INDIAN PREMIER LEAGUE
SEASONS OVERVIEW

Basic Information

Members:

Sai Krishna Dontaraju - u1426879
Raahul Vallurupalli - u1419533
Tirumal Reddy Ramidi - u1414933

Github Repository:
https://github.com/tirumalramidi/IPL_SeasonOverview.git

Background and Motivation

All three of us grew up in the Indian subcontinent, where cricket is considered a religion. Likewise, cricket is emotionally connected to us in many ways. However, there isn't a website that shows the stats of the teams and players intuitively. We wanted to bridge that gap by choosing this project and further helping the people who watch cricket. Additionally, we also wanted our visualization to be relatable and understandable to any user, even without technical experience. For instance, cricbuzz is used more than espncricinfo for cricket information, be it live scores or archival information, because cricbuzz is easier to use and understand than espncricinfo for naive users.

Project Objectives

Primary questions we intend to answer:

1. The winner and the league finish of every IPL season from 2009 until 2022.
2. The league table position after every game week.
3. The scorecard overview of the specific match played upon selection.
4. The team and team players' details upon selection on the given map.
5. Displaying the league positions of every year on default before choosing the year

Through this project, we intend to intuitively help several cricket watchers perceive and understand the statistics of the teams, players, and matches. Additionally, as a team, we aim to learn how to visualize complex data, select the best features to represent, visualize metrics, data mining, and different viewpoints. Furthermore, working on this project could further help us understand the emphasis on teamwork, leadership, and creative thinking.

Data

The dataset we considered until this point for the project is https://www.kaggle.com/datasets/rajsengo/indian-premier-league-ipl-all-seasons. Additionally, we need to scrape data from cricbuzz.com and espncricinfo.com for the league positions of the teams after every game week. The dataset has data from 2009 up to 2022. The dataset contains the league position of the teams for every year, the batting and bowling statistics of players of all years, and the summary of all matches of every year. Again, by scraping the data from the mentioned websites, we intend to get the league positions of every team after the game week. Alternatively, we could use an API for the same, and one such example is CricAPI.

Data Preprocessing

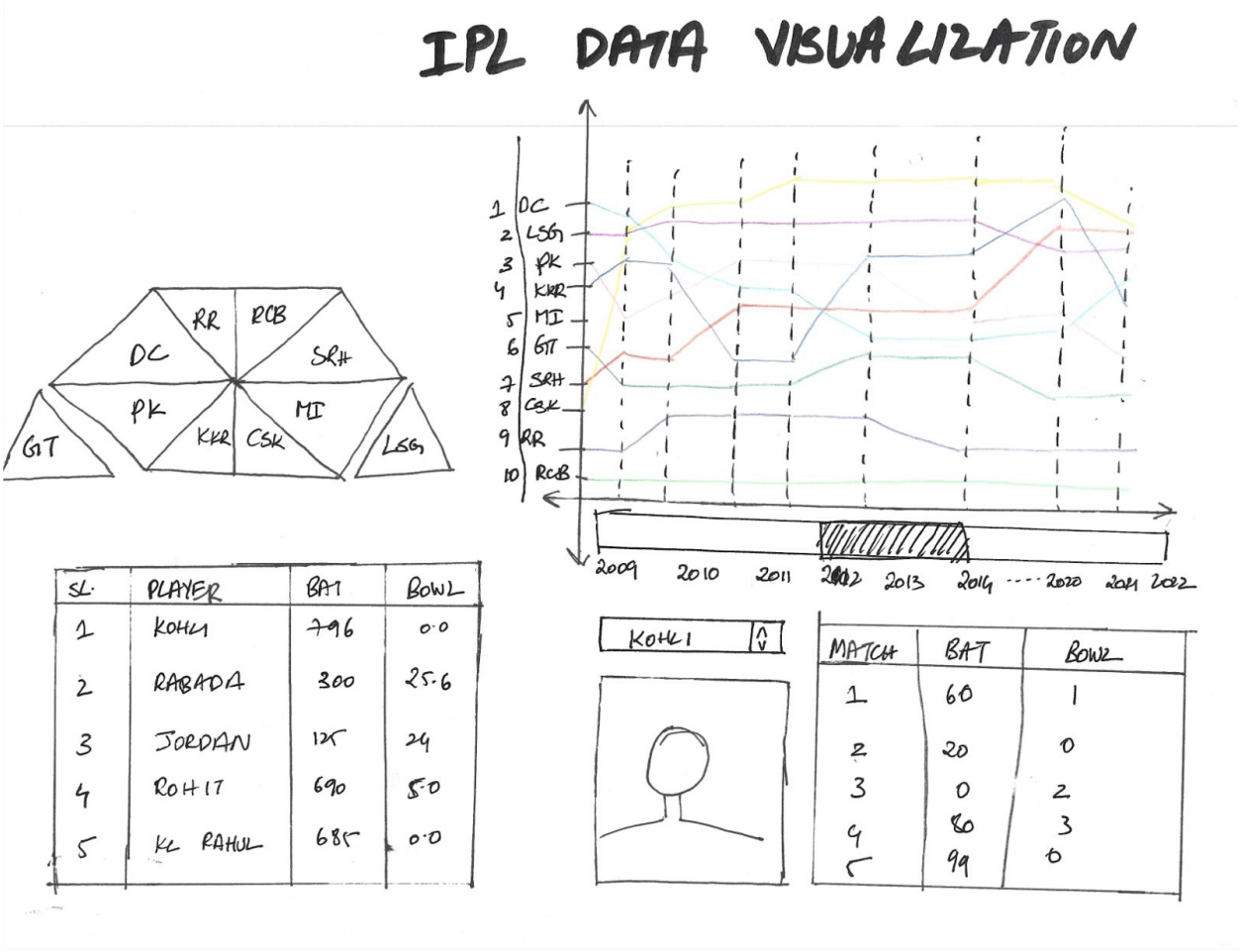
The data will be exported as CSV from the Kaggle dataset mentioned. The data is clean, with not many missing values. The dataset has 0.2% missing values, and there are several ways to tackle missing values in a dataset:

1. Replacing with an arbitrary value
2. Replacing with the mean or mode
3. Dropping the missing values row altogether

Here, we intend to drop the missing values rows altogether. Additionally, there are many columns that are trivial to the visualization. For the sake of making the project uncomplicated to understand from a neutral perspective, we intend to remove the columns altogether. Furthermore, we intend to add more columns to the dataset to solve the problem of the dataset not having the league positions of every team after every game week.

Visualization Design

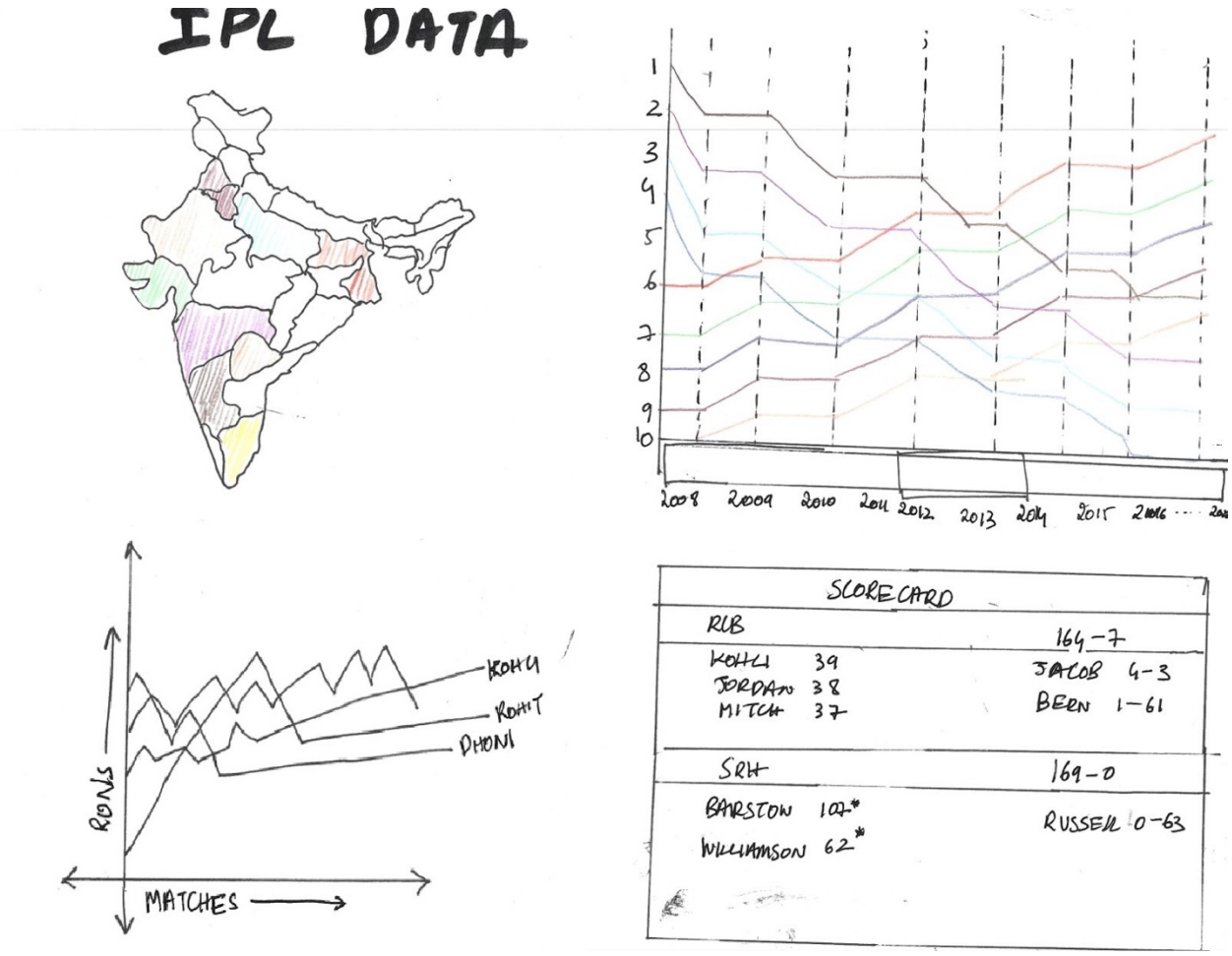
Prototype 1



In this prototype design, we are using a polygon where each tile corresponds to one team. Clicking on any of the tile updates data in other visualizations and shows a detailed view of the data of any team. A table visualization will be displayed for that team with various measurements. We also included a dropdown menu to select a player from a particular team to view his statistics.

We can select any tile on the polygon to choose any team to display corresponding statistics in various visualizations. We have a dropdown menu to drill down into any player's data.

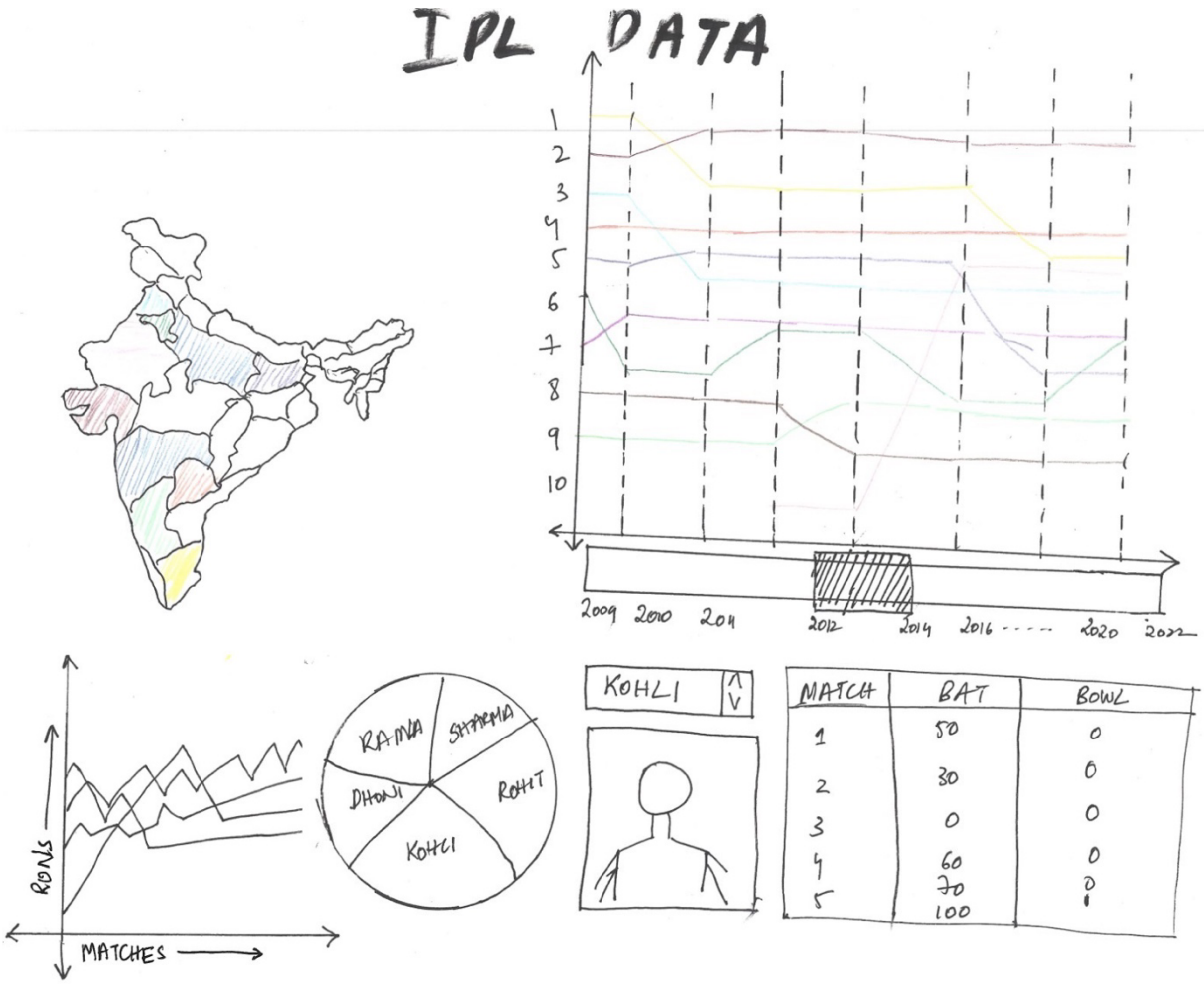
Prototype 2



In this prototype design, we replaced the polygon with a geographical map. It will have specific portions highlighted with different colors. These portions represent all teams in the league in different seasons in the tournament's history since its inception. Users can select a state from the country map, highlighting the corresponding bar and lines. Line chart visualization will be displayed for that team with various measurements.

Selecting any state representing a team highlighted in the map, highlighting corresponding lines in the adjacent line chart and table illustrating the related team data, and pie charts also should represent that particular team's player's statistics.

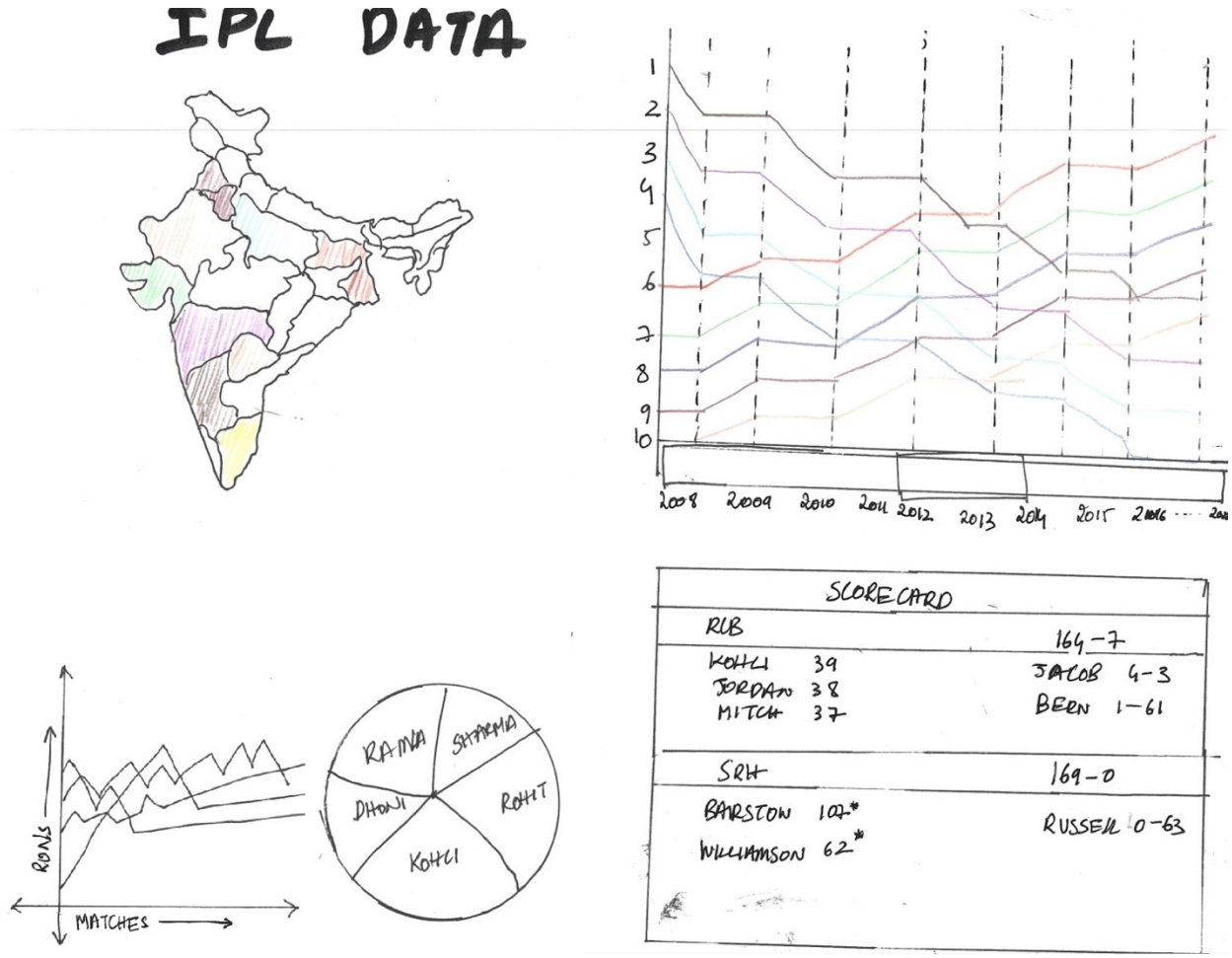
Prototype 3



In this prototype design, we replaced the line chart data of the players with more charts. This would enable the user to have more information at his hand while viewing the visualizations. Again, searching the name of the player would show the statistics of the player and change the rest of the visualizations dynamically.

Selecting any state representing a team highlighted in the map, highlighting corresponding lines in the adjacent line chart and table illustrating the related team data, and pie charts also should mean that particular team's player's statistics.

Final Design



In the final design, a geographical map of India is used where every portion represents a state, and every state further represents a team. Upon the selection of the state, the other visualizations are further updated dynamically. Upon the selection of a particular team on the map, the main chart highlights the league positions of the team after every game week. Several charts are used to display the statistics of the team; for instance,

1. Highest run scorer of the team
2. Highest wicket-taker of the team
3. Most catches taken in the team
4. Most fours and sixes hit in the team

This enables the user to perceive the information even when the viewer is naive. Additionally, upon selecting any node on the main chart, the specific match overview is displayed to the user. The match overview predominantly consists of the two best batsmen and the two best bowlers from each team, and the team which won and by how many runs or wickets.

To summarize, there are 4 main visualizations to display:

1. Geographical Map i.e., India Map - Displays the teams via states of the country.
2. Main Chart - Displays the league position of the teams after every game week.
3. Statistical Charts - Displays the statistics of the teams and the players after the specific teams are selected.
4. Match Overview - Displays the match overview after the node is selected on the Main Chart

Moreover, all the four visualization parts are interactive and serve the purpose of helping the user understand the information. The geographical map allows the user to categorize based on the state and, therefore, the teams. The nodes on the main chart are interactive, through which the match overview and the team and player statistics are displayed dynamically.

Must-Have Features

1. Overlay of the Indian Map representing the corresponding states of each team.
2. Overview of the timeline-based statistics of all teams during the current season.
3. High-level representation of the statistics of players.
4. When clicked on a particular team, corresponding team statistics must be displayed in the visualizations, such as max run scorers for a team and top wicket takers of the group.
5. Match overview portion that displays the match overview of the match upon the selection of the node in the main chart.

6. A dropdown box with every year from 2009 to 2022, where the visualization updates dynamically based on the year upon selection.

Optional Features

1. Displaying all members of a particular team in one specific season along with their statistics.
2. Toggle between a detail and a summary view.
3. Displaying the complete points table after every game week.
4. Radial range selector.

Project Schedule

Week 1: October 29 – November 5
Pre-processing dataset

Week 2: November 6 - November 11
Create and populate the country's map view

November 11: Project Milestone Due

Week 3: November 12 - November 18
Implementing the overall season's statistics visualization
Discussing with the project TA or mentor to finalize the statistics view

Week 4: November 19 - November 25
Implementing the individual team's statistics views

Week 5: November 25 - December 2
Wrapping up things, attempting optional features
Preparing for the final project presentation

December 2: Final Project Presentation