**Final Project Proposal**

*Ojaas H, Meet V, Vijeet S*

**Dataset:**

Indian Premier League 2008-2019 (Ball by Ball Dataset)

**Description of the Dataset:**

There are two files in this dataset namely deliveries.csv and matches.csv

Matches.csv includes –

* Id
* Season
* City
* Team 1
* Team 2
* Toss\_winner
* Toss\_decision
* Winner
* Win\_by\_runs
* Win\_by\_wickets

Deliveries.csv includes –

* Match\_id
* Inning
* Batting\_team
* Bowling\_team
* Total\_runs

**Research Questions:**

1. Basic insights into the data to reveal total runs scored in each season, total wickets taken in each season, performance of teams by seasons.
2. What is the relation between winning the toss and winning the match, based on the factor like season, city and the participating teams? This basically answers the question whether winning the toss has any advantages in winning the match.
3. Finding out if home advantage is a real thing in the IPL considering data across the seasons.