Aaron Wu
STAT 410

A Comprehensive Statistical Analysis of Pitching Performance Trends in Major League Baseball

This project aims to undertake an extensive statistical analysis of pitching performance in professional baseball, examining data spanning the last year. Drawing from a compiled dataset that includes official MLB statistics, detailed Pitch Info data, and advanced metrics from MLB Statcast, this study will explore a wide array of variables, including pitchers' Earned Run Average (ERA), strikeout-to-walk ratios, pitch velocity and type effectiveness, situational success rates, and performances under varied game conditions.

My objectives are to analyze how different variables contribute to a lower ERA and other indicators of effective pitching, assess how external factors such as weather conditions, stadium characteristics, and game timing influence pitching outcomes, and explore how pitchers adjust their strategies in response to different game situations and opponent strategies.

The dataset will be sourced from several reputable platforms that provide comprehensive statistics and in-depth analytics, including seasonal and career data for a wide range of pitchers. This will ensure a robust framework for predictive modeling. Multiple linear regression models will be developed to predict continuous outcomes, such as ERA and WHIP (Walks + Hits per Inning Pitched), while logistic regression will be used to model binary outcomes like game wins and high-leverage strikeout success.

I could visualize these using scatter plots for visualizing pitch velocity and ERA, box plots to compare ERA distributions across different pitch types and pitcher demographics, and heat Maps for various pitching metrics and performance outcomes.

My predictors include

- Base ERA: Foundational metric of a pitcher's overall effectiveness.
- Opponents' Batting Average: Reflects the quality of competition.
- Pitch Type Frequency and Effectiveness: Assesses which pitch types lead to successful outcomes.
- Strikeout-to-Walk Ratio: Measures a pitcher's control and tactical execution.
- Situational Performance Metrics: Analyzes outcomes in different game scenarios, such as bases loaded or late innings.

By bridging detailed statistical analysis with practical insights into pitching strategy, this project intends to offer actionable recommendations for improving pitching strategies. The results will likely interest a broad audience, including team coaches, player development analysts, and sports statisticians, all of whom seek to enhance understanding and performance at the highest levels of baseball.