****

**VIRGINIA COMMONWEALTH UNIVERSITY**

**Statistical analysis and modelling (SCMA 632)**

**A1b:** **Analysis of IPL dataset - R**

**AMBALA GOWTHAM REDDY**

**V01107254**

**Date of Submission: 18-06-2024**

**CONTENTS**

|  |  |  |
| --- | --- | --- |
| **Sl. No.** | **Title** | **Page No.** |
| **1.** | Introduction | **1** |
| **2.** | Results | **2** |
| **3.** | Interpretations | **3** |
| **4.** | Recommendations | **4** |
| **5.** | Codes | **-** |
| **6.** | References | **5** |

**Introduction**

The Indian Premier League (IPL) is a professional Twenty20 cricket league in India, which has become one of the most popular and lucrative cricket leagues globally. Players in the IPL are compensated with substantial salaries, which are often reflective of their skills, experience, and marketability. However, the relationship between a player's salary and their on-field performance is a topic of ongoing interest for team management, stakeholders, and analysts.

This report aims to investigate the correlation between IPL players' performance metrics and their salaries for the year 2024. The primary objectives of this analysis are:

1. To calculate and summarize the total runs scored and wickets taken by each player in 2024:
   * Runs scored and wickets taken are essential indicators of a player’s contribution to the team. These metrics are crucial for evaluating a player's performance in both batting and bowling.
2. To merge the performance data with salary data and analyze the relationship:
   * By integrating performance data with salary information, we aim to identify any patterns or correlations that might exist between a player’s earnings and their on-field contributions.
3. To perform a specific analysis for Surya Kumar Yadav:
   * Surya Kumar Yadav, a prominent player in the IPL, is selected for a detailed individual analysis to provide insights into his specific performance metrics in relation to his salary.

The data utilized for this analysis includes:

* IPL Ball-by-Ball Data: This dataset contains detailed information on each ball bowled in IPL matches, including the date, striker, bowler, and runs scored.
* IPL Salary Data: This dataset provides the salary details of IPL players for the year 2024.

The methodology involves:

* Loading and processing the ball-by-ball and salary data.
* Calculating total runs scored and wickets taken by each player for the year 2024.
* Merging the performance metrics with the corresponding salary data.
* Analyzing the correlation between player salaries and their performance metrics.
* Conducting a specific analysis for Surya Kumar Yadav to highlight his contributions and salary.

**Results**

2.1 Data Processing

The IPL ball-by-ball data was loaded from IPL\_ball\_by\_ball\_updated till 2024.csv.

The IPL salary data was loaded from IPL SALARIES 2024.xlsx.

The year information was extracted from the match dates in the ball-by-ball data.

2.2 Performance Metrics

Total runs scored and wickets taken by each player in 2024 were calculated.

Data was filtered to include only the year 2024.

Runs and wickets data were merged into a single dataframe, filling any missing values with zero.

The total performance metric was created by summing runs scored and wickets taken.

2.3 Correlation Analysis

Player names were matched between the performance and salary data using a string distance method.

The dataframes were merged based on the matched player names.

The correlation between player salaries and their total performance was calculated.

2.4 Specific Player Analysis: Surya Kumar Yadav

Performance data for the player "SA Yadav" was extracted.

Salary data for "Surya Kumar Yadav" was extracted.

**Interpretations**

3.1 Correlation

The correlation coefficient between player salaries and their total performance (sum of runs scored and wickets taken) in the IPL 2024 season is calculated to be 0.4658806. This value provides several insights:

* Moderate Positive Relationship: The positive value indicates that there is a direct relationship between salary and performance. In other words, players who are paid higher salaries tend to perform better in terms of runs scored and wickets taken.
* Magnitude of Correlation: The correlation coefficient of 0.4658806 suggests a moderate relationship. While it is not a strong correlation, it is significant enough to imply that performance is a factor in determining player salaries. However, other factors likely contribute to salary decisions as well, such as player marketability, experience, and role within the team.
* Predictive Power: The moderate strength of the correlation means that while there is a trend of higher salaries being associated with better performance, it is not a definitive predictor. There are instances where players with high salaries may not perform as expected, and vice versa. This variability highlights the complexity of salary determination in the IPL.
* Implications for Team Management: For team management, this correlation suggests that investing in higher-paid players can generally lead to better on-field performance. However, it also emphasizes the need for a balanced approach that considers other attributes and not just performance metrics when negotiating salaries.

3.2 Specific Player Analysis: Surya Kumar Yadav

A focused analysis was conducted on Surya Kumar Yadav, a notable player in the IPL, to understand his performance in relation to his salary.

* Total Performance: Surya Kumar Yadav's total performance metric, combining both runs scored and wickets taken, is 176 for the year 2024. This performance highlights his significant contribution to his team's success in the league.
* Salary: His salary for the year 2024 is 800 (in the unit of currency used in the dataset, presumably lakhs or thousands of dollars).
* Alignment of Performance and Salary: The analysis shows that Surya Kumar Yadav's salary is relatively high, which aligns with his substantial on-field performance. This alignment suggests that his compensation is justified based on his contributions to the team. It reinforces the notion that key players who consistently perform well are rewarded with higher salaries.
* Value for Money: From a financial perspective, the return on investment for Surya Kumar Yadav appears to be positive. His high salary correlates with significant on-field performance, indicating good value for the team. Such players are critical assets for any team aiming for success in the IPL.
* Broader Context: While Surya Kumar Yadav's case aligns well with the overall trend observed in the correlation analysis, it is essential to consider that individual player evaluations should also factor in aspects such as consistency, injury history, and potential future performance. This holistic view ensures a well-rounded approach to salary and performance assessment.

In conclusion, the interpretations of the correlation analysis and the specific case study of Surya Kumar Yadav provide valuable insights into the relationship between player salaries and performance in the IPL. These insights can guide team management in making informed decisions regarding player contracts and salary negotiations, ultimately contributing to the team's strategic and financial success.

**4. Recommendations**

Based on the analysis, the following recommendations can be made:

1. **Performance-based Salary Adjustments**:
   * Consider implementing a more performance-based salary structure to incentivize players. Given the moderate correlation, salaries should be more closely tied to performance metrics to ensure fair compensation and motivation.
2. **Player Development Programs**:
   * Invest in player development programs to enhance performance, which could, in turn, justify higher salaries and improve overall team success.
3. **Regular Performance Reviews**:
   * Conduct regular performance reviews and adjust salaries based on up-to-date performance data to maintain a fair and competitive environment within the team.

**References:**

IPL\_Ball\_by\_Ball dataset

IPL\_Salaries dataset

ChatGPT