CRICKET ANALYTICS - LEVERAGING DATA FOR SPORTS DECISION-MAKING

ROBERT MCCOY, MBA

4WI2024 FOUNDATIONS OF DATA ANALYTICS

DR. HUNT

MARCH 3, 2025

Introduction

This presentation demonstrates a framework to analyze player and team performance in One-Day International (ODI) cricket using statistical calculations and visualizations.

NOTE: This Analysis Processed over 1/5 Million Lines of Data

Image by Open AI, 2025

ODI CRICKET MATCH DATA (MALE) DATASET (TOMAR, 2023)

Image by Open AI, 2025

METHODOLOGY

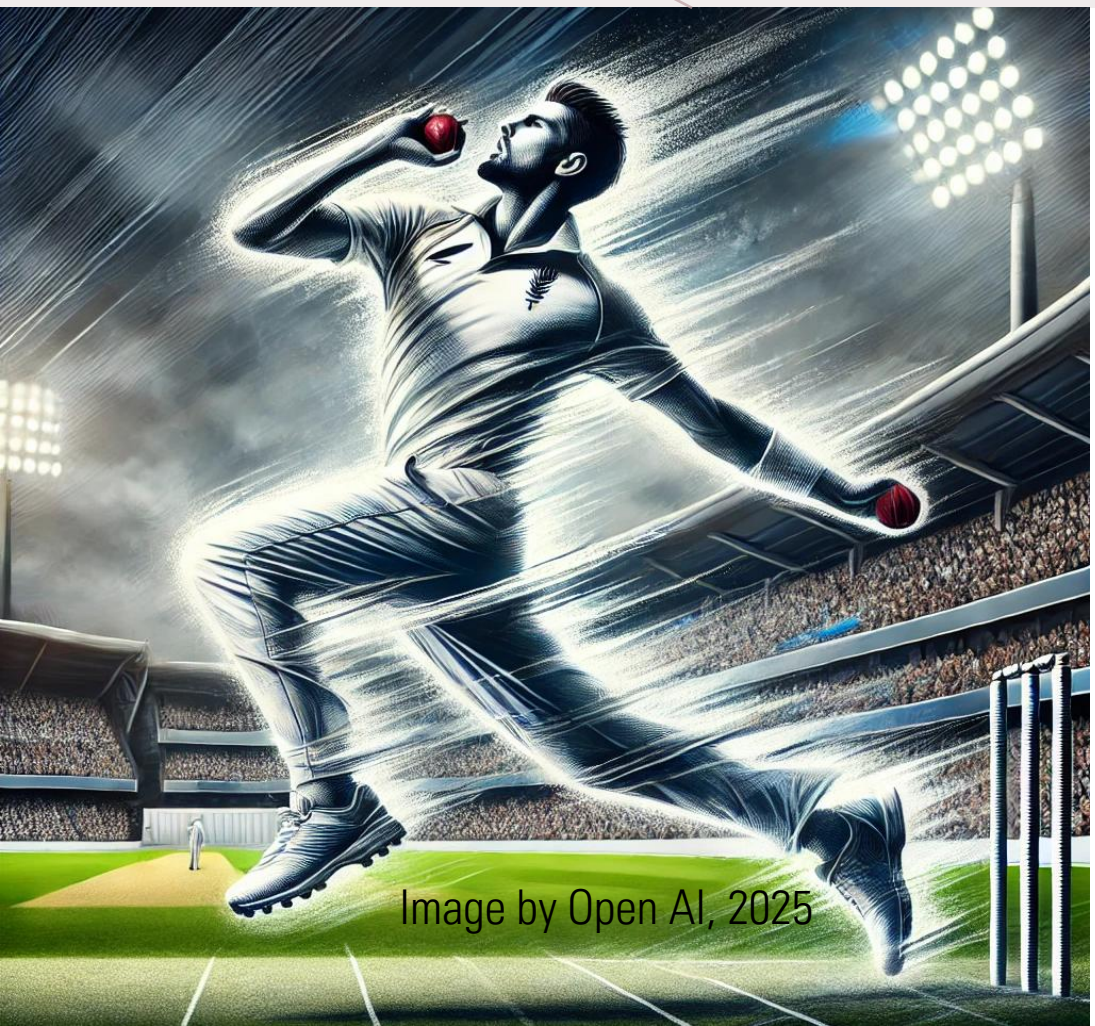

Image by Open AI, 2025

Data Source Kaggle Open Data - **ODI Men's Cricket Match Data (2002-2023) (Tomar, 2023)**

Analysis Techniques - Data analysis was conducted using a combination of Excel formulas, Pivot Tables, and statistical functions to extract insights from the ODI cricket dataset. Key performance metrics such as total runs, batting averages, strike rates, and economy rates were calculated using SUM(), AVERAGE(), and IF() functions. Pivot Tables were employed to summarize team and player statistics, allowing for trend identification across matches and seasons. Conditional formatting was used to highlight top-performing players, while filtering and sorting functions helped categorize match outcomes. Additionally, ChatGPT was utilized for cricket-specific data interpretation, terminology validation, and ensuring statistical accuracy in performance comparisons

METHODOLOGY

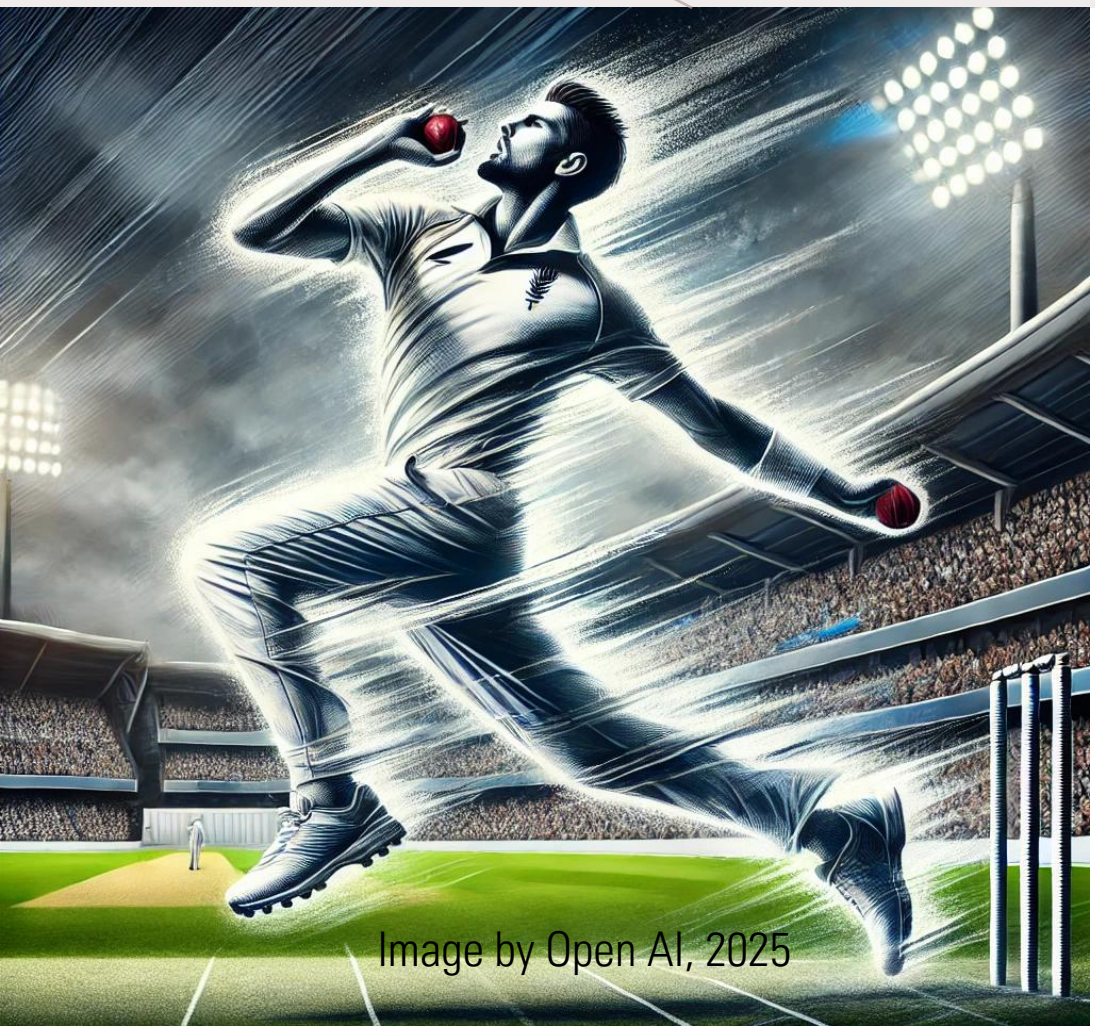

Image by Open AI, 2025

Data Cleaning & Preprocessing – A primary function of this assignment was the cleaning of the dataset. Data inconsistencies such as missing city and venue information were manually corrected. Advanced Excel tools such as Excel's "Power Query", were used to standardize cell structure between data sets and to prepare the files to be merged.

A filtering operation was applied to remove all pre-2014 matches from the DATA sheet. This aligned the data sets. This was done using Power Query. Additionally, column headers were standardized across both sheets, and necessary transformations were applied to ensure accurate mapping between match details and statistical data. ChatGPT was leveraged for cricket-specific knowledge, especially during the data cleaning of cities and venues which was extensive including player roles, match formats, and performance metrics.

A Special note on "winners" there were many null cells in this area and my research determined that this is not unusual for the game of Cricket do to: Draw (test cricket), Tie (limited Overs), abandoned matches (more common in rain-affected areas), no results (limited overs), and stalemate situations (Tcricketworldcupofficial, 2024) . The decision was made to mark these null cells "N/A". A full understanding of the governing law (law 16) that defines this topic and the game of cricket world-wide can be found at the Melborne Cricket Club (Lord's, 2018).

KEY METRICS

Image by Open AI, 2025

1. Identify patterns you see in player performance, match outcomes, and team composition.

2. Which players have the best performance metrics?

3. What strategies have led to successful matches?

4. How does team composition affect match outcomes?

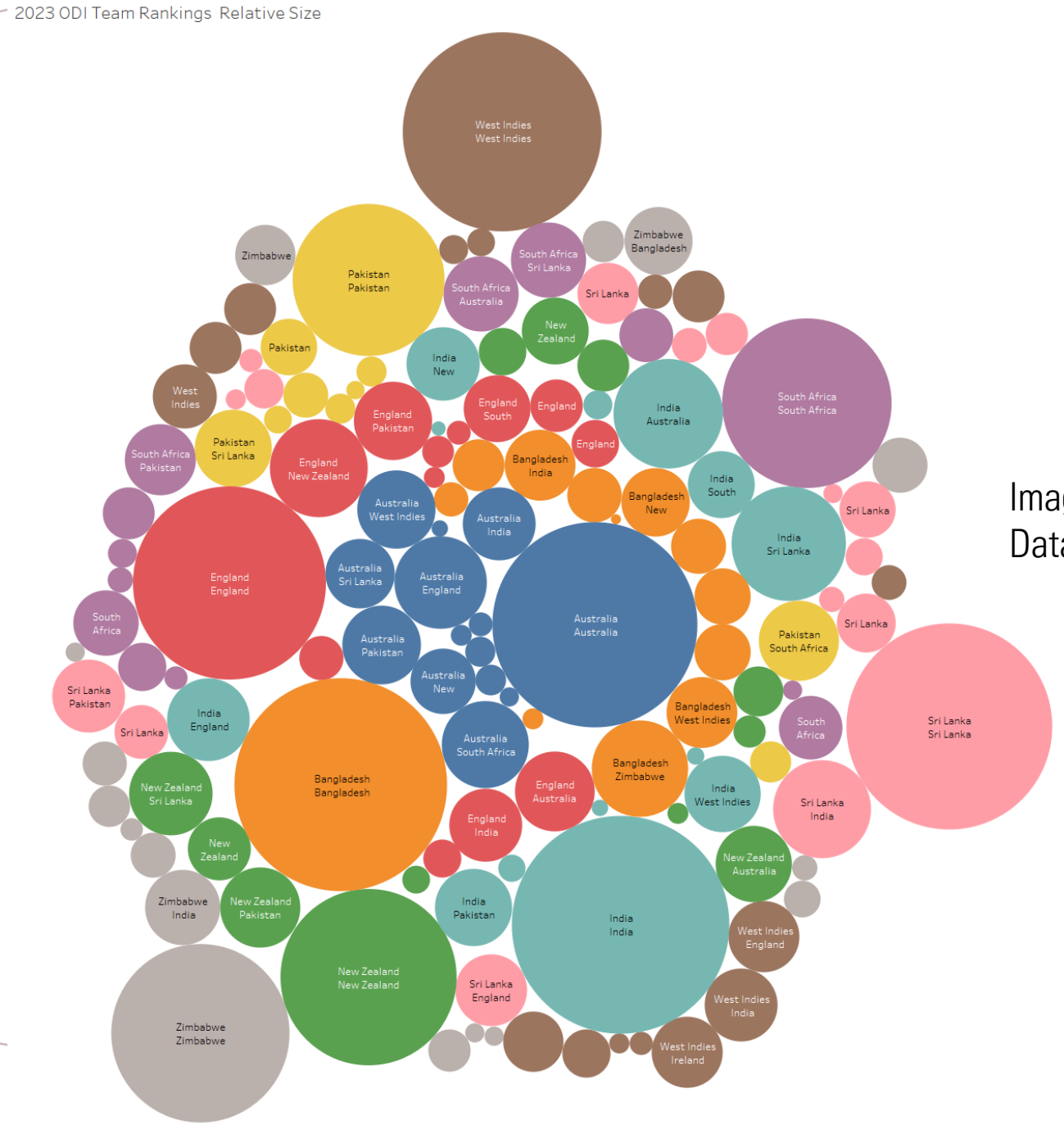

Image created in Tableau Utilizing the same Dataset created in Excel (Tomar, 2023)

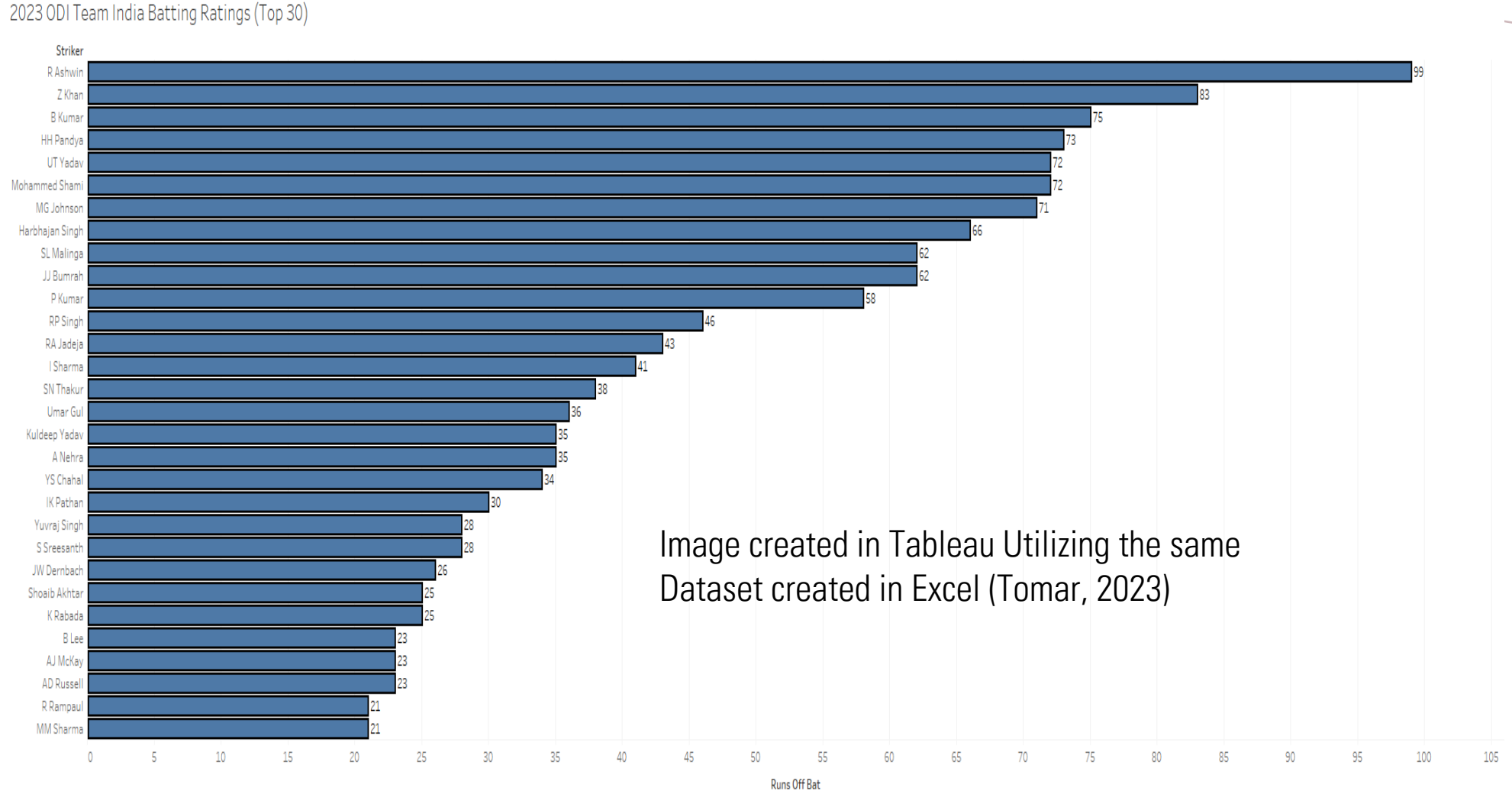

Image created in Tableau Utilizing the same Dataset created in Excel (Tomar, 2023)

India's Power Batting Lineup as Strategy

1. The Majority of Batters Score 0 or 1 Runs Per Game (81.88%)

52.69% of strikers **never score a run** ("ducks").

29.19% of strikers **score only one run** before getting out or ending their innings.

Inference: This suggests that **a large proportion of batters either fail to score or get out quickly**, which could include **lower-order batters, tail-enders, or players facing difficult bowling conditions**.

Possible explanations:

- ✔ Strong **bowling dominance** in certain conditions.
- ✔ High number of **specialist bowlers** batting at the end.
- ✔ Early dismissals due to **aggressive playing styles (T20 impact)**.

Runs per Game		Percentage
7	2	0.00%
6	14979	1.37%
5	156	0.01%
4	76757	7.00%
3	53017	4.84%
2	53700	4.90%
1	319900	29.19%
0	577411	52.69%
Total Points	1095922	100.00%

2. Only ~20% of Batters Score Multiple Runs Per Game

- **4+ runs per game** occur in only **13.86% of cases**.
- Players scoring **6+ runs per game** are **extremely rare (1.37%)**.
- **Inference:** Consistently hi**only top-order or world-class players regularly contribute big scores**.gh-scoring batters are a **minority**, which reinforces the idea that

3. Teams Like India Likely Have More High-Scoring Players

✔️ **Elite batters (e.g., Virat Kohli, Rohit Sharma, Shubman Gill) consistently contribute high scores.**

✔️ India's **depth in batting** ensures multiple players perform consistently.

✔️ Teams with a **strong domestic system (Ranji Trophy, IPL)** develop high-quality batters.

4. Comparing Elite Teams vs. Weaker Teams

- In strong teams like **India, Australia, and England**, the **percentage of high-scoring batters (>4 runs per game)** is **higher**.
- **Weaker teams (lower-ranked ICC nations, associate teams)** have a much higher **0 or 1-run percentage**, reflecting **less batting depth**.

References

Lord's. (2018, October 17). *The result Law*. Lords.org; Lord’s. https://www.lords.org/mcc/the-laws-of-cricket/the-result

Morgan, C. (2006, October 10). *Understanding Cricket Statistics | Cheryl’s Mewsings*. Www.cheryl-Morgan.com. https://www.cheryl-morgan.com/writing/sport/cricket-for-baseball-fans/understanding-cricket-statistics/

OpenAI. (2025). AI-generated illustration of a cricket batsman. OpenAI DALL·E.

OpenAI. (2025). AI-generated illustration of a cricket bowler. OpenAI DALL·E.

OpenAI. (2025). AI-generated illustration of a cricket stadium. OpenAI DALL·E.

Tcricketworldcupofficial. (2024, February 10). *What are the Possible Results of a Cricket Match? - Tcricketworldcupofficial - Medium*. Medium. https://medium.com/@t20cricketworldcupofficial/what-are-the-possible-results-of-a-cricket-match-d7bd010c5c04

Utkarsh Tomar. (2023). ODI Men's Cricket Match Data (2002-2023) [Data set]. Kaggle. https://doi.org/10.34740/KAGGLE/DS/3780212