

# CAPSTONE PROJECT 02

## English Premiere League

BY NOEL RODRIGUES



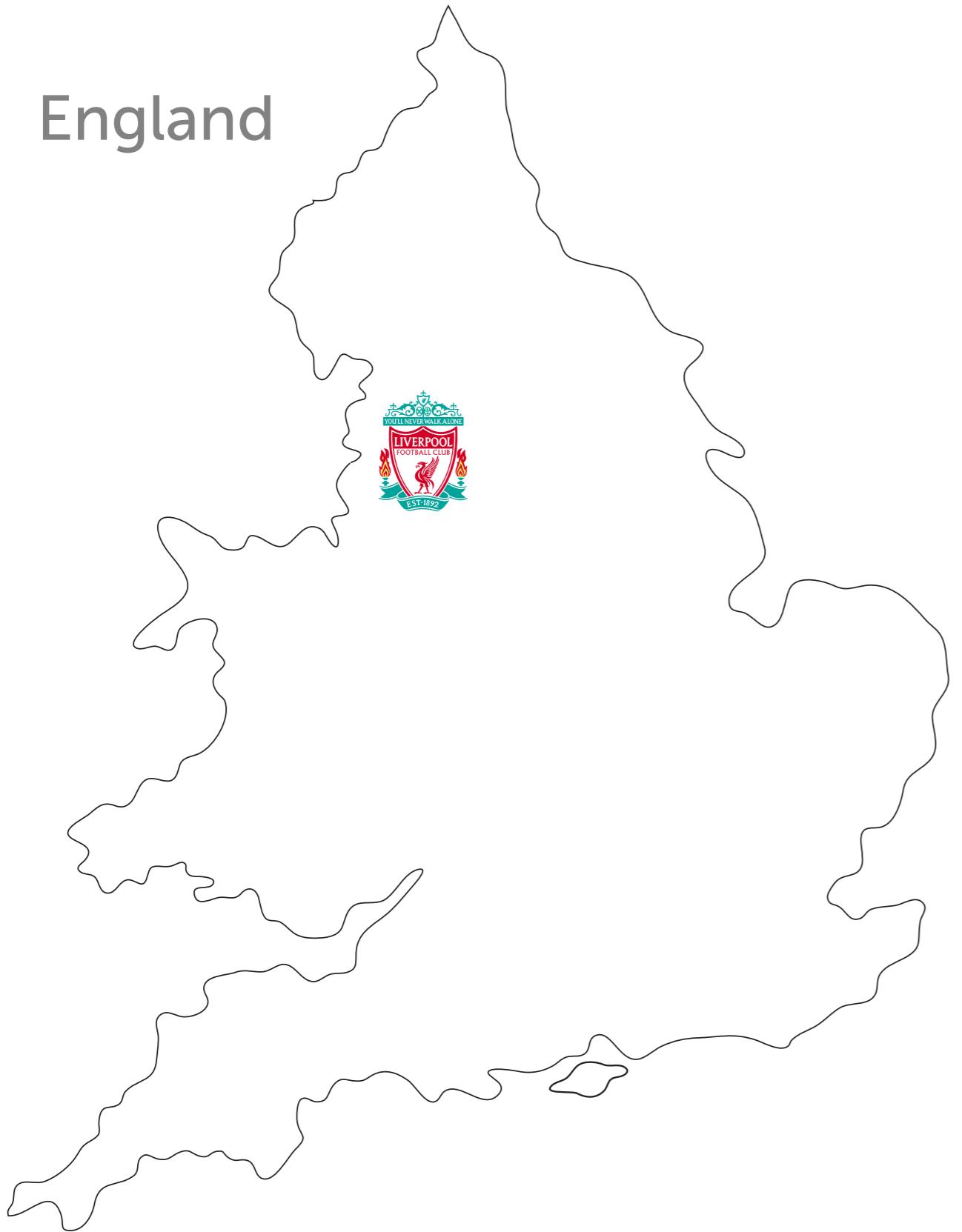
22 June 2020

**LIVERPOOL FC**  
**EPL Champions**  
**Season 2019/2020**

Last won in 1990



England





**EPL CHAMPIONS:**  
**1993, 1994, 1996, 1997, 1999,  
2000, 2001, 2003, 2007, 2008,  
2009, 2011, 2013**



**EPL CHAMPIONS:**  
**1991, 1998, 2002, 2004**



**EPL CHAMPIONS:**  
**2005, 2006, 2010, 2015, 2017**



**EPL CHAMPIONS:**  
**2016**



**EPL CHAMPIONS:**  
**2012, 2014, 2018, 2019**



England

## WEST HAM UNITED

- Had its hey days in the 60s and 70s.
- Last won a trophy — the FA Cup — in 1980.



Premier League

*2019/20 SEASON*

		P	GD	PTS
1	• Liverpool	38	52	99
2	• Manchester City	38	67	81
3	• Manchester United	38	30	66
4	• Chelsea	38	15	66
5	• Leicester City	38	26	62
6	• Tottenham Hotspur	38	14	59
7	• Wolverhampton Wanderers	38	11	59
8	• Arsenal	38	8	56
9	• Sheffield United	38	0	54
10	• Burnley	38	-7	54
11	• Southampton	38	-9	52
12	• Everton	38	-12	49
13	• Newcastle United	38	-20	44
14	• Crystal Palace	38	-19	43
15	• Brighton and Hove Albion	38	-15	41
16	• West Ham United	38	-13	39
17	• Aston Villa	38	-26	35
18	• Bournemouth	38	-25	34
19	• Watford	38	-28	34
20	• Norwich City	38	-49	21

### TABLE STANDING:

- 1 - Liverpool
- 2 - Manchester City
- 3 - Manchester United
- 4 - Chelsea
- 5 - Leicester City
- 6 - Tottenham Hotspur
- 
- 
- 
- 
- 
- 
- 
- 
- 16 - West Ham United



+

Data  
Science

==



???

# **CONCERNS**

## **THREE CONCERNS FROM WEST HAM UNITED**

- 1. Do expensive, big-money players win the league for the club?**
- 2. If the club can't afford big-money players, is there a superior playing style they should adopt?**
- 3. Which areas are West Ham players lacking in, compared to the players in the top teams?**



# **DATA SEARCH & DATA CLEANING**

kaggle™

Google  
Dataset Search



epl\_clubs\_2019\_2020



epl\_matches\_2019\_2020



epl\_players\_stats



player\_transfer\_value

- Standardize club names across all tables.

Screenshot of a database interface showing a table of transfer data and two SQL queries for standardizing club names.

**Table Data:**

	club_name	player_name	age	position	club_involved_name	fee	transfer_movement	fee_cleaned	league_name	year	season
1	AFC Bournemouth	Amaut Danjuma	22	Left Winger	Club Brugge	£16.20m	in	16.2	Premier League	2019	2019/2020
2	AFC Bournemouth	Philip Billing	23	Central Midfield	Huddersfield						
3	AFC Bournemouth	Lloyd Kelly	20	Left-Back	Bristol City						
4	AFC Bournemouth	Jack Stacey	23	Right-Back	Luton						
5	AFC Bournemouth	Harry Wilson	22	Right Winger	Liverpool						
6	AFC Bournemouth	Nnamdi Oforoborh	19	Central Midfield	Bournemouth U21						
7	AFC Bournemouth	Brad Smith	25	Left-Back	Sounders FC						
8	AFC Bournemouth	Sam Surridge	21	Centre-Forward	Swansea						
9	AFC Bournemouth	Nnamdi Oforoborh	20	Central Midfield	Wycombe						
10	AFC Bournemouth	Asmir Begovic	32	Goalkeeper	Qarabag Agdam						
11	AFC Bournemouth	Jemaine Defoe	37	Centre-Forward	Rangers						
12	AFC Bournemouth	Kyle Taylor	20	Central Midfield	Forest Green						
13	AFC Bournemouth	Emerson Hyndman	23	Central Midfield	Atlanta United						
14	AFC Bournemouth	Harry Arter	30	Central Midfield	Fulham						
15	AFC Bournemouth	Tyrone Mings	26	Centre-Back	Aston Villa						
16	AFC Bournemouth	Lys Mousset	23	Centre-Forward	Sheffield Utd.						
17	AFC Bournemouth	Connor Mahoney	22	Right Winger	Millwall						
18	AFC Bournemouth	Marc Pugh	32	Left Winger	QPR						
19	AFC Bournemouth	Harry Arter	29	Central Midfield	Fulham						
20	AFC Bournemouth	Nnamdi Oforoborh	19	Central Midfield	Wycombe						
21	AFC Bournemouth	Sam Surridge	21	Centre-Forward	Swansea						
22	AFC Bournemouth	Asmir Begovic	32	Goalkeeper	Qarabag Agdam						
23	AFC Bournemouth	Kyle Taylor	20	Central Midfield	Forest Green						
24	AFC Bournemouth	Emerson Hyndman	23	Central Midfield	Atlanta United						

**SQL Queries:**

```

SELECT *
FROM transfers_epl
ORDER BY club_name;

UPDATE transfers_epl
SET club_name = REPLACE(club_name, 'AFC Bournemouth', 'Bournemouth');
    
```

**Completion time:** 2020-11-11T17:30:58.3833695+08:00

**Results:**

club_name	player_name	age	position	club_involved_name	fee	transfer_movement
Aston Villa	Birkir Bjarnason	31	Central Midfield	Without Club	-	out
Aston Villa	Anwar El Ghazi	24	Right Winger	LOSC Lille	End of loan	Jun 30, 2019
Bournemouth	Amaut Danjuma	22	Left Winger	Club Brugge	£16.20m	in
Bournemouth	Philip Billing	23	Central Midfield	Huddersfield	£14.85m	in
Bournemouth	Lloyd Kelly	20	Left-Back	Bristol City	£13.32m	in
Bournemouth	Jack Stacey	23	Right-Back	Luton	£4.01m	in
Bournemouth	Harry Wilson	22	Right Winger	Liverpool	Loan fee: £2.43m	in
Bournemouth	Nnamdi Oforoborh	19	Central Midfield	Bournemouth U21	-	in
Bournemouth	Brad Smith	25	Left-Back	Sounders FC	End of loan	Dec 31, 2019
Bournemouth	Sam Surridge	21	Centre-Forward	Swansea	End of loan	May 31, 2020
Bournemouth	Nnamdi Oforoborh	20	Central Midfield	Wycombe	End of loan	May 31, 2020
Bournemouth	Asmir Begovic	32	Goalkeeper	Qarabag Agdam	End of loan	Dec 31, 2019
Bournemouth	Jemaine Defoe	37	Centre-Forward	Rangers	End of loan	May 31, 2020
Bournemouth	Kyle Taylor	20	Central Midfield	Forest Green	End of loan	May 31, 2020
Bournemouth	Emerson Hyndman	23	Central Midfield	Atlanta United	End of loan	Dec 31, 2019
Bournemouth	Harry Arter	30	Central Midfield	Fulham	End of loan	May 31, 2020
Bournemouth	Tyrone Mings	26	Centre-Back	Aston Villa	£20.07m	out
Bournemouth	Lys Mousset	23	Centre-Forward	Sheffield Utd.	£9.99m	out
Bournemouth	Connor Mahoney	22	Right Winger	Millwall	£990k	out

- Change the data type to be more accurate, such as changing the Transfer Fee column from INT to MONEY.

```
--Change datatype for Market_Value from INT to MONEY
ALTER TABLE player_transfer_value
ALTER COLUMN market_value_Euros MONEY;
```

dbo.player\_transfer\_value

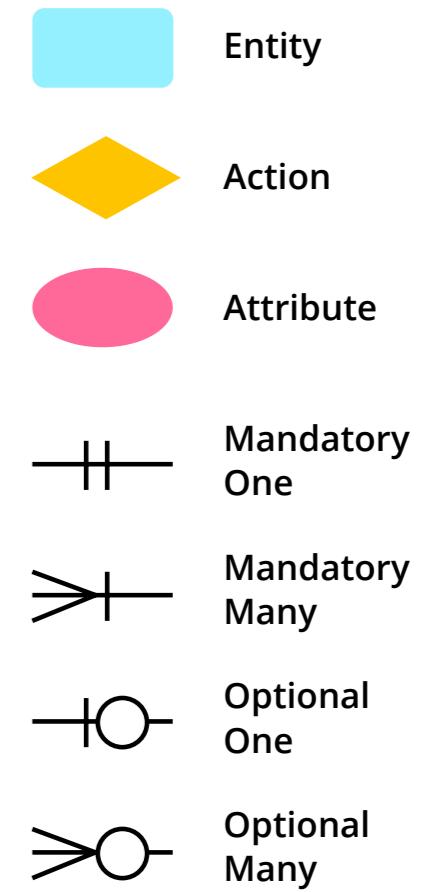
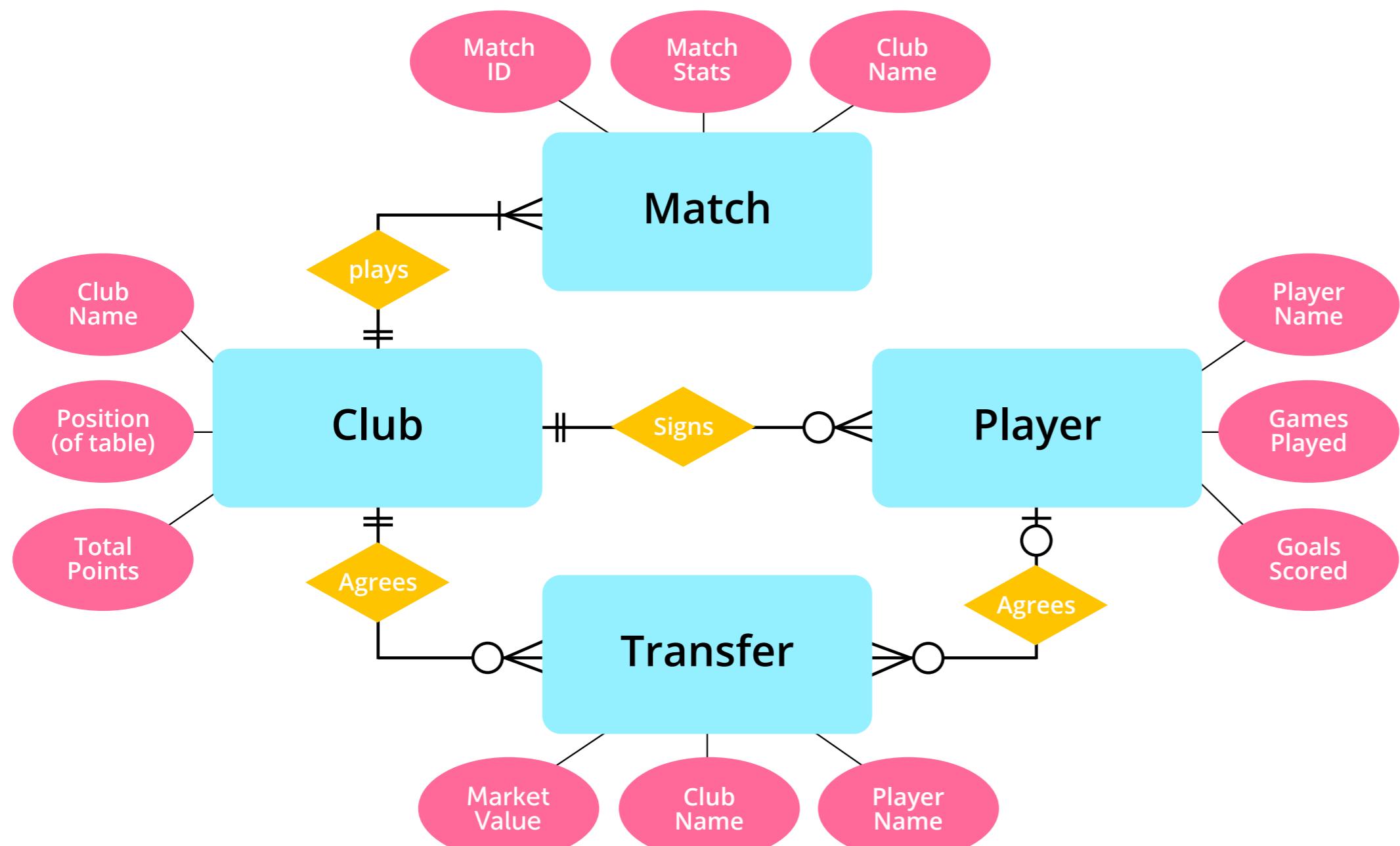
Columns

- player\_name (nvarchar(50), not null)
- player\_id (int, not null)
- FIFA\_id (int, not null)
- player\_position (nvarchar(50), not null)
- club\_name (nvarchar(50), not null)
- birthdate (datetime2(7), not null)
- age (int, not null)
- birth\_place (nvarchar(50), not null)
- height (nvarchar(50), not null)
- weight (int, not null)
- market\_value\_Euros (money, null)
- date\_joined (datetime2(7), not null)
- sponsor (nvarchar(50), null)
- wage (int, not null)
- all\_positions (nvarchar(50), not null)
- foot (nvarchar(50), not null)

# **ENTITY RELATIONSHIP DIAGRAM**

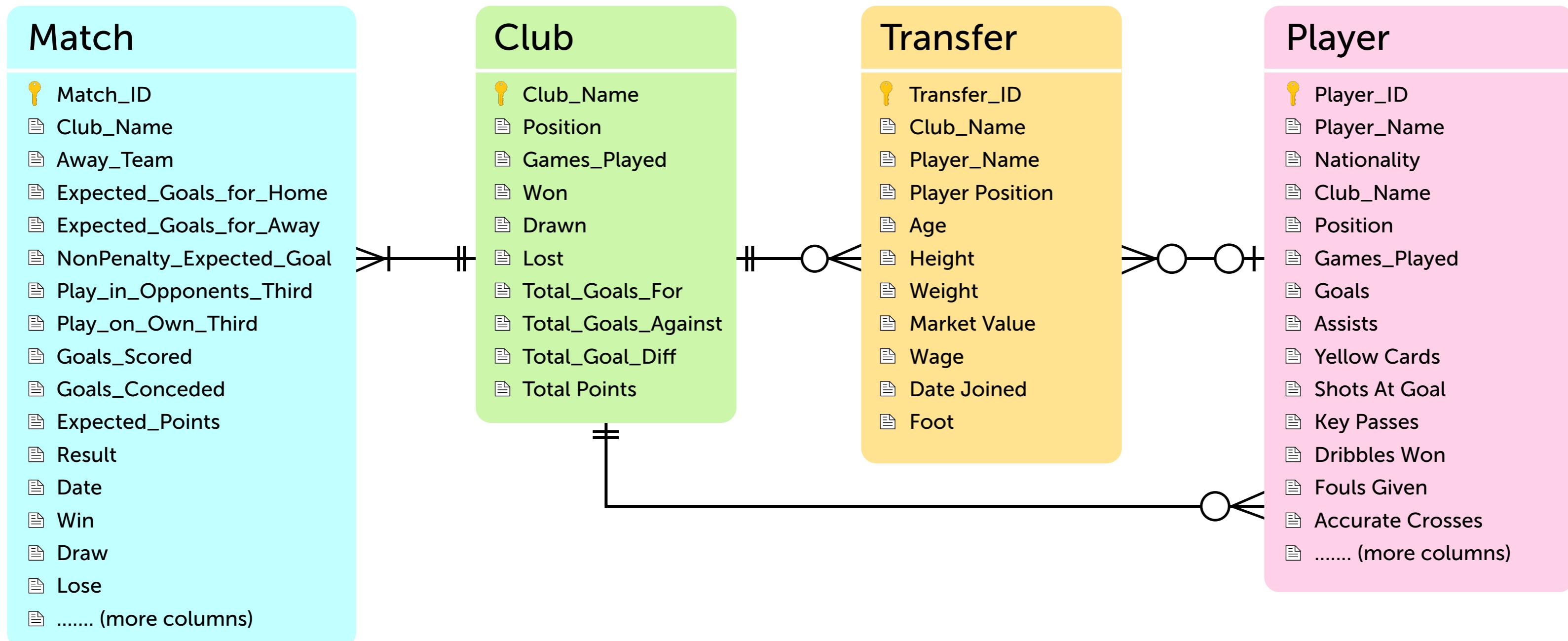
# ENTITY RELATIONSHIP DIAGRAM

English Premiere League (Season 2019 / 2020)



# ENTITY RELATIONSHIP DIAGRAM

English Premiere League (Season 2019 / 2020)



# **FIRST CONCERN / SQL QUERIES**

## 1. Do expensive, big-money players win the league for the club?

- First find out what is an “expensive” player.
- By finding median market value of all 515 players.

```
--Finding the Median Market Value of 500 players
SELECT (
(
    SELECT MAX(market_value_Euros)
    FROM (
        SELECT TOP 50 PERCENT market_value_Euros
        FROM player_transfer_value
        ORDER BY market_value_Euros) AS bottomhalf
)
+
(
    SELECT MIN(market_value_Euros)
    FROM (
        SELECT TOP 50 PERCENT market_value_Euros
        FROM player_transfer_value
        ORDER BY market_value_Euros DESC) AS tophalf
)
)
/ 2 AS median_value_of_player;
```

120 %

	median_value_of_player
1	10000000.00

## 1. Do expensive, big-money players win the league for the club?

- Now we know that the median market value of an EPL player is \$10 million, let's find out if the top clubs consist of expensive players.
- JOIN Club table with Transfer table.

```
--Finding the Top Most Expensive Players (Market Value) in relation to their club position  
SELECT TOP 50 t.player_name, t.club_name, t.market_value_Euros, c.Position AS 'table_position'  
FROM player_transfer_value AS t  
INNER JOIN epl_clubs_2019_2020 AS c  
ON t.club_name = c.club_name  
ORDER BY t.market_value_Euros DESC;
```

	player_name	club_name	market_value_Euros	table_position
1	Mohamed Salah	Liverpool	150000000.00	1
2	Harry Kane	Tottenham Hotspur	150000000.00	6
3	Raheem Sterling	Manchester City	140000000.00	2
4	Kevin De Bruyne	Manchester City	130000000.00	2
5	Sadio Mané	Liverpool	120000000.00	1
6	Virgil van Dijk	Liverpool	100000000.00	1
7	N'Golo Kanté	Chelsea	100000000.00	4
8	Paul Pogba	Manchester United	100000000.00	3
9	Christian Eriksen	Tottenham Hotspur	100000000.00	6
10	Bernardo Silva	Manchester City	100000000.00	2
11	Leroy Sané	Manchester City	100000000.00	2
12	Dele Alli	Tottenham Hotspur	90000000.00	6
13	Alisson	Liverpool	80000000.00	1
14	Heung-min Son	Tottenham Hotspur	80000000.00	6
15	Roberto Firmino	Liverpool	80000000.00	1
16	Rodrigo	Manchester City	80000000.00	2
17	Trent Alexander-Arnold	Liverpool	80000000.00	1
18	Marcus Rashford	Manchester United	80000000.00	3
19	Aymeric Laporte	Manchester City	75000000.00	2
20	Nicolas Pépé	Arsenal	75000000.00	8
21	Alexandre Lacazette	Arsenal	70000000.00	9

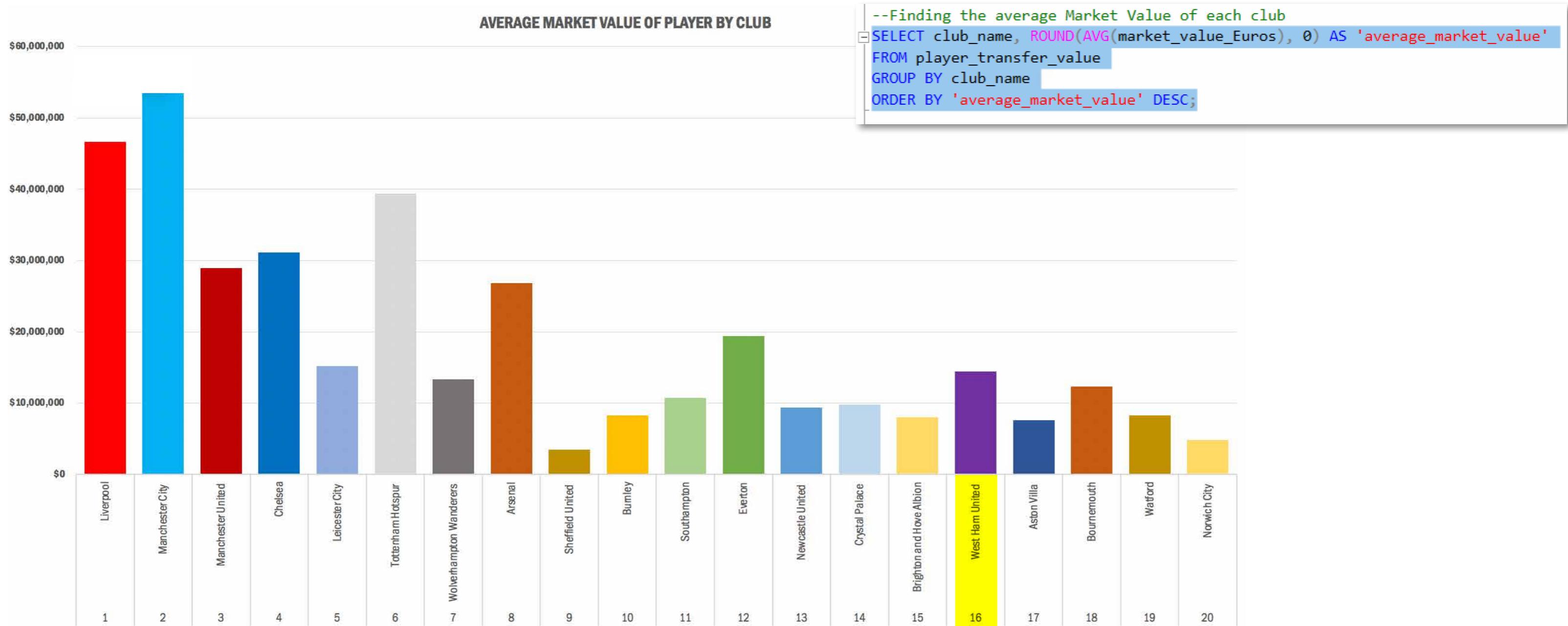
# 1. Do expensive, big-money players win the league for the club?

- The top 50 players are all valued above the median of \$10 million.
- 41 out of 50 of the most expensive players are with the top six clubs.

player_name	club_name	market_value_Euros	table_position
Mohamed Salah	Liverpool	\$150,000,000.00	1
Harry Kane	Tottenham Hotspur	\$150,000,000.00	6
Raheem Sterling	Manchester City	\$140,000,000.00	2
Kevin De Bruyne	Manchester City	\$130,000,000.00	2
Sadio Mané	Liverpool	\$120,000,000.00	1
Virgil van Dijk	Liverpool	\$100,000,000.00	1
N'Golo Kanté	Chelsea	\$100,000,000.00	4
Paul Pogba	Manchester United	\$100,000,000.00	3
Christian Eriksen	Tottenham Hotspur	\$100,000,000.00	6
Bernardo Silva	Manchester City	\$100,000,000.00	2
Leroy Sané	Manchester City	\$100,000,000.00	2
Dele Alli	Tottenham Hotspur	\$90,000,000.00	6
Alisson	Liverpool	\$80,000,000.00	1
Heung-min Son	Tottenham Hotspur	\$80,000,000.00	6
Roberto Firmino	Liverpool	\$80,000,000.00	1
Rodrigo	Manchester City	\$80,000,000.00	2
Trent Alexander-Arnold	Liverpool	\$80,000,000.00	1
Marcus Rashford	Manchester United	\$80,000,000.00	3
Aymeric Laporte	Manchester City	\$75,000,000.00	2
Nicolas Pépé	Arsenal	\$75,000,000.00	8
Alexandre Lacazette	Arsenal	\$70,000,000.00	8
Pierre-Emerick Aubameyang	Arsenal	\$70,000,000.00	8
Ederson	Manchester City	\$70,000,000.00	2
Harry Maguire	Manchester United	\$70,000,000.00	3
Gabriel Jesus	Manchester City	\$70,000,000.00	2
David de Gea	Manchester United	\$65,000,000.00	3
Sergio Agüero	Manchester City	\$65,000,000.00	2
Jorginho	Chelsea	\$65,000,000.00	4
Tanguy Ndombélé	Tottenham Hotspur	\$65,000,000.00	6
Andrew Robertson	Liverpool	\$60,000,000.00	1
Kepa Arrizabalaga	Chelsea	\$60,000,000.00	4
Riyad Mahrez	Manchester City	\$60,000,000.00	2
John Stones	Manchester City	\$60,000,000.00	2
Anthony Martial	Manchester United	\$60,000,000.00	3
Christian Pulisic	Chelsea	\$60,000,000.00	4
Wilfried Zaha	Crystal Palace	\$55,000,000.00	14
João Cancelo	Manchester City	\$55,000,000.00	2
Davinson Sánchez	Tottenham Hotspur	\$55,000,000.00	6
Lucas Torreira	Arsenal	\$55,000,000.00	8
Antonio Rüdiger	Chelsea	\$50,000,000.00	4
Fabinho	Liverpool	\$50,000,000.00	1
Georginio Wijnaldum	Liverpool	\$50,000,000.00	1
Ilkay Gündogan	Manchester City	\$50,000,000.00	2
Kyle Walker	Manchester City	\$50,000,000.00	2
Naby Keïta	Liverpool	\$50,000,000.00	1
Giovani Lo Celso	Tottenham Hotspur	\$50,000,000.00	6
Rúben Neves	Wolverhampton Wanderers	\$50,000,000.00	7
Richarlison	Everton	\$50,000,000.00	12
Declan Rice	West Ham United	\$50,000,000.00	16
Felipe Anderson	West Ham United	\$45,000,000.00	16

# 1. Do expensive, big-money players win the league for the club?

- The top six clubs' players do have a higher market value.



## 1. Do expensive, big-money players win the league for the club?

- In conclusion, big-money players do help elevate the success of the clubs.
- It does not, however, guarantee success as evidenced by the spending by Arsenal, Everton and West Ham United.

player_name	club_name	market_value_Euros	table_position
Mohamed Salah	Liverpool	\$150,000,000.00	1
Harry Kane	Tottenham Hotspur	\$150,000,000.00	6
Raheem Sterling	Manchester City	\$140,000,000.00	2
Kevin De Bruyne	Manchester City	\$130,000,000.00	2
Sadio Mané	Liverpool	\$120,000,000.00	1
Virgil van Dijk	Liverpool	\$100,000,000.00	1
N'Golo Kanté	Chelsea	\$100,000,000.00	4
Paul Pogba	Manchester United	\$100,000,000.00	3
Christian Eriksen	Tottenham Hotspur	\$100,000,000.00	6
Bernardo Silva	Manchester City	\$100,000,000.00	2
Leroy Sané	Manchester City	\$100,000,000.00	2
Dele Alli	Tottenham Hotspur	\$90,000,000.00	6
Alisson	Liverpool	\$80,000,000.00	1
Heung-min Son	Tottenham Hotspur	\$80,000,000.00	6
Roberto Firmino	Liverpool	\$80,000,000.00	1
Rodrigo	Manchester City	\$80,000,000.00	2
Trent Alexander-Arnold	Liverpool	\$80,000,000.00	1
Marcus Rashford	Manchester United	\$80,000,000.00	3
Aymeric Laporte	Manchester City	\$75,000,000.00	2
Nicolas Pépé	Arsenal	\$75,000,000.00	8
Alexandre Lacazette	Arsenal	\$70,000,000.00	8
Pierre-Emerick Aubameyang	Arsenal	\$70,000,000.00	8
Ederson	Manchester City	\$70,000,000.00	2
Harry Maguire	Manchester United	\$70,000,000.00	3
Gabriel Jesus	Manchester City	\$70,000,000.00	2

2019/20 SEASON			
	P	GD	PTS
1 • Liverpool	38	52	99
2 • Manchester City	38	67	81
3 • Manchester United	38	30	66
4 • Chelsea	38	15	66
5 • Leicester City	38	26	62
6 • Tottenham Hotspur	38	14	59
7 • Wolverhampton Wanderers	38	11	59
8 • Arsenal	38	8	56
9 • Sheffield United	38	0	54
10 • Burnley	38	-7	54
11 • Southampton	38	-9	52
12 • Everton	38	-12	49
13 • Newcastle United	38	-20	44
14 • Crystal Palace	38	-19	43
15 • Brighton and Hove Albion	38	-15	41
16 • West Ham United	38	-13	39
17 • Aston Villa	38	-26	35

# **SECOND CONCERN / SQL QUERIES**

## 2. If the club can't afford big-money players, is there a superior playing style they should adopt?

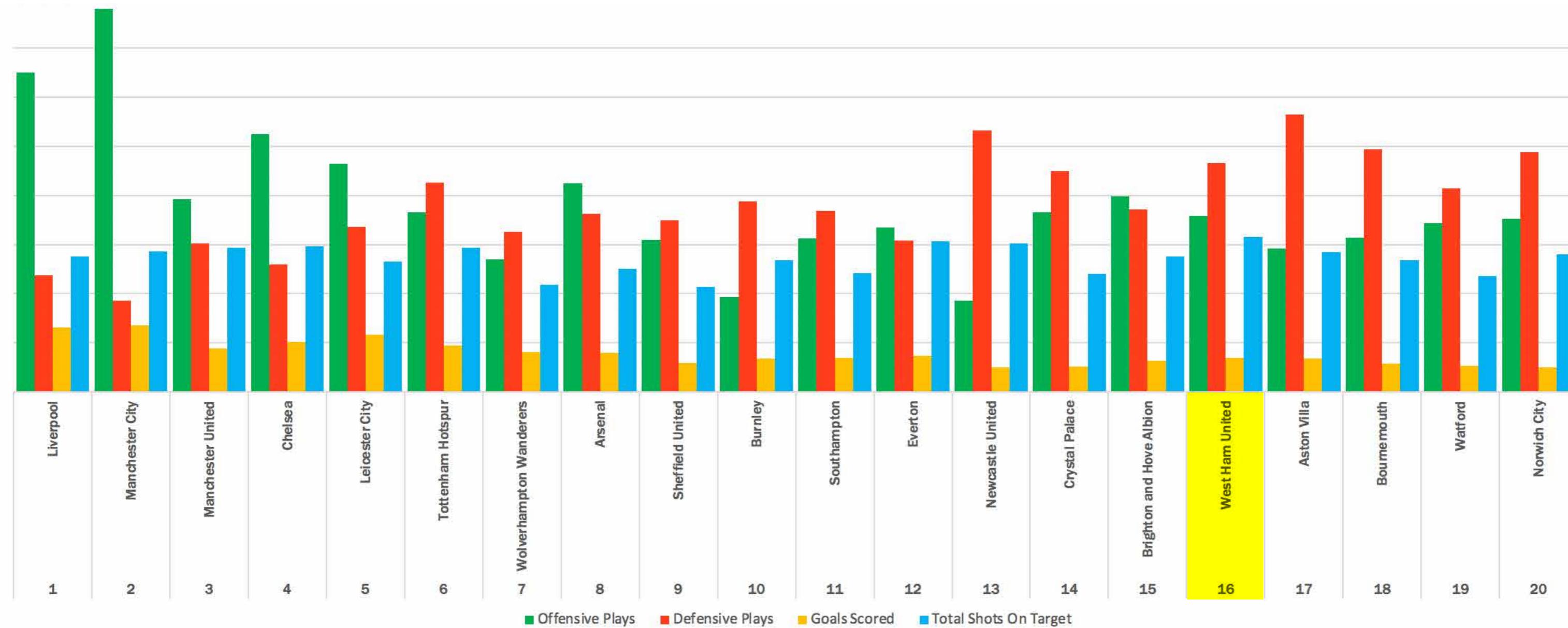
- Use the Match table to retrieve relevant match stats for comparison by club.

```
--Find match stats of each club
SELECT
    club_name AS 'Club Name',
    SUM(deep) AS 'Offensive Plays',
    SUM(deep_allowed) AS 'Defensive Plays',
    SUM(scored) AS 'Goals Scored',
    SUM(HS_x) AS 'Total Shots',
    SUM(HST_x) AS 'Total Shots On Target',
    ROUND(SUM(ppda_cal), 4) AS 'Measure of Pressing Play'
FROM epl_matches_2019_2020
GROUP BY club_name;
```

Results

Club Name	Offensive Plays	Defensive Plays	Goals Scored	Total Shots	Total Shots On Target	Measure of Pressing Play
Arsenal	212	181	40	386	125	330.9926
Aston Villa	146	282	34	481	142	357.9497
Bournemouth	157	247	29	376	134	405.6173
Brighton and Hove Albion	199	186	32	410	138	275.7663
Burnley	97	194	34	400	134	360.8156
Chelsea	262	130	51	412	148	272.6223
Crystal Palace	183	225	26	400	120	410.7902
Everton	167	154	37	386	153	296.2872
Leicester City	232	168	58	373	133	277.9457
Liverpool	325	119	66	384	138	258.4836
Manchester City	390	93	68	404	143	239.2871
Manchester United	196	151	44	416	147	306.0677
Newcastle United	93	266	25	394	151	662.2712
Norwich City	176	244	25	395	140	379.0199
Sheffield United	155	175	30	352	107	389.4123
Southampton	156	184	35	358	121	260.4134
Tottenham Hotspur	183	213	47	433	147	380.2675
Watford	172	207	27	410	118	401.2819
West Ham United	179	233	35	400	158	416.3616
Wolverhampton Wanderers	135	163	41	366	109	404.5678

## 2. If the club can't afford big-money players, is there a superior playing style they should adopt?



## Exporting from SQL (Virtual Space) to EXCEL

The screenshot shows the context menu for a column named 'player\_name' in the 'dbo.player\_transfer\_value' table. The 'Tasks' section is highlighted, and the 'Export Data...' option is selected.

The series of screenshots illustrates the process of running an SSIS package:

- Choose a Data Source:** Selects 'SQL Server Native Client 11.0' as the data source and 'AdventureWorksLT' as the database.
- Choose a Destination:** Sets the destination to 'Microsoft Excel' and specifies the file path as 'D:\Users\agunitd3\Documents\QueryResultsForEXCEL\FindAverageWagesOfEachClub.xls'.
- Provide a Source Query:** Enters the SQL query to select average market value by club name.
- Select Source Tables and Views:** Preview data from the source query.
- Review Data Type Mapping:** Maps columns from the source to the destination.
- Save and Run Package:** Chooses to run immediately and saves the package.
- Complete the Wizard:** Verifies the choices and clicks 'Finish'.
- The execution was successful:** Shows a summary of the execution results with 11 rows transferred.

Action	Status	Message
Initializing Data Row Task	Success	
Initializing Connectors	Success	
Setting SQL Command	Success	
Setting Source Connection	Success	
Setting Destination Connection	Success	
Validating	Success	
Preparing for Execute	Success	
Pre-execute	Success	
Executing	Success	
Copying to 'Query'	Success	20 rows transferred
Post-execute	Success	

# **THIRD CONCERN / SQL QUERIES**

### 3. Which areas are West Ham players lacking in, compared to the players in the top teams?

- Use Players table and join with the Transfer table to retrieve the Top 20 most expensive players and their respective stats.
- Similarly, retrieve the stats of the West Ham players for comparison

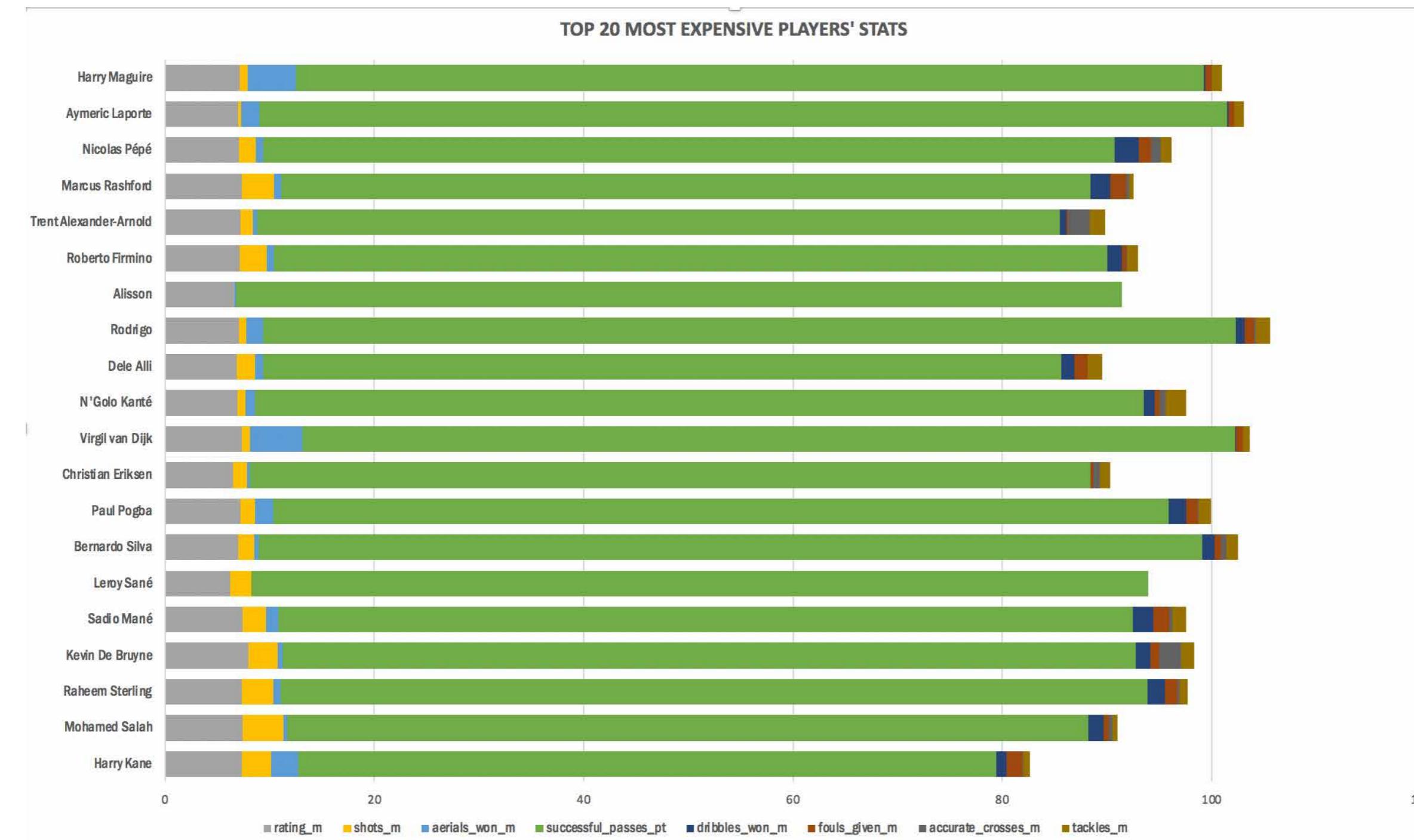
```
--Find West Ham average player stats
SELECT rating_m, shots_m, aerials_won_m, successful_passes_pt, dribbles_won_m, fouls_given_m, accurate_crosses_m, tackles_m
FROM epl_players_stats
WHERE club_name = 'West Ham United';
```

	rating_m	shots_m	aerials_won_m	successful_passes_pt	dribbles_won_m	fouls_given_m	accurate_crosses_m	tackles_m
1	6.62	0.5	1.5	74.6	0.4	0.5	0.8	1.4
2	6.67	0.5	0.5	82.8	0.8	0.9	0.4	1.6
3	7.14	1.5	5.5	70.1	0	0.7	0	2.2
4	6.79	0.9	0.6	77.8	0.8	1.3	1.6	1.2
5	6.48	1	0.3	78.8	0.4	0.4	0.3	1.3
6	6.43	0	0.2	40.5	0	0	0	0
7	6.55	1	0.2	86.9	1.4	1.1	0.6	0.6
8	6.76	0.3	2.3	72.7	0.3	0.3	0.7	2.7
9	6.42	0	0.8	75.5	0.2	0.9	0.1	1.8
10	6.6	0.3	0.7	80.3	1.6	0.5	0.3	2.4
11	6.74	0.4	3.3	76.2	0.2	0.5	0	1.2
12	6.3	1.5	0	85	0	0	0	0
13	7.29	2.8	3.8	68.4	2.7	1.5	0.4	0.5
14	6.99	1.8	0.4	77	1.2	1.5	1.2	1.5
15	6.51	0.6	3.2	71.6	0.1	0.1	0	1
16	7.08	1.8	6.8	65.8	0.8	1.2	0	0.7

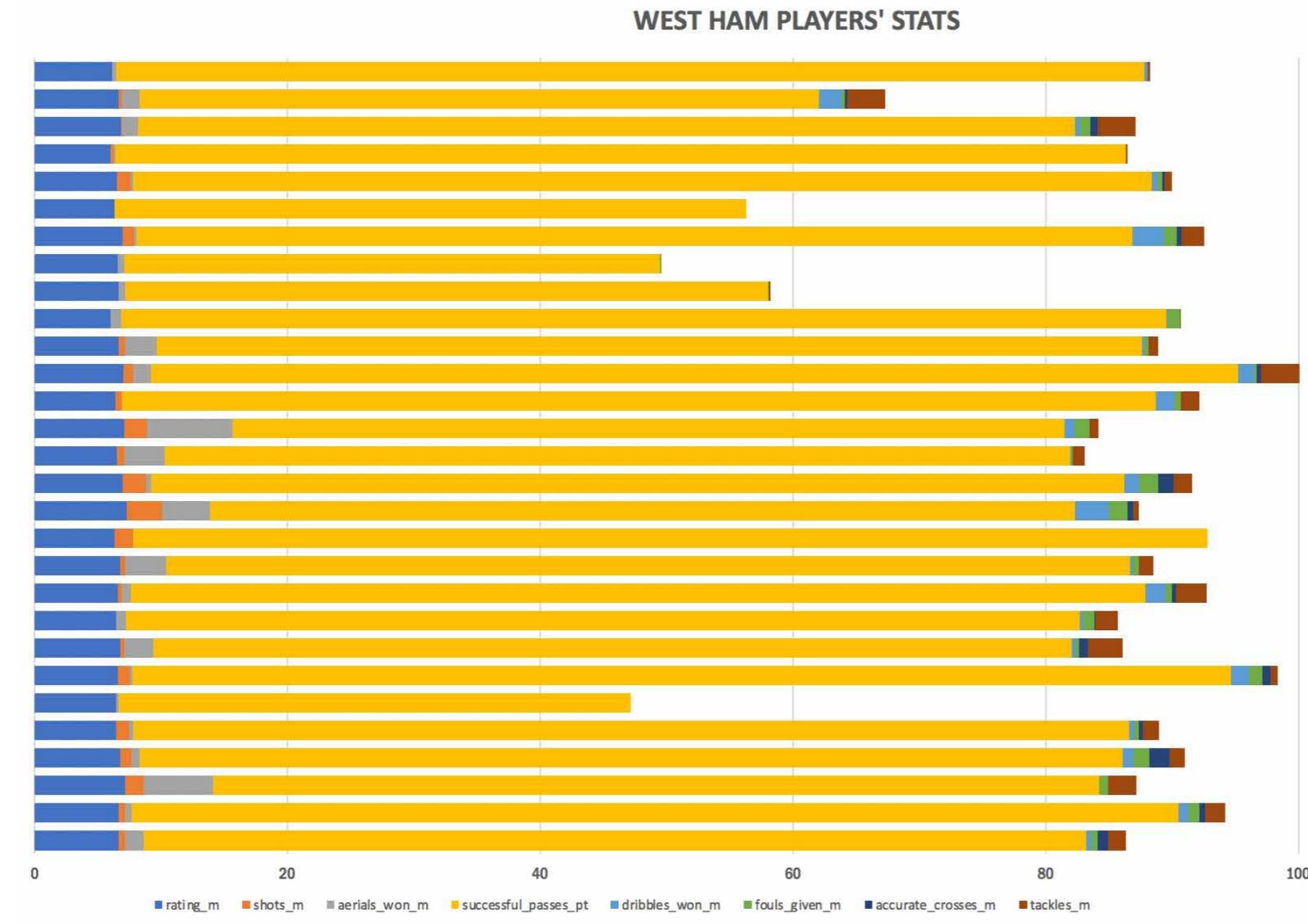
```
SELECT TOP 20 p.player_name, t.market_value_Euros, p.club_name, p.rating_m, p.shots_m, p.aerials_won_m, p.successful_passes_pt, p.dribbles_won_m, p.fouls_given_m, p.accurate_crosses_m, p.tackles_m
FROM epl_players_stats AS p
INNER JOIN player_transfer_value AS t
ON (p.player_name = t.player_name)
ORDER BY t.market_value_Euros DESC;
```

player_name	market_value_Euros	club_name	rating_m	shots_m	aerials_won_m	successful_passes_pt	dribbles_won_m	fouls_given_m	accurate_crosses_m	tackles_m
1 Harry Kane	150000000.00	Tottenham Hotspur	7.33	2.8	2.6	66.7	1	1.5	0.1	0.6
2 Mohamed Salah	150000000.00	Liverpool	7.4	3.9	0.4	76.5	1.5	0.5	0.3	0.5
3 Raheem Sterling	140000000.00	Manchester City	7.35	3	0.7	82.8	1.7	1.2	0.2	0.8
4 Kevin De Bruyne	130000000.00	Manchester City	7.97	2.8	0.5	81.5	1.4	0.8	2.1	1.3
5 Sadio Mané	120000000.00	Liverpool	7.45	2.2	1.2	81.6	2	1.5	0.3	1.3
6 Leroy Sané	100000000.00	Manchester City	6.23	2	0	85.7	0	0	0	0
7 Bernardo Silva	100000000.00	Manchester City	7.03	1.5	0.4	90.2	1.2	0.6	0.5	1.1
8 Paul Pogba	100000000.00	Manchester United	7.18	1.4	1.8	85.5	1.7	1.1	0.1	1.2
9 Christian Eriksen	100000000.00	Tottenham Hotspur	6.51	1.3	0.3	80.3	0	0.3	0.6	1
10 Virgil van Dijk	100000000.00	Liverpool	7.32	0.8	5	89.1	0.1	0.7	0	0.6
11 N'Golo Kanté	100000000.00	Chelsea	6.9	0.8	0.9	84.9	1.1	0.5	0.5	2
12 Dele Alli	90000000.00	Tottenham Hotspur	6.87	1.7	0.8	76.3	1.2	1.3	0	1.4
13 Rodrigo	80000000.00	Manchester City	7.09	0.7	1.6	92.9	0.9	0.9	0.1	1.4
14 Alisson	80000000.00	Liverpool	6.62	0	0.2	84.6	0	0	0	0
15 Roberto Firmino	80000000.00	Liverpool	7.14	2.6	0.7	79.6	1.4	0.4	0.1	1
16 Trent Alexander-Arnold	80000000.00	Liverpool	7.2	1.2	0.4	76.7	0.6	0.2	2.1	1.4
17 Marcus Rashford	80000000.00	Manchester United	7.35	3.1	0.7	77.3	1.9	1.5	0.3	0.4
18 Nicolas Pépé	75000000.00	Arsenal	7.06	1.6	0.7	81.4	2.3	1.2	0.9	1
19 Aymeric Laporte	75000000.00	Manchester City	7	0.3	1.7	92.5	0.1	0.6	0	0.9
20 Harry Maguire	70000000.00	Manchester United	7.12	0.8	4.6	86.7	0.2	0.6	0	1

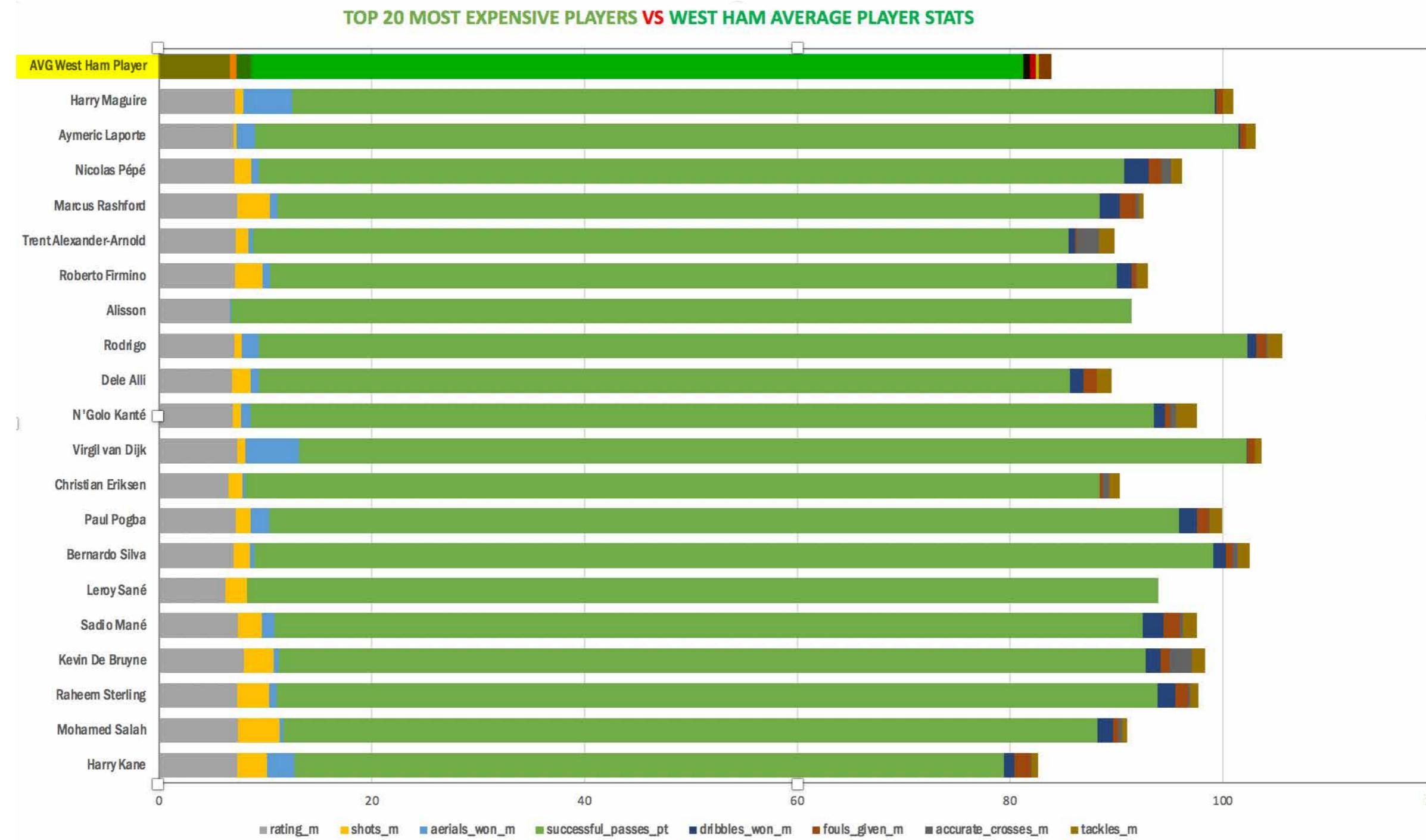
### 3. Which areas are West Ham players lacking in, compared to the players in the top teams?



### 3. Which areas are West Ham players lacking in, compared to the players in the top teams?



### 3. Which areas are West Ham players lacking in, compared to the players in the top teams?



3. Which areas are West Ham players lacking in, compared to the players in the top teams?

- Compared to the top 20 most expensive players, West Ham players are, on average, not too far behind in terms of skills.
- The areas that they need to improve on are:
  - Shots at goals (take more shots)
  - Headers (aerial contests)
  - Passing (need to be more accurate)
  - Dribbles (need to be more skillful)



# **CONCLUSION**

## THREE CONCERNS FROM WEST HAM UNITED

1. Do expensive, big-money players win the league for the club?

Yes, but doesn't guarantee success.

2. If the club can't afford big-money players, is there a superior playing style they should adopt?

Be more offensive and accurate in shooting.

3. Which areas are West Ham players lacking in, compared to the players in the top teams?

Shots at goals, headers, accurate passing and effective dribbles



## **ACKNOWLEDGMENT:**

Zaleha Ali for her kind, genuine assistance

## **SOURCES:**

- <https://www.kaggle.com/idoyo92/epl-stats-20192020>
- <https://www.kaggle.com/irkaal/english-premier-league-results>
- <https://www.kaggle.com/saife245/english-premier-league>
- <https://www.kaggle.com/cashncarry/epl-players-deep-stats-20192020>
- <https://www.kaggle.com/aricht1995/premier-league-epl-player-information>
- <https://www.kaggle.com/hugomathien/soccer>
- <https://www.premierleague.com/>
- <https://www.21stclub.com/>
- <https://towardsdatascience.com/visualizing-the-2019-20-english-premier-league-season-with-matplotlib-and-pandas-fd491a07cfda>
- [https://en.wikipedia.org/wiki/West\\_Ham\\_United\\_F.C.](https://en.wikipedia.org/wiki/West_Ham_United_F.C.)
- <https://wwwexasol.com/what-difference-can-data-make-for-a-football-team/>
- <https://www.ft.com/content/84aa8b5e-c1a9-11e8-84cd-9e601db069b8>
- <https://www.intel.co.uk/content/www/uk/en/it-management/cloud-analytic-hub/data-powered-football.html>



# TRANSCRIPT

On 22 June this year, Liverpool Football Club finally won the English Premiere League after 30 years.

This north-western English club was a major force in the 70s and 80s but since then,

other football clubs like Manchester United, Arsenal and Chelsea have overtaken them in winning trophies.

Moving down to East London, West Ham United is another football club in the EPL (which is short for English Premier League). West Ham had its heydays in the 60s and 70s. Sadly, their last silverware was 40 years ago (when Anting, Michelle and Billy weren't even born).

The top six teams in the last season were Liverpool, Manchester City, Manchester United, Chelsea, Leicester City, and Tottenham Hotspur. West Ham came in 16th out of 20 clubs.

Even after buying scores of players and changing managers several times, the club owners of West Ham, are finally betting on data science to change their game, so that they can go back to winning ways.

## CONCERNS

Here are three concerns from West Ham United:

- 1. Do expensive, big-money players win the league for the club?**
- 2. If the club can't afford big-money players, is there a superior playing style they should adopt?**
- 3. Which areas are West Ham players lacking in, compared to the players in the top teams?**

## DATA SEARCH & CLEANING

I went through Kaggle and Google Dataset Search and found dozens of datasets.

After sieving through them, I decided on four (4) datasets for analysis.

Because each dataset was from a different source, I did quite a bit of cleaning up with SQL such as standardising the names of each club.

For example, one dataset has the club name has AFC Bournemouth while the other dataset had it has Bournemouth. So I use this command to clean up.

I also changed the data type for a few columns to be more accurate. For example, the market value of a player was INTEGER. So I changed it to MONEY using this command.

## ENTITY RELATIONAL DIAGRAMS

After data cleaning, I create my ERDs with my four tables. Here's what it look like:

I have: Match table, Club table, Transfer table and Player table.

Under Match table, the primary key is Match ID, followed by match stats like Number of Goals Scored, Number of Goals Conceded, Expected Points, etc.

Under Club table, there are attributes like Club Name, Position of the club at the end of the season, and total points accumulated.

Under Transfer table, there is Market Value of the player, the Club Name and Player Name.

And finally under Player table, there is the Player Name, the stats of his playing like Total Number Of Minutes Played, Average Passes, Average Number of Shots At Goal.

Here's the full details of the ERD. We can refer back to it if need be as we move forward to the SQL queries.

## SQL QUERIES / 1ST CONCERN

Alright, let's head back to the three concerns, starting with the first one.

Do expensive, big-money players win the league for the club?

But before we tackle this concern, we need to ask "What makes a player expensive?".

So using SQL, I found the median market value of all 515 players. Sadly, there is no quick SQL function to find median. We can find average, maximum, minimum with a simple command. But for median, this is the command.

So the median market value of all the players in EPL is \$10 million dollars.

Now we know that the median market value of an EPL player is \$10 million, let's find out if the top clubs consist of expensive players.

For this query, I joined the Club table with Transfer table with the common column of Club\_Name. And as there were 515 rows, I choose the top 50 rows to get only the 50 most expensive players.

And this is the results.

Exporting to EXCEL, we get this table that indicates that the most expensive players do in fact belong to the top clubs. As revealed here, 41 out of 50 of the most expensive players are with the top six clubs.

The exception being Arsenal, which had the 5th highest spending on players but ended up in 8th position.

Interestingly, West Ham United spent quite a bit too on their players. They spent more than Wolves, Sheffield United, Burnley, Southampton, Newcastle. And based on the average market value of their players West Ham should be between 7th to 10th position. But ended up in 16th position at the end of the season.

In conclusion, big-money players do help win trophies as evidenced by the top two highest-spending clubs, Liverpool and Manchester City, occupying the top two table positions.

However, spending on expensive players do NOT guarantee success as evidenced by the spending by Arsenal, Everton and West Ham. [GO BACK TWO PAGES]

## 2ND CONCERN

The second concern was:

If the clubs like West Ham can't afford big-money players, is there a superior playing style they should adopt?

For this query, I use the Match table to see the stats of each team during each game.

The green bars represent Offensive Play. And it is clear that the top clubs are much more attacking than the lower clubs. Conversely, the red bars

which represents Defensive Plays, are employed more by the lower clubs.

In terms of total shots at goal, represented by the blue bars, West Ham has the highest total number of shots. The problem is that they are not converting the shots into goals, as shown by the yellow bars. This means that West Ham strikers are hardworking but not very accurate in scoring.

All in all, West Ham needs to work on being more aggressive and offensive, as well as work on their accuracy shooting.

Just a quick note on how I export my SQL queries to EXCEL. I could NOT do the direct export to EXCEL as I'm on the Virtual Space and don't have SSMS on my Mac. So, I exported the data by providing a Source Query and then saving it on the cloud as an XLS file.

## 3RD CONCERN

The third and last concern of West Ham is:

3. Which areas are West Ham players lacking in, compared to the players in the top teams?

For this concern, I used the Player table and join it with the Transfer table to retrieve the Top 20 most expensive players and their respective stats.

It's a bit blur but here are the top 20 most expensive players players, with their individual ratings.

Here's the query converted into a bar chart in EXCEL.

For ease of comparison, I did the same query for all the West Ham players as I did in the previous query.

I used just columns such as the distance covered, number of accurate passes.....

If you merge both barcharts together, you get the top 20 most expensive players versus the West Ham average player.

The West Ham players aren't too far off with the top guys. But they can improve on these areas:

- Passing (green)
- Dribbles (dark blue)

So, if West Ham players improve on these four or five areas, the entire club can raise its standards and move up the league. In return, the higher ranking will attract more money from advertisers, broadcasters and additional fans.

## CONCLUSION

To recap, we have addressed the three concerns of West Ham United.

- 1. Do expensive, big-money players win the league for the club?**

To a large extent, yes. But it does not guarantee success.

- 2. If the club can't afford big-money players, is there a superior playing style they should adopt?**

Be more offensive and accurate in shooting and adopt a more pressing game.

- 3. Which areas are West Ham players lacking in, compared to the players in the top teams?**

West Ham players need to improve on their shots at goals, headers, accurate passing and effective dribbles.

With these changes, West Ham will significantly improve their table position of 16th to a higher table position.

Lastly, I wish to thank Zaleha for helping me out on this project. She volunteered to help me look into my tables and offered invaluable advice on how I should tackle them. So thank you Zaleha.

And thank you everyone for your time. Cheers!

**Endnote:** Today, data has allowed football clubs to quickly find new players with using advanced clustering methods. This cuts down the number of players clubs need to review, reducing the burden of classical approaches and allowing for deeper analysis of a small number of potential picks. Of course, this form of player recruitment still require lots of improvement to cover hard-to-measure statistics and include anomalies.