

PLAYER VALUATION SOCCER

Presented By:

- *Abhijit Nelli*
- *Jash Patel*
- *Karthik Iyer*
- *Mohit Jain*



The Global Football Marketplace: The Challenge of Player Valuation

Football isn't just a game; it's a billion-dollar global industry.

- Clubs compete fiercely for top talent, breaking transfer records annually.
- Yet, determining a player's market value remains a complex challenge.

Key Questions:

- What drives a player's market value?
- How do clubs make valuation decisions amidst fierce competition?

The Stakeholders:

- Who? Clubs, agents, analysts, and fans are eager to evaluate player worth.
- What? Valuation blends skills, performance, marketability, and potential.
- Why? Smarter investments mean competitive advantage.

The Challenge:

Balancing subjective assessments with data-driven insights in a fast-changing industry.

Our Data: Building a Foundation for Player Valuation

Data Source:

- Kaggle's FIFA dataset: A comprehensive repository of player attributes, performance metrics, and market values.

Steps Taken:

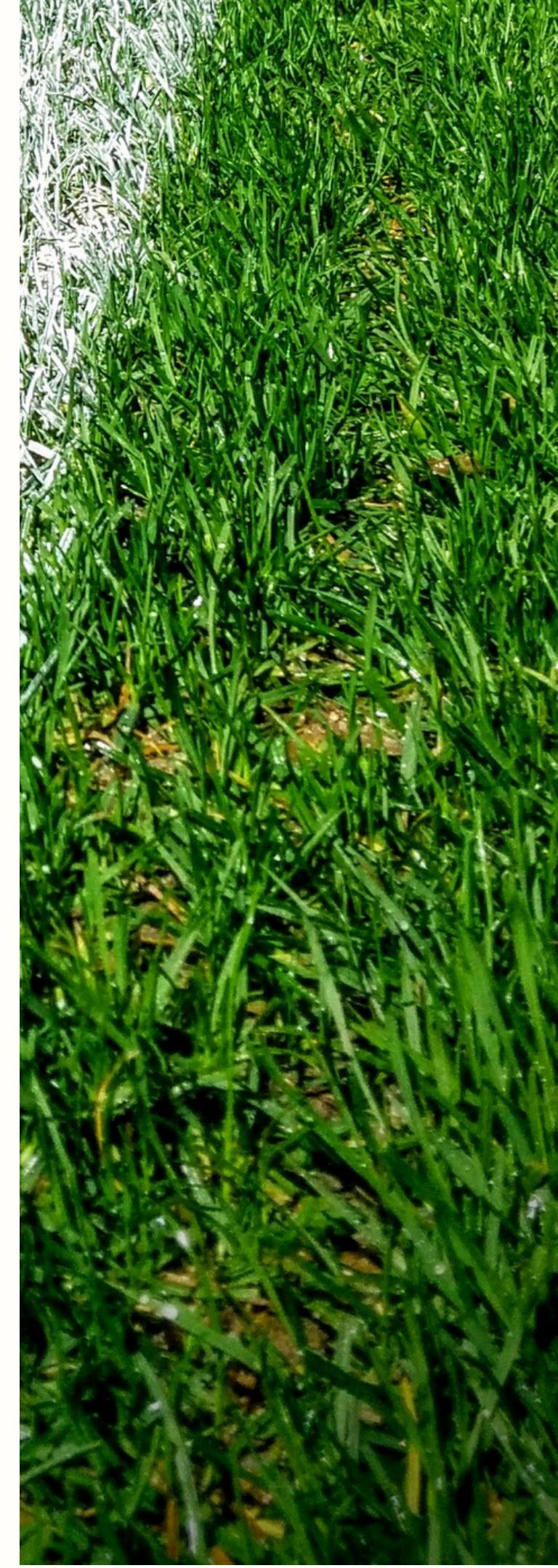
1. Data Cleaning: Addressing missing values, removing redundancies, and normalizing formats.
2. Feature Engineering: Created new variables like age difference from peak performance and position-specific multipliers.
3. Standardization: Aligned values across versions to ensure comparability.

Unique Challenges:

- Merging datasets from different FIFA versions.
- Handling categorical variables like player position and league context.
- Ensuring consistency in currency conversions for market values.

Why This Matters:

Clean, well-structured data is essential for reliable predictions and actionable insights.



Understanding Our Dataset at a Glance

	player_id	fifa_version	short_name	long_name	player_positions	overall	potential	value_eur	wage_eur	pace	shooting	passing	dribbling	defending	physic	age	height_cm	weight_kg	league_id
0	158023	15	L. Messi	Lionel Andrés Messi Cuccittini	CF	93	95	100500000.0	550000.0	93.0	89.0	86.0	96.0	27.0	63.0	27	169	67	53.0
1	20801	15	Cristiano Ronaldo	Cristiano Ronaldo dos Santos Aveiro	LW, LM	92	92	79000000.0	375000.0	93.0	93.0	81.0	91.0	32.0	79.0	29	185	80	53.0
2	9014	15	A. Robben	Arjen Robben	RM, LM, RW	90	90	54500000.0	275000.0	93.0	86.0	83.0	92.0	32.0	64.0	30	180	80	19.0
3	41236	15	Z. Ibrahimović	Zlatan Ibrahimović	ST	90	90	52500000.0	275000.0	76.0	91.0	81.0	86.0	34.0	86.0	32	195	95	16.0
4	167495	15	M. Neuer	Manuel Peter Neuer	GK	90	90	63500000.0	300000.0	NaN	NaN	NaN	NaN	NaN	NaN	28	193	92	19.0



Understanding Our Dataset at a Glance

Raw dataset:

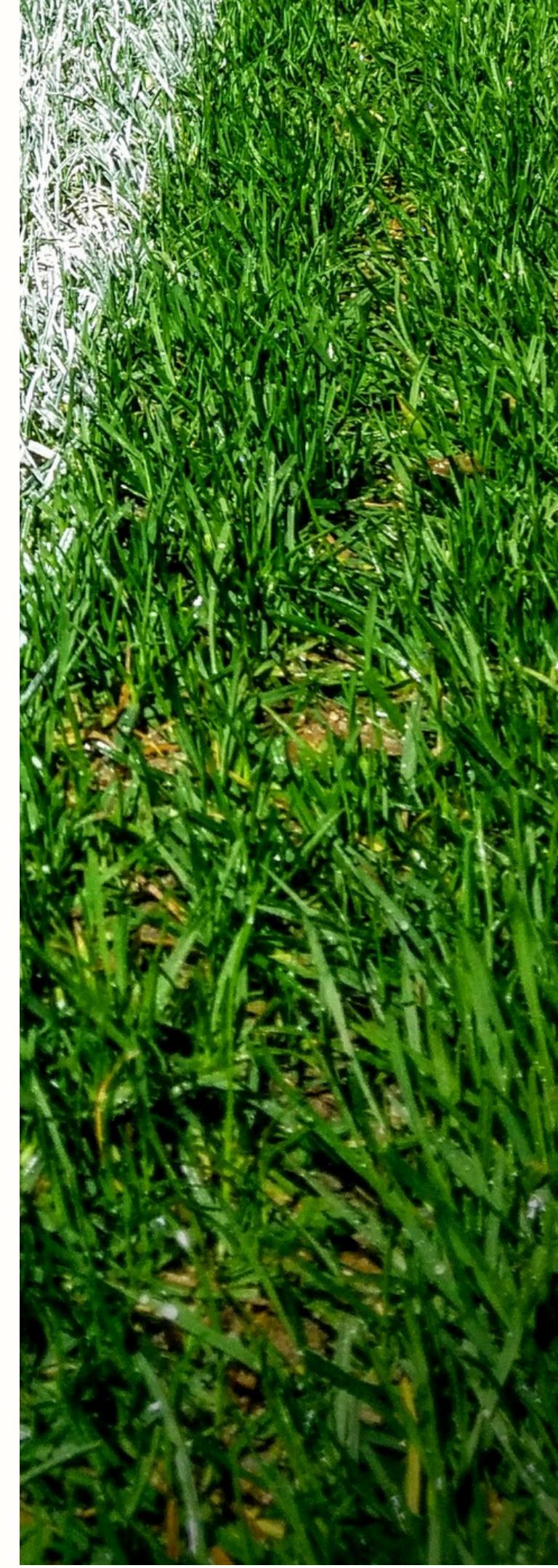
(161583, 110)

unique players ids: 49699

unique clubs ids: 1056

unique players urls: 161583

unique players long names: 50360



Understanding Our Dataset at a Glance

Dataset Highlights:

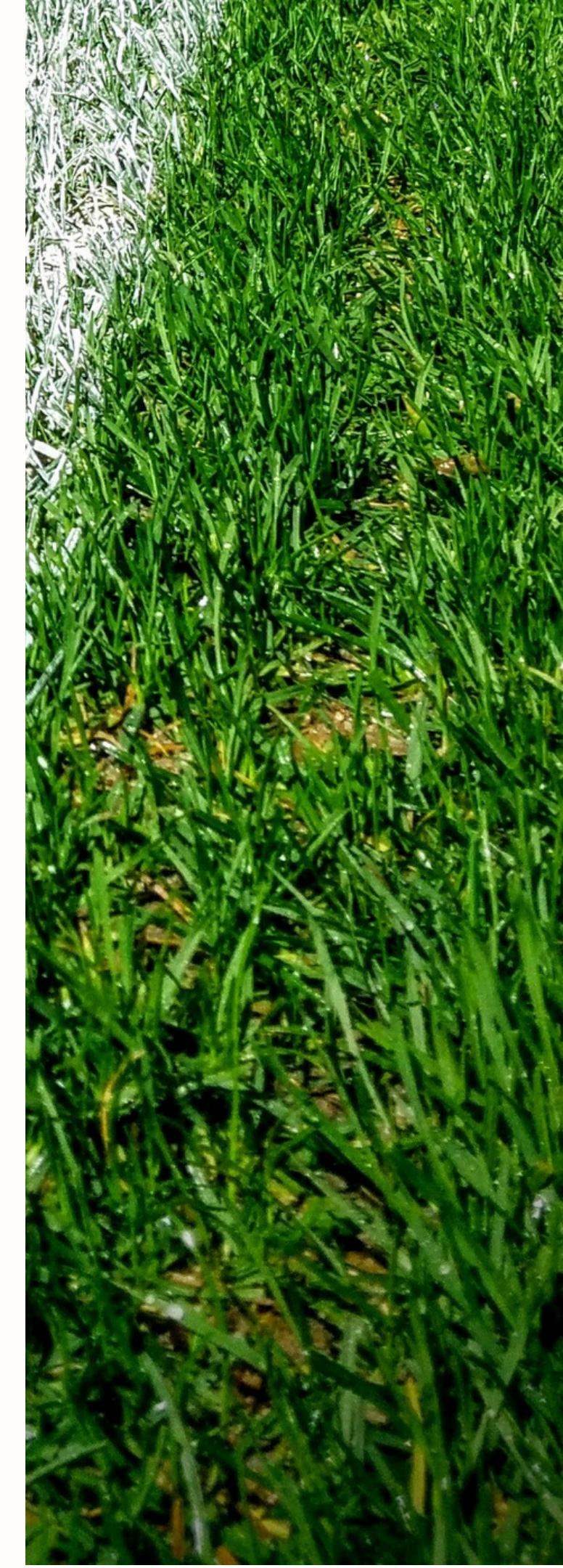
- Player metrics cover performance, potential, and physical characteristics.
- Core attributes include:
 - General stats: Overall rating, potential, age, height, weight, value, wage.
 - Skills: Shooting, passing, dribbling, defending, and physical strength.
 - Specialized metrics: Goalkeeping abilities, attacking skills, and movement precision.
 - Year-wise data: Tracks progression across different seasons.
- Comprehensive data helps analyze players' current abilities, future potential, and positional suitability.



A Deep Dive into Our Dataset

```
['short_name',  
 'player_positions',  
 'overall',  
 'potential',  
 'value_eur',  
 'wage_eur',  
 'age',  
 'height_cm',  
 'weight_kg',  
 'pace',  
 'shooting',  
 'passing',  
 'dribbling',  
 'defending',  
 'physic',  
 'attacking_crossing',  
 'attacking_finishing',  
 'attacking_heading_accuracy',  
 'attacking_short_passing',  
 'attacking_volleys',  
 'skill_dribbling',  
 'skill_curve',  
 'skill_fk_accuracy',  
 'skill_long_passing',  
 'skill_ball_control',  
 'movement_acceleration',  
 'movement_sprint_speed',  
 'movement_agility',  
 'movement_reactions',  
 'movement_balance',  
 'power_shot_power',  
 'power_jumping',  
 'power_stamina',  
 'power_strength',  
 'power_long_shots',  
 'mentality_aggression',  
 'mentality_interceptions',  
 'mentality_positioning',  
 'mentality_vision',  
 'mentality_penalties',  
 'defending_marking_awareness',  
 'defending_standing_tackle',  
 'defending_sliding_tackle',  
 'goalkeeping_diving',  
 'goalkeeping_handling',  
 'goalkeeping_kicking',  
 'goalkeeping_positioning',  
 'goalkeeping_reflexes',  
 'goalkeeping_speed',  
 'year']
```

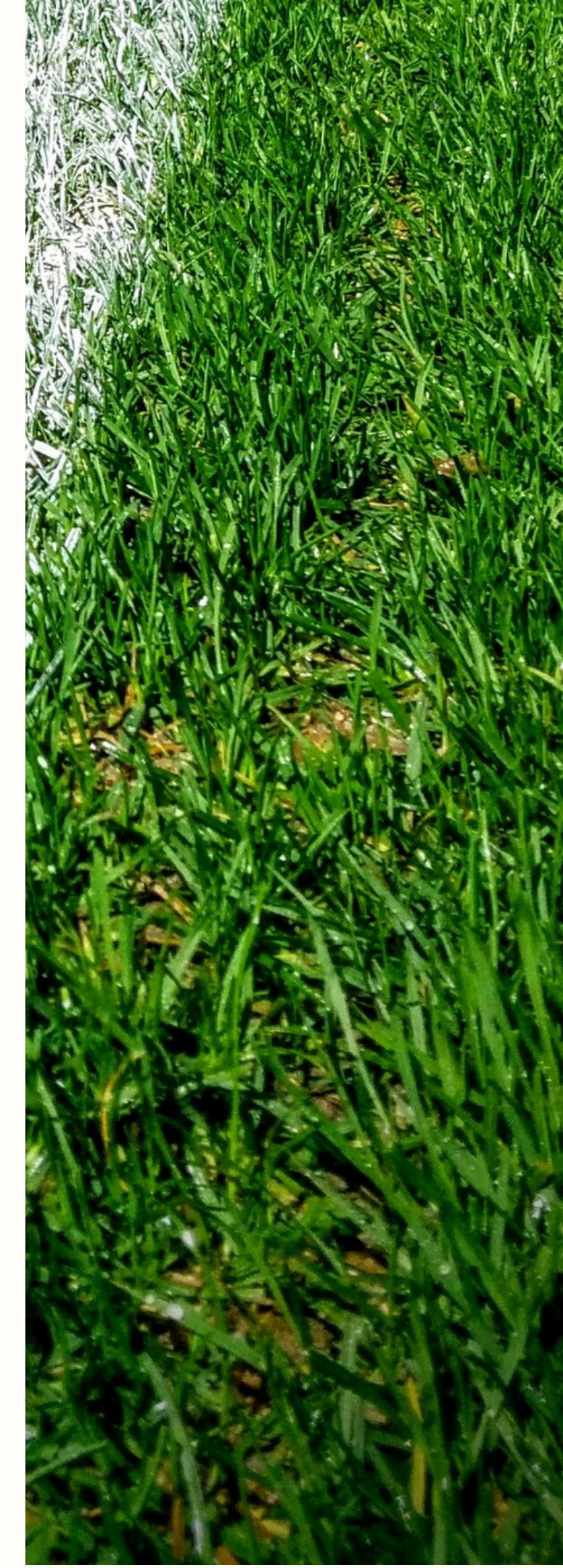
Dataset Columns



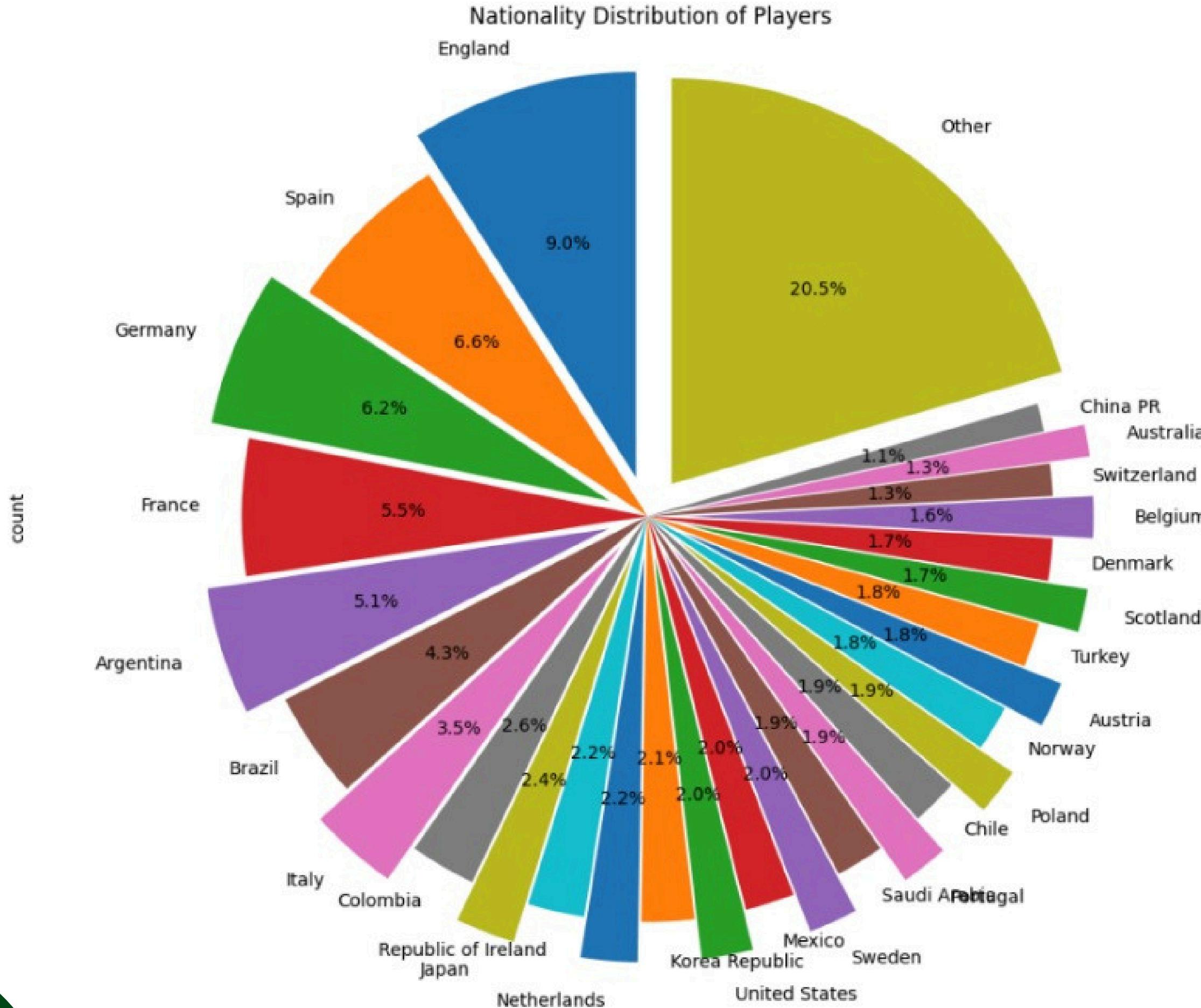
Player Position Distribution

player_positions	count
CM	3303
CB	3272
ST	2979
CDM	2244
LM	2204
RM	2095
CAM	2041
RB	2016
LB	1869
GK	1775
RW	911
LW	839
CF	667
LWB	100
RWB	82

- The dataset covers 13 distinct positions, showcasing the diversity of roles in football.
- Top 3 positions by count:
 - Central Midfielders (CM): 3,303 players.
 - Centre-Backs (CB): 3,272 players.
 - Strikers (ST): 2,979 players.
- Specialized positions like Right Wing-Backs (RWB) and Left Wing-Backs (LWB) are less common.
- Players often have the flexibility to play in multiple positions, but for analysis, we evaluate each position individually.

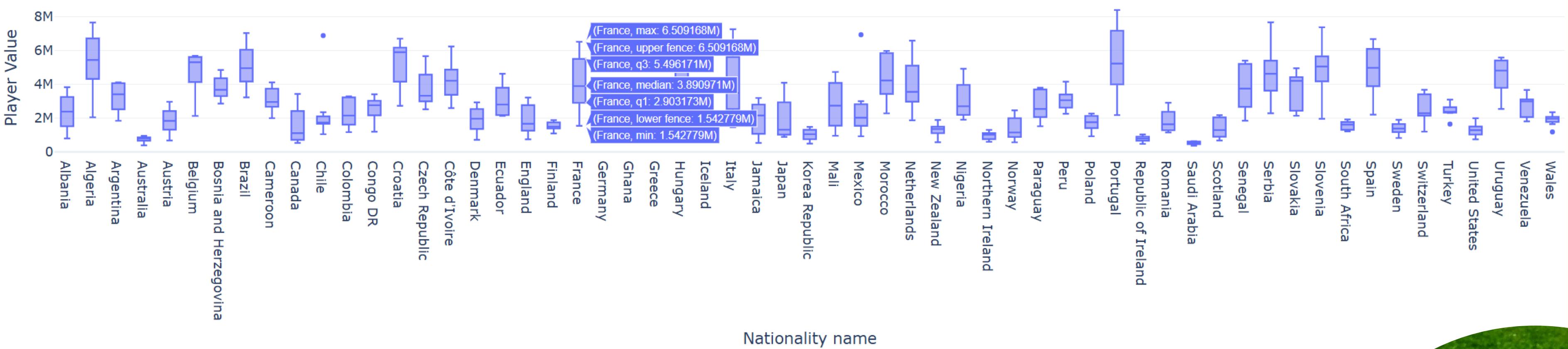


Mapping the Landscape of Player Data



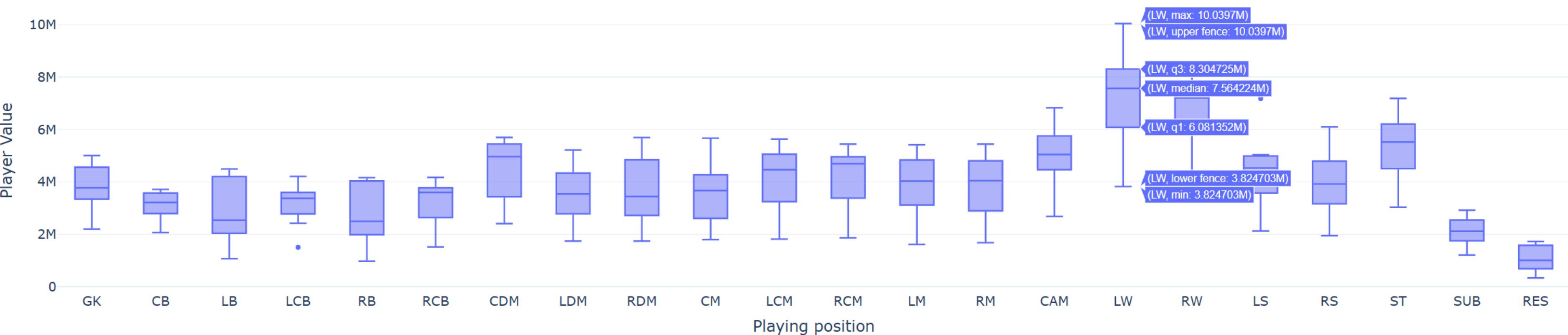
Unveiling Insights: Nationality, Age, and Position Impact on Player Value

Distribution of player value by nationality name



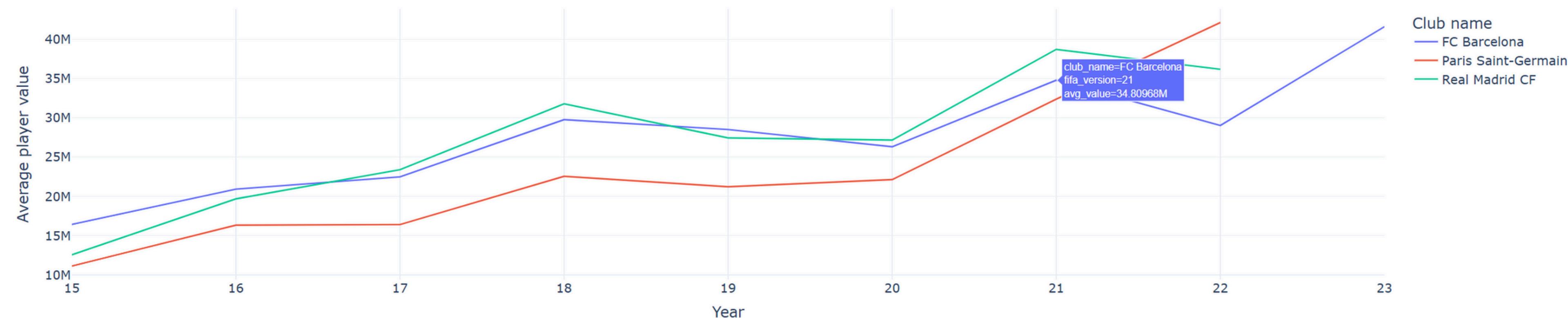
Unveiling Insights: Nationality, Age, and Position Impact on Player Value

Distribution of player value by playing position

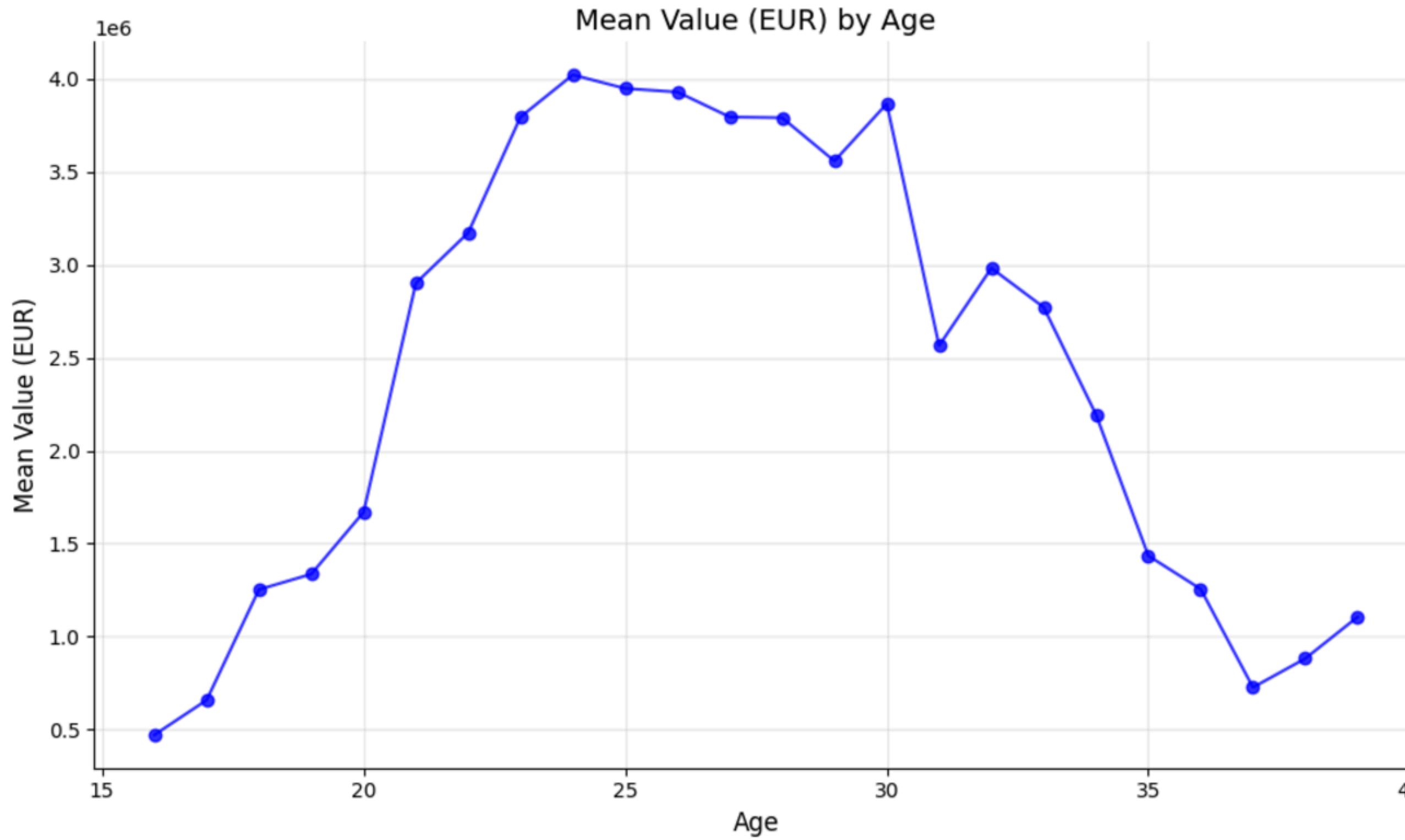


Unveiling Insights: Nationality, Age, and Position Impact on Player Value

Average player value over years for multiple clubs

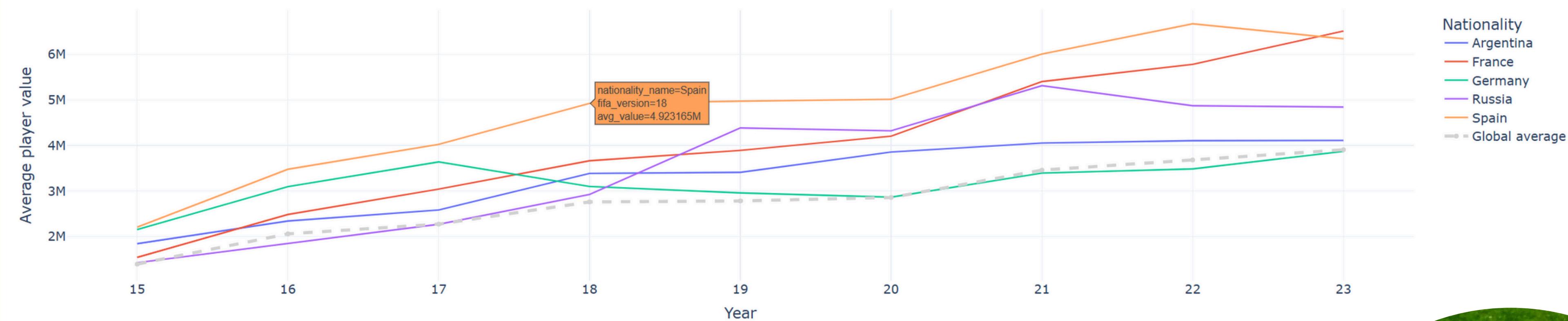


Unveiling Insights: Nationality, Age, and Position Impact on Player Value

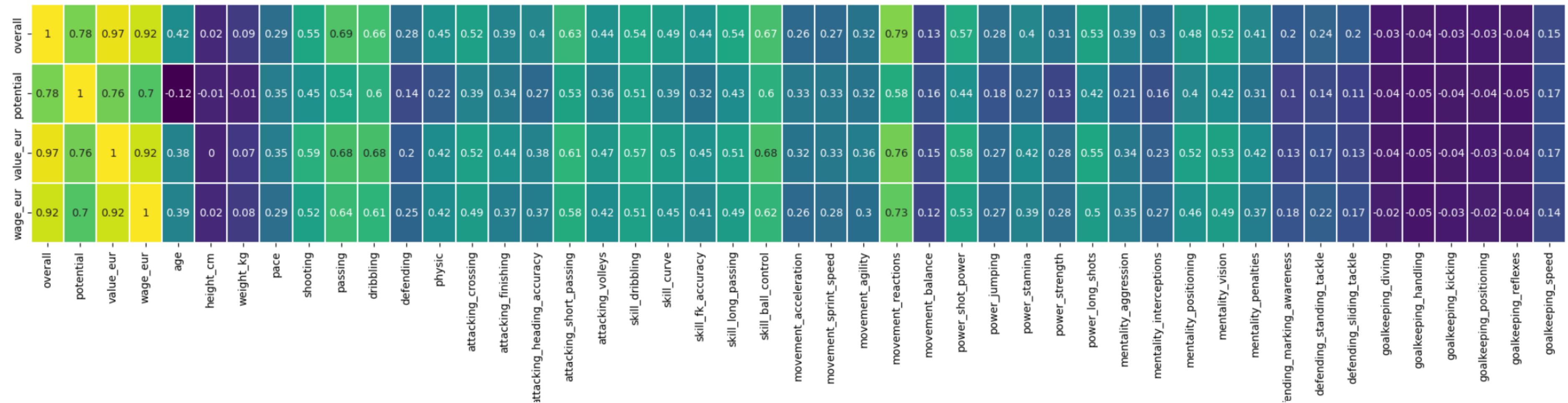


Unveiling Insights: Nationality, Age, and Position Impact on Player Value

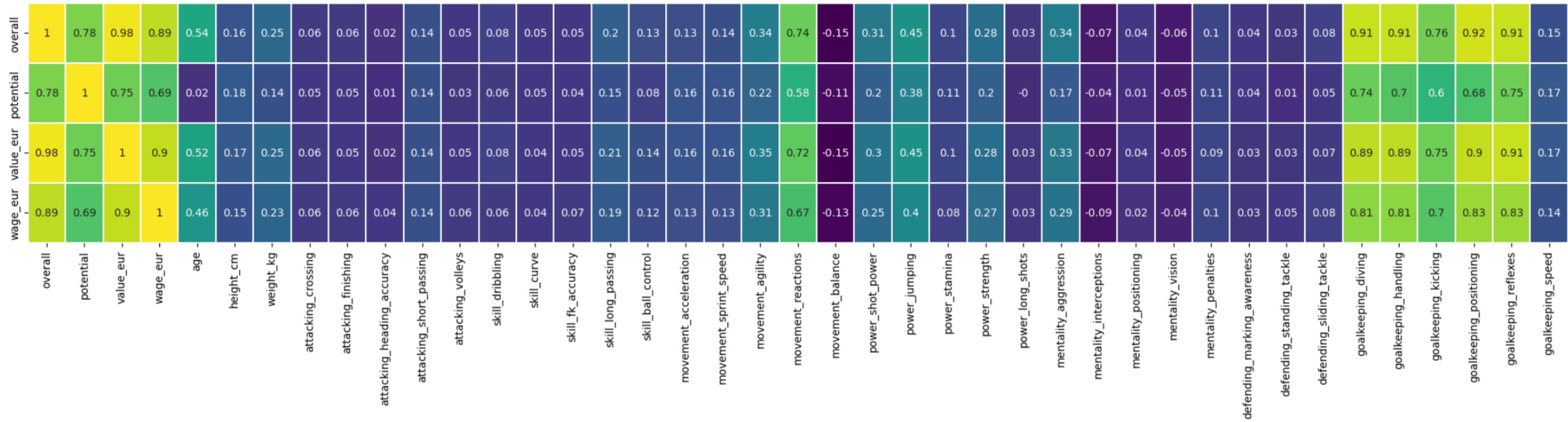
Average player value over years for multiple nations



Exploring Relationships: Attribute Correlation Analysis



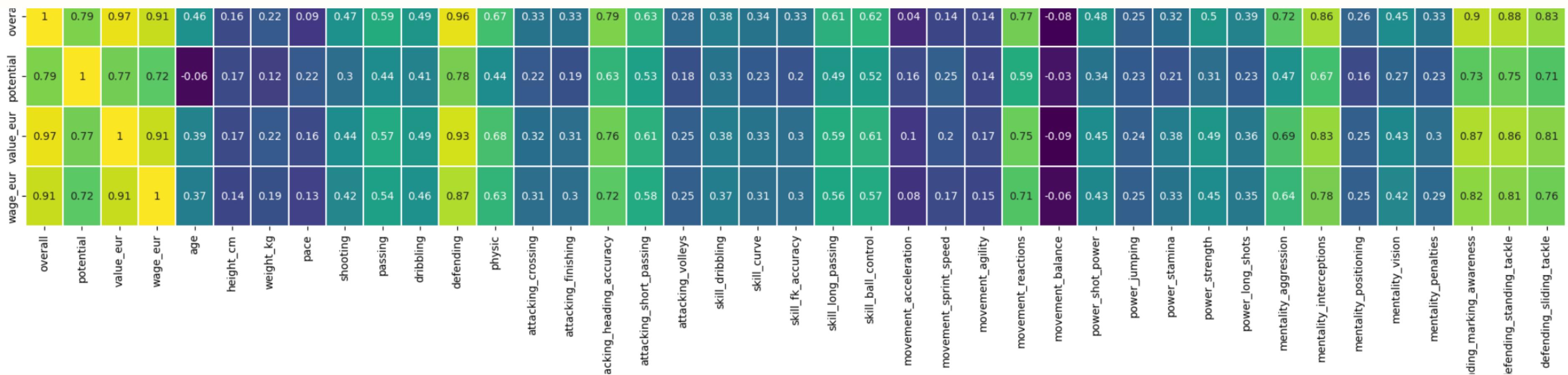
Exploring Relationships: Attribute Correlation Analysis



Goalkeeper Correlation Matrix:

- Focused on attributes critical for goalkeepers (e.g., diving, reflexes, positioning).
- Highlights how goalkeeping-specific metrics interrelate.

Exploring Relationships: Attribute Correlation Analysis



Center Back (CB) Correlation Matrix:

- Attributes relevant to defensive roles (e.g., strength, interceptions, standing tackles).
- Evaluates defensive performance factors.

Exploring Relationships: Attribute Correlation Analysis

	overall	potential	value_eur	wage_eur	age	height_cm	weight_kg	pace	shooting	passing	dribbling	defending	physic	attacking_crossing	attacking_finishing	attacking_heading_accuracy	attacking_short_passing	attacking_volleys	skill_dribbling	skill_curve	skill_fk_accuracy	skill_long_passing	movement_acceleration	movement_sprint_speed	movement_agility	movement_reactions	movement_balance	power_shot_power	power_jumping	power_stamina	power_strength	power_long_shots	mentality_aggression	mentality_interceptions	mentality_positioning	mentality_vision	mentality_penalties	mentality_marking_awareness	defending_standing_tackle	defending_sliding_tackle	
overall	1	0.76	0.98	0.93	0.51	0.09	0.19	0.34	0.92	0.75	0.82	0.44	0.54	0.61	0.83	0.6	0.75	0.79	0.77	0.64	0.57	0.55	0.87	0.31	0.35	0.35	0.83	0.06	0.82	0.37	0.49	0.39	0.75	0.45	0.29	0.86	0.65	0.56	0.1	0.2	0.1
potential	0.76	1	0.74	0.68	-0.04	0.05	0.07	0.38	0.7	0.55	0.7	0.23	0.29	0.43	0.66	0.39	0.59	0.57	0.67	0.48	0.39	0.41	0.72	0.36	0.38	0.34	0.59	0.1	0.61	0.26	0.32	0.2	0.57	0.23	0.13	0.62	0.45	0.41	0	0.09	0.05
value_eur	0.98	0.74	1	0.93	0.48	0.1	0.19	0.36	0.9	0.72	0.8	0.43	0.55	0.58	0.81	0.58	0.72	0.76	0.76	0.61	0.54	0.52	0.84	0.33	0.37	0.36	0.81	0.06	0.8	0.37	0.52	0.39	0.73	0.44	0.28	0.84	0.62	0.53	0.09	0.19	0.09
wage_eur	-0.93	0.68	0.93	1	0.5	0.09	0.18	0.32	0.85	0.7	0.75	0.42	0.52	0.56	0.76	0.55	0.71	0.74	0.71	0.59	0.53	0.52	0.8	0.29	0.33	0.33	0.77	0.05	0.76	0.34	0.48	0.37	0.7	0.44	0.28	0.8	0.62	0.51	0.09	0.19	0.09

Striker (ST) Correlation Matrix:

- Concentrates on offensive metrics (e.g., finishing, positioning, sprint speed).
 - Identifies key drivers for scoring and attacking play.



Breaking Down OBPM & DBPM: A New Lens on Player Value

- OBPM highlights a player's offensive prowess, focusing on skills that directly influence attacking plays.
- Key Attributes Assessed:
 - Pace: How quickly a player can move with the ball, creating opportunities.
 - Shooting: The accuracy and power of shots, critical for goal-scoring.
 - Passing: The ability to distribute the ball effectively and create scoring chances.
 - Dribbling: Skill in maneuvering past opponents, maintaining possession, and opening up spaces.
- DBPM emphasizes defensive contributions, assessing attributes that mitigate the opposition's threats.
- Key Attributes Assessed:
 - Defending: Tackling, interceptions, and overall defensive positioning.
 - Physic: Physicality in duels, stamina, and strength to maintain defensive stability.
 - Pace: Speed in recovery and tracking back to defend against counterattacks.

Offensive Box Plus-Minus (OBPM) vs Defensive Box Plus-Minus (DBPM)

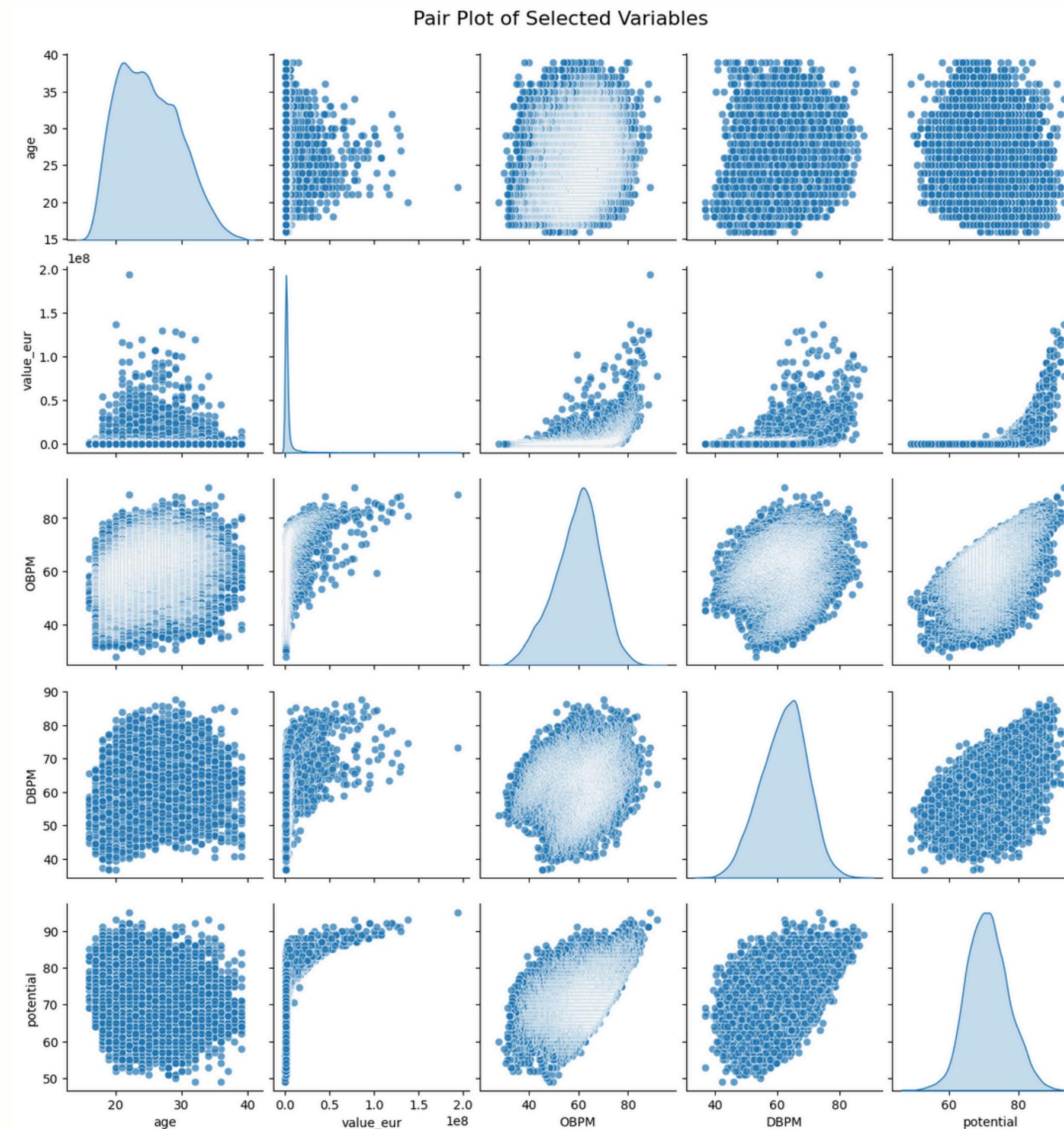
	CAM	CB	CDM	CF	CM	LB	LM	LW	LWB	RB	RM	RW	RWB	ST
pace	0.10	0.16	0.09	0.13	0.09	0.17	0.13	0.14	0.16	0.16	0.12	0.13	0.16	0.12
shooting	0.27	0.22	0.26	0.28	0.26	0.23	0.25	0.28	0.22	0.22	0.26	0.28	0.20	0.32
dribbling	0.32	0.31	0.33	0.31	0.33	0.30	0.32	0.31	0.32	0.30	0.32	0.30	0.33	0.30
passing	0.30	0.31	0.33	0.28	0.33	0.31	0.29	0.28	0.30	0.31	0.30	0.28	0.31	0.27

OBPM : Evaluates offensive contribution

	CAM	CB	CDM	CF	CM	LB	LM	LW	LWB	RB	RM	RW	RWB	ST
defending	0.31	0.53	0.53	0.35	0.46	0.47	0.21	0.22	0.44	0.49	0.25	0.27	0.43	0.32
physic	0.38	0.31	0.32	0.30	0.34	0.28	0.37	0.35	0.30	0.28	0.37	0.34	0.31	0.38
pace	0.30	0.17	0.15	0.35	0.20	0.24	0.42	0.43	0.26	0.23	0.38	0.39	0.26	0.30

DBPM : Measures defensive impact

Exploring the Interplay: Age, Value, OBPM, DBPM, and Potential



Highlights correlations, trends, and outliers, aiding in identifying high-value, high-potential, and balanced players for strategic decision-making.

Prediction Model for Player Value

Objective:

- To predict player value for FIFA 23 using a regression model trained on FIFA 22 data. Key features include Age, OBPM, and DBPM, reflecting offensive and defensive contributions.

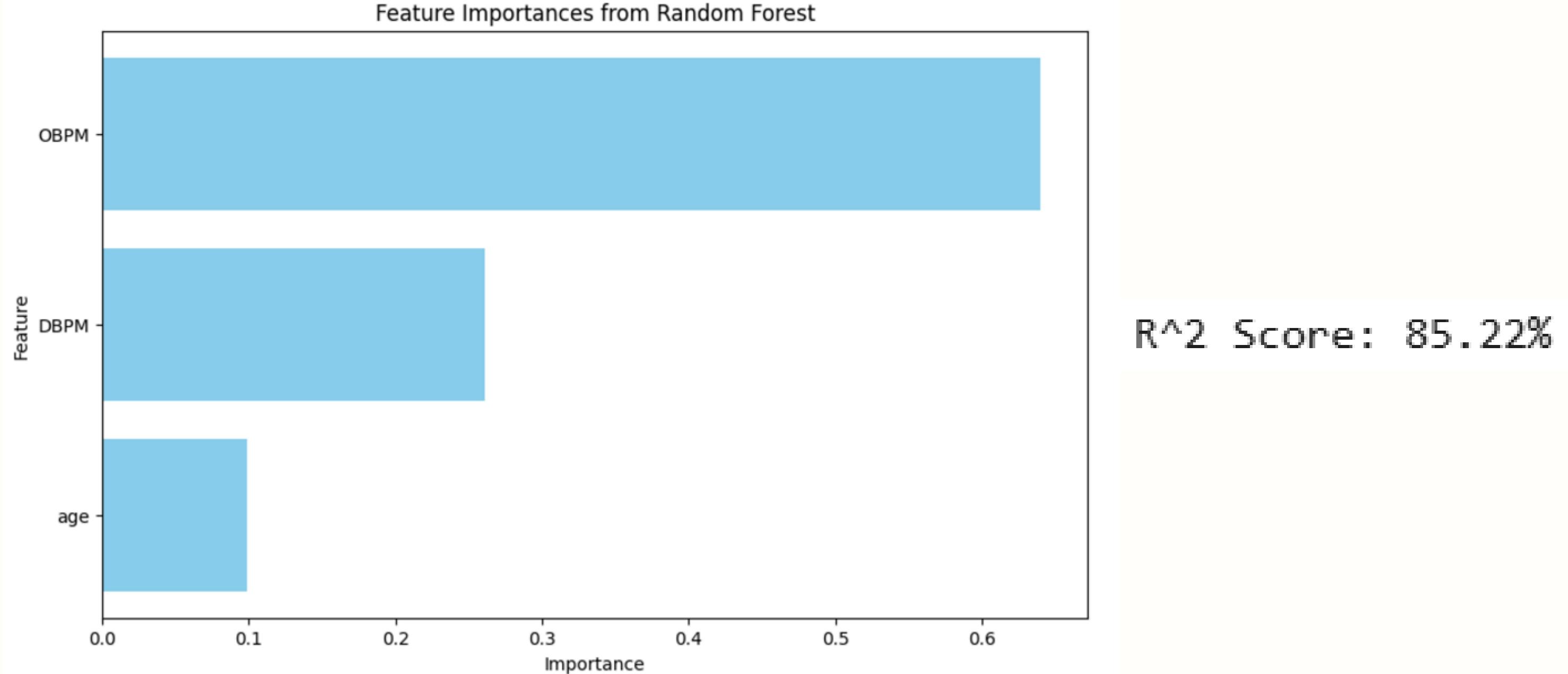
```
1 x_train.head(5)
```

	age	OBPM	DBPM
1642	27.0	74.855000	68.510000
13166	27.0	69.266667	69.343333
3072	23.0	51.650000	72.420000
810	18.0	41.280000	55.990000
7196	24.0	57.630000	72.920000

```
1 y_train.head(5)
```

	value_eur
1642	15000000.0
13166	23000000.0
3072	5500000.0
810	190000.0
7196	7000000.0

Understanding Feature Importance in Player Valuation



The feature importance analysis reveals that OBPM is the key driver in player value prediction. With an accuracy of 85.22%, the model demonstrates strong predictive power.

Unveiling the Value: Model Predictions vs Reality

long_name	value_eur
Kylian Mbappé Lottin	190.5
Erling Braut Haaland	148.0
Frenkie de Jong	116.5
Pedro González López	116.0
Mohamed Salah Ghaly	115.5
Philip Foden	109.5
Vinícius José Paixão de Oliveira Júnior	109.0
Kevin De Bruyne	107.5
Rúben dos Santos Gato Alves Dias	107.5
Harry Kane	105.5
Joshua Walter Kimmich	105.5
손흥민 孙兴慜	101.0
Trent Alexander-Arnold	100.5
Sadio Mané	99.5
Neymar da Silva Santos Júnior	99.5

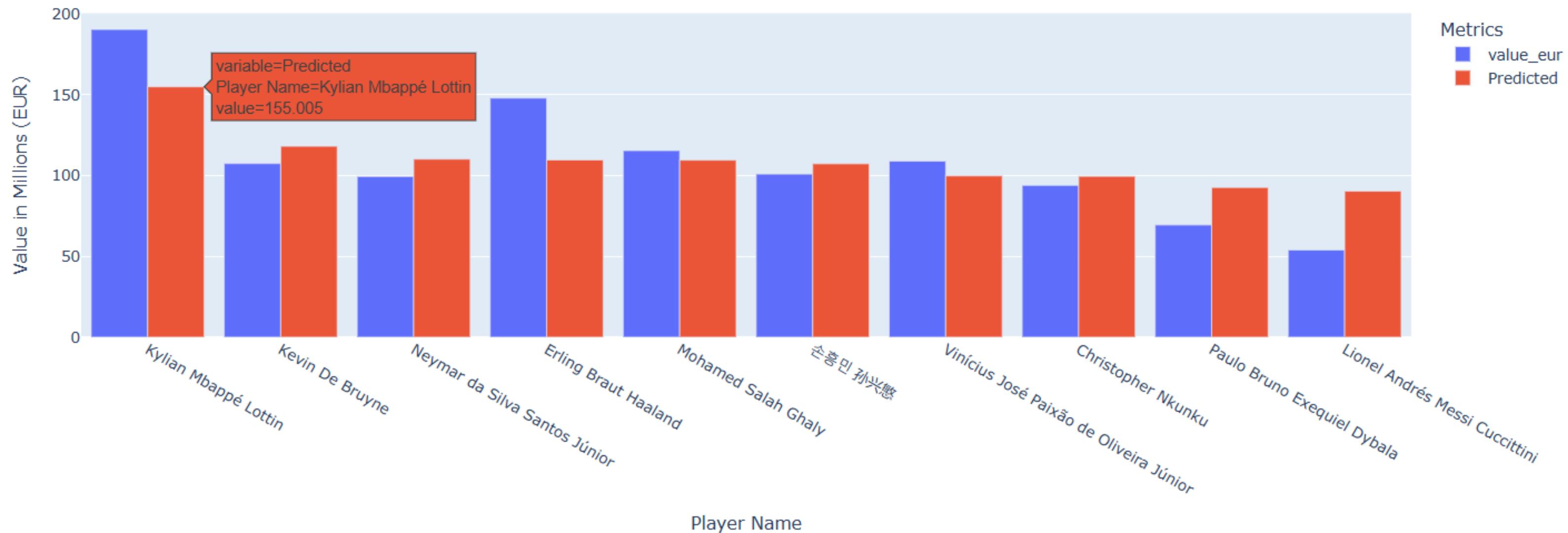
Top Real Value

Top Predicted Values

long_name	Predicted
Kylian Mbappé Lottin	155.005
Kevin De Bruyne	118.220
Neymar da Silva Santos Júnior	110.195
Erling Braut Haaland	109.695
Mohamed Salah Ghaly	109.565
손흥민 孙兴慜	107.365
Vinícius José Paixão de Oliveira Júnior	99.880
Christopher Nkunku	99.610
Paulo Bruno Exequiel Dybala	92.690
Lionel Andrés Messi Cuccittini	90.365
João Félix Sequeira	88.840
Karim Benzema	85.145
Presnel Kimpembe	81.840
Jude Victor William Bellingham	80.260
Mikel Merino Zazón	79.290

Unveiling the Value: Model Predictions vs Reality

Comparison of Actual and Predicted Values



The dual bar plot showcases the difference between real and predicted player values for each player, allowing us to quickly assess the model's accuracy.

Conclusion and Recommendation

- **Accurate Player Valuation:**

- Our model predicts player value based primarily on their skills and age, ensuring a more objective and realistic valuation.

- **Mitigating Biases:**

- By excluding factors like leagues, nationality, and ethnicity, the model addresses common biases in current player valuations, such as overvaluation or undervaluation due to these external factors.

- **Objective Basis for Decision-Making:**

- This model provides a more reliable base valuation for players, allowing managers, club owners, and stakeholders to make more informed decisions on player acquisitions, purely based on performance attributes and age.

- **Recommendation for Use:**

- **We recommend using this system** as a core tool for evaluating player market value. While factors like player popularity can influence a club's decision for revenue purposes, our model ensures that the valuation process remains focused on skills and age, providing a fairer and unbiased starting point.

- **Practical Application:**

- Using the model's valuation, clubs can better determine how much to invest in players, aligning financial decisions with a player's inherent abilities and potential, rather than being swayed by external factors.

THANKS FOR LISTENING

- OUR MODEL MIGHT NOT PREDICT THE NEXT MESSI,
BUT IT WILL DEFINITELY PREDICT HIS VALUE!