**BIS581 Business Data Analytics**

**–**

**Final Project: Cleaning, Visualization, Modelling, and Analysis of a Dataset using RStudio**

**Professor: Vishal Shah**

**E-Mail: Gampa2g@cmich.edu**

**Phone: (616)2747889**

**Semester: 4 Th**

**CMICH ID#: Gampa2g**

**Data Collection and Initial Setup**

* **Source and Structure**: The dataset was sourced from Kaggle, featuring comprehensive ball-by-ball data for each match played during the IPL 2021 season. The dataset includes detailed fields like innings, teams, players, and run specifics, which are pivotal for deep analysis.
* **Tools and Libraries**: Key R libraries used were **readr** for reading CSV files, **dplyr** for data manipulation, **ggplot2** for visualization, **tidyr** for data tidying, and **lubridate** for handling date formats. These tools were essential for efficient data preprocessing and analysis.

**Data Cleaning and Preprocessing**

* **Data Inspection**: Initial data exploration involved checking for completeness, outliers, and errors. This helped in understanding the dataset’s quality and determining the necessary cleaning steps.
* **Data Transformation**: Tasks included formatting date columns, splitting complex columns like **match\_name** into more usable components (**teams** and **match\_date**), and renaming columns to ensure consistency and clarity across the dataset.

**In-depth Analysis**

* **Performance Metrics Calculation**: Detailed calculations were performed for individual players and teams. For players, metrics like total runs, strike rate, and number of wickets were computed. For teams, overall wins, losses, and win rates were calculated.
* **Focus on Sunrisers Hyderabad**: Special attention was given to analyzing performance metrics specifically for Sunrisers Hyderabad. This included identifying key players, their contributions, and areas where the team could improve strategically to enhance their chances of winning.

**Predictive Modeling**

* **Logistic Regression Model**: This model was used to predict the outcomes of matches based on historical performance data. It helped in understanding the factors that most significantly impact match outcomes and provided a probability score for Sunrisers Hyderabad's chances against various teams.
* **Simulation of Match Outcomes**: By simulating different playing conditions and matchups, the model could predict how changes in team composition and strategy could potentially improve outcomes for Sunrisers Hyderabad.

**Strategic Recommendations**

* **Team Composition**: Based on the analysis, strategic recommendations were made regarding player selection and match tactics specifically for Sunrisers Hyderabad, aimed at maximizing their strengths and minimizing weaknesses.
* **Data-Driven Decisions**: The project provided insights that could help team strategists and coaches make informed decisions on player line-ups and game strategies, particularly focusing on critical matches that could define the season for Sunrisers Hyderabad.

**Visualization and Reporting**

* **Insightful Visuals**: Advanced graphical representations were used to depict trends, player performances, team comparisons, and predictive outcomes. These visuals were integral in presenting complex data in an understandable format, aiding strategic discussions.
* **Detailed Reports**: The project culminated in a comprehensive report that not only detailed the analyses and findings but also outlined actionable strategies tailored for Sunrisers Hyderabad, supported by data-driven evidence.

**Conclusion**

The project not only highlighted the analytical possibilities within sports data but also showcased how targeted data analysis could be used to strategically enhance team performance in competitive sports leagues like the IPL. The detailed approach, from data handling to strategic insights, specifically aimed at boosting Sunrisers Hyderabad’s performance, sets a foundation for future analytical projects in sports analytics.

This expanded detail provides a thorough understanding of how your project was structured and executed with the specific goal of enhancing the performance of Sunrisers Hyderabad. If there are any additional details or specific areas you would like further expanded, please let me know!