The Chinese University of Hong Kong
Department of Systems Engineering and Engineering Management

SEEM3550A Fundamentals of Information Systems

2021-22 Assignment 1: SQL

This is an individual assignment. Everyone must submit it individually.1 This assignment is for understanding SQL (Chapter 3), which is worth of 10% of the total assessment. The due date for the assignment 1 is **17:00pm, March 11, 2022**. Submissions need to be submitted to the online Blackboard system. The late penalty will be 10% per day. A submission will not be accepted five days after the deadline.

Departmental Guideline for Plagiarism (Department of Systems Engineering and Engineering Management): If a student is found plagiarizing, his/her case will be reported to the Department Examination Panel. If the case is proven after deliberation, the student will automatically fail the course in which he/she committed plagiarism. The definition of plagiarism includes copying of the whole or parts of written assignments, programming exercises, reports, quiz papers, mid-term examinations and final examinations. The penalty will apply to both the one who copies the work and the one whose work is being copied, unless the latter can prove his/her work has been copied unwittingly. Furthermore, inclusion of others' works or results without citation in assignments and reports is also regarded as plagiarism with similar penalty to the offender. A student caught plagiarizing during tests or examinations will be reported to the Faculty office and appropriate disciplinary authorities for further action, in addition to failing the course. For more details, see the website http://www.cuhk.edu.hk/policy/academichonesty/.

You do not need to use MySQL to answer the questions for this assignment. But it helps, if you try to use MySQL.

It is important to note that all questions need to be answered by a single SQL query. By a single SQL, we mean basic SQL queries (slides 3.6 – 3.27), SQL queries using "set operations" (slides 3.28 – 3.29), "aggregate function" with "group by" and "having" (slides 3.33 – 3.37), and "nested subqueries" (slides 3.40 – 3.51). On the other hand, the following are not considered as a single SQL such as an SQL query using "with" (slide 3.52) or "views" (slides 4.12 – 4.16), and a

sequence of basic SQL queries. Also, join operators such as "left join", "right join" and "full join" (Chapter 4) are prohibited.

Questions

A relational database schema for an application is a set of relation schemas. Consider a database for a university. Its relational database schema is given on the slide 3.13, in ch3.ppt, which consists of 11 relation schemas. The relation names are shown in the blue rectangles. The corresponding attribute names are shown in the white box below the relation names. The underlined attribute names in a relational schema form the primary key. An arrow "A → B" from a set of attributes, A, in a relation schema R to a set of attributes, B, in another relation schema S shows that B is the primary key for S and A is a foreign key for R. The relation names and attribute names are self-explained. In the following, you are requested to answer questions using SQL statements. Assume no NULL value exists in the relations. Note that an answer is considered correct only if it addresses all possible cases.

- **Question 1 (30 marks):** Consider the 4 relations, namely, student, takes, section, and course in the relational database schema (the slide 3.13 in ch3.ppt). There are foreign keys defined: from takes to student, from takes to section, and from section to course. Answer the following questions while keeping the efficiency and answer correctness in mind.

  a) Suppose that you are requested to find student information (i.e., ID and name) and the course information (i.e., course id and course title) they took in Spring Semester ('Spring') in year 2022.

    a. (5 marks) Give a SQL to find the above information using 4 relations: student, takes, section, and course.

    b. (5 marks) Give a SQL to find the above information using 3 relations: student, takes and course.

    c. (5 marks) Explain if SQLs in a) and b) give the same answer. Justify your answer.

    d. (5 marks) Besides finding above information, you also need to find the building(s) of each course in which it is taken. Do SQLs in a) and b) give the same answer? Justify your answer.

  b) (5 marks) Give a SQL to find student IDs of students from the department of SEEM ('SEEM') and the course ids they took as well as the grades they got in each taken course.

c) (5 marks) Give a SQL to find IDs of students who took the course with id 12345 in year 2022.

- **Question 2 (25 marks):** Consider the following query: "Find IDs of students who take the courses offered in Spring Semester ('Spring') in year 2022 and get at least one grade 'A' or grade 'B'.

  a) (5 marks) Answer this query using set operations.

  b) (5 marks) Answer this query using joins and selections. You should not use nested SQL or derived relation.

  c) (5 marks) Answer this query using a nested query with correlation variables.

  d) (5 marks) Answer this query using a nested query without correlation variables.

  e) (5 marks) Answer this query using derived relations.

- **Question 3 (15 marks):** Consider a query to find the course IDs, course names in Spring Semester ('Spring') in year 2022 and number of students in each course.

a) (5 marks) Give your answer using aggregate function.

b) (10 marks) Consider a query to find the department names and the number of students in each department. One possible answer is as follow. Is it correct? Give your reason and give the correct query if it isn't.

```
SELECT dept_name, count (DICTINCT name)
FROM student
GROUP BY dept_name
```

- **Question 4 (30 marks):** Answer the following questions using SQL statements.

a) (5 marks) Show student names, if she/he does not have an advisor assigned yet, assuming that a student may not always have a advisor.

b) (5 marks) Show the names of the instructors whose name include the substring "abc" or end with "d".

c) (5 marks) Show the ID and name of all students having the highest total credits in the "SEEM" department.

d) (5 marks) For all buildings having total capacity (sum of the capacities of rooms in a building) greater than 1000, show their names and total capacity of each of them.

e) (5 marks) Show student names, if she/he does not has an advisor from 'SEEM' department. Assuming that a student may not always have an advisor.

f) (5 marks) For each department, show both the department budget and the remaining department budget after paying all its instructor's salaries.