Swing Science: Using 2024 Player Data to Predict 2025 Batting Averages

May Mishima and Rex Schiffmann (Group R)

12-15-2025

Abstract

This study investigates whether commonly used offensive performance metrics from the 2024 Major League Baseball season can be used to predict player batting averages in the 2025 season. Using data from all qualified MLB hitters in 2024, we examined the relationships between batting average and measures of contact quality, swing outcomes, and plate discipline. Exploratory analysis revealed strong linear associations between batting average and total hits, as well as moderate negative associations with strikeout-related measures. Multiple linear regression models were fitted and compared using diagnostic checks, variable selection procedures, and nested model testing to identify an efficient and effective predictive model. The final model included total hits, strikeout percentage, and high-quality swing counts, explaining approximately 68% of the variation in batting average. When applied to 2025 data, the model demonstrated reasonable predictive accuracy, suggesting that prior-season contact and strikeout tendencies provide meaningful information about future batting performance.

Background & Meaning

Batting average remains one of the most widely recognized statistics used to evaluate offensive performance in baseball, despite the growth of more advanced metrics. Batting average reflects a player's ability to convert plate appearances into hits, making it a natural outcome of both contact skill and swing quality. Modern tracking technologies, such as Statcast, allow analysts to decompose batting outcomes into measurable components related to swing mechanics, contact quality, and plate discipline, offering new opportunities to understand what drives variation in batting average.

Players who strike out less and make more consistent contact tend to sustain higher batting averages, while measures such as hard-hit rate and barrel frequency capture the quality of that contact. However, many of these metrics are correlated with one another, raising questions about which measures provide independent explanatory value. Additionally, while year-to-year stability in batting average is imperfect, prior-season performance remains a strong predictor of future outcomes.

The goal of this study is to determine whether a small set of offensive metrics from the 2024 season can meaningfully predict batting average in 2025. Specifically, we aim to identify which aspects of offensive performance contribute most to explaining variation in batting average, while avoiding unnecessary model complexity.

Methods

Data

The data consist of Major League Baseball hitters who met the minimum qualification threshold for plate appearances during the 2024 season (an average of 3.1 per scheduled game). There were 129 players who qualified for this dataset. All observations were obtained from publicly available Statcast leaderboards hosted by Baseball Savant. Because the dataset includes all qualified players rather than a random sample, the analysis is descriptive of this population. It focuses on modeling relationships rather than inference to a larger group. A separate dataset containing 2025 batting averages for the same players was used exclusively for our 2025 batting average prediction.

Variables

This study uses one response variable and several explanatory variables drawn from Statcast-based offensive performance metrics. Variable names from the dataset are reported once for transparency and reproducibility, followed by plain-language descriptions used throughout the remainder of the report.

Response Variable:

batting_avg (Batting Average)

Batting average is defined as the proportion of official at-bats that result in hits. This is the primary outcome of interest because it is a widely recognized summary measure of offensive success and directly reflects a hitter's ability to convert plate appearances into hits. All modeling efforts in this study aim to explain variation in batting average across qualified Major League hitters.

Explanatory Variables:

hit (Total Hits)

This variable measures the total number of hits recorded by a player during the 2024 season.

k_percent (Strikeout Percentage)

Strikeout percentage represents the proportion of plate appearances that end in a strikeout.

whiff_percent (Whiff Percentage)

Whiff percentage measures how often a player swings and misses when attempting to make contact.

blasts_swing (Blast Swings)

This variable counts swings in which the ball is struck with both high bat speed and optimal contact quality. This variable is calculated using bat speed and the percentage of squared-up swings.

squared_up_swing (Squared-Up Contact Swings)

Squared-up swings measure instances where the ball is struck on the barrel of the bat.

hard_hit_percent (Hard-Hit Rate)

Hard-hit rate represents the percentage of batted balls hit at high exit velocities.

barrel_batted_rate (Barrel Rate)

Barrel rate measures the percentage of batted balls that meet an optimal combination of exit velocity and launch angle.

bat (Handedness)

Handedness is a categorical variable indicating whether a player bats left-handed, right-handed, or as a switch hitter.

Statistical Methods

Exploratory data analysis was conducted using univariate density plots and bivariate scatterplots to assess distributional shape, linearity, and potential outliers. Associations between predictors were examined using correlation matrices to identify potential multicollinearity.

Modeling proceeded in stages. First, a one-way ANOVA was used to test whether batting average differed by handedness, with a randomization-based F-test employed due to unequal group variances. Next, multiple linear regression models were fit using combinations of offensive metrics. A full model containing all candidate predictors was evaluated for significance, multicollinearity (as indicated by variance inflation factors), and model assumptions. Variable selection procedures and nested model comparisons were used to identify a parsimonious final model that balanced explanatory power and interpretability.

Model diagnostics included residual plots, normal Q-Q plots, and Cook's distance to assess homoskedasticity, normality, and influence. Finally, the selected model was applied to 2025 data to generate predicted batting averages, and predictive performance was assessed using mean squared error and predicted-versus-actual plots.

Results

Univariate Analysis

In the first step of exploratory data analysis (EDA), each individual variable was analyzed to visualize distributions and note values that could skew the data unreasonably. The quantitative variables were analyzed through density plots and the categorical variable was analyzed with a bar graph.

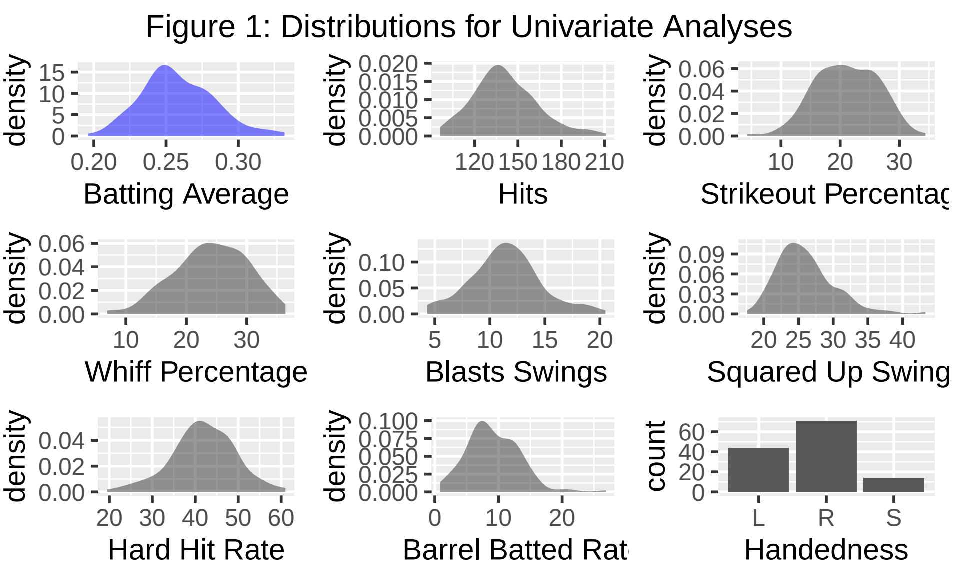


Figure 1:

Batting Average (batting_avg): The distribution of batting average is relatively normally distributed with tendencies of a slight right skew. The mean batting average of the 129 players who qualified in 2024 is 0.258 with a standard deviation of 0.03. The shape is expected because most qualified hitters cluster around league-average performance, while a small number of elite hitters create a right skew.

Hits (hit): The distribution of hits is relatively normally distributed with tendencies of a slight right skew. The mean number of hits of the 129 players who qualified in 2024 is 141 hits with a standard deviation of 22.4 hits. The same logic from batting average can be applied here.

Strikeout Percentage (k_percent): The distribution of strikeout percentage is normally distributed with a mean strikeout percentage of 21% and a standard deviation of 5.3%. This reflects natural variation in hitter approach, with most players clustering near the league average and fewer players exhibiting extremely high or low strikeout rates.

Whiff Percentage (whiff_percent): The distribution of whiff percentage is normally distributed with a slight tendency of left skewness. The mean whiff percentage of 24% and a standard deviation of roughly 6%. The slight left skew occurs because elite contact hitters are relatively rare, while there is a practical upper limit on how often a hitter can miss.

Blasts Swing (blasts_swing): The distribution of blasts swing is roughly normally distributed with a mean of 11.5 swings and a standard deviation of 3.2 swings. Most players generate a similar number of high-quality swings over a season, while differences in power and swing mechanics create modest variability.

Squared Up Swings (squared_up_swing): The distribution of squared of swings is right skewed with a median of 25.4 swings and an IQR of 4.9 swings. This right skew is expected because consistently squaring up the ball is difficult, and only a small number of hitters achieve very high counts.

Hard Hit Rate (hard_hit_percent): The distribution of hard hit rate is relatively normally distributed with a mean of 42% and a standard deviation 7%. Most qualified hitters fall within a narrow range of contact quality, while differences in strength and swing efficiency produce moderate variation.

Barrel Batted Rate (barrel_batted_rate): The distribution of barrel batted rate is bimodally right skewed with a median of 8.8% and and IQR of 5.6%. This shape likely reflects a separation between contact-oriented hitters and power hitters, with a small subset of players generating barrels at a much higher rate.

Handedness (bat): There are an overwhelming number, almost twice the number of right handed batters compared to left handed batters. This imbalance reflects the general population's right-handed dominance and the difficulty of developing and maintaining effective swings from both sides of the plate.

Bivariate Analysis

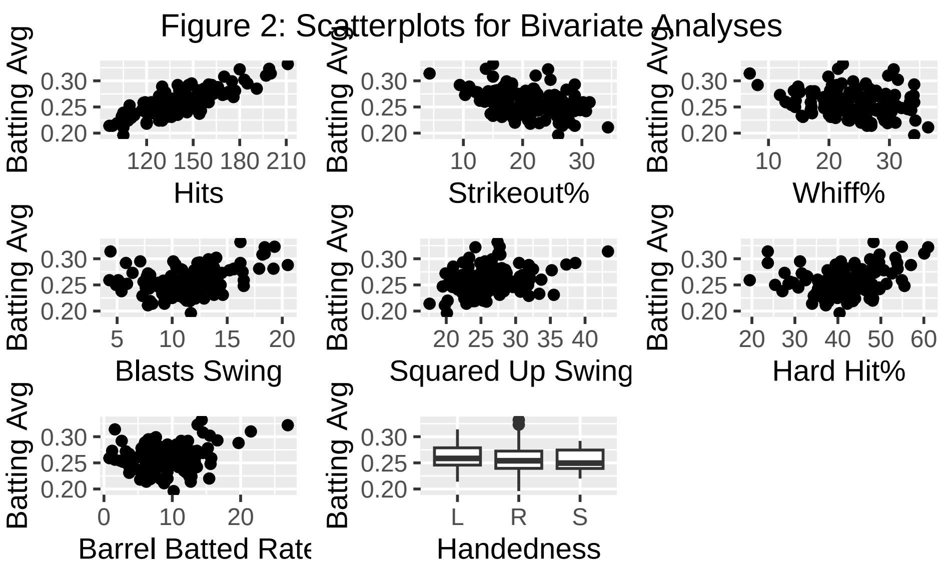


Figure 2:

For the bivariate analysis, we explored each predictor's relationship with the response variable, batting average, by looking at their scatterplots. For the one categorical variable, handedness, boxplots were used to analyze the predictor's relationship with batting average. It is important to note that when deciding what variables to include in this study, some preliminary analysis was done to pick the predictors that had the strongest correlations with batting average, so exploring another set of variables could have yielded weaker results.

Hits vs Batting Average: The relationship between hits and batting average is positive, linear, and strong. The correlation between the two is 0.8. This makes sense because getting more hits directly helps your batting average.

Strikeout Percentage vs Batting Average: Strikeout percentage and batting average has a negative, moderate, and relatively linear association, with a correlation of -0.4. This makes sense because striking out less, leads to more positive outcomes, which helps your batting average.

Whiff percentage vs Batting Average: The association between whiff percentage and batting average is negative, weak, and somewhat linear, with a correlation of -0.2. This is valid because not missing the ball can help your chances with experiencing success at the plate, but it is not directly affected.

Blasts Swing vs Batting Average: The relationship between blasts swings and batting average is positive, weak, and somewhat linear, with a correlation of 0.4. This makes sense because hitting a ball on the sweet spot and with a fast swing, can result to more success if the ball finds a gap on the field.

Squared Up Swings vs Batting Average: The relationship between squared up swings and batting average is positive, weak, and somewhat linear, with a correlation of 0.3. This is the same idea as blasts swing that hitting the ball on the barrel can increase your batting average but not directly.

Hard Hit Rate vs Batting Average: The association between hard hit rate and batting average is positive, weak, and somewhat linear, with a correlation of 0.3. Hitting the ball hard can increase your batting average, but can also lead to more outs depending on where the ball is hit.

Barrel Batted Rate vs Batting Average: The association between barrel batted rate and batting average is positive, weak, and somewhat linear, with a correlation of 0.2. This is the same idea as hard hit rate where hitting the ball can be beneficial or detrimental depending on where the ball is hit.

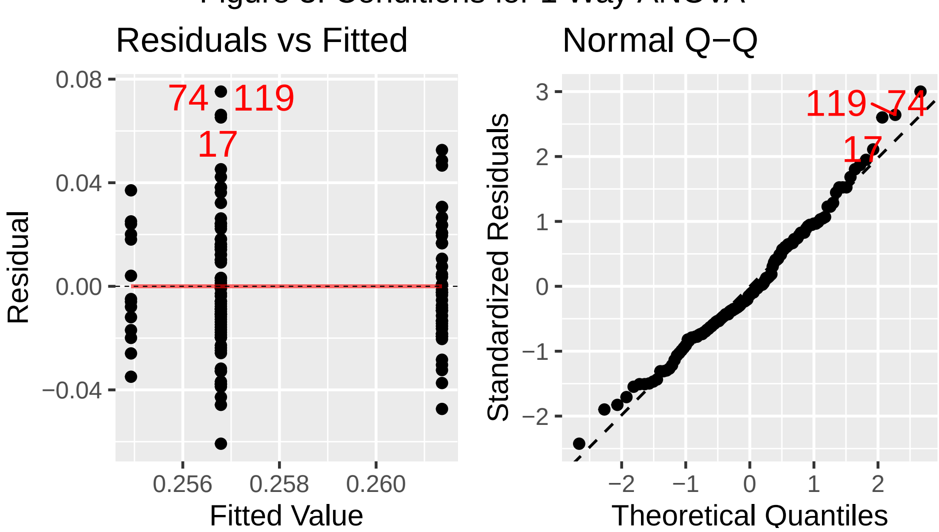
Handedness vs Batting Average: Right handed batters have much more variability than any of the other sides of the plate which makes sense because right handed batters dominate the sport, compared to left handed hitters or switch hitters. However, the median looks similar for all three handedness.

One Way ANOVA

Table 1. One-way ANOVA Results for Batting Average by Handedness

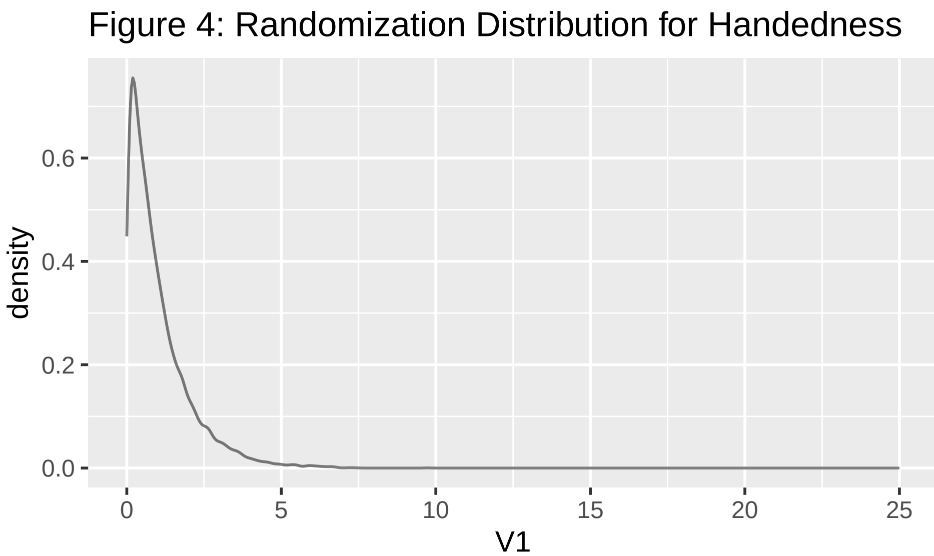
```
## Analysis of Variance Table
##
## Response: batting_avg
##            Df   Sum Sq    Mean Sq F value Pr(>F)
## bat         2 0.000731 0.00036563  0.5738 0.5649
## Residuals 126 0.080291 0.00063723
```

Figure 3: Conditions for 1 Way ANOVA




```r
origres <- anova(modanova)$"F value"
set.seed(230)
randomizedsamples <- do(10000) * (anova(lm(shuffle(batting_avg) ~ bat, data = baseball24))$"F value")

randomizedsamples <- data.frame(randomizedsamples)
gf_dens(~ V1, data = randomizedsamples) |>
  gf_lims(x = c(0, 25)) |>
  gf_labs(title = "Figure 4: Randomization Distribution for Handedness")
```




```r
pdata(~ V1, origres[1], data = randomizedsamples, lower.tail = FALSE)
```

```
## [1] 0.575
```

Using our one categorical variable, handedness, we conducted a one way ANOVA (Table 1) to predict batting average. Based on figure 3, the normality condition is met based by the QQ plot although there are some deviation at the tails, but overall the data is relatively normal. We can assume independence since player performance does not affect each other and randomness does not work for this dataset since all players who meet the minimum plate appearances are included. Figure 3 also revelas that the equal variance is not met because the variability for right handed batters is larger than switch hitters and left handed batters. This logically makes sense because there are many more right handed batters than switch handed batters in the league. Since the equal variance condition is not met, a randomization F test in ANOVA, as seen in Figure 4, was conducted to evaluate the likelihood of obtaining results such as those from this experiment.

Multiple Linear Regression

Table 2: Kitchen Sink Model

```
##                    Estimate Std. Error t value Pr(>|t|)    
## (Intercept)       1.451e-01  2.393e-02   6.061 1.58e-08 ***
## hit               7.715e-04  6.994e-05  11.031  < 2e-16 ***
## k_percent        -6.721e-04  5.846e-04  -1.150  0.25256    
## whiff_percent    -1.190e-04  5.021e-04  -0.237  0.81312    
## blasts_swing      1.265e-03  4.477e-04   2.826  0.00551 ** 
## squared_up_swing  3.094e-04  5.315e-04   0.582  0.56150    
## batR             -1.150e-03  2.794e-03  -0.412  0.68121    
## batS             -7.734e-03  4.406e-03  -1.756  0.08170 .  
## 
## Residual standard error: 0.0142 on 121 degrees of freedom
## Multiple R-squared:  0.699,  Adjusted R-squared:  0.6816 
## F-statistic: 40.14 on 7 and 121 DF,  p-value: < 2.2e-16
```

```
##                      GVIF Df GVIF^(1/(2*Df))
## hit              1.559725  1        1.248889
## k_percent        6.099547  1        2.469726
```

```
## whiff_percent    5.691974  1        2.385786
## blasts_swing     1.311900  1        1.145382
## squared_up_swing 3.121555  1        1.766792
## bat              1.066145  2        1.016141
```

We first tried a kitchen sink model (See Table 2) to determine what predictors were significant. F(7, 121) = 40.14. Since p is very small (2.2*10^-16 < 0.05), we can state that the overall model is significant and at least one of the predictors is significant. Looking at the individual predictors, we determined that hits and blasts swing were the only significant predictors. The R^2 is 0.68, meaning that 68% of the variability in batting average is accounted for in this model. Additionally, the residual standard error is 0.0142 which is small and shows a pretty strong model. Looking at VIFs below the table, it can be seen that strikeout percentage and whiff percentage raised some concerns of multicollinearity. This model is overall not bad, but it is important to look for a simpler model that accounts for the same or more variation in batting average.

Stepwise Variable Selection Method Results

```
##                 rsq     adjr2        cp        rss hit k_percent whiff_percent
## 1  ( 1 ) 0.6446870 0.6418893 17.837306 0.02878824   *                        
## 2  ( 1 ) 0.6685391 0.6632778 10.248667 0.02685569   *         *              
## 3  ( 1 ) 0.6888790 0.6814121  4.071953 0.02520771   *         *              
## 4  ( 1 ) 0.6970995 0.6873285  2.767288 0.02454167   *         *              
## 5  ( 1 ) 0.6984079 0.6861480  4.241297 0.02443566   *         *              
## 6  ( 1 ) 0.6988685 0.6840587  6.056131 0.02439834   *         *              
## 7  ( 1 ) 0.6990081 0.6815954  8.000000 0.02438702   *         *             *
##          blasts_swing squared_up_swing batR batS
## 1  ( 1 )                                        
## 2  ( 1 )                                        
## 3  ( 1 )            *                           
## 4  ( 1 )            *                          *
## 5  ( 1 )            *                *         *
## 6  ( 1 )            *                *    *    *
## 7  ( 1 )            *                *    *    *
```

Table 3: Overall Best Model for Predicting Batting Average

```
##                 Estimate Std. Error t value Pr(>|t|)    
## (Intercept)   1.580e-01  1.136e-02  13.903  < 2e-16 ***
## hit           7.566e-04  6.633e-05  11.407  < 2e-16 ***
## k_percent    -1.009e-03  2.635e-04  -3.830 0.000201 ***
## blasts_swing  1.266e-03  4.428e-04   2.859 0.004986 ** 
## 
## Residual standard error: 0.0142 on 125 degrees of freedom
## Multiple R-squared:  0.6889, Adjusted R-squared:  0.6814 
## F-statistic: 92.26 on 3 and 125 DF,  p-value: < 2.2e-16

##          hit    k_percent blasts_swing 
##     1.401774     1.238589     1.282185
```

Table 4: Interaction Term Model Testing

```
##                 Estimate Std. Error t value Pr(>|t|)    
## (Intercept)   1.580e-01  1.136e-02  13.903  < 2e-16 ***
## hit           7.566e-04  6.633e-05  11.407  < 2e-16 ***
```

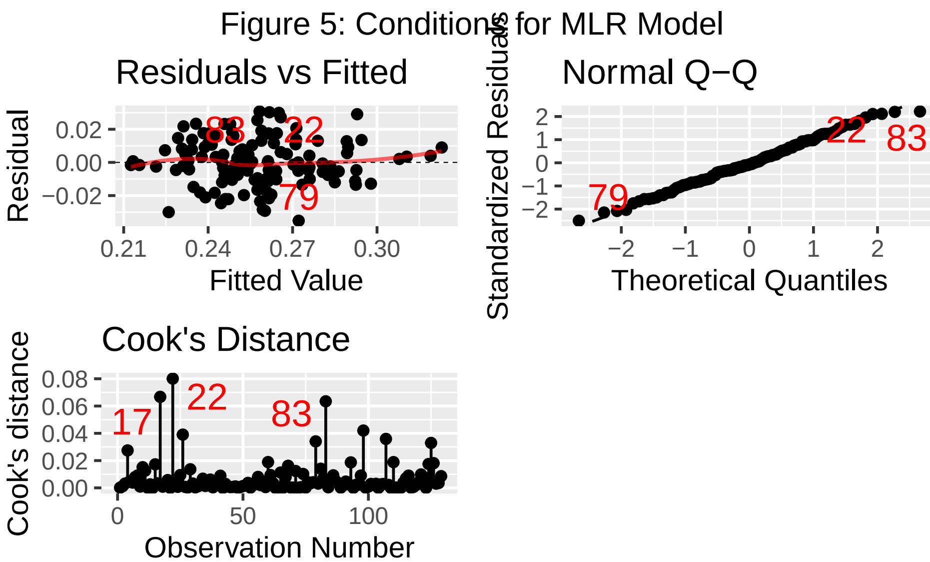
```
## k_percent    -1.009e-03  2.635e-04  -3.830 0.000201 ***
## blasts_swing  1.266e-03  4.428e-04   2.859 0.004986 **
##
## Residual standard error: 0.0142 on 125 degrees of freedom
## Multiple R-squared:  0.6889, Adjusted R-squared:  0.6814
## F-statistic: 92.26 on 3 and 125 DF,  p-value: < 2.2e-16
```

Next, we tried a variable selection method (stepwise, backward, and forward) to determine which model was the strongest based on the best balance of lowest Cp Mallow Values, highest adjusted R^2, and parsimony. In all three variable selection methods, the best model included hits, strikeout percentage, blasts swing, and switch hitters, as revealed in Table 4. It had the lowest Cp value at 2.77 and an adjusted r^2 of 0.687, meaning that 68.7% of the variability was explained by this model. This is slightly higher than the kitchen sink model, but much simpler with a lower Cp value. However, we determined in the one way ANOVA that handedness did not have a significant effect on predicting batting average (Figure 4), so it is best if we remove switch hitters from the model for significance and parsimony. Our best model now includes hits, strikeout percentage, and blasts swing (Table 3). This model is overall significant F(3, 125) = 92.26, p = 2.2*10^-16, has all significant predictors, a low residual standard error at 0.01, and an adjusted r^2 of 0.6814, meaning that 68% of the variability in batting average was explained by this model. Checking the VIFs below Table 3, we see that all the predictors have a VIF below 5 so there is no multicollinearity. We also tried an interaction term between strikeout percentage and blasts swing in Table 4, but still resulted to the model with hits, strikeout percentage, and blasts swing.

Table 5: Nested Testing with Our Previous Best Model and Manual Models

```
## Analysis of Variance Table
##
## Model 1: batting_avg ~ hit
## Model 2: batting_avg ~ hit + blasts_swing
## Model 3: batting_avg ~ hit + k_percent + blasts_swing
##   Res.Df      RSS Df  Sum of Sq       F    Pr(>F)
## 1    127 0.028788
## 2    126 0.028166  1 0.00062182  3.0835 0.0815413 .
## 3    125 0.025208  1 0.00295871 14.6717 0.0002014 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Finally, we did some manual model building based on the results of the kitchen sink model, the variable selection methods, and intuition to compare which model would be the best. Model 1 above is a simple linear regression of just hits since it had the strongest correlation with batting average. Model 2 was the model of the significant predictors from the kitchen sink model. Both of these models were compared with the best model from the variable selection methods in Table 5 to determine which model would be the strongest but balanced with simplicity.



We also checked the conditions to confirm that we can use this model, as seen in Figure 5. The equal variance condition was met based on the homoskedasticity in the residual vs fitted plot. The normality condition is met based on the QQ plot. Although there is minimal deviation at the tails, overall, the plot looks relatively normal. Independence is met and randomization is not applicable as this is not a sample. Additionally, there are no unusual points based on the Cook's Distance as none of the distances are greater than 0.5.

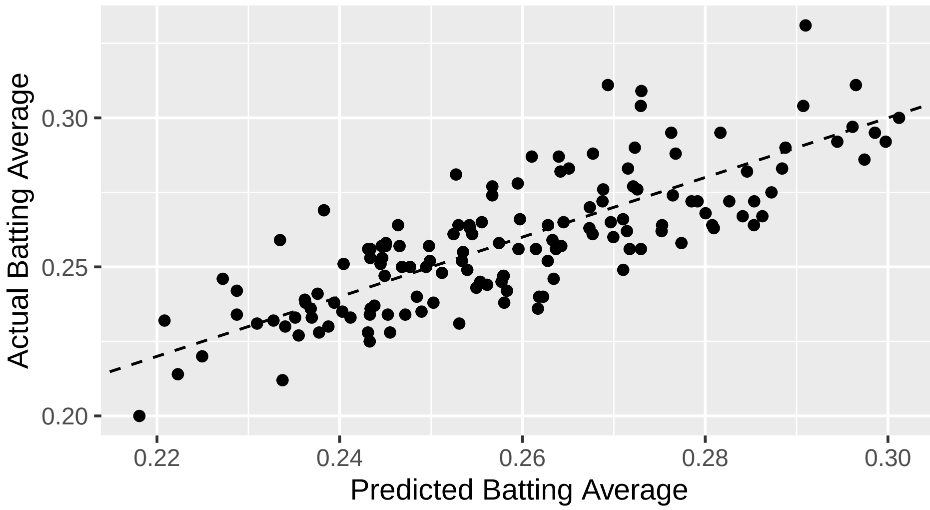
```r
#prediction vs actual batting average plot
baseball25$pred_ba <- predict(modbest, newdata = baseball25)

mse <- mean((baseball25$batting_avg - baseball25$pred_ba)^2)
mse
```

```
## [1] 0.0001920227
```

```r
ggplot(baseball25, aes(pred_ba, batting_avg)) +
  geom_point() +
  geom_abline(slope = 1, intercept = 0, linetype = "dashed") +
  labs(
    title = "Figure 6: Predicted vs Actual Batting Average (2025)",
    x = "Predicted Batting Average",
    y = "Actual Batting Average"
  )
```

Figure 6: Predicted vs Actual Batting Average (2025)



To evaluate our model's performance, we created a prediction vs actual plot for the 2025 batting average data. We first, took our model to predict their batting averages for 2025 and created a new variable, pred_ba. Then we obtained the 2025 data for batting average and created this plot to see how similar our predictions were to the empirical data, as seen in Figure 6. The dashed line on the model represents perfect prediction, with points closer to the line indicating greater predictive accuracy. Overall, the plot shows a positive relationship between predicted and actual batting averages, suggesting that the model captures the general trend in performance. The mean squared was also calculated to determine the prediction error. For this model, the mean squared errors was 0.0019.

Interpretations

One Way ANOVA

For the one way ANOVA, the null hypothesis would be that there is no difference in predicted batting average by handedness and the alternate hypothesis would be that there is a difference in predicting batting average based on handedness. Since the equal variance condition was not met, we obtain the p value from the randomization F test. The p value was 0.575, meaning that there is a 57.5% chance of observing an F statistic like this one (0.5738) or more extreme. This is a very high probability, so we fail to reject the null, meaning that there is no statistically significant evidence that batting average differs by handedness. This supports the conclusion we would have seen if we proceeded with the ANOVA.

Multiple Linear Regression

We first tried a kitchen sink model to get a feel for which predictors were significant and as a starting point. We saw that only two of the predictors were significant and although the overall model was significant, there were also some issues of multicollinearity that showed that a simpler model would be stronger.

Using all three variable selection methods, the best model we found used hits, strikeout percentage, blasts swing, and handedness. However, after the results of the ANOVA showing its non significance, we decided to not include it in the final model.

After some manual testing through this ANOVA, we determined that the model including hits, strikeout percentage, and blasts swing (the variable selection model) was still the strongest and adding strikeout percentage added valuable "juice" to the model. We did not find a signficiant interaction term when manually building this model, which makes sense because the effect of one predictor does not logically, significantly change depending on another predictor since they are not really related. Many of these predictors are similar measuring metrics. For example, strikeout percentage and whiff percentage, or hard hit rate and blasts swing both measure similar aspects of hitting. Creating an interaction with these variables would result in multicollinearity and overall complication of the model for not enough valuable "juice".

predicted batting average = 0.158 + 0.0008(hits) - 0.001(strikeout%) + 0.001(blasts swing).

intercept: When a player has no hits, zero percent strikeout percentage and no blasts swing, their predicted batting average would be 0.158. This is a highly unlikely scenario and is not applicable to interpret.

hits: For every additional hit, the predicted batting average for a player increases by 0.008 points.

strikeout percentage: For every additional one percent increase in strikeout percentage, there is a 0.001 point decrease in a player's batting average.

blasts swing: For every additional barrel batted ball with a fast swing, there is a 0.001 point increase in a player's batting average.

Conclusion

Our goal through this project was to determine a model using the 2024 batting average data to predict 2025 batting averages. In this project we have determined that hits, strikeout percentage, and blasts swing were the best predictors in predicting batting average. This all logically makes sense that getting hits directly increases batting average and not striking out increases your chance of getting a hit, increasing your batting average. Blasts swings intuitively increases batting average because squaring up a ball with a fast swing can increase your chances at getting a hit, increasing your batting average. To answer our research question, we used this model and tested it against the 2025 batting average data. When the 2024 regression model was applied to the 2025 batting average data, the predicted vs actual plot for 2025 batting averages showed a positive, linear, strong relationship. The mean squared error was 0.0019, meaning that the root mean squared error was roughly 0.01. On average, the model's predicted 2025 batting averages are off by around 0.014 points. This suggests that the model trained on 2024 data generalizes relatively well to the 2025 season. However, the noise around the line indicates player to player variation, highlighting the idea that it is impossible to completely predict an individual's batting performance. Overall, we were able to create a relatively strong model using 2024 batting average data to predict 2025 batting averages.

This study has several limitations that should be considered when interpreting the results. First, the model relies on a limited set of offensive metrics from one prior season (2024), and batting performance is known to fluctuate year to year due to many factors. For example, Major League Baseball can sometime adjust the composition of the baseball based on the season's result to balance pitchers and hitters' skills. Since, only one year of data was accounted for in building this model, year to year variation may be more prevalent. Incorporating multiple seasons of data in future analyses would allow the model to account for longer-term trends and reduce the influence of season-specific anomalies, potentially improving predictive stability. It is very important to note that even with a strong model, it is not possible to fully capture an individual player's outcomes.

Another limitation is that the several of the explanatory variables used in the analysis are closely related measures in terms of contact and swing quality, which may introduce multicollinearity. The selection of predictors involved a preliminary process to only include the variables that had the best correlation or linearity with batting average. This narrowed our predictors to be overlapping and very standard measures for batting average. In the future, it would be interesting to see whether any of the special, baseball-savant statistics such as, sprint speed, would have any sort of unexpected impact of predicting batting average. Future models could incorporate a broader set of less correlated variables or use dimension-reduction techniques to better isolate unique sources of variation in batting average.

Finally, this study did not account for any contextual factors such as defensive shifts, ballpark environments, or strength of schedule, all of which can influence batting average. Defensive shifts can take away from batting average even if it is a hard hit ball off the barrel. Ballpark environments such as the Colorado Rockies Field, Coors Field, can improve batting average as the ball flies further in higher altitudes. The strength of schedule and specifically, the pitchers faced can have a direct impact on a player's batting averages. Incorporating contextual variables into future models would allow for more nuanced predictions by separating player skill from environmental and competitive influences.

References

- *Player Batting 2024 Custom Leaderboard* . baseballsavant.com. (n.d.). https://baseballsavant.mlb.com/leaderboard/custom?year=2024&type=batter&filter=&min=q&selections=pa%2Ck_percent%2Cbb_percent%2Cwoba%2Cxwoba%2Csweet_spot_percent%2Cbarrel_batted_rate%2Chard_hit_percent%2Cavg_best_speed%2Cavg_hyper_speed%2Cwhiff_percent%2Cswing_percent&chart=false&x=pa&y=pa&r=no&chartType=beeswarm&sort=xwoba&sortDir=desc

- *Player Batting 2025 Custom Leaderboard* . baseballsavant.com. (n.d.). https://baseballsavant.mlb.com/leaderboard/custom?year=2025&type=batter&filter=&min=q&selections=pa%2Ck_percent%2Cbb_percent%2Cwoba%2Cxwoba%2Csweet_spot_percent%2Cbarrel_batted_rate%2Chard_hit_percent%2Cavg_best_speed%2Cavg_hyper_speed%2Cwhiff_percent%2Cswing_percent&chart=false&x=pa&y=pa&r=no&chartType=beeswarm&sort=xwoba&sortDir=desc