Alissa Trujillo

Professor Denton

DSC 680

28 September 2023

The Importance of a Good Quarterback – Milestone 1

Topic

How important is having a good quarterback to a team's success? Which factors are most statistically significant in predicting a team's success?

Business Problem

This project looks at the importance of an NFL team having a reliable quarterback. Does having a quarterback with high QBR, lots of passing touchdowns, or good ball control make a significant impact on the team's ability to make it to the playoffs or even the Superbowl? We will discuss each of these factors and aim to discover they key to a team's success.

This information can provide insights to coaches and managers looking to maximize their efficiency with cap space, draft picks, and player salaries. Knowing how much value a quarterback with these qualities adds to a team, they base their signing decisions based on quantifiable data.

Datasets

This project will pull data from Amazon's Next Gen Stats API through a Python Package called nfl-data-py. This package allows the user to import a large variety of NFL data, including play-by-play data, player stats, team rosters, draft pick data and more. This data can be pulled from the API source and formatted into a data frame.

For this project specifically I will be looking at the player data, specifically as it pertains to QB stats. I also intend to look at play-by-play data as well as overall team win stats. I will have multiple data frames created with different sets of information, and I am planning on implementing an app that constantly refreshes the populated data so it stays up to date as games play out in the NFL.

Methods

I intend to build a model that uses all the play-by-play and historical game data to determine which factors are most integral to a team being successful. I would like to focus on the question of whether having a top-tier quarterback is necessary, however I am sure I will find lots of interesting stats throughout the journey so it might flesh out a bit wider than that. I would like to create a regression model powered by a neural network. I have not begun the process of building it, so if it is not feasible to use neural networks I will likely regress to more simple forms of regression models. I am still working to decide my definition of "success" whether that be making the playoffs, winning the Superbowl, or just having consistently good games over a long period of time. Whichever definition I choose will change the way I analyze the resulting data.

The second piece of the project is that I would like to create an app where users can interact with player data from their browser. I would like this app to display QB stats in a user-friendly and interactive format. I have begun working on this portion of the project using a package called Streamlit, and so far I have QBR rating displayed, with starting QB's ranked from best to last displayed with their photo icon pulled from the Next Gen API. This is interactive and allows the user to select which year they would like to see ratings from. It is also set to refresh when the link is clicked so that it is always up to date (and therefore on Sunday after the games it should reflect any changes). I have begun working on a second piece where the user can select a QB and it will

display their change over the years. As I flesh out the model and find out which stats are most crucial to team success, I hope to add a few more elements to this dashboard.

Ethical Considerations

When it comes to sports analysis, there is a huge emphasis on predicting scores, Vegas odds and betting in general. I aim to stay away from predictions of individual game outcomes and rather focus on the concept of success as a whole. Focusing on a team's ability to be a playoff contender or possibly just vastly improve their offense will allow me to draw conclusions that are helpful to the business model but do not promote gambling.

Challenges

I have never created a user interface, so that so far has been the biggest learning challenge. I have the framework of a very basic app, but I will need to do more learning and tuning in order to get it up and running the way I'd like. In terms of the model and answering the business question, there really are an incredible number of variables contained in this data so weeding through them to make meaningful connections will be the most difficult part. Focusing on QB stats will help in that vein, however it may also leave out some key components to success, so I must be careful to be clear with my assumptions.

References

API reference - streamlit docs. API Reference - Streamlit Docs. (n.d.). https://docs.streamlit.io/library/api-reference

Carr, C. (n.d.). *NFL Stats Dashboard*. GitHub. https://github.com/CharlesCarr/nfl_nextgen_stats/tree/main