NFL Home & Away Performance Analysis

Lan Cao, Matt Carswell

Xingrui Huo, Alex Pattarini

Data Science & Analytics, Georgetown University

DSAN 5100: Probability Modeling/Statistics Computing

Group 17 Final Project Report

Professor James Hickman

Dec 6, 2023

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Objective

In the dynamic and competitive realm of professional sports, the concept of "home-field advantage" has frequently sparked debate and called for analyses. This phenomenon is particularly pertinent in the National Football League (NFL), where teams are theorized to perform better in games the team plays in their home arena compared to when they travel to their opponent's. This project is motivated by the desire to unravel the intricacies of team and individual performances in varying game environments. An intriguing piece of empirical evidence supporting this exploration is a visual representation of the home win percentage over time in the NFL. As illustrated in **Figure 1**, the home win percentage consistently hovers slightly above 50%, offering a quantifiable perspective to the discussion.

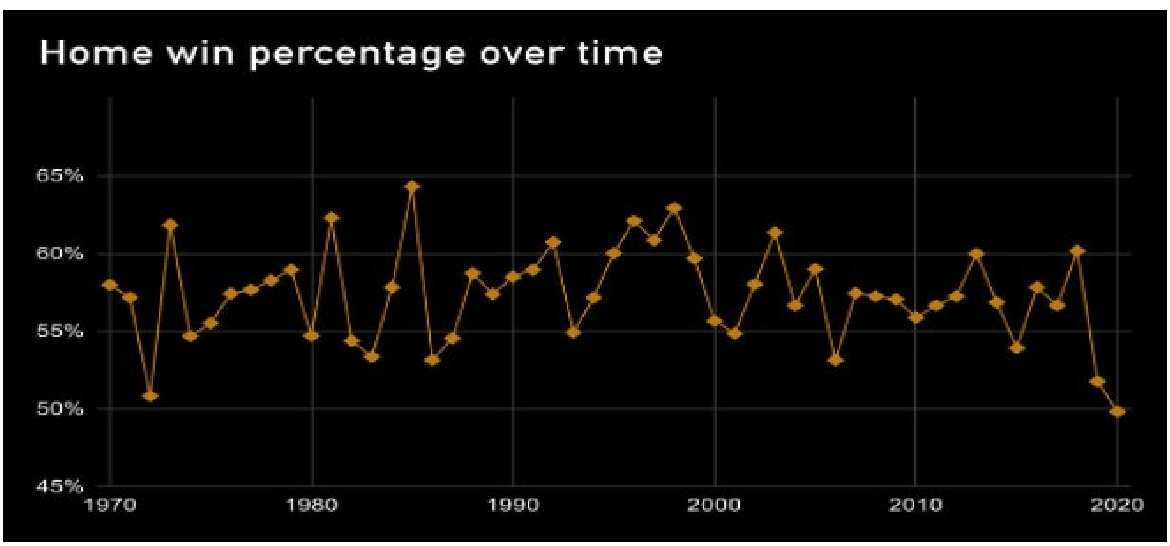

Figure 1: Historical NFL Home Win Percentage (Princiott)

Our primary research question revolves around a pivotal aspect of professional football: **does a home field advantage exist in the NFL and if so, to what extent?** Our analysis will focus on metrics that concern individual player performance and overall team performance, and it will employ only data collected during the **2022 NFL season**. The goal is to unravel patterns and correlations within the NFL, providing insights for strategic decision-making and enhancing our understanding of the role environmental factors play in sports performance.

1.2 Report Structure

This report will be addressed and analyzed from two perspectives: team and individual. From the team standpoint, our analysis will be three-pronged, encompassing points scored, Expected Points Added (EPA), and Win Probability (WP). These metrics will offer a comprehensive view of a team's performance, ranging from the basic yet crucial scoreline to the more intricate EPA, which reflects the contribution of plays to the team's scoring potential. Win Probability, on the other hand, offers a dynamic measure, capturing the likelihood of a team securing victory at any point during a game. This triad of metrics will enable us to assess whether playing at home or away significantly sways a team's chances of success.

From the individual standpoint, this report will shine a light on the most pivotal position in the NFL - the quarterback (QB). Quarterbacks, often recognized as the linchpins of their teams, can dramatically sway the course of a game. Each quarterback is often identified as the "face of the franchise (team)" and is generally regarded as a leader, so their performance is naturally very influential.

There are several potential reasons why QB performance may be impacted by being at home or traveling away. One of the most frequently considered reasons is that the home crowd tends to be extremely loud while the away QB is operating the offense. This leads to more difficult communication between offensive players, reducing the QB's ability to adjust the offense in real-time (Martin). Additionally, NFL QBs may play at a given away stadium once every few years, so there is likely some unfamiliarity each QB has with certain stadiums, possibly affecting their performance. By analyzing their performances through numerical statistics in the context of home and away games, we aim to ascertain if the game environment tangibly influences the effectiveness of quarterbacks on the field.

To ensure the rigor and validity of our findings, we will employ advanced statistical methodologies, including regression analysis and the Elo rating by FiveThirtyEight. FiveThirtyEight's NFL Elo ratings are a statistical system designed to estimate the relative strength of NFL teams (Paine, 2015). The Elo rating system, originally developed for chess, has been adapted for various sports, including the NFL. Regression analysis upon this Elo data will enable us to control various confounding variables, thereby isolating the effect of the 'home advantage', allowing us to quantify an exact 'edge' a homefield team might have over an away team.

Chapter 2: Data Overview

2.1 Section Overview

In this section, we describe the nature of the data used in this project and how it was used. This exploration of the nature of the NFL data itself will allow us to subsequently move into descriptions of more specific statistical methods we employed to answer our research question.

2.2 About the Data

For the first part of our analysis, we utilized two data sources for our project. The first data source used is the **nflreadr dataset**, and we focused on the most recent NFL regular season of 2022. This dataset is a treasure trove of detailed information, meticulously capturing the intricacies of the game through its expansive coverage. It offers a granular view into each play, presenting a play-by-play analysis that reveals the strategic and dynamic nature of football. Additionally, the dataset provides in-depth player information, encompassing statistics that span from basic biographical data to advanced performance metrics.

In our project, a significant portion of our analysis is derived from the play-by-play sub-dataset within nflreadr, specifically tailored for team performance analysis. In the play-by-play dataset, we approached our examination from three critical perspectives: score, Expected Points Added (EPA), and win probability. The score analysis allowed us to understand the direct impact of plays on the game's outcome, giving us insights into how teams strategically accumulated points. EPA is a sophisticated tool that measures the contribution of each play toward scoring, and the examination of win probability offered a dynamic perspective, showcasing how the likelihood of winning evolved throughout the game and how each play swayed this probability.

The other data source that was used is **FiveThirtyEight's NFL Elo data**. FiveThirtyEight's NFL Elo ratings data contains the Elo ratings of teams from every game going back to 1920, totaling over 17,000 observations. As previously mentioned, FiveThirtyEight's NFL Elo ratings are a system designed to estimate the relative strength of NFL teams over time. At the beginning of each NFL season, all teams are given an initial Elo rating. The initial ratings are typically based on the team's performance in the previous season and, to a lesser extent, historical performance. An Elo score of 1500 is considered to be the 'baseline.' The Elo system relies on head-to-head outcomes to update team ratings. If Team A defeats Team B, Team A's Elo rating is adjusted upward, and Team B's rating is adjusted downward. The magnitude of the adjustment depends on the difference in the teams' ratings and the margin of victory.

Most important to the Elo rating is what is known as the K-factor. The K-factor represents the weight or sensitivity of each game's impact on team ratings. FiveThirtyEight uses a dynamic K-factor, meaning it adjusts the K-factor based on the number of games played by a team. Early in the season, the K-factor is higher, allowing for more significant rating adjustments, while later in the season, it stabilizes. For more information on how Elo is derived and the history behind FiveThirtyEight's Elo rating, please check out FiveThirtyEight's website: https://fivethirtyeight.com/methodology/how-our-nfl-predictions-work/ .

For this project, we will be using FiveThirtyEight's 'nfl_elo.csv' file that can be found on their GitHub or on the following link: https://projects.fivethirtyeight.com/nfl-api/nfl_elo.csv . The dataset contains 33 variables. For the scope of this project (which is NOT to create a highly effective probabilistic model that "picks" winners highly successfully), we will just be interested in just a subset of these variables. Some of our variables of interest include:

- **season**: Year of season (we will just be looking at 2022)
- **neutral**: Whether a game was on a neutral site or not (we will be not be including games that are labeled as being played on a neutral site)
- **playoff**: Whether the game was a playoff game or not (we we not be including games that are labeled as a playoff game)
- **elo1_pre**: Home team's Elo rating before the game
- **elo2_pre**: Away team's Elo rating before the game
- **elo1_post**: Home team's Elo rating after the game
- **elo2_post**: Away team's Elo rating after the game
- **qbelo1_pre**: Home team's quarterback-adjusted Elo rating before the game
- **qbelo2_pre**: Away team's quarterback-adjusted Elo rating after the game

A full description of all variables in the dataset can be found in the aforementioned link.

Chapter 3: Statistical Methods

3.1 Section Overview

In this section we explore the potential for an NFL home-field advantage through the three lenses described prior: team performance, quarterback performance, and Elo rating. We explore a variety of statistical methods including regression analysis, hypothesis testing, and bootstrapping.

3.2 Team Performance Analysis

First, let's embark on some fundamental visualization to distill the vast datasets into comprehensible insights. Sports betting is an integral part of the NFL in the modern era, and oddsmakers assign a "points spread" to each game, describing the number of points either the home or away team is expected to lose/win by. Our first visualization, **Figure 2** captures a critical dimension of home-field advantage by comparing the average points spread of each team during the 2022 season, segregated into home and away statuses.

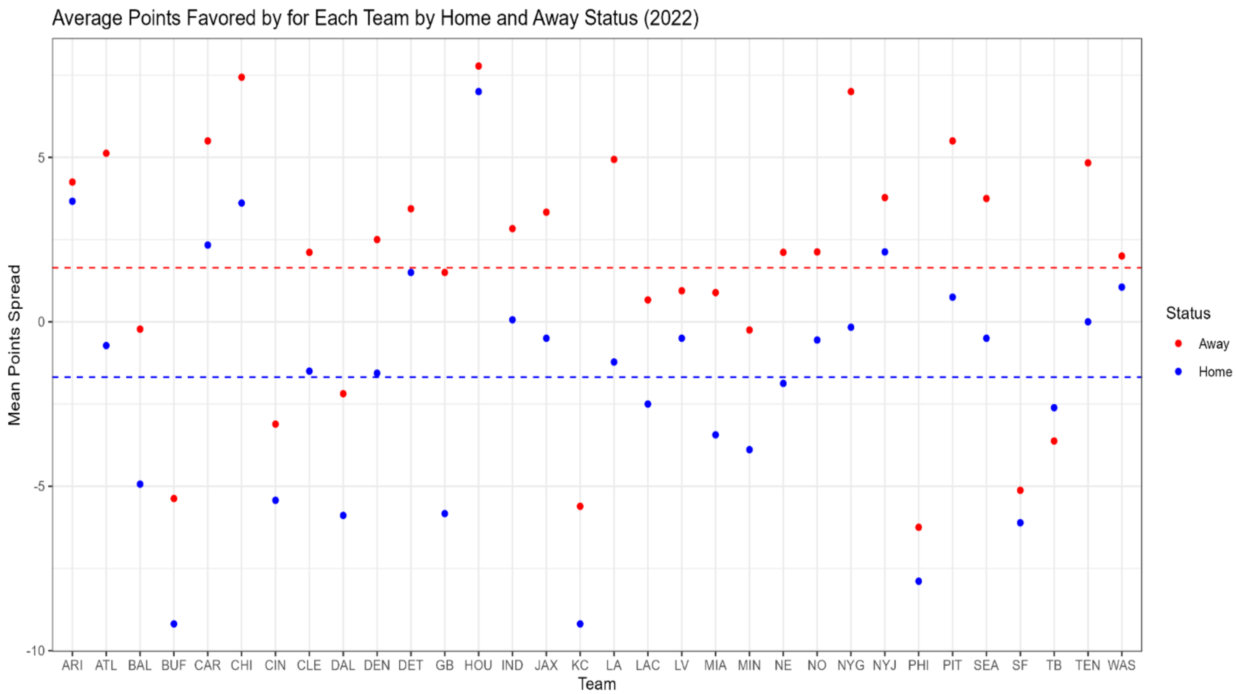

Figure 2: Points Spreads by Home and Away Teams

To initiate our exploration, each vertical line represents one NFL team. The y-axis measures the average point spread, indicating the extent to which a team is expected to win or lose. Positive values suggest that a team is predicted to lose by that number of points, while negative values indicate that a team is favored to win by the corresponding margin. Ideally, for a team demonstrating home-field advantage, we would observe these data points skewing towards the lower end of the axis. A cursory examination reveals a discernible pattern—nearly every home team has lower (more favorable) average point spreads, hence better scores, when they play at home.

We ran a Welch Two Sample t-test to analyze whether this disparity in point spread in home and away games was statistically significant. This test returned a statistically significant p-value of 0.0009553. **This strongly suggests that teams are significantly favored by a wider margin when playing at home**, a finding consistent with the observed pattern where home teams generally have lower, and thus more favorable, average point spreads. The negative mean value for home teams in comparison to the positive mean for away teams indicates that there is a *perceived* home-field advantage, at least with respect to oddsmakers.

Next, our analysis will probe deeper into the components of Score, Expected Points Added (EPA), and Win Probability to evaluate whether home-field advantage extends its influence to these aspects. To establish empirical evidence, we will construct hypothesis tests that quantify the effects. Our null hypothesis posits that there is no difference in the mean performance metrics of teams between home and away games. Conversely, the alternative hypothesis contends that there is a difference in the mean performance metrics of teams between home and away games. This hypothesis format will be applied to all three of the metrics of interest we will describe subsequently. These hypotheses are summarized below in **Table 1**.

Hypothesis	**Explanation**
$H_0: \bar{X}_h = \bar{X}_a$	There is no difference in the performance of teams between home and away games
$H_A: \bar{X}_h > \bar{X}_a$	There is a difference in the performance of teams between home and away games.

Table 1: Team Performance at Home vs. Away Hypotheses

Before delving into the intricate visual analyses that will illuminate our understanding of the data, it is crucial to first grasp the core metrics that will be employed: **Expected Points Added (EPA)** and **Win Probability (WP)**. Expected Points (EP) serve as the foundation, representing the anticipated average points a team is expected to score from a given field position during a drive. EPA enhances this concept by quantifying the contribution of each individual play to the chances of a team scoring. When a play results in a positive EPA, it indicates an uptick in scoring potential; conversely, a negative EPA reflects a reduction in those expected points. Similarly, Win Probability provides a percentage-based forecast of a team's likelihood of victory at any given moment within the game. This metric adapts fluidly, reflecting the game's changing momentum, with 0% implying no chance of victory, and 100% signifying a guaranteed win. Both metrics are pivotal—they lay the groundwork for the detailed statistical analysis to follow, aiming to shed light on the potential existence and possible strategic intricacies of home-field advantage in the NFL.

Within the broader context of team performance analysis, the actual score of the games remains a fundamental and direct success measure. It encapsulates the culmination of strategic plays, reflecting the efficacy of both offensive and defensive efforts. Scores and their margins demonstrate the totality of the competitive dynamics at play, offering a transparent and quantifiable gauge of a team's dominance. We commence our quantitative exploration by examining the differences in team scores. This exploration can be seen below in **Figure 3** where the y-axis represents the divergence in

points scored at home compared to away games. A positive value on this axis indicates a team's heightened scoring at home, while a negative value denotes greater scoring away from home. **Empirical observations reveal that the majority of teams indeed score higher in home games.**

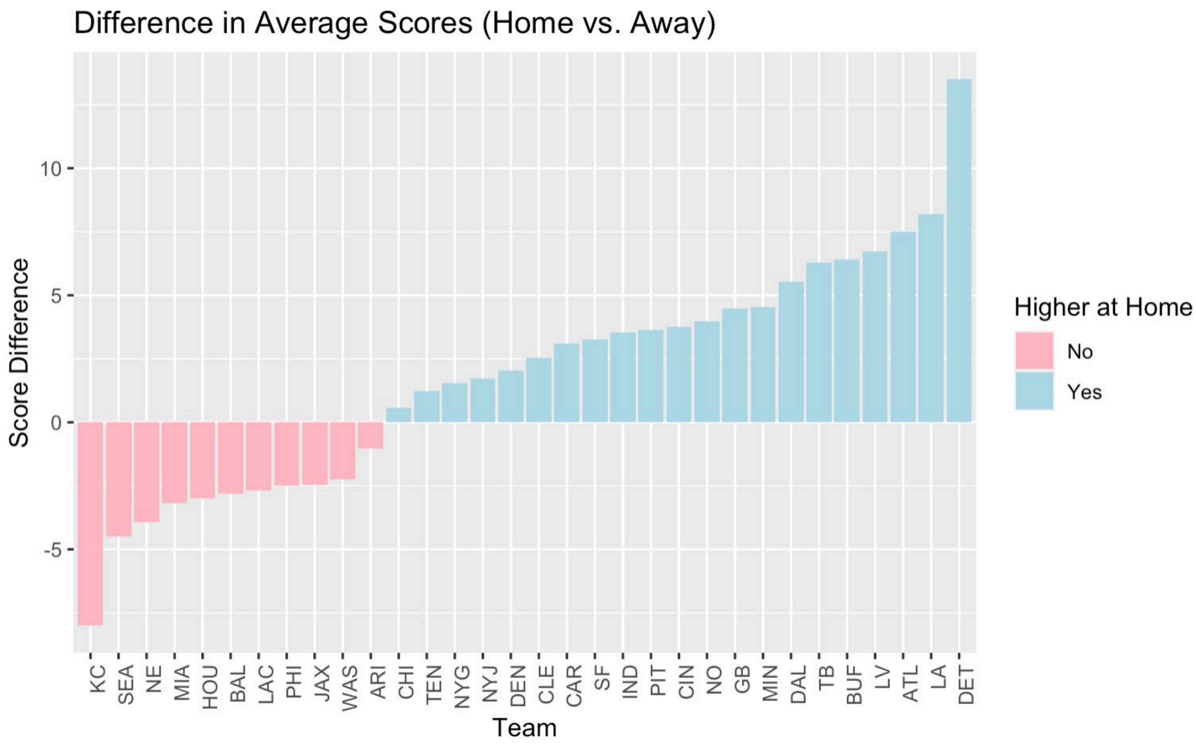

Figure 3: Score Difference for Each Team (Homes vs. Away)

To probe deeper into a team's performance, however, one must consider the richer narratives provided by EPA and WP. EPA offers a play-by-play dissection, which is indicative of the quality and impact of each action beyond the final score. The analytical exploration of Expected Points Added (EPA) yields a robust understanding of how team performance is influenced by the game's location. This exploration is conveyed below in **Figure 4.**

The visualization at hand delineates the average EPA for teams in both home and away contexts. The plot bifurcates into two segments: one where a positive EPA is more desirable and another where a negative EPA is observed. In an ideal scenario for the positive segment, teams would exhibit a higher light-blue bar (home EPA) compared to the light-pink (away EPA), signaling a superior performance in familiar environments. Conversely, in the negative segment of the plot, a shorter blue bar relative to the red would be preferable, indicating a lesser detraction in home games compared to away. At first glance, the graphical representation provides an overarching view, but the intricacies of team performance relative to EPA necessitate a more numerical analysis to justify any conclusions. A similar sentiment is echoed in **Figure 5,** portraying the average win probabilities by team in home and away games.

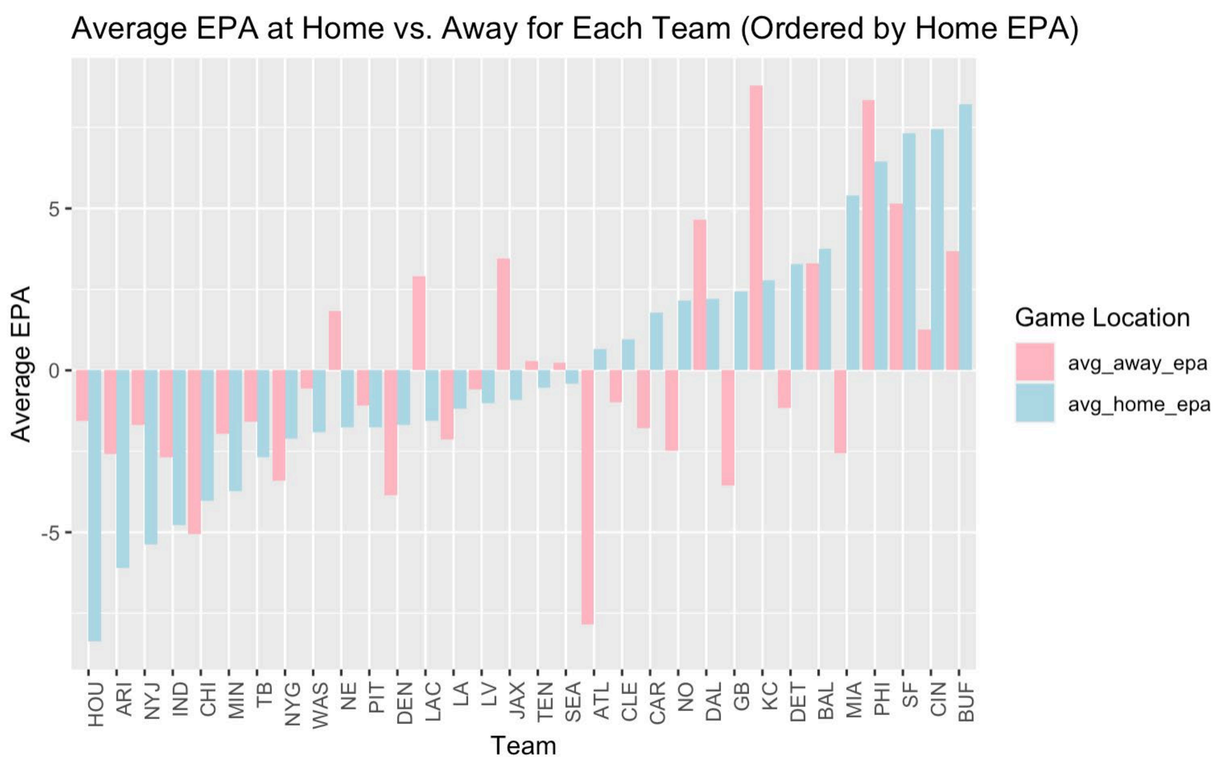

Figure 4: Average EPA at Home vs. Away for Each Team

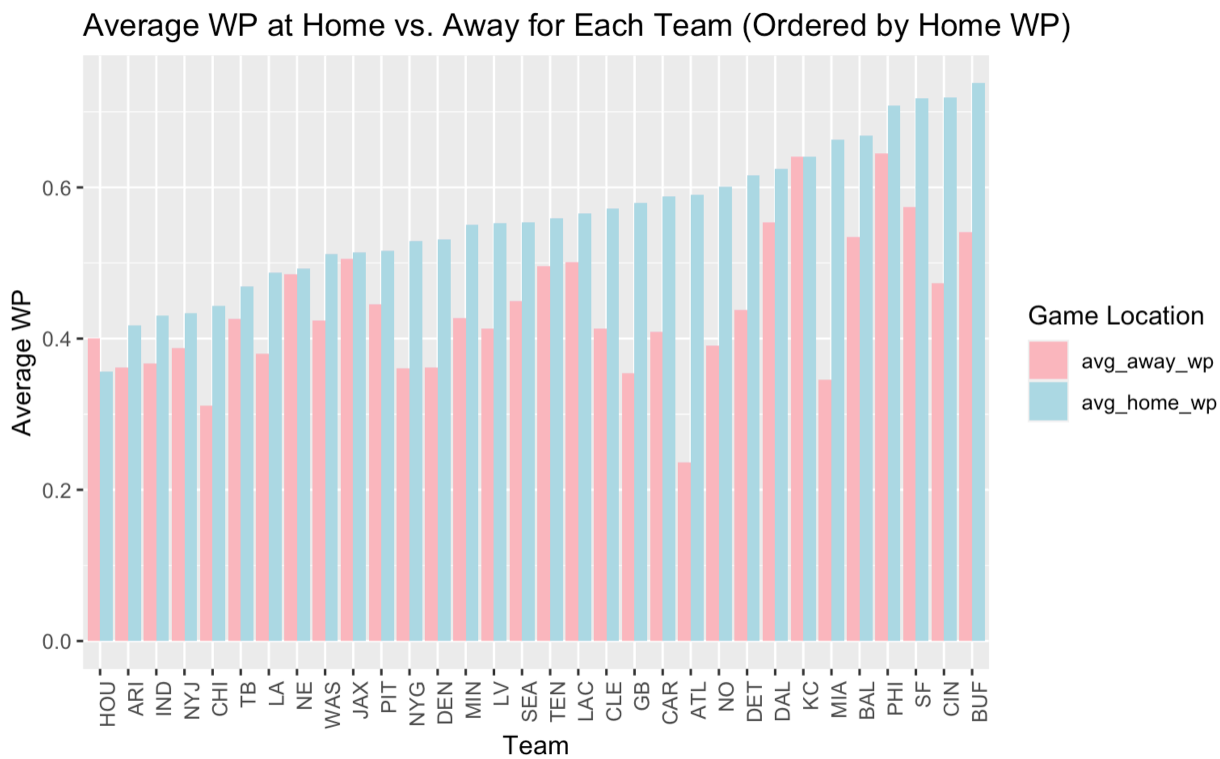

Figure 5: Win Probability for Each Team(Homes vs. Away)

WP captures the ebb and flow of a contest, reflecting how certain plays can swing the odds of victory. It encapsulates not just the actions taken but also the situational pressures of time and score. WP's real-time narrative allows teams to gauge the impact of their strategies within the evolving context of the game. This forward-looking metric complements the historical insights provided by EPA and they all together form an effective set of potential indicators of home-field advantage that thoroughly encapsulate the complex nature of the game of football at the NFL level.

3.3 Quarterback Analysis

To investigate and analyze the potential effects of the purported "home field advantage," we employed a two-sample hypothesis test regarding the average EPA generated by quarterbacks in home and away games. The "players" dataset (subsetted to the 2022 season) contained within the nflverse repository contains the game-wise statistics of each player that took part in any plays in the season. We then subsetted this data set to quarterbacks only and determined which statistics were collected in a home or away game for each QB. Various measures for each QB were calculated for all of their home and away games, including mean passing yardage and EPA per home/away game and the number of passing attempts. Subsequently, any quarterbacks whose mean passing yardage was less than 50 were removed to exclude backup quarterbacks who saw extremely limited play and may have unusual statistics that interfere with the overall analysis.

Using this data set, we performed an exploratory analysis to identify any potential trends/patterns to be investigated further. In **Figure 6** below, a scatter plot of the mean quarterback passing yardage and offensive EPA is depicted and stratified by home and away games. Each quarterback is represented by two bubbles, one for their home averages and one for their away averages, with their size determined by the total number of passing attempts they had in the 2022 season.

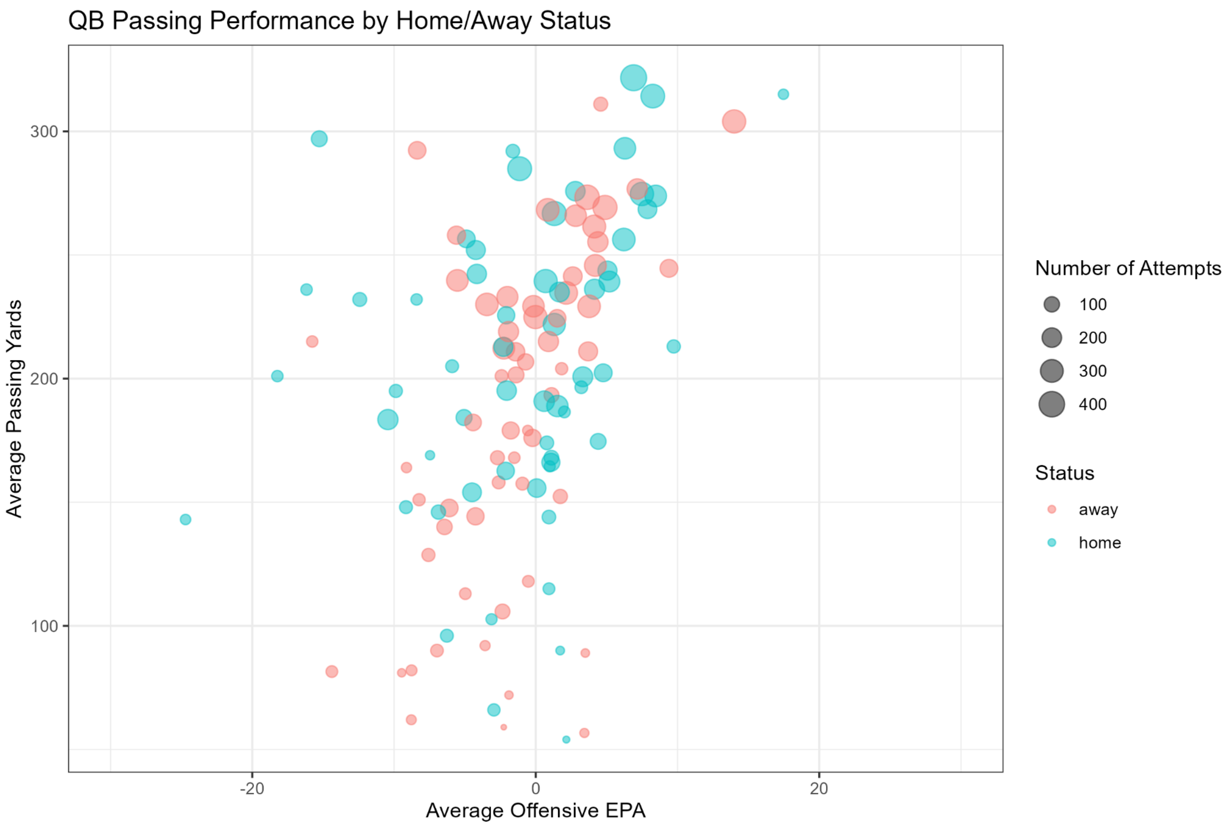

Figure 6: Quarterback Passing Yardage and Offensive EPA by Home and Away Performance

In this plot above, there is not a very easily discernible overall difference between the general performance of home and away quarterbacks by passing yardage and EPA, although the poorest passing yardage outputs generally occur during away games. We ultimately decided to focus on EPA as the primary, more cumulative measure of QB performance as the effect a QB has on a play is not purely just the passing yardage they created. EPA is a more comprehensive measurement of the QB's contribution to a given play. Thus, the data were subsetted further to only include the EPA of home and away games, which ultimately returns a data set where each QB has their own row with a mean EPA at home and away games, and a sample of this data set is shown in **Table 2.**

EPA by QB in Home/Away Games

Player	Away EPA	Home EPA	EPA Difference
M.White	-8.37	17.47	-25.84
S.Darnold	-2.71	9.73	-12.44
P.Walker	-7.57	0.93	-8.50
S.Ehlinger	-14.39	-5.90	-8.49
M.Ryan	-5.53	2.79	-8.33
T.Brady	-0.03	6.90	-6.93
R.Tannehill	-1.77	5.05	-6.82
K.Murray	-5.60	0.59	-6.19
T.Huntley	-4.97	0.93	-5.90
A.Dalton	-1.43	3.32	-4.74

Table 2: QB EPA Dataset

Using these data, we can perform further exploratory analysis to ascertain whether there is some appreciable difference between home and away quarterback performance. **Figure 7** below depicts the difference in the mean EPA of QB's in home and away games, calculated as $(EPA_{away} - EPA_{home})$ to return how much lower a given QB's mean away performance is compared to their home performance. A more positive bar means a QB generally performs better in away games than home games in the 2022 season.

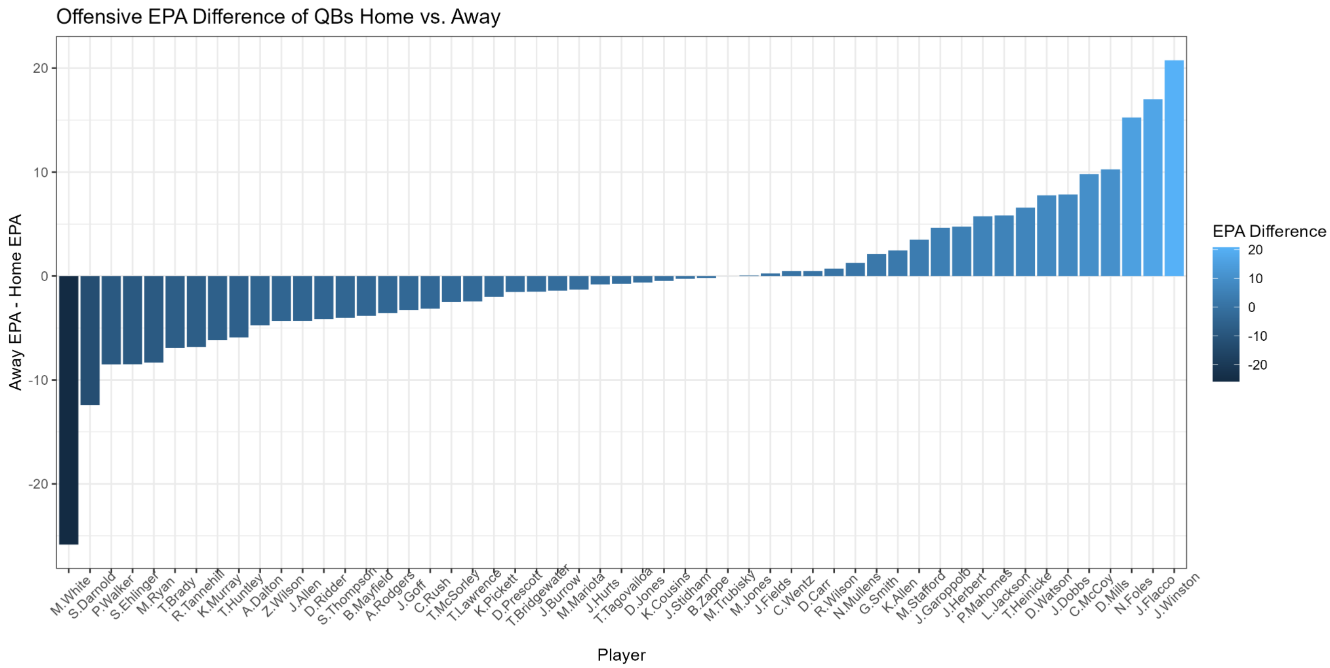

Figure 7: EPA Difference by Quarterback

Although there is certainly an appreciable number of quarterbacks that perform better in away games (i.e., those on the right), **most quarterbacks appear to perform better at home (as determined by EPA)**, with about two-thirds of the quarterbacks doing so. Thus, we continue to investigate EPA as an indicator of QB performance and potentially of a home-field advantage.

With this filtered and subsetted data set in hand, we applied a Welch two sample t-test for a difference in means to ascertain whether there is a statistically significant difference between the means of QB EPA in home and away games. A t-test is a method used to test for statistical significance when the population mean and standard deviation are unknown. A Welch two sample t-test evaluates whether there is a statistically significant difference in the means of two populations, in this case the QB performances in home and away games given unequal variances. To evaluate a t-test, there must be a null hypothesis (H_0) and an alternative hypothesis (H_A). The former consists of what the truth is assumed to be, and the latter is what is hypothesized to be the truth. In this analysis, we define the following hypotheses (let $\bar{X}$ be the mean QB EPA) in **Table 3.** The results of this t-test are presented and analyzed later in this paper.

Hypothesis	**Explanation**
$H_0: \bar{X}_h = \bar{X}_a$	There is no difference in the means of quarterback EPA in home and away games.
$H_a: \bar{X}_h > \bar{X}_a$	The mean EPA of quarterbacks in home games is greater than in away games.

Table 3: QB EPA Hypotheses

3.4 Elo Analysis

As enumerated earlier, our dive into FiveThirtyEight's Elo ratings is going to involve a regression analysis of how Elo differences between home and away teams can be used to explain point differential between the home and away team. Most importantly, **we want to know when we hold the Elo rating between a home and away team equal, is there a score differential in favor of the home or away team?** To answer this question, we will look at a linear regression of pre-game QB-adjusted Elo rating differential of a home and away football team in the 2022 NFL season on the score differential between the two teams.

It should be noted that the goal of this analysis is NOT to develop a perfectly predictive model for estimating score differential with pinpoint accuracy. Our goal is to create a parsimonious model that attempts to quantify the relationship between differences in how good teams are and how these differences affect expected outcomes. Creating a regression model for score differential will allow us to develop a specific quantity (with a confidence interval) for what the score differential between two teams will be with all else being equal.

To perform this analysis, once again, we will be looking strictly at pre-game QB-adjusted Elo. This variable was derived by FiveThirtyEight through the pre-game Elo and QB raw Elo variables. This variable takes into account both the importance of team ability and the historical play of the most consequential player on the field, the quarterback. FiveThirtyEight has found that adding weight to the QB Elo provides a more accurate Elo rating relative to the probability a team wins a given game. According to FiveThirtyEight, the "quarterback-adjusted Elo model incorporates news reports to project likely starters for every upcoming game" which accounts for factors like a backup quarterback being projected to start in a game, but being a worse player than the team's top starter (Paine, 2015).

Before we even create a model, we need to check our assumption of linearity. In other words, we need to check if there even is a linear relationship between QB-adjusted Elo and score differential that would justify even performing a linear regression analysis. **Figure 8** below was created to assess the potentiality of such a relationship.

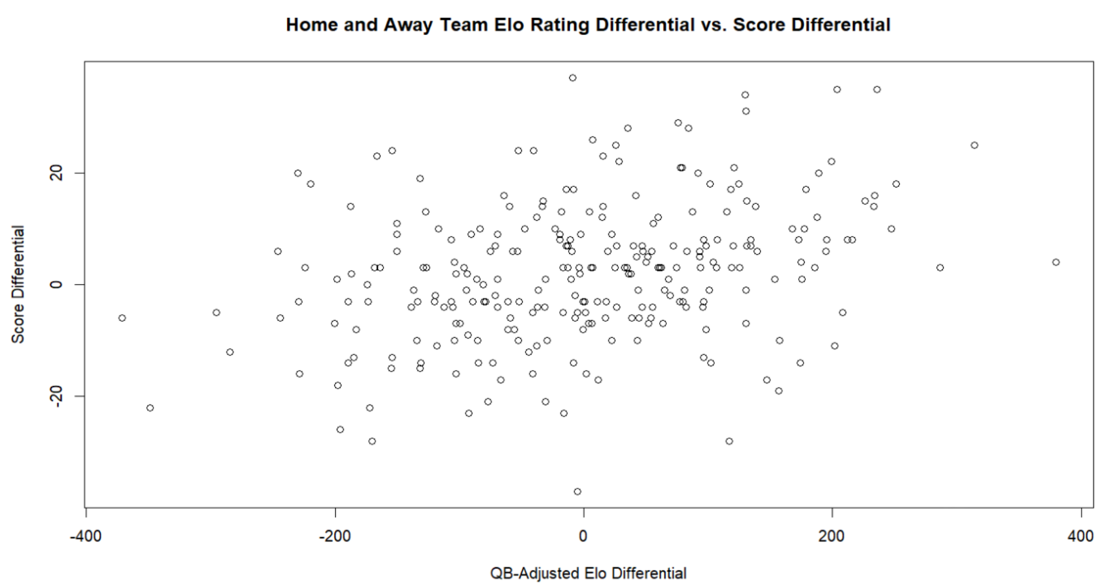

Figure 8: Elo Rating Differential vs. Score Differential

Looking at this plot, we can definitely see there is a presence of some linearity, at least enough to justify performing linear regression.

Now we can go ahead and create our parsimonious linear regression model. Again, if making a nuanced predictive model was of interest, doing a test-train split of the data might be of interest, but once again, we are just using this regression analysis to try and quantify the actual edge a home team might have, so no data split will take place. Our model will take the following form:

$$\text{Score Differential} = \beta_0 + \beta_1(\text{QB-Adjusted Elo Rating}) + \epsilon$$

Where: Beta_0 is the intercept, Beta_1 is the coefficient for the QB-Adjusted Elo Rating, epsilon represents the error term.

To interpret this model, we must first of all perform a couple hypothesis tests. First we will perform a global F test to determine if our model is statistically useful. It should be noted that because our model only has one predictor coefficient, such a hypothesis test is essentially the same as conducting an individual hypothesis test on the QB-Adjusted Elo Coefficient, beta_1. If the global F test leads to the rejection of the null hypothesis, it suggests that the model is statistically significant in explaining the variability in the response variable. The hypotheses for our global F test can be stated as follows:

Hypothesis	**Explanation**
$H_0 : \beta_1 = 0$	Our coefficient for our predictor, QB-Adjusted Elo Rating is equal to zero (there is no difference between an intercept-only model and a model with our predictor).
$H_a: \beta_1 \neq 0$	Our coefficient for the predictor QB-Adjusted Elo Rating is not equal to zero, indicating that the model as a whole is significant.

Table 4: Global F Test Hypotheses

We will also perform the important step of interpreting the significance of our intercept coefficient, beta_0 The hypotheses for testing the significance of the intercept coefficient can be stated as follows:

Hypothesis	Explanation
$H_0 : \beta_0 = 0$	The intercept coefficient is equal to zero, implying that the model intercept is not statistically different from zero.
$H_a: \beta_0 \neq 0$	The intercept coefficient is not equal to zero, indicating that there is a significant intercept in the model.

Table 5: Coefficient t-Test Hypotheses

It should be stated that before these hypotheses tests are conducted and interpretation of our linear model is conducted, we will have to check the assumptions of linear regression in order to ensure we are making statistically valid interpretations of our model. We will touch on these assumptions and subsequent diagnostics in the results section.

Chapter 4: Results

In this section we describe the results of the individual analyses, which includes hypothesis testing and regression analysis.

4.1 Team Analysis Results: Scoring

A rigorous statistical evaluation was conducted utilizing a t-test accompanied by a bootstrap analysis for the difference in mean scoring for home and away teams. The p-value obtained from the t-test stands at 2.2e-16, markedly below the conventional alpha threshold of 0.05, thereby denoting a high level of statistical significance.

Test Statistic	Degrees of Freedom	P-value	Confidence Interval	Bootstrap $\mathbf{2.5\% - 97.5\%}$
$t = 18.578$	95460	$2.2 * 10^{-16}$	(1.015,1.254)	1.013919 - 1.253325

Table 6: Team Score Hypothesis Testing Results

Consequently, we reject the null hypothesis, affirming that there is a statistically significant difference in the mean scores of teams at home versus away games. **The magnitude of this difference is quantified as falling between 1.015 and 1.255 points (within our level of confidence), as denoted by the 95% confidence interval**. This interval not only confirms the statistical significance of our findings but also suggests practical significance in the context of the game's scoring dynamic. The 95% confidence interval gives us a high degree of certainty that the true mean difference in scores between home and away contexts resides within this range.

Further, the bootstrap analysis and accompanying 95% bootstrap percentile interval corroborates these findings, proposing a very similar true mean score difference to the traditional analysis, ranging between approximately 1.016 and 1.253 points. The proximity of this range to the confidence interval derived from the t-test enhances the reliability of our initial analysis, providing a consistent narrative of the home-field advantage phenomenon as reflected in team scoring.

4.2 Team Analysis Results: Expected Points Added

Moving to EPA, we again applied a traditional two-sample t-test to analyze the difference in mean EPA by home and away team. The t-value of 2.8247 indicates a standardized difference in the mean EPA between home and away games, with a tilt towards the home scenario.

Test Statistic	Degrees of Freedom	P-value	Confidence Interval
$t = 2.8247$	95680	0.004734	(0.064,0.352)

Table 7: Team EPA Hypothesis Testing Results

This observation is accompanied by a p-value of 0.004734, which is below the conventional significance threshold of 0.05. This result suggests that we have evidence to reject the null hypothesis and can claim that **there is a statistically significant difference in the mean EPA of home and away teams**. The magnitude of this EPA difference, as denoted by the 95% confidence interval, ranges from approximately 0.0637 to 0.3523 expected points per play. This quantifiable difference could hold significant sway over tactical in-game decisions, pre-game preparations, and even broader considerations, such as fan involvement strategies and the calibration of betting odds. Analysts and team strategists might find this differential critical when evaluating past performances and in predicting future ones.

Ultimately, this section outlines how teams are likely to accrue more expected points when playing at home than away, a phenomenon substantiated by both visual and numerical analysis. The synthesis of these findings delivers a potent message: **there is some quantifiable divergence in home and away performances of NFL teams**, which warrants attention from all football stakeholders. Whether it is the fans passionately cheering in the stands or the analysts dissecting play strategies, the influence of the home setting ought to be considered and should be a factor in tactical decision making and in every cheer that reverberates through the stadium.

4.3 Team Analysis Results: Win Probability

A two-sample t-test comparing the home WP and away WP yields a t-value of 65.912, with a very large number of degrees of freedom (df = 95,680), indicating a vast dataset, and produces a p-value of less than 2.2e-16 as summarized in **Table 8** below.

Test Statistic	**Degrees of Freedom**	**P-value**	**Confidence Interval**	**Bootstrap 2.5% - 97.5%**
$t = 65.912$	95680	$2.2 * 10^{-16}$	$(0.115, 0.122)$	$(0.115, 0.122)$

Table 8: Team WP Hypothesis Testing Results

This p-value is substantially lower than the conventional threshold of 0.05 for rejecting the null hypothesis, signifying that the differences in mean WP between home and away games in the 2022 season are not merely by random chance but are statistically significant. The mean

WP for home games stands at 0.5596453, while for away games it is 0.4403547, **substantiating the notion that teams have a higher chance of winning when playing at their home venue**.

A 95% confidence interval obtained from the t-test ranges between 0.1157433 and 0.1228379, which does not include zero, further reinforcing the conclusion that home games are associated with an increased win probability. This finding is in line with the relatively preconceived notion of home-field advantage that many hold. To validate the t-test results, a bootstrap analysis was conducted, which provides an empirical approach to estimating the confidence interval by resampling the data. The bootstrap results yielded a 95% bootstrap percentile interval from 0.1157628 to 0.1229125, closely mirroring the interval from the t-test and lending additional credibility to the initial findings. The consistency between the t-test and bootstrap percentile intervals suggests a high level of confidence in the robustness of the observed home-field advantage.

In a broader context, the clear delineation between home and away WP has implications for various stakeholders within the NFL. For team management and coaching staff, these results can inform strategic decisions, suggesting a potential focus on leveraging home advantage to the fullest or developing strategies to mitigate the challenges of away games. For sports analysts and enthusiasts, this data provides a quantifiable measure of home advantage, contributing to richer, data-driven narratives about the game. For those in the sports betting industry and bettors themselves, the implications are direct, as WP can serve as a critical factor in odds-making and betting decisions.

In summary, the substantial sample size and the rigorous statistical approach employed in this team-based analysis firmly establish that the home advantage phenomenon is both real and quantifiable within the NFL**. Teams playing at home are more likely to have a higher WP than when they play away, and this effect is statistically significant across the league**. This statistical evidence can be harnessed by teams in strategic planning and by analysts in their assessments, ensuring that the home-field advantage is not just a matter of conjecture but a demonstrable factor in the success of NFL teams.

4.4 Quarterback Analysis Results

As outlined in the Statistical Methods section, we employed a Welch two sample t-test to ascertain whether there is any statistically significant difference between the mean expected points added (EPA) by quarterbacks in home and away games. At the $\alpha = 0.05$ significance level, the following statistics and one sided 95% confidence interval were computed as shown in **Table 9** below:

Test Statistic	Degrees of Freedom	P-value	Confidence Interval
$t = 0.19335$	94.271	0.4236	$(-1.873, \infty)$

Table 9: QB EPA Two Sample t-Test Results

From the results in **Table 9,** we can conclude that **we do not have evidence to reject the null hypothesis that there is no difference in the mean EPA for QBs in home and away game**s. The p-value of 0.4236 is greater than the $\alpha = 0.05$ significance level, so this result is not statistically significant. Additionally, 0 is contained within the 95% confidence interval, another indicator of the lack of statistical significance. This is further supported by a visualization of the distribution of QB EPA by home and away game, which are shown in **Figure 9a and 9b.**

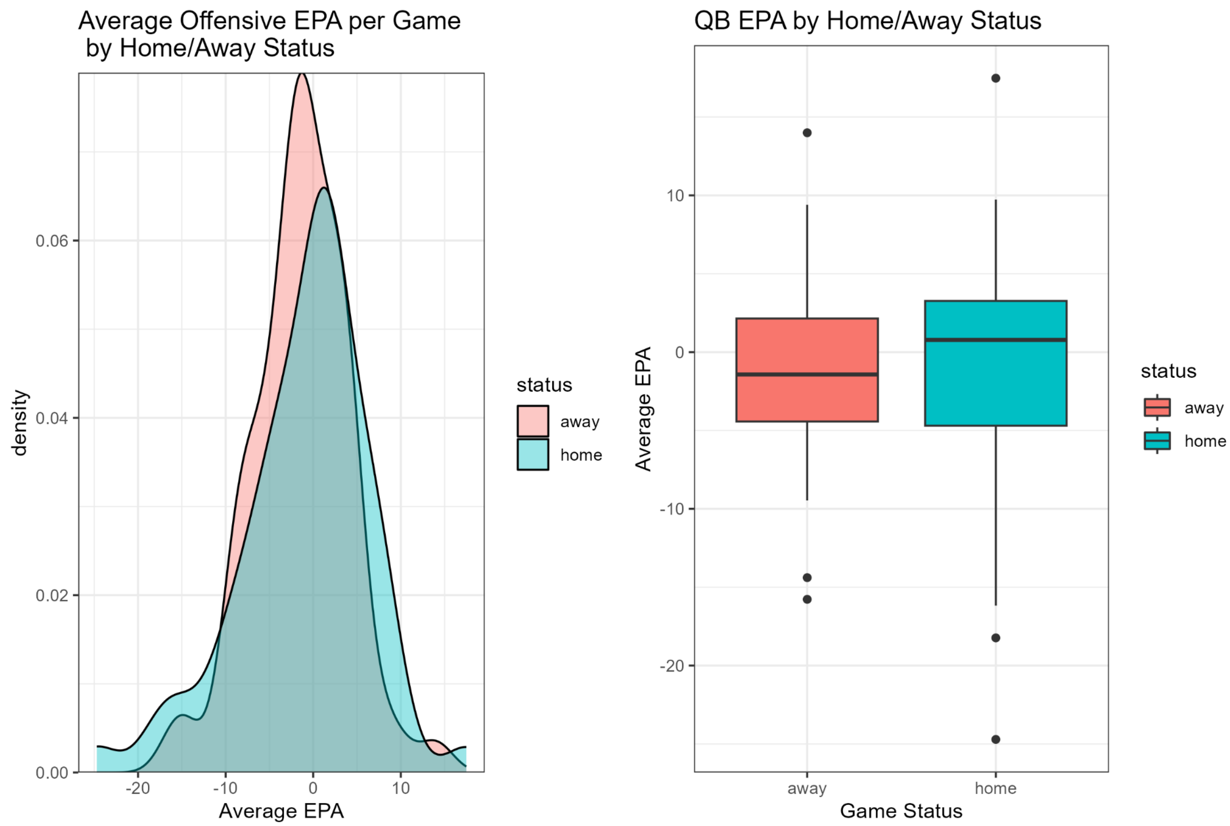

Figure 9a and 9b: Distribution of QB EPA by Game Status

While the mean and median EPA of quarterbacks are slightly higher in home games than away games for the 2022 season as shown in Figure X and Y, there is no evidence to suggest that this difference is statistically significant. There is significant overlap between these two distributions, and the application of the two sample t-test only quantifies and enunciates their similarities. **While there may be a demonstrative difference in overall home and away *team* performance as evidenced earlier in this paper, an analysis of the quarterback position yields an inverse result.**

4.5 Elo Results

As stated earlier, before we begin assessing our linear model and its coefficients that we created as a part of our analysis of FiveThirtyEight QB-Adjusted Elo ratings, we need to check the additional assumptions of linear regression to ensure our model is valid. We already checked the assumption of linearity, but now we need to assess if our residuals are independent, have constant variance (homoscedasticity), and are normally distributed. We also check if we have any outliers in our data that may be impacting our model. With a fitted model, we will first take a look at our residuals in the figure below.

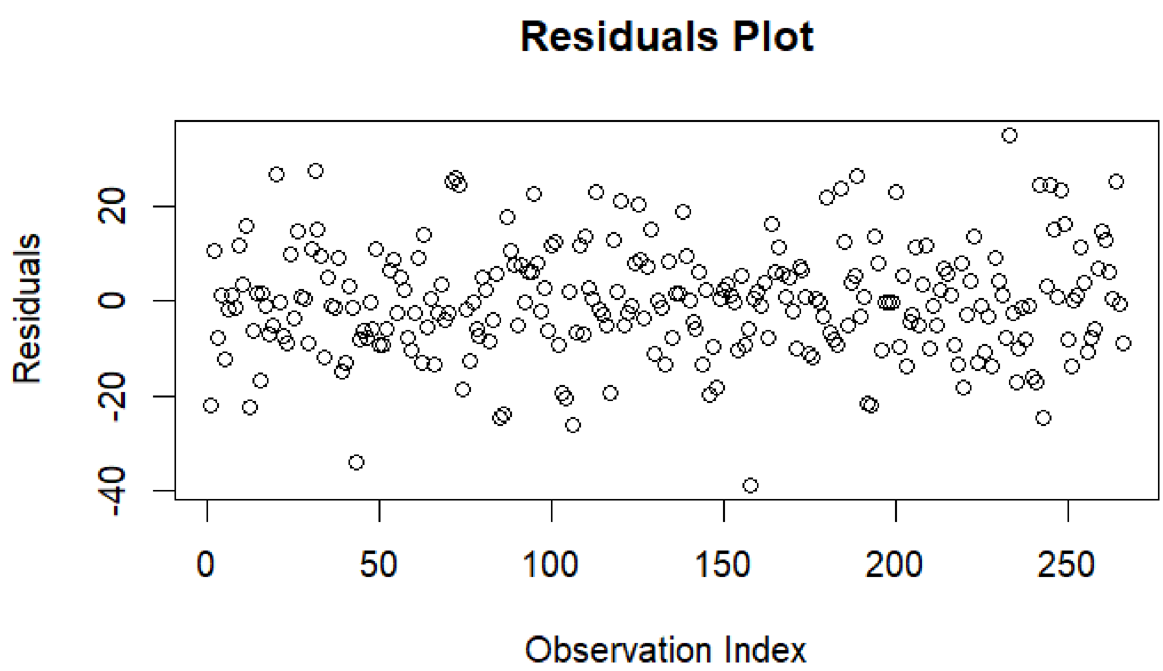

Figure 10: Residuals Plot

In this plot, we can see that our residuals appear to present homoscedasticity and be independently distributed with a mean of 0, important assumptions to linear regression. We will go ahead and take a look at some additional diagnostic plots below.

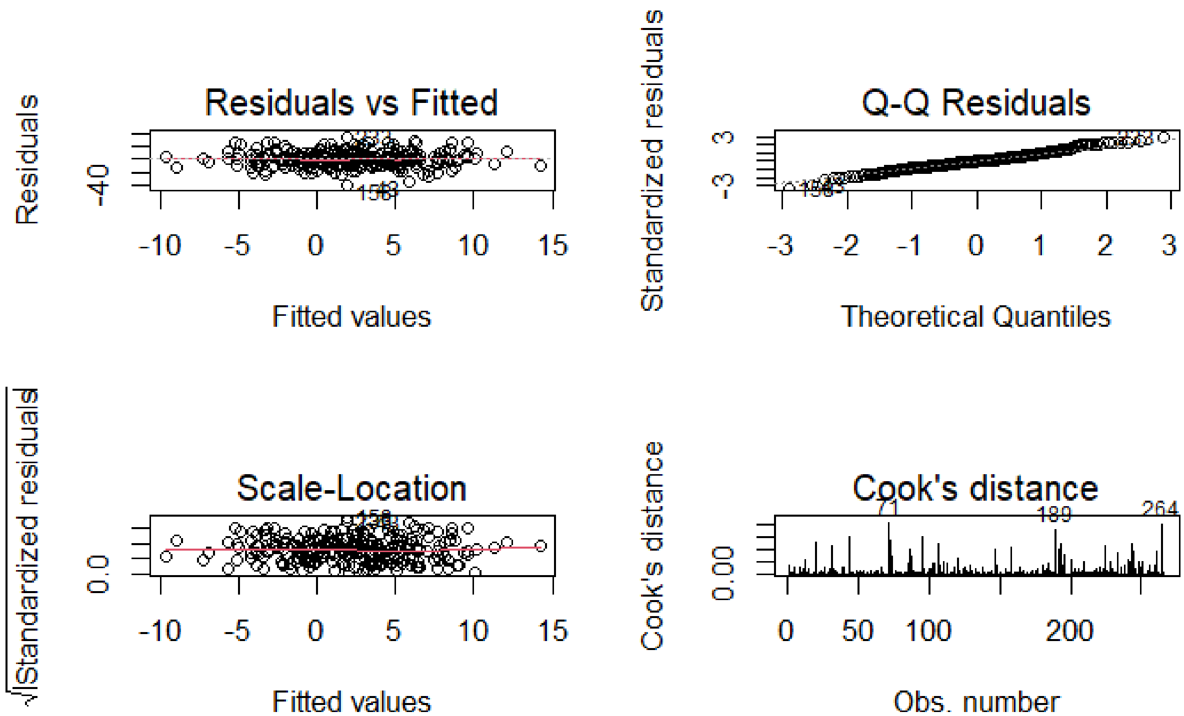

Figure 11a, 11b, 11c, and 11d: Diagnostic Plots

In the above plots, we can once again affirm all of our assumptions in linear regression. The Residuals vs Fitted plot once again highlights that there is not a pattern to our residuals, implying independence. Our Scale-Location plot helps us visualize more of the homoscedasticity of our residuals, helping us visualize the "slope" between the fitted values and residuals. Our slope is, all things considered, fairly constant, allowing us to affirm the validity of this assumption.

In our QQ plot, we can see that our residuals definitely follow a normal distribution which is an important assumption as we begin to delve into hypothesis testing and interpretations of confidence intervals. Our plot that displays Cook's distance tells us that we have outliers in observations 71, 189, and 264. Because our model is not heavily emphasizing prediction, we are going to go ahead and just leave these points in. Overall, all the assumptions of linear regression hold true for our model.

With the assumptions of linear regression validated, we can begin to interpret our model. A summary of our model can be found below.

```
Call:
lm(formula = score.diff ~ elo.adj.diff, data = nfl.clean)

Residuals:
    Min      1Q  Median      3Q     Max
-38.971  -7.617  -0.172   6.460  35.138

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.141523   0.715992   2.991  0.00304 **
elo.adj.diff 0.031821   0.005774   5.511 8.48e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 11.68 on 264 degrees of freedom
Multiple R-squared:  0.1032,    Adjusted R-squared:  0.09977
F-statistic: 30.37 on 1 and 264 DF,  p-value: 8.481e-08
```

Figure 12: Fitted Linear Regression Model Summary

We will first perform a global F test to determine if our model is statistically useful.

Test Statistic	**Degrees of Freedom**	**P-value**
$F = 30.37$	264	8.481e-08

Table 10: Global F Test Summary

With a p-value of 8.481e-08 and a significance level of 0.05, we can reject the null hypothesis and suggest that there is strong evidence to say that our coefficient associated with QB-Adjusted Elo Rating is not equal to 0 and our model is useful.

Next, we will perform the important step of interpreting the significance of our intercept coefficient. The hypothesis test results can be found below.

Test Statistic	Coefficient Estimate	P-value
t = 2.991	2.141523	0.00304

Table 11: : t-Test of the Intercept Coefficient Summary

With a p-value of 0.00304 and a significance level of 0.05, we reject the null hypothesis. We have strong evidence to suggest that the intercept coefficient, beta_0, is not equal to 0.

Now we begin to analyze what our model is telling us. First of all, our coefficient for Elo rating, 0.031821, tells us that **for every unit increase in Elo difference between the home team and away team, the score differential between the home and away team will increase by approximately 0.032**. This is interesting to see, but not really surprising that a higher Elo rating would lead to a higher score differential.

What is really of interest is our intercept coefficient. If our home and away team both have the exact same Elo rating, then our model tells us that the home team will still score 2.141 more points than the away team. **With everything between two teams being exactly equal,**

which we know is not even possible in the real world, the home team would still win by 2.141.

We can create a confidence interval of our intercept coefficient to further pin down this home "edge". Our 95% confidence interval for our intercept coefficient is (0.7317418, 3.551304). We can say that we are 95% confident that **when a home and away team have equal Elo ratings, the home team will score between approximately 0.73 and 3.55 more points than the away team.**

Chapter 5: Conclusion

The analysis presented in this paper suggests that the primary research question "does a home field advantage exist in the NFL and if so, to what extent?" has neither a simple nor absolute answer. Through both traditional and bootstrapped hypothesis testing of team-wise statistics, significant evidence for a home-field advantage was discovered across all three metrics of interest (scoring, EPA, WP). However, when focusing on individual players, particularly the most important position on the field, the quarterback, no significant evidence was found to suggest an advantage. While the overall performance of a team seems to improve at home compared to away (at least for the 2022 season), the individual quarterback performance does not seem to vary as much or as significantly.

In our Elo analysis, we found that there is a consistent advantage held by the home team, as reflected in the intercept coefficient in our linear regression model. This coefficient indicates that, under equal Elo ratings, the home team is anticipated to score 2.141 more points than the away team. The confidence interval for the intercept coefficient further refines this insight, suggesting a 95% confidence that, when teams have equal Elo ratings, the home team's scoring advantage falls between approximately 0.73 and 3.55 points. Overall, our validation process and subsequent interpretation shed light on the model's reliability and offer meaningful insights into the relationship between Elo ratings and score differentials in football games.

Ultimately, there is no clear answer to the question of whether there is a true home-field advantage in the NFL. While home winning percentage and yardage have generally been on the decline in recent years (Princiotti), there still appears to be some significant difference in overall team performance at home and when traveling away. The specific causes/factors that play into this difference are not obvious, and future research can and should focus on discovering these potential phenomena. Although the analysis presented in this paper is multifaceted and examines various potential indicators of home-field advantage, this can always be expanded upon. Future research could consider other measures of what a “good” performance is, focus on other individual positions/position groups, perform other analytical methods, among others. While there certainly is room for other analyses, our analyses suggest that the existence/absence of an NFL home field advantage is difficult to ascertain, and there is no straightforward or simple way to characterize it.

References

Baldwin, B., Carl. S., Edwards, J., Ho, T., and Sharpe, L. (2023). *nflverse: Easily Install and Load the 'nflverse'*. https://nflverse.nflverse.com/, https://github.com/nflverse/nflverse

Martin, C. (2023, December 4). *What is a silent count in football? (full explanation).* Football Advantage. https://footballadvantage.com/silent-count-football/

Paine, N. (2015, September 10). *NFL Elo ratings are back!*. FiveThirtyEight. https://fivethirtyeight.com/features/nfl-elo-ratings-are-back/

Princiotti, N. (2021, January 6). *What happened to NFL Home-field advantage?* The Ringer. https://www.theringer.com/nfl/2021/1/6/22216167/nfl-playoffs-home-field-advantage-covid-19-restrictions