Ghent University - TELIN - DDCM
2021-2022

SQL: Introduction

During the SQL lab sessions, you will extensively learn to write correct `SELECT`-queries in SQL. The reason for this is that correctly analyzing data in a relational database by means of `SELECT`-queries is one of the most essential skills for each database manager, statistician or software developer. Therefore, we expect you to apply a self-learning approach during the first part of the semester in order to learn and (finally) master SQL. Besides that, you will also have to prove, by means of two SQL assignments, that you indeed manage to analyze data correctly. In order to get familiar with our way of teaching and our expectations, we give a detailed overview of the practicalities related to the SQL lab sessions and assignments. We recommend you to read carefully throught this overview.

1 Lab sessions

We will organize 6 supervised lab sessions (1.25 hours per lab session) in total, during which you can solve SQL exercises and ask questions related to these exercises. The exercises can be solved on our self-developed exercise platform, named Q'exr, which is accessible via `qexr.ugent.be`. To access the exercises, you need an account on Q'exr, but normally, you have already received an email, in which we have provided a link that redirects you to a registration page (if not, please contact us). It is important that you register by using the link in the email and not directly by means of the registration page, otherwise, we are not able to give you access to the exercises. Your account remains available as long as you want (you can contact us to remove your account) and you can access the platform and the exercises at any time.

Once logged in, you will see different statistics related to the exercises. In the upper-left corner, you can press the 'Exercise' button that redirects you to the exercise module. Here, you will find an overview of all courses that you have access to. The course that we are going to use during the lab sessions is called 'Master Statistics Radio PostgreSQL'. This course consists of 4 series of exercises, all related to a database storing data about radio shows. Each of these exercise series focuses on a new SQL topic, as listed below.

1. Basic functionality
2. Combination of multiple tables
3. Subqueries
4. Aggregation, grouping and having

For each exercise, you have to submit a `SELECT`-query that provides an answer to the given question. Therefore, it is important to know the relational schema of the `radio` database, which can be found in the `relationalschema.pdf` document in the directory of the SQL exercises on Ufora. Besides that, you can find for each series an introductory presentation in which all concepts that you have to (or can) use are explained and, after each session, an FAQ file containing an overview of the most common questions. You can type your `SELECT`-query in the Query-box of the corresponding exercise and when you press the 'Validate' button, Q'exr will tell you if your answer is correct or not. Moreover, it will also provide you with some feedback about your answer and a comparison between the data retrieved by your answer and the expected data.

2 Assignments

As already pointed out in the introduction, you will have to solve two SQL-related assignments in order to test your SQL knowledge. Below, we provide a detailed overview of the practicalities related to the assignments.

General

- The first assignment will be made available on Ufora after the third hour of SQL lab sessions (February 25th, 2022) and will focus on concepts related to schema interpretation and basic SQL (i.e. basic functionality and combination of multiple tables). The second assignment will be made available on Ufora right after completion of the SQL lab sessions (March, 11th, 2022) and will mainly focus on concepts related to advanced SQL (i.e. subqueries, grouping and aggregation). Although, it is possible that some concepts learned during

the first three hours of lab sessions should also be used during the second assignment.

- The assignments have to be solved individually. If any form of irregularities are detected (e.g. plagiarism, collaboration,...), we will intervene in an appropriate way.
- To solve the exercises, you can use any of the provided (course) material (online documentation, presentations, solved exercises,...).

Preparation

- The SQL exercises are of the same type as those that were encountered during the lab sessions, as we will mainly ask to solve a problem by means of writing a `SELECT`-query. Important to know is that the exercises are not related to the `radio`-database, but to a second and previously unseen database, which stores data about car rentals and which is called `rollsrobin`. Again, the relational schema of this database can be found on Ufora (`relationalschema.pdf`).
- Because it is undesirable that you receive immediate feedback on your answers (it is part of the assignment to learn to assess both the correctness of the SQL query and the results), Q'exr will not be used for solving these exercises. Instead, you can use any tool of your choice, but we recommend to use pgAdmin 4.
- We have already set up a `rollsrobin` database containing data. To connect to this database, you can use the following connection parameters.
  - Hostname/IP-address: ddcmstud.ugent.be
  - Port: 8088
  - Database: rollsrobin_databases
  - Username: sql_exerciser
  - Password[1]: 7UCVuJeLCGcQbk2M

  This database is implemented on a PostgreSQL instance on one of our servers, and not on your own computer. Therefore, it is not necessary to have a local PostgreSQL server running, only pgAdmin 4 is needed. However, you can only connect with the database if you are within the UGent network (or by means of a VPN-connection to the UGent network).

[1] Do not copy the password from this file, but insert it manually to avoid any mistakes.

Submission

- Submitting your solutions can be done via the 'Assignments' module on Ufora. For each assignment, we will open a separate space in which you can submit your answers before the deadline. We will only look into the final submission before the deadline.
- For each assignment, we expect you to upload one .zip file (so not .tar, .7z,...) with filename `studentcode_firstname_lastname_assignmentX.zip`. In this filename, you have to substitute 'studentcode', 'firstname' and 'lastname' by your studentcode, first- and lastname (without spaces) respectively, and 'X' by the number of the assignment.
- In the .zip file, you may add two types of files. First, we can ask for a short report (in any format of your choice) containing textual answers to some exercises. Besides that, we can ask for `SELECT`-queries providing an answer to a `SQL`-related exercise. These queries should be added to separate (one for each exercise) .sql files (so not .txt,...) with `studentcode_firstname_lastname_X.sql` as filenames. Again, you have to substitute 'studentcode', 'firstname' and 'lastname' by your studentcode, first- and lastname (without spaces) respectively, and 'X' by the number of the corresponding SQL exercise.
- The deadlines of both assignments are listed below.
  - Assignment 1: March 11th, 2022, 10pm
  - Assignment 2: March 25th, 2022, 10pm

Correction

- Correction of the SQL exercises will be done by execution of an automatic correction tool (which is the same tool as the one adopted in Q'exr). This tool will verify whether the table that is returned as a result of your `SELECT`-query equals the expected table. Specifically, this means that, only when the following requirements are fulfilled, you receive all grades for the corresponding exercise.
  - The query does not feature any errors (e.g. syntax errors,...) or, in other words, the query will execute.
  - All columns in the result table have a correct datatype, as requested.
  - All columns in the result table have a correct name, as requested.
  - All rows in the result table contain correct data, as requested. No rows are absent or there are no additional rows.
  - All rows in the result table are unique, unless otherwise stated.

  - The order of the rows in the result table is as requested, but only when this is stated in the question. If not stated, the order does not matter.

- It is possible that the data on which your solution is verified does not equal the data that is provided to you. Therefore, your query should work correctly for each dataset within the `rollsrobin` database.

- The SQL assignments are part of the non-periodic evaluation (NPE). The NPE will account for 70% of the total score.

- To pass this course, you should obtain a minimum score of 10/20 for the NPE. Partial exemption for the NPE (retake) is possible if you have passed this part in the first evaluation period. If not, you should complete a modified version of the entire NPE during the summer.

- If any of the stated requirements are not met, you will lose all points for the assignment in which you have made a mistake. This is, for example, the case if filenames are not correct, files are not executable or unreadable, solutions are submitted after the deadline,... Therefore, make sure to check everything before submitting an assignment. If any form of irregularities are detected, further steps will follow.

Contact

Below, you find an overview of different ways to contact us, if you have some issues during any of the lab sessions.

- Send an email to toon.boeckling@ugent.be, yoram.timmerman@ugent.be and rihem.nasfi@ugent.be. In order to get a fast response, we recommend to send the email to multiple supervisors.

- Drop by our office (Sint-Pietersnieuwstraat 25, Technicum, floor -T, `https://soleway.ugent.be/routes/4453`) after making an arrangement. Make sure to contact us before dropping by such that we can make sure that somebody is present.

- Ask for an online meeting with one of the supervisors.

- Ask for help during the lab sessions.