**Methods:** This is a very high effort post. I built a model for play-by-play simulating NFL games. That is, my model simulates games by simulating each aspect of the game: Was there a punt? Was there a FG? Did they pass the ball or run it? Was there a fumble and if so, who recovered it? Was there an interception? How many yards did this play get? How much time did the play require? Notably, within the model, the team which is playing was used as a variable allowing me to make predictions about how well a specific team would do in some situation rather than

how often the Bears disappoint everybody against the Packers.

The modeling is done through a series of modified linear and logistic regressions (i.e., correlations but with multiple variables). I detail the challenges associated with this and the statistical procedure I employed below. However, here I want to say that I spent the most effort into developing something which accurately models the number of yards gained in a play, and I think I found a good way to do this using, relatively simple, regressions.

**Results:** From a cursory glance (e.g., checking what type of plays teams run at different times of the game), my model seems to be better than the ones available online. Furthermore, I did some preliminary tests showing that my model outperforms opening Vegas ATS lines X% of the time [95% confidence interval] and opening OU lines Y% of the time []. The closing Vegas lines are a bit better,

In either case, I think these

I aimed to develop something where, if the offense needs 9 yards to go on third down, then gaining 9 or 10 yards has a relatively high likelihood compared to getting 6 yards, even though, on average, 6 yards is the most common amount gained.

**Future Directions:**

Implementing penalties. This would have been a non-trivial degree of effort as it would require finding out what type of penalty causes an automatic first down or not. Penalty distances are discrete units which is unlike the rest of the things in the model. I think I would also need to make something that predicts among all types of penalties all at once. I may also need to make something that determines whether penalties were accepted or rejected seems tricky, since that is generally a surefire accept or reject. Each of these steps isn’t too difficult but the effort adds up, and I wasn’t able to spend another Sunday morning building this. Potentially some teams benefit particularly much from penalties (e.g., GB), and the element of randomness certainly has some effect, but I don’t imagine adding this would change the results. I doubt the spread or over-under lines would shift by more than 1.

Joke about this not working for the Pro Bowl.