

FOOTBALL ANALYTICS

Karthik Iyer
Visualization of Complex Data

ABOUT THE PROJECT

This project aims to design a creative Visualization dashboard for Manager, Coaches and Scouts to understand the teams, players and the different aspects that can enhance the performance of their teams.



WHATS THE GAME



A group of 22 players in 2 teams who

- Run
- Dribble
- Shoot
- Snatch
- Defend...

and in general take actions to put a ball in the back of the opposing net and not concede one in their own...

ANALYTICAL COURSE

Generating Data

The data comes from the official FIFA website.

We drop unnecessary column and impute where needed.

Cleaning

Feature Engineering

We add columns to the data, for gender, year, positions and more.

We can now move on to understand our data.... ->

Visualizing



THE DATA

01

Positions

We discover the different roles a player can take in the team

02

Attributes

Pace, Shooting Dribbling and the other traits that a player is evaluated on

03

Nationality

The nationality of football/soccer players across the world

04

Gender

The gender-wise distribution of players in the sport.

05

League

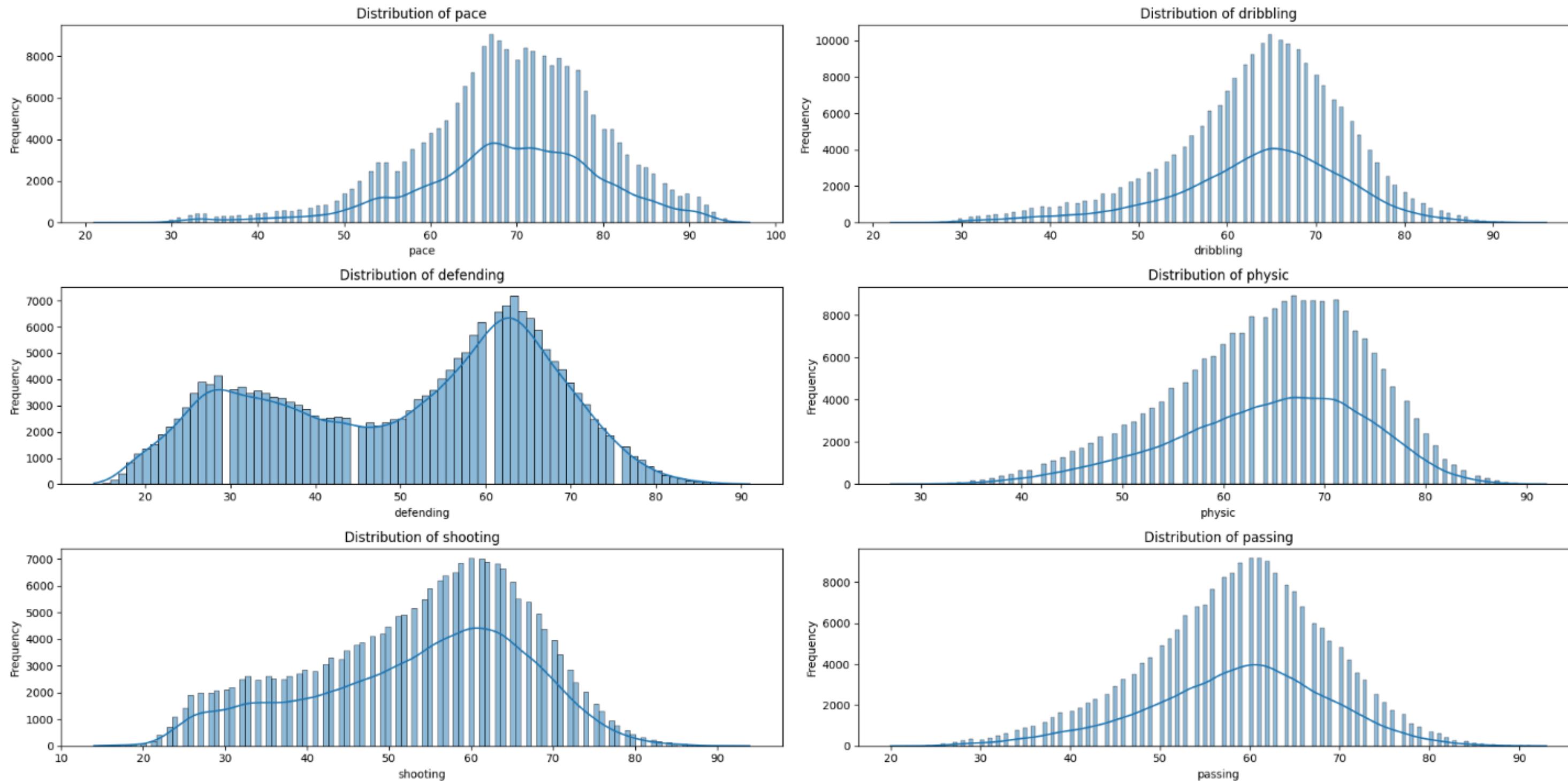
The distribution of players at the different levels of Football.

06

Market Value

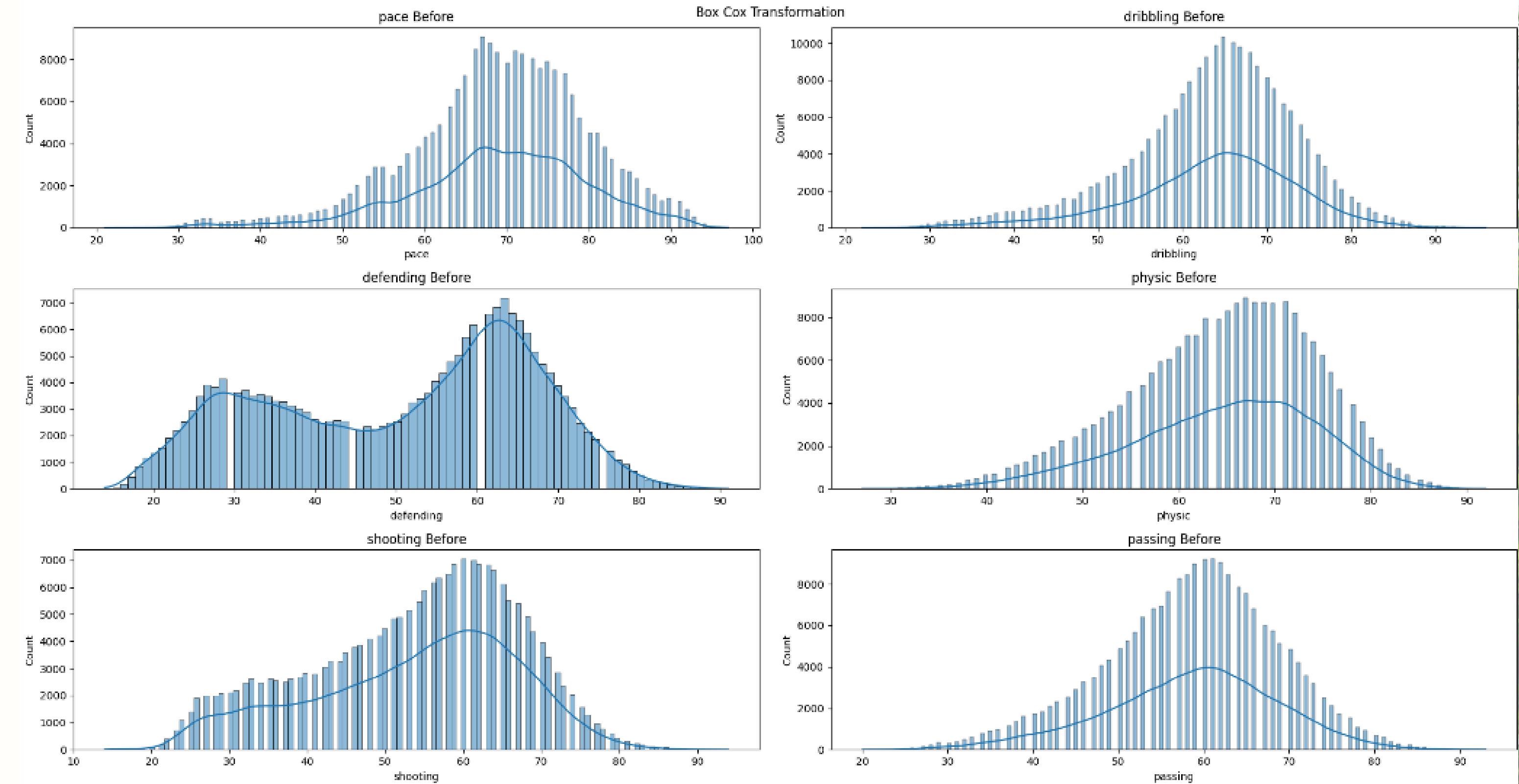
An analysis of the market value of the players.

DATA OVERVIEW

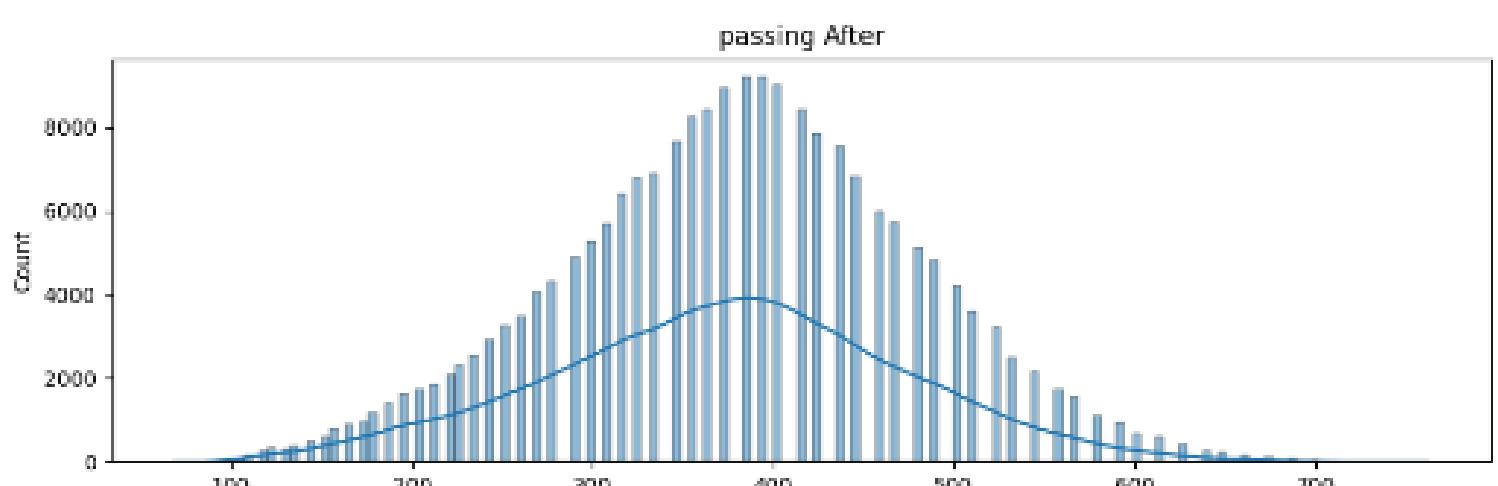
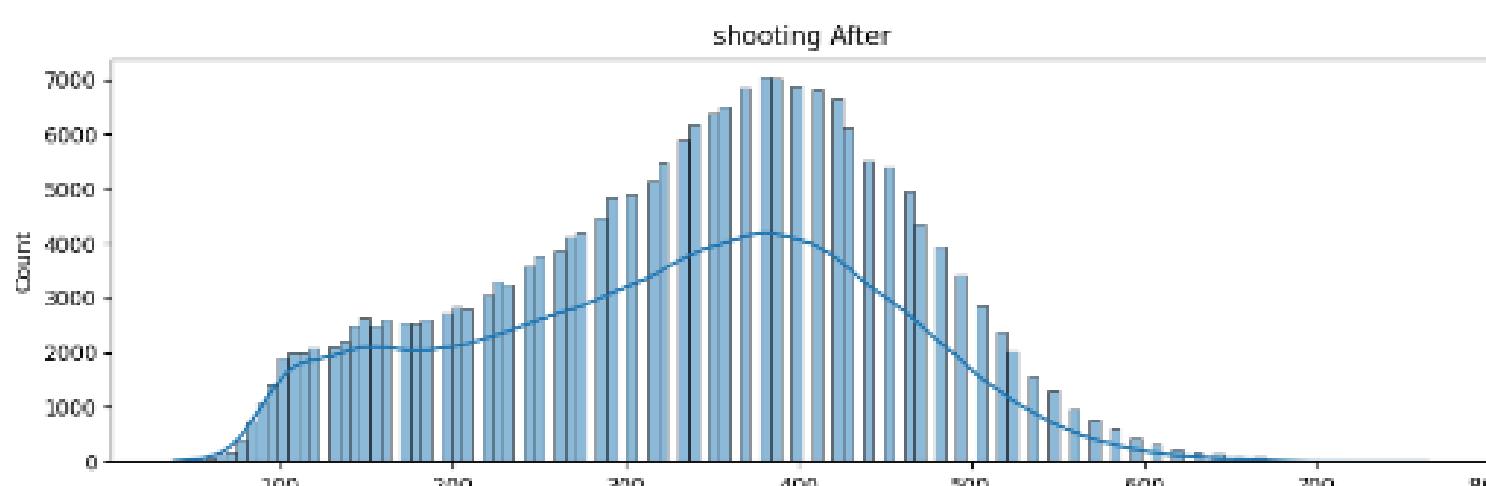
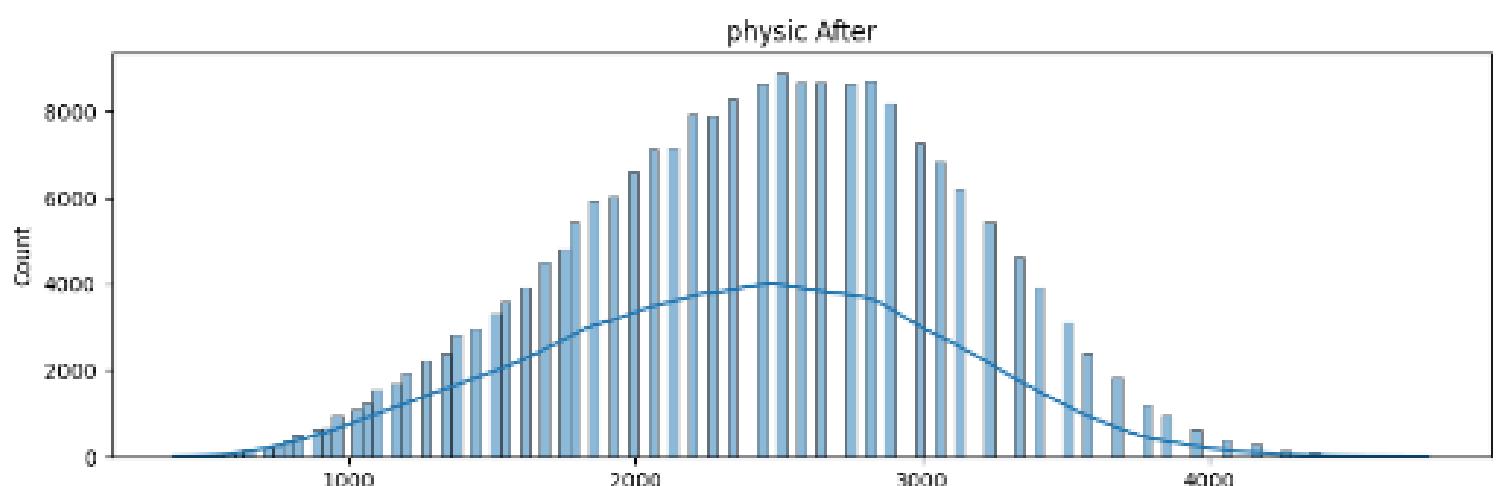
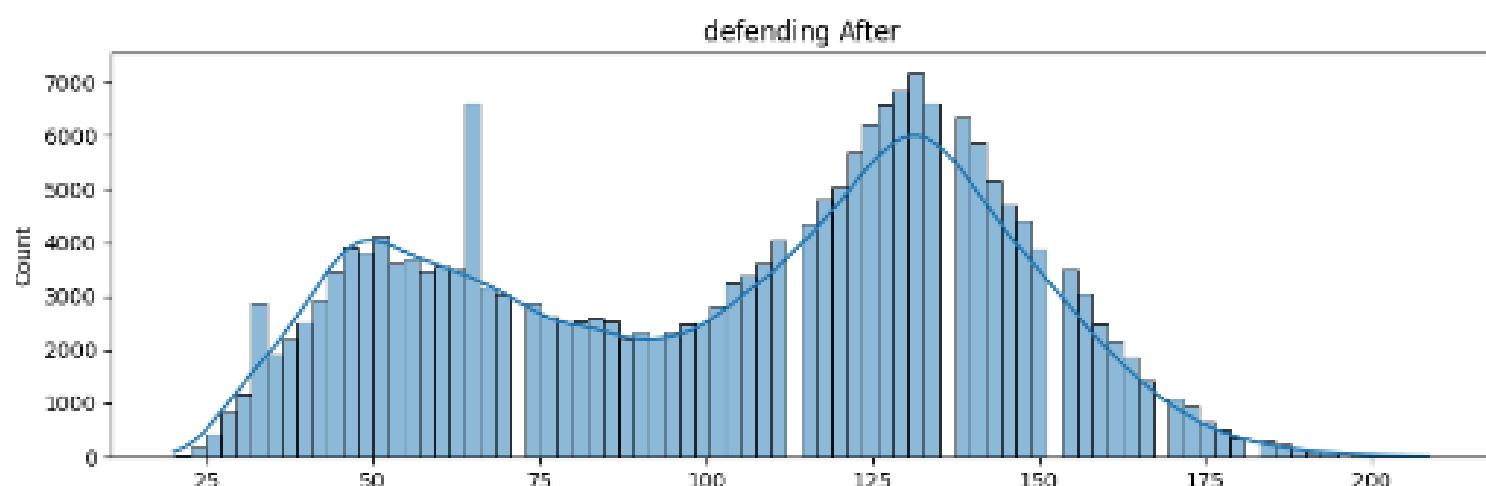
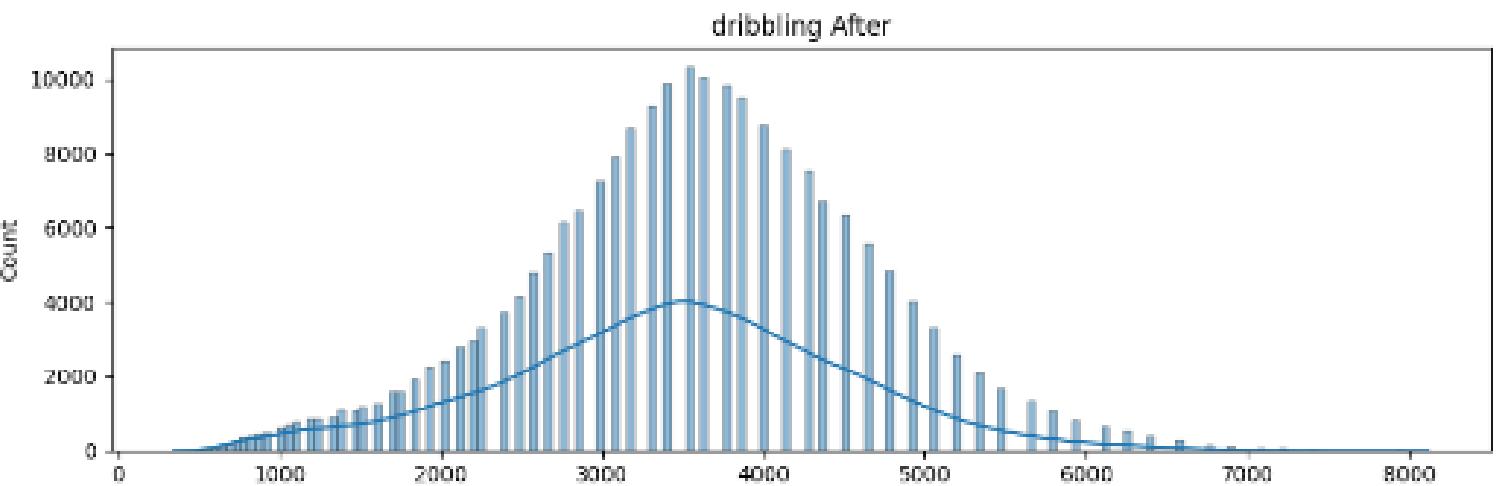
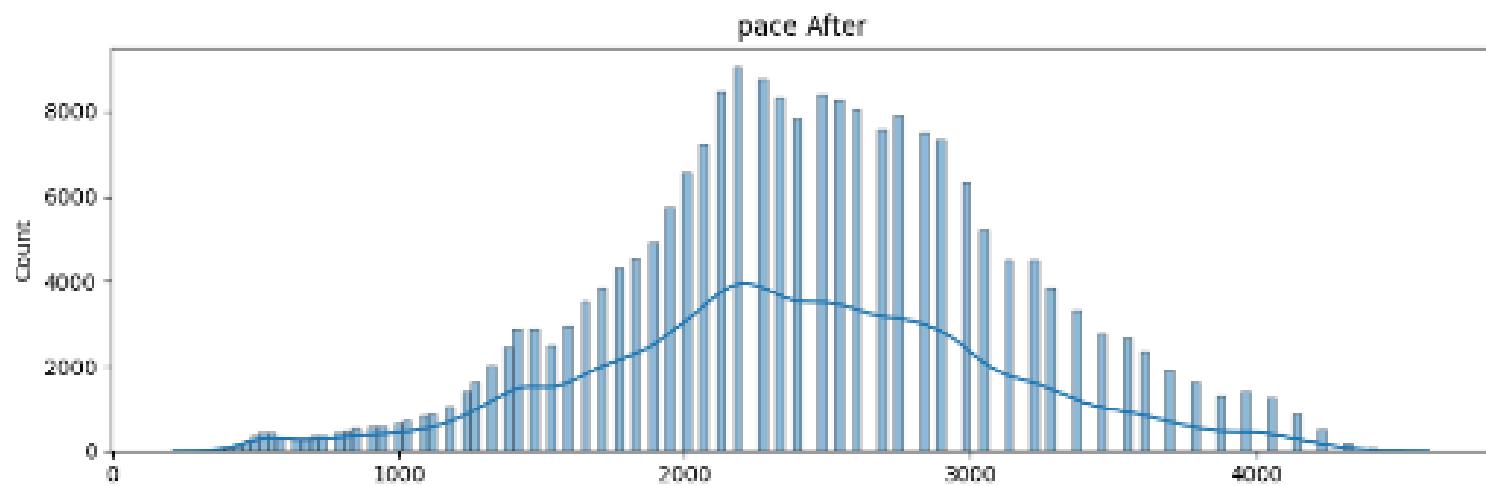


We notice that although data follows a Gaussian distribution, there is a need to check for normality.

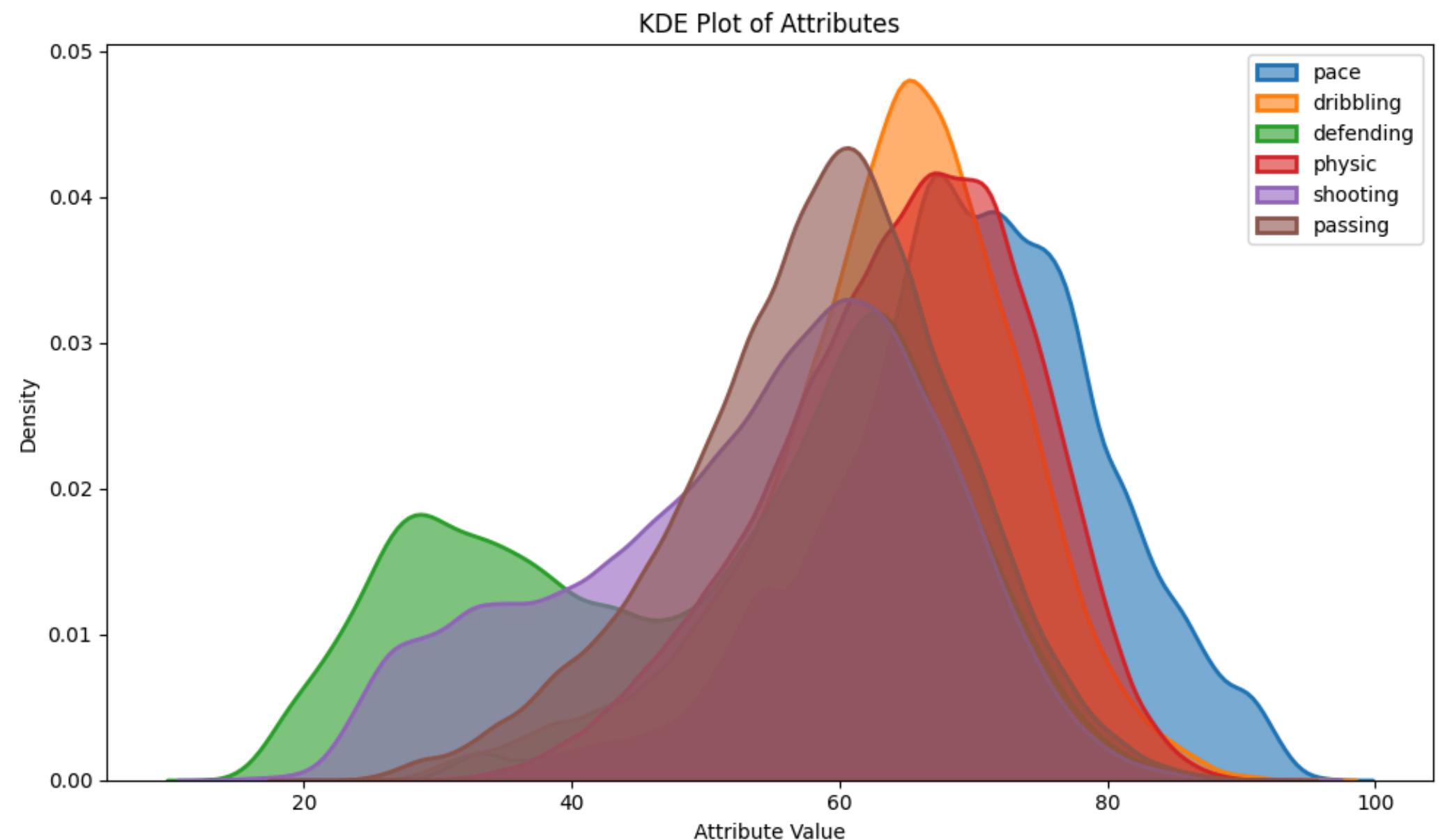
BEFORE BOX-COX



AFTER BOX-COX



LEAGUE IMPACT



NORMALITY CHECK

Column	Lambda	P-Value after
pace	1.99	0.0
dribbling	2.14	0.0
defending	1.23	0.0
physic	2.03	0.0
shooting	1.56	0.0
passing	1.56	0.0

The different attributes all have a lambda value of 1.5 or higher.

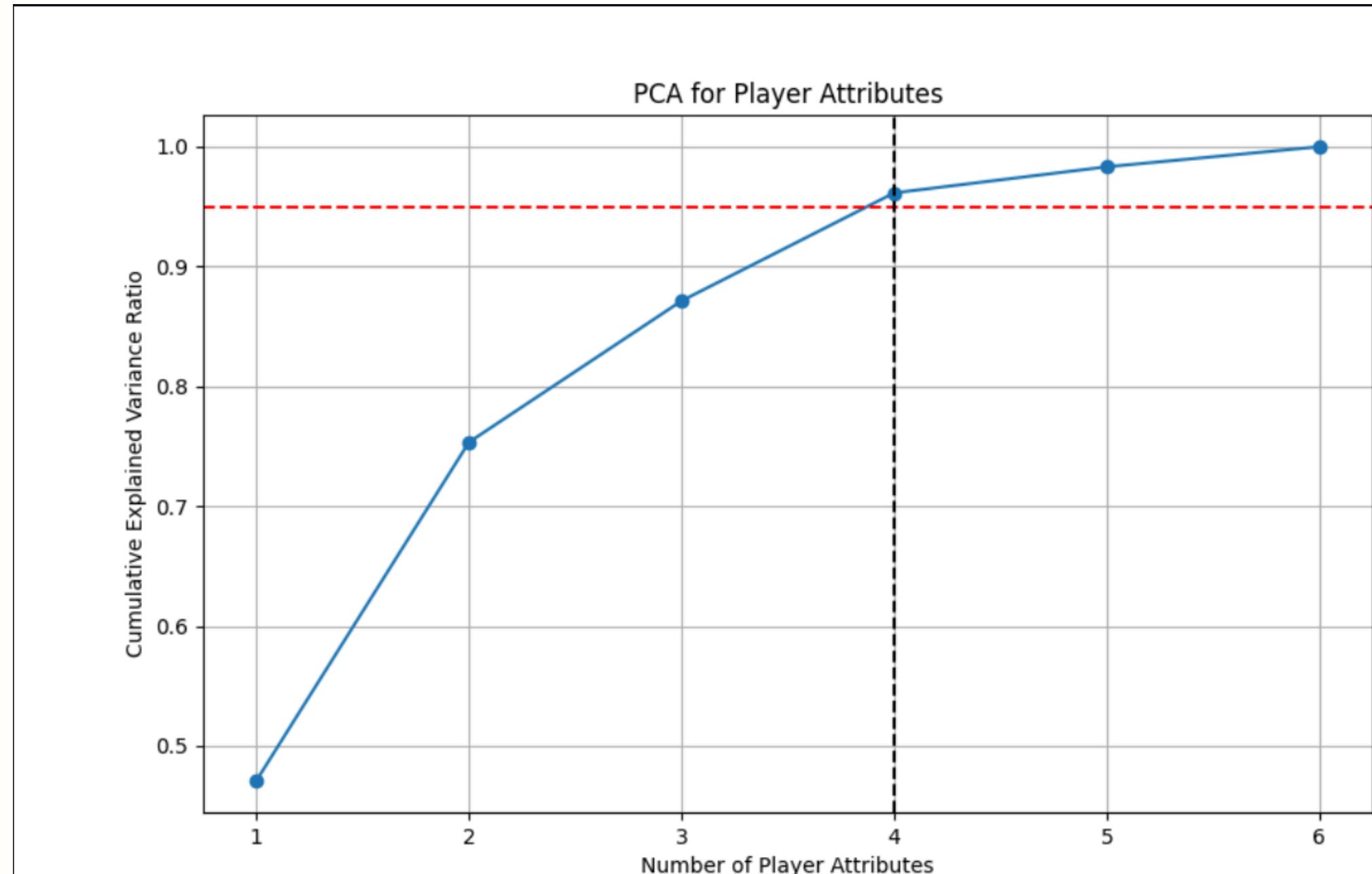
This indicates that there is a skew in the data.

Additionally the P-values are less than zero when checked using the KS-test indicating that data is not normal even after transformation.

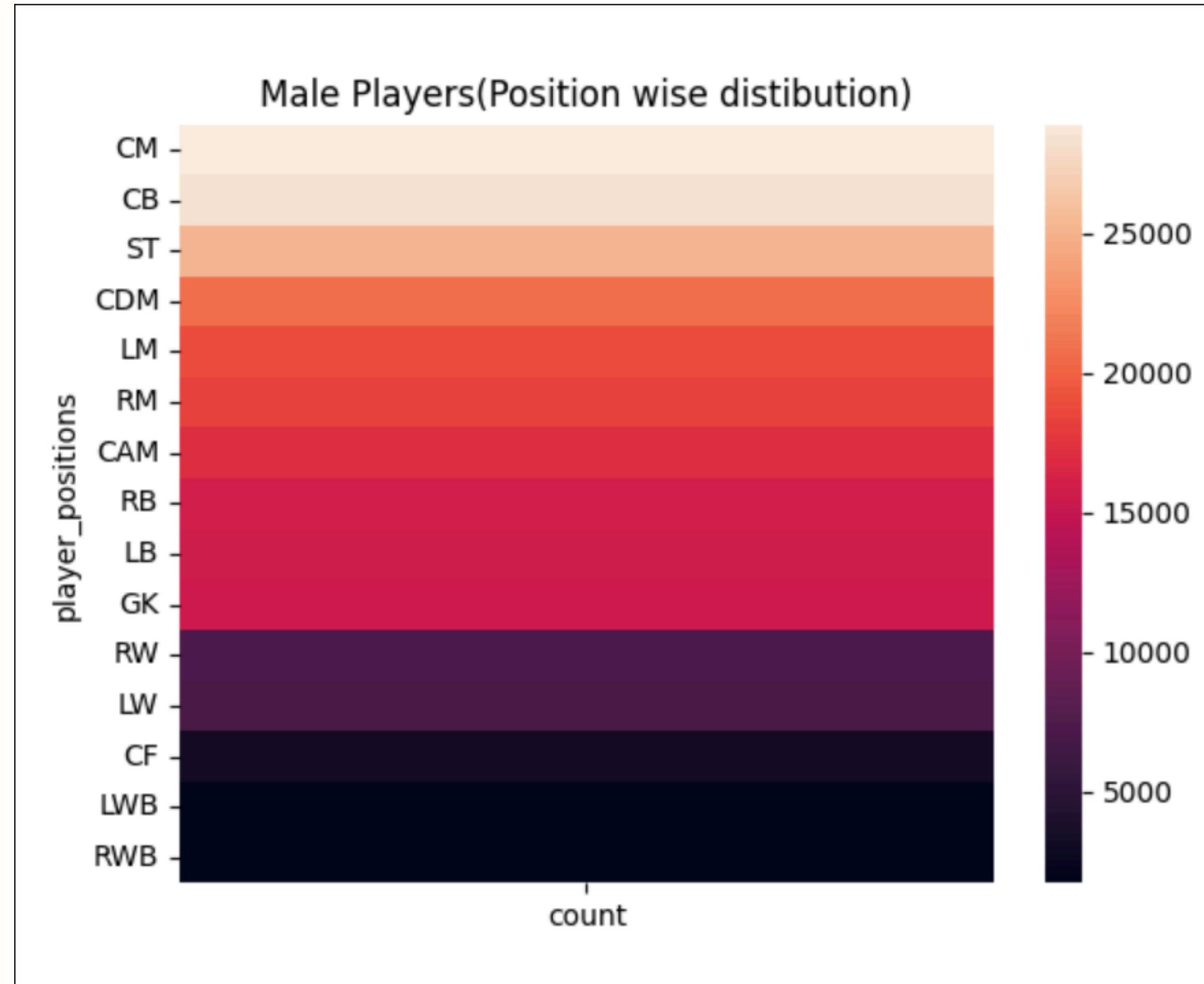
PCA ANALYSIS

For the PCA Analysis we take into consideration the 6 attributes for the player -'pace', 'shooting', 'passing', 'dribbling', 'defending', 'physic'

With the analysis we can see that 4 components would suffice in explaining 95% of the variance in the 'overall' performance of players.



POSITION DISTRIBUTION

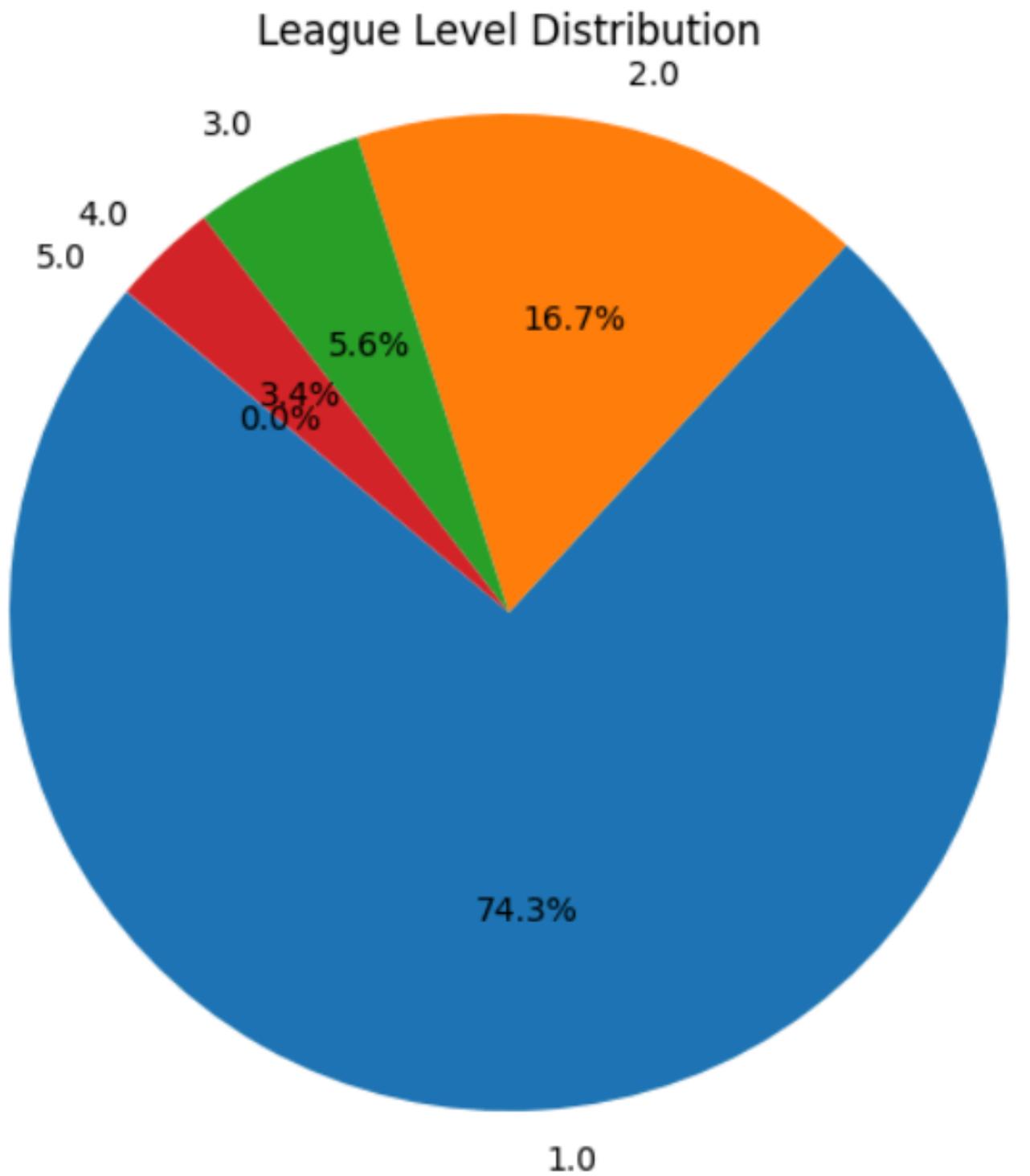


There are more players playing in the center of the field
CM, CB, ST, CDM...
>20000

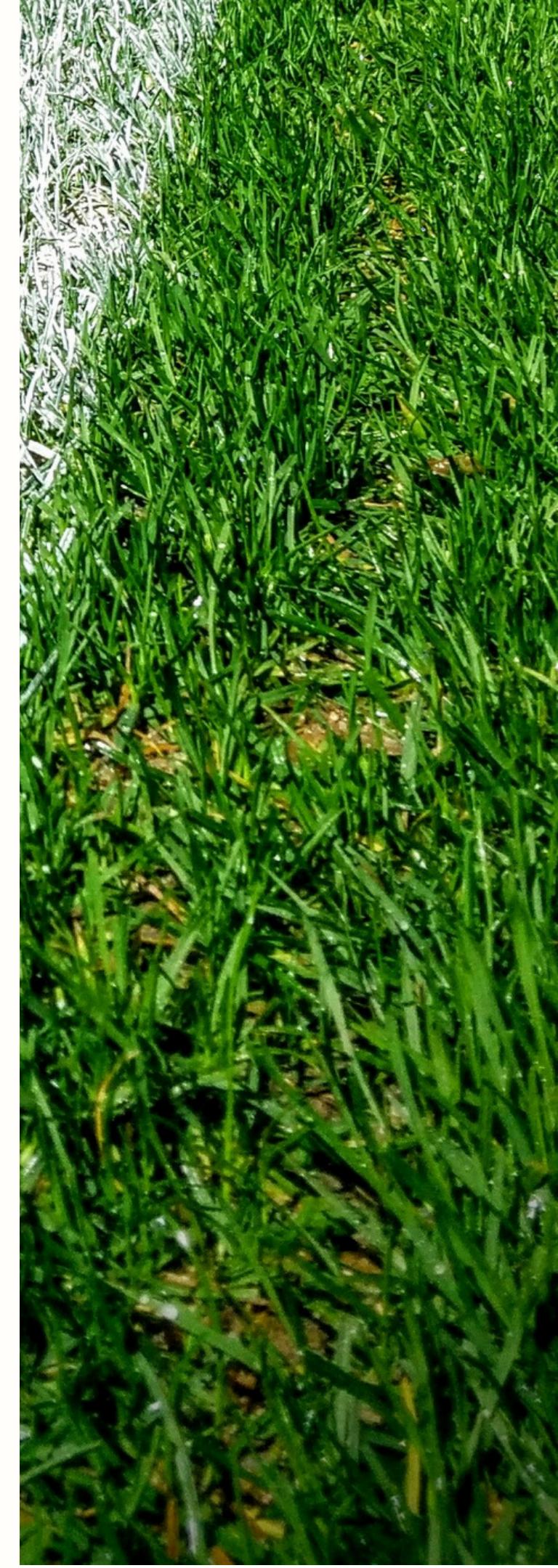
as compared to the edges of the field
LWB, RWB, LW, RW...
<5000



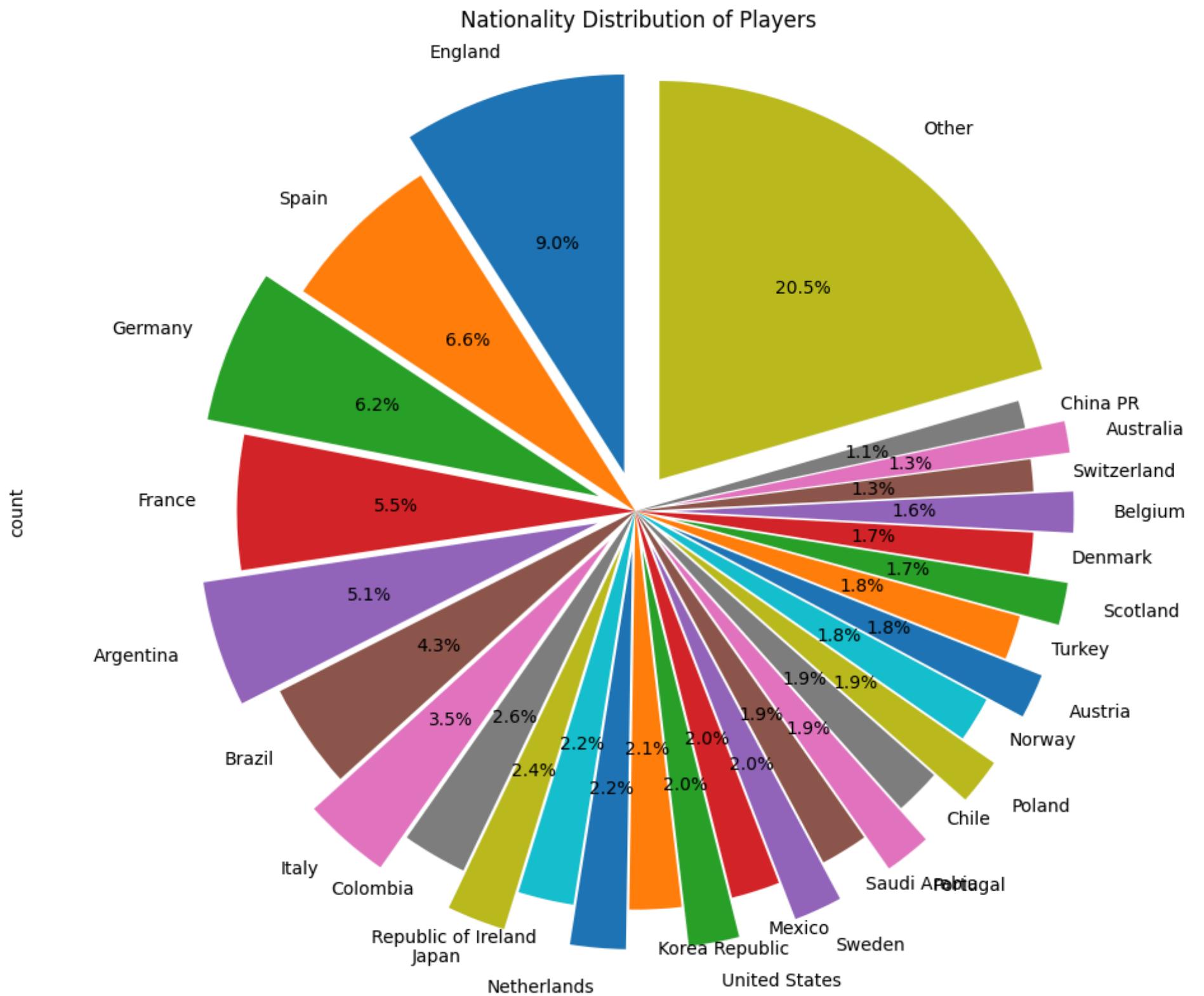
LEAGUE IMPACT



Majority of the players in FIFAs purview play in the top league in all nations



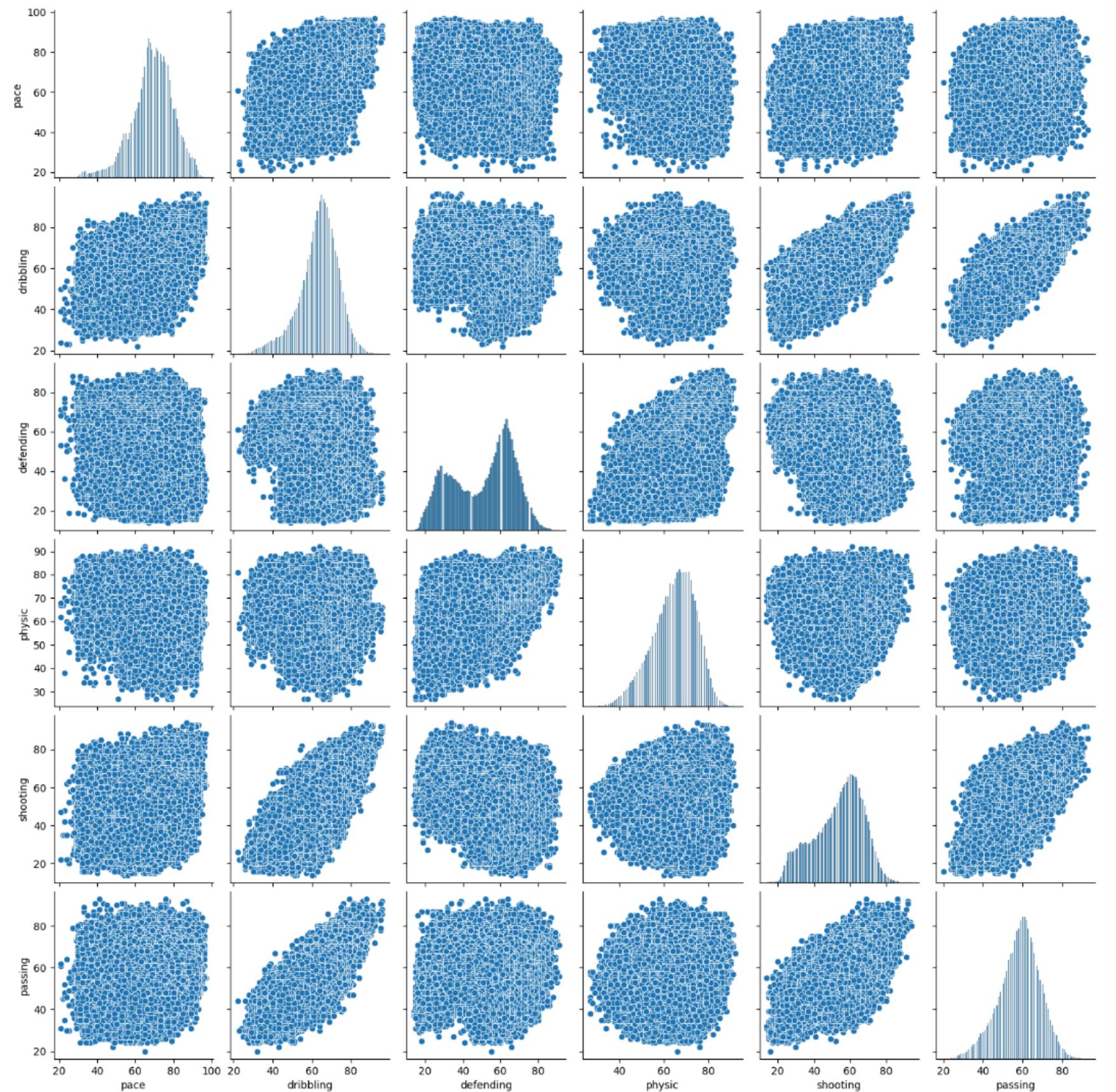
UNDERSTANDING NATIONALITY



We can notice here, that England, Spain, Germany and France have the majority of players.

**Closely followed by
Argentina and Brazil.**

ATTRIBUTE PAIR



The pair plot, shows us how the different attributes are correlated amongst each other.

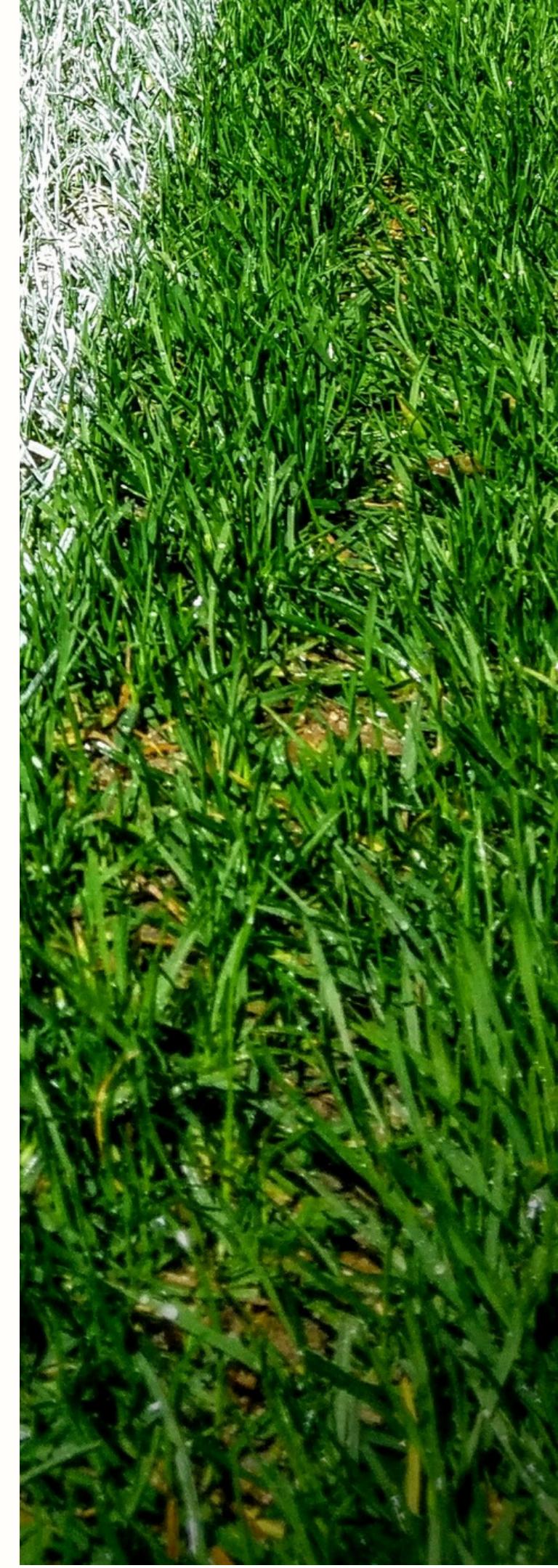
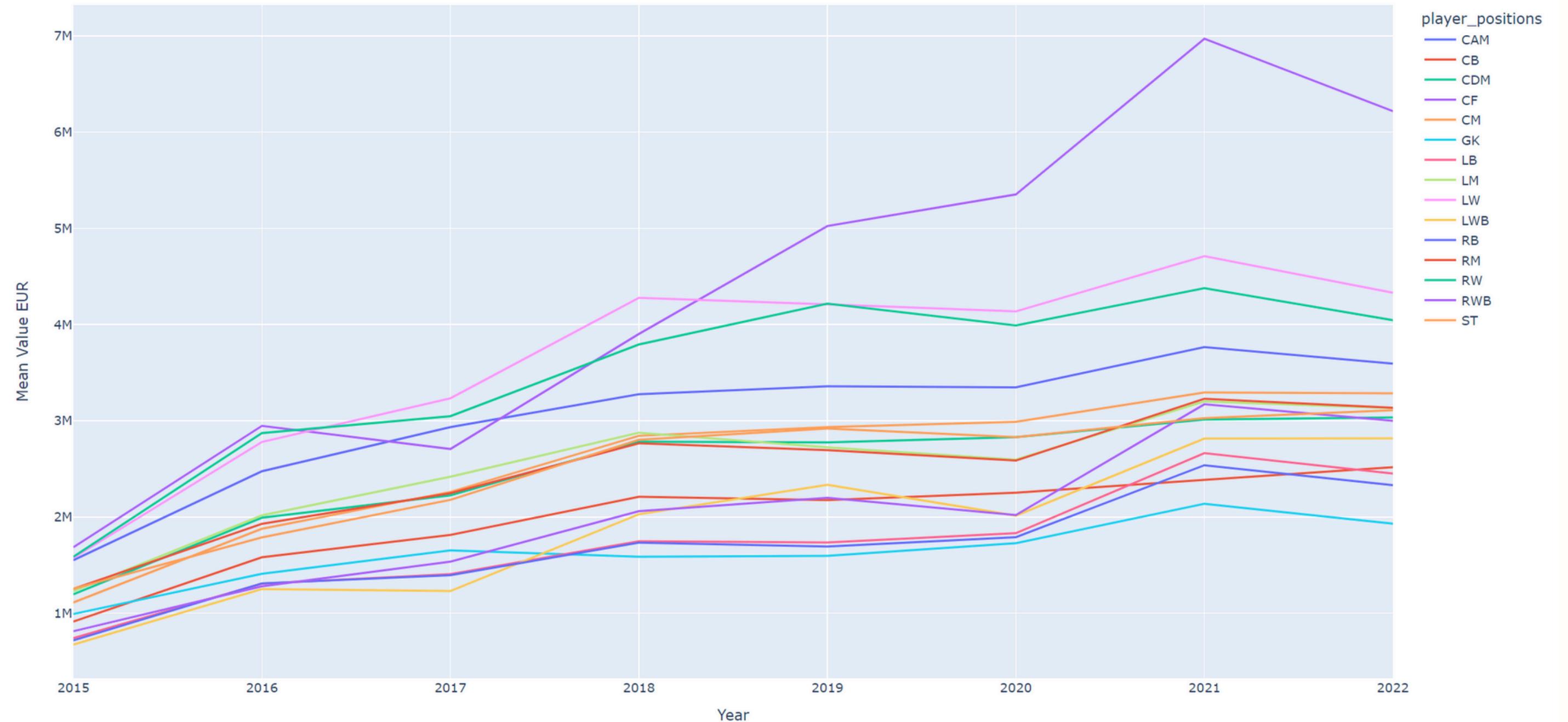
As we can see from the plots, dribbling has some positive correlation with shooting and passing.

But there is no real correlation other wise

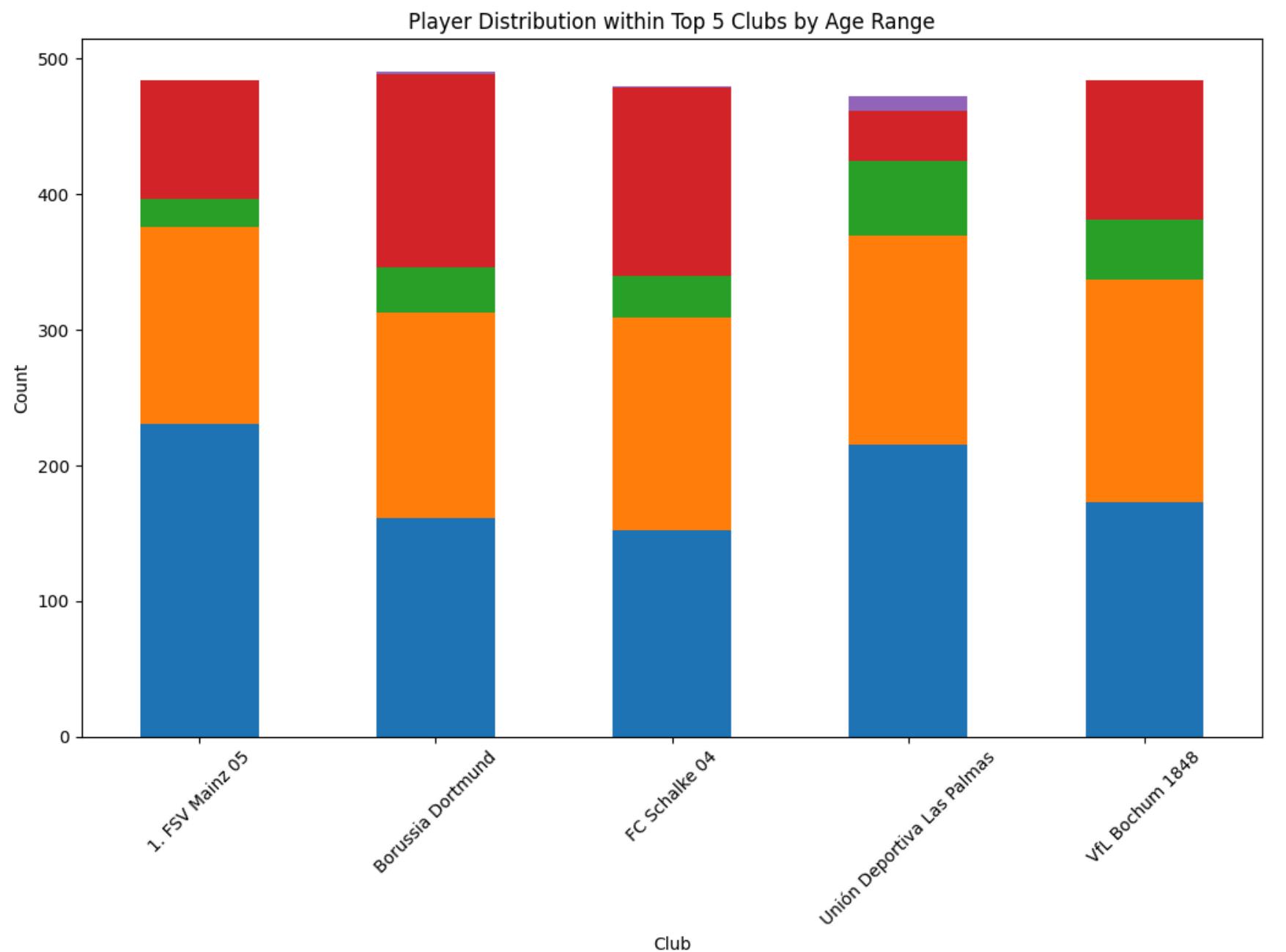


VALUE AS PER POSITION

Mean Value EUR Over the Years for Both Genders



AGE OF TOP 5 CLUBS

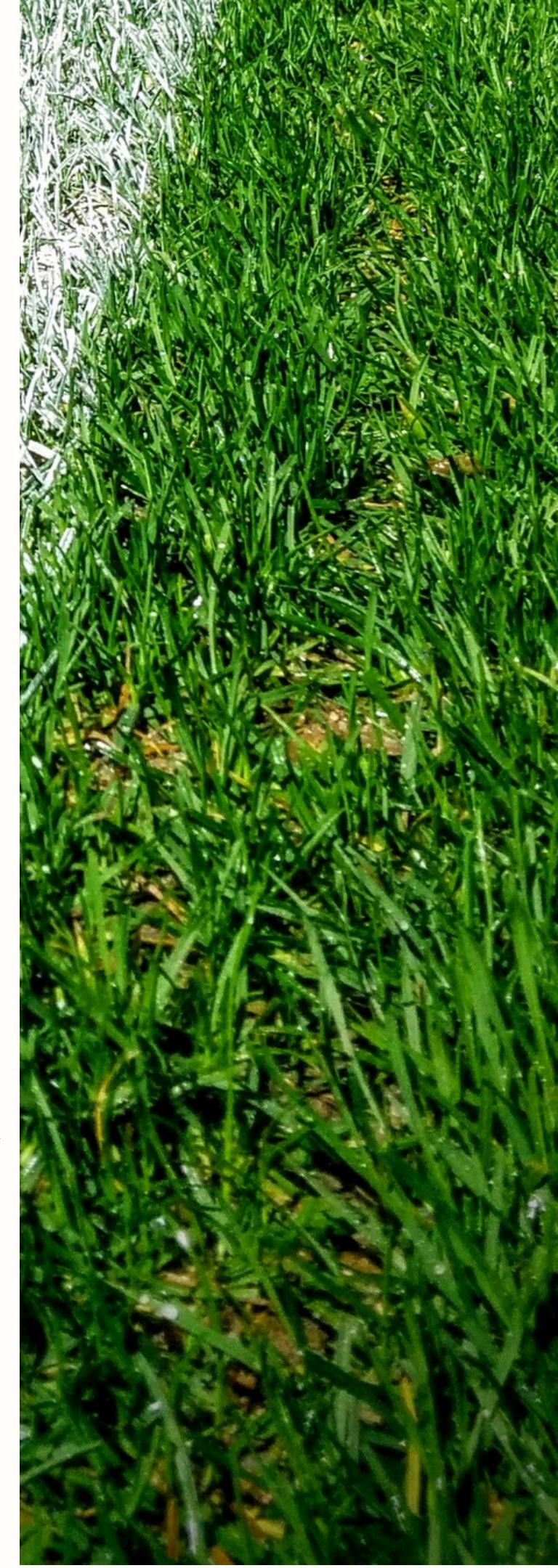
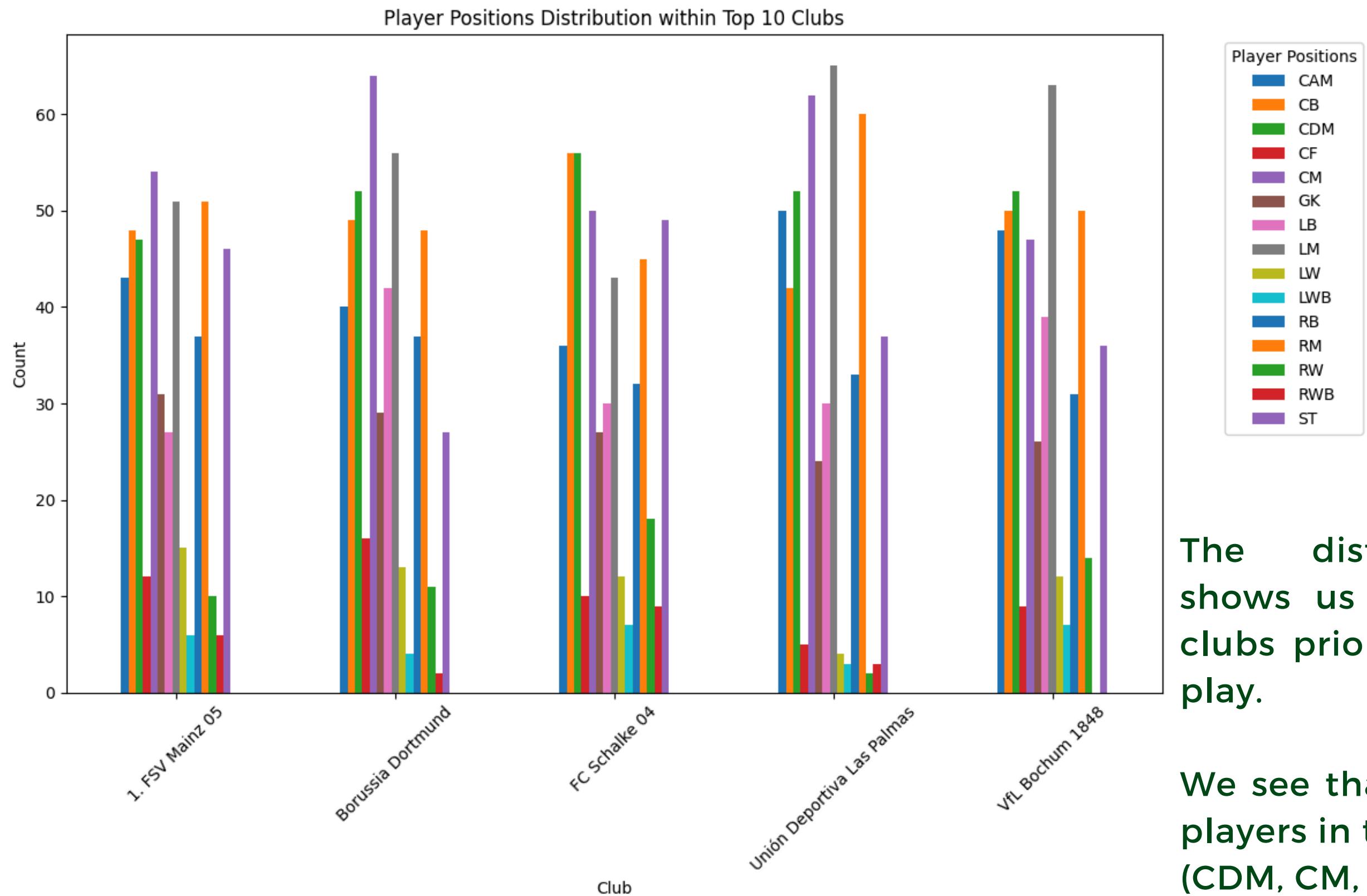


We take a look at the top 5 clubs, and can see that a majority of players in all clubs are in the 21-30 age range

With the next majority being below 20yrs of age.

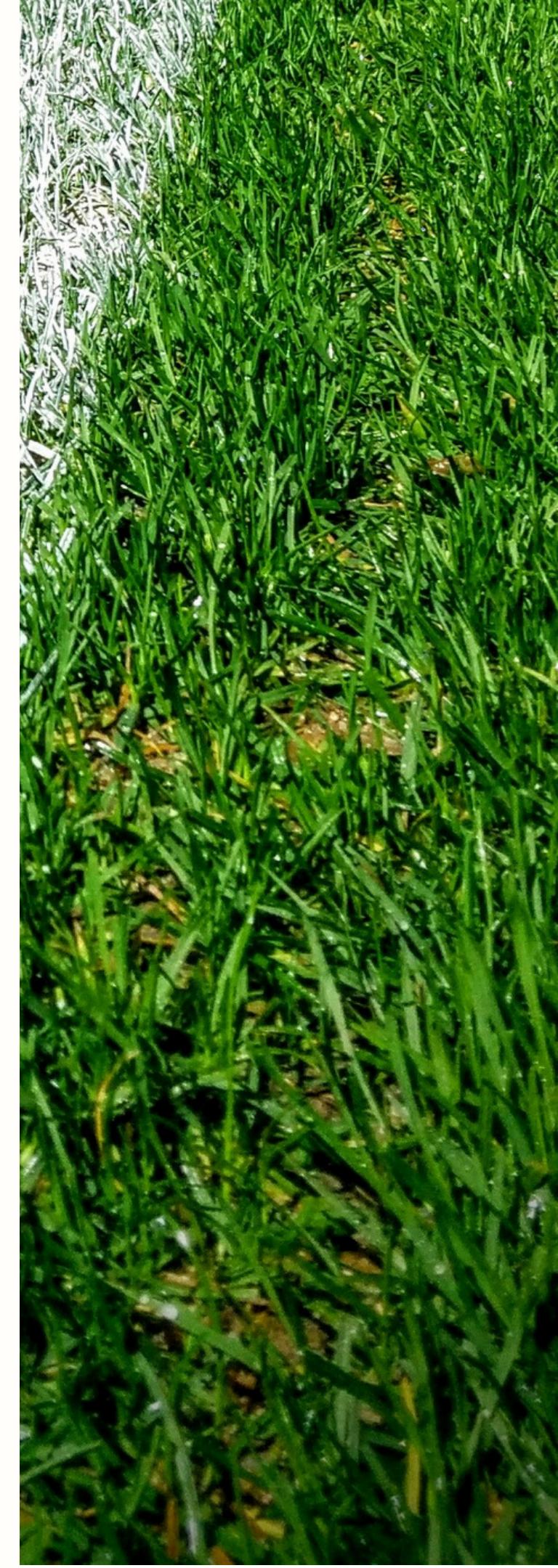
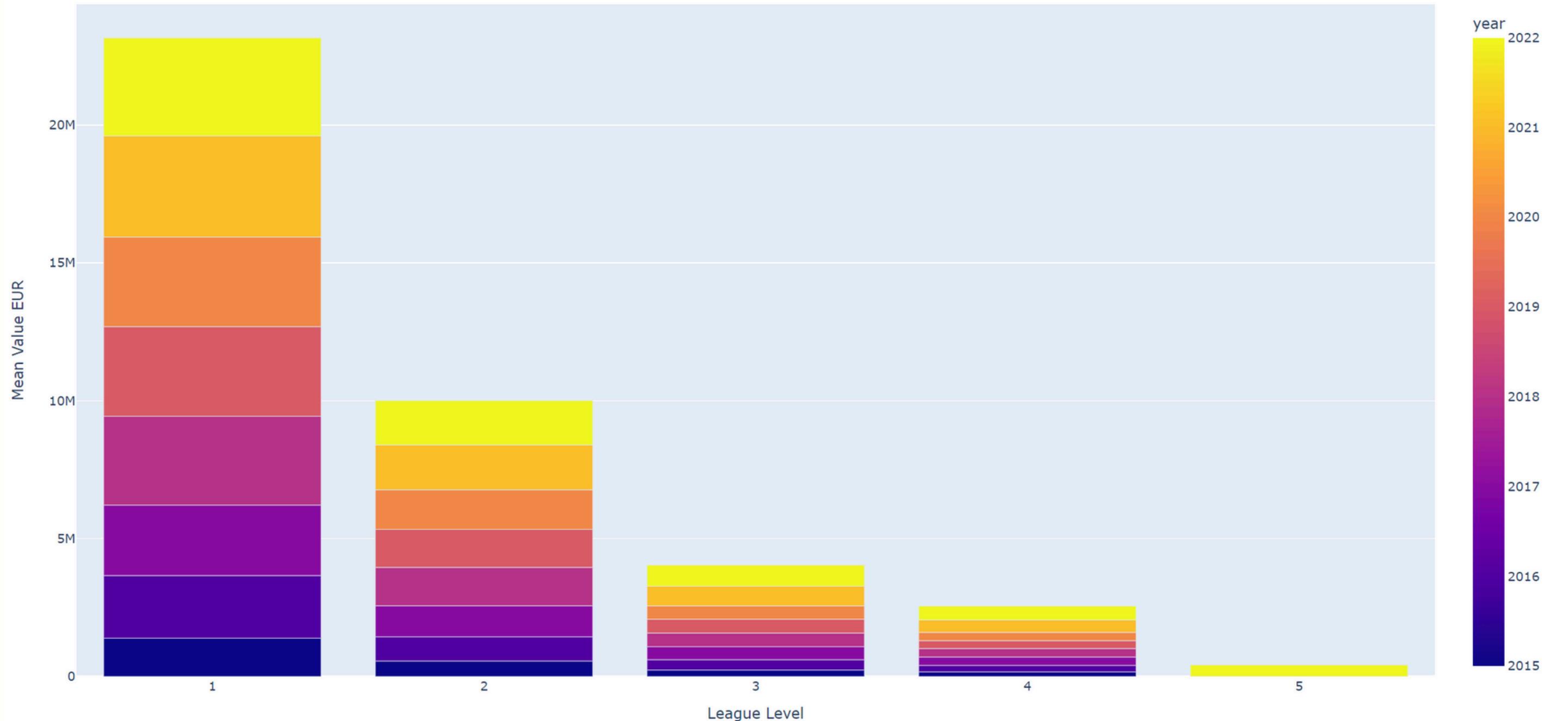


POSITIONALITY BY TEAM



MARKET OVER THE YEARS

Mean Value EUR Across Different Leagues Over the Years



ANALYTICS DASHBOARD

DASHBOARD SNIPPET

European Football Dashboard

Dynamic player evaluation based on player attributes

Overall Data Evaluation

In-depth Analysis

|Choose a evaluation tab|

Male Players Data

Female Players Data



[DOWNLOAD DATA SET](#)

DASHBOARD SNIPPET

European Football Dashboard

Dynamic player evaluation based on player attributes

Overall Data Evaluation

In-depth Analysis

Data Overview - Understanding the features of the Data

Select from following options to evaluate

Male Players Female Players

Outlier Check Normality Check

CAM

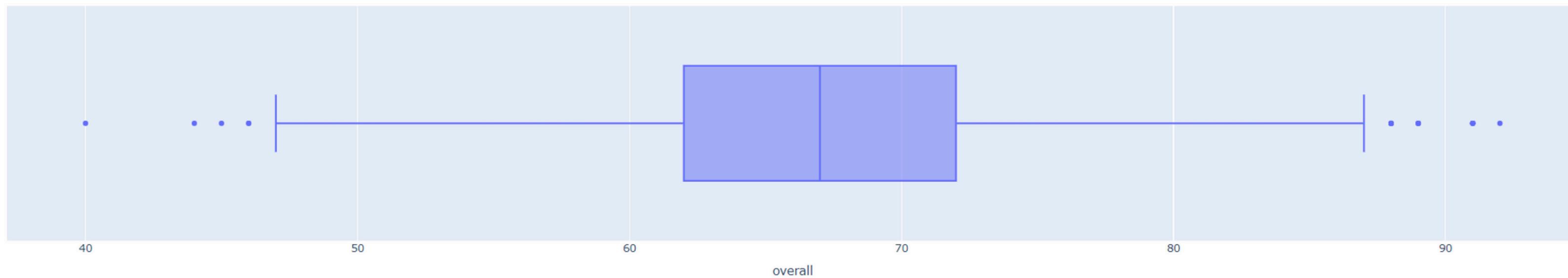
x ▾

Overall

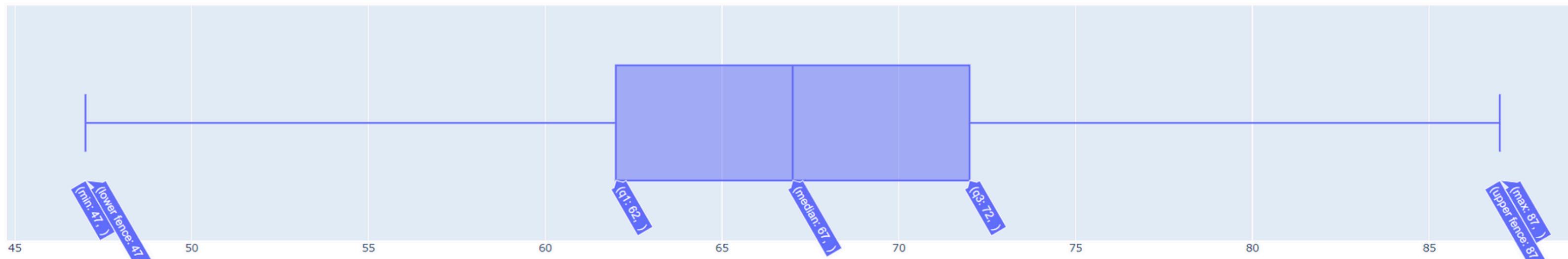
x ▾

DASHBOARD SNIPPET

Boxplot before outliers

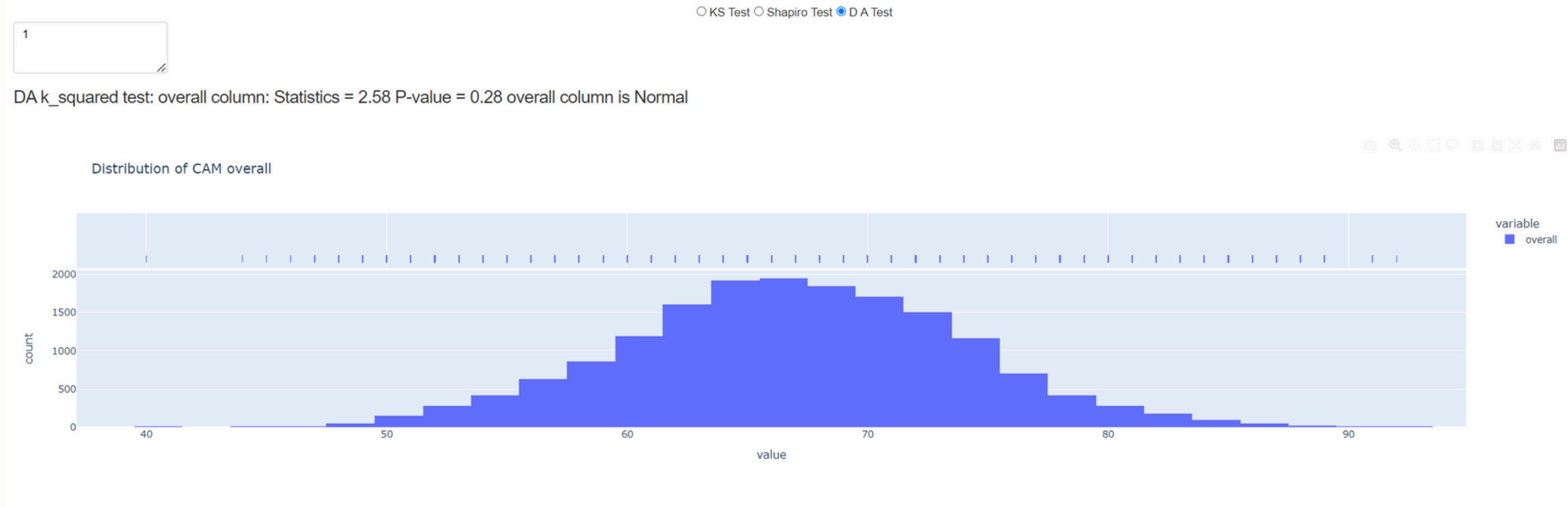


Boxplot after removing outliers



DASHBOARD SNIPPET

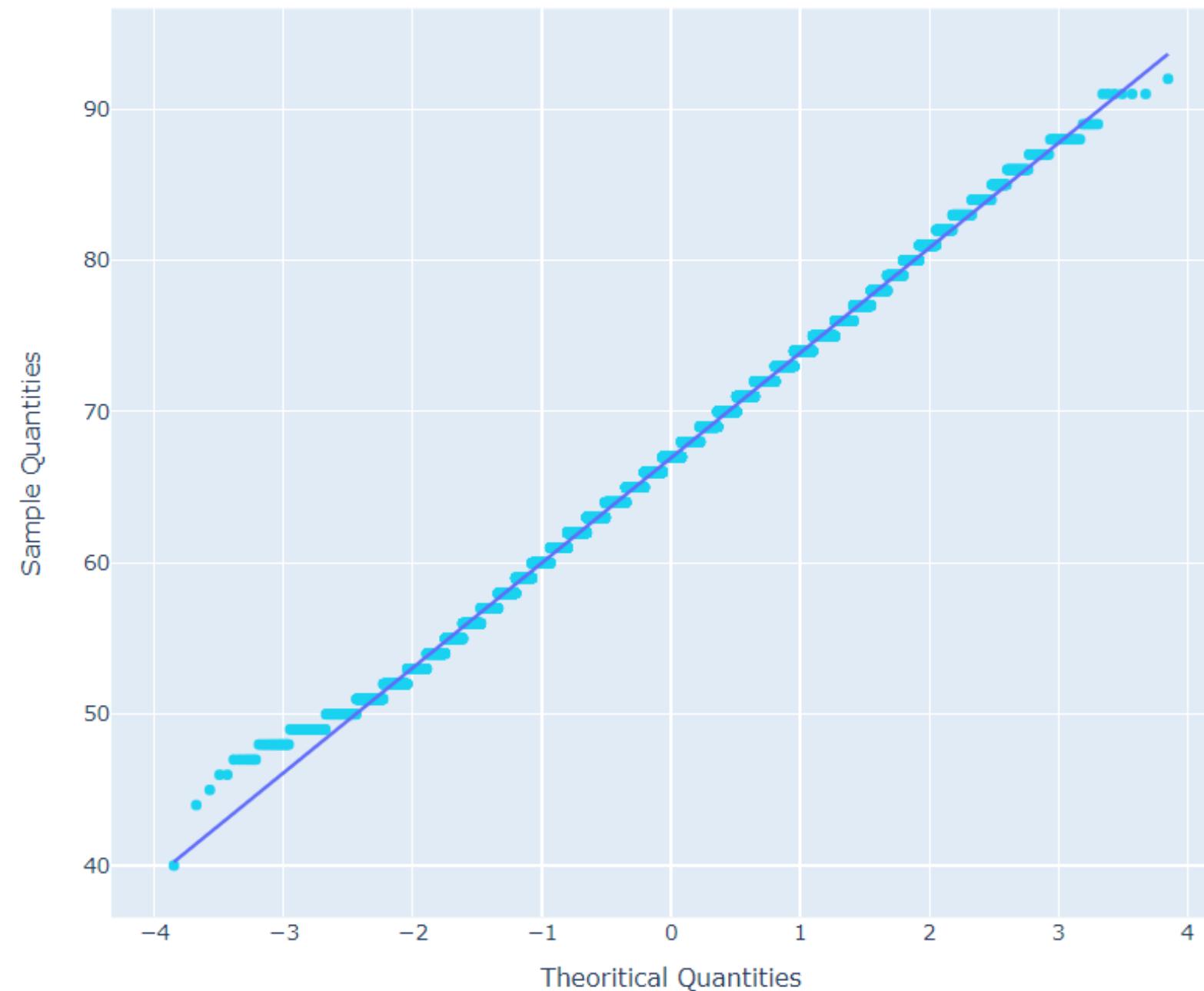
Based on the normality test:



DASHBOARD SNIPPET



Quantile-Quantile Plot



DASHBOARD SNIPPET

European Football Dashboard

Dynamic player evaluation based on player attributes

Overall Data Evaluation

In-depth Analysis

Player vs Player evaluation

Male Players Female Players

Player Names

▼

Player Attributes





**THANK
YOU!**