Wharton
University of Pennsylvania

Wharton Moneyball Academy

Standard Units and the Normal Curve

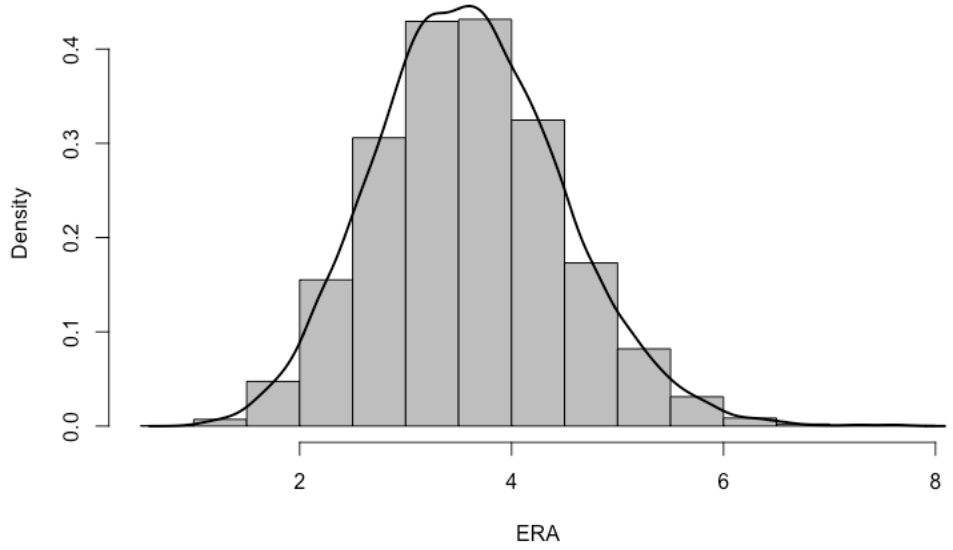
Professor Abraham Wyner, Department of Statistics,
The Wharton School of the University of Pennsylvania

The Normal Curve applied to data

Many datasets follow a "Bell Shaped" curve quite well. For these datasets the empirical rule holds precisely. In fact, every quantile can be calculated using only the mean and SD.

Example 1: Consider the 9833 individual seasons pitched by MLB starters (min 150 IP).

Distribution of ERAs



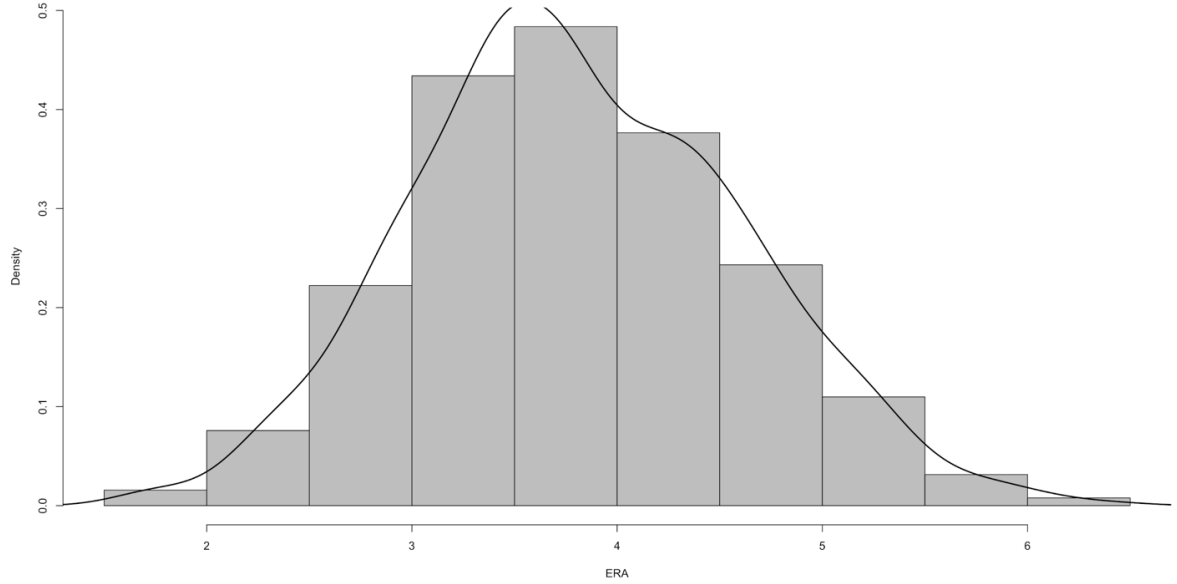
Mean	3.597
SD	0.8861
N	9833

Source: Lahman's Baseball Database

The normal curve applied to data – example pitching

765 seasons for starting pitchers since 2010.

Distribution of ERAs



Mean	3.807
SD	0.8015
N	765

You can "look-up" the frequency under a normal curve between any two points.

Source: Lahman's Baseball Database

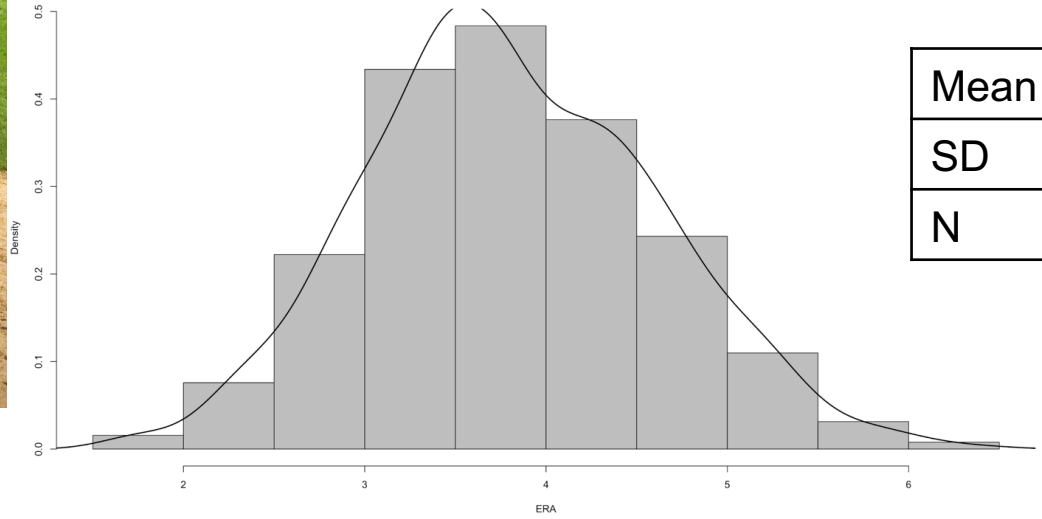
The normal curve applied to data – example pitching

So, for example, how rare has it been (in last 5 years) for a starter to have a 2.50 ERA or below?

If 2.50 was 1 SD then only 16% of pitchers would have a lower ERA.
If 2.50 was 2 SD then only 2.5% of pitchers would have a lower ERA.
2.50 is about 1.6 SDs less than the mean. It is closer to 2.5% than 16%.

Jake Arrieta, 2014
2.53 ERA

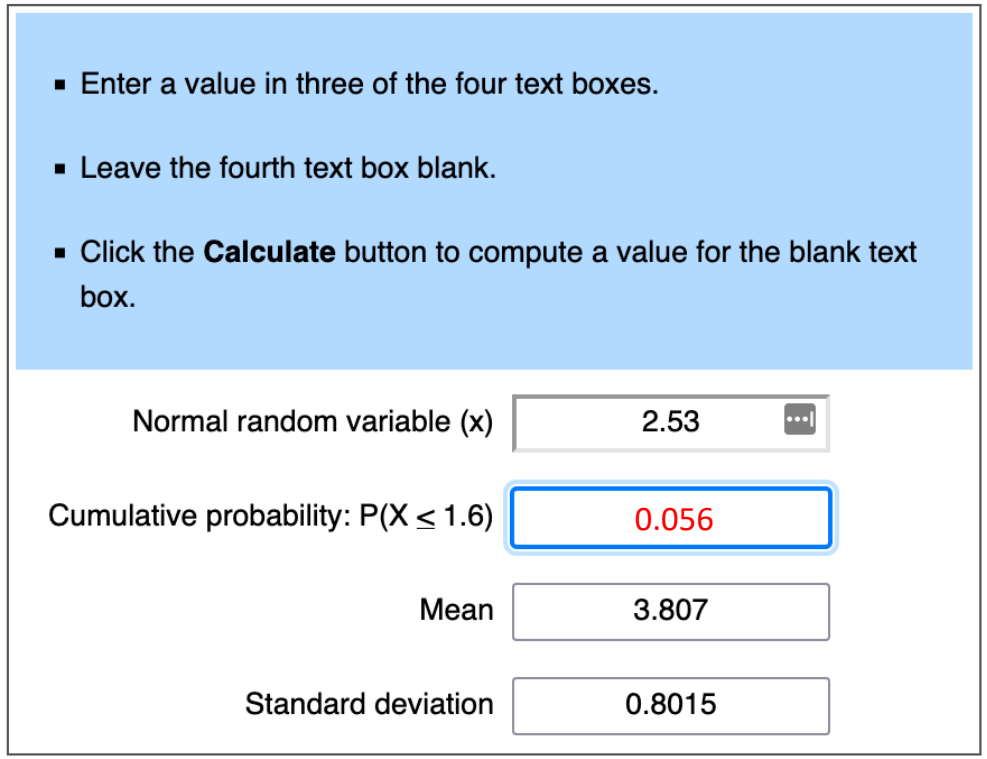
Distribution of ERAs



Mean	3.807
SD	0.8015
N	765

Normal distribution calculator

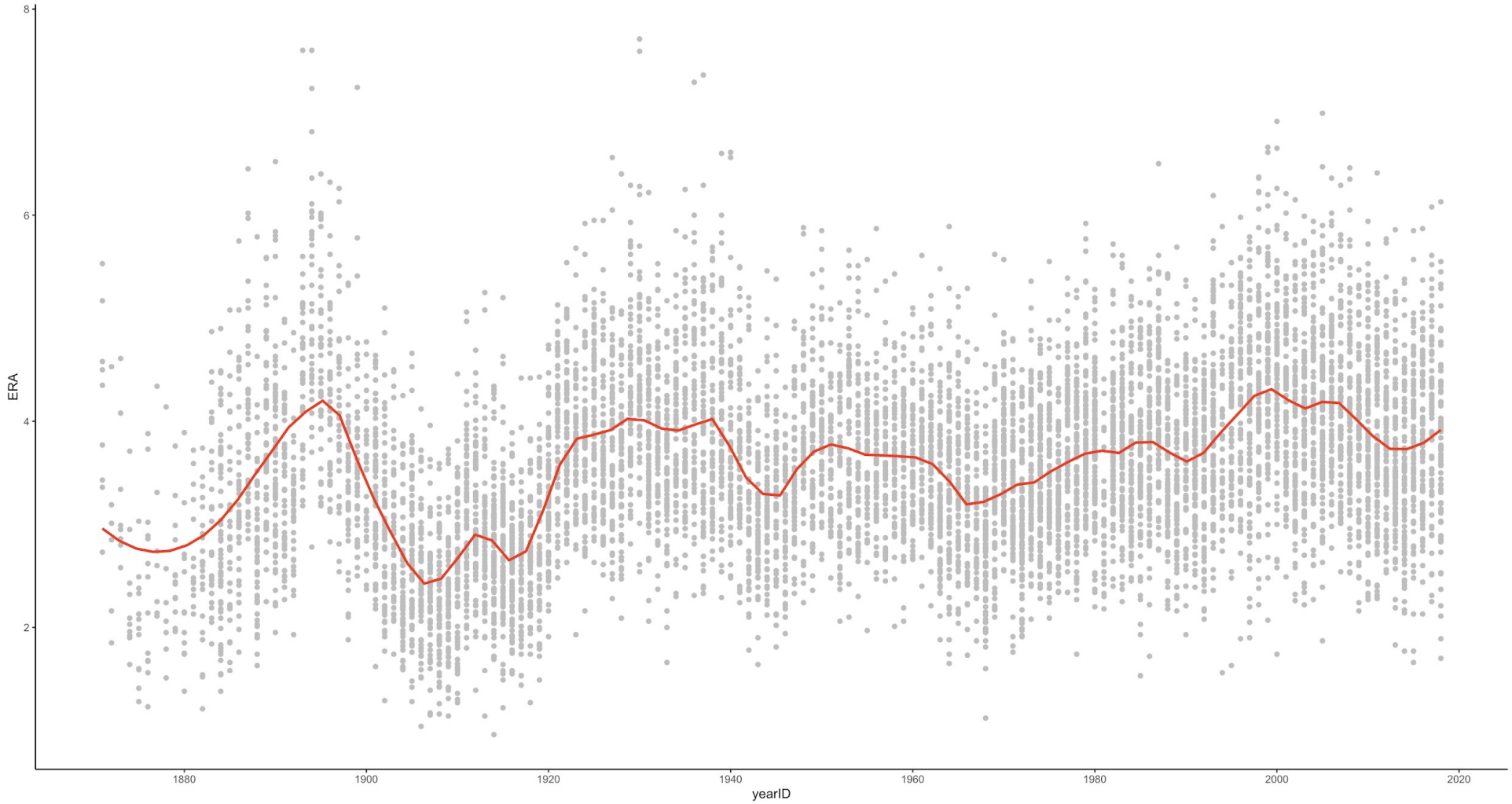
Use a calculator: http://stattrek.com/online-calculator/normal.aspx



$x_{Arrieta,\ 2014}$	2.53
Mean	3.807
SD	0.8015
N	765

You can of course use R or a calculator.

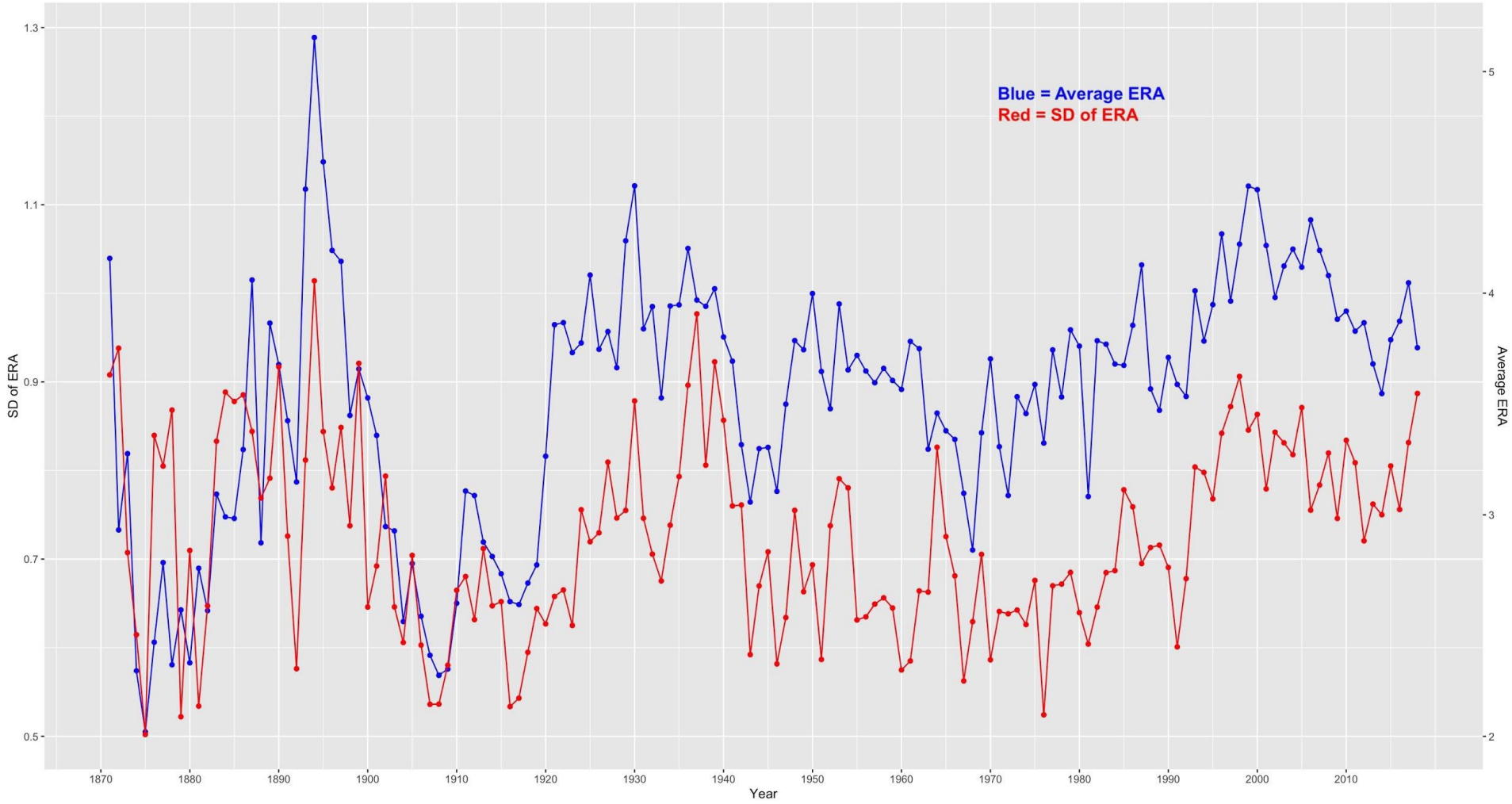
Which pitcher had the best year of all time?



Source: Lahman's Baseball Database

Which pitcher had the best year of all time?

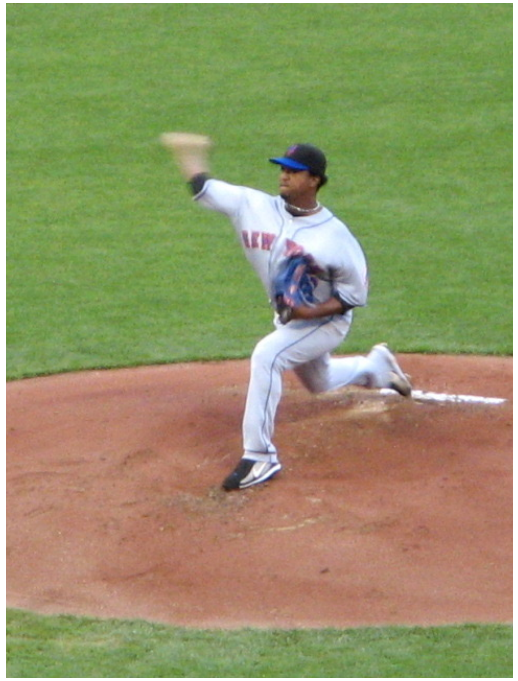
Adjust the comparison for ERA by subtracting



Source: Lahman's Baseball Database

Which pitcher had the best year of all time?

Player	Year	ERA	Standardized ERA (in SU)
Pedro Martinez	2000	1.74	-3.151
Dwight Gooden	1985	1.53	-2.998
Mark Eichorn	1986	1.72	-2.938
Greg Maddux	1994	1.56	-2.929
Greg Maddux	1995	1.63	-2.874
Dolph Leonard	1914	0.96	-2.858
Bob Gibson	1968	1.12	-2.854
Kevin Brown	1996	1.89	-2.822
Roger Clemens	2005	1.87	-2.757
Ron Guidry	1978	1.74	-2.756
Pedro Martinez	1999	2.07	-2.729
Dolf Luque	1923	1.93	-2.696
Walter Johnson	1913	1.14	-2.670
Cart Hubbel	1933	1.66	-2.599
Whitey Ford	1958	2.01	-2.583
Roger Craig	1959	2.06	-2.538
Lefty Grove	1931	2.06	-2.536



Pedro Martinez

Photo "Pedro Martinez warms up" by tingley

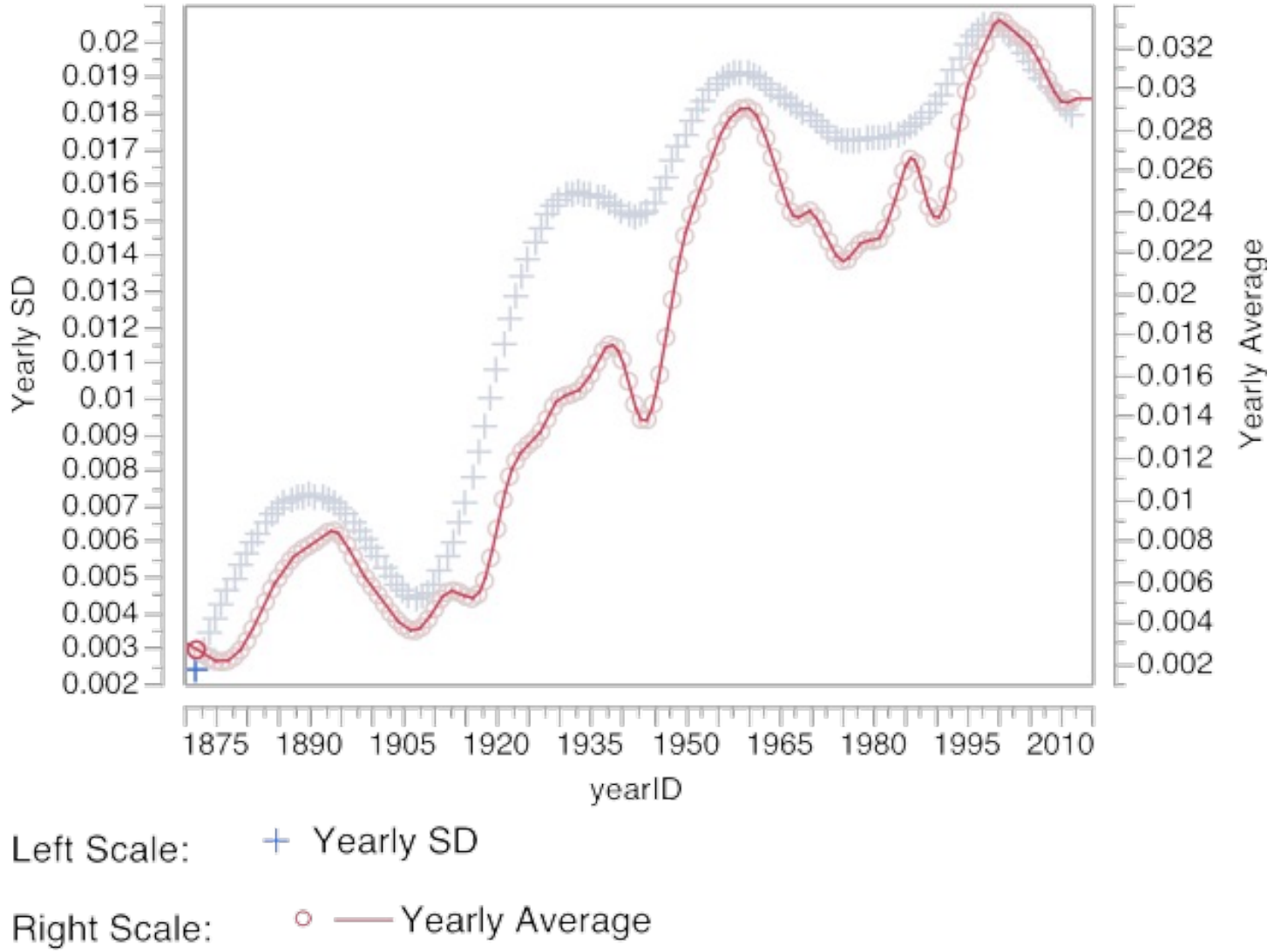
How about WAR as a measure of best season ?

Year	Pitcher	Team	GWAR	Z-score
1966	Sandy Koufax	LAN	11.543	4.298
1968	Bob Gibson	SLN	11.045	4.032
1985	Dwight Gooden	NYN	11.039	4.029
1997	Roger Clemens	TOR	10.97	3.993
1972	Steve Carlton	PHI	10.712	3.855
1953	Robin Roberts	PHI	10.429	3.705
1963	Sandy Koufax	LAN	10.405	3.692
1978	Ron Guidry	NYA	10.332	3.653
2000	Pedro Martinez	BOS	10.294	3.633
1972	Gaylord Perry	CLE	9.997	3.474
1964	Dean Chance	LAA	9.782	3.360
1971	Wilbur Wood	CHA	9.733	3.334
1971	Tom Seaver	NYN	9.67	3.300
1971	Vida Blue	OAK	9.67	3.300
1965	Sandy Koufax	LAN	9.595	3.260-

Why do you think modern pitchers are not appearing on this list?

(reminder: WAR is ERA, league and park adjusted)

Baseball: What about most HRs in a season?

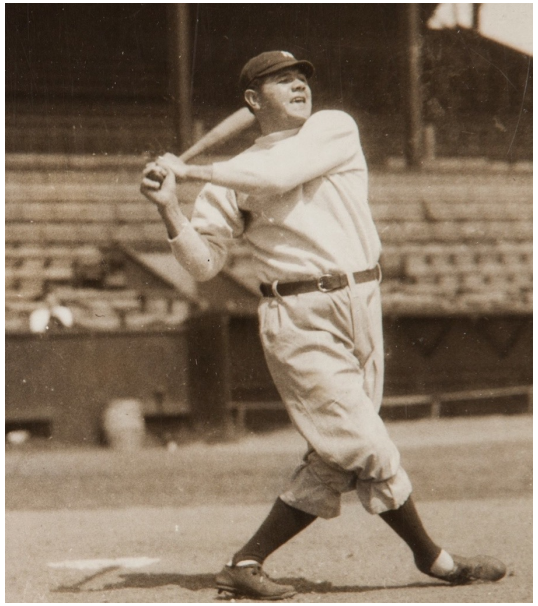


Source: Lahman's Baseball Database

Baseball: What about most HRs in a season?

Who had the best HR season ever, relative to the players at the time?

Ruth in 1927 ?
Bonds in 2001?
Maris in 1961?



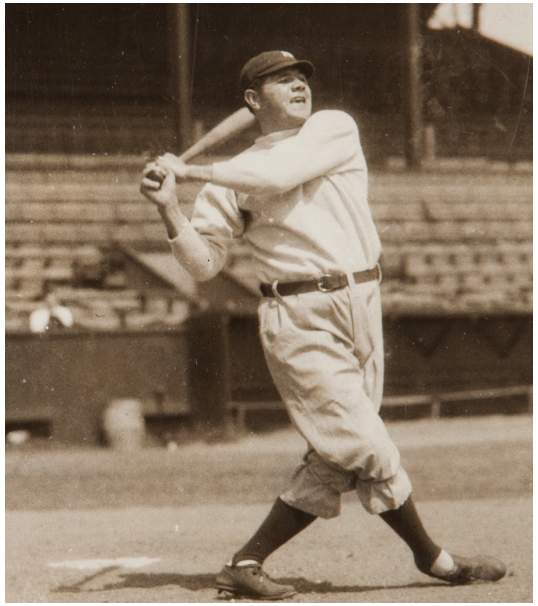
Babe Ruth

Source: Lahman's Baseball Database

Baseball: What about most HRs in a season?

HR Rate, Standardized

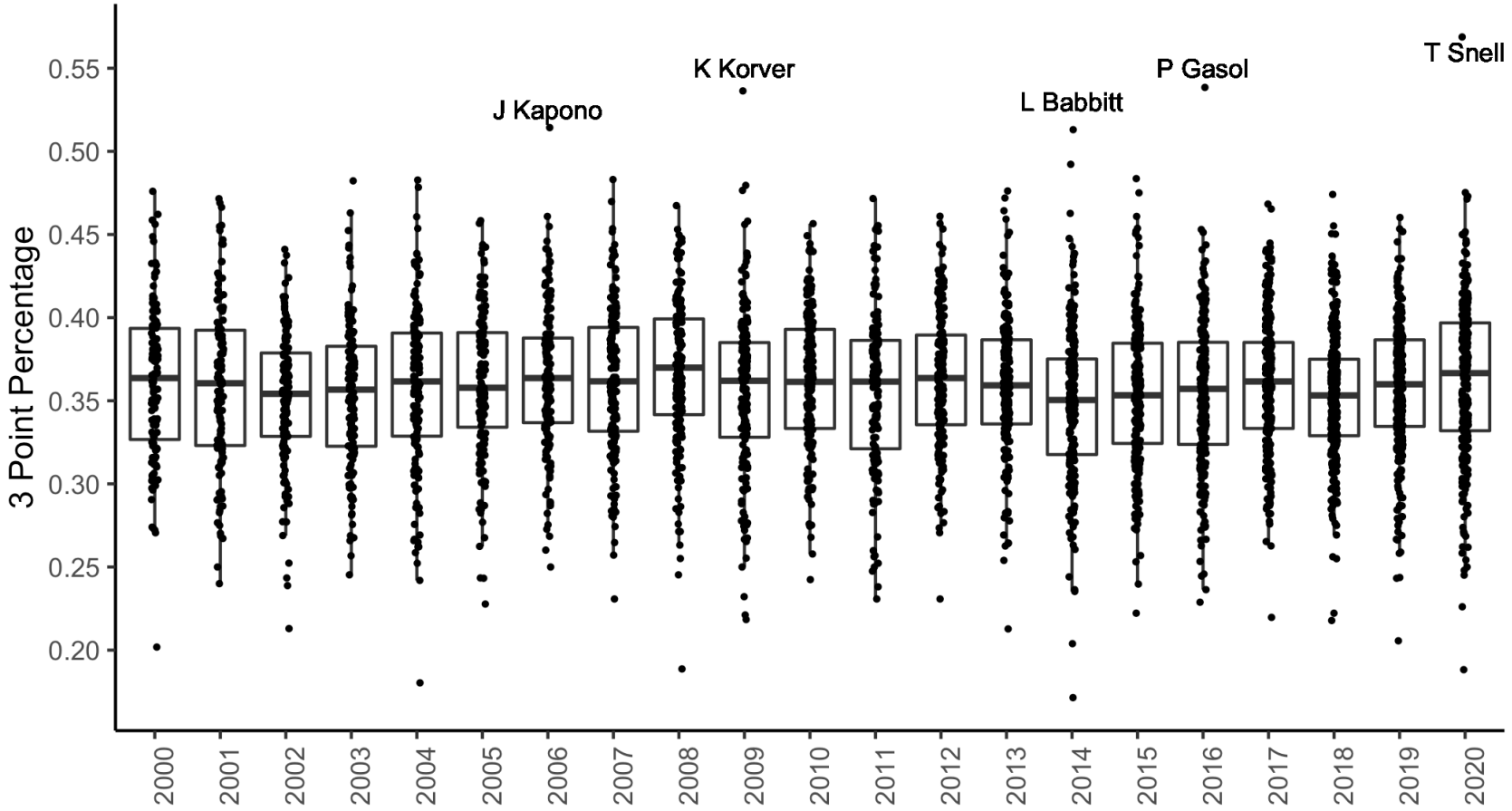
Player	year	team	HR	Standardized HR rate
Babe Ruth	1920	NYA	54	10.16
Babe Ruth	1921	NYA	59	8.634
Babe Ruth	1927	NYA	60	6.619
Babe Ruth	1922	NYA	35	6.178
Babe Ruth	1919	BOS	29	5.971
Barry Bonds	2001	SFN	73	5.916
Babe Ruth	1928	NYA	54	5.749
Babe Ruth	1926	NYA	47	5.709
Hank Greenberg	1938	DET	58	5.632
Babe Ruth	1924	NYA	46	5.586



Babe Ruth

Source: Lahman's Baseball Database

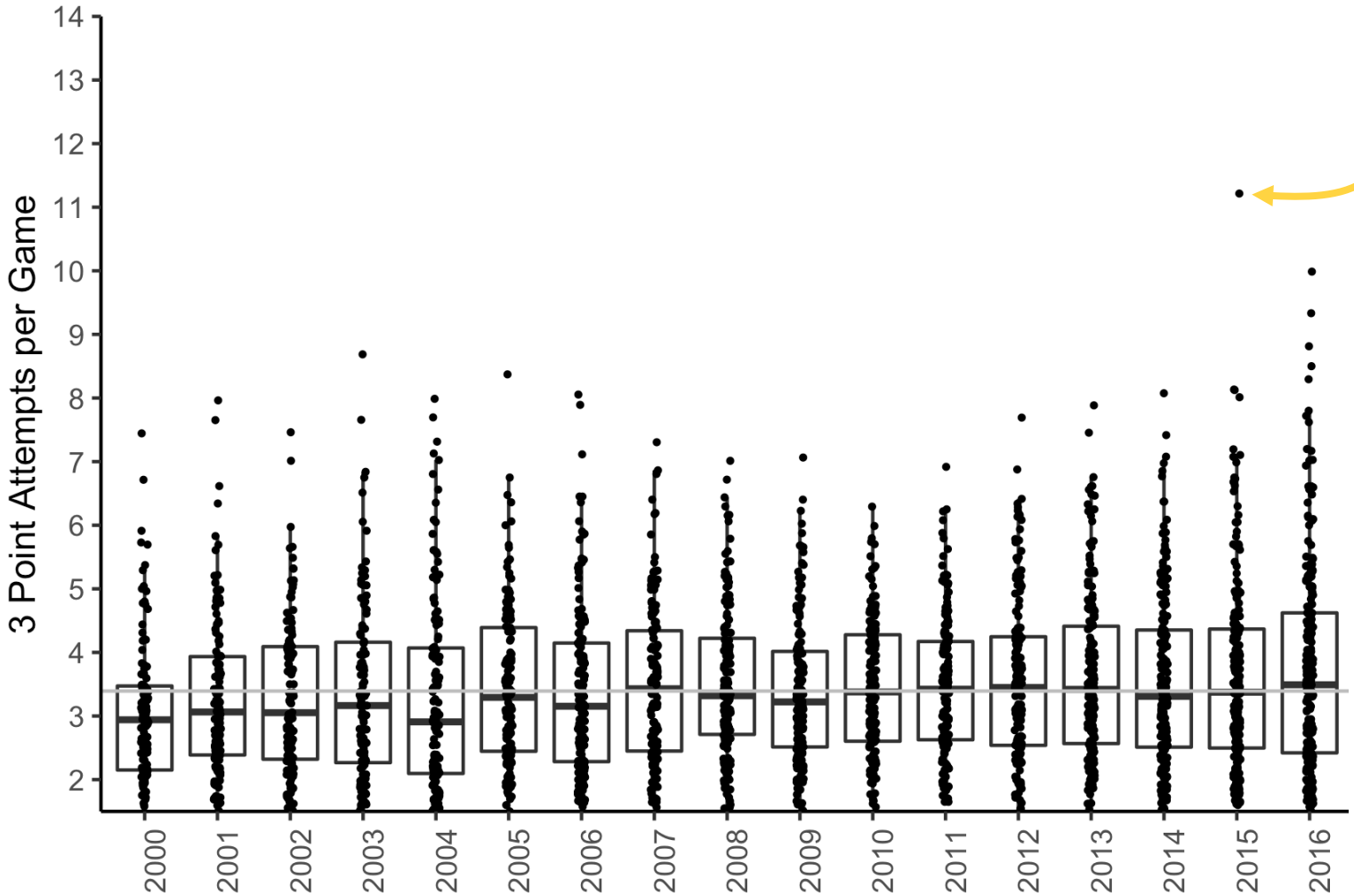
Basketball: How about 3pt success percentage ?



Pretty stable....
Surprising?

Source: www.nba.com/stats

Basketball: 3pt attempts per game 2000 - 2016



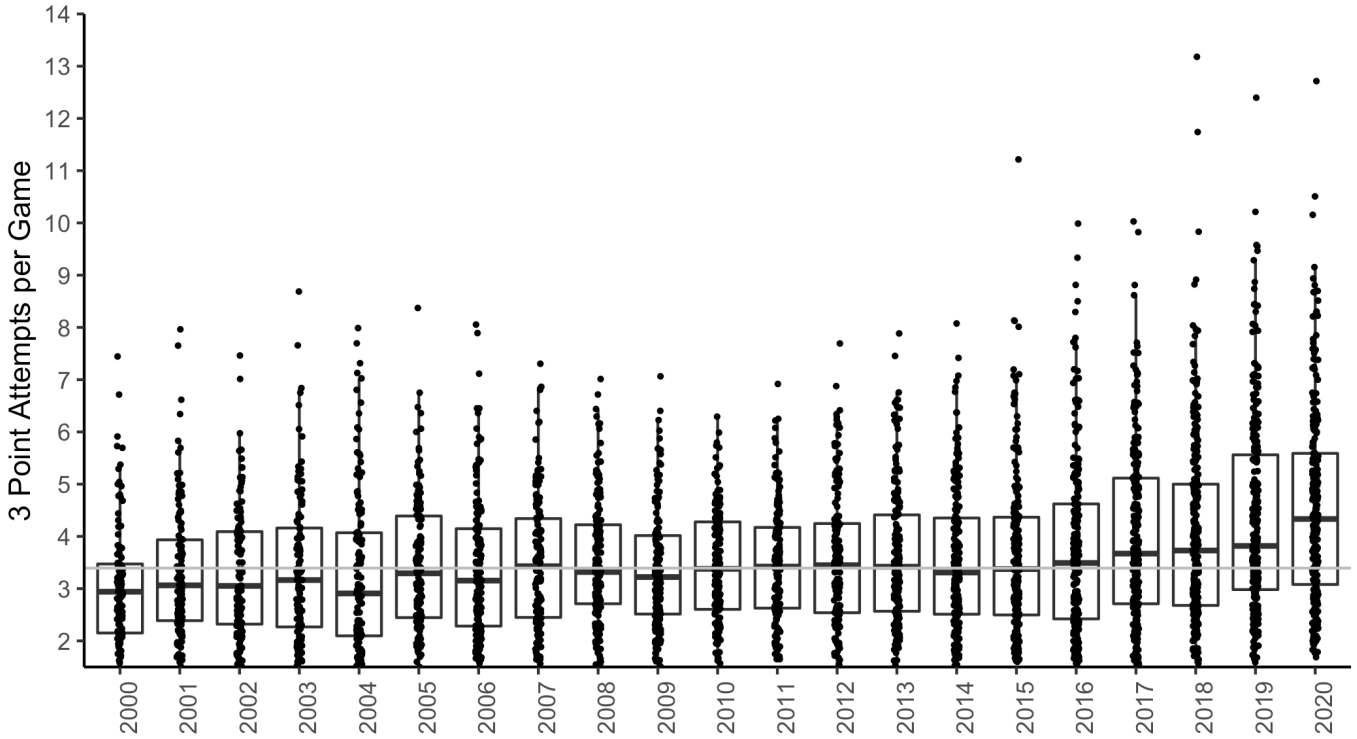
Steph Curry was a big outlier here.

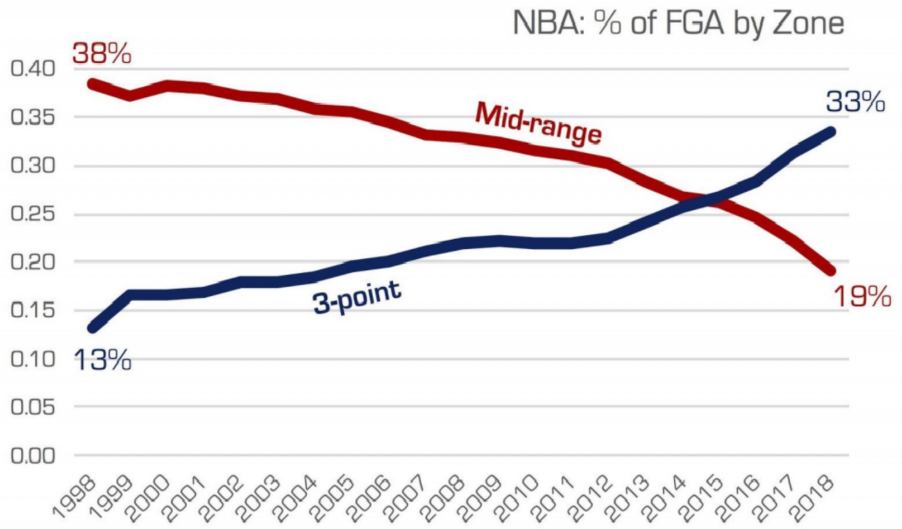
In the recent seasons, other players have followed his trend of shooting a lot of threes.

Stephen Curry

Source: www.nba.com/stats
Photo by Keith Allison from Hanover, MD, USA, CC BY-SA 2.0

Basketball: 3pt attempts per game 2000 - 2020

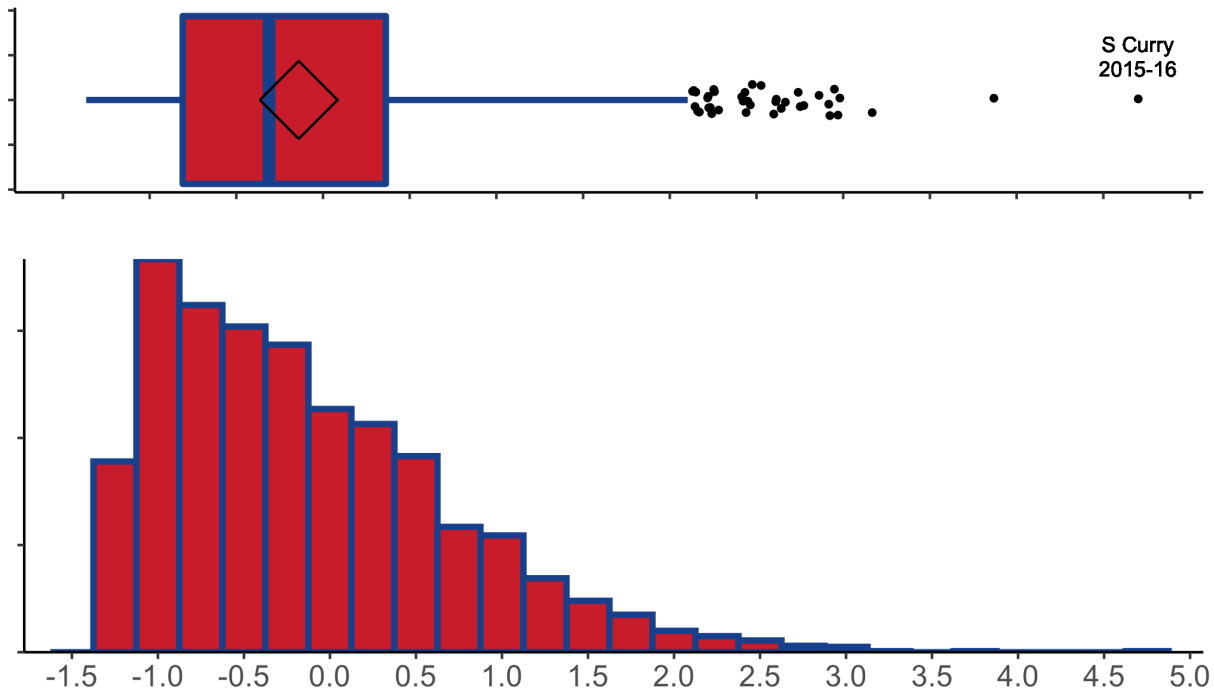




Source: www.nba.com/stats

Basketball: 3pt attempts per game 2000 - 2016

3pt Attempts, Standardized by Season



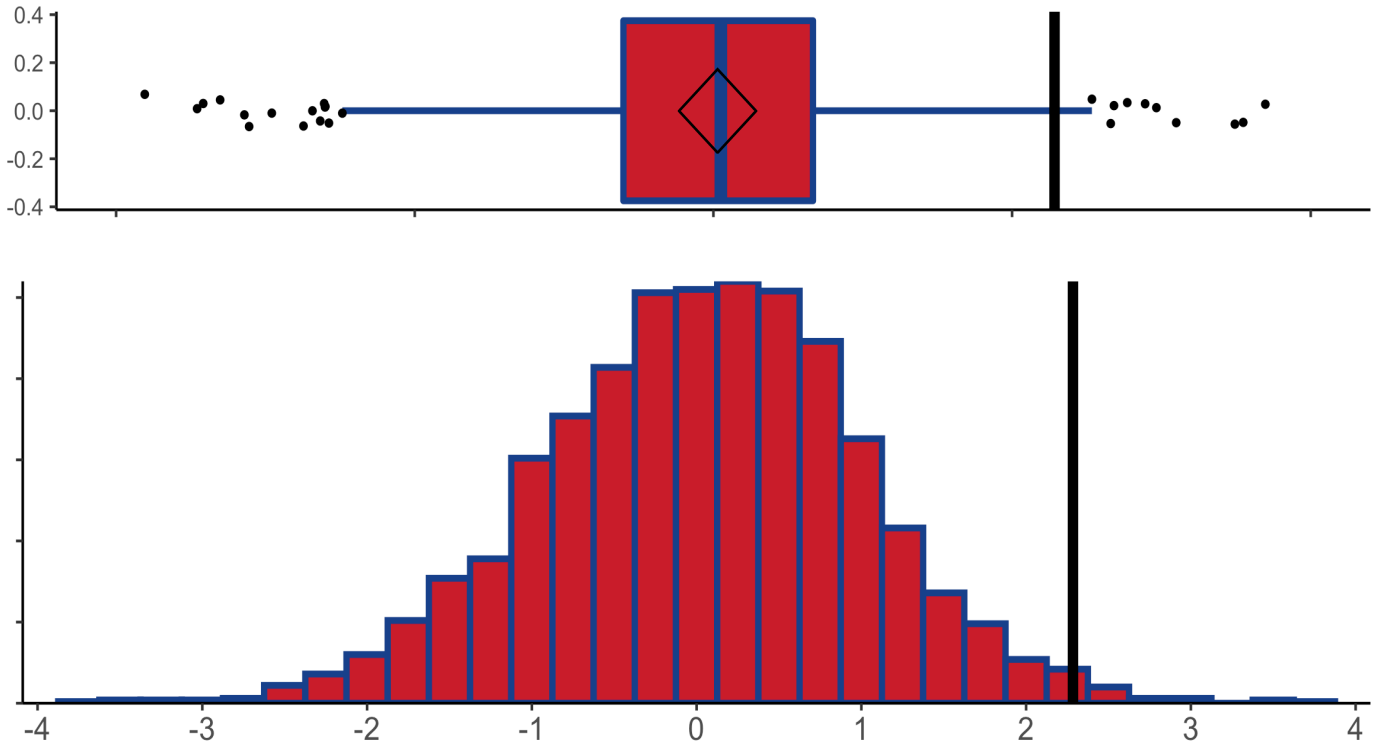
Stephen Curry

Steph Curry 2015-16: 5.96 Standard Units

Source: www.nba.com/stats
Photo by Keith Allison from Hanover, MD, USA, CC BY-SA 2.0

Basketball: 3pt shooting % 2000 - 2016

3 Pt Percentage, Standardized by Season (minimum 100 attempts)



Stephen Curry

Source: www.nba.com/stats
Photo by Keith Allison from Hanover, MD, USA, CC BY-SA 2.0

Basketball: 3pt attempts per game in standard units

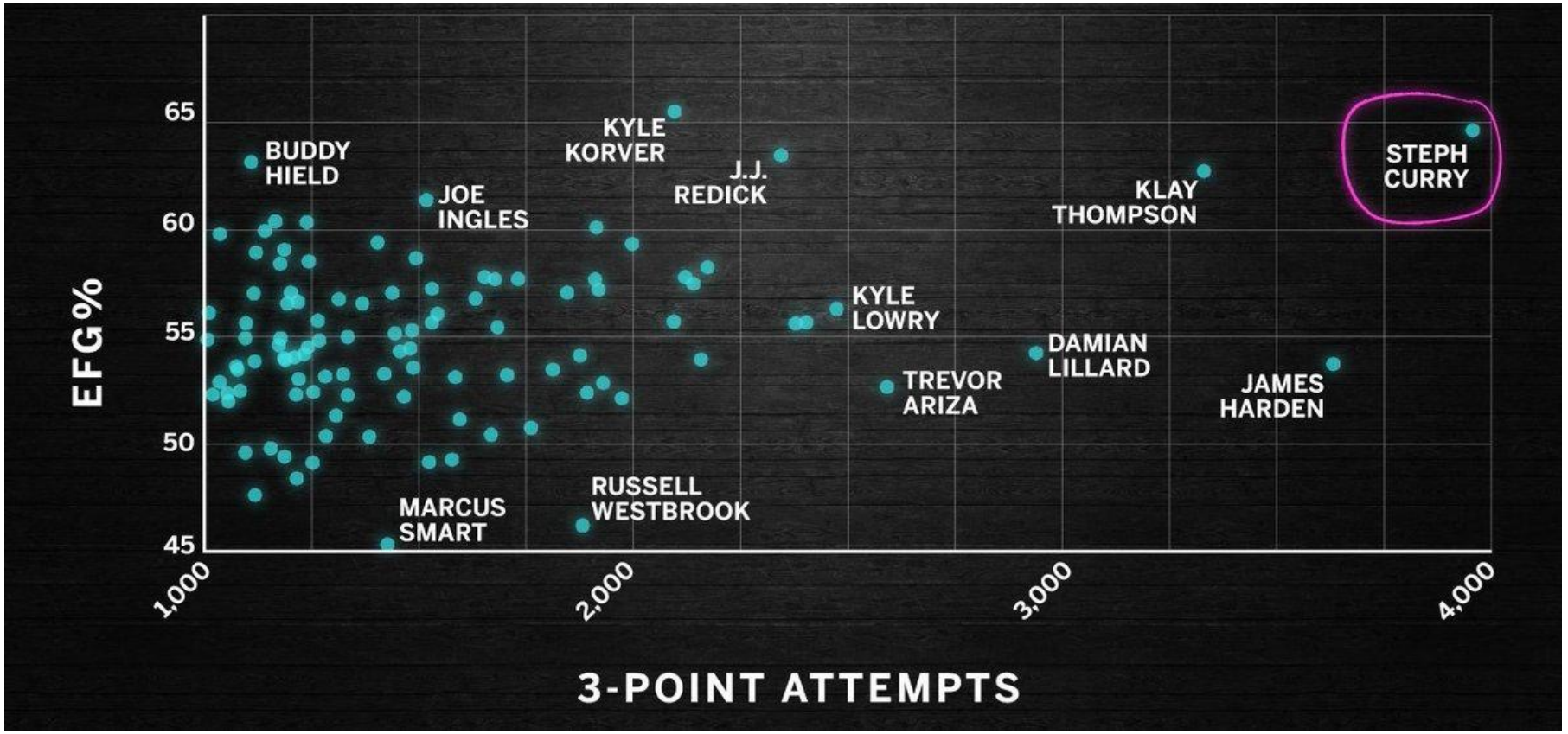
Player	Season	3Pt made/ game	Standardized 3Pt made / game
Stephen Curry	2015	5.09	6.01
Stephen Curry	2018	5.13	5.24
James Harden	2018	4.85	4.84
Stephen Curry	2020	5.35	4.66
Stephen Curry	2014	3.58	4.12
Ray Allen	2005	3.45	4.07
Stephen Curry	2012	3.49	4.06
Stephen Curry	2016	4.10	4.05
Ray Allen	2001	3.32	4.04
Stephen Curry	2017	4.16	3.95
James Harden	2019	4.40	3.73
Stephen Curry	2013	3.35	3.56
Antoine Walker	2000	2.73	3.48
Klay Thompson	2015	3.45	3.43

Steph Curry 2015-16 season
2.4 Standard Units above average in 3pt %
5.1 Standard Units above average in 3pt attempts

Source: www.nba.com/stats

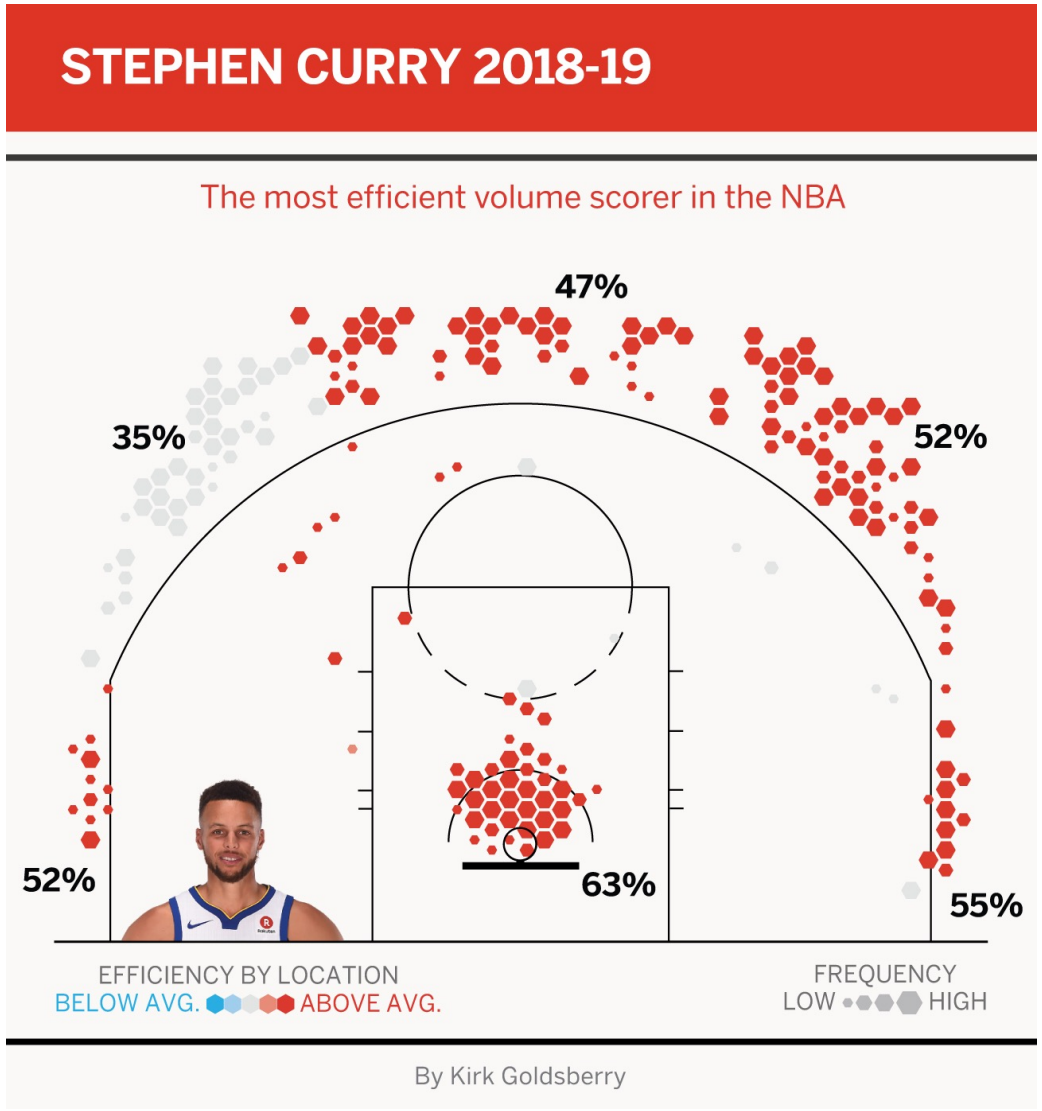
Basketball: Trend in 3pt attempts

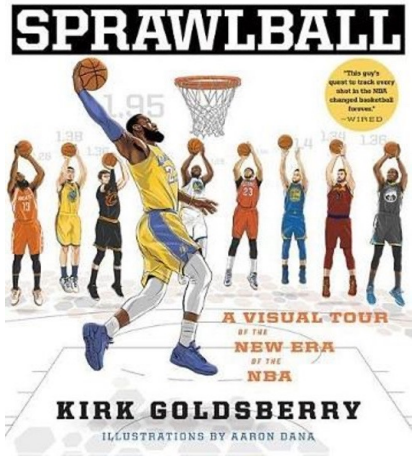
3-Point Volume and Efficiency
2014/15 through 2019/20 Seasons



Kirk Goldsberry , Steph Curry is still completely unfair, ESPN.com, January 16, 2019. https://www.espn.com/espn/print?id=25771897#.

Basketball: Trend in 3pt attempts, Steph Curry





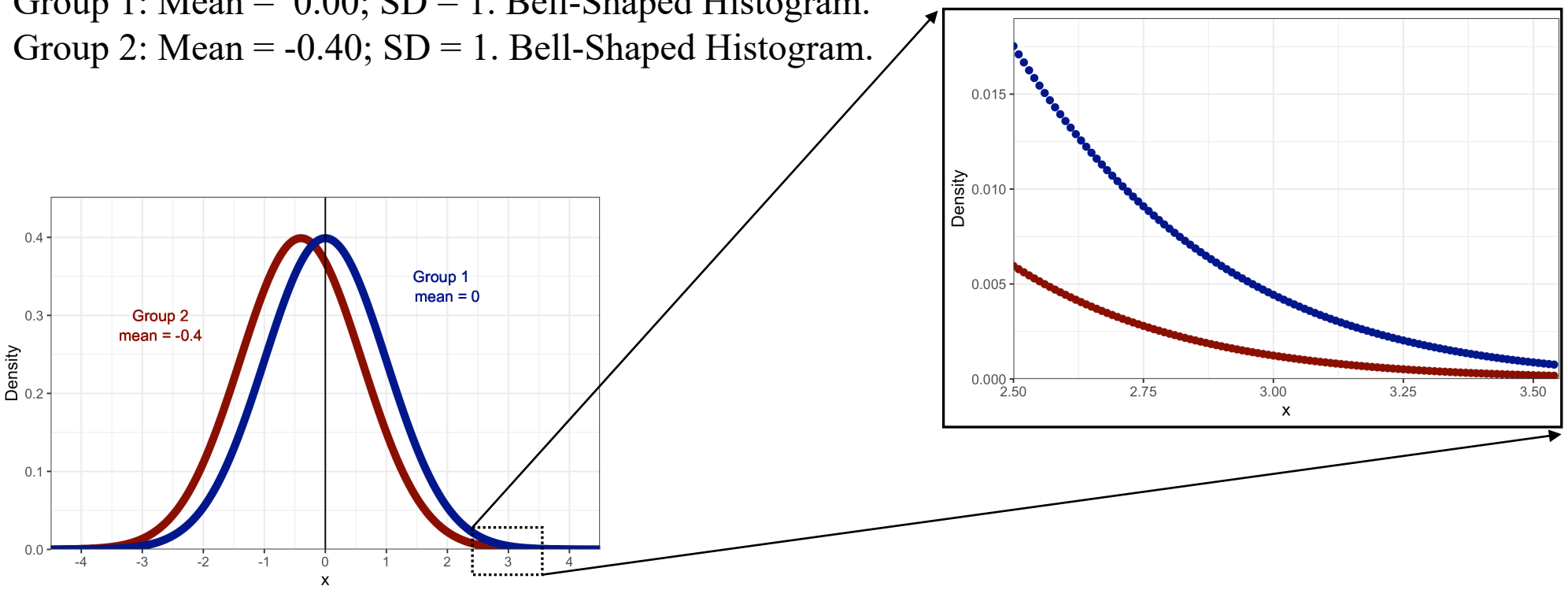
Kirk Goldsberry , Steph Curry is still completely unfair, ESPN.com, January 16, 2019. https://www.espn.com/espn/print?id=25771897#
Goldsberry, K. (2019). Sprawlball: A visual tour of the new era of the NBA. Mariner Books.

Hockey: Points

Player	Year	Points	Z-score
gretzwa01	1983	205	5.0708723474
gretzwa01	1981	212	4.8036146919
gretzwa01	1984	208	4.7077055474
gretzwa01	1985	215	4.6274918697
gretzwa01	1982	196	4.5941712478
gretzwa01	1986	183	4.1574799431
esposph01	1973	145	3.9736521169
lemiema01	1988	199	3.8199737718
lemiema01	1987	168	3.7530036969
tardima01	1975	148	3.7238005495
lemiema01	1995	161	3.7172098566
gretzwa01	1980	164	3.6129981189
gretzwa01	1990	163	3.6093656725
esposph01	1972	130	3.3406854586
gretzwa01	1993	130	3.3224356659

Small differences in means can mean big difference in extremes

Consider two distributions, with slightly different means:
Group 1: Mean = 0.00; SD = 1. Bell-Shaped Histogram.
Group 2: Mean = -0.40; SD = 1. Bell-Shaped Histogram.

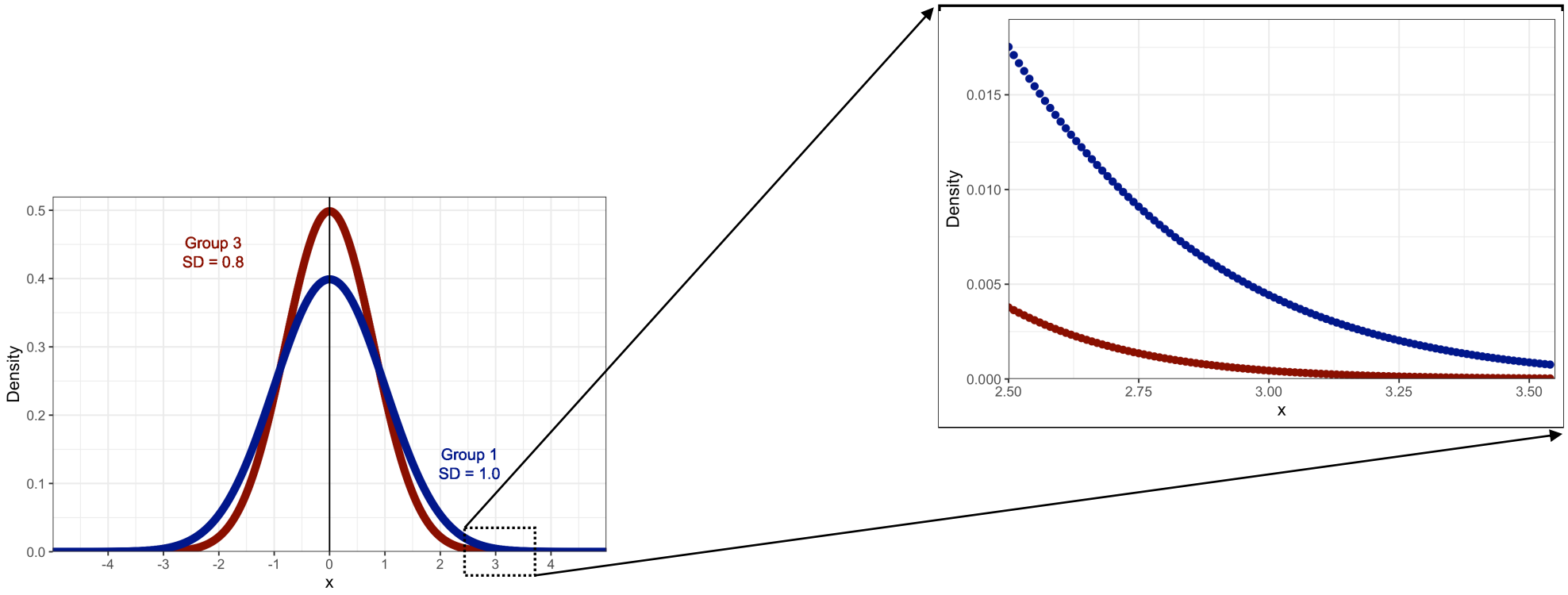


A randomly chosen individual from Group 1 is 60% likely to have larger value than a randomly chosen individual from Group 2.
The distributions are not practically that different.

But among the largest members, Group 1 dominates.

Small differences in SD can mean big difference in extremes

Another example
Group 1 and Group 3 have the same mean, but different SD; the SD for Group 3 is 80% of Group 1.
Group 1: Mean = 0.0; SD = 1.0. Bell-Shaped Histogram.
Group 3: Mean = 0.0; SD = 0.8. Bell-Shaped Histogram.



Implications

Increasing variance by only a small amount increases the tail probabilities.

Applications

- Underdogs should take risks!
- Teams that are behind.. Should take risks!

A case study: Fumbles and the New England Patriots

Deflategate 2014/15

Warren Sharp's Hypothesis: An under-inflated football is easier to grip, particularly in bad weather.

He then conjectures that if you look at the Patriot's fumbles compared to other teams (who also play in open roofed stadiums) the Patriots should stand out,

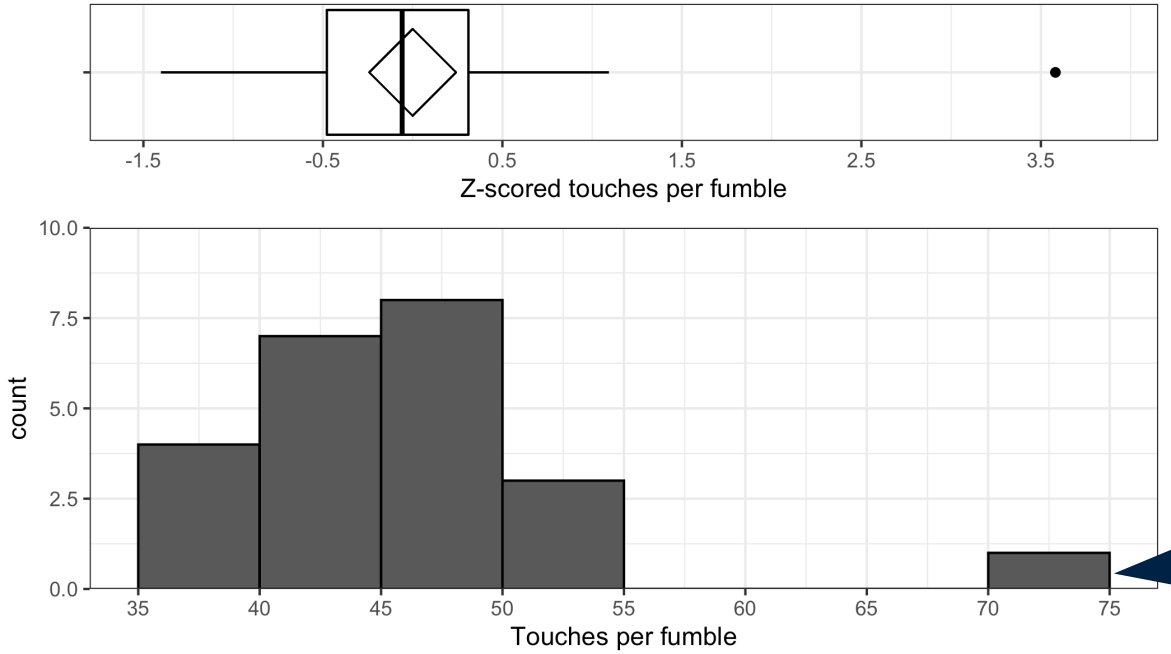
"looking at the fumble data… the Patriots' prevention of fumbles was nearly impossible"

Bill Belichick

Sharp, Warren, 'Stats show Patriots became nearly fumble-proof after '06 rule change proposed by Tom Brady', *SportingNews*, Jan 27, 2015.
www.sharpfootballanalysis.com

A case study: Fumbles and the New England Patriots

Histogram, Box Plot, Mean, and SD for Touches/Fumble (All NFL)



73.1 touches/fumble

Mean	46.41
SD	7.45
N	23
5% Trimmed Mean	46.65

N = 23; there are 23 NFL teams that play home games in outdoor stadiums. Teams with indoor stadiums were excluded. Means per team over the 2007 – 2014 seasons
Source: Sharp Football Analysis.

A case study: Fumbles and the New England Patriots

New England Patriots - avg of 73 touches per fumble.

Next closest competitor?

Chargers - avg of 54 touches per fumble.

In Standard Units?

$$Z = \left(\frac{x_i - \bar{x}}{s}\right) = \left(\frac{73.1 - 46.41}{7.45}\right) = 3.6$$

73.1 touches per fumble is 3.6 standard deviations above average.
According to the Empirical Rule, this should happen in fewer then 1 in 1,000 times over this period.

Mean	46.41
SD	7.45
N	23
5% Trimmed Mean	46.65

Source: Sharp Football Analysis.

A case study: Fumbles and the New England Patriots

What about the Patriot's Fumbles/Touch?

Using the exact same data, we can also look at it as Fumbles per Touch

People tend to prefer large numbers over dealing with fractions

- 73.1 touches per fumble
- 0.0137 fumbles per touch

73.1 is 3.6 in Standard Units, this is very large.

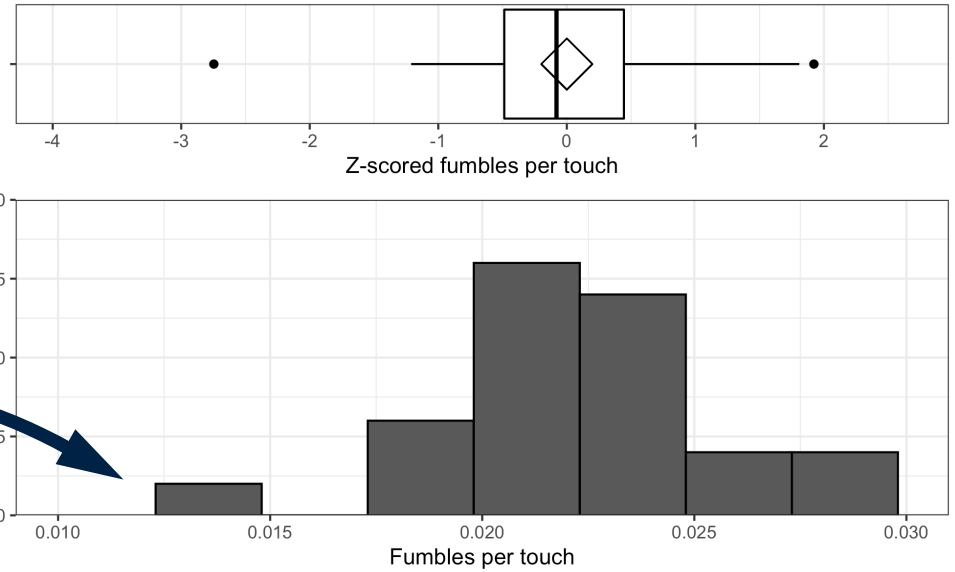
How big is 0.0137 in Standard Units?

A case study: Fumbles and the New England Patriots

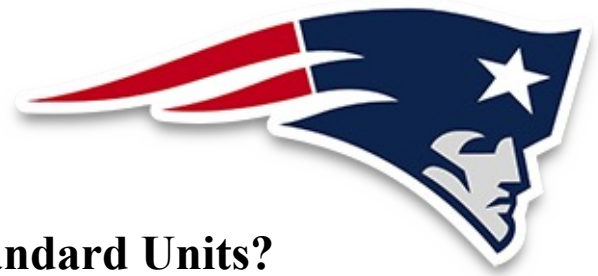
Patriot's Fumbles/Touch

Mean	0.02199
SD	0.00303
N	23

Histogram, Box Plot for Fumbles/Touch



.0137 fumbles/touch.



In Standard Units?

$$Z = \left(\frac{x_i - \bar{x}}{s}\right) = \left(\frac{.0137 - .02199}{0.00303}\right) = -2.7$$

0.0137 fumbles/touch is = -2.7 standard deviations below average.
It ends up being unusual, more like 1 in 100 (not 1 in 1,000).

Source: Sharp Football Analysis.

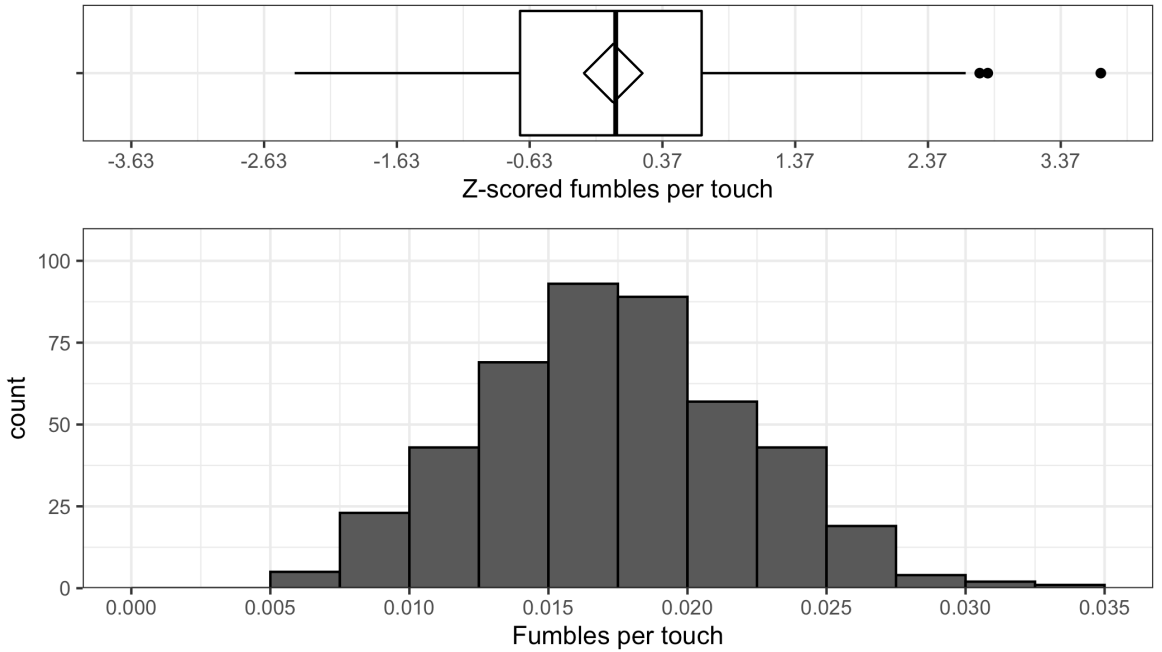
A case study: Fumbles and the New England Patriots

Here is the fumbles/touch data - it appears to follow the empirical rule quite well.

Mean = .01767

SD = .00523

N = 480

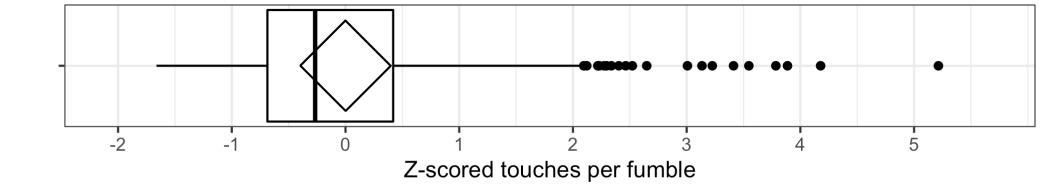


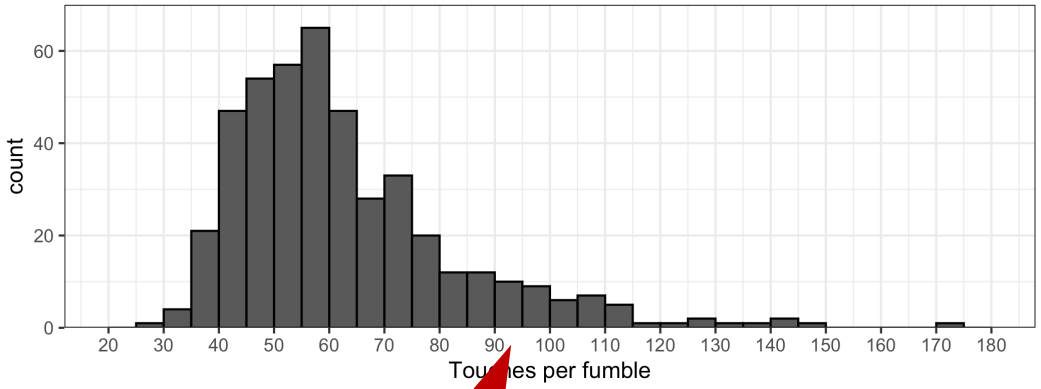
by team, by season

N=480, by team, by season from 2007 through 2021
Source: nflfastR

CAUTION: The Empirical Rule Doesn't Always Work- Especially in The Tails.

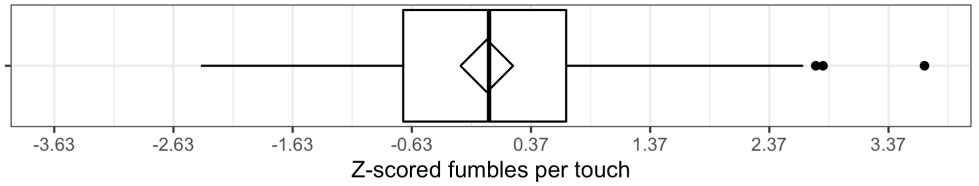
Touches per fumble

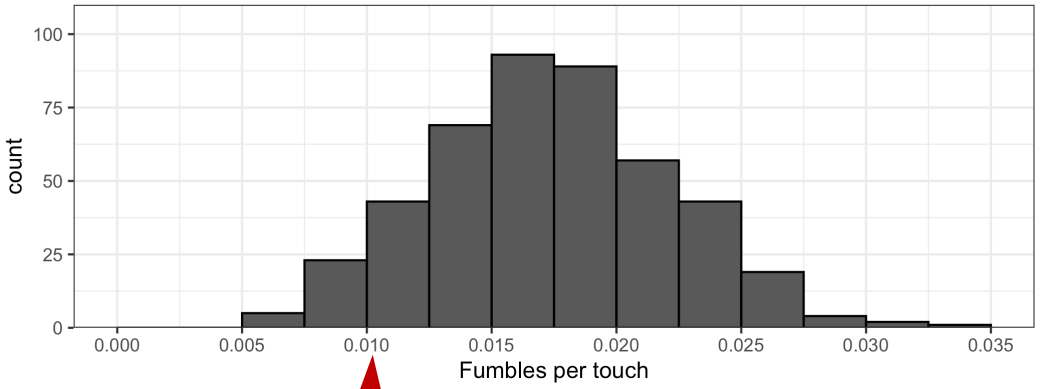




Mean = 60.08
SD = 20.64
N = 480
NE in 2014: touches/fumble = 93.58
Z = -2.27

Fumbles per touch





Mean = .01767
SD = .00523
N = 480
NE in 2014: fumbles/touch = 0.01068
Z = -1.69

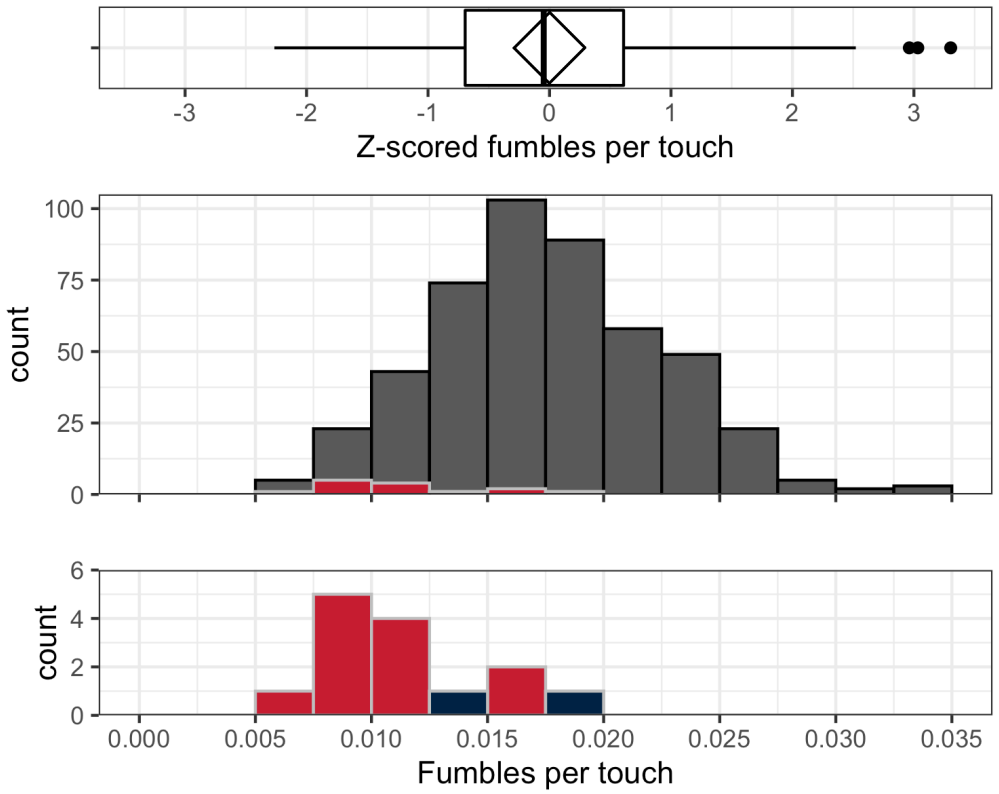
The team, the same data (NE in 2014): 93.58 touches/fumble = 1/(0.01068touches/fumble), but very different result in standard units.

Source: nflfastR

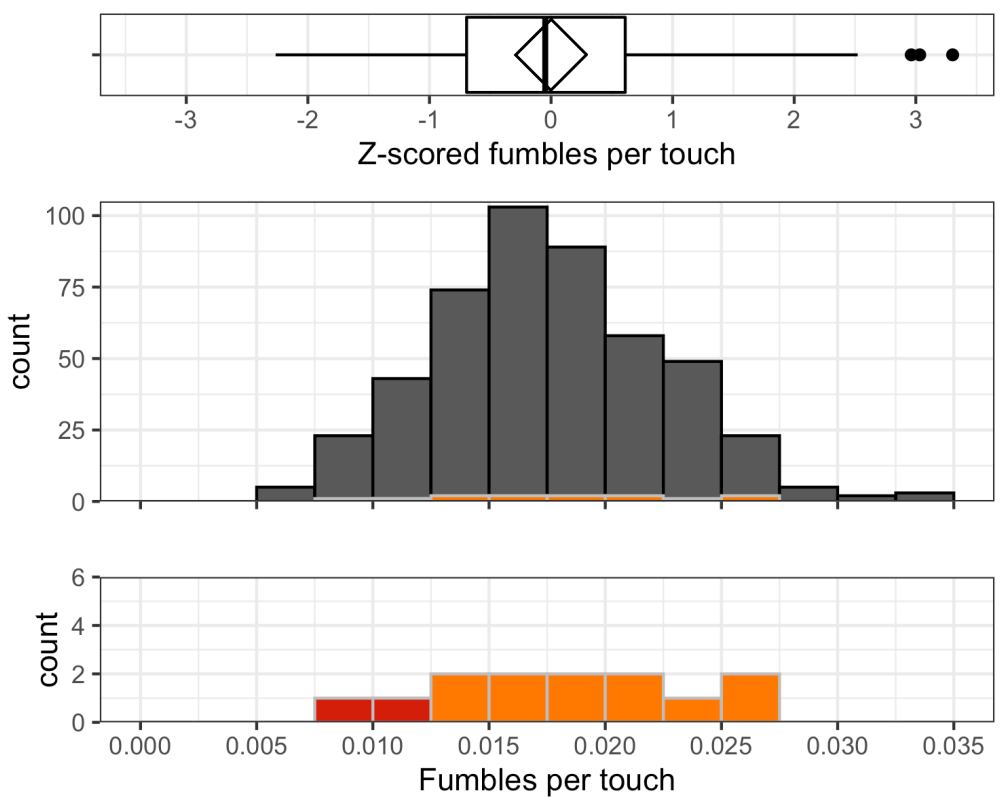
A case study: Fumbles and the New England Patriots

Other Considerations for "Deflategate": Why would the Patriots have a lower fumble rate that had nothing to do with the air pressure of the football?
In 2020 Tom Brady moved from New England (NE) to Tampa Bay (TB), creating a natural experiment:

NE: with Brady / post-Brady



TB: pre-Brady / with Brady



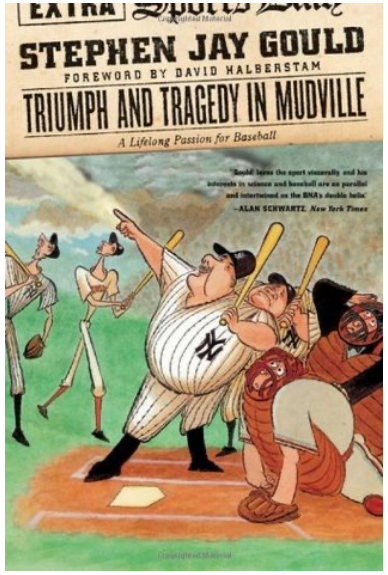
Source: nflfastR

Standard Units and the Normal Curve – Summary

1. Basic Summary Tools
2. Empirical Rule
3. Normally distributed data
4. Controlling and Confounding

Bonus:
Enjoy a conversation with the late Stephen Jay Gould on why no one will ever hit .400 again. Why No One Hits .400 Anymore

https://youtu.be/2ZBh2ZbEo0I?t=2m6s



Gould, S. J. (2010). Triumph and tragedy in Mudville: A lifelong passion for baseball. WW Norton & Company.
Gould, S. J. (1998). Full house: the spread of excellence from Plato to Darwin. Senior Managing Editor, 5(2), 68.

Standard Units and the Normal Curve