International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

CRICKET STATISTICS ANALYSIS AND INTERACTIVE DASHBOARD DEVELOPMENT USING PYTHON AND POWER BI

[1] J Priyanka, [2] G Teja, [3] G Upender, [4] U Nandishwar, [5] Malavath Nuthan Kumar
[1]AssistantProfessor, [2345]Students
Department of CSE(Data Science)
Siddhartha Institute of Technology & Sciences, Narapally

Jalagapriyanka.cse@siddhartha.co.in, 24tq1a6719@siddhartha.co.in, 24tq1a6701@siddhartha.co.in, 24tq1a6745@siddhartha.co.in, 24tq1a6729@siddhartha.co.in,

Abstract

Cricket is one of the most popular sports worldwide, generating vast amounts of data from matches, players, and tournaments. The Cricket Statistics Analysis and Interactive Dashboard Development project focuses on analyzing this data using modern analytics tools to extract meaningful insights and support better decision-making in the sport. The objective is to evaluate player performance, team strategies, and match outcomes through systematic data analysis.

The project utilizes Python programming language for data preprocessing, cleaning, and statistical analysis, along with Microsoft Power BI for interactive data visualization. Various datasets containing details of players, matches, scores, and performance metrics are processed using Python libraries such as Pandas and NumPy. The cleaned and structured data is then imported into Power BI to develop dashboards and visual reports that effectively present trends and patterns. Through this analysis, key insights such as player consistency, strike rates, win probabilities, and team performance comparisons are identified. Visualizations including bar charts, line graphs, and heat maps help simplify complex datasets and improve interpretability. The project also examines factors influencing match outcomes, enabling better evaluation of player contributions and team strategies.

Overall, this project demonstrates how data analytics tools can be effectively applied in sports analytics. The results provide valuable insights for coaches, analysts, and cricket enthusiasts, supporting data-driven decision-making. Furthermore, the approach can be extended to predictive analytics and real-time performance monitoring in cricket and other sports.

I. Introduction

Cricket is one of the most popular sports in the world, followed by millions of fans and generating a vast amount of data from matches, players, and tournaments. With advancements in technology, detailed information such as player performance, team statistics, and match results is continuously recorded. This data provides an opportunity to gain valuable insights and improve the overall understanding of the game. In recent years, data analytics has become an essential part of modern sports. Teams, coaches, and analysts increasingly rely on data to evaluate player performance, design strategies, and make informed decisions. Cricket analytics helps identify

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

strengths and weaknesses of players and teams, enabling performance improvement and strategic planning. It also allows comparison of players based on metrics such as runs scored, wickets taken, and strike rates.

This project focuses on analyzing cricket data using Python programming language and Microsoft Power BI. Python is used for data cleaning, preprocessing, and statistical analysis, while Power BI is used to create interactive dashboards and visualizations. These tools transform raw cricket data into meaningful and easy-to-understand insights. The dataset used in this project includes key attributes such as runs, wickets, strike rates, player performance, and match outcomes. By processing and analyzing this data, patterns and trends can be identified, helping to understand how players and teams perform under different match conditions. Visual elements such as charts, graphs, and dashboards simplify complex data and make it accessible even to non-technical users.

II. Literature Survey

In recent years, data analytics has played a significant role in the field of sports, especially in cricket. Earlier, cricket analysis relied mainly on basic statistics such as batting average, strike rate, and bowling economy. However, with advancements in technology, researchers have started using more sophisticated analytical techniques to process large datasets and extract deeper insights into player performance and team strategies. These modern approaches provide a more comprehensive understanding of the game compared to traditional methods.

Several studies have utilized Python programming language for cricket data analysis due to its powerful libraries such as Pandas, NumPy, and Matplotlib. These tools enable efficient data cleaning, transformation, and visualization. Researchers have applied these techniques to evaluate player performance, compare team statistics, and identify trends across matches and tournaments. The findings from these studies highlight Python’s effectiveness in handling large-scale sports data and performing detailed statistical analysis.

In addition to data processing, visualization tools like Microsoft Power BI have gained importance in recent years. Power BI allows users to create interactive dashboards and reports that present complex data in a simple and intuitive format. Studies emphasize that visualization plays a crucial role in helping users quickly understand patterns, trends, and relationships within the data, thereby improving decision-making.

Some research has also focused on predicting match outcomes using historical cricket data. These studies apply machine learning algorithms to analyze factors such as player form, pitch conditions, and team composition. Although prediction accuracy varies, these approaches demonstrate the potential of analytics in forecasting match results and enhancing strategic planning.

Other works concentrate on player performance analysis, including identifying consistent performers, analyzing strike rates, and comparing batting and bowling

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

efficiency. These insights help teams and coaches in player selection, strategy formulation, and performance improvement. Additionally, data analytics has contributed to talent identification by evaluating players based on objective performance metrics.

Despite the availability of advanced research, there remains a need for simple, integrated, and user-friendly systems that combine data analysis and visualization. Many existing solutions focus either on analysis or visualization, but not both in a unified framework. This project addresses this gap by integrating Python for data processing and Power BI for visualization, providing an accessible system for analyzing cricket data.

In conclusion, the literature indicates that data analytics has become an essential component of modern cricket. The use of tools like Python and Power BI enables effective analysis and visualization of complex datasets, supporting data-driven decision-making and improving overall performance in the sport.

III. System Analysis

Cricket generates a large amount of data from matches, players, and tournaments, making it important to analyze this data effectively. Traditional analysis methods are not sufficient to handle large datasets and extract meaningful insights. There is a need for a system that can process, analyze, and visualize cricket data efficiently. Data analytics helps in evaluating player performance and team strategies. Understanding metrics like strike rate, average, and win probability is essential. Visualization tools help simplify complex data. The system should support interactive exploration of data. It should provide insights into player consistency and team performance. Real-time or near real-time analysis can improve decision-making. The system must be scalable to handle large datasets. Integration of analysis and visualization is necessary. This project addresses these needs using modern tools.

Existing System

In the existing system, cricket data analysis is mostly done manually or using basic statistical tools. Analysts rely on simple metrics like averages and totals. Data is often stored in spreadsheets without proper structure. There is limited use of advanced analytics techniques. Visualization is minimal or static. There is no integration between data processing and visualization tools. Performance comparisons are difficult to perform. Large datasets are hard to manage and analyze. Decision-making is based on limited insights. There is no interactive dashboard for exploring data. Trend analysis is not properly implemented. Overall, the existing system lacks efficiency and depth.

Disadvantages of Existing System

- Limited data analysis capabilities
- No advanced statistical insights
- Lack of interactive visualization

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

- Manual and time-consuming processes
- Difficulty in handling large datasets
- No integration between tools
- Limited performance comparison
- Poor decision-making support
- Static reports with no interactivity
- No real-time analysis
- Inefficient data management
- Reduced accuracy in insights

Proposed System

The proposed system is an advanced cricket analytics solution using Python programming language and Microsoft Power BI. Python is used for data cleaning, preprocessing, and statistical analysis. Libraries such as Pandas and NumPy handle large datasets efficiently. The processed data is then imported into Power BI for visualization. The system provides interactive dashboards for exploring cricket data. It analyzes player performance metrics such as runs, wickets, and strike rates. Team comparisons and match outcome analysis are included. Visualization tools such as charts and graphs simplify data interpretation. The system supports trend analysis over time. Users can filter and drill down into data. It is user-friendly and efficient. Overall, it improves data-driven decision-making in cricket.

Advantages of Proposed System

- Advanced data analysis capabilities
- Interactive dashboards and visualizations
- Efficient handling of large datasets
- Integration of analysis and visualization
- Better player and team performance insights
- Supports trend and pattern analysis
- Easy comparison of players and teams
- User-friendly interface
- Faster and accurate decision-making
- Scalable and flexible system
- Improved data management
- Enhanced understanding of cricket data

IV. Methodology

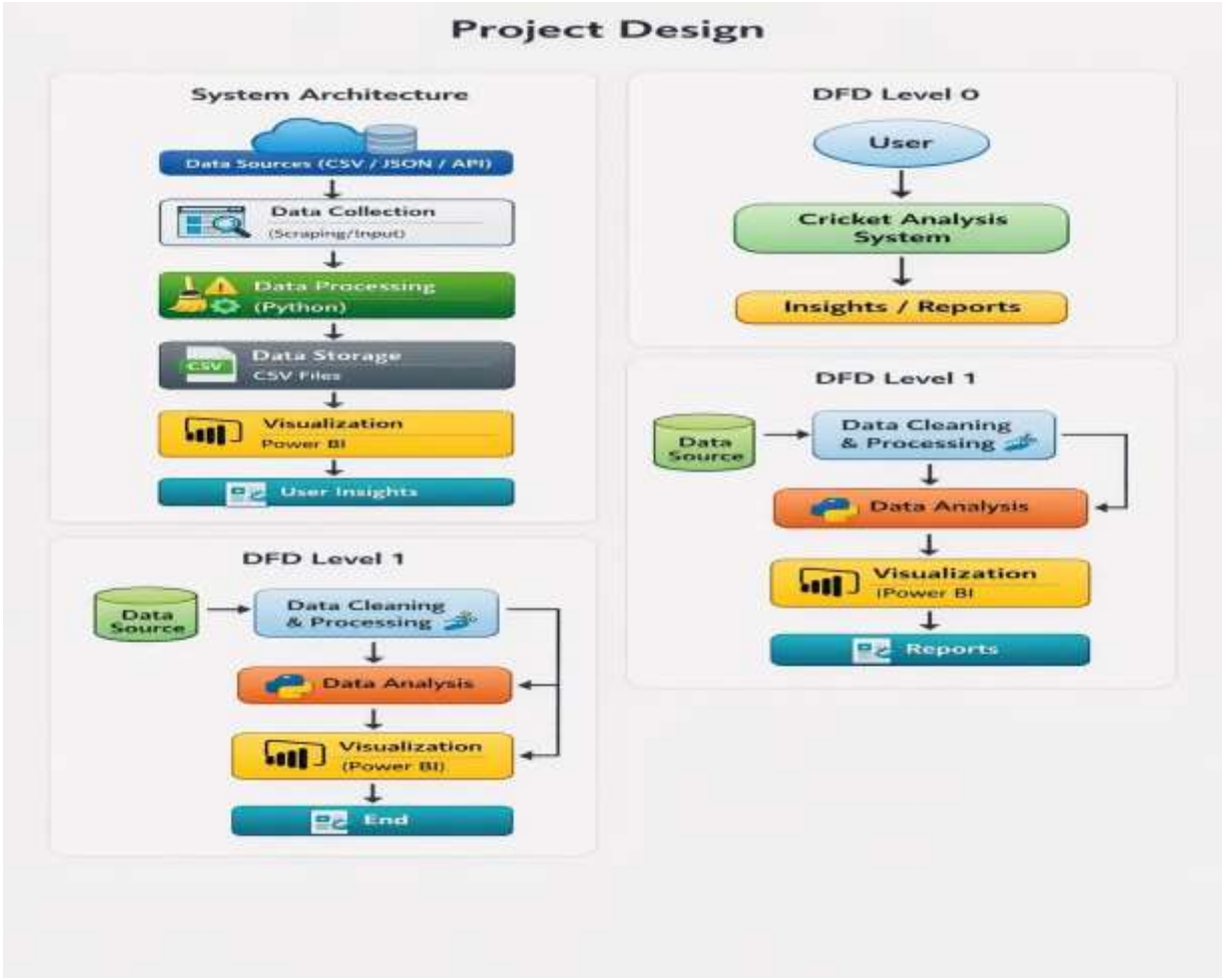
The project begins with collecting cricket datasets from reliable sources. The data is cleaned to remove errors and missing values. Python libraries like Pandas are used for preprocessing. Data transformation is performed to structure it properly. Statistical analysis is carried out to calculate key metrics. The processed data is imported into Power BI. Data modeling is performed to establish relationships. KPIs such as average, strike rate, and performance metrics are defined. Interactive dashboards are created using charts and graphs. Filters and slicers are added for dynamic exploration. Trend analysis is performed to identify patterns. The system is tested for accuracy and

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

usability. Finally, the dashboard is deployed for users. The project begins with collecting cricket datasets from reliable sources such as match records and player statistics. The raw data is inspected to understand its structure and identify missing or inconsistent values. Data cleaning is performed using Python programming language libraries like Pandas, where null values are handled and duplicate records are removed.

System Architecture

The system architecture consists of multiple layers for efficient data processing. The first layer is the Data Source Layer, which includes cricket datasets. The second layer is the Data Preprocessing Layer, where cleaning and transformation are performed. The third layer is the Data Storage Layer, where processed data is stored. The fourth layer is the Data Analysis Layer, where statistical analysis is conducted. The fifth layer is the Data Integration Layer, connecting Python and Power BI. The sixth layer is the Visualization Layer, where dashboards are created. The seventh layer is the User Interaction Layer, where users explore insights. The system supports scalability and flexibility. It ensures smooth data flow between layers. Real-time updates can be incorporated. Overall, it provides a complete analytics solution.



International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

Result and Output

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

V. Conclusion

The Cricket Statistics Analysis Project developed using and successfully demonstrates how data analytics can be applied to extract meaningful insights from cricket datasets. The project focuses on analyzing player performance, team statistics, and match outcomes, providing a comprehensive understanding of patterns and trends in the game. Through the use of Python, efficient data preprocessing, cleaning, and analysis were carried out, ensuring the accuracy and reliability of the dataset. The integration of Power BI enabled the creation of interactive and visually appealing dashboards, which made it easier to interpret complex data and present insights in a user-friendly manner. The project highlights key metrics such as batting performance, bowling efficiency, and team success rates, helping users to make informed observations.

The results of this project show that data visualization and analytical techniques play a crucial role in sports analytics. By transforming raw data into meaningful information, the system allows users to identify trends, compare performances, and gain deeper insights into cricket matches. It also demonstrates the importance of combining programming and visualization tools to build effective analytical solutions. In conclusion, the project not only fulfills its objective of analyzing cricket statistics but also serves as a foundation for further advancements in the field of sports analytics. With the integration of advanced technologies such as machine learning and real-time data processing, the system can be further enhanced to provide predictive insights and smarter decision-making capabilities.

References

1. Kumar, R. D., Prudhviraj, G., Vijay, K., Kumar, P. S., & Plugmann, P. (2024). Exploring COVID-19 through intensive investigation with supervised machine learning algorithm. In Handbook of Artificial Intelligence and Wearables (pp. 145-158). CRC Press.
2. Swathi, B., Vijay, K., Sushanth Babu, M., & Dinesh Kumar, R. (2024, November). Machine Learning Techniques in Cloud Based Intrusion Detection. In The International Conference on Artificial Intelligence and Smart Environment (pp. 557-564). Cham: Springer Nature Switzerland.
3. Sv satyakrishna, shirisha rangu ,bhargavi nalacheruve.(2024) Prospective investigation on colorectal cancer with SMOTE on machine learning Algorithm
4. Dr.G.Vishnu Murthy, BhargaviNalacheruve 1Professor, Department of computer Science & engineering, Anurag University, TS, India. 2Student, Department of computer Science & engineering, Anurag University, TS, India.
5. N. S. Manaswini, K. K, C. Nigam, S. S. Ali, R. Niranjana, and Suman, “Real-Time Object Detection in Drone Surveillance Using YOLOv5,” in Proc. 2025 3rd Int. Conf. IoT, Communication and Automation Technology (ICICAT), Gorakhpur, India, 2025, pp. 1–6, doi: 10.1109/ICICAT68430.2025.11414670.
6. B. Soundarya, V. N. S. Manaswini, M. Ayyakrishnan, R. D. Kumar, “Contextual Analysis of Big Data Analytics in Intelligent Transportation Frameworks,” in Intersection of Artificial Intelligence, Data Science, and Cutting-Edge Technologies: From Concepts to Applications in Smart

Environment, Lecture Notes in Networks and Systems, vol. 1353, Cham: Springer, 2025, doi: 10.1007/978-3-031-88304-0_79.
7. R. D. Kumar, V. N. S.Manaswini, "Applications of blockchain in smart cities: detecting fake documents from land records using blockchain technology," in Blockchain for Smart Cities, Elsevier, 2021, pp. 105–117, doi: 10.1016/B978-0-12-824446-3.00017-X.
8. Tejavath Veeramma, Badarla Anil, Guguloth Ravinder, "An advanced movie recommender using collaborative filtering and sentiment analysis," International Research Journal of Modernization in Engineering Technology and Science, vol. 7, no. 7, July 2025, doi: 10.56726/IRJMETS81618.
9. Ravi Kumar Banoth, Ramana Murthy B V, "Automatic crop recommendation system using LightGBM and decision tree machine learning models," Journal of Machine and Computing, vol. 5, no. 1, pp. 343, Jan. 2025, doi: 10.53759/7669/jmc202505026.
10. Ravi Kumar Banoth, Dr. B.V. Ramana Murthy, "Smart agriculture through IoT and machine learning for analyzing carbon footprints," in Proc. Int. Conf. Computer Science and Communication Engineering (ICCSCE), Apr. 2025.
11. Ravi Kumar Banoth, B. V. Ramana Murthy, "Soil image classification using transfer learning approach: MobileNetV2 with CNN," SN Computer Science, vol. 5, art. no. 199, 2024, doi: 10.1007/s42979-023-02500-x.
12. Purmani, S. S. R. (2025). Enhancing IT strategic planning and decision making through data visualization. International Journal of Enhanced Research in Management & Computer Applications, 14(4), 75–81
13. Mudusu, S. K. (2025). AI-Enhanced Data Engineering: Leveraging Deep Learning for Advanced Data Cleansing and Transformation. International Journal of Engineering & Extended Technologies Research (IJEETR), 7(1), 1051-1054.
14. Mudusu, S. K. (2024, August). Designing self-healing data pipelines for autonomous and continuous AI operations. Journal of Computational Analysis and Applications, 33(2), 1238–1247.
15. Gajula, S. (2025, December). Ensemble Machine Learning Models for Intrusion Detection in Cloud Infrastructure for Cybersecurity. In 2025 International Conference on Artificial Intelligence, Blockchain, Cloud Computing, and Data Analytics (ICoABCD) (pp. 1-6). IEEE.
16. Maturi, S. Y. (2022). Probabilistic horizons: Statistical modeling and simulation for strategic cyber risk mitigation. Journal of Information Systems Engineering and Management, 7(2).
17. P. Venkata Ramana. (2024). AI-driven predictive analytics in ERP systems for proactive supply chain optimization. Eudoxus Press Journal.
18. Pavan Kumar Adabala. (2026). Best Practices for Enterprise System Integration in Modern Organizations. Journal of Information Systems Engineering and Management, 11(2s), 1137–1146. https://doi.org/10.52783/jisem.v11i2s.14558
19. Gajula, S. (2025). Cloud transformation in financial services: A strategic framework for hybrid adoption and business continuity. International Journal of Scientific Research in Computer Science, Engineering and Information technology.

20. Majumder, R. Q. (2025). A Review of Anomaly Identification in Finance Frauds Using Machine Learning Systems. SSRN Electronic Journal. https://doi.org/10.2139/ssrn.5267287
21. Gaddam, S. Integrating Analytics into the Development Process: Bridging the Gap between Data Insights and Design Execution.
22. Reddy, S. K. R. Developing a Modular AI Framework to Enhance Scalability and Personalization in Next-Generation Reward Platforms.
23. Poojari, R. Frameworks for Data Management and Lineage in Large-Scale Healthcare Data Systems.
24. Purmani, S. S. R. (2025). Streamlining IT operations and service management with agile frameworks. European Journal of Advances in Engineering and Technology, 12(4), 76–81.
25. Viswanathan, V. (2024). Embedding Ethical Principles into Generative AI Workflows for Project Teams.
26. Mudusu, S. K. (2026, April 15). The secure intelligence framework: Architecting AI systems for a data-driven world. CIO (Foundry Expert Contributor Network).
27. Viswanathan, V. (2024). Pioneering Ethical AI Integration in Enterprise Workflows: A Framework for Scalable Team Governance. Available at SSRN 5375619.
28. Mudusu, S. K. (2025, June 3). Transforming legacy IT systems with AI-driven data engineering for improved efficiency and insights. Hampton Global Business Review (HGBR).
29. Gajula, S. (2026, March). Two Pillars of Banking Intelligence: A Comparative Analysis of AI Techniques for Fraud Prevention and Churn Mitigation. In 2026 14th International Symposium on Digital Forensics and Security (ISDFS) (pp. 1-6). IEEE.
30. Maturi, S. Y. (2023). Crowdsourced frontier: Unveiling autonomous adversarial cybercapabilities via open AI competition. International Journal of Intelligent Systems and Applications in Engineering, 11(1s), 275–284.
31. Chowdhury, A. K., Muhit, M. M. I., & Islam, M. M. (2023). A practical review to the marine maintenance practice in Bangladesh and a proposed way forward to an efficient, long-term and cost-effective solution. In Proceedings of the 13th International Conference on Marine Technology (MARTEC 2022). https://doi.org/10.2139/ssrn.4445071
32. Manoharan, D. (2025). Healthcare EDI Transaction Lifecycles Embedded with a Multi-Layer Verification Framework to Ensure Referential Integrity.
33. Ravishankara, M. (2026, February). CircuChain: Disentangling Competence and Compliance in LLM Circuit Analysis. In SoutheastCon 2026 (pp. 1-7). IEEE.
34. Doragacharla, V. R. (2023). Comprehensive Benchmarking Analysis of Auto Scaling Approaches in Cloud Native Streaming Pipelines During Flash Sales and Holiday Traffic Peaks. Available at SSRN 6566479.
35. P. Venkata Ramana. (2024). AI-driven predictive analytics in ERP systems for proactive supply chain optimization. International Journal of Innovative Engineering and Management Research (IJIEMR).
36. Kumar Adabala, P. (2021). Optimizing ERP Modernization: A Smart Data Migration Framework Approach. International Journal of Enhanced Research

in Science, Technology &Amp; Engineering, 10(07), 61–72. https://doi.org/10.55948/ijerste.2021.0708

37. Kavuri, S. (2025). Critical Review of Software Testing Problems in the Current Decade. International Journal on Science and Technology, 16(2). https://doi.org/10.71097/ijsat.v16.i2.9469
38. Srikanth Kavuri. (2024). Probabilistic Generative Modeling for Synthesizing High-Coverage Test Data in Safety-Critical Software Applications. Computer Fraud and Security, 633–642. https://doi.org/10.52710/cfs.838
39. Venkata Pavan Kumar Gummadi. (2024). API Design and Implementation: RAML and OpenAPI Specification. Journal of Electrical Systems, 16(4), 76–85. https://doi.org/10.52783/jes.9329
40. Venkata Pavan Kumar Gummadi. (2025). MuleSoft's Role in Advancing Sustainable Digital Infrastructure: An Enterprise Integration Perspective. Journal of Information Systems Engineering and Management, 10(62s), 1313–1321. https://doi.org/10.52783/jisem.v10i62s.13783