WONDER WOMEN

TEAM BLUE

Code First Girls
Intro to Data MOOC Challenge
August 2023

Hania Rehman
Sharon Wanjiru
Pamela Seale
Chiedza Nhuvira

TEAM BLUE

“You cannot win a game of football on your own; it's about the entire squad working together to achieve something. That's how football works: it has always been about the group, not the individual.”

- Lucy Bronze

Quest Brief: Wonder Women

You've been assigned to develop a database system to manage information about the UK Women's Football Leagues.
You will need to track teams, players, matches and league standings. Wonder Women require the following:

1. Team Management: Store the name, coach and home stadium
2. Player Management: Store their name, position, number and team
3. Match Management: Store the teams involved, match date, location and final score

Meeting the Brief:

We created a database system to manage information about an imaginary International Women's Football League.
We included the following information:

1. Team Management: the team name, it's coach, home stadium and league.
2. Player Management: their name, position, number and team.
3. Player Performance Statistics: goals, assists, yellow & red cards for each match.
4. Matches: Store the teams involved, match date & day of the week, location, attendance and final score.
5. League Standings: wins, losses and draws across six matches.
6. League Information: the season, it's official title, and it's start and end dates.
7. Stadium Information: each stadium's name, it's location (country), year opened, surface type and capacity.

Our Approach:

This year's FIFA Women's World Cup gripped fans and onlookers across the world, drawing record-breaking audiences both at home and in stadiums. We looked at real data that was available for the tournament to find inspiration for our own (imaginary) Women's Football League.

We created data on:

- sixteen players
- across four teams,
- all playing each other
- in a total of twelve matches.

We then did some research to find out what kind of data a league like ours might generate.

Attendance and viewership statistics have been prominently reported for the 2023 World Cup matches as women's football gains in popularity.

A wealth of information is also available on **performance data** for teams and players including:

- goals scored & conceded
- attempts on goal
- assists
- yellow & red cards

Most notably, the 2023 World Cup Golden Boot went to **Miyazawa Hinata**, who scored five goals despite her team (Japan) going out in the quarter finals.

We used aggregate functions and queried our database for both attendance and performance data.

Our Code: DDL & DML

Deliverable:
You are expected to highlight some of the most crucial commands and talk through them in the presentation. Do not include your WHOLE code in the presentation but add bits that showcase your understanding of **DDL & DML commands** and how to properly **assign keys** and use **aggregate functions.**

```
CREATE DATABASE WonderWomenFootball;

USE WonderWomenFootball;
```

DATA DEFINITION LANGUAGE

CREATE
USE
ALTER
DROP

```
CREATE TABLE team (
team_id INT AUTO_INCREMENT PRIMARY KEY,
t_name VARCHAR(50) NOT NULL,
coach VARCHAR(50),
home_stadium INT,
league_id INT,
FOREIGN KEY (home_stadium) REFERENCES stadium (stadium_id),
FOREIGN KEY (league_id) REFERENCES league(league_id)
);
```

```
ALTER TABLE Matches
ADD COLUMN day_of_week VARCHAR (20) AFTER match_date;
UPDATE matches
SET day_of_week = dayname(match_date);
SELECT * FROM matches;
```

SELECT
UPDATE
INSERT
DELETE

DATA MANIPULATION LANGUAGE

```
INSERT INTO stadium
VALUES
(101,'Rose Bowl', 'USA', '1922', 'Grass', 92542),
(201,'Olympiastadion', 'Germany', '1936', 'Grass', 74475),
(301,'Wembley Stadium', 'England', '2007', 'Turf', 90000),
(401,'Parc des Princes', 'France', '1972', 'Turf', 47929);
SELECT * FROM stadium;
```

```
SELECT * FROM league_standings;
UPDATE league_standings ls
SET ls.points = (ls.wins * 3) + ls.draws;
SELECT * FROM league_standings;
```

Deliverable:
You are expected to highlight some of the most crucial commands and talk through them in the presentation. Do not include your WHOLE code in the presentation but add bits that showcase your understanding of **DDL & DML commands** and how to properly **assign keys** and use **aggregate functions.**

Our Code:
Keys & Aggregate Functions

ASSIGNING KEYS
PRIMARY
FOREIGN
REFERENCES

```
— Create the Team table
CREATE TABLE team (
team_id INT AUTO_INCREMENT PRIMARY KEY,
t_name VARCHAR(50) NOT NULL,
coach VARCHAR(50),
home_stadium INT,
league_id INT,
FOREIGN KEY (home_stadium) REFERENCES stadium (stadium_id),
FOREIGN KEY (league_id) REFERENCES league(league_id)
);
```

```
-- average % attendance at matches by day of week
SELECT m.day_of_week, ROUND(AVG(m.attendees/s.capacity*100))
AS average_stadium_percentage_full
FROM matches AS m
LEFT JOIN stadium AS s ON
m.stadium_id = s.stadium_id
GROUP BY day_of_week
ORDER BY average_stadium_percentage_full DESC;
```

```
-- List of teams that got red and yellow cards
SELECT t.t_name, SUM(ps.yellow_cards) AS yellow_cards,
SUM(ps.red_cards) AS red_cards
FROM team AS t
JOIN playerstats AS ps
ON t.team_id = ps.team_id
WHERE ps.yellow_cards > 0 OR red_cards > 0
GROUP BY t.t_name;
```

SUM( )
AVG( )
MIN( )
MAX( )
COUNT( )
AGGREGATE FUNCTIONS

Deliverable:
Screenshots or Snippets of the tables within your database: you are expected to showcase the schema you've created.

Our Tables

League

league_id	l_name	season	start_date	end_date
1	English Diva cup	Season 1	2023-03-19	2023-05-28

League Standings

standings_id	league_id	team_id	matches_played	wins	draws	losses	points
S001	1	1	6	3	1	2	10
S002	1	2	6	2	1	3	7
S003	1	3	6	2	0	4	6
S004	1	4	6	3	2	1	11

Stadium

stadium_id	s_name	location	year_opened	surface_type	capacity
101	Rose Bowl	USA	1922	Grass	92542
201	Olympiastadion	Germany	1936	Grass	74475
301	Wembley Stadium	England	2007	Turf	90000
401	Parc des Princes	France	1972	Turf	47929

Matches

match_id	match_date	day_of_week	stadium_id	attendees	team1_id	team2_id	team1_score	team2_score
1	2023-03-19	Sunday	101	87000	1	2	3	2
2	2023-03-27	Monday	201	40200	2	1	1	2
3	2023-04-06	Thursday	101	42000	1	3	1	2
4	2023-04-15	Saturday	301	89500	3	1	1	0
5	2023-04-16	Sunday	401	47900	4	1	1	1
6	2023-04-24	Monday	101	40000	1	4	2	0
7	2023-05-04	Thursday	301	38000	3	2	1	3
8	2023-05-13	Saturday	201	73900	2	3	2	0
9	2023-05-14	Sunday	401	47920	4	2	1	0
10	2023-05-22	Monday	201	32000	2	4	1	1
11	2023-05-25	Thursday	301	40650	3	4	4	5
12	2023-05-28	Sunday	401	47927	4	3	2	1

Team

team_id	t_name	coach	home_stadium	league_id
1	USA Falcons	Alex Morgan	101	1
2	German Tigresses	Martina Mueller	201	1
3	English Lionesses	Karen Taylor	301	1
4	French Sirens	Eugenie Le Sommer	401	1

Player Stats

playerstats_id	player_id	match_id	team_id	goals_scored	assists	yellow_cards	red_cards
PS01	US1	1	1	2	1	0	0
PS02	US2	1	1	1	2	0	0
PS03	US3	1	1	0	0	0	0
PS04	US4	1	1	0	0	0	0
PS05	GR1	1	2	2	0	1	0
PS06	GR2	1	2	0	2	0	0
PS07	GR3	1	2	0	0	0	0
PS08	GR4	1	2	0	0	1	0
PS09	US1	2	1	2	0	0	0

Player (excerpt)

player_id	p_name	p_position	p_number	team_id
EN1	Ava Turner	Forward	9	3
EN2	Lily Anderson	Midfielder	7	3
EN3	Ella White	Defender	4	3
EN4	Isabella Harris	Goalkeeper	1	3
FR1	Amelia Walker	Forward	6	4
FR2	Mia Roberts	Midfielder	9	4
FR3	Charlotte Turner	Defender	2	4
FR4	Grace Smith	Goalkeeper	1	4

Our Queries: Match Attendance

```
-- Stored procedure to calculate the attendance percentage per match in every stadium
DELIMITER $$
CREATE PROCEDURE `CalculateAttendancePercentage`()
BEGIN
    SELECT
        m.match_id,
        m.match_date,
        m.day_of_week,
        m.stadium_id,
        m.attendees,
        s.capacity,
        ROUND((m.attendees / s.capacity) * 100)
            AS AttendancePercentage
    FROM
        matches m
    JOIN
        stadium s ON m.stadium_id = s.stadium_id;
END $$
DELIMITER ;
CALL CalculateAttendancePercentage();
```

match_id	match_date	day_of_week	stadium_id	attendees	capacity	AttendancePercentage
1	2023-03-19	Sunday	101	87000	92542	94
3	2023-04-06	Thursday	101	42000	92542	45
6	2023-04-24	Monday	101	40000	92542	43
2	2023-03-27	Monday	201	40200	74475	54
8	2023-05-13	Saturday	201	73900	74475	99
10	2023-05-22	Monday	201	32000	74475	43
4	2023-04-15	Saturday	301	89500	90000	99
7	2023-05-04	Thursday	301	38000	90000	42
11	2023-05-25	Thursday	301	40650	90000	45
5	2023-04-16	Sunday	401	47900	47929	100
9	2023-05-14	Sunday	401	47920	47929	100
12	2023-05-28	Sunday	401	47927	47929	100

```
-- average % attendance at matches by day of week
SELECT m.day_of_week, ROUND(AVG(m.attendees/s.capacity*100))
AS average_percentage_attendance
FROM matches AS m
LEFT JOIN stadium AS s ON
m.stadium_id = s.stadium_id
GROUP BY day_of_week
ORDER BY average_percentage_attendance DESC;
```

day_of_week	average_percentage_attendance
Saturday	99
Sunday	98
Monday	47
Thursday	44

Player & Team Performance

player_name	team	player_position	goals	assists
Amelia Walker	French Sirens	Forward	8	3
Ava Turner	English Lionesses	Forward	6	1
Emma Johnson	USA Falcons	Forward	5	1
Lucy Wilson	German Tigresses	Forward	4	1
Mia Taylor	German Tigresses	Midfielder	4	3
Sophie Williams	USA Falcons	Midfielder	3	3
Lily Anderson	English Lionesses	Midfielder	2	4
Olivia Smith	USA Falcons	Defender	1	0
Chloe Roberts	German Tigresses	Defender	1	0
Ella White	English Lionesses	Defender	1	0
Mia Roberts	French Sirens	Midfielder	1	2

Top scorers and assists by position

Goals conceded

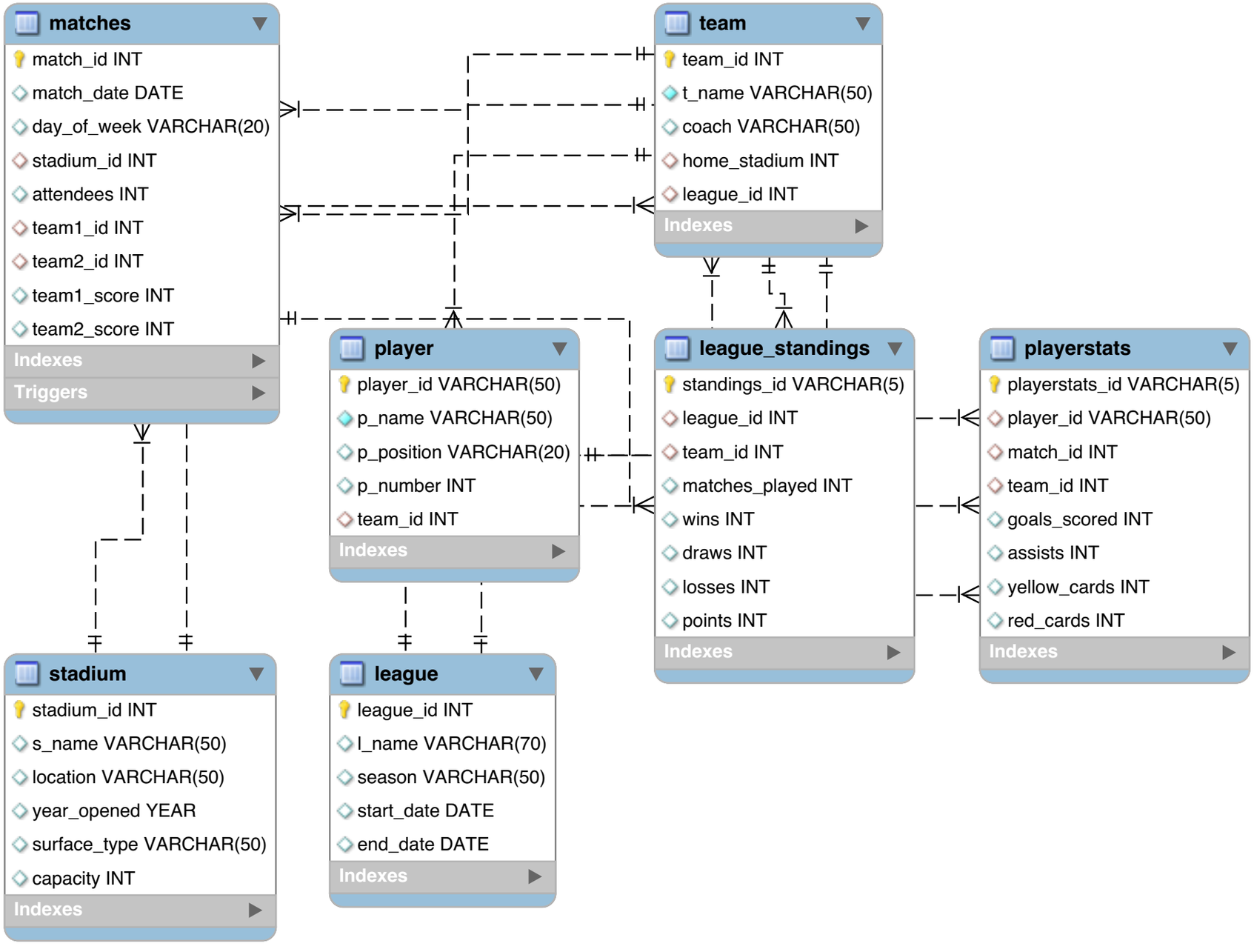
team_name	team_points	goals_conceded
English Lionesses	6	13
French Sirens	11	9
German Tigresses	7	8
USA Falcons	10	7

Discipline

t_name	points	yellow_cards	red_cards
French Sirens	11	8	1
USA Falcons	10	3	0
German Tigresses	7	4	1
English Lionesses	6	2	0

Deliverable:
EER diagram of the tables within their database: noting the relationships and fields included.

EER Diagram



Deliverable:
Evidence of teamwork: How did you adhere to the challenge and split the work between you? What challenges did you encounter and how did you overcome them?

Our Process

- 1st meeting to agree on which challenge and our approach
- Initial draft of the code prepared
- Code reviewed by team
- 2nd meeting to discuss code review
- Second draft of the code prepared
- 3rd meeting to discuss the updated code and next steps
- Updates to code from team members: queries and aggregate functions
- 4th meeting to agree finishing touches

Deliverable:
***Evidence of teamwork**:* How did you adhere to the challenge and split the work between you? What challenges did you encounter and how did you overcome them?

Challenges

CHALLENGE	STRATEGY
We did not use a collaboration platform - code was posted to slack. Up-to-date code therefore available only at intervals.	Development split into 2 phases. Phase 1: Creation and refinement of database. Significant updates to the database code were done by one team member at any given time. Phase 2: Preparation of queries. Did not affect database code. Frequent meetings and good communication!
Different team members reviewing code concurrently and adding points to slack: there was an element of duplication and it was difficult for them to be addressed methodically	A list of the team's review points and any responses was created from the slack history. This was used in the code review meeting to make sure everyone's contribution was registered.

“I’ve never scored a goal in my life without getting a pass from someone else.”

- Abby Wambach