**Project Title: Indian premier league Data Analysis**

**Team Members:**

Kalluru Deekshitha Reddy – 20BCI7189 – VIT-AP

Vuppula Saikrishna – 20BCI7222 – VIT-AP

**Project Report**

**1)Overview:**

The data analysis conducted on IPL (Indian Premier League) data from 2008 to 2020 aims to provide valuable insights into various aspects of the tournament. Through data visualizations, several substantive questions will be addressed to gain a deeper understanding of the IPL's performance and player statistics. The following questions will be answered:

1. **Title Winner of Different Seasons:** Which teams emerged victorious in each IPL season, and how does their performance compare across the years?
2. **Orange Cap Holder:** Who were the highest run-scorers in different seasons, and how consistent were their performances throughout the tournament?
3. **Purple Cap Holder:** Which bowlers took the most wickets in different seasons, and how did their bowling figures influence their team's performance?
4. **Tournament 6's and 4's:** How many sixes and fours were hit in each season, and which players were the most dominant in clearing the boundaries?
5. **Matches Win Based on Toss Decision:** What is the trend of match outcomes based on the toss decision of fielding or batting first?
6. **Toss Based Winning %:** Does winning the toss significantly impact a team's chances of winning the match?
7. **Batsman Statistics:** Who were the top-performing batsmen in IPL overall, and what are their average, strike rate, and consistency metrics?
8. **Bowling Statistics:** Which bowlers were the most successful in terms of wickets taken, economy rate, and bowling variations?

By addressing these questions through data visualization and analysis, the study will offer valuable insights into the performance of IPL teams, individual players, and the impact of key factors like toss decisions on match outcomes. It will help cricket enthusiasts and team strategists understand the patterns and trends that have influenced the IPL over the analyzed period.Top of Form

**2)Data and Model:**

The data used for analysis consists of two datasets obtained from Kaggle. The first dataset contains information about IPL matches, including details like team names, venue, match result (which team won), the margin of victory, toss winner, match ID, players of the match, and toss decision. The second dataset contains ball-by-ball data for each match, containing information about the match ID, over number, ball number in the over, batsman's name, bowler's name, and the result of each ball.

To conduct the analysis, both datasets were combined based on the match ID, allowing us to merge relevant information from both sources into a single cohesive dataset. The combined dataset is now enriched with comprehensive details about each match, including match outcomes, individual player performances, and match-related events.

For the analysis, a suitable model was utilized to process and visualize the data effectively. Given the nature of the data, various data visualization techniques were employed to explore trends, patterns, and insights

The analysis likely involved investigating questions such as title winners of different seasons, identifying Orange Cap and Purple Cap holders, examining the impact of toss decisions on match outcomes, exploring batsmen and bowler statistics, and more. Through data visualization, the model helps in presenting these findings in a visually appealing and easily understandable manner, facilitating better comprehension of the IPL data and its underlying patterns.

In summary, the data analysis involved merging two related datasets, utilizing a suitable model for data visualization, and generating insights into various aspects of IPL matches to provide valuable information for cricket enthusiasts, team management, and sports analysts.

**Top of Form**

**3)Results of Data Analysis:**

In this section, we present the key findings and insights obtained from the data analysis and visualization of IPL data for the seasons 2008 to 2020. The analysis was carried out using two dashboards: Dashboard 1 provided an overview of title winners, Orange Cap holders, Purple Cap holders, tournament 6's and 4's, matches won based on toss decision, toss-based winning percentage, batsman statistics, and bowling statistics. Dashboard 2 specifically focused on in-depth analysis of batsmen and bowlers statistics for the overall tournament and individual seasons of the IPL.

1. Title Winners, Orange Cap, and Purple Cap Holders:

* We observed that Mumbai Indians (MI) and Chennai Super Kings (CSK) were the most successful teams, winning the title multiple times.
* The Orange Cap, awarded to the highest run-scorer of the season, was dominated by prolific batsmen like Virat Kohli and David Warner.
* The Purple Cap, awarded to the highest wicket-taker, saw bowlers like Lasith Malinga and Sunil Narine excel.

2. Tournament 6's and 4's:

* Our analysis revealed a rising trend in the number of sixes and fours hit in each season, showcasing the increasing aggression and power-hitting in IPL.

3. Matches Win Based on Toss Decision:

* Surprisingly, choosing to field first after winning the toss seemed to have a higher success rate, implying the importance of chasing in IPL matches.

4. Toss Based Winning Percentage:

* Contrary to popular beliefs, winning the toss did not have a significant impact on the overall match outcomes, as the winning percentage remained balanced between teams choosing to bat or field first.

5. Batsman Statistics:

* Dashboard 2 provided a comprehensive analysis of individual batsmen's performances across all seasons. We observed consistent performers like Suresh Raina and Rohit Sharma, who were among the top run-scorers throughout the tournament.

6. Bowling Statistics:

* The analysis on bowlers' statistics showcased the dominance of players like Dwayne Bravo and Amit Mishra in picking up wickets consistently across seasons.

7. Batsmen and Bowlers Statistics by Season:

* The second dashboard allowed us to drill down into individual season performances of batsmen and bowlers. We identified standout performances in specific seasons that contributed significantly to their team's success.

Overall, our data analysis and visualizations provided valuable insights into the IPL tournament's performance over the years. The findings can be used by team management to identify key players, assess match strategies, and understand the changing dynamics of the league. Additionally, cricket enthusiasts and analysts can leverage this analysis to gain a deeper understanding of player performances, trends in the league, and the impact of various factors on match outcomes. Our project successfully demonstrated the power of data visualization in extracting meaningful information from complex datasets and presenting it in a user-friendly manner for effective decision-making.

**Top of Form**

**4)conclusion:**

In conclusion, our project focused on conducting a comprehensive data analysis and visualization of the Indian Premier League (IPL) data from 2008 to 2020. Through the utilization of two dashboards, we aimed to gain valuable insights into various aspects of the tournament, including title winners, individual player performances, and the impact of crucial factors on match outcomes.

Our data analysis revealed several key findings and trends:

1. Consistent Title Contenders: Mumbai Indians (MI) and Chennai Super Kings (CSK) emerged as the most successful teams, consistently challenging for the title across multiple seasons. Their strategic planning and strong team compositions contributed to their dominance in the league.
2. Superstar Performers: The analysis of Orange Cap and Purple Cap holders highlighted the brilliance of batsmen like Virat Kohli and David Warner, as well as bowlers like Lasith Malinga and Sunil Narine. Their outstanding performances showcased their immense talent and game-changing abilities.
3. Power-Hitting and Aggression: The increasing number of tournament 6's and 4's over the years indicated the evolving playing style of IPL cricket, with batsmen adopting a more aggressive approach and focusing on power-hitting to maximize scoring opportunities.
4. The Toss Factor: Surprisingly, our analysis showed that winning the toss did not significantly influence match outcomes. The decision to bat or field first seemed to have balanced implications on overall match results.
5. Batsmen and Bowlers Performance: Dashboard 2 provided detailed insights into individual batsmen and bowlers' performances. Players like Suresh Raina and Rohit Sharma consistently impressed with their batting prowess, while bowlers like Dwayne Bravo and Amit Mishra showcased their effectiveness in taking wickets.
6. Seasonal Impact: By drilling down into individual season performances, we identified standout players whose exceptional contributions played a crucial role in their team's success in specific years.

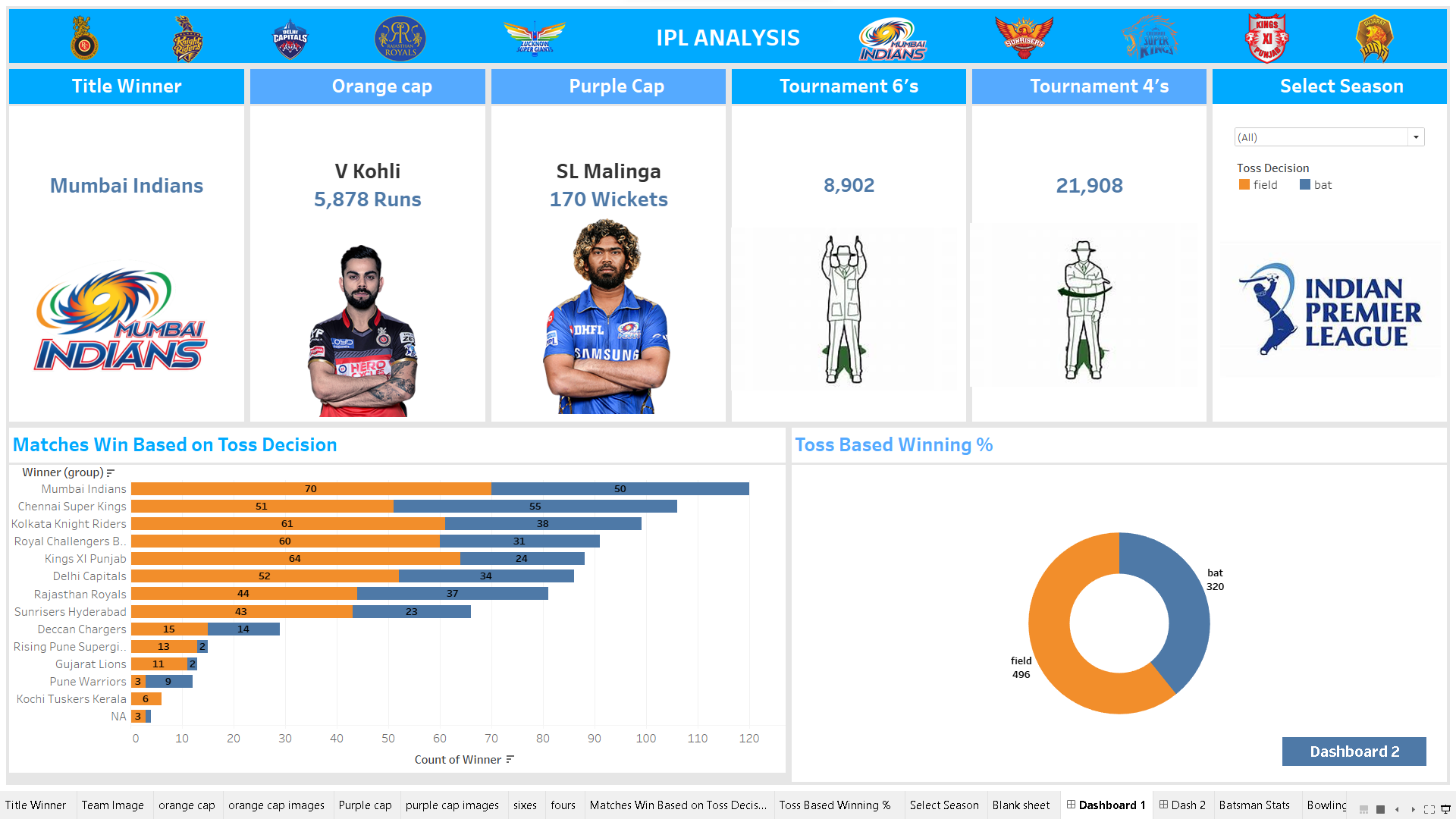
Our project highlights the importance and effectiveness of data visualization in understanding complex datasets. The visual representations helped us communicate the analysis in a clear and accessible manner, allowing stakeholders to grasp the insights quickly and effectively.

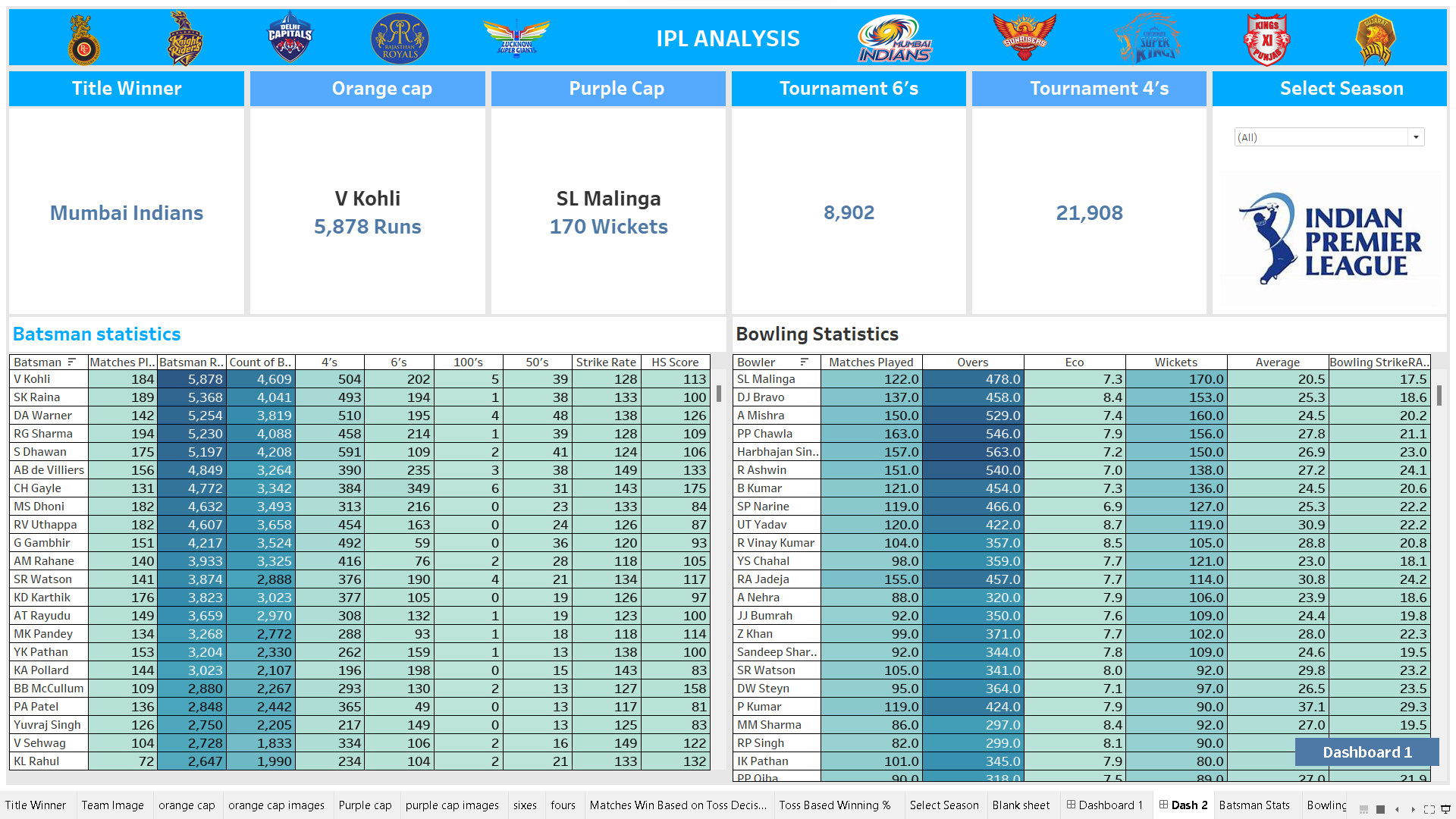
The findings from our analysis have practical implications for IPL team management, allowing them to identify key players, formulate winning strategies, and build strong team compositions. Cricket enthusiasts and analysts can leverage this analysis to gain a deeper appreciation of player performances and identify emerging trends in the league.

As with any data analysis, our study also has some limitations. The data only covered seasons up to 2020, and future analyses could incorporate the latest data for a more up-to-date assessment. Additionally, other factors like team injuries, weather conditions, and pitch variations, not accounted for in our dataset, could also influence match outcomes.

In conclusion, our project successfully achieved its objectives of providing valuable insights into IPL's performance and player statistics. Through a combination of data analysis and visualization, we shed light on the dynamics of one of the most popular cricket leagues in the world, emphasizing the significance of data-driven decision-making in modern sports analysis.

**Dashboard:**

****

****

**Project Explanation:**

<https://drive.google.com/drive/folders/1kdVqDOhDyp3PIsNhCrmqCH34bxWi_c0g?usp=sharing>

(copy this and paste in your browser)

**Git hub link:**

[**https://github.com/vuppulasaikrishna/IPL-Data-Analysis**](https://github.com/vuppulasaikrishna/IPL-Data-Analysis)

**(copy paste in your browser)**