Aaron Rodgers' Free Plays - a Project in Exploratory Data Analysis

Topic: Aaron Rodgers (Super Bowl champion and MVP, 2-time NFL MVP) is considered one of the greatest quarterbacks of all time. Although many of his statistics are remarkable, one of his most unique skills is his ability to provide his offense with opportunities by taking advantage of free plays, which are generated when the defense commits one of a few penalties. If the defense is called for offsides, lining up in the neutral zone, twelve men on the field, or illegal formation, the offense has a 'free play' in which they can run a play and then either decline the penalty and take the yards gained by the play, or accept the yards from the penalty if that provides greater benefit. On such occasions, the offense usually attempts high-risk, high-reward plays because there is no real penalty for mistakes; even if the play results in a turnover, for example, that result will be negated as the offense can simply choose to accept the penalty and replay the down.

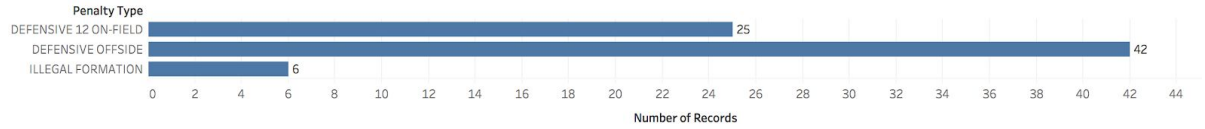
Dataset: I began with a dataset of all NFL plays beginning in September 2013 up to the beginning of October 2017. The data included penalties by name, which is important in identifying free plays; it also had descriptions of every play in a standard format, which I knew I could dissect for certain strings if I needed specific data. It initially had about 40 dimensions and 20,000 rows. I was able to cut it down significantly by removing all plays that did not have the Packers on offense, and all the plays that wouldn't be relevant to my free plays question, such as punts and kickoffs. I also removed all of the columns which did not provide relevant information (about 15 columns in total). I did this with Trifacta Wrangler. Data source here.

Guiding Questions:

1. Does Aaron Rodgers actually draw significantly more free plays than quarterbacks on other NFL teams?
2. Do these free plays actually cause a difference in Packers games?
3. What would the Packers' season look like if Aaron Rodgers drew free plays at the rate of other quarterbacks?

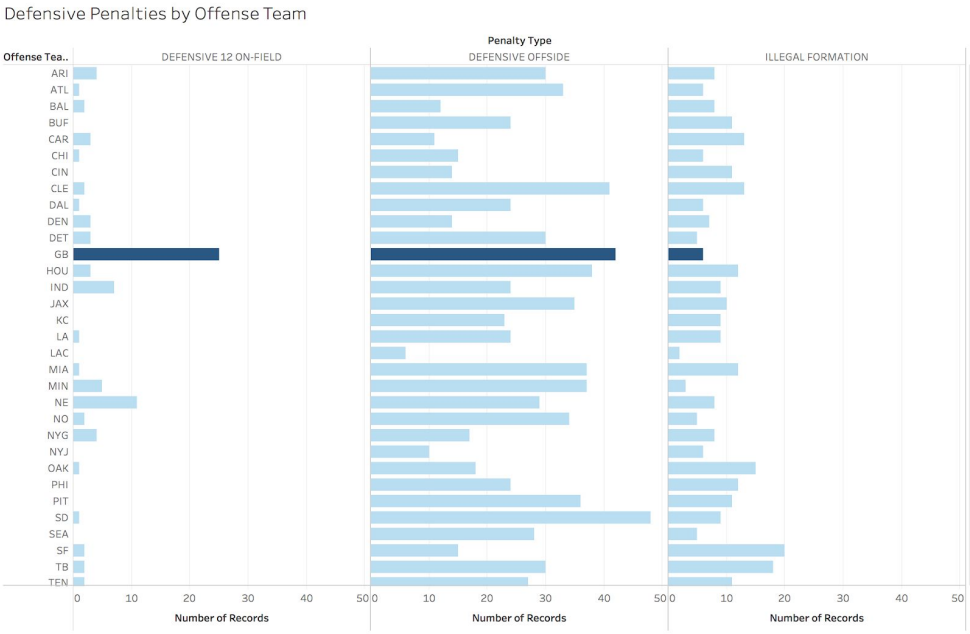
Frequency of Free-Play Inducing Defensive Penalties

First I had to find all of the possible penalties that would have resulted in a free play, and remove the ones that weren't relevant. I had no idea how many of these penalties there would be, or which ones would make the biggest difference. So I constructed this visualization. However, this raises the question of perspective: are these penalties consistent with my question of whether or not the Packers used free plays in a meaningful way? Or, more concretely, are the Packers any different than other teams at drawing such penalties?



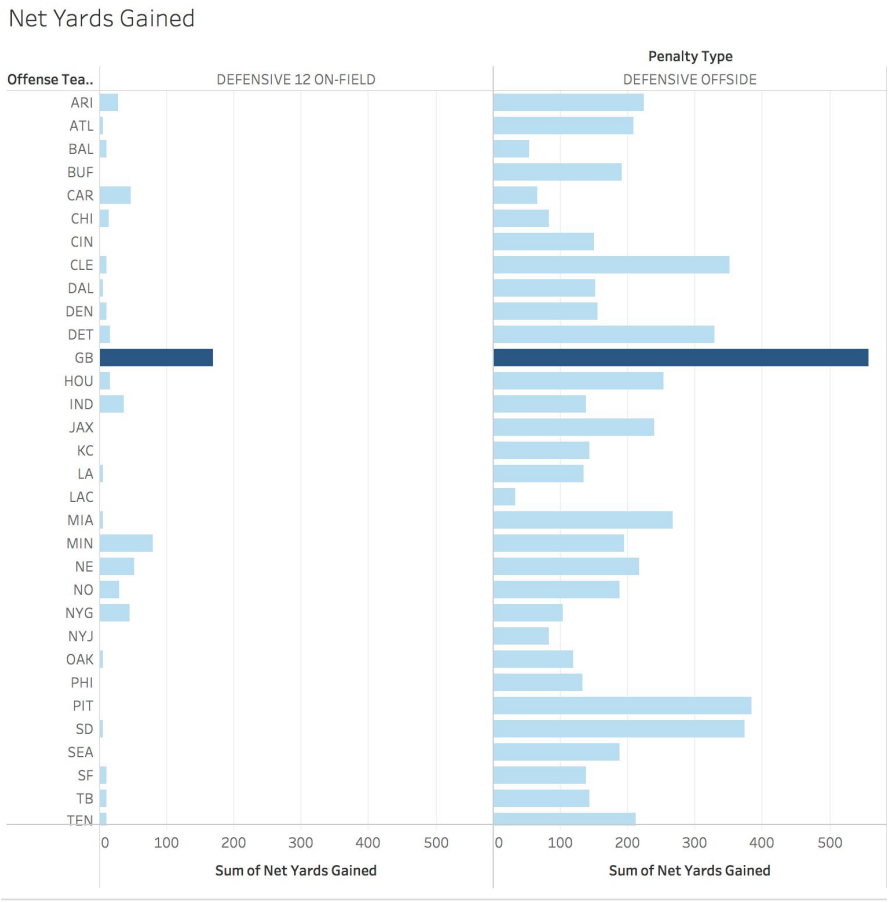
Frequency of Free-Play Inducing Defensive Penalties by Team

Comparing these defensive penalties across all offenses, the Packers clearly draw the most twelve men on the field penalties. Looking closer, you can see they draw the second-most defensive offsides penalties, and slightly below-average illegal formation penalties.



Net Yards Gained from Free-Play Inducing Defensive Penalties

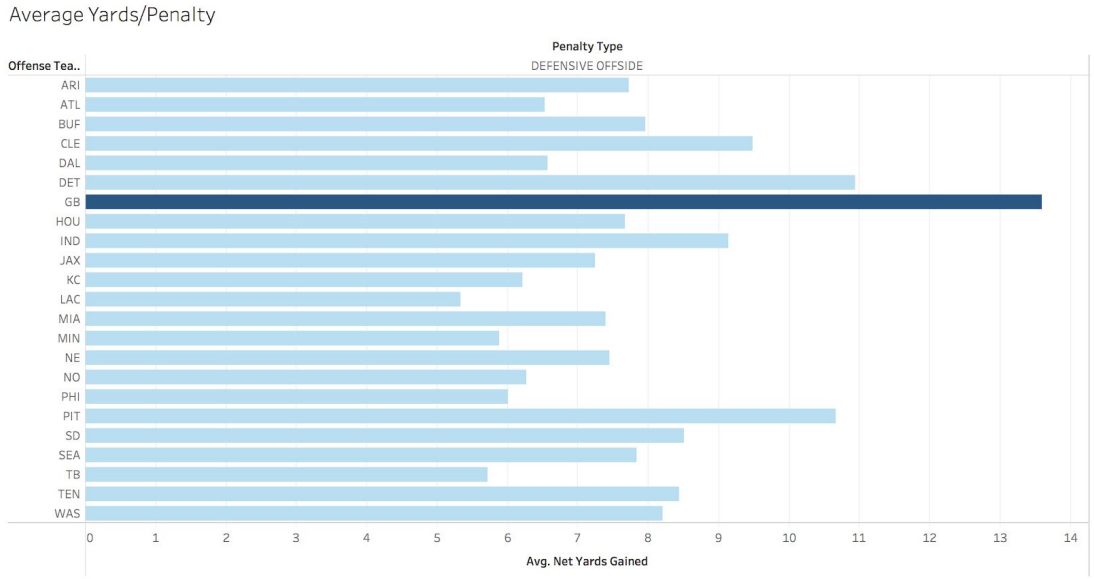
While the number of penalties is one thing, it is important to know the total yards gained off of these penalties to see just how much benefit is derived overall. Instead of using the count of those penalties, I graphed the net yards gained off of the penalties using a simple filter and combining two of the data fields. Here, the difference is even more striking. However, it is not immediately obvious whether this is purely the result of the Packers' advantage in twelve men on the field penalties, or whether it extends to defensive offside and illegal formation penalties. Splitting this by penalty type, it's clear that the Packers have gotten more yards from not only twelve men on the field but also many more yards from defensive offsides than any other team. I removed illegal formation penalties since the benefit was negligible compared to yards gained from the other two penalties.



Average Yards Gained per Game from Defensive Offsides Penalties

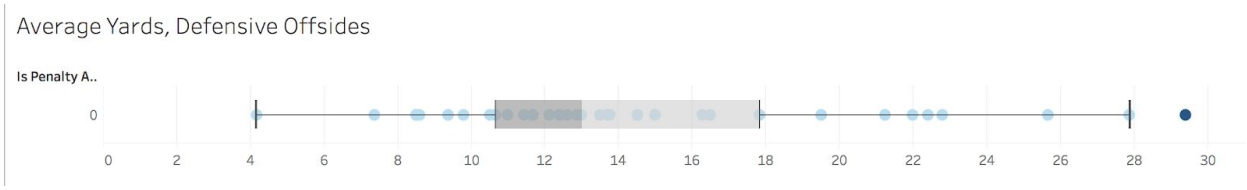
Given that the Packers haven't drawn the most defensive offsides penalties, the fact that they have gained the most total yards from defensive offsides penalties illustrates that they are

averaging more yards per penalty than the other teams. I decided to visualize this effect to provide a better comparison amongst all teams. Graphing all yards per penalty together is misleading, however, since different penalties draw different yardage. Separating the penalties yields a much weaker balance in favor of the Packers; however, this falls prey to the statistics trap in which values with few data points are prone to wild variation. Removing the ten teams with the fewest drawn penalties, the trend becomes more clear.



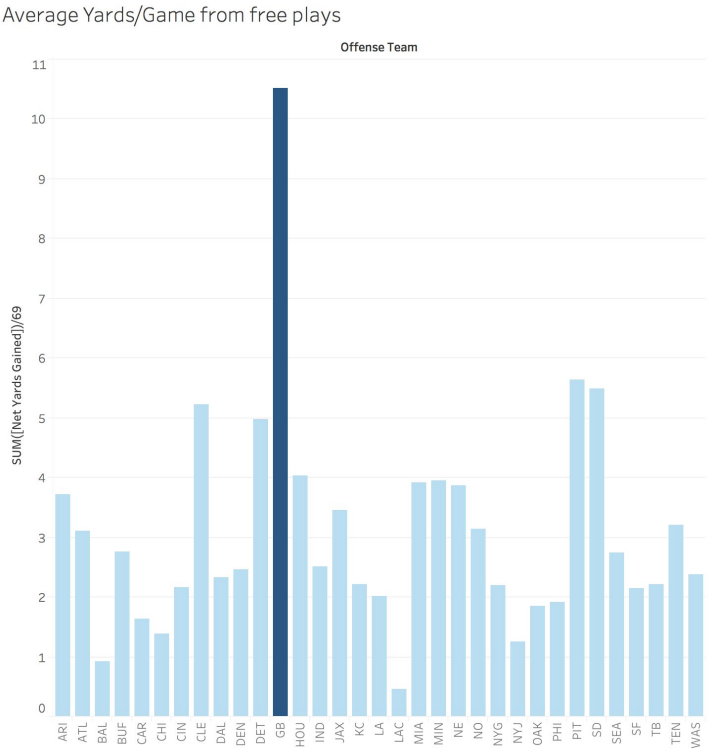
Average Yards Gained per Accepted Free Play

The Packers average about 13.6 yards per defensive offsides penalty (which, when accepted, gives the offense 5 yards). Filtering against the plays where the penalty is accepted, the Packers average 29.4 yards per free play caused by defensive offsides, the highest of any team, as is visible in this box-whisker plot. The Packers are the dark blue dot on the far right.



Average Yards Gained from Free Plays per Game

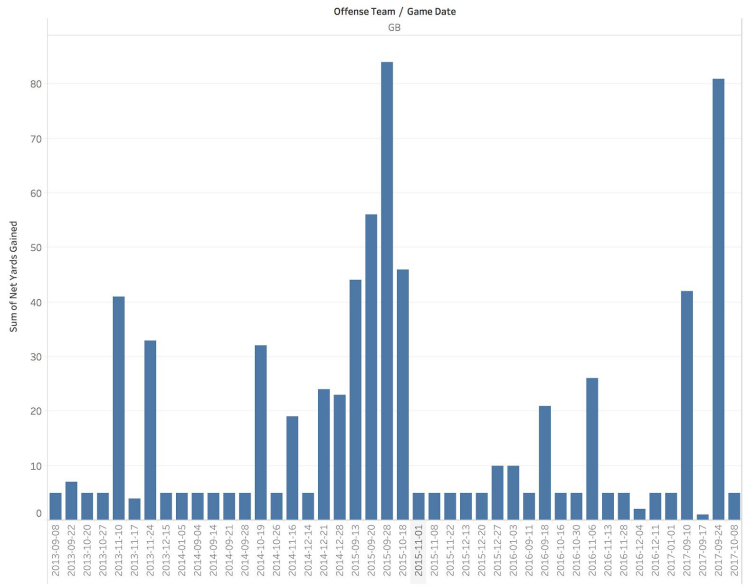
However, all of these findings only show that the Packers gain significantly more from free plays than other teams do. To fully answer the initial question, we must discover whether these gains are significant to the outcomes of games. (Is it possible that despite their comparatively large gains from free plays, free plays still don't make a relevant difference in game outcome?) The dataset contains 16 games per year per team, plus approximately 5 games per team from 2017 (not accounting for a few bye weeks) = 69 total games per team. Taking the total yards from free plays per team and dividing by 69 games gives this distribution of average yards off of free plays per game.



Packers Yards from Free Plays per Game, 2013-17

The above graph shows that even the Packers average only about 10.5 yards per game off of free plays, an insignificant difference in any single game. However, plotting this game-by-game for the Packers shows that these yards are not evenly distributed. In fact, almost half of all games had no free-play yards at all (and thus Tableau doesn't display them in the plot), and another third had only about five to ten yards, the benefit from a single accepted penalty. There were a few games, however, in which the yardage gains were significant.

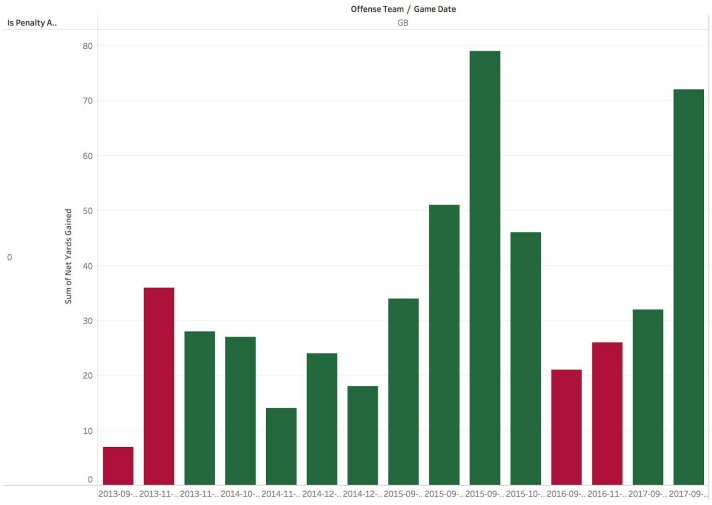
Packers Free Play Yards per Game



Packers Games with Accepted Free-Play Yardage

Looking only at plays where the gains were because of yards gained on the free play, not by accepting the penalty yardage, yields only 15 games. I decided to check whether or not these large gains could have had any impact on win/loss by encoding a Packers win as green and a loss as red. Under these conditions, the Packers were 11-4 (a 73.33 winning percentage). For comparison, the Packers were 44-25 for all games over that period (a 63.77 winning percentage).

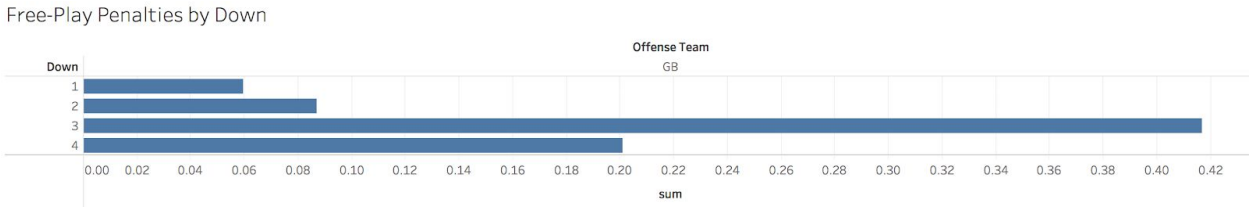
Packers Free Play Yards per Game



———

Packers Free-Play Penalties Drawn by Down

However, ignoring plays where the penalty was accepted does not fully capture the benefits of drawing penalties to cause free plays. Often, Aaron Rodgers uses a hardcount to draw defenders offsides during critical moments in the game, especially on 3rd and 4th down. In these cases, the benefit is not derived from the free play but from the penalty yardage, regardless of whether the penalty is accepted. Controlling for total number of plays on each down, we can see that free play penalties are drawn disproportionately high on 3rd down (and 4th).

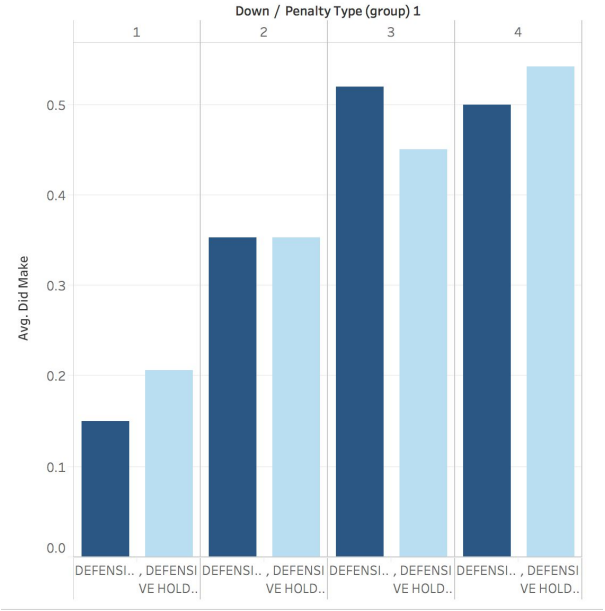


———

Packers Down Success Rate, with and without Free Plays

For every one of these converted 3rd and 4th downs, we can see whether or not they were statistically likely to convert that down without the free play by visualizing the success rate of all Packers attempted 3rd and 4th downs. This chart shows the success rate of the Packers on every down, with and without free plays. Somewhat surprisingly, the success rate for converting is only higher for free plays on 3rd down. Looking into the data shows that the sample size for 4th down is so small that there were only 2 values in the 4th down free play column (one success and one failure), which led to the 0.5 result. However, for the failure, accepting the yards from the free play turned a 4th-and-6 into a 4th-and-1 (thus not encoding as a success), although the Packers converted the following 4th-and-1.

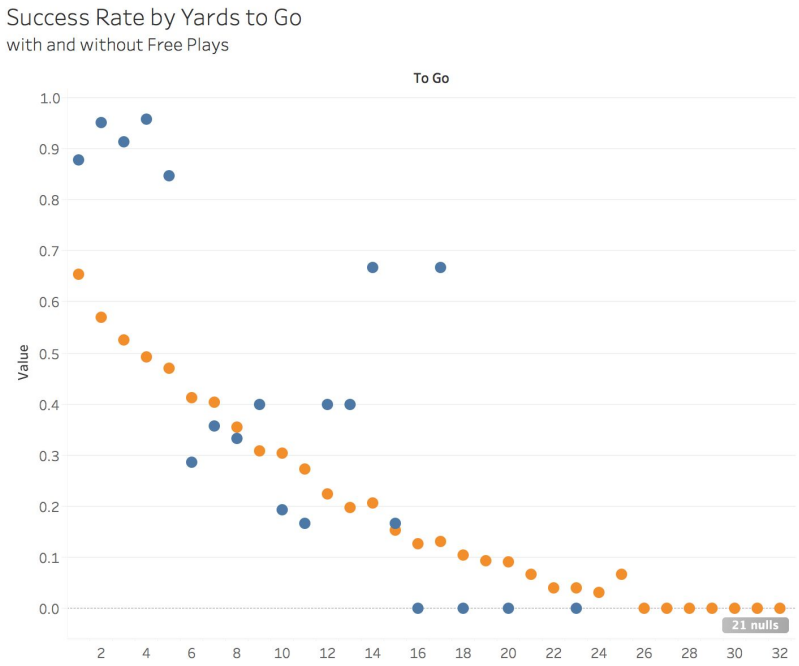
Down Success Rate, with and without Free Plays



Packers Down Success Rate by Yards to Go, with and without Free Plays

The most important case from the above graph was the 3rd down column; it had the most potential to make an impact in game results. Additionally, the lack of benefit of free plays on 1st and 2nd down does not decrease the overall benefit of the free play, because as mentioned above, free plays cannot cause negative effects for the offense. So, I decided to focus in on 3rd-down benefits. On cases without a free play, the Packers were only 45% likely to convert the 3rd down; with a free play, 52% likely. But now, is it possible to predict whether or not the Packers would have made each third down if they hadn't drawn a free play? It's impossible to create a perfect predictor of whether or not the Packers would have converted the down. I first created a plot of down success rate by yards to go, then added on the 3rd-down free play average success rate for each value of yards to go.

Judging from this graph, drawing a free play drastically increases success rate for 5 or fewer yards to go, makes little difference from about 6-15 yards to go, and then jumps around for values greater than that. Since there are very few plays in which there are 3rd downs with both greater than 15 yards to go and a free play (two unlikely occurrences), those values were probably inaccurate measures. In fact, those points lie on fairly common values (0.667, etc.), indicative of a lack of data.



Final Visualization: Projecting Packers Scores without Free Plays

The success rate of converting 3rd downs without free plays is above 50% for 3rd downs of only 3 or fewer yards. We can reconstruct each of these games if every third down converted with a free play was instead converted with the average success rate of non-free play 3rd downs. For the final visualization, I decided to filter only for games with free plays that led to either the penalty being declined (and thus a greater yardage gain), or the penalty occurring on a 3rd or 4th down. I also narrowed my scope to 2015 and 2016 for a few reasons: 2014 had an incomplete data set, Rodgers sat out most of the 2013 season due to injury (and I didn't want a gap in the data), and the 2017 season is still ongoing.

From here, I had to do some data transforming but was able to generate a table which showed, for each score of each game: the time on the clock, the Packers' score, the opponents' score, each free play, the likelihood that that free play caused a scoring difference, the Packers' score if not for the free play (the opponents' score wouldn't have changed under these conditions, so I didn't need to generate a field to update it).

Then, after trying a few different line charts, and adjusting colors, labels, and ranges, I arrived at this design for showing score over time and how it would have been different without free plays. I also added circles where the free plays occurred. Most of them lead to a difference in score, which is visible in the way the lines diverge at those locations.

Actual and Theoretical Outcomes

Packers games with 3rd-down or declined penalty free plays, 2015-2016