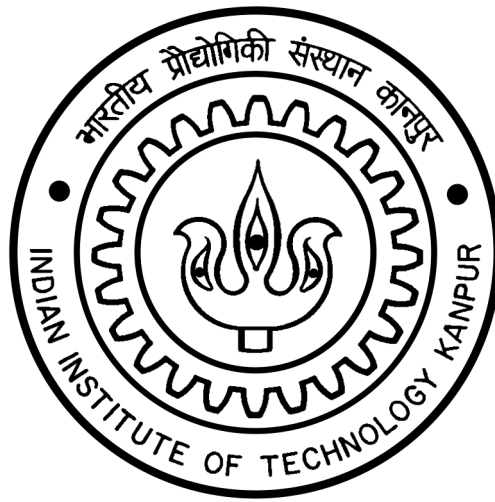# Ranking the Greatest Test Cricket Batsmen Across Generations

Prepared by:
SHANTAPRASAD KAMAT (231080081)
INDRAJ PRAJAPAT (231080044)
DIVYA (210353)
SUNITA(221106)

Supervised by:
DOOTIKA VATS

19, January, 2025

Contents

1   Acknowledgment

2   Abstract

This project explores a statistical methodology for ranking the greatest Test cricket batsmen across generations. We aim to develop a statistical methodology to rank the greatest batsmen across generations. This is not as straightforward as organizing players by batting averages due to drastic changes in playing conditions, rules, ground sizes, player physiology, and general skills over time.

3   Idea

To compare players across generations, we normalize a batsman's performance against the average performance of their peers from the same generation. The method involves normalizing players' performances against the average performance of their peers from their respective eras, using a metric called **Runs Over Others (ROO)**. The study aims to provide an objective assessment by analyzing data spanning decades of Test cricket, factoring in generational differences, changes in playing conditions, and evolution in gameplay strategies.

4   Data

The dataset is scraped from *ESPNCricinfo*, containing the following information for each player:

- **Name**
- **Country**
- **Player code** (unique identifier)
- **Batting average** of the player
- **Total Runs scored by the top seven batters** (excluding the player in question) in all the matches he played
- **Total Dismissals of the top seven batters** (excluding the player in question) in all the matches he played
- **Generation average batting statistics** for other players during the same period

The final dataset includes calculated values for runs scored by a player, total dismissals, the average batting performance of other players in their era, and the resultant ROO metric.

5   Methodology for Ranking Players

Calculation of ROO Statistic

For each player, the ROO is calculated as follows:

1. For every match the player participated in, compute the total runs scored and total dismissals of the top seven batters excluding the player. Repeat this for all matches in the player's career.

2. Calculate the sum of all top-seven runs and dismissals:

$$\text{total\_top7\_runs} = \sum \text{top7\_runs},$$
$$\text{total\_top7\_dismissals} = \sum \text{top7\_dismissals}.$$

3. Compute the average batting performance of other players in the player's era:

$$\text{total\_top7\_average} = \frac{\text{total\_top7\_runs}}{\text{total\_top7\_dismissals}}.$$

4. Normalize the player's performance by taking the difference between their career batting average and total_top7_average:

$$\text{avg\_diff} = \text{player\_average} - \text{total\_top7\_average}.$$

5. Multiply the avg_diff by the total number of times the player has been dismissed to compute ROO:

$$\text{ROO} = \text{avg\_diff} \times \text{player\_dismissals}.$$

Rank the players in descending order of ROO, with a higher ROO indicating a greater relative impact.

5.1 Example: Sachin Tendulkar

Consider the match India vs. Australia, 2nd Test, 2001:

- In the 1st and 3rd innings, the top 7 Australian batsmen scored 335 and 162 runs, respectively.

- In the 2nd and 4th innings, the top 7 Indian batsmen (excluding Sachin) scored 129 and 582 runs, respectively.

Total runs by top 7 batsmen: $335 + 162 + 129 + 582 = 1208$ runs.
Total dismissals of top 7 batsmen: 26.

We calculate these for all the 200 matches that Sachin played and then, compute **total_top7_average**:

$$\text{total\_top7\_average} = \text{total\_top7\_average} = \frac{\sum \text{top7\_runs}}{\sum \text{top7\_dismissals}} = \frac{164278}{4330} \approx 37.94$$

Sachin's career batting average is 53.79, and he was dismissed 296 times. Calculate ROO:

$$\text{Normalized Average} = 53.78 - 37.94 = 15.84,$$
$$\text{Runs Over Others (ROO)} = 15.84 \times 296 \approx 4689.$$

Which means that Sachin made 4689 runs more than the players of his era.

6 Rankings

Rank	Name	Country	Inns	Batting	others	Norm_avg	ROO Score
1	Jacques Kallis	South Africa	278	55.25	34.28	20.97	5011.23
2	Sachin Tendulkar	India	329	53.78	37.94	15.84	4688.79
3	Joe Root	England	278	50.87	32.86	18.01	4591.57
4	Donald Bradman	Australia	80	99.94	34.56	56.38	4576.55
5	Brian Lara	West Indies	230	53.17	32.92	20.25	4536.23
6	Kumar Sangakkara	Sri Lanka	233	57.40	37.37	20.03	4326.67
7	Steven Smith	Australia	204	55.86	32.78	23.08	4131.36
8	Ricky Ponting	Australia	287	51.85	36.63	15.22	3925.65
9	Rahul Dravid	India	284	52.63	37.98	14.65	3691.31
10	Shivnarine Chanderpaul	West Indies	280	51.37	35.63	15.74	3635.75

Table 1: Top 10 Cricket Players by ROO Score

7 Why This Method Works

- Unlike traditional statistics like average, this method adjusts a player's performance relative to his peers, accounting for differences in playing conditions and eras.

- By incorporating the performance of other batters, we create a benchmark representing the generation's overall batting standards.

- Only top 7 batsmen are chosen as generally the top 7 players that come out to bat are specialist batsmen, and others are all-rounders or bowlers. The analysis would make less sense if bowlers are included as well.

- As we are multiplying by the no. of times the player got out in the end, the longevity of a batsman is considered. And hence, players like Sachin are preferred over Bradman.

- The ROO metric is based on clear, mathematically defined steps, ensuring the ranking is unbiased and replicable.

8 Reproducibility to Compare other Players

- Bowling average can be used similarly to compare bowlers across generations.

- As strike rate is not that important in test cricket, it is not being considered for this problem. But for other formats like ODI and T20I cricket, strike rate becomes much more significant. Hence, we can use similar methodology to compute **Strike Rate Over Others (SROO)** and then take a linear combination of ROO and SROO.

- Similar method can be used to compare players across generations in other sports as well. For e.g., In chess, FIDE Rating of other players in a particular chess player's era can be used.

9 References

- https://www.cricketcalculator.info/cricket-stats-guides.html

- https://www.cricketanalytics.com/player-performance

- https://www.cricketarchive.com/Archive/Articles/1/1326.html

- https://www.sciencedirect.com/science/article/pii/S0313592609500380