ICC Men's World Cup 2023 Team Performance Analysis

About Dataset:

team: 2 or 3 letter code matching that used in player name description.

player: Standard name used in statsguru innings database. Note: this may not match standard name used in statsguru for teamlists.

bat_or_bowl: indicates whether row represents a batting or bowling innings

bb_bf: Ball Bowled or Balls Faced. Provides a consistent and clean statistic. For bowlers, this is a cleaned version of number of overs bowled. So 1.5 overs becomes 11 in the bb_bf column.

runs: Either runs score by batsman or conceded by bowler. If a batsman produces a not out innings, the data is cleaned to only show the score (eg 100, rather than 100*). Reference can be made to the not_out column to determine if the batsman was out or not.

wkts: Number of wickets taken by a bowler in this innings.

wicketball_prob: Number of wickets taken (or lost by batsman) divided by number of balls bowled or faced. Will be zero for a not out batsman. Can be used to represent the probability of taking or losing a wicket in any given delivery. [This stat developed by JDau].

runs_per_ball: Number of runs score or conceded divided by the number of balls bowled or faced. Represents the average runs score or conceded per ball in this innings. [This stat developed by JDau].

opposition: The team this innings was played against.

ground: Which ground in India was the game played

start_date: Which date the game was played

overs: Represents the number of overs delivered by each bowler. This is raw data. Compare to the cleaned bb_bf column. A full over is usually 6 deliveries. A partially completed over will be shown as a decimal point where 0.1 represents 1 delivery. So 1.5 overs means the bowler 1 complete over and 5 additional balls for 11 deliveries.

mdns: Number of maidens a bowler bowled. A maiden is an over of 6 balls that does not concede any runs.

econ: The average number of runs conceded by the bowler per over in this innings. Compare to runs_per_ball column.

inns: 1 means this was the first innings of the day. 2 represents the 2nd innings of the day. So a row that includes the value 'bowl' in the bat_or_bowl column and 1 in the inns column indicates the innings in this row occurred when that team bowled first and batted second.

4s: How many 4's did the batsman score.

6s: How many 6's did the batsman score.

sr: The batsman strike rate. This has been converted to runs_per_ball by diving the sr by 100.

not_out: Whether the batsman's innings was a not out or not. This column removes the need for a *
beside the batsman's score.

mins: Duration of a batsman's inning in minutes.

Acknowledgement:

I would like to express my gratitude to the original owners of this dataset, as it is not owned

by me. The dataset used in this data analysis project is attributed to its rightful owners. All

credit and recognition go to them for providing such valuable data.