

# Executive Summary

Lance Brady

Due Date: 2024-12-16

## Introduction and context

This project aimed to predict the likelihood of Frontier League baseball players advancing to higher levels of professional baseball, including affiliated minor leagues and, potentially, Major League Baseball. The work integrated advanced performance metrics collected in the Frontier League through Yakkertech with freely available Statcast data from a comparable Single-A league. By contrasting these two data sources, the analysis sought to determine how Frontier League performance translated into future opportunities at higher competitive levels.

## Data and Predictors

The analyses combined pitch-by-pitch datasets from the 2023 and 2024 Frontier League seasons with Statcast metrics from the Florida State League. Key predictors included traditional statistics (e.g., on-base percentage, slugging), rate statistics (strikeout and walk percentages), expected metrics (xwOBA+), and hit quality measures (exit velocity, launch angle). Players were “qualified” based on having at least 130 plate appearances. The data underwent extensive cleaning, SQL-based querying, and the creation of run value-based wOBA weights to ensure robust, comparable statistics across leagues and seasons.

## Modeling Approaches

The modeling process relied primarily on multinomial logistic regression to estimate players’ probabilities of reaching Single-A, Double-A, Triple-A, MLB, or remaining in lower levels, including independent leagues. There were two distinct analyses: Analysis1.R compared 2024 Frontier League players to those in the Florida State League in 2021, and Analysis2.R used 2024 Frontier League outcomes of 2023 Frontier League players to predict 2025 outcomes for 2024 Frontier League players. Multiple models were tested, each focusing on different predictor sets. The process also included calculating log loss and confusion matrices to identify which predictor sets yielded more accurate forecasts. Traditional metrics consistently performed best, outpacing expected and hit quality metrics in predicting advancement.

## Key Results and Performance

The best-performing models consistently used traditional statistics, even though these metrics did not fully capture underlying player potential indicated by advanced measures. The resulting probabilities demonstrated that players with higher xwOBA+ generally had better chances of reaching higher levels, but the system also revealed that many Frontier League players, despite strong performance, had limited opportunities to move into affiliated leagues. The Statcast models suggested that plenty of Frontier League players are dominant enough hitters to move to higher levels of affiliated baseball. None of the models comparing 2023 to 2024 Frontier League data significantly surpassed a no-information baseline by a large margin, suggesting that factors beyond on-field performance influenced player promotions.

## Implications

These findings suggested that Frontier League players exhibited skill levels similar to Single-A counterparts, yet they often remained overlooked by MLB-affiliated organizations. Traditional performance indicators still appeared to drive advancement decisions, indicating a lag in fully incorporating advanced analytics at lower levels. To improve predictive power, acquiring more comprehensive data and encouraging MLB organizations to consider richer metrics would likely be necessary. Overall, while the developed models provided useful forecasts, they highlighted persistent structural barriers and the importance of more sophisticated data usage in evaluating and promoting independent league talent.