

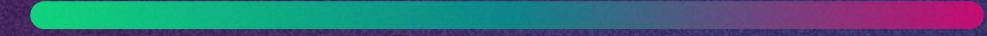


ENGLISH  
PREMIER LEAGUE.

2023-2024

EPL STATISTIC ANALYSIS

# Insights Player Performance



Breaking Down the Stats that Shaped the Season

2023-2024

Presented by

KRITSANAPONG TITAYANUWAT  
TAKDANAI ONWAN

DATA SCINCETIS

# Content

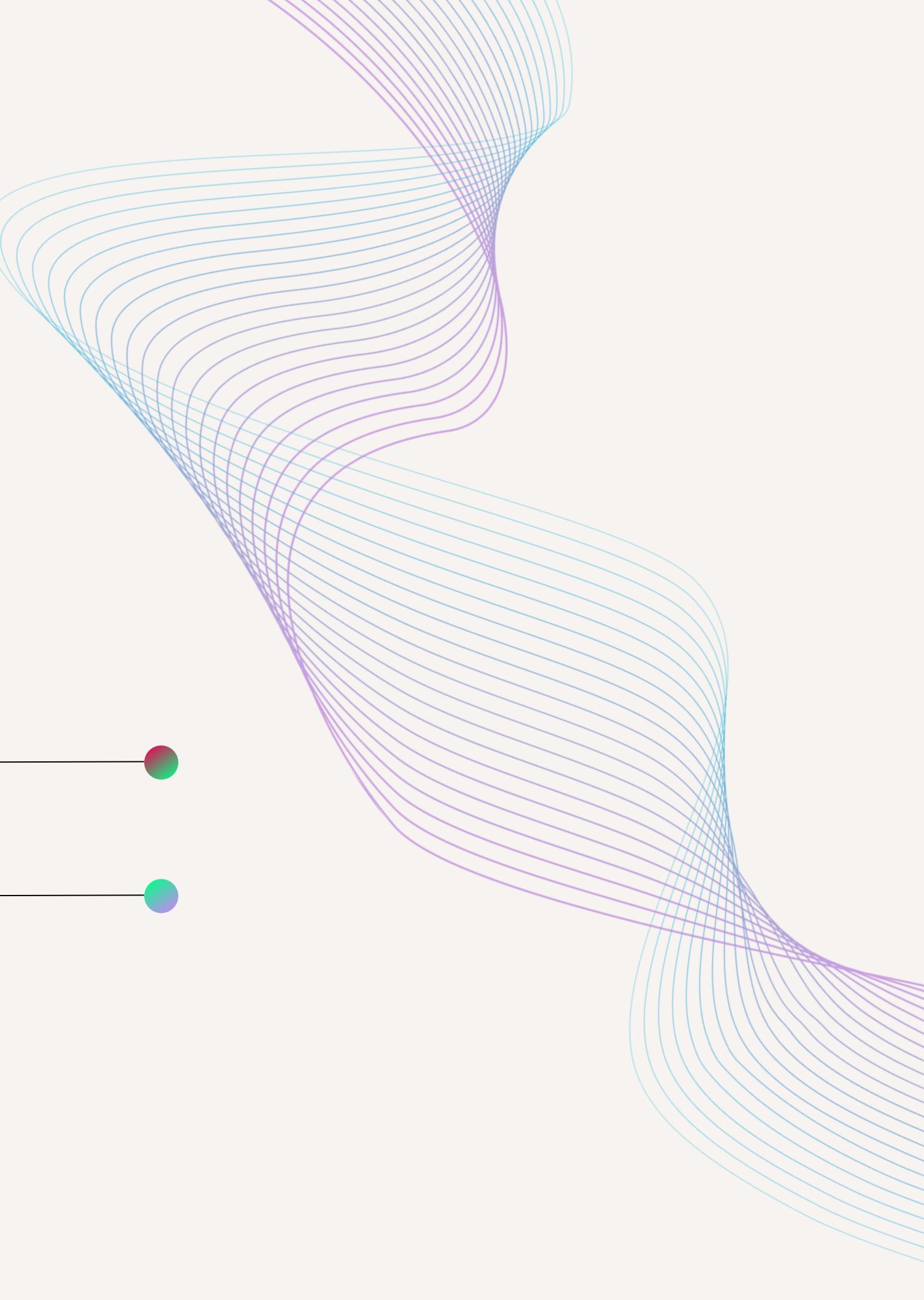
---

ABOUT DATA

DESCRIPTIVE

VISUALIZATION

PREDICTIVE





ENGLISH  
PREMIER LEAGUE.

2023-2024

# Abstract.

This study analyzes player statistics from the 2023-2024 season, providing insights into key performance metrics across various positions. Data collected includes goals scored, assists, passing accuracy, and overall impact on match outcomes.

This study concludes by outlining potential future trends in player performance metrics and their implications for tactical strategies in elite football competitions.

DATA SCINCETIS



# ENGLISH PREMIER LEAGUE.

# A BRIEF OF Data

# player effectiveness and team dynamics in the 2023-2024 EPL season

Player	Nation	Pos	Age	MP	Starts	Min	90s	Gls	Ast	G+A	G-PK	PK	PKatt	CrdY	CrdR	xG
Rodri	es ESP	MF	27.0	34	34	2931.0	32.6	8.0	9.0	17.0	8.0	0.0	0.0	8.0	1.0	4.1
Phil Foden	eng ENG	FW,MF	23.0	35	33	2857.0	31.7	19.0	8.0	27.0	19.0	0.0	0.0	2.0	0.0	10.3
Ederson	br BRA	GK	29.0	33	33	2785.0	30.9	0.0	0.0	0.0	0.0	0.0	0.0	5.0	0.0	0.0

npxG	xAG	npxG+xAG	PrgC	PrgP	PrgR	GIs_90	Ast_90	G+A_90	G-PK_90
4.1	3.9	8.0	76.0	376.0	55.0	0.25	0.28	0.52	0.2
10.3	8.4	18.7	93.0	168.0	269.0	0.6	0.25	0.85	0.
0.0	0.1	0.1	0.0	4.0	0.0	0.0	0.0	0.0	0.

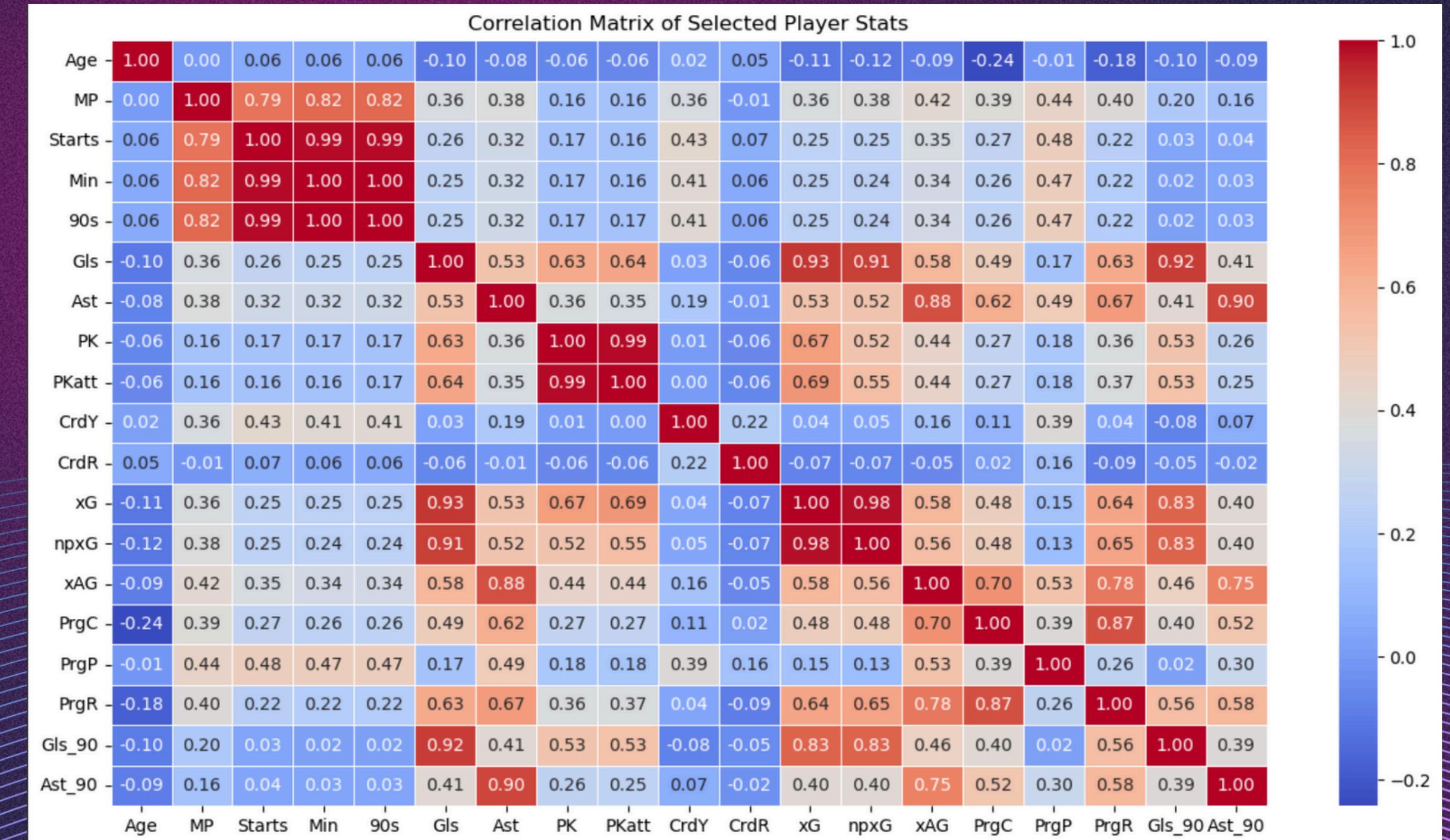


ENGLISH  
PREMIER LEAGUE.

2023-2024

# Visualization

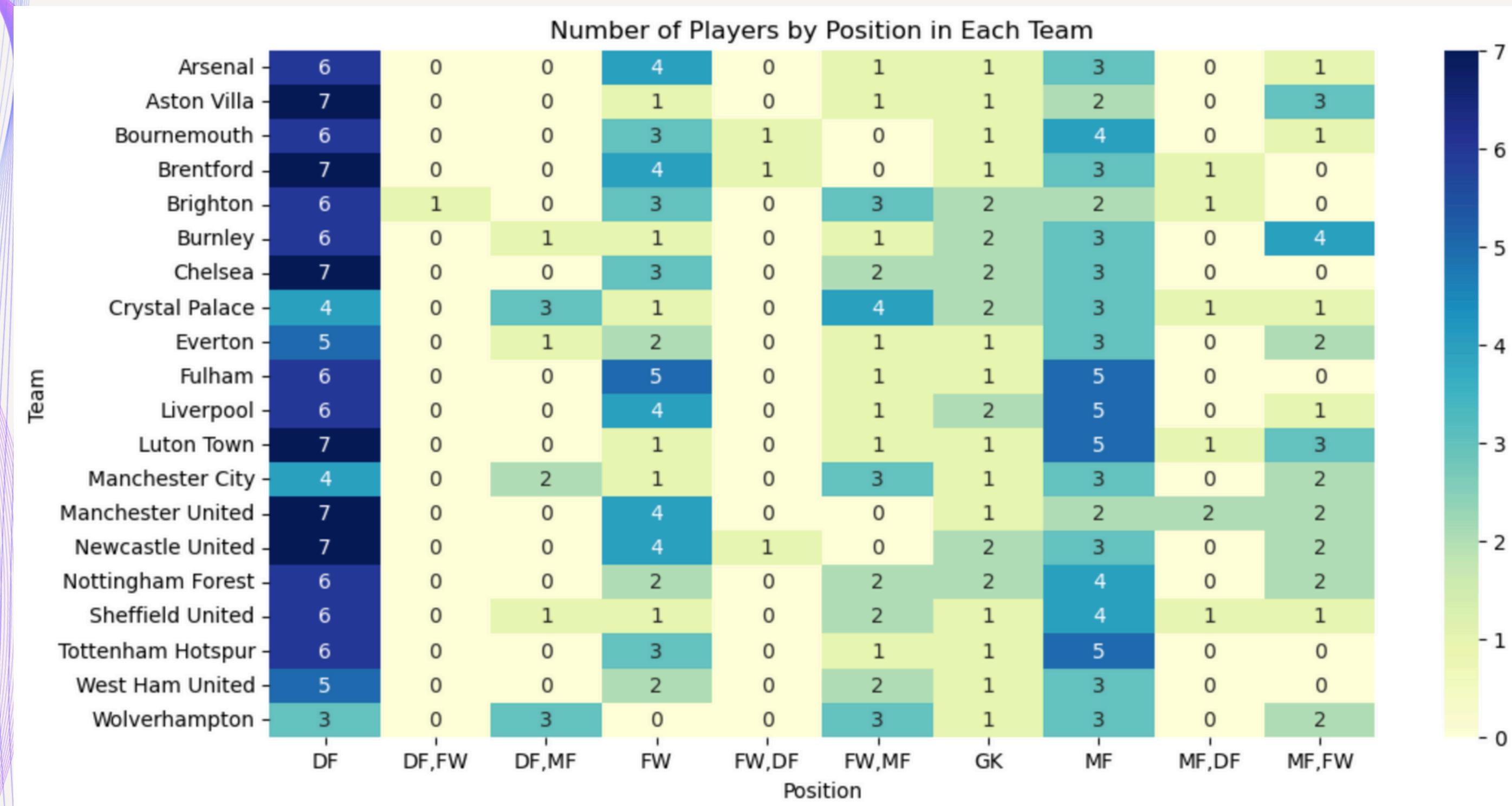
## Heat map Correlation





# ENGLISH PREMIER LEAGUE.

## Number of Player Position in each team

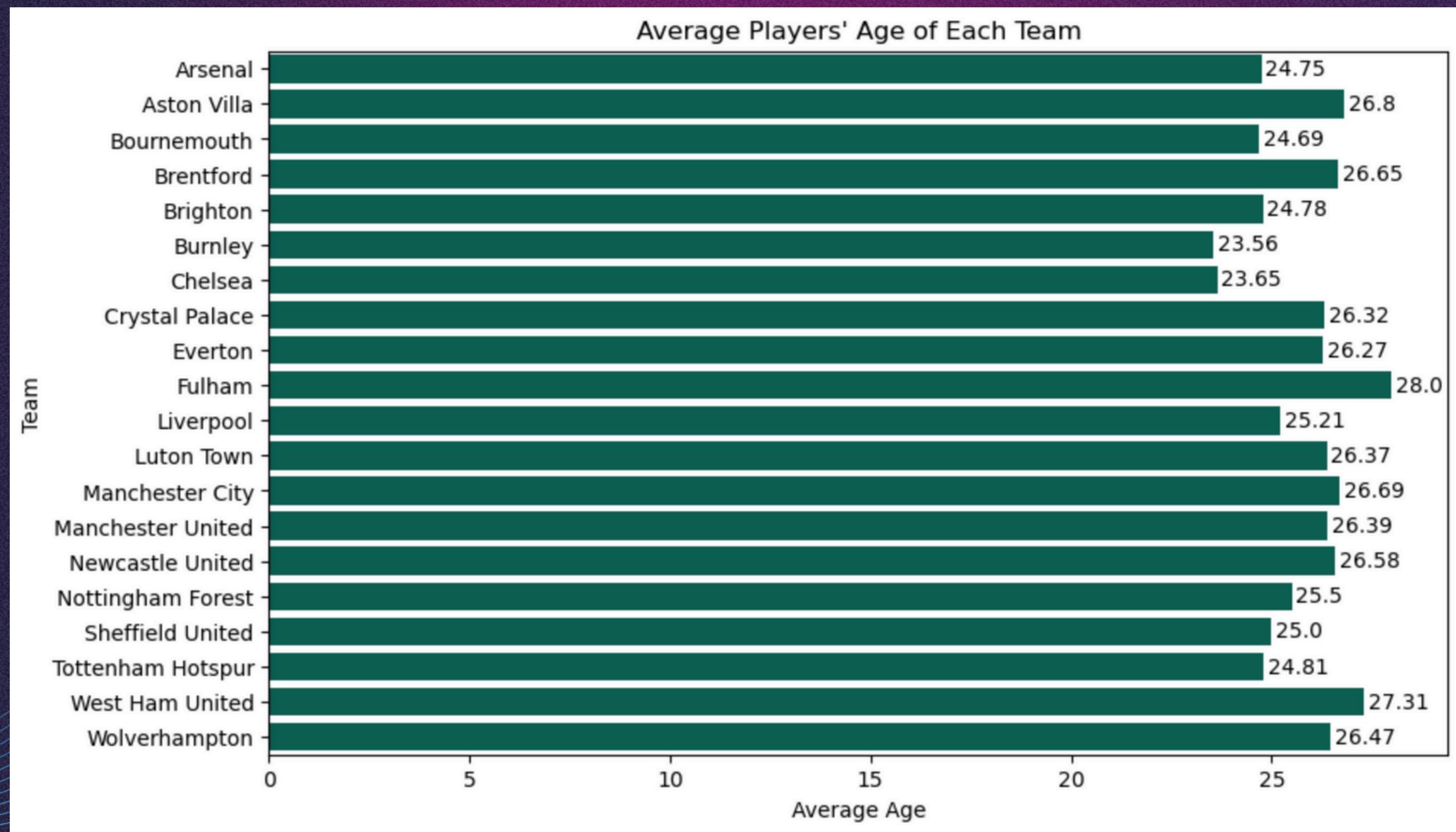




ENGLISH  
PREMIER LEAGUE.

2023-2024

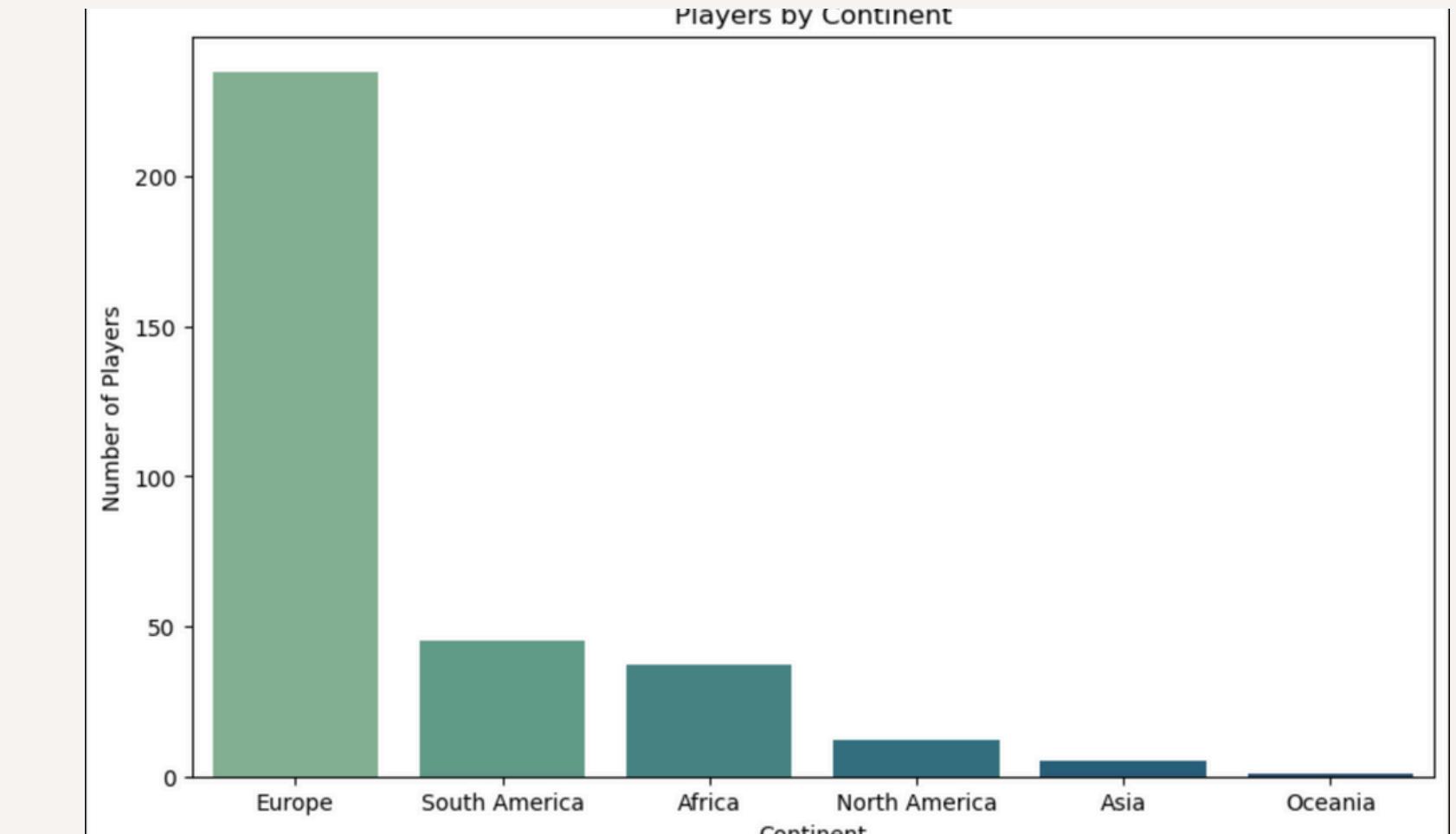
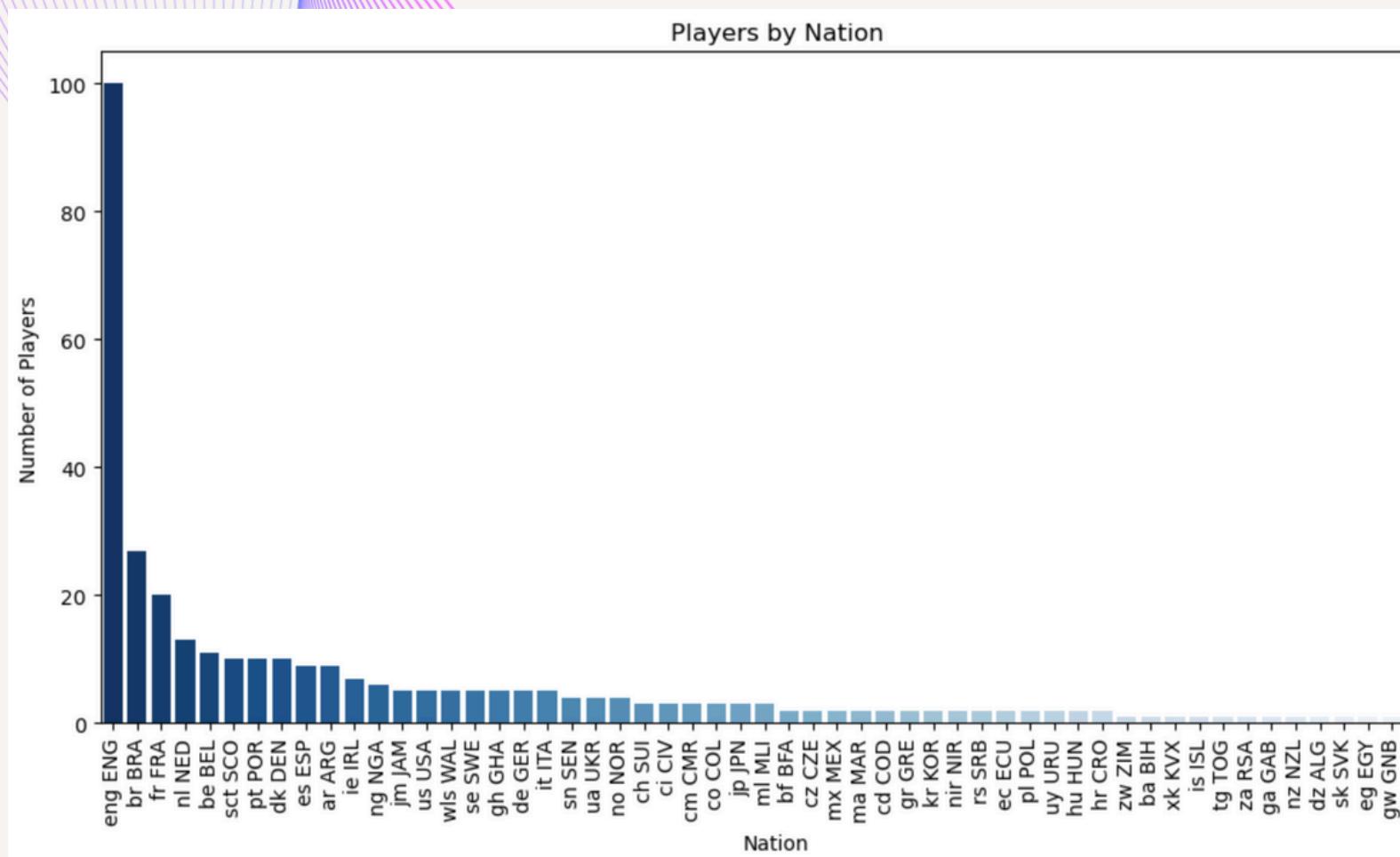
## Average Player Age In Each Team





# ENGLISH PREMIER LEAGUE.

## Player National



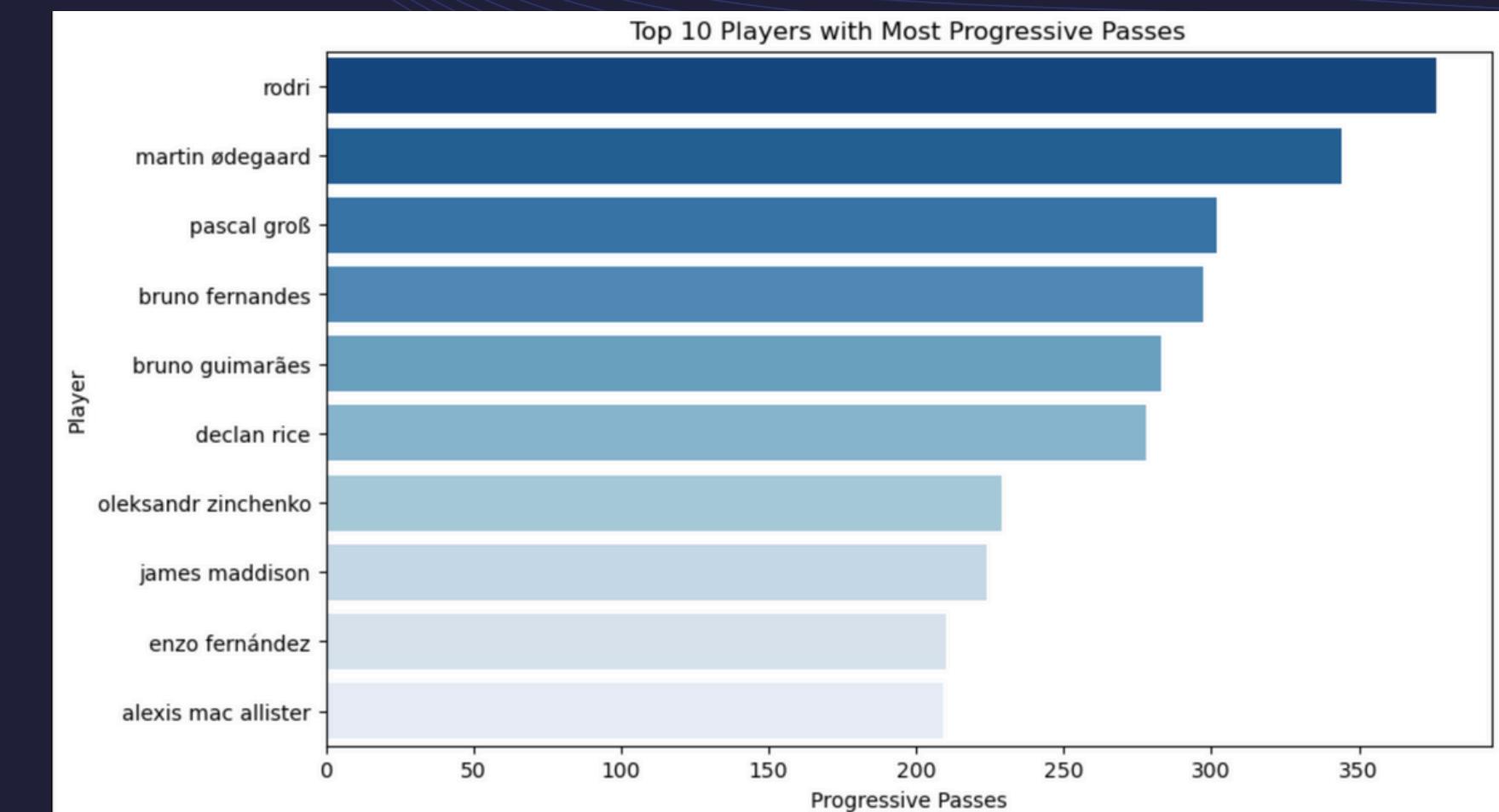
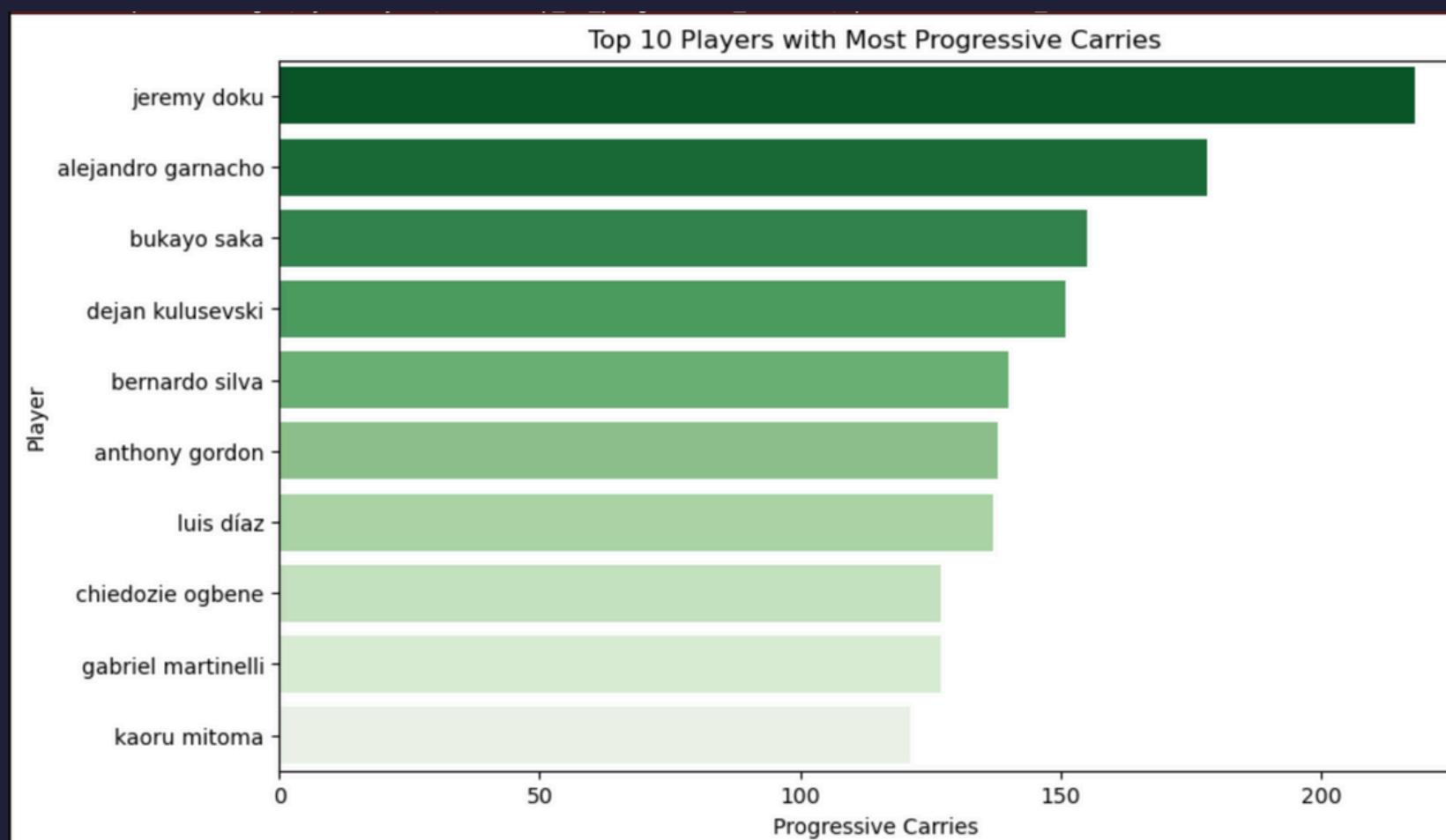
โดยการแบ่งตามประเทศข้อมูลจะมีเยอะเกินไปไม่เหมาะสมแก่  
การทำตาราง จึงปรับข้อมูลเป็นทวบทำให้ Scope ข้อมูล  
กว้างขึ้นเหมาะสมแก่การทำ Visualization



ENGLISH  
PREMIER LEAGUE.

2023-2024

## Progressive



แบ่งเป็น 2 ประเภท ได้แก่  
ผู้เล่นที่เลี้ยงบอลขึ้นข้างหน้า และ ผู้เล่นที่ส่งบอลยาวไปข้างหน้า



## ENGLISH PREMIER LEAGUE.

### Goals Scored by Age

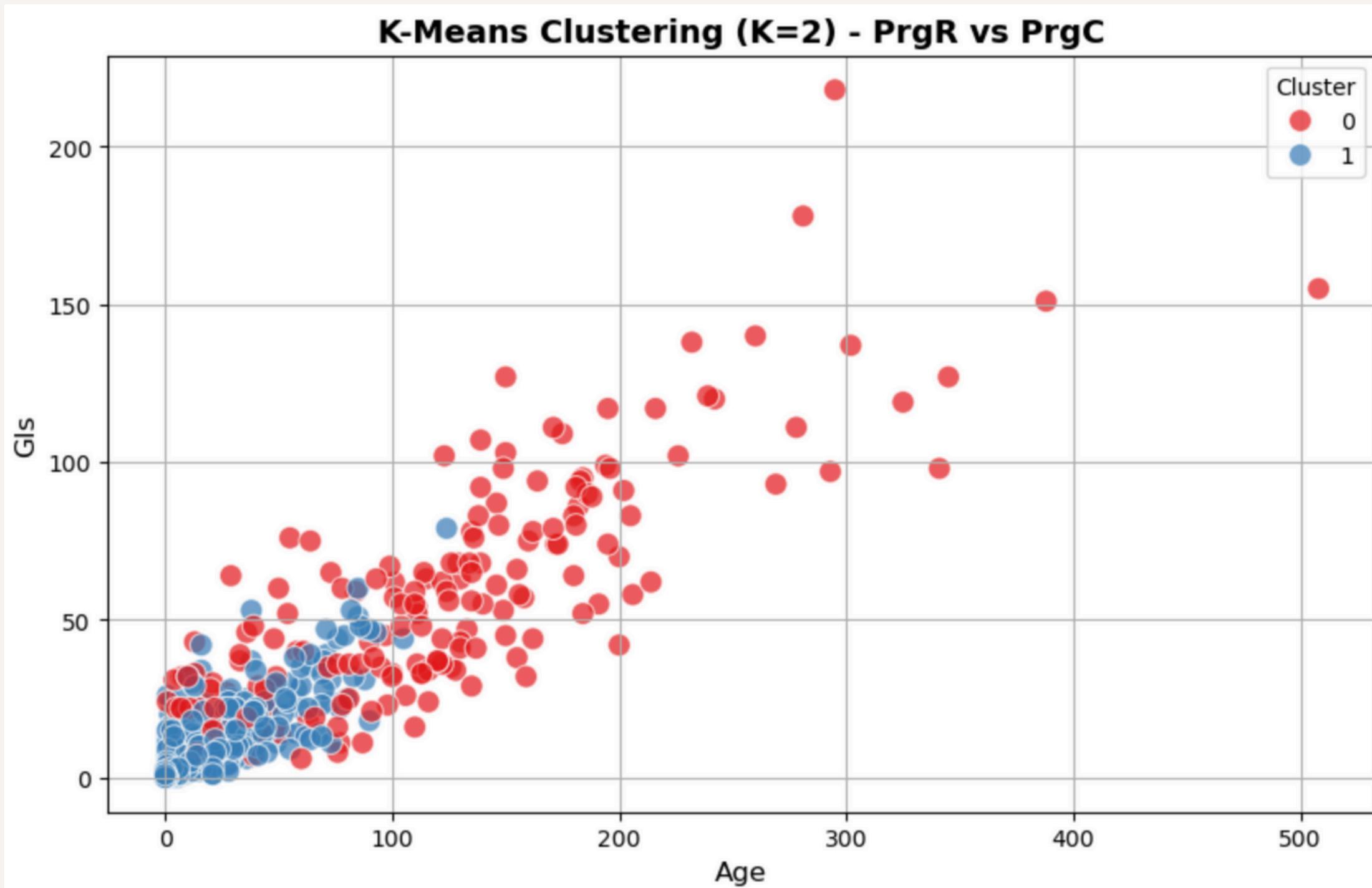
Cluster 0 (สีแดง):

ผู้เล่นในกลุ่มนี้มีจำนวน Progressive Runs (PrgR) สูง  
อาจเป็นผู้เล่นในตำแหน่ง ปีก (Winger) หรือ  
กองกลางตัวรุก (Attacking Midfielder) ที่  
เน้นพาบอลไปข้างหน้า

Cluster 1 (สีฟ้า):

ผู้เล่นในกลุ่มนี้มีจำนวน Progressive Carries (PrgC) สูงแต่ Progressive Runs (PrgR) ต่ำ  
อาจเป็นผู้เล่นในตำแหน่ง กองกลางตัวรับ  
(Defensive Midfielder) หรือ กองหลัง  
(Defender) ที่เน้นการครองบอลและสร้างเกม  
จากแดนหลัง

K-Means Clustering (K=2) - PrgR vs PrgC





ENGLISH  
PREMIER LEAGUE.

## DATA CLUSTERING

```
cluster_summary = df_cluster.groupby('Cluster')[features].mean()  
cluster_summary['count'] = df_cluster['Cluster'].value_counts()  
cluster_summary
```

	90s	Gls	Ast	PK	PKatt	CrdY	CrdR	xG	npxG	xAG	PrgC	PrgP	PrgR	Gls_90	Ast_90
32353	2.262443	1.819005	0.045249	0.054299	4.737557	0.190045	2.284615	2.239819	1.906787	33.959276	80.895928	63.502262	0.114027	0.087014	
96026	0.357616	0.321192	0.006623	0.009934	1.158940	0.046358	0.462583	0.454636	0.372517	7.261589	16.102649	14.334437	0.082086	0.066987	
64912	10.333333	6.315789	1.473684	1.614035	4.491228	0.035088	9.815789	8.542105	5.833333	80.701754	115.754386	187.245614	0.397544	0.240000	



ENGLISH  
PREMIER LEAGUE.

DATA MINING

# Descriptive

TOP 5 PLAYER THAT HAS TRANSFER CHANCE

```
potential_transfers = df_cluster[(df_cluster['G+A_90'] > 0.5) & (df_cluster['Min'] < 1100)]
potential_transfers = potential_transfers[['Player', 'Age', 'Pos', 'Min', 'G+A_90', 'xG_90', 'xAG_90']]
top_5_transfers = potential_transfers.sort_values(by='G+A_90', ascending=False).head(5)
print(top_5_transfers)
```

	Player	Age	Pos	Min	G+A_90	xG_90	xAG_90
129	callum wilson	31.0	FW	991.0	0.91	0.80	0.05
470	taiwo awoniyi	25.0	FW	1042.0	0.78	0.41	0.15
95	noni madueke	21.0	FW	1053.0	0.60	0.21	0.23
69	eddie nketiah	24.0	FW	1082.0	0.58	0.38	0.07



ENGLISH  
PREMIER LEAGUE.

## PROGRESSIVE ANALYSIS

เพื่อหาผู้เล่นที่มีอัตราการเลี้ยงบอลไปแทนตรงข้าสูงในอายุไม่เกิน 23 ปี เพื่อวิเคราะห์การซื้อผู้เล่น

```
df_cluster[(df_cluster['PrgC']>=200)&(df_cluster['Age']<=23)]
```

Player	Nation	Pos	Age	MP	Starts	Min	90s	Gls	Ast	...	G+A-PK_90	xG_90	xAG_90	xG+xAG_90	npxG_90	npxG+xAG_90	Team
11 jeremy doku	be BEL	FW,MF	21.0	29	18	1595.0	17.7	3.0	8.0	...	0.62	0.14	0.26	0.4	0.14	0.4	Manchester City

เพื่อหาผู้เล่นที่มีอัตราการเปิดบอลไปแทนตรงข้าสูงในอายุไม่เกิน 23 ปี เพื่อวิเคราะห์การซื้อผู้เล่น

```
df_cluster[(df_cluster['PrgR']>=400)&(df_cluster['Age']<=23)]
```

Player	Nation	Pos	Age	MP	Starts	Min	90s	Gls	Ast	...	G+A-PK_90	xG_90	xAG_90	xG+xAG_90	npxG_90	npxG+xAG_90	Team
59 bukayo saka	eng ENG	FW	21.0	35	35	2919.0	32.4	16.0	9.0	...	0.59	0.48	0.32	0.8	0.33	0.65	Arsenal



ENGLISH  
PREMIER LEAGUE.

# Predictive

TOP 5 PLAYER THAT HAS TRANSFER CHANCE

Accuracy: 1.0						
	precision	recall	f1-score	support		
0	1.00	1.00	1.00	168		
1	1.00	1.00	1.00	6		
accuracy			1.00		174	
macro avg		1.00	1.00	1.00	174	
weighted avg		1.00	1.00	1.00	174	
Mean Absolute Error (MAE): 0.0000						
	Player	Age	Min	G+A_90	xG_90	xAG_90 \
307	adama traoré	27.0	377.0	1.19	0.36	0.17
447	kevin schade	21.0	333.0	0.81	0.32	0.19
223	jhon durán	19.0	475.0	0.95	0.38	0.06
388	enes ünal	26.0	328.0	1.10	0.80	0.28
448	shandon baptiste	25.0	229.0	0.79	0.35	0.08
Transfer_Probability						
307		0.99				
447		0.98				
223		0.98				
388		0.97				
448		0.96				



ENGLISH  
PREMIER LEAGUE.

## PREDICT PLAYER GOAL PER MATCH

ทำนายการยิงประตูของผู้เล่น จากสถิติการเล่น โดยใช้

CROSS VALIDATION : TRAIN 80 | TEST 20

Cross Validation Scores (MSE): [1.35445256 1.57103467 2.31105668 1.33356356 1.78656411]

Average MSE: 1.67

Root Mean Squared Error (RMSE): 1.29

Mean Squared Error (MSE): 1.35

Root Mean Squared Error (RMSE): 1.16

R-squared (R<sup>2</sup>): 0.93

	Player	Actual	Predicted	Error
0	sam curtis	0.0	-0.072453	0.072453
1	nathaniel clyne	0.0	0.106243	-0.106243
2	ian maatsen	0.0	0.694573	-0.694573
3	marcus rashford	7.0	7.287379	-0.287379
4	ryan yates	1.0	2.560460	-1.560460
5	bernd leno	0.0	-0.374227	0.374227
6	luca koleosho	1.0	2.094824	-1.094824
7	mikkel damsgaard	0.0	0.914575	-0.914575
8	ivan perišić	0.0	0.042549	-0.042549
9	Łukasz fabiański	0.0	-0.192322	0.192322



ENGLISH  
PREMIER LEAGUE.

## PREDICT OVER PERFORMANCE PLAYER

กำหนดว่าผู้เล่นจะทำประตูได้เกินค่าที่คาดการว่าจะทำได้  
หรือไม่(XG)

CROSS VALIDATION : TRAIN 80 | TEST 20

Accuracy: 87.07%					
	precision	recall	f1-score	support	
0	0.87	0.97	0.92	86	
1	0.86	0.60	0.71	30	
accuracy			0.87	116	
macro avg	0.87	0.78	0.81	116	
weighted avg	0.87	0.87	0.86	116	
Top 10 Overperformers:					
	Player	Gls	xG		
83	cole palmer	22.0	18.2		
117	alexander isak	21.0	20.3		
1	phil foden	19.0	10.3		
28	mohamed salah	18.0	21.2		
146	son heung-min	17.0	12.0		
265	jean-philippe mateta	16.0	10.9		
59	bukayo saka	16.0	15.5		
458	chris wood	14.0	11.9		
62	kai havertz	13.0	12.3		
404	matheus cunha	12.0	9.5		

# Thanks You

IF YOU HAVE ANY QUESTION  
PLEASE RAISE YOUT HAND

DATA SOURCE

DATA SOURCE<https://www.kaggle.com/datasets/orkunaktas/premier-league-all-players-stats-2324>