

# Seattle Mariners Baseball Projects 2023 Application: Round Two



The Mariners Baseball Projects team combines baseball analytics, project management, and software development to create innovative tools for use within the organization. While the Projects team does assist with traditional R&D tasks, its primary focus is to provide analytical tools to streamline decision-making processes.

A representative example of the type of work the Baseball Projects team does is collaborating with our coaching staff to design, implement, and monitor player development plans for each player.

As a hypothetical example, imagine you're a member of the Baseball Projects team and are presented with the following scenario:

---

Towards the end of the 2022 season, Pitcher X approached one of our coaches and said that his goal in 2023 is to maximize the number of swing-and-misses (whiffs) that he generates on a per-pitch basis. The coach has asked you to provide recommended data-driven strategies that Pitcher X can pursue to optimize his whiff rates, based on his data from the last season.

You have been provided with 2022 data of Pitcher X's pitch-by-pitch performance. We've created a column for you, `whiff_prob`, that is the probability that the given pitch will result in a swing-and-miss based on its velocity, movement, location, and release. Based on the provided data, identify strategies that Pitcher X can work towards implementing to maximize whiff probability moving forward.

*Note: For the purposes of this exercise, neglect any pitch sequencing-based strategies, i.e., each pitch may be considered independent of the pitch(es) that preceded it.*

Please create a 10-minute screenshare presentation for the coach sharing the recommendations you unearth from the data.

---

Logistics for the Round 2 Interview:

- Expect a 30-45 Teams video call interview.
- After introductions, you'll screenshare & give your 10-minute presentation to members of our team, who will be listening as the hypothetical coach in the above scenario.
- After your presentation is complete, we'll go through some debrief questions and further discuss the design decisions that you made throughout your analysis and presentation.
- The dataset is designed to be self-contained as a whiff-optimization breakdown given the provided whiff probability.
- Further details about the data are provided on the next page.

If there are any unclear elements of the exercise above, please do not hesitate to reach out ahead of time to David Hesslink at [dhesslink@mariners.com](mailto:dhesslink@mariners.com).

# Seattle Mariners Baseball Projects 2023 Application: Round Two



Column	Description
Date	Date pitch was thrown on
Year	Season that pitch was from
Pitcher	Pitcher identity
PitcherThrows	Pitcher handedness
Inning	nth inning of the game
PAofInning	nth batter of the inning
BatterSide	Batter handedness
PitchofPA	nth pitch thrown to this batter in this plate appearance
Balls	Balls in count before pitch is thrown
Strikes	Strikes in count before pitch is thrown
PitchType	The classification of the pitch thrown
ReleaseSpeed	Velocity of pitch
InducedVertBreak	Vertical movement of pitch, in inches, where zero is a "gravity-only" trajectory
HorzBreak	Horizontal movement of pitch, in inches
ReleaseHeight	Distance above ground pitch was released from, in feet
ReleaseSide	Distance from center of mound pitch was release from, in feet, positive numbers indicate closer to 3B side of diamond
Extension	Distance in front of the mound (towards the plate) that pitch was released, in feet
PlateHeight	Distance above ground that pitch crosses the plate, in feet; zero is at ground level
PlateSide	Distance from centerline of plate that pitch crosses, in feet; zero is over the center of the plate and positive numbers indicate closer to right-handed batter's box
SpinRate	How fast the pitch was spinning, in rotations per minute
SpinAxis	The angle or "tilt" that the pitch was spinning at; 0 is true topsin; 180 pure backspin; 90 sidespin towards left-handed batter's box; 270 sidespin towards right-handed batter's box
swing_prob	The probability that given the metrics above, this pitch will result in a swing
whiff_prob_gs	The probability that given a swing, the hitter will miss (whiff probability conditional on swing); short for "whiff probability given swing"
whiff_prob	The probability that given the metrics above, this pitch will result in a swing and miss; i.e., $sw\_prob * whiff\_prob\_gs$