2025 MLB Player-Pitch Rankings

By Cole Geller

Objective

Identify the most effective pitch during the 2025 MLB season using a publicly available performance-metric dataset with a player-pitch unit-of-observation.

Methods

Using the **Baseball Savant Pitch Arsenal Stats Leaderboard (2025)**, I analyzed all 508 player-pitches thrown at least 405 times (Baseball Savant's qualifier). To evaluate overall effectiveness, each pitch was ranked across these key performance metrics:

- Run Value ***(RV)*** and Run Value per 100 pitches ***(RV/100)***
- Batting Average ***(BA)*** and Expected Batting Average ***(xBA)***
- Slugging Percentage ***(SLG)*** and Expected Slugging Percentage ***(xSLG)***
- Weighted On-Base Average ***(wOBA)*** and Expected Weighted On-Base Average ***(xwOBA)***
- Whiff Percentage ***(Whiff %)*** and Strikeout Percentage ***(K%)***
- Put-Away Percentage ***(Put Away %)*** and Hard-Hit Percentage ***(Hard Hit %)***

Data Wrangling

```
# Loading dataset
player_pitch <- read.csv('player_pitch.csv')

# Removing unnecessary columns
player_pitch_ranked <- subset(player_pitch, select = -c(team_name_alt, player_id, pitch_name, pitches, p

# Unranked DataFrame
head(player_pitch_ranked[,c("last_name..first_name", "pitch_type", "run_value", "slg", "woba", "k_percen
```

```
##    last_name..first_name pitch_type run_value   slg  woba k_percent put_away
## 1         Gausman, Kevin         FF        20 0.380 0.306      19.0     17.7
## 2       Parker, Mitchell         FF        -6 0.447 0.358       8.9     10.1
## 3            Ray, Robbie         FF         9 0.357 0.300      26.0     20.6
## 4          Soriano, José         SI         2 0.380 0.357       7.8     13.3
## 5           Nelson, Ryne         FF        23 0.360 0.281      22.7     17.4
## 6        Peralta, Freddy         FF         9 0.381 0.314      23.5     17.6
## 7     Senzatela, Antonio         FF       -22 0.542 0.437      10.4     11.9
## 8        Valdez, Framber         SI        15 0.378 0.330       8.5     19.8
## 9    Sánchez, Cristopher         SI        19 0.342 0.313      14.9     23.1
## 10         Pivetta, Nick         FF        21 0.326 0.263      30.3     25.8
```

ˆ A subset of the ***player_pitch*** DataFrame with ten player-pitch observations, along with raw, unranked statistical values for several metrics.

Rankings

Depending on the metric, a higher value can imply better or worse performance. For example, a higher RV indicates better performance, whereas a higher Hard Hit % indicates worse performance. This factor was considered when generating rankings.

```
library(dplyr)

# Metrics where higher is better
higher_is_better_metrics <- c("run_value_per_100", "run_value", "whiff_percent", "k_percent", "put_away

# Metrics where lower is better
lower_is_better_metrics <- c("ba", "slg", "woba", "est_ba", "est_slg", "est_woba", "hard_hit_percent")

# All metrics
performance_metrics <- c(lower_is_better_metrics, higher_is_better_metrics)


# Generating rankings
player_pitch_ranked <- player_pitch_ranked %>%

  mutate(
    across(all_of(higher_is_better_metrics), ~ rank(-.x, ties.method = "min")),
    across(all_of(lower_is_better_metrics), ~ rank(.x, ties.method = "min"))
  )

#Ranked DataFrame
head(player_pitch_ranked[,c("last_name..first_name", "pitch_type", "run_value", "slg", "woba", "k_percen
```

```
##    last_name..first_name pitch_type run_value slg woba k_percent put_away
## 1         Gausman, Kevin         FF         7 228  216       314      315
## 2       Parker, Mitchell         FF       438 365  384       487      496
## 3            Ray, Robbie         FF        66 186  194       169      191
## 4          Soriano, José         SI       242 228  379       494      450
## 5          Nelson, Ryne          FF         2 192  142       239      328
## 6        Peralta, Freddy         FF        66 234  236       224      317
## 7     Senzatela, Antonio         FF       507 471  500       473      471
## 8        Valdez, Framber         SI        22 225  285       489      225
## 9    Sánchez, Cristopher         SI        11 154  232       396      102
## 10         Pivetta, Nick         FF         4 124   85       111       55
```

ˆ A subset of the ***player_pitch_ranked*** DataFrame with ten player-pitch observations, along with rankings for several metrics.

Overall Performance Score

Rankings across all twelve metrics were summed to produce an overall performance score, where lower values represent stronger effectiveness. A pitch that ranked first in every metric would receive a score of 12 x 1(st) = 12, while a pitch that ranked last in every metric would receive 12 x 508(th) = 6096.

```r
# Generating overall performance scores
player_pitch_ranked$overall_performance_score <- rowSums(player_pitch_ranked[, performance_metrics])
player_pitch_ranked %>%
  select(last_name..first_name, pitch_type, overall_performance_score) %>%
  head(10)
```

```
##    last_name..first_name pitch_type overall_performance_score
## 1          Gausman, Kevin         FF                      3175
## 2        Parker, Mitchell         FF                      5301
## 3             Ray, Robbie         FF                      2245
## 4           Soriano, José         SI                      4495
## 5            Nelson, Ryne         FF                      2532
## 6         Peralta, Freddy         FF                      2901
## 7      Senzatela, Antonio         FF                      5719
## 8         Valdez, Framber         SI                      3866
## 9     Sánchez, Cristopher         SI                      3060
## 10          Pivetta, Nick         FF                      1952
```

ˆ A subset of the ***player_pitch_ranked*** DataFrame with ten player-pitch observations, along with their overall performance scores.

Results

Section 1 - Overall Patterns

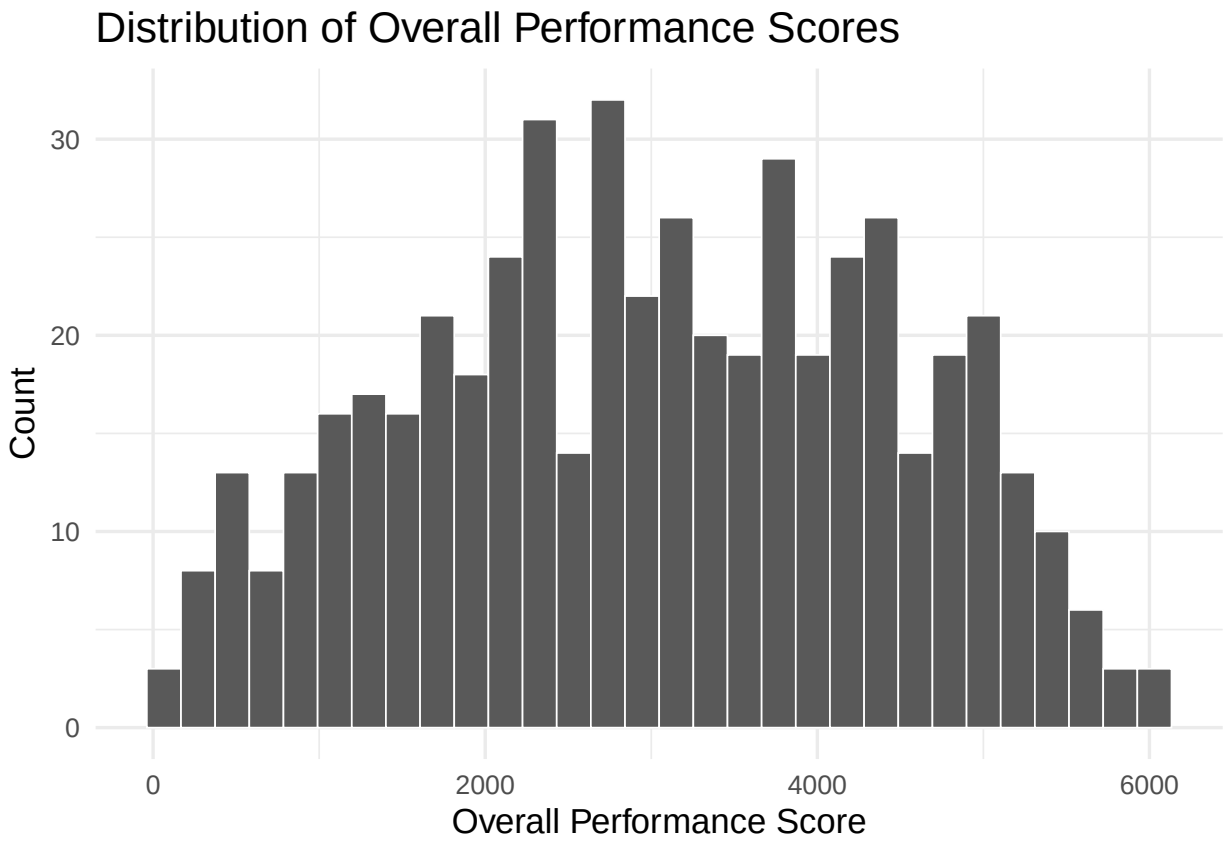
The distribution of overall performance scores appears approximately uniform to slightly mound-shaped, with scores spread relatively evenly across the full range. There is no strong skew, and no single cluster dominates the distribution, which suggests that pitch effectiveness varies widely but without a strong concentration of highly exceptional or extremely poor performers.

There are more observations near the center of the distribution (roughly 2,000–4,000), consistent with many pitches being "middle of the pack," while far fewer pitches fall at the extreme low (elite) or extreme high (weak) ends of the spectrum. This aligns with expectations in a large dataset of MLB pitches: only a small number of pitches perform exceptionally across all metrics, and only a small number perform poorly across all metrics.

Overall, the histogram reflects a broad and evenly dispersed performance landscape, with the best pitches being clear outliers relative to the majority.

```r
library(ggplot2)

ggplot(player_pitch_ranked, aes(x = overall_performance_score)) +
  geom_histogram(bins = 30, color = 'white', linewidth = 0.3) +
  labs(
    title = "Distribution of Overall Performance Scores",
    x = "Overall Performance Score",
    y = "Count"
  ) +
  theme_minimal(base_size = 13)
```




```
# Top 10 Overall Performance Scores
sorted <- player_pitch_ranked %>%
  select(last_name..first_name, pitch_type, overall_performance_score) %>%
  arrange(overall_performance_score)

head(sorted, 10)
```

```
##    last_name..first_name pitch_type overall_performance_score
## 1          Miller, Mason         SL                         65
## 2          Muñoz, Andrés         SL                        124
## 3       Crochet, Garrett         ST                        137
## 4            Sale, Chris         SL                        193
## 5          Skubal, Tarik         CH                        199
## 6           Uribe, Abner         SL                        270
## 7          Helsley, Ryan         SL                        299
## 8    Sánchez, Cristopher         CH                        306
## 9          Okert, Steven         SL                        316
## 10        Luzardo, Jesús         ST                        326
```

ˆ Best overall performance scores

```
# Bottom 10 Overall Performance Scores
tail(sorted, 10)
```

```
##     last_name..first_name pitch_type overall_performance_score
```

```
## 499         Helsley, Ryan         FF                    5610
## 500       Buehler, Walker         FF                    5640
## 501       Márquez, Germán         SI                    5667
## 502    Senzatela, Antonio         FF                    5719
## 503        Freeland, Kyle         FF                    5728
## 504       Cabrera, Edward         SI                    5750
## 505        Rogers, Taylor         SI                    5751
## 506       Márquez, Germán         FF                    5962
## 507       Chivilli, Angel         FF                    5984
## 508      Blalock, Bradley         FF                    6030
```

ˆ Worst overall performance scores

Section 2 - Top Performer: Mason Miller's Slider

Mason Miller's slider ranked as the top-performing pitch among qualifiers by generating the minimum overall-performance score.

```
# Best Player-Pitch (Minimum Overall Performance Score)
best_player_pitch <- player_pitch_ranked %>%
  filter(overall_performance_score == min(overall_performance_score))

cat(
  paste(
    "Best Pitch:", best_player_pitch$last_name..first_name, "-", best_player_pitch$pitch_type)
)
```

```
## Best Pitch: Miller, Mason - SL
```

Mason Miller's Slider Rankings:

- Run Value: (17)
- RV/100: (1)
- Batting Average: (1)
- Slugging Percentage: (1)
- Weighted On-Base Average: (1)
- Strikeout Percentage: (3)
- Put-Away Percentage: (1)
- Expected Batting Average: (10)
- Expected Slugging Percentage: (4)
- Expected Weighted On-Base Average: (3)
- Hard-Hit Percentage: (20)
- Whiff Percentage: (3)

Conclusion

Mason Miller's slider stands out as the most effective pitch of the 2025 MLB season, ranking at or near the top across both contact-quality and swing-and-miss metrics. Its combination of velocity, late vertical depth, and sharp glove-side movement allows it to suppress damage while generating elite strikeout rates. The consistency of its performance across all twelve evaluated metrics underscores its dominance relative to other high-usage pitches league-wide. Future work could incorporate pitch-tracking characteristics (e.g. vertical drop, horizontal break) to better understand why Miller's slider performs at such an exceptional level and how its underlying traits compare to other top-tier MLB pitches.