

Capstone Project

EPL 24/25 Player Analysis & Scouting Report

Kevan Tay
28-July-2025



Personas



James

Football Scout at Newly Promoted EPL Club

Goal: Discover underrated players in the league to bolster squad strength

Challenge: Limited time and budget for comprehensive global scouting



Omar

Director of Football at a Championship Club

Goal: Sign undervalued players now to flip for profit after promotion

Challenge: Need high ROI on tight budget

Introduction

Problem Statement:

- Identify the most value-for-money players in the English Premier League 24/25 season

Project Objective:

- To assess and rank players on how much impact/contribution they make with respect to their market value.
- Provide recommendations to stakeholders to identify **value-for-money** players for recruitment.
- Develop a **predictive model** to aid in player valuation.



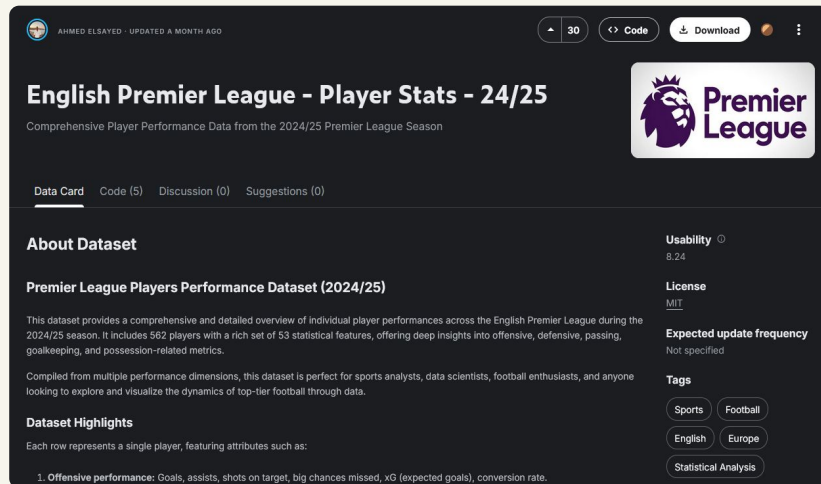
Dataset

Dataset 1: English Premier League - Player Stats - 24/25

- Kaggle dataset containing player statistics for 562 players with 53 statistical features
- Contains statistics used to measure offensive performance, defensive contribution, passing & creativity, possession and ball movement, and disciplinary records.
- Source:
<https://www.kaggle.com/datasets/aesika/english-premier-league-player-stats-2425>

Dataset 2: Football Data from Transfermarkt

- A collection of CSVs scraped from reputable football statistics site Transfermarkt
- Source:
<https://www.kaggle.com/datasets/davidcariboo/player-score-s/data>



English Premier League - Player Stats - 24/25
Comprehensive Player Performance Data from the 2024/25 Premier League Season

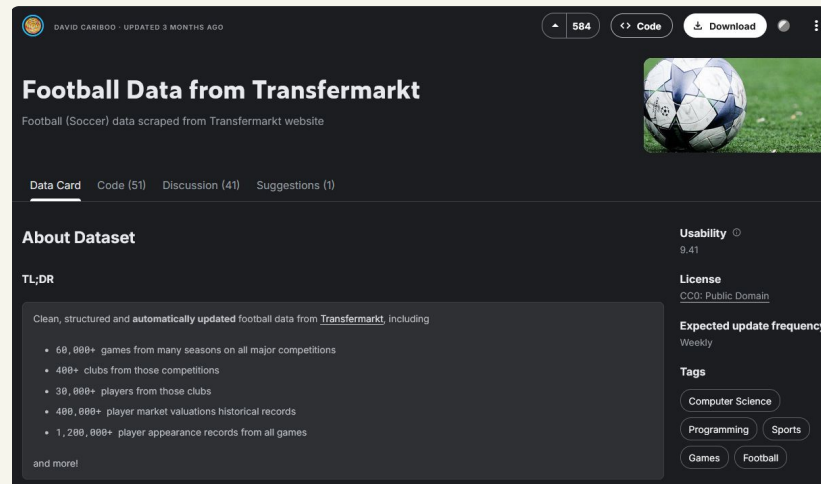
Usability: 8.24
License: MIT
Expected update frequency: Not specified
Tags: Sports, Football, English, Europe, Statistical Analysis

About Dataset
Premier League Players Performance Dataset (2024/25)
This dataset provides a comprehensive and detailed overview of individual player performances across the English Premier League during the 2024/25 season. It includes 562 players with a rich set of 53 statistical features, offering deep insights into offensive, defensive, passing, goalkeeping, and possession-related metrics.

Compiled from multiple performance dimensions, this dataset is perfect for sports analysts, data scientists, football enthusiasts, and anyone looking to explore and visualize the dynamics of top-tier football through data.

Dataset Highlights
Each row represents a single player, featuring attributes such as:

1. **Offensive performance:** Goals, assists, shots on target, big chances missed, xG (expected goals), conversion rate.



Football Data from Transfermarkt
Football (Soccer) data scraped from Transfermarkt website

Usability: 9.41
License: CC0: Public Domain
Expected update frequency: Weekly
Tags: Computer Science, Programming, Sports, Games, Football

About Dataset
TL;DR
Clean, structured and **automatically updated** football data from **Transfermarkt**, including

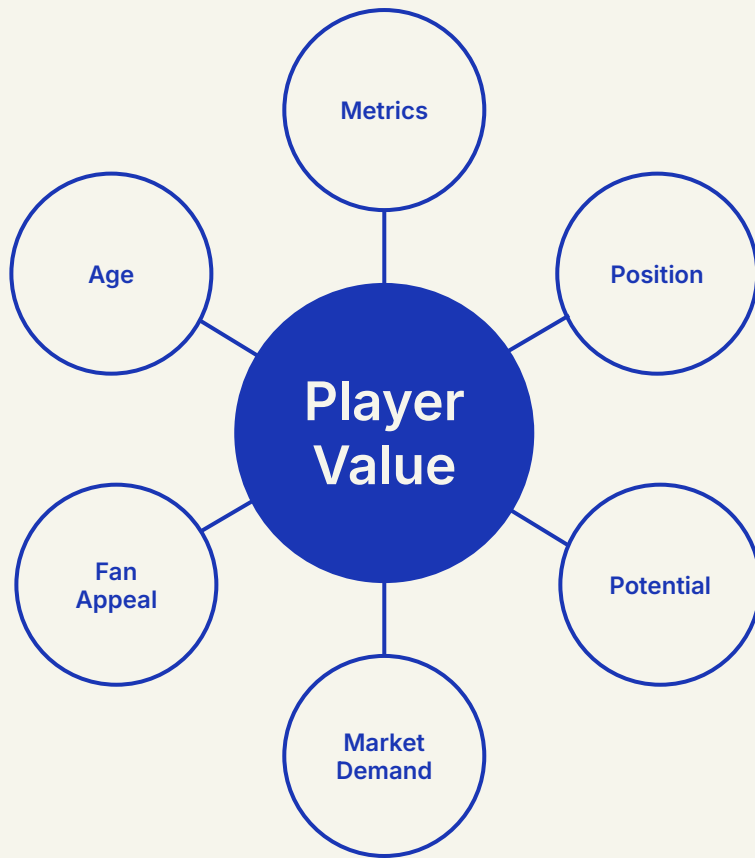
- 60,000+ games from many seasons on all major competitions
- 490+ clubs from those competitions
- 30,000+ players from those clubs
- 400,000+ player market valuations historical records
- 1,200,000+ player appearance records from all games

and more!

How is Player Value Determined?

Player valuations are determined through a multifaceted process influenced by both **on-field performance** and **off-field considerations**. Factors such as player performance, position, age and marketing factors like media hype, fan appeal can all greatly influence a player's valuation.

Note: We will use the closest valuations on Transfermarkt prior to start of EPL 24/25 Season for this analysis.

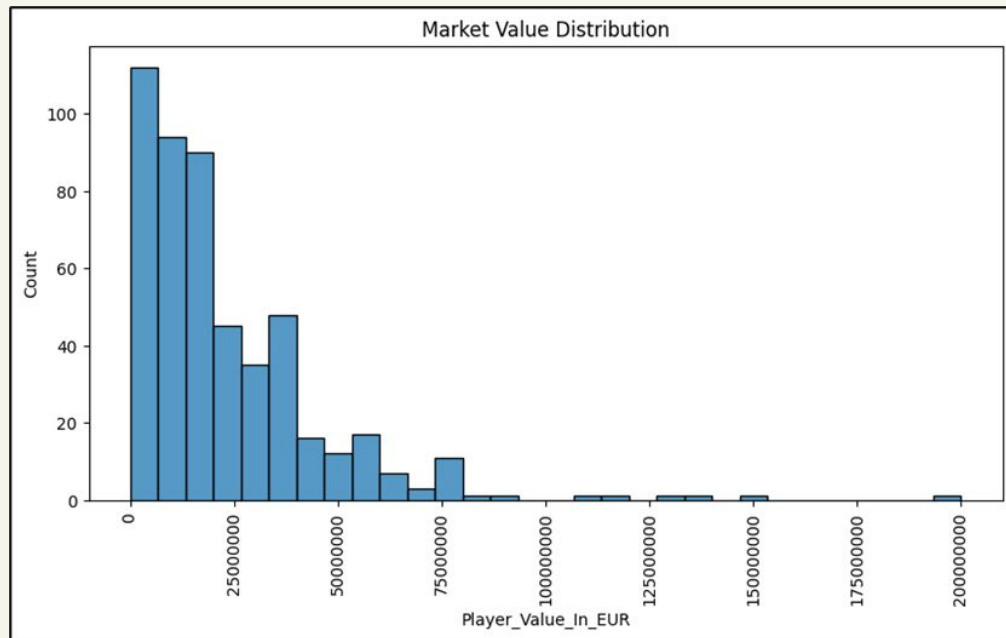


Exploratory Data Analysis



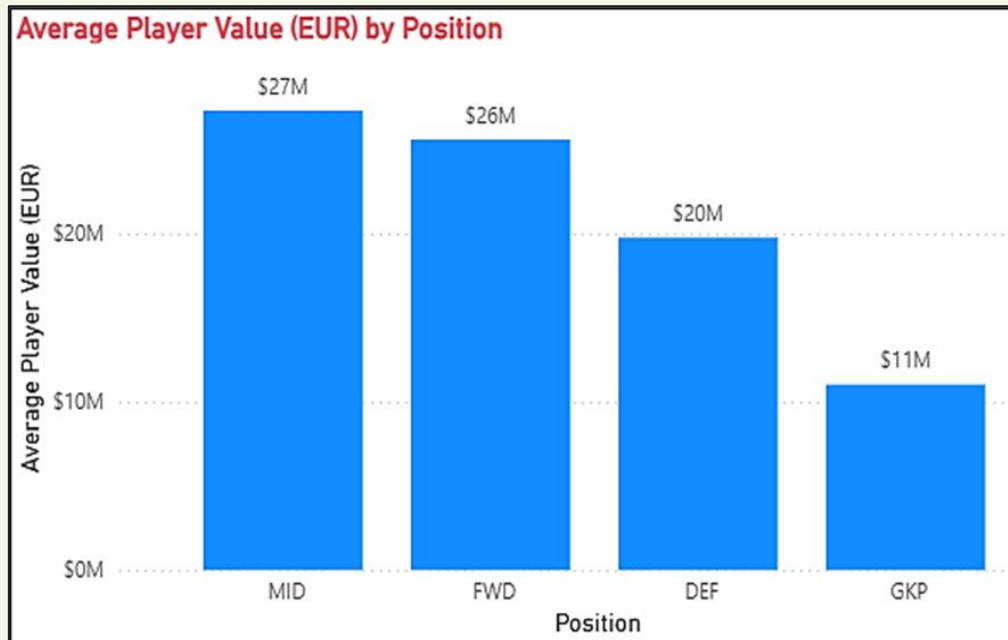
Player Market Value Distribution

- Majority of players are valued below €50,000,000
- Top players like Erling Haaland are considered outliers with a market valuation of >€100,000,000.



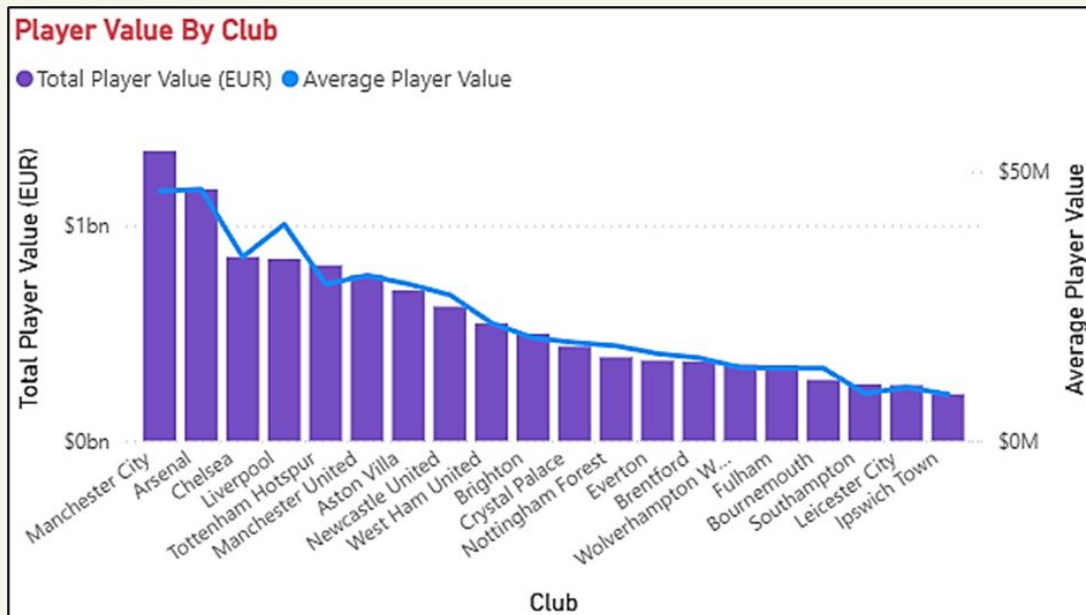
Player Positions

- Attacking positions such as Forwards and Midfielders have a higher valuation compared to defensive positions such as Defenders and Goalkeepers.
- Attacking players directly influence the outcome of the match (i.e. Goals).
- Fans are more attracted to players with attacking flair → more marketable assets.



Club Squad Value

- Established teams in the league have the highest average player value and total player value.
- Newly promoted teams like Southampton, Leicester City and Ipswich Town have the lowest overall squad value.



Player Performance

From a football perspective, players of different positions should be evaluated on different statistical feature. To evaluate each player on a common ground, we would create a calculated **Performance Score** using a weighted score of each feature

$$\text{Performance Score} = 0.5A + 0.3B + 0.2C$$

Goalkeepers

A: Saves
B: Clean Sheets
C: Goals Prevented

Defenders

A: Tackles
B: Interceptions
C: Clean Sheets

Forwards

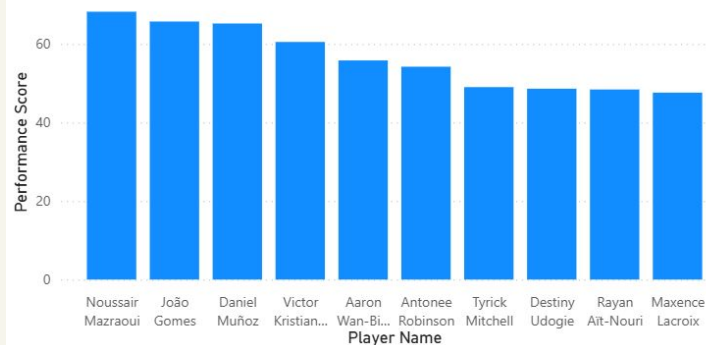
A: Goals
B: Assists
C: Conversion %

Midfielders

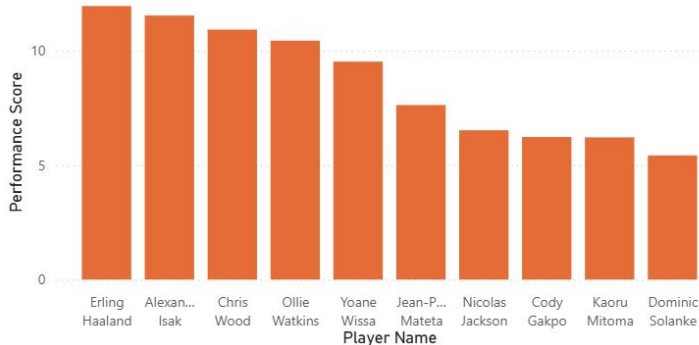
A: Assists
B: fThird Passes%
C: Tackles

Top 10 Performers By Positions

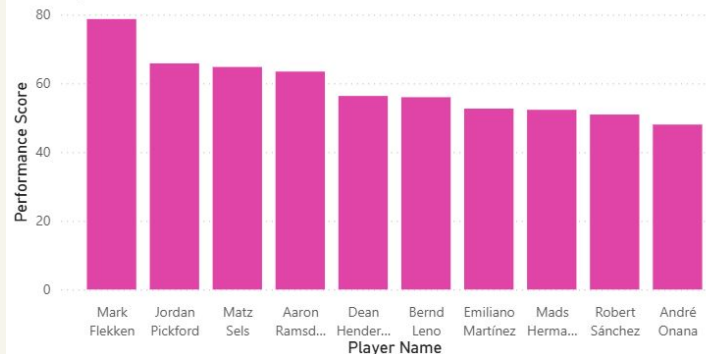
Defenders



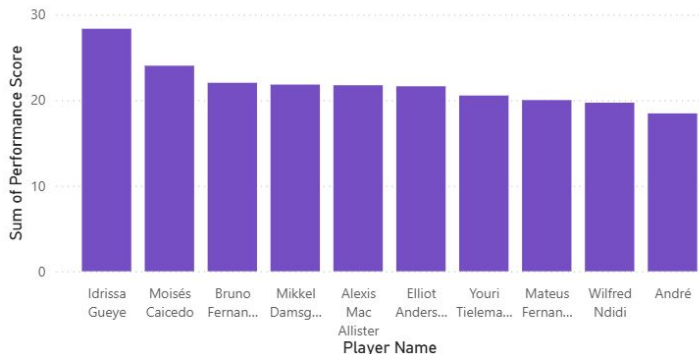
Forwards



Goalkeepers



Midfielders



Top Forwards: Erling Haaland, Alexander Isak, Chris Wood

Top Midfielders: Idrissa Gueye, Moises Caicedo, Bruno Fernandes

Top Defenders: Noussair Mazraoui, Joao Gomes, Daniel Munoz

Top Goalkeepers: Mark Flekken, Jordan Pickford, Matz Sels

Value-For-Money Player Analysis



Value-For-Money Player Analysis

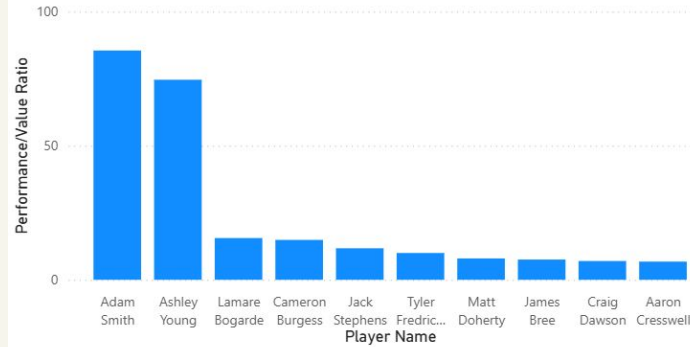
As the manager/team scout for a newly promoted team, the budget allocated for player recruitment is often very small in comparison to other established teams.

Hence, we want to identify players that are able perform on the pitch at a high-level while being reasonable/cheaply priced. We will do this by looking at a ratio between their performance score against their player value (per million euros).

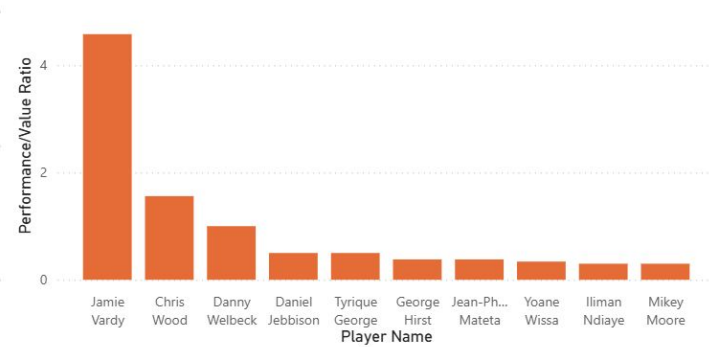
$$\text{Performance/Value Ratio} = \text{Performance Score} \div \text{Market Value}/1000000$$

Top 10 Value-For-Money Players By Positions

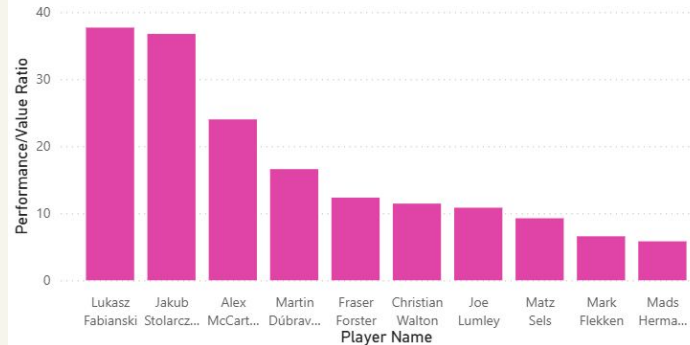
Defenders



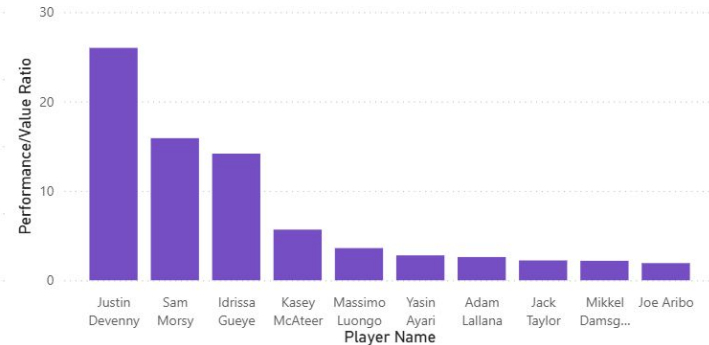
Forwards



Goalkeepers



Midfielders



Top Forwards: Jamie Vardy, Chris Wood, Danny Welbeck

Top Midfielders: Justin Devenny, Sam Morsy, Idrissa Gueye

Top Defenders: Adam Smith, Ashley Young, Lamare Bogarde

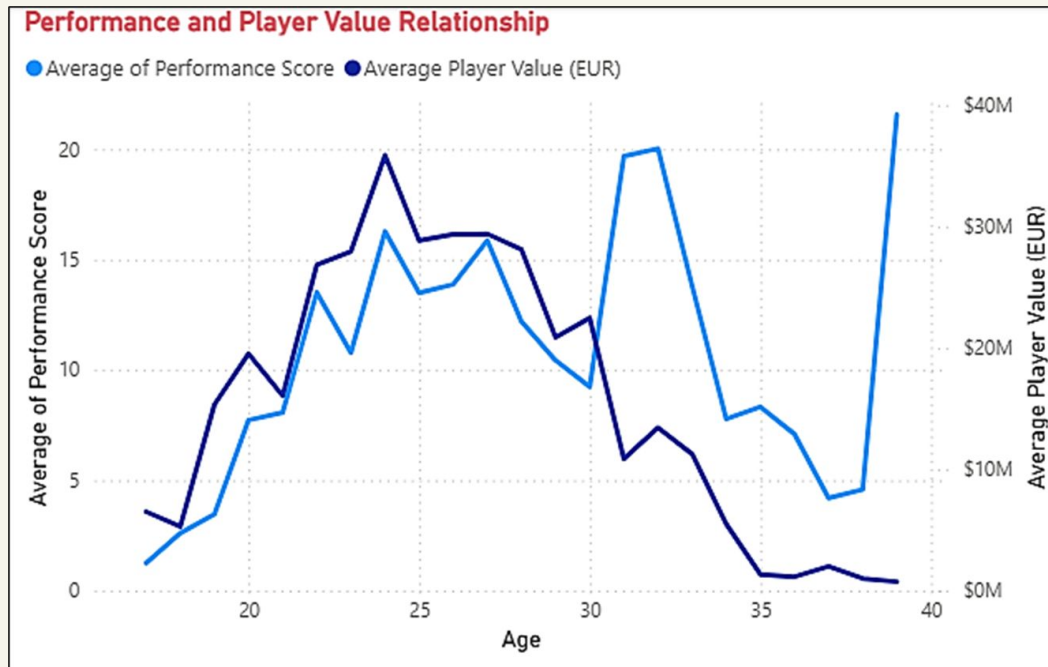
Top Goalkeepers: Lukasz Fabianski, Jakub Stolarczyk, Alex McCarthy

Recruitment Strategy



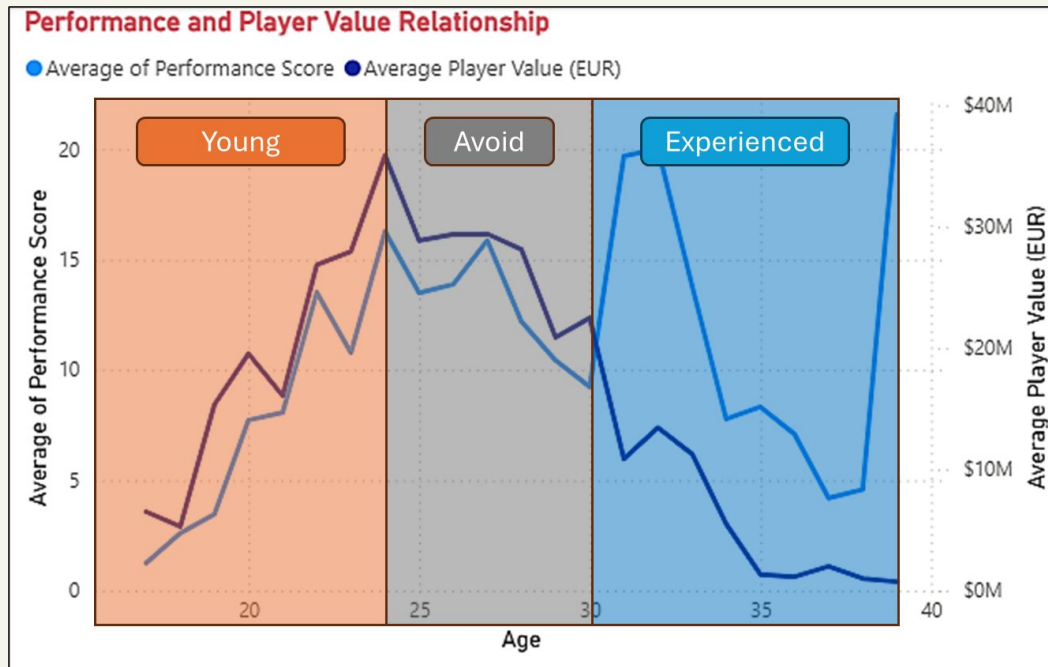
Player Recruitment

- Players are considered valuable assets much like commodities such as gold or silver.
- Player valuations may fluctuate based on an array of factors:
 - Player performance
 - Position
 - **Age**



Investment Strategy

- For **short-term** investments:
 - Experienced players that are past their peak valuations.
 - High performers.
- For **long-term** investments:
 - Young, inexperienced players below the age of 24.
 - The club can get a return on their investment when the player is sold at a potentially higher price.

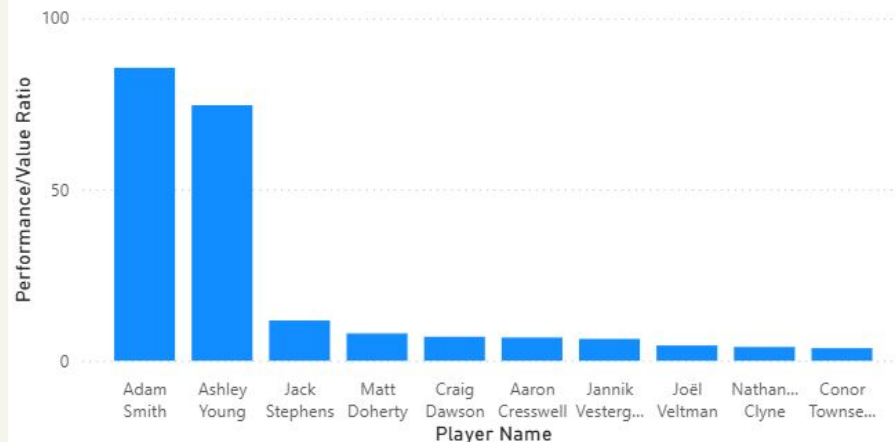


Player Recommendations (Experienced)

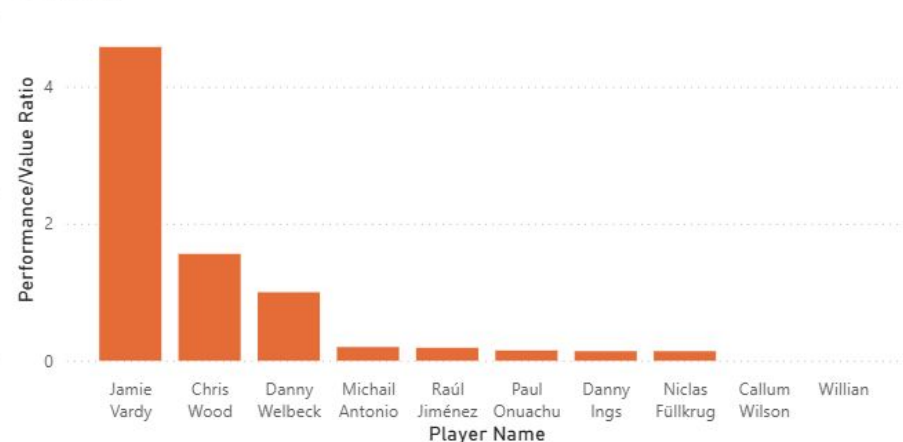


Top 10 Experienced Value-For-Money Players By Positions

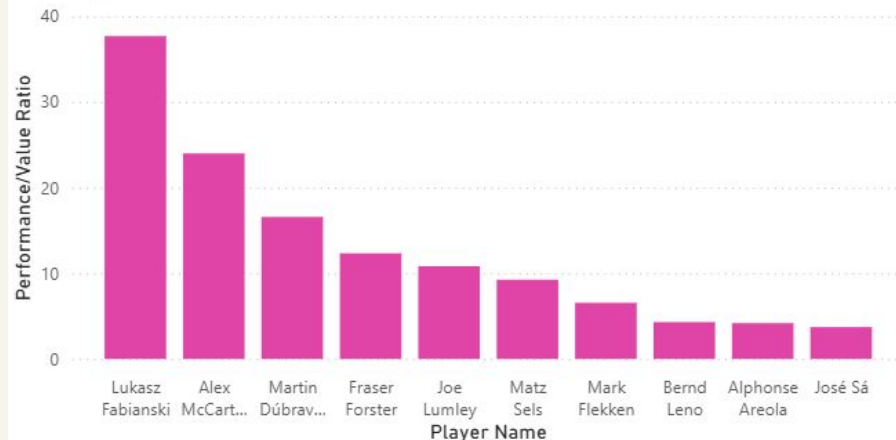
Defenders



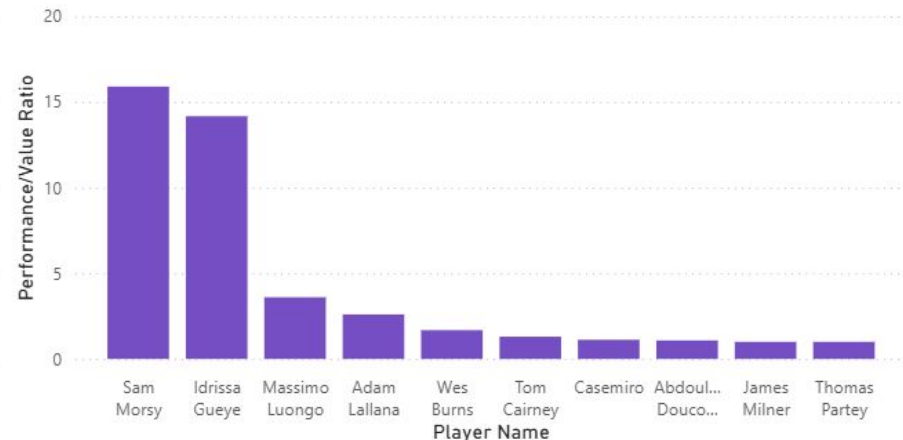
Forwards



Goalkeepers



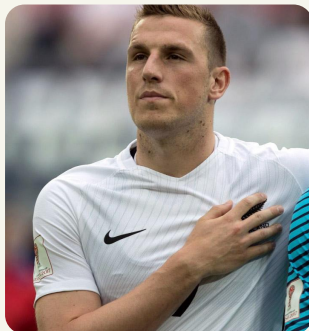
Midfielders



Forwards



**Jamie Vardy,
38**



**Chris Wood,
33**



**Danny
Welbeck, 34**

- All 3 players are seasoned EPL forwards with plenty of experience under their belt.
- Jamie Vardy provides the best value but considering his age of 38, we might only be able to secure his service for 1-2 years.
- Else, consider to secure Chris Wood and Danny Welbeck instead for a slightly longer stint (e.g. 3 years).

Midfielders



**Idrissa Gueye,
35**

- One of the top performing midfielders in the league for the season.
- Valued at only €2,000,000, you are guaranteed a high performing player that is a bargain for value.

Defenders



**Adam Smith,
32**

- 32-year-old defender valued at only €200,000.
- Able to bring aggression to the defence with his clearances and aerial duelling prowess.
- Decent on the ball for a defender with 79% passing accuracy and ability to contribute to the offense with his final third passing.

Goalkeepers



**Alex
McCarthy, 35**

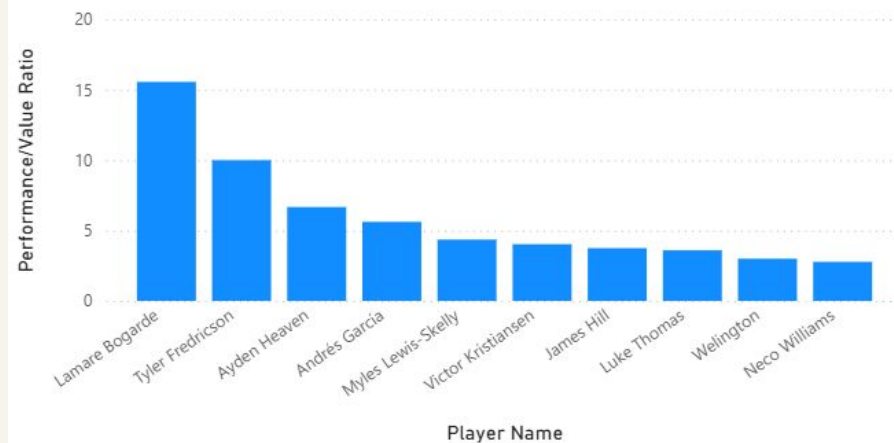
- Veteran PL goalkeeper bringing leadership and solid performances as a backup.
- Cheap backup goalkeeper option at 35 years old and valued at €500,000.

Player Recommendations (Young)

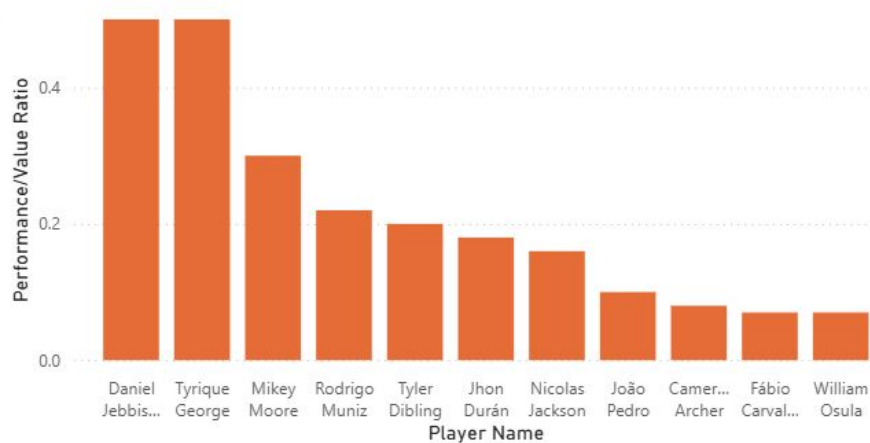


Top 10 Young Value-For-Money Players By Positions

Defenders



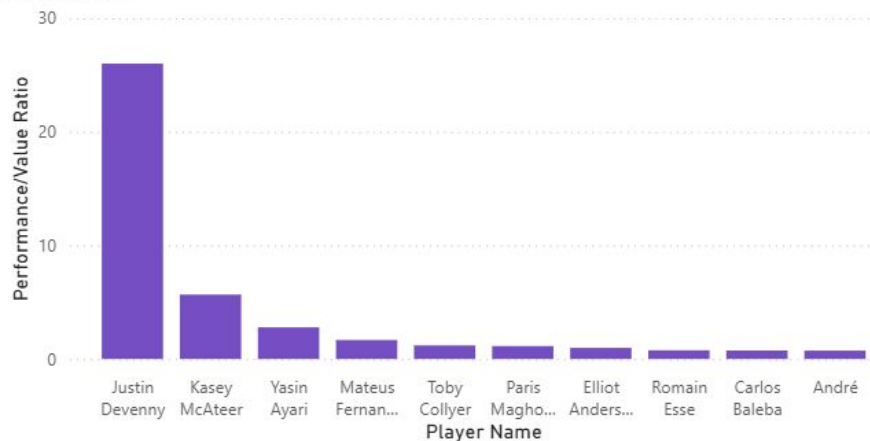
Forwards



Goalkeepers



Midfielders



Forwards



**Daniel
Jebbison, 21**



**Tyrique
George, 19**

- Both are young talents aged 21 and 18 respectively that are valued at €1,000,000 which could be cheap investments for the future.
- Both players are averaging ~10 mins of game time per appearance
- Consider convincing their current clubs to send them on loan for development/more game time opportunities.

Midfielders



**Justin
Devenney, 21**

- One of the top young midfielder prospects in the league
- Valued at only €150,000, he is a low risk investment with potentially high reward

Defenders



**Lamare
Bogarde, 21**



**Tyler
Fredricson, 19**

- Cheap and young defenders who are valued at <€500,000.
- Performing at a high performance/value comparable to senior players
- Any growth and development almost guarantee a positive ROI.

Goalkeepers



**Bart
Verbruggen, 22**

- While Antonin Kinsky has an overall high performance/value ratio, Bart Verbruggen is an experience starter in the league with 36 appearances and 3240 minutes played this season.
- At only 22 years old and currently valued at €22,000,000, he could prove to be a valuable asset as he approaches his prime.

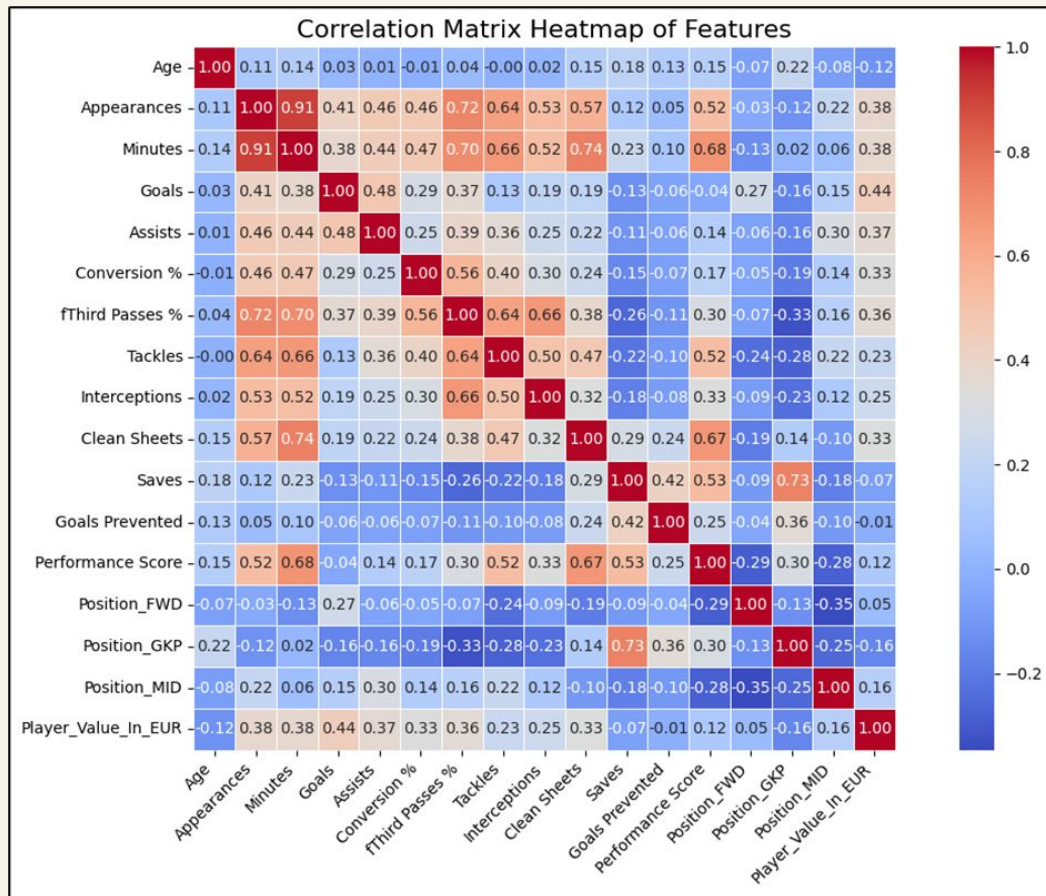
Player Value Prediction Model



Feature Selection

Features

'Age', 'Appearances', 'Minutes', 'Goals', 'Assists',
'Conversion %', 'fThird Passes %', 'Tackles',
'Interceptions', 'Clean Sheets', 'Saves', 'Goals
Prevented', 'Performance Score', 'Position_FWD',
'Position_GKP', 'Position_MID'



Model and Features Testing

DecisionTreeRegressor

RandomForestRegressor

LinearRegression

PolynomialRegression

XGBoost

KNNRegressor

GridSearchCV

Remove Outliers

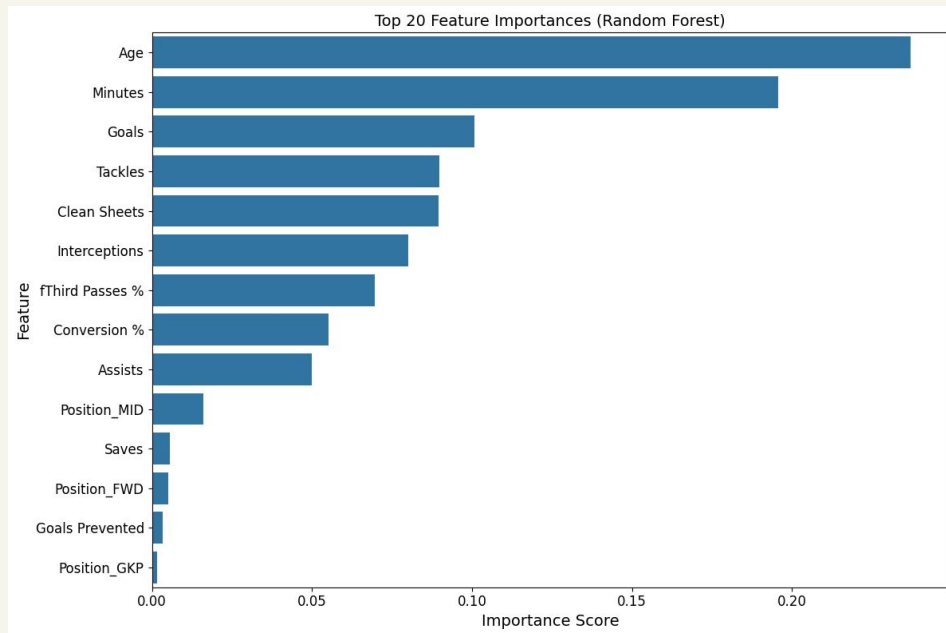
Features Tuning

Final Model

Final Features

'Age', 'Minutes', 'Goals', 'Assists', 'Conversion %',
'fThird Passes %', 'Tackles', 'Interceptions', 'Clean
Sheets', 'Saves', 'Goals Prevented', 'Position_FWD',
'Position_GKP', 'Position_MID'

Model	RMSE	R2
RandomForestRegressor	13098958	0.340



Moving Forward

Large variance expected with model due to high RMSE

Enhance model by expanding the dataset to other leagues

Include past historical data into consideration

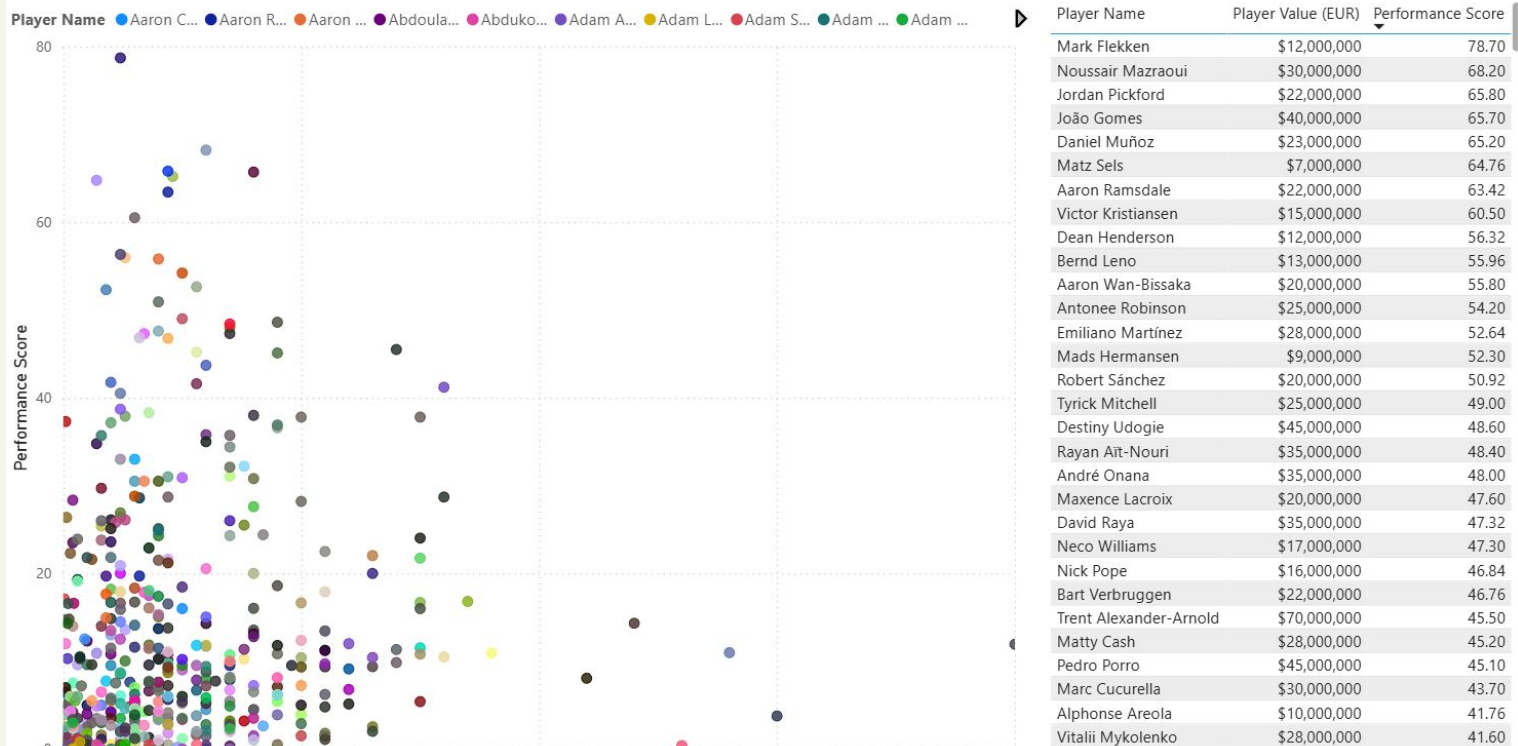
Re-evaluate metrics used for Performance Score

Questions?



Model	Feature Set	RMSE	R ²	Remarks
DecisionTreeRegressor	feature_cols	23737968	-0.394	-
DecisionTreeRegressor	feature_cols_2	24670514	-0.505	-
RandomForestRegressor	feature_cols	16115671	0.358	-
RandomForestRegressor	feature_cols	15333680	0.419	With GridSearchCV hyperparameter tuning
RandomForestRegressor	feature_cols_2	15978510	0.369	-
RandomForestRegressor	feature_cols_2	15303500	0.421	With GridSearchCV hyperparameter tuning
Removed outliers from dataset				
RandomForestRegressor	feature_cols	13428963	0.306	
RandomForestRegressor	feature_cols	13415802	0.307	GridSearchCV hyperparameter tuning
LinearRegression	feature_cols	14226678	0.221	
LinearRegression	feature_cols_2	14198151	0.224	
PolynomialRegression	feature_cols	14226678	0.221	
PolynomialRegression	feature_cols_2	14198151	0.224	
XGBoost	feature_cols	14208326	0.223	
XGBoost	feature_cols	13953906	0.251	GridSearchCV hyperparameter tuning
RandomForestRegressor	feature_cols_3_refined	13737990	0.274	
RandomForestRegressor	feature_cols_3_refined	13430079	0.306	GridSearchCV hyperparameter tuning
KNNRegressor	feature_cols_3_refined	13777584	0.269	
KNNRegressor	feature_cols_3_refined	14063634	0.239	GridSearchCV hyperparameter tuning

Performance Score vs Player Value (EUR)



1. Mark Flekken, Noussair Mazraoui and Jordan Pickford forms the top 3 performers for the 24/25 season.
2. When looking at the overall distribution of performance against the player value, we can see that they are high performers who are greatly outperforming their market value.
3. In comparison, highly valued players (>€100,000,000) like Bukayo Saka, Declan Rice and falls in the lower end of the spectrum.