

CS699 Project Report

ICC Men's Ranking Report Generator

Advisor: Professor Bhaskaran Raman

1 Objective:

The objective of this project is to automate the process of collecting and analyzing cricket player rankings from the International Cricket Council (ICC) website, creating a comprehensive report in LaTeX format.

2 Group Members

SL No.	Name	Roll No.
1	Arijeet De	23M0742
2	A Asish	23M0759

3 Dependencies

- requests (2.31.0)
- beautifulsoup4 (4.12.2)
- pandas (2.1.1)
- matplotlib (3.8.0)
- pdflatex (0.1.3)

4 Overview

Cricket enthusiasts often seek up-to-date information on player rankings to understand the current standings in the cricketing world. This project aims to simplify this process by automating the retrieval of player rankings from the ICC website and generating a detailed report with insightful visualizations.

Steps involved in this project:

- Fetching Current Date
- Web Scraping
- Merging CSV Files
- Data Plotting
- PDF Report Generation

4.1 Fetching Current Date

The current day, month and year is retrieved:

- For keeping track of csv files
- To use as date inside the latex file

4.2 Web Scraping

The project leverages web scraping techniques to extract the latest ICC Men's rankings directly from the official ICC website. The data is collected from total '9 webpages' as shown below:

- [T20i Batting Ranking Data](#)
- [T20i Bowling Ranking Data](#)
- [T20i All-Rounder Ranking Data](#)
- [ODI Batting Ranking Data](#)
- [ODI Bowling Ranking Data](#)
- [ODI All-Rounder Ranking Data](#)
- [Test Batting Ranking Data](#)
- [Test Bowling Ranking Data](#)
- [Test All-Rounder Ranking Data](#)

The Data collected is stored inside a csv folder and data from each webpage is in separate csv files.

- t20i_batting_DD-MM-YYYY.csv
- t20i_bowling_DD-MM-YYYY.csv
- t20i_all-rounder_DD-MM-YYYY.csv
- odi_batting_DD-MM-YYYY.csv
- odi_bowling_DD-MM-YYYY.csv
- odi_all-rounder_DD-MM-YYYY.csv
- test_batting_DD-MM-YYYY.csv
- test_bowling_DD-MM-YYYY.csv
- test_all-rounder_DD-MM-YYYY.csv

Note: Day, Month and Year are also specified to represent the date of web scraping.

4.2.1 About The Data

Column names and their description:

- Pos: Rank of Player
- Player: Name of Player
- Team: Team the Player belongs to
- Rating: Rating of Player (out of 1000)
- Career Best Rating: Career Best Rating of Player (out of 1000)

4.3 Merging CSV Files

For each $\langle cricket_format \rangle$ (t20i, odi, test) we have 3 csv files:

- $\langle cricket_format \rangle_batting_DD-MM-YYYY.csv$
- $\langle cricket_format \rangle_bowling_DD-MM-YYYY.csv$
- $\langle cricket_format \rangle_all-rounder_DD-MM-YYYY.csv$

For each $\langle cricket_format \rangle$ we are merging these csv files into one. So we will have 3 merged csv files:

- overall_odi_DD-MM-YYYY.csv
- overall_t20i_DD-MM-YYYY.csv
- overall_test_DD-MM-YYYY.csv

So in total we now have '12 csv files'.

4.4 Data Plotting

Using the csv files bar plots are created. From each csv file 2 plots are created.

- **Total Players Plot:** Number of Players for each Cricket Team
- **Average Rating Plot:** Average Rating of Players for each Cricket Team

So in total we now have ‘24 bar plots‘.

4.5 PDF Report Generation

Using ‘pdflatex‘ command the DataReport.tex file was compiled and converted to a pdf file. The latex file is compiled twice so that the table of contents gets displayed properly. The pdf file gets generated in the output folder along with some auxiliary files.

5 Conclusion

This project can be leveraged for ongoing analysis by consistently gathering data and subsequently conducting comprehensive evaluations to gain insights into players’ long-term performance trends.