

# Strategy for the Association of Tennis Professionals (ATP)

## Tournament Play

Ian Freed

# Agenda



Project Overview



ATP Tennis Data



Exploratory Data Analysis



Regression Modeling



Recommendations

# Project Overview

## Problem Statement

# Problem Statement

How should men's professional tennis players focus their training to yield the best possible outcomes (winnings, wins) for yearly tournament play?

# ATP Tennis Data

Scraping and Cleaning

# Scraped stats from top 1500 ranked players in April 2018 from 2009 through 2017

Ranking ^	Move ^	Country ^	Player ^	Age ^	Points ^	Tourn Played ^	Points Dropping ^	Next Best ^
1	-		Rafael Nadal	31	8,770	14	1,000	0
2	-		Roger Federer	36	8,670	17	0	0
3	-		Marin Cilic	29	4,985	20	180	45
4	-		Alexander Zverev	20	4,925	24	0	0
5	-		Grigor Dimitrov	26	4,635	22	0	0
6	-		Juan Martin del Potro	29	4,470	20	0	0
7	-		Dominic Thiem	24	3,665	25	90	45
8	-		Kevin Anderson	31	3,390	22	0	0
9	-		John Isner	32	3,125	25	0	0
10	-		David Goffin	27	3,110	24	360	45



Scores Stats Rankings Players Tournaments News Video Photos

Watch Listen

Rafael Nadal

Overview Bio Activity Win/Loss Titles and Finals Player Stats Rankings History Rankings Breakdown

Infosys ATP SCORES & STATS

2017

All Surfaces

#### Singles Service Record

Aces	286
Double Faults	123
1st Serve	68%
1st Serve Points Won	74%
2nd Serve Points Won	61%
Break Points Faced	340
Break Points Saved	70%
Service Games Played	939
Service Games Won	89%
Total Service Points Won	70%

#### Singles Return Record

1st Serve Return Points Won	35%
2nd Serve Return Points Won	56%
Break Points Opportunities	718
Break Points Converted	41%
Return Games Played	908
Return Games Won	33%
Return Points Won	43%
Total Points Won	56%

#### 2017 Player Activity

Singles

Doubles

2017

Tournament

Go



SGL

67-11  
W-L

6  
Titles

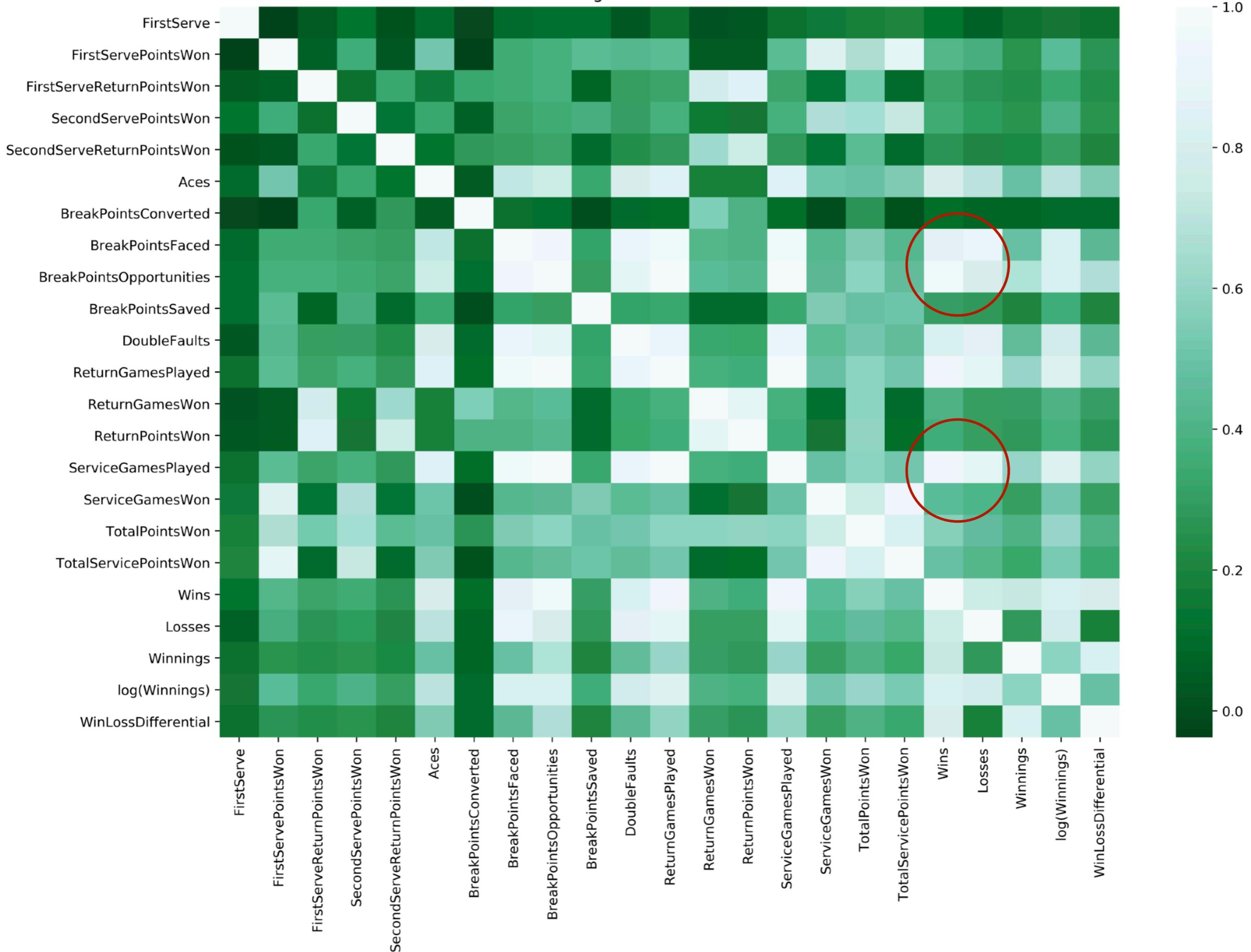
\$12,691,341  
Prize Money

- 18 different features
- 4 outcomes
- 1,893 player-years with data

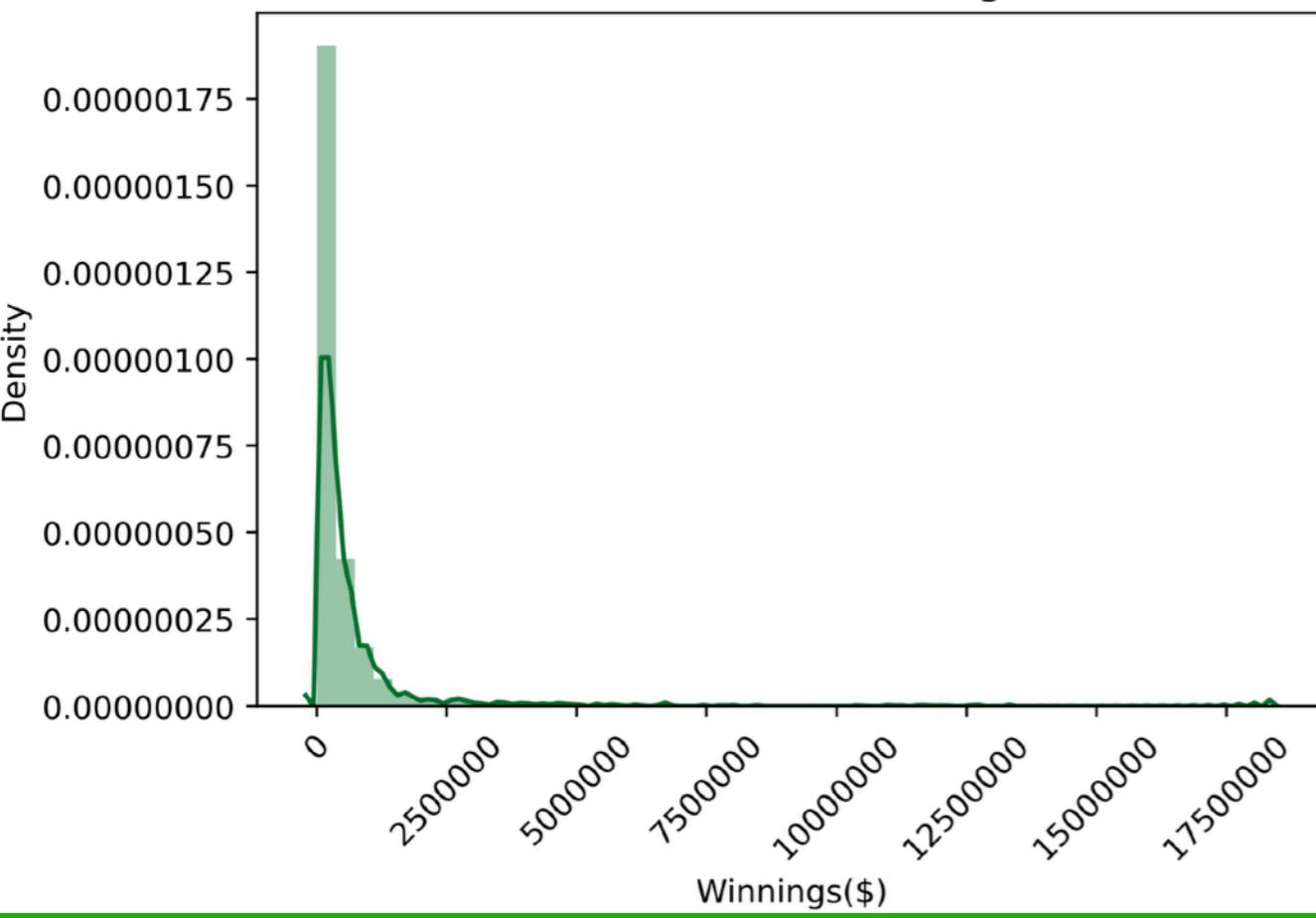
# Exploratory Data Analysis

Insights and Observations

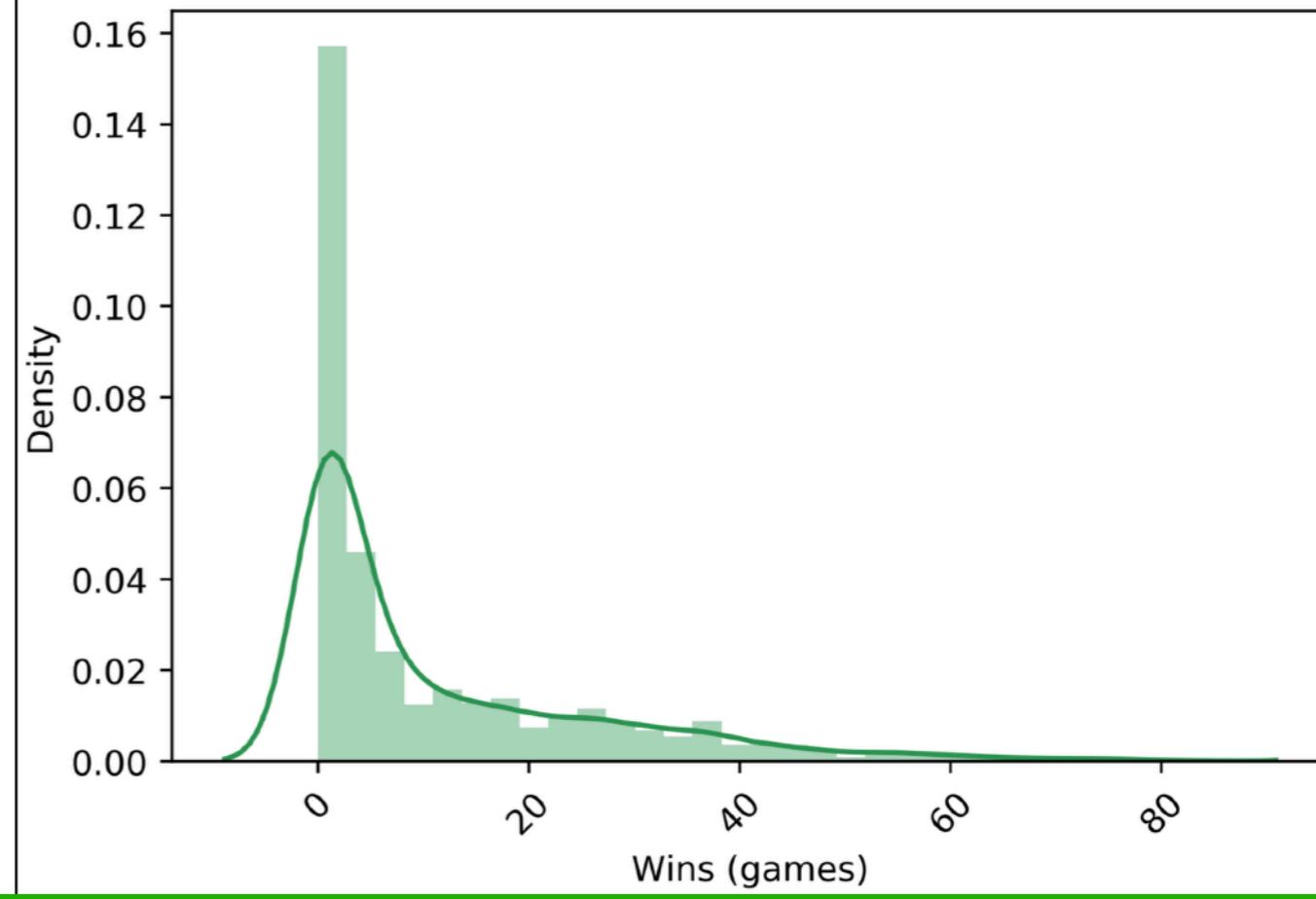
Correlation Amongst Features and Outcome Variables



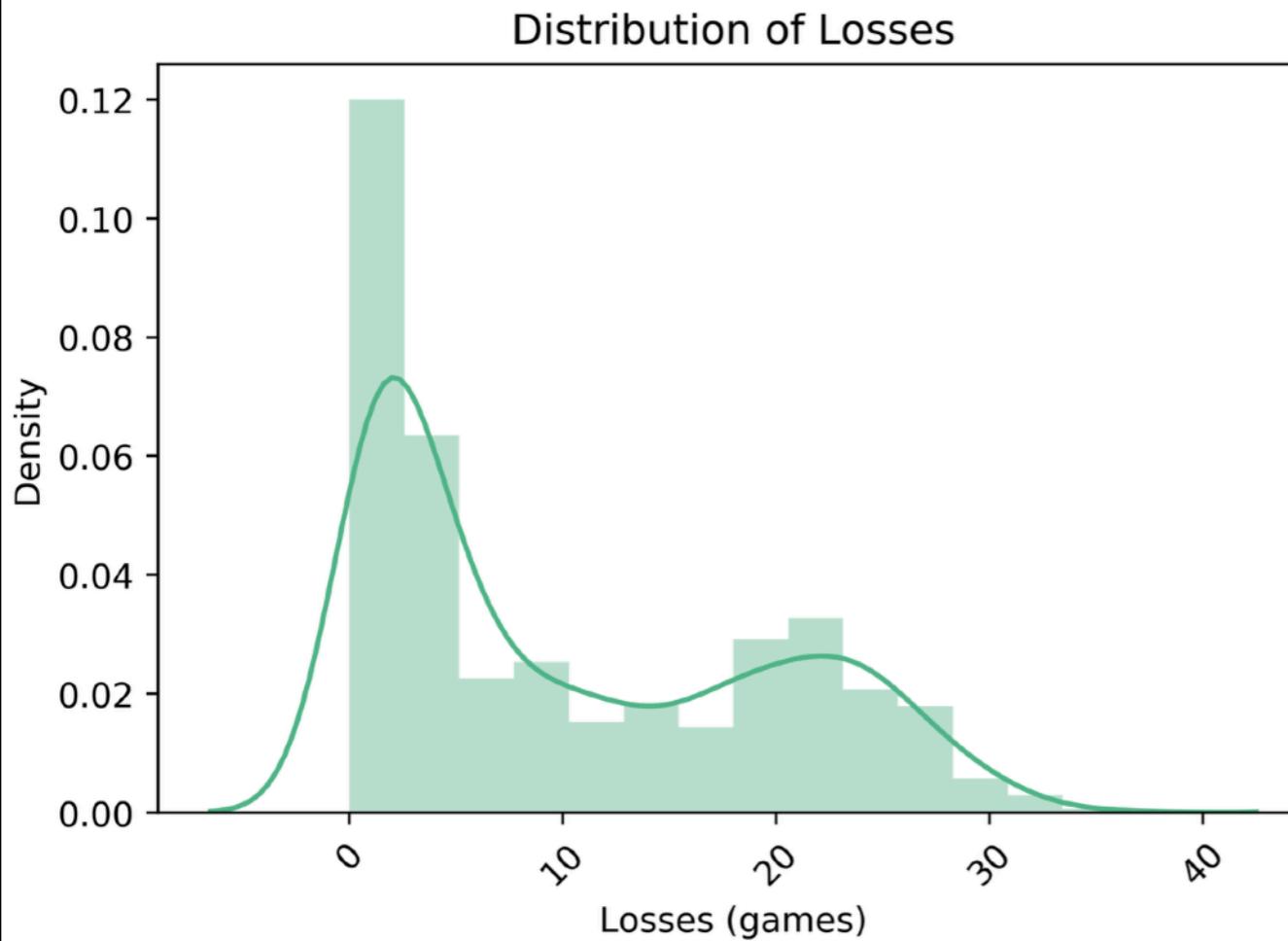
Distribution of Winnings



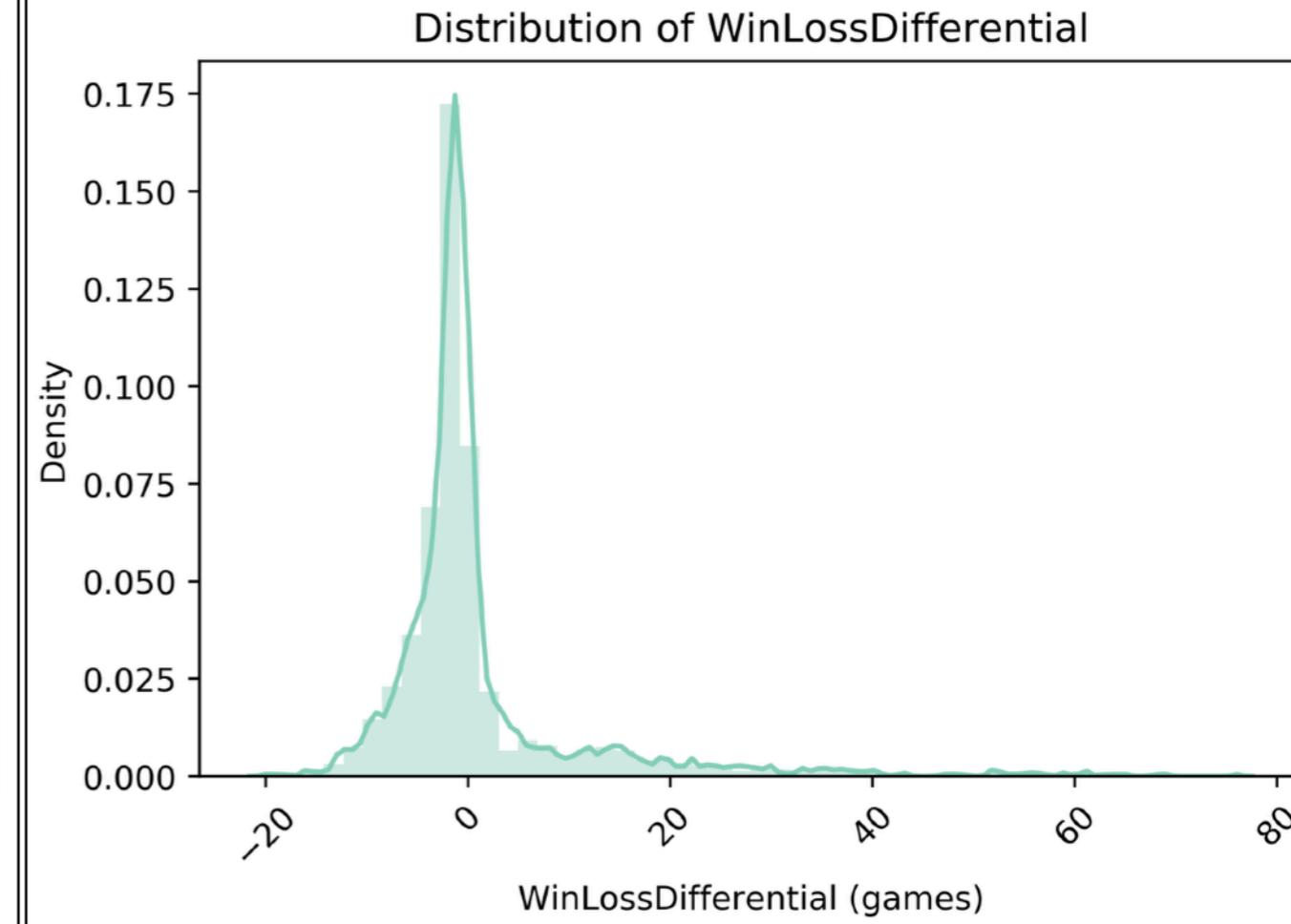
Distribution of Wins



Distribution of Losses



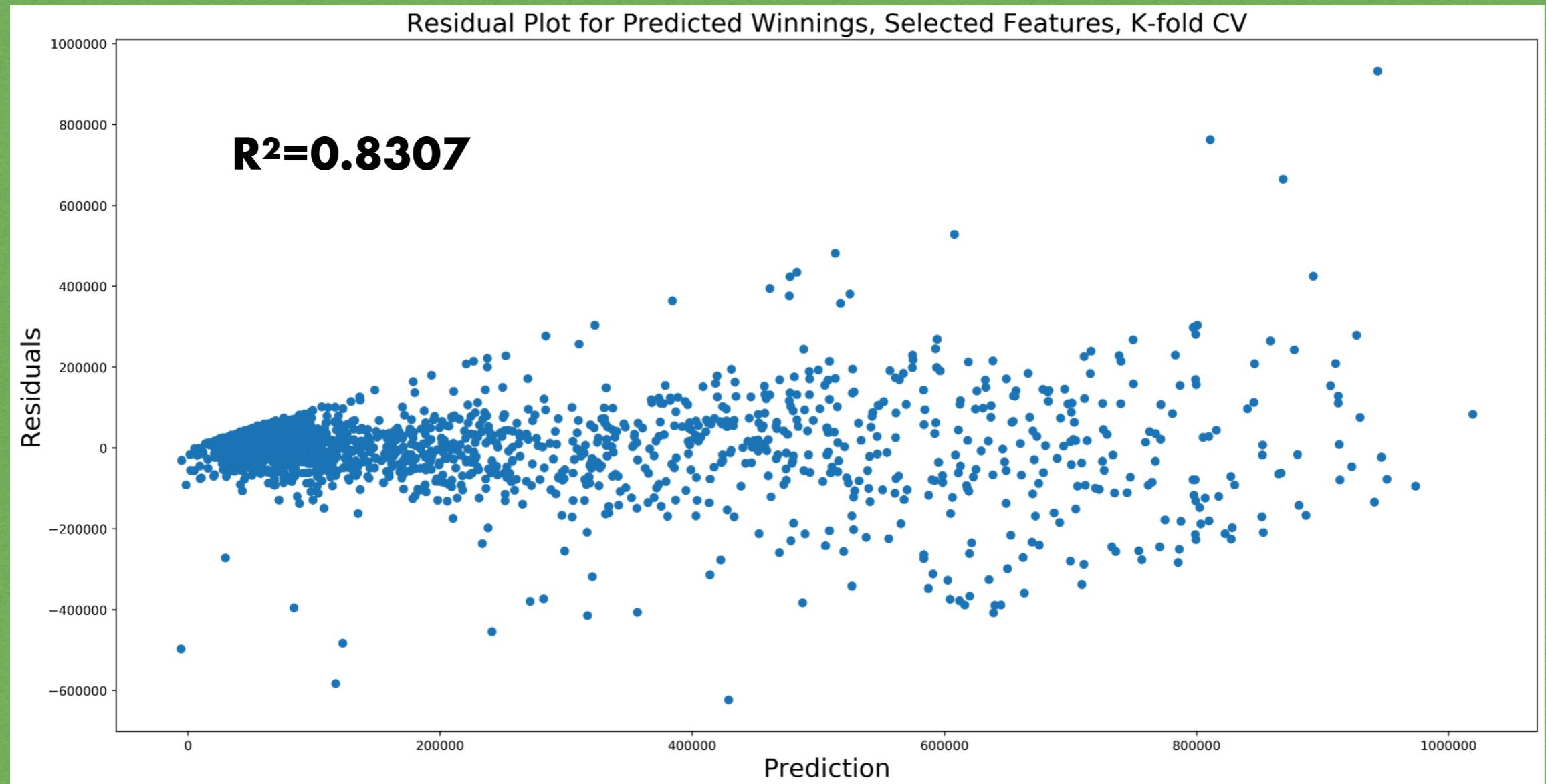
Distribution of WinLossDifferential



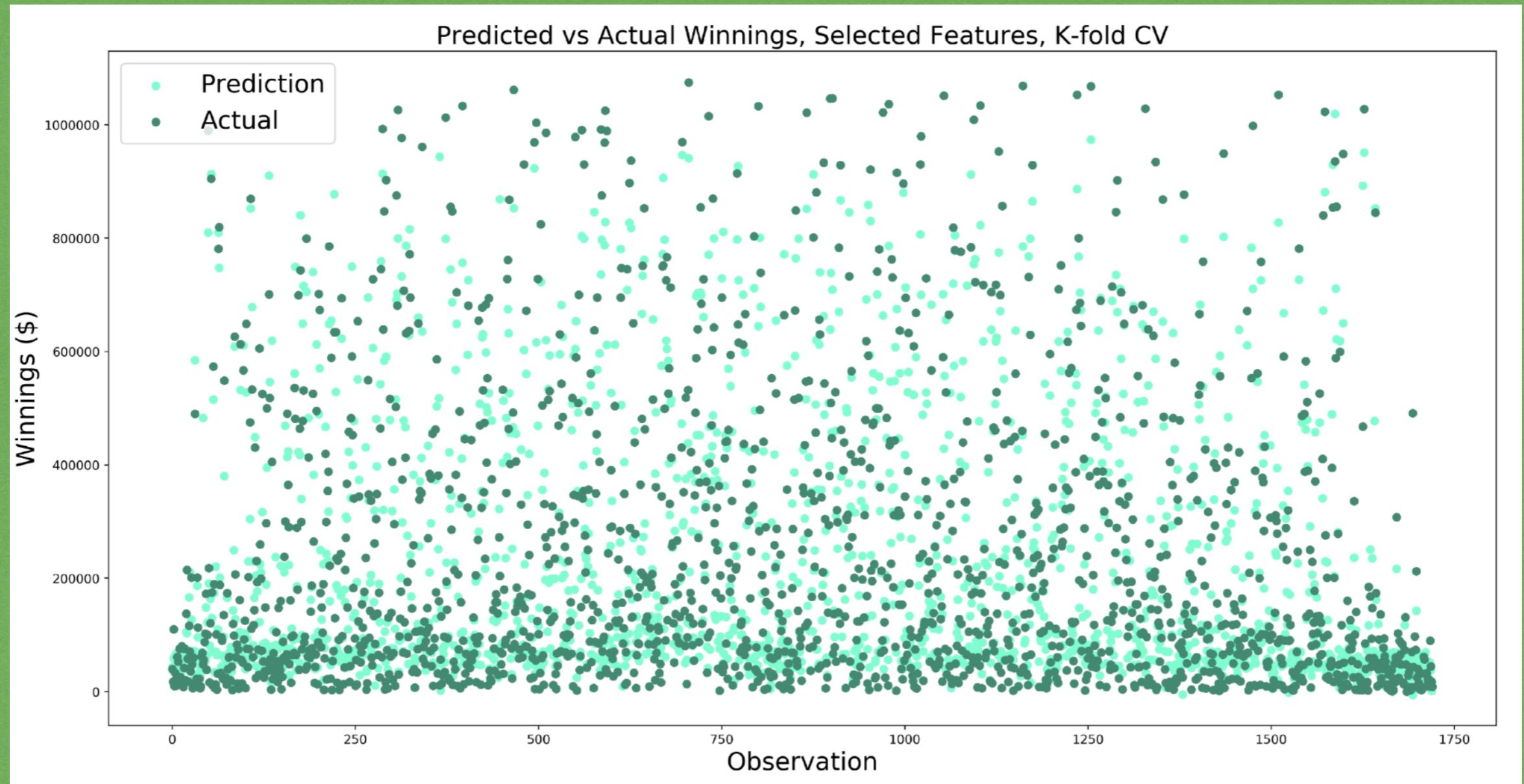
# Regression Modeling

Winnings, Losses, Wins and Win Loss Differential

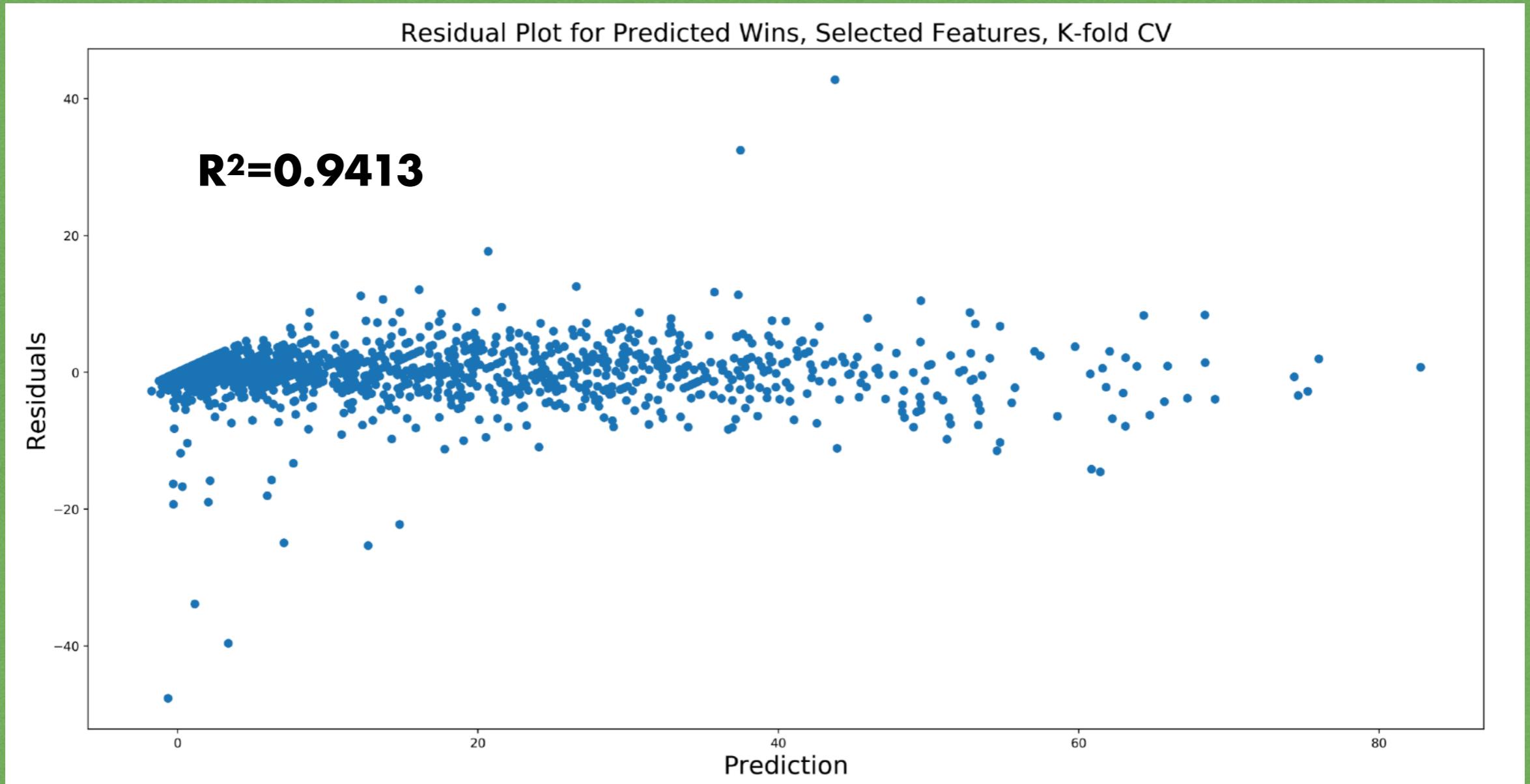
# Winnings Residuals



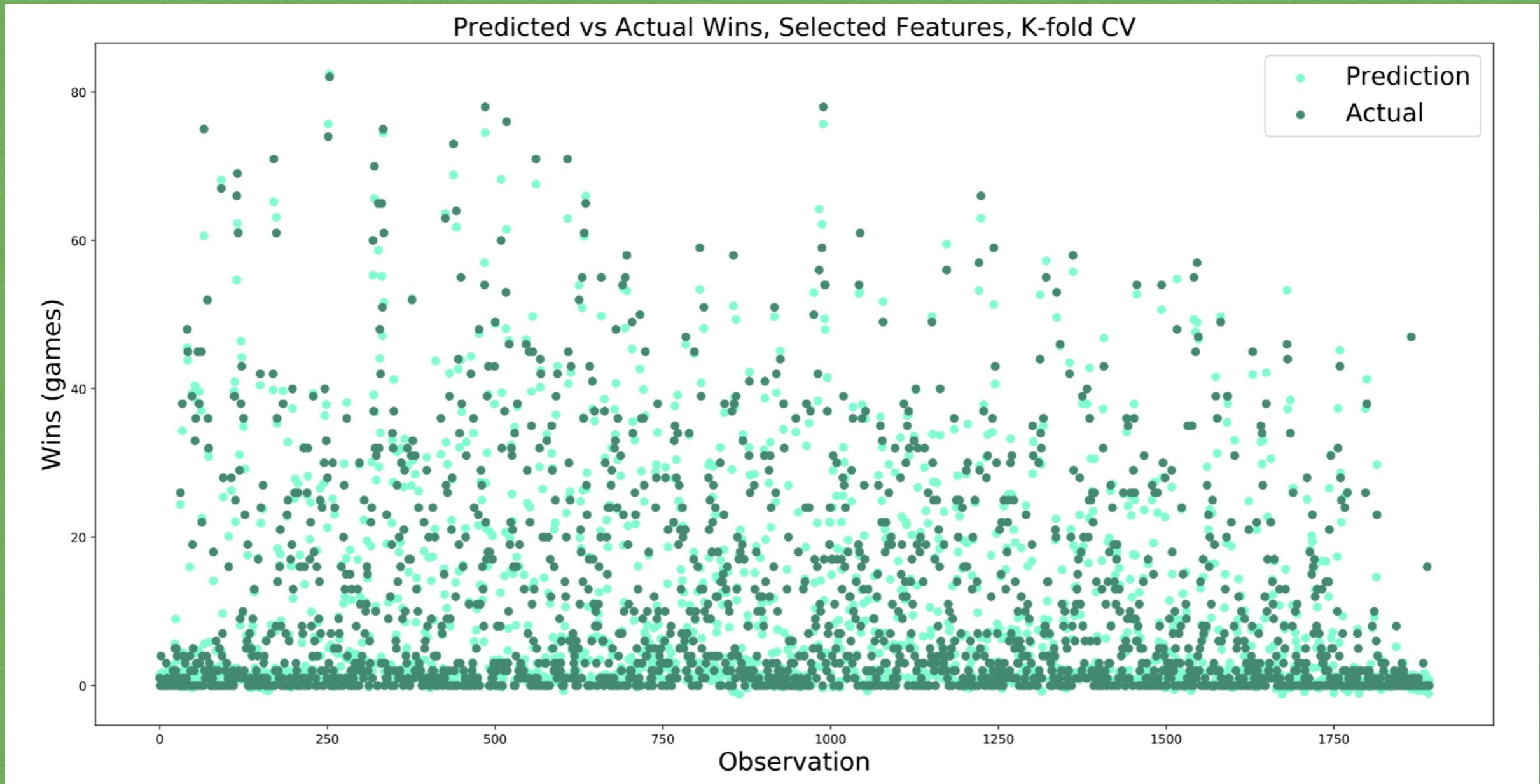
# Predicted vs Actual Winnings



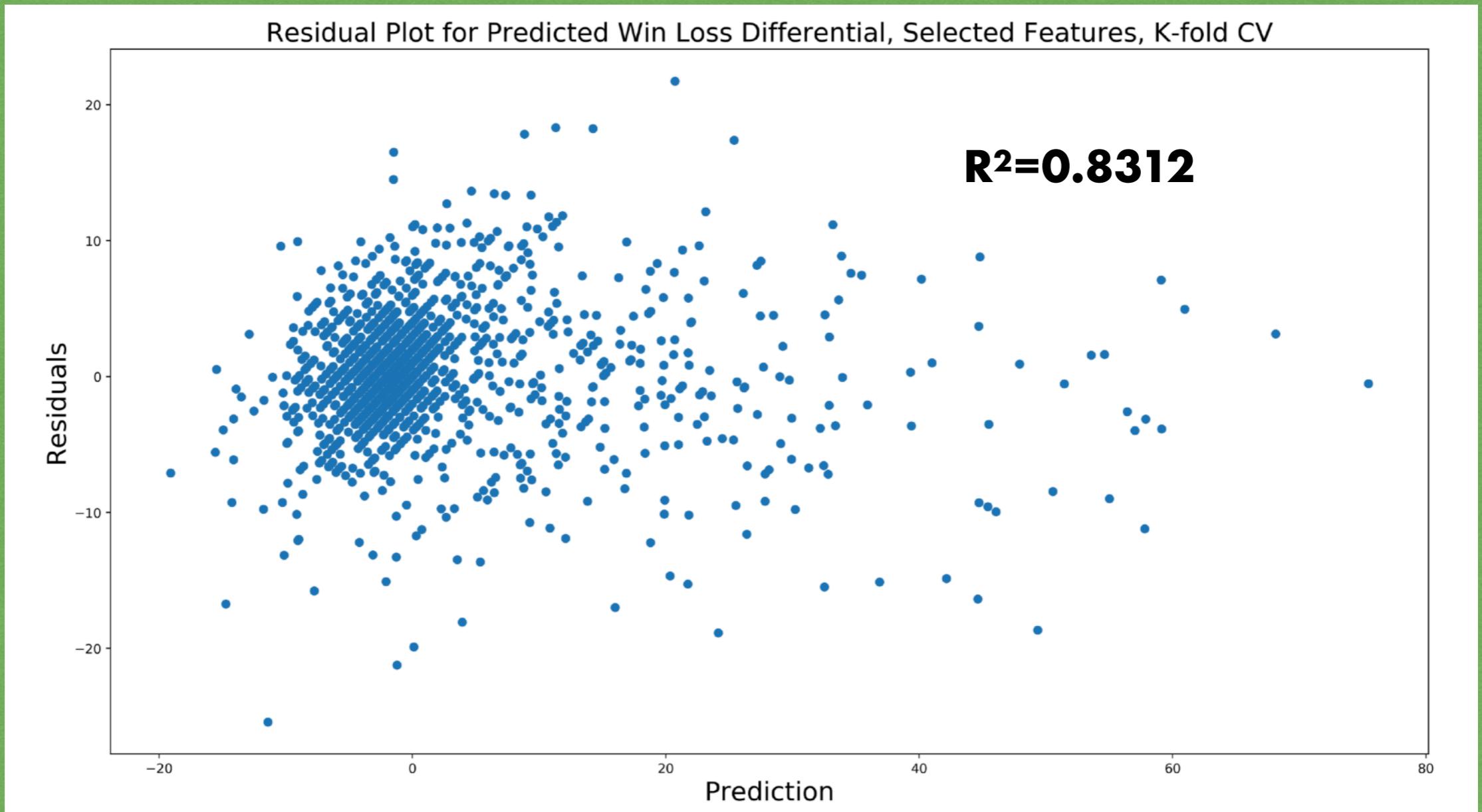
# Wins Residuals



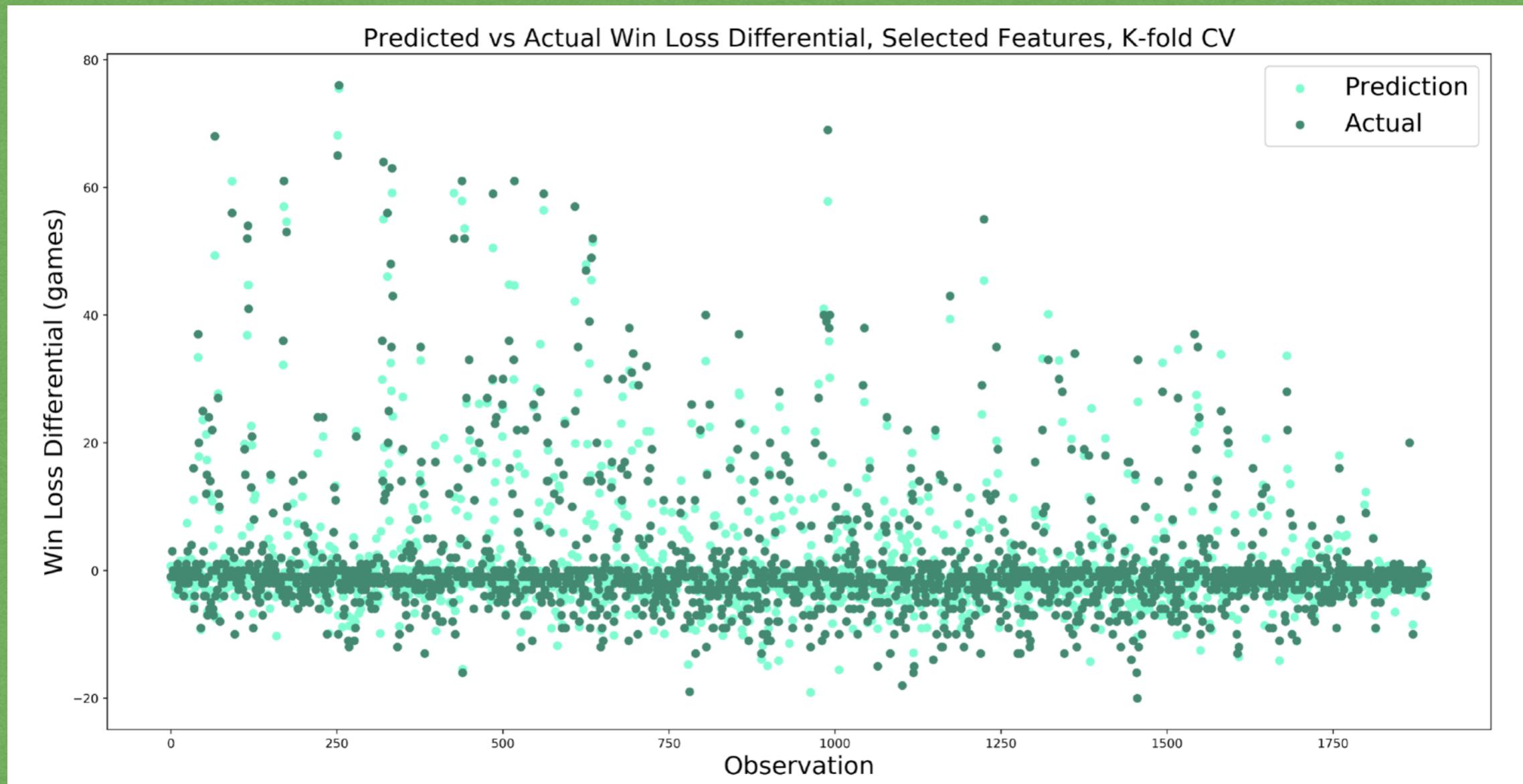
# Predicted vs Actual Wins



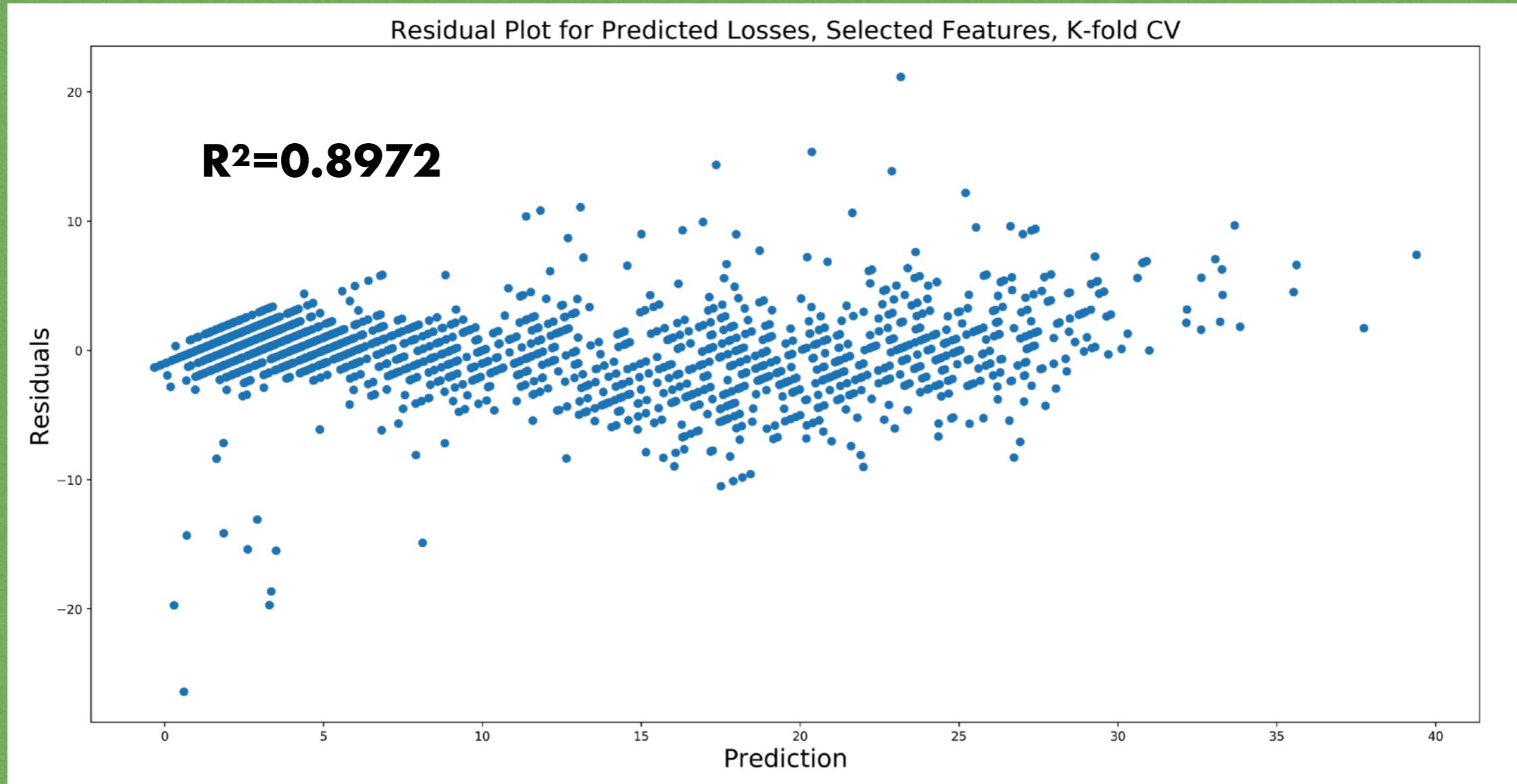
# Win Loss Differential Residuals



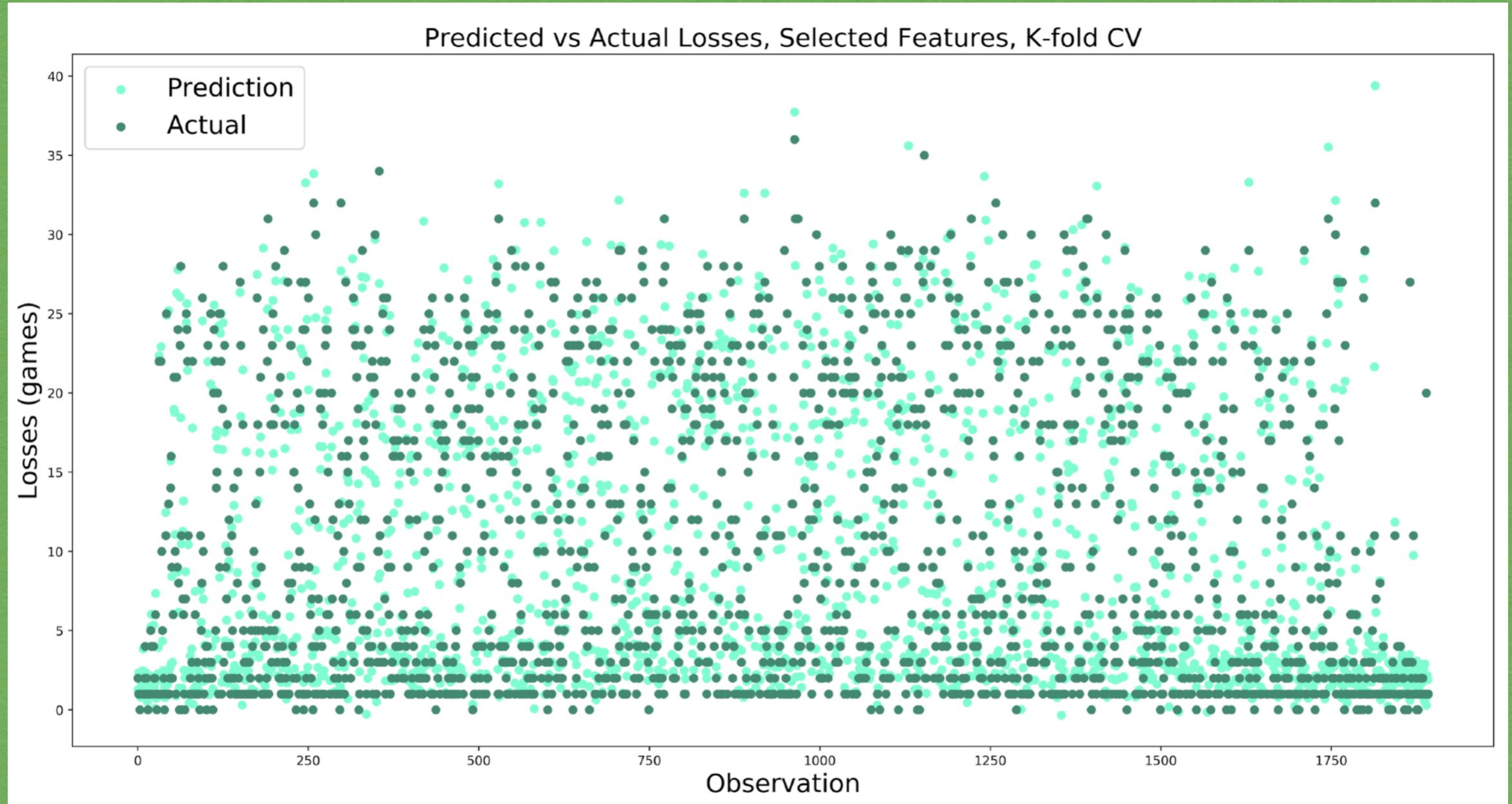
# Predicted vs Actual Win Loss Differential



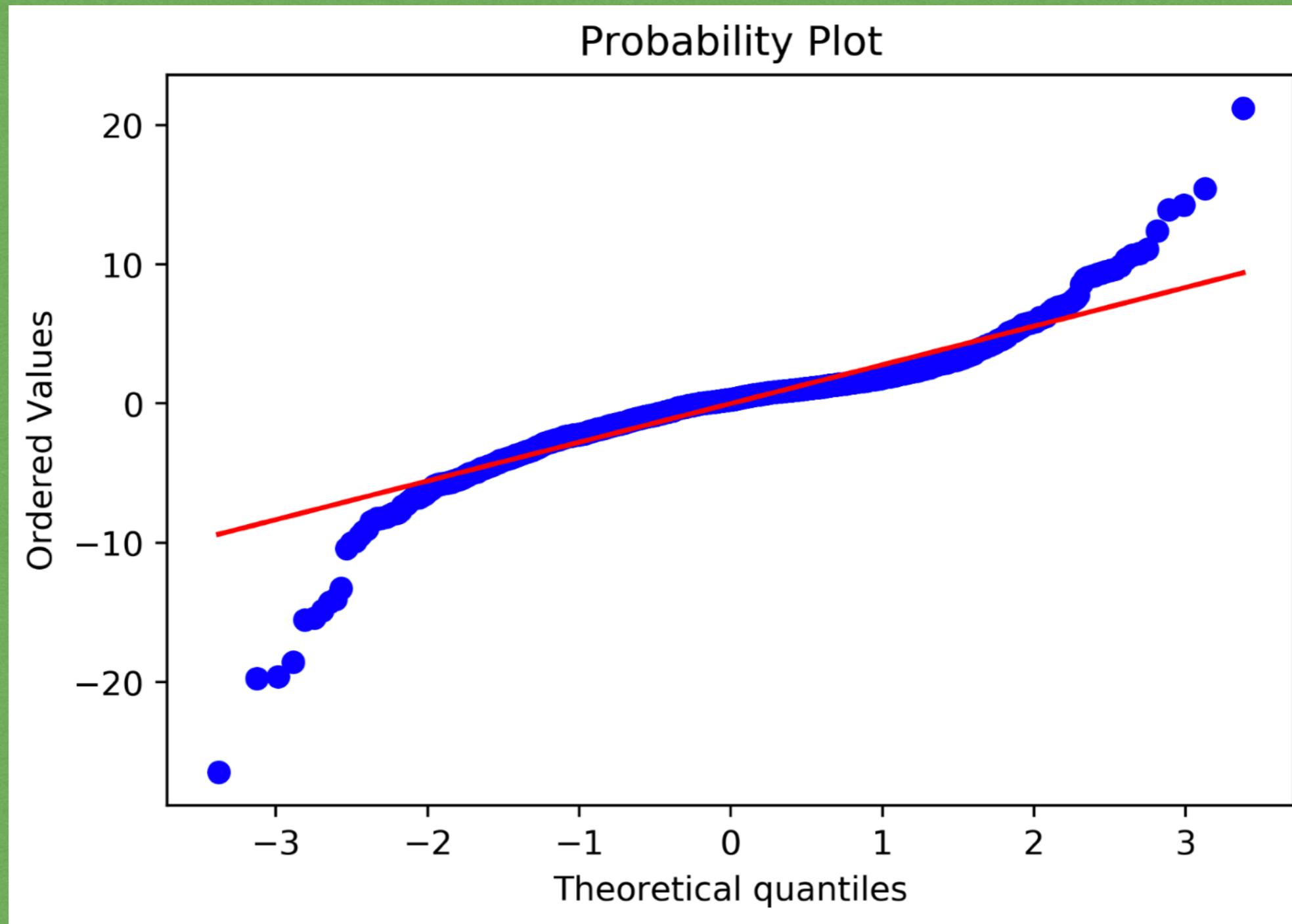
# Losses Residuals



# Predicted vs Actual Losses



# Q-Q Plot



# Statsmodels Regression

## OLS Regression Results

<b>Dep. Variable:</b>	Losses	<b>R-squared:</b>	0.898
<b>Model:</b>	OLS	<b>Adj. R-squared:</b>	0.897
<b>Method:</b>	Least Squares	<b>F-statistic:</b>	766.4
<b>Date:</b>	Fri, 27 Apr 2018	<b>Prob (F-statistic):</b>	0.00
<b>Time:</b>	09:17:47	<b>Log-Likelihood:</b>	1423.2
<b>No. Observations:</b>	1325	<b>AIC:</b>	-2814.
<b>Df Residuals:</b>	1309	<b>BIC:</b>	-2731.
<b>Df Model:</b>	15		
<b>Covariance Type:</b>	nonrobust		

	coef	std err	t	P> t	[0.025	0.975]
Intercept	0.0400	0.053	0.751	0.453	-0.064	0.144
FirstServe	-0.0124	0.043	-0.285	0.775	-0.097	0.073
FirstServePointsWon	0.0565	0.132	0.429	0.668	-0.202	0.315
FirstServeReturnPointsWon	0.0118	0.063	0.186	0.853	-0.112	0.136
SecondServePointsWon	0.0020	0.126	0.016	0.987	-0.245	0.249
SecondServeReturnPointsWon	-0.0153	0.065	-0.234	0.815	-0.144	0.113
Aces	0.3747	0.034	11.183	0.000	0.309	0.440
BreakPointsConverted	-0.0399	0.019	-2.084	0.037	-0.077	-0.002
BreakPointsFaced	1.5609	0.036	43.535	0.000	1.491	1.631
BreakPointsOpportunities	-0.9973	0.040	-24.961	0.000	-1.076	-0.919
BreakPointsSaved	-0.1410	0.024	-5.786	0.000	-0.189	-0.093
DoubleFaults	-0.0084	0.031	-0.275	0.783	-0.068	0.052
ReturnGamesWon	-0.0203	0.040	-0.502	0.615	-0.100	0.059
ReturnPointsWon	0.0131	0.114	0.115	0.909	-0.211	0.237
ServiceGamesWon	0.0112	0.053	0.212	0.832	-0.092	0.114
TotalServicePointsWon	0.0779	0.215	0.363	0.717	-0.343	0.499
Omnibus:	281.757	Durbin-Watson:	2.054			
Prob(Omnibus):	0.000	Jarque-Bera (JB):	5685.408			
Skew:	0.427	Prob(JB):	0.00			
Kurtosis:	13.112	Cond. No.	271.			

# Recommendations And Next Steps

# Recommendations

- Practice getting out of tough situations  
(opponent's break point opportunities)
- Convert on break point opportunities
- Essentially be **CLUTCH**

# Next Steps

- Evaluate regularization for decreased complexity
- Consider stats by court type (clay, grass, hard)

# Thank You

Questions?

# References

- Icon made by Freepik <<https://www.flaticon.com/authors/freepik>> from [www.flaticon.com](https://www.flaticon.com)
- Icon made by mynamepong <<https://www.flaticon.com/authors/mynamepong>> from [www.flaticon.com](https://www.flaticon.com)
- Icon made by Smaishicons <<https://www.flaticon.com/authors/smashicons>> from [www.flaticon.com](https://www.flaticon.com)
- Icon made by Pause08 <<https://www.flaticon.com/authors/pause08>> from [www.flaticon.com](https://www.flaticon.com)