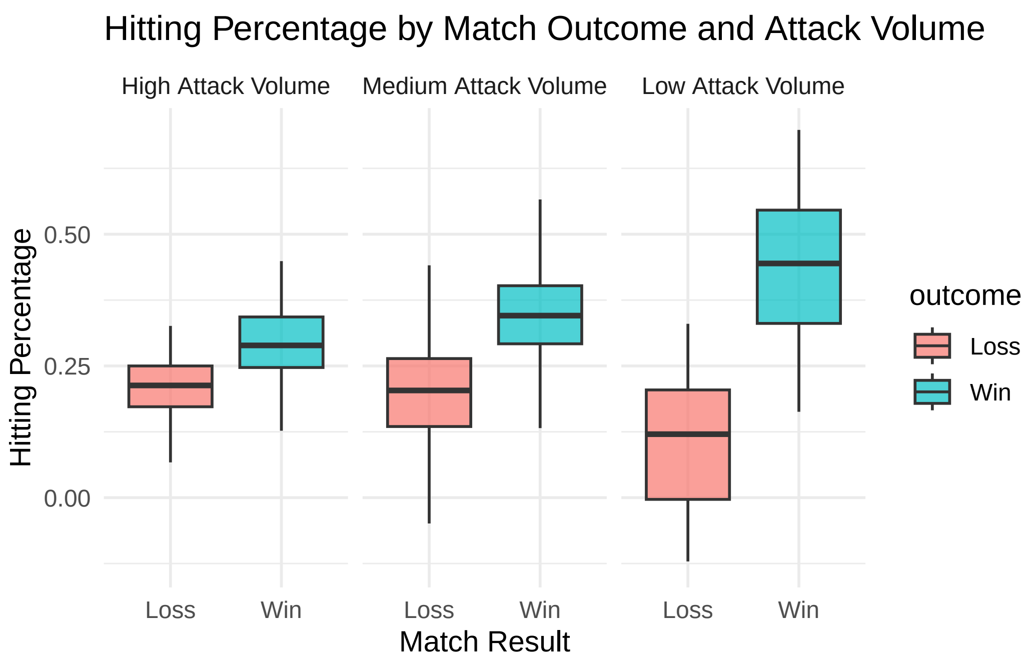

Figure 5: Hitting percentage by match outcome, grouped by attack volume per set.

Across all attack volume groups (low, medium, high), winners have higher hitting percentages than losers. Interestingly, the biggest separation between winners and losers occurs in low and medium attack volume matches, suggesting that efficiency becomes even more critical when attacking opportunities are limited. When attack volume is high, the gap narrows, possibly because volume starts to outweigh minor efficiency differences.

Transition

Our exploratory analysis highlighted key differences between winning and losing teams, particularly in offensive production, defensive plays, and error management. However, to quantify the combined effects of these factors and better predict match outcomes, we now turn to statistical modeling.

Methods

Modeling Approach

To predict match outcomes, we used a multilevel logistic regression model. The response variable is binary, coded as 1 for a win and 0 for a loss. Predictor variables include key match-level statistics such as kills, aces, digs, errors, and other performance metrics reflecting offensive, defensive, and error-related aspects of play.

Because each team appears in multiple matches across the season, observations are not independent. To account for this, we included team ID as a random effect, allowing the model to capture unobserved, team-specific differences in baseline performance. This hierarchical structure better reflects the real-world data-generating process and improves the validity of our coefficient estimates for match-level factors.

The main assumptions of the model are:

- The outcome variable follows a binomial distribution (win/loss).
- Observations are conditionally independent, given the fixed effects and random effects.
- Random intercepts for teams are normally distributed.

Overall, a multilevel logistic regression is appropriate because it directly models the probability of a team winning while properly accounting for repeated measurements from the same teams.

Model Comparison and Evaluation

To evaluate the importance of predictors and check model robustness, we compared alternative specifications:

- Models with and without highly correlated predictors (e.g., excluding assists to avoid multicollinearity).
- Simpler models with only offensive variables versus models including offense, defense, and error terms.

Quantifying Uncertainty

We used bootstrapping to quantify the uncertainty of both fixed and random effect estimates. Specifically, we performed 100 bootstrap resamples of the full dataset, refitting the multilevel logistic regression model each time. For each fixed-effect coefficient, we calculated the bootstrap mean estimate and constructed 95% confidence intervals across the resamples. We visualized this uncertainty using coefficient plots, which highlight the range and stability of each predictor's estimated effect on match outcomes.

Additionally, we also bootstrap the random effect estimates numerous times. This allows us to understand how the team-level variance fluctuates across all teams. From there, we create a density ridge plot to visualize the uncertainty of the top 10 teams' random effects with median lines to show the difference between each team's performances

Results

We fitted a multilevel logistic regression model to predict match outcomes based on team-level performance statistics. The response variable was binary (1 = win, 0 = loss), and a random intercept for team ID was included to account for repeated measurements across teams.

The table summarizes the estimated effects of the match statistics on winning probability. Positive coefficients suggest that higher values of the predictor increase the odds of winning, while negative coefficients suggest the opposite.

Table 1: Table 1: Fixed Effects Estimates with 95% Confidence Intervals

effect	term	estimate	std.error	statistic	p.value	conf.low	conf.high
fixed	(Intercept)	-0.489	0.522	-0.938	0.348	-1.512	0.533
fixed	Kills	0.168	0.049	3.460	0.001	0.073	0.263
fixed	Aces	0.422	0.046	9.259	0.000	0.332	0.511
fixed	Assists	-0.084	0.049	-1.708	0.088	-0.180	0.012
fixed	Digs	0.054	0.015	3.555	0.000	0.024	0.083
fixed	Blocks	0.178	0.031	5.734	0.000	0.117	0.239
fixed	Errors	-0.297	0.025	-11.886	0.000	-0.346	-0.248
fixed	SErr	-0.069	0.024	-2.835	0.005	-0.116	-0.021
fixed	RErr	-0.445	0.044	-10.025	0.000	-0.532	-0.358
fixed	BErr	-0.054	0.073	-0.735	0.462	-0.197	0.089

Table 1: Fixed effects estimates from the multilevel logistic regression model, with 95% confidence intervals.

From the fixed effects estimates, we observe that kills, aces, and blocks have strong positive associations with winning, while receive errors (RErr) and general errors show strong negative associations. Serve errors (SErr) have a small, less consistent relationship with winning, indicating that aggressive serving may introduce risk but can still be beneficial.

To better account for uncertainty in our coefficient estimates, we applied a bootstrap resampling procedure. Figure 6 shows the 95% bootstrap confidence intervals for each predictor.

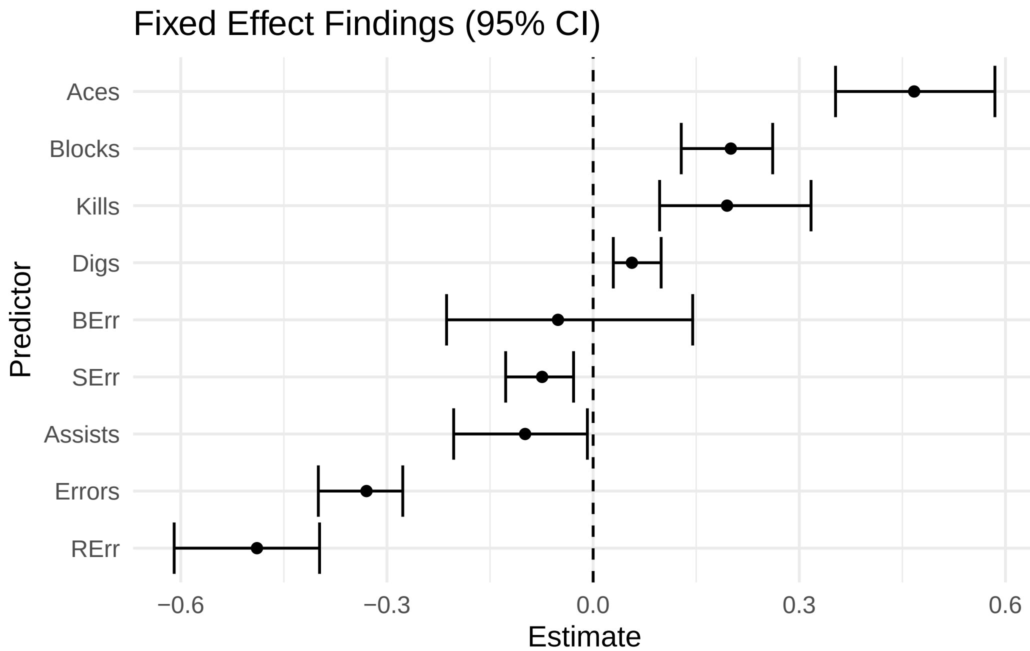

Figure 6: 95% bootstrap confidence intervals for the fixed effect predictors of match outcomes.

From Figure 6, we find that nearly all match statistics are statistically significant predictors of match outcomes, except for block errors (BErr) and assists. Both BErr and assists have small estimates (approximately -0.047) and wide confidence intervals that include zero, suggesting they are weak predictors of winning or losing a match.

Among the significant predictors, receive errors (RErr, -0.441) and aces (0.429) show the greatest impact on match outcome. A high number of receive errors greatly reduces a team's chance to win, and a high number of aces greatly increases the winning likelihood. Attack errors (Errors, -0.293) have the second largest negative impact. On the other hand, serve errors (SErr) are statistically significant but have a relatively small estimate (-0.065). Thus, the model suggests that not all errors are equally detrimental.

Interestingly, not all offensive or defensive plays traditionally emphasized in volleyball are strong predictors of match success. For example, kills have a smaller estimate (0.084) compared to blocks (0.179), and a wider confidence interval skewed toward lower values. This reflects variability in the impact of kills towards a smaller effect, which is surprising as kills are often displayed as decisive plays. Similarly, digs (0.05) show a small positive effect. These findings show that while kills, assists, and digs are often highlighted in games, their independent contributions to match outcome are less impactful compared to serving ability and careful error minimization, particularly towards receive and attack errors.

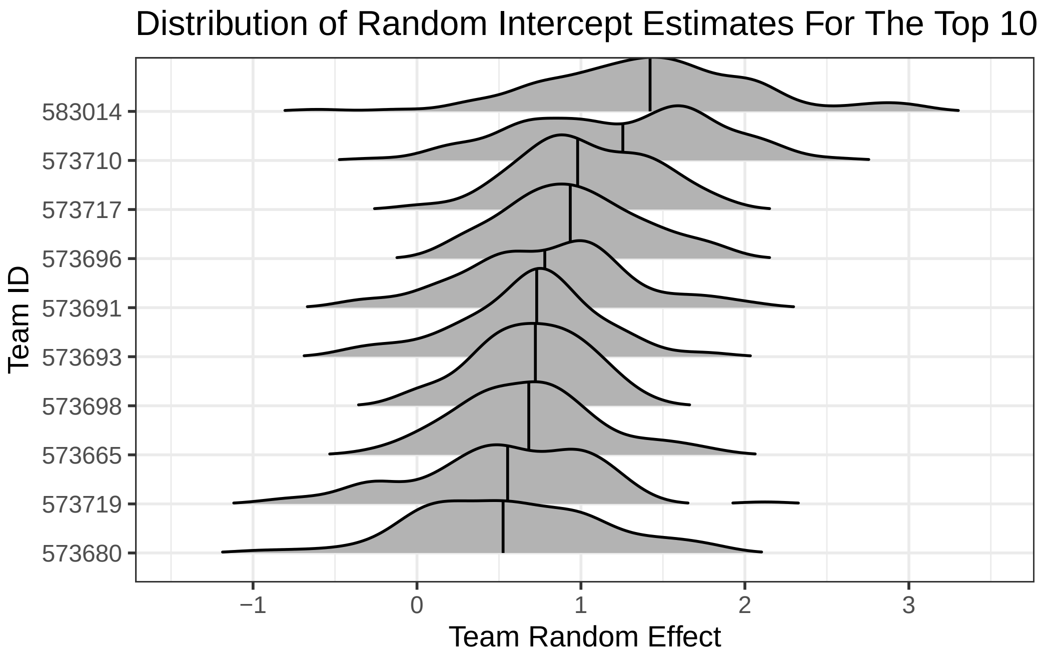

Figure 7: Distribution of bootstrapped random effects for the top 10 teams.

Figure 7 displays the distribution of team-level random effects for the top 10 teams. The random effects adjust for variability in team performance that is not explained by match predictors such as aces, errors, and kills. A higher median random effect means a team consistently performs better than average, even after controlling for game performance. For example, Team 573717 (Maryville) has the greatest positive random effect, suggesting a strong baseline performance beyond measurable game statistics. In contrast, Team 583014 (Thomas More) has the widest distribution, indicating greater uncertainty in the estimation of their team-specific factors. Teams like 573698 (North Greenville), 573666 (Long Beach St.), and 573671 (UC San Diego) showcase high, narrow peaks, meaning the model is more confident in the estimation of their baseline performance.

From this analysis, we can infer that while in-game predictors are great indicators of match success, team identity and underlying factors such as coaching quality, teamwork, experience, and others also influence the outcomes. As a result, not all winning or losing patterns in the NCAA Division I Men's Volleyball can be explained by in-game plays.

However, there is one important limitation of the model. When fitting the model using glmer and using team ID as our random effect, we encountered a convergence issue. This suggests the model may not properly estimate all coefficients due to the complexity of calculating random intercepts for each team. The model still offers valuable insights, but this limitation indicates that careful interpretation is necessary, and that in the future, it may benefit from more data per team.

Discussion

So what really drives winning in NCAA Men's Division I? From both exploratory data analysis (Figures 1-5) and our multilevel logistic regression (Figure 6-7), we find that winning is not just about scoring, defense matters equally as much.

Serving and receiving define the game. Aces (winning points on a serve) and receive errors (mistakes when receiving a ball) have the largest impact on match outcomes (Figure 6). Teams that serve aggressively to win points while minimizing their own receiving mistakes have a competitive advantage. In fact, serve errors have a small (Figure 6) or even a slightly positive association with winning (Figure 3), suggesting that aggressive serving strategies pay off significantly if high aces with controlled errors can be made.

Not all errors are created equally. Receive errors and attack errors have strong negative effects on winning, while serve and block errors have a much smaller impact. In fact, block error might not affect winning at all (Figure 6).

Attack smarter, not harder. More total attacks do not necessarily mean more wins. In fact, teams with fewer total attacks often had higher win rates (Figure 5), and it is having a high hit percentage that matters (Figure 4). While both the media and Figures 1 and 3 highlight kills, assists, and digs as important, the model suggests their effects are more modest compared to other predictors. Factors like hit percentage, precise serving, and effective defensive plays are more reliable predictors of winning. In other words, teams that maximize aggressive point-winning plays while minimizing high-impact errors outperform others. Training should prioritize precise serving (Aces $\uparrow$), efficient attacking (Errors $\downarrow$ while Kills $\uparrow$), and strong recieving (RErr $\downarrow$), as these predictors drive volleyball success.

Through utilizing team-level random intercepts (Figure 7), we identified that team-specific attributes have an affect on match performance beyond in-game statistics. This suggests that coaching, experience, and other strengths play influential roles in predicting success.

Model Limitations and Future Work

Though we avoid attributing match success or failure to in-game statistics alone, our model encountered a convergence issue when estimating random effects for each team. Thus, caution is crucial when interpreting the estimated results, as dependency across team-level factors must be fully captured. Further work could benefit from expanding a larger dataset for more team-level information.

Additionally, evaluating model fit statistcs such as AIC and BIC would provide more insight into refining the model. We can also explore potential interaction effects (for example, whether hitting efficiency is more critical for teams with fewer blocks) to uncover patterns within team strategies.

With a more concrete model, future analysis can provide coaches strategies based not only on lineup decisions but also designing for training programs that maximize effectiveness.