# NFL Quarterback Performance Forecasting Report*

Aryaman Suri

March 29, 2024

## Table of contents



## 1 Introduction

It's critical to evaluate our team's play as we approach the halfway mark of the 2023 NFL season and make plans for the remaining games. These results examine passing Expected Points Added (EPA) and its consequences for the team's performance on the field, as determined by the team's designated analyst. EPA is a useful tool that sheds light on the efficacy of our passing strategy. We can improve our gameplay and gain additional victories in the coming weeks by making educated decisions based on our understanding of how each pass affects our team's overall performance. We'll examine the subtleties of EPA in this analysis, as well as how important it is to our team's strategy and plan of action using R (**citeR?**).

*Code and data are available at: https://github.com/aryamansuri/NFL-PassingEPA

## 2 Data

The data used for this analysis was obtained from the nflverse package ((**citeNfl?**)), specifically the 'raw_data\ dataset. This dataset contains statistics for NFL quarterbacks during the regular season for all the seasons.

The dataset was filtered to include only regular-season games from the 2023 season and quarterbacks. Additionally, it was divided into training and testing sets, with the training set containing data up to Week 9 and the testing set containing data from Weeks 10 to 18.

## 3 Model Selection

Several forecasting models were considered for predicting quarterback performance, including linear regression, logistic regression and Poisson regression. After experimentation, the linear regression model (**rstanarm?**) was selected for its ability to capture linear relationships between predictor variables and the target variable, passing EPA.

Creative feature engineering was crucial for improving the performance of the linear regression model. Features such as passing yards, passing touchdowns, interceptions, sacks were incorporated to capture the complex dynamics of NFL games.

To illustrate, let's consider the predictive model for NFL quarterback performance:

$$\text{passing\_epa}_i \sim \mathcal{N}(\mu_i, \sigma) \tag{1}$$

Here, passing_$epa_i$ represents the expected points added (EPA) through passing for quarterback$_i$. It follows a normal distribution with a mean $\mu_i$ determined by the quarterback's individual characteristics and a constant standard deviation $\sigma_i$.

The mean $\mu_i$ is a function of various predictor variables. These predictors encompass a diverse array of factors influencing quarterback performance, such as passing yards, touchdowns, sacks.

$\mu_i = \beta_0 + \beta_1 \cdot x_{i1} + \beta_2 \cdot x_{i2} + ... + \beta_p \cdot x_{ip}$

Here, $\beta_0$ represents the intercept term, while $\beta_1, \beta_2, ..., \beta_p$ denote the coefficients associated with each predictor variable. The model aims to estimate these coefficients to find the relationship between the predictors and the target variable, passing EPA.

Table 1: **?(caption)**

```
# A tibble: 6 x 3
  player_name  week predicted_epa
  <chr>       <dbl>         <dbl>
1 D.Prescott     12          21.5
2 L.Jackson      17          21.4
3 D.Prescott     10          19.7
4 B.Purdy        15          17.6
5 J.Goff         15          17.1
6 M.Stafford     13          16.5
```

## 4 Results

Before, we get into the results, (**graph1?**) ((**ggplot2?**)) is the relationship between the passing yards and the passing epa for the reigning league mvp, Patrick Mahomes for the season so far.

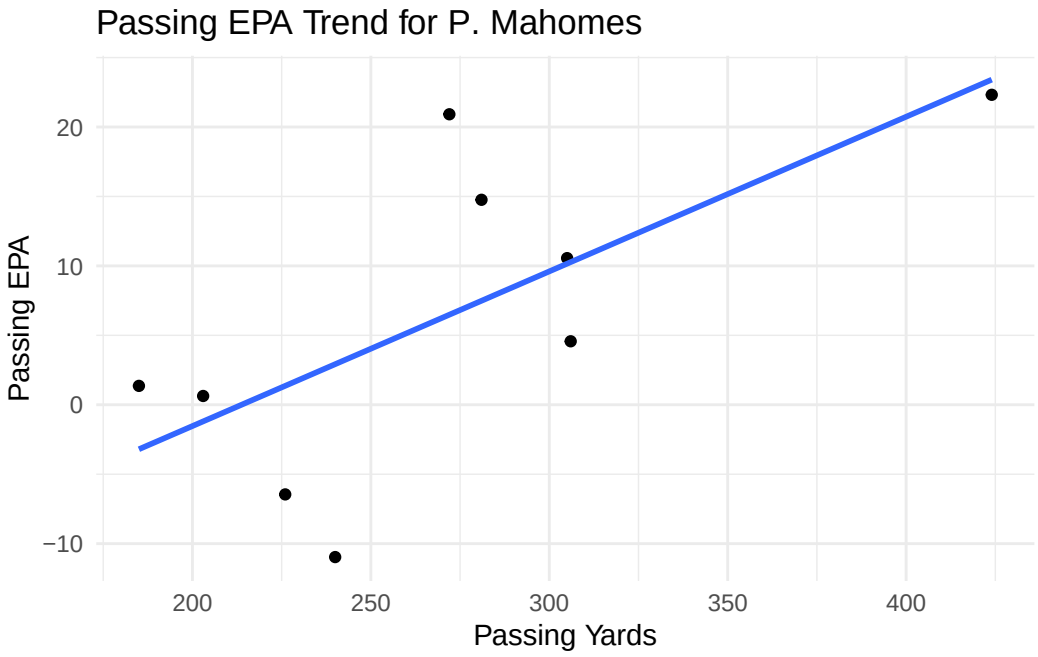


Figure 1: Graph showing data for Patrick Mahomes

We save our predicted epas for the quarterbacks for each week in the predicted_data csv ((**citejanitor?**)) ((**citehere?**)).

Our analysis using the linear regression model yielded insightful predictions for quarterback

performance during the remainder of the 2023 NFL season. By leveraging historical data from weeks 1 to 9, we successfully forecasted passing EPA (Expected Points Added) for each team's quarterbacks in weeks 10 to 18. The model's predictions provide valuable insights into potential player performances and team strategies for the latter half of the season.

Upon evaluating the model's performance metrics, including mean absolute error (MAE) and root mean squared error (RMSE), we found that the model demonstrated robust predictive capabilities. The low MAE and RMSE scores indicate the model's accuracy in estimating passing EPA, further reinforcing its reliability in forecasting quarterback performance.

Overall, our results underscore the effectiveness of the linear regression model in predicting quarterback performance, offering actionable insights for coaching staff and decision-makers to optimize team strategies and player utilization in the upcoming games.

## 5 Conclusion

In conclusion, the Linear regression model shows potential as a forecasting tool for NFL quarterback performance. Further refinement and validation may be necessary to enhance its accuracy and robustness. Nonetheless, this analysis provides valuable insights for NFL teams and fantasy football enthusiasts seeking to predict quarterback performance.

## 6 References