

A quest to find some absurd yet accurate predictors of Super Bowl success

PNIS Editorial Staff¹

1 - PNIS

INTRODUCTION

Two days ago, we determined the probability of a team winning the Super Bowl based on its ranking in 24 criteria. For example, the team with the better regular season record won the Super Bowl 62.5% of the time. [See the full article here.](#)

In that article, we focused exclusively on criteria that were relevant to a team's performance, such as Total Points Allowed, Total Points Scored, and Total Yards Allowed on Defense (see Table 1 [here](#)). According to our numbers, the two criteria that best predicted a team's chances of winning the Super Bowl were 1st Downs Allowed on Defense and Defensive Simple Rating System (DSRS). The team ranked higher in either criterion went on to win the Super Bowl 66.7% of the time.

This exercise got us thinking about the predictive success of non-football-relevant criteria. You should be able to compare two teams in just about anything and then determine how that comparison predicted Super Bowl success.

Thus, the goal of this paper is to find absurd and irrelevant comparisons of Super Bowl teams. Could any of them actually be accurate predictors of Super Bowl success? And, more importantly, could we find criteria that actually performed better than our 66.7% benchmark we set with the relevant data? The hunt was on.

THE PREDICTORS

Just like in our prior article, we looked for 24 criteria. The stipulations for these criteria were that: 1) the data must exist, and 2) they should ideally have nothing to do with playing football. We should also point out that we did not determine the accuracy of these predictors until after we collected all the data, and we did not get rid of any predictors because they turned out to be of poor quality. The 24 predictors are, ranked in order of absurdity from least to most:

- Average team height¹
- Length of starting quarterback's full name – middle initials not included
- Sum of the retired jersey numbers of all players who retired before the particular Super Bowl game was played
- Distance of the team's stadium from the Super Bowl Venue – measured as straight line distance
- Date in which team's city was incorporated
- Number of sister cities of the team's city – from: [Sister Cities International](#)
- Racial diversity for the team's city
- Average area code for the team's city
- Bible-mindedness of the team's city - based on

¹ okay, this might have something to do with playing football, but no one ever bets on games with average height in mind

the rankings given by the [American Bible Society](#)

- Percent of vote that the team's US State gave to the victor in the most recent US presidential election
- Average of the Red Green Blue (RGB) number codes of the team's main colors
- Opinion of a person that we can attest knows nothing about football
- Color that matches the color from [this random color generator from Random.org](#)
- Number of Facebook likes given to the team's home page on pro-football-reference.com
- Cost to sponsor the team's home page on pro-football-reference.com

Okay, these next few are a bit involved, so stay with us here:

- In real life, would the team name (i.e., the "Cowboys") own or use a gun? – After much discussion, the team names that could conceivably have used a gun are: 49ers, Patriots, Buccaneers, Raiders, Cowboys, Bills, Redskins, Chiefs, and Jets.
- Current population size of the team name – For some teams, we made slight substitutions: Seahawks (we used the current population size of ospreys), Broncos (feral horses), Patriots (Tea Party members), Colts (domesticated horses), Buccaneers (pirates), Rams (bighorn sheep), and Packers (meatpackers). We considered 49ers, Giants, Raiders and Titans as extinct.
- Genome size of the team name – measured in c-value, obtained from the [Animal Genome Size Database](#). Substitutions here included: Broncos & Colts (horse), Seahawks (sharp-shinned hawk), Panthers (cougar) and Rams (sheep).
- Normality of the points scored by the team in each regular season game – Normality was assessed using the Shapiro-Wilk's Test, and we used exact *P*-values to make the comparison of more/less normal (*there's equal sample sizes, so don't freak out about us using P-values for comparison purposes*)
- Alphabetical distance of the Coach's last name to the last name of the Nobel Literature Prize winner of that particular year – for example, how did the Coaches of the 2013 Super Bowl teams match up with Alice Munro, alphabetically?

- Weight of the player with the jersey number that was closest to the absolute value of the temperature anomaly (in deg C) for that particular year – [climate anomalies available here](#)
- Alphabetical distance of the team name to the element whose atomic number corresponded to whatever number Super Bowl it was – for example, Super Bowl I would match with Hydrogen.
- Atomic mass of the element whose symbol is closest alphabetically to the Coach's initials

If we didn't provide a URL, then the information was gathered from a combination of [pro-football-reference.com](#) and [Wikipedia](#).

We then determined if being greater or lesser in each criteria led to more Super Bowl wins (for example, did taller teams or shorter teams win the highest percentage of Super Bowls?). Ties were omitted. Table 1 shows the resulting predictive ability of all 24 criteria.

THE RESULTS

The main conclusion is that, we did it!! We were able to find 4 criteria that were better than the 66.7% standard set by the football-relevant criteria. However, two of these criteria (Number of Facebook Likes, and Cost to Sponsor Homepage) are extremely suspect. It's clearly evident that people were liking a particular team's homepage on pro-football-reference.com because they won the Super Bowl. We're proud to say we figured this out when we realized Facebook didn't exist until 2004. Likewise, the cost to sponsor a page on pro-football-reference.com is probably dependent on the number of views that page gets, which again is probably dependent on if that team won the Super Bowl. So, we can throw out our top 2 predictors.

But there's nothing suspect about our next two: genome size and alphabetical distance of team name to element name. It turns out that the team name with the larger genome wins almost 70% of the Super Bowls. Thus, we highly recommend the next NFL expansion city should choose the 'Marbled Lungfish' as their team name.

Also, if your team name is closer alphabetically to the element whose atomic number matches the Super Bowl number, then you have a 68.8% chance of winning the Super Bowl. Tennessee Titans fans just got a glimmer of hope for next year.

Table 1. Quality of the 24 criteria used for rankings. Percentage refers to percent of time the higher ranked team won the Super Bowl. Numbers in parentheses are number of times better team won and number of instances after accounting for ties.

Criterion	Direction That Won Higher %age of Super Bowls	% of Time Better Team Won Super Bowl
Cost to Sponsor Home Page	Higher Cost	91.5% (43/47)
Number Facebook Likes	More Likes	89.6% (43/48)
Genome Size of Team Name	Larger Genome	69.7% (23/33)
Alphabetical Distance of Team Name to Element whose Atomic Number Matches the Super Bowl Number	Closest Alphabetically	68.8% (33/48)
Opinion of Person who Knows Nothing about Football	Trust this Person	64.6% (31/48)
Atomic Mass of Element whose Symbol is Closest Alphabetically to Coach's Initials	Closest Alphabetically	63.8% (30/47)
Weight of Player with Jersey Number Closest to Temperature Anomaly	Heavier Player	61.7% (29/47)
Could Team Name Have a Gun?	Has Gun	58.6% (17/29)
Distance to Super Bowl Venue	Closer to Venue	58.3% (28/48)
Average Team Height	Taller Team	58.1% (25/43)
Population Size of Team Name	Larger Population Size	56.5% (26/46)
Bible-mindedness of Team's City	More Bible-Minded	56.3% (27/48)
Normality of Points Scored by Team during Regular Season	Less Normal	56.3% (27/48)
Racial Diversity of Team's City	More Diverse	56.3% (27/48)
Average of RGB Numeric Codes of Team's Colors	Lower Average RGB	56.3% (27/48)
Length of Quarterback's Name	Shorter Name	55.0% (22/40)
Alphabetical Distance of Coach's Name to Nobel Literature Laureate	Closest Alphabetically	54.5% (24/44)
Date that Team's City was Incorporated	Later Date	54.2% (26/48)
Population Size of Team's City	Larger City	54.2% (26/48)
Number of Sister Cities for Team's City	More Sister Cities	53.2% (25/47)
Average of Area Codes for Team's City	Lower Average	52.1% (25/48)
Sum of Retired Jersey Numbers	Lower Sum	51.1% (23/45)
% Vote Given to Victor in Last US Presidential Election	Higher % to Victor	50.0% (23/46)
Team Color Matches Random Color	Closer Match to Random Color	50.0% (21/42)

Most of the other predictors, though, fell between 50 and 60%. The average accuracy of these absurd predictors (without the Facebook likes and Cost to Sponsor criteria) was 57.2%, while our average for the football-relevant predictors was 59.2%, so overall we did a bit worse.

The larger message, though, is that you can find just about any accurate predictor of Super Bowl success. In fact, we wouldn't be surprised if there was a

criterion that has predicted all 48 Super Bowls correctly. In a way, our paper is somewhat conceptually similar to this [Sasquatch niche modelling paper](#), which cautions researchers against using just any old data to build ecological niche models.

The even larger takeaway from this study, however, is that we have got to get our Person-who-Knows-Nothing-about-Football out to Vegas. We may or may not have uncovered [the next Balki](#).