

## What Makes an NFL Team Playoff Bound?

By Hitesh Boinpally and Alan Liu

### Research Questions:

1. Which category of scoring (offense, defense, or special teams) did the past five years' playoff teams differentiate the most from the rest of the league?

**Answer:** Playoffs were clearly scored more on offense than non-playoff teams, but were less differentiated on defense and special teams.

2. How did the past five years' playoff teams' quarterback play rank compared to the rest of the league?

**Answer:** Playoff team's quarterbacks were consistently better than non-playoff teams' quarterbacks, with the median playoff quarterback often ranking above 75% of non-playoff quarterbacks for a given year.

3. Did playoff teams from the past five years have any statistics that they were significantly better in than the rest of the league?
  - a. If so, what were these statistics?
  - b. If so, were these statistics more on offense, defense, special teams, or some sort of combination?

**Answer:** Playoff teams were consistently better on offense across the board, particularly in Yards per Play and 1st Downs. Playoff teams also generally performed better on defense, forcing more turnovers and allowing less third downs to be converted. Playoff teams did not differentiate much from non-playoff teams in special teams.

### Motivation and Background:

While this might not be seen as that important in the current pandemic, typically people place great importance on sports. Particularly, coaches, NFL executives, and players are all directly tied to how their team performs every season and whether or not they should be kept by an organization for the next season. Their lives depend on how their team performs and this analysis hopes to see how much they really can control/what things they should prioritize versus whether it really just comes down to whether or not your team was lucky a certain year. Essentially, the goal here is to find out more about the NFL and look to see how teams can improve themselves and make themselves playoff contenders.

### Data Sets:

All datasets that will be used are within the following link:

[FootballCSVs](#)

(All the datasets besides abbreviations come from Pro Football Reference -

<https://www.pro-football-reference.com/>)

(NFL abbreviation data - <https://gist.github.com/cnizzardini/13d0a072adb35a0d5817>)

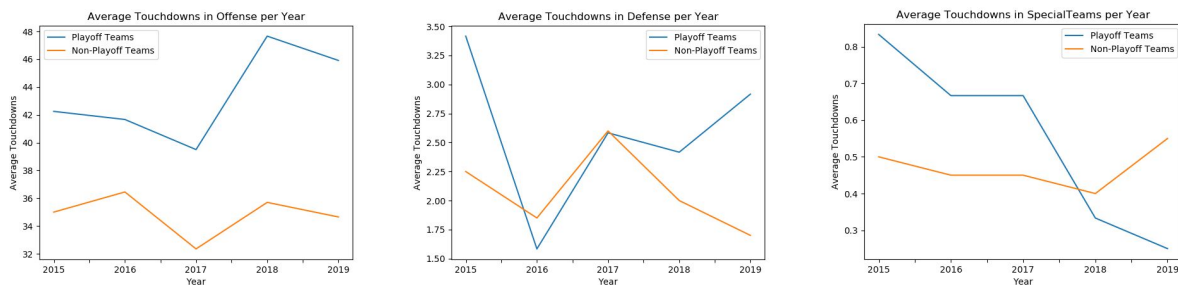
## **Methodology (Algorithm or Analysis):**

- 1. Methodology for Research Q1:** We will repeat the same algorithm for every year and then plot each year's data as a graph over time. In order to calculate the yearly average scoring for each category (offense, defense, or special teams) between playoff teams and non-playoff teams we will do the following. First, we will use the AFC Playoff Standings and NFC Playoff Standings datasets to create a list of playoff teams for that particular year, based on the top six teams from each table. Next, we will split the Scoring Offense datasets up based on whether the team was a playoff team or not from the list we already generated. Next, we will find the average offensive touchdowns scored for playoff teams by using the RshTD and RecTD columns. Then, we will find the average defensive touchdowns scored for playoff teams by using the FblTD and IntTD columns. Lastly, we will find the average special teams touchdowns scored for playoff teams by using the PR TD and KR TD columns. Now that we have the data for playoff teams in each category for that year, we will repeat the process to find the data for non-playoff teams in each category for that year. Once we have the data for both playoff and non-playoff teams in each category for a particular year, we will repeat the entire process for the next year until all the years have had their individual scoring averages calculated. We will then plot this data with three separate lines, one for offensive scoring, one for defensive scoring, and one for special teams scoring with average scores as the y-axis and year as the x-axis, showing us if there are any trends over time or any sort of overall pattern.
- 2. Methodology for Research Q2:** First, we will use the quarterbacks' stats from 2015 and use the same old AFC and NFC Playoff Standings to differentiate quarterbacks in playoff-bound teams from the non-playoff teams. Repeat the same process for all 5 years. Now, merge the quarterback datasets of different years. The result should contain the year of that quarterback's stats, whether the player belongs to a playoff team or not, and his quarterback rating (the column named "Rate"). With this data, we will create a box plot with a year on the x-axis, quarterback rating on the y-axis, and color it based on if the quarterback belongs to a playoff team. Now, we will see if there is any correlation between the quarterback rating and playoff or not in all 5 years.
- 3. Methodology for Research Q3:** We will be looking at a variety of statistics and looking to see if there is any sort of trend for them between playoff and non-playoff teams. Specifically, these statistics are Offensive Yards per Play (Y/P in Team Offense dataset); Offensive 1st Downs (1stD in Team Offense dataset under Tot Yds & TO); Penalty Yards committed by the team (Yds in Team Offense dataset under Penalties); Yards per Kickoff Returns (Y/Rt in Kick & Punt Returns dataset under Kick Returns); Yards per Punt Returns (Y/R in Kick & Punt Returns dataset under Punt Returns); Defensive 3rd Down Percentage (3D% in Conversions Against dataset under Downs); and Turnover Percentage (TO% in Drive Averages Against dataset). With each statistic, we will

average how playoff teams did versus non-playoff teams using the AFC Playoff Standings and NFC Playoff Standings datasets again to separate. Then we will make 7 separate box plots, one for each of these statistics, grouped by their category (offense, defense, and special teams) with years as the x-axis. Specifically, Y/P, 1stD and Yds will all be offensive categories, Y/Rt and Y/R are defensive, and 3D% and TO% are special teams. We can then look at the plots to see whether playoff teams excel at any of the three categories in metrics outside of scoring (differentiating from research question 1).

## Results:

### Results for Research Question 1:



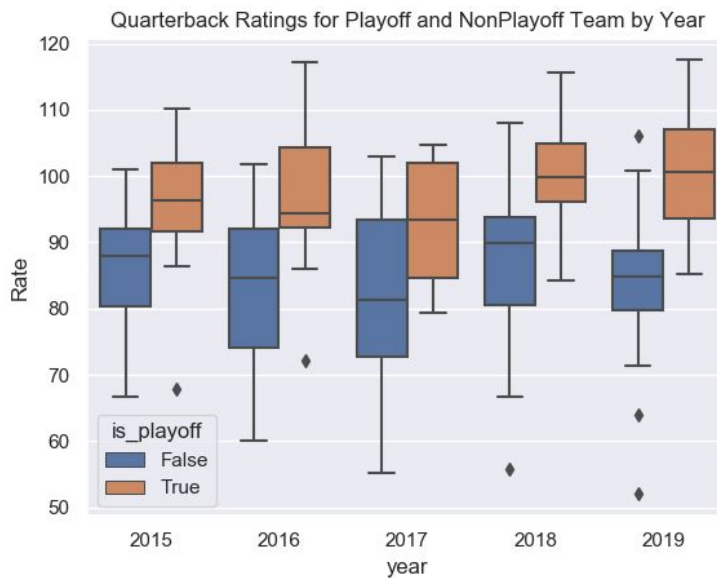
From the three graphs above we can see that there is clear differentiation between average offensive touchdowns between playoff and non-playoff teams. Consistently, the average playoff teams score 8-10 more offensive touchdowns than the average non-playoff team. This is not really too surprising, clearly solidifying the fact that offensive touchdowns are a key aspect of success in the NFL.

However, there is no such clear difference on defense and special teams. For the former, playoff and non-playoff teams differentiated at most by a little over 1 touchdown, with playoff teams actually ranking below non-playoff teams in 2016. Special teams follow a similar pattern, with neither the average playoff nor non-playoff team scoring more than 1 touchdown in the category, as well as playoff teams again ranking below non-playoff teams in 2018 and 2019.

This questions the traditional saying that “Defense wins Championships”, instead showing that offense is far more important and the definitive key factor. The lack of disparity in defensive and special team touchdowns most likely comes from the randomness of the two aspects of the game, with scoring really just being pure luck there. Both require a touchdown to occur within a single play (a blocked punt, an interception, a kickoff, etc), whilst offensive touchdowns allow for a culmination of multiple plays. Because of this singular opportunity, defensive and special team scoring is far more sporadic and randomized, leading to a lack of difference between defensive and special teams scoring.

We can only conclude that playoff teams are much better than non-playoff teams at scoring on offense, and that Defense and Special Teams scoring is far too chance-based to see a difference.

## Results for Research Question 2:



This result is mostly not surprising. But it is interesting that there is one playoff quarterback who is very under-averaged for both 2015 and 2016. It is Peyton Manning in 2015, who is a famous quarterback near the end of his career, so it makes sense that he is not playing well. On the other hand, Brock Osweiler is the one in 2016. He is just an average quarterback, but somehow his team Houston Texans makes the playoff that year.

Other than that, as we have noted in our summary, the medians for playoff quarterbacks are higher than the 75th percentile for non-playoff quarterbacks for a given year. Since there are 12 playoff teams and 20 non-playoff teams each year, this means that the top 8 playoff quarterback ratings are higher than at least 10 non-playoff quarterbacks. But we can see this is pretty reasonable. Let's break down the quarterback rating.

$$\left( \frac{\frac{C}{A} - 0.3}{0.2} + \frac{\frac{Y}{A} - 3}{4} + \frac{\frac{T}{A}}{0.05} + \frac{0.095 - \frac{I}{A}}{0.04} \right) \cdot \frac{100}{6}$$

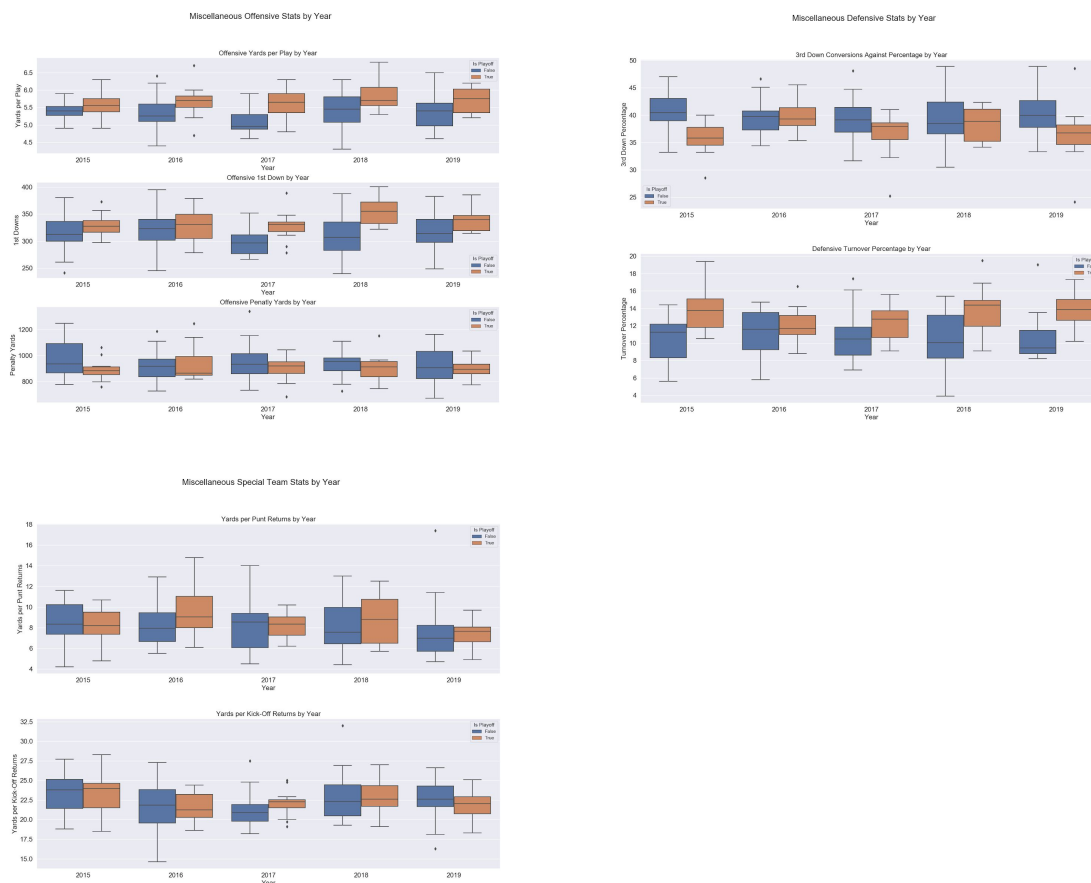
Source: <https://web.archive.org/web/20130918021025/http://www.bluedonut.com/qbrating.htm>

Here, C refers to pass completions, A refers to pass attempts, Y refers to the passing yardage, T refers to the number of touchdown passes, and I refers to interceptions. We know that since

playoff teams have more wins, they generally have more touchdowns and less interceptions than non-playoff teams. Thus, it makes sense for the rating for playoff quarterbacks to be generally higher than non-playoff.

Secondly, we can see that the spread of non-playoff quarterbacks is higher than that of the playoff quarterbacks. Excluding the outliers, the range in non-playoff quarterbacks is higher than that in the playoff quarterbacks. This should also make sense because playoff quarterbacks are usually competitive and thus have fewer variations among them, while non-playoff quarterbacks can range from very good to much below-average.

### Results for Research Question 3:



This data tells a similar story to that of Research Question 1. We can see again that playoff teams beat non-playoff teams in offensive categories across the board, particularly, average yards per play, first downs, and penalty yards. All but one year (2018), the top 50% of playoff teams had greater average yards per play than the bottom 75% of non-playoff teams. The median for playoff teams typically ranged between 5.5 - 6.0 yards per play, while non-playoff teams ranged between 5.0-5.5. 1st downs didn't have as clear of a split, but playoff teams still

typically tended to be higher ranked than non-playoff teams. Interestingly, we can see that some non-playoff teams actually ranked had more first downs than the best playoff team for that year, showing that 1st downs might not have too strong of a correlation to success. Penalty yards were fairly close together, but we can see that the range of penalty yards for playoff teams was far smaller than non-playoff teams, rather all staying much closer to the median.

On defense, we see some trends that show playoff teams being better, but there isn't a clear split. We can generally see that the spread is smaller in playoff teams for third downs converted (where smaller numbers are better), but we also see that non-playoff teams sometimes boast the best team against conversions, rather than playoff teams. However, turnover percentage has a clear split with playoffs teams garnering significantly more turnovers than non-playoff teams. We can see that in four out of five years, the top 50% of playoff teams performed better than 75% of non-playoff teams and in two of those years, the top 25% performed better than all non-playoff teams. This shows a clear connection between turnover percentage and a team's chances to make the playoffs, affirming the concept of "defense wins championships", outside of scoring.

Lastly, special teams seem completely random. The spread is extremely similar between playoff and non-playoff teams, with the medians of each ranking very close together. In fact, most box plots look borderline identical between the two categories for any given year. This is slightly surprising, but also further cements the randomness of special teams, with greater success obviously helping, but not being a key contributor. Something interesting, is that the worst team in these categories for any year was most often a non-playoff team, showing that while being good at special teams may not significantly help you, being bad at it can significantly hinder you.

Overall, we found that offensive capabilities clearly increase a team's chance to reach the playoffs, with defense having a less clear, but still beneficial impact, and special teams being extremely random.

### **Challenge Goals:**

- 1. Multiple Datasets:** We are planning to meet this goal with our project since we are going to have to utilize and merge many datasets together. For example, we will be looking at the Quarterbacks' stats to answer one of our questions, while also using the defensive players' stats to solve another question. This will help us better understand what kind of data we hope to look at and develop a more thorough analysis.
- 2. Many Perspectives:** Our research theme is to analyze in what categories did the playoff teams stand out compared to the non-playoff teams. Within this theme, we cultivate 3 questions. In this first question, we dive into the macro-perspective of scoring by various types of units (offensive, defensive, special teams). In the second question, we look from a micro-perspective by comparing each team's core player - quarterback. In the third question, we back up to the macro-perspective and examine statistics other than scoring.

## Work Plan

- High-Level Task:
  - a. Task 1: Identify Playoff teams from Nonplayoff teams (1hr)
    - Use AFC/NFC Playoff Standing to create a list of playoff and non playoff teams
    - Filter all data of interest from the designated dataset in all research questions for all 5 years

**Evaluation:** Took around 1.5 hours, it was challenging deciding exactly how we wanted to split the data.
  - b. Task 2: Create Summary Statistics (0.5-1hr)
    - Calculate the average of the specified statistics for all 5 years for research question 1 and 3

**Evaluation:** Took around 3 hours, it was hard to navigate the data and parse it correctly
  - c. Task 3: Combine the desired dataframe from each year (1hr)
    - Combine the data from all 5 years into 1 single dataframe for each research question

**Evaluation:** Took about 1 hour
  - d. Task 4: Create a visualization for each (1hr)
    - Create the specific visualization for each research question

**Evaluation:** Took about 1 hour
  - e. Task 5: Write out report (2-3 hr)
    - We will take turns working on this part, in a similar operation to pair programming but instead writing out the report

## Testing:

We will test our code by 1) having each other look over it to check for errors as we go, 2) print out results as we go to make sure the numbers are in the proper ballpark, and 3) use smaller test sets to see if the resulting plots match against manually created plots.

Particularly, we used the smaller “test” data sets, which for question 2 and 3 only consisted of one year and for all questions consisted of only 4 teams, 2 of each of playoff and non-playoff. These datasets were small enough where we could easily calculate the way the graphs should look manually, allowing us to make sure that our code ran as we expected. It also tested to make sure that the program was splitting the playoff and non-playoff teams properly.

## Collaboration Strategy:

- Pair programming
  - a. We alternate between tasks, so each person is responsible for 2 tasks total.

- b. During the process, the person coding will share his screen on Zoom, and we will discuss and write codes together
- Code management
  - a. We will create a repository for this project on CSE GitLab to manage the versions of code and resolve merge conflicts
- We did not work with anyone else
- We used some online resources to solve simple code questions. These were typically documentation, such as the pandas doc and seaborn doc. We also looked at stack overflow for some help.