

MSDS-460 Final Project

Major League Baseball Roster and Lineup Optimization



Northwestern
University

Group 5

Alexander Wang, Andrew Stevens, Joshua Roberts, Kris Heller

MSDS 460: Decision Analytics

December 5, 2021

Agenda

- Introduction
- Objectives and Challenges
- Assumptions and Methods
- Approach and Model
- Results and Sequels

Introduction

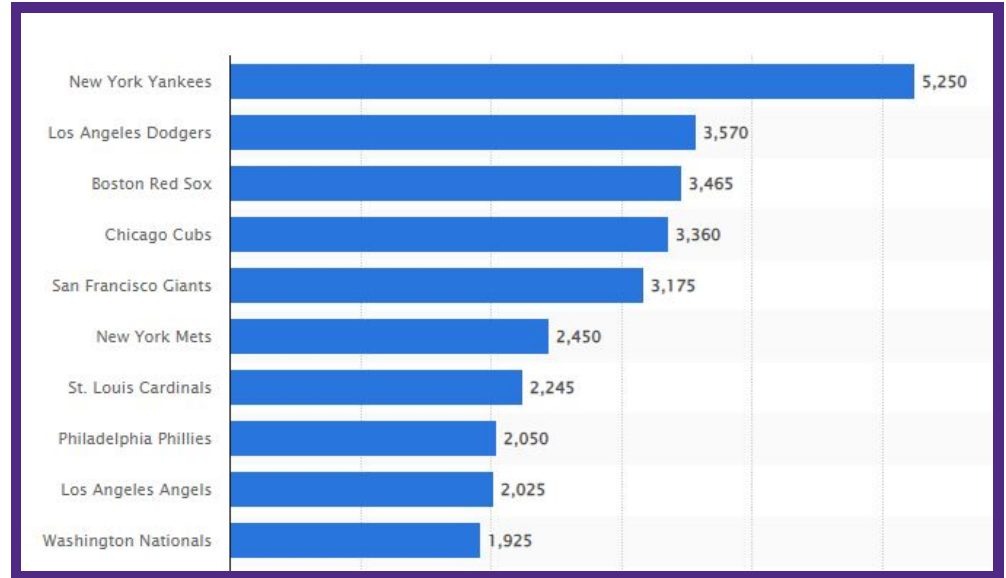
Introduction

- Group 5: Alex, Andrew, Josh, Kris
- SABRmetrics
 - Bill James
 - Society for American Baseball Research
- Workforce Optimization Versatility
 - Utilities, retail, financial services



Introduction

- MLB = big business
- Assets + revenue sources + expenses + teams = \$80.7B



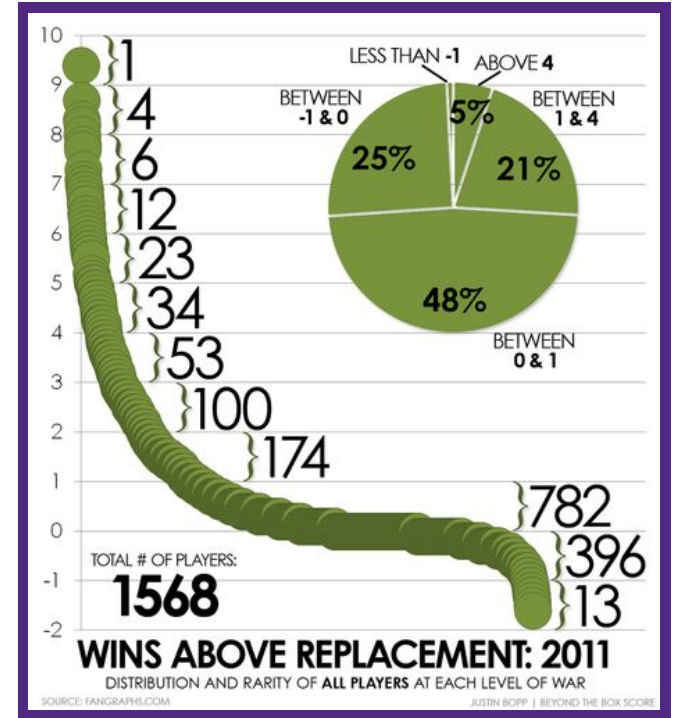
MLB Franchise values in \$M

Sources: [statista](#), [beyondtheboxscore.com](#)

Objectives and Challenges

Objectives and Challenges

- Build the best team
- Shoot for realism
- Maximize season WAR
- $\sum_{\text{Games } 1:162} \text{WAR}_{\text{player}} * \text{Contribution}_{\text{player}}$
- Challenges
 - Specialized Constraints
 - MIP due to integer & binary constraints
 - Sensitivity



Sources: beyondtheboxscore.com

Assumptions and Methods

Assumptions and Methods

- Anticipated Performance
 - Average Season vs. Player Trends
- Roster Size
 - 40 man vs. 26 man
- Full Game Simulation vs. Per-Inning Simulation
 - Fielders
 - Pitchers
- Static Roster
 - Trades + Rookies

Approach and Model

Approach and Model


- Variable Classes
 - Roster
 - Game
 - Innings
- Objective - Maximize Total WAR Contributions
- Constraints
 - Fatigue
 - Full Team Taking The Field
 - Salary Cap
 - Roster Size
 - Main Position
 - “Helpers”

Results and Future Work

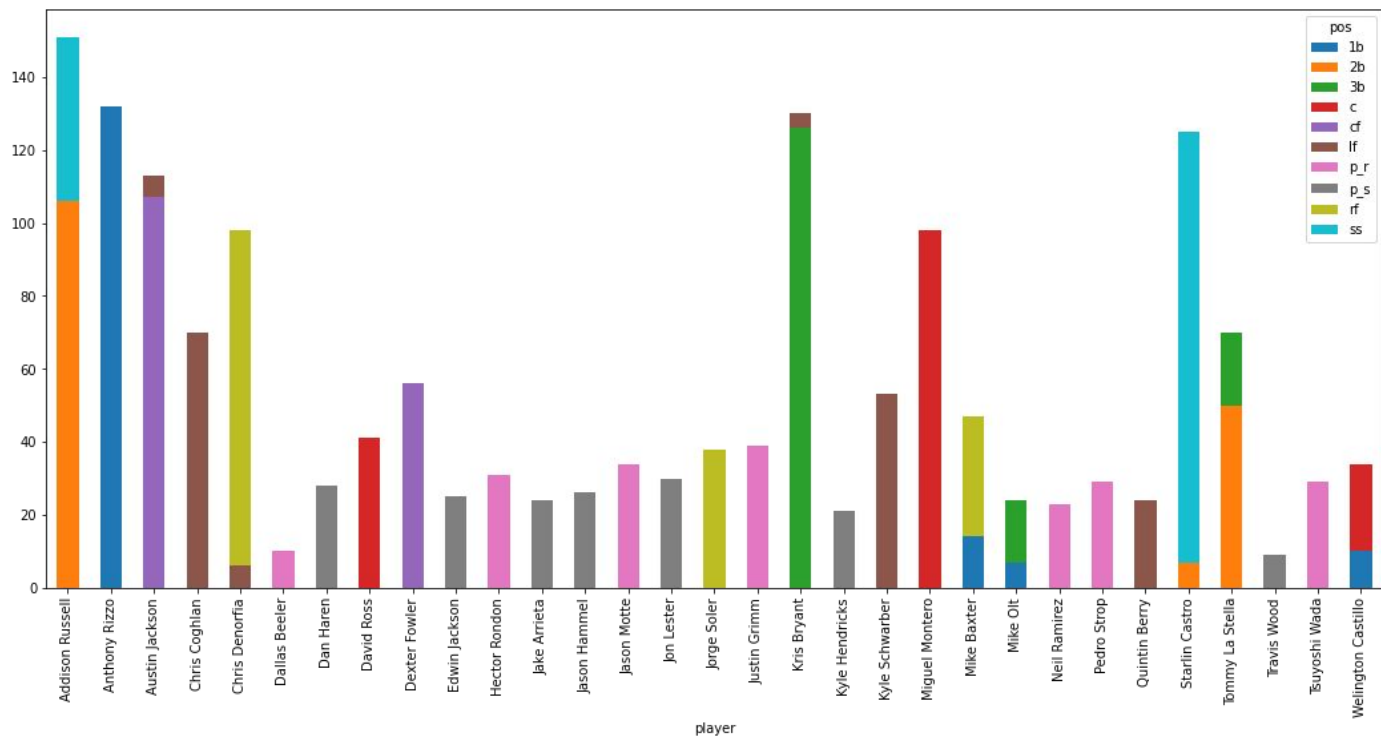
Results

- WAR of 53.66
- Calculated wins this season: $52 + 53.66 = 105.66$
- Absolute percentage error: 2.58%

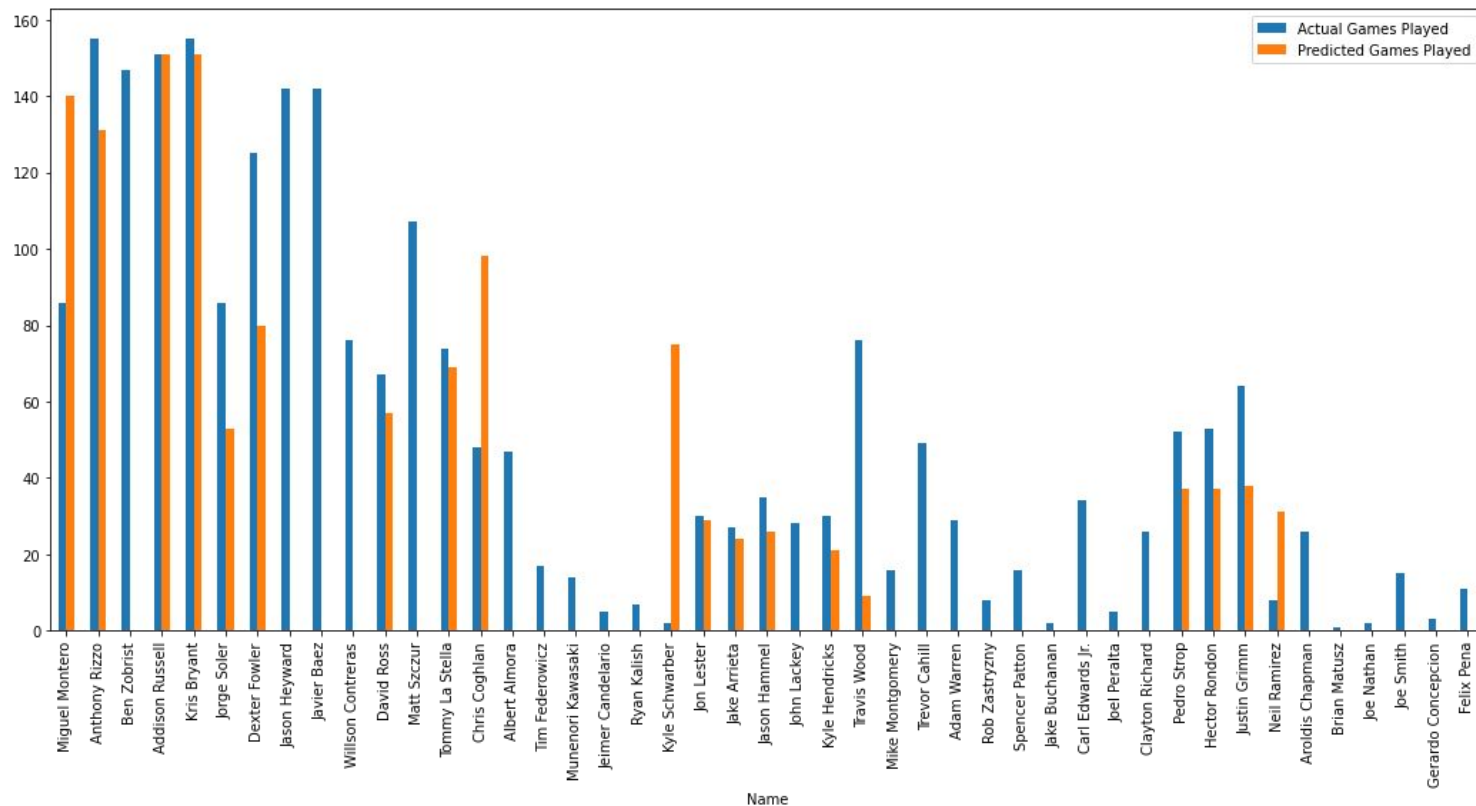
2016 Cubs Season Record

Team	W	L	Pct
 Cubs	103	58	.640

Results



Results



Sensitivity Analysis

- Tested sensitivity of a crucial player
 - Full season unavailability: No valid solution
 - Roster should be more robust
 - “Injury” season midpoint
 - 4.33 decrease in final season WAR

Table 2: Sensitivity Analysis

playerID	Name	Pos	Player WAR	Time injured	Season WAR
-	Baseline - all available		-	-	53.66
bryankr01	Kris Bryant	3B	6.1	Full Season	No Solution
bryankr01	Kris Bryant	3B	6.1	Half Season	49.33
lestejo01	Jon Lester	P	4.8	Half Season	52.6
solerjo01	Jorge Soler	RF	.50	Half Season	53.55
roachdo01	Donn Roach	P	-.14	Half Season	53.67

Future Work

- Other variables/factors:
 - Trading players
 - Better statistic or metric to gauge rookies or older players
 - Cost benefit analysis for players
 - Sensitivity analysis for salary cap

