INTUICRIC -VISUAL ANALYTICS OF CRICKET

Anegama Lahari
2286598
lanegama@cougarnet.uh.edu

Nitish Dadapuram
2304669
ndadapur@cougarnet.uh.edu

Narayan Soni
2301240
nsoni@cougarnet.uh.edu

***Abstract*— Cricket is hugely popular in the Indian subcontinent, yet most cricket websites fail to present team and player statistics in an intuitive, easy-to-understand manner. This project aimed to create a cricket data visualization that bridges this gap by enabling both technically skilled and casual fans to explore metrics of their interest. The visualization covers seasonal data on Indian Premier League (IPL) matches, displaying information on overall team performance, individual player statistics, progression of matches, and ball-by-ball scoring trends. Users can interact with integrated visualizations like bar charts, scatter plots, and line graphs to uncover insights. The navigation allows viewers to drill down from a global view of the IPL season into team and player specifics. Dropdown menus give control over axes displayed and teams focused on. Hovering over elements reveals further data. Clear linkage between multiple visualizations provides connections to spot patterns. This project ultimately transforms complex cricket data into an intuitive format for anyone to analyze, making statistics more engaging and match developments easier to grasp. The goal was for a broad audience, regardless of analytical proficiency, to find the visualization both usable and delightful.**

Keywords— Data-Driven Sports Insights, Intuitive Statistics Exploration, Integrated Multi-Graph Format, Audience centered

I. INTRODUCTION

Cricket boasts tremendous popularity across the Indian subcontinent, with millions tracking performance of teams and players. However, statistics presented on most cricket websites employ complex, technical formatting unsuitable for intuitive analysis. This frustrates both experts desiring to conduct detailed analytics and casual fans trying to interpret results. Our project attempted to transform dense cricket data into an interactive, visually inviting website promoting simplicity. Focus narrowed specifically to Indian Premier League (IPL) seasonal statistics to enable seamless cross-sectional investigation ranging from macroscopic performance to microscopic ball-by-ball dynamics. The resulting modular dashboard featuring integrated graphics like bar charts and scatter plots enables users to explore layers aligning with their background and interest.

Core project motivations involved: 1) Constructing a publicly accessible analytics portal satisfactory to experts in functionality while also proving fun and instructional even for novices personally analyzing data, and 2) Inspiring greater enthusiasm for engaging with statistics across user segments often intimidated by numbers alone. Success metrics consequently centered on accessibility, engagement, and educational stimulation judged qualitatively from structured user testing.

II. RELATED WORK

Interactive sports visualizations is increasingly common for major leagues. Prior work has explored visualizing details for sports including football, baseball, and basketball. Common capabilities include showing team performance over a season, integrating statistics per player and per match, and using coordinated multiple views. While some related cricket visualizations exist, they lack the sophistication, customization, and interactivity of our approach.

A number of prior efforts have explored cricket data analytics and visualizations. Cricbuzz is an Indian cricket news website owned by Times Internet. It provides news, articles, live match coverage, and basic team ranking and player statistics [1]. Cricbuzz is easier to use and understand than espncricinfo for naive users. While Cricbuzz proves more intuitive for casual users than alternatives like ESPN Cricinfo, visualization functionality remains limited.

The Premier League Season Explorer by B. Eisner, K. Wood, and J. Johnson from Hall of Fame Projects has inspired us in many ways [2]. For instance, the bump chart has encouraged us to select a line chart to display the positions of the teams after every match. Additionally, the visualizations on the Hall of Fame helped us understand the different ways the users can perceive the data. This enabled us to create better visualizations for native users.

The visualizations on IPLT20.com helped us understand the data in a better way [3]. For instance, the idea to use a scatter plot stems from the basic understanding that a specific player can have excellent or poor stats when using different metrics to view the player's statistics. Savle et al. previously visualized Indian cricket team player performance using parallel coordinates. Their work focused on Test and One-Day International (ODI) cricket for a single national team. In contrast, our work encompasses the IPL Twenty20 league with its franchise-based teams and international players. Another related work by Saikia et al. developed a system called CricketVis for exploring Indian cricket data [4]. This also focused on Test and ODI formats over IPL seasons. While having some common capabilities, our work is specialized for IPL cricket and provides greater interactivity through coordinated brushing and linking.

More broadly, sports analytics platforms have adopted a variety of visualization techniques for enabling insights into performance. Tableau differentiates itself through advanced graph customization, aggregation flows, and predictive analytics. PowerBI integrates interactive reports, natural language queries, and AI collation of insights. Finally, academic literature has produced various conceptual guidelines around effective sports data presentation.

III. Proposed System

The proposed interactive visualization dashboard offers rich insights into team and player performance over an Indian Premier League (IPL) season. Multiple coordinated plot types of trace progression at the franchise and individual levels while enabling drilling down into specifics.

Core components and capabilities include the following major components.

Points Table Tracking: A line chart displays teams' rankings throughout the season as matches occur. Position displays the rankings of each franchise over every gameweek in the IPL season based on the latest net run rates, so it reflects net run rates after each gameweek. This reveals a macro-level flow of performance against other franchises. As wins, losses and draws occur and run rates shift, the chart visually conveys how every team's standing progress compared to others in the competitive league table. Users can mouseover any point to view exact placement and the team. selecting a team filters all data.

Match Scorecards: These represent run rates per over in the form of bar charts visualize the progression of individual matches, with heights representing scoring intensity and milestones. Icons indicate key events like wickets falling. Line series overlay to connect over-by-over rates for readability. Tapping into specific games provides further details like batting lineups. The scorecards reveal team performance game-by-game and let users analyze contributing factors. Users Can mouseover on any bar that represents an over and get the entire summary of that over. Filtering produces cards for specific matches of interest. Line charts and scorecards are interlinked and enable you to view any particular match scorecard from the line chart match selection.

Player Analytics: Flexible 2D scatterplots leverage customizable x and y axes along with coordinated colors to compare player batting and bowling indicators over the season, such as runs scored, strike rate, dots bowled, economy rate and more. Interactive legends filter by statistics of interest. Selecting marks on players reveals statistical profiles. The graphs facilitate deeper inquiry exploring factors impacting individual cricketer performance for any franchise. Hovering displays the player, runs, batting average and team they belong to. Selecting any player displays the player card with more information aside the visualization.

Season Summary: Aggregate franchise results over months offer historical baseline for comparison. Win/loss differentials versus opponents indicate seasonal challenges to address. Selected line chart results in displaying team summary using some historic data.

Data flows from curated ball-by-ball, team standing, player stats and auction deal datasets. Interconnected graphical formats tuned for varied expertise levels answer strategic performance queries to study. Integrated filtering and details-on-demand offer customizable analysis without overloading beginners. Evaluations will assess effectiveness across fan levels.

IV. Implementation

The methodology began with a comprehensive understanding of the project's objectives and requirements. This involved defining the scope, identifying key features, and establishing the overall architecture of the IPL 2022 project.

Requirement Analysis:

Requirement Analysis was conducted to gather specific requirements, including the types of visualizations needed, data sources, user interaction expectations, and the technology stack suitable for implementation. Upon viewing many websites like Cricbuzz and EspnCricInfo we wanted the visualizations to be easy plots or graphs that the viewer can comprehend easily. Hence, Line charts, bar graphs and scatter plots were our choice, as they perfectly fit the data.

Design Planning:

This phase involved conceptualizing the user interface, navigation flow, and visual representation of data. Wireframes, mockups, and flowcharts were created to visualize the user journey and interface design.

Fig. 1. Initial Prototype Design on Paper

In this prototype design, we had a geographical map. It will have specific portions highlighted with different colors. These portions represent all teams in the league in the tournament's history since its inception. Users can select a state from the country map, highlighting the corresponding bar and lines. Line chart visualization will be displayed for that team with various measurements. Selecting any state representing a team highlighted in the map, highlighting corresponding lines in the adjacent line chart and table illustrating the related team data, and pie charts also should represent that team's player's statistics.

In the final design, the India map has been replaced by a ledger and a navigation bar. The teams are all represented in hexagons and with their respective logos. To aid user comprehension, a legend is provided, showcasing team colors and facilitating easier interpretation. Additionally, this section elucidates the unique aspects that made the IPL 2022 season exceptional, along with a summary of the season, contributing to a richer understanding of the tournament's significance. Upon the selection of the team, the other visualizations are further updated dynamically. The main chart highlights the league positions of the team after every game week. Several charts are used to display the statistics of the team.

Data Collection and Preprocessing:

A crucial aspect was the collection of IPL 2022 data from reliable sources. The collected data underwent preprocessing stages, including cleaning, normalization, and structuring, ensuring its suitability for visualization. For our IPL 2022 project, we utilized a primary dataset sourced from Kaggle, encompassing various aspects of the tournament across different seasons. Additionally, data for team positions after each game week was gathered from cricbuzz.com, espncricinfo.com, or through CricAPI. The primary dataset comprised five CSV files: Auction data, Ball-by-Ball details, Bowling Statistics, Batting Statistics, and Team Positions. This comprehensive dataset offered insights into player performances, match details, and team statistics for the 2022 season.

Our data preprocessing aimed to ensure data quality and streamline the dataset for effective visualization. The initial dataset had minimal missing values (0.2%). To maintain data integrity and simplify analysis, we opted to drop rows with missing values. Additionally, certain columns considered trivial for visualization clarity were removed to focus on pertinent information. A crucial enhancement involved enriching the dataset with additional columns to address the absence of league positions for teams after every game week. This augmentation aimed to provide a holistic view of team standings throughout the season, facilitating more insightful visualizations and analysis. The exported dataset was in CSV format, ensuring compatibility and ease of use across various analytical tools. Our decision to drop missing values and refine columns aimed to optimize the dataset for straightforward and robust analysis of IPL 2022.

Through these data collection and preprocessing strategies, we tailored the dataset to suit the project's visualization needs.

Technology Selection and Development:

Selection of appropriate technologies and tools was based on compatibility, scalability, and ability to handle real-time updates. Consideration was given to frontend frameworks, and visualization libraries that align with project requirements.

D3.js, or Data-Driven Documents, is a JavaScript library renowned for its prowess in creating dynamic and interactive data visualizations on the web. It excels in binding data to the Document Object Model (DOM) and facilitating seamless updates and transitions based on changes in the dataset. Its key features include robust DOM manipulation, scalability through SVGs, and the creation of smooth transitions and animations, enhancing user experience. While its flexibility allows for highly customized visualizations, mastering its capabilities might pose a steep learning curve. It is known for being a powerful choice for developers seeking to craft sophisticated and responsive data-driven visualizations.

Agile methodologies were employed for iterative development. The project was divided into sprints, allowing for continuous improvement, feedback incorporation, and adjustments based on evolving needs.

Implementation is done starting with building the front-end using HTML, CSS and JavaScript. Following this was Visualizations Implementation, Data Integration, Testing and finally Deploying in GitHub Pages.

Contribution:

Lahari played a pivotal role in the project by overseeing the design and implementation of the visualizations. Leveraging a deep understanding of data visualization principles and tools, she meticulously crafted the visual representations that formed the cornerstone of the project. Her expertise in visualization libraries such as D3.js ensured the creation of dynamic, interactive, and aesthetically pleasing visualizations. Lahari's dedication to user experience and design principles significantly enriched the project, enhancing its visual appeal and facilitating comprehensive data exploration for users.

Nitish's expertise in front-end development and design was instrumental in shaping the core structure and aesthetics of the project. With proficiency in HTML, CSS, and front-end frameworks, he meticulously constructed the body of the webpage, ensuring a seamless and visually engaging user interface. Nitish's attention to detail and commitment to design principles played a pivotal role in creating an intuitive and user-friendly web environment. His contributions laid a solid foundation for an aesthetically pleasing and functionally robust platform.

Narayan made substantial contributions by procuring, organizing, and preparing the datasets integral to the project's success. His meticulous approach to data collection and management aligned seamlessly with the project's design and requirements. The data sourcing and preprocessing ensured the datasets met the project's specific needs, maintaining data integrity and relevance which were critical in providing the team with comprehensive and structured datasets that formed the backbone of the project's analysis and visualizations.

V. RESULTS

The interactive Indian Premier League (IPL) data visualization web page is implemented as a single page web application using HTML, CSS and JavaScript. The primary interface is designed to welcome users with an introductory section providing an overview of the project and details about the contributing authors. The interface features two prominent buttons, offering users the choice between exploring visualizations or accessing an in-depth document detailing the project's entirety.

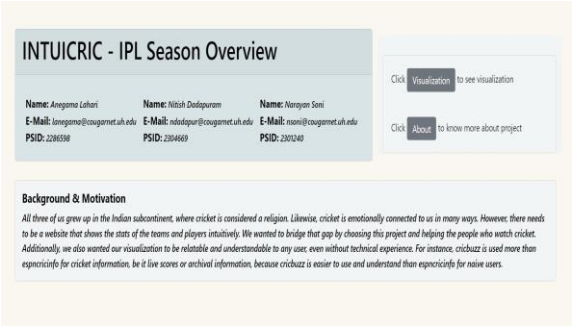
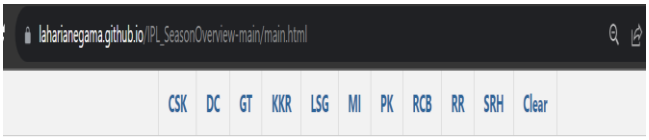


Fig. 2. Initial Page

Navbar:

A pivotal component of the webpage is its dynamic Navigation Bar, empowering users to select specific teams. Upon selection, the visualizations are dynamically updated, presenting hexagon representations of all IPL 2022 teams alongside their respective logos. This feature serves as a gateway to delve deeper into team-specific analyses.



anchises

Fig. 3. Navbar

The Franchise Overview visualization is a crucial feature, displaying a comprehensive depiction of all participating IPL 2022 teams. To aid user comprehension, a legend is provided, showcasing team colors, and facilitating easier interpretation. Additionally, this section elucidates on the unique aspects that made the IPL 2022 season exceptional, contributing to a richer understanding of the tournament's significance.

Fig. 4. Ledger

Fig. 5. Overall Navbar

Points Progression module:

Meticulously designed to showcase the standings of each team after every match in the season. Users can dynamically view wins, losses, and Net Run Rate (NRR) for comparative analysis. Interactive features include tooltips displaying team names upon hovering over lines and match summaries accessible by clicking on nodes. The navigation bar integration further allows for instant access to specific team summaries for the 2022 season.

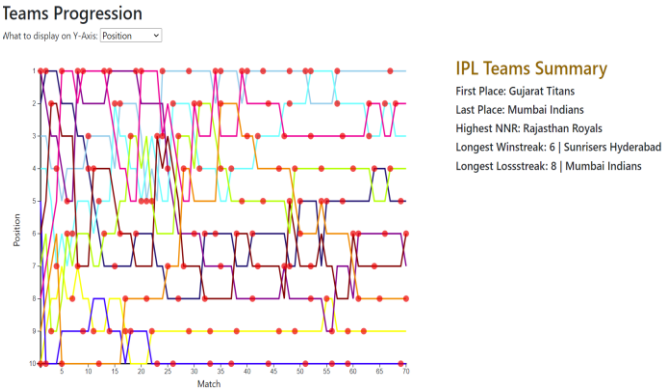


Fig. 6. Teams Progression

Below are few potential questions we can answer from this:

1. How did a specific team's rankings trend over the course of the IPL season?
2. Which match was most impactful for a team's season results?
3. Who were the most/least consistent teams in terms of points progression?
4. Did wins have a stronger impact than losses or net run rate for rankings?
5. Which qualifying race was the closest by the season end?

Individual Match Overview:

This visualization presents a detailed overview of runs scored in every over by both competing teams. Interactive elements allow users to hover over specific overs to reveal the runs scored by every batsman in the over along with the number of extras and wickets. Access to match specifics, including summaries (won/loss), is facilitated through dropdown lists offering users an immersive understanding of each match.The option to select the matches until the finals are provided through the dropdown list. The specific information about the match, i.e., the match summary is displayed on the side for the user to better understand the match.

Fig. 7. Bar plots

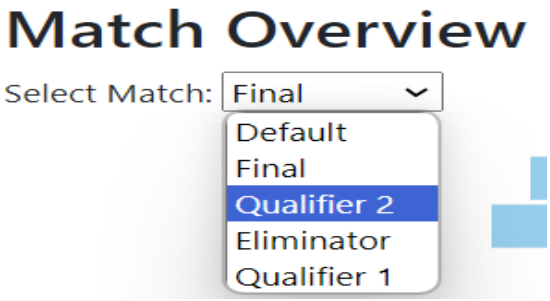


Fig. 8. Dropdown options

The Individual Match Overview visualization enables answering several key questions to deeply analyze matches:

1. Which overs saw momentum shifts between teams?
2. Who were the highest impact batsmen in accelerating run rates?
3. How did bowlers contain batsmen at different phases of the innings?
4. When did matches become effectively one-sided?
5. How evenly was the scoring distributed between batsmen?

ScatterPlot:

The Player Statistics visualization meticulously captures individual player performance throughout IPL 2022. Users can flexibly select variables for both the X and Y axes, fostering unbiased analysis. Interactive features include hover-over data displays for individual players, with additional player details accessible through clickable circles. The visualization also incorporates a legend to elucidate the significance of circle radius, denoting the number of matches played by each player.

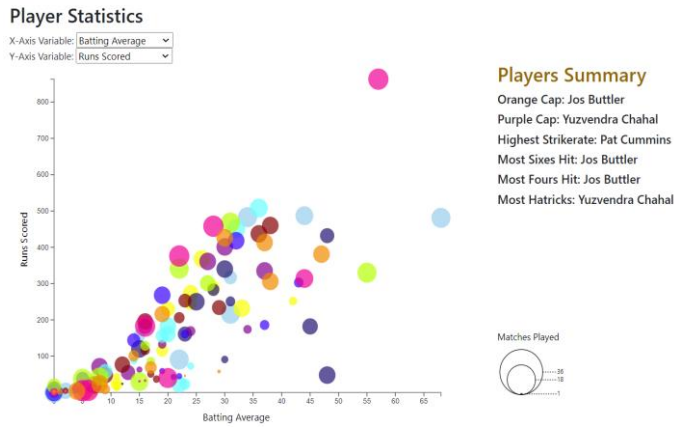


Fig. 9. Player Statistics plot

This Scatterplot helps us visualize the data in 2-Dimensions.

Here are some key questions that could be answered from the Player Statistics visualization:

1. Who was the most consistent player across different statistical measures?
2. Which players delivered the best returns for the number of matches played?
3. How evenly were runs distributed across teams or formats?
4. Who were the emerging breakthrough stars of the season?
5. What new records were achieved by players this season?

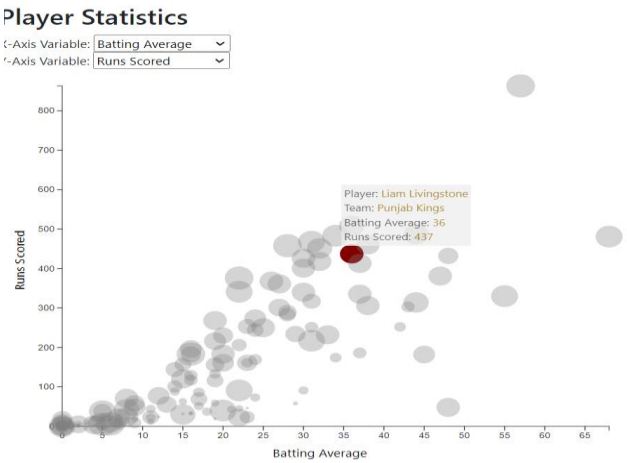


Fig. 10. Hover on player display

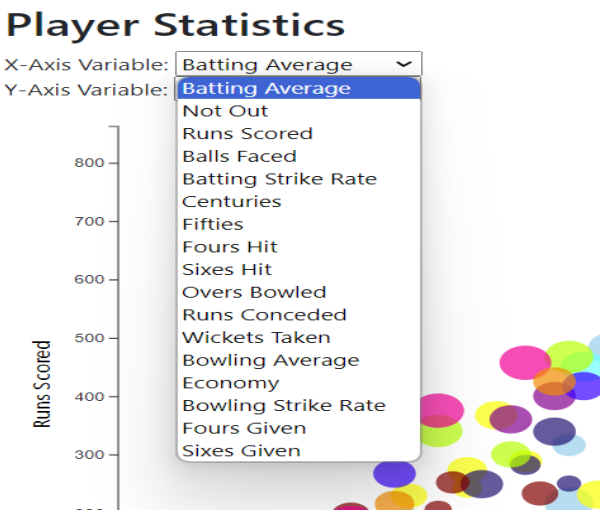


Fig. 11. X-axis Dropdown Options

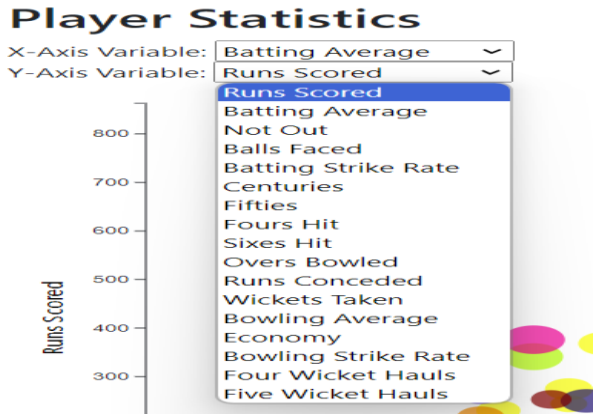


Fig. 12. Y-axis Dropdown Options

We show how each team performed versus other teams and as a function of time over the entire season. We also allow the user to select any individual games and see the significant events during each game, such as when the runs scored exceed a certain limit or when a wicket falls.

Also, we tell stories about a team in a particular season. We have shown data related to transfer of players between different seasons across the seasons. We also show the detailed statistics of a particular player like the number of wickets taken, dot balls bowled etc.

We intend to answer multiple questions, such as:

1. The position of every team after every match in the points table.
2. The statistics of several players compared with other players.
3. The scorecard of every game, which shows the runs scored by the team in every over
4. Using a scatter plot to display players' statistics where the user can select the X-axis, and the Y-axis enables the user to understand the statistics more clearly. For instance, the user can use the scatter plot to understand players who score low runs but play with a high strike rate.
5. Checking the match summary to check where the teams usually lost their matches and further improving their weaknesses can improve their performances in the future.

.

VI. Conclusion

The IPL 2022 project stands as a testament to the pursuit of graphical excellence in data visualization, aiming to elevate both vision and understanding through innovative design and meticulous execution.

Our endeavor to improve vision revolves around crafting visualizations that marry clarity and aesthetics. By prioritizing clear and uncluttered designs, we aim to captivate users' attention while conveying information effectively. This commitment to visual excellence ensures an engaging and visually appealing platform. Simultaneously, our focus on improving understanding embodies simplifying complex data and providing contextualization. Through intuitive visual representations and contextual information, users can effortlessly comprehend intricate data patterns, empowering them to derive meaningful insights.

Interactivity serves as a cornerstone for achieving deeper insights. By incorporating interactive features, users are empowered to explore and interact with data points, enabling personalized and dynamic data exploration. This interactive experience fosters a deeper understanding of IPL 2022 statistics and performances.

Moreover, the project emphasizes storytelling and consistency to guide users through a coherent narrative. Visualizations are harnessed not only to convey data but also to tell a compelling story, ensuring a seamless and engaging user experience across the platform.

Ultimately, our project centers on empowering users through visualization. By harnessing the power of data visualization as a tool, we enable users to gain meaningful insights and make informed decisions. The project's success lies in its commitment to delivering a user-centric, visually compelling, and insightful platform that facilitates comprehensive exploration and understanding of IPL 2022 data. Through a harmonious integration of these principles, the IPL 2022 project embodies a commitment to graphical excellence, leveraging visualization as a means to enrich user experiences, unlock insights, and enable informed decision-making.

VII. Future Scope

Few Limitations of the project are Limited historical data - Since the visualizations focus on IPL 2022 season. Narrow focus - The current scope looks only at IPL cricket data. Widening support to other domestic leagues, international cricket, or additional sports could broaden appeal. Lack of predictive insights - The analytics presented are more descriptive based on past match data. Incorporating some machine learning to unlock predictive capabilities could be more powerful. Difficulty keeping data updated - With fast paced sports schedules, new data gets created daily. The backend data pipelines would need to reliably sync to have latest information. Web-only access - Building mobile applications alongside the web app could allow for native platform optimizations and offline usage for intermittent connectivity. Visualization simplicity - Advanced users may want more customizable and complex visual charting capabilities than available out-of-the-box options.

Hence the Future Work could be:

Advanced Visualization Techniques: Explore 3D, VR, or AR for immersive experiences.

Real-time Data Updates: Implement live match updates and automate data retrieval for timely information.

Predictive Analytics: Utilize machine learning or statistical models for match outcome predictions based on historical data.

Enhanced User Interactivity: Develop customizable views and comparative analysis tools for deeper user engagement.

Mobile Optimization: Ensure seamless usability on mobile devices to broaden accessibility.

Social Integration: Integrate social media elements and create a community space for user engagement and discussions.

Expanded Data Sources: Consider adding sentiment analysis or weather data to enhance insights.

Accessibility Focus: Adhere to accessibility standards and incorporate features for inclusivity.

User Feedback Loop: Continuously improve the platform based on user feedback and analytical insights.

These potential avenues represent opportunities for further development and enhancement, allowing the project to evolve and offer more comprehensive and engaging experiences to users exploring IPL 2022 data.

VII. References

[1] https://www.cricbuzz.com/

[2] Eisner, B., Wood, K., and Johnson, J. (2023). The Premier League Season Explorer. Hall of Fame Projects.

[3] https://www.iplt20.com/points-table/men/2023

[4] Savle, et al. "Analyzing Performance of Indian Cricket Team Players Using Parallel Coordinates."