Are Madden Ratings Accurate Indicators of Player Talent?

Owen Galinac

DS 325: Final Report

Introduction

Madden is a popular video game available for PC and consoles which allows players to control an NFL team in games against their opponents. The game advertises itself as an "immersive, simulation-based, authentic NFL interactive experience" and the "most realistic digital football experience gaming has to offer." Backing this claim are the extensive ratings Madden assigns all NFL players to try to replicate their skills and specialties. While this works fine for a video game, New York Jets owner Woody Johnson recently came under fire for blocking a trade for receiver Jerry Jeudy, due to a perceived low Madden rating of 81. I hypothesize that Madden ratings are not adequate indicators of player skill to be used to evaluate trades. With data from the 2021-2023 NFL seasons and Madden ratings from Madden 25 (released in 2024), I fit linear regression models to try to predict Madden ratings based on player performance. Additionally, I used Madden ratings to try to predict EPA (expected points added) for offensive players, which I used as an indicator of success in the league. I found that while player stats can somewhat reliably predict Madden ratings, Madden ratings are mostly unable to successfully predict player EPA, providing strong support that they should not be used when making trading decisions.

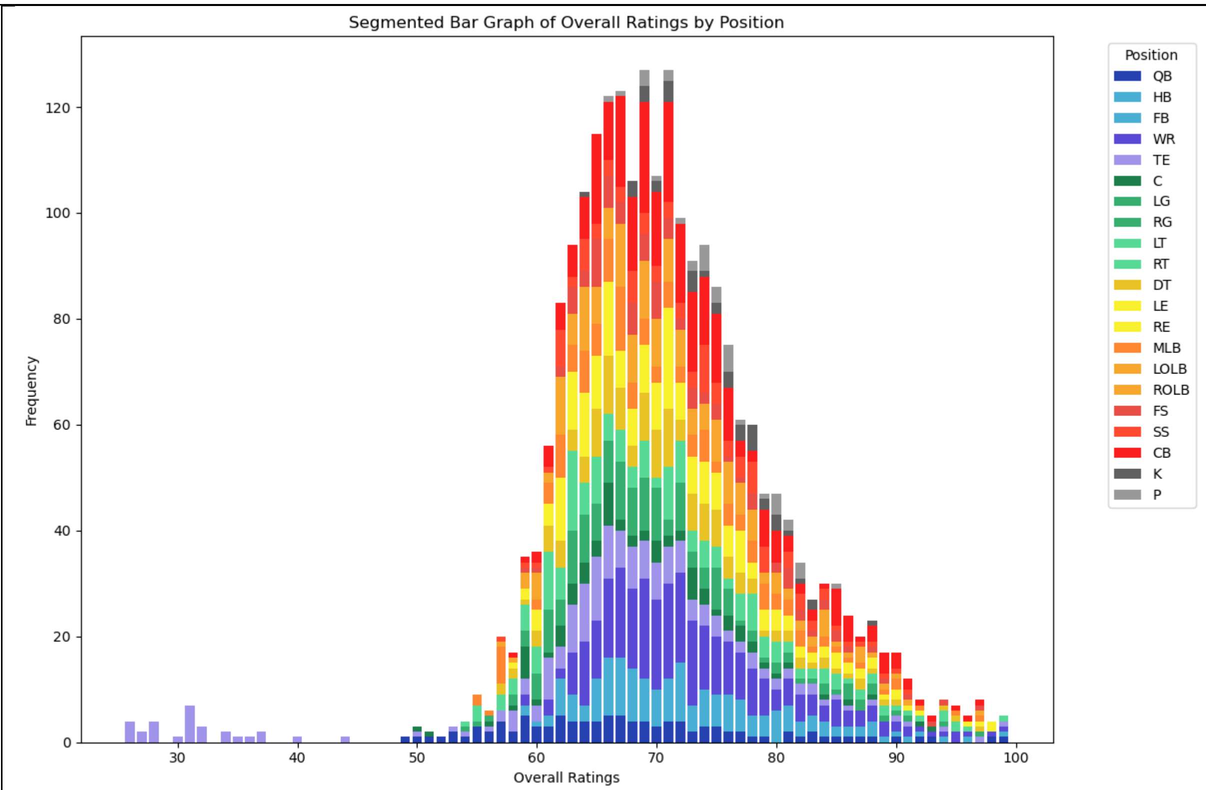

Figure 1. Distribution of Madden ratings across each position in Madden 25. Color coded by position and position group.

Methods

Datasets

Two datasets containing player stats, one for offensive players and one for defensive players. One dataset containing Madden ratings for players in Madden 25.

Data cleaning/preprocessing

Names of players with the same name were changed manually in the CSV files (ex. Josh Allen -> Josh Hines-Allen). Dataframes were filtered to just the 2021-2023 seasons, summing or averaging stats based on what made more sense (ex. sum total yards, average EPA). The Madden rating dataframe was merged with the stats dataframe by player name.

Dataframes were separated into individual dataframes for each position group (QB, RB, WR/TE, DB, DL, LB). This way the model was able to compare players with relatively similar roles and archetypes, rather than all players in the sport, who vary wildly in talents and statistics.

Models used

Linear regression models with Lasso regularization (to reduce overfitting and collinearity) were used to predict Madden ratings in part one, and EPA for offensive players in part two. This model was picked because it predicts a number, rather than classifies each case. Linear regression models were fit to predict Madden ratings for the QB, RB, WR/TE, DB, DL, and LB groups. Features that were relevant to each position were used to predict Madden ratings.

The general approach for each regression is outlined below:

1. Created features and targets, using most features from their respective dataframes.
2. Split the data into training data and testing data, with a test size of 0.2.
3. Standardized the features using StandardScaler.
4. Fit the Lasso model using the scaled training data.
5. Made predictions using the scaled training and testing data.
6. Assessed the models by calculating R^2 and MSE.
7. Plotted the models, placing the actual value on the X-axis and the predicted value on the Y-axis.
8. Plotted the residuals to see the relative difference in residuals between the different positions.

Results

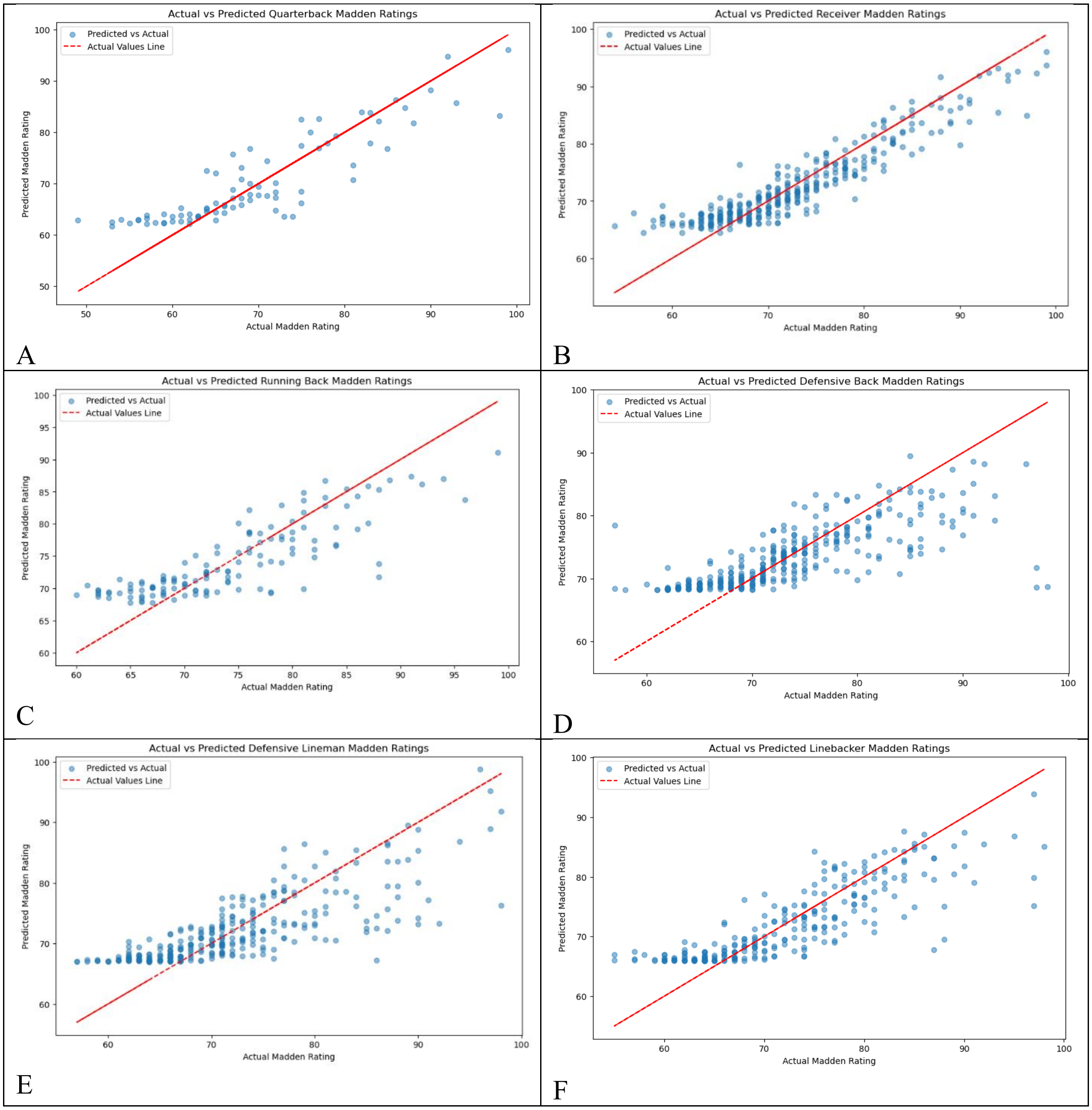

Figure 2. Linear regression lines generated for the different position groups, predicting their Madden ratings from their performances in the 2021-2023 NFL seasons. 1A: Quarterback, R^2 = .70, MSE = 27.92 1B: Running back, R^2 = .69, MSE = 21.27. 1C: Receiver/Tight end, R^2 = .85, MSE = 11.53 1D: Defensive back, R^2 = .44, MSE = 28.88 1E: Defensive linemen, R^2 = 0.69, MSE = 27.76. 1F Linebacker, R^2 = .64, MSE = 24.61.

Table 1. R^2 and MSE for regression models predicting NFL players' Madden ratings per position group

Position	R^2	MSE
Quarterback	0.70	27.92
Running back	0.69	21.27
Wide receiver/tight end	0.85	11.53
Defensive back	0.44	28.88
Defensive linemen	0.69	27.76
Linebacker	0.64	24.61

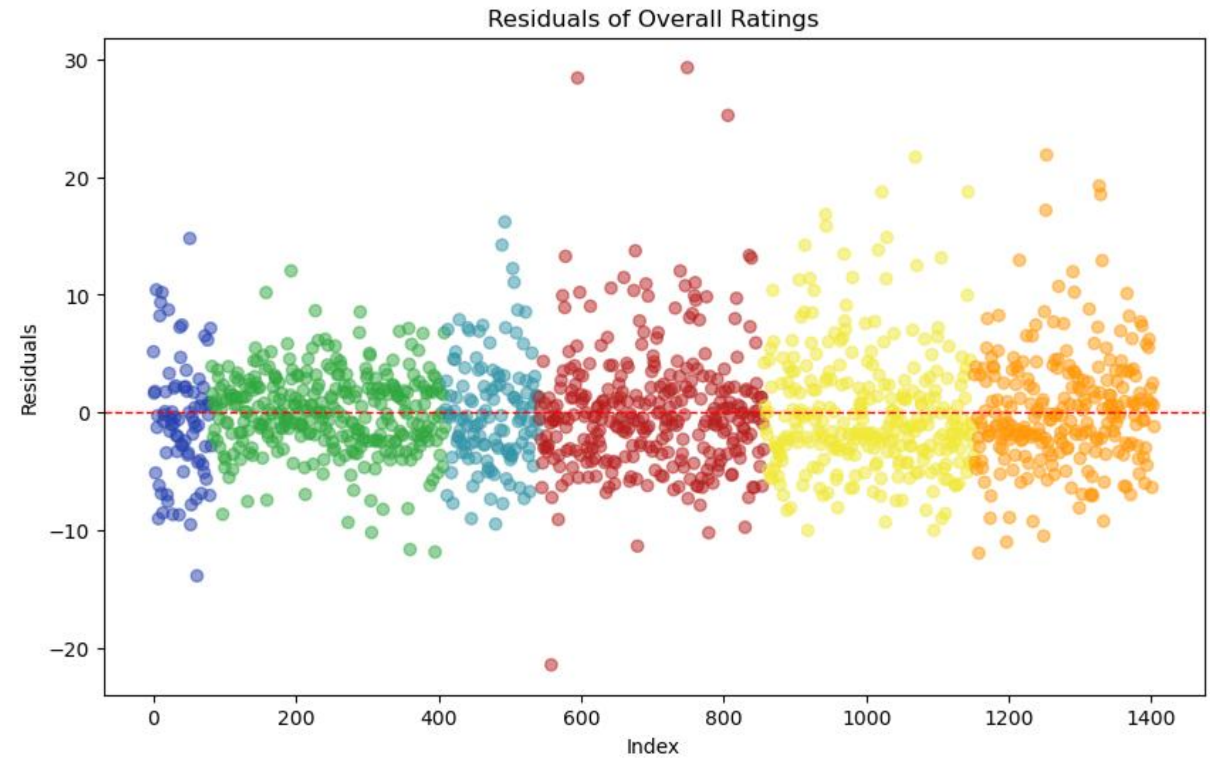

Figure 3. Residuals of Madden ratings for QB (blue), WR/TE (green), RB (light blue), DB (red), DL (yellow), and LB (orange).

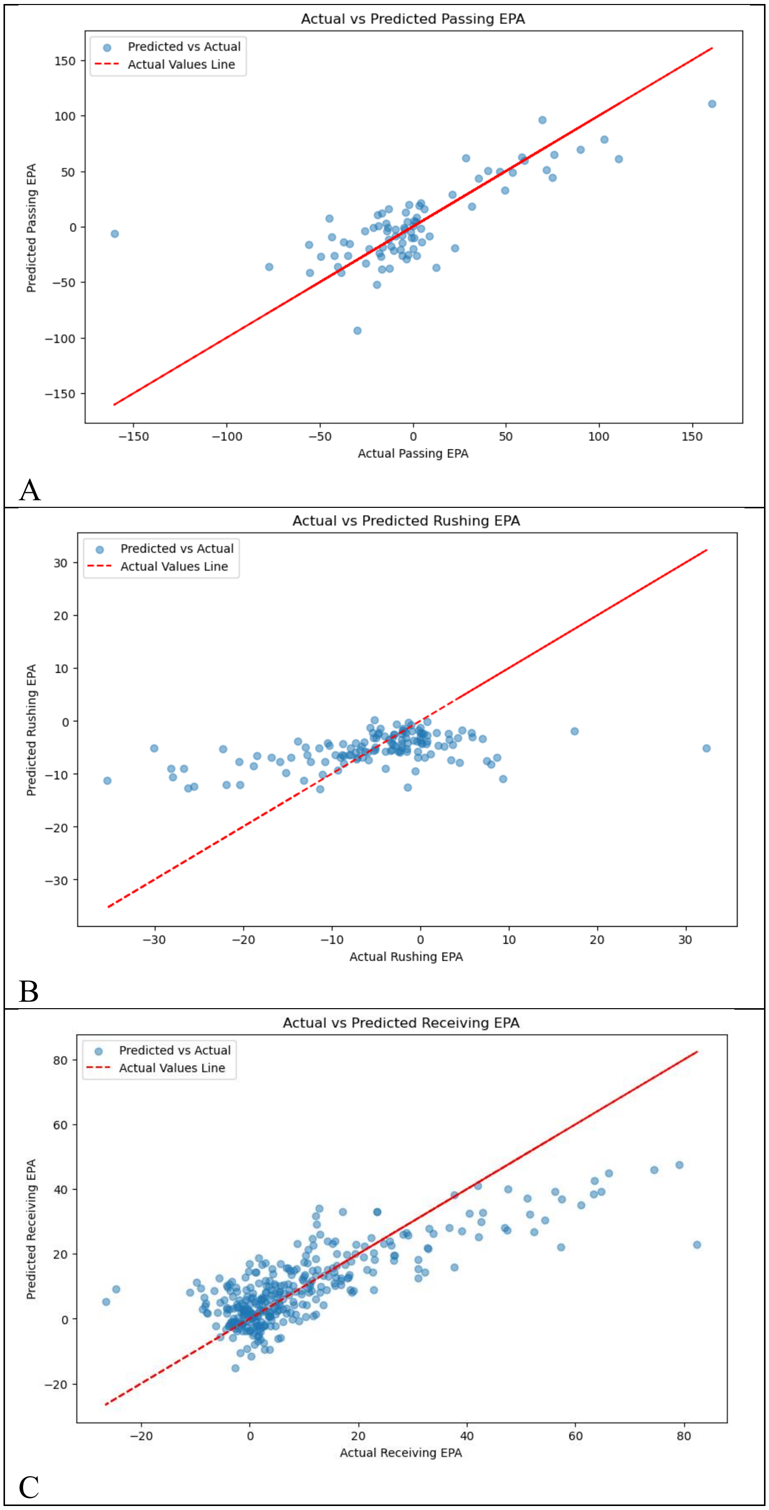

Figure 4. Linear regression lines predicting player EPA from Madden 25 ratings. 4A: Quarterback: $R^2 = 0.09$, MSE = 772.6. B: Running back, $R^2 = 0.19$, MSE = 66.21. 4C: Receivers, $R^2 = 0.60$, MSE = 100.26.

Table 2. R^2 and MSE for linear regression models predicting offensive player EPA from their Madden ratings.

Position	R^2	MSE
Quarterback	0.09	772.6
Running back	0.19	66.21
Wide receiver/tight end	0.60	100.26

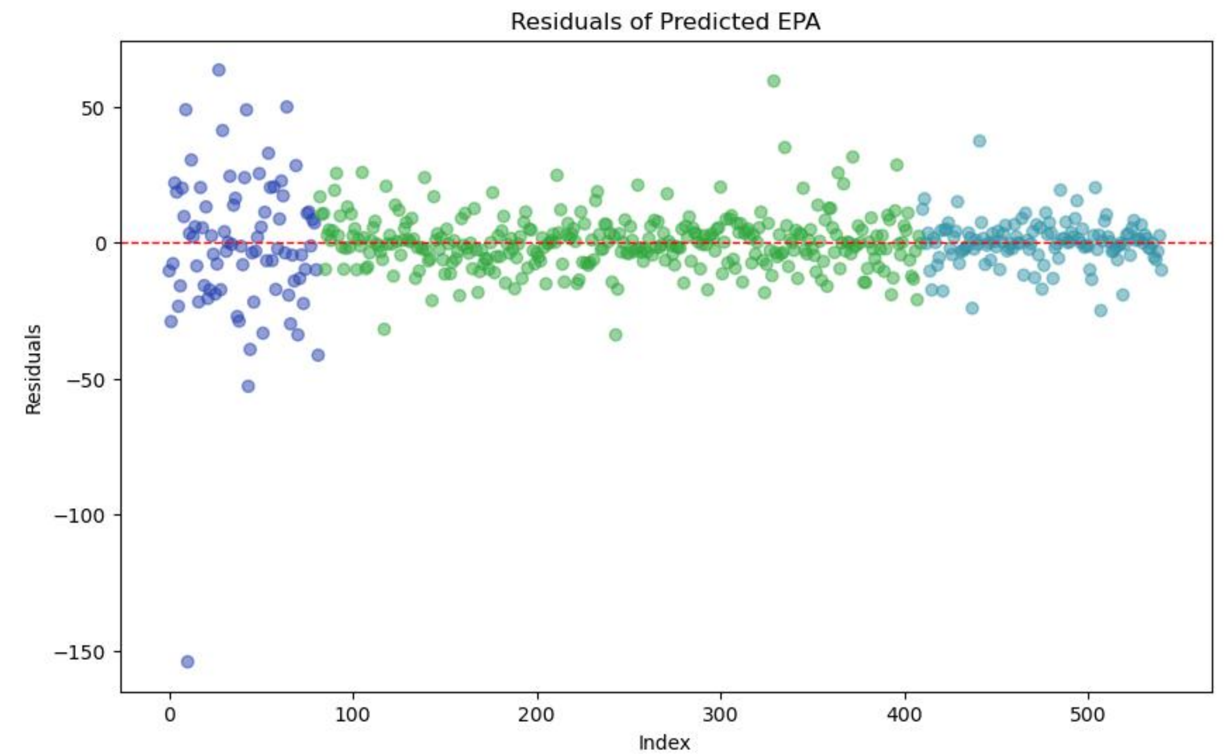

Figure 5. Residuals of predicted EPA for QB (blue), WR/TE (green), and RB (light blue).

The linear regression models predicting Madden ratings from game statistics produced six trendlines with varying R^2 values (figure 2, table 1). The model with the highest R^2 value was for wide receivers/tight ends, with an R^2 of 0.85 and MSE of 11.53. The model with the lowest R^2 value was for defensive backs, at 0.44 and MSE of 28.88. The other four models ranged from R^2 values of 0.6 to 0.7. The residuals of the Madden ratings were plotted to compare their relative sizes and see outliers in the data (figure 3). Visually, the wide receiver/tight end group appears to have the lowest residual spread, while all the other groups have large outliers and more widely spread residuals.

The linear regression models predicting player EPA from their Madden ratings produced three trendlines, two of which had R^2 values near 0 (figure 4, table 2). The quarterback model had an incredibly low R^2 of 0.09, while the running back model had an R^2 of 0.19. The receiving model was much more successful, with an R^2 of 0.60. The plotted residuals (figure 5) show some large outliers in the quarterback dataset, while wide receivers and running backs have much lower residuals and only a few stronger outliers.

Discussion

The results show that depending on the position, Madden ratings can serve as a decent indicator of success in the NFL. However, they should still not be used solely to justify trading for or not trading for a player.

The wide receiver/tight end models were the most successful at predicting both Madden ratings from statistics and EPA from Madden ratings. This suggests that Madden's receiver ratings are the most accurate compared to their other positions. However, their performance with the other positions leaves much to be desired. Statistics largely were able to predict Madden ratings with some success, suggesting that there is at least a decent connection between the ratings Madden assigns its players and the numbers they put up in game. The one exception was defensive backs, with a lower R^2 value. This, however, could be a fault of the modeling process, as I lumped together cornerbacks and safeties, two fairly different positions.

Conversely, Madden ratings were mostly unable to predict the EPA of both quarterbacks and running backs. The R^2 values near 0 reveal the model was unable to predict EPA with any measure of success. Part of this is likely due to EPA not being the best metric of player success. However, there was not a metric readily available that could better represent success for multiple positions. I could have used QBR or passer rating for quarterbacks, but these too have been much-maligned indicators of player success and are only available for quarterbacks.

While some models were relatively successful in their predictions, there is still far too much uncertainty for Madden ratings to be used as sole evaluations of player talent. Many factors go into evaluating the success a player may have on a team which a single number cannot evaluate, such as their fit into a scheme, chemistry with other players, and potential.

This study was not without room for improvement. A future study could be more focused in its targets, attempting to predict only quarterback Madden ratings using next-gen stats passing stats, which are more specific and focused than pure game numbers. This could possibly improve the model, which could show a better relationship between skill and Madden ratings.

Bibliography

No AI was used while typing this report, but Github Copilot was used in the coding process. It is noted in the markdown where Copilot was used.

Conahan, J. (2024, December 19). *Jets owner Woody Johnson cited "Madden" ratings in evaluating players: Report.* New York Jets On SI. https://www.si.com/nfl/jets/news/jets-owner-woody-johnson-cited-madden-ratings-evaluating-players

EA. *Madden 25*. https://www.ea.com/games/madden-nfl

Github Copilot. (2025). *GPT-4o* [Large langauge model]. https://github.com/features/copilot

StorymyGalaxy4. (2024). *Madden 25 Ratings Spreadsheet*. Reddit. https://www.reddit.com/r/Madden/comments/1eir4tr/madden_25_ratings_spreadsheet/

Tanho63 (2023). *Player_stats_def_season.* NFLverse [Github repository]. https://github.com/nflverse/nflverse-data/releases?page=2

Tanho63 (2023). *Player_stats_season.* NFLverse [Github repository]. https://github.com/nflverse/nflverse-data/releases?page=2

Specific Copilot prompts used:

"Hello, I would like to take the predicted values from the linear regression and put them into a dataframe containing the player's names they correspond to and their actual values"

"I would like code to see the most common fullName in the dataframe"

"Plot a histogram of the overall ratings based on their positions, with their color coming from the color df""

"do the same thing but make it a segmented bar graph where the x value is every overall value"

"order the positions and associated hex colors in this order QB HB FB WR TE C LG RG LT RT DT LE RE MLB LOLB ROLB FS SS CB K P"

"positions = ['HB', 'QB', 'TE', 'LT', 'WR', 'RE', 'DT', 'MLB', 'FS', 'LE', 'CB', 'LOLB', 'RG', 'RT', 'C', 'SS', 'LG', 'ROLB', 'K', 'FB', 'P'] hex_colors = ['#49aed4', '#2540af', '#9f94e9', '#58da97', '#5a49d4', '#f9f02d', '#e9c325', '#d57215', '#ad231c', '#e9c325', '#f91d1d', '#f6a62d', '#35ae6f', '#58da97', '#1c7e4b', '#e84e46', '#35ae6f', '#f6a62d', '#5f5f5f', '#49aed4', '#999999'"

"just give me the reordered hex codes"

"reorder the bar plot so it is in that order"

"I would like to plot the residuals of the Overall rating, and color code the dots based on the HexColor value"