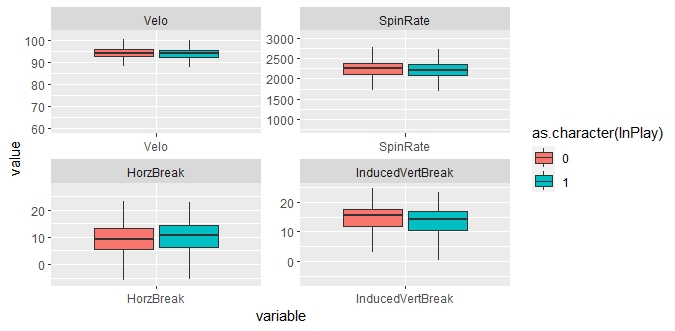
The process for answering question 1 began with looking at the data to see what the best method for answering the question was. Seeing that 1 in 0 were the response variables, my mind went straight to logistic regression since they are binary variables. Next came importing the data into R then checking the variables and seeing that SpinRate was a character because of the NULL value. A quick fix to make it numeric and then it was time to make the model. The first model used all variables and when looking at the summary showed a SpinRate p-value of 0.096. I usually use α = 0.95, so there was no reason to keep SpinRate in the model since it was not significant. The new model looked good with significance, so it was then time to test it with the deploy data. Then it was time to add the predictions in a new column and create the .csv file.

With a 4 seam fastball throwing hard with a high spin rate leading to more induced vertical break and avoiding too much arm side run nearing a dead zone, will make for a more effective fastball in relation to the ball being put in play.



I believe the next step for this model would be looking at data against left-handed hitters and seeing what can get more swing and misses.