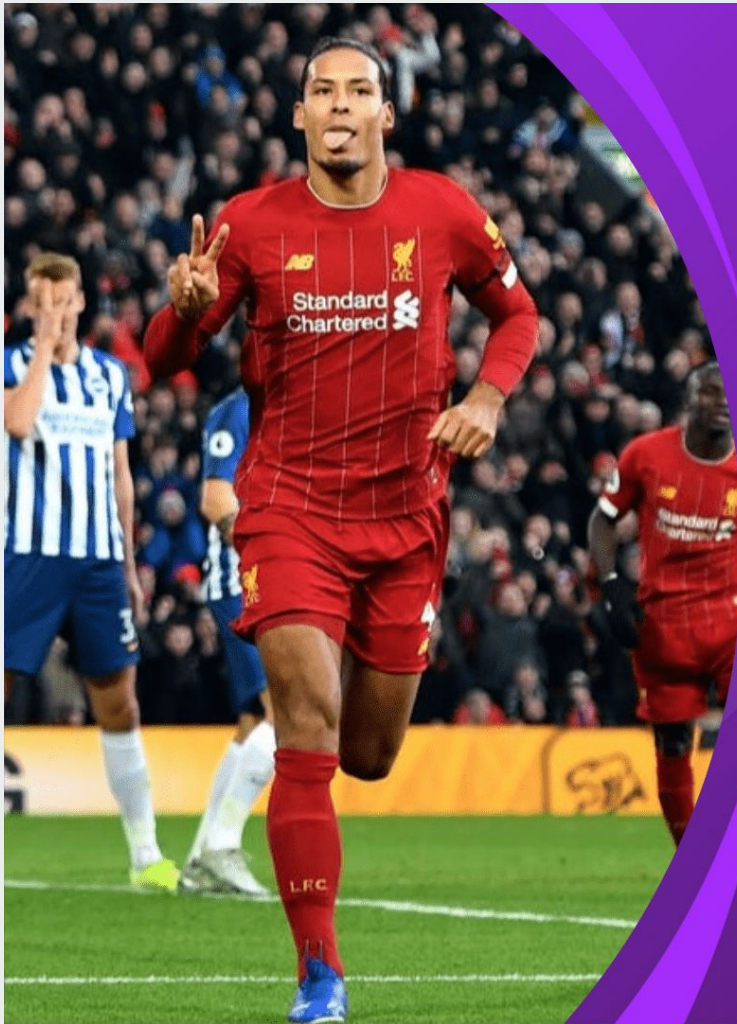




# Modeling for Predictive Analysis in Soccer

Using Data from the 2019-2020 Premier League  
Season

Arthur Lennard



# 2019/20 PREMIER LEAGUE

---

# SPORTS BETTING LEGALIZED IN NORTH CAROLINA

CAN WE PREDICT GAME  
OUTCOMES BETTER  
THAN THE BETTING  
ODDS?

WHAT ARE BETTING ODDS?







**THE DATASET WE USED WAS  
GATHERED BY VAASTAV  
ANAND**

**LOSS - 0 POINTS**

**DRAW- 1 POINT**

**WIN- 3 POINTS**

**WHAT IS XG AND WHY IS IT  
IMPORTANT FOR PLACING  
BETS?**

# VARIABLES USED IN OUR MODELS

- CUMULATIVE XG AND XGA
- CUMULATIVE MEAN XG AND XGA
- XG DIFFERENCE
- PLAYS IN FINAL 3RD
- XPOINTS
- BETTING ODDS  
(FAVORITE TEAM WINS)





# ORDINAL REGRESSION MODELS

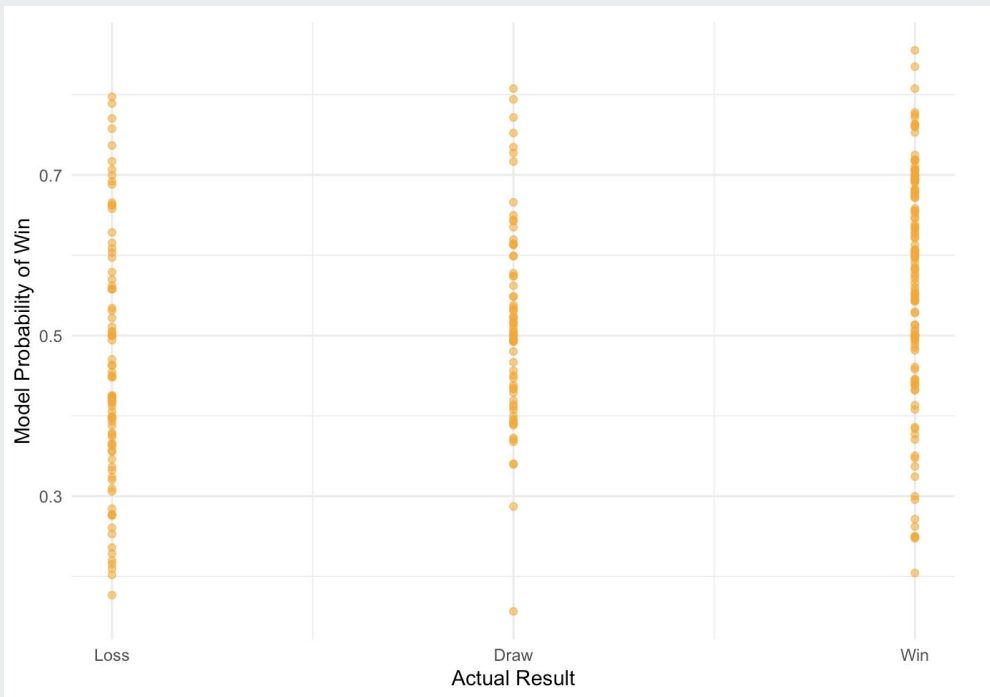
8 TOTAL

PREDICTIONS	RESULTS		
	0	1	3
0	45	20	23
1	0	0	0
3	42	52	106

MOST ACCURATE MODEL (MODEL 3): ~52%  
(WITHOUT PREDICTING ANY DRAWS)



# MEAN CUMULATIVE: XG + XGA + OPPONENT XG + OPPONENT XGA

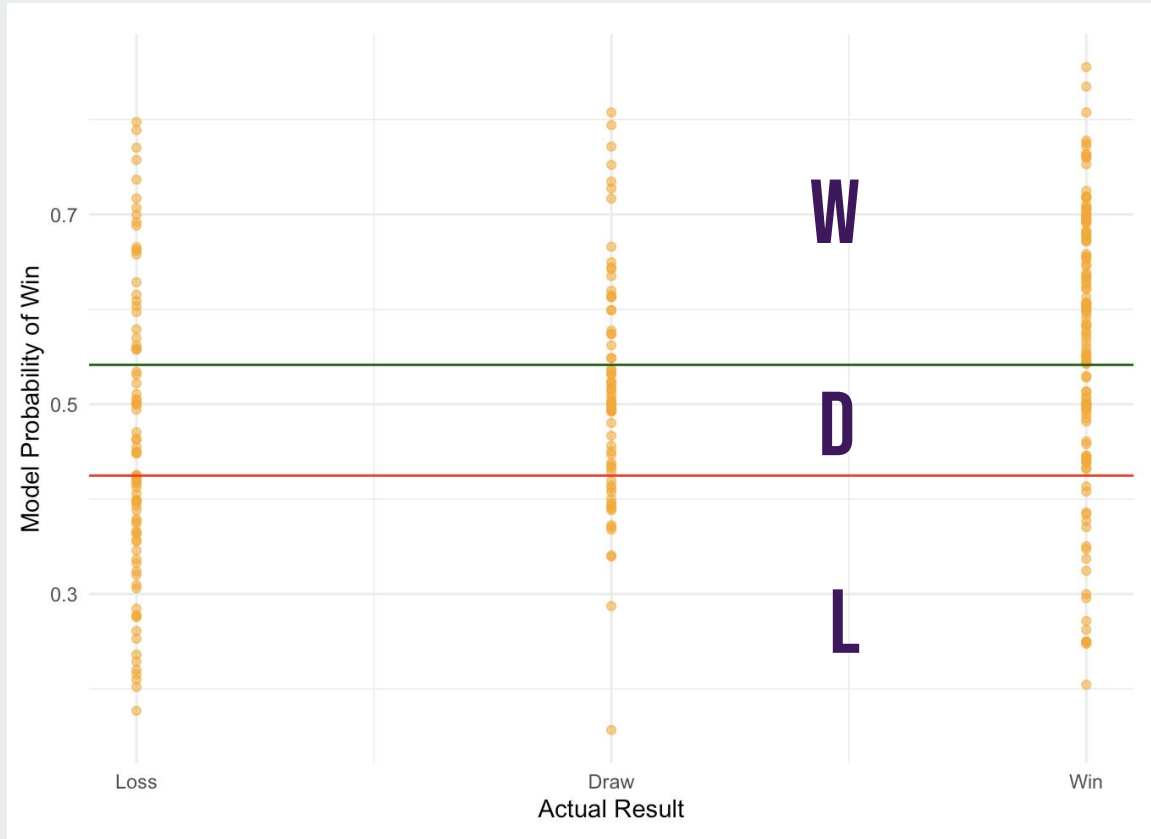


MODEL 3





# MANUALLY DRAWN INTERVALS





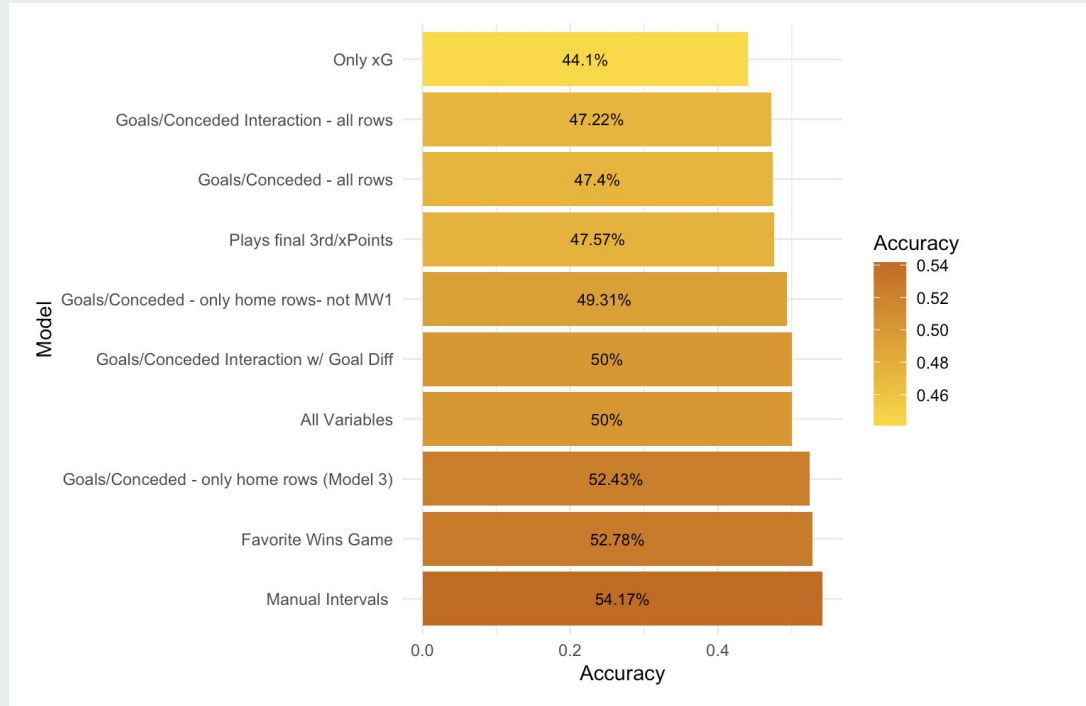
# NUMBER OF EACH PREDICTED OUTCOME MUCH CLOSER TO TRUE NUMBER OF EACH OUTCOME

Result	Model Predictions	True Outcomes
0	80	87
1	78	72
3	130	129





**MANUAL INTERVALS: 54.17%**  
**BETTING ODDS: 52.78%**



- SLIGHTLY HIGHER  
THAN BETTING ODDS
- ODDLY LARGE  
NUMBER OF CORRECT  
UPSETS  
—>PROFIT ?





# THANK YOU!

