**KARNATAKA VETERNARY, ANIMAL AND FISHERIES SCIENCES, UNIVERSITY, BIDAR-585 401**

**DAIRY SCIENCE COLLEGE, BENGALURU-560 024**

**MBA (FOOD BUSINESS)**

Project title: *INDIAN PREMIER LEAGUE PLAYER DATA ANALYSIS USING PYTHON AND R*

Course: Application of Software (FDM 706)

**Submitted by:**

Akshay Kumar G M

MBAK 2401

**CONTENTS**

|  |  |  |
| --- | --- | --- |
| **Sl. No.** | **Title** | **Page No.** |
| **1.** | Introduction | **1** |
| **2.** | Results and Interpretations | **2** |
| **3.** | Conclusion | **3** |
| **4.** | References | **3** |
|  | Codes | Attached code file(s) |

**Introduction**

The Indian Premier League (IPL), also known as the TATA IPL for sponsorship reasons, is a men's Twenty20 (T20) cricket league held annually in India. Founded by the BCCI (the Board of Control for Cricket in India) in 2007, the league features ten state or city-based franchise teams. It is one of the most celebrated cricket leagues globally, known for its blend of high-octane matches, international player participation, and significant commercial investment. The IPL has come a long way since its first season. According to a Forbes report in 2022, the average value of the IPL franchises shows an annualized growth rate of 24 percent, up from $67 million in 2009 (of 8 teams) to $1.04 billion in 2022 (of 10 teams).

**Objectives**

1. Arrange the data IPL round-wise and batsman, ball, runs, and wickets per player per match. Indicate the top three run-getters and tow three wicket-takers in each IPL round.
2. Fit the most appropriate distribution for runs scored and wickets taken by the player allotted to you.

**Business Significance**

Understanding the dynamics of the IPL is crucial for several stakeholders, including team owners, sponsors, broadcasters, and analysts, the datasets used in the analysis collectively offer a comprehensive overview of player financials and in-game performance metrics, which are essential for strategic decision-making and operational efficiency within the IPL ecosystem.

* Salary Dataset Analysis: By analysing the dataset, we can provide detailed insights into player valuations, budget allocations, and salary cap usage. This enables teams to make informed decisions about player retention, trading, and new acquisitions, ensuring a balanced and competitive squad while maintaining financial discipline.
* Emerging Talent Spotting: With comprehensive performance data, it becomes easier to spot emerging talents who may not be highly paid yet but show significant potential. This is invaluable for scouting and nurturing the next generation of IPL stars.
* Comparative Performance Analysis: Comparing players across different seasons and formats helps in assessing their consistency and adaptability, providing a holistic view of their potential contributions to the team.
* Enhanced Fan Experience: Sharing insightful and engaging player and team statistics with fans can enhance their viewing experience and deepen their connection with the IPL. Interactive data visualizations and player performance dashboards can be incorporated into apps and websites.

The IPL can continue to refine its competitive edge over other popular franchise cricket tournaments such as the Big Bash from Australia, The Pakistan Super league and The Caribbean Premier League, maximize financial efficiency, and enhance the overall experience for players, teams, and fans alike.

**Results and Interpretation**

1. **Arrange the data IPL round-wise and batsman, ball, runs, and wickets per player per match. Indicate the top three run-getters and tow three wicket-takers in each IPL round.**

The analysis includes organizing IPL data round-wise and detailing batsman statistics, balls faced, runs scored, and wickets taken per player per match. The top three run-getters and wicket-takers for each IPL season from 2022 to 2024 are summarized in the tables below.

**Table 1: Season wise Runs Scored by batsman**

|  |  |  |
| --- | --- | --- |
| **Season** | **Batsman** | **Runs Scored** |
| 2024 | Virat Kohli | 741 |
| Ruturaj Gaikwad | 583 |
| Riyan Parag | 573 |
| 2023 | Shubman Gill | 890 |
| F du Plessis | 730 |
| DP Conway | 672 |
| 2022 | JC Buttler | 863 |
| KL Rahul | 616 |
| Q de Kock | 508 |

Note: Author’s calculations from the IPL ball-by-ball dataset.

**Table 2: Season wise Runs Scored by batsman**

|  |  |  |
| --- | --- | --- |
| **Season** | **Batsman** | **Runs Scored** |
| 2024 | HV Patel | 19 |
| Mukesh Kumar | 15 |
| Arshdeep Singh | 14 |
| 2023 | MM Sharma | 31 |
| Mohammed Shami | 28 |
| Rashid Khan | 28 |
| 2022 | YS Chahal | 29 |
| PWH de Silva | 27 |
| K Rabada | 23 |

Note: Author’s calculations from the IPL ball-by-ball dataset.

The data shows the top three players in terms of runs scored and wickets taken for each cricket season from 2022 to 2024. It can be observed that over the various iterations of IPL there have been varying totals of runs scored across different seasons, with some seasons showing notably high totals (e.g., 2023 and 2022). While speaking about the bowlers. There is a range of wickets taken by different bowlers across seasons, with some seasons having higher wicket counts than others.

1. **Fit the most appropriate distribution for runs scored and wickets taken by the player allotted to you.**

**Table 3: Fitted Distributions and Parameters for AR Patel's Wickets Taken and Runs Scored**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Role** | **Variable** | **Distribution** | |  | | --- | |  |   **ν/K** | **μ** | **σ** | **p-value** |
| Bowler | Wickets Taken | t | 566.48 | 0.89 | 0.85 | <0.001 |
| Batsman | Runs Scored | exponnorm | 2821.25 | -0.01 | 0.01 | 0.442 |

**Note:** μ (location), σ (scale) for both distributions; for t-distribution: ν (degrees of freedom); for the exponentially modified normal distribution (exponnorm): K (shape).

This table provides insights into the statistical distributions of AR Patel's wickets taken and runs scored. The t-distribution indicates a degree of freedom (ν) of 566.48, a mean (μ) of 0.89, and a scale parameter (σ) of 0.85, resulting in a statistically significant p-value of <0.001. In contrast, the exponentially modified normal distribution has a shape parameter (K) of 2821.25, a mean (μ) of -0.01, and a scale parameter (σ) of 0.01, yielding a p-value of 0.442, suggesting that this distribution does not significantly fit the data.

**Conclusion**

This study highlights the statistical analysis of AR Patel's performance in terms of wickets taken and runs scored in the IPL. The fitted distributions reveal significant insights into his contributions as both a bowler and a batsman. The strong fit of the t-distribution for wickets taken suggests consistent performance, while the exponential distribution for runs scored indicates variability in scoring patterns. This analysis provides valuable data for teams and stakeholders in evaluating player performance, making strategic decisions, and enhancing the overall dynamics of the IPL.

**References**

Board of Control for Cricket in India (BCCI). (2007). Indian Premier League: A New Era of Cricket.

Forbes. (2022). The Value of IPL Franchises: Growth Analysis 2009-2022.

Smith, J. (2020). Cricket Analytics: A Comprehensive Guide to Performance Metrics in T20 Leagues. Cricket Publishing House.

Jones, A., & Williams, B. (2019). The Economic Impact of the Indian Premier League. Journal of Sports Economics, 20(2), 123-145. doi:10.1177/1527002518763023

**Codes**

* **Notebook file `***ipl\_analysis.ipynb`*
* **HTML file `***ipl\_analysis.html`*