

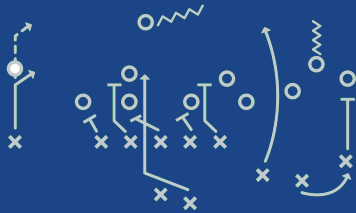
A Better Way to Evaluate a Football Team's Performance



By Sam Ouimet

Goal

Use regression analysis to identify what team statistics contribute to most to most and least to a teams win total, and then predict what a team's current win total "should" be based on in-season and historical team based statistics.



What the product is:

- A tool for NFL team personnel to leverage in order to quantitatively identify WHY they are winning or losing, rather than leaving it up to gut feel or hours reviewing tape.
- A tool for bettors to identify under and over performing teams in order to profit from reversion.

Data (2002-2018)

NFL.com

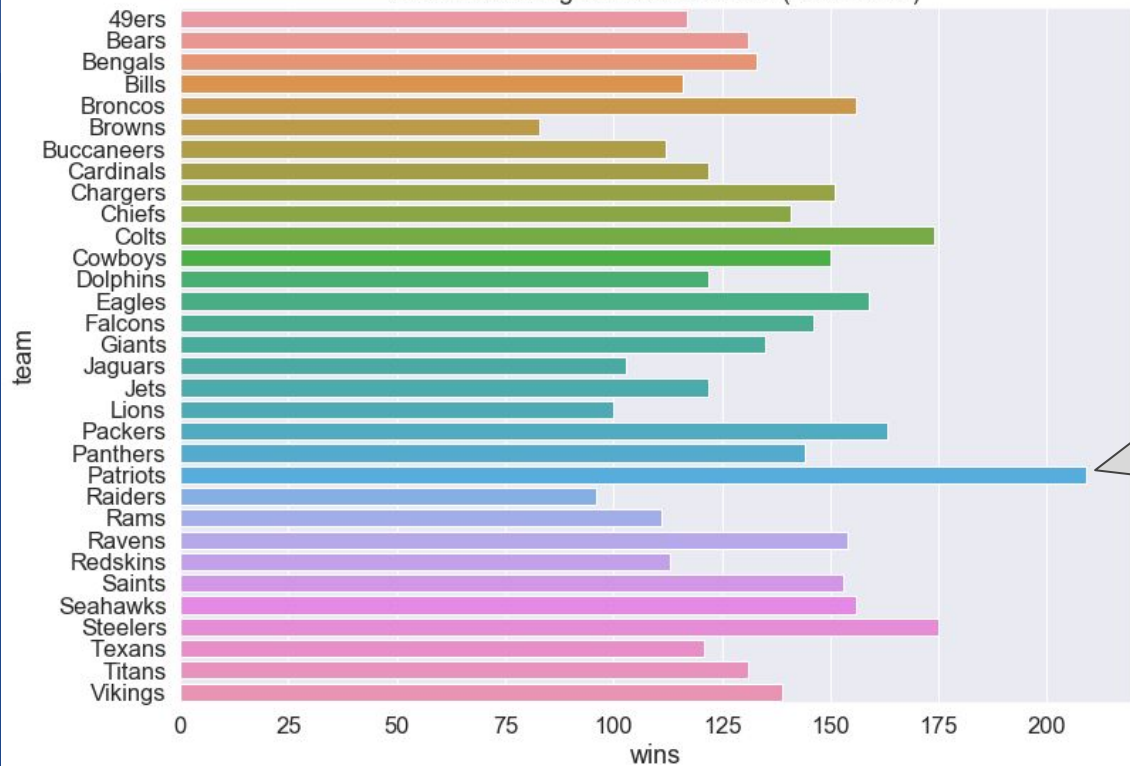
- Standard offensive and defensive statistics including:
 - Average points per game (PTS/G)
 - Average time of play per game (TOP/G)
 - Average rushing attempts per game (RUSH_ATT/G)

Footballoutsiders.com

- More advanced and propriety team based metrics including:
 - Defense-adjusted Value Over Average (DVOA)
 - Turnovers per play (TO/P)
 - Average number of 1st downs per game (1ST/G)

First Look at the Data

Cumulative Regular Season Wins (2002-2018)



Patriots are clearly the best team since 2002 with 209 wins.

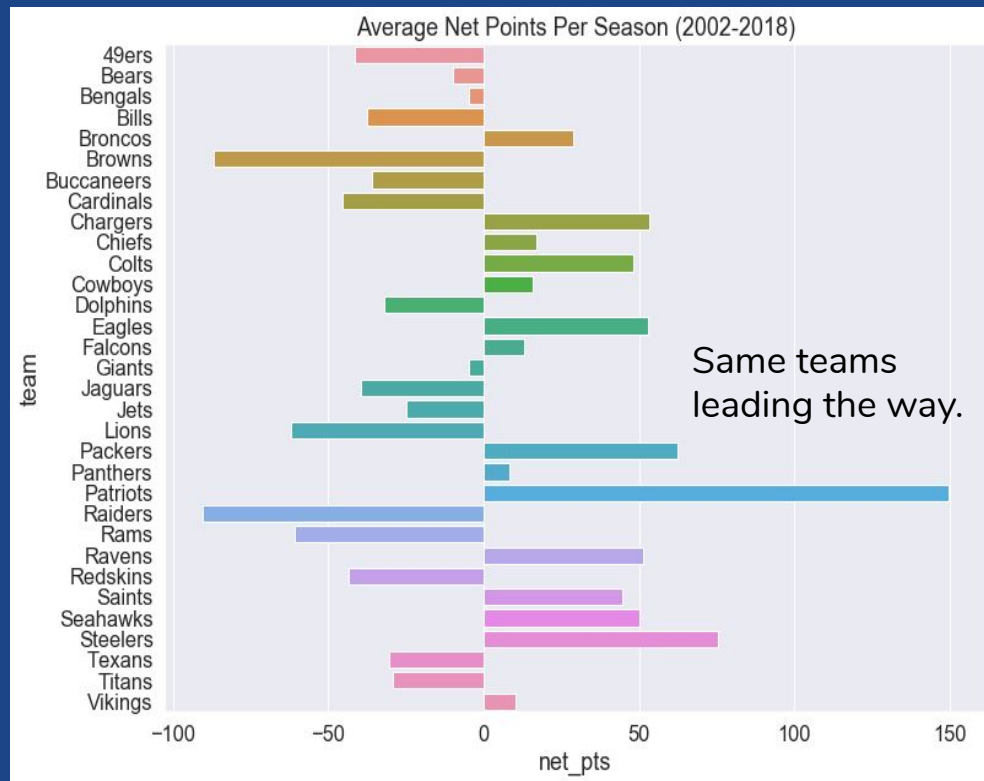
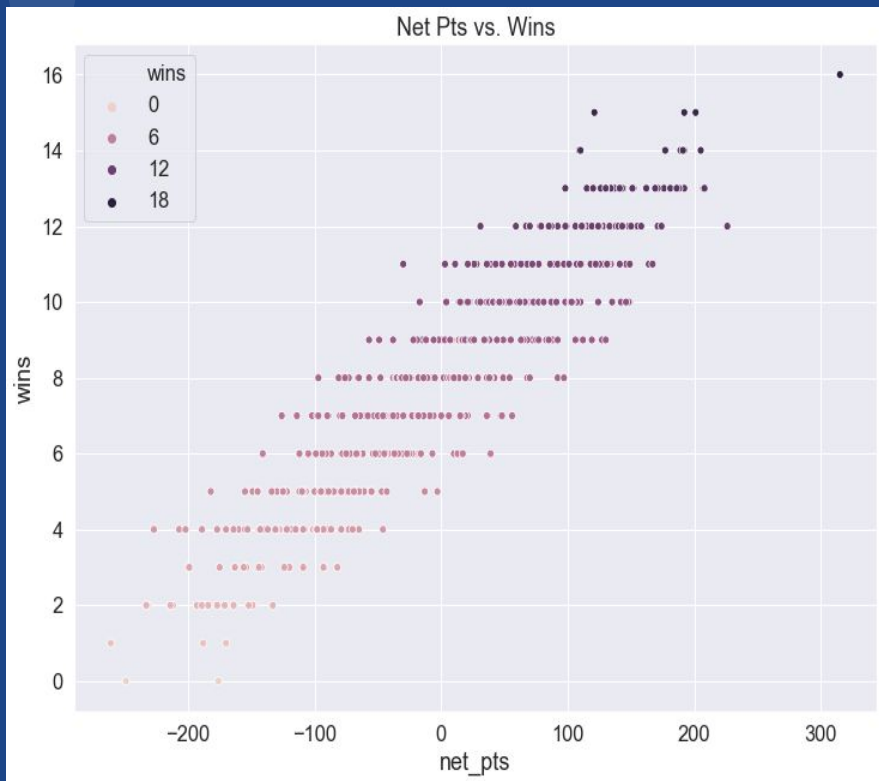
Average = 135 wins

Worst = Browns, 83 wins

But why?

A Closer Look

Net Points = total points scored - total points allowed

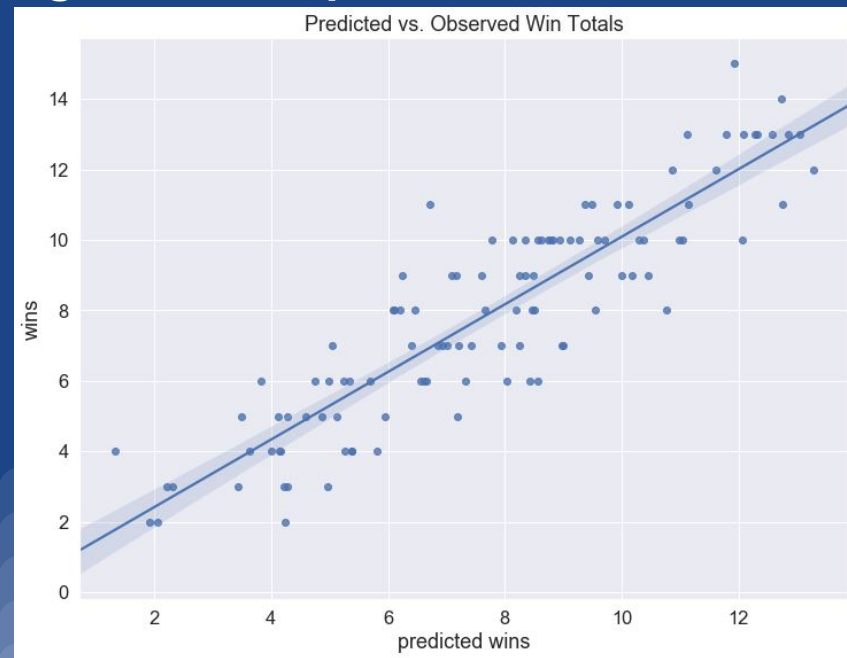


Model Building

- **Can't rely just on net wins - it is not uncommon for games to be decided by a single point**
- **Incorporated statistics from three core aspects the game:**
 - **Offense, defense & special teams**
- **Engineered interaction features including:**
 - **Points per minute with offense on the field to measure offensive potency**
 - **Sack to tackle ratio to measure aggressiveness and effectiveness of a defense**
 - **1st downs/g multiplied by 3rd down percentage to show ability to continue drives and not give the ball to the other team**

Results

- Cross validation (KFold) revealed the Ridge regression model provided the highest score on validation data with all 54 features. ($r^2 = 0.847$)
- After tuning the model more and finding the best alpha, r^2 rose to 0.87.
- Final Test Results:
 - $r^2 = 0.797$
 - MAE = 1.12 wins



How this can help today!

<u>Contributes most to wins</u>	<u>Contributes least to wins</u>
Net Points (0.705)	Total points allowed (-0.899)
Points per minute offense (0.283)	4th down attempts (-0.299)
Yards per drive (0.241)	Turnovers per drive (-0.148)

(numbers represent model
coefficients after standardizing)

NFL Personnel:

- Do not neglect 1 side of the ball (offense or defense)
- Sustained success is achieved through the ability to outscore opponents with a high margin
- Focus on ball security - turnovers are win killers

Bettors:

- Bet **against** a team with an actual win total more than the MAE above predicted total
- Bet **on** a team with an actual win total more than the MAE less predicted total

Future Work



1. Expand analysis to predict how a team should perform the following year
2. Scrape more data (player stats & individual game data)
3. Research more NFL trends and metrics
4. Injuries, personnel changes, team ownership, team valuations, weather and more

Thank you for listening!



Go Giants!