What Makes a National Football League Team Successful? an Analysis of Play by Play Data

Josef Ur
Department of Computer Science
MacEwan University
Edmonton, Canada
urj@mymacewan.ca

Mathew Craner
Department of Computer Science
MacEwan University
Edmonton, Canada
cranerm2@mymacewan.ca

Rehab El-Hajj
Department of Computer Science
University of Alberta
Edmonton, Canada
relhajj@ualberta.ca

***Abstract*—The National Football League (NFL) is one of the most popular sports in North America. The league showcases many strong athletes and winning is very important to all teams; however, for every winning team, there is a losing team, and it is the coaches' responsibility to decide what plays to call in order to help their teams win. Data mining play data can help show trends and areas where the top teams in the NFL excel by looking at questions like how often certain plays are run, how many yards do the plays get and where on the field are touchdown scored. With the help of data mining intelligent tools such as Naive Bayes, decision tree algorithms and association rules, we worked to isolate the areas where the best teams in the league separate themselves and produce winning franchises. By classifying the teams into two categories - the top teams and bottom teams - we were able to compare the two classes for differences which explain what results in more success in the league. Although we found that most teams - both top and bottom teams - use similar plays, there were also factors that distinguished the two types of teams. This research highlights these specific factors and some overall distinctions found between the more successful versus less successful teams to the NFL community.**

***Index Terms*—Algorithms, Management, Measurement, Design, Experimental, Sport data Analysis**

I. Introduction

During a National Football League (NFL) season, a coach makes over 1000 play calls on offence [3]. Coaches spend a lot of time researching and preparing for each game in order to maximize the plays they run and yield the best results for their teams. A data set from Kaggle.com [10] contains every play ran from 2009 to 2018 for every team. This data set covers information such as which team had possession, what play was ran, how many yards were gained, if points were scored on the play and the amount of time left in the game. Through using SQL Business Intelligence Tools for Visual Studio [1], we will analyze this data set and compare teams that won more vs. teams that lost more. Our goal is to try and find answers to some of the following questions: how do different teams call plays? Are some teams better at gaining yards in certain situations? Where are points scored? Are third downs important? Are penalties a factor in winning?

II. Related Works

Statistical analysis has always been an important aspect in professional sports. In Back to back evaluations on the gridiron, David J. Berri [4] states the two reasons why teams look at stats are to see what players contributed the most to a victory and which players are the best investment in the future. Player stats are a very important factor when creating contracts and many of them have performance bonuses which give the player a certain amount of money conditional to whether they reach certain milestones. By looking for tendencies in other teams combined with their own players strengths and weaknesses they try and figure out what players would be the best for the type of offence a teams wants to run.

978-1-7281-3885-5/19/$31.00 ©2019 IEEE

Unlike the National Basketball Association (NBA) where there are only 5 players per team on the court at any time, or the Major League Baseball (MLB) where one player can steal the show, the NFL is a true team sport. This is because while the MLB and NBA have multiple players on the field at once, it is significantly easier for a single player to take over a game. A single home run by a baseball player can end a game, while a basketball player can dominate by keeping the ball for them self and scoring the majority of the points, this is a lot harder to do in the NFL. On almost every NFL play at least three different players touch the ball. In order for a play to be successful, every player on the field has to contribute in some way. This makes the NFL a multiple-component sport and means that there are multiple aspects to determining a winner - which increases the difficulty of performing analysis. Player skill at each position, time remaining in the game, and overall team cohesion are some of the many variables that have an effect on how the game will proceed and determining outcomes [15].

III. The Game

Below we are going to talk about how the NFL is played and some key concepts to know. The National Football League (NFL) is the highest level of football played in North America. Made of 32 teams, each team comprises of a 53 player roster. Of those 53 players, 11 start on offense, 11 start on defence and the remaining 31 players are backups, with various other specialty positions. The objective of the sport is simple, try to score more points than your opponent in a 60 minute game. There are two main ways to score points: the first is a touchdown (TD), worth 6 points and the second is a field goal (FG), worth 3 points. Before the game starts, teams meet with referees to determine who will get first possession of the game, this is done by a coin toss. Once this is done, the team who lost the toss punts (kicks) the ball to the winning team, this starts the game. The length of a NFL field is 120 yards, 100 yards for the field of play and 10 yards for each teams end zone. The goal of the team who receives the ball is to go down the field and score a TD. To do this, the team needs to move the football 10 yards in 4 chances (downs). If a team gains more than 10 yards, they are given another 4 downs to move the ball. The team can move the ball in two ways: by running the ball or by passing the ball. If the team fails to do this, they must give the ball over to the other team. In order score points, a team must enter the end zone with possession of the football, a TD, or punt the ball through two large yellow uprights located in the back of each end zone, a FG. This goes back and forth for the whole game until time runs out and a winner is determined. [12]

IV. Data Mining Overview

The data set we used came from a open data source on Kaggle.com [10]. Entitled *Detailed NFL Play-by-Play Data 2009-2018*, it is a collection of every NFL play between the years 2009 and 2018. We looked at all 10 years of data which has over 449,000 plays with 255 different attributes. The data had many important dimensions such as:

- **Play Type:** A strategy used to move.
- **Yards Gained:** The amount of yards gained during a play.
- **Down:** 4 chances to move the football 10 yards.
- **Kicking Distance:** The distance in yards for kicking the ball.
- **Game Seconds Remaining:** Time remaining in seconds after each play.
- **Pass Location:** Location of the player when the pass occurred.
- **Run Location:** The location of the run in the field.
- **Touchdown:** A scoring play that occurs by advancing the ball into the opponent's end zone.
- **Third Down Converted:** The likelihood of a third-down conversion.

We removed the description column along with any column that had to do with a players name. The description was removed as it was qualitative data that could not be used in the analysis itself - though we did refer to it to help us understand some of the flags used by the data. The names were excess data as we already had unique player ids. For the data we collected, we had to add some new columns such as pre-computed columns like absolute score differential. This column took the absolute value of the home score minus the away

score and it doesn't change between negative or positive depending on who has possession of the ball. While cleaning the data, we had to apply three main techniques. First, any records with a lot of missing data were removed. Second, we changed all the "NA" in the flags for plays that they don't apply to, to be Null. This allowed the flags to be properly recognized as Booleans and changed the data to work better with our analysis in a way that would not change the results. Last, we changed the data type of some of the distance fields to floats from text based data types. For many attributes we had to smooth their integer values into ranges that facilitated the mining process [11].

A. Table Creation

Initially, the data set collected was very large, with 449,000 plays. As this initial data was very large and did not lead to any meaningful trends, we decided rather than comparing all the teams across all the years we would separate the teams into two classes, the top NFL teams versus the bottom NFL teams, per season. We would then select a set amount of each teams based on their win total. We were required to do this due to the fact that the league is split into two conferences (the American Football Conference and the National Football Conference) and they don't rank the two directly against each other even though the teams do play across the conferences. To identify the top and bottom teams for each year, we first looked at the yearly regular season standings for the league as a whole. The first aspect we looked at was the wins of the teams for that year. For the top teams it was the most wins and for the bottom teams it was the least wins. If there was a tie between teams, we then looked at points scored by the team - as our analysis is focused on offense plays. Again, for the top teams we were looking for the teams with the most points and for the bottom we were looking for the teams with the least points. Overall, the separation of top and bottom teams was done per season. This meant that if a team that was successful in a single season, it could be classified as a top team and the plays for that team in that particular season were kept. However, if that same team also had an unsuccessful season in the past it could be classified as a bottom team for that season, so the same team could appear as both a top and bottom team depending on how they played in a certain season. After filtering through this data, we had 42,020 data plays left for the top teams and 40,260 data plays for the bottom teams.

V. Methods

The tools used in the mining and analysis of the data [9] are SQL Server [14] to store the data and Visual Studio Business Intelligence Analysis Services [13] for the mining and analysis of the data. The combination of the two tools allowed for easy integration and work flow. Three types of models were used in the analysis of the data: Decision Trees [16], Naïve Bayes [5] and Association Rules [2]. The use of different models allowed for different angles to be looked at throughout the process and to help guide the process.

Decision Tree analysis [16] is a tree-like support supervised learning that takes one or more attributes as input and tries to predict possible consequences, including the chance an event may occur. This kind of algorithm is especially good for small data sets, which the NFL data became after running it through the pre-processing steps. Using Microsoft SQL data tools, the decision tree algorithm provided in this tool [7] required a key, input columns as well as predictable columns. The key is used to uniquely identify the record in the data set with the input being used to try and predict the predictable columns. If a specific input occurred multiple times with the output, that would have a high rate of being associated and the decision tree would find and display those correlations. When the tree is being built, a node was added for each input which is significantly correlated to a predictable column. This network of nodes is then displayed and trends can be found by studying which input is associated with the predictable columns. The decision tree also uses feature selection to take only the columns that provide the most information and removes the values that only occur a fraction of the time.

Next is another supervised learning algorithm, Naïve Bayes. We used this Microsoft tool [8] mainly for the initial modelling and to see portions within various inputs. The reason for this primarily being used as the initial modelling point is that Naïve Bayes treats all inputs as independent and then

Teams	Run	Punt	Kneel	Pass	Field Goal
Top	35.6%	5.2%	1.4%	48.5%	2.7%
Bottom	33.5%	7.2%	0.5%	49.8%	2.3%

TABLE I
PLAY DISTRIBUTION

predicts a result. With the NFL, the various different inputs can not be assumed independent as it is a very interconnected sport.

Finally is Association Rules, which was used for discovering interesting relations between variables. This intelligent tool supported by Microsoft [6] was used with the intention to identify strong rules discovered in data using some measures of interest such as support and confidence.Support is a term that describes the frequency of an item to occur with another item. Confidence is a percentage of confidence that two or more items would occur together. In our work, we used Association Rules for modelling penalties and their effect on the results of the games. This resulted in a large number of rules with varying importance and support. Many rules have an importance that is large enough to equate it with known occurrences or events.

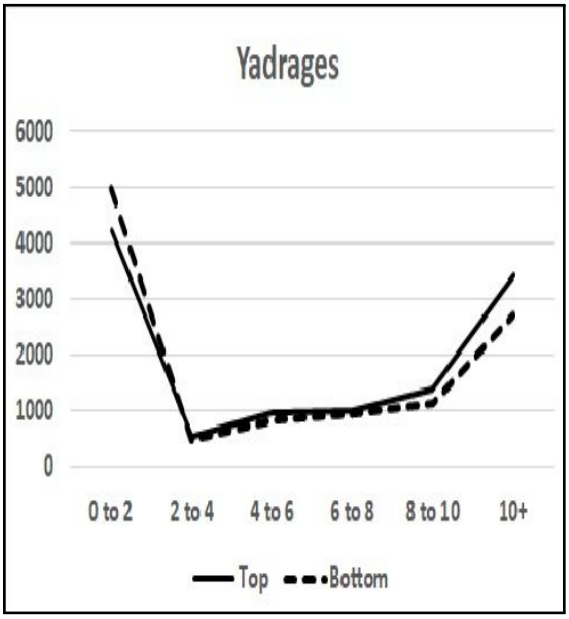

Fig. 1. Amount of Passing Plays each group completed

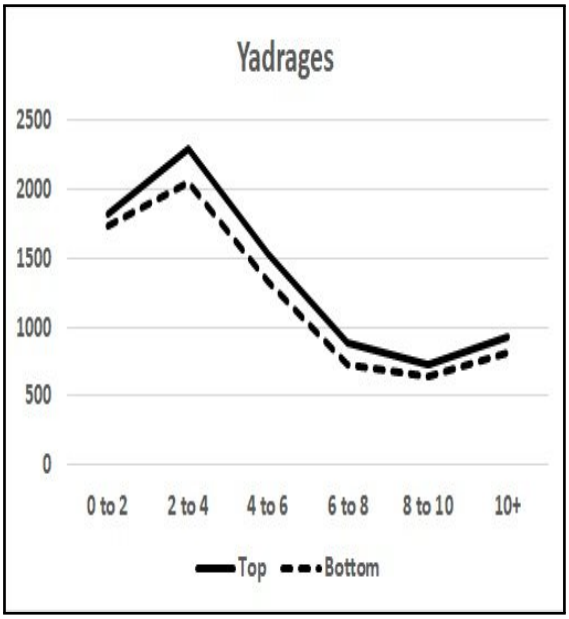

Fig. 2. Amount of Running Plays each group completed

VI. DATA ANALYSIS

We began looking into types of runs per team with our goal being to find patterns that differentiate each class from the other. To take this further, we asked some research questions such as: Do better teams run a different amount of plays? Where do teams gain yards? Where do teams lose yards? We also studied the significance of each of the 4 downs and how teams score points as well as the significance of specific score types - touch down vs goal field - and how they distinguish a top team from a bottom team.

A. Do better teams run a different amount of plays?

The play types chosen by the top and bottom teams was among the first questions we asked. The distribution we came across for the whole game is shown by Table I. In the table, it can be seen that the top and bottom teams choose similar plays and therefore, they potentially play similar games. These similarities suggest that even though the top and bottoms teams run similar percentages of plays, the top teams must being doing something different during the plays that help make them more

successful. Though the Kneel play, which is when the winning team kneels down, is an exception in that it is used to run the clock by the winning team. So as the top teams win more, they will Kneel more.

B. Where do teams gain yards?

Since we figured out how each set of teams differed in their play choices we decided to look at how those choices affected gaining yards. We looked at passing and running plays primarily as they accounted for over 80% of the total plays ran by each group as shown in Table I. We found that for both categories the majority of the plays ran gain 10 yards or less. So when we drilled down to investigate further we broke the data into 2 yard increments between 0 an 10 yards, and then put everything over 10 yards into its own category. As you can see in Figure 1 the data is very similar. The areas where the top teams have a noticeable advantage is the 10+ range. This suggest that the top teams are more adapt at both the short gains and longer gains. This is important as if the top teams find themselves in a situation where they need to get a large chunk of yardage they are able to do so more often than the bottom teams. Another important area to note is the 0 to 2 yard range, the bottom teams ran significantly more plays from that range. This is suggests that the bottom teams are more ineffective because their plays goes for less yards more often.

For the running plays there is a bit more of a gap. Figure 2 shows that the top teams run plays occur more often and they gain more yards from them. The largest difference occurs in the 2 to 6 yard range, where the top teams ran more than 400 more running plays than the bottom teams. This trend shows that the top teams running plays are more effective because they are able to gain more yards per attempt.

C. Where do teams lose yards?

In the association rules algorithm, supported by Microsoft business tools, we have to provide the input variables and determine the predicted values. In our example, we started with three input variables to test with: the inputs for the progress of the game, the chosen play type and the score differential. The predicted variable was the penalty type. With multiple runs, we found that no obvious rules were

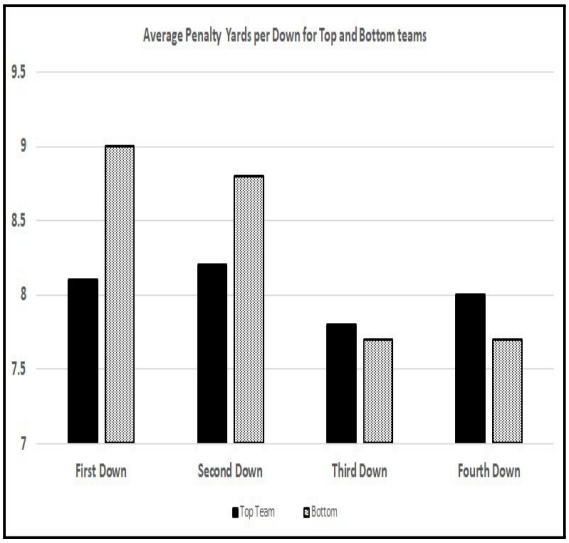

Fig. 3. Top teams(Orange) and Bottom teams(Blue) Average Penalty Yards

able to be generated. This suggests that the type of penalty assigned is not affected by the current status of the game at a larger level and may be more affected by the direct situation on the field that cannot be represented through our data set. The remaining rules are all situations in which penalties do not occur. The analysis of penalties leads into the analysis of yards lost through the penalties and their effect on the offensive plays of the teams.

Yards lost by penalties for the teams as shown in Figure 3 show that the top teams get more penalty yards in the earlier downs while the bottom teams get more penalties in the latter downs. As well, the latter downs account for less yards on average. If the top teams are getting more penalty yards but still winning then they seem to be better at moving down the field as they score more points and win more games. When looking at the Decision Trees for the teams and penalty yards it can be seen that the only factor that matters when given the same inputs as the association rules is penalty type in determining the yards lost by the teams. Then looking at the values given before any penalty types are taken into consideration, the majority of penalty yards for both sets of teams is less than 9 yards, followed by 9-12 penalty yards coming in second. The top teams

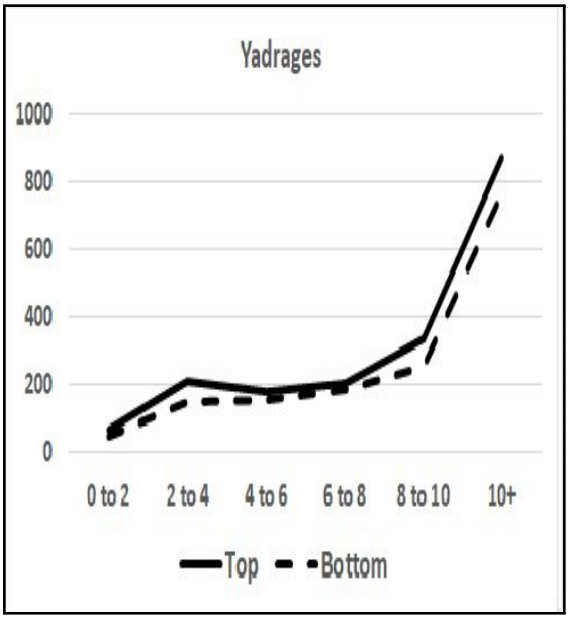

Fig. 4. Amount of Third Downs Converted

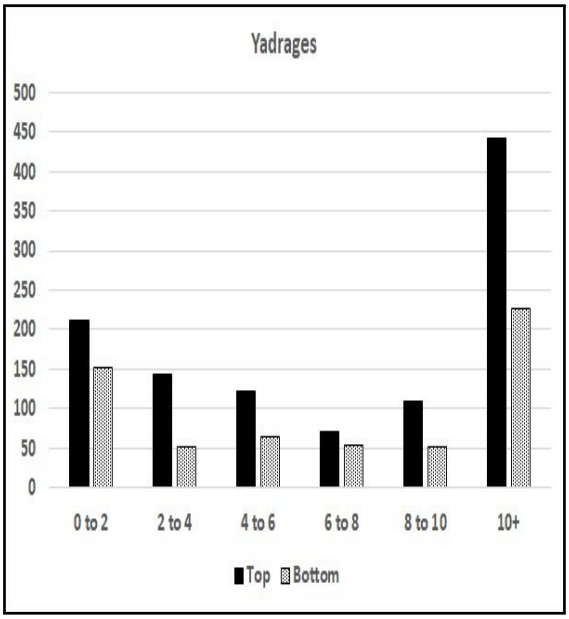

Fig. 5. Amount of Touchdowns

have a slightly increased probability of their penalty giving less than 9 yards vs the bottom teams.

D. Significance of Third Down Conversions

Of all the four downs, the third down is critical. It is the last chance an offence has before they have to give the ball away and the teams who convert more third downs have a higher chance of winning [17]. By keeping the same yardage categories as early, we can see in Figure 4 that the top teams convert more third downs across all categories. The top teams pull ahead in the 2 to 4 yard range and the 8 to 10+ range suggesting that they are able to convert both short yardage and long yardage with greater ease than the bottom teams. Being well versed in both areas help the top teams win more often because they are able to stay on the field longer and put themselves in a position to score more points.

E. How do teams score points?

Scoring is the main objective of any sport, and the NFL is no exception. The more points you score the higher chance you have of winning. Figure 5 highlights this concept very well. In almost all yard ranges, the top teams performance is over double of the bottom teams. This means that they are twice as likely to score a touch down (TD). This large difference is really where the top teams separate themselves from the bottom teams. Looking at the data, the top team scored an average of 83 more touchdowns per category. Over the course of a season, that number really adds up and helps give the top teams the edge.

Fields goals are another important aspect of scoring in the NFL. Figure 6 shows that while the top teams don't have as much of an advantage across the board, they still out perform the bottom teams in every category. Again since the top teams are scoring more field goals all the time, an average of 12 more field goals per category, they are able to put themselves in a position to win more often and by a larger margin.

VII. Conclusion

In conclusion, we were able to find that even though the top and bottom teams run similar amounts of plays, their effectiveness of the plays greatly differ. The top teams ran more plays at longer yardages making it easier for them to convert third downs. We found that even if the

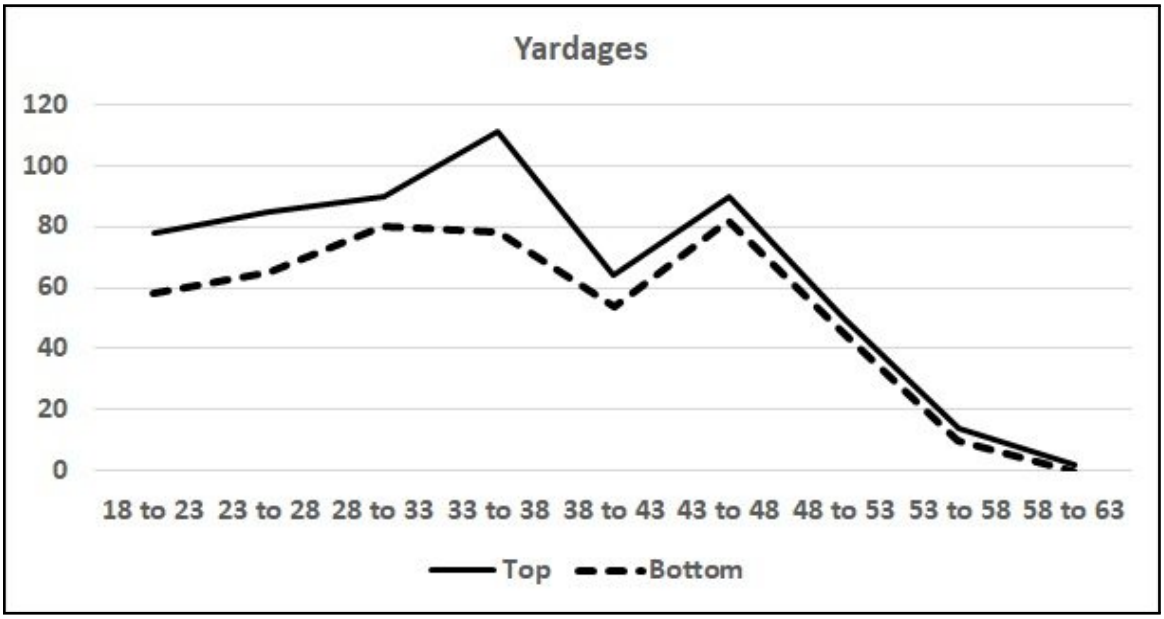

Fig. 6. Amount of Field Goals Made

top teams get into longer yardages on the third down, they were able to convert them more often than the bottom teams. Penalties that the top team committed were on earlier downs vs. the latter downs for the bottom teams. Having penalties on the early downs gives the top teams more chance to gain that yardage back and move down the field. Finally, we discovered that touch down scores are more critical for defining top team behaviour and winning more games.

Acknowledgement

The authors thank Max Horowitz and Kaggle.com for providing the data for this research.

References

[1] Empower your people with self-service business intelligence white paper. *Microsoft Dynamics AX*, May 2008.
[2] Rakesh Agrawal and Ramakrishnan Srikant. Fast algorithms for mining association rules. In *Proceedings of the International Conference on Very Large Databases*, pages 487–499. ACM Press, 1994.
[3] Benjamin C Alamar. Measuring risk in nfl playcalling. *Journal of Quantitative Analysis of Sports*, 6(2), 2010.
[4] David J. Berri. Back to back evaluations on the gridiron. In Jim Albert and Ruud H. Koning, editors, *Statistical Thinking in Sports*, chapter 14, pages 241–263. Publisher: Chapman and Hall/CRCs, 2007.
[5] Pedro Domingos, Pedro Domingos, and Michael Pazzani. On the optimality of the simple bayesian classifier under zero-one loss. *Mach. Learn.*, 29(2-3):103–130, November 1997.
[6] Owen Duncan, Craig Guyer, and Carl Rabeler. Microsoft association algorithm, 2018.
[7] Owen Duncan, Craig Guyer, and Carl Rabeler. Microsoft decision trees algorithm, 2018.
[8] Owen Duncan, Gene Miller, and Carl Rabeler Craig Guyer. Microsoft naive bayes algorithm, 2018.
[9] Jiawei Han, Micheline Kamber, and Jian Pei. *Data Mining Concepts and Techniques*. Morgan Kaufmann, 3rd edition, 2012.
[10] Max Horowitz. Detailed nfl play-by-play data 2009-2018, 2019.
[11] Ivy Liu and Alan Agresti. The analysis of ordered categorical data: An overview and a survey of recent developments. *Test*, 14(1):1–73, Jun 2005.
[12] Tim McDonald. A beginner's guide to football, 2017.
[13] Microsoft. Sql business-intelligence, 2017.
[14] Microsoft. Sql server, 2017.
[15] B Ofoghi, J Zeleznikow, C MacMahon, , and M Raab. Data mining in elite sports: A review and a framework. *Measurement in Physical Education and Exercise Science*, 17(3):171–186, July 2013.
[16] Craig Guyer Owen Duncan, Bruce Hamilton. Microsoft decision trees algorithm, 2016.
[17] T. Troy Russell. Third down efficiency, three and outs and other stats i hate, 2017.