

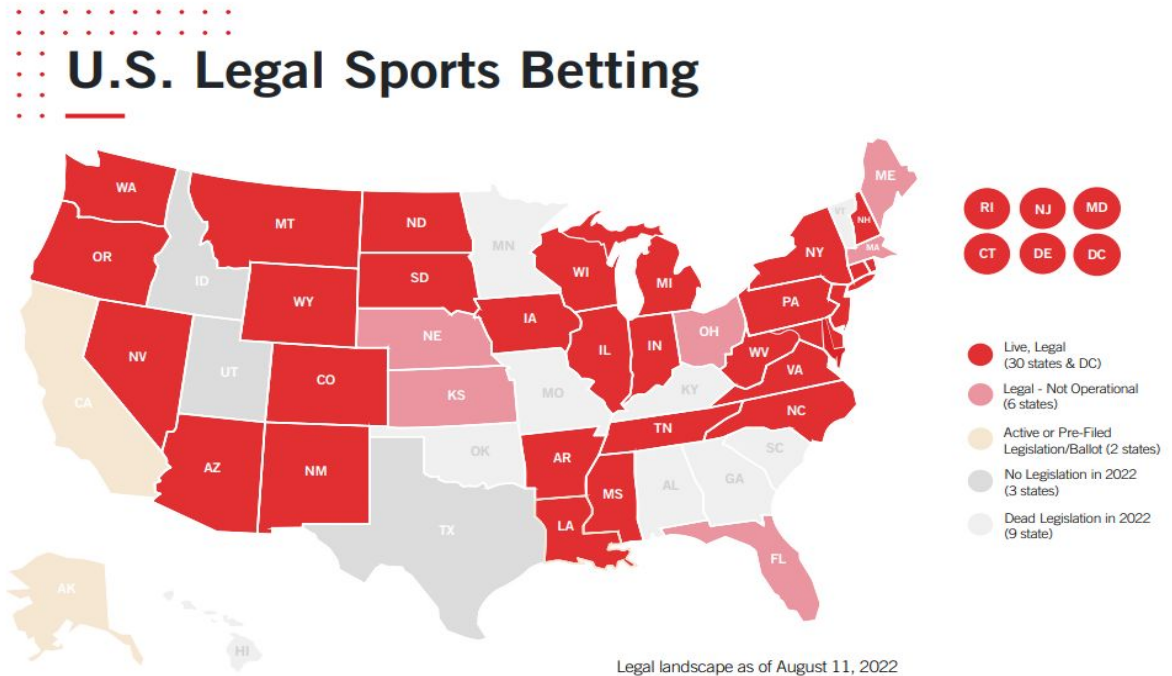


# MLB Betting Model

Ryan Pedersen

# Background Info

- Professional and Amateur Sports Protection Act passed in 1992 that made sports gambling illegal in all states except Nevada
- Federally legalized in 2018 by the Supreme Court
- Sports betting has grown massively in popularity the past few years.



# Betting Lines Example

		Runline	Win	Total
8/16/22 3:40 PM	Philadelphia Phillies Kyle Gibson (R) ————— + 110 Bets >	-1.5 (-120)	-180	O 9.0 (-115)
	Cincinnati Reds T.J. Zeuch (R)	+1.5 (EVEN)	+148	U 9.0 (-105)
8/16/22 3:40 PM	San Diego Padres Sean Manaea (L) ————— + 110 Bets >	-1.5 (+115)	-155	O 7.0 (-125)
	Miami Marlins Edward Cabrera (R)	+1.5 (-135)	+130	U 7.0 (+105)
8/16/22 4:05 PM	Boston Red Sox Nick Pivetta (R) ————— + 111 Bets >	-1.5 (+120)	-145	O 8.5 (-105)
	Pittsburgh Pirates Mitch Keller (R)	+1.5 (-140)	+122	U 8.5 (-115)

# Process

**01**

---

## Data Acquisition

How all of the data was gathered

**02**

---

## Modeling

Type of model used, features

**03**

---

## Utilization

What to do with the predictions



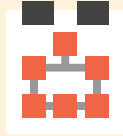
**01**

---

# Data Acquisition



# Data Sources



## Fangraphs

Used pybaseball wrapper to pull game logs, batting and pitching stats



## MLB.com

Scraped mlb.com probable pitchers page to get daily games and the starting pitchers.



## Betting sites

Used the-odds-api to pull odds from various betting sites



# Sample of Statistics Used

Stat name	Description	Player type
BA	Hits divided by at-bats	Batter
wRC+	Weighted runs created plus	Batter
ERA	Earned Runs allowed per 9 innings	Pitcher
WHIP	Walks plus hits per inning pitched	Pitcher

# 02 Modeling





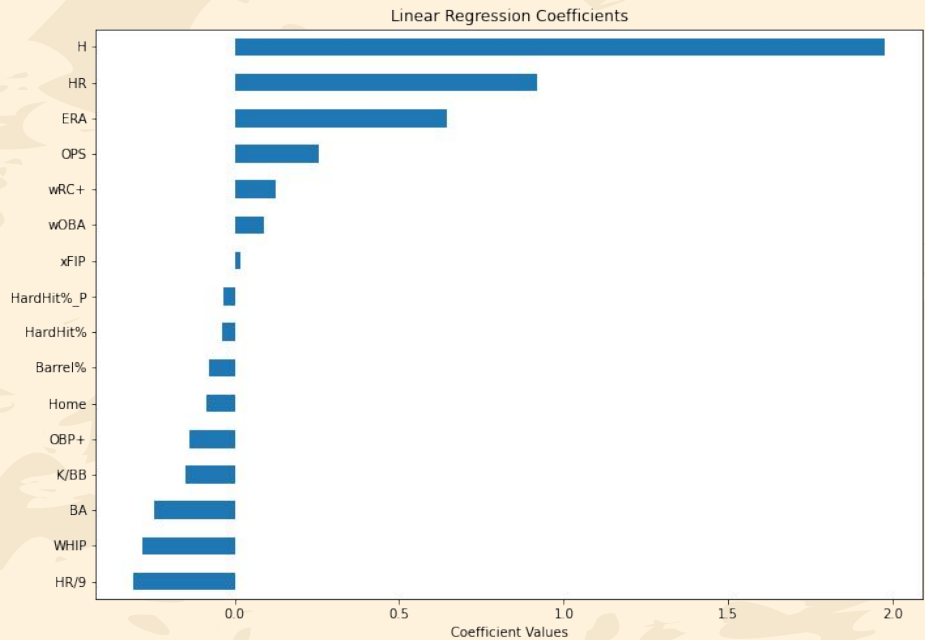
# Modeling Info

- Model chosen: Linear Regression
- Random Forest also considered
- 21 features used
- Use game logs from Fangraphs to train model



# Model Metrics

Metric	Value
Training r2	0.69
Testing r2	0.71
Testing MSE	2.95

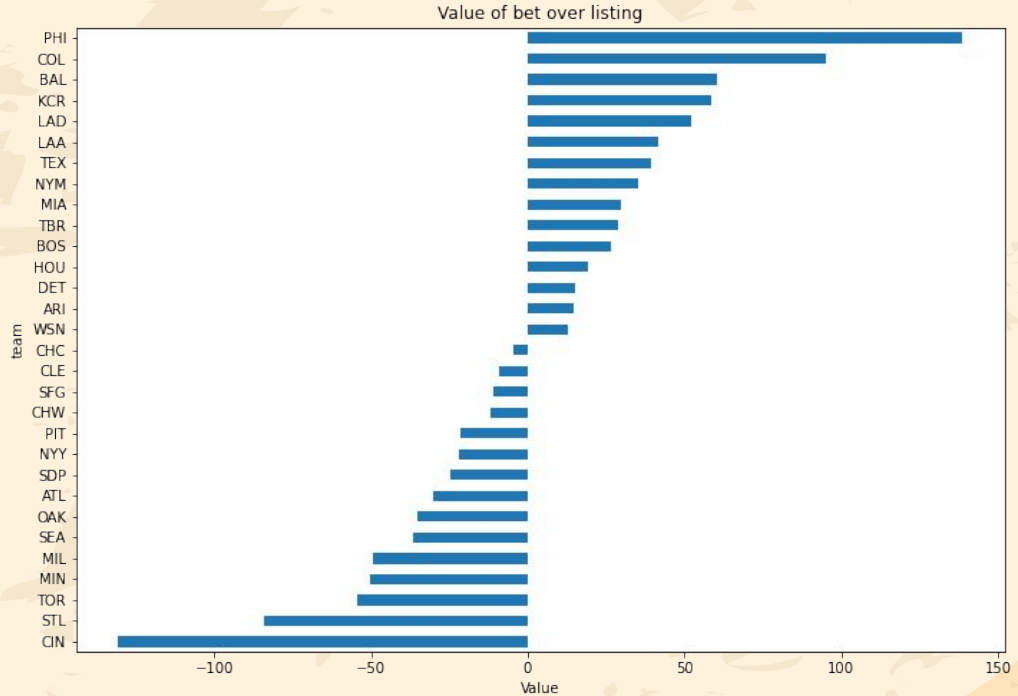


# 03 Utilization



# Moneyline bets

- Scrape daily games and probable pitchers from mlb.com
- Build prediction dataframe
- Get predicted scores
- Generate odds
- Compare to sportsbook odds pulled from the-odds-api







**7/10 Bets Won**

The background features a large, light-yellow oval with a cracked, stone-like texture. The entire scene is set against a light-orange background with white diagonal lines and several small red stars scattered around.

So far.....



# Conclusion and Next Steps

- Time is needed to fully understand the performance of the model relative to bookmakers
  - Early returns are good, sample size too small.
  - Generate daily predictions
  - Assess model performance over the course of a month
  - Add more features
  - Player value coefficient
- 
- 

# Thanks!

Any questions?

CREDITS: This presentation template was created by Slidesgo, including icons by Flaticon and infographics & images by Freepik

Please keep this slide for attribution

