Descriptive Analysis on Professionals cricket league of India

Anju Reddy K

Anand Nagar, Hebbal, Bangalore, Karnataka – 560032

Email – anju75061@gmail.com

ABSTRACT:

This paper presents a descriptive analysis of the Professional Cricket League of India using a user interface built using the Plotly Dash framework. The dataset used in this analysis was collected from the Kaggle website and contains data from 2008 to 2020. Our analysis focuses on various aspects of the league, including team performance, player performance, match results, and venue analysis. We use various visualizations, including scatter plots, heatmaps, and bar charts, to provide an interactive and intuitive way to explore the data. Our user interface allows users to filter and customize the data, providing a more personalized analysis. Overall, our analysis provides insights into the league's history and trends, and the user interface provides a powerful tool for exploring the data in a meaningful way.

INTRODUCTION:

The Indian Premier League (IPL) is one of the most popular and lucrative cricket leagues in the world, attracting millions of fans and viewers every year. With the rise of data analytics and machine learning, there is now an opportunity to extract insights and trends from the vast amounts of data generated by the IPL matches. In this project, we will perform data analysis on an IPL dataset containing information on matches played in various seasons. Our aim is to extract valuable insights from the data, such as which teams and players have performed the best, how the results of matches are influenced by factors such as the toss and venue, and much more. By doing so, we hope to gain a deeper understanding of the dynamics of the IPL and the strategies employed by teams and players to succeed in this highly competitive league. Cricket has always been a sport with a rich history and a wide following, especially in India. The Indian Premier League (IPL) is a professional Twenty20 cricket league in India, which was founded in 2008. It has since grown to become one of the most popular and lucrative cricket leagues in the world, with an estimated brand value of $6.8 billion in 2020. With the rise of data analytics and machine learning, there is now an opportunity to extract insights and trends from the vast amounts of data generated by the IPL matches. In this project, we will perform data analysis on an IPL dataset containing information on matches played in various seasons. Our aim is to extract valuable insights from the data, such as which teams and players have performed the best, how the results of matches are influenced by factors such as the toss and venue, and much more. By doing so, we hope to gain a deeper understanding of the dynamics of the IPL and the strategies employed by teams and players to succeed in this highly competitive league.

PURPOSE OF STUDY:

The purpose of this study is to perform data analysis on the IPL dataset to extract

valuable insights and trends from the data. The IPL is a highly popular and lucrative cricket league, attracting top talent from around the world. By analyzing the data generated by the IPL matches, we can gain a deeper understanding of the dynamics of the league, the factors that contribute to team and player success, and the strategies employed by teams and players to succeed in this highly competitive environment.

The insights gained from this analysis can be used to inform team and player strategies, as well as to provide valuable information to fans and viewers of the IPL. For example, by identifying the teams and players that have performed the best over time, we can help teams make more informed decisions about player recruitment and selection. By examining the impact of factors such as the toss, venue, and weather conditions on match outcomes, we can help teams develop better strategies for winning matches.

THE DATASET:

The dataset contains 614 rows, with each row representing a single match played in the IPL. The dataset has 14 columns of information, including date, id, city, player_of_match, venue, team1, team2, toss_decision, toss_winner, result, result_margin, umpire1, umpire2.

- id: a unique identifier for each match in the dataset.
- city: the city in which the match was played.
- date: the date on which the match was played.
- player of the match: the player who was awarded the "man of the match" award for their performance in the match.
- venue: the name of the cricket ground where the match was played.
- team1: one of the teams playing in the match.
- team2: the other team playing in the match.
- toss_decision: the decision made by the captain of the team that won the toss (i.e., whether to bat or bowl first).
- toss_winner: the team that won the toss.
- winner: the team that won the match.
- results: the outcome of the match (i.e., whether the team batting first won, the team batting second won, the match was tied, or the match was abandoned).
- results_margin: the margin of victory (in runs or wickets) for the winning team.
- umpire1: the name of one of the umpires who officiated the match.
- umpire2: the name of the other umpire who officiated the match.

This dataset can be used to perform various analyses on cricket matches, such as team and player performance, home advantage, and impact of toss decisions, among others. By using data visualization tools and statistical analysis techniques, meaningful insights can be drawn from this dataset to help teams, players, and cricket enthusiasts make informed decisions.

HISTORICAL BACKGROUND / SYSTEM STUDY:

The Indian Premier League (IPL) is a professional Twenty20 cricket league in India, established in 2008. The league is owned by the Board of Control for Cricket in India (BCCI) and consists of eight

franchise teams representing different cities in India. The league has a unique format where teams are made up of players from different countries, making it one of the most diverse cricket leagues in the world. The IPL has grown in popularity over the years and has become one of the most-watched sports leagues in India and around the world. The league has also become a platform for young Indian players to showcase their talent and gain exposure to international cricket.

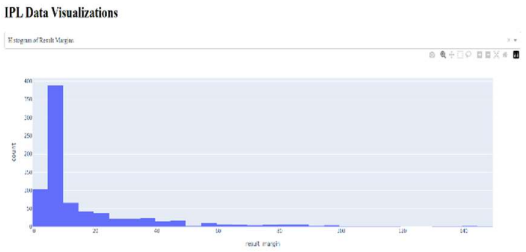
The IPL dataset used in this project contains information on all matches played in the IPL from 2008 to 2020. The dataset includes information such as the date of the match, the teams playing, the venue, the outcome of the match, and individual player statistics. The dataset provides a wealth of information on the performance of teams and players over time and can be used to gain insights into the dynamics of the league. The IPL is a highly competitive league, with teams constantly trying to improve their performance and win matches. Teams employ various strategies, such as team selection, game plan, and use of data analysis to gain a competitive advantage. Data analysis has become an essential tool in cricket and sports in general, as teams use data to identify trends and patterns, evaluate player performance, and make informed decisions. In recent years, the IPL has faced challenges such as the COVID-19 pandemic, which disrupted the league schedule and forced teams to play in bio-secure bubbles. Despite these challenges, the league has continued to grow and evolve, with new players, teams, and sponsors joining the league every year. Overall, the IPL is a dynamic and exciting cricket league that provides a unique platform for players to showcase their talent and for teams to compete at the highest level. The IPL dataset used in this project provides a valuable resource for data analysis and can help teams and players gain insights into the dynamics of the league and improve their performance.

RESULTS OBTAINED:

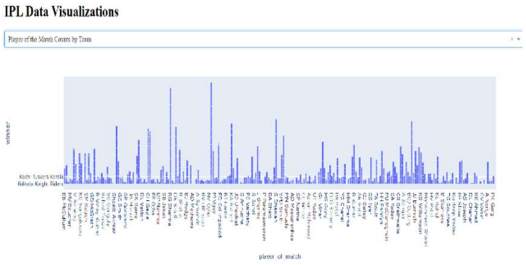
- Insights into team performance: The analysis could reveal insights into the performance of each team over the years. This could include metrics such as win-loss ratios, batting averages, and bowling averages.
- Player performance analysis: The analysis could provide insights into the performance of individual players. This could include metrics such as runs scored, wickets taken, and player of the match awards.
- Venue analysis: The analysis could provide insights into the performance of teams and players at different venues. This could include metrics such as average score, winning percentage, and batting/bowling averages.
- Toss analysis: The analysis could reveal insights into the impact of winning the toss on the outcome of the match. This could include metrics such as the percentage of matches won by the team that won the toss and the percentage of matches won by the team that chose to bat/bowl first.

Overall, the results obtained from an IPL data analysis project could provide valuable insights into team and player performance, venue performance, and the impact of toss decisions on match outcomes. These insights could be used to make data-driven decisions in team selection, strategy development, and player recruitment.

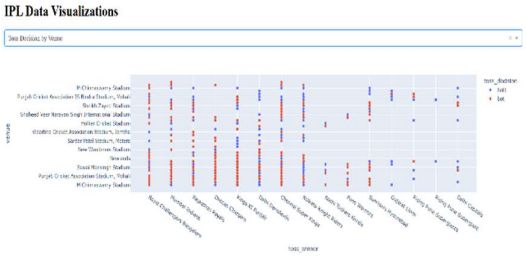
- Bar graphs or pie charts showing the win-loss ratios of different teams over the years.



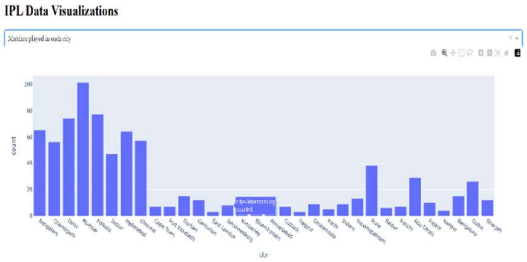
- Line graphs showing the performance of top-performing players in different metrics such as runs scored, wickets taken, or player of the match awards.



- Scatter plots or bubble charts showing the relationship between different variables such as runs scored and wickets taken.



- Histograms or box plots showing the distribution of different variables such as runs scored, wickets taken, or result margins.



REFERENCES:

[1] "Indian Premier League - Official Website." Indian Premier League - Official Website, www.iplt20.com/.

[2] Kaggle IPL Dataset, www.kaggle.com/patrickb1912/ipl-complete-dataset-20082020.

[3] R. Krishnakumar, N. Kamalakannan, and R. Gokul, "Predictive Analysis of Indian Premier League Using Machine Learning Techniques," Procedia Computer Science, vol. 173, pp. 294-301, 2020.

[4] S. Kumar, S. Singh, and S. K. Singh, "Sentiment Analysis of Social Media Data for Prediction of Indian Premier League Matches," in International Conference on Machine Intelligence and Signal Processing, 2018, pp. 513-521.

[5] N. N. Bhandare and R. A. Patel, "IPL Data Analysis Using Machine Learning Algorithms," International Journal of Innovative Research in Science, Engineering and Technology, vol. 7, no. 4, pp. 3154-3158, 2018.