

# Batting Through Numbers: A Data-Driven Analysis of Cricket

## GROUP-25

Fayzan Ali Akhtar 25100157

Mustafa Ashfaq 25100216

Qayyum Ahmed 26100289

Muhammad Zaeem Rizwan 25100261

# Data Creation

## 1. Data Prep:

- Loaded, cleaned, and simplified batting/bowling datasets.
- Handled null values and dropped unnecessary columns.

## 2. Data Transformation:

- Developed specific functions for dataset generation.
- Imputed numeric nulls, mapped categorical data, and standardized columns.

## 3. Summarized Data:

- Condensed ball-by-ball details into summarized player stats per match.
- Key metrics include runs, balls faced, batting position, etc.

## 4. Result:

- Created a concise dataset focusing on player performance in matches.
- Columns: batter, runs, ballsPlayed, howOut, battingPosition, teamName, batFirst, targetScore, extraRuns, winner, matchId.

# Statistical Analysis and Causal Inference



# Association between a player's batting position and the likelihood of their dismissal

## RESULTS:

### Statistical Association

Chi-squared: 3753.28

P-Value: 0.0

### Average Treatment Effect (ATE) Analysis

ATE of Batting Position on Runs: -1.46 (p-value: 0.027)

ATE of Batting Position on Balls Played: -1.38 (p-value: 0.040)

# Inferences

- The statistical analysis reveals a strong link between batting position and dismissal likelihood.
- Specific batting positions seem to influence runs scored and balls faced, suggesting a potential impact on player performance.

# Distribution of dismissal types for Top-Scoring Batsmen

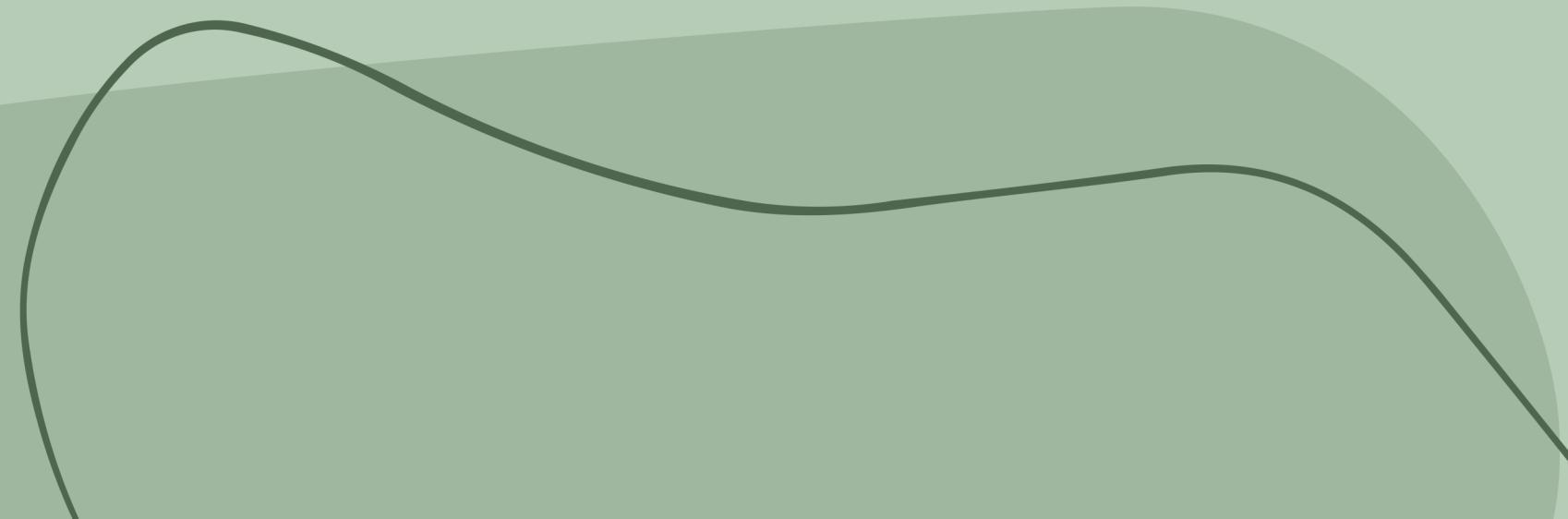
## RESULTS:

- Statistical Test:
  - Chi-square statistic: 324.20
  - P-value: 3.84e-62
- Average Treatment Effect (ATE):
  - ATE: 0.177
  - Inference: Top-scoring batsmen exhibit a higher average dismissal type compared to non-top-scorers.

# Inferences

**Reject the null hypothesis.** Significant differences exist in dismissal types for top-scoring batsmen.

By checking the **ATE**, we came to the inclusion that Top-scoring batsmen exhibit a higher average dismissal type compared to non-top-scorers.



# Forecasting Batting Performance: Machine Learning with Historical Similar Batting Positions"

## Results for LUMS Cricket Team

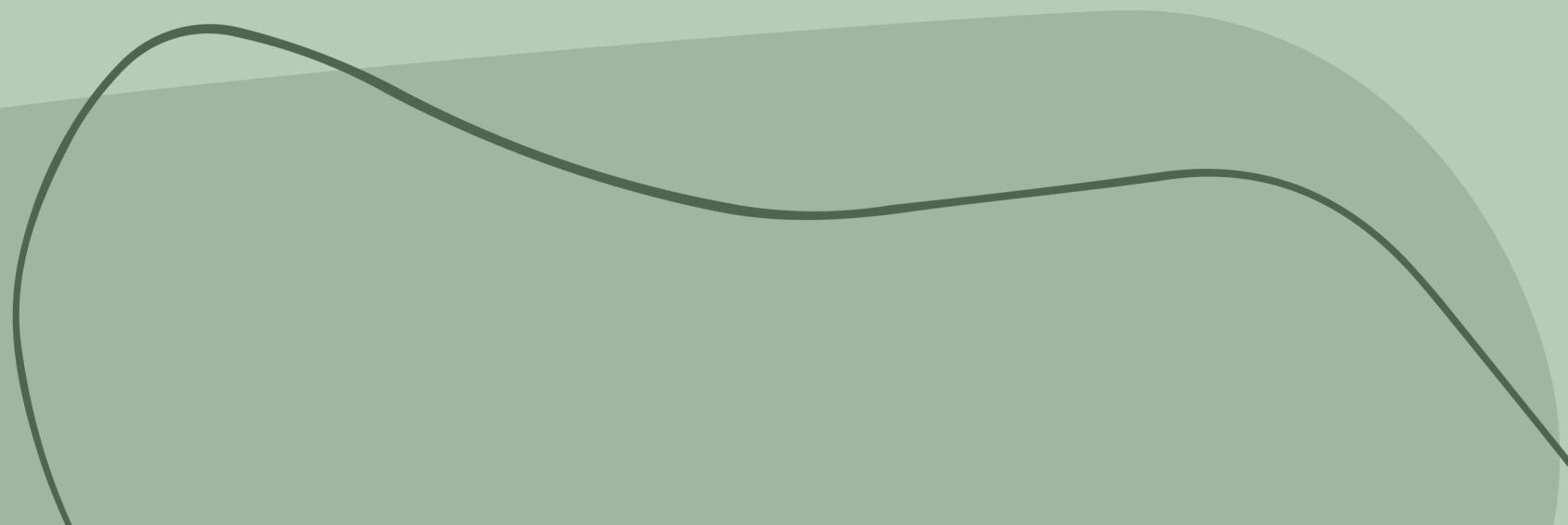
LUMS CRICKET TEAM

Position 1

1 Player: Owais, Predicted Score: 13  
2 Player: Mustafa, Predicted Score: 57  
3 Player: Awais, Predicted Score: 13  
4 Player: Umer, Predicted Score: 34  
5 Player: Mustafa, Predicted Score: 58  
6 Player: Adeen, Predicted Score: 45  
7 Player: Adeen , Predicted Score: 3  
8 Player: Mustafa, Predicted Score: 6  
9 Player: Mustafa, Predicted Score: 1  
10 Player: Sheheryar, Predicted Score: 5  
11 Player: Imran, Predicted Score: 45  
12 Player: Sheheryar, Predicted Score: 1  
13 Player: Iftikhar, Predicted Score: 12  
14 Player: Qaiser, Predicted Score: 0  
15 Player: Imran, Predicted Score: 86  
16 Player: Sheheryar, Predicted Score: 14  
17 Player: Mahad, Predicted Score: 44  
18 Player: Sherry, Predicted Score: 11  
19 Player: Sheheryar, Predicted Score: 14  
20 Player: Omar azeem, Predicted Score: 25  
21 Player: Sheheryar, Predicted Score: 24

# Inferences

Our analysis employs machine learning to enhance batting strategies for the cricket team. By examining historical performance and using a RandomForestRegressor model on 19 teams, we predict the impact of batting positions on player performance, aiming to optimize the batting lineup for strategic advantage.



# Thank You

