# Reds Data Challenge

Reece Calvin, Tim Clay, Vivek Divakarla, Jacob Gottesman, Daniel Vonderwalde

## Introduction

In 2023, starting pitchers with over 30 IP averaged 1.4 fWAR. Qualified relievers average 0.5 fWAR over the same span. Tanner Scott and Felix Bautista, the 2023 leaders in reliever fWAR (2.8 fWAR), would have ranked 37th overall in pitcher fWAR. We believe this fact shows that unless a reliever can consistently perform at an elite level, think Josh Hader, they hold more value as a starter. For this reason, our goal was to identify relievers who showed signs they could be a serviceable starting pitcher.

## Preliminary Inquiries

To identify transition candidates, we wanted to first understand the quantitative differences between starters and relievers. The benefit would be to see which indicators exist that would lead a team to typically pin a pitcher as a starter as well as see if there are tangible differences when a given pitcher moves between starting and coming from the bullpen.

One accepted notion about starters is they generally throw a larger variety of pitches. Relievers threw 3.22 unique pitches, on average, while starters threw 4.16. This is a clear indication that starters should have more unique pitches. A deeper dive into the data can help us find out if that is managerial preferences or a true symptom of pitchers with fewer unique pitches struggling as starters. Looking at Stuff+ grades for relievers, we saw significantly higher numbers than for starters (104.6 vs. 96.0). We identified two possible reasons for this, relievers will eliminate their worst pitch, say a 40 grade CH (20/80 scale), even if it only leads them with their 70 SL and 65 FB. Meanwhile, a starter may have to throw all 3 pitches to last longer in games, bringing down their average. A second possibility is pitchers may throw harder when they know they will be in for a shorter amount of time. This would typically increase Stuff+ grades.

Before we began modeling, we wanted to define a metric for success. As previously stated, we find it very rare for a pitcher with even average SP upside to be better off as a reliever. For this reason, we decided that we would seek out league-average starting-caliber pitchers. Think 2023 Yu Darvish. We defined this as a SIERA at or below 4.00.

Initially, we attempted to directly predict a pitcher's SIERA given only their pitch attributes. Our model returned an $R^2$ score of .63. The most important features we found were botERA, Pitching+, OSwing%, OContact%, Contact%, and Stuff_plus.

# Initial Findings

We first looked at the pitchers tagged as *Hybrids*, we defined Hybrids as pitchers who threw at 500 pitches as a starter and 500 pitches as a reliever since 2021.

In Fig. A, we plotted pitch characteristics for pitchers in each of the two roles. We compared their metrics with their average for each given pitch type. The most clear change in relievers and starters is when pitchers come out in relief, they throw harder. In the general sense, we are confident in taking that as fact. Knowing this, determining a pitcher's success when transitioning from reliever to starter can not be simply looking at their pitch grades and expecting it to translate. Their grades will change following their changing pitch characteristic.

In Fig. B, we show the results of our logistic regression function. With it, we were able to predict if a hybrid pitcher was in a starter role or a reliever role, based purely on their delta velocity. This only reinforces our previous findings.
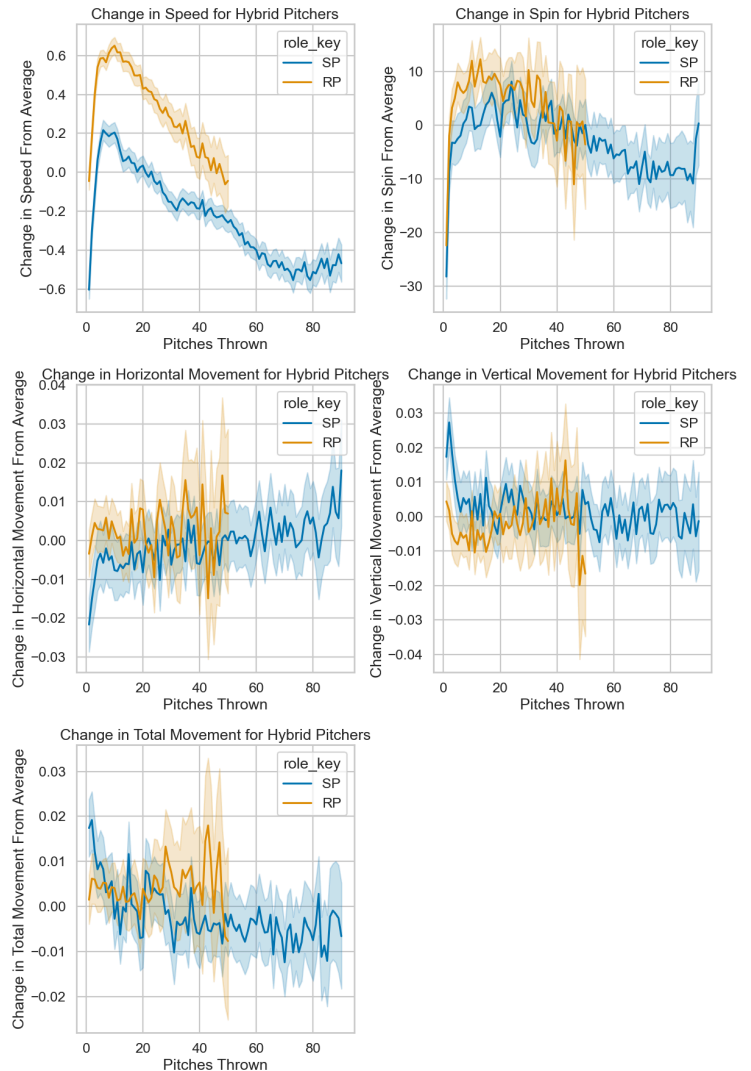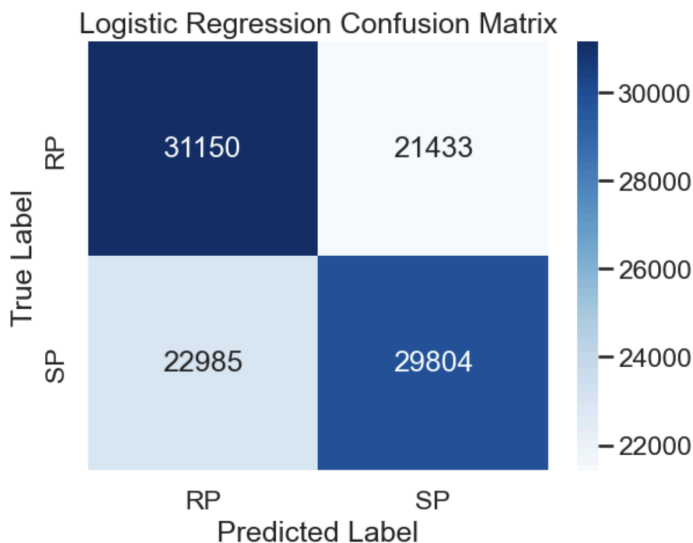
## Fig. A



## Fig. B



Next, we looked into the opener and follower roles. While we know relievers and starters behave differently, is the effect different for openers and followers?

We initially tagged a pitcher as an opener if they were pulled in fewer than 7 PAs and allowed fewer than 6 runs. We then ruled a pitcher as a follower if they came in after an opener and faced 15 PAs. We then retagged openers to see if they were replaced by a follower in an effort to mitigate false tags for injuries or other reasons.

In Fig. C, one can see the value of this approach. Again, only looking at delta velocity, it was able to identify followers really well. Essentially, if a pitcher comes in as a reliever but is

## Fig. C



Logistic Regression Confusion Matrix
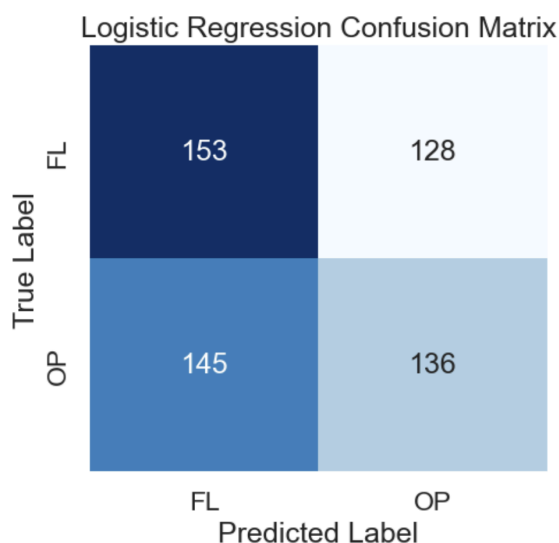
not throwing as hard as we have seen, they probably are a follower. We believe our issue with tagging openers is that it is hard to know if a pitcher was pulled for another reason. We believe it is much more common for a true starter to be pulled early than a true reliever to be left in for 15 PAs.

# Modeling

We went with an XGBoost classifier model. We used a binary SIERA under 4.00 tag as our target variable. Our features include the pitcher's arsenal, pitch grades, CSW%, Hard Hit Rate, and Barel %. We fit our model parameters using bayesian optimization.



Top Combined Feature Importance
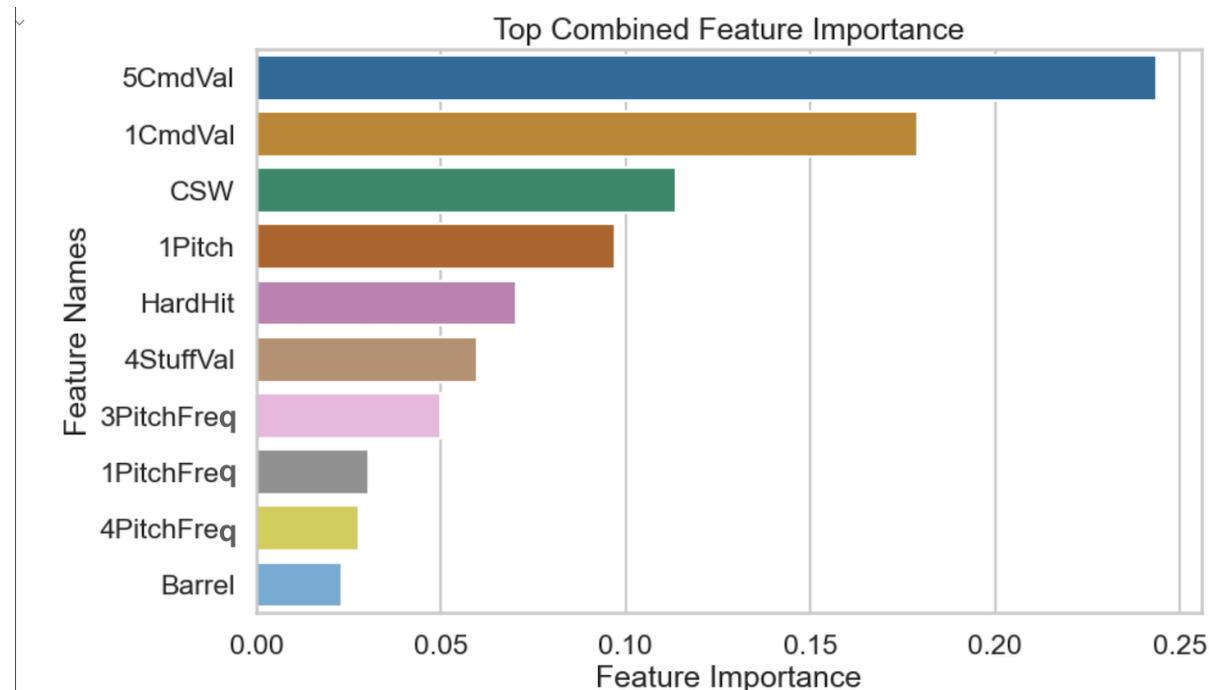
# Suggestions

We chose to highlight 3 of the pitchers our model suggested:

Our first candidate for a transition to starter is Carmen Mlodzinski. He has a 4-pitch mix with a 46-grade fastball, 60-grade slider, 64-grade ChangeUp, and a 52-grade Cutter. He has an impressive Barrel Rate of 4% and a Chase Rate of 28%. His GB% is also above average. Our model sees these as good indicators of a successful transition to starter. His K% leaves something to be desired at 22% and his BB% is 9%. With only 34.2 innings pitched, it's hard to say if these numbers are sustainable. He's a candidate for a transition to starter, but it's not a slam dunk.

Someone we feel more excited about is Jordan Hicks. While he only has 3 pitches, they all grade out as 60 or better. He has an advanced K rate and GB%. His BB% is a little high, but he's a candidate for a transition to starter. What makes him even more exciting is his 100 MPH fastball. Of course, with that comes injury risks but we also know that his velo. will likely fall a little. Despite this, his stuff should still grade out well. He's a candidate for a transition to starter.