Assignment 1: Visualisation in 7PAM2000 Applied Data Science 1

Rohith Changayil Pallathu

SRN: 22079600

Repository Link: https://github.com/RVdude/ADS_Assignment.git

Introduction

There are different types of visualisation methods which can be used for programming among which only three methods are used here. The "pyplot" function is used to plot all graphs, and the "pandas" tool is used to read datasets. Detailed description of data points is provided for each section.

Dataset

The dataset was acquired from the "Kaggle" database. Since the primary focus of this project is to explore visualization methods, the authenticity of these datasets has not been verified. FIFA 23 stands as a football simulation video game published by Electronic Arts, constituting the 30th and final instalment in the FIFA series. Developed by EA Sports, it was globally released on September 30, 2022, across various such as PC, PlayStation 5, Xbox.
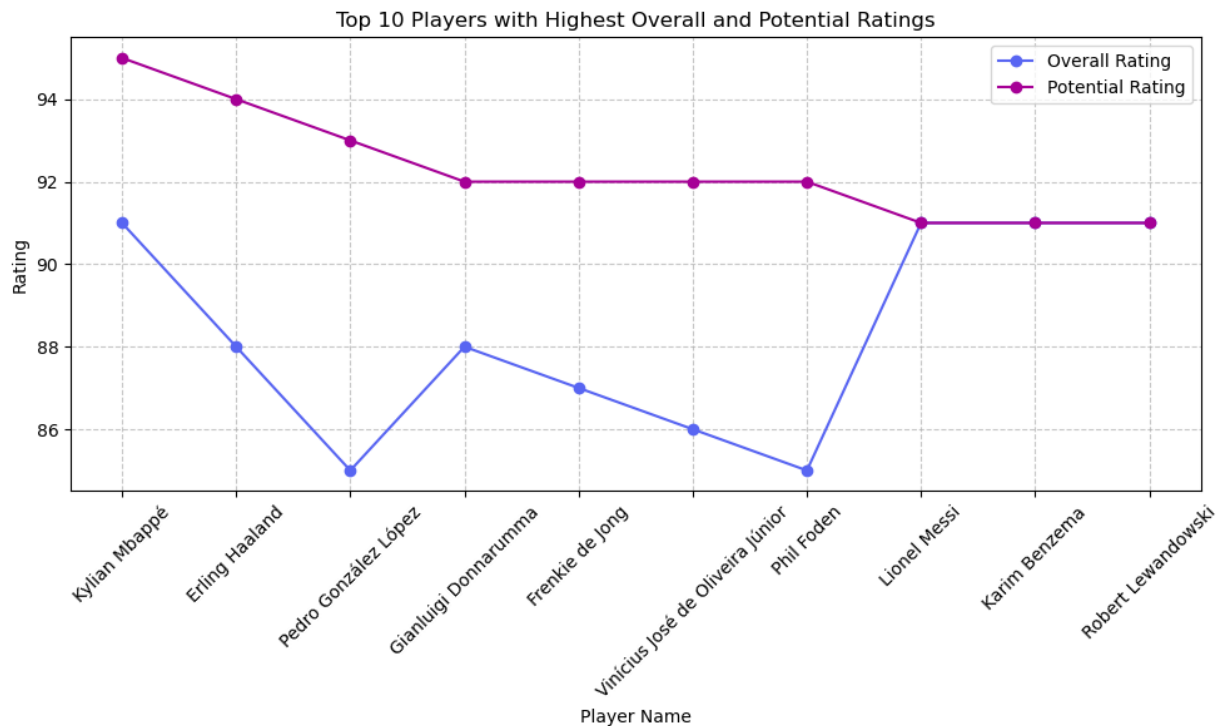
In this realm of football, lot of factors can be considered for the analysis of the given data set such as skills, ranking, age, potential ratings, overall rating, Nationality etc. As there is lot of data available, I have only considered potential and overall ratings of the players, nationality and age categorization for better analysis and graphical representation.

Dataset is taken from the following link: https://www.kaggle.com/datasets/bryanb/fifa-player-stats-database?select=FIFA23_official_data.csv

Line Plot

"pyplot" function is used to produce this line plot. The line graph shows the top 10 players with highest overall and potential ratings. Two graphs are plotted here. The blue line represents the overall rating, and the purple line represents the potential rating.

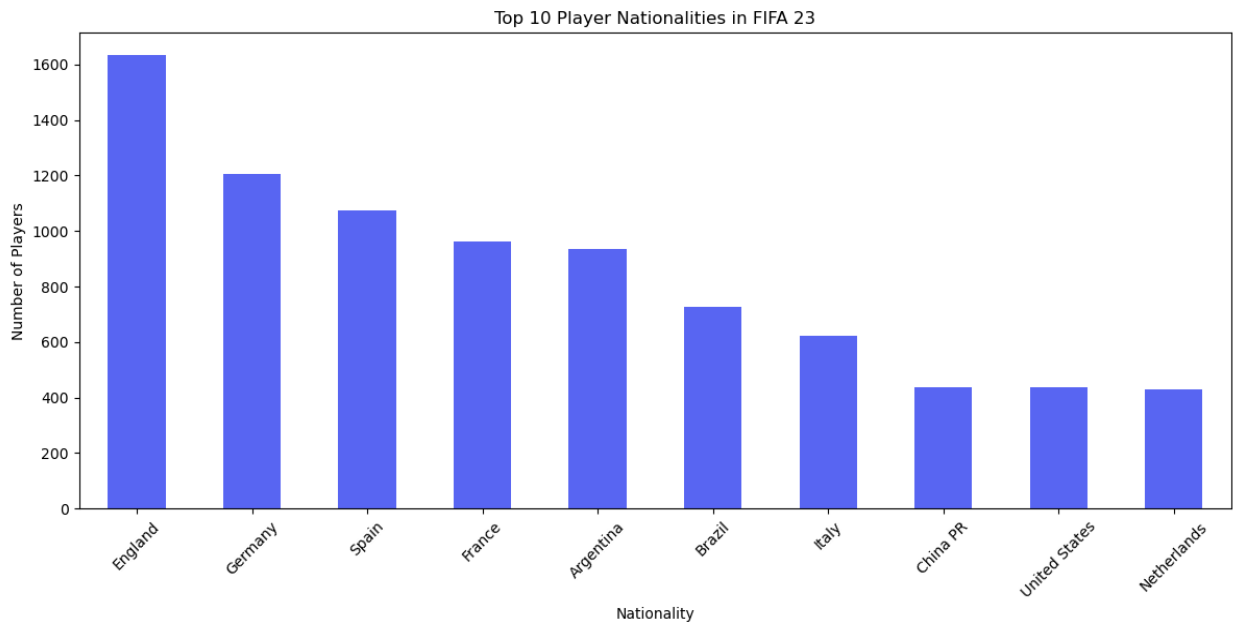Top 10 Players with Highest Overall and Potential Ratings

Kylian Mbappe is the player with highest overall and potential rating. Lionel Messi, Karim Benzema and Robert Lewandowski are the players with lowest potential rating but has high overall rating. Pedro Gonzalez Lopez, Phil Foden are the two players with lowest overall rating but has high potential rating than three players. Gianluigi Donnarumma, Frenkie de jong, Vinicius jose de Oliveira junior and Phil Foden ha the same potential rating but the overall rate has decreased respectively. Erling Haaland and Gianluigi Donnarumma shares the same overall rating.

Bar graph

This bar graph is produced by using pyplot function.
In this graph, nationality of the top 10 players in FIFA 23 is considered. That is, X-axis indicates the Nationality of the players and Y-axis indicates the number of players.
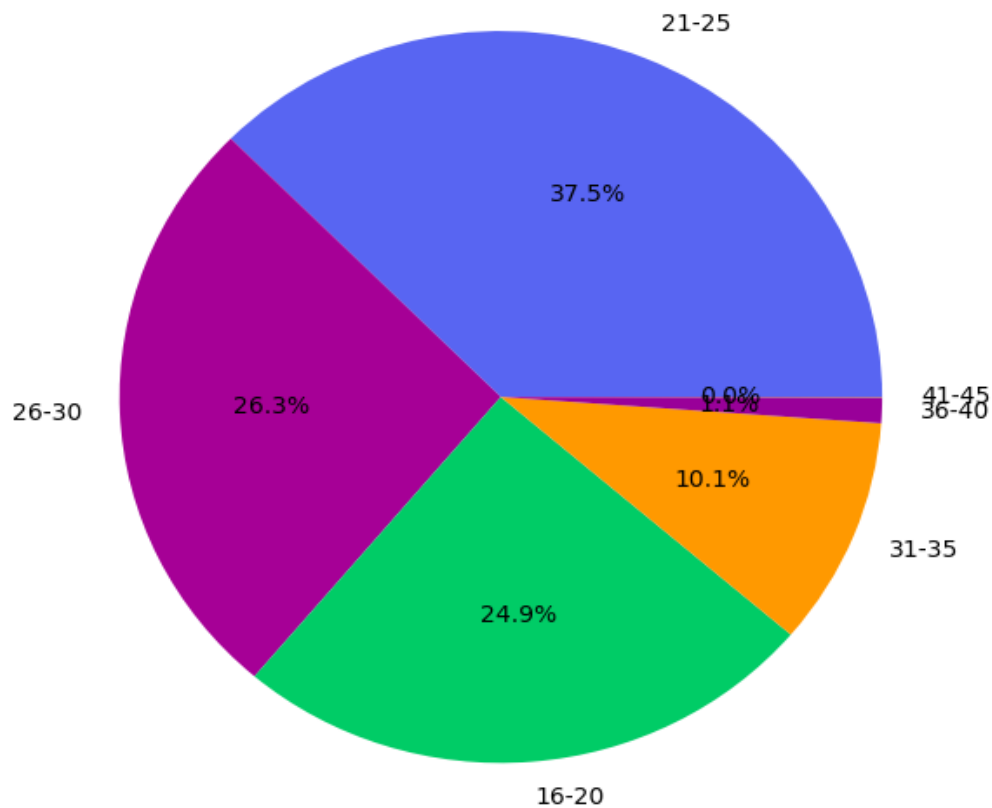
From the graph, the country with the highest number of players is England. Netherlands is the country with lowest number of players in FIFA 23. Even though Netherlands has the least number of players, China and United states are very close to Netherlands in numbers. Germany has the second highest number of players followed by Spain and France. Argentina comes on fifth position in this ranking but has higher number of players than Brazil and Italy. The graph gives better and clear representation of these data.

Pie chart

This pie chart shows the Age distribution of Players in FIFA 23. The age groups of the players are sliced into five categories. That are, 16-20, 21-25, 26-30, 31-35, 36-40 and 41-45.

Player Age Distribution in FIFA 23

The age group 21-25 holds the highest numbers of players. The are no players in the age group 41-45. The age group with least number of players is 36-40 with 1.1%. The youngest age category is 16-20 and 24.9% of the players are from this age category. 26-30 age category has 26.3% of the players and 31-35 contains 10.1% players.