UEFA DATA ANALYSIS PROJECT

EXPLORING THE DYNAMICS OF EUROPEAN FOOTBALL USING SQL

PRESENTED BY MEGHANA L

ABOUT UEFA

- Founded: June 15, 1954, in Basel, Switzerland
- Scope: Administrative and controlling body for European football
- Members: 55 member associations (countries)
- Key Responsibilities:
- Organizing competitions (e.g., UEFA Champions League, UEFA Euros)
- Promoting football development and fair play

OBJECTIVES OF THE PROJECT

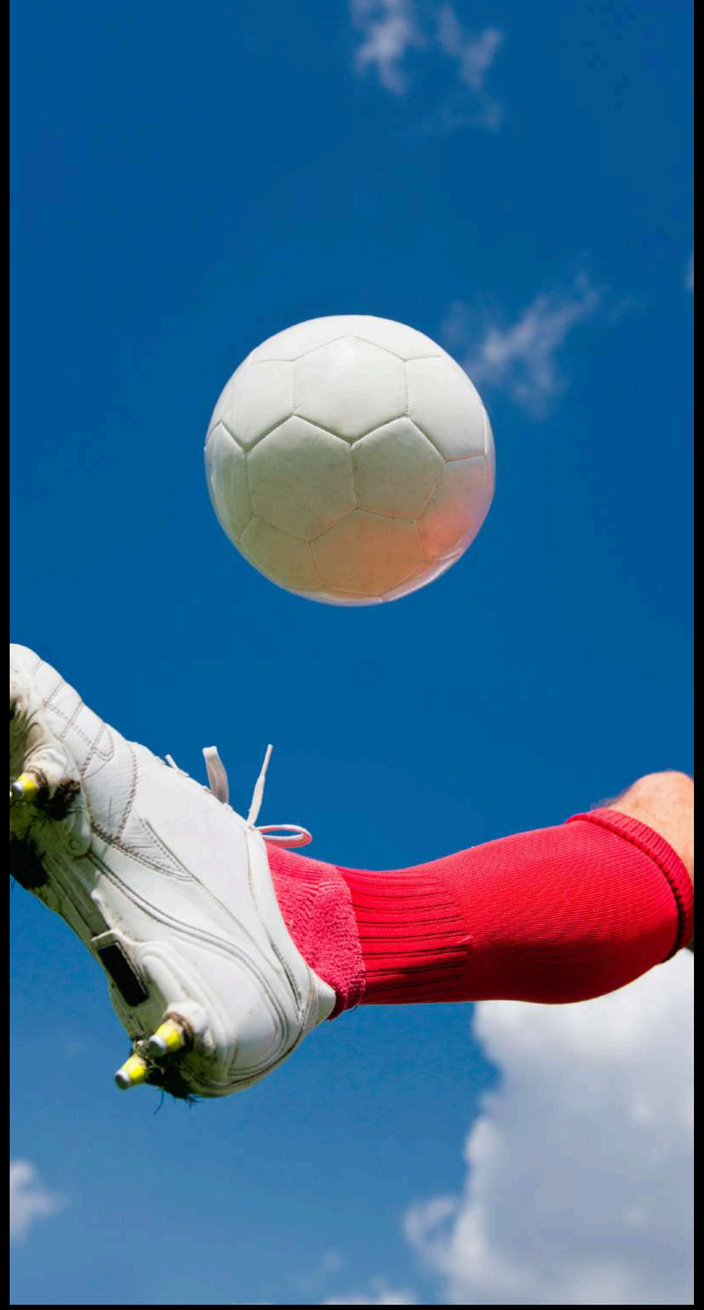
- To analyze UEFA-related datasets using SQL
- To answer football-specific questions and provide actionable insights

DATASET FILES:

- Goals.csv
- Matches.csv
- Players.csv
- Teams.csv
- Stadium.csv

KEY FOCUS AREAS:

Teams, players, matches, goals, and stadium analysis



OVERVIEW OF DATASET STRUCTURE

- **Goals.csv:** Details of goals scored
- **Matches.csv:** Match details
- **Players.csv:** Player information
- **Teams.csv:** Team data
- **Stadium.csv:** Stadium details

APPROACH TO DATA ANALYSIS

- **Tools Used:** PostgreSQL for query execution
- **Steps Taken:**

1. Loaded data into PostgreSQL
2. Formulated queries to answer specific questions

KEY INSIGHTS

1. **Total Number of Teams:**

select count(team_name) total_teams from teams;

total_teams bigint
74

2. Find the Number of Teams per Country

select count(team_name) total_teams from teams;

	country character varying (50)	number_of_teams bigint
1	Germany	8
2	England	7
3	Italy	7
4	Spain	6
5	Russia	6
6	France	6
7	Netherlands	3
8	Portugal	3
9	Turkey	3

KEY INSIGHTS

3.Calculate the Average Team Name Length

```
select round(avg(length(team_name)))
avg_team_name_length
from teams;
```

	avg_team_name_length numeric
1	12

4.Calculate the Average Stadium Capacity in Each Country round it off and sort by the total stadiums in the country.

```
select      country,round(avg(capacity),0)
avg_stadium_capacity,count(distinct(nam
e)) as total_stadiums
from stadium
group by country
order by avg_stadium_capacity desc;
```

	country character varying (50)	avg_stadium_capacity numeric	total_stadiums bigint
1	Wales	74500	1
2	Scotland	60832	1
3	England	58802	8
4	Portugal	57217	3
5	Romania	55611	1
6	Serbia	55538	1
7	Italy	54923	5

KEY INSIGHTS

5.Calculate the Total Goals Scored.

```
select count(distinct goal_id) total_goals
from goals;
```

	total_goals bigint
1	2279

6.Find the total teams that have city in their names

```
select distinct(team_name),
count(distinct(team_name))
from teams where
team_name like '%City%'
group by team_name;
```

	team_name character varying (50)	count bigint
1	Leicester City	1
2	Manchester City	1

KEY INSIGHTS

7. Use Text Functions to Concatenate the Team's Name and Country

```
select concat(team_name,' - ',country)
team_country
from teams;
```

	team_country text
1	AC Milan - Italy
2	AFC Ajax - Netherlands
3	Atalanta - Italy
4	Atlético Madrid - Spain
5	Bayern München - Germany
6	Beşiktaş - Turkey
7	Borussia Dortmund - Germany
8	BSC Young Boys - Switzerland
9	Chelsea FC - England

8. What is the highest attendance recorded in the dataset, and which match (including home and away teams, and date) does it correspond to?

```
select home_team,away_team,date,attendance from
matches
where attendance=(select max(attendance) from
matches);
```

	home_team character varying (50)	away_team character varying (50)	date character varying	attendance integer
1	FC Barcelona	Liverpool FC	01-05-2019	98299

KEY INSIGHTS

9.What is the lowest attendance recorded in the dataset, and which match (including home and away teams, and date) does it correspond to set the criteria as greater than 1 as some matches had 0 attendance because of covid.

```
select
home_team,away_team,date,attendance
from matches
where attendance=(select
min(attendance) from(select * from
matches where attendance >1 order by
attendance));
```

	home_team character varying (50)	away_team character varying (50)	date character varying	attendance integer
1	FC Midtjylland	Atalanta	21-10-2020	132
2	FC Midtjylland	AFC Ajax	03-11-2020	132

10. Identify the match with the highest total score (sum of home and away team scores) in the dataset. Include the match ID, home and away teams, and the total score.

```
select
match_id,home_team,away_team,home_team_score+a
way_team_score total_score
from matches
order by total_score desc
limit 1;
```

	match_id character varying (20)	home_team character varying (50)	away_team character varying (50)	total_score integer
1	mt688	Borussia Dortmund	Legia Warszawa	12

KEY INSIGHTS

11.Find the total goals scored by each team, distinguishing between home and away goals. Use a CASE WHEN statement to differentiate home and away goals within the subquery

```
select a.team_name,sum(case when
b.home_team=a.team_name then home_team_score else 0 end)
as home_goals,
sum(case when b.away_team=a.team_name then
away_team_score else 0 end) as away_goals
from teams a
left join matches b on a.team_name=b.home_team or
a.team_name=b.away_team
group by a.team_name;
```

	team_name character varying (50)	home_goals bigint	away_goals bigint
1	AFC Ajax	27	36
2	İstanbul Başakşehir	5	2
3	Legia Warszawa	4	5
4	FC Sheriff	3	4
5	Feyenoord	3	2
6	RB Leipzig	26	26

12. windows function - Rank teams based on their total scored goals (home and away combined) using a window function.In the stadium Old Trafford.

```
select away_team as team_name,
rank() over(order by total_goals desc) as rank
from (select
away_team,sum(home_team_score+away_team_score)
as total_goals
 from matches where stadium like 'Old Trafford'
 group by away_team);
```

	team_name character varying (50)	rank bigint
1	Paris Saint-Germain	1
2	Atalanta	2
3	İstanbul Başakşehir	2
4	RB Leipzig	2
5	Villarreal CF	5
6	FC Basel	5
7	Sevilla FC	5

KEY INSIGHTS

13. TOP 5 players who scored the most goals in Old Trafford, ensuring null values are not included in the result (especially pertinent for cases where a player might not have scored any goals).

```
select player_name, count(distinct(goal_id)) as total_goals from
(select concat(a.first_name, ' ', a.last_name) as player_name ,b.*
from players as a
 right join (select goals.goal_id, goals.match_id, goals.pid,
matches.stadium from goals left join matches on goals.match_id
= matches.match_id) as b
 on a.player_id = b.pid where b.stadium like 'Old Trafford') group
by player_name having count(distinct(goal_id)) is not null order
by total_goals desc limit 5;
```

	player_name text	total_goals bigint
1	Marcus Rashford	8
2	Romelu Lukaku	3
3	Neymar	2
4	Bruno Fernandes	2
5	Cristiano Ronaldo	2

14. Write a query to list all players along with the total number of goals they have scored. Order the results by the number of goals scored in descending order to easily identify the top 6 scorers.

```
select player_name,total_goals from(
 select p.player_id,concat(p.first_name,' ',p.last_name) as
player_name,
 count(g.goal_id) as total_goals
 from players p
 right join goals g on p.player_id=g.pid
 group by p.player_id,player_name having p.player_id is not null
 order by total_goals desc
 limit 6);
```

	player_name text	total_goals bigint
1	Robert Lewandowski	54
2	Cristiano Ronaldo	47
3	Lionel Messi	42
4	Karim Benzema	40
5	Kylian Mbappé	33
6	Mo Salah	33

KEY INSIGHTS

15. Identify the Top Scorer for Each Team - Find the player from each team who has scored the most goals in all matches combined. This question requires joining the Players, Goals, and possibly the Matches tables, and then using a subquery to aggregate goals by players and teams.

```
select player_name,team,total_goals from
 (select player_name,team,total_goals,row_number() over(partition by
team order by total_goals desc)as rank
 from
 (select player_name,team,count((goal_id)) as total_goals
 from
 (select p.player_id,concat(p.first_name,' ',p.last_name) as
player_name,p.team,g.goal_id
 from goals g
 left join players p on g.pid=p.player_id
 where p.team is not null)
 group by player_name,team)) where rank=1;
```

	player_name text	team character varying (100)	total_goals bigint
1	Andrej Kramarić	1899 Hoffenheim	5
2	Olivier Giroud	AC Milan	8
3	Steven Zuber	AEK Athen	1
4	Dušan Tadić	AFC Ajax	10

16. Find the Total Number of Goals Scored in the Latest Season - Calculate the total number of goals scored in the latest season available in the dataset. This question involves using a subquery to first identify the latest season from the Matches table, then summing the goals from the Goals table that occurred in matches from that season.

```
select count(g.goal_id) as total_number_of_goals
from goals as g
inner join matches m on g.match_id=m.match_id
where m.season=(select max(season) from matches);
```

	total_number_of_goals bigint
1	380

KEY INSIGHTS

17. Find Matches with Above Average Attendance - Retrieve a list of matches that had an attendance higher than the average attendance across all matches. This question requires a subquery to calculate the average attendance first, then use it to filter matches.

```
select *
from matches
where attendance > (select avg(attendance) from
matches);
```

	match_id character varying (20)	season character varying (25)	date character varying	home_team character varying (50)	away_team character varying (50)	stadium character varying (50)	h
1	mt1	2021-2022	15-09-2021	Manchester City	RB Leipzig	Etihad Stadium	
2	mt3	2021-2022	28-09-2021	Paris Saint-Germain	Manchester City	Parc des Princes	
3	mt6	2021-2022	19-10-2021	Paris Saint-Germain	RB Leipzig	Parc des Princes	

18.Find the Number of Matches Played Each Month - Count how many matches were played in each month across all seasons. This question requires extracting the month from the match dates and grouping the results by this value. as January Feb march

```
select to_char(date::date,'Month') as months,
count(distinct(match_id)) as number_of_matches
from matches
group by months
order by number_of_matches desc;
```

	months text	number_of_matches bigint
1	November	184
2	October	152
3	September	128
4	December	112
5	February	48
6	April	47

THANK YOU