

McMaster University
SFWR ENG 3DB3 Fall 2020
Assignment 2
Due: November 13, 2020 at 6:00pm

October 16, 2020

For this assignment, you will continue to work with the *International soccer tournament database*, and perform data analytics over the data. On Avenue, under Assignments, you will find:

- a) the DDL for CREATE TABLE statements, `createTable.ddl` (to create the necessary tables), and
- b) the INSERT statements `loadData.ddl` (to load data into the tables).

These files correspond to a simplified schema shown in the E-R diagram `asg2ER.pdf`. Please drop all previous tables created by you for this database and run these two scripts on your database to create tables and then to populate them. You will use this schema for the questions below.

I. SQL (37 marks)

Write and provide SQL statements for the following queries. Execute each of your SQL queries against your international soccer tournament database, and give the result of each query.

- q1) [1 marks] Get the total number of records in the players table.
- q2) [1 marks] Display all records in the players table.
- q3) [2 marks each] Get the number of records from a table that meet a criteria.
 - i) Get the number of countries that participated in Euro Cup 2016.
 - ii) How many goals were scored during extra time (ET) in the entire tournament?
 - iii) How many goals were scored during stoppage time (ST) in the entire tournament?
 - iv) How many goals were scored during normal time (NT) in the entire tournament?
 - v) Get the number of bookings that happened during stoppage time.
 - vi) Get the number of bookings that happened during extra time.
- q4) [2 marks each] Use of MIN/MAX.

- i) Get the date when the first match was played in Euro Cup 2016.
 - ii) Get the date when the last match was played in Euro Cup 2016.
- q5) [2 marks each] More than one conditions.
- i) Get names of countries whose teams played the first match in Euro cup 2016.
 - ii) Get name of the country whose team won the final match of Euro cup 2016.
- q6) [3 marks each] Nested queries and use of operators:
- 1) Get names of countries with the number of penalty shots by their teams.
 - 2) Get names of countries whose teams played the final match of Euro cup 2016.
 - 3) Use a subquery with IN operator to get the dates when matches with penalty shootouts were played.
 - 4) Get names of venues where penalty shootout matches were played.
 - 5) Get the total number of players of the French team that participated in the final match.

II. Relational Algebra (6X3=18 marks)

For the SQL queries q3 in Part I, give the corresponding relational algebra expressions.

III. Indexes (5 marks)

For the workload given in Part I, several queries have been showing poor performance (i.e., increasing response times). Your task is to improve the performance of this workload as much as possible by recommending four indexes that should be defined on the tables. What four indexes would you recommend? For each index, state:

- The attribute(s) the index is defined on.
- Properties of the index (e.g., type of index, clustered/unclustered, etc.)
- Which queries (q1 - q6) you think this index will help, and why.

Grading

This assignment is worth 10% towards your final grade. A quiz based on this assignment will be given in tutorials during the week starting Nov 4 which is worth 4% towards your final grade.

Submission

All files are to be submitted using the Avenue system. Please ensure you submit all files with the correct names and your student ID prefixed to each file name in the drop box created for this assignment and your tutorial. For example, if your student ID is 1234567, the file name for Part III will be 1234567-index.pdf. Upload four files with the indicated file extensions (no compression based .tar, .zip, .rar files).

- For Part I: Submit your SQL statements and the result for each query in two files. Submit your SQL statements in a script file called `queries.sql`, and the corresponding query results in a file called `queries.results`. Ensure your SQL statements are syntactically correct and that they are executable on the DB2 servers. Non-executable queries will not be marked. Clearly label, with comments, which query the result tuples correspond to in `queries.results`.
- For Part II: Submit your relational algebra expressions in a file named `ra.pdf`.
- For Part III: Submit your index recommendations in a file named `index.pdf`.