# Project Proposal

**Executive Summary:**
Sports fans love talking about ranking players among their contemporaries. NFL fans in particular love to rank their team's QB amongst the rest of the league. However, are the stats they use to support their arguments good stats to supplement their points with or are their better stats out there that helps say, determine a QB's efficiency relative to league average? I intend to try to compare QBs with more thorough metrics through public play by play data from 2009-2018 derived from Kaggle.

**Target Audience ("Who" and the "Why"):**
The target audience is any NFL fan who is interested in comparing their QBs relative to their peers. While they don't need to know about the numbers that constitute the formula, they are interested to learn about the intended usage of these advanced metrics and why several of the more "simpler" stats to compare players in public discourse is flawed.

**Data ("What"):**
The data I'm using constitutes of a diverse and detailed play-by-play data from the years 2009 to 2018 that is publicly available in Kaggle. Each play has various metrics assigned to it from efficiency(EPA/Estimated Points Added, WPA/Win Probability Added) and more "simpler" stats like completion%, result of the play, etc.

**Presentation("How"):**
The data will be presented through two visualizations. One will be a story that initially begins with comparisons between QBs by starting out with more volume-based stats like passing yards and Total TDs before narrating through more advanced metrics like efficiency, third down success, etc. This will greatly help give a more accurate indication of the best QBs in the NFL. The other is an interactive dashboard that interested stakeholders can manipulate with as they please.

**Limitations and Challenges:**
The dataset I'm using is populated with a voluminous amount of data and it will have to be combined for a more concise dataset. This can be challenging, and multiple aggregate functions and calculated fields may have to be created to help satisfy the requirements of my project proposal.

# Persona 1


**Name:** Samantha Wilks

**Age:** 27

**Gender:** Female

**Goals:**

- Interested in having a more thorough understanding of the game and is interested in statistics beyond the traditional passing yards, Touchdowns, etc.
- Interested in using these more advanced metrics to supplement Samantha's viewpoints and analysis about QBs

**Challenges:**

- May be unfamiliar with some stats even with the explanation my visualization will provide
- A large volume of data, so chance of being overwhelmed
- May not be interested in these new metrics
- May have a limited attention span

**Context:**

- Will interact with dashboard or read and experience the story during her down hours when wanting to learn more about the league and to a larger extent, the NFL players, she passionately follows