Cricket Analytics

Business Requirement Document

- **A problem statement**

  The owners and the management of Mumbai Indians want to identify the best possible players suitable for them for the next auction. They have hired you as an Analyst to help them identify the best IPL team for the 2022 auction. You need to analyze the IPL data from 2008 to 2020, and perform the following tasks:

  1. Rank batsmen by year and overall based on ability
  2. Rank bowlers by year and overall based on ability
  3. Rank the most valuable player in IPL over all the years
  4. Visualize this in an intuitive Power BI Dashboards
  5. Find Similar Batsmen and Bowler

- **Project objectives**
  - Data Preprocessing & Exploratory Data Analysis
  - Complete the stats creation of KPIs for batsmen, bowler, team and all the matches. Also set up Power BI and import summarized data
  - Create the dashboard and publish it to be shared with the public

- **Project Scope**
  - Create a dashboard to show player-wise charts and analysis against a particular player or a team or aggregate.
  - Noting the KPIs for Batsman, Bowler and Team to find out the value of that player for the team.
  - We will also find which players have the quality of the required skill set.
  - Filtering the players performance according to a season/year.
  - Win and toss analysis of a particular team (Win or lose and option choose after winning or losing the toss)

- **Out of scope**
  - Not taking fielders and wicket keepers into consideration
  - We can not take retirement or injury of a player into consideration

- **Dataset Details**
  - **id:** Unique number for Matches
  - **inning:** Inning attribute (1 or 2)
  - **over:** Over count (0 to 19)
  - **batsman:** Batsman on the Strike
  - **non_striker:** Batsman on the Non-strike
  - **bowler:** Bowler bowling that particular over
  - **batsman_runs** : runs scored by batsman (0 to 6)
  - **extra_runs:** Extra runs scored in that ball
  - **total_runs:** runs scored by batsman + extra runs

- **non_boundary**: Boundaries (4s or 6s)
- **is_wicket :** 0 - not out / 1 - out
- **dismissal_kind :** Type of dismissal (Wicket)
- **Player_dismissed:** Name of Dismissed Player
- **fielder :** fielder's name

- **Key Performance indicators**
  - Batsman
    - Runs against team or player = $Count\ (runs\ for\ a\ batsman)\ :\ if\ player/team/total$
    - Strike rate = $\frac{Runs\ Scored\ by\ a\ Batsman}{Total\ Balls}\ x\ 100$
    - Average number balls taken to hit a Boundary = $\frac{Count\ of\ Boundaries\ hit\ by\ a\ batsman}{Total\ number\ of\ balls\ faced\ by\ the\ batsman}$
    - Average number balls taken to hit a Six = $\frac{Count\ of\ Sixes\ hit\ by\ a\ batsman}{Total\ number\ of\ balls\ faced\ by\ the\ batsman}$
    - %age of times dismissed (Out/Wicket) = $\frac{No.\ of\ times\ dismissed}{No.\ of\ times\ played}\ x\ 100$
    - Most Runs scored - total or against a team/player = $Max\ (No.\ of\ score\ against\ a\ team\ or\ player)$
    - Least Runs scored - total or against a team/player = $Min\ (No.\ of\ score\ against\ a\ team\ or\ player)$
    - Average Runs scored - total or against a team/player = $\frac{Total\ no.\ of\ runs\ made\ by\ a\ Batsman}{Total\ no.\ of\ matches\ played\ by\ the\ batsman}$
  - Bowler
    - Wickets = $Count\ wickets\ \ for\ a\ bowler)\ :\ if\ player/team/total$
    - Economy = $\frac{Total\ no.\ of\ Runs}{Total\ number\ of\ Overs\ Played\ (6\ Balls\ each)}$
    - Strike Rate (Runs per Ball) =

$$\frac{Total\ no.\ of\ Runs}{Total\ number\ of\ Balls}$$

- Maiden Overs
- Average
- Hattricks

- Teams:
  - %age win = $\frac{Total\ no.\ of\ matches\ won}{Total\ no.\ of\ matches\ played}\ x\ 100$
  - No of toss won/lost = $\frac{Total\ no.\ of\ toss\ won\ or\ lost}{Total\ no.\ of\ matches\ played}\ x\ 100$
  - Highest/lowest Score = $Max\ (total\ score\ by\ a\ team\ in\ an\ inning)$
  - %age win against a particular team = $\frac{Total\ no.\ of\ matches\ won\ against\ a\ team}{Total\ no.\ of\ matches\ played\ against\ a\ team}\ x\ 100$