Data analysis and visualization of Indian Premier League(IPL) matches

Sandeep Khandelwal
skhande@iu.edu

Abhishek Gupta
abhigupt@iu.edu

ABSTRACT

This exploratory analysis will help us visualize the best players and team in previous IPL matches. There will also be a breakdown on whether home pitch was favorable in terms of wins. It will also depict how the players performed and eventually led to a match winning effort. For example if a team has best players but they never performed together in a single match. If they all play well, it will result in a match win. Hence, these facts can be used to predict the winners in upcoming match. Other analysis can be done are players by number of run, wickets, maximum number of six's, four's etc. Man of the match by each season. Average runs scored by each team over by over. It may also have analysis on poorly performing players for each team. Poor players can be based on their bowling or batting scores.

KEYWORDS

ipl, analysis, python, packages, bowling, batting, fielding, match strategy

1 INTRODUCTION

We would like to analyze IPL match for last decade for all IPL matches played. This analysis will be done on ball by ball data avaiable from previous IPL matches. Some important visualization we plan to create per match season are

- Top batsman - Player who had maximum run
- Top bowler - Player who had maximum wicket
- Player by maximum man of match - Player who had maximum man of the match
- Winner team - Team who won the season
- Player by max six's - Player who made maximum six
- Player by maximum fours's - Player who made maximum four
- Top Team by match venue - Team who won maximum match per venue

These are key visualizations we plan to make but may not be limited to these visualization and may have more added to the list above.

2 TECHNOLOGY

We plan to use python as programming language and use D3 or Plotly as charting library. We plan to incorporate all charts in a jupyter notebook. We may also use other python modules like matplotlib and pandas if needed.

3 DATA SOURCES

We will be using datasources from IPL [1] official website and Kaggle challenge [2].

4 RELATED WORK

- iplt20 website - [1] This website display insights related to each IPL match. It also shows statistics per year by batting and bowling as well as all time records but doesn't analyze how players performed and contributed to winning match. If fact it doesn't give analysis of data which can be used to infer a winning team based on previous match performance. It also doesn't show good visualizations rather shows just tabular view of all reports.
- Reflecting Against Perception: Data Analysis of IPL Batsman [3]. It talks about analyzing performance of an IPL batsman comparing in different formats of the game like test, one-day. Its just talks about performance of batsman and doesn't consider any other aspects of the IPL series.

5 FURTHER ENHANCEMENT

Future enhancements can be done in-terms of building predictive model to predict the winning team based on the venue of the match, player who are going to play in the match etc.

6 ACKNOWLEDGEMENTS

The authors thank Prof. YY Ahn for his technical guidance. The authors would also like to thank TAs of Data Visualization class for their valued support.

7 REPO

All project and report document can be found at github project.

REFERENCES

[1] https://www.iplt20.com. IPL T20, official website. Web Page. Accessed: 2018-05-26.
[2] Manas Garg - Kaggle. Indian premier league (cricket) ball-by-ball cricket data. Web Page. Accessed: 2018-05-26.
[3] Amit Kumar and Ritu Sindhu. Reflecting against perception: Data analysis of ipl batsman. 2014.