Analyzing NFL Quarterback Decisions and Play Performance Under Pressure

Prathik Addepalli, Matthew Sitto & Utku Yurday

Winter 2025

Motivation

The NFL is the top professional football league in the U.S., where teams compete in high-intensity games. The quarterback (QB) is the team's leader on offense, responsible for making play decisions. Being "under pressure" refers to situations where the opposing team's defense closes in on the QB, forcing hurried decisions, increasing the risk of sacks, interceptions, or incomplete passes. This is defined by Next Gen Stats as when the pressure probability on a play exceeds 75%.

According to Alex Vigderman's study "Under Pressure: Using Timing Data To Evaluate QBs Under Duress", about 3 seconds of the QB holding the ball will increase likelihood of pressure from 20% to 80%. In these situations, the QB only has two decisions: run (scramble) or attempt a pass.

While QB data has been available in recent years, many studies have solely focused on specific plays.

For example Matt Kelley, Taylor Bechtold and Ana Gonzalez's report titled "Pressure Points: Performance Under Duress Separates the Best QBs From the Rest of the NFL" addresses QB pass ratings under pressure, but omits other play events, such as sacks and scrambles.

Arjun Menon and Judah Fortgang's report, "Evaluating NFL Quarterbacks' Decision-Making Process", analyzed QB decision making as it related to EPA (Expected Points Added) per play; however, this does not focus on plays where the QB was under pressure.

Our project aims to explore 2022 NFL data to investigate how QBs perform under pressure vs. no pressure. The goal is to identify trends in decision-making and search for correlations with resulting QB performance.

Project Value

- Analysis of QB decision making under pressure can provide insights for teams to improve play calling.
- Findings of high-risk/low-reward play decisions can be used to reduce the risk of sacks or injuries.

Key Questions

- Are QBs more likely to attempt a pass or scramble when under pressure?
- What are the pass vs. scramble play outcomes for QBs under pressure vs. no pressure?

Data Sources

Primary Dataset

The primary dataset is called **Player Play Data** (52.5 MB, *player_play.csv*) and it contains player-specific play data for every NFL player in every game during the 2022 season (i.e., each row is a unique player/play combination).

- Players are identified by *nflId*
- Plays are identified by *playId*
- Games are identified by *gameId*

This dataset allows us to identify all plays where the QB was under pressure by using the *causedPressure* field.

When combined with the secondary datasets, we are able to extract a dataset of QB stats that we can analyze, where the QB was under pressure.

Secondary Datasets

The secondary datasets are:

Players Data (105 KB, *players.csv*) contains player data for all the NFL players in the 2022 season.

- Contains the *nflId* key for linking with the Player Play Data.
- The player name and position are the fields used for our analysis, in order to filter the data for QBs and extract name data for the individual QB analysis.

Plays Data (6.9 MB, *plays.csv*) contains play-specific data for all plays in the 2022 season.

- Contains the *playId* key for linking with the Player Play Data.
- The fields of interest for our analysis are: pass result, pass length, time to throw, time to sack, time in the tackle box, and pre snap score.

Games Data (7 KB, *games.csv*) contains information regarding each game of the 2022 season, such as the date, home/away teams.

- Contains the *gameId* key for linking with the Player Play data.
- We used home team and visitor team information to determine the side for each QB and ultimately, calculate the pre-snap score differential.

All of the datasets can be found on Kaggle: https://www.kaggle.com/competitions/nfl-big-data-bowl-2025/data

Data Manipulation Methods

Preparation of Data Sources

- All four CSV files were downloaded from Kaggle and added to the project repository. This is the only manual step, as all processing is done within a single python file. The files are *player_play.csv*, *players.csv*, *plays.csv*, and *games.csv*.
- After uploading the data to the project repo, the files were read, cleaned and manipulated using the Pandas library.

Data Filtering & Merging

- Once the data is loaded, we filter for QBs by selecting players labeled as “QB” in the **Players Data** and extract their unique *nflId* numbers. We then create our own numerical identifiers (*game_play_id* and *game_play_nfl_id*) that uniquely links plays across datasets, as the *playId* identifier available to us is not unique across games. These are created using *gameId*, *playId* and *nflId*.
- We then identify plays where defensive pressure was applied by filtering **Player Play Data** for instances where *causedPressure* value is “True” and use the *playId* associated with those plays to find QB plays under pressure. This is done to create 2 datasets, one with plays where QBs face pressure and one where they do not, allowing for a comparative analysis of performance under different conditions.
- With the QB plays under pressure filtered, we then merge the data with the **Plays Data** to add play details to the dataset. We also merge the data with **Games Data** to include broader game context. A new metric called *scoreDelta* is calculated, representing the score differential before each play. This is calculated by comparing the home and visitor scores based on the QB’s team, to help analyze how pressure situations correlate with the game’s state. Finally, unnecessary columns are dropped to remove noise.
- Once all datasets have been merged and the scoreDelta is calculated, the *process_qb_pressure.py* script selects relevant features to create a structured dataset focused on QB decision making. The script can be called from within the *qb_pressure_analysis.ipynb* Notebook to create the pre-processed dataset or can also be saved as a new CSV file. We use the same script to generate a pre-processed dataset for plays with and without pressure.

Data Wrangling Challenges

- Incorrect data was found in the final dataset after **Player Play Data** was merged with **Plays Data** using *playId*. This is because *playId* is not unique for each play across all games, it is only unique for each play in a single game.

  To ensure proper merging, the *game_play_id* key was created by combining *gameId* and *playId* to create unique identifiers for each play across all games. When **Player Play Data** was merged with **Plays Data** using the new key, the dataset was merged properly.

- Duplicate plays with different QBs were found in the final dataset when multiple QBs were listed as players in a unique play, regardless of which QB on the team actually played.

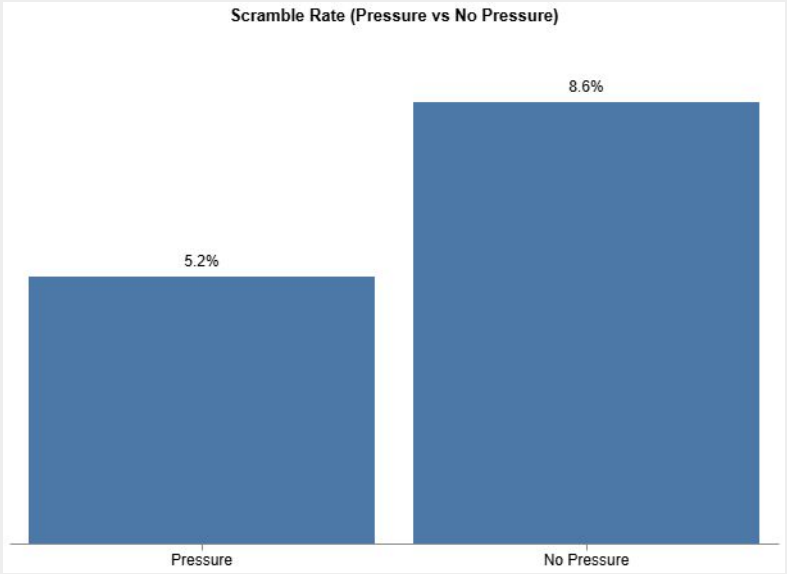
  To remove these incorrect QB plays, we validated QB assignments using the QB name found in the **Plays Data** *playDescription* field and creating the *game_play_nfl_id* key to identify all plays in all games for each QB. If the QB was not found in the *playDescription*, they would be removed from the dataset by using the *game_play_nfl_id* key.

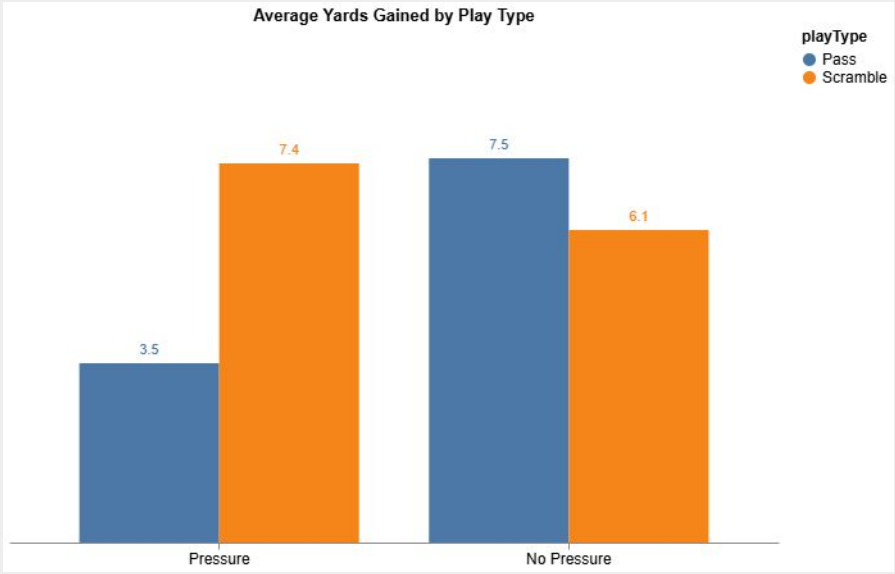
- Some plays had null values for *passResult*, which is the key field indicating the result of each play (scramble, complete pass, incomplete pass, interception or sack).

  This was predominantly a problem for “no pressure” plays, most of which were plays involving a running play by design, which would be out of the scope of our analysis. We dealt with null values using *hadRushAttempt*, where if this field equalled 1, we set *passResult* to “Scramble”. Otherwise, the null value *passResult* rows were dropped. For the handful of “pressure” plays specifically, we also reviewed the individual *playDescription* values to confirm that the choice of “Scramble” for *passResult* was appropriate.

Scramble Rate Comparison

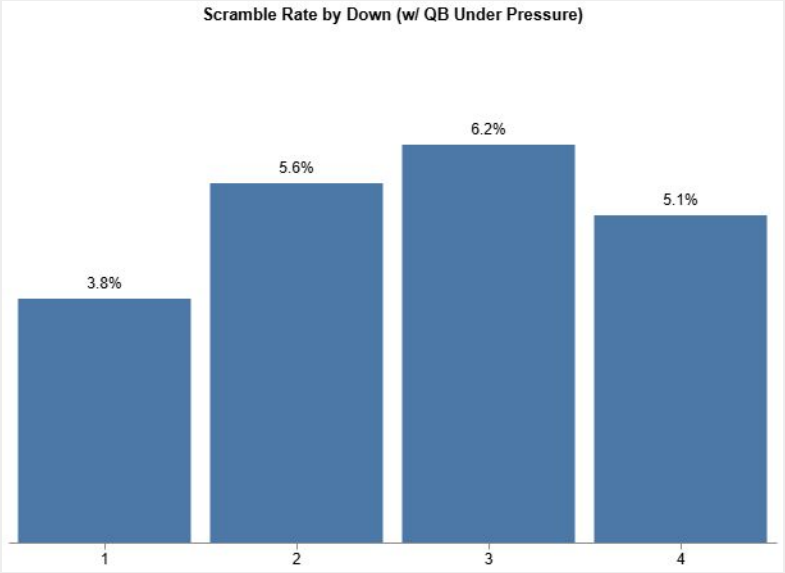
One of the fundamental decisions a QB can make when facing pressure is whether to find a quick pass or to scramble. Our analysis showed that 1) QBs overwhelmingly choose to pass when under pressure (c.95% of the time), and 2) Relative to “no pressure” situations, QBs are less likely to scramble under pressure. Interestingly, QBs under pressure gain more yards if they choose to scramble, which went against our intuition and is the opposite of what we observed for “no pressure” plays. This is likely primarily driven by the much higher completion rate that we observe in “no pressure” passing plays.

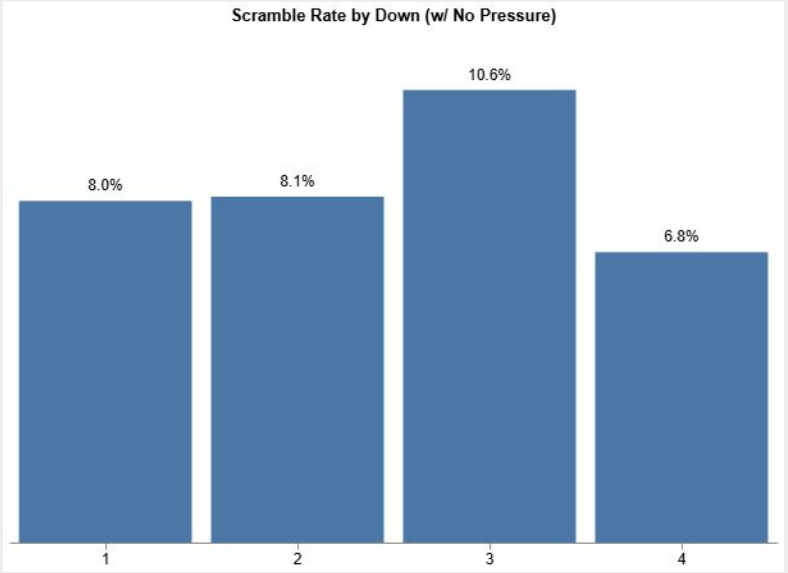




Scramble Rate Comparison by Downs

The likelihood of a scramble is much higher for 3rd down plays, which is true for both “pressure” and “no pressure” situations. This is likely somewhat influenced by fewer yards remaining to 1st down, even though the yardage difference is not that great. The lower likelihood of a scramble on a 4th down seems less meaningful, given the lower number of observations (e.g., our data has <100 pressure plays on a 4th down vs. >1,000 for each of the first 3 downs).

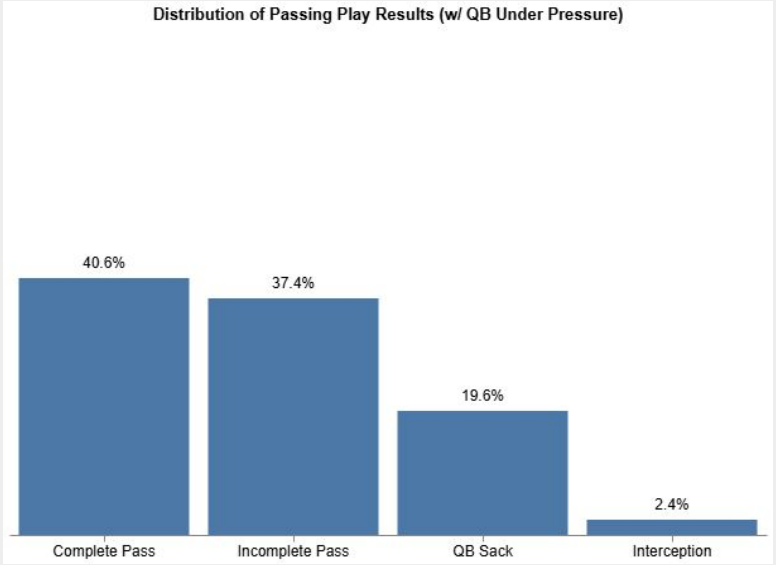


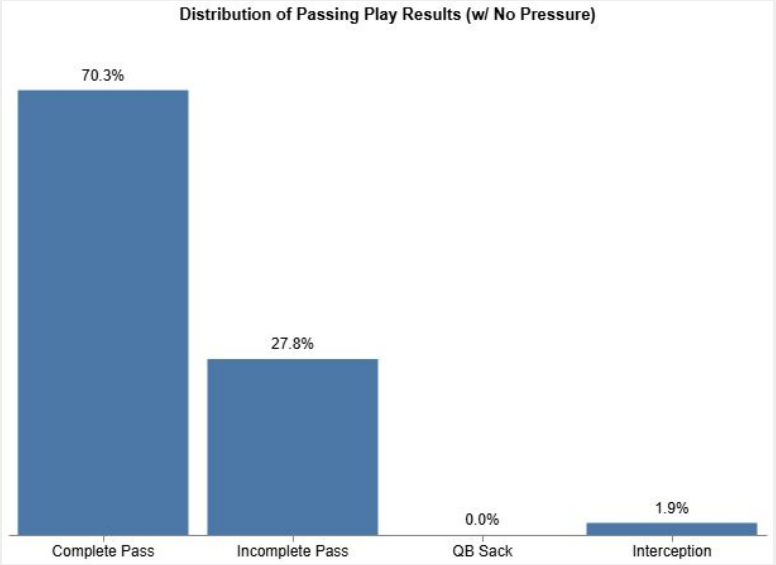


Avg. yards to 1st down	1	2	3	4
	10.2	8.7	7.8	5.4
	10.2	8.4	7.3	4.2

Passing Play Result Comparison

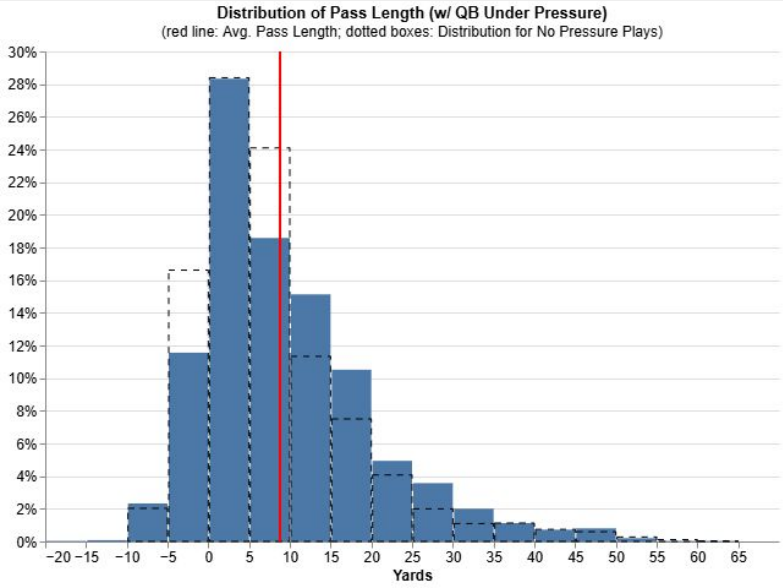
Since the majority of “pressure” plays result in a QB deciding to pass the ball, we dug deeper into results of passing plays. Unsurprisingly, completion rates are only c.40% when under pressure, notably lower than the c.70% rate without pressure. There is a slightly higher risk of an interception when under pressure. Most interestingly, nearly 20% of the time, the play results in a sack, with an average 6.8 yards lost. On average, QBs hold the ball for 3.8 seconds (vs. 2.9 seconds for a complete pass). This would suggest that QBs should think to scramble or force a pass, even if less ideal, in order to avoid a sack for a loss of yardage.

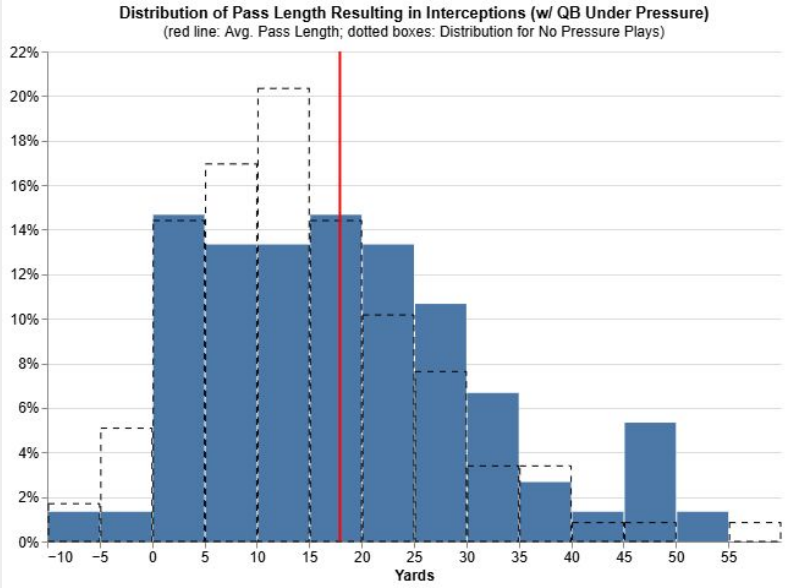




Pass Length Impact on Play Results

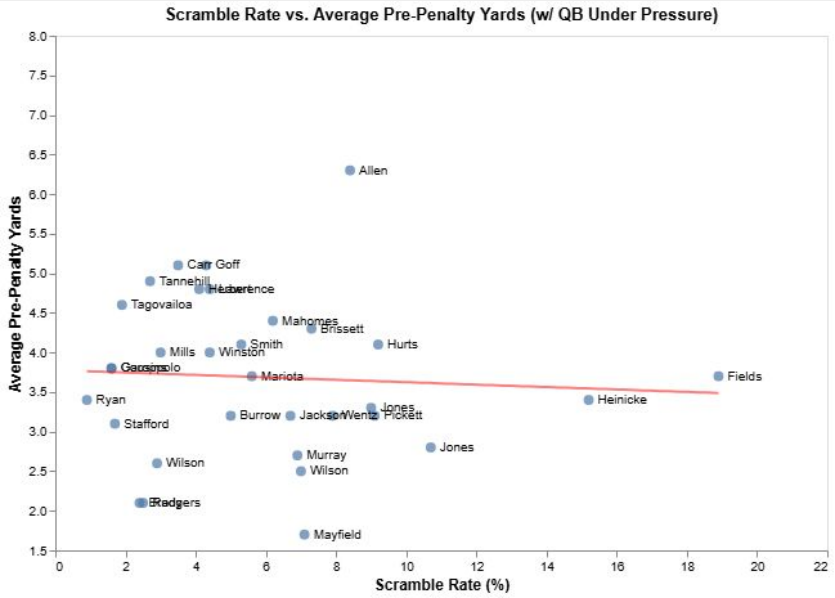
We also analyzed a range of pre-snap conditions that may influence the QB decision, including yards to first down, score delta and yards to endzone, none of which had meaningful differences in distribution between “pressure” vs. “no pressure” plays. Yet, for passing plays with non-null passing length stats (i.e., excluding sacks), we noticed a notably higher average length of pass in the case of pressure plays (8.8 vs. 7.2 yards for “no pressure”), which is also evident in the histogram below. Even more interestingly, the difference for interceptions was 17.9 vs. 13.9 yards and the histogram for “pressure” plays more right-tail skewed as well. This would suggest that QBs under pressure tend to force longer passes with a higher propensity of getting intercepted.

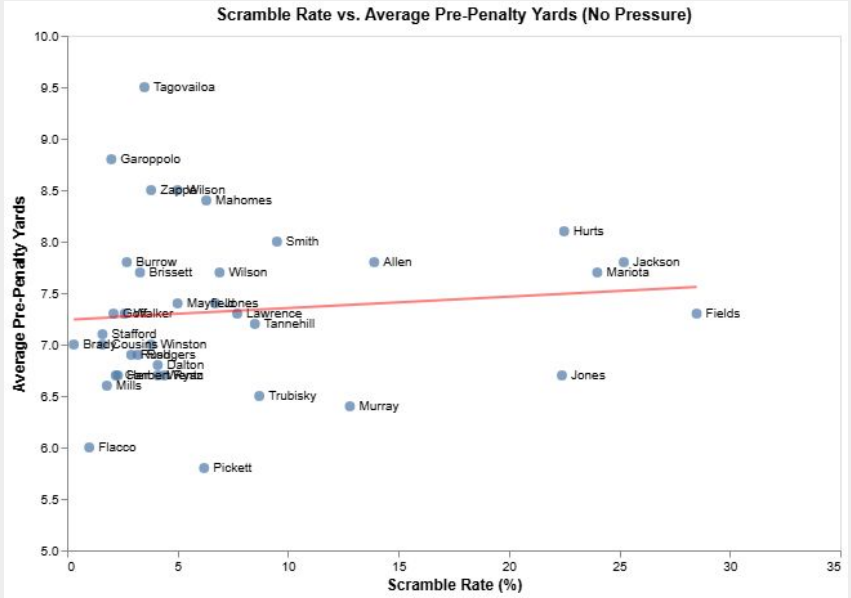




Scramble Rate vs. Yards Gained

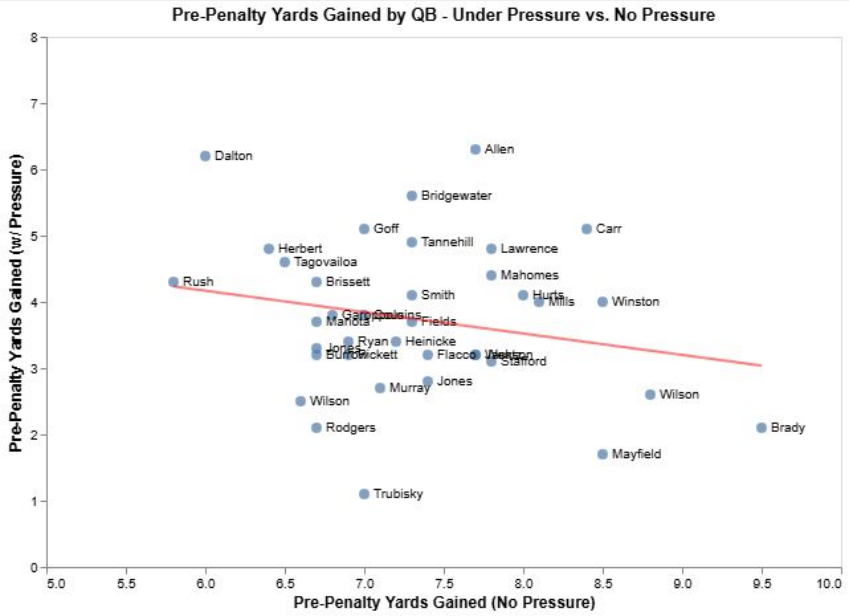
Since the aggregate level data suggests there is a yardage gain advantage to scrambling when under pressure, we wanted to analyze individual QB data and see if there is a correlation between QBs that choose to scramble more often vs. the resulting yards gained on the play. There is no strong correlation in the case of pressure plays, suggesting that the higher yardage gain when QBs choose to scramble is likely driven by smart choices around when to scramble (i.e., scrambling more often does not mean more yards gained). We repeated the analysis for “no pressure” plays and found a similar lack of correlation, albeit a slightly positive relationship relative to the “pressure” plays.

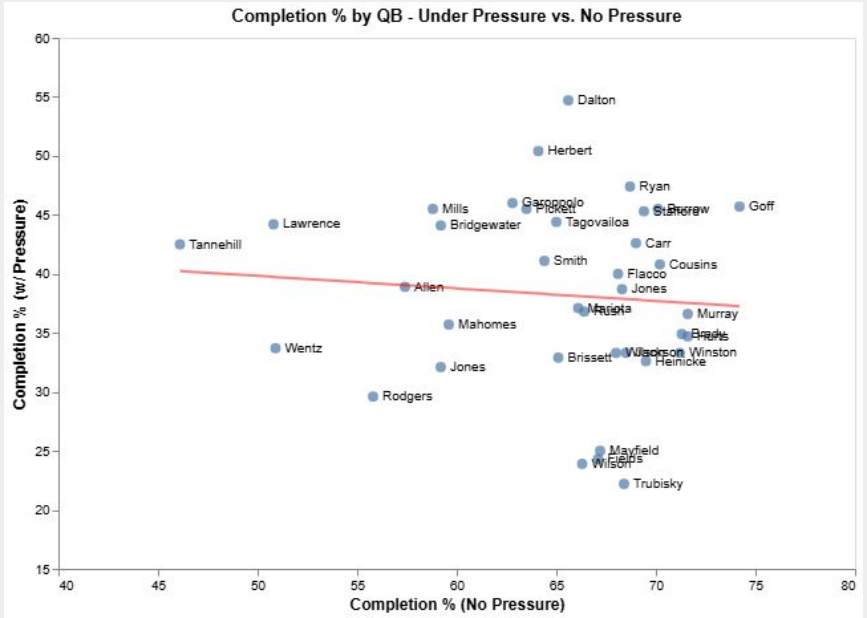




QB Comparison Under Pressure vs. No Pressure

Finally, we analyzed how individual QBs perform with and without pressure. Specifically, we looked at pre-penalty yards gained and % of plays that result in a completed pass for each QB, comparing the results for “pressure” vs. “no pressure” plays. Unfortunately, our analysis did not uncover a strong correlation. If anything, there is a slight negative correlation, especially for yards gained. That is, QBs who have achieved higher average yards gained when there is no pressure, tend to do more poorly when under pressure (e.g., Tom Brady).





References & Statement of Work

References

Kelley, M., Bechtold, T., & Gonzales, A. (2025, January 3). *Pressure points: Performance under duress separates the best QBs from the rest of the NFL*. The Analyst. https://theanalyst.com/2025/01/nfl-quarterback-performance-under-pressure

Menon, A., & Fortgang, J. (2022, July 6). *Evaluating NFL quarterbacks' decision-making process*. Pro Football Focus. https://www.pff.com/news/nfl-evaluating-quarterbacks-decision-making-process

National Football League. (2023, September 21). Next Gen Stats: Introduction to Pressure Probability. NFL.com. https://www.nfl.com/news/next-gen-stats-introduction-to-pressure-probability

Vigderman, A. (2020, October 23). *Under pressure: Using timing data to evaluate QBs under duress*. Sharp Football Analysis. https://www.sharpfootballanalysis.com/analysis/pressure-timing-data-patrick-mahomes-josh-allen-lamar-jackson/

Statement of Work

The work for this project was conducted in a highly collaborative fashion, with weekly pre-scheduled Slack huddles to update all team members on progress, as well as ad hoc check-ins via Slack several times each week to discuss any questions or concerns.

We also held regular catch-up sessions with our Faculty Mentor to update him on our progress and to discuss any questions.

For coding, the team primarily used the Coursera environment with manual version control. In future projects, we aim to use GitHub for more efficient code collaboration.

While all aspects of the project were handled as a joint responsibility, the following is a breakdown of which team member(s) took the lead on each of the major deliverables:

- Data source investigation - Group
- Team project proposal - Group
- Data processing, cleaning, merging - Matthew
- Data analysis - Utku/Prathik
- Visualizations - Utku
- Final report - Group