

Group 9 - Manchester City

Cassandra Albright
Charlie Mixson
Divya Mathew
JT Warren

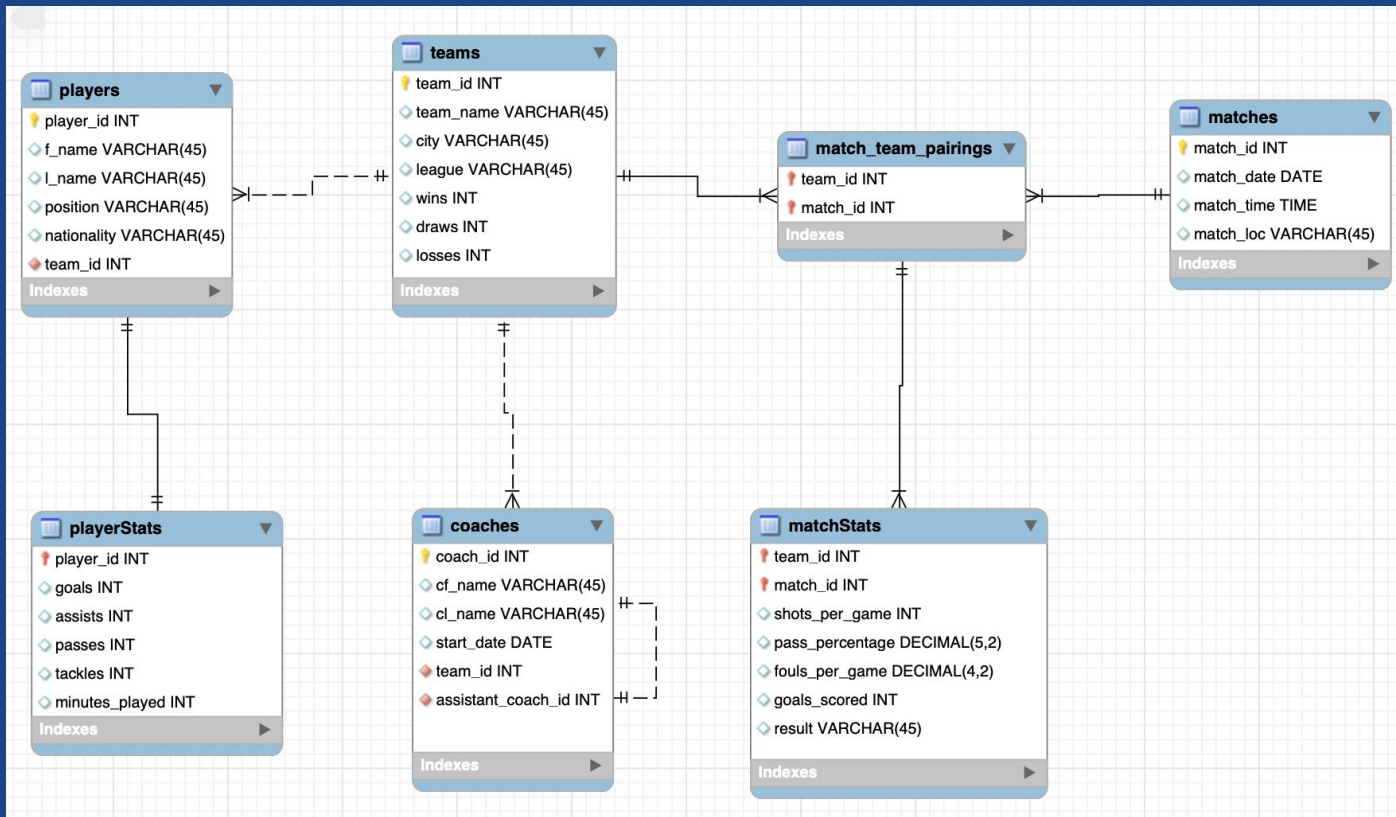




Scenario Description

- We are building our model for Manchester City to scout future potential opponents in the Champions League
- We included real teams, players, and coaches from the Premier League, Bundesliga, La Liga, and the MLS
- Our tables give Man City insight on the performance of other teams
 - Can track goals scored, wins, matches against common opponent, etc.
- Advanced statistics and data modeling has taken over the professional sports industry, and this a tiny example of what they do on a weekly basis

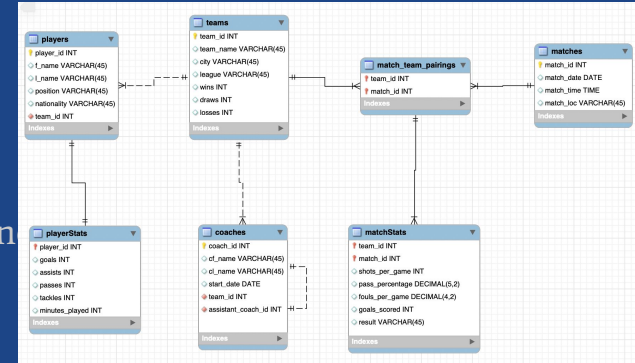
Data Model





Relationships

- One-to-One: We have 1 one-to-one relationship.
 - players to playerStats: Each player's overall statistics in their lifetime.
- One-to-Many: We have 3 one-to-many relationship.
 - teams to players: Each team has multiple players.
 - teams to coaches: Each team has multiple coaches (e.g. head and assistant).
 - match_team_pairings to matchStats: Each match pairing has statistics for both teams for all of their matches.
- Many-to-Many: We have 1 many-to-many relationship.
 - teams to matches: Each team has many matches and each match has two teams. The match_team_pairings entity is a weak entity that connects the other two.
- Recursive: We have 1 recursive relationship.
 - coach_id to assistant_id, in coaches entity: A coach can have an assistant coach.



Data Dictionary



Table: Teams

Column Name	Data Type	Description
team_id	INT	Unique identifier for each team
team_name	VARCHAR(45)	Name of the team
city	VARCHAR(45)	City where the team is based
league	VARCHAR(45)	League the team participates in
wins	INT	Number of wins
draws	INT	Number of draws
losses	INT	Number of losses

players Table

Column Name	Data Type	Description
player_id	INT	Unique identifier for each player
name	VARCHAR(45)	Player's full name
position	VARCHAR(45)	Player's primary position (e.g., forward, defender)
nationality	VARCHAR(45)	Player's country of origin
team_id	INT	Foreign key referencing the teams table

matches Table

Column Name	Data Type	Description
match_id	INT	Unique identifier for each match
match_date	DATE	Date of the match
match_time	TIME	Time of the match
match_loc	VARCHAR(45)	Location of the match

Data Dictionary



Table: MatchStats

Column	Data Type	Description
match_id	INT	Foreign key referencing the match
shots_per_game	INT	Average number of shots per game for the team
pass_percentage	DECIMAL(5,2)	Pass completion percentage for the team
fouls_per_game	DECIMAL(4,2)	Average number of fouls per game for the team
goals_scored	INT	Total goals scored by the team in the match
result	VARCHAR(45)	Result of the match (e.g., win, loss, draw)

playerStats Table

Column Name	Data Type	Description
player_id	INT	Foreign key referencing the <code>players</code> table
goals	INT	Number of goals scored
assists	INT	Number of assists provided
passes	INT	Total number of passes
tackles	INT	Number of tackles made
minutes_played	INT	Total minutes played

coaches Table

Column Name	Data Type	Description
coach_id	INT	Unique identifier for each coach
coach_name	VARCHAR(45)	Coach's name
start_date	DATE	Date the coach started with the team
team_id	INT	Foreign key referencing the <code>teams</code> table
assistant_coach_id	INT	Foreign key referencing another coach (if applicable)

match_team_pairings Table

Column Name	Data Type	Description
team_id	INT	Unique identifier for the team that participated in the match.
match_id	INT	Unique identifier for the match in which the team participated.



Complex Query 1

- This query shows matches where a team scored more goals than their average and what was the matches result, their total shots and the pass percentage
- Could be used to understand specific contexts or conditions that might have led to the higher performance by the opponent or by themselves
- For example, if there were many shots taken then the team could implement more aggressive tactics into their gameplans

match_id	team_name	result	shots_per_ga...	pass_percenta...
1	Manchester United	Win	15	85.30
16	Real Madrid	Win	14	85.80
21	Real Madrid	Win	16	88.20
12	Liverpool	Win	17	86.30
17	Barcelona	Win	17	89.50
13	Chelsea	Draw	14	82.50
18	Atletico Madrid	Win	15	83.60
6	RB Leipzig	Win	18	89.80
26	RB Leipzig	Win	16	86.50

```
SELECT matches.match_id, teams.team_name, matchStats.result,
matchStats.shots_per_game, matchStats.pass_percentage
FROM matches
JOIN matchStats ON matches.match_id = matchStats.match_id
JOIN teams ON teams.team_id = matchStats.team_id
WHERE matchStats.goals_scored > (
    SELECT AVG(matchStats.goals_scored)
    FROM matchStats
    WHERE teams.team_id = matchStats.team_id
);
```



Complex Query 2

- Shows which mentee has a better record than the mentor
- Could be used to determine which coach to hire

cf_name	cl_name	wins	cf_name	cl_name	wins
Carlo	Ancelotti	28	Greg	Vanney	18
Erik	ten Hag	25	Jürgen	Klopp	23
Edin	Terzić	26	Ronny	Deila	17
Carlo	Ancelotti	28	Graham	Potter	22
Carlo	Ancelotti	28	Diego	Simeone	Potter
Graham	Potter	22	Marco	Rose	21

Result 2 x

Output

#	Time	Action	Message
1	12:47:53	select f_name, l_name, team_name, tackles from players join playerStats on players.pl...	Error Code: 1046. No data
2	12:47:59	use ha_cjm75885	0 row(s) affected
3	12:48:03	select f_name, l_name, team_name, tackles from players join playerStats on players.pl...	Error Code: 1054. Unknow
4	12:48:12	select f_name, l_name, team_name, tackles from players join playerStats on players.pl...	10 row(s) returned
5	12:53:44	select mentee.cf_name, mentee.cl_name, mentee_team.wins, mentor.cf_name, ment...	12 row(s) returned

```
-- Show every coach that has a more wins than his mentor
```

```
select mentee.cf_name, mentee.cl_name, mentee_team.wins, mentor.cf_name, mentor.cl_name, mentor_team.wins
from coaches as mentee
join coaches as mentor on mentee.coach_id = mentor.assistant_coach_id
join teams as mentee_team on mentee.team_id = mentee_team.team_id
join teams as mentor_team on mentor.team_id = mentor_team.team_id
where mentee_team.wins > mentor_team.wins;
```


Complex Query 3

- This query is used for finding each club's top scorer. It lists their name, team, and their number of goals.
- This is helpful for us to know as we would be better able to plan defensive strategies.
- For example, assigning our best defenders to mark the top scorer closely, reducing their impact on the game.

f_name	l_name	team_name	goals
Erling	Haaland	Borussia Dortmund	40
Lionel	Messi	Barcelona	30
Karim	Benzema	Real Madrid	25
Mohamed	Salah	Liverpool	22
Javier	Hernandez	LA Galaxy	12
Bruno	Fernandes	Manchester United	10
João	Felix	Atletico Madrid	8
Mason	Mount	Chelsea	7
Maxi	Moralez	New York City FC	5



```
select players.f_name, players.l_name, teams.team_name, playerStats.goals
from players
join playerStats on players.player_id = playerStats.player_id
join teams on teams.team_id = players.team_id
WHERE
    playerStats.goals = (
        SELECT MAX(playerStats.goals)
        FROM playerStats
        JOIN players ON playerStats.player_id = players.player_id
        WHERE teams.team_id = players.team_id)
order by playerStats.goals desc;
```

[illegible]

Thank You