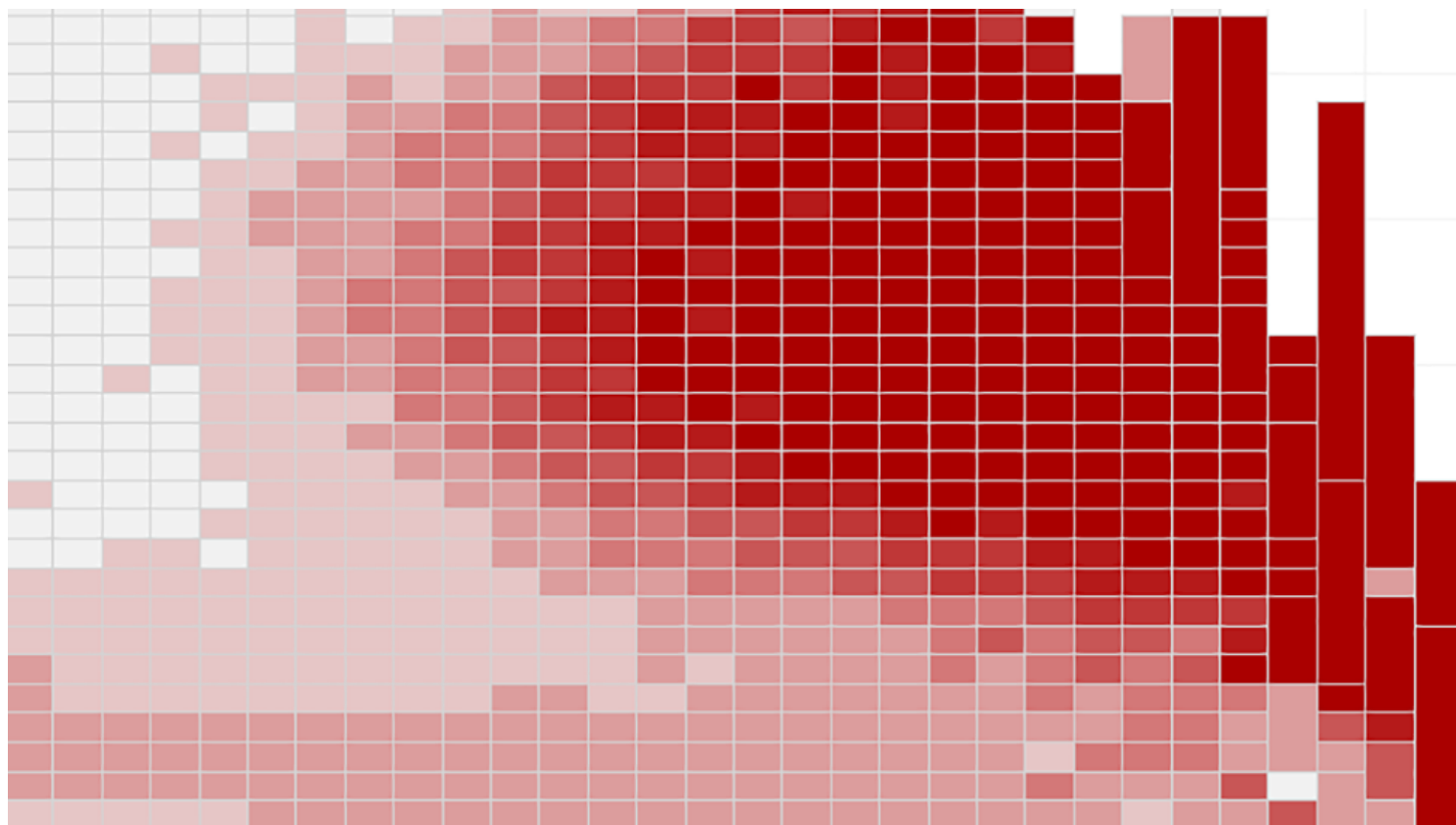


# SANTA: A Binary Approach to Pitcher Evaluation

by Eli Ben-Porat



**To no one's surprise, harder-hit balls with higher launch angles result in more extra base hits.**

**I**n 1980, Michael Porter outlined his **three generic strategies** for business: cost leadership, differentiation and focus. This simple, yet powerful model for business became the de-facto lens through which many business decisions were evaluated. Similarly, FIP and xFIP are simple, yet powerful models that allow us to evaluate a pitcher's performance with basic ingredients and delicious results. Any truly powerful model should be able to be understood by a lay person yet be sophisticated enough to explain complex phenomena.

The author began this piece looking to build a model that would assign a value to each pitch based on the batter, location of the pitch and pitch type. However, while

playing with the data, a far simpler approach surfaced—simply classifying each pitch as either “good” or “bad.” A “bad” pitch refers to a pitch that resulted in a ball or was hit well (measured by **StatCast near-barrels**). All other pitches, including those that can be considered neutral, are lumped into the “good” bucket.

The rationale behind this approach is that there is never a negative run outcome from a foul ball, unless the defending team allows a baserunner to score on a sacrifice. On the flip side, foul balls often will result in outs.

We’ve moved The 2018 Hardball Times Annual online to give the entire Internet the chance to read the research and insight The Hardball Times Annual has brought to print for the past 15 years.

Please consider supporting the 2018 Hardball Times Annual by becoming a FanGraphs member. Your membership will help fund future Hardball Times Annuals, the thousands of articles FanGraphs publishes each year, and our ever growing database of baseball stats.

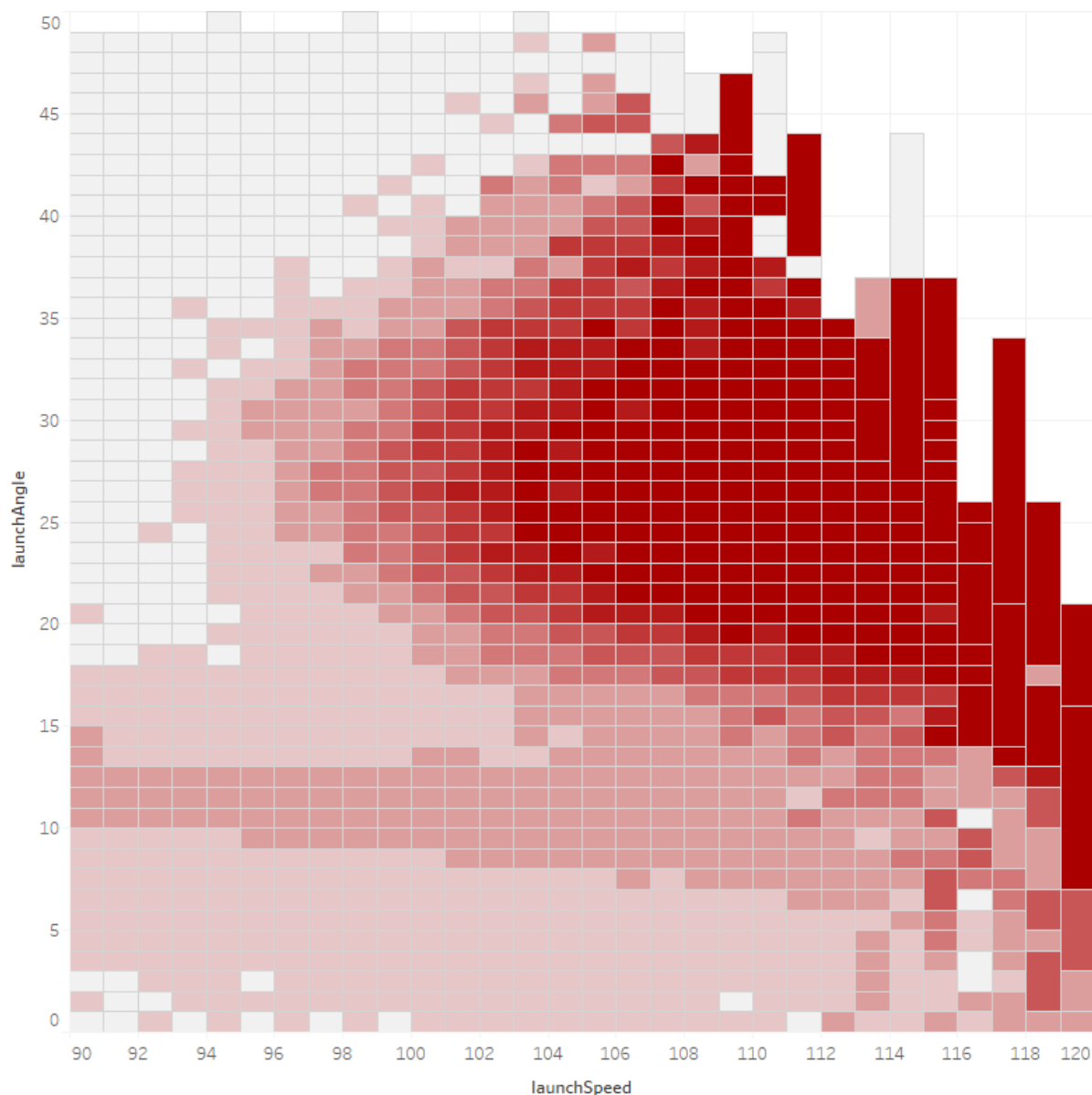
Support FanGraphs

We'll explore today how a tweaked version of this very simple metric works exceptionally well to describe pitcher performance, more so than with batters. We're calling it SANTA (Somewhat Arbitrary Name, Thanks Anyway) since it is essentially a measure of pitches as either Naughty or Nice.

## A Short Discussion on Barrels

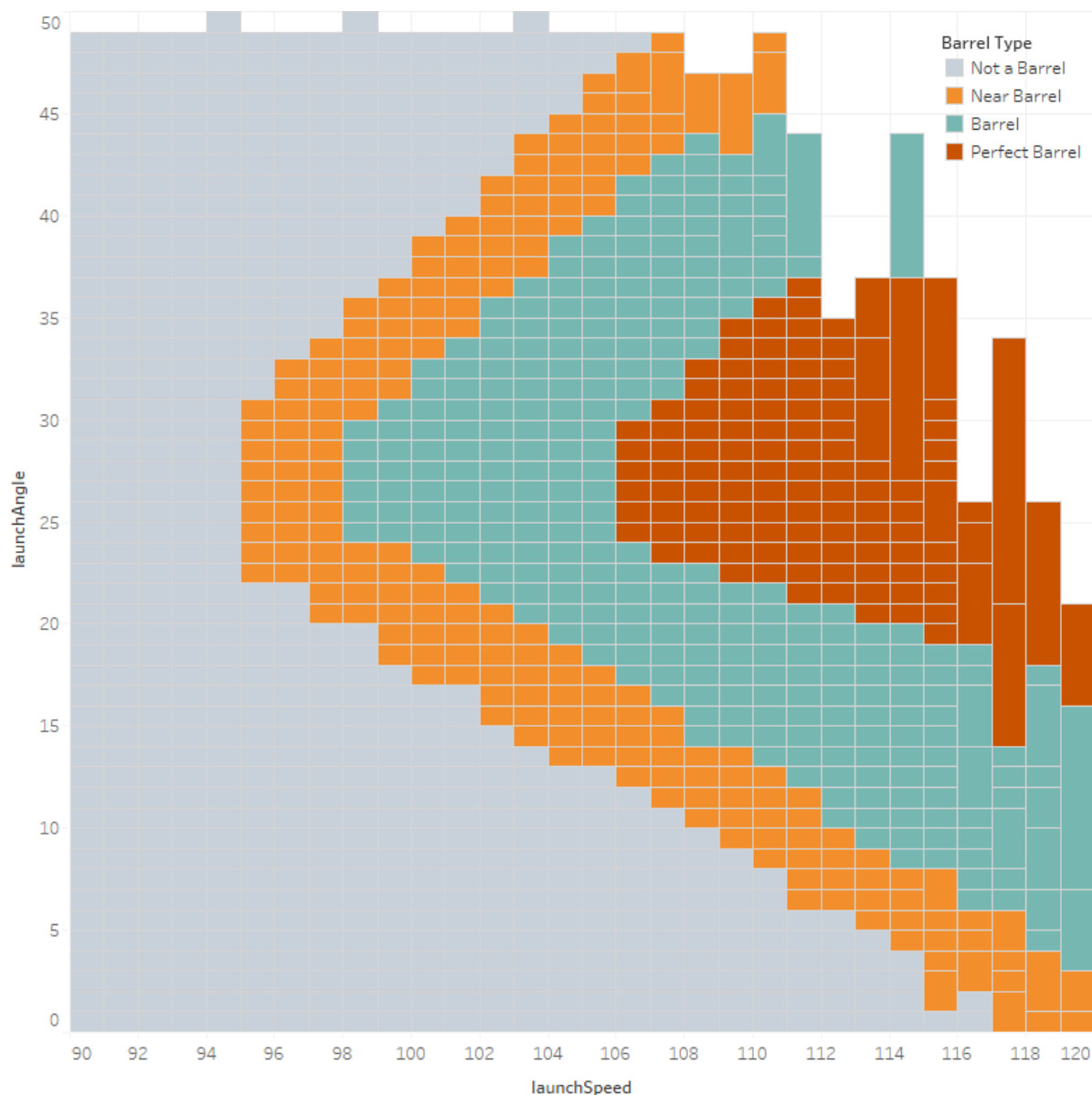
Tom Tango **describes three classes of barrels**, Near-Barrels (NB), Barrels and Perfect Barrels (PB), with near-barrels as the superset, barrels a subset of NB, and PB a subset of Barrels. For a detailed explanation on the theory behind the metric, read Tango's blog post on the subject. For the purposes of this article, we're going to use it to classify contact as either good (Near-Barrel or better) or not good, using the two metrics hitters shoot for—launch speed and launch angle. There is some debate whether spray angle should be used. For the purposes of this article, we'll simplify and stick to speed and angle.

Slugging Percentage by Launch Angle and Launch Speed:



This author prefers to use the less-than-perfect slugging percentage metric over the more accurate wOBA, for the simple reason that it translates better for storytelling. In this case, the story the data are telling is that—starting in the 20 to 30 degree launch angle window at about 96 mph exit velocity, then spreading out in a triangle pattern—we see increasingly stronger contact. Additionally, executing a launch angle between 10 and 15 degrees will give you pretty good results, even without elite exit velocity. How does this translate to barrels? Let's take a look.

Barrel Type by Launch Angle and Launch Speed:



Here we see the superset concept visualized; the Near-Barrels are a superset of Barrels, which are a superset of Perfect Barrels. Basically, if you're a pitcher, you want to avoid giving up near-barrel type contact. Using this metric allows us to simplify the first visual's complexity into a blissfully simple metric: Did the batter barrel up the pitch or not?

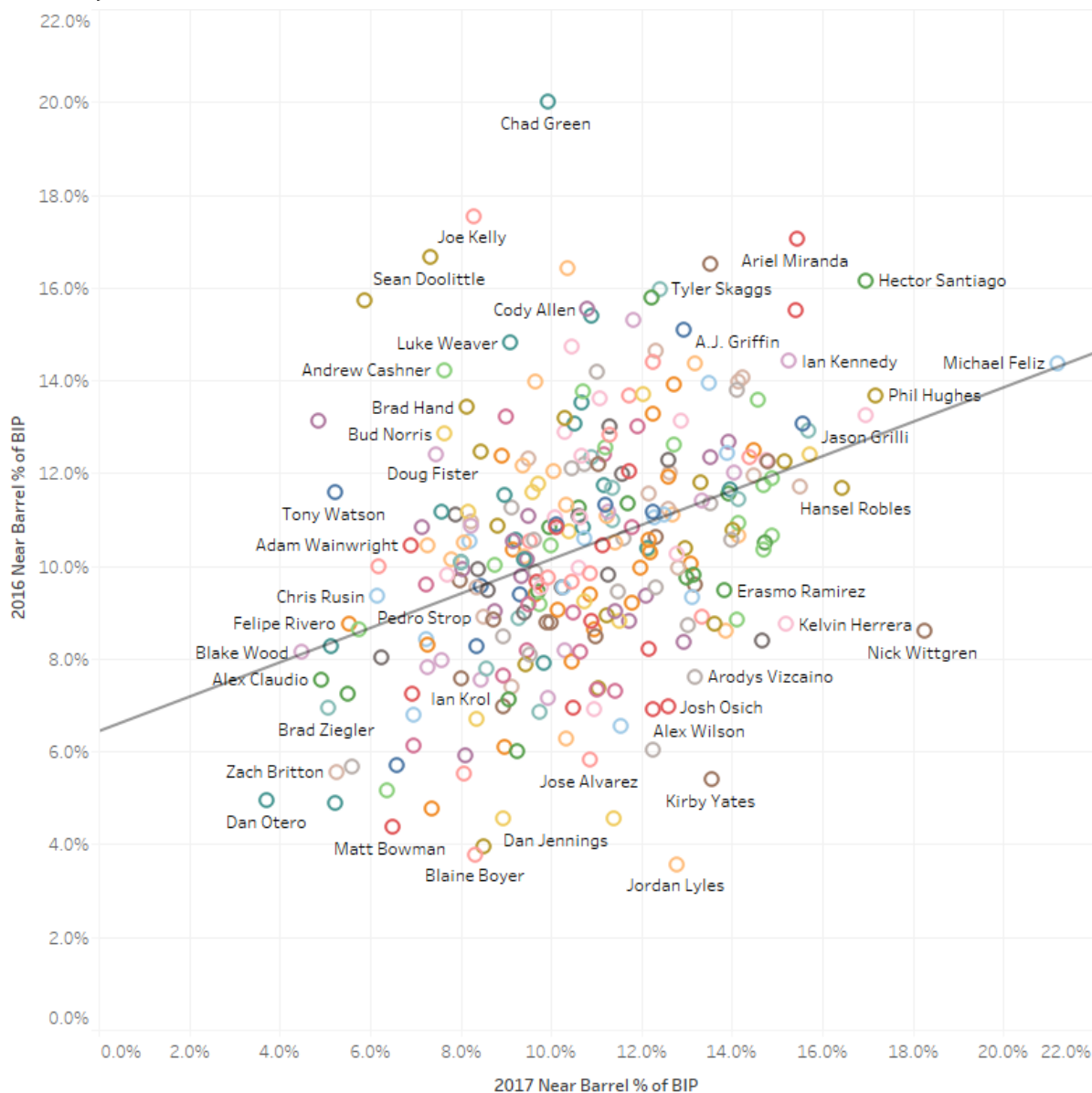
## Is Preventing Barrels a Repeatable Skill?

The FIP revolution in pitching evaluation re-wired our collective psyche to heavily discount a pitcher's ability to manage contact. However, if we are going to use the metric above, it would only make sense to use it if it was a repeatable skill from year to year; i.e. if a pitcher performs well in one year, will that have predictive power on the next season?

To evaluate this as a pitcher skill, we will look at the percentage of NBs a pitcher allows in relation to the total number of pitches he threw. While in theory the correct denominator would be balls in play or at-bats, in this case the metric we're working towards building relates to all pitches; i.e. if a pitcher threw 100 pitches, how many of them were "good" and how many were "bad."

The weakness behind this approach lies in the fact that a pitcher who throws a higher percentage of balls will have a lower percentage of NBs given up, so in some respect the ability (or inability) to throw strikes will play a big factor in predicting year-to-year NB percentage. Let's see what the data have to say on the subject and determine how predictive the skill is, using both pitches and balls in play as the denominator.

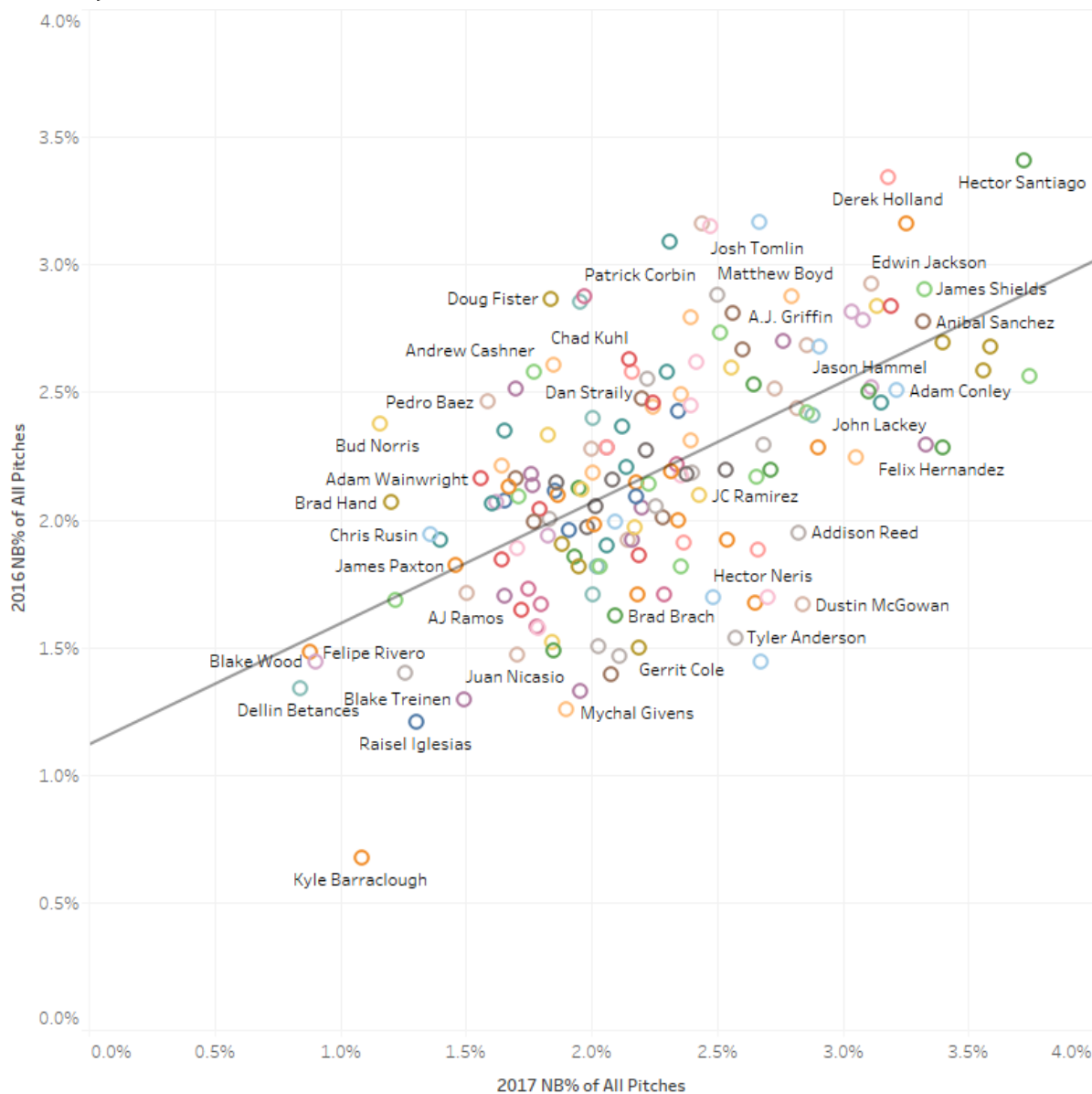
Near Barrels as % of Balls in Play–2016 to 2017 |  $R^2 = .14$  (Min. 100 BIPs in both seasons)



Not only do we get a very strong signal, we also get data that makes a lot of sense.

**Zach Britton** and **Brad Ziegler** are top-notch barrel preventers; **Phil Hughes** and **Ian Kennedy** not so much. **Chad Green** made a big jump in barrel prevention by moving into the bullpen. Let's now take a look at the same metric, this time with NBs as a percentage of all pitches.

Near Barrels as % of all Pitches – 2016 to 2017 |  $R^2 = .33$  (Min. 1,000 pitches in both seasons)



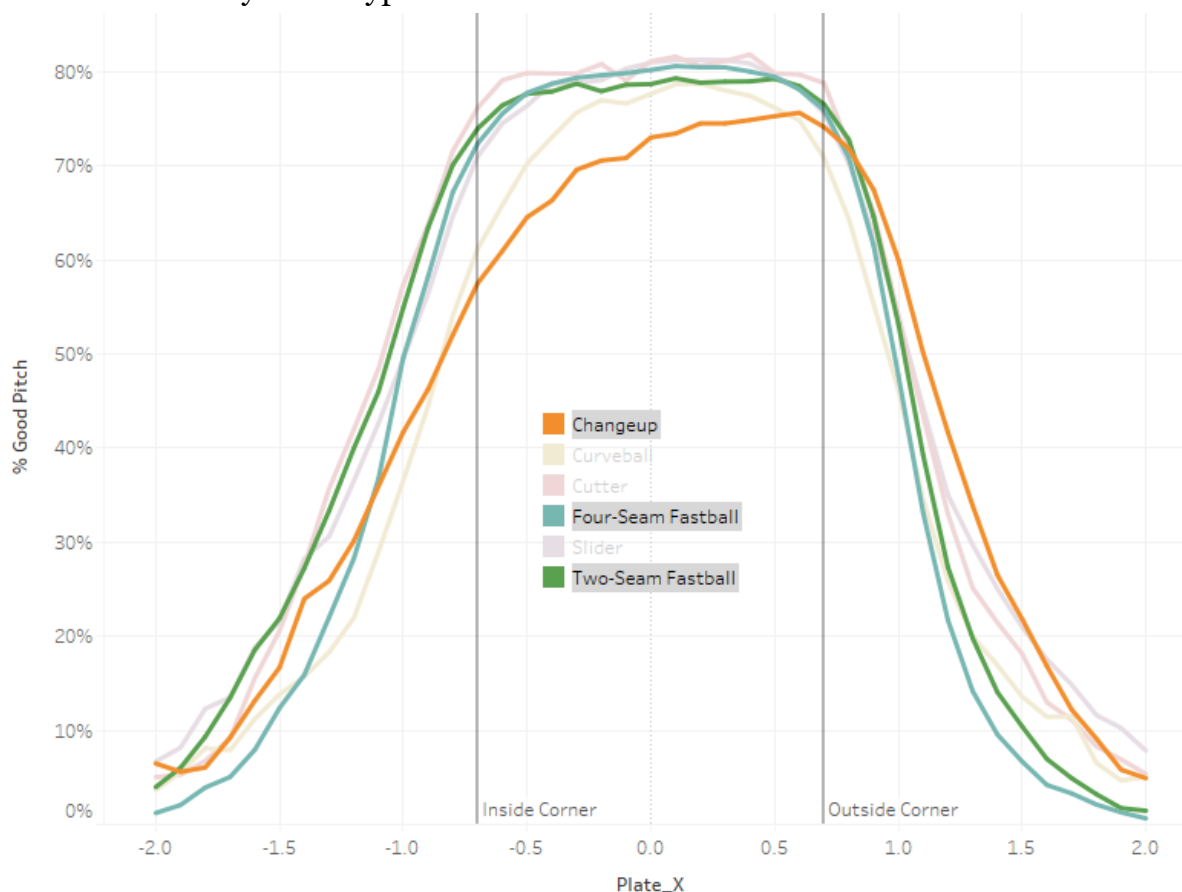
Here we get an incredibly strong signal where, given a big enough sample size, pitchers generally cluster very close to the trend line. Pitchers such as **Kyle Barraclough** and **Dellin Betances**, who walk a lot of batters, are also less likely to give up good contact. It is this balance that the metric we're developing is attempting to capture—the always changing game of how much strike zone to aim for; balls are bad, but so is good contact, so what is the right balance?

Now that we've established that pitchers definitely can impact the quality of contact (as measure by NBs), we're going to add one other simple ingredient to the mix—the pitcher's ability to avoid throwing called balls. Before we delve into how this translates to pitcher performance, let's look at the metric and see how it translates into various pitch types within the strike zone.

## Good % = Percentage of Pitches that DO NOT result in a Called Ball or Near Barrel

We're going to tweak the metric at the end to give different weights to the Ball% and NB% metrics. However, before we do that, let's take a short detour to see how this simple metric describes a few different pitch types and how they perform in different parts of the strike zone.

% Good Pitches by Pitch Type and Horizontal Location

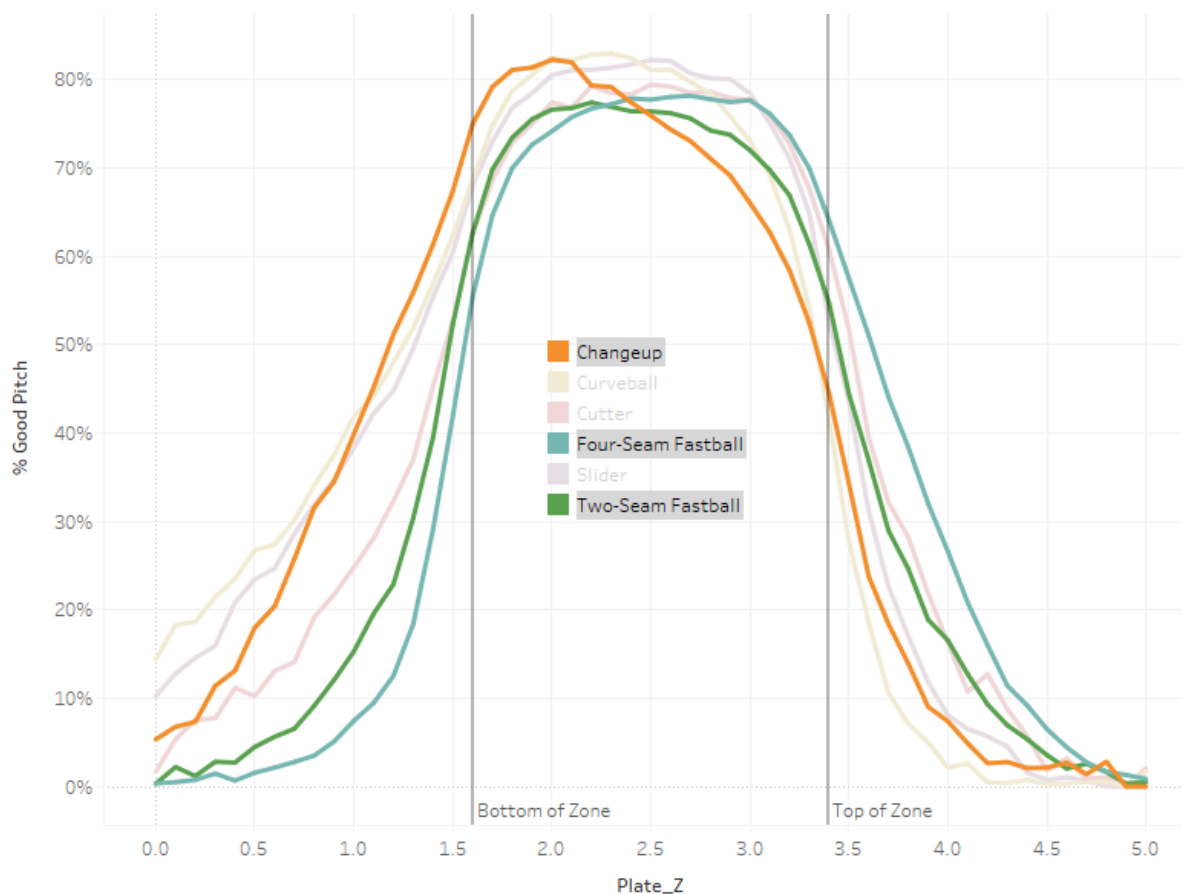


The graph fixes the horizontal location to the batter's perspective, with negative being inside and positive being outside, with change-ups (orange color) compared to four-seam and two-seam fastballs. We see that change-ups have a very distinct curve, where they are far more effective when thrown to the outside, peaking at the outside corner. Fastballs are pretty neutral in terms of horizontal location, though



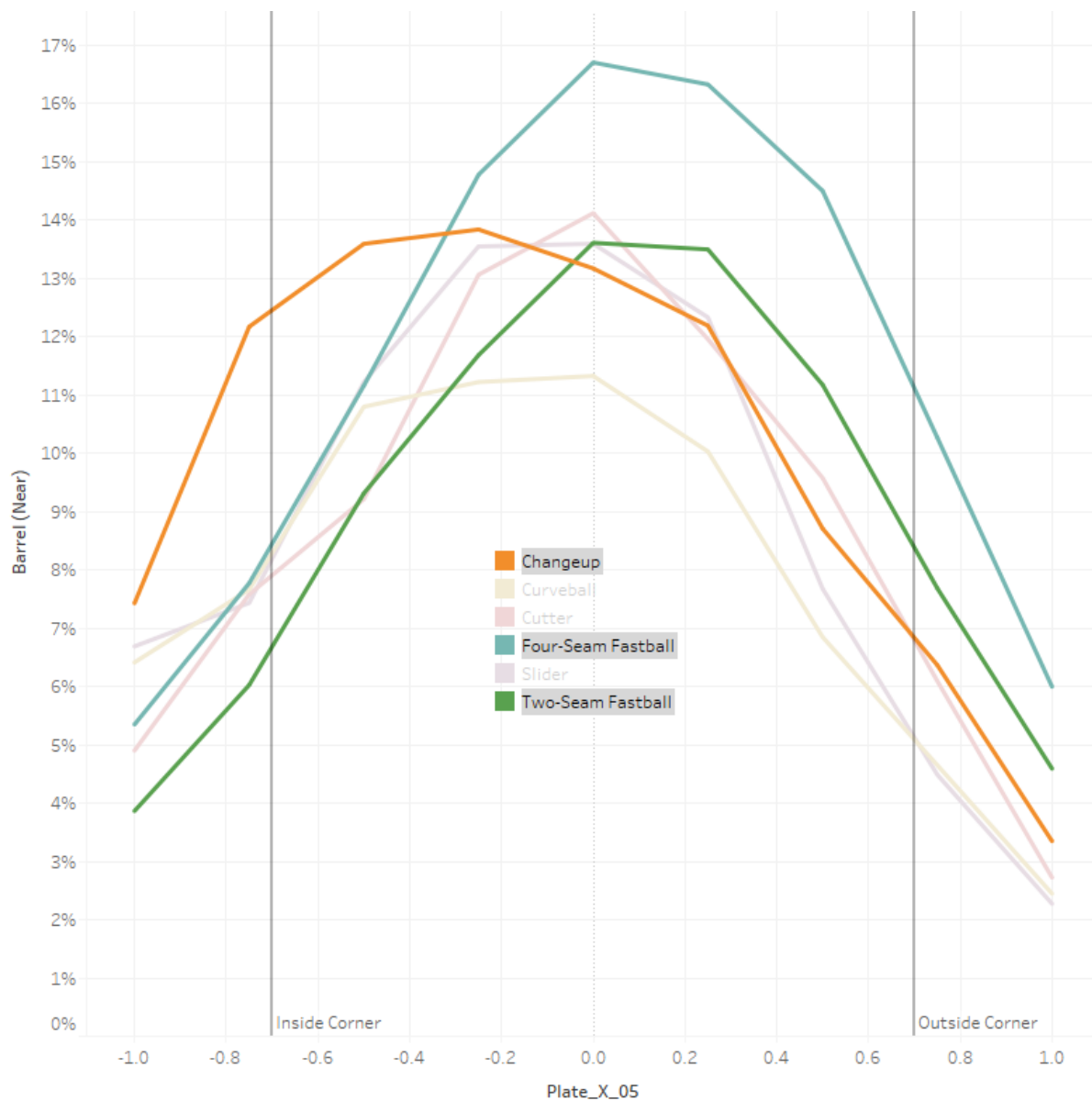
with steep declines as they get farther outside (which result in disproportionately more balls). This measure is quite imperfect, as it is assuming throwing a called ball is equally as bad as yielding a NB. However, it is useful for setting the stage for our more refined SANTA metric.

### % Good Pitches by Pitch Type and Vertical Location



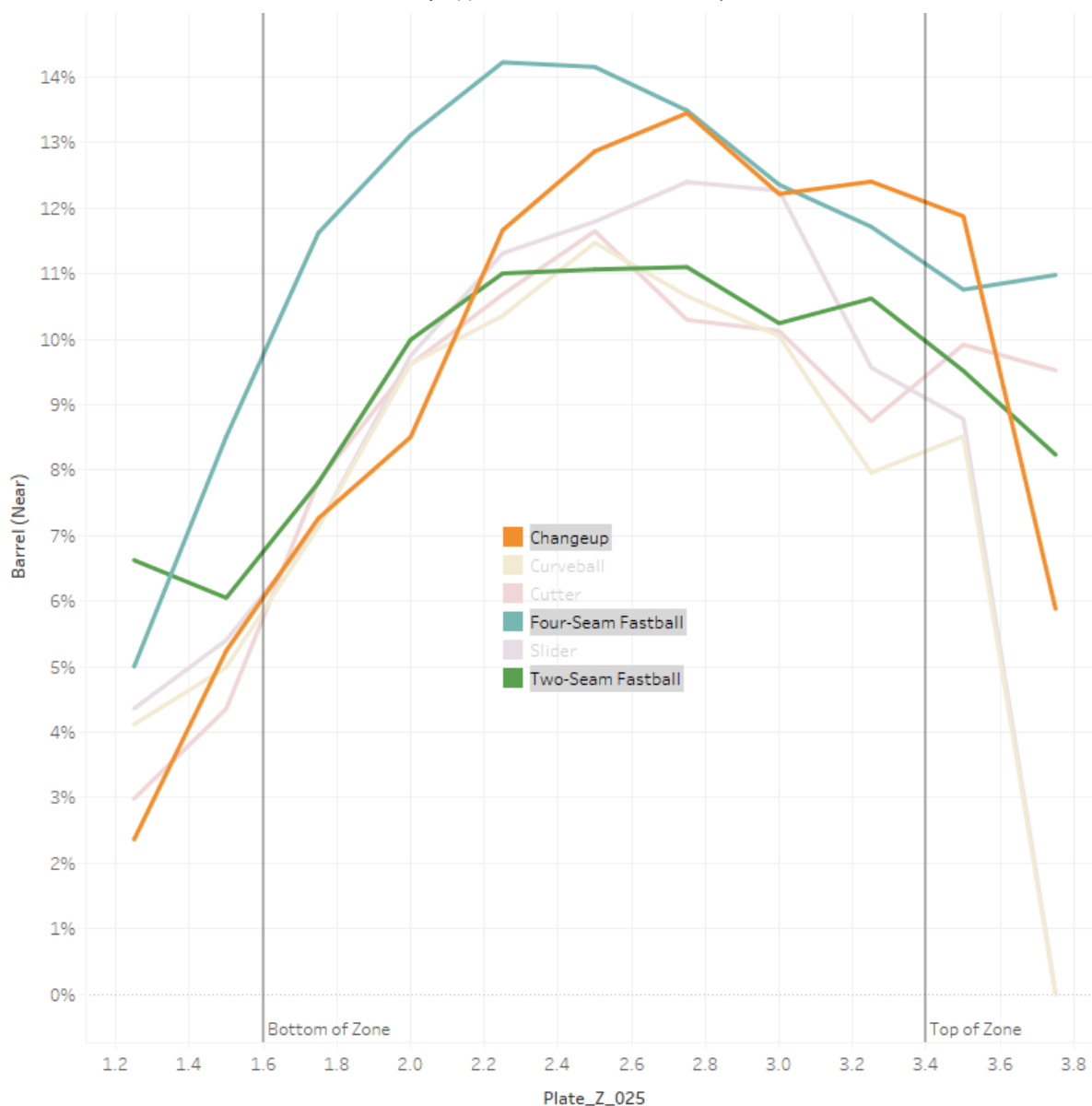
This simple metric demonstrates that change-ups should be thrown low in the zone (and away as we saw in the previous chart). Comparatively, four-seam fastballs are effective higher in the zone than two-seam fastballs. This shows how important command is for change-up pitchers; it is the only pitch that has dramatic fluctuations within the strike zone. Let's take a deeper dive into change-ups, looking at them from the lens of NB% of Balls in Play.

### NB% of BIP by Horizontal Location



Four-seam and two-seam fastballs have a very predictable pattern. Specifically, balls thrown over the middle of the plate are barreled up the most, with a fairly linear decline inside and outside. Change-ups, thrown over the middle third of the inside of the plate are hit better (when put into play) than two-seam fastballs thrown right over the heart of the plate. This highlights the tremendous importance command plays with respect to change-ups, moreso than other pitch types, which partially explains why very few pitchers can throw them well.

NB% of BIP by Vertical Location



A change-up high in the zone is twice as likely to be barreled as one at the bottom of the zone. Pretty much anything over 2.2 feet above the ground is going to get crushed, magnifying the penalty for a poorly commanded change-up. Not surprisingly, four-seam fastballs are harder to barrel up high in the zone. However, two-seam fastballs appear to have a small window of NB-prevention effectiveness, in the lower third of the strike zone.

## Building SANTA from Two Ingredients (Not Cookies and Milk)

Let's begin building SANTA by mixing our two ingredients (missing barrels and not throwing balls) into the blender with equal weight, measuring very simply the percentage of pitches that are "good," and see which pitchers grade out as the best over the past three seasons (minimum 1,000 pitches in a season, 1,500 pitches for 2015 to 2017).

## 2015

Pitcher	% Good	% NB		Pitcher	% Good	% NB
Max Scherzer	68.90%	2.20%		Matt Harvey	66.50%	1.50%
Liam Hendriks	68.20%	1.50%		Jordan Zimmermann	66.20%	1.90%
Bartolo Colon	67.60%	2.70%		Carlos Carrasco	66.10%	1.90%
Alex Wilson	67.10%	1.50%		Corey Kluber	66.10%	2.20%
Jacob deGrom	67.00%	1.60%		Hisashi Iwakuma	66.10%	2.50%
Mark Melancon	67.00%	1.40%		Chris Sale	66.10%	1.70%
Clayton Kershaw	67.00%	1.00%		Michael Pineda	66.10%	2.30%
John Lackey	66.80%	2.00%		Phil Hughes	66.00%	3.30%
David Price	66.70%	2.00%		Tony Watson	65.70%	1.60%
Fernando Salas	66.70%	2.70%		Joakim Soria	65.70%	1.30%

## 2016

Pitcher	% Good	% NB		Pitcher	% Good	% NB
Addison Reed	69.90%	2.00%		David Price	65.70%	2.00%
Andrew Miller	69.00%	1.60%		Tyler Anderson	65.70%	1.50%
Roberto Osuna	68.70%	2.00%		Chris Devenski	65.70%	1.90%
Clayton Kershaw	67.80%	1.50%		Jeurys Familia	65.50%	0.90%
Blaine Boyer	67.10%	0.90%		Bartolo Colon	65.50%	2.70%
Noah Syndergaard	66.40%	1.70%		Steven Matz	65.50%	1.70%
Max Scherzer	66.10%	1.90%		Mark Melancon	65.40%	1.40%
Liam Hendriks	66.00%	1.80%		Ken Giles	65.10%	2.00%
Seung Hwan Oh	66.00%	1.40%		Justin Verlander	64.80%	2.20%
Kelvin Herrera	65.90%	1.50%		Masahiro Tanaka	64.80%	2.50%

## 2017

Pitcher	% Good	% NB		Pitcher	% Good	% NB
Addison Reed	67.80%	2.80%		Joe Musgrove	66.10%	2.20%
Seung Hwan Oh	67.50%	2.10%		Brent Suter	66.10%	1.40%
Alex Claudio	67.10%	1.20%		Josh Tomlin	65.70%	2.50%
Felipe Rivero	66.60%	0.90%		Stephen Strasburg	65.60%	1.70%
Craig Kimbrel	66.50%	1.80%		Juan Nicasio	65.60%	1.70%
Chris Sale	66.50%	1.90%		Alex Wood	65.60%	1.90%
Clayton Kershaw	66.50%	2.00%		Jacob deGrom	65.50%	1.80%
Corey Kluber	66.40%	1.60%		Brandon McCarthy	65.10%	1.40%
Max Scherzer	66.30%	1.80%		Archie Bradley	65.10%	1.60%
Anthony Swarzak	66.20%	0.90%		Raisel Iglesias	65.10%	1.30%

## 2015-2017

Pitcher	% Good	% NB		Pitcher	% Good	% NB
Kenley Jansen	71.40%	1.60%		Seung Hwan Oh	66.70%	1.70%
Pat Neshek	70.50%	1.90%		Koji Uehara	66.70%	2.30%
Sean Doolittle	69.30%	2.10%		Joe Smith	66.40%	2.00%
Andrew Miller	67.70%	1.20%		Tony Watson	66.00%	1.80%
Addison Reed	67.60%	2.10%		Roberto Osuna	65.90%	1.90%
Nick Vincent	67.50%	2.10%		Alex Claudio	65.90%	1.60%
Dan Otero	67.40%	1.50%		Shawn Kelley	65.90%	2.50%
Matt Belisle	67.30%	1.50%		Bartolo Colon	65.80%	2.90%
Max Scherzer	67.10%	2.00%		Tommy Hunter	65.80%	1.80%
Clayton Kershaw	67.00%	1.40%		Chris Sale	65.80%	1.90%

We see a few lists that sport the names we'd expect at the top (**Clayton Kershaw, Chris Sale, Max Scherzer, Kenley Jansen**), but also **Bartolo Colon**.

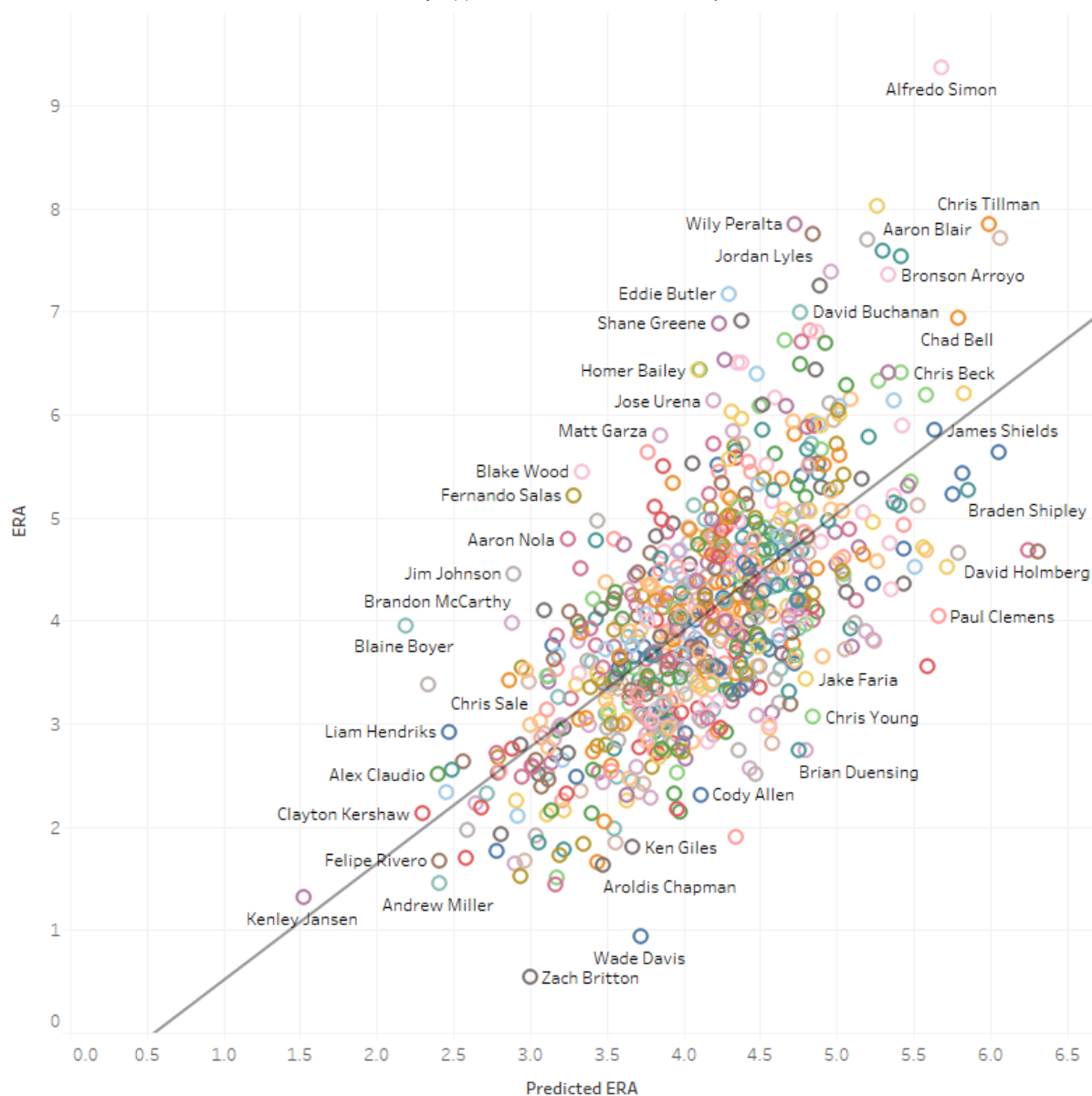
Surprisingly, even with even weights given to strike-throwing and barrel-missing, we still can produce a pretty good list of the best pitchers in baseball. The surprising thing isn't that avoiding throwing balls is good, but that we can impute success while knowing just a pitcher's strike% and NB%, without any direct knowledge of strikeout percentage (though strike% will correlate strongly to K%-BB%).

## Assigning Weights to SANTA (No Pun Intended)

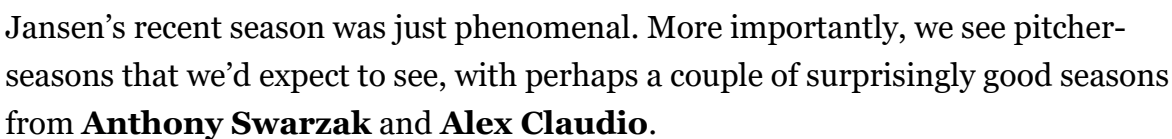
The classical scientific approach to assigning weights to our two variables would be to throw them into a predictive model against one of SIERA, ERA, FIP, xFIP or other to determine what the negative weights should be for Ball% and Near Barrel%. The author has his personal reservations with some applications of this methodology, since it can at times result in a soup strategy: Throw a whole bunch of variables into a soup and see what the model spits out. We'll begin with a basic multiple linear regression to see how well using a mixture of Ball% and NB% can be used to estimate ERA in a given season, as well as how well it does in predicting ERA in a subsequent season.

$$\text{SANTA} = -4.18 + 17.3 * \{\text{Ball Percent}\} + 96.7 * \{\text{Near\_Barrel\_Percent}\}$$

SANTA to Same year ERA | R2 = 0.42 (Min. 1,000 pitches thrown)

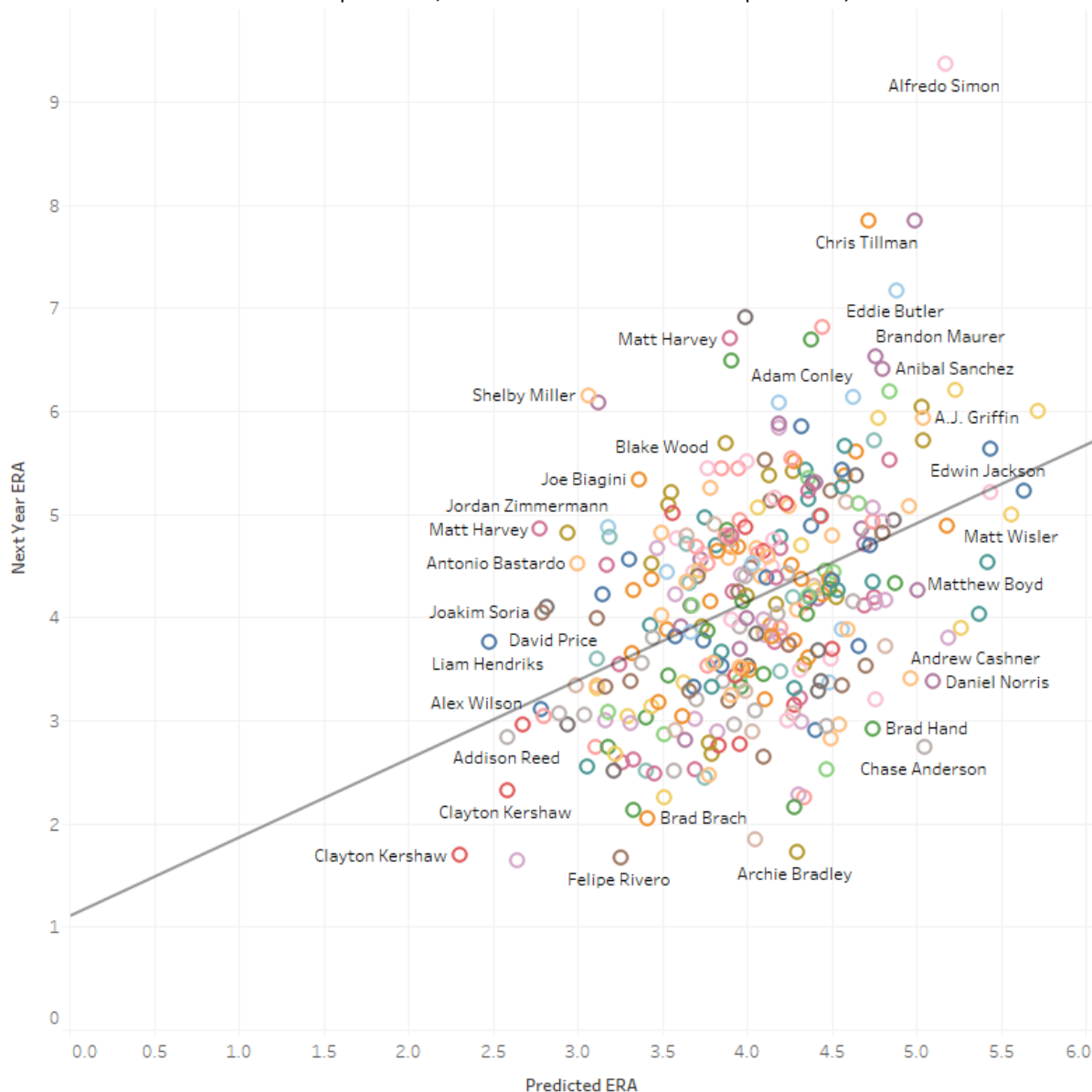


SANTA does a very decent job of identifying legitimate dominant seasons, as well as mediocre and abysmal seasons, with very few outliers. This isn't altogether surprising given a key ingredient is "How much hard contact did the pitcher give up?" But is encouraging in that it tells a very compelling story. Let's zoom in on the top end.



The magic algorithm spit out a slightly different formula for next-year ERA, but the results weren't materially different, so we'll use the same weights as above to predict a pitcher's next-year ERA. For this exercise, we had 327 pitcher seasons in which a pitcher threw at least 1,000 pitches in both seasons. With this significant sample size, we get an outstanding  $R^2$  of 0.17, which puts it in the same league as SIERA for predicting ERA.

## SANTA to Next-Year ERA | Min. 1,000 IP in Each Season | R2 0.17



The relationship holds up, even when we lower the threshold to 500 pitches in each season ( $R^2$  goes down to 0.11), indicating we can quickly measure a pitcher's talent rapidly (i.e. using a sample size as small as five full starts). However, it's clearly far more reliable with larger individual sample sizes.

## SANTA Leaderboards

Pitcher	2017		Pitcher	2016		Pitcher	2015
Corey Kluber	2.91		Clayton Kershaw	2.58		Clayton Kershaw	2.30
Max Scherzer	3.06		Tyler Anderson	2.94		Matt Harvey	2.78
Chris Sale	3.09		Noah Syndergaard	3.01		Jake Arrieta	2.78
Stephen Strasburg	3.11		Jose Fernandez	3.08		Jacob deGrom	2.80
Jacob deGrom	3.18		Rich Hill	3.11		Max Scherzer	2.94
James Paxton	3.19		Steven Matz	3.12		Chris Sale	2.99



Luis Severino	3.20		Max Scherzer	3.22		Shelby Miller	3.06
Clayton Kershaw	3.24		Chris Devenski	3.22		Stephen Strasburg	3.11
Alex Wood	3.24		Aaron Nola	3.24		David Price	3.12
Jimmy Nelson	3.32		David Price	3.31		John Lackey	3.12
Joe Musgrove	3.43		James Paxton	3.31		Carlos Carrasco	3.16
Aaron Nola	3.47		Carlos Martinez	3.32		Madison Bumgarner	3.18
Rich Hill	3.50		Gerrit Cole	3.33		Jordan Zimmermann	3.18
Brad Peacock	3.53		Juan Nicasio	3.33		Clay Buchholz	3.19
Clayton Richard	3.55		Kyle Hendricks	3.4		Noah Syndergaard	3.26
Minimum 1,500 pitches							

## Conclusion

SANTA organically rated Kershaw's 2015 and 2016 seasons as the best in the majors and thus passed this author's litmus test: Does it accurately capture Kershaw's greatness? In other words, if this model were to have a high correlation with ERA or Next-Year ERA, but for some reason the best pitchers in baseball didn't rank well, it wouldn't lend a lot of credence to the model. With SANTA, especially in 2017, we see a near-flawless list of top pitchers, with only a few surprises mixed in such as **Tyler Anderson** and **Joe Musgrove**. I'm not sure what to make of Musgrove's 3.43 SANTA compared to his 4.77 actual ERA, except that if he is really good next year, SANTA is a genius; and if he isn't, well, you know, random variation.

## References & Resources

- Tom Tango, Tangotiger Blog, "**Statcast Lab: Barrels**"
- Wikipedia, "**Porter's generic strategies**"
- FanGraphs, Library, "**FIP**"

[Facebook](#)
[Twitter](#)
[Reddit](#)
[Email](#)


---

Eli Ben-Porat is a Senior Manager of Reporting & Analytics for Rogers Communications. The views and opinions expressed herein are his own. He builds data visualizations in Tableau, and preps data in Alteryx. Follow him on Twitter [@EliBenPorat](#).

---