

Greatness in Sports: Understanding and Visualizing its Patterns Article Draft, V2

Benjmamin Leinwand and Matthew Murray

Introduction

Background

DON'T TAKE GREATNESS FOR GRANTED.

It's a common phrase regurgitated by sports fanatics and enthusiasts alike. Intuitively, it makes sense. While the sports themselves are (hopefully) lasting, the athletes come and go, with their careers seeming like a brief blip in time when it is all said and done. Baseball fans listen to stories of the Great Bambino from their elders, wondering if they will ever see a player or a talent burst on to the scene again. Meanwhile, basketball fans seem to engage in an everlasting debate between Michael Jordan and Lebron James.

Nonetheless, with the seemingly endless abundance of talent on display today, coupled with the advances being made in athlete performance and recovery that have athletes performing at the highest levels ever, one has to call to question: *can* we take greatness for granted? Are there more greatest of all time (GOAT) candidates than one may intuitively believe there to be? Are there patterns in which greatness occurs in one sport, and a lack thereof in another? Do sports with certain qualities (e.g: contact vs. non-contact) tend to attract more GOAT candidates than another? In this paper, we seek to provide the answers to these questions and test whether the people's perceptions of greatness are mostly misnomers.

Disclaimer - What we are NOT Trying to Do

One point that we would like to *avidly* emphasize before getting into the details of the paper is that we are *NOT* looking to answer the question of who is the GOAT in any sport. Said question - while incredibly interesting and thought provoking - has been discussed ad nauseam at every bar, barber shop, and online forum imaginable. In fact, the ubiquity of the question is a testament to the fact that it can almost never be answered objectively. We are instead looking to examine the frequency, patterns, and cadence in which legitimate GOAT candidates appear in the most prominent and popular sports. In doing so, we had to inevitably - find GOAT candidates and sought to do so by using very reputable and defensible sources and metrics.

Additionally, we are aware and cognizant of the fact that if a reader were to conduct a similar analysis, the results and data may be very different. However, our goal in this project is not to conduct an analysis that is extremely robust/resistant to sensitivity tests, but rather, one that is informative, insightful, and interesting. If 25 different researchers were to conduct this analysis, each dataset and approach may be very different, and we are okay with that fact. Like many statistical analyses, answering our questions of interest is an art, not a science.

Data

Due to the unique nature of this project, we had to impute our own data. While time consuming, doing so allows us to be fully transparent and elucidate on all the assumptions that we inevitably had to make while creating the datasets.

We decided to compile lists of men’s GOAT candidates from the following sports:

“Subjective” Sports

1. Baseball
2. Basketball
3. Cricket
4. Football
5. Golf
6. Hockey (Ice)
7. Soccer
8. Tennis

“Objective” Sports

9. Chess
10. Running
11. Swimming

The rationale for including these sports is as follows:

- Baseball, Basketball, Football, Hockey (Ice) - They are the 4 of the 5 most popular sports in the United States (Gallup 2017).
- Cricket, Soccer - Included due to their immense popularity and following worldwide (Media 2017).
- Golf, Tennis - Included because professionals in both sports compete in four “major” tournaments each year.
- Chess, Running, Swimming - We decided to include these more “objective” sports that have very well-accepted, objective metrics that are more-than-often referred to when discussing the best player in that sport. The inclusion of these sports serves as litmus test from the results of the “subjective” sports.

The data spans from 1945 to the present. The choice of 1945 can be attributed to the fact that at this time, the MLB (1876) and NFL (1920) were already founded, while the Basketball Association of America (which eventually became the NBA) was founded shortly thereafter in 1946. Therefore, one can argue that the three most popular sports in North America (in addition to soccer) were already up and running by that time period. Additionally, 1945 marked the end of World War II and the advent of a period of economic flourishing in the Western world.

Out of our valid concern of bogging down - and slowly lulling to sleep - most readers, we have not included the full list of statistical liberties - including how we determined a player’s active years, the different criteria we used for including players in our dataset, etc. - that we took while creating the dataset in this article. These can instead be found at our [GitHub repository](#). Please refer to the file “Data Collection Assumptions.pdf.”

Notable Findings

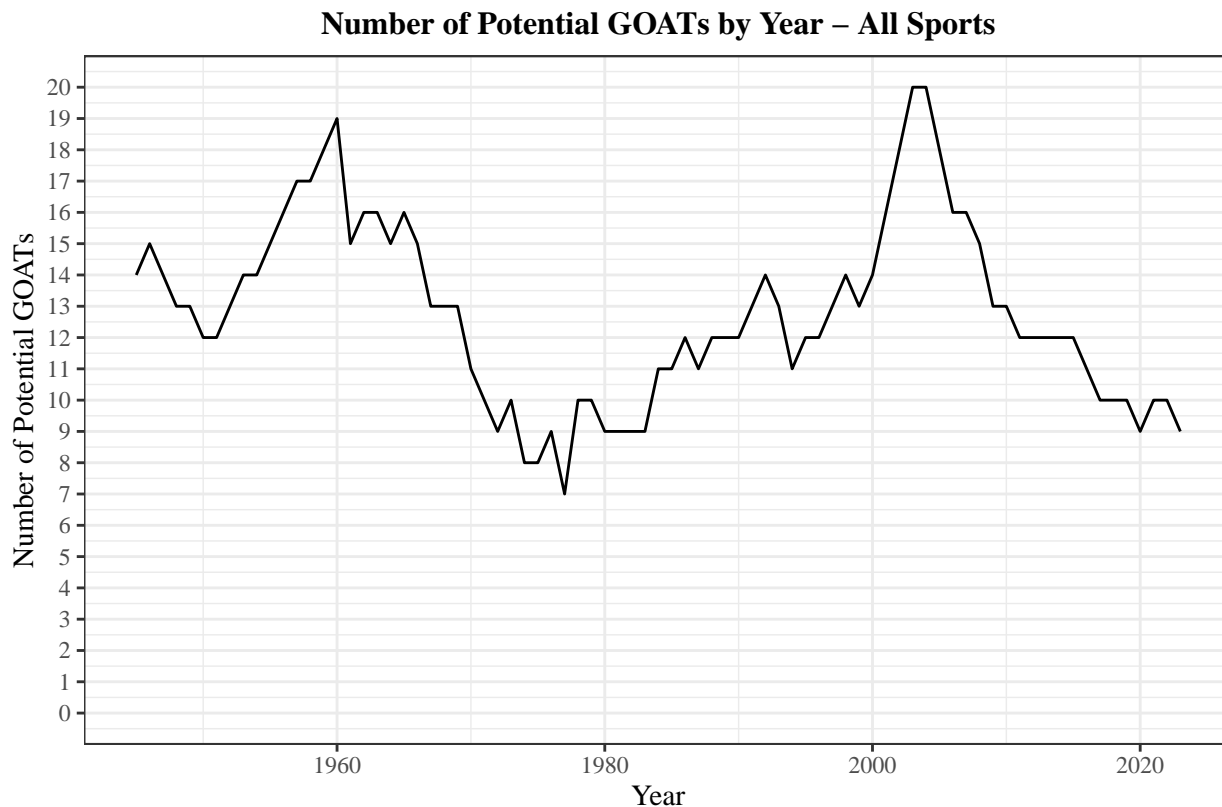


Figure 1

Subjective Sports

Main Finding 1

There are 3 main cadences at which GOAT candidates appear. They - as well as the sports that fall into each grouping - are as follows:

1. GOAT candidates are essentially one-offs, or statistical anomalies - Baseball, Cricket
2. GOAT candidates occur sporadically without any legitimate pattern, which aligns with our prior beliefs - Hockey, Golf
3. GOAT candidates either occur more frequently OR in larger quantities than we previously expected (corroborating our hypothesis that greatness can be taken for granted) - Basketball, Football, Soccer, Tennis

Intuition vs. Results

Main Finding 2

In general, sports with longer career lengths tend to less candidates because there is a higher mountain. The opposite is true for sports with shorter career lengths, which lower the barrier to entry for greatness.

##	Sport	N
## 1	Baseball	9
## 2	Basketball	7
## 3	Chess	4
## 4	Cricket	3
## 5	Football	10
## 6	Golf	7
## 7	Hockey	6
## 8	Soccer	8
## 9	Swimming	16
## 10	Tennis	11

Intuition vs. Results

Toy Model

Objective Sports - Can Our Findings be Validated?

Potential Explanations that Lack Credence

Conclusion

References

Gallup. 2017. “Sports.” <https://news.gallup.com/poll/4735/sports.aspx>; Gallup.
Media, News Members, News Freeview. 2017. “The World’s Most Watched Sports.” <https://sportforbusiness.com/the-worlds-most-watched-sports/>; Sport for Business.