

Please past below

(1) the SQL code you have used to create the schema of your database (only create table and alter table statements (if any), not statements for inserting values)

(2) the SQL code of the queries (possibly with an explanation)

(3) the SQL code used for query optimization for HW2. For each query, indicate the un-optimized version and the optimized one. In case the optimization has been realized through indexes, insert the SQL code for the index creation; in case you have modified the schema (e.g. changed the domain of a field, or constructed a new materialized table, etc.), insert the code you have used for this modification.

(1) the SQL code you have used to create the schema of your database (only create table and alter table statements (if any), not statements for inserting values).

```
SET @OLD_UNIQUE_CHECKS=@@UNIQUE_CHECKS, UNIQUE_CHECKS=0;
SET @OLD_FOREIGN_KEY_CHECKS=@@FOREIGN_KEY_CHECKS, FOREIGN_KEY_CHECKS=0;
SET @OLD_SQL_MODE=@@SQL_MODE,
SQL_MODE='ONLY_FULL_GROUP_BY,STRICT_TRANS_TABLES,NO_ZERO_IN_DATE,NO_ZERO_DATE,ERROR_FOR_DIVISION_BY_ZERO,NO_ENGINE_SUBSTITUTION';
```

```
-- -----
-- Schema Transfermarkt
-- -----
```

```
CREATE SCHEMA IF NOT EXISTS `Transfermarkt` ;
USE `Transfermarkt` ;
```

```
-- -----
-- Table `Transfermarkt`.`leagues`
-- -----
```

```
CREATE TABLE leagues (
    id VARCHAR(25),
    name VARCHAR(45) NOT NULL,
    confederation VARCHAR(45) NOT NULL
);
```

```
-- -----
-- Table `Transfermarkt`.`clubs`
-- -----
```

```
CREATE TABLE clubs (
    id VARCHAR(40) NOT NULL,
    name VARCHAR(45) NOT NULL,
    pretty_name VARCHAR(45) NOT NULL,
    leagues_id VARCHAR(25) NOT NULL,
    total_market_value FLOAT,
    squad_size INT NOT NULL,
    average_age FLOAT,
    foreigners_number INT,
    national_team_players INT NOT NULL,
    stadium_name VARCHAR(100) NOT NULL,
    stadium_seats INT NOT NULL,
```

```
coach_name VARCHAR(45) NOT NULL,  
url LONGTEXT NOT NULL  
);
```

```
-- -----  
-- Table `Transfermarkt`.`players`  
-- -----
```

```
CREATE TABLE players (  
    id VARCHAR(25) NOT NULL,  
    last_season CHAR(4) NOT NULL,  
    clubs_id VARCHAR(20) NOT NULL,  
    name VARCHAR(45) NOT NULL,  
    pretty_name VARCHAR(45) NOT NULL,  
    country_of_birth VARCHAR(45),  
    country_of_citizenship VARCHAR(45),  
    date_of_birth DATE NOT NULL,  
    position VARCHAR(45) NOT NULL,  
    sub_position VARCHAR(45) NOT NULL,  
    foot VARCHAR(45) NOT NULL,  
    height_in_cm INT NOT NULL,  
    market_value_in_gbp DECIMAL,  
    highest_market_value_in_gbp DECIMAL,  
    url LONGTEXT NOT NULL  
);
```

```
-- -----  
-- Table `Transfermarkt`.`competitions`  
-- -----
```

```
CREATE TABLE competitions (  
    id VARCHAR(5),  
    name VARCHAR(45) NOT NULL,  
    type VARCHAR(45) NOT NULL,  
    country_id VARCHAR(45) NOT NULL,  
    country_name VARCHAR(45),  
    leagues_id VARCHAR(15),  
    confederation VARCHAR(45) NOT NULL,  
    url LONGTEXT NOT NULL  
);
```

```
-- -----  
-- Table `Transfermarkt`.`games`  
-- -----
```

```
CREATE TABLE games (  
    id VARCHAR(45) NOT NULL,  
    competitions_id VARCHAR(5) NOT NULL,  
    season CHAR(4) NOT NULL,  
    round VARCHAR(45) NOT NULL,  
    date DATE NOT NULL,
```

```

        home_club_id VARCHAR(20) NOT NULL,
        away_club_id VARCHAR(20) NOT NULL,
        home_club_goals INT NOT NULL,
        away_club_goals INT NOT NULL,
        home_club_position INT,
        away_club_position INT,
        stadium LONGTEXT NOT NULL,
        attendance INT,
        referee VARCHAR(45),
        url LONGTEXT NOT NULL
    );

-- -----
-- Table `Transfermarkt`.`Appearances`
-- -----

CREATE TABLE appearances (
    players_id VARCHAR(25) NOT NULL,
    games_id VARCHAR(25) NOT NULL,
    id VARCHAR(45) NOT NULL,
    competitions_id VARCHAR(5) NOT NULL,
    clubs_id VARCHAR(20) NOT NULL,
    goals INT NOT NULL,
    assists INT NOT NULL,
    minutes_played INT NOT NULL,
    yellow_cards INT NOT NULL,
    red_cards INT NOT NULL
);

SET SQL_MODE=@OLD_SQL_MODE;
SET FOREIGN_KEY_CHECKS=@OLD_FOREIGN_KEY_CHECKS;
SET UNIQUE_CHECKS=@OLD_UNIQUE_CHECKS;

## *****
## Data Cleaning
## *****

## Games Table
## Drop column url
ALTER TABLE games
DROP url;

## Clubs Table
## Drop column url
ALTER TABLE clubs
DROP url;
## Drop column name
ALTER TABLE clubs
DROP name;

```

```
## Rename column pretty_name
ALTER TABLE clubs
RENAME COLUMN pretty_name TO Name;
```

```
## Players Table
## Drop column url
ALTER TABLE players
DROP url;
## Drop column name
ALTER TABLE players
DROP name;
## Rename column pretty_name
ALTER TABLE players
RENAME COLUMN pretty_name TO Name;
```

```
## Competitions Table
## Drop column url
ALTER TABLE competitions
DROP url;
```

(2) the SQL code of the queries (possibly with an explanation).

```
## 01
## Select the name of the top 10 players that have scored more goals
for Napoli.
SELECT p.Name AS Player, SUM(a.goals) AS Goals
FROM players AS p, appearances AS a
WHERE p.id = a.players_id AND a.clubs_id = (SELECT id FROM clubs WHERE
Name = 'Ssc Neapel')
GROUP BY p.id
ORDER BY goals DESC
LIMIT 10;
```

```
## 02
## Select the name of the player with the highest market value, his
club, his country of citizenship, his date of birth and his position.
SELECT c.Name AS Club, p.Name AS Player,
MAX(highest_market_value_in_gbp) AS Max_Value, country_of_citizenship
AS Country, date_of_birth AS Date, position AS Position
FROM players AS p, clubs AS c
WHERE highest_market_value_in_gbp = (SELECT
MAX(highest_market_value_in_gbp) FROM players)
AND c.Name = (SELECT Name FROM clubs WHERE id =
(SELECT clubs_id FROM players WHERE highest_market_value_in_gbp =
(SELECT MAX(highest_market_value_in_gbp) FROM players)));
```

```

## 03
## Select the italian clubs whose players have an average market value
greater than 10 million.
SELECT c.Name AS Club, ROUND(AVG(p.market_value_in_gbp), 2) AS
Average_Market_Value
FROM clubs AS c JOIN players AS p ON c.id = p.clubs_id
WHERE c.leagues_id = 'IT1' AND p.last_season = '2021'
GROUP BY Club
HAVING AVG(p.market_value_in_gbp) > 10000000
ORDER BY AVG(p.market_value_in_gbp) DESC;

```

```

## 04
## Select the players that had taken more than 4 red cards since 2014
until now.
SELECT p.Name AS Player, SUM(a.red_cards) AS Red_Cards
FROM players p JOIN appearances AS a ON p.id = a.players_id
GROUP BY p.id
HAVING SUM(a.red_cards) > 4
ORDER BY Red_Cards DESC;

```

```

## 05
## Select the players with the highest number of appearances in Serie A
(Top 10).
SELECT Name AS Player, COUNT(Name) AS Appearances
FROM players AS p, appearances AS a
WHERE p.id = a.players_id
      AND a.games_id IN (SELECT id
                        FROM games
                        WHERE competitions_id = 'IT1')
GROUP BY p.id
ORDER BY Appearances DESC
LIMIT 10;

```

```

## 06
## Select the names of the Italian players with the highest number of
appearances in Champions League.
SELECT p.Name AS Player, COUNT(p.Name) AS Appearances
FROM players AS p, appearances AS a
WHERE p.id = a.players_id AND p.country_of_citizenship = 'Italy' AND
a.competitions_id = 'CL'
GROUP BY p.id
ORDER BY Appearances DESC;

```

```

## 07
## Select the names of the Italian players with the highest number of
victories in Champions League matches.
SELECT p.Name, COUNT(p.Name) AS Wins
FROM players AS p, appearances AS a, (SELECT g.id AS game, c.id AS
club, c.Name
FROM games AS g,
clubs AS c
WHERE
g.competitions_id = 'CL' AND

((g.home_club_goals > g.away_club_goals) AND c.id = g.home_club_id OR

(g.away_club_goals > g.home_club_goals) AND c.id = g.away_club_id)) AS
cl
WHERE p.id = a.players_id AND a.games_id = cl.game AND a.clubs_id =
cl.club
AND p.country_of_citizenship = 'Italy'
GROUP BY p.id
ORDER BY Wins DESC;

## 08
## Select the players who have scored in all the Roma Lazio derby since
2014 and their number of goals (Top 10).
SELECT p.Name AS Player, SUM(a.goals) AS Goals
FROM players AS p, appearances AS a
WHERE p.id = a.players_id AND a.goals > 0 AND a.games_id IN (SELECT id

FROM games

WHERE (home_club_id = (SELECT id FROM clubs WHERE
Name = 'As Rom') AND

away_club_id = (SELECT id FROM clubs WHERE
Name = 'Lazio Rom')

or

(home_club_id = (SELECT id FROM clubs WHERE
Name = 'Lazio Rom') AND

away_club_id = (SELECT id FROM clubs WHERE
Name = 'As Rom'))))

GROUP BY p.Name
ORDER BY Goals desc
LIMIT 10;

```

```

## 09
## Select the clubs that have obtained the maximum number of victories
with a goal difference equal to 1.
SELECT c.Name AS Club, COUNT(c.Name) as Counter
FROM games AS g, clubs AS c
WHERE ((g.home_club_goals - g.away_club_goals) = 1 AND c.id =
g.home_club_id) OR
      ((g.away_club_goals - g.home_club_goals) = 1 AND c.id =
g.away_club_id)
GROUP BY c.id
ORDER BY Counter DESC;

```

```

## 10
## Select the match with the maximum number of goals: the teams that
played that match, the final result, the name of the competition, the
round of the competition, the date when the match was played, the
stadium and the attendance.
SELECT c_home.name AS Home_Team, c_away.name AS Away_Team,
      home_club_goals AS Home_Goals, away_club_goals
AS Away_Goals, MAX(home_club_goals + away_club_goals) AS Num_Max_Goal,
      round AS Round, c.Name AS Competition, date AS Date,
stadium AS Stadium, attendance AS Attendance
from clubs AS c_home, clubs AS c_away, games AS g, competitions AS c
WHERE c_home.id = (SELECT home_club_id
                  FROM games
                  WHERE (home_club_goals + away_club_goals) =
(SELECT MAX(home_club_goals + away_club_goals) FROM games))
AND c_away.id = (SELECT away_club_id
                 FROM games
                 WHERE (home_club_goals + away_club_goals) =
(SELECT MAX(home_club_goals + away_club_goals) FROM games))
AND home_club_goals = (SELECT home_club_goals
                       FROM games
                       WHERE (home_club_goals+ away_club_goals) =
(SELECT MAX(home_club_goals + away_club_goals) FROM games))
AND away_club_goals = (SELECT away_club_goals
                       FROM games
                       WHERE (home_club_goals+ away_club_goals) =
(SELECT MAX(home_club_goals + away_club_goals) FROM games))
AND c.id = g.competitions_id;

```

```

## 11
## Select the match with the largest number of spectators, the teams
that played the match, the competition and the round, the date when the
match was played, the stadium, the final result and the players who
scored.

## Set the external variables
# Select the match ID with the largest number of spectators
SET @Match_Id := (SELECT id FROM games where attendance = (SELECT
max(attendance) FROM games));
# Select the name of the first scorer
SET @p.Name := (SELECT Name FROM players where id in (SELECT players_id
FROM appearances where games_id = @Match_Id and goals > 0) limit 1);
# Select the name of the second scorer
SET @p2.Name := (SELECT Name FROM players where id in (SELECT
players_id FROM appearances where games_id = @Match_Id and goals > 0)
limit 1,1);
# Select the name of the third scorer
SET @p3.Name := (SELECT Name FROM players where id in (SELECT
players_id FROM appearances where games_id = @Match_Id and goals > 0)
limit 2,2);

## Query
SELECT c_home.Name AS Home_team, c_away.Name AS Away_Team,
competitions_id AS Competition, round AS Round, stadium AS Stadium,
        date AS Date, home_club_goals AS Home_Goals, away_club_goals
AS Away_Goals, MAX(attendance) AS Attendance,
        p.Name AS Soccer_Player1, p2.Name AS Soccer_Player2, p3.Name AS
Soccer_Player3
FROM games AS g, clubs AS c_home, clubs AS c_away, players AS p,
players AS p2, players AS p3
WHERE attendance = (SELECT max(attendance) FROM games) AND
        c_home.Name = (SELECT Name FROM clubs WHERE id = (SELECT
home_club_id FROM games WHERE attendance = (SELECT MAX(attendance) FROM
games))) AND
        c_away.Name = (SELECT Name FROM clubs WHERE id = (SELECT
away_club_id FROM games WHERE attendance = (SELECT MAX(attendance) FROM
games))) AND
        p.Name = @p.Name AND
        p2.Name = @p2.Name AND
        p3.Name = @p3.Name;

```



```

## 12
## Select the name of the players who scored more goals in Champions
League finals during their career.
SELECT p.Name AS Player, goals as Goals, c.Name AS Club
FROM players AS p JOIN appearances AS a ON p.id = a.players_id, games
AS g, clubs AS c
WHERE a.competitions_id = (SELECT id FROM competitions WHERE Name =
'uefa-champions-league') AND
      a.games_id IN (SELECT id
                      FROM games
                      WHERE competitions_id = 'CL' AND round =
'Final') AND
      a.goals IN (SELECT goals FROM appearances WHERE a.games_id IN
(SELECT id
                      FROM games
                      WHERE competitions_id = 'CL' AND round =
'Final') AND a.goals > 0) AND
      c.id = a.clubs_id
GROUP BY p.id
ORDER BY Goals DESC;

```

```

## 13
## Select the name of the player that scored more goals in a single
match, the club of the player, the match ID, the final result of the
match, the competition, the round and the date when the match was
played.
SELECT p.Name AS Player, c.Name AS Club, MAX(goals) AS Num_Max_Goals,
games_id AS Games_id, g.home_club_goals AS Home_Goals,
g.away_club_goals AS Away_Goals,
      g.competitions_id AS Competition, g.date AS Date, g.round AS
Round
FROM appearances AS a, players AS p, clubs AS c, games AS g
WHERE p.Name = (SELECT Name
                FROM players
                WHERE id = (SELECT players_id FROM appearances
WHERE goals = (SELECT MAX(goals) FROM appearances)))
      AND c.Name = (SELECT Name
                    FROM clubs
                    WHERE id = (SELECT clubs_id
                                FROM appearances
                                WHERE goals = (SELECT MAX(goals) FROM
appearances)))
      AND games_id = (SELECT games_id FROM appearances WHERE goals =
(SELECT MAX(goals) FROM appearances))
      AND home_club_goals = (SELECT home_club_goals FROM games WHERE id
= (SELECT games_id FROM appearances WHERE goals = (SELECT max(goals)
FROM appearances)))
      AND away_club_goals = (SELECT away_club_goals FROM games WHERE id
= (SELECT games_id FROM appearances WHERE goals = (SELECT max(goals)
FROM appearances)))
      AND g.competitions_id = (SELECT competitions_id FROM games WHERE
id = (SELECT games_id FROM appearances WHERE goals = (SELECT max(goals)
FROM appearances)))
      AND g.date = (SELECT date FROM games WHERE id = (SELECT games_id
FROM appearances WHERE goals = (SELECT max(goals) FROM appearances)))
      AND g.round = (SELECT round FROM games WHERE id = (SELECT
games_id FROM appearances WHERE goals = (SELECT max(goals) FROM
appearances)));

```

(3) the SQL code used for query optimization for HW2. For each query, indicate the un-optimized version and the optimized one. In case the optimization has been realized through indexes, insert the SQL code for the index creation; in case you have modified the schema (e.g. changed the domain of a field, or constructed a new materialized table, etc.), insert the code you have used for this modification.

```
-- -----  
-- Set keys and constraints on tables  
-- -----  
  
# Table leagues  
ALTER TABLE leagues  
ADD PRIMARY KEY (id);  
  
# Table competitions  
ALTER TABLE competitions  
ADD PRIMARY KEY (id);  
  
# Table clubs  
ALTER TABLE clubs  
ADD PRIMARY KEY (id),  
ADD FOREIGN KEY (leagues_id) REFERENCES leagues(id);  
  
# Table players  
ALTER TABLE players  
ADD PRIMARY KEY (id),  
ADD FOREIGN KEY (clubs_id) REFERENCES clubs(id);  
  
# Table games  
ALTER TABLE games  
ADD PRIMARY KEY (id),  
ADD FOREIGN KEY (competitions_id) REFERENCES competitions(id),  
ADD FOREIGN KEY (home_club_id) REFERENCES clubs(id),  
ADD FOREIGN KEY (away_club_id) REFERENCES clubs(id);  
  
# Table appearances  
ALTER TABLE appearances  
ADD PRIMARY KEY (id),  
ADD FOREIGN KEY (players_id) REFERENCES players(id),  
ADD FOREIGN KEY (games_id) REFERENCES games(id),  
ADD FOREIGN KEY (competitions_id) REFERENCES competitions(id),  
ADD FOREIGN KEY (clubs_id) REFERENCES clubs(id);
```

```

-- -----
-- Optimized Queries
-- -----

## 07 - Optimized by adding keys and constraints on tables
## Select the names of the Italian players with the highest number of
victories in Champions League matches
SELECT p.Name, COUNT(p.Name) AS Wins
FROM players AS p, appearances AS a, (SELECT g.id AS game, c.id AS
club, c.Name
                                     FROM games AS g,
clubs AS c
                                     WHERE
g.competitions_id = 'CL' AND

((g.home_club_goals > g.away_club_goals) AND c.id = g.home_club_id OR

(g.away_club_goals > g.home_club_goals) AND c.id = g.away_club_id) AS
cl
WHERE p.id = a.players_id AND a.games_id = cl.game AND a.clubs_id =
cl.club
      AND p.country_of_citizenship = 'Italy'
GROUP BY p.id
ORDER BY Wins DESC;

## 09 - Optimized by adding keys and constraints on tables
## Select the clubs that have obtained the maximum number of victories
with a goal difference equal to 1.
SELECT c.Name AS Club, COUNT(c.Name) as Counter
FROM games AS g, clubs AS c
WHERE ((g.home_club_goals - g.away_club_goals) = 1 AND c.id =
g.home_club_id) OR
      ((g.away_club_goals - g.home_club_goals) = 1 AND c.id =
g.away_club_id)
GROUP BY c.id
ORDER BY Counter DESC;

```

12 - Optimized by creating a view

Create View

DROP VIEW `transfermarkt`.`final`;

CREATE VIEW final AS

SELECT id

FROM games

WHERE competitions_id = 'CL' AND round = 'Final';

Select the name of the players who scored more goals in Champions League finals during their career.

SELECT p.Name AS Player, goals AS Goals, c.Name AS Club,

a.minutes_played AS Minutes_Played

FROM players AS p JOIN appearances AS a ON p.id = a.players_id, games AS g, clubs AS c, competitions AS co, final AS f

WHERE a.competitions_id = co.id AND co.Name = 'uefa-champions-league' AND

a.games_id = f.id AND

a.goals > 0 AND

c.id = a.clubs_id

GROUP BY p.id

ORDER BY Goals DESC;

```

## 13 - Optimized by creating a view.
# Create View
DROP VIEW `transfermarkt`.`match_id_max_goal`;
CREATE VIEW match_id_max_goal AS
SELECT games_id
FROM appearances
WHERE goals = (SELECT MAX(goals) FROM appearances);

## Select the name of the player that scored more goals in a single
match, the club of the player, the match ID, the final result of the
match, the competition, the round and the date when the match was
played.
SELECT p.Name AS Player, c.Name AS Club, MAX(goals) AS Num_Max_Goals,
a.games_id AS Games_id, g.home_club_goals AS Home_Goals,
g.away_club_goals AS Away_Goals,
      g.competitions_id AS Competition, g.date AS Date, g.round AS
Round
FROM appearances AS a, players AS p, clubs AS c, games AS g,
match_id_max_goal as mmg
WHERE p.Name = (SELECT Name
                  FROM players
                  WHERE id = (SELECT players_id FROM appearances
WHERE goals = (SELECT MAX(goals) FROM appearances)))
      AND c.Name = (SELECT Name
                     FROM clubs
                     WHERE id = (SELECT clubs_id
                                  FROM appearances
                                  WHERE goals = (SELECT MAX(goals) FROM
appearances)))
      AND a.games_id = mmg.games_id
      AND home_club_goals = (SELECT home_club_goals FROM games WHERE id
= mmg.games_id)
      AND away_club_goals = (SELECT away_club_goals FROM games WHERE id
= mmg.games_id)
      AND g.competitions_id = (SELECT competitions_id FROM games WHERE
id = mmg.games_id)
      AND g.date = (SELECT date FROM games WHERE id = mmg.games_id)
      AND g.round = (SELECT round FROM games WHERE id = mmg.games_id);

```

-- EER Diagram

