

# **Washington Huskies Football CFP Analysis**

By Jack Arbuckle, Reuben Dayal, and Ian Pezzella

## **Motivation**

In this project, the 2023 Washington Huskies' football team data will be analyzed to see how they would match up against the top 4 teams in college football. Since the college football regular season is winding down and teams are jockeying for playoff positioning and potential bowl games, this is a very relevant and interesting problem. Washington currently sits at number 5 with less than a month left and is in a prime spot to work into the top 4. Washington was selected as the team in question because of its chance of making the college football playoff. It is also not one of the "blue bloods" of college football so they get less attention than some of the other top teams in the national media. Washington also does not play the vast majority of CFP title contenders, such as Georgia or Ohio State, over the course of the regular season, so there is a lot of question about how a potential matchup between Washington and these other contenders would turn out. It is a question that can help Washington see how they stack up if they do in fact make the CFP.

The question of ranking teams vs other teams without having head to head or an overlapping matchup is not new. Metrics known as "Strength of Schedule" and FPI have been invented to attempt to solve this problem. ESPN and other sports outlets have analyzed this metric and many other metrics to judge the caliber of opponents that teams have faced. The CFP committee that releases the rankings are supposed to take strength of schedule and the quality of their top wins into account but often recent bias, tradition, and other factors get in the way of ranking teams. Washington has a win vs the #6 team in the nation (the highest ranked win by any team) and is undefeated but yet is still behind four teams. This project looks to analyze strength of schedule based on Expected Points Added by their top opponents and judge how Washington looks on both sides of the ball against the other top teams with their opponents taken into account.

## **Problem Framing**

To analyze the matchups between Washington and the other four members of the top 5 in the latest CFP rankings, rushing and passing EPA per play for offense and defense will be used for our main analysis. The strengths and weaknesses of each team will be looked at using EPA. Strength of schedule will also be taken into account, as these five teams played a vastly different regular season schedule and had differing levels of opponent toughness. (We are still trying to figure out how to calculate EPA for each team while weighting them based on their opponents EPA) Observing the trends in the graphs created will allow the strengths and weaknesses of each team compared to their respective schedules to be shown. This will help show

where Washington stands compared to the rest of the teams on a more “even playing field” Not all the opponents that each team played were used in the opponent analysis. The analysis focused on more meaningful opponents (FBS rather than FCS).

Creating visuals that compare Washington’s strengths and weaknesses to the rest of the teams with the proper weights, a rough plan of attack can be devised for Washington against each of the teams. For example, if FSU has a high passing yards EPA allowed and a low rushing yards EPA allowed, it might be wise for Washington to throw the ball more especially when they have a high EPA per play for passing. Insights gained from the data like these will help to come up with the best plan of attack for Washington if they do get the opportunity to play one of these teams.

## Data Overview

This project is using the college football play by play data for analysis. This data looks at every play for every game for each college team. We will filter the data for the top five teams in the country this year (Ohio State, Georgia, Michigan, Florida State, and Washington) and also filter it for only the teams that the above-mentioned teams played.

The data describes all the dynamics of the play to include the play type, the period, the amount of time on the clock, offensive and defensive epa per play, and offensive and defensive wpa per play. EPA is the expected points added while the WPA is the win probability added. The epa and wpa changes for each play depending on the flow of the game. The data is very well structured but is broken down for the project into team summaries and certain teams. The team summaries take all the play by play data for a certain team and create a row of stats just for that certain team.

How to weight (experiment)(Numbers may have changed on R file)

Weighting the Offense (Higher Weight the better the offense)

Teams	Passing EPA Per Play (Offense)	Opponents Passing EPA DEF	Formula (Plus 1 gets rid of negative values)	Weighted Passing Per Play (Offense) (New Metric For comparison)
Georgia	0.3987114	0.15815626	$10^{*(1/(0.15815626+1))}*(0.3987114+1)$	12.0770525387
Michigan	0.5077118	0.06057973	$10^{*(1/(0.06057973+1))}*(0.5077118+1)$	14.2159213245

### Weighting the Defense (Lower the weight the better the defense)

Teams	Passing EPA Per Play DEF	Opponents Passing EPA	Formula	Weighted Passing Defense
Georgia	-0.07572475	0.18072007	$10^{*(1+(-0.07572475))} * (1/(1+0.18072007))$	7.8280705434
Michigan	-0.15126227	0.01259745	$10^{*(1+(-0.15126227))} * (1/(1+0.01259745))$	8.3817888342

I had trouble weighting the EPA with negative values so I added 1 to all the EPA values to get them away from being negative. The new metric is not EPA or anything that has a name. It is a new metric whose purpose is only for comparing the teams selected offense and defense to each other based on the strength of their opponents. The times 10 makes it easier to visualize on the graph.

### Contribution

Reuben: 33%: Motivation and Problem Framing

Ian: 33%: Data Overview

Jack: 33% Editing and R Code