# Thomas Thornton's CS251 Project 10

22 Added by Thomas W. Thornton, last edited by Thomas W. Thornton on May 09, 2014 (view change)

**Abstract:**

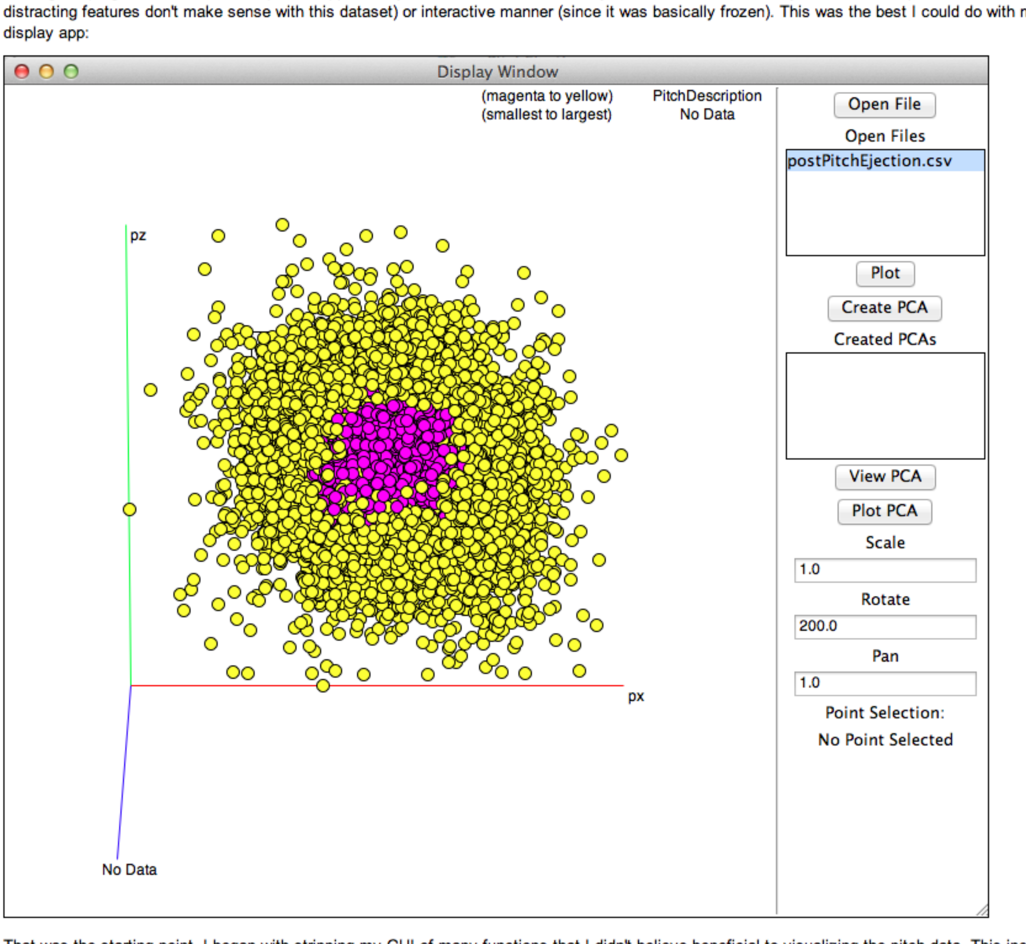
The strike zone dataset, provided by Travis Carter, consists of over 170,000 different pitches with over 30 different features of information about each pitch, including it's coordinates relative to the strike zone, whether it was called a strike or a ball, the batter's height who the pitch was thrown to, and more. For my area of investigation, I was only interested in pitches that had to be judged and subsequently called (either a strike or a ball) by the umpire, so this excluded a pitch thrown in the dirt, a pitch a batter swung at, etc. And, I only looked at pitches that had been thrown during games that had a pitch-related ejection (of either a player or manager), because a team member berating an umpire about pitches should affect the umpire's judgement more than an unrelated complaint/ejection. After pruning with these criteria, the dataset was down to about 39,000 pitches. From here, I divided the remaining pitch data into two groups: pre ejection and post ejection pitches. I further divided these two groups in regard to team, as either the team that had the ejected player/manager, or the team that did not. I used the provided pitch coordinates to differentiate called strikes from actual strikes, and called balls from actual balls. These can be understood as the miscalls the umpires made. This is where the majority of my project is focused: comparing miscalls, whether it be an actual strike called a ball (favorable for the team at bat), or an actual ball called a strike (not favorable). For visualization purposes, I generated random samples from the data, so a different subset is plotted each time the application launches. The visualization canvas is divided in half, each side has a strike zone and the pitches thrown to that team, plotted on their side. Each pitch, as a data point, is plotted based off of its coordinates and colored based off of its call as either a ball or a strike. Additional colors can be activated to see the miscalls. Pitches can also be selected by clicking, in order to see the numeric information. For my analysis, I worked with the entire data files, rather than the randomly selected subset. I began with computing base rates of miscalls between teams, before and after ejections. With these statistics, I used Bayes Theorem to formalize the actual probabilities of all possible pitch outcomes for each team, pre and post ejection. In light of the potential trends revealed from this, I conducted two ANOVAs specifically targeted at the unfavorable calls (actual balls called as strikes) to determine if these miscalls differ significantly between the teams, and if the miscalls differ pre and post ejection for the team with the ejected team member.

**Problem Statement:**

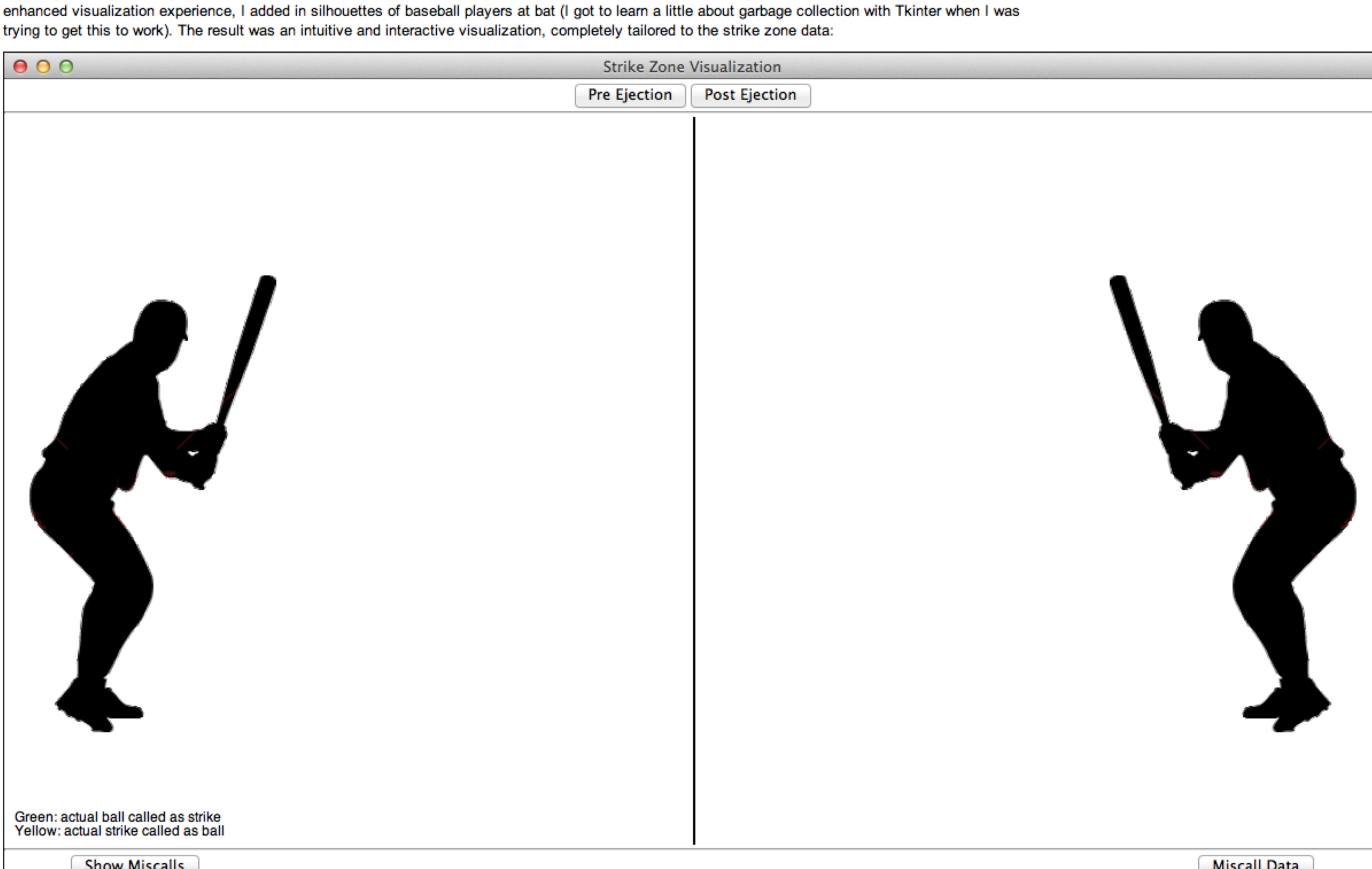
On many pitches throughout a baseball game, it is solely the judgement of the umpire that determines whether a pitch is treated as a strike or a ball. The consequences of the umpire's calls are huge, and it can even lead certain team members, whether it be the managers or the players, to lash out verbally against an umpire if they feel the call is unjust. After an umpire ejects a team member because of one of these heated verbal attacks, does this interaction leave some residual effects on the umpire's judgement? For many years, it was almost impossible to answer this question. When trying to compare the calls of an umpire before and after an ejection, there was no way to know for certain if a pitch was an actual miscall, or just a close, but correct call. The dataset provided by Travis Carter contains the information necessary to this area of investigation. Along with the aforementioned question, does an ejection affect the frequency of unfavorable calls for either team? Seeing an actual ball called as a strike is what provokes the berating of an umpire, but does getting ejected for this berating actually affect the umpire's subsequent calls? Is berating an umpire an effective strategy to combat unfavorable miscalls for the ejection-affiliated team?

**Methods:**

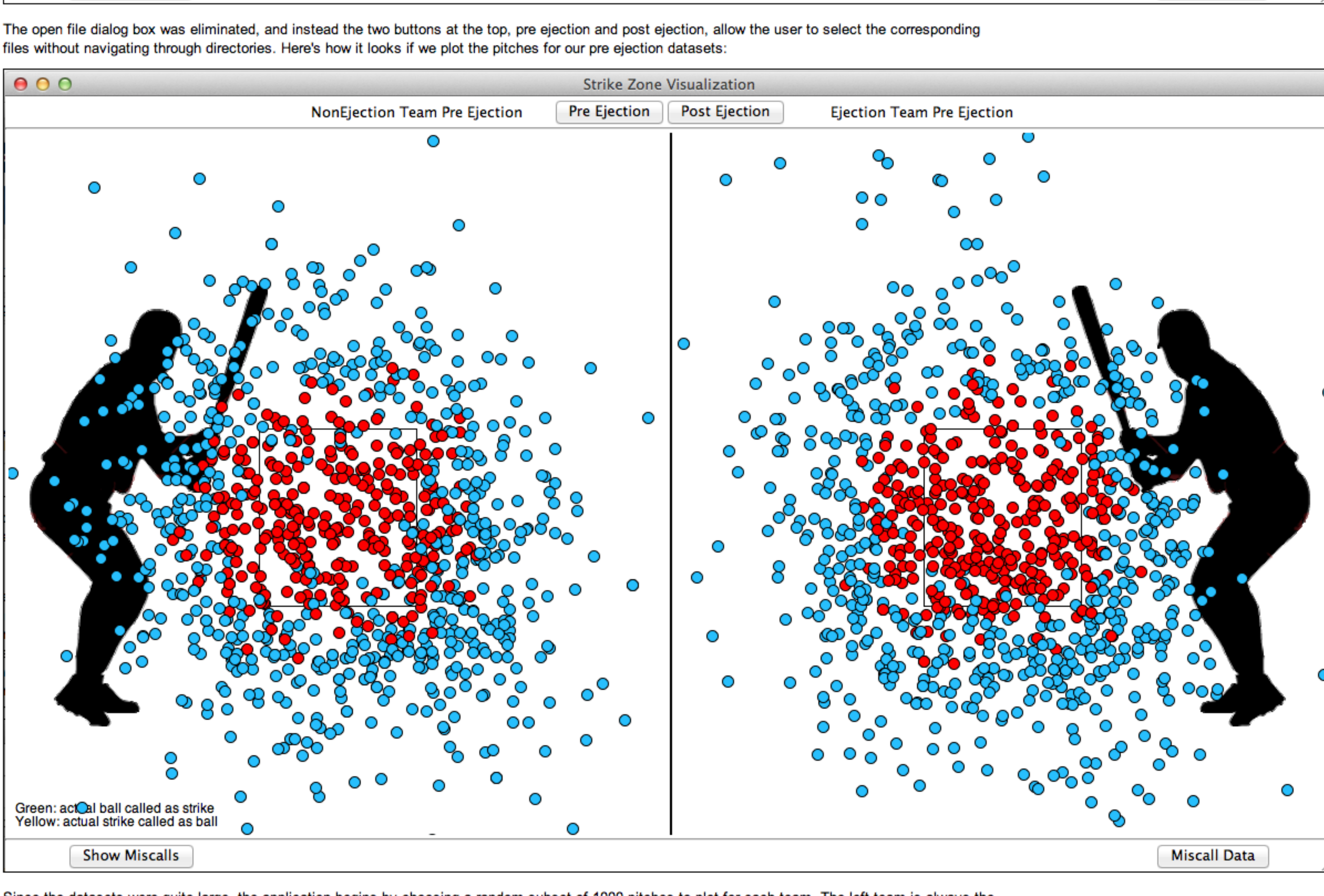
Once the pruning of the dataset was complete, it was time to move onto visualizing the data. The resulting files were still too large for plotting (without slowing the program to almost a complete stop), and my current GUI couldn't convey the information in an easy-to-understand (axes and some of the other distracting features don't make sense with this dataset) or interactive manner (since it was basically frozen). This was the best I could do with my current display app:



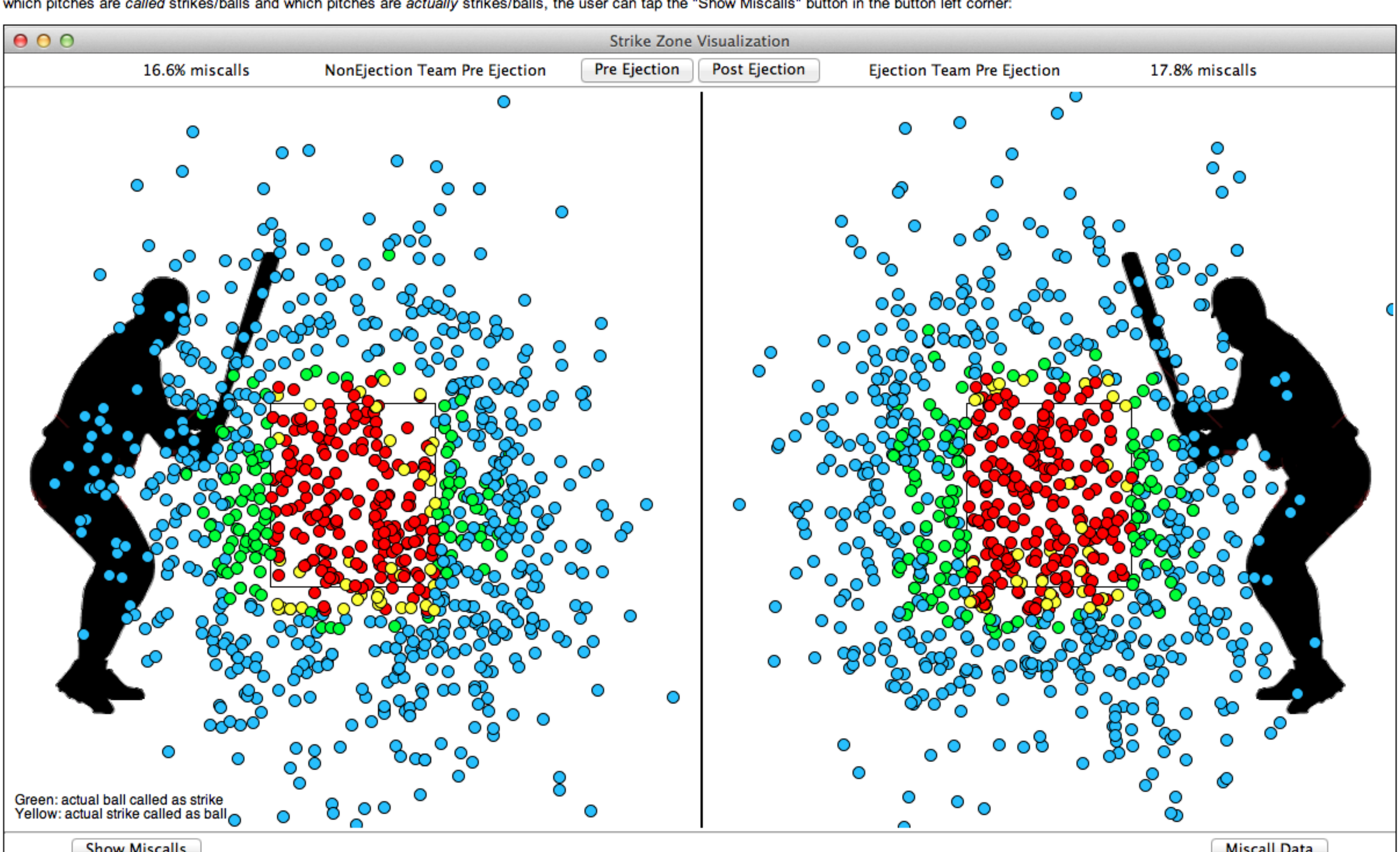

That was the starting point. I began with stripping my GUI of many functions that I didn't believe beneficial to visualizing the pitch data. This included disabling rotating, panning, PCA, constant controls, and more. I also wanted to maximize the space, so I redesigned the control frame. Additionally, for an enhanced visualization experience, I added in silhouettes of baseball players at bat (I got to learn a little about garbage collection with Tkinter when I was trying to get this to work). The result was an intuitive and interactive visualization, completely tailored to the strike zone data:



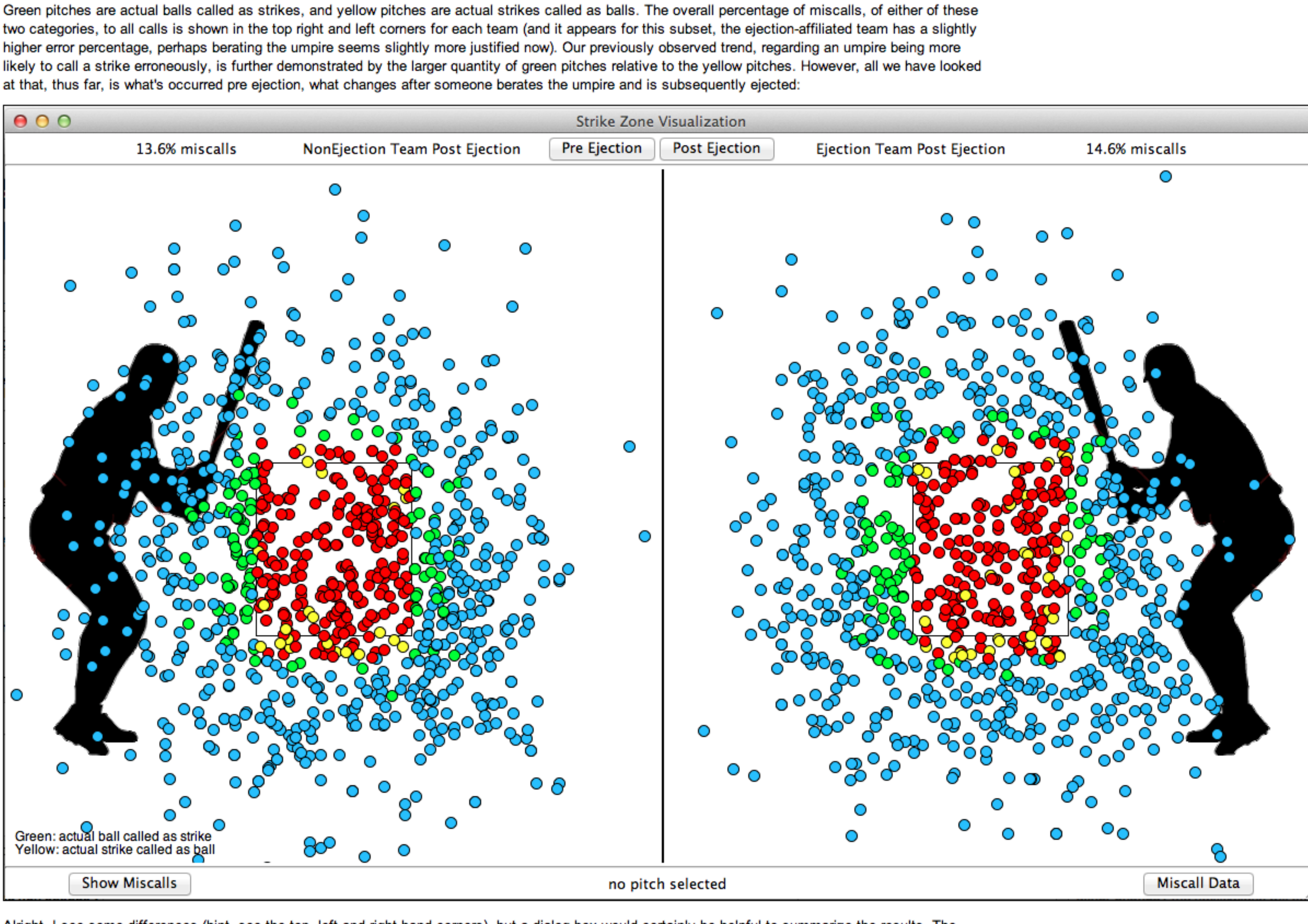

The open file dialog box was eliminated, and instead the two buttons at the top, pre ejection and post ejection, allow the user to select the corresponding files without navigating through directories. Here's how it looks if we plot the pitches for our pre ejection datasets:




Since the datasets were quite large, the application begins by choosing a random subset of 1000 pitches to plot for each team. The left team is always the non ejection team, and the right is the ejected team (meaning which team the ejected team member was part of). In baseball, the strike width remains constant at 17 inches (the width of home plate) while the height adjusts for each batter. In order to produce the strike zone rectangle above, I averaged the strike zone heights of all of the batters within the currently displayed random subset. The coloring of the pitches as red indicates that they were *called* strikes by the umpire, and blue that they were *called* balls. As we can see, there's a decent amount of spill over with the strike zone. While the height spill over is expected (since this is the average strike zone height a pitch far above that was called a strike could have just been pitched at a taller batter), the width spill over is not, since the 17 inch width is constant for all batters. Another trend that is easily observable from this visualization is that an umpire is far more likely to call a ball a strike, than he is to call a strike a ball (though the analysis will prove this more concretely later). In order to highlight which pitches are *called* strikes/balls and which pitches are *actually* strikes/balls, the user can tap the "Show Miscalls" button in the button left corner:



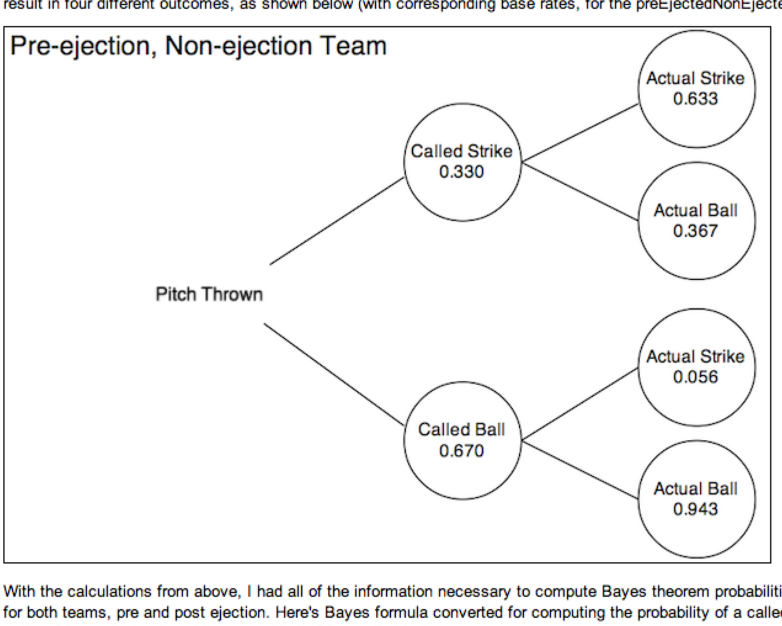

Green pitches are actual balls called as strikes, and yellow pitches are actual strikes called as balls. The overall percentage of miscalls, of either of these two categories, to all calls is shown in the top right and left corners for each team (and it appears for this subset, the ejection-affiliated team has a slightly higher error percentage, perhaps berating the umpire seems slightly more justified now). Our previously observed trend, regarding an umpire being more likely to call a strike erroneously, is further demonstrated by the larger quantity of green pitches relative to the yellow pitches. However, all we have looked at that, thus far, is what's occurred pre ejection, what changes after someone berates the umpire and is subsequently ejected:




Alright, I see some differences (hint, see the top, left and right hand corners), but a dialog box would certainly be helpful to summarize the results. The "Miscall Data" button does just that.


Before examining this information, let's remember that the current statistics are simply for these thousand pitches that were randomly selected. That was for visualization purposes, now, for the data analysis the "Miscall Data From Entire Files" will give us the same statistics computed for every pitch we have data on (not just those that are being displayed):

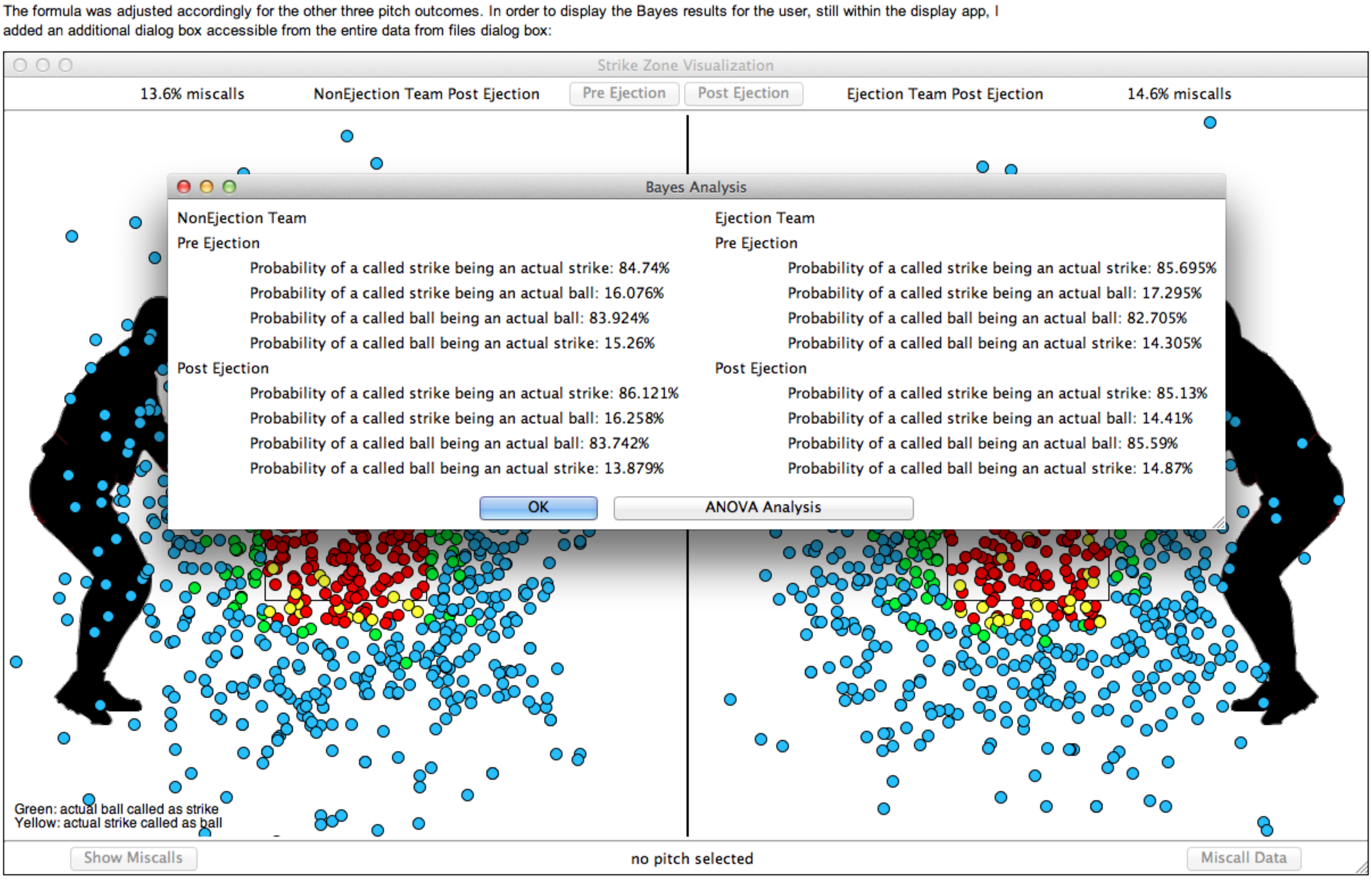

The miscalled strikes/balls percentages are interesting to examine between teams before and after ejections. Before an ejection, the ejection-affiliated team has a higher overall percentage of error, and after berating the umpire and being ejected, their error percentage drops below that of the non-ejection-affiliated team. While these statistics are revealing, the probability of a pitch being called a ball or a strike (and actually being one), given these base rates, has yet to be examined. In order to analyze each team's chances of receiving favorable and unfavorable miscalls, I used Baye's theorem. Each pitch can result in four different outcomes, as shown below (with corresponding base rates, for the preEjectedNonEjectedTeam file):




With the calculations from above, I had all of the information necessary to compute Bayes theorem probabilities for each of these four outcomes of a pitch for both teams, pre and post ejection. Here's Bayes formula converted for computing the probability of a called strike being an actual strike. AS = actual strike, CS = called strike, AB = actual ball, CB = called ball:

$$P(CS|AS) = \frac{P(AS|CS)\,P(CS)}{P(AS)} = \frac{P(AS|CS)\,P(CS)}{P(CS)\,P(CS|AS) + P(CB)\,P(CB|AS)}$$

The formula was adjusted accordingly for the other three pitch outcomes. In order to display the Bayes results for the user, still within the display app, I added an additional dialog box accessible from the entire data from files dialog box:




The probabilities hold potential implications, but I was interested in the unfavorable calls. Specifically, I wanted to investigate the called strikes that were actually balls. The probabilities for the team not associated with the ejection stayed roughly the same (about 16%), while, for the ejection team, before the ejection it was about 17% and after 14%. To examine this relationship further, in regard to the unfavorable calls for both teams, pre and post ejection, I conducted an analysis of variance (ANOVA). I followed the methodologies of:

https://www.khanacademy.org/math/probability/statistics-inferential/anova/v/anova-3--hypothesis-test-with-f-statistic

and

https://people.richland.edu/james/lecture/m170/ch13-1wy.html

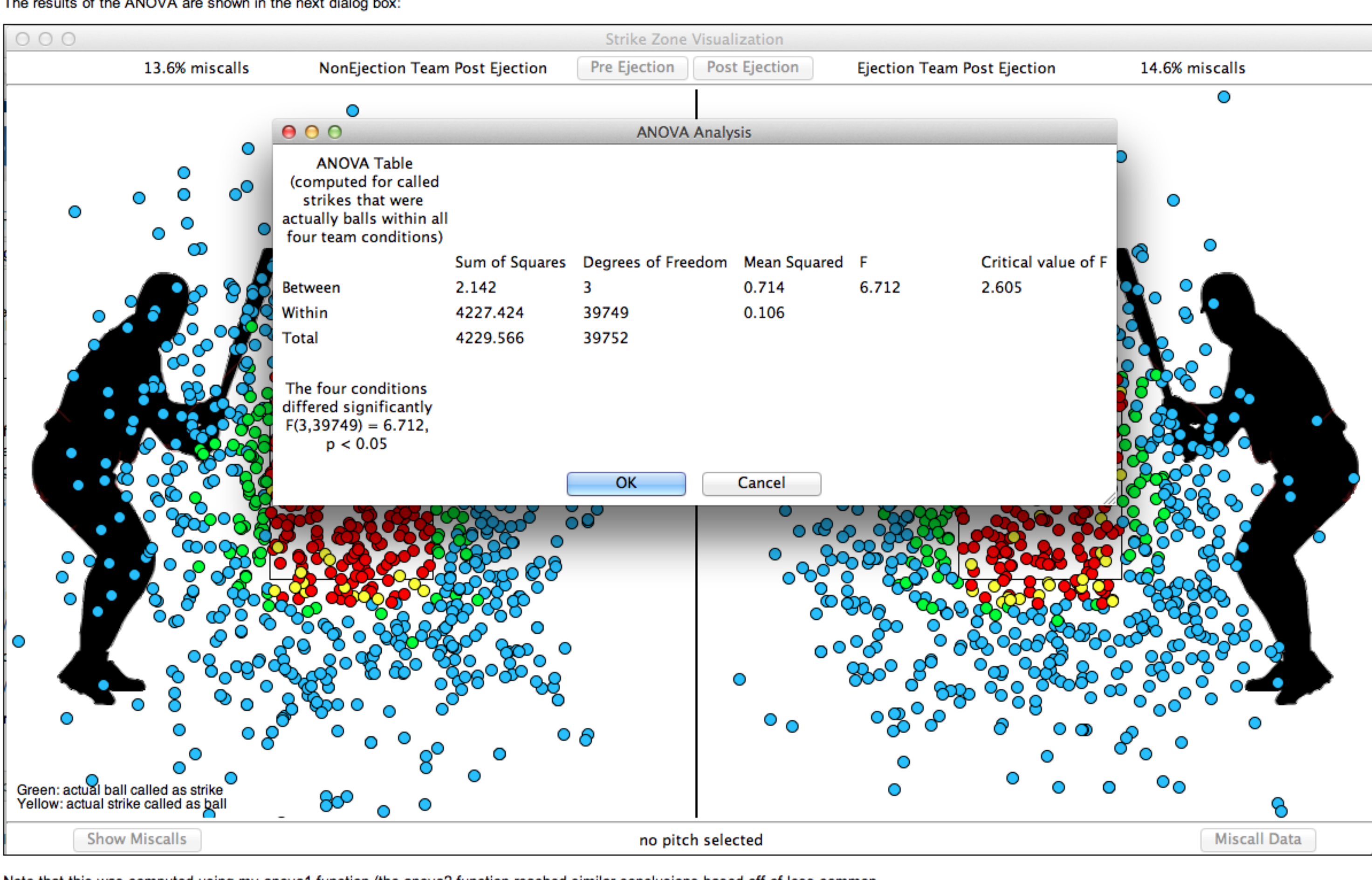
and in order to find my critical value (for what ended up being (3,397499) with an alpha value of .05) I used:

http://scistatcalc.blogspot.com/2013/11/online-critical-f-value-calculator.html

and checked it with

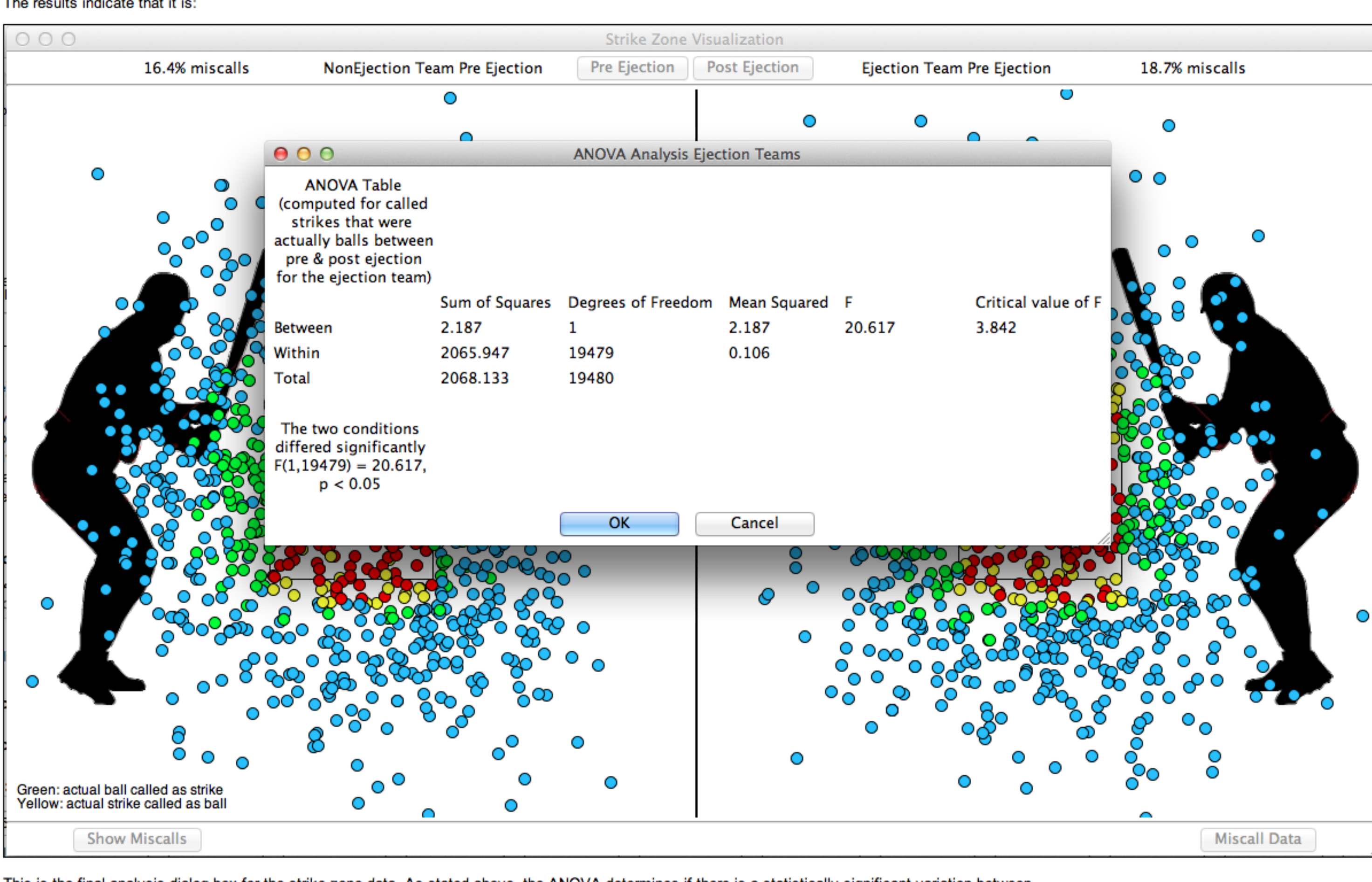
http://epitools.ausvet.com.au/content.php?page=f_dist&t=6.712&df1=3&df2=39749&Conf=0.05

The results of the ANOVA are shown in the next dialog box:




Note that this was computed using my anova1 function (the anova2 function reached similar conclusions based off of less common methodologies/formulas relative to the sources above, but it is still included in my code).

Above, we observe that the F value of 6.712 is in fact greater than the critical value of F of 2.605, meaning that there is a statistically significant variation between the unfavorable miscalls between the two teams, pre/post ejection. However, based off of these findings, we are left with one final question: is berating an umpire and being ejected effective? Presumably, the team affiliated with the pitch-related ejection, received unfavorable calls (or at least calls that they perceived as unfavorable). Following this, a team member berates the umpire about the pitch calls and is subsequently ejected from the game. Did this verbal attack on the umpire actually affect the post-ejection calls? Essentially, what we are asking, is whether or not the original driving motive behind berating an umpire is actually fulfilled after the ejection? Is having a team member ejected in protest of pitch calls an effective tactic for a team? The results indicate that it is:




This is the final analysis dialog box for the strike zone data. As stated above, the ANOVA determines if there is a statistically significant variation between pre and post ejection of the ejection team (with only two conditions this is essentially a t-test). The product of these analyses indicate that there is, in fact, a statistically significant variation of unfavorable calls before and after an ejection, for the ejection-affiliated team.

**Conclusion:**

The formal calculations of Bayes Theorem and the ANOVAs appear to confirm many of the initial suspicions and trends observed earlier with the visualization and initial dialog boxes. The base rates computed for the entire dataset pointed to a possible varying of frequency of miscalls between teams, pre/post ejection. Following this, the use of Bayes formula combined these known base rates to produce probabilities for each of the four outcomes of any pitch thrown for either team. The actual balls called as strikes, or unfavorable miscalls, showed differences in probability between the teams pre/post ejection. This potential trend revealed by the Bayes theorem calculations fueled the final steps of the analysis. In order to determine if the unfavorable calls of the teams actually statistically differed, or if instead this difference in probability wasn't due to the separate conditions the teams were in, an ANOVA of the unfavorable miscalls was conducted for both of the teams pre & post ejection. This ANOVA confirmed that the miscalls of the teams do in fact differ. Specifically, the actual balls called as strikes were shown to have a statistically significant variation between the four conditions. Finally, the conditions of the analysis were narrowed, and the unfavorable miscalls were analyzed pre and post ejection for the ejection-affiliated team. The results showed that there was a statistically significant variation for the team before and after the ejection. Thus, demonstrating that berating and umpire and getting ejected accomplishes its goal of affecting the unfavorable miscalls that the umpire makes.

**Labels:** cs251s14project10

Add Comment