NCAA Pitching Dashboard Read Me

Matthew Adams

2025-07-17

# NCAA Pitching Analysis Dashboard

## Overview

This project consists of a comprehensive **Shiny dashboard** built to analyze, explore, and visualize NCAA Division I baseball **pitching data**. It is part of a broader effort to construct and maintain a **custom NCAA pitching dataset** that does not currently exist in a centralized or accessible format.

The app is designed for baseball analysts, researchers, and fans interested in evaluating historical and team-level pitching performance across multiple dimensions.

## Live Dashboard Prototype

➡️ [**Click here to explore the NCAA Pitching Analysis Dashboard**](https://matthew-apps.shinyapps.io/baseball_data_analysis/)  
*The app is hosted via* [*shinyapps.io*](https://www.shinyapps.io)*.*

## License

This project is proprietary and all rights are reserved by Matthew Adams.

Unauthorized copying, modification, distribution, or commercial use is strictly prohibited.  
For collaboration inquiries or licensing requests, contact: [mdadams626@outlook.com](mailto:mdadams626@outlook.com).

## Dashboard Features

### Leaderboard Tab

Quickly explore top pitching performers with built-in filters: - **Innings Pitched Slider**: Adjust the minimum and maximum IP threshold - **Top 5 SO (Team-Year & Player-Year)** - **Top 5 ERA (Team-Year & Player-Year)** - **Top 5 WHIP (Team-Year & Player-Year)**  
Displayed using interactive DataTables for easy comparison.

### Data Tab

Dive into the raw player-level pitching data: - **Filter by School & Year** (with select/deselect all toggle) - **Dynamic IP Slider** for custom filtering - **Reset Button** to return to default view - **Interactive Table** shows all available player records matching filter criteria

### Analysis Tab

Split into two panels for deep-dive reporting and plotting:

#### Report Tab

* Filter data by school
* Dynamically select which numeric metrics (e.g., ERA, WHIP, SO/9) to summarize
* Generate grouped **summary stats** by school and year (mean, median, standard deviation)
* Switch between raw data view and summary output

#### Plot Tab

* Select a school and variable to analyze over time
* Choose display mode:
  + **Raw Only** (mean value by year)
  + **Summary Only** (mean, median, and SD)
  + **Both**
* Output is rendered as a grouped **interactive bar chart** via Plotly

### Player Ratings Tab

Comprehensive system for evaluating pitchers across multiple dimensions: - **Customizable prior shrinkage factors** (Bayesian smoothing by IP) - **User-defined composite rating weights** for ERA, FIP, SO/9, BB/9, etc. - League-year averages and rescaling to **0–100 rating scale** - **Dynamic recalculation of ratings** with adjustable inputs - Two views: **Player Ratings** and full **Methodology Table** with raw/shrunk/rescaled stat breakdown - Expanded Info tab with full explanation of rating logic and design

## Data Structure

Data is loaded from a local directory (/Data/) and cleaned using the janitor package. Each file typically includes:

* school, year, name, ip, era, whip, bb9, so9, h9, hr9, etc.
* Only pitching stats are used. Hitting stats are excluded.

⚠️ **Note**: Large-scale data ingestion is ongoing. Full D1 coverage is a long-term goal.