

Brandon Mather  
DSC 540-T302  
Milestone 1

Project Subject Area: What I want to try and accomplish with my project is to determine which team is most likely to win the 2023 super bowl based off 2022 statistics. Then I will rank each NFL team, 1 to 32, 1 being the most likely to win the super bowl, 32 being the least likely.

Data Sources:

CSV File:

Description: I will be using the 2022 play by play data. This is every play ran in the NFL in the entire 2022 season with their results.

Link: <http://nflsavant.com/about.php>

API:

Description: This is an API from RapidAPI that has the following NFL team base stats: wins, losses, ties, win-loss percentage, points scored, points scored against, point differential and margin of victory.

Link: <https://rapidapi.com/eternal-la-mancha-eternal-la-mancha-default/api/nfl-statistics/>

Website:

Description: This website keeps track of more advanced passing, rushing, and receiving statistics for the 2022 NFL season.

Link: <https://www.pro-football-reference.com/years/2022/advanced.htm>

Relationship: All sources are related by NFL teams. The CSV file has a one-to-many relationship with the API by NFL teams and has a one-to-many relationship to the Website by NFL teams as well.

I plan to complete this project by mainly focusing on offensive production and defensive efficiency. I believe I can use the data that I found to create an idea of who the better teams were. I'm going to focus mainly on the 2022 season, since it's the most recent. It won't be perfect, especially going into a new season where things can change, such as: Players, Coaches, offensive/defensive schemes. This will more give an idea of what teams may have a better head start to be successful more than a perfect prediction. The ethical implications of the project would be related to sports betting. With sports betting becoming so large and everyone looking for an edge that data could be used unethically. I think the most challenging part of the project is making sure I use the correct data to convey my message on why it makes a team successful. I don't want everything, because not all of it matters to a win or loss and may not convince someone that the data is helpful. The other thing I want to make sure I focus on is creating good visuals that makes sense and can compare all teams together. This will be very interesting but may be too large scale, if it gets too much I won't worry about the rankings as much, my main goal is to come away with the number 1 team based off statistics that could come away the champion in the next season.