NFL Dashboard Writeup

Tommy Burkett

College of William & Mary

I. Project Overview

This project uses National Football League data to create an interactive dashboard that visualizes key NFL team statistics across multiple seasons. Even though there are many categories under season statistics, this project emphasizes offensive and defensive performances during the 17game regular season. The National Football League, or sports in general, are increasingly becoming data-centered, with individual and team performances at the forefront. Teams and analysts use these statistics for game plans, player analysis, and predictions. This dashboard will provide users with an organized and simplified interface to visually represent these performances.

The visualizations in this dashboard allow for quick and easy comparisons between teams across each of the last five seasons, saving fans, like me, from having to scroll through dense and disorganized spreadsheets. Additionally, for those who play Fantasy Football each season, this interactive dashboard will highlight teams that perform well offensively, which will help guide you through your fantasy draft and in-season trades. Further, through these visualizations, users can uncover team trends and patterns due to the project's five-year data range. Lastly, this project prioritizes reproducibility, so this dashboard will be updated at the end of each season, always displaying the previous five-year interval.

II. Data

The data utilized for this project consists of National Football League play-by-play data summed up across an entire 17-game season. This data was aggregated by Pro Football Reference, which is a reliable website for NFL data [1]. Each row in the dataset represents one of the 32 NFL teams for that season.

For this project, there are two datasets for each season. The first dataset is Team Offense with key offensive statistics like total yards, total touchdowns, passing/rushing yards, passing/rushing attempts, passing/rushing yards per attempt, passing completions, penalties, and turnovers (fumbles and interceptions). The Team Defense dataset is formatted the same, with variables such as passing/rushing yards against, passing/rushing attempts against, passing/rushing touchdowns against, and penalties against [1].

There are a few variables that are not obvious, such as turnover percentage, scoring percentage, and net yards gained per pass attempt, because they involve in-depth calculations based on team performance. However, these variables are not significant to the project, so providing extraneous details or explanations is unnecessary.

The data preprocessing for this project was pretty straightforward and not too intricate. The data is separated by seasons, so all of the Team Offense and Team Defense files had to be

downloaded separately for each season. After they were downloaded, they were consolidated and cleaned up. Some rows were removed for formatting purposes, such as the “Tot Yds & TO”, “Passing”, “Rushing”, “Penalties”, “Avg Team”, “League Total”, and “Avg Tm/G” because including them was unnecessary for the scope of my project. It also created a weird dataframe with odd column names, so I found it best to remove the rows entirely.

After the rows were removed and the datasets were combined, I renamed many variables to fit their description, so “Yds” under the “Passing” header became “Pass Yds” and so on. This allows me and the user to easily understand what the columns represent, instead of having to refer to the section headers and locate the cutoff columns.

Up to this point in the project, I have used the Pandas, NumPy, os, Dash, and Plotly libraries. The Pandas library allows me to load and clean the CSV files and transform them into dataframe objects [3]. The NumPy library allows for easy mathematical computing, especially if I consider adding any calculated rows to the datasets [2]. The os library allows me to check for existing files in the “data” folder, so that I can upload them and condense them into a singular CSV file. The most important library implemented in my project is Dash, which allows me to create an interactive dashboard with buttons and drop-down menus for the user to select [5]. Lastly, Plotly is just as important as Dash, since it powers the Dash visualizations and allows me to create interactive labels and graphs [4].

Table 1. Cleaned Data

Rk	Team	Games Played	Points Allowed	Total Yards	Yards Per Play	Total Turnovers	Passing Attempts	Passing Yards	Passing Touchdowns	Interceptions	Rushing Attempts	Rushing Yards	Rushing Touchdowns	Year	Side	Points For
1.0	Buffalo Bills	17.0	289.0	4637.0	4.6	30.0	530.0	2771.0	12.0	19.0	442.0	1866.0	19.0	2021	Defense	
2.0	New England Patriots	17.0	303.0	5284.0	5.1	30.0	538.0	3181.0	21.0	23.0	464.0	2103.0	9.0	2021	Defense	
3.0	Denver Broncos	17.0	322.0	5544.0	5.3	19.0	562.0	3652.0	22.0	13.0	441.0	1892.0	9.0	2021	Defense	
7.0	Philadelphia Eagles	17.0		6024.0	5.4	28.0	563.0	3834.0	24.0	16.0	510.0	2190.0	22.0	2023	Offense	433.0
8.0	Los Angeles Rams	17.0		6108.0	5.6	18.0	583.0	4063.0	26.0	13.0	477.0	2045.0	18.0	2023	Offense	404.0
9.0	New Orleans Saints	17.0		5732.0	5.1	18.0	606.0	3990.0	28.0	11.0	480.0	1742.0	13.0	2023	Offense	402.0
30.0	Philadelphia Eagles	17.0	428.0	6054.0	5.5	18.0	652.0	4296.0	35.0	9.0	410.0	1758.0	13.0	2023	Defense	
31.0	Arizona Cardinals	17.0	455.0	6047.0	5.7	17.0	506.0	3613.0	32.0	11.0	521.0	2434.0	19.0	2023	Defense	
32.0	Washington Commanders	17.0	518.0	6612.0	5.9	18.0	598.0	4457.0	39.0	8.0	480.0	2155.0	15.0	2023	Defense	

Table 1: Sample of cleaned NFL season statistics used in this project. The 9 rows were randomly selected from the full dataset to illustrate the structure and variables used for the project.

Figure 1. Offensive Team Performance (2024) Graph from Dashboard

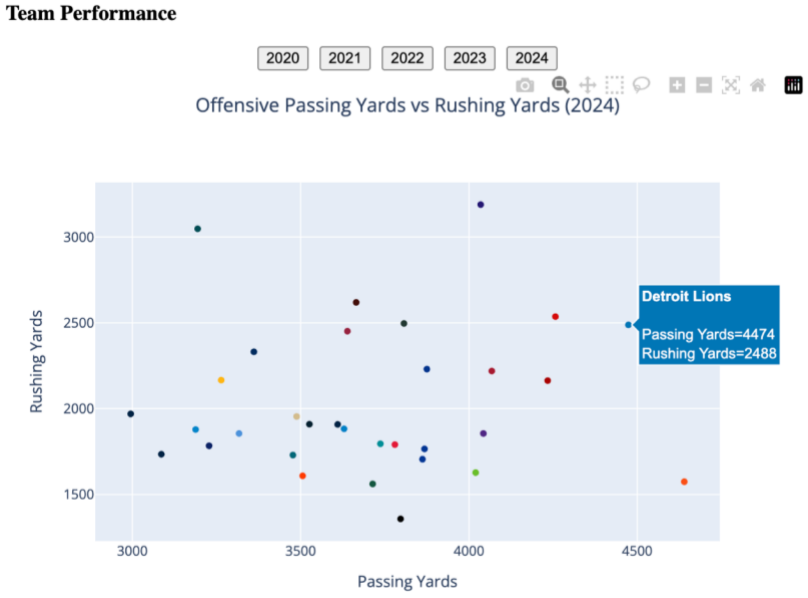


Figure 2. Defensive Points Allowed by Team (2022) Graph from Dashboard

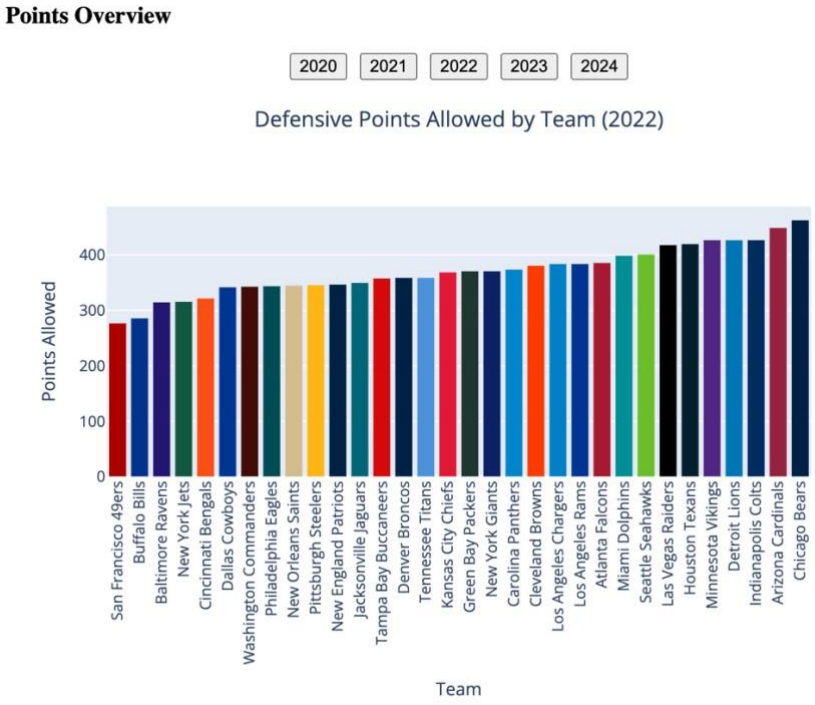
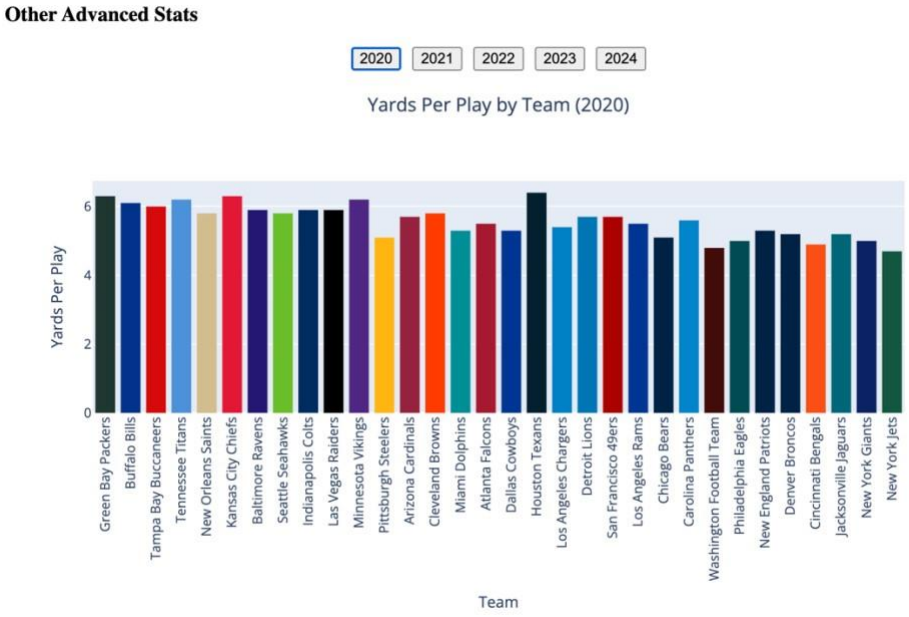


Figure 3. Offensive Yards Per Play by Team (2020) Graph from Dashboard



III. Conclusion

I will use this project to view seasonal offensive and defensive statistics in an organized, interactive dashboard. This dashboard allows me, and others, to quickly compare teams across the five included seasons and uncover trends and team performances through the visualizations. This dashboard will remove the need to scroll through unorganized and dense spreadsheets. Not only will it reveal team trends and patterns, but it will also serve as a tool for Fantasy Football users. For example, some visualizations will identify the top-scoring offenses in the league, guiding Fantasy Football managers to select certain players from these high-scoring teams. Additionally, with the defensive statistics, Fantasy Football managers will see which defenses to avoid, especially because most fantasy leagues constantly add and drop defenses, depending on each team's weekly matchup.

IV. References

Data Source:

[1] Pro Football Reference. 2020-2024 NFL Standings & Team Stats.

2024 Team Offense. https://www.pro-football-reference.com/years/2024/#all_team_stats

2024 Team Defense. https://www.pro-football-reference.com/years/2024/opp.htm

2023 Team Offense. https://www.pro-football-reference.com/years/2023/#all_team_stats

2023 Team Defense. https://www.pro-football-reference.com/years/2023/opp.htm

2022 Team Offense. https://www.pro-football-reference.com/years/2022/#all_team_stats

2022 Team Defense. https://www.pro-football-reference.com/years/2022/opp.htm

2021 Team Offense. https://www.pro-football-reference.com/years/2021/#all_team_stats

2021 Team Defense. https://www.pro-football-reference.com/years/2021/opp.htm

2020 Team Offense. https://www.pro-football-reference.com/years/2020/#all_team_stats 2020 Team Defense. https://www.pro-football-reference.com/years/2020/opp.htm

[2] NumPy. https://numpy.org/

[3] Pandas. https://pandas.pydata.org/

[4] Plotly. https://plotly.com/

[5] Plotly. Explore Dash. https://plotly.com/examples/

[6] Plotly. Explore Dash. Sports Analytics. Alfredo Gutierrez. https://nflstatsdashboard.pythonanywhere.com/