

# A Machine Learning Approach to MLB Catcher Framing

Nathan Hemenway and Matthew Boyd

12/10/2021

## Introduction

- ▶ “Catcher framing is the art of a catcher receiving a pitch in a way that makes it more likely for an umpire to call it a strike – whether that’s turning a borderline ball into a strike, or not losing a strike to a ball due to poor framing.” - MLB.com Glossary

## Motivation

- ▶ Baseball catchers can influence the call of a ball or strike on how they catch the ball
- ▶ Some catchers are better than others at this skill
- ▶ Baseball teams are aware of this and are acquiring players good at this skill to win more games
- ▶ We want to quantify the best catcher's at framing for the 2021 season
- ▶ There are several factors that influence whether a pitch will be a strike or ball
- ▶ Catchers getting more strikes translates to more outs, and fewer points for the opposing team

## Data

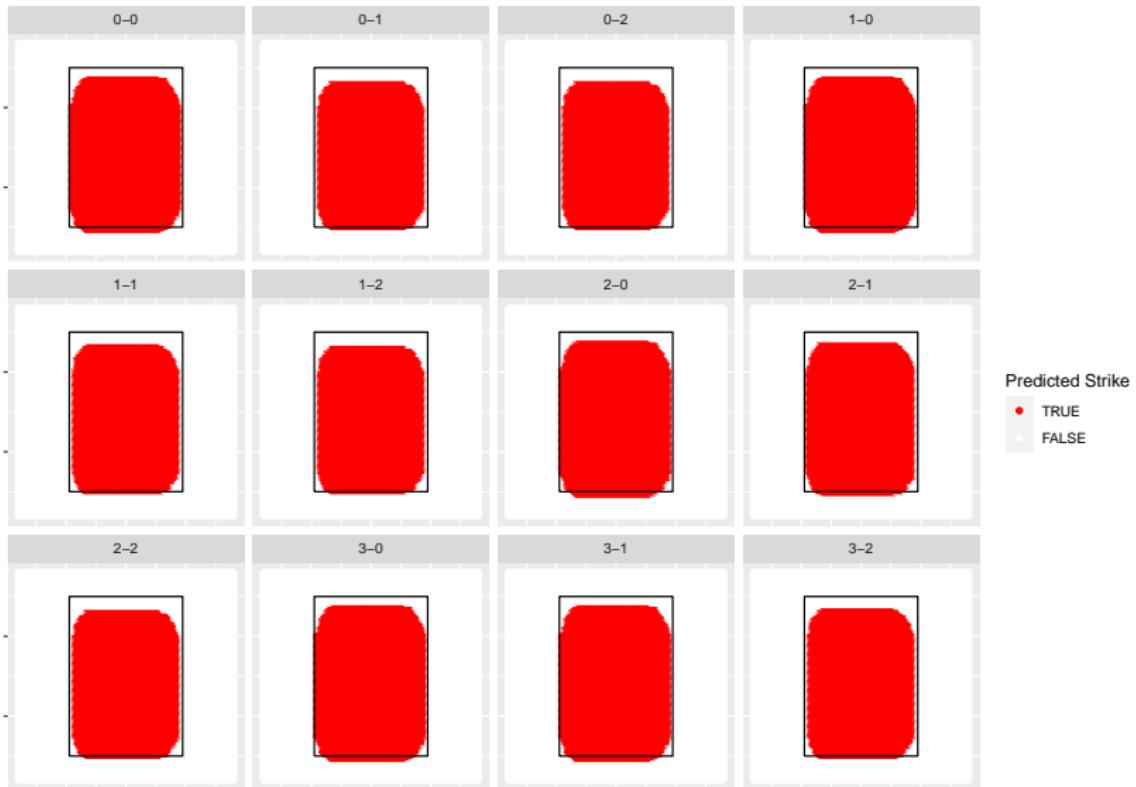
- ▶ 2021 pitch data scraped from Baseball Savant through baseballr package
- ▶ ~700,000 rows (each for a single pitch)
- ▶ Wanted to look at pitches that were not swung at by the batter (called strike or ball)
- ▶ ~350,000 rows remain

## Variables

- ▶ We used:
- ▶ Pitch type and pitch release speed, position, and spin rate
- ▶ Whether or not the pitcher and batter are right or left handed
- ▶ Count, number of outs during the at-bat, and inning number
- ▶ Where the pitch landed
- ▶ Whether the game was played home or away
- ▶ How tall the batter is

# Example

## ► Strike Probability by Location and Count



# Logistic Regression Model

```
##  
## Call:  
## glm(formula = strike ~ ., family = "binomial", data = data_no_catchers[train,  
##   ])  
##  
## Deviance Residuals:  
##       Min      1Q  Median      3Q     Max  
## -1.6091 -0.9918 -0.6061  1.2274  2.4843  
##  
## Coefficients:  
##                               Estimate Std. Error z value Pr(>|z|)  
## (Intercept)           1.865e+00  6.897e-01  2.705  0.006835 **  
## pitch_typeCurveball  4.188e-01  3.862e-02 10.845 < 2e-16 ***  
## pitch_typeCutter    2.829e-01  3.182e-02  8.890 < 2e-16 ***  
## pitch_typeFastball  4.249e-01  2.865e-02 14.829 < 2e-16 ***  
## pitch_typeSinker   6.128e-01  2.705e-02 22.652 < 2e-16 ***  
## pitch_typeSlider   3.491e-01  2.818e-02 12.392 < 2e-16 ***  
## release_speed      -1.018e-02  1.791e-03 -5.682 1.33e-08 ***  
## release_pos_x       8.904e-03  7.481e-03  1.190  0.233984  
## release_pos_z      -3.874e-02  1.075e-02 -3.604  0.000313 ***  
## standR              7.218e-02  1.150e-02  6.275 3.50e-10 ***  
## p_throwsR            3.429e-02  3.134e-02  1.094  0.273856  
## count0-1             -9.788e-01  1.820e-02 -53.783 < 2e-16 ***  
## count0-2             -2.091e+00  3.570e-02 -58.585 < 2e-16 ***  
## count1-0             -1.136e-01  1.686e-02 -6.741 1.57e-11 ***  
## count1-1             -7.870e-01  2.007e-02 -39.217 < 2e-16 ***  
## count1-2             -1.835e+00  2.916e-02 -62.922 < 2e-16 ***  
## count2-0             6.662e-02  2.597e-02  2.565  0.010309 *  
## count2-1             -5.223e-01  2.608e-02 -20.029 < 2e-16 ***  
## count2-2             -1.481e+00  3.003e-02 -49.318 < 2e-16 ***  
## count3-0             7.445e-01  3.688e-02 20.183 < 2e-16 ***  
## count3-1             -2.489e-01  3.656e-02 -6.809 9.85e-12 ***  
## count3-2             -1.045e+00  3.694e-02 -28.281 < 2e-16 ***  
## pfx_x                -2.131e-03  7.356e-03 -0.290  0.771989  
## pfx_z                2.260e-02  1.539e-02  1.468  0.142011
```