Cincinnati Reds

Pitch Mix Report

Summary

This report will consist of 4 parts:

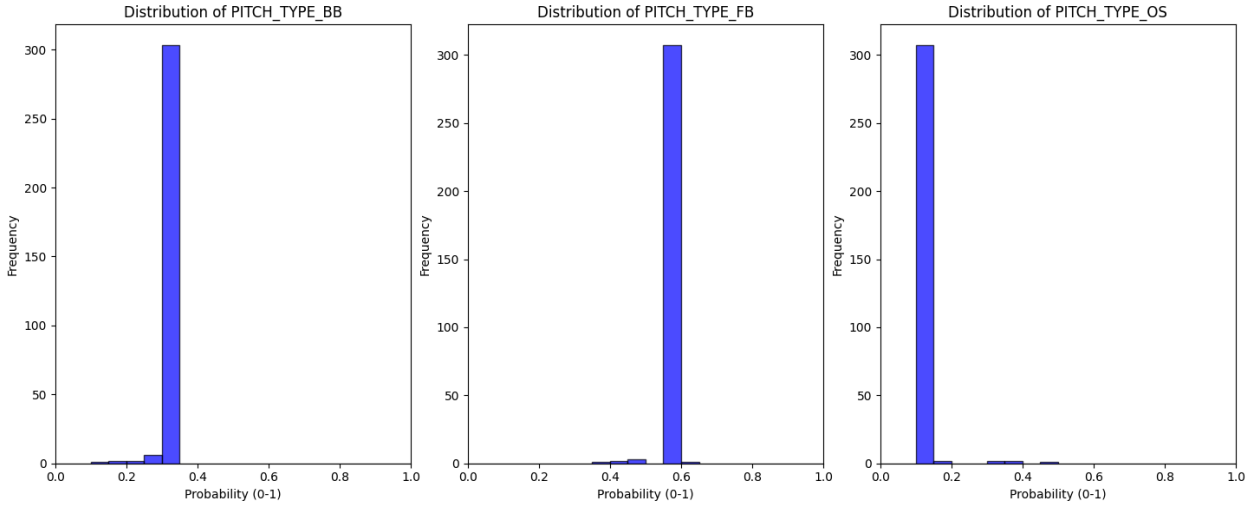
1. **A short non-technical explanation of the model used**
2. **A discussion of the results from a statistics lens**
3. **A deep-dive into the implications of the results from a baseball oriented lens**
4. **A discussion of nuance in the model and its results**

For those who don't have time to read this entire report please skip to the Closing Summary on the final page.

The Model

The model I have created takes a batters stats from a season and predicts the next season's percent distribution of fast balls, breaking balls, and off-speed balls they should expect to see the following season. I used this model to predict the pitch mix for all batters in 2024 that received at least 1,000 pitches in 2021-2023. These predictions can be found in this github in the predictions.csv file which can be explored in both Excel for windows users or Numbers for mac users.

The Results



We can see from the above chart that for a large majority of players, our model predicted the following approximate pitch mix – 32% Breaking Balls, 57% Fast Balls, 11% Off-Speed Balls. Since these results were so prominent, it may be safe to assume that for batters with less than 1,000 pitches in 2021-2023, or new batters to the league, we should assume this breakdown of pitches is a safe starting point. When it comes to outliers, we can see one major group arise. This group is those that receive a smaller percentage of fastballs but a much greater percentage of off-speed balls. Some players that fall into this category are Abraham Toro of the Athletics, Earnie Clement of the Blue Jays, Otto Lopez of the Blue Jays, and Joey Ortiz of the Brewers. These players are all predicted to receive above 30% of off-speed pitches in 2024. This may point to the fact that these players have struggled against timing mix-ups in their pitches in the past, and should work on dealing with off-speed pitches in specific for the coming seasons. This also may indicate to the Reds that they should instruct their pitchers to mix in these pitches more against these players, as they may prove greatly effective.

In general due to the overall small percentage of outliers, they are in general safe to ignore. Instead the Reds could benefit by training on similar percentages of pitches to the 32%-57%-11% breakdown above which seems to be greatly unaffected in the predictions for the majority of batters. For those that these results do vary, it presents a good opportunity for pitchers to abuse these weaknesses of these players by pitching a different pitch mix to their usual mix.

Nuance of Results

Despite the strong statistical model behind these predictions, this doesn't mean they are perfect. First and foremost I'd like to bring attention to the fact that the model having similar predictions for almost all batters may be a result of these predictions just being the "safest bet". This takeaway definitely does have its positives as for batters we don't have data on this "safe bet" is still useful, but for batters we do have data on it may be too "safe" and overgeneralizes most batters. The second major thing to note about this model is that results may vary if run again. The software behind this model may draw different conclusions from the data if run again and may result in predictions that are different each time. Despite this, over 10 identical test runs of the model we saw no greater than a 2% change in any given percentage of the results (i.e. fast balls could range from 55% - 59%). These might indicate that it is fair to assume that our true pitch mixes for the 2024 season may fall somewhere within these 2% ranges for each pitch type.

Closing Summary

Despite flaws in this predictive model, it definitely gives a good insight into a general pitch mix to expect for most players (32% Breaking Balls, 57% Fast Balls, 11% Off-Speed Balls). This can be used most effectively in percentage of pitches thrown in training, or when predicting newer players expected pitch mixes. The model also recognized multiple batters who would receive pitch mixes drastically different to this, which may be useful to exploit. Overall the model is a good starting point for our pitch mix analysis, and could be improved with greater time and data.