PITCHER DROPOFF ANALYSIS

In the realm of baseball analytics, understanding the differences between starting pitchers and relievers can help strategize and optimize team performance. In this investigation, we introduce the concept of "dropoff" values– a metric we propose as a significant indicator that distinguishes the success patterns between starters and relievers. Our research is anchored in the hypothesis that starting pitchers exhibit a more consistent performance throughout the game, while relievers, despite a potentially strong first at-bat, face a sharper decline in effectiveness after a certain number of at-bats.

In the initial phase of our analysis, we conducted a comprehensive evaluation of the datasets provided to us, identifying the inherent strengths and weaknesses within each. The FanGraphs table provides seasonal data, which is useful for looking at holistic statistics and checking the significance of stats, but is less useful for comparing individual players and does not provide very detailed observations of a pitcher's performance. Alternatively, the BaseballSavant table is much more extensive, making it very useful for generating models because we can analyze every single detail of a pitcher's performance, including location data, spin rate, and speed. For the majority of the process, we operated our research on the BaseballSavant pitch by pitch data, and then we incorporated the FanGraphs seasonal data towards the end.

Due to the volume of the data we reduced the number of parameters, keeping [`“player_name”, “game_date”, “pitcher_at_bat_number”, “pitch_type”`] to identify pitches, [`“role_key”`] for identifying starters and relievers, and [`“effective_speed”, “release_spin_rate”, “pfx_x”, “pfx_z”`] to make up the proposed "dropoff" stat that we will utilize.

We grouped the dataset by its own unique identifier: by player, date, bat number, and pitch type. Because of the variance of the volume of observations in the resulting groups, we averaged each statistic within the group so that there was only one observation representing each unique combination of these indicators, reducing the row count to `562445`.

Next, we aimed to create the table that presented how a pitcher's performance (for each type of pitch) declined from inning to inning. To simplify the drop-off, we picked one singular stat for each type of pitch to measure the dropoff. To select that stat, we analyzed each of the four stats (speed, spin, x, z) on each pitch type to determine which stat on that pitch deviated the most from an average pitch of any type. If a stat deviated more aggressively than the other stats for said pitch type, we chose it as the important stat for that pitch, and thus used it to calculate the drop-offs. Here is the code of the results:

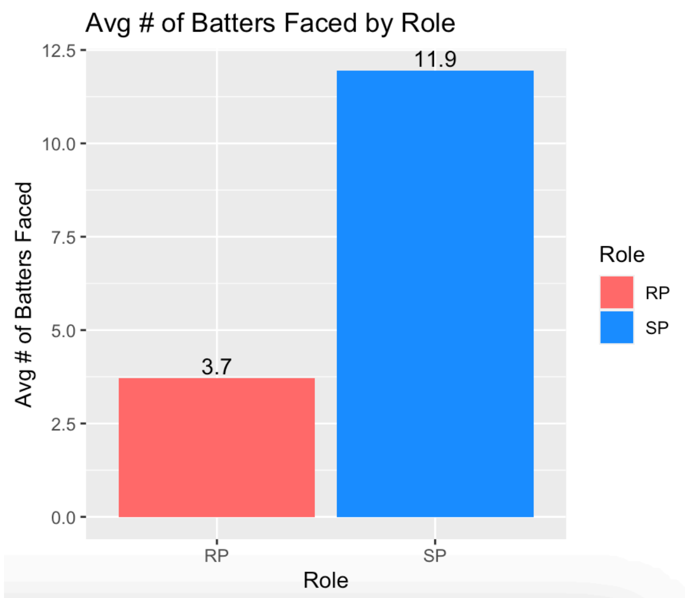
```
if row['pitch_type'] == 'FC' or row['pitch_type'] == 'SV' or
row['pitch_type'] == 'ST':
        return row['release_spin_rate']
    elif row['pitch_type'] == 'CU' or row['pitch_type'] == 'KC' or
row['pitch_type'] == 'SC':
        return row['pfx_z']
    elif row['pitch_type'] == 'SL' or row['pitch_type'] == 'CS' or
row['pitch_type'] == 'FS':
        return row['pfx_x']
    elif row['pitch_type'] == 'FF' or row['pitch_type'] == 'SI' or
row['pitch_type'] == 'CH':
        return row['effective_speed']
```

The average value of the selected statistic across all initial at-bats was designated as the player's baseline for that particular pitch. The "[“`dropoff`”]" column was calculated by subtracting this baseline value from the value of the pitch in the current at-bat, grouped by the variables [“`player_name`”, “`game_date`”, “`pitch_type`”]. This calculation thus reflects the decline in a pitcher's effectiveness over time. For the first at-bat, the "dropoff" is automatically set to 0, indicating no change from the baseline. While our model accommodates up to 35 at-bats to ensure comprehensive coverage, our primary focus was on the initial at-bats, given their richer data availability and relevance to assessing performance trends.

Next, we had to separate the results to differentiate between starters and relievers. We engineered the table “DropoffRole”, which analyzed the average dropoff for each pitch type for starters and relievers. While the averages for starters and relievers ended up being relatively similar for all the pitch_types, it's crucial to contextualize these findings within our initial hypothesis, which assumes that relievers tend to be substituted out before any significant dropoff in performance occurs– a strategy that might explain why they are not typically utilized as starters. This idea was further explained by the number of at bats thrown decreasing drastically for relievers past 10 at bats.

By analyzing the differences in the first couple dropoffs, we can bypass this problem. We decided to only look at the first four at bats because this is the average number of batters relief pitchers faced in the dataset (Figure 1). As such, we took the average of the second, third, and fourth at bat for each player and pitch type, and subtracted the corresponding average drop-off of the opposite role key. This quantifies the numerical difference between a single player’s dropoff versus the population dropoff of the opposite role key.

FIGURE 1



For the final computation, we introduced a new column ["dist"] to compute the Euclidean distance between the drop-off values of each player and the mean drop-off values of the opposite role key, across the three drop offs. A smaller distance suggests a player's performance closely aligns with the typical performance patterns of the opposite role. Conversely, a larger distance implies a significant deviation from those patterns, highlighting a clear distinction between the player's performance and that of the alternate role.

Following this calculation, we segmented the data according to role key and pitch type, ranking each player based on their distance scores across all pitch types. Players receiving the lowest scores are identified as potentially being mismatched in their current roles. We then went through all the top players to make sure they were eligible to be chosen: some top-ranking players had a low volume of data, are not listed as pitchers, or are not currently active players.

We then pivoted to the FanGraphs dataset, where we ran three different regression models to determine the most indicative stat for starter vs reliever success. For each regression model, we used a different dependent variable (ER for the first one, ERA for the second one, and WAR for the third one), and set the independent variables to all predictive seasonal statistics in the dataset (found through research and general baseball knowledge). We determined the ERA model to be most accurate with a 0.8155 adjusted R-squared value (0.9164 for starters, 0.7686 for relievers). The summary of the ERA regression model for both starters (data = selectedFanGraphsTestSP) and relievers (data = selectedFanGraphsTestRP) is shown below.

FIGURE 2 (Starting Pitchers)

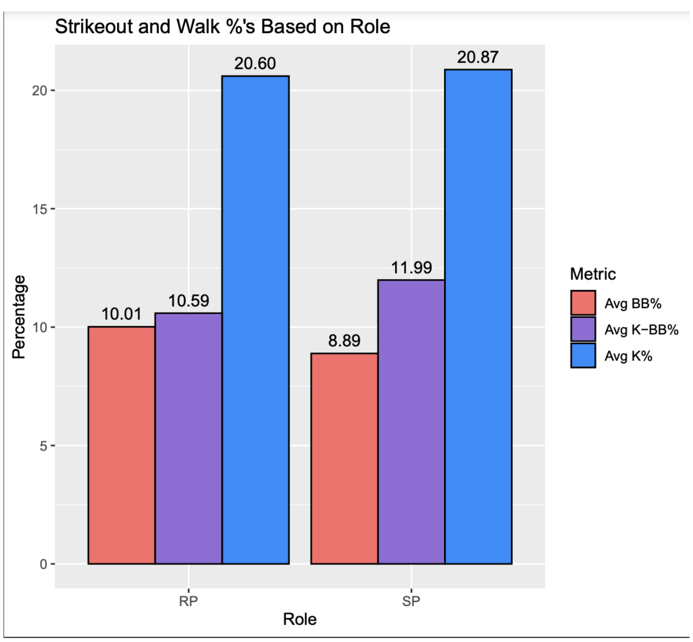
```
Call:
lm(formula = ERA ~ . - ER - WAR - Name, data = selectedFanGraphsTestSP)

Residuals:
     Min       1Q   Median       3Q      Max 
-25.5419  -0.3871   0.0480   0.5170  16.9675 

Coefficients:
               Estimate Std. Error t value             Pr(>|t|)    
(Intercept)    9.129068   1.551645   5.883     0.000000005287 ***
K_pct_plus    -0.013077   0.006551  -1.996              0.04616 *  
BB_pct_plus   -0.032769   0.002135 -15.351 < 0.0000000000000002 ***
SIERA          0.062356   0.139223   0.448              0.65432    
WHIP_plus      0.108171   0.001738  62.222 < 0.0000000000000002 ***
BABIP_plus    -0.057690   0.002535 -22.753 < 0.0000000000000002 ***
K_per_9_plus   0.041461   0.005885   7.045     0.00000000000322 ***
CSW_pct       -3.395320   1.902972  -1.784              0.07466 .  
Stuff_plus    -0.008611   0.004975  -1.731              0.08371 .  
Location_plus -0.015855   0.011445  -1.385              0.16624    
LOB_pct       -8.470295   0.411332 -20.592 < 0.0000000000000002 ***
GB_to_FB      -0.238142   0.083148  -2.864              0.00426 ** 
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.979 on 1130 degrees of freedom
  (4 observations deleted due to missingness)
Multiple R-squared:  0.9172,	Adjusted R-squared:  0.9164 
F-statistic:  1139 on 11 and 1130 DF,  p-value: < 0.00000000000000022
```

FIGURE 3 (Relief Pitchers)

```
Call:
lm(formula = ERA ~ . - ER - WAR - Name, data = selectedFanGraphsTestRP)

Residuals:
    Min      1Q  Median      3Q     Max 
-40.294  -0.667  -0.013   0.740  45.587 

Coefficients:
                Estimate Std. Error t value             Pr(>|t|)    
(Intercept)    13.230478   1.102335  12.002 < 0.0000000000000002 ***
K_pct_plus     -0.006317   0.007096  -0.890              0.37347    
BB_pct_plus    -0.023256   0.001744 -13.339 < 0.0000000000000002 ***
SIERA           0.101788   0.094043   1.082              0.27922    
WHIP_plus       0.077609   0.001667  46.567 < 0.0000000000000002 ***
BABIP_plus     -0.039444   0.003045 -12.954 < 0.0000000000000002 ***
K_per_9_plus    0.019606   0.006739   2.909              0.00366 ** 
CSW_pct         4.021259   1.777640   2.262              0.02379 *  
Stuff_plus     -0.001678   0.001720  -0.976              0.32930    
Location_plus  -0.037962   0.008460  -4.487            0.0000076 ***
LOB_pct       -12.338457   0.407722 -30.262 < 0.0000000000000002 ***
GB_to_FB       -0.305800   0.072735  -4.204            0.0000273 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.035 on 2097 degrees of freedom
  (70 observations deleted due to missingness)
Multiple R-squared:  0.7698,	Adjusted R-squared:  0.7686 
F-statistic: 637.5 on 11 and 2097 DF,  p-value: < 0.00000000000000022
```

From there, we determined that BB%+ was more indicative of starter success while a mix of BABIP+, CSW%, and GB/FB% was more indicative of reliever success. We decided to use BB%+ compared to similar metrics such as K-BB% and K% based on our findings from the visual below.

FIGURE 4



We then closely examined the select group of pitchers (who according to their Euclidean distance scores were highly ranked as potentially being better suited for the opposite pitching role) to see if their FanGraphs stats fit the dropoff stats of the role we wanted them to switch to. Through this comparison, we narrowed our focus to three pitchers who emerged as prime candidates for a role change based on the data from both BaseballSavant and FanGraphs: Madison Bumgarner and Marcus Stroman were considered for a transition from starters to relievers, while Caleb Thielbar was identified as a candidate moving from reliever to starter.

All three of these players have a close proximity to top performers in the opposite role based on dropoff data. Immediately, this flags them as candidates for transitioning, but all of them have season-level stats that support these notions. Beginning with Bumgarner, his proximity to top relievers in his pitch-by-pitch data combined with his above-average BABIP+ and close to league average CSW% hint at a potential suitability for a relief role. As for Stroman, while he is a great starter (as evidenced by his All-Star selection this year as well as his above league average BB%+), he also showed attributes that could make him an exceptional reliever. His impressive CSW% and elite GB/FB% (2.10, 61% higher than the league average) support this notion. Finally, we believe Thielbar would be better suited as a starter due to his great BB%+ of 79.53, nearly 21% better than league average. This is further evidenced by him having a below average BABIP+ and significantly below average GB/FB% (0.66, nearly 50% below league average), indicating he may not be in his best fit as a relief pitcher.

In this analysis of starting pitchers and relievers, we created novel dropoff statistics to identify pitchers whose statistical profiles suggest they might be better suited to opposite roles. Regression analyses on season-level stats helped further reduce our population and pinpointed key indicators for success in these roles, guiding our recommendations for potential role adjustments. Our findings can help bring enhanced competitive advantage for MLB teams through the alignment of pitcher roles with their inherent performance strengths.