**Assignment 1**

In this essay, were going to take a closer look at Tom Brady’s and the New England Patriots’ record breaking 2007 regular season. Brady’s 2007 season saw him break records such as most consecutive games with three touchdowns thrown (9), most touchdowns throw in a regular season (50), best touchdown to interception ratio (50:8) and matched Hall of Fame quarterback Roger Staubach for the most wins ever by a starting quarterback in their first 100 games (76). Brady also had several personal milestones such as throwing his most ever touchdowns in one game (6) and his first game with a perfect passer rating. He led the Patriots to their first ever undefeated season (16-0), directed an offence that scored a then-record 589 points and 75 total touchdowns, 50 of which were his passes and was named the NFL’s Most Valuable Player and Offensive Player of the Year for the first time. We are going to look at the passing stats behind this legendary season and get an insight into how Brady liked to use his top 3 wide receivers in Randy Moss, Wes Welker and Donte’ Stallworth. The data comparing Brady and the other quarterbacks in the league that year will be demonstrated on a line graph using parallel coordinates. I chose this method as it allows the viewer to distinguish different pieces of information based on colour and it also shows the gulf between the statistics on the plots in the graph.

The reasoning behind using a parallel coordinate plot instead of something such as a pie chart is the fact that in using the plot, it will allow me to demonstrate multiple statistics across one graph as it allows for direct comparisons of multiple variables across parallel axes. A layout like this makes it easier for the viewer to see patterns and trends and allows them to see if multiple statistics are related to one another, none of which would be possible on a pie chart. I feel is the best way of comparing a dense dataset as it allows for great readability regardless of your knowledge on the topic of discussion. In terms of sports stats, it is the most straightforward graph to use to represent most comparisons between individuals under any set of statistics.

In this interest of fairness, I will only be comparing Brady against other quarterbacks that played all 16 games of the regular season (except Brad Johnson who played all 16 games as a backup for Tony Romo in Dallas) as well, so the statistics don’t get skewed. To demonstrate the data, I used a line chart with multiple different statistics as I felt this was the best method of showing how great Brady was visually as it compares all the quarterbacks’ stats for the season across multiple metrics. I feel like it also allows for mitigation of outliers in data as one bad game won’t allow the data to deviate from the mean too far which would happen if the data shown was on a game-by-game basis. Below you can see the graph of all the quarterbacks’ stats and another graph with the stats of Brady highlighted.

In the graph we see a direct comparison of Brady against his peers. Each axis on the graph displays a performance metric, clearly showing how far above Brady was in each individual stat, giving instant insight into the level of performance he was at. There are certain metrics where another quarterback is close to Brady, for example Drew Brees is close to him in terms of completion percentage, but when you compare Brees’ touchdown per games to Brady’s, its not even close. The plot allows for this advantage to be clearly displayed, not allowing the viewer to be confused as to what they’re seeing. In other areas, Brady is head and shoulders above the other and in my opinion the plot does a great job at showing how big this gap is in stats such as Pass Rating and Touchdown/Percentage.A graph of different colored lines

Description automatically generated

Choosing the parallel coordinate plot to show this dataset allows me to illustrate how the plot simplifies the complex relationships between the different stats. For example, someone may not know what any of the abbreviations for the stats stand for, but just by looking at the graph, most people would be able to tell you that Tom Brady had the best season out of all of them just down to how high above (or below depending on the stat – i.e. interceptions) the competition he is. Adding in the interactivity to this graph is also something that adds to this as it allows the viewer to focus directly on any quarterback they like and see an unimpeded view of their stats.



Of course, no quarterback can be as good as Brady was in that season without some great running mates, or in Brady’s case, his trio of wide receivers in the for of Randy Moss, Wes Welker and Donte’ Stallworth. In the 2007 season Wes Welker and Randy Moss both finished in the top 10 in terms of receptions at number 2 (Welker) and 9 (Moss) respectively. Both were number 2 (Moss) and number 11 (Welker) in yards gained as well. Moreover, Randy Moss still holds the record for the most touchdowns in season by a wide receiver of all time with 23, surpassing a record that had stood for 30 years before that. Below is a table of their stats that season.



To put these stats into perspective, of Brady’s 398 completions in 2007, these 3 guys made up for 256 (64.32%) of them and of his 50 touchdowns, they made up 34 of them (68%). While Moss and Welker were doing most of the scoring, it was Donte’ Stallworth who was a bit of an unsung hero in this team, being responsible for converting a lot of 3rd down opportunities allowing the team to continue pushing up the field and opening the door for the other two guys to shine. Below is another parallel coordinate plot of the top wide receivers in the league.

To show the level both Welker and Moss were performing at, the plot with their data compared to the rest of leagues best wide receivers will also only be against those who only played all 16 regular season games (some of them didn’t start all 16 games but did play in all of them). The data used to plot the graph is contained in the following table.

*A graph with lines and numbers

Description automatically generated*

If you look for Brady’s leading 2 receivers on this graph (Moss and Welker) you’ll be able to see there are certain stats they are miles clear of the opposition in – Moss with 23 touchdowns and 0 fumbles and Welker in Ctch% and Succ%. Using the parallel coordinate plot, the gap in these statistics is fully on show and it is clear that because of the graph we chose, after looking at what data is being presented in the table above, the viewer can easily distinguish what’s going on in the graph and it would allow them to understand what the graph is displaying. Using the parallel coordinate plot in this instance allows for comparisons of different players across various performance metrics all in one place as opposed to changing between multiple charts. This holistic view is very important in the context of sports analytics as the total impact of one player cannot be defined by a one statistic.

In conclusion, the parallel coordinates plot has proven to be a highly effective method for visualizing the multifaceted data related to quarterback performance in the NFL. This technique has allowed for us to draw comprehensive comparisons between Tom Brady and his competition in the quarterback position, allowing us to highlight the interconnectivity of various performance metrics. The visual clarity the parallel coordinate plot allows us to present allows a deeper understanding that overcomes the limitations of regular tabular data presentation.

On the same note, taking a closer look at Brady’s wide receivers, and comparing them to the league’s best wide receivers, it lets us gauge the individual greatness but also their chemistry on the field with Tom Brady and showing how far apart from the rest of the league it set them apart. The use of the parallel coordinates plot allows us to demonstrate how well Randy Moss and Wes Welker were able to perform while having Brady at the helm.

On a personal level, I feel like d3.js is a great tool for people like sports analysts to use to display data. In the professional world I believe that the parallel coordinate plot is underutilized and can be very descriptive with the right dataset attached to it.

References

The data used in this essay can be found at the following locations:

<https://www.pro-football-reference.com/years/2007/receiving.htm>

<https://www.statmuse.com/nfl/ask/what-are-the-most-receiving-tds-in-a-season-by-a-wide-receiver>

<https://www.statmuse.com/nfl/ask/tom-brady-2007>

<https://www.pro-football-reference.com/players/W/WelkWe00.htm>

<https://www.pro-football-reference.com/players/M/MossRa00.htm>

<https://www.pro-football-reference.com/players/S/StalDo00.htm>

**Please note** that some of the stats have been divided by 16 to calculate the per game value.

Video Link

<https://drive.google.com/file/d/1xm40VdLI4sDVbvvatfvUwqkTudnXJ88L/view>