FOOTBALL

NTI

Firstly: web_scraping

Salford & Co.

https://www.premierleague.com/en/tables?competition=8&season=2025&round=L_1&matchweek=1&ha=-1

Premier League

INTRODUCTION

Premier League

Premier League | Matches | Table | Statistics | Fantasy | News | Players | Clubs | Video | Sign

Table

Premier League | 2025/26 | Matchweek 1 | Home/Away | Reset

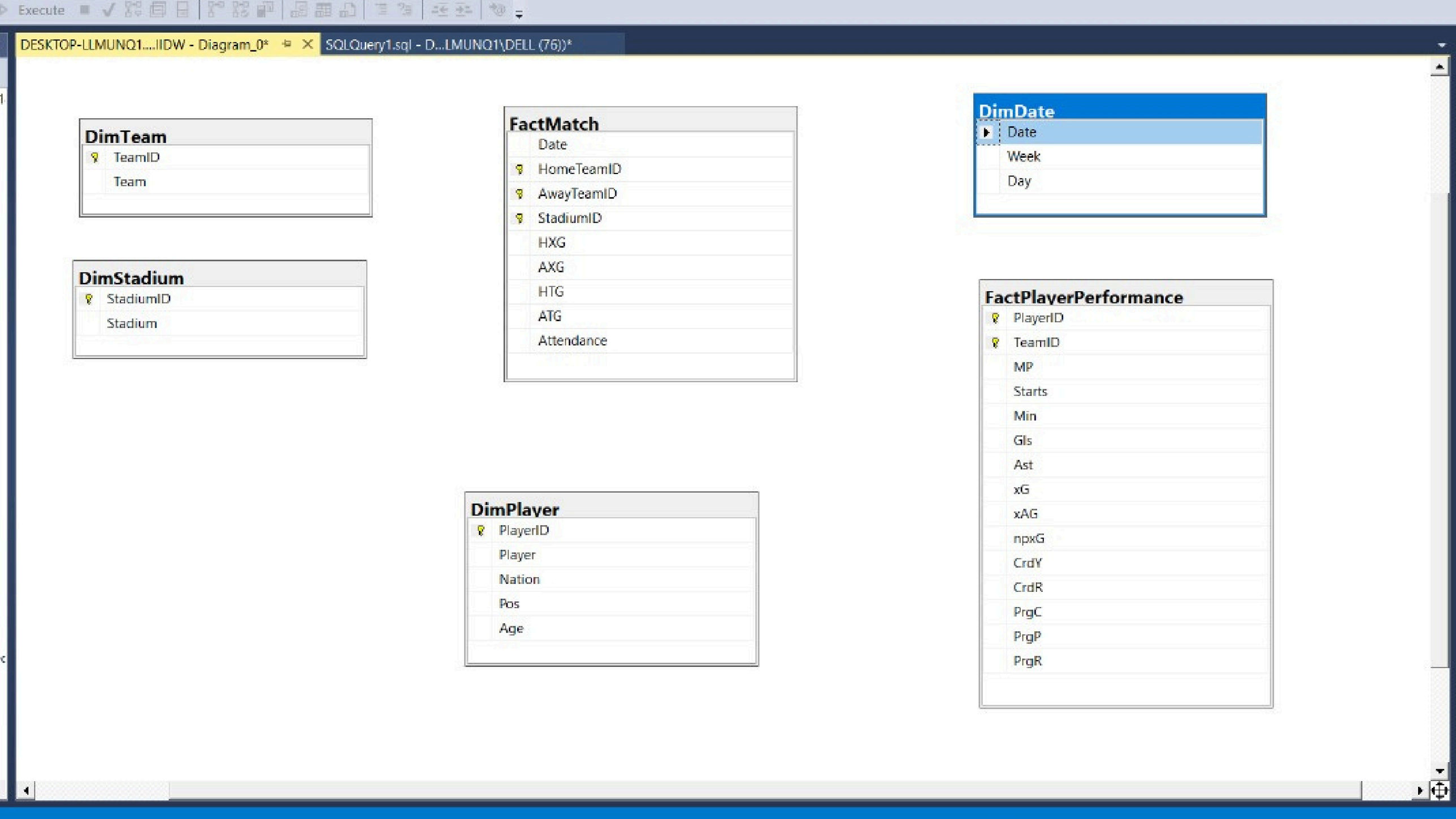
Pos	Team	Pl	W	D	L	GF	GA	GD	Pts	Next
1	Arsenal	0	0	0	0	0	0	0	0	
2	Aston Villa	0	0	0	0	0	0	0	0	
3	Bournemouth	0	0	0	0	0	0	0	0	
4	Brentford	0	0	0	0	0	0	0	0	
5	Brighton and Hove Albion	0	0	0	0	0	0	0	0	
6	Burnley	0	0	0	0	0	0	0	0	

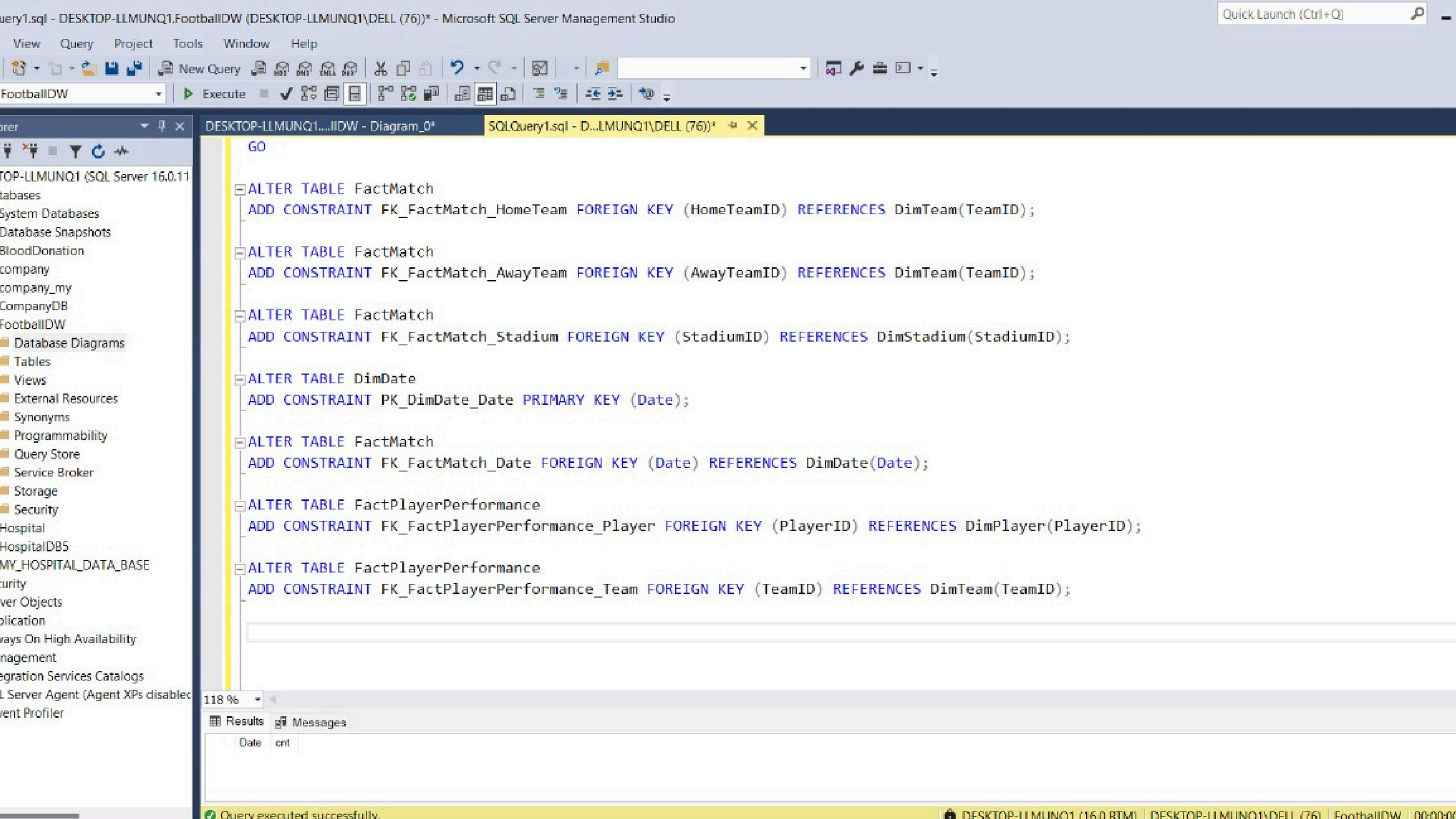
```
rt pandas as pd


atches = pd.read_csv('match_scrapp.csv')
layers = pd.read_csv('premier-player-23-24.csv')


eate DimDate
date = df_matches[['Date', 'Week', 'Day']].drop_duplicates().reset_index(drop=Tr

eate DimStadium
stadium = df_matches[['Stadium']].drop_duplicates().reset_index(drop=True)
stadium['StadiumID'] = dim_stadium.index + 1

eate DimTeam
s = pd.concat([df_matches['HomeTeam'], df_matches['AwayTeam'], df_players['Team'
team = teams.drop_duplicates().reset_index(drop=True).to_frame(name='Team')
team['TeamID'] = dim_team.index + 1

eate DimPlayer
player = df_players[['Player', 'Nation', 'Pos', 'Age']].drop_duplicates().reset_
player['PlayerID'] = dim_player.index + 1

eate FactMatch
_match = df_matches.copy()
_match = fact_match.merge(dim_team, left_on='HomeTeam', right_on='Team', how='le
_match = fact_match.rename(columns={'TeamID': 'HomeTeamID'}).drop(columns='Team'
_match = fact_match.merge(dim_team, left_on='AwayTeam', right_on='Team', how='le
_match = fact_match.rename(columns={'TeamID': 'AwayTeamID'}).drop(columns='Team'
_match = fact_match.merge(dim_stadium, on='Stadium', how='left')
_match = fact_match[['Date', 'HomeTeamID', 'AwayTeamID', 'StadiumID', 'HXG', 'AX


eate FactPlayer
_player = df_players.copy()
_player = fact_player.merge(dim_player, on=['Player', 'Nation', 'Pos', 'Age'], h
_player = fact_player.merge(dim_team, on='Team', how='left')
_player = fact_player[['PlayerID', 'TeamID', 'MP', 'Starts', 'Min','Gls', 'Ast',
```

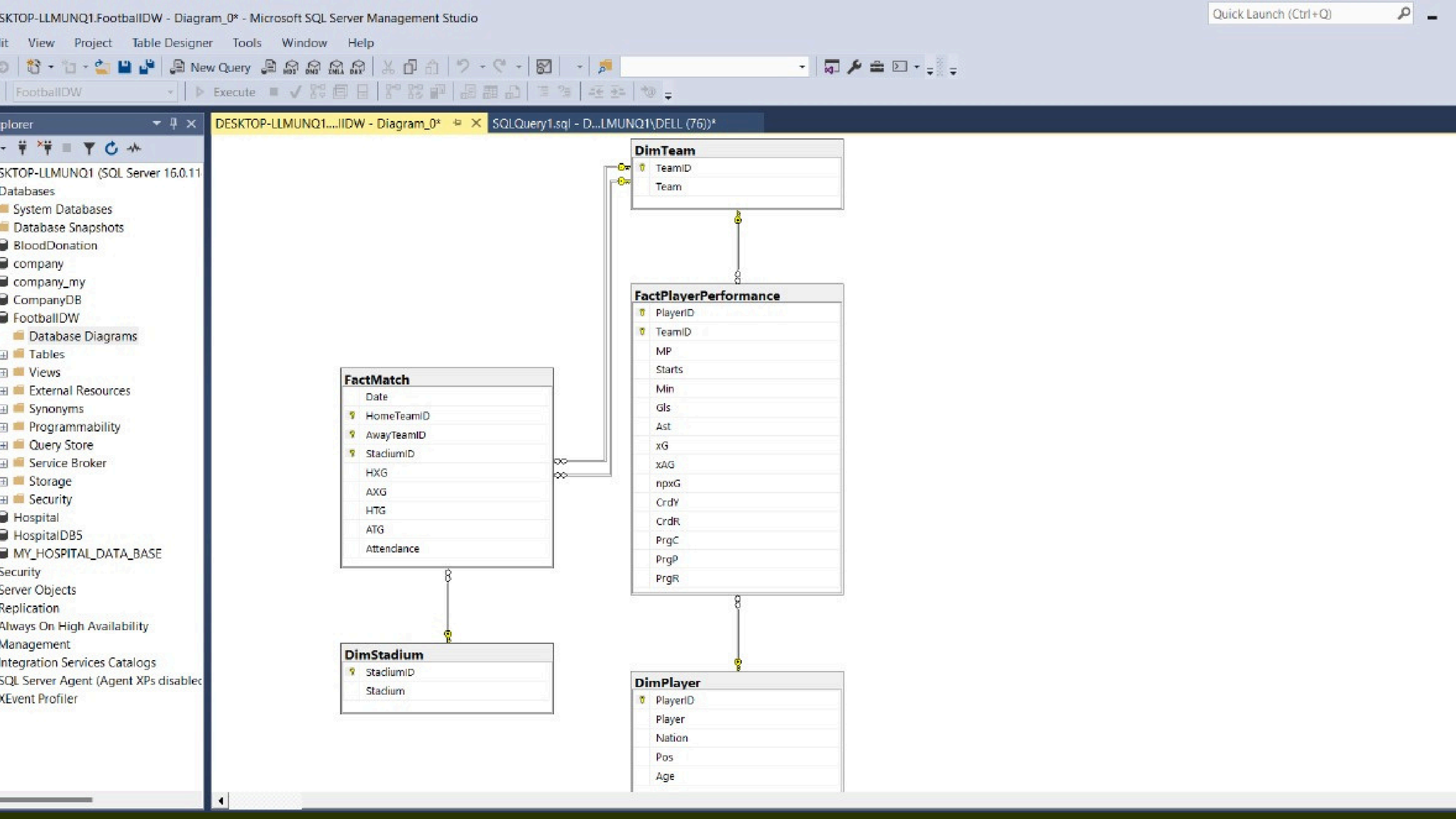
```
from selenium import webdriver
from selenium.webdriver.chrome.service import Service
from bs4 import BeautifulSoup
import pandas as pd
import time

    if len(cols) > 8:
        date = cols[0].text.strip()
        time_ = cols[1].text.strip()
        home_team = cols[2].text.strip()
        score = cols[3].text.strip()
        away_team = cols[4].text.strip()
        attendance = cols[5].text.strip()
        xg = cols[6].text.strip()
        wk = cols[7].text.strip()

        matches.append({
            'Date': date,
            'Time': time_,
            'HomeTeam': home_team,
            'AwayTeam': away_team,
            'Score': score,
            'Attendance': attendance,
            'xG': xg,
```

DATA STORAGE AND MODELING

SQL Server


Execute
DESKTOP-LLMUNQ1....IIDW - Diagram_0*
SQLQuery1.sql - D...LMUNQ1\DELL (76))*
DimTeam
TeamID
Team
DimStadium
StadiumID
Stadium
FactMatch
Date
HomeTeamID
AwayTeamID
StadiumID
HXG
AXG
HTG
ATG
Attendance
DimDate
Date
Week
Day
FactPlayerPerformance
PlayerID
TeamID
MP
Starts
Min
Gls
Ast
xG
xAG
npxG
CrdY
CrdR
PrgC
PrgP
PrgR
DimPlayer
PlayerID
Player
Nation
Pos
Age


uery1.sql - DESKTOP-LLMUNQ1.FootballDW (DESKTOP-LLMUNQ1\DELL (76))* - Microsoft SQL Server Management Studio
Quick Launch (Ctrl+Q)
View Query Project Tools Window Help
New Query
FootballDW
Execute
TOP-LLMUNQ1 (SQL Server 16.0.11
tabases
System Databases
Database Snapshots
BloodDonation
company
company_my
CompanyDB
FootballDW
Database Diagrams
Tables
Views
External Resources
Synonyms
Programmability
Query Store
Service Broker
Storage
Security
Hospital
HospitalDB5
MY_HOSPITAL_DATA_BASE
curity
ver Objects
plication
ways On High Availability
nagement
egration Services Catalogs
L Server Agent (Agent XPs disablec
vent Profiler
DESKTOP-LLMUNQ1....IIDW - Diagram_0*
SQLQuery1.sql - D...LMUNQ1\DELL (76))*
GO
ALTER TABLE FactMatch
ADD CONSTRAINT FK_FactMatch_HomeTeam FOREIGN KEY (HomeTeamID) REFERENCES DimTeam(TeamID);
ALTER TABLE FactMatch
ADD CONSTRAINT FK_FactMatch_AwayTeam FOREIGN KEY (AwayTeamID) REFERENCES DimTeam(TeamID);
ALTER TABLE FactMatch
ADD CONSTRAINT FK_FactMatch_Stadium FOREIGN KEY (StadiumID) REFERENCES DimStadium(StadiumID);
ALTER TABLE DimDate
ADD CONSTRAINT PK_DimDate_Date PRIMARY KEY (Date);
ALTER TABLE FactMatch
ADD CONSTRAINT FK_FactMatch_Date FOREIGN KEY (Date) REFERENCES DimDate(Date);
ALTER TABLE FactPlayerPerformance
ADD CONSTRAINT FK_FactPlayerPerformance_Player FOREIGN KEY (PlayerID) REFERENCES DimPlayer(PlayerID);
ALTER TABLE FactPlayerPerformance
ADD CONSTRAINT FK_FactPlayerPerformance_Team FOREIGN KEY (TeamID) REFERENCES DimTeam(TeamID);
118 %
Results
Messages
Date
cnt
Query executed successfully.
DESKTOP-LLMUNQ1 (16.0 RTM)
DESKTOP-LLMUNQ1\DELL (76)
FootballDW


SKTOP-LLMUNQ1.FootballDW - Diagram_0* - Microsoft SQL Server Management Studio
Quick Launch (Ctrl+Q)
it View Project Table Designer Tools Window Help
New Query
FootballDW
Execute
plorer
SKTOP-LLMUNQ1 (SQL Server 16.0.11
Databases
System Databases
Database Snapshots
BloodDonation
company
company_my
CompanyDB
FootballDW
Database Diagrams
Tables
Views
External Resources
Synonyms
Programmability
Query Store
Service Broker
Storage
Security
Hospital
HospitalDB5
MY_HOSPITAL_DATA_BASE
Security
Server Objects
Replication
Always On High Availability
Management
Integration Services Catalogs
SQL Server Agent (Agent XPs disablec
XEvent Profiler
DESKTOP-LLMUNQ1....IIDW - Diagram_0*
SQLQuery1.sql - D...LMUNQ1\DELL (76))*
DimTeam
TeamID
Team
FactPlayerPerformance
PlayerID
TeamID
MP
Starts
Min
Gls
Ast
xG
xAG
npxG
CrdY
CrdR
PrgC
PrgP
PrgR
FactMatch
Date
HomeTeamID
AwayTeamID
StadiumID
HXG
AXG
HTG
ATG
Attendance
DimStadium
StadiumID
Stadium
DimPlayer
PlayerID
Player
Nation
Pos
Age

OVERVIEW

```sql
SELECT TOP 10
    Stadium , SUM(Attendance) AS TotalAttendance
FROM
    DimStadium AS DS
JOIN
    FactMatch AS FM ON FM.StadiumID = DS.StadiumID
GROUP BY
    Stadium
ORDER BY
    TotalAttendance DESC;
```

	Stadium	TotalAttendanc
1	Old Trafford	1397127
2	London Stadium	1188782
3	Tottenham Hotspur Stadium	1168163
4	Emirates Stadium	1144488
5	Anfield	1063606
6	Etihad Stadium	1007227
7	St James' Park	990369
8	Villa Park	795310
9	Stamford Bridge	750953
10	Goodison Park	741793

	Player	TotalGoals
1	Erling Haaland	27
2	Cole Palmer	22
3	Alexander Isak	21
4	Dominic Solanke	19
5	Ollie Watkins	19
6	Phil Foden	19
7	Mohamed Salah	18
8	Son Heung-min	17
9	Bukayo Saka	16
10	Jarrod Bowen	16

```sql
SELECT TOP 10
    Player , SUM(Gls) AS TotalGoals
FROM
    DimPlayer AS DP
JOIN
    FactPlayerPerformance AS FPP ON FPP.PlayerID = DP.PlayerID
GROUP BY
    Player
ORDER BY
    TotalGoals DESC;
```

	Player	TotalAssists
1	Ollie Watkins	13
2	Cole Palmer	11
3	Anthony Gordon	10
4	Brennan Johnson	10
5	Kevin De Bruyne	10
6	Kieran Trippier	10
7	Martin Ødegaard	10
8	Mohamed Salah	10
9	Morgan Gibbs-White	10

	TotalAssists
1	859

Premier League

Premier League Overview

Premier League

Premier League

Pos

- DF
- DF,FW
- DF,MF
- FW
- FW,DF

Week

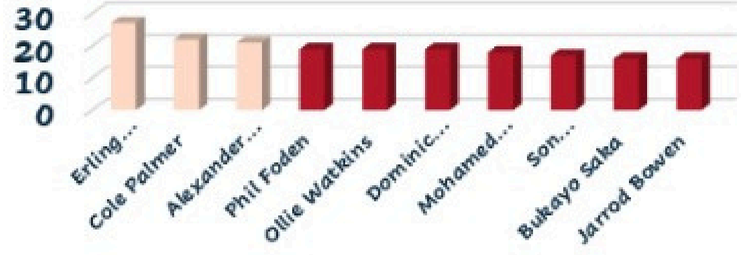
- 1
- 2
- 3
- 4

Team

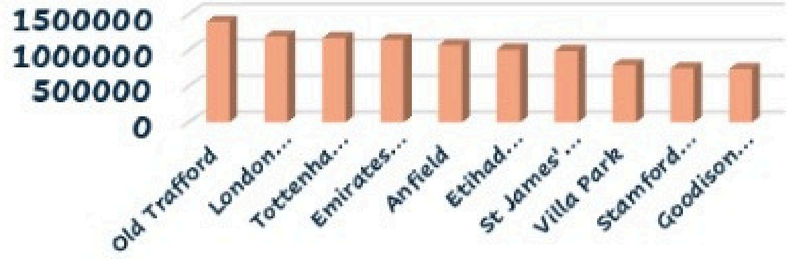
- Arsenal
- Aston Villa
- Bournemouth
- Brentford
- Brighton

Total Goals	Total Assists	Total Attendance	non-Penalties XG
1,197	859	14,673,059.00	1,119.20

Top 10 Scorers



Top 10 Stadiums By Attendance



Total Goals By Weeks

Top 10 Assisters

INSIGHTS:

- - Very small margin between #goals & #assists implies that goals rarely come from a single effort/player.

- - Goals per weeks stays between 20-40 further emphasizing the expected/repetitive nature of the league.

- - Old Trafford ranks 1st in attendance which implies that name recognition matters more than winning or playing decent games.

- - About 50% of players only come from English which imply a diverse set of players.

Premier League

SQLQuery_Nti_final_project.sql - DESKTOP-LLMUNQ1.FootballDW (DESKTOP-LLMUNQ1\DELL (59)) - Microsoft SQL Server Management Studio

Quick Launch (Ctrl+Q)

Edit View Query Project Tools Window Help

New Query | Execute | FootballDW

SQLQuery_Nti_fina...LMUNQ1\DELL (59))

```sql
SELECT
    t.Team,
    SUM(CASE WHEN t.TeamID = m.HomeTeamID THEN m.HTG ELSE 0 END +
        CASE WHEN t.TeamID = m.AwayTeamID THEN m.ATG ELSE 0 END) AS TotalGoals
FROM
    DimTeam t
JOIN FactMatch m
    ON t.TeamID = m.HomeTeamID OR t.TeamID = m.AwayTeamID
GROUP BY
    t.Team
ORDER BY
    TotalGoals DESC;
    ------------------------------
```

118 %

Results | Messages

	Team	TotalGoals
1	Manchester City	96
2	Arsenal	91
3	Liverpool	86
4	Newcastle Utd	85
5	Chelsea	77
6	Aston Villa	76
7	Tottenham	74
8	West Ham	60
9	Manchester Utd	57
10	Crystal Palace	57
11	Brentford	56
12	Brighton	55
13	Fulham	55
14	Bournemouth	54
15	Luton Town	52
16	Wolves	50
17	Nott'ham Forest	49
18	Burnley	41
19	Everton	40

Query executed successfully. DESKTOP-LLMUNQ1 (16.0 RTM) | DESKTOP-LLMUNQ1\DELL (59) | FootballDW | 00:00:00

SQLQuery_Nti_final_project.sql - DESKTOP-LLMUNQ1.FootballDW (DESKTOP-LLMUNQ1\DELL (59)) - Microsoft SQL Server Management Studio

Quick Launch (Ctrl+Q)

Edit View Query Project Tools Window Help

New Query | FootballDW | Execute

SQLQuery_Nti_fina...LMUNQ1\DELL (59))

```sql
SELECT
    t.Team,
    SUM(p.CrdY) AS TotalYellowCards,
    SUM(p.CrdR) AS TotalRedCards
FROM
    FactPlayerPerformance p
JOIN
    DimTeam t ON p.TeamID = t.TeamID
GROUP BY
    t.Team
ORDER BY
    TotalRedCards DESC, TotalYellowCards DESC;

-----------------------------------------------
```

118 %

Results | Messages

	Team	TotalYellowCards	TotalRedCards
1	Burnley	77	7
2	Sheffield United	101	5
3	Liverpool	69	5
4	Chelsea	108	4
5	Wolverhampton	102	4
6	Tottenham Hotspur	92	4
7	Fulham	83	4
8	Brighton	90	3
9	West Ham United	82	3
10	Nottingham Forest	82	3
11	Bournemouth	79	3
12	Aston Villa	94	2
13	Brentford	89	2
14	Crystal Palace	73	2
15	Arsenal	64	2
16	Manchester City	53	2
17	Manchester United	83	1
18	Everton	82	1
19	Newcastle United	79	1

Query executed successfully.

DESKTOP-LLMUNQ1 (16.0 RTM) | DESKTOP-LLMUNQ1\DELL (59) | FootballDW | 00:00:0

uery_Nti_final_project.sql - DESKTOP-LLMUNQ1.FootballDW (DESKTOP-LLMUNQ1\DELL (59)) - Microsoft SQL Server Management Studio

Quick Launch (Ctrl+Q)

View Query Project Tools Window Help

New Query | Execute | FootballDW

uery_Nti_fina...LMUNQ1\DELL (59))

```
-----------------------------------------------------------------------

SELECT
    SUM(CrdR) AS TotalRedCards
FROM
    FactPlayerPerformance;

SELECT
    SUM(CrdY) AS TotalYellowCards
FROM
    FactPlayerPerformance;
-----------------------------------------------------------------------
```

esults | Messages

TotalRedCards
58

TotalYellowCards
1653

uery executed successfully.

DESKTOP-LLMUNQ1 (16.0 RTM) | DESKTOP-LLMUNQ1\DELL (59) | FootballDW | 00:00:00

Premier League Report
(Team Analysis)

2.08
Average of xG

1653
Yellow Cards

58
Red Cards

Team Analysis

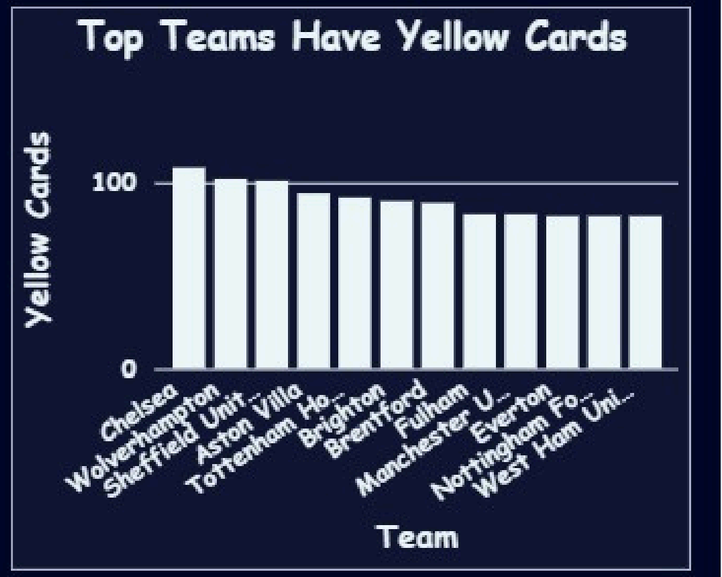
Stadium Analysis

Player Analysis

Age Group
All

Player
All

Nation
All



INSIGHTS:

- Man City scored the most goals and had the highest average xG per game.
- Liverpool had the highest total xG and led in progressive passes.
- Everton wasted the most goals, showing poor finishing.
- Chelsea received the most yellow cards.
- Top attacking teams: Liverpool, Tottenham, Arsenal (high progressive passes).
- Aggressive teams: Chelsea and Wolves (most yellow cards).

Premier League

```sql
-- Top Stadiums Have XG
SELECT
    Stadium , ROUND(SUM(HXG + AXG) , 1) AS total_xg
FROM
    DimStadium AS DS
JOIN
    FactMatch AS FM ON FM.StadiumID = DS.StadiumID
GROUP BY
    Stadium
ORDER BY
    total_xg DESC;
-- Top Teams Have AVG HTG
SELECT
    Team , SUM(HTG) AS home_team_goals
FROM
    DimTeam AS DT
JOIN
    FactMatch AS FM ON FM.HomeTeamID = DT.TeamID
GROUP BY
    Team
ORDER BY
    home_team_goals DESC;
```

```sql
-- Total Foreign Players
SELECT COUNT(*) AS total_foreign_players FROM DimPlayer
WHERE Nation <> 'ENG';


-- Total Players Under 20
SELECT COUNT(*) AS total_players_under20 FROM DimPlayer
WHERE Age < 20;


-- Total Players Above 30
SELECT COUNT(*) AS total_players_above30 FROM DimPlayer
WHERE Age > 30;


-- Total Progressive Runs by Players
SELECT SUM(PrgR) AS total_progressive_runs FROM FactPlayerPerformance;
-- Total HTG
SELECT SUM(HTG) AS home_team_goals FROM FactMatch;

-- Total ATG
SELECT SUM(ATG) AS away_team_goals FROM FactMatch;

-- Total Attendance
SELECT SUM(Attendance) AS total_attendance FROM FactMatch;

-- Total HTG
SELECT ROUND(SUM(HXG + AXG) , 1) AS total_xg FROM FactMatch
```

Premier League Report

(Player Analysis)

377	74	71	50.06
Total Foreign Players	Players Under20	Players Above30	Progressive Runs

Team Analysis

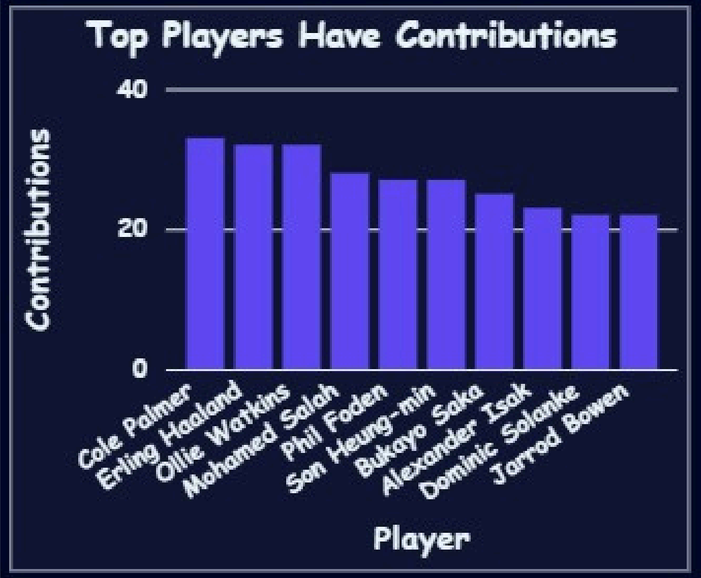
Stadium Analysis

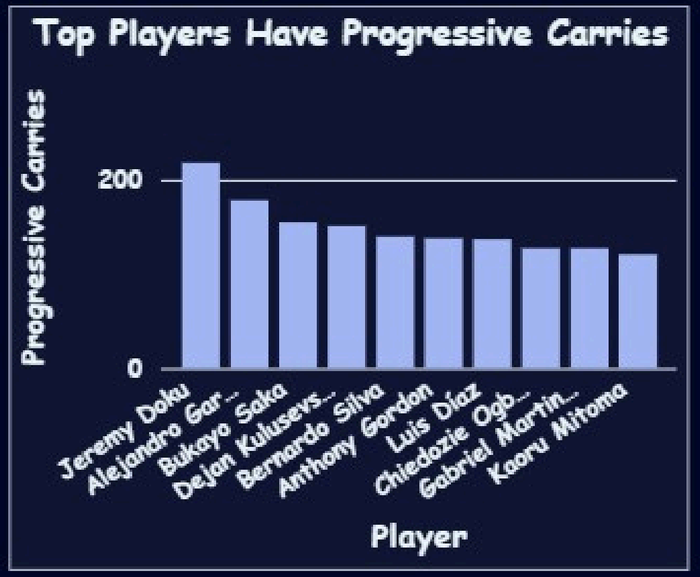
Player Analysis

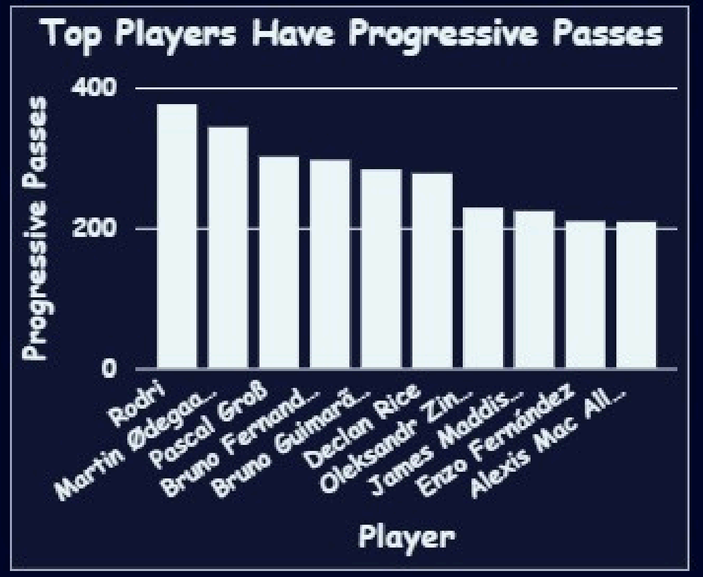
Age Group: All

Team: All

Pos: All







INSIGHTS:

- Anfield (Liverpool) had the highest expected goals, showing strong home performance.
- Old Trafford (Manchester United) had the highest average attendance, indicating strong fan support.
- Peak attendance happened mid-season.
- Manchester City had the highest home scoring rate, followed by Liverpool and Arsenal.
- Despite high attendance, Manchester United had one of the lowest home scoring rates.
- Teams scored more at home than away, showing the advantage of home field and fans.
- Fan attendance remained high all season, reflecting the Premier League's strength and competitiveness.

Premier League

Premier League Report
(Stadium Analysis)

684	562	15M	1.18K
Home Team Goals	Away Team Goals	Total Attendance	Total XG

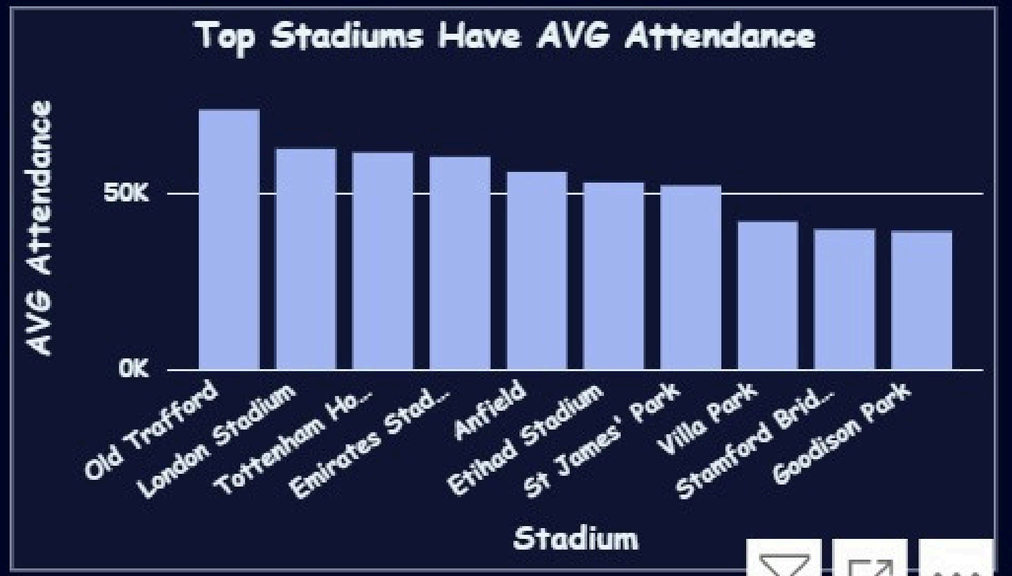
Team Analysis

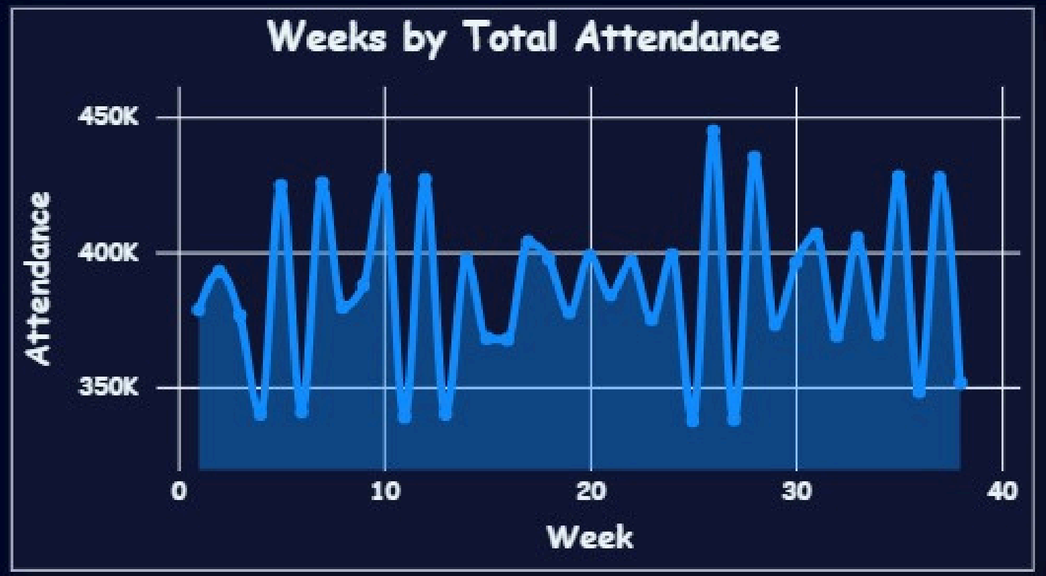
Stadium Analysis

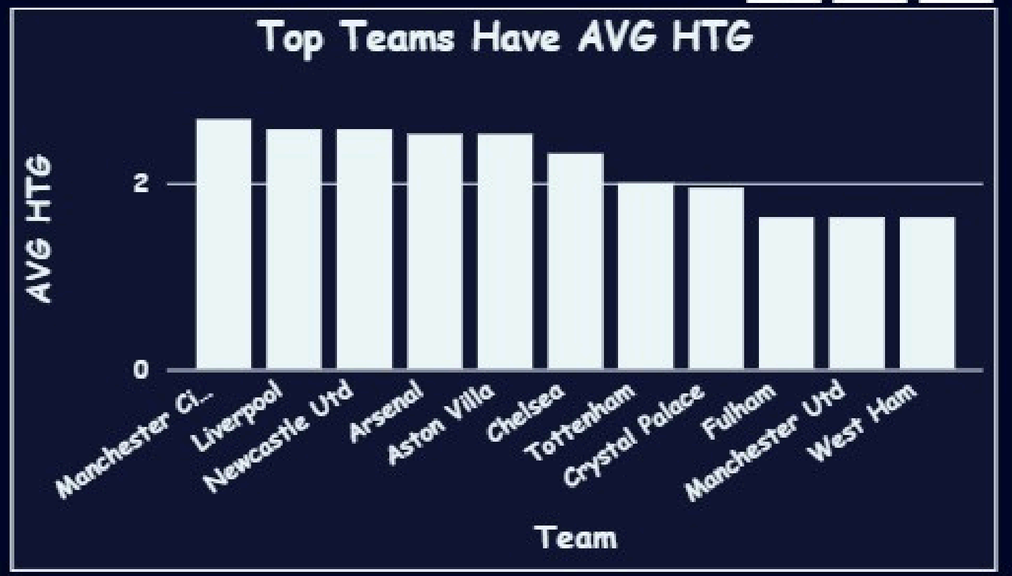
Player Analysis

Week
All

Team
All







INSIGHTS:

- Cole Palmer and Mohamed Salah were key contributors in both goals and assists.
- Doku led in progressive carries, showcasing his dribbling and movement.
- Rodri led in progressive passes, highlighting his line-breaking passing ability.
- Pascal Groß was the top defensive progressive passer, contributing to attack from the back.
- 3 of the top 10 big chance missers were Liverpool players, affecting their title hopes and leading to a 3rd-place finish.
- Phil Foden scored 8 goals over expected, playing a crucial role in Manchester City’s title win.
- Haaland had the highest non-penalty expected goals (xG), proving his elite scoring instinct.
- Premier League clubs heavily depend on foreign players, with Brazilians being the most common.

Premier League

OUR INTERACTIVE_DASH_BOARD

World Tournament

Continental Championship

Domestic Leagues

https://app.powerbi.com/groups/me/reports/a5c8a14c-047f-4c39-87b1-bee8e1406804/41d097455ffed5bee28c?experience=power-bi

Salford & Co.

THANK YOU

Final Thoughts

Thank you for your time! Football is more than just a game; it's a global phenomenon that unites people across cultures. Enjoy the game and keep the passion alive!