Individual Technical Deliverable #1

2.980 -- MLB Automatic Strike Zone Project

Devon Goetz

Problem Statement:

Choose a set of 10 games and visualize balls called, strikes called, and mistakes made by the umpire.

Solution Approach:

To achieve these plots, I broke out data from April 21st, 2018, a date on which 15 games were played, for a total of 904 pitches.. I knew I wanted to examine the difference between calls made on right-handed and left-handed batters, so I divided the data based on batter handedness. I took each pitch call and sorted these into called balls and strikes, and then filtered which ones were correct calls and which were incorrect. I wanted to compare the different batter sides to see if there was a notable difference in what umpires were calling against them.

Graphs and Data:

In my two strike zone plots, we see anything called a strike represented by a diamond, and anything called a ball represented by a circle. All wrong calls are in red, while correct calls are in blue.

The strike zone was plotted and used to determine what made a strike versus a ball. The width is consistent, and anything outside of the 17-inch plate was a ball. However, the strike zone changes vertically based upon the height and stance of the batter, so chose to make it from 1.5 ft to 3.5 ft, a loose approximation.

In the future, I would plan on using the batter specific strike zone for pitch classification, but would likely maintain a generic strike zone for visualization (unless we were visualizing just one batter).

Figures 1 and 2 below are strike zone visualizations divided based on batter handedness, in order to try and see the difference between the calls that left-handed and right-handed batters are seeing.

Figure 3 is a bar chart that shows the percentage of pitches that batters saw divided into four categories: Correct Strikes, Correct Balls, Balls as Strikes (called strikes that should have been balls), and Strikes as Balls (called balls that should have been strikes).

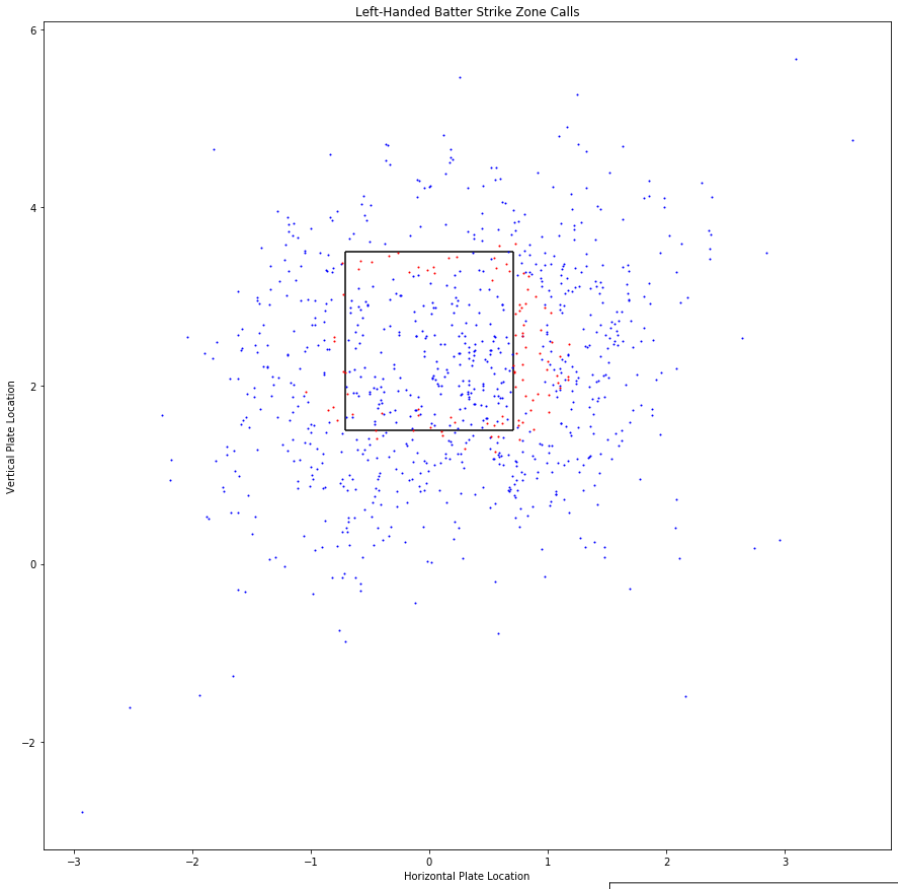

Figure 1. Strike zone splay chart for Left Handed Batters (pitches on 21APR2018).

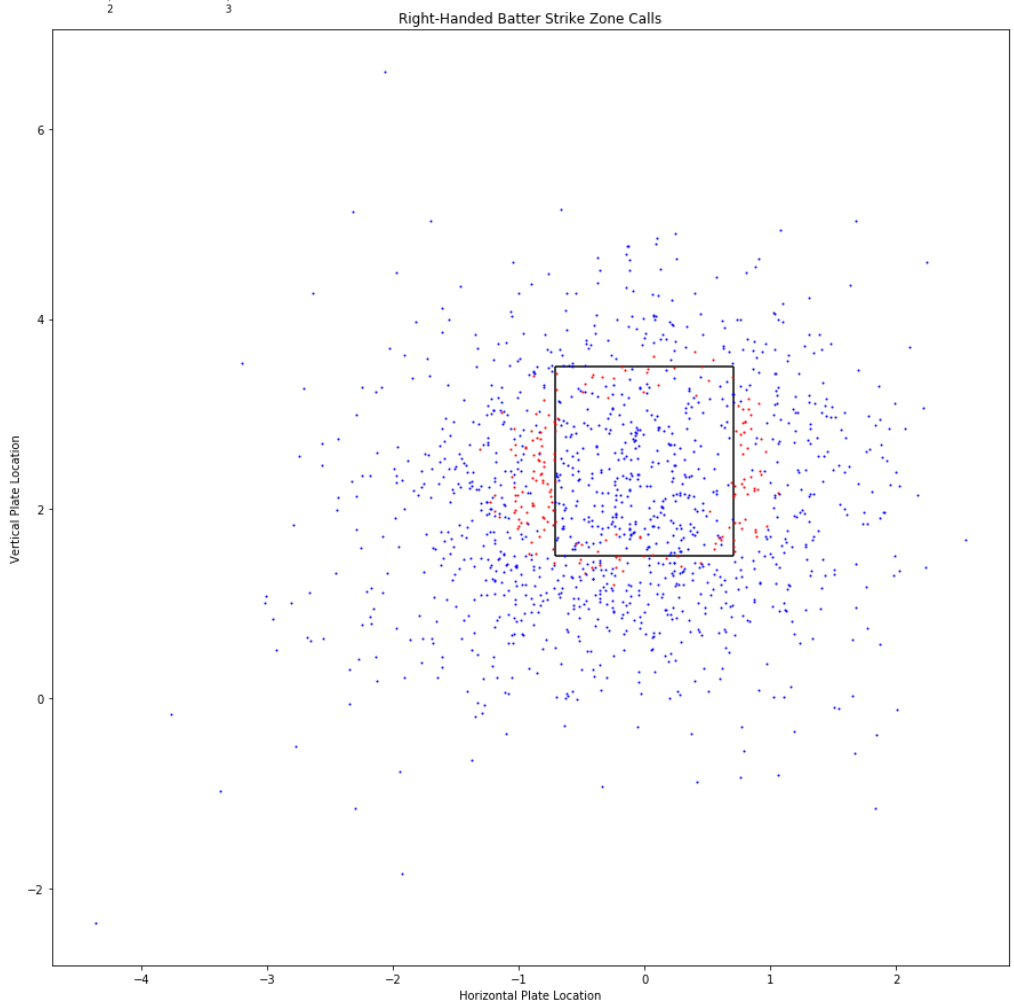

Figure 2. Strike zone splay chart for Right Handed Batters (pitches on 21APR2018).

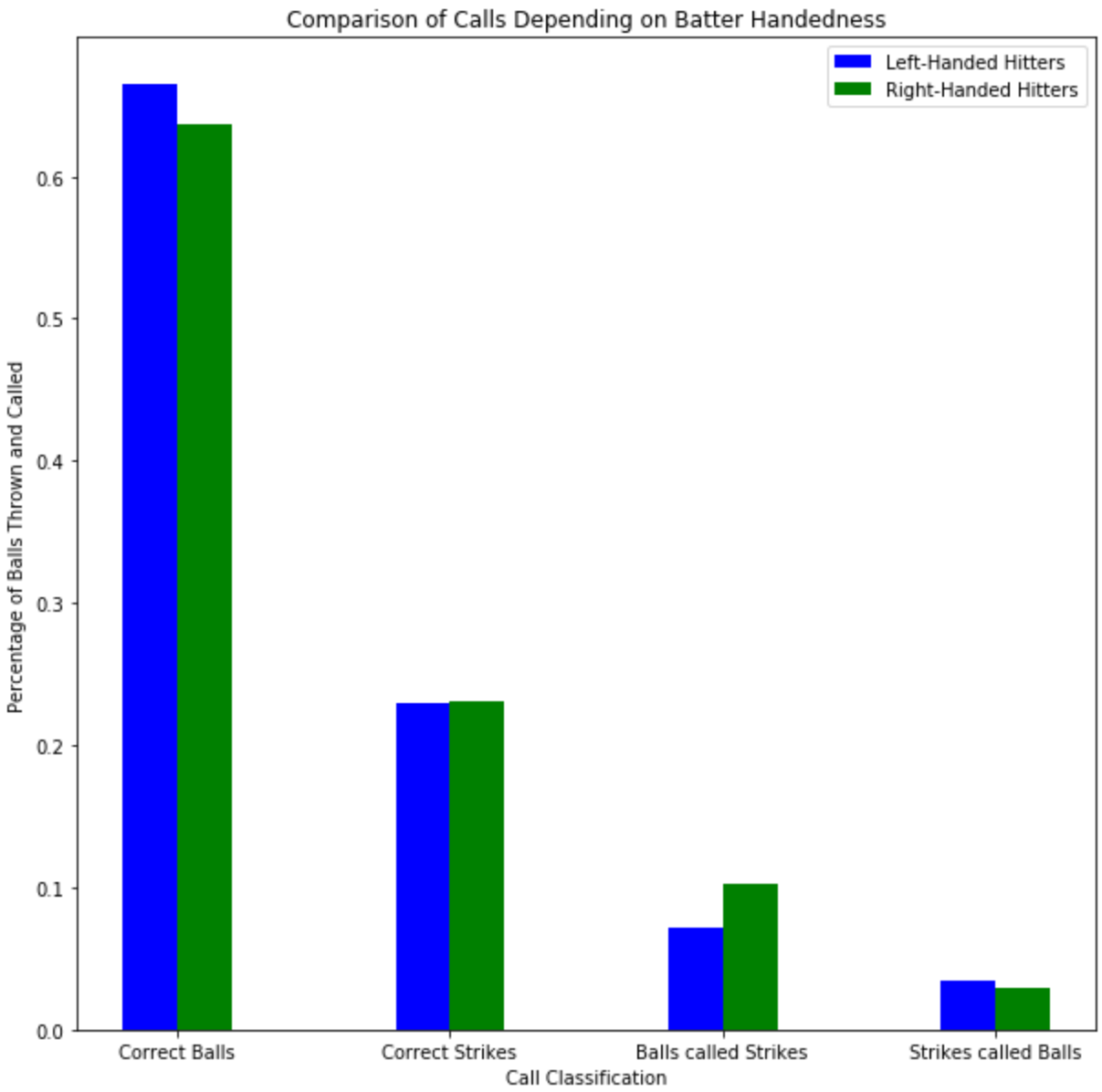

Figure 3. A bar chart showing the percentage of pitches that each kind of batter saw.

Insights Gained:

From a quick glance at the pitching sprays, we can see umpires are missing in the horizontal directions more than the vertical direction. Further, it appears that the umpires are missing more outside than inside on batters, regardless of handedness. It looks like some of the left-handed incorrect outside strikes are farther than the right-handed incorrect outside strikes, but I do not have statistical analysis to back up that observation.

An interesting next step would be to see if there is a statistical difference in what we're seeing, and also if there is a difference on *where* these missed calls are happening. It would be

exciting to continue forward with this examination, because it might reveal an inherent umpiring bias for or against one particular group of batters in a particular part of the zone.

Based on the bar chart, it would seem that righties and lefties are relatively close in the percentages they're seeing, but I think it is interesting to note that left-handed batters see more true balls. This could imply that pitchers struggle more to throw to left-handed batters. It would be interested to cross correlate this with the handedness of the pitcher.

Additionally, right-handed batters get more balls called mistakenly as strikes, while left-handed batters get slightly more strikes mistakenly called as balls. Although both of those margins are small, they would appear to prefer left-handed batters.

Suspicious Data:

Overall, I would say there is not much suspicious data, which is great news! There are a few pitches that are strikes but were called balls that seem to be very much in the strike zone which seems weird, but my hypothesis is that these were extreme breaking balls that crossed the front of the plate in the zone and move a lot between the front of the zone and the catchers mitt, causing the catcher to move laterally and bias the umpire away from calling the strike.