

# How Our Club Soccer Predictions Work

See our latest predictions

---

## references

[Engsoccerdata](#) / [Opta](#) / [Transfermarkt](#)

[Harmonic mean](#) / [Massey's method](#) / [Monte Carlo method](#) / [Poisson process](#) / [Ranked probability score](#)

## the details

We first published FiveThirtyEight's [club soccer predictions](#) in January 2017 with six leagues. Since then, we've steadily expanded the number of leagues we forecast, added features to our interactive graphics, tweaked our predictive model to perform better and published our [global club soccer rankings](#).

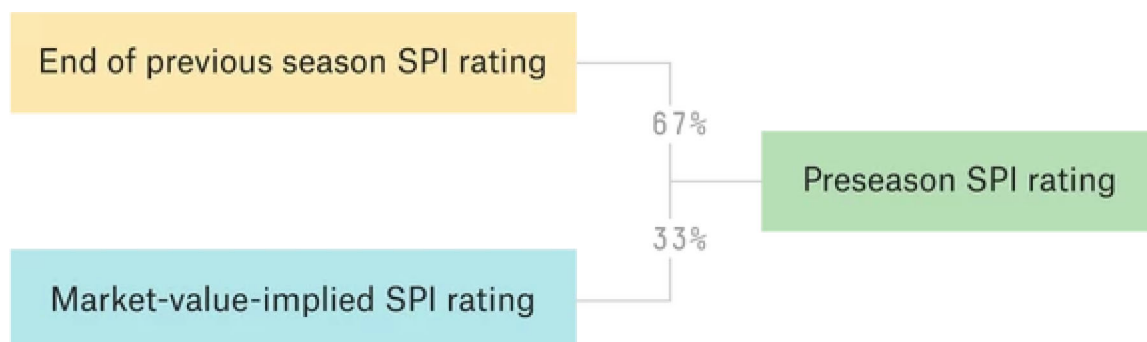
The forecasts are based on a substantially revised version of ESPN's Soccer Power Index (SPI), a rating system [originally devised](#) by FiveThirtyEight editor-in-chief Nate Silver in 2009 for rating international soccer teams. We have updated and adapted SPI to incorporate club soccer data going back to 1888 (from more than 550,000 matches in all) that we've collected from ESPN's database and the [Engsoccerdata GitHub repository](#), as well as from [play-by-play data produced](#) by [Opta](#) that has been available since 2010.

## SPI ratings

At the heart of our club soccer forecasts are FiveThirtyEight's SPI ratings, which are our best estimate of a team's overall strength. In our system, every team has an offensive rating that represents the number of goals it would be expected to score against an average team on a neutral field, and a defensive rating that represents the number of goals it would be expected to concede. These ratings, in turn, produce an overall SPI rating, which represents the percentage of available points — a win is worth 3 points, a tie worth 1 point, and a loss worth 0 points — the team would be expected to take if that match were played over and over again.

Given the ratings for any two teams, we can project the result of a match between them in a variety of formats — such as a league match, a home-and-away tie or a cup final — as well as simulate whole seasons to arrive at the probability each team will win the league, qualify for the Champions League or be relegated to a lower division.

Before a season begins, a team's SPI ratings are based on two factors: its ratings at the end of the previous season, and its market value as calculated by [Transfermarkt](#) (a site that assigns a monetary value to each player, based on what they would fetch in a transfer). We've found that a team's market value — relative to their league's average value — is strongly correlated with its end-of-season SPI rating. Thus, we use these market values to infer each team's preseason SPI rating.



As a season plays out, a team's ratings are adjusted after every match based on its performance in that match and the strength of its opponent. Unlike with the [Elo](#) rating system we use in [several other sports](#), a team's rating doesn't necessarily improve whenever it wins a match; if it performs worse than the model expected, its ratings can decline.

## Match performances

Soccer is a tricky sport to model because there are so few goals scored in each match. The final scoreline will fairly often disagree with many people's impressions of the quality of each team's play, and the low-scoring nature of the sport will sometimes lead to prolonged periods of luck, where a team may be getting good results despite playing poorly (or vice versa).

To mitigate this randomness, and better estimate each team's underlying quality of play, we use three metrics to evaluate a team's performance after each match: *adjusted goals*, *shot-based expected goals* and *non-shot expected goals*.<sup>1</sup>

The first, *adjusted goals*, accounts for the conditions under which each goal was scored. For *adjusted goals*, we reduce the value of goals scored when a

team has more players on the field,<sup>2</sup> as well as goals scored late in a match when a team is already leading.<sup>3</sup> After downweighting these goals, we increased the value of all other goals to make the total number of adjusted goals generally add up to the total number of actual goals scored over time.

*Shot-based expected goals* are an estimate of how many goals a team “should” have scored, given the shots they took in that match. Each shot is assigned a probability of scoring based on its distance and angle from the goal, as well as the part of the body the shot was taken with, with an adjustment for which specific player took the shot.<sup>4</sup> These individual shot probabilities are added together to produce a team’s *shot-based expected goals* for that match, which may be bigger or smaller than the number of goals it actually scored.

*Non-shot expected goals* are an estimate of how many goals a team “should” have scored based on non-shooting actions they took around the opposing team’s goal<sup>5</sup>: passes, interceptions, take-ons and tackles. For example, we know that intercepting the ball at the opposing team’s penalty spot results in a goal about 9 percent of the time, and a completed pass that is received at the center of the six-yard box leads to a goal about 14 percent of the time. We add these individual actions up across an entire match to arrive at a team’s *non-shot expected goals*. Just as for *shot-based expected goals*, there is an adjustment for each action based on the success rates of the player or players taking the action (both the passer and the receiver, in the case of a pass).

Since all three metrics represent the number of goals a team either scored or could have been expected to score during the match, they’re directly comparable. So a team’s composite offensive score for that match is an average of its performance across the three metrics, and its composite defensive score is an average of the three metrics for its opponent.

1/15	 <b>Everton 4</b>	22%	25%
	 <b>Man. City 0</b>	54%	
		EVE	MNC
Adjusted goals		3.5	0.0
Shot-based xG		0.4	0.7
Non-shot xG		0.7	2.7

Take the [January 2017 match](#) between Everton and Manchester City, for example. Although Everton won 4-0, our model didn't see the match as nearly so lopsided: Two of Everton's goals came with the lead after the 70th minute. Furthermore, Everton took only six shots. Our shot-based expected goals model would expect only about 0.4 of those shots to go in the net, rather than the four that did. Man City also was the better team according to our non-shot based expected goals model. In all, our composite scores saw the final result as a 1.53-1.13 win for Everton — much narrower than 4-0.

## Forecasting matches

Given two teams' SPI ratings, the process for generating win/loss/draw probabilities for a given match is three-fold:

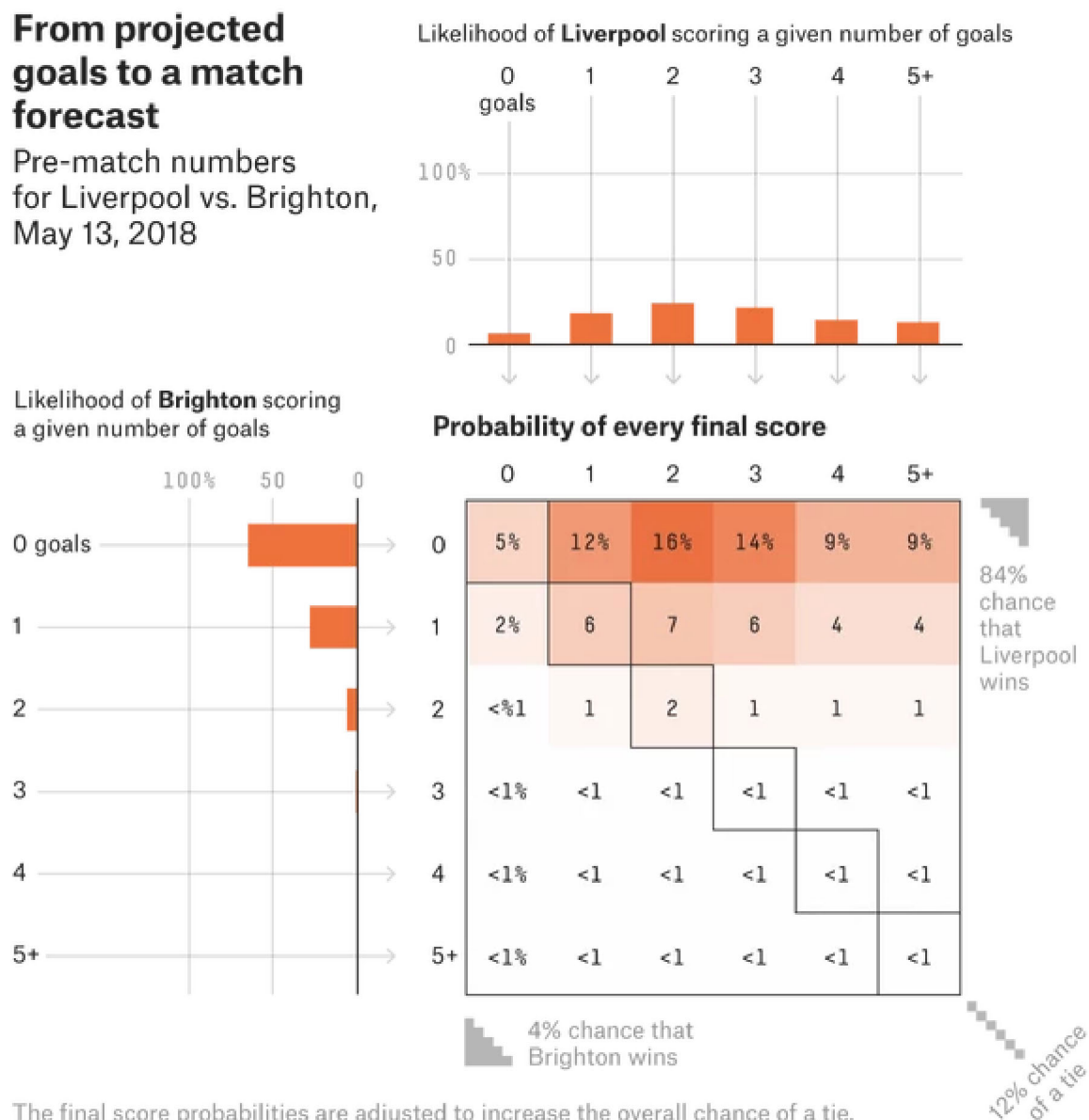
1. We calculate the number of goals that we expect each team to score during the match. These projected match scores represent the number of goals that each team would need to score to keep its offensive rating exactly the same as it was going into the match, and they are adjusted for a league-specific home-field advantage and the importance of the match to each team ([described below](#)).
2. Using our projected match scores and the assumption that goal-scoring in soccer follows a [Poisson process](#), which is essentially a way to model random events at a known rate, we generate two Poisson distributions around those scores. These give us the likelihood that each team will score no goals, one goal, two goals, etc.

3. We take the two Poisson distributions and turn them into a matrix of all possible match scores, from which we can calculate the likelihood of a win, loss or draw for each team. To avoid undercounting draws, we increase the corresponding probabilities in the matrix to reflect the actual incidence of draws in a given competition.<sup>6</sup>

Take, for example, the [May 2018 Premier League match](#) between Liverpool and Brighton, which Liverpool won 4-0. Before the match, our model was very confident that Brighton would score either no goals or one goal. Liverpool's distribution, however, was much wider, leading to it being a significant favorite (84 percent) in the match. Here's a visual interpretation of how our model uses those distributions to generate each team's chance of winning a match:

### From projected goals to a match forecast

Pre-match numbers for Liverpool vs. Brighton, May 13, 2018



## Forecasting seasons

Once we have probabilities for every match, we then run [Monte Carlo simulations](#) to play out each league's season 20,000 times using those

forecasts. As with our other projections, we run our Monte Carlo simulations “hot,” meaning that instead of a team’s ratings remaining static within each simulated season, the ratings can rise or fall based on the simulated matches the team plays. In effect, this widens the distribution of possible outcomes by allowing a weak team to go on a winning streak and increase its ratings substantially, or providing for the possibility that a strong team loses its first few games of a simulated season and is penalized accordingly.

## **Match ratings and importances**

On any given week during peak soccer season, FiveThirtyEight offers projections for dozens of club soccer matches across the globe. The sheer volume of matches taking place at some times of the year can be paralyzing. With that in mind, we have a feature on our interactive graphics to [rate upcoming matches](#) on their *quality* and *importance*.

*Quality* is simply a measure of how good the teams are. Specifically, it’s the [harmonic mean](#) of the two teams’ SPI ratings.<sup>7</sup> Because every team has an SPI rating between 0 and 100, match quality also ranges from 0 to 100.

*Importance* is a measure of how much the outcome of the match will change each team’s statistical outlook on the season. This outlook considers different factors depending on which league the match is being played in; for some leagues, the outlook only considers winning the league, while other leagues incorporate the possibility of being promoted or relegated, or qualifying for the Champions League. To calculate the importance of a match to a team, we generate probabilities for each factor conditional on winning (or losing) the match, and then find the difference between those two possible numbers. We take the factor with the maximum range of difference for each team and scale the result to between 0 and 100. Finally, we average the match’s importance to both teams to find the overall match importance. All leagues are treated equally when calculating importance, so a match to decide the winner of the Swedish Allsvenskan would rate just as high as a match to decide the winner of the English Premier League.

The overall *match rating* is just the average of quality and importance.

As of 2018, our match predictions incorporate *importance* in two ways :

1. When a match is more important to one team than the other, that team tends to outperform expectations, with its boost in performance relative to how much more important the match is to them.



2. If a match isn't important to either team, uncertainty in the outcome of the match increases.

To understand the magnitude of these importance adjustments, consider a match that is equally important to the two teams, where the home team has a 50 percent chance of winning the match, the away team has a 25 percent chance of winning the match, and the remainder is the chance of a draw.

If, instead, we assume that it is a hugely important match for the home team, and a meaningless match for the away team, the home team's chances of winning would go up to 58 percent, and the away team's chances would go down to 18 percent.

On the other hand, if the match was meaningless to both teams, the home team's chances of winning would go down to 43 percent, and the away team's chances would go up to 30 percent.

The improvement we see in our match forecasts when incorporating match importance is about one-third the size of the improvement we saw when we added expected-goals metrics in 2016, and about one-half the size of the improvement we saw when we incorporated market values in preseason ratings for 2017.<sup>8</sup>

## League strengths

Most club soccer matches are played against teams from the same domestic league, but some matches — like those in the UEFA Champions and Europa leagues — can be played against teams from different countries.

To assess the relative strength of domestic leagues, we use recent matches played between teams from different leagues, supplemented with league market values from [Transfermarkt](#), to assign a strength rating to every league for which we have data.

To generate these league strength ratings, we've set up a system where we first assume that all leagues are of equal strength and determine how far above or below expectation each league has performed over the past five years. In order, we:

1. Run through all domestic matches in history and calculate domestic team Soccer Power Index (SPI) ratings throughout time.
2. Look at each inter-league match from the past five years and calculate the expected score of the match based purely on each team's domestic rating at

the time.

3. Take the difference between our expected score of the match and the actual score and run these results through [Massey's Method](#) to find a rating for each league, expressed in how many goals better or worse than the global average that league is.
4. Regress these calculated ratings toward market-value based ratings, weighted by how many inter-league matches we have for each league.
5. Run through all matches in history one more time, incorporating league strengths into the predictions for any inter-league matches to improve the final team ratings.

After going through that process, our league strengths can be interpreted as a bonus (in goals) given to each team in an inter-league match.

There are club soccer leagues being played year-round; follow dozens of them on our [club soccer predictions](#).

*Editor's note: This article is adapted from [previous articles](#) about how our club soccer predictions work.*

model creator

### **Jay Boice**

A computational journalist for FiveThirtyEight.

✉ | [@jayboice](#)

version history

### **1.3**

Match importance incorporated into forecasts; forecasts for 35 leagues updated for 2018-19 season.

Aug. 10, 2018

### **1.2**

Match quality and importance added to interactive.

Feb. 14, 2018

### **1.1**

Improvements to relative league strengths; forecasts for 26 leagues updated for 2017-18 season.



Aug. 10, 2017

## 1.0

Model and forecast launched for 2016-17 season with six leagues.

Jan. 19, 2017

related articles

## Global Club Soccer Rankings

---

### Footnotes

1. Prior to Aug. 10, 2018, we also included the actual score of the match to calculate match performances.
2. These are worth about 0.8 goals apiece. This and all other weights were chosen in order to optimize the model for predicting future match outcomes.
3. Specifically, after the 70th minute, the value of a goal when a team is leading decreases linearly to the end of the game, when a real-life goal is worth half a goal in the eyes of our model. So a 70th-minute goal when leading is worth a full goal to our model, an 80th-minute goal is worth 0.75 goals, and a goal in the 90th minute or later is worth 0.5 goals.
4. All players who have enough shots in our database to qualify are given a modifier based on their historical conversion rates (the number of goals they've actually scored, given the quality of the shots they've had). For example, [Lionel Messi](#) has historically converted a shot into a goal about 1.4 times as often as expected, so the probability of any shot he takes is multiplied by 1.4.
5. That is, within an area slightly larger than the 18-yard box.
6. There has been [some debate](#) about what kind of distribution best models scoring in soccer. We've found that two independent Poisson distributions work well with the addition of diagonal inflation. That is, we generate the two distributions independently, but increase the value of each cell in the matrix where the scores are equal by some constant (somewhere around 9 percent, but this differs by league and is based on the degree to which we would have undercounted draws had we not inflated the diagonal).
7. We're using the harmonic mean instead of merely averaging the two ratings, because in lopsided matches it limits the impact of very high or low ratings, resulting in a more balanced number.
8. If we look at the model's [Ranked probability score](#) (RPS) over every match in the five strongest leagues — Spain, England, Germany, Italy and France — during the past three seasons, adding our expected-goals metrics reduces our model's RPS by 0.0018, adding market values to preseason team ratings reduces it by a further 0.0011, and incorporating match importance reduces it by another 0.0006, to 0.1957.