Fall 2024 Baseball Analytics Projects Abstracts

Project #1: Creating a Frontier League Player Projection System
For: Yale University S&DS 425 (Statistical Case Studies) Semester-Long Capstone Project

Abstract: This project is developing two models using Yakkertech tracking data of the Frontier League, in conjunction with supplemental Statcast data from the Florida State League (FSL). The first model is a logistic regression designed to determine the probability that players in the Frontier League will advance to higher levels of affiliated baseball leagues—Single-A, Double-A, Triple-A, or even Major League Baseball. The second model will project expected performance statistics for both pitchers and hitters at those higher levels. By analyzing data such as exit velocity, launch angle, pitch velocity, and spin rate, this approach uniquely integrates independent league data with historical advanced performance analytics in the FSL to provide new insights into player development and potential career trajectories.

Project #2: Quantifying a Catcher's Impact on Pitcher's Performance Over a Game
For: Yale School of Management MGT 817 (Sports Analytics) Semester-Long Project

Abstract: This project introduces the "Catcher Calling Statistic," a new metric designed to quantify a catcher's impact on a pitcher's performance over the course of a game. While the influence of catchers on pitching outcomes has been explored in previous research, there is a noticeable gap in recent literature addressing this topic. Traditional statistics often focus on a catcher's framing or throwing ability, neglecting other crucial contributions like game management, pitch selection, and strategic decision-making that affect pitcher performance, particularly in terms of runs allowed. The Catcher Calling Statistic aims to provide a more comprehensive assessment of a catcher's influence by utilizing an "expected runs allowed" model. This model predicts the number of runs a pitcher is likely to concede based on factors such as the opposing team's offensive strength, the pitcher's historical performance, and ballpark characteristics. By comparing these expected runs to the actual runs allowed when different catchers are behind the plate, the metric isolates the catcher's effect on pitcher performance. This approach seeks to enhance understanding of a catcher's role, aiding teams in optimizing pitcher-catcher pairings and improving overall performance.

Project #3: Data Challenge: Exploring Pitcher-Batter Dynamics Using New Bat Speed and Swing Length Data
For: Connecticut Sports Analytics Symposium (April 11-12 2025)

Abstract: This project aims to analyze newly available Major League Baseball data on bat speed and swing length to explore various aspects of the pitcher-batter interaction during an at-bat. Utilizing pitch-level data from Baseball Savant covering 346,250 plate appearances between April 2, 2024, and June 30, 2024, the study will incorporate relevant Statcast metrics alongside the new swing measurements. The objective is to investigate how bat speed and swing length

influence batter performance, pitcher strategies, and their mutual interactions. Potential analyses include examining the relationship between swing characteristics and plate discipline, assessing how these metrics affect a batter's decision to swing, and determining whether pitchers adjust their pitching approach based on a batter's swing profile. By tapping into this new data, the project hopes to uncover fresh insights into the game, helping us better understand player behavior and decision-making on the field.