

All T20 Internationals Dataset (2005 - 2023)

Avinash Bunga

Information Systems and Business Analytics, Park University

CIS622DLAF2P2023 Data Architecture for Business Analytics

Professor: Gulnoza Khakimova

Oct 25, 2023

All T20 Internationals Dataset (2005 - 2023)

With the increasing prominence of T20 International cricket, a deep comprehension of its multifaceted components has become imperative. This study undertakes an exhaustive analysis of comprehensive datasets from 2005 to 2023. These datasets encapsulate various dimensions of T20 International cricket, encompassing match details, batting, and bowling performances. (PRASAD,2023)

Source of the Dataset:

The dataset was sourced from the Kaggle website, specifically a contribution by Bhuvanesh Prasad. The dataset can be accessed at: [All T20 Internationals Dataset \(2005 - 2023\)](#). (PRASAD,2023)

Data Files and Their Description:

- **t20i_Matches_Data.csv:** This file contains general match details and consists of 2,325 rows of data. Attributes encapsulated include Match ID, Series ID, names of the competing teams, match date, team captains, runs scored, wickets taken, and other pivotal metrics. A snapshot of this file provides a comprehensible layout with individual columns representing each attribute.
- **t20i_Batting_Card.json:** Presented in JSON format, this file provides detailed batting performances for each T20I match and comprises 45,479 rows of data. Key attributes delineate the batsman's name, batting style, runs scored, and more. The structured format of JSON ensures that data hierarchies are seamlessly maintained, thus enhancing data comprehension.
- **t20i_Bowling_Card.txt:** This TXT file offers insights into individual bowling performances in T20 international matches, housing 24,922 rows of data. Vital data points embrace the bowler's name, number of overs bowled, wickets taken, and

economy rate, among other relevant metrics. The tabulated structure guarantees clarity and readability. (PRASAD,2023)

Common Attribute:

A consistent "**Match ID**" is a unique identifier for each match across all datasets. This ensures uniformity and enables easy linking and comparative analysis of datasets. (PRASAD,2023)

Conclusion:

With the datasets combined, a panoramic view of T20 International cricket spanning almost two decades emerges. Catering to various analytical requirements, these datasets enable insights into general match outcomes, individual batting prowess, or intricate bowling performances. Future investigations can harness these datasets to discern patterns, forecast outcomes, or even innovate advanced metrics that astutely evaluate player and team performances.

Reference

PRASAD, B. (2023, October 23). *All T20 Internationals Dataset (2005 - 2023)*. Kaggle.

https://www.kaggle.com/datasets/bhuvaneshprasad/all-t20-internationals-dataset-2005-to-2023?select=t20i_Matches_Data.csv