Sydney Lockwood

Analysis Memo #1

Question and Outcome

My main research question is how do defensive sacks and defensive turnovers influence a team's likelihood of winning an NFL game. Coaches, analysts, and stakeholders care deeply about this because they want any way to increase their likelihood of winning. Identifying the value of sacks and turnovers can inform defensive strategies and help decide what risks to take in a game. The outcome used is a win (1) or a loss (0).

Data Used

In the cleaned file used for data exploration, one row represents a team in a single game, season, total sacks, total turnovers caused, and a win or loss. The data exploration is all regular-season games from 2020-2024 with no major filters applied. Final sample size is 2,540 rows.

Method Choice

The analysis method I chose is logistic regression. This model can help estimate the strength of each defensive input and is ideal for predicting a binary outcome like winning or losing. This allows me to estimate the individual impact of each defensive metric while holding others constant. The baseline model includes only turnovers and sacks to keep the analysis simple and focused.

Analysis Spec

Component	Description
Outcome	Win or loss: binary
Predictors	Turnover, Sacks
Row Definition	Team-Game
Sample	All regular-season games from 2020-2024
Formula	Predicted Win = B0 + B1(Sacks) + B2 (Turnovers)

Results

Predictor	Coefficient	Interpretation	Statistical Significance
Intercept	0.1308	A team with 0 sacks and 0 turnovers has a ~13% win probability	3.57E-12

Sacks	0.0716	Each additional sack increases win probability by 7.2 percentage points	1.37E-41
Turnovers	0.0925	Each takeaway increases win probability by 9 percentage points	1.26E-45

This linear regression shows that both sacks and turnovers are statistically significant predictors of winning an NFL game (p<0.0001). Turnovers have the strongest effect. Each turnover increases a team's win probability by approximately 9.3 percentage points. Sacks also have a positive impact on win probability by about 7.2 percentage points per sacks. This model explain that overall 16.6% of the variation in wins (R2=0.166), which is reasonable given the limitations.

Checks
1. Correct Scale: Sacks and turnovers are small integers between 1-10
2. Stability Check: Ran the regression on only 2023-2024 to produce nearly identical relationships. This helps confirm that the regression

Interpretation
1. Turnovers matter the most
   a. Each turnover increased win probability by about 9%, making it the strongest predictor in the model
2. Sacks are also meaningful predictors
   a. Each sack lifts win probability by about 7%. While it is not as powerful as turnovers, sacks are still a measurable competitive advantage
3. Teams that generate pressure on the offense and take the ball away win more often
   a. Creating pressure can lead to offensive mistakes that opens the door for turnovers and sacks more often

Limitations
- The model treats all turnovers as equal, even though interceptions and fumbles differ in value
- Only two defensive inputs are used in this model while NFL wins depend on a wide variety of factors

Next Steps
- Add another variable such as sack rate or points allowed
- Add the win probability layer to each game and re-run the regression