# Design and Development of a Sports Analytics Visualization Dashboard

[1]Prof. Dhaval Chandarana, [2]Prof. Hirenkumar Bhatt, [3]Parikshit Mevada,
[4]Dishant Chudasama, [5]Parth Parmar
*[1]Head of Department, Department of Information Technology GMIU, Bhavnagar, India*
*[2] Prof. Hirenkumar Bhatt*
*[2]Professor, Department of Information Technology GMIU, Bhavnagar, India*
*[3,4,5] B. Tech – Information Technology (Semester 6) GMIU, Bhavnagar, India*

***Abstract*—Recent advancements in modern data collection technologies in sports have created a necessity for creating efficient tools to analyse and comprehend the gathered data. Sports Analytics Visualization Dashboards present an exciting chance to enhance comprehension of intricate sports data through visual representations and insights. The Sports Analytics Visualization Dashboards collect and offer an interface to different data sources, presenting player performance, game statistics, and historical data through charts, heat maps, and key performance metrics. Additionally, the efficiency of data visualization dashboards, through storytelling, reveals insights into concealed patterns, trends, and relationships in the existing data, which traditional statistical tools and methods cannot uncover. The use of data visualization dashboards in sports enhances efficiency, while also being crucial for bridging the divide between data and understanding.**

## I. INTRODUCTION

Cricket is one of the most widely followed sports globally, attracting billions of fans across different continents. Over the years, the sport has evolved into a highly data-intensive domain, generating vast amounts of structured and unstructured data through scorecards, match statistics, tracking technologies, and textual commentary. The increasing availability of such data has led to the emergence of sports analytics as a significant research area. Big Data Analytics in sports has transformed the way performance evaluation, strategy formulation, and decision-making are conducted [11].

One-Day International (ODI) cricket, in particular, provides a well-structured statistical framework that makes it suitable for analytical modeling and visualization. Researchers have applied predictive modeling techniques to ODI matches to forecast match outcomes while the

game is in progress [10]. Similarly, optimization-based approaches have been utilized for cricket team selection and strategic decision-making, highlighting the importance of quantitative analysis in cricket performance evaluation [12]. Furthermore, statistical rating models have been proposed to assess individual player performance in international cricket formats [13].

The main goal of this project is to create a web application that allows users to visualize and analyze batting statistics in men's ODI cricket. The dataset consists of comprehensive career statistics of male players in the history of ODI cricket. By converting raw statistical information into significant visual formats, the system improves the understanding of data and analytical insights.

This system is aimed at cricket fans, analysts, and scholars who wish to create insightful visualizations of ODI batting statistics and examine the historical progression of player performance. Although the primary emphasis continues to be on the ODI format, the system's architecture has been designed to provide flexibility, allowing for the incorporation of data from other cricket formats like Test and T20 cricket in the future.

## II. SYSTEM REQUIREMENTS

The system requirements are categorized into non-functional requirements and functional requirements, outlining the operational and performance expectations of the proposed cricket analytics dashboard.

1. Non-Functional Requirements

R1. Web-Based Application:

The system needs to function as a web application that can be accessed via all major web browsers without the need for extra plugins or external installations. This guarantees independence from platforms and wide accessibility.

R2. User-Friendly Interface:

The app must offer an easy, straightforward, and user-friendly experience. The design should accommodate cricket fans who lack advanced technical skills, guaranteeing straightforward navigation and engagement[8], [9].

2. Functional Requirements

R3. Accessing Data from Cricket Data Source:

The system should be able to obtain data from an online cricket data source (e.g., ESPN Statsguru) and save it in an internal database. Furthermore, it must execute regular updates to ensure cricket statistics remain current[3], [11].

R4. Dataset Choice:
Users should be able to choose from several datasets available for visualization within the system. The selection of datasets will rely on two main criteria: Cricket formats (ODI, Test, T20) Gender classification (Men's or Women's cricket) [13].

R5. Chart Development:
Users need to create charts by querying chosen datasets. By defining parameters like performance indicators or player data, the system will generate significant visual outputs[8], [9].

R6. Filtering Charts:
The application must allow users to utilize filters to narrow down the data shown. Filtering options will enable users to concentrate on particular players, time frames, or statistical metrics, improving data investigation[8].

R7. Chart Analysis:
The system should enable the concurrent visualization of several charts, permitting users to juxtapose players or performance metrics for enhanced analytical insights[9].

R8. Functionality for Searching Players:
Users must be allowed to look for individual players in the dataset, allowing fast retrieval of particular statistical records.

R9. Player Profile Overview:
The system should offer an in-depth player profile display that showcases extensive data about a player's ODI career achievements.

## III. USER PROFILES

The intended users of the system are cricket enthusiasts eager to acquire knowledge. additional information regarding cricket and its extensive heritage. I have established two users. personas representing the ideal final users of this system: Ricky Baker and Raj Singh.

*Primary User*
Name: Raj Singh
Age: 32
Profession: Data Analyst and Cricket Writer
Bio:
Raj Singh is a skilled data analyst employed at a global IT firm. In addition to his full-time job, he manages a cricket blog where he shares match previews, post-match analyses, and reviews of player performances. Raj is deeply intrigued by statistics and takes pleasure in analyzing numerical

information to reveal significant insights. He frequently views international cricket games in all formats, particularly ODI and T20, and participates in online cricket forums.

Expectations:
Raj needs a more analytical and data-oriented tool that enables him to examine detailed cricket statistics effectively. He favors interactive dashboards that enable filtering, comparison, and advanced visualization capabilities. In contrast to casual fans, Raj frequently requires historical data trends, player consistency evaluations, strike rate comparisons, and performance metrics for defined time frames.
In the process of creating blog posts or social media content, Raj aims to swiftly produce visually appealing charts that can bolster his analysis and reinforce his points. He prioritizes precision, current statistics, and the capability to compare several players or teams at once. The system needs to enable him to export or take screenshots of visualizations for inclusion in his blog posts.

*Secondary user*
Name: Raj Singh
Age: 40
Occupation: PE teacher
Bio:
Raj Singh teaches physical education at St. Patrick's high school. University in Silverstream. Raj has a strong enthusiasm for fitness, activity and physical health. He puts in effort and takes his work with great dedication because he understands the significance of learning and bodily wellness. During his early years, Raj engaged in play. extensive cricket sessions with his friends, and he fostered a love for cricket. Even though Raj no longer plays cricket, he takes pleasure in studying and tracking the sport while teaching his students concerning it.

Expectations: Raj has recently launched a social media account with the aim of informing his students and broader community about cricket. On this page, he shares entertaining and intriguing facts about records, updates, and data in cricket. The web offers a abundant tabular data on cricket statistics, yet this information is dull and unsuitable for Raj's viewers. He is searching for a fast and engaging tool that offers straightforward depictions of cricket statistics.

Raj acknowledges the intricacy of cricket statistics, so he seeks a user-friendly system that will allow him to create representations of various types of cricket data analysis. Raj plans to utilize these visuals on his social media. posts to enhance the engagement of his content, educational and enlightening. Raj will utilize this tool to collect data for his material and he has no plans to invest significant quantities on it hence the instrument should be straightforward and easy to understand.

## IV. CONNECTED WORK

The project's background research and related work are covered in this part. We will examine a few current sports data visualisation tool solutions and compare them to this system. We will also go over some fundamental visualisation theory, which is essential for figuring out the system's architecture.

Existing System and Limitations:-
Cricket is a multifaceted sport that offers a vast array of data. (via ESPN cricinfo) yet this information is not easily obtainable or consumable in its provided format [3], [6]. Statsguru (by ESPN Cricinfo) is an internet-based tool that enables users to retrieve data from its database for cricket statistics [11].

Once the user sends a request, Statsguru provides a data table that includes all the resulting lines. Statsguru hosts the most extensive online repository of officially data documented in ODI cricket. Its information spans from the start of ODI cricket (1971) until now. This instrument is effective in the feeling that it enables users to select from multiple filters and it includes all of the information in the history of ODI cricket.

Despite being a solid data querying tool, Statsguru has its constraints. is that it lacks any functionalities for data visualization and hence restricts the possibilities for data evaluation and findings. Statsguru provides a table of unprocessed data, which can be challenging to examine and understand due to the intricacy and elements engaged in the activity. Owing to the absence of data visualization (such as charts) within this tool, users are missing a great deal of knowledge and results derived from this information. Over the past 52 years of ODI cricket, numerous elements have shifted and developed. the activity.

Cricket Visualizer by Lascarides:
Cricket Visualiser is a straightforward cricket visualization tool created by Michael Lascarides as a personal IT endeavour. This tool primarily concentrates on games held during the 2015 ICC World Cup. It enables users to track the progression of each match by displaying comprehensive details about each ball’s events.

Nonetheless, this instrument differs significantly from the system suggested in this study. Cricket Visualiser addresses only a single tournament (2015 World Cup), whereas our system emphasizes batting performance throughout the entire history of One-Day International (ODI) cricket [8], [9]. Due to this variation in emphasis, each system utilizes distinct types of data and fulfils different functions.

Cricket Visualiser displays highly detailed events of the match. It employs Manhattan graphs to illustrate each ball bowled in a match, assisting users in grasping the progression of the game over each over. Conversely, our system does not concentrate on specific match occurrences [7]. Rather, it utilizes comprehensive career statistics and team performance metrics to provide a broader perspective on performance trends.

Our system utilizes visualizations such as scatter plots to evaluate players and examine their career performance throughout different periods. This allows users to grasp long-term performance, easily contrast players, and acquire valuable insights, instead of focusing on a single match exclusively.

## V. METHODOLOGY

### A. Data Collection

Using web scraping techniques, the data for this study was gathered from ESPN Cricinfo, the official cricket statistics portal. Structured statistical data was extracted using tools like Selenium, Beautiful Soup (a Python module), and Bright Data [11]. For additional analysis and visualisation, the gathered data was cleaned, arranged, and saved in a database [3], [6].

The dataset includes the following performance indicators:

Batting Metrics: Total Runs, Batting Average, Strike Rate, Boundary Percentage

Bowling Metrics: Total Wickets, Economy Rate, Dot Ball Percentage

Fielding Metrics: Catches, Stumpings, Direct Run-Outs

Advanced Performance Metrics: Impact Index (overall influence of a player in matches), Player Consistency Score (performance stability across matches), Win Contribution Score (player's contribution in winning matches), Pressure Handling Index (performance in high-pressure situations)

AI-Based Analytical Metrics:
Sentiment Analysis of player performance based on match reports and commentary, Context-Based Performance Evaluation to assess decision-making in match situations [7].

Data Tidying and Feature Development

- Standardization of batting and bowling averages depending on match circumstances. Development of a Performance Index measure merging several performance metrics [12], [13].

• Grouping players with K-Means for classification based on roles (e.g., batsman, bowler, all-rounder).

• Recognition of player types based based on previous performances and historical patterns.

• Inclusion of outside elements like field conditions and opponent strength.

## V. DATA ANALYSIS

Initially, a single team's score and match result were analysed using straightforward visualisations like bar charts. Likewise, different kinds of bar charts were utilised to identify winning trends.

When deciding whether to bat first or second, another method for analysing the relationship between team score and run rate is linear regression [10].

These provided information regarding the high score and victory of the match when batting first. The linear relationship is depicted in Figure 1.

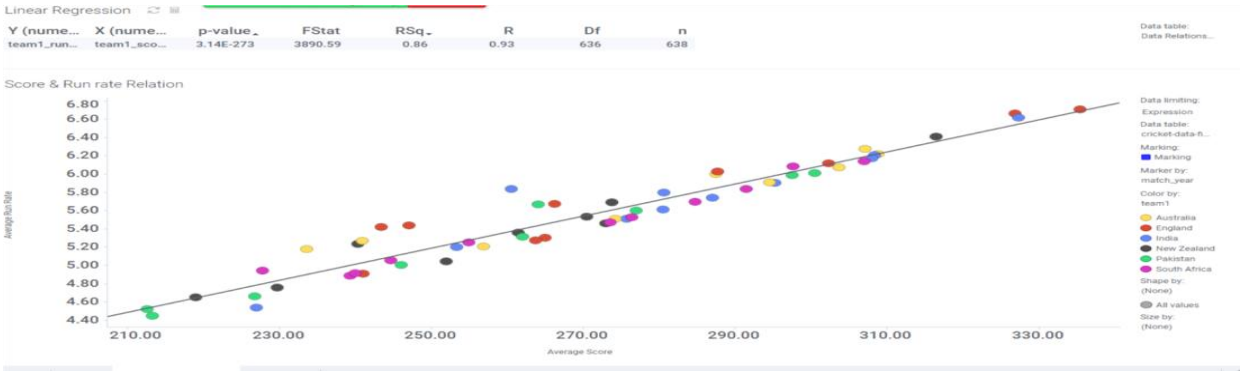



FIG.1 - AVERAGE TEAMS SCORE AND RUN RATE

## VI. VISUALISATION THEORY

The Eyes Have it

Ben Shneiderman's "The Eyes Have it" [8] is a foundational work that I read as part of the project's background research. Important background information was given in this text, which gave me the knowledge I needed to plan and create powerful visuals.

This book is a vital resource for learning how to design and produce graphical user interface visualizations that are efficient, interesting, and serve a purpose. Based on the kinds of data being displayed and the goals that users hope to accomplish, Schneiderman presents and explains his taxonomy for information visualizations in the text.

The objective is to offer a framework that enables researchers and designers to produce powerful visualizations [9]. Shneiderman notes that selecting the right visualization techniques requires an understanding of and classification of the data type (in the system).

The tasks that users typically complete when interacting with visualizations are also covered in the text.

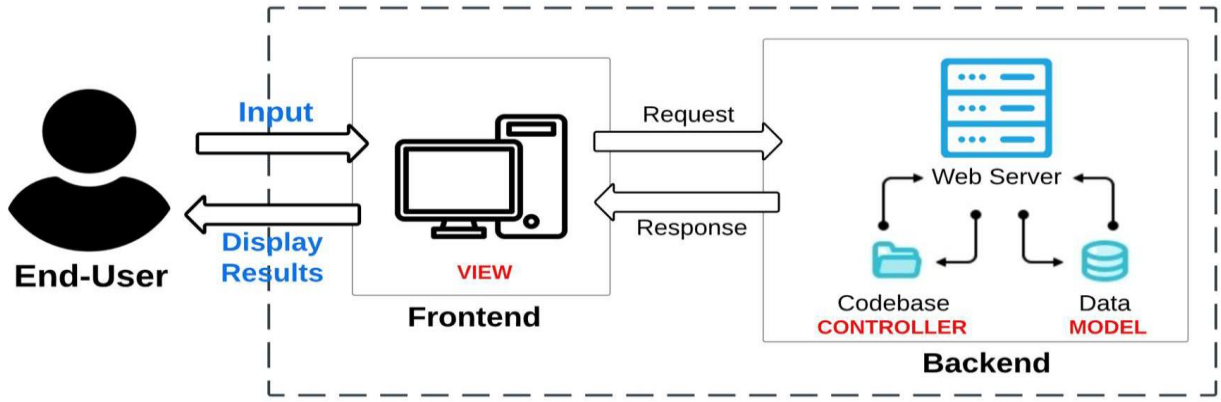



FIG.2 - SYSTEM MVC ARCHITECTURE

## VII. EVALUATION

In the final stages of the project, I conducted user study testing to evaluate the implemented system. The purpose of this evaluation was to:

- Assess the system's usability – the quality of the user experience
- Validate whether the software meets real user needs and expectations
- Identify common errors and obstacles
- Collect user feedback – opinions, suggestions, criticisms

*Testing Participants*

Seven user-testing participants participated in the evaluation. This group's demographic consisted of male fourth-year engineering students, all of whom were between the ages of 21 and 23. The results of the pre-test questionnaire showed an unpredictable distribution among the participants (see the following page). In particular, 57% of respondents said they were very confident in their knowledge of cricket, whereas 43% said they knew very little [8].

Additionally, there was a comparable divide in terms of interest, with 43% showing less interest in cricket and 57% showing a strong affinity for it. When looking at sports-related interest from a wider angle, it's interesting to note that 71% of participants showed a strong passion for sports, while 29% fell into the low-to-moderate interest category.

These results were crucial in shedding light on the participants' backgrounds, especially their connections to sports and cricket.

The charts' visual representations highlight the wide range of participant interest and knowledge in cricket. This diversity was crucial to guaranteeing a more thorough and fair assessment of the system. It made it possible to gather opinions and findings from users with different levels of expertise, providing a more thorough understanding of the system's functionality.

*User Testing Process*

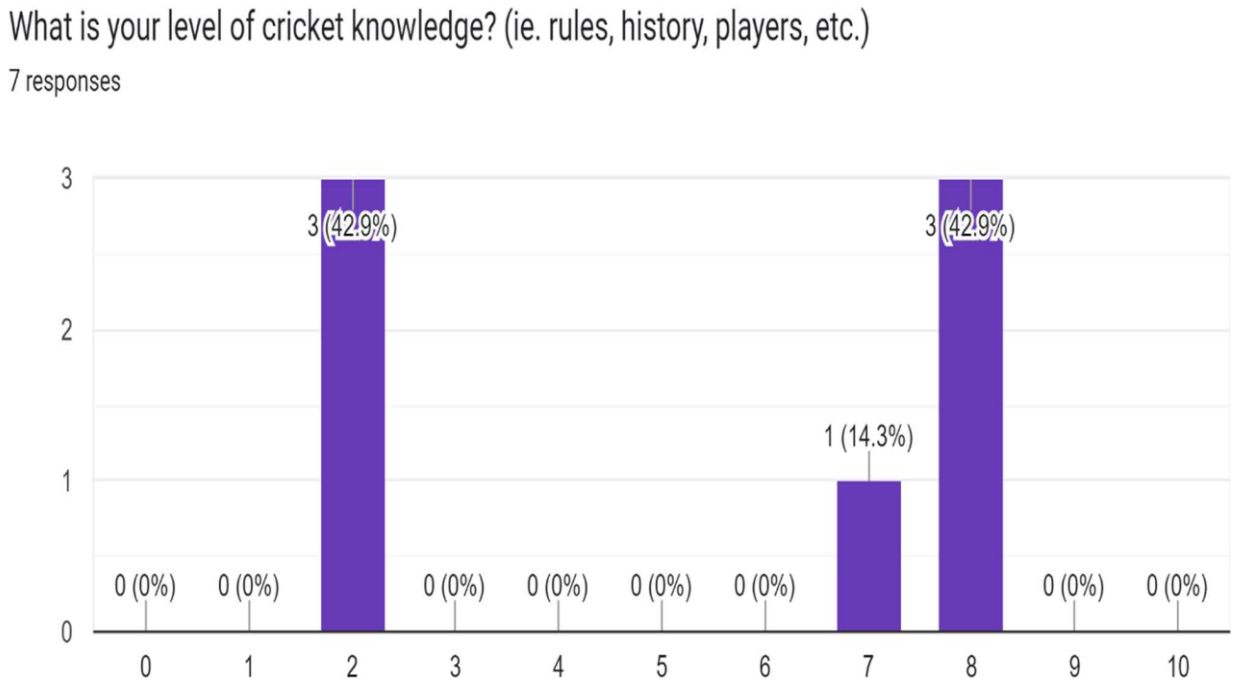
I'll now describe the steps I took to perform user testing on every participant. Three phases comprised the user testing: the system trial, the post-test questionnaire, and the pre-test questionnaire. For the seven participants, I conducted seven testing sessions in total. Each session lasted up to forty-five minutes. A Linux computer in lab room CO228 was used for the testing.

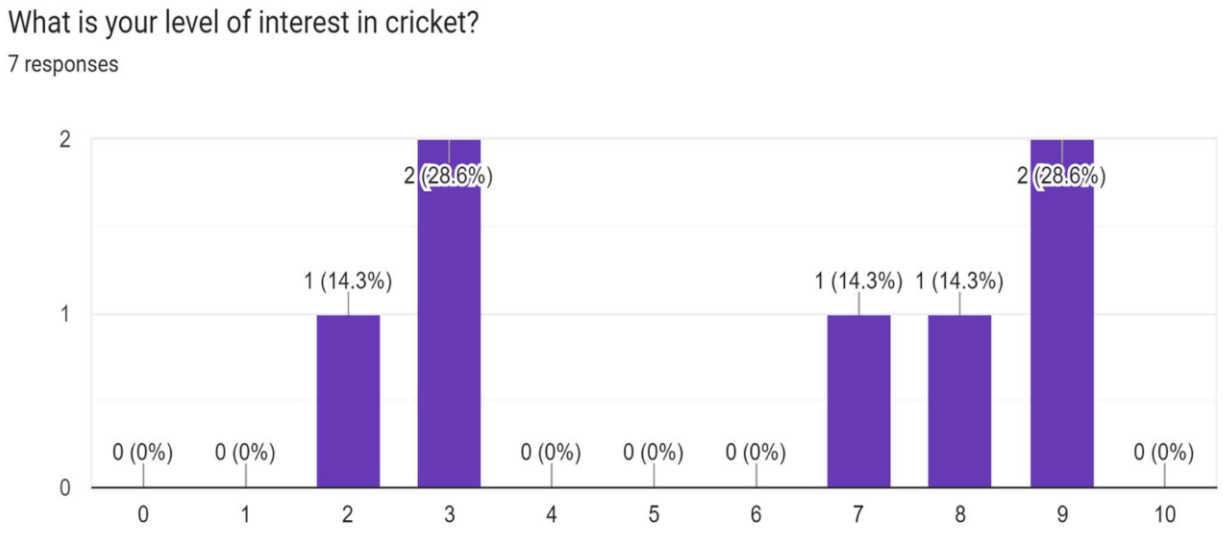
Test Questionnaire

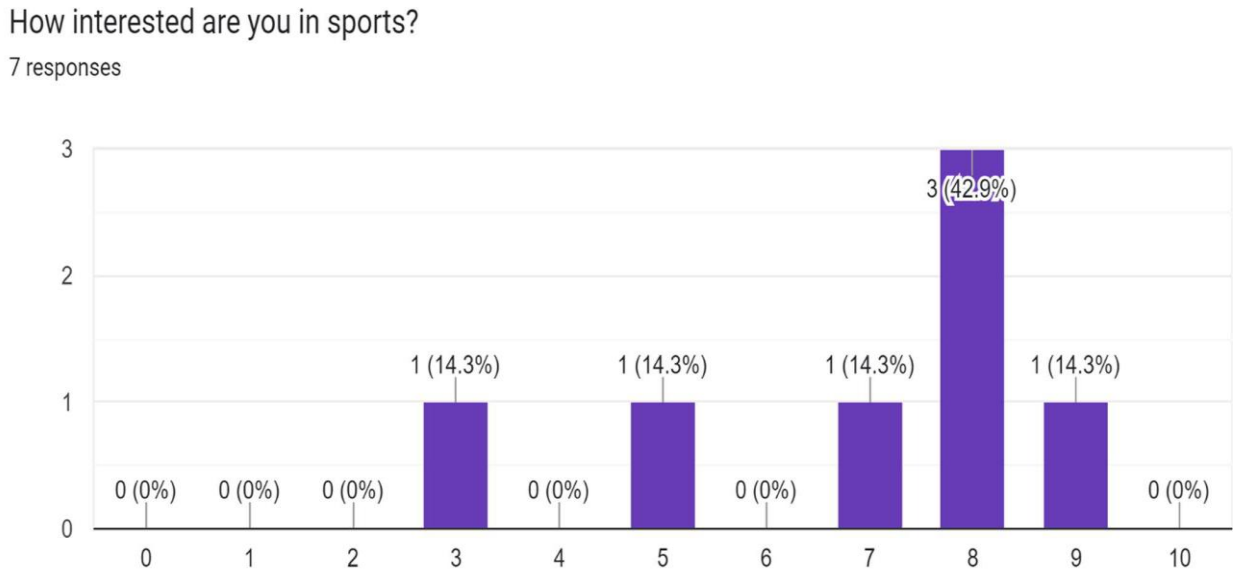
Before the session began, participants used Google Forms to complete a pre-test questionnaire. The purpose of this survey was to evaluate their backgrounds, particularly their involvement in cricket and other sports. To measure the following, this survey included three Likert scale questions.

- User's level of knowledge in cricket
- User's level of interest in cricket
- User's level of interest in sports

Test Questionnaire Results






System Trial

The system trial was the most important aspect of the session. I gave a user briefing at the beginning, outlining the goal of the application and giving background information on the data at play.

I provided the user with a notepad for comments and suggestions in order to guarantee thorough feedback. I also prepared a document to capture the user's answers to questions that were asked aloud.

There were two rounds in the trial. In order to acquaint the user with the system, the first round provided a simple, introductory user experience.

The user was led by thorough, detailed instructions through the process of creating a simple chart and utilizing two essential features for data exploration. To make sure the user understood the basic functions of the system, I asked three straightforward questions.

A more difficult task with two in-depth user journeys for in-depth data analysis was presented in the second round. The user had to rely on the interface design and heuristics to complete tasks because the instructions were less explicit.

An analysis of player strike-rate performance over several decades was part of the first journey. The second journey looked at players from four different countries and their century-scoring performances. Analysis and judgment were required because participants in these journeys were asked verbal questions. Seeing the various strategies used by participants was fascinating. This exposed parts of the system that gave rise to varying interpretations and methods, illuminating features that may require enhancement for improved user comprehension.

## VIII. RESULTS AND ANALYSIS

TABLE I BEST PLAYER PERFORMANCE

| Player Name | Team | Bat Avg | Bat S/R | Order |
|---|---|---|---|---|
| Jos Buttler | ENG | 45.00 | 144.23 | 1 |
| Rilee Rossouw | SA | 35.25 | 169.88 2 | 2 |
| Virat Kohli | IND | 98.67 | 136.41 | 3 |
| Suryakumar Yadav | IND | 59.75 | 189.68 | 4 |
| Glenn Phillips | NZ | 40.20 | 158.27 | 5 |

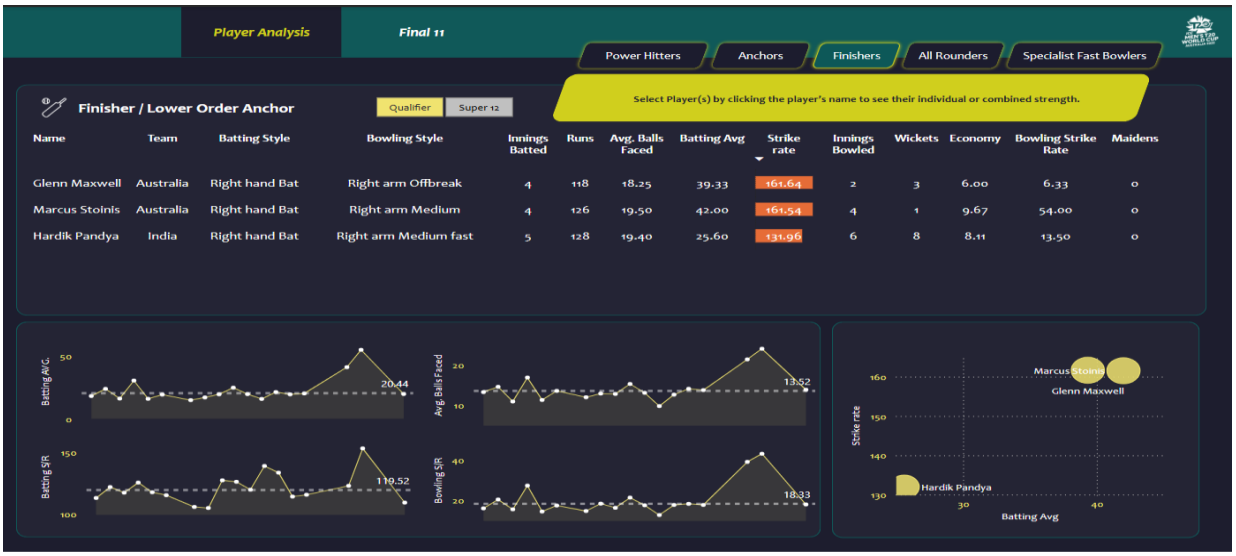



Fig. 1. power bi dashboard


FIG. 4. BEST PLAYER PERFORMANCE

*Key Findings*

• Jos Buttler (England) is a top-order wicketkeeper batter with a high strike rate (144.23) and a solid batting average (45.00).

• Rilee Rossouw (South Africa) is a lefthanded top-order batter with an exceptional strike rate (169.88).

• Virat Kohli (India) is a consistent performer with a high batting average (98.67) and a strike rate of 136.41.

• Suryakumar Yadav (India) stands out with an extraordinary strike rate (189.68), making him a key finisher.

• Glenn Phillips (New Zealand) is a versatile wicketkeeper batter with a strike rate of 158.27.

## IX. BACKGROUND, MOTIVATION, AND DESIGN GOALS

The history of the different types of cricket data is presented in this section. Unstructured short text commentary is used to discuss the motivation behind the current work. The system overview comes after the identification of fundamental requirements and the derivation of related design goals.

Background

Box-score information in cricket (e.g. G. scorecards) are the earliest type of data to be documented (in 1772). Only a few locations show a change in the box-score data, despite the fact that it is nearly 247 years old.

As a result, the modification to the box-score data visualization advanced the available supplementary data. Coaches and broadcasters were able to gain new insights into each shot, run, and wicket by using sophisticated data capture techniques in cricket. Hawk-Eye technology tracks the trajectory of the ball, the delivery speed, the position of the pitched ball, and the ball's bounce.

The data visualizations are utilized by analysts, experts, and umpires in their decision-making process. The primary obstacle with tracking data is that it is not publicly accessible. They also cost a lot of money to use in every game.

Meta data in cricket primarily takes into consideration circumstances that are not related to the game's context. Moisture conditions are among the examples.

Only a small amount of information unique to the events in matches is captured by the structured data used for box-score and tracking. Unstructured data, including text commentary, audio commentary, and video broadcast, records every detail of matches and events unique to those matches.

News articles and social media posts supplement the three sources of unstructured data mentioned above. Nonetheless, there are a few distinct benefits to text commentary data, which is a type of brief text. It is more accessible and uses less storage space and processing power than audio or video data. In contrast to posts on social media, it is reliable, keeps a consistent tone, and maps each delivery to a single textual commentary.

In contrast to news reports, it offers a close-up look at every detail of the game. Despite being around for a while, text commentary data is still an unexplored and untapped resource.

Motivation

There are several different visualization techniques available for tracking and box-score data. Nevertheless, there are no such visualization methods for unstructured data. In this study, unstructured data—such as box-score and tracking data—is utilized to uncover information not found in the structured data.

The purpose of this work is to add a collection of new visualizations utilizing unstructured data to the body of visualization techniques. According to this viewpoint, player-specific strategies are created through the analysis of shorttext commentary data in the form of visualization.

Design Goals

The following two design objectives are derived from the requirements that have been identified.

D1: Visualization of each player's strengths and weaknesses**.** A visualization technique is presented with the express purpose of illustrating the rules of strength and weakness. The flexibility of the suggested approach lies in its ability to extract and visualize the rules of strength and weakness in a range of scenarios, including

(i) a player's strength and weakness against a specific player,
(ii) a player's strength and weakness against a specific team,
(iii) a player's strength and weakness in aseries, etc.

D2: Visualization of similar players based on their strengths and weaknesses**.** The To visualize players who are similar, the design goal D1 is expanded.

## IX. DATA PROCESSING AND CONFRONTATION MATRIX CONSTRUCTION

The unstructured portion of short text commentary is dominated by domain-specific technical vocabulary. Two major issues with this technical vocabulary, however, are as follows:

• Stop words: The majority of this vocabulary refers to stop words found in the traditional literature on information retrieval. Table 1 displays the non-exhaustive list of technical stop words.

• Sparsity: One delivery of cricket commentary contains up to 50 words of unstructured content, despite having a large vocabulary.

Domain-specific bowling and batting features are found to help overcome these two obstacles. A set of unigrams and bigrams is used to represent each feature; this set is known as feature definition FD and is found to correspond to the feature in question. For instance, table 2 shows the corresponding unigrams and bigrams for the batting feature beaten.

Cricket experts were consulted in order to obtain this unigram / bigram to feature mapping. Table 3 lists the 19 batting features and 12 bowling features that have been identified. Nineteen (batting features) and twelve (bowling features) sets of unigrams and bigrams are derived in accordance with these features. The stop word-related issue has been resolved by this feature-gathering technique. Every unigram and bigram is mapped to one of the features in order to address the sparsity.

The procedure for obtaining the complete commentary database was described in Section 4. One must extract a subset of text commentary from this document in order to meet the design objectives.

- Player: The one player for whom strength rules and weakness rules are to be extracted.
- Opponent Player: The opponent player may be a single player or a group of players; for a batsman, the opponent is the bowler, and vice versa.
- Time: The element time can be per session, per day ,per innings, per match, per series, or an entire career.
- Type: Depending on whether we wish to examine the player in question's batting or bowling, it can be either.

## X. CONCLUSION

In this study, a web-based Sports Analytics Visualization Dashboard for evaluating One-Day International (ODI) cricket performance was designed and put into use. Large-scale, intricate cricket datasets are effectively converted by the system into insightful visual representations that improve comprehension and decision-making.

The dashboard goes beyond conventional tabular systems and offers more profound performance insights by fusing structured statistical data with sophisticated and AI-based analytical metrics.

The study shows that for a variety of user groups, including analysts, writers, educators, and cricket enthusiasts, interactive data visualization greatly enhances the usability, accessibility, and interpretability of cricket statistics. The system facilitates thorough examination of cricket data with features like filtering, player comparison, performance trend analysis, and strength-weakness visualization.

According to the findings of user evaluations, the dashboard successfully satisfies practical analytical requirements while remaining usable by both novice and expert users. By revealing contextual and strategic insights not apparent in structured datasets alone, unstructured commentary analysis further deepens the analysis.

Overall, the developed system bridges the gap between raw cricket data and actionable intelligence. It contributes to the growing field of sports analytics by combining visualization theory, machine learning techniques, and domain-specific feature engineering into a unified platform.

REFERENCES

[1] Jose, Kevin. "Cricket Data Visualisation App." *Wellington Faculty of Engineering Symposium*. 2023.

[2] Babu, Bincy, and Haripriya Surendran Nair."Project Report on Sports Analytics Visualization with Cricket Data."

[3] Vora, Shubham. "Cricket Data Analytics: Building the Best XI Players from T20 World Cup." *Available at SSRN 5180123* (2025).

[4] Tyagi, S., Kumari, R., Makkena, S. C., Mishra, S. S. and Pendyala, V. S.(2020) Enhanced Predictive Modeling of Cricket Game Duration Using Multiple Machine Learning Algorithms. In: *International Conference on Data Science and Engineering.* Kochi, India, 3rd-5th December.

[5] The evolution of ODI cricket | Andy Bull (2015) The Guardian. Guardian News and Media. Available at: https://www.theguardian.com/sport/2015/feb/10/the- evolution of- odi - cricket (Accessed: March 31, 2023).

[6] Tyagi, Jahanvi, et al. "BOUNDRY BOX: Cricket Data Analytics & Fantasy Team Model." Available at SSRN 5192677 (2025).

[7] Behera, Swarup Ranjan, and Vijaya V. Saradhi. "Visualization of Unstructured Sports Data--An Example of Cricket Short Text Commentary." arXiv preprint arXiv: 2404.00030 (2024).

[8] Losada, Antonio G., Roberto Therón, and Alejandro Benito. "Bkviz: A basketball visual analysis tool." IEEE computer graphics and applications 36.6 (2016): 58-68.

[9] Fu, Yu, and John Stasko. "Hoop insight: Analyzing and comparing basketball shooting performance through visualization. " IEEE Transactions on Visualization and Computer Graphics 30.1 (2023): 858-868.

[10] Bailey, M.; Clarke, S.R. Predicting the match outcome in one-day international cricket matches, while the game is in progress. J. Sports Sci. Med. 2006, 5, 480.

[11] Kaur, A.; Kaur, R.; Jagdev, G. Analyzing and Exploring the Impact of Big Data Analytics in Sports Sector. SN Comput. Sci. 2021, 2, 1–19. [CrossRef].

[12] Ahmed, F., Deb, K. and Jindal, A. (2013). Multi-objective optimization and decision-making approaches to cricket team selection. Applied Soft Computing, 13, 402-414.

[13] Akhtar, S., Scarf, P.A. and Rasool, Z. (2015). Rating players in test match cricket. Journal of the Operational Research Society, 66, 684-695.