

Sydney Lockwood

Analysis Memo #2

Quick Reminder

This project examines how defensive sacks and defensive turnovers influence a team’s likelihood of winning an NFL game. This matters for coaches and analysts who want to understand which defensive actions lead to a higher win probability.

Data Used

I used the same data set from checkpoint 4. 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.

CP6 Baseline

In checkpoint 6, I ran a linear regression using sacks and turnovers as predictors. The baseline model showed that both sacks and turnovers were statistically significant, while turnovers had the strongest effect on wins. This model explained about 16.6% of the variation in wins.

Upgrades in CP7

For checkpoint 7, I added one key feature, whether the team was home or away. I chose this upgrade because historically, home teams win more often. This allows for a direct comparison to checkpoint 6 and tests whether contextual features improve the prediction.

Model or Analysis Spec

Component	Description
Outcome	Win or loss: binary
Inputs	Turnover, Sacks, Home/Away Game
Row Definition	Team-Game
Sample	All regular-season games from 2020-2024
Formula	Predicted Win = $B_0 + B_1(\text{Sacks}) + B_2(\text{Turnovers}) + B_3(\text{Home})$
Expected Direction	Home/Away piece is expected to increase the win probability

Model or Analysis Spec

Model	R2	Notes
-------	----	-------

CP6 (Baseline)	0.16598	Sacks and turnovers
CP7 (Upgrade)	0.16688	Added home/away piece

Improvement: +0.00090 (0.09 Percentage Points)

Results and Comparison

- Stability
 - Coefficients for sacks and turnovers barely change from CP6 to CP7
- Time-window check
 - Running a model on only 2023-2024 produced a similar home/away effect
- No post-game data leakage
 - Home or away field advantage is known before the game and does not depend on the game outcome
- Reasonable error
 - R2 changes slightly but in the expected direction

Interpretation

1. Home teams win more often
 - a. The home advantage is worth about 5 percentage points
2. Turnovers remain the strongest predictor
 - a. Adds 9% win probability per takeaway
3. Sacks continue to matter by increasing win probability by about 7% per sack
4. Adding game context improves prediction

Limits

My two main concerns that still worry me:

1. Turnovers are split into fumbles vs. turnovers
2. The model still excludes offensive stats that could dramatically change the weight defensive levers hold, such as pass attempts or completion percentage

Next Steps

1. Keep the 3 already tested variables
2. Add the win probability, given before the game, as another variable to test
3. Consider adding sack rate or points allowed to better capture defensive efficiency