

Sydney Lockwood

The Impact of Defensive Pressure on NFL Win Probability

Snapshot of Growth

At the start of this semester, I could analyze data once it was already cleaned, but I did not fully understand how much preprocessing and handling the data required. Now I can take raw data, define the correct variables, clean and document it carefully, and carry out an analysis to a stakeholder-ready recommendation. The biggest change in my approach to analytics is that I now think first about the decision the analysis will inform, rather than jumping straight into analysis or charts.

Skills and Evidence

One skill I significantly improved at is data cleaning and merging. I demonstrated this in checkpoint 4, where I had to merge multiple data sets using Python and then cleaned the data through Excel to aggregate on a team-game level. A second skill I developed is modeling and interpretation, shown through the logistic regression analysis, where I translated the coefficients into a win probability analysis. A third skill I was able to improve is data communication and visualization. I haven't often used graphs to convey analysis, so it was certainly a learning curve and required diligent work for labeling and organization of the graphs.

Challenges and Fixes

One challenge I encountered was deciding which variables to include without overcomplicating the model. Early versions of the analysis included too many features and made interpretation difficult. I fixed this by narrowing down the variables to compare only a few, creating clearer insights. This taught me the value of focus and direction when applying analytics.

Data Practice

CP4 fundamentally changed how I approach raw data. I learned to clearly define each row and the importance of data preparation. Being diligent in every step through CP4 (preprocessing, merging, etc.) made the later stages much smoother as I was able to confidently explain how metrics were constructed.

Communication

My writing and visuals improved by shifting from describing what I did to explaining what the results mean. My strongest visuals for clear communication were the bar charts showing the effect of turnover, sack, and home-field advantage. With clear labels, axis titles, and one-line takeaways, it was beneficial to aid in explaining how influential each variable was. The biggest improvement was learning to frame insights in stakeholder language, such as percentage-point changes, instead of the technical model output.

Use of AI Tools

I used AI tools to help with editing and calculation problems, but I always verify the analytical logic and make data interpretations myself. AI was most helpful for improving clarity and structure. I learned to treat AI as a support tool but not a source of truth, and I will continue to double-check any output that influences conclusions or decisions.

Career Readiness

This portfolio helps me present myself as someone who can connect the data to real-world decisions. The two artifacts I would share with a hiring manager are the cleaned raw CP4 dataset and the final project report with visuals. Together, they tell a story that I can manage messy data, apply appropriate analytical methods, and communicate insights clearly.

Goals and Next Steps

Over the next three months, I plan to strengthen my model validation skills, improve documentation to support reproducibility, and incorporate NFL offensive metrics to better test the accuracy and robustness of the analysis.