**Atlanta Braves – R Questions 2019**

For these exercises, you will be using your knowledge of R 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. **Prepare a scatterplot that shows each player’s average fastball velocity and spin rate from the 2018 season, as if you were presenting it to a baseball executive. Plot velocity on the x-axis and spin rate on the y-axis.**
  2. **Find all pitchers in the dataset with a slider usage rate of at least 20%. Arrange those pitchers in a data frame from the highest slider usage to the lowest. Print a data frame of the top 5 pitchers in slider usage, including how many sliders and total pitches they threw.**
  3. **Using Trackman data from before July 1st of the season, build TWO different pitch classification models to classify all pitches in time period from July 1 to the end of the season: Fastball, Changeup, Curveball, Slider and Cutter. Evaluate and compare the performances of your models using any method(s) you’d prefer. Explain your results in 500 words or less.**
  4. **Create a model to predict the likelihood of a swing and miss based on the characteristics of a curveball. Create a visualization to display the most important characteristics of a curveball in recording a swing-and-miss. Explain the results of your model and 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) |