Data-Driven Goals: A Comprehensive EDA of Football

Soumya Chauhan
Naksh Jain Sonigara

Introduction

About the Project

This EDA offers a comprehensive exploration of football analytics, leveraging Python libraries to analyze players, club games, and various other aspects of the sport. Through statistical analysis and visualization, we uncover insights into player performance, club strategies, and league dynamics providing valuable insights for football enthusiasts

About Football

Football, also called soccer, is the most popular sport in the world. Known for its excitement and global reach, it captivates millions of fans. Its widespread popularity and dynamic nature make it a great choice for data analysis, offering many opportunities to discover useful insights.

Objective

Player Analysis

- Analyze player statistics and performance metrics to identify top performers across various positions and roles.
- Understand the factors influencing player market valuation and transfer fees.
- Discover trends and patterns in player market valuation and moves.

Club Analysis

- Evaluate team performance and tactics based on aggregated player data.
- Evaluate the effectiveness of player recruitment strategies and their impact on team performance.
- Analyze the correlation between squad composition (including player nationalities, squad size, and squad age), and team success.

Competition Analysis

- Explore the competitive landscape and dynamics across various leagues and competitions.
- Identify factors contributing to the success or failure of teams in different competitions.
- Analyze historical trends and patterns in league standings, championship wins, and tournament performances.

Information for Coaches and Fans

- Provide data-driven insights and visualizations that enhance the understanding of the game for coaches and fans.
- Offer engaging and interactive content that allows coaches to explore and analyze player and team statistics for tactical purposes.
- Create compelling narratives and stories around player performances, team tactics, and league dynamics using statistical evidence to engage fans.

Data

Source - [Football Data: Competitions, Clubs, Players](#)

Data Summary

- The dataset compiles comprehensive football data from Transfermarkt, offering precise and regularly updated information sourced from over 60,000 games across major global competitions.

- It includes data from more than 400 clubs, covering performance metrics and benchmarks to provide in-depth insights into various aspects of the game.

- Detailed player statistics for over 30,000 players are included, featuring attributes like physical characteristics, team affiliations, contract statuses, and individual performances such as goals scored and assists provided.

- Historical valuation records for players are maintained, spanning several years to showcase how performance impacts value over time and across different instances like transfers.

- The dataset also highlights specific player appearances, with around 1.2 million records detailing competition-level performance patterns, including games played and cards earned during play.

- Each CSV file is structured neatly, containing entity-specific information and unique IDs that facilitate comprehensive analysis across different datasets.

Data analysis approach and tools

- **Data Manipulation**: Pandas and NumPy libraries were utilized for efficient data wrangling and preprocessing tasks.

- **Interactive Visualization**: Matplotlib, Seaborn, Plotly, and Dash frameworks were employed to create engaging and interactive visualizations, enabling effective communication of insights.

- **Statistical Analysis**: SciPy and Scikit-learn libraries facilitated robust statistical analyses, including hypothesis testing through techniques like the t-test.

- **Complex Data Merging**: Extensive data merging operations were performed to integrate and consolidate various data sources, enabling comprehensive analysis.

- **Advanced Grouping and Aggregation**: Various grouping functions, such as pivot tables, groupby, and crosstab, were utilized to uncover patterns and trends within the dataset.

Appearances

Dataset name: appearances

	appearance_id	game_id	player_id	player_club_id	player_current_club_id	date	player_name	competition_id	yellow_cards	red_cards	goals	assists	minutes_played
0	-1_172127	-1	172127	3999	3999	NaN	Tyronne	GRP	0	0	0	0	90
1	-1_210988	-1	210988	3999	553	NaN	Christos Eleftheriadis	GRP	0	0	0	0	90
2	-1_231471	-1	231471	3999	3999	NaN	Theofanis Tzandaris	GRP	0	0	0	0	90
3	-1_256819	-1	256819	3999	3999	NaN	Nikolaos Tsoukalos	GRP	0	0	0	0	90
4	-1_281919	-1	281919	3999	3999	NaN	Adam Tzanetopoulos	GRP	0	0	0	0	90
...	...	...	...	...	...	...	...	...	...	...	...	...	...
1166210	3869240_701325	3869240	701325	2503	2503	2023-04-17	Bruno Onyemaechi	PO1	0	0	0	0	90
1166211	3869240_71830	3869240	71830	6912	6912	2023-04-17	Bruno Wilson	PO1	0	0	0	0	90
1166212	3869240_739438	3869240	739438	6912	6912	2023-04-17	Matheus Pereira	PO1	0	0	0	0	75
1166213	3869240_818252	3869240	818252	2503	2503	2023-04-17	Pedro Malheiro	PO1	0	0	0	0	90
1166214	3869240_95429	3869240	95429	6912	6912	2023-04-17	Osama Rashid	PO1	1	0	1	0	15

1166215 rows × 13 columns

Clubs

```
Dataset name: clubs
```

	club_id	club_code	name	domestic_competition_id	total_market_value	squad_size	average_age	foreigners_number	foreigners_percentage	national_team_players	stadium_name	stadium_seats	net_transfer_record	coach_name
0	127	sc-paderborn-07	NaN	L1	NaN	26	25.7	5	19.2	1	Home Deluxe Arena	15000	€-450k	NaN
1	192	roda-jc-kerkrade	NaN	NL1	NaN	25	23.9	9	36.0	0	Parkstad Limburg Stadion	19979	+€1.30m	NaN
2	2239	us-cremonese	US Cremonese	IT1	NaN	26	26.0	12	46.2	7	Giovanni Zini	16003	€-33.80m	NaN
3	2423	cd-santa-clara	CD Santa Clara	PO1	NaN	31	25.8	20	64.5	2	Estádio de São Miguel	15000	+€5.85m	NaN
4	24245	umraniyespor	Ümraniyespor	TR1	NaN	24	28.4	10	41.7	3	Ümraniye Belediyesi Şehir Stadı	3513	+-0	NaN
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
406	750	fc-sochaux-montbeliard	NaN	FR1	NaN	25	26.2	11	44.0	4	Stade Auguste-Bonal	20005	+€4.00m	NaN
407	82	vfl-wolfsburg	VfL Wolfsburg	L1	NaN	25	24.6	15	60.0	10	Volkswagen Arena	30000	€-12.05m	NaN
408	940	celta-vigo	Celta de Vigo	ES1	NaN	23	26.3	10	43.5	6	Abanca Balaídos	29000	€-5.15m	NaN
409	976	cf-uniao-madeira	NaN	PO1	NaN	0	NaN	0	NaN	0	Centro Desportivo da Madeira	2500	+-0	NaN
410	987	motherwell-fc	Motherwell FC	SC1	NaN	29	25.9	19	65.5	4	Fir Park	13750	+€225k	NaN

411 rows × 16 columns

Club games

```
Dataset name: club_games
```

	game_id	club_id	own_goals	own_position	own_manager_name	opponent_id	opponent_goals	opponent_position	opponent_manager_name	hosting	is_win
0	2221747	440	0	-1	Sven Köhler	52	1	-1	Oliver Reck	Home	0
1	2221760	113	0	-1	Petrik Sander	15	4	-1	Sascha Lewandowski	Home	0
2	2222539	60	1	11	Christian Streich	39	1	10	Thomas Tuchel	Home	0
3	2222541	41	0	15	Thorsten Fink	4	1	6	Dieter Hecking	Home	0
4	2222542	24	2	5	Armin Veh	15	1	12	Sascha Lewandowski	Home	1
...	...	...	...	...	...	...	...	...	...	...	...
123071	4031397	12	0	-1	José Mourinho	681	0	-1	Imanol Alguacil	Away	0
123072	4034848	11127	1	-1	Viktor Goncharenko	232	1	-1	Guillermo Abascal	Away	0
123073	4036081	338	2	4	Mircea Lucescu	18105	0	16	Anatoliy Bezsmertnyi	Away	1
123074	4036087	3592	2	6	Yuriy Vernydub	26459	1	12	Yuriy Virt	Away	1
123075	4042704	11127	0	11	Viktor Goncharenko	964	2	1	Sergey Semak	Away	0

123076 rows × 11 columns

Competitions

Dataset name: competitions

	competition_id	competition_code	name	sub_type	type	country_id	country_name	domestic_league_code	confederation	url
0	BE1	jupiler-pro-league	jupiler-pro-league	first_tier	domestic_league	19	Belgium	BE1	europa	https://www.transfermarkt.co.uk/jupiler-pro-le...
1	SCI	supercoppa-italiana	supercoppa-italiana	domestic_super_cup	other	75	Italy	IT1	europa	https://www.transfermarkt.co.uk/supercoppa-ita...
2	SFA	sfa-cup	sfa-cup	domestic_cup	domestic_cup	190	Scotland	SC1	europa	https://www.transfermarkt.co.uk/sfa-cup/starts...
3	UKRS	ukrainian-super-cup	ukrainian-super-cup	domestic_super_cup	other	177	Ukraine	UKR1	europa	https://www.transfermarkt.co.uk/ukrainian-supe...
4	FAC	fa-cup	fa-cup	domestic_cup	domestic_cup	189	England	GB1	europa	https://www.transfermarkt.co.uk/fa-cup/startse...
5	KLUB	fifa-klub-wm	fifa-klub-wm	fifa_club_world_cup	other	-1	NaN	NaN	europa	https://www.transfermarkt.co.uk/fifa-klub-wm/s...
6	BESC	belgian-supercup	belgian-supercup	domestic_super_cup	other	19	Belgium	BE1	europa	https://www.transfermarkt.co.uk/belgian-superc...
7	CGB	efl-cup	efl-cup	league_cup	other	189	England	GB1	europa	https://www.transfermarkt.co.uk/efl-cup/starts...
8	DK1	superligaen	superligaen	first_tier	domestic_league	39	Denmark	DK1	europa	https://www.transfermarkt.co.uk/superligaen/st...
9	ECLQ	uefa-europa-conference-league-qualifikation	uefa-europa-conference-league-qualifikation	uefa_europa_conference_league_qualifiers	international_cup	-1	NaN	NaN	europa	https://www.transfermarkt.co.uk/uefa-europa-co...
10	ELQ	europa-league-qualifikation	europa-league-qualifikation	europa_league_qualifying	international_cup	-1	NaN	NaN	europa	https://www.transfermarkt.co.uk/europa-league-...
11	NLSC	johan-cruijff-schaal	johan-cruijff-schaal	domestic_super_cup	other	122	Netherlands	NL1	europa	https://www.transfermarkt.co.uk/johan-cruijff-...
12	CDR	copa-del-rey	copa-del-rey	domestic_cup	domestic_cup	157	Spain	ES1	europa	https://www.transfermarkt.co.uk/copa-del-rey/s...
13	USC	uefa-super-cup	uefa-super-cup	uefa_super_cup	other	-1	NaN	NaN	europa	https://www.transfermarkt.co.uk/uefa-super-cup...
14	CL	uefa-champions-league	uefa-champions-league	uefa_champions_league	international_cup	-1	NaN	NaN	europa	https://www.transfermarkt.co.uk/uefa-champions...
15	DFB	dfb-pokal	dfb-pokal	domestic_cup	domestic_cup	40	Germany	L1	europa	https://www.transfermarkt.co.uk/dfb-pokal/star...
16	FR1	ligue-1	ligue-1	first_tier	domestic_league	50	France	FR1	europa	https://www.transfermarkt.co.uk/ligue-1/starts...

Games

Dataset name: games

	game_id	competition_id	season	round	date	home_club_id	away_club_id	home_club_goals	away_club_goals	home_club_position	away_club_position	home_club_manager_name	away_club_manager_name	stadium	atten
0	2222539	L1	2012	1. Matchday	2012-08-25	60	39	1	1	11	10	Christian Streich	Thomas Tuchel	Dreisamstadion	2
1	2222541	L1	2012	1. Matchday	2012-08-25	41	4	0	1	15	6	Thorsten Fink	Dieter Hecking	Volksparkstadion	5
2	2222542	L1	2012	1. Matchday	2012-08-25	24	15	2	1	5	12	Armin Veh	Sascha Lewandowski	Commerzbank Arena	2
3	2222595	RU1	2012	3. Matchday	2012-08-05	2700	4128	1	0	3	12	Guus Hiddink	Rustem Khuzin	Dinamo	1
4	2222636	RU1	2012	4. Matchday	2012-08-12	1083	2696	1	2	15	6	Miodrag Bozovic	Andrey Kobelev	Olimp-2	
...	...	...	...	...	...	...	...	...	...	...	...		...	...	
61533	2888300	NL1	2017	26. Matchday	2018-03-03	383	200	3	0	1	4	Phillip Cocu	Jean-Paul de Jong	Philips Stadion	3
61534	3450596	NL1	2020	12. Matchday	2020-12-13	383	200	2	1	2	10	Roger Schmidt	René Hake	Philips Stadion	
61535	3597835	NL1	2021	6. Matchday	2021-09-22	467	200	0	3	10	5	Rogier Meijer	René Hake	Goffertstadion	
61536	2250211	NL1	2012	5. Matchday	2012-09-16	200	383	1	0	8	5	Jan Wouters	Dick Advocaat	Stadion "Galgenwaard"	1
61537	4031353	UKR1	2022	17. Matchday	2023-03-12	48726	48332	0	1	12	6	Leonid Kuchuk	Yaroslav Vyshnyak	Arena Lviv	

61538 rows × 21 columns

Game events

Dataset name: game_events

	game_id	minute	type	club_id	player_id	description	player_in_id
0	2221747	70	Substitutions	52	2125	NaN	43984.0
1	2221747	80	Substitutions	440	23566	NaN	40679.0
2	2221747	86	Substitutions	52	8639	NaN	42749.0
3	2221747	88	Substitutions	440	43822	NaN	83191.0
4	2221760	68	Substitutions	15	61087	NaN	39102.0
...	...	...	...	...	...	...	...
559957	4000515	86	Goals	37125	707891	, 3. Tournament Goal	NaN
559958	3921099	80	Substitutions	50728	1055537	NaN	324360.0
559959	3986024	80	Substitutions	2503	314942	NaN	685807.0
559960	4035007	51	Goals	1237	648046	, Right-footed shot, 2. Tournament Goal	NaN
559961	3971175	69	Substitutions	3725	623553	NaN	220650.0

559962 rows × 7 columns

Players

```
Dataset name: players
```

	player_id	first_name	last_name	name	last_season	current_club_id	player_code	country_of_birth	city_of_birth	country_of_citizenship	date_of_birth	sub_position	position	foot	height_in_cm	market_value_in_eur	highest
0	597	Aleksandr	Hleb	Aleksandr Hleb	2016	2696	aleksandr-hleb	UdSSR	Minsk	Belarus	1981-05-01	Attacking Midfield	Midfield	NaN	NaN	NaN	NaN
1	1428	Mike	Hanke	Mike Hanke	2013	60	mike-hanke	Germany	Hamm	Germany	1983-11-05	Centre-Forward	Attack	right	185.0	NaN	NaN
2	1560	Youssef	El Akchaoui	Youssef El Akchaoui	2012	306	youssef-el-akchaoui	Netherlands	Dordrecht	Morocco	1981-02-18	Left-Back	Defender	NaN	NaN	NaN	NaN
3	1587	Mario	Eggimann	Mario Eggimann	2012	42	mario-eggimann	Switzerland	Brugg	Switzerland	NaN	Centre-Back	Defender	right	190.0	NaN	NaN
4	1726	Heiko	Westermann	Heiko Westermann	2016	610	heiko-westermann	Germany	Alzenau	Germany	1983-08-14	Centre-Back	Defender	right	190.0	NaN	NaN
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
28454	199337	Rifat	Erdem	Rifat Erdem	2013	2293	rifat-erdem	Turkey	Konya	Turkey	1990-02-28	Right-Back	Defender	right	NaN	NaN	NaN
28455	209569	Mücahit	Atalay	Mücahit Atalay	2020	2293	mucahit-atalay	Turkey	Trabzon	Turkey	1991-10-26	Goalkeeper	Goalkeeper	right	186.0	NaN	NaN
28456	221118	Patrick Friday	Eze	Patrick Friday Eze	2017	2293	patrick-friday-eze	Nigeria	Kaduna	Nigeria	1992-12-22	Centre-Forward	Attack	right	187.0	200000.0	NaN
28457	273053	Burak	Saban	Burak Saban	2018	2293	burak-saban	Netherlands	Hoorn	Netherlands	1994-04-20	Right Winger	Attack	both	177.0	NaN	NaN
28458	621911	Ekrem	Kayilibal	Ekrem Kayilibal	2021	2293	ekrem-kayilibal	Turkey	Konya	Turkey	1999-09-22	Right Winger	Attack	both	176.0	50000.0	NaN

```
28459 rows × 23 columns
```

Player valuations

```
Dataset name: player_valuations
```

	player_id	last_season	datetime	date	dateweek	market_value_in_eur	n	current_club_id	player_club_domestic_competition_id
0	3132	2013	2003-12-09 00:00:00	2003-12-09	2003-12-08	400000	1	126	TR1
1	6893	2012	2003-12-15 00:00:00	2003-12-15	2003-12-15	900000	1	984	GB1
2	10	2015	2004-10-04 00:00:00	2004-10-04	2004-10-04	7000000	1	398	IT1
3	26	2017	2004-10-04 00:00:00	2004-10-04	2004-10-04	1500000	1	16	L1
4	65	2015	2004-10-04 00:00:00	2004-10-04	2004-10-04	8000000	1	1091	GR1
...	...	...	...	...	...	...	...	...	...
424318	158071	2015	2023-05-17 00:00:00	2023-05-17	2023-05-15	250000	1	2448	ES1
424319	187191	2016	2023-05-17 00:00:00	2023-05-17	2023-05-15	400000	1	192	NL1
424320	279110	2015	2023-05-17 00:00:00	2023-05-17	2023-05-15	250000	1	1533	ES1
424321	863829	2022	2023-05-17 00:00:00	2023-05-17	2023-05-15	100000	1	2414	DK1
424322	405676	2018	2023-12-19 00:00:00	2023-12-19	2023-12-18	50000	1	23826	L1

424323 rows × 9 columns

Player Analysis

Market valuation trend

Mohamed Salah's Valuation Over Years
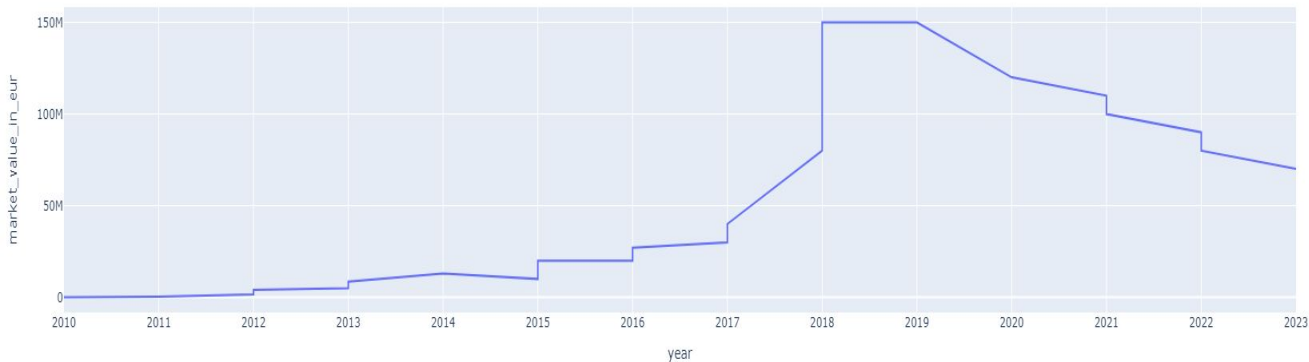
- Players' trends can be analyzed through searches.
- Mohamed Salah's trajectory over the years reveals fluctuations.
- His market value reached its highest point in 2018 coinciding with Liverpool FC's Champions League victory.
- This pattern is observed in the majority of players.

Players with most matches played for a single club

player_name	name	No. of Matches
Sergio Busquets	FC Barcelona	421
Thomas Müller	Bayern Munich	403
Koke	Atlético de Madrid	403
Toni Kroos	Real Madrid	397
David de Gea	Manchester United	393
...	...	...
Rafael Forster	Fenerbahce	2
Orest Lebedenko	Shakhtar Donetsk	2
Tarik Elyounoussi	SC Freiburg	2
Giovani Lo Celso	Crystal Palace	2
Rafael Forster	Manchester United	2

193637 rows × 1 columns

- These players have the longest tenures with a single club.

- The amount of loyalty which these players have shown has made them Iconic Figures for their respective clubs and fan- favourites.

Most transfers

Highest number of transfers
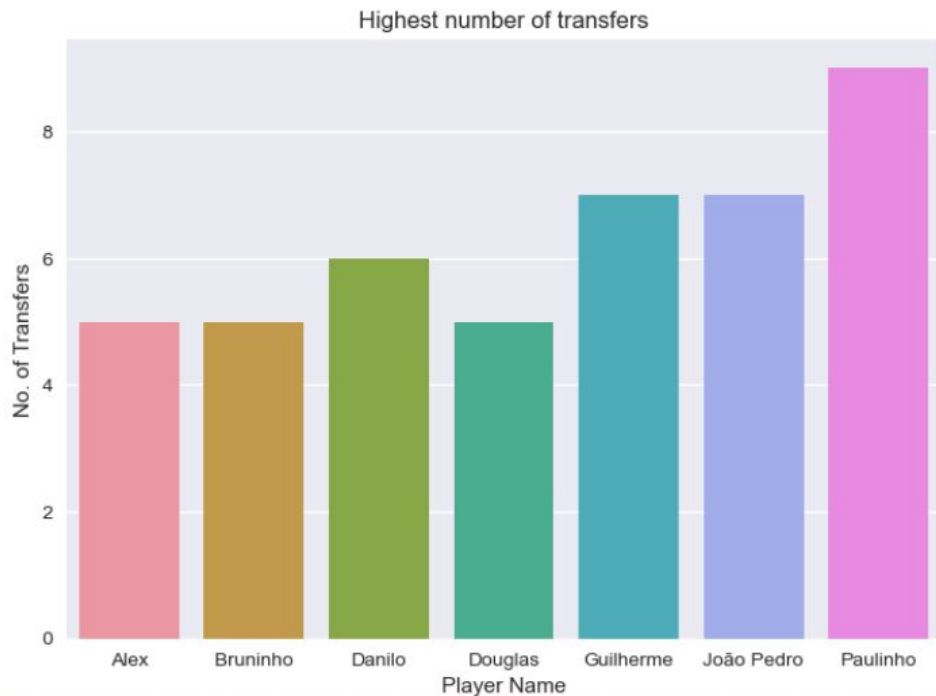
Paulinho's Transfer History:

- Paulinho has undergone 9 transfers throughout his entire career, indicating a level of mobility and adaptability within the footballing landscape.

Consistent Transfers Every 2 Seasons:

- Interestingly, Paulinho's career trajectory reveals a pattern of transferring to a different club approximately every 2 seasons, suggesting a deliberate strategy or response to changing circumstances.

Dominance at Corinthians Football Club:

- Paulinho's allegiance lies prominently with Corinthians Football Club, where he has spent the majority of his career and currently continues to play. This enduring connection signifies a deep-rooted association with the club and its values.

Commonality of Multiple Transfer:

- It is observed that over 98% of players undergo at least 1 transfers during their career.

Height-Distributions of Center-Forwards

col_0	Percentage of players	Count of players
height_bin		
(162.959, 171.2]	2.433281	93
(171.2, 179.4]	17.425432	666
(179.4, 187.6]	52.982732	2025
(187.6, 195.8]	25.327054	968
(195.8, 204.0]	1.831502	70

- The data suggests that the height range of 179 - 184 cm is the most common among Centre-Forwards, with over half (52%) of players falling within this bin. This could indicate that this height range is considered optimal for the physical and tactical demands of the Centre-Forward position.

- When scouting and recruiting Centre-Forwards, teams may prioritize players within the 179 - 184 cm height range, as this seems to be the most prevalent height among successful players in this position. However, it's important to consider other factors beyond just height.

Club Analysis

Average Age Distribution across Clubs

Average Age Analysis
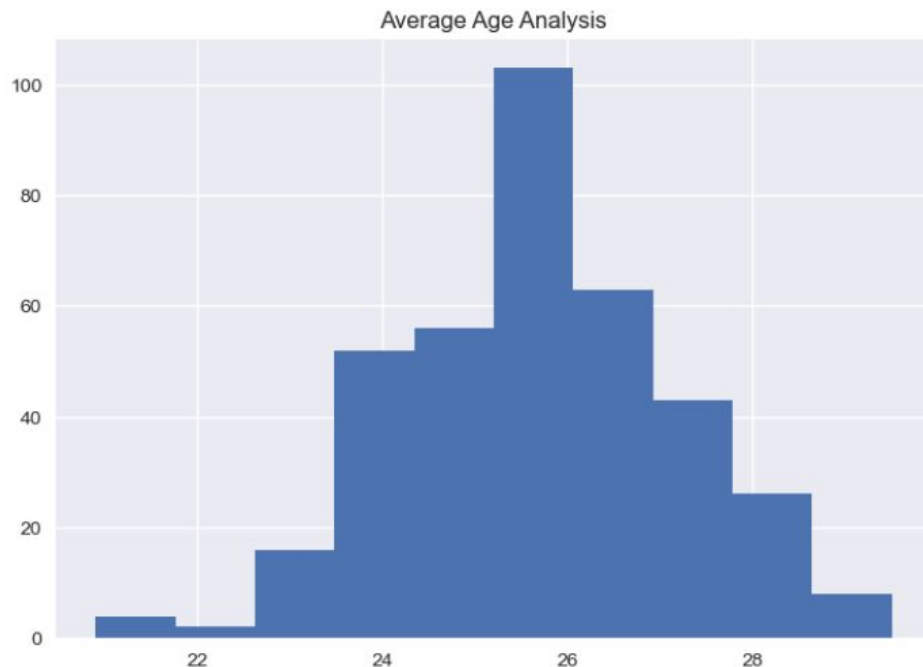
- More than 100 clubs have teams with an average age around 26.

- A team with an average age around 26 can be considered youthful.

- There are relatively few teams with average ages significantly younger or older than 26, suggesting that 26 is an ideal age for clubs.

Squad Size and Wins

```
[[1.        0.39450477]
 [0.39450477 1.       ]]
```
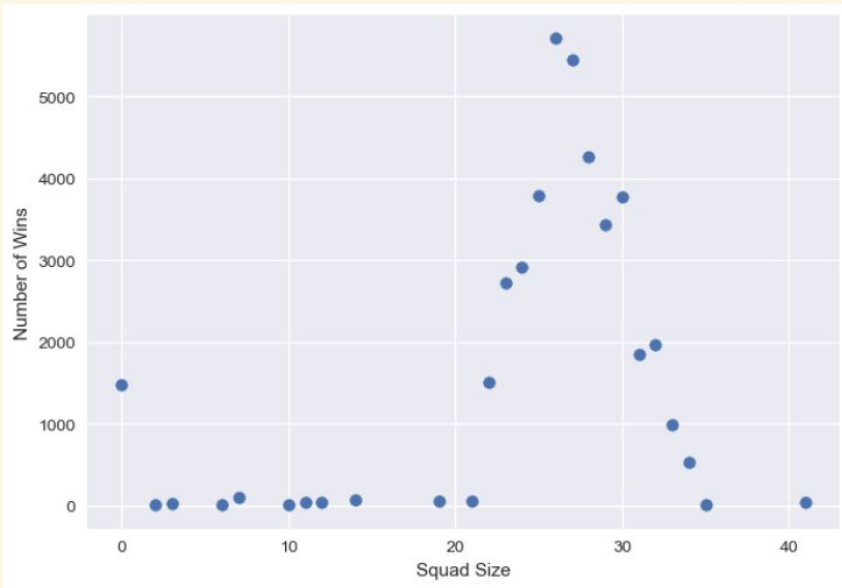

- Squad size and no. of wins have a moderate positive correlation 0.39.

- Teams with smaller squad sizes (below 20 players) generally have fewer wins.

- A significant concentration of data points is observed for squad sizes ranging between 25 and 30 players.

- Beyond a squad size of 30, the number of wins tends to decrease.

- The phenomenon of diminishing returns is evident in the data.

- There are a few notable outliers with very high or very low wins at certain squad sizes. These outliers can indicate exceptional cases or special circumstances.

Relationship between Squad Size and Win Rate

Probability Plot for Squad Size

Probability Plot for Wins
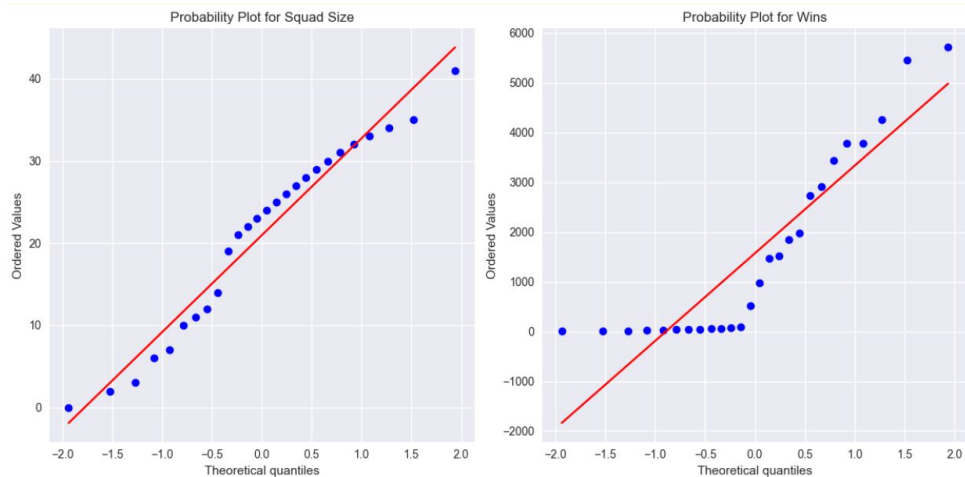
- The assumptions of the t-test were examined, and it was found that win rates did not follow a normal distribution.

- To enhance the normality of the data distribution, a natural logarithm transformation was applied to the wins data, aiding in achieving the assumptions required for the t-test.

- The results of the t-test indicated a significant association between squad size and win rates, suggesting that the size of a squad has a meaningful influence on the number of wins achieved.

- Based on the p-value, the null hypothesis of no association between squad size and wins was rejected.

Top 5 players for individual club

	Club name	Top players
0	FC Barcelona	[Robert Lewandowski, Sergio Agüero, Raphinha, ...
1	Paris Saint-Germain	[Lionel Messi, Neymar, Kylian Mbappé, Sergio R...
2	Manchester United	[Cristiano Ronaldo, Wout Weghorst, Bruno Ferna...
3	Tottenham Hotspur	[Harry Kane, Heung-min Son, Ivan Perisic, Luca...
4	Real Madrid	[Karim Benzema, Gareth Bale, Eden Hazard, Marc...
...	...	...
232	Torpedo Moscow	[Igor Lebedenko, Ilya Kukharchuk, Aleksandr Ry...
233	Gaziantep FK	[Torgeir Börven, Marko Jevtovic, André Sousa, ...
234	Metalist Kharkiv	[Dmytro Korkishko, Maksym Pryadun, Yegor Kartu...
235	Istanbulspor	[Valon Ethemi, Ertugrul Ersoy, Sidrit Guri, Ja...
236	FK Orenburg	[Joel Fameyeh, Vladimir Obukhov, Blagoy Georgi...

237 rows × 2 columns

- These are the top 5 players for individual clubs based on the amount of goals they score for that club.

Player nationality analysis
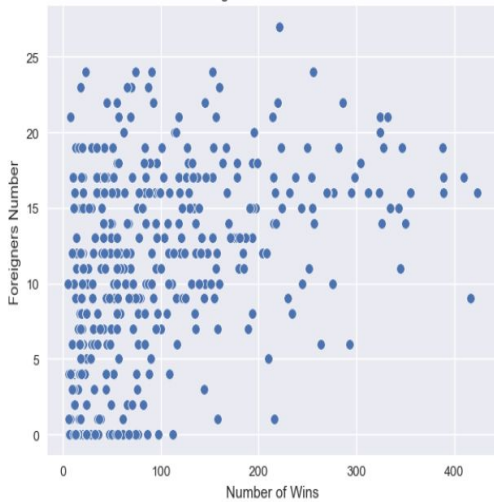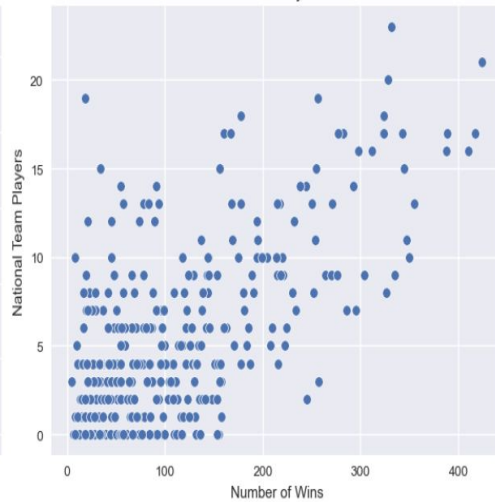
- The plots suggest that the presence of national team players in a squad has a more pronounced positive impact on securing wins compared to the presence of foreign players.

- While the correlation between winning outcomes and the presence of national team players was significant, the t-test did not find a statistically significant positive association between the presence of foreign players and improved club performance..

Ranking of Clubs

	name	score	rank
9	Bayern Munich	0.216166	1.0
64	Celtic FC	0.191035	2.0
101	Olympiacos Piraeus	0.187729	3.0
93	Ajax Amsterdam	0.175238	4.0
50	Manchester City	0.168389	5.0
...	...	...	...
118	Celta de Vigo	-0.021305	232.0
72	Udinese Calcio	-0.022362	233.0
129	UC Sampdoria	-0.022968	234.0
128	Bologna FC 1909	-0.026504	235.0
207	US Sassuolo	-0.045464	236.0

236 rows × 3 columns

- To rank the clubs based on their overall performance, the data was preprocessed, and a weighted score was calculated for each club. Weights were assigned to numerical features, such as goals scored, cards received, wins, and opponent statistics to perceived importance, with higher weights given to factors like goals scored and wins. For each club, a weighted score was calculated by combining the scaled feature values with their respective weights.

- This mean weighted score represents the club's overall performance across the considered features. Finally, the clubs were ranked based on these mean weighted scores. The higher the weighted score, the better the club's rank, indicating superior performance compared to lower-ranked clubs.

Yellow and Red Card Analysis

Interactive Bubble Chart
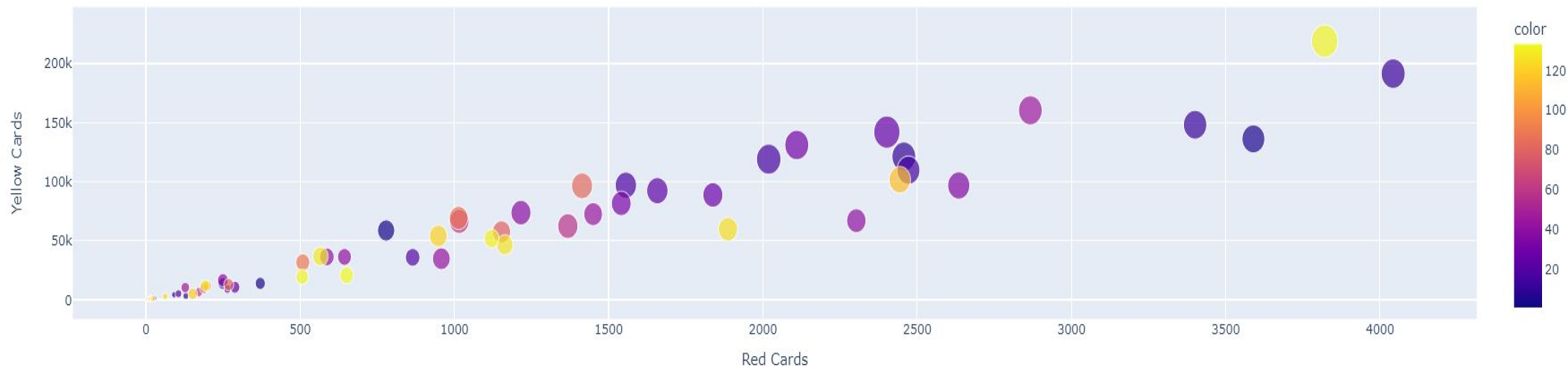
This bubble chart provides insights into how football clubs deal with player discipline, focusing on yellow and red cards ,By combining club game data with player event details.
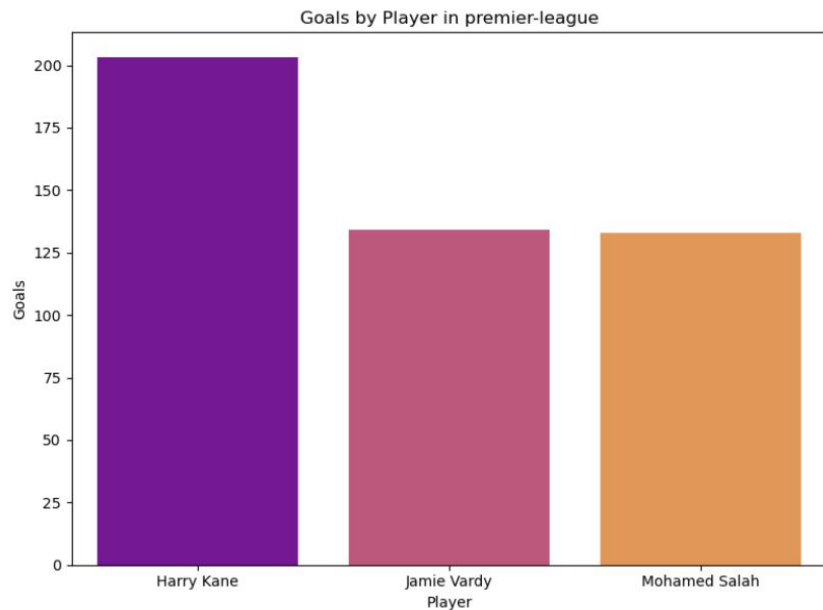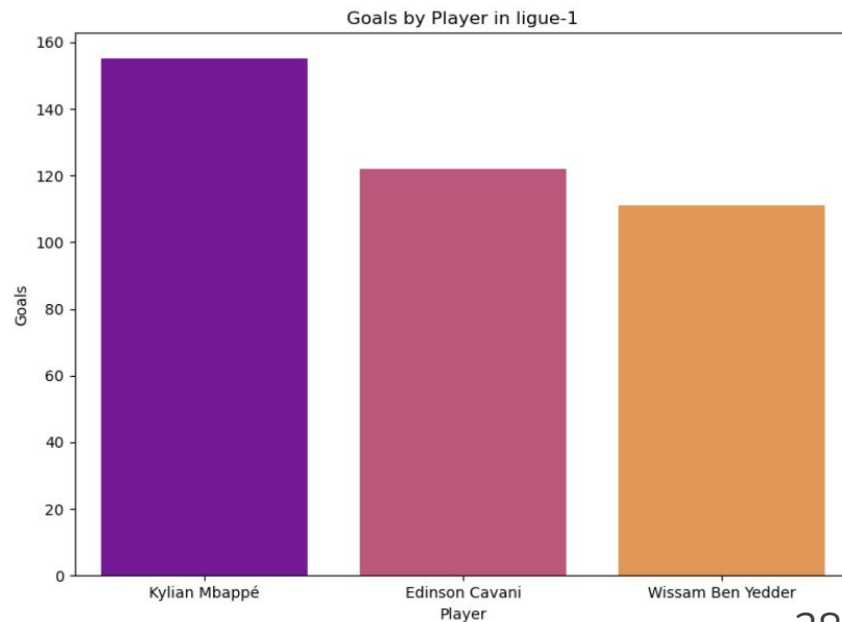
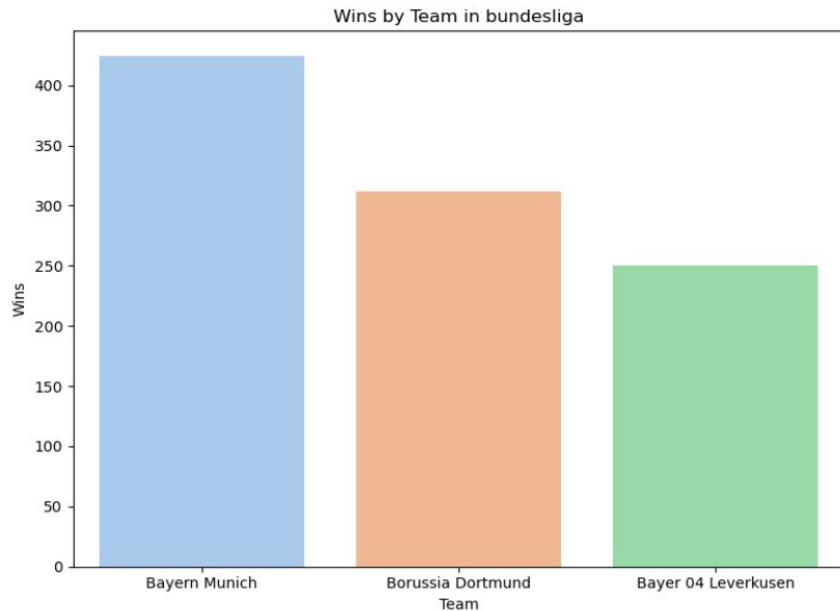# Competition Analysis

Top Scorers in Each League

Top Teams in Each League

- Bayern Munich has significantly more wins compared to the other two teams.

- This makes sense as Bayern Munich has maintained an unbroken winning streak for the past 11 consecutive years.

- Leagues can be selected with the Dropdown and the teams will change accordingly.

Home / Away Wins

Home Vs Away Club Performance Across Various Competitions
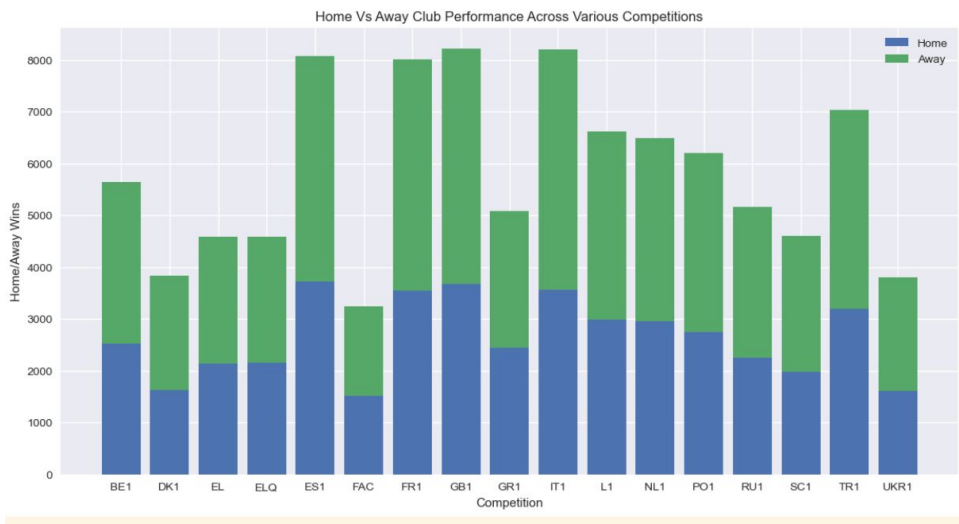
- The analysis compares clubs' performance in home versus away matches across major competitions. Findings reveal a notable difference in the number of "Home" and "Away" wins across well-represented competitions. This suggests the home team advantage plays a significant role, though some clubs excel at overcoming it in away matches.

- In La Liga (ES1), visiting teams tend to secure more wins.

Conclusion

Key takeaways and Future Prospects

Benefits for Coaches and Fans:

- This project offers invaluable insights for coaches and fans alike. By examining player attributes and performance metrics, it provides coaches with actionable information to enhance training programs and scouting efforts. Meanwhile, fans gain deeper insights into player dynamics, fostering a more engaging and informed experience of the sport.

Future Prospects:

- Future analysis could expand to include more comprehensive datasets, such as player performance metrics (e.g., goals, assists, passes), injury histories, and fitness levels. This would enable a more holistic understanding of player impact and career longevity.
- Machine learning models could be developed to predict future star players based on historical data, helping clubs invest in potential talent early.

The End

Dive deeper into the data and explore the full code and insights from this football EDA project. Visit our GitHub repository for detailed notebooks, visualizations, and further analyses.

🔗 [Check out the full project on GitHub](#)