

# War against WAR: Applying Machine Learning to Baseball Analytics

George Briggs, Matthew Loomis, Otis Peterson

# INTRO - ROADMAP

*$WAR = (Batting\ Runs + Base\ Running\ Runs + Fielding\ Runs + Positional\ Adjustment + League\ Adjustment + Replacement\ Runs) / (Runs\ Per\ Win)$*



Here are the full FanGraphs positional adjustments used in WAR:

Catcher: +12.5 runs (all are per 162 defensive games)  
First Base: -12.5 runs  
Second Base: +2.5 runs  
Third Base: +2.5 runs  
Shortstop: +7.5 runs  
Left Field: -7.5 runs  
Center Field: +2.5 runs  
Right Field: -7.5 runs  
Designated Hitter: -17.5 runs

# ML Methods

- Linear Regression
- Random Forest Regression
- KNN Regression
  - Low Dimensional
  - Control Variation



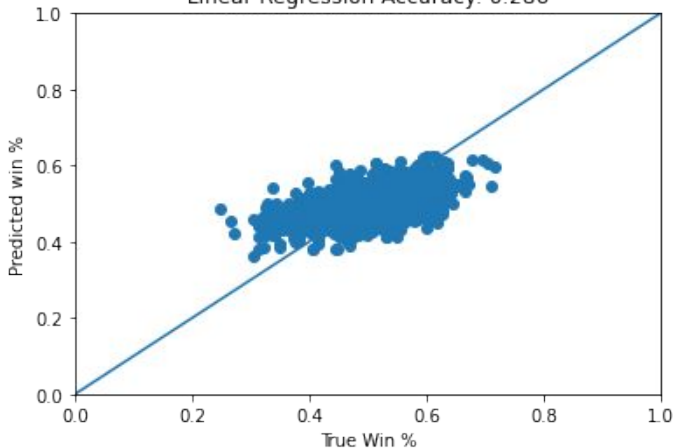
# Main Stat

- Stats we settled on
- Weights generate make sense
  - SB vs CS is an interesting comparison
- Regression struggled to predict extremes
- WAR predicts similarly, with exceptions

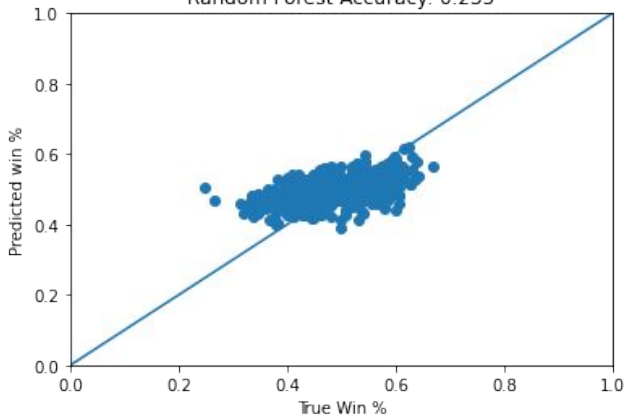
## Weights:

```
((0.1437930963450888, 'H'),  
(0.00981612120859117, '2B'),  
(0.04565299228158608, '3B'),  
(0.19303170024445593, 'HR'),  
(0.09924158594991374, 'BB'),  
(-0.03904679363780258, 'SO'),  
(-0.07461915212315173, 'CS'),  
(0.15227352843513156, 'SB'))
```

Linear Regression Accuracy: 0.286

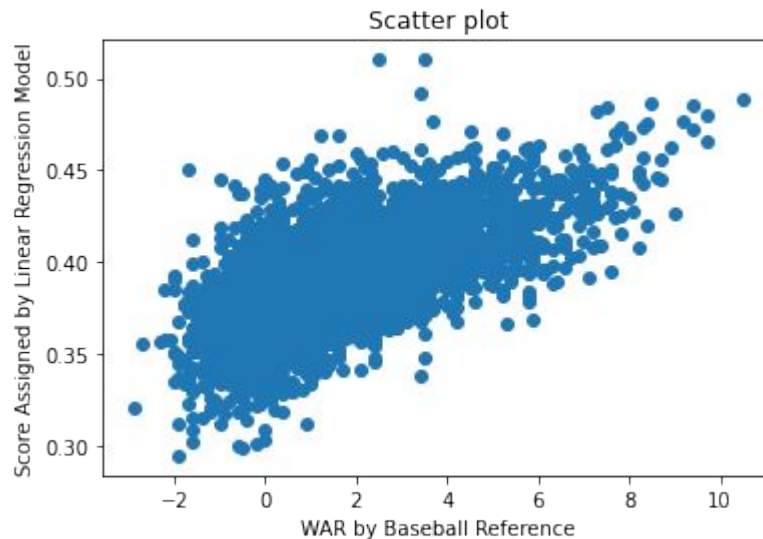


Random Forest Accuracy: 0.235



# Main Stat

- Stats we settled on
- Weights generate make sense
  - SB vs CS is an interesting comparison
- Regression struggled to predict extremes
- WAR predicts similarly, with exceptions



# Pitching Extension

- Similar to hitting model
- Stats: H, K, BB, HBP, ER
- Takeaways

Difference in accuracy

Importance of Earned Runs



Pictured: Gerrit Cole

Figure 1

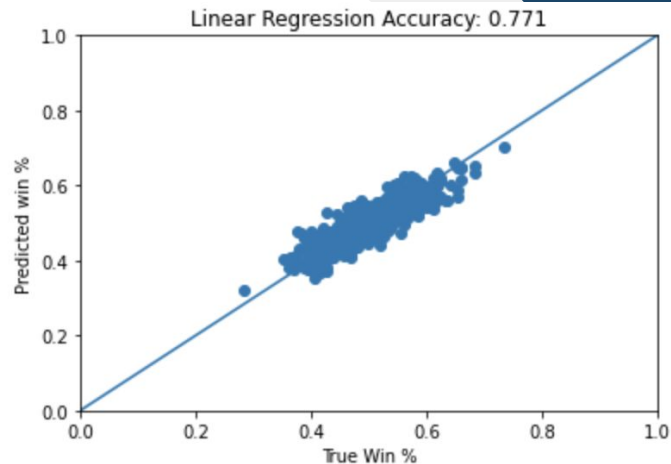
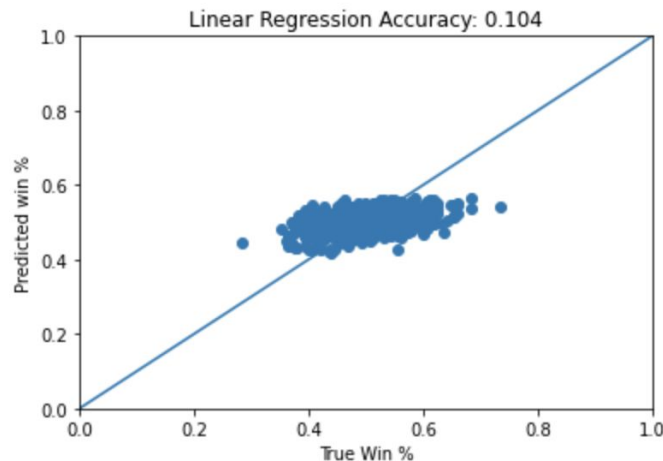


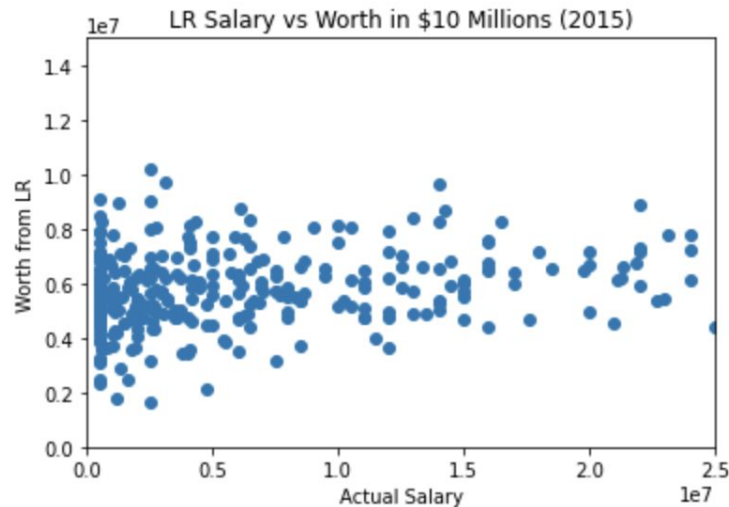
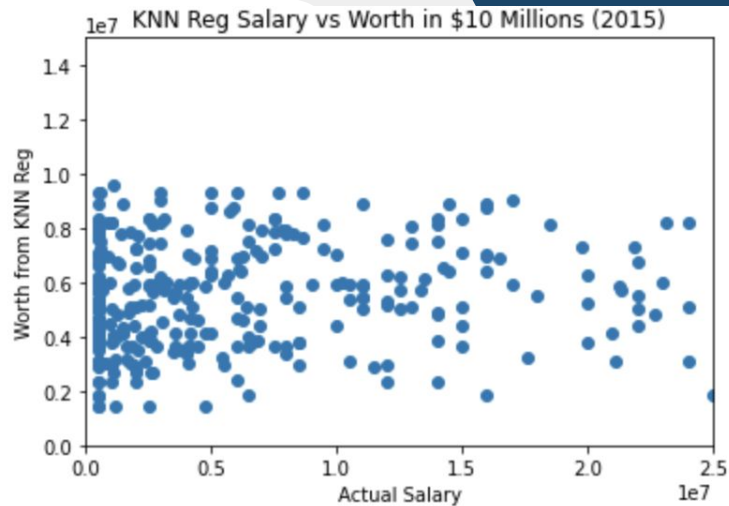
Figure 2



# Salary

- Data importing/cleaning
- Linear Regression & KNN Regression
- Correlation Coefficient
- Interpreting Results

	score	playerID	yearID	sal
7659	0.419911	aguaylu01	1985	237000
7660	0.371442	almonbi01	1985	255000
7662	0.401720	armasto01	1985	915000
7663	0.415972	ashbyal01	1985	416667
7664	0.392923	backmwa01	1985	200000
...	...	...	...	...
18736	0.373947	youngde03	2015	2250000
18737	0.376446	zimmer01	2015	14000000
18738	0.383162	zobribe01	2015	7500000
18739	0.393640	zobribe01	2015	7500000
18740	0.320524	zuninmi01	2015	523500





## Stakeholders

- Players
- Management
- Fans

## Future Research

- More Comprehensive Model
- Predictive Statistics
  - Draft Picks
  - Contract Offers



### Trevor Bauer signs 3-year, \$102 million contract with Dodgers

Reigning NL Cy Young Award winner joins reigning World Series winner





**THANK YOU FOR LISTENING**

**HOW CAN YOU NOT BE ROMANTIC ABOUT BASEBALL**