

PS6 *Gillingham*

Andrew Gillingham

March 2019

1 Visualizations

In this problem set I wanted to create three different and interesting visualizations that would be both pleasing to the eye and have an impact on my data moving forward. After scouring the internet for ideas on what I could do I had a plan to create some incredible visualizations using the programs pybaseball and baseballr that I had recently downloaded. My first visual was going to be a "bump" chart that followed the AL West and showed how they were ranked throughout the season. After following the code word for word for almost 2 hours, my final code, for whatever reason, didn't work and I had no clue as to why. So I scaled down my expectations and tried looking at the visualization chapter in "Analyzing Baseball Data in R". After downloading the required files, I found out that they did not have all of the files and that I would end up having to use my own.

The first scatterplot I created actually used the data from the books data file. I used Hall of Fame voting data to show the increase of votes needed in order to get into the Hall of Fame. I also wanted to illustrate that there have been several years where it has actually been easier to get in because the class was weaker. These blue dots can be seen significantly lower than the regression line. The bar chart was not used using ggplot but was to illustrate how launch speed can have an affect on hit outcome. For this I ended up making two new data frames. The first one split the data into outcomes and the second found the mean launch speed of each outcome. I tried figuring out an easier way to do that directly in the original data and ggplot but it didn't work so I improvised.

The final visualization may not be as clear as the others, but this is a heatmap of balls that were put into play. The lighter blue areas show where in the strikezone Astros hitters have been swinging and making contact with pitches. Although this one is a bit rough, I believe that by the end of the semester I will be able to make individual heat maps for players that I am analyzing to get a sense of where they like to make contact.