Improving Baseball Team Performance and Fan Engagement

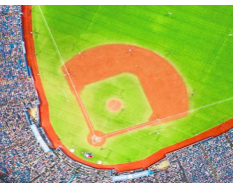## A General Manager's Perspective

Problem Statement

General Manager of a Last-Place Baseball Team

- Improve team performance
- Enhance fan experience

Constrained by Limited Resources ($$)
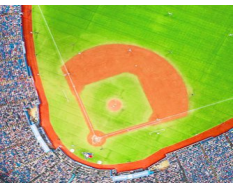
- Small market team
- Unlikely to sign superstars

Strategy

Team Performance:

- Identify lesser-known / "under the radar" players
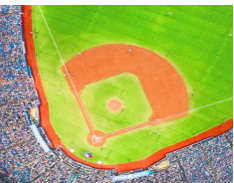- Not superstars; but good enough

Enhance Fan Experience:

- App that shows the home team's win probability

Key Questions

- What features most strongly correlate with team wins?

- Which players should I add to increase team wins?

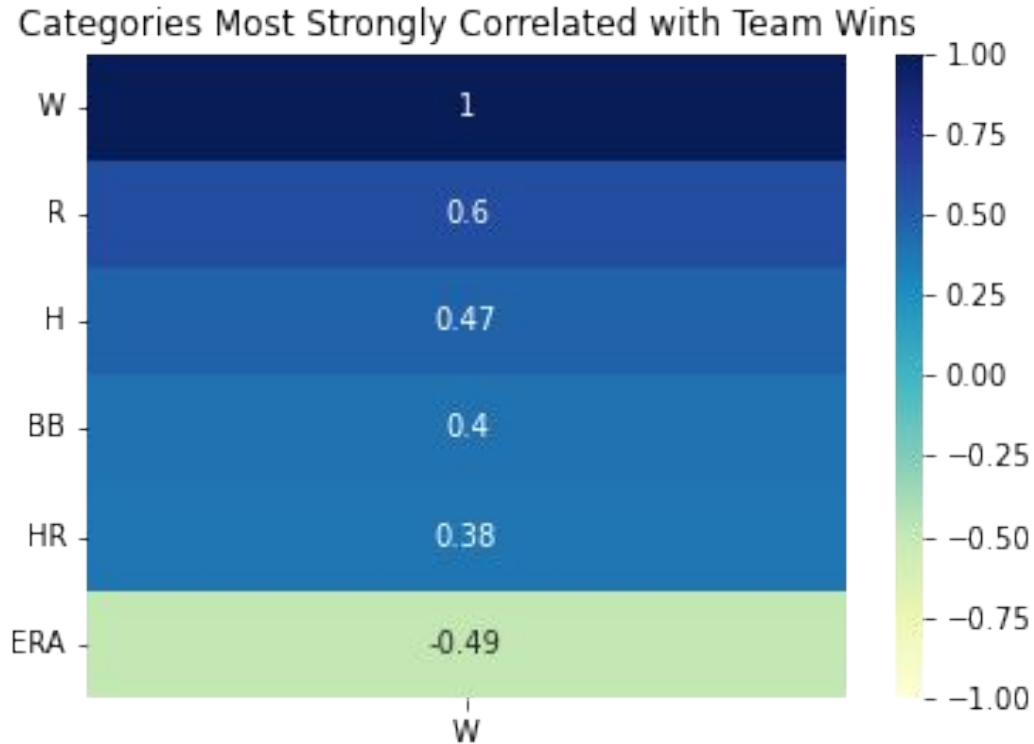# The Data

Lahman Baseball Database

-   Offensive, defensive, and pitching stats for every player and team, 1871 - 2021.

Retrosheet Game Logs

-   Record of every Major League game played
-   161 features in each record

Analyzed data after 1945 (start of Modern Baseball Era)

Categories Most Strongly Correlated with Team Wins

	W
W	1
R	0.6
H	0.47
BB	0.4
HR	0.38
ERA	-0.49
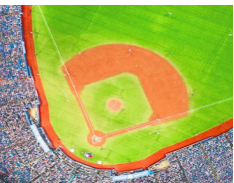
Features Correlated with Team Wins

- Positively correlated: runs (R), hits (H), walks (BB), homeruns (HR)

- Negatively correlated: Earned Run Average (ERA)

Number of Team Wins vs. Number of Above Average Players Per Team (2021 Season)
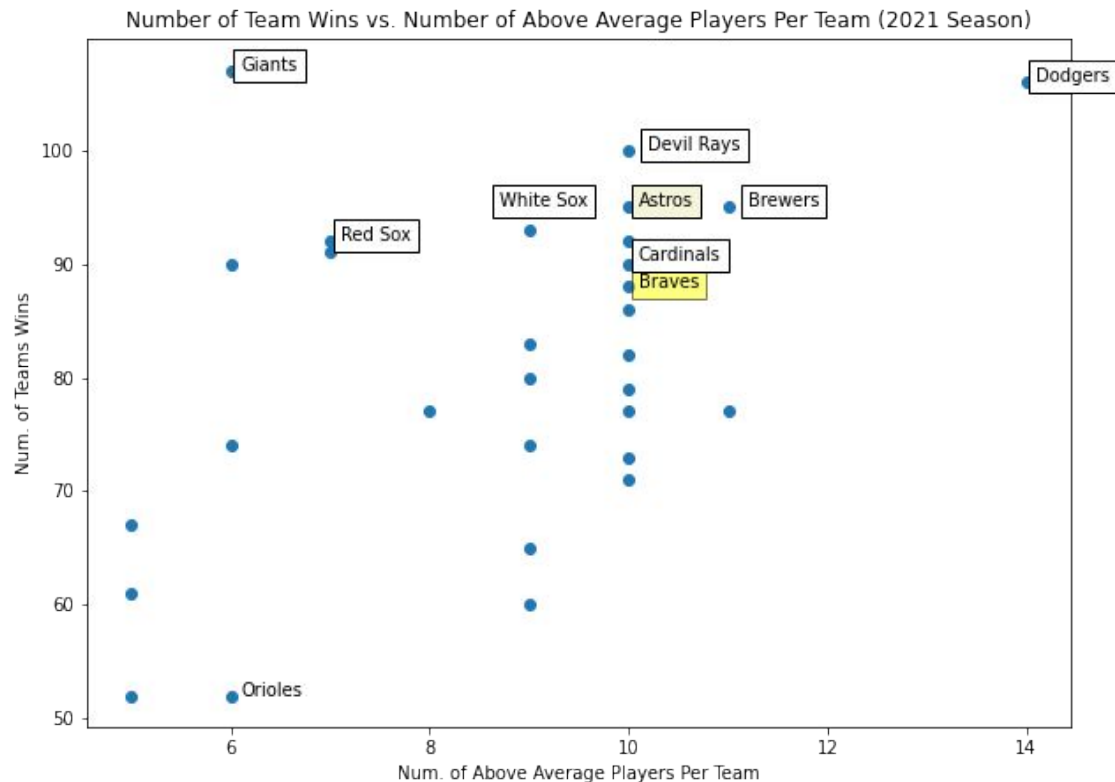
Team Wins and # of Above Average Players Per Team (2021)

- Generally a positive relationship

- Teams with more above average players have more wins

- But not always the case (For example: Giants and Orioles)

Potential Batters to Add

Lesser-known hitters whose average numbers in aforementioned offensive categories over the past five seasons have been better than the league average.

Ryan McMahon #24
3B | B/T: L/R | 6' 2"/219 | Age: 27

FOLLOW

Ryan Patrick McMahon
Nickname: Ry Mac
Born: 12/14/1994 in Yorba Linda, CA
Draft: 2013, Colorado Rockies, Round: 2, Overall Pick: 42

| Summary | Stats | Charts | News | Awards | SHOP |

Year	AB	R	H	HR	RBI	SB	AVG	OBP	OPS
2022 Regular Season	529	67	130	20	67	7	.246	.327	.741
Career Regular Season	1909	259	466	81	282	19	.244	.324	.751

Yoán Moncada #10
3B | B/T: S/R | 6' 2"/225 | Age: 27

FOLLOW

Yoán Manuel Moncada
Nickname: Yoyo
Born: 5/27/1995 in Cienfuegos, Cuba
Debut: 9/02/2016
Follow @whitesox

| Summary | Stats | Charts | News | Awards | SHOP |

Year	AB	R	H	HR	RBI	SB	AVG	OBP	OPS
2022 Regular Season	397	41	84	12	51	2	.212	.273	.626
Career Regular Season	2424	333	613	82	299	30	.253	.334	.759

Travis Shaw #23
3B | B/T: L/R | 6' 4"/230

FOLLOW

Travis R. Shaw
Nickname: Mayor-DDC
Born: 4/16/1990 in Washington Court House, OH
Draft: 2011, Boston Red Sox, Round: 9, Overall Pick:

| Summary | Stats | News | Awards | SHOP |

Year	AB	R	H	HR	RBI	SB	AVG	OBP	OPS
2022 Regular Season	19	0	0	0	0	0	.000	.000	.000
Career Regular Season	2374	310	563	114	366	20	.237	.319	.756

Potential Pitchers to Add

Cal Quantrill #47
P | B/T: L/R | 6' 3"/195 | Age: 27
FOLLOW

Summary   Stats   Charts   News   Awards   SHOP

Logan Webb #62
P | B/T: R/R | 6' 1"/220 | Age: 25
FOLLOW

Summary   Stats   Charts   News   Awards   SHOP

Adrian Houser #37
P | B/T: R/R | 6' 3"/222 | Age: 29
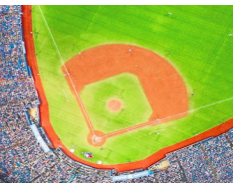FOLLOW

Summary   Stats   Charts   News   Awards   SHOP

Lesser-known starting pitchers whose average ERAs over the past five seasons have been better than the league average.

Models

- Trained Classification Models:
    - Data from 146,691 games (1946-2021)
    - Predict whether the home team will win or lose

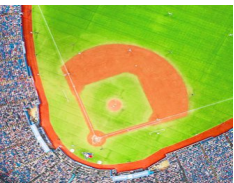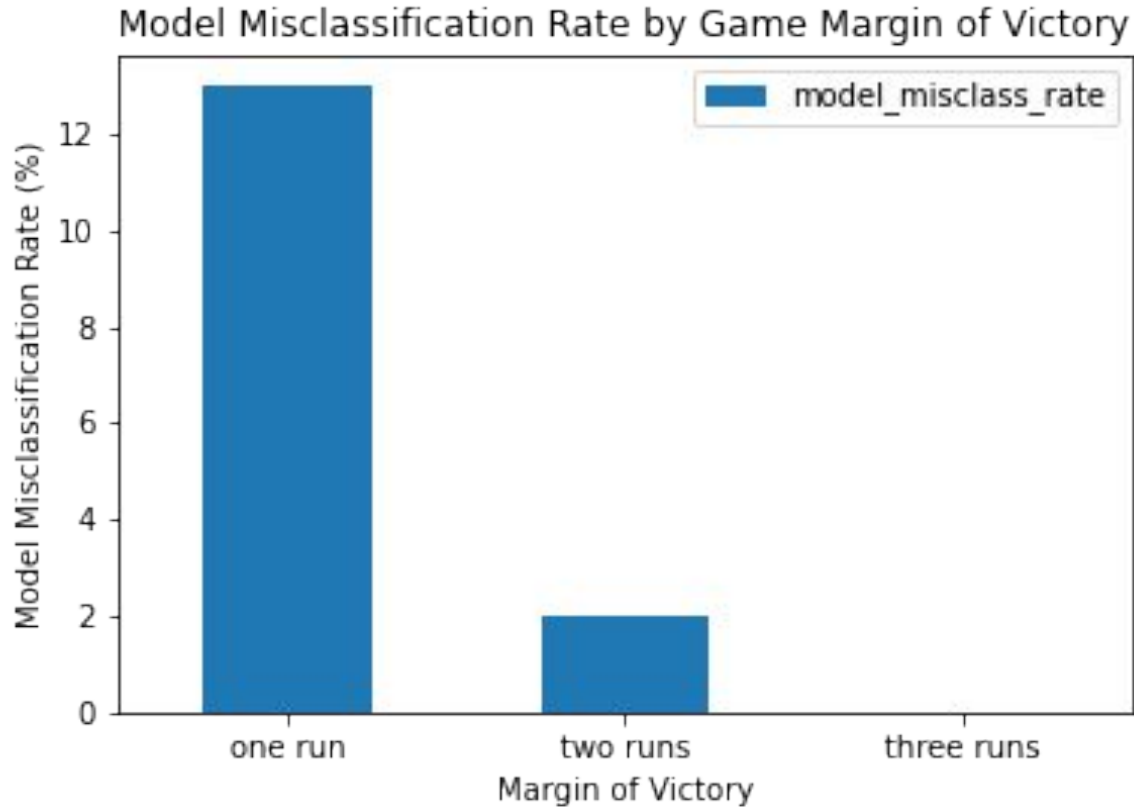- 14 Predictor Variables: 7 for the home team, 7 for the visiting team

Predictor Variables

Best model: Logistic Regression; predicted with 95% accuracy the winning team based on these variables for each team:

- Hits

- Walks

- Left on base

- Caught stealing

- Double plays

- Errors

- Batters hit-by-pitch

Model Misclassification Rate by Game Margin of Victory

Logistic Regression Model Evaluation

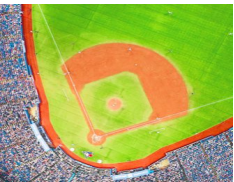- Misclass rate: higher for close games

Findings/Recommendations

- Talent helps you win, but not always guaranteed.
    - Other factors to consider: injuries to key players, quality of coach, team chemistry

- Recommend pursuing the "lesser-known" players.

    - Risks: Shorter track record of sustained performance, greater chance of "bust"
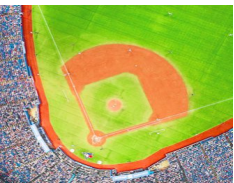
Thank you for your time!

References/Resources/Sources

Lahman Baseball Database https://www.seanlahman.com/baseball-archive/statistics/

Retrosheet Game Logs https://www.retrosheet.org/gamelogs/index.html

Baseball Rules https://www.youtube.com/watch?v=57mc7Df7Arw

Baseball Glove Image
https://depositphotos.com/122268474/stock-photo-baseball-bat-with-glove-and.html

Baseball Diamond Image
https://www.slidemembers.com/en_US/view/Free-Slides/simple-google-slides-templates-baseball-game-3504

Player Information https://www.mlb.com/players