FootballMetric.AI

1. Introduction

Football (soccer) is the most popular sport in the world, and as the game evolves, so does the need for intelligent, data-driven analysis. Behind every pass, goal, and tackle lies a story told by numbers — and the ability to interpret those numbers brings a whole new dimension to both fans and professionals alike.

Football is no longer just a sport — it's a data-driven battlefield.

FootballMetric.AI is a comprehensive football analytics project powered by Python, machine learning, and visual storytelling. From comparing players and analyzing positional performance to building predictive models and discovering tactical insights, this project shows how data science can transform our understanding of the beautiful game.

2.General Exploratory Data Analysis

The dataset consists of 2,782 raw records and 267 columns, containing a wide range of variables. These include demographic, behavioral, and contextual features relevant to the analysis. To simplify the exploration and reduce complexity, a light version of the dataset with 160 selected columns was also prepared. This reduced version focuses on the most relevant or high-quality features, ensuring more efficient analysis while preserving the essential structure and information.

During this initial exploratory phase, we aim to:

- Understand the structure and types of data present
- Identify missing values and anomalies
- Examine basic statistical summaries (mean, median, standard deviation, etc.)
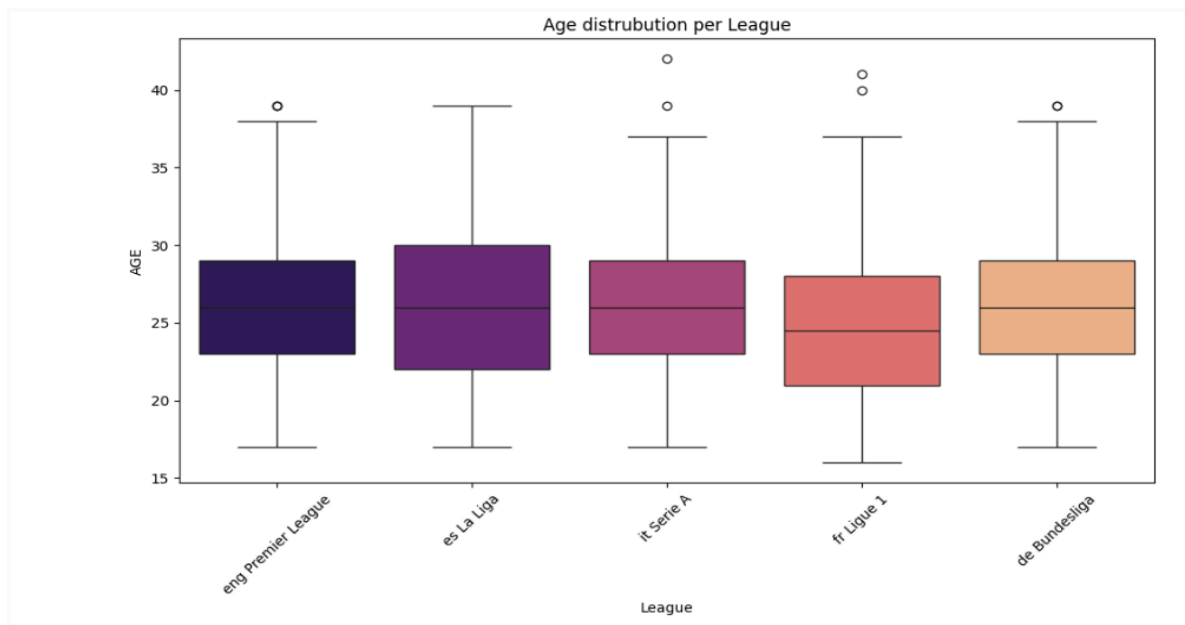- Visualize distributions and relationships between key variables

3.Comparative analysis of football leagues

A comparative study between the top 5 leagues in the world can shows the points of power of each league by it's player and the profile of it's players

1. Age comparaison (youth players) :

The French Ligue 1 shows an interesting number of youth players with 270 players under the average age (25 y.o) and 153 under 22 y.o . (less than 75% of players).

In the other hand the spanish league La Liga has more experienced players with 305 player above the average age (>25 y.o) and 147 player over 29 y.o (older than 75% of players)

1.Age distribution per league plot (boxplot)

Youngest and oldest players in each league :

```
Youngest Players:
                  Comp                 Player   Age Pos
962         de Bundesliga          Viggo Gebel  17.0  MF
1694  eng Premier League      Shumaira Mheuka   17.0  MF
307           es La Liga           Marc Bernal  17.0  MF
1805          fr Ligue 1  Djylian N'Guessan     16.0  FW
50            it Serie A         Honest Ahanor  17.0  DF

Oldest Players:
                  Comp                 Player   Age Pos
1491        de Bundesliga    Anthony Losilla   39.0  MF
850   eng Premier League   Łukasz Fabiański    39.0  GK
75            es La Liga          Raúl Albiol   39.0  DF
649           fr Ligue 1               Dante   41.0  DF
2138          it Serie A           Pepe Reina   42.0  GK
```
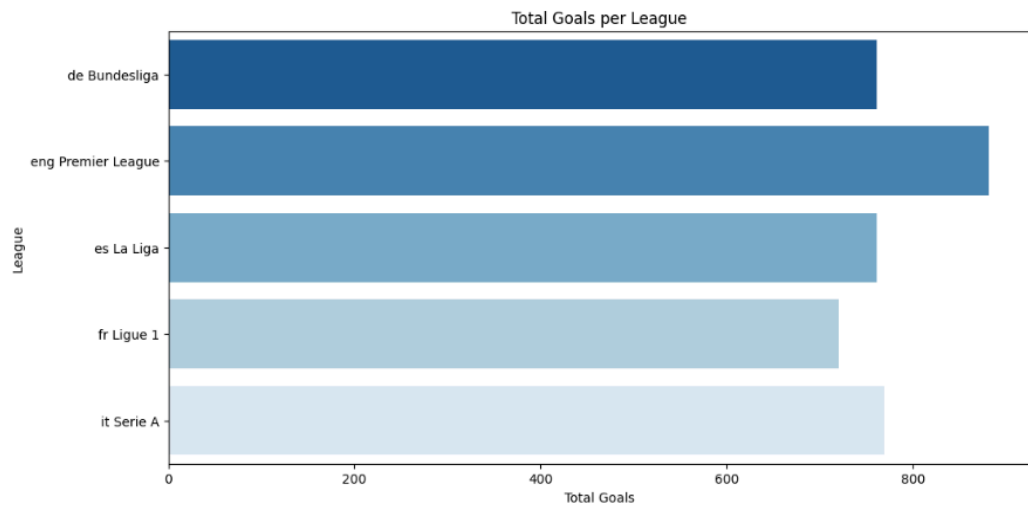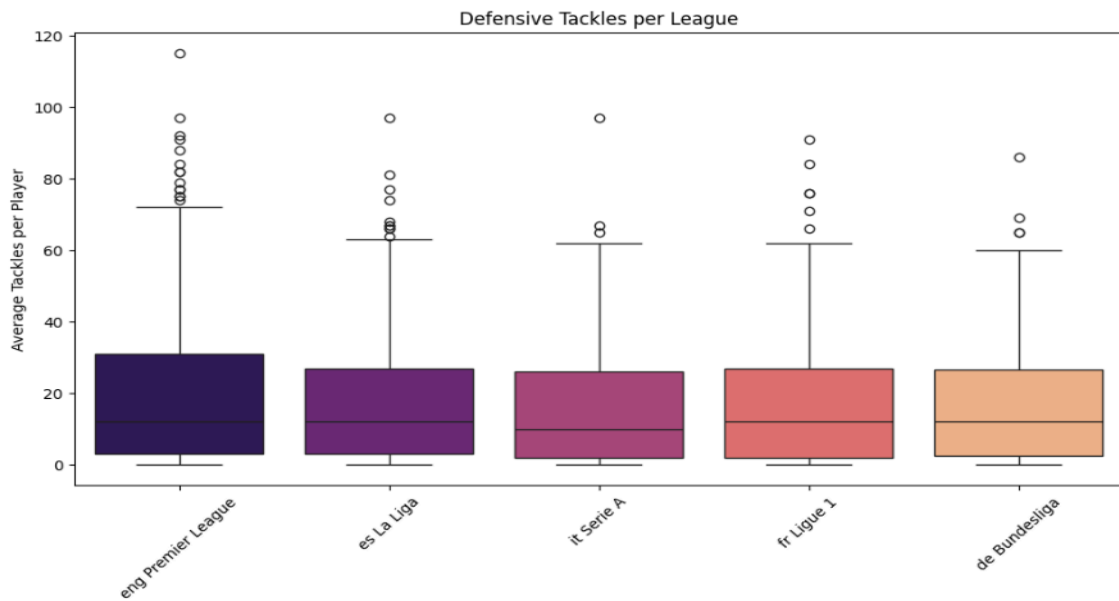
2. Performance, Passing, and Defense Summary

The English Premier League leads the race when it comes to performance showing the good quality of its players .

2.Total Goals per league plot

4.Defensive Tackles per league plot (boxplot)

3. Players :

Top players from each league :

```
Top Scorers:
                  Comp              Player  Gls   Age Pos
1287        de Bundesliga       Harry Kane   23  31.0  FW
2243  eng Premier League    Mohamed Salah   27  32.0  FW
1449          es La Liga  Robert Lewandowski   25  36.0  FW
683           fr Ligue 1   Ousmane Dembélé   21  27.0  FW
2145          it Serie A    Mateo Retegui   22  25.0  FW

Top Assist Difference Players:
                  Comp              Player  A-xAG   Age     Pos
116         de Bundesliga   Mohamed Amoura    4.6  24.0      FW
2166  eng Premier League  Antonee Robinson    6.0  27.0      DF
1858          es La Liga      Saúl Ñíguez    3.6  30.0      MF
1843          fr Ligue 1       João Neves    4.1  20.0   MF,DF
1502          it Serie A    Romelu Lukaku    4.9  31.0      FW

Top xG Players:
                  Comp              Player    xG   Age Pos
807         de Bundesliga    Hugo Ekitike   18.5  22.0  FW
2243  eng Premier League    Mohamed Salah   22.5  32.0  FW
1449          es La Liga  Robert Lewandowski   24.0  36.0  FW
823           fr Ligue 1   Emanuel Emegha   15.2  22.0  FW
1299          it Serie A      Moise Kean   15.6  25.0  FW

Top Tackling Players:
                  Comp              Player   Tkl   Age     Pos
331         de Bundesliga     Tom Bischof   86.0  19.0   MF,FW
939   eng Premier League  Idrissa Gana Gueye  115.0  35.0      MF
1135          es La Liga   Omar El Hilali   97.0  21.0      DF
2285          fr Ligue 1    Andrey Santos   91.0  20.0      MF
914           it Serie A  Morten Frendrup   97.0  24.0      MF
```

Players_comaprative_analysis