Introduction

Predicting the outcome of sports seasons has always been a topic of immense interest among statisticians, sports enthusiasts, and analysts. The National Football League (NFL) is no exception, its dynamic and highly competitive environment provides fertile ground for predictive modeling. The goal of this project is simple: Use data from the 2023 NFL season, to predict how many games each team will win in the 2024 season.

The backbone of our predictive model is the data from the 2023 NFL season, which includes the number of games each team won and their respective performance metrics, as evaluated by Pro Football Focus (PFF). PFF scores offer a comprehensive assessment of a team's offense, defense, and overall performance, and in my opinion, is the most quantified metric to assess the performance of any specific player, and any specific team.

By analyzing the relationship between these 2023 metrics and the number of games won, I developed a linear regression model to predict the outcomes for the 2024 season. My approach involved:

- 1. **Data Collection and Preparation:** We gathered and cleaned data for the 2023 season, ensuring all relevant metrics were numeric and ready for analysis.
- 2. **Model Development**: Using linear regression, we established a predictive model based on the 2023 season wins and performance metrics.
- 3. **Prediction and Visualization:** The model was then applied to the 2024 PFF metrics to forecast each team's win totals. The results were visualized to highlight the predicted standings and performance trends.

This writeup details the steps taken to build the model, the insights gained from the analysis, and the predicted outcomes for the 2024 NFL season. Our findings provide a data-driven perspective on how teams are expected to perform, offering valuable insights for fans, analysts, and stakeholders in the NFL community.

The Data:

The data used in this model was meticulously gathered from the Pro Football Focus (PFF) website. This dataset includes various columns, each integral to calculating my proprietary metric for predicting team performance. The data collection process was particularly challenging. I had to subscribe to PFF Pro, manually record scores for nearly every starting player in the NFL, and compile overall scores for offense, defense, and other metrics.

Web scraping techniques were not viable due to the dynamic nature of team rosters. To predict the 2024 season, it was crucial to account for roster changes that PFF had not yet updated from the 2023 season. This included tracking every rookie drafted, as well as players who were traded or signed during free agency. Consequently, I manually maintained an updated roster for each team, ensuring the accuracy of each player's PFF score. This labor-intensive process was essential for the reliability and precision of my predictive model.

Longa morris SS. 6	Jake nations 71.2	Them Elemone 67.4
Joe 14-104 74-9	Marken Bosson 59.1	Danier cons 59.6
	Dere Onlinen 82.3	Aucho Cabell 47.9
cices washing 81.4	Chris Lindsform 87.6	Robert Lint 76.4
Tree Smith 74.6	Kaleb menera 75.5	Tonylor morten 74.6
Januar toold 49.8	teres in the same	. 1 4
Beinord Reinon 82.7	Andrew themas 74.1	Chukawa okrafor bo. Y
Dentin Mison 70.8	Jermane Eunlor 68.7	cole stronge 64.6
Ryon kelly 77.2	John Awhard seamer 41.4	Jake Ardrews 53
Lill Fries GI, L	son rentan 56.5	Sidy sow 64.4
	fron deal 34.8	Mike onen 71.5
Braden Smith 83.3		1 100 miller 20.)
Brown colonon 60	Com Robason 65.4	hollen miller 80.2 Dylan parken 60.4
NICIT Almert, US-8	Ezra circurions 59.5	Andre Somes 74.6
I Tule Birdist 69.2	mitch none 63.7	2102 68.
Som cosni 80.6	12:0000	honer motord 74.1
Andrew Lylic 61.2		
Trice burion 67	Thron Smith 83.7 Al	eric Sackson bt. 6
Tuler Sairl 74.7	John SIMPSON 56.3 Jo	man, Jackson 61
Brock wolfen, 53.7	Joe Tramon 61 5	feer Avila 60-5
Zoch norsen 68	ALT 71.7 K	evin Dotson 84.4
Terme Steele 523	wowen work 80.4 150	6 Hoursten 79.8
The state of the s	- sales of 1 Ro	mie Stoniery 68-1
Terron Arnstral 74.3		frem Vorhees SS
15aih wan 47.8	Graham Grasson (1)	er Linderbaum 78.5
Aaron Brewer 71.6	Fronk 1200000 71.6	
nobert somes 57.6	Brown Sevell 92.8 Rods	Clevelone 65.8 per Rosensartan 60
Apron Brewer 71.6 Robert Jones 57.6 Austin Jackson 66.1	De la Russ	Range 73.5
Tordon mailara 84.8	leasteed halmer v. 3	e train 69
- 1 andon Dickenson by	1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1. 1	Salderus 52.1
- lan Jurains bi.	Seen Rhyan 49.3 Enth	maio4 79.4
Lane Johnson 80.1		
Lane Johnson 80.1	Tach tom 79.7 lesar	1(1) 51.7

Stors (2.6 821 82 68.1 Stors (2.6 77.1 63.7 69.5 Stors (2.6 77.1 63.7 69.5 Stors (3.6 70.8 71.3 60.6 Stors (3.6 70.8 69.8 Stors (3.6 70.7 69.8 Stors (3.6 70.7 70.7 70.7 Store (3.7 70.7 70.7 10.7 Store (3.7 70.7 70.7 Store (3.7 7	(em	03	1513	~	V	DL		I
TOTALS 41 80.9 73.6 64.6 ILLS 87.2 70.8 71.3 63.6 ILS 87.2 70.8 71.3 63.6 ILS 87.2 70.8 71.3 63.6 ILS 87.2 10.5 64.2 67.4 ILS 87.2 10.5 64.7 ILS 82.6 70.5 70 ILS 70.7 82.6 70.5 70 ILS 70.7 15. 60.7 66 ILS 70.7 15. 60.7 77.8 ILS 82.6 71.7 77.8 ILS 82.6 71.7 77.8 ILS 82.6 71.7 77.8 ILS 72.9 84.9 64.2 64.7 ILS 72.9 84.9 64.2 64.7 ILS 72.9 84.9 65.5 66.9 ILS 70.9 16.5 66.8 66 ILS 70.9 16.5 66.8 66 ILS 61.5 61.6 62.2 ILS 61.5 61.6 62.1 ILS 61.5 61.7 70.1 ILS 61.7 10.7 10.7 ILS 61.7	tres	75.1	82-1	8.	2	68.2		-
11 80.9 73.6 64.6 11 87.2 70.8 71.3 63.6 11 87.2 70.8 71.3 63.6 12 10.7 14.3 14.2 17.4 13 15 15 17.7 14.7 17.7 14 15 15 16 15 16 15 16 16 16 16 16 16 16 16 16 16 16 16 16	cors	67.6	77.1	63	.7	69.5		
Ally 87.2 70.8 71.3 60.6 Alsono 64.5 64.3 64.2 67.4 Alsono 51 40.7 68.7 73.3 Bacc 52.2 66.5 67 668 Abores 76.7 82.6 76.5 70 Alsono 65.1 76.1 54.8 62.4 Abores 76.7 82.6 76.5 70 Alsono 67.8 68.2 70 Alsono 67.8 68.3 Alsono 68.2 70 Alsono 68.2 70 Alsono 68.2 70 Alsono 72.4 68.3 Alsono 72.9 84.9 64.2 64.7 Alsono 72.9 84.9 64.2 64.7 Alsono 72.9 84.9 64.2 64.7 Alsono 72.9 62.8 62.1 Alsono 72.9 76.4 68.1 63.2 Alsono 72.9 76.4 68.1 63.2 Alsono 72.9 76.4 68.2 Alsono 72.9 76.4 68.3 Alsono 72.9 76.4 68.3 Alsono 72.9 76.4 68.1 63.2 Alsono 72.9 76.5 66.8 66 Alsono 72.9 76.5 62.8 66 Alsono 72.9 76.1 62.9 65.1 Alsono 72.9 76.1 62.9 66.1 Alsono 73.7 76.2 64.8 68.1 Alsono 74.0 65.2 Alsono 75.9 64.8 68.1 Alsono 75.9 64.8 68.1 Alsono 76.9 64.8 68.1	mag (80.9	100	and a state of the latest	64.6		-
ASTRONAL CY.S 64.3 64.2 67.4 ASTRONAL SI 70.7 68.7 73.3 BUS B2.2 66.5 67 66.8 BUS B2.1 76.1 54.8 67.4 BUS TO. 7 82.6 76.5 75 ASTRONAL SI. 76.1 76.5 75 ASTRONAL SI. 76.1 76.5 75 ASTRONAL SI. 76.7 76.7 ASTRONAL SI. 76.7 76.7 ASTRONAL SI. 76.7 76.7 ASTRONAL SI. 76.7 76.2 ASTRONAL SI. 76.7 76.7 ASTRONAL SI. 76.7 76.2 ASTRONAL SI. 76.8 16 ASTRONAL SI. 76 ASTRON			70.8	71.	3	65.6		
SI 10.7 68.7 73.3 Sues 52.2 66.5 67 66.8 Sues 52.2 66.5 67 66.8 Sues 76.7 82.6 76.5 70 Theres 76.7 82.6 76.5 70 Theres 67 71.5 65.4 66.8 Theres 68.2 90 71.4 68.3 Theres 68.2 90 71.4 68.3 Theres 68.3 77 77.8 Theres 83.8 83.2 77 77.8 Theres 72.9 84.9 64.2 64.7 Theres 72.9 84.9 64.2 64.7 Theres 72.9 84.9 64.2 Theres 72.9 84.9 64.2 Theres 72.9 84.9 64.2 Theres 72.9 84.9 64.2 Theres 72.9 76.1 Theres 72.9 76.1 Theres 72.9 Theres 73.0 Theres 74.1 Theres 75.7 Theres 75.7	7	-		69	.2	67.4		
Secs 52.2 66.5 64 668 Colors 65.1 76.1 59.8 67.9 Colors 70.7 \$2.6 76.5 75 Colors 70.7 \$2.6 76.5 75 Colors 70.1 76.7 75 76.9 Colors 65.2 95 76.7 665 Colors 65.2 95 76.7 68.3 Colors 83.8 83.2 77 77.8 Colors 67.9 84.9 64.2 64.7 Colors 72.9 84.9 64.2 64.7 Colors 74.1 82.9 65.5 66.4 Colors 76.1 69.4 74.7 71.7 Colors 71.6 69.4 74.7 71.7 Colors 71.6 69.4 68.2 Colors 72.9 84.9 64.2 68.2 Colors 71.6 69.4 74.7 71.7 Colors 71.6 69.4 66.2 Colors 72.9 86.8 66.8 Colors 67.5 91.9 72.2 67.5 Colors 67.5 91.9 72.2 67.5 Colors 67.5 91.9 72.2 67.5 Colors 65.4 77.7 76.1 65.2 Colors 65.4 77.7 76.1 65.2 Colors 73.7 84.8 75.6 65.2 Colors 73.7 75.2 69.8 66.4 Colors 75.7 84.8 75.6 65.2 Colors 75.7 84.8 75.6 65.2 Colors 75.7 76.2 69.8 66.4 Colors 75.7 76.8 69.8 Colors 75.7 76.8 69.8 Colors 75.9 76.8 Colors 75.8 76.8 Colors 75.8 76.8 Colors 75.8 76.8 Colors 7				1 68	.7	73.3		
China's 65.1 76.1 59.8 67.9 Decres 76.7 82.6 76.5 75 Decres 95.1 76.2 75 76.7 Decres 67.7 76.2 75 76.7 Decres 68.2 90 71.4 68.3 Decres 82.8 83.2 77 77.7 Decres 67.9 84.9 64.2 64.7 Decres 72.9 84.9 64.2 64.7 Decres 72.9 84.9 64.2 64.7 Decres 76.9 68.8 68.1 67.0 Decres 76.1 68.1 67.0 Decres 76.1 68.1 67.0 Decres 76.1 68.1 68.2 Decres 76.1 68.1 68.2 Decres 76.1 68.8 66 Decres 76.1 67.0 68.8 66 Decres 76.1 67.0 68.8 66 Decres 76.1 67.0 76.1 65.2 Decres 66.4 77.7 74.1 66.2 Decres 76.7 84.8 76.0 65.2 Decres 66.4 77.7 74.1 66.2 Decres 76.7 84.8 76.0 65.2 Decres 66.4 77.7 74.1 66.2 Decres 76.7 84.8 76.0 65.2			66.5	6.	7	61.8		
14.5 75 71.5 65.4 66 14.5 75 71.5 65.4 66 14.5 75 71.5 65.4 66 14.5 75 71.5 65.4 66 14.5 75 71.4 68.3 14.5 88.2 75 71.4 68.3 14.5 83.8 83.2 77 77.8 14.5 83.8 83.2 77 77.8 14.5 83.8 83.2 77 77.8 14.5 83.8 83.2 77 77.8 14.5 83.8 83.2 77 77.8 14.5 83.8 83.2 77 77.8 14.5 83.8 83.2 77 77.8 14.5 83.8 83.2 77 77.8 14.5 83.8 83.2 77 77.8 14.5 83.8 83.2 77 77.8 14.5 83.8 83.2 77 77.8 14.5 83.8 83.2 77 77.8 14.5 83.8 83.2 77 77.8 14.5 83.8 83.2 77 77.8 14.5 83.8 83.2 77.7 14.5 83.8 83.2			76.	1 59.	8	67.9		
11.5 70 71.5 60.4 68 Tenders 67 81 70 61.7 Less 68.2 90 71.4 68.3 Less 68.2 90 71.4 68.3 Less 67 74 73 76.2 Less 67 74 73 76.2 Less 67 74 73 76.2 Less 67 74 68.1 61.6 Less 72.9 84.9 64.2 64.7 Less 72.9 84.9 64.2 64.7 Less 75.9 76.4 68.1 61.6 Less 76 76.4 68.1 61.6 Less 76 76.4 68.1 68.2 Less 69.7 86.5 66.8 66 Less 69.7 86.5 66.8 66 Less 69.7 76.7 76.2 Less 69.7 76.1 65.5 Less 69.7 76.1 60.5 Less 69.7 77.7 76.2 Less 75.7 84.8 70.6 65.2	pracis	76.7	82.0	76.	5	70		
115 70 71.5 65.4 68 TOTALS 67 81 70 61.7 Loss 68.2 90 71.4 68.3 Loss 83.8 83.2 77 77.8 Loss 67 74 73 76.2 Loss 72.9 84.9 64.2 64.7 25.9 75.9 65.5 66.4 25.9 76.1 68.1 61.6 25.9 76.1 68.1 61.6 25.9 76.1 69.4 74.7 26.7 76.1 68.2 26.8 64.6 63.2 26.8 64.6 63.2 26.8 64.6 63.2 26.8 64.6 63.2 26.8 64.6 63.2 26.8 64.6 63.2 26.8 64.6 63.2 26.8 65.4 77.7 74.1 66.2 26.4 77.7 74.1 68.1 26.4 77.7 74.1 68.1 26.4 77.7 74.1 68.2 26.4 77.7 74.1 68.1 26.4 77.7 74.1 68.2 26.4 77.1 74.1 68.2 26.4 77.1 74.1 68.2 26.4 77.1 74								
10 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1		70	71.5	60.	4	68		
1695 68.2 93 71.4 68.3 1 1011 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	montag	67	81	75	16	1.7		****
10 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1	-6416		95	71.4	68	3.3		
1605 67 74 73 76.2 215 72.9 84.9 64.2 64.7 227 74.1 82.9 65.5 66.4 23 75.9 76.4 68.1 61.6 24 75.9 10.1 67.9 68.7 24 10.1 67.9 68.2 24 25.5 66.8 66 24 25.5 66.8 66 24 25.5 66.8 66 24 25.5 67.6 67.2 67.5 25 67.6 67.6 67.2 74.1 26 70.2 74.1 27 70.2 66.3 70.2 74.1 27 70.3 70.2 74.1 28 18 70.6 65.2 28 18 70.6 65.2 28 18 70.6 65.2 28 18 70.6 65.2		-	74.4	68.4	6	.3		
1625 67 74 73 76.2 1625 72.9 84.9 64.2 64.7 1625 74.1 82.9 65.5 66.4 163 75.9 76.4 68.1 61.6 163 75.9 76.4 68.1 61.6 164 74.1 74.1 71.7 164 72 12.1 67.9 68.2 165 69.8 61.6 63.2 165 69.7 86.5 69.8 66 165 69.7 86.5 69.8 66 165 67.6 67.6 67.2 74.1 166 65.2 75.1 167 75.1 80 75.2 74.1 168 75.1 84.8 72.6 65.2 168 75.1 84.8 72.6 65.2 169 73 75.1 64.8 68.4	9175	83.8	83.2	77	77	7. 7		
72.9 84.9 64.2 64.7 74.1 82.9 65.5 66.4 75.9 76.4 76.1 68.1 61.6 75.9 71.6 69.4 74.1 71.7 76.10 69.4 74.1 68.7 76.10 69.8 64.6 63.2 11.10 69.7 81.5 60.8 66 11.10 69.7 81.5 60.8 66 11.10 69.7 81.5 67.5 71.5 11.10 69.7 69.9 65.5 11.10 69.4 77.7 74.1 61.2 11.10 69.4 77.7 74.1 61.2 11.10 69.4 75.2 69.8 68.1 11.10 69.4 75.2 69.8 68.1 11.10 69.4 75.2 69.8 68.1 11.10 69.4 75.2 69.8 68.1		1000	74	73	7	6.2		
75.9 76.4 68.1 67.6 75.9 76.4 68.1 67.6 25.7 76.4 68.1 67.6 26.8 72.4 10.1 67.9 68.7 26.8 64.6 63.2 21.20 75 81.5 68.8 66 21.20 67.8 16.9 65.5 21.21 67.5 21.22 67.6 22.2 67.8 66.4 23.2 66.8 66.8 24.2 66.8 66.8 25.2 66.8 66.9 26.3 76.7 77.7 74.2 61.2 26.3 75.7 84.8 70.6 65.2 26.3 75.7 84.8 70.6 65.2 26.3 75.7 84.8 70.6 65.2		72.9	84.9	64.2	64	.7		
75.9 76.4 68.1 61.6 75.9 71.6 69.4 74.1. 71.7 76.414 72.1 67.9 68.7 11.015 69.7 81.5 68.8 66 11.015 69.5 91.9 73.2 69.5 11.015 67.6 67 68.9 66.5 11.015 67.6 67 68.9 66.5 11.015 65.4 77.7 74.2 61.2 11.015 65.4 77.7 74.2 61.2 11.015 75.4 84.8 70.6 65.2 11.015 75.4 84.8 70.6 65.2		74.1	82.9	65.5	1 66	.4		
75 91.6 69.4 74.1 71.7 76.6 724 10.1 67.9 68.7 1000 75 91.8 64.6 63.2 1000 69.7 81.5 68.8 66 1000 69.5 91.9 72.2 67.5 1000 67.6 67 68.9 65.5 1000 70.2 74.1 1000 70.2 74.1 1000 70.2 74.1 1000 70.2 74.1 1000 70.2 74.1 1000 70.2 74.1 1000 70.2 74.1 1000 70.2 74.1 1000 70.2 74.1 1000 70.2 74.1 1000 70.2 70.2 1000 70.2 70.2 1000 70.2 70.2 1000 70.		75.7	76.4	68.1	61.	6		
75 61.8 64.6 63.2 11005 69.7 81.5 68.8 66 11005 69.5 91.9 73.2 69.5 11005 67.6 67 68.9 65.5 11005 67.6 67 70.2 74.1 11005 65.4 77.7 74.2 61.2 11005 75.4 84.8 70.6 65.2 11005 75.4 84.8 70.6 68.4 11005 75.4 64.8 68.4		71.6	67.4					
1105 69.7 81.5 68.8 66 1105 69.7 81.5 68.8 66 1105 67.5 91.9 72.2 69.5 125 67.6 67 68.9 65.2 126 65.4 77.7 7-1.2 61.2 126 73 75.7 84.8 70.6 65.2 126 73 75.2 69.8 68.7	ochers	721	10.1	Phillips of Contract Contract				
1105 69.7 81.5 01.8 66 1105 67.5 01.9 73.2 67.5 120 67.6 67 69.9 65.5 120 70.7 80 70.2 74.1 120 65.4 77.7 71.2 61.2 120 73 75.7 84.8 70.6 65.2 120 73 75.2 69.8 68.7	11/205	75	81.8			-		
16.5 41.9 75.2 66.8 68.4 75.2 66.8 68.2 66.8 66.8 66.8 66.8 66.8 66		69.7	81.5	1		_		112
10.5 70.7 80 70.2 74.1 10.5 65.4 77.7 74.2 61.2 10.6 65.4 77.7 74.2 61.2 10.6 65.2 10.6 65.2 10.6 65.2 10.6 65.2 10.6 66.8 66.1 10.6 65.2		67.5	1					
75.7 74.2 66.2 66.4 77.7 74.2 66.2 66.4 75.7 84.8 70.6 65.2 66.8 68.7 75.2 69.8 68.7 75.2 69.8 68.7		67.6	- Annual Contract of the Contr	Contract Services and Services				
75.7 84.8 70.6 65.2 106.645 73 75.2 69.8 68.71						Name and Address of the Owner, Street, or other		-
160 73 75.2 69.8 68.1 160 73 75.2 69.8 68.1	.ak	and the same of th					The state of the s	
16 13 15 15 15 15 15 15 15 15 15 15 15 15 15								1
	reces		84.8	65		1		
71.4 86.7 68.6 58.6	12%	Contract of the Contract of th	The second secon	-	-	A PART TRANSPORT CONTRACTOR OF THE		A.
110x 77.1 84.2 61.7 71.1	1100	1		67.7	71.	1		H

Previous Season's Wins: Although the model will only be predicting on the amount of games a team will win in the regular season, I am inputting even the playoff wins, because it's a good indicator of how the team will perform for the future season. For example, the *Green Bay Packers*, the youngest team in the postseason, only won 9 games during the regular season, just finishing with a winning record, but they defeated the Dallas Cowboys in the first round of the playoffs, which was the largest playoff upset last postseason. I believe that win is imperative for the model to know, it is a strong indicator of what is to come of a better, more experienced Packers team.

2023 PFF Offense: To determine the offensive PFF score of each team, I took a weighted average of the four main offensive position groups. QB (40%), RB (15%) WR (20%) and OL (25%).

To find the PFF score for the **quarterback** position, I just found the starters PFF passing grade, but if I determined that he's a "mobile" quarterback, I took a weighted average of his passing grade (80%) and rushing grade (20%).

To find the PFF score for the **running back** position, I found the starters PFF rushing grade, but if I determined that the team used a committee at the running back position, then I used a weighted average of the RB1 (75%) and RB2 (25%)

To find the PFF score for the **wide receiver** position, I took a straight average of the three starting receiver's PFF receiving grade, the starting running back's receiving grade, and the tight end's receiving grade

To find the PFF score for the **offensive line** position, I took a straight average of the five starting offensive lineman's overall blocking grade.

Once I found all of the individual position grades, I computed the weighted average mentioned above, and each team was given an offensive grade.

2023 PFF Defense: To determine the defensive PFF score of each team, I took a weighted average of the four main defensive position groups. Edge Rusher (25%), IL (15%), LB (24%), CB (20%), S (16%).

To find the PFF score for the **Edge Rusher** position, I took a straight average of the two starting edge rushers.

To find the PFF score for the **Interior Lineman** position, I took a straight average of the two starting interior lineman.

To find the PFF score of the **linebacker** position, I took a straight average of the three starting linebackers.

To find the PFF score for the **cornerback** position, I took a straight average of the two starting cornerbacks, and the nickel cornerback.

To find the PFF score for the **safety** position, I took a straight average of the two starting safeties.

Once I found all of the individual position grades, I computed the weighted average mentioned above, and each team was given a defensive grade.

2023 PFF Average: I took a straight average of the scores given for 2023 PFF Offense and 2023 PFF Defense for every team.

2023 Opponent PFF Average: Now that I have the average PFF score for every team, I can find the average PFF score of everyone's opponents. For example, let's say the Pittsburgh Steelers played three games this season. They play the Jets, Giants, and Panthers. I would take an average of the 2023 PFF score of those three teams. Whatever that number turns out to be, will be the 2023 Opponent PFF Average. So for all 32 teams, I am finding the average PFF score of all 17 of their opponents, and averaging all of those scores.

2023 Metric: The metric is simple, yet it is the most important part of this project. This is what I am predicting the 2024 season outcome on. The metric is just the average PFF score of any team, minus the average PFF score of all of their opponents. The formula is **2023 PFF Average** - **2023 Opponent PFF Average.** This will give me an output for every team that I will use to predict the amount of games every team will win in 2024.

2024: The process for the 2024 columns are the same as 2023.

Now that we know how the data works. let's talk about the model.

The Model:

A linear regression model is being used. We are predicting on the 2024 metric, and we are predicting the amount of games a team will win in 2024. We look at the 2023 metric, and how many games each team won in 2023, then we will use those same data points to predict the 2024 results. The model was coded in R, and I hope to code it in Python in the near future.

Areas of Improvement:

The biggest issue I see with this project is the lack of data. If I am trying to predict how many games a team will win, I should be looking at much more data than just one season's worth. The issue is, gathering one season's worth of data takes a tremendously long time, especially when I am working a 9 to 5 everyday. Also, the NFL recently switched from a 16 game season to a 17 game season, meaning the data will be skewed once we reach the 16 game seasons. I plan on

improving this for future seasons by making the 2024 data into training data for the 2025 season, and then making the 2025 data into training data for the 2026 season, etc etc.

Determining the Credibility of the Model:

To determine whether the results from my model are valid, I am comparing them to the DraftKings team wins over/under line. If my model is within half a game of the DraftKings line I will give myself 3 points, if my model is within 1.5 games of the DraftKings line, I will give myself 1 point.

Team	Model	DraftKings	Points Received
49ers	12	11.5	3
Bears	9	8.5	3
Bengals	8	10.5	0
Bills	9	10.5	1
Broncos	6	5.5	3
Browns	9	8.5	3
Buccaneers	7	7.5	3
Cardinals	6	6.5	3
Chargers	8	8.5	3
Chiefs	10	11.5	1
Colts	7	8.5	1
Commanders	7	6.5	3
Cowboys	12	10.5	1
Dolphins	11	9.5	1
Eagles	9	10.5	1
Falcons	10	9.5	3
Giants	6	6.5	3
Jaguars	9	8.5	3
Jets	11	9.5	1
Lions	11	10.5	3
Packers	8	9.5	1
Panthers	7	5.5	1
Patriots	8	4.5	0
Raiders	8	6.5	1

Rams	9	8.5	3
Ravens	12	10.5	1
Saints	10	7.5	0
Seahawks	10	7.5	0
Steelers	9	8.5	3
Texans	8	9.5	1
Titans	7	6.5	3
Vikings	8	6.5	1

75% of my model's predictions were within at least 1.5 games of Draftkings' predictions, scoring at least 1 point. 47% of my model's predictions were within half a game of DraftKings' predictions, scoring 3 points. Out of the 96 points I could have scored, I scored 58, achieving 60% accuracy.

This is much better than I expected. The folks at draftKings have access to billions of more data points than I do. They are constantly working on their predictive models, and they are all far more experienced than me, it is literally their job to do this. I am a chump compared to these guys. I had access to one data set (which I created myself), I used one of the most basic types of models, and I actually paid to work on this, only after I got home from my 9 to 5 as an actuary. My model being 60% as "correct" as the DraftKings model is a testament to the creativity and sharpness of the model.

Acknowledgements:

I would like to thank my good friend Jonah Bierig for his help gathering the data and PFF scores necessary for the completion of the project. I would not have been able to finish it without his help. <u>Jonah's LinkedIn</u>

Code:

title: "NFL Win Predictor" author: "Joshua Eisner" date: "07/16/2024"

```{r}

# Load necessary libraries library(readr) library(ggplot2)

# Load the dataset

```
nfl data = read.csv("~/NFL Data CSV.csv")
Display column names to check for exact matches
print(names(nfl data))
Rename columns to ensure there are no issues with spaces or special characters
names(nfl data) = make.names(names(nfl data))
View the renamed columns
print(names(nfl data))
Convert relevant columns to numeric
nfl data$X2023.Season.Wins = as.numeric(nfl data$X2023.Season.Wins)
nfl_data$X2023.Metric = as.numeric(nfl_data$X2023.Metric)
nfl data$X2024.Metric = as.numeric(nfl data$X2024.Metric)
Create the linear model using 2023 data
model 2023 = Im(X2023.Season.Wins ~ X2023.Metric, data = nfl data)
Summary of the model
summary(model 2023)
Predicting the wins for the 2024 season using 2024 metric
nfl data$Predicted Wins 2024 = predict(model 2023, newdata = data.frame(X2023.Metric =
nfl data$X2024.Metric))
Round the predicted wins to 2 decimal places
nfl data$Predicted Wins 2024 = round(nfl data$Predicted Wins 2024, 2)
View the results
print(nfl data)
Plot the data and the regression line using 2023 data
ggplot(nfl data, aes(x = X2023.Metric, y = X2023.Season.Wins)) +
 geom_point(color = "blue", label = "2023 Actual Wins") + # Scatter plot for 2023
 geom smooth(method = "Im", col = "blue") + # Regression line for 2023
 geom_point(aes(x = X2024.Metric, y = Predicted_Wins_2024), color = "red", label = "2024
Predicted Wins") + # Scatter plot for 2024 predictions
 #geom_text(aes(label = Team, x = X2024.Metric, y = Predicted_Wins_2024), vjust = 1, color =
"black", size = 3) + # Labels for 2024 predictions
 labs(title = "2023 Season Wins vs 2023 Metric with 2024 Predictions",
 x = "Metric",
 y = "Wins") +
 theme minimal() +
```

```
scale_color_manual(values = c("blue", "red")) +
theme(legend.position = "top")

Output the predicted wins for each team in 2024 to the console, sorted by
Predicted_Wins_2024 in descending order
predicted_wins = nfl_data[, c("Team", "Predicted_Wins_2024")]
predicted_wins = predicted_wins[order(-predicted_wins$Predicted_Wins_2024),]
print(predicted_wins)
```