

Classifying GOATs (like Brady, Russell and Ruth) by Measuring Their Tails.

Rick Cleary, Babson College
Steve Miller, Williams College

rcleary@babson.edu, sjm1@williams.edu
MathFest: August 5, 2021



BIG PICTURE SPORTS QUESTIONS...

-Who is the GOAT (Greatest Of All Time) in a particular sport?

-Who is GOAT of GOATs across sports?

We recognize that this is not a well posed question ... but fans and media try to answer it. To do so they make mathematical and statistical arguments that lead them to particular metrics ... sometimes without even realizing it!



Examples abound...

- A student in the school paper at Nova Southeastern University makes the case for Tom Brady in football: [TomBradyGOATArgument](#)
- Justin Quinn in USA Today says it's Bill Russell in basketball: [BillRussellGOATArgument](#)
- And at Quora.com Mike Berard makes the case for Babe Ruth: [BabeRuthGOATArgument](#)

The Brady and Russell arguments are largely team oriented, while Ruth's case is more about his individual excellence.



OUR GOAL TODAY...

This is a *preliminary* report on our work to date. We want to:

- Show that metrics matter.

- Give examples of ways that assumptions lead to metrics.

- Choose explicit metrics first and use them to evaluate ‘something like’ a GOAT argument; maybe a ‘best teammate’.

Note: The ideas here can be applied in teaching many applications besides sports. Consistent with ideas in social choice, finance, economics and other fields. Can adjust the technical level to be anything from a first year seminar to a capstone project!












METRICS MATTER - Who's IN first?

How can we tell who the GOAT if we can't even decide GORN (Greatest of Right Now)!

Consider June 4, 2018. Boston Red Sox were 41-19, winning percentage .683. New York Yankees were 37-17, winning percentage .685.

ESPN *correctly* reports Red Sox in first (HURRAH!) since they are one "game ahead."

GOOGLE reports Yankees in first (BOOOO!) since they have higher winning percentage.

American League					American League				
EAST					AL East				
	W	L	PCT	GB		W	L	Pct	GB
 Boston Red Sox	41	19	.683	-	 Yankees	37	17	.685	1.0
 New York Yankees	37	17	.685	1	 Red Sox	41	19	.683	-
 Tampa Bay Rays	28	30	.483	12	 Rays	28	30	.483	12.0
 Toronto Blue Jays	26	33	.441	14.5	 Blue Jays	26	33	.441	14.5
 Baltimore Orioles	17	41	.293	23	 Orioles	17	41	.293	23.0

Or ... Cross country example

Order of finish:

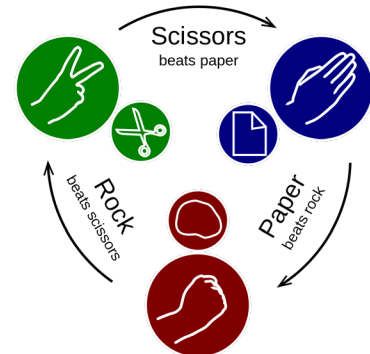
A - B - B - A - A - C - C - B - C - C - C - C - C - B - A - B - A - A - B - B

“Invitational” Scoring: A wins!!

A: $1+4+5+15+17 = 42$; C: $6+7+9+10+11 = 43$, B: $2+3+8+14+16 = 43$. (C gets second since their sixth runner beat B's sixth runner.)

“Dual Meet” Scoring: Three different races:

B beats A 27-28; A beats C 29-30; and C beats B 28-31! Everybody goes 1-1!



WHICH COMES FIRST? THE METRIC OR THE GOAT?

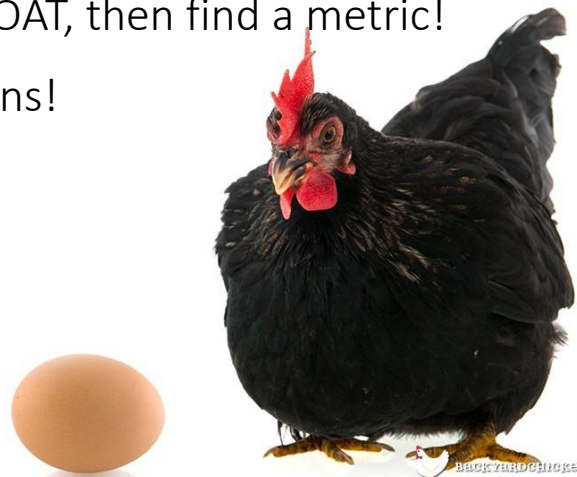
You don't simply choose the GOAT ... you choose a metric!

Scientific method ... Choose the metric and see who's the GOAT;

OR

As a fan/writer with deadline for a column... Choose the GOAT, then find a metric!

We will pick some reasonable metrics and see what happens!



Let's have a vote!

In the chat, tell us who you think is the best candidate for GOAT of GOATs:

BRADY ... NFL



RUSSELL ... NBA

RUTH ... MLB



OTHER ... Provide Name/Sport



Let's follow up with some data

In the chat, we had votes for the best candidate for GOAT of GOATs:

BRADY ... NFL

RUSSELL ... NBA

RUTH ... MLB

OTHER ... Provide Name/Sport

AN ALTERNATE PLACE WE CHOSE TO START: Who is the BOAT (Best Of All Teammates)? Measured by team success relative to league size/quality/playoff format. Preliminary research suggests Brady and Russell is the place to start looking.



The BOAT must ... Get to the playoffs!

Under simplest assumptions, with goal of 'make the playoffs'. Each team equally likely to qualify each year, years are independent, so a binomial model.

Brady: 18 times to playoff in 20 year career in 32 team league with 12 playoff qualifiers. Appearances \sim binomial (20, 12/32).

$P(18 \text{ or more in 20 years}) = 3.02 * 10^{-9} = .00000000302.$

Brady vs. Russell playoff appearances

Brady: 18 times to playoff in 20 year career in 32 team league with 12 playoff qualifiers. Appearances \sim binomial (20, 12/32).

$$P(18 \text{ or more in 20 years}) = 3.02 * 10^{-9} = .00000000302.$$

Russell: Thirteen for thirteen in making playoffs ... but in small league where more than half of teams made playoffs!

$$P(13 \text{ for 13 in making playoffs}) = 6.29 * 10^{-3} = .00629.$$

A point in Brady's favor here!



But if we do titles instead of appearances:

Under simplest assumptions: Each team equally likely to win each year, years are independent, so a binomial model:

Brady: Seven titles in 20 years in 32 team league.

$X \sim \text{binom}(20, 1/32)$

$P(X \text{ at least } 7) = 1.49 * 10^{-6} = .00000149.$

Russell: Eleven titles in 13 years in (approx.) 10 team league.

$X \sim \text{binom}(13, 1/10)$

$P(X \text{ at least } 11) = 6.44 * 10^{-10} = .000000000644.$

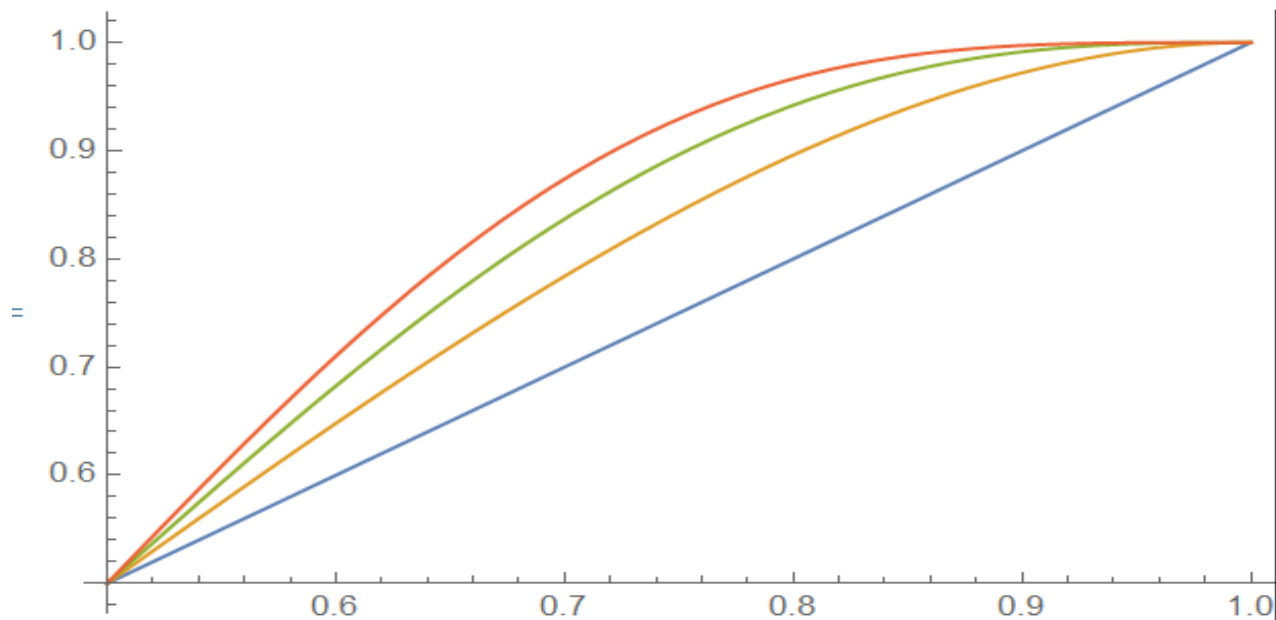
Russell looks better here ... again, metrics matter!!



BUT Brady played *GAMES*; Russell *SERIES*

Horizontal axis: $P(\text{stronger team wins any one game})$.

Vertical axis: $P(\text{stronger team wins series of 1,3,5,7 games})$ under independence.

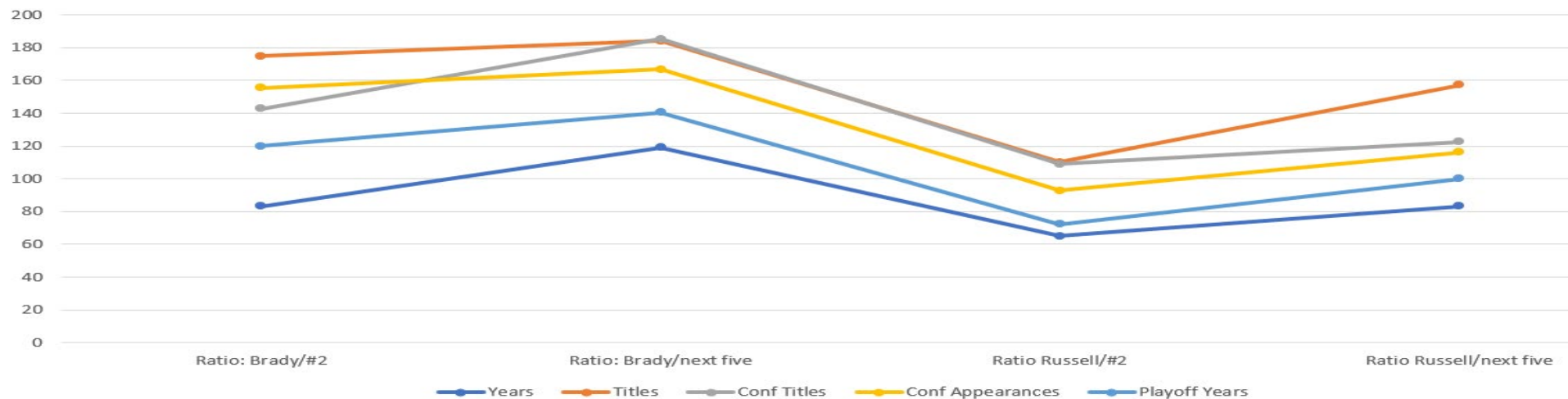


BUT Brady played *GAMES*; Russell *SERIES*

Interpretation: Series help better teams avoid upsets. Also Brady had to win 3-4 games while Russell usually played two series. Does that make up the difference? IT MIGHT!

	Years	Titles	Championship	Conferences	Playoffs
Brady/#2*	83	175	143	156	120
Brady/next five	119	184	185	167	141
Russell/#2*	65	110	109	93	72
Russell/next five	83	157	122	116	100

Ratios: Brady and Russell to Top Competitors



The log 5 rule (due to Bill James)

Suppose team A wins games with probability p , and team B wins with probability q .

A good estimate of $P(\text{A wins a game vs. B})$ is given by

$$\frac{p(1-q)}{p(1-q)+(1-p)q} = \frac{p-pq}{p+q-2pq}$$

Example: In a playoff round we might have $p = .8$ and $q = .6$

So $P(\text{A wins a game vs. B}) = (.8 - .48) / (1.4 - .96) \approx .727$.

(See SJM's paper at

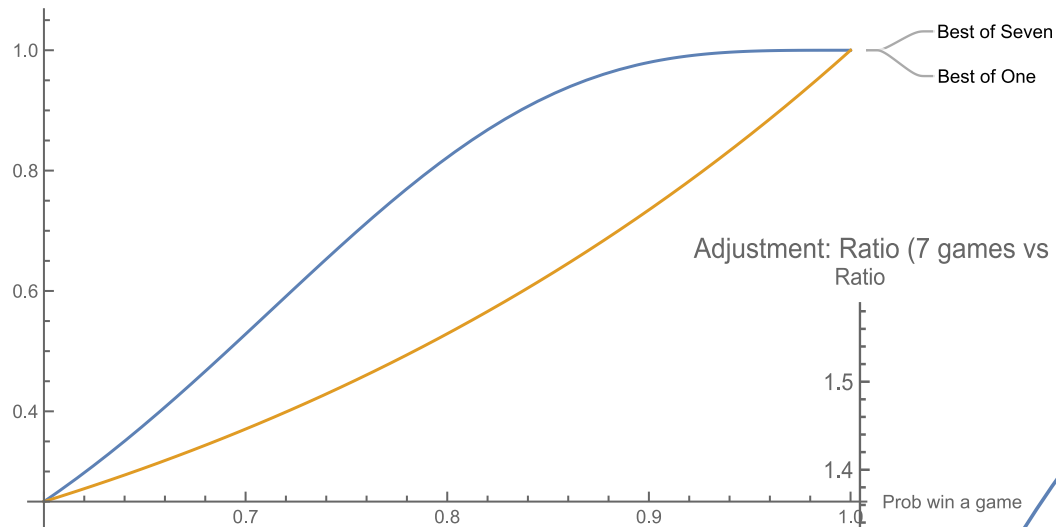
https://web.williams.edu/Mathematics/sjmiller/public_html/399/handouts/Log5WonLoss_Paper.pdf)

A Log-5 Adjustment for series vs. games

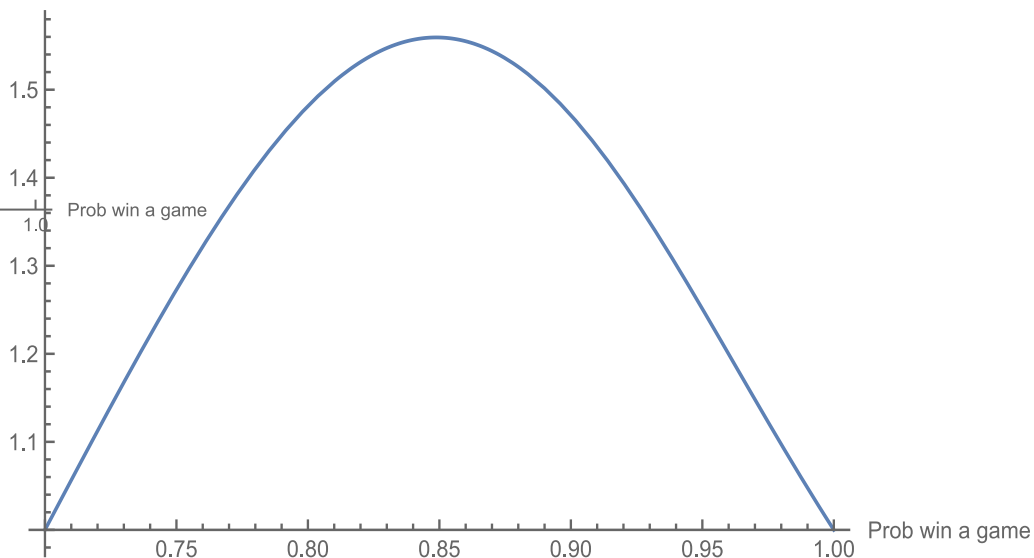
Team A, $P(\text{win})=p$, plays B with $q=0.6$. Maximum ratio is about 1.6. Two series.

James Log-5 Adjustment: Playing a team that wins 60%

Prob win two series



Adjustment: Ratio (7 games vs 1 game) expected titles (two series) playing a team that wins 70%

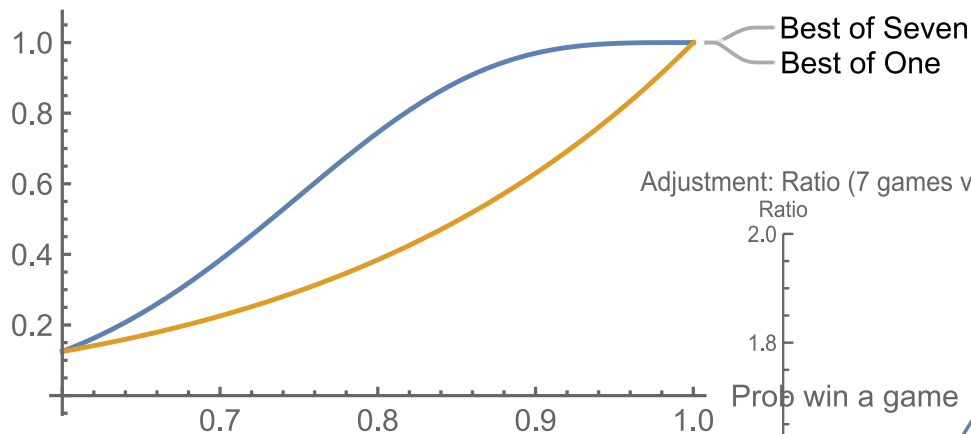


A Log-5 Adjustment for series vs. games

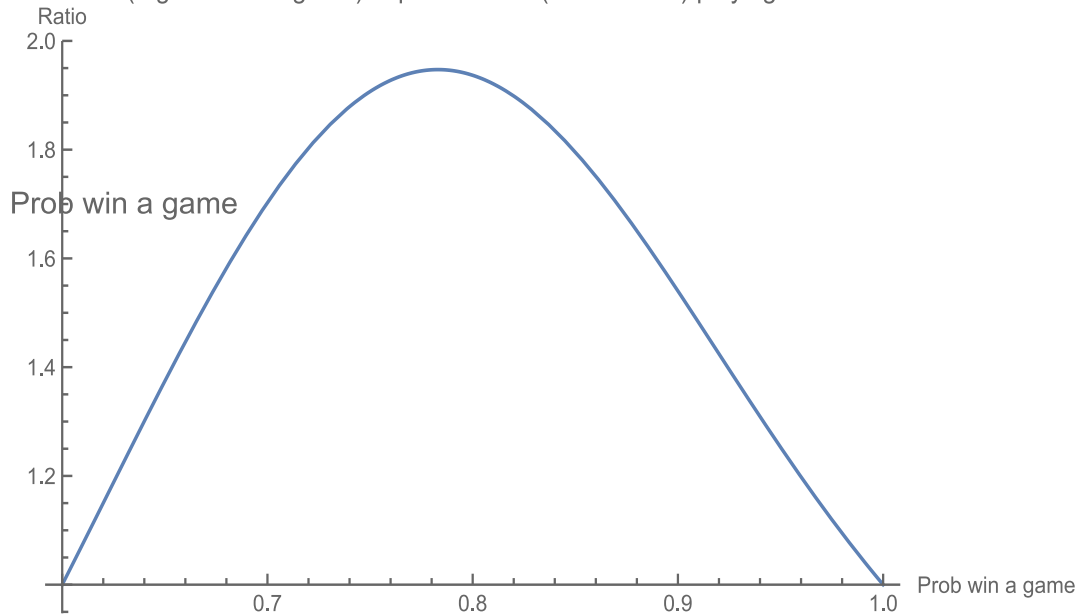
Team A, $P(\text{win})=p$, plays B with $q=0.6$. Maximum ratio is about 1.9. Three series.

James Log-5 Adjustment: Playing a team that wins 60%

Prob win three series



Adjustment: Ratio (7 games vs 1 game) expected titles (three series) playing a team that wins 60%



So here's an argument for Brady...

RUSSELL ... 11 titles.

BUT NFL titles might be about 1.8 times more difficult because of the games/series issue!

So here's an argument for Brady...

RUSSELL ... 11 titles.

BUT NFL titles might be about 1.8 times more difficult because of the games/series issue!

So BRADY's seven titles might be worth about:

$$7 * 1.8 = 12.6 \text{ RUSSELL titles!}$$

-Some next steps ...

We have begun considering (and will recruit students to help us consider):

- More sophisticated metrics with playoff wins models (Poisson vs. binomial);
- Championship round success (proponents of Michael “six for six” Jordan and Joe “four for four” Montana
- Similar argument for longevity vs. high peak. Is four title in eight year career more or less impressive than four in 20?
- Influence by sport (Russell one of five, on floor 80% of time; Brady one of eleven, of field 40% of time ... but Brady key in all of those plays while Russell might go some time without a touch.)
- We'll think of more!

One more vote ...

Based on our discussion, please use the chat to vote again for BOAT:

BRADY

RUSSELL

RUTH

OTHER

Did we change any minds?
(We did change the question!)

THANKS for your attention ...

-Any questions please contact us at rcleary@babson.edu or sjm1@williams.edu



Slides below here are for another day...

Players to consider... NFL

		Yrs	Titles	Fnals	Semis	Playoffs
Tom Brady	2002-21	20	7	10	14	18
Joe Montana	1980-94	15	4	4	8	11
Terry Bradshaw	1971-84	14	4	4	6	9
Jerry Rice	1986-05	20	4	4	8	14
Steve Young**	1986-00	15	3	7	7	11
Roger Staubach	1970-80	11	2	4	7	10
John Elway	1984-99	16	2	5	6	10
Adam Vinatieri	1997-20	24	4	5	6	15

Players to consider... NBA

	Seasons	Yrs	Titles	Final	Semis	Playoffs
Bill Russell	1958-70	13	11	12	13	13
Michael Jordan	Various	15	6	6	8	13
Sam Jones	1958-69	12	10	11	12	12
K Abdul-Jabbar	1971-90	20	6	10	14	18
LeBron James	2004-21	18	4	10	11	9
Tom Heinsohn	1957-65	9	8	9	9	9
Magic Johnson	Various	13	5	9	10	13

GOAT METRICS IN TEAM SPORTS ...

- Individual statistics?
 - Counting stats like Points, Touchdown Passes, Home Runs, Goals, Assists
 - Derived stats like Wins Above Replacement, Adjusted Plus/Minus
- Team statistics?
 - Titles won (consider time periods, league size)
 - Consistent excellence (high winning percentage over a long time.)

Measuring GOATS by their tails...

Principle: What a GOAT does is accomplish something that is *least likely* among a set of candidates using a particular metric. There are many variables to consider when computing these probabilities, but as an exercise...

Positives:

- Great modeling project that requires careful consideration of inputs and assumptions.
- Adaptable; can be used for both team and individual performance.

Measuring GOATS by their tails...

Principle: What a GOAT does is accomplish something that is *least likely* among a set of candidates. There are many variables to consider when computing these probabilities, but as an exercise...

Positives:

- Great modeling project that requires careful consideration of inputs and assumptions.
- Adaptable; can be used for both team and individual performance.

Negatives:

- Sort of like a p-value. Why right tail only?
- Almost certainly not definitive.

The “What Abouts...”

- What about teammates? Russell played 80% of the time and was one fifth of the team at any time. Brady played 40% of all plays and was 1/11 of his team at any time.

Another one we might not have time to explore in talk

The “WhatABOUTs...”

- What about influence? Russell played 80% of the time and was one fifth of the team at any time. Brady played 40% of all plays and was 1/11 of his team at any time.
- What about league parity? The NBA in Russell’s era had teams that were never truly competitive.

The “What Abouts...”

- What about influence? Russell played 80% of the time and was one fifth of the team at any time. Brady played 40% of all plays and was 1/11 of his team at any time.
- What about league parity? The NBA in Russell's era had teams that were never truly competitive.
- What about quality of teammates? (And many more...)

So stay tuned: We hope to attract students to build these more sophisticated interpretations!