

# Final project proposal

World Cup is happening in the recent months. And I would like to make my project focus on the matches, teams, and players in this year's World Cup. The project will help people unfamiliar with soccer to gather information about the matches happening.

I will scrap one or more of the soccer websites. Obtain the data of matches, teams, and players.

For the scraping process, I will obtain the raw HTML string form websites. And I will utilize some HTML parser (maybe beautiful soup) to get the python friendly data (like json).

After gathering all the data required, my program will show the matches recently. They could be finished, ongoing or in the future.

Also, the program will allow the user to have more information about the matches shown.

For example, user can enter string to make different selections, which made browsing information of teams or players available.

I plan to use a tree liked data structure to store all the data of matches, teams, and players.

In addition, different player can be related through the same football club that they attended during their career.

The below graph shows the tree liked structure planned to use.

