Atlanta Braves R&D Questionnaire

**Analysis Questions**

Please limit each answer to a maximum of 500 words.

1. On 8/24/2021, the Cardinals trailed the Tigers 4-3 going into the top of the 9th. To begin this inning, Daz Cameron doubled, Akil Baddoo struck out, and Jonathan Schoop grounded out, moving Cameron to 3rd. The batter is now Robbie Grossman. Assume Luis Garcia will pitch through the 5th spot in the batting order, Jeimer Candelario. Should the Cardinals intentionally walk Grossman? Describe what your process would be to determine whether to pitch to him. The following link contains the box score information for this game: <https://www.mlb.com/gameday/tigers-vs-cardinals/2021/08/24/632781#game_state=final,lock_state=final,game_tab=box,game=632781>
2. You are running a generic mid-market team and are exploring the idea of signing Cody Bellinger this offseason. What contract would you be willing to offer him? Please explain your thought process and discuss any important considerations.
3. Pitcher A walks half the batters he faces and strikes out the other half. Pitcher B doesn’t walk or strike out any of the batters he faces. Which pitcher would you prefer? What ratio of strikeouts to walks would make you indifferent between the two pitchers?
4. Briefly explain how you would go about estimating the effect of catcher framing at the major league level? Assume you only have access to the identities of the people involved, information about the pitch (location, characteristics, etc.), and information about the game (count, inning, score, etc.).

**Modeling Questions**

For these exercises, you will be using your knowledge of R or Python to answer a few baseball-related questions.

Use the attached Trackman pitch by pitch data of Braves pitchers from the 2018 season to answer the following prompts. It is not necessary to utilize every column in the attached file; only use those you feel are necessary. There is a GLOSSARY defining the columns in the PitchData.csv file on the second page of this document.

This exercise should not take more than a few hours. Please include all of your code with your responses. If you do not know how to complete one or more of the questions, feel free to leave them blank.

* 1. **Create TWO models to predict the likelihood of a swing and miss based on the characteristics of a curveball. Evaluate and compare the performances of your models using any method(s) you’d prefer. Explain your results in 500 words or less.**
  2. **Using your preferred model from Question #3, create a visualization to display the most important characteristics of a curveball in recording a swing-and-miss. Explain your visualization in 500 words or less.**

**Note:** Models in this exercise will be less accurate due to small samples of pitches and pitchers, so proceed with your evaluations and conclusions as if there were a complete set of 2018 data.

**PitchData.csv Glossary**

|  |  |
| --- | --- |
| **Variable** | **Definition** |
| Pitcher\_ID | The pitcher’s MLBAM ID |
| Pitcher | The pitcher’s full name |
| Pitcher\_Throws | The pitcher’s handedness |
| Batter\_ID | The batter’s MLBAM ID |
| Batter | The batter’s full name |
| Batter\_Hits | The batter’s handedness |
| Game\_Date | The date the game occurred |
| Top\_Bot | Whether it is the top or bottom of the inning (1 signifies the top and 2 signifies the bottom) |
| Inning | The inning the pitch was thrown |
| Balls | The number of balls when the pitch was thrown |
| Strikes | The number of strikes when the pitch was thrown |
| Outs | The number of outs when the pitch was thrown |
| Pitch\_Outcome | The outcome after the pitch was thrown |
| Pitch\_Type | The pitch type (4-Seam and 2-Seam are grouped as fastballs) |
| release\_speed | The pitch’s velocity (mph) |
| x\_movement | The pitch’s horizontal movement (inches) |
| z\_movement | The pitch’s vertical movement (inches) |
| release\_spin\_rate | The pitch’s spin rate (rpm) |
| spin\_dir | The pitch’s spin axis (degrees) |
| release\_pos\_z | The horizontal release point for that pitch (ft) |
| release\_pos\_z | The vertical release point for that pitch (ft) |
| release\_extension | The release extension for that pitch (ft) |
| plate\_x | The horizontal location of the ball when it crosses home plate (ft) |
| plate\_z | The vertical location of the ball when it crosses home plate (ft) |