

NFL Fourth Down Decision Factors

Cody Mitchell

Colorado State University Global

MIS 581: Capstone - Business Intelligence and Data Analytics

Dr. Kimberly Ford

July 4th, 2021

Abstract

In recent years, the National Football League has begun shifting culture towards data influencing the decision-making process across different levels of the company. However, this has not been a smooth process, as some coaches are reluctant to use data driven decisions to make play calling adjustments. This paper analyzes fourth down decision making in the NFL between winning and losing coaching groups to allow for peer comparison in decision factors for game time decisions. This allows a comparative analysis for losing coaches to adjust their game based on the most successful coaches in the league.

To compare decision factors between coaching groups, a regression model was conducted first on all of the coaches to see what the general trend was in the NF from available play by play information from the 2015-2020 seasons. Then the method was applied to the winning and losing coaching groups. The resulting significant factors were then compared to see what the differences were between groups. The suggest losing and winning coaches do have different decision-making factors with their decision to attempt a conversion or kick on fourth down.

NFL Fourth Down Decision Factors

The National Football League, or NFL, is an American football federation in the United States that most everyone in the country has at least heard of or watched a game at some point. Recently, teams in the NFL have taken it upon themselves to implement a larger amount of data collection in their day-to-day operations. This can look similar to the decisions that other companies across the world are using data for such as setting the team ticket prices, fan giveaway prizes, or other marketing goals.

However, the NFL has also been moving to incorporate the data process during the games that are being played. Since the 2017 season, the NFL partnered up with AWS to help provide more live analytics function on players using RFID chips placed on all players and balls. (Grandinetti, 2019) This has helped leaders see where each of their players are at during the game, how fast they are moving and other movement patterns. Having this information has assisted in the process of making game time decisions.

These game-oriented decisions can focus on what players should be in the game in different situations, expected timelines for players to return from injuries, nutrition and exercise programs, performance, directing what plays the team should run, and even what plays their opponent is likely to try. However, these come with their own challenge with how fast the results must be generated during the game to be useful. Since the NFL play clock has a maximum time of 40 seconds, predictions must be generated in seconds, to be useful for the coaches making final decisions and relaying the decision in time. To help with this time issue, data can be used ahead of time to help tell the coaches of any trends they have which they may not be aware of.

Even though there have been some coaches leaping in to utilize data driven decision making for game time decisions, there are still some coaches hesitant to utilize data into their game time decisions. One area that many coaches have been more welcome to insight is whether it is worth it to attempt to go for it on fourth down. Other studies have been conducted to compare the coach's fourth down decision making compared to an optimal strategy already, however these have not explored the factors that different groups of coaches consider when deciding whether they should go for it (Yam & Lopez, 2019). These differences between winning and losing coaches could set forward a learning experience for coaches to see how they compare. Rather than only pointing to the correct decision making, this focuses on how they differ from other actual coaches, making the transition to data more welcomed.

Objectives

The objectives for this study are to analyze the different factors that go into coaching decisions on fourth down of NFL games between winning and losing coaches. To first start this, an understanding will need to be done to see how often fourth downs come up and if one of the groups of coaches is more likely to attempt a conversion than the other. After this, the factors of the coaching decisions for all coaches will be examined to see if a model is reasonable to make for each of the two groups. If this is successful, a model will be created to examine the different factors correlating with the coaches' decision to attempt a conversion on fourth down and see the significant factors. These factors can then be used to compare coaches to each other with the potential of using them as an educational tool for coaches to adapt their playstyle more like the winning coaches' decisions. Having an education focused route for implementing data decisions into the NFL focusing on fellow coaching decision can also help welcome the use of analytics to more hesitant coaching staffs.

Overview of Study

The NFL was founded in 1920 and has seen large amounts of evolution as technology and players have changed over the years. Currently having a revenue of over \$12 billion and nearly 4000 employees, excluding team personnel, they have been successful in retaining relevancy through American culture and are continuing to grow. There are currently 32 teams across the United States and on some occasions those teams will travel throughout the world for games. Each team has their own owner but is required to follow a set of rules set in place by the league. In most situations, once a team has started to successfully implement something new, other teams will follow suit to not be at a competitive disadvantage.

Fourth down decision making is an area of hot topic as more analytics are being used across the league. Most studies have centered around comparing the efficiencies of coaching decisions to that of optimal strategy through models, but not what correlates with the decisions the coaches are making throughout the game. This study seeks to have a greater understanding of the different factors that correlate with a coach's decision to attempt to convert on fourth down, and how these decisions may differ between coaching groups.

Looking at the different factors around coaching decisions will help give more of an insight to what the coaches are considering at the moment of these game changing decisions. Since the coach will ultimately have to make their decision in a very short amount of time based on all the factors of the game, having a quick look into their tendencies can be both beneficial to them and their opponents.

For the coach, they will be able to see what the winning coaches in the league and considering when deciding what to do on fourth down. This can then be compared to the

strategies they consider when deciding, or even compared to an analysis on a coach-by-coach basis to see how the weighted factors are different between them and the rest of the group. For the opposition, they will be able to have more of a predictive power if the opposing coach will attempt a conversion and have plans in place before it even gets to fourth down. This could be accomplished through having a greater understanding of the factors different coaching groups consider when making fourth down decisions, then adapted to the needs and desires of each organization.

Research Hypothesis

There are several different approaches this project will take in analyzing the tendencies that coaches have in their fourth down decision making. The proposed business questions and hypothesis statements are designed to be used in combination with each other to paint the full picture of the analytics project. The questions focus on framing the decision-making process for the coaching groups and how they differ from each other.

RQ 1: Do winning and losing coaches have different rates in going for it on fourth down in the NFL?

H0 1: The NFL coaches all have the same go rate for fourth down in the NFL.

Ha 1: The winning and losing coaches have different rates of going for it on fourth down in the NFL.

The goal of this question is to see the differences between the two groups of coaches. Either result of the research question will result in interesting results from the following questions but is an important step to understand. If the coaches are found to attempt to convert the same amount of time, it furthers the interest in the different factors that each group considers

when attempting to convert to look for differences. If the two groups have different rates of attempting on fourth down, it will be interesting to see how the groups differ that causes this discrepancy.

RQ 2: Can variables related to field position, score differential, game performance previous to the current play, and winning probabilities be used as predictors for coaches going for it on fourth down or not?

H0 2: None of the variables can be used in accurately predicting if a coach will be going for it on fourth down.

HA 2: There is at least one variable that can be used to predict if a coach will go for it on fourth down.

This research question will provide more of a direction for the final goal of the project and show that is it possible at all to model the factors correlating to a coach's decision to convert a fourth down. Is the null is rejected, insight will be gained for some of the factors that apply to both coaching groups and compared to the groups in the next question. The selected variable groups add up to approximately 30 of the starting 350 attributes and were selected based on background information on the topic and what could reasonably be known by the coach at the time they make the decision. The selected groups of variables give a good image of the situation the team is in when making their decision and should be able to be used to make a reliable model. If the null for this question is rejected, the project will proceed to the final question.

RQ 3: Are there different predictors between winning coaches and losing coaches on what they decide to do on fourth down?

H0 3: There are no differences between winning and losing coaches' predictors for fourth down decision making.

HA 3: There is at least one different predictor between winning and losing coaches' decision making on fourth down.

This question is the heart of the project. The results of the previous two research questions will provide the context of this framework and how the two groups of coaches will be compared to each other. This question will use the same group of variables that were used in the second research question to ensure that the models are comparable, and only relevant variables are used for the model. For this, coaches will be determined as winning if they have a lifetime record of 0.500 or above and losing if their lifetime record is below 0.500. Limiting variables will help limit the false relationships that could come from one coach seemingly only going for it on fourth down when indoors, while the reality they faced weaker teams indoors, or other similar explanations. Conducting this data trimming will also allow for quicker computations on the data as the amount to process should be decreased significantly when filtering is used. The specific for analysis on this will be outlined shortly.

Literature Review

With the high level of player talent and coaching in the NFL, many games will come down to the decisions the play callers are making throughout the game. One of these opportunities for play callers to stand out from their peers is when to convert or not on fourth down. Merely having efficient coaching can be enough to contribute to an additional three or four wins for a team in a season (Hadley et al., 2000). In a league where eleven wins will all but guarantee a playoff berth, coaches should do all they can to maximize their decision making, and

their wins. Through looking at how the coach compares to the top percent of coaches' decision making on fourth down, these coaches will be able to improve upon their calling ability, leading to more wins.

Each coach will have their own philosophies on what risk they want to take on when debating whether they should attempt a conversion. Comparing coach's decisions to that of an algorithm has shown that the coaches are trying to be more conservative in their actions (Romer, 2006). Knowing that most of the NFL coaches are likely to be conservative in their play calling in these situations, it can bring more weight into when they end up deciding it is an appropriate time to convert the down. This conservative play calling will be able to be revealed through an analysis of the factors that go into what the coaches are considering when going for it on fourth down. Looking at the overall conversion attempt rates for both coaching groups will help reveal if winning or losing coaches are more likely to be conservative in their play calling nature. Additionally, through looking at the factors going into the decisions for each of the groups, more insight can be gathered into the decision-making process for the groups. If there are substantial differences in the groups, it could show coaches are weighting the factors differently, leading to more conservative calling for one of the groups.

Coaches in the NFL are the top play callers in the country, but this does not mean that they are optimally making these decisions. Since coaches are not following a necessarily optimal strategy, it becomes more important to see how these factors align for different coaching groups (Kovash & Levitt, 2009). These factors will ultimately make up each of the NFL coaches' unique styles, and some may not be consciously aware of the factors they appear to be considering when deciding to convert the down. Looking at different coaching groups, steps will be taken to see if winning and losing coaches are factoring in the same information when

deciding whether they convert. Accepting some level of data into decision making could allow for the coaches to make more optimal decision making, thus improving their team's success.

Regarding when to convert or not, there are some models that have used simple rules for making the decisions. When looking at models to go for it every time it is fourth and less than five, the models found that teams had more positive results compared current coaching decisions (Palmquist et al., 2020). Although the models used for these simulations were not factoring in many different aspects that can be found in a real game scenario, it can still be an interesting place for coaches to grow from. If coaches were able to increase their aggressiveness or see on paper what factors they are using when deciding to go for it, they could improve their calling.

Overall, it will not be enough to build models for the coaches to help with play calling. The coaches have had data analytics tools available to them for many seasons but have not been changing their overall habits. Looking at data since 2014, coaches have been consistently going for it on fourth down between 12 percent and 17 percent with little evidence of change over time (Yam & Lopez, 2019). Since analytics began to be more heavily involved in the NFL in the past several seasons, one could expect the coaches to have some different habits on such an influential part of the game. Because the coaches have not changed in their play calling on fourth down, having coaches look at their own fourth down tendencies rather than optimal strategy could help encourage them to make adaptations and see more success.

Research Design

Methodology

This project focuses on predicting the differences between coaching decisions on fourth down between winning and losing coaches. The overall methodological approach for this project

will be through using quantitative data collected from NFL plays using R. The data used for this project was collected from web scraping on Pro Football Reference records. This website maintains records on every NFL game available including information such as game stats, player stats, records, and more.

The data included for the purpose of this project includes approximately 340 variables of information gathered from every NFL play from the 2015 season through 2020 season. This timeframe was chosen to use relevant data as coaching tendencies will naturally change over time, and this range allows for enough observations to analyze. Since fourth down will only occur on less than one fourth of all NFL plays, and coaches attempting a conversion accounts for an even smaller subsection of data, utilizing several years of data will help make sure there are enough conversion attempts to make an accurate model. The other piece of data utilized for this project is a set also from Pro Football Reference is coaches overall win loss records since their time in the league began.

The data used for this project was stored on a personal computer protected through username and password. The dataset for winning and losing coaches was stored on the computer and was already deemed acceptable for use through the NFL, so no confidential or personal information was included. The play-by-play dataset was never stored on the computer, only used through the RStudio interface. This information was queried each time the entire script was ran, ensuring up to date play by play information can be obtained if done during the season. This was obtained through the NFLFastR library, again without personal information or other security concerns.

Methods

The first step in preparing the data is to clean it for use. This process will be trimming the total set from all plays in the seasons down to only the plays that fit the criteria needed for the study. This includes limiting the plays to only those that happened on fourth down, as well as plays where no penalty or other non-important result came of it. From this subset, background knowledge was used to eliminate columns not relevant to the study, or other information that would only mislead the model in looking at what correlates with coaching decisions. The final variables included dropped to approximately 40 variables consisting of groups such as play identification, location on field, score, team performance to that point, and time left in the game. The dataset used for prediction consists of over 23,000 entries over 44 variables.

The first test to be used for the project is a chi test comparing how often coaches in each of the groups are deciding to attempt a conversion or not on fourth down. Datasets for this test are created to help make sure the test is as straight forward as possible and will use a confidence level of 95%. The results of this test will be used to gain insight if winning and losing coaches have different proportions of attempted fourth down conversions.

The next research questions will be answered through looking at significant variables in a regression model. The first regression model will be created to look at the predictors in all the coaches in the NFL. The minimum confidence used to evaluate significant variables is 90% to help have a wide range of variables to evaluate, and since the model will not actually be used to predict more flexibility can be given on accepted variables. This model will use the same dataset mentioned above.

The last question will use also be conducted using a regression model to observe the correlations of factors between the coaching groups. The groups will be separated using an already included attribute and each get a model built. The resulting models will also use a 90% confidence level. The resulting models will be compared with each other based on which variables were included for predicting, their alpha levels, and their magnitude of correlation. This will allow each of the variables to be included to give an idea of the weighted factors involved in each of the coaching groups' decisions.

Limitations

This study will only be able to look at the data from the provided source, and without any comparison to the aspects a coach ranks as important to them. Without having available survey data on what each of the coaches involved considers important in their decisions, the study will only be able to analyze correlations with the decision-making process. Additionally, the losing coach sample is smaller than the sample of winning coaches, although both have substantial play counts. This makes sense in the scope of the study from the high turnover of head coaches in the NFL and a team will likely not have a coach long term who is not generating wins for the program.

Another limitation of the study will be factoring in a coach's performance on a season-by-season basis. Since coaches may be winning in one season and losing the next, their lifetime win loss record is used to categorize them. This will help solve some issues, but if a coach were to make large changes in their decision making from one season to the next and improve their record, the data will be unable to show it. However, with a large sample size of plays and coaches this should not impact the study largely.

Ethical Considerations

This project is conducted through using a publicly available dataset provided from NFL sources. The NFL has been involved in hosting competitions for data analytics and will often provide datasets to be used for analytics research. Because of this, there are no issues in using the dataset for research. Additionally, coach names are used throughout the research, but again all within the realm of publicly available information, and the only personal data associated with this is each coaches' win loss record. However, this is not a personal piece of information that would be of concern in some other types of data sets.

Findings

Data Description

As stated in the methods for the project, only fourth down plays were used for the purpose of building the models and comparing strategy between coaching groups. After trimming down to fourth down plays and variables that were deemed appropriate for analysis, there were 23,018 plays with 29 available attributes to work with for creating the models. In this group, any play that resulted in a penalty or otherwise did not have a play ran was also excluded from analysis. Because this is a relatively small set, all data was included for the analysis portion of the project. The frequencies that different coaching groups attempted a conversion are shown in Figure 1. Coaches were considered winning if their lifetime win loss record was greater than 0.500 and winning if less than or equal to 0.500. The attributes considered for the final models can be found in Figure 2.

Figure 1

How often the coaching groups attempted a conversion on fourth down.

Coaching Decision by Group			
Decision	Losing Coaches	Winning Coaches	All Coaches
Kick	6981	12653	19634
Attempt	1203	2181	3384

Figure 2

Attributes used for the prediction table.

```
[1] "season_type"      "week"
[3] "yardline_100"    "half_seconds_remaining"
[5] "game_seconds_remaining" "game_half"
[7] "qtr"             "goal_to_go"
[9] "ydstogo"         "ydsnet"
[11] "timeout"         "posteam_timeouts_remaining"
[13] "defteam_timeouts_remaining" "score_differential"
[15] "no_score_prob"   "opp_fg_prob"
[17] "opp_safety_prob" "opp_td_prob"
[19] "fg_prob"         "safety_prob"
[21] "td_prob"         "total_home_rush_epa"
[23] "total_away_rush_epa" "total_home_pass_epa"
[25] "total_away_pass_epa" "wp"
[27] "def_wp"          "div_game"
[29] "fourth_attempt"
```

Results

The first test conducted was a chi-squared test on the proportion that the two coaching groups attempted a conversion on fourth down. The data used for this test is similar to that shown in Figure 1. The output for this test is found in Figure 3 below. The p-value is a way to see the probability of obtaining the observed result if the null hypothesis is assumed to be correct. As shown in the test results window, the p-value for the test is 1, meaning the null

hypothesis is failed to be rejected since there is a 100% chance the result could be seen with the null hypothesis being correct. This suggests that the two coaching groups does not have a significant difference in the proportion of attempts versus kicks on fourth down.

Figure 3

Chi-squared test to see if coaching decisions differ.

```
Pearson's Chi-squared test with Yates' continuity correction
data: winning.v.Losing
X-squared = 3.8519e-28, df = 1, p-value = 1
```

When thinking of the structure of an NFL game, this suggestion makes some sense. For most of the fourth down opportunities in a game, they will be happening in relatively calm or seemingly uneventful times of the game. Since NFL coaches have been found to be risk adverse when making these decisions, it should not be surprising that both groups have similar numbers (Hadley et al., 2000). A more aggressive strategy where coaches would go for it in the first and third quarter, or with ample time left in the game could allow for a bigger separation, but this is not the reality of how the game is played today. Them having similar attempted conversion rates can draw some more interest into the factors coaches appear to be considering in their actions though. Since they go for it the same proportion of the time, what are the differentiating factors between the two groups?

The next research question to investigate is if a model could be built looking into the factors causing all NFL coaches to either attempt a conversion or kick on fourth down using the variables listed above in Figure 2. For this test, a logistic model was generated using 0 for kicking the ball and 1 for attempting a conversion. The performance of the model generated in

this step is not important to the scope of the study, only the variables that have a significant contribution to the model. The output of the model is found in Figure 4.

Figure 4

Logistic model results for all NFL coaching fourth down decisions since the 2015 season.

Coefficients: (4 not defined because of singularities)					
	Estimate	Std. Error	z	value	Pr(> z)
(Intercept)	1.505e+06	1.200e+06	1.254	0.20985	
season_typeREG	3.013e-01	1.674e-01	1.800	0.07186	.
week	2.809e-02	6.572e-03	4.275	1.91e-05	***
yardline_100	-2.760e-03	5.874e-03	-0.470	0.63845	
half_seconds_remaining	-3.287e-04	2.425e-04	-1.356	0.17525	
game_seconds_remaining	-2.993e-04	3.093e-04	-0.968	0.33322	
game_halfhalf2	1.094e+00	3.788e-01	2.888	0.00388	**
game_halfovertime	NA	NA	NA	NA	
qtr	4.953e-02	1.400e-01	0.354	0.72355	
goal_to_go	-4.980e-01	1.611e-01	-3.090	0.00200	**
ydstogo	-5.325e-02	7.115e-03	-7.485	7.18e-14	***
ydsnet	5.529e-02	2.083e-03	26.548	< 2e-16	***
timeout	2.460e+00	4.232e-01	5.812	6.18e-09	***
posteam_timeouts_remaining	-3.067e-01	4.465e-02	-6.868	6.51e-12	***
defteam_timeouts_remaining	9.184e-02	4.443e-02	2.067	0.03873	*
score_differential	7.674e-03	6.496e-03	1.181	0.23745	
no_score_prob	-1.505e+06	1.200e+06	-1.254	0.20985	
opp_fg_prob	-1.505e+06	1.200e+06	-1.254	0.20985	
opp_safety_prob	-1.505e+06	1.200e+06	-1.254	0.20987	
opp_td_prob	-1.505e+06	1.200e+06	-1.254	0.20985	
fg_prob	-1.505e+06	1.200e+06	-1.254	0.20985	
safety_prob	-1.505e+06	1.200e+06	-1.254	0.20979	
td_prob	-1.505e+06	1.200e+06	-1.254	0.20986	
total_home_rush_epa	3.764e-03	5.540e-03	0.680	0.49681	
total_away_rush_epa	NA	NA	NA	NA	
total_home_pass_epa	2.472e-03	2.548e-03	0.970	0.33187	
total_away_pass_epa	NA	NA	NA	NA	
wp	-3.811e+00	2.460e-01	-15.491	< 2e-16	***
def_wp	NA	NA	NA	NA	
div_game	8.835e-02	6.638e-02	1.331	0.18320	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

The resulting model has ten different attributes that appear to be significant to an alpha of 0.10. This value means that there is a 10% chance or less that the observed correlations were significant in the coaching decisions result. The variables that are significant in the model include: if it is a regular season game, the week in the season, if the play is in the second half, if they are fourth and goal, the yards remaining, the net yards of the drive, if there was a timeout, how many timeouts the possession team has, how many timeouts the defensive team has, and the winning probability of the possession team. Since there were variables able to be used to predict if a coach was going to attempt a conversion, the null hypothesis is rejected for the second

research question, suggesting there are attributes correlating with a coaches fourth down decision.

This model suggests the all coaches group is more likely to attempt a conversion during the regular season, a situation where in the long run is fairly low risk if losing. This could also be due to the difference in number of games between regular and post season. Each year there were a total of 272 regular season games and 11 post season games. Contrary to this is that games later in the season have a higher likelihood of coaches attempting a conversion. This means that as the season advances, including post season, coaches are more likely to attempt a fourth down conversion. As the season continues, teams may be more desperate to win games, thus causing coaches to take the risk and go for it more often.

The game being in the second half was also found to be a significant factor. In close games, taking a risk in the second half of a game is one of the last chances a team may have to win the game. This is further supported from teams going for it more often when not currently favored to win the game given by wp. Teams losing late in the game and not being favorable to win could make the decision easy for coaches to attempt a fourth down conversion with not much to lose in a failure. In this case, it is also interesting that fourth and goal situations make the coach less likely to go for it. They seem to be more interested in the more probable field goal conversion rather than the possibility of a touchdown.

The last variables have to do with how the game has been going and what the coach expects of the other team. Teams having a longer drive are more expected to go for it, indicating the coach may believe they have the momentum and wanting to continue to build. Additionally, the all coach group model suggests that if a timeout is taken the team is more likely to attempt a conversion. Teams were also less likely to convert when having more timeouts, and more likely

to convert when their opponent had more timeouts remaining. This could be from a game control standpoint, where the defensive team having more timeouts remaining could improve their clock management potential in the rebuttal drive, leading to a higher chance of points being scored.

The final research question in this project is to see if winning and losing coaches have different predictors associated with their decision to attempt a conversion on fourth down. Both coaching groups had models generated for them similar to the model created for research question two. The created model for the winning coaches is shown in Figure 5 and the model for losing coaches shown in Figure 6. The two models have some differences in the variables that were used to create them using a logistic model. Because of this, the null hypothesis can be rejected, suggesting that winning and losing coaches have different factors going into their decision to attempt a conversion or kick on fourth down.

Figure 5

Significant predictors for winning NFL coaches fourth down decision.

Coefficients: (4 not defined because of singularities)				
	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	1.676e+06	1.516e+06	1.105	0.26902
season_typeREG	5.199e-01	1.848e-01	2.814	0.00490 **
week	3.475e-02	8.297e-03	4.188	2.81e-05 ***
yardline_100	-1.351e-02	7.385e-03	-1.829	0.06736 .
half_seconds_remaining	-8.528e-05	3.128e-04	-0.273	0.78516
game_seconds_remaining	-5.208e-04	3.957e-04	-1.316	0.18809
game_halfHalf2	6.583e-01	4.932e-01	1.335	0.18197
game_halfovertime	NA	NA	NA	NA
qtr	2.614e-02	1.755e-01	0.149	0.88164
goal_to_go	-5.532e-01	1.986e-01	-2.785	0.00536 **
ydstogo	-6.038e-02	9.205e-03	-6.559	5.40e-11 ***
ydsnet	5.606e-02	2.565e-03	21.855	< 2e-16 ***
timeout	2.892e+00	6.020e-01	4.803	1.56e-06 ***
posteam_timeouts_remaining	-2.778e-01	5.522e-02	-5.031	4.88e-07 ***
defteam_timeouts_remaining	8.917e-02	5.408e-02	1.649	0.09918 .
score_differential	1.117e-02	8.255e-03	1.353	0.17591
no_score_prob	-1.676e+06	1.516e+06	-1.105	0.26902
opp_fg_prob	-1.676e+06	1.516e+06	-1.105	0.26901
opp_safety_prob	-1.675e+06	1.516e+06	-1.105	0.26904
opp_td_prob	-1.676e+06	1.516e+06	-1.105	0.26901
fg_prob	-1.676e+06	1.516e+06	-1.105	0.26901
safety_prob	-1.676e+06	1.516e+06	-1.105	0.26895
td_prob	-1.676e+06	1.516e+06	-1.105	0.26902
total_home_rush_epa	1.491e-03	7.005e-03	0.213	0.83144
total_away_rush_epa	NA	NA	NA	NA
total_home_pass_epa	3.605e-03	3.302e-03	1.092	0.27493
total_away_pass_epa	NA	NA	NA	NA
wp	-3.706e+00	3.054e-01	-12.134	< 2e-16 ***
def_wp	NA	NA	NA	NA
div_game	3.100e-02	8.361e-02	0.371	0.71080

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

Figure 6

Significant predictors for losing NFL coaches fourth down decisions.

Coefficients: (4 not defined because of singularities)				
	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	6.184e+05	2.000e+06	0.309	0.75716
season_typeREG	-7.002e-01	4.994e-01	-1.402	0.16085
week	1.675e-02	1.096e-02	1.528	0.12645
yardline_100	1.672e-02	1.014e-02	1.649	0.09920
half_seconds_remaining	-6.099e-04	3.877e-04	-1.573	0.11565
game_seconds_remaining	-4.368e-05	5.025e-04	-0.087	0.93073
game_halfHalf2	1.527e+00	5.945e-01	2.569	0.01019 *
game_halfovertime	NA	NA	NA	NA
qtr	1.135e-01	2.354e-01	0.482	0.62967
goal_to_go	-4.132e-01	2.798e-01	-1.477	0.13977
ydstogo	-4.556e-02	1.151e-02	-3.958	7.54e-05 ***
ydsnet	5.446e-02	3.657e-03	14.891	< 2e-16 ***
timeout	1.903e+00	6.438e-01	2.956	0.00312 **
posteam.timeouts_remaining	-3.435e-01	7.758e-02	-4.427	9.54e-06 ***
defteam.timeouts_remaining	9.823e-02	8.022e-02	1.224	0.22077
score_differential	6.157e-03	1.073e-02	0.574	0.56615
no_score_prob	-6.184e+05	2.000e+06	-0.309	0.75716
opp_fg_prob	-6.184e+05	2.000e+06	-0.309	0.75715
opp_safety_prob	-6.183e+05	2.000e+06	-0.309	0.75717
opp_td_prob	-6.184e+05	2.000e+06	-0.309	0.75715
fg_prob	-6.184e+05	2.000e+06	-0.309	0.75716
safety_prob	-6.185e+05	2.000e+06	-0.309	0.75709
td_prob	-6.183e+05	2.000e+06	-0.309	0.75717
total_home_rush_epa	6.470e-03	9.217e-03	0.702	0.48271
total_away_rush_epa	NA	NA	NA	NA
total_home_pass_epa	1.431e-03	4.084e-03	0.350	0.72599
total_away_pass_epa	NA	NA	NA	NA
wp	-4.486e+00	4.366e-01	-10.274	< 2e-16 ***
def_wp	NA	NA	NA	NA
div_game	1.657e-01	1.115e-01	1.486	0.13715

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

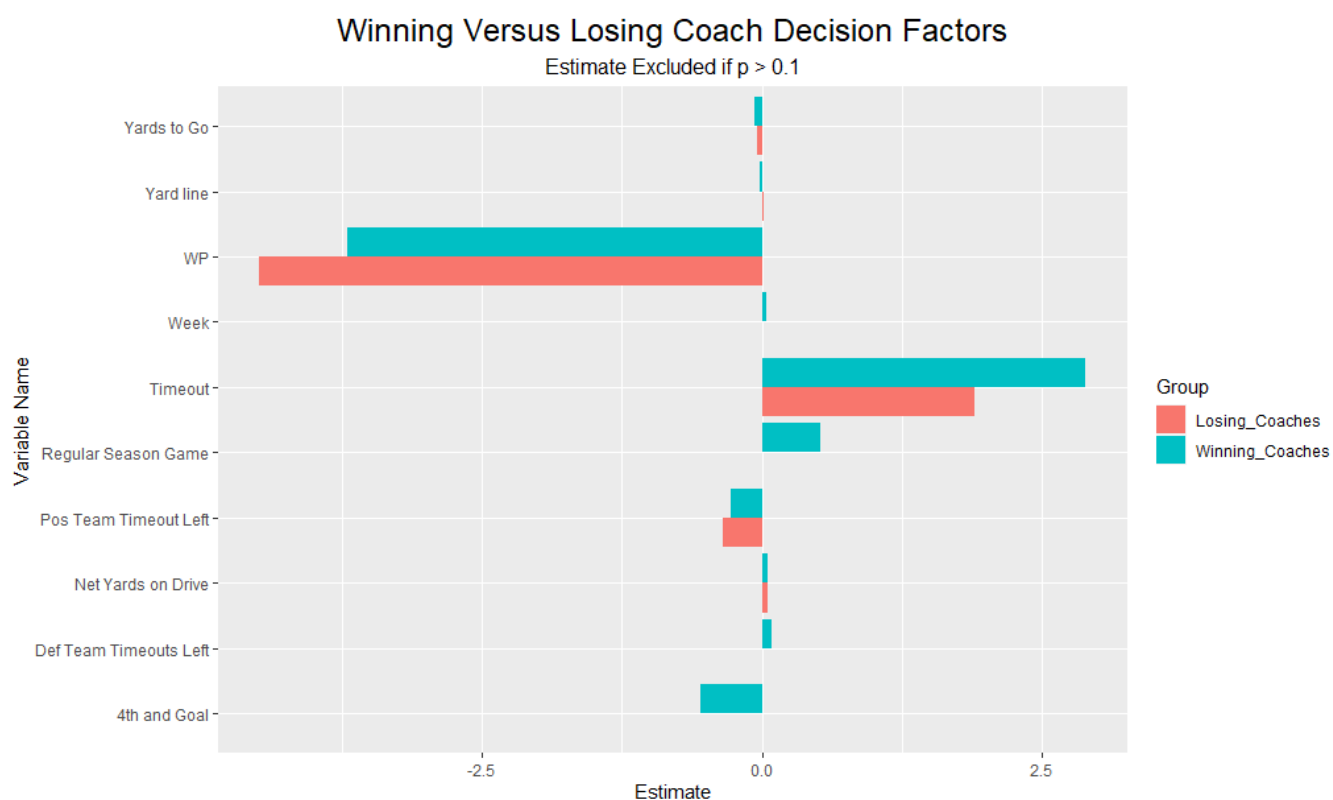
Since the significant attributes of coaching decisions were discussed in the overall model, the differences between winning and losing coaches will now be highlighted. The differences between the variables can be seen in Figure 7, with estimate values being set to 0 when the variable was not significant to an alpha of 0.10. First, the difference in the appearance of variables will be discussed.

Winning coaches did not appear to consider the half the game was in, while losing coaches did. This could show that winning coaches are more open to risk than initially thought, while losing coaches are not, or that losing coaches are losing at the end of games and more desperate to pull off the win. Losing coaches were found to not consider if it was a regular season game, the week, and the defensive timeouts remaining. The first two may be from the coaches not making it to the post season, thus not having additional motivation to win late in the

season games to clinch wins, but also could show a lack of building up their team. Not considering defensive timeouts remaining indicates these coaches may also be more focused on their own team, rather than the management of the other team.

Figure 7

Winning versus Losing Coach Decision Factors



There were also some differences in the estimates for each of the variables found in the model, as shown in Figure 7. Only one variable had a positive effect in one model, and a negative in the other, which was yard line. Although a smaller factor, it is still interesting that winning coaches are more likely to attempt a conversion the closer to the goal line they are, while losing coaches are slightly more likely as they are farther. All other variables were affecting the model in the same way, only with a change in the magnitude. Overall, there are

significant differences between the factors winning and losing coaches consider in their fourth down decisions.

Conclusion

The NFL has a high ceiling for the data collection and analysis using the systems already in place. This information can be a powerful tool for the coaches who decide to factor it in their decision-making process, but the reality is many are hesitant to use data. Through looking at fourth down decision making, coaches could compare their important factors to others around the league to see how their peers are making decisions. Implementing data in this way could help some coaches see the benefits of data for their team, as instead of a black box making decisions for them, they are compared to others around the league.

Through this fourth down analysis, winning and losing coaches were suggested to use different factors when deciding whether to attempt a fourth down conversion in a game. Although the two groups had similar rates of attempts, the models suggesting they have different decision-making processes is significant. This information can then be used to help increase the performance of coaches in the losing group, improving the record and profit of their team.

Recommendations

This project was able to provide a base look into if coaching groups considered different factors on their fourth down decisions from an already collected dataset of NFL play-by-play data. Because this project relied on already collected data, there is room for improvement with data collected designed for the purpose of looking into the differences between coaching groups. Some of this data could include key player performances, such as quarterback or kickers, surveys conducted on the coaches to see what their beliefs are, or other similar information collection.

From the data provided, one way to improve the dataset would be to look at the coaches win-loss record on a season-by-season basis, or even look at each coach over time. This would allow insight into the performance of the year's impact on the decision-making process, or if coaches changed their decision-making process as they gained experience. This could potentially unveil a difference between coaches with high and low experience and could be conducted using a similar dataset and methodology. Another improvement in this area could be to look at all the coaches' decisions made throughout their career from the same base used in this study. This would open more data for analysis and more potential for insight.

References

- Grandinetti, J. (2019). Welcome to a new generation of entertainment: Amazon web services and the normalization of big data analytics and RFID tracking. *Surveillance & Society* 17(1). Pg. 169-175. Doi: 10.24908/ss.v17i1/2.12919.
- Hadley, L. Poitras, M. Ruggiero, J. & Knowles, S. (2000). Performance evaluation of national football league teams. *Managerial and Decision Economics* 21. Page 63-70. Retrieved from <https://www-jstor-org.csuglobal.idm.oclc.org/stable/3108334>.
- Kovash, K. & Levitt, S. (2009). Professionals do not play minmax: Evidence from major league baseball and the national football league. *National Bureau of Economic Research*. Doi: 10.3386/w15347
- Palmquist, W. Elmore, R. & Williams, B. (2020). Fourth down decision making: Challenging the conservative nature of NFL coaches. *University of Denver Undergraduate Research Journal*. Retrieved from <https://www.duurjportal.com/index.php/duurj/article/view/93/60>.
- Romer, D. (2006). Do firms maximize? Evidence from professional football. *Journal of Political Economy* 114(2). Page 340 – 365. Doi: 0022-3808/2006/11402-0006\$10.00.
- Yam, D. & Lopez, M. (2019). What was lost? A casual estimate of fourth down behavior in the National Football League. *Journal of Sports Analytics* 5. Page 153 – 167. Doi: 10.3233/JSA-190294.