**2021-22** Assignment 1: SQL

This is an individual assignment. Every one must submit it individually.[[1]](#footnote-1) This assignment is for understanding SQL (Chapter 3,4,5), which is worth of 20% of the total assessment. The due date for the assignment 1 is 11:59pm, February 24, 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 de

In the website for the textbook http://codex.cs.yale.edu/avi/db-book/, the solutions to practice exercises in most chapters are given. It is strongly recommended for you to try the practice exercises and see whether your answers are correct by checking the answers provided at the website. It will indeed help you to understand the course materials.

1

There is a well-known free database management system, called MySQL. On the eLearning system, in addition to tutorial 1, you can find how to install MySQL

(installxampposx.pptx for Max OS and installxamppwindows.pptx for Windows), how to create relations for the database schema given in ch3.pptx slide 3.13 by SQL (create.sql), and how to insert tuples into a relation (insert.sql). Start practicing SQL as soon as possible.

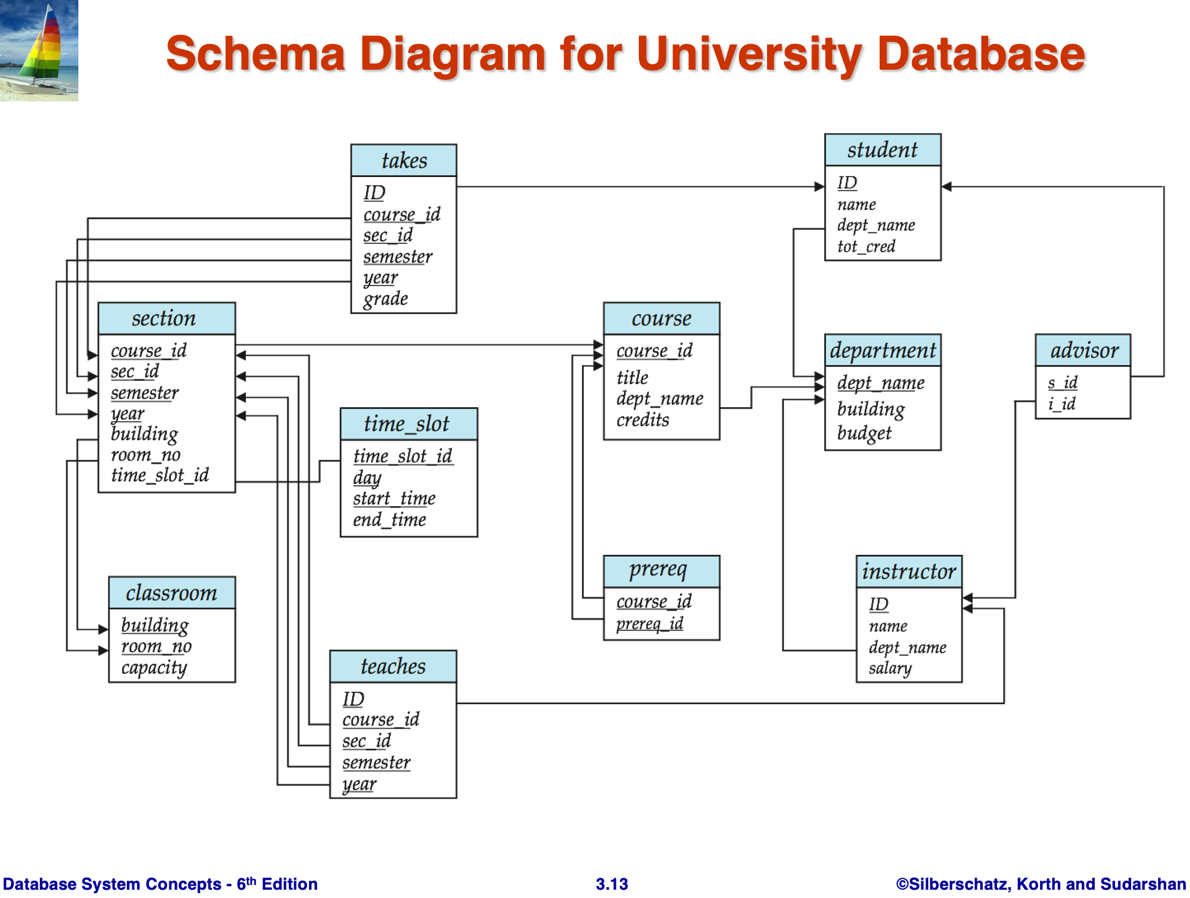
**You do not need to use MySQL to answer the questions for this assignment. But it helps, if you try to use MySQL.**

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 – 3.26), SQL queries using “set operations” (slides 3.27 – 3.28), “aggregate function” with “group by” and “having” (slides 3.32 – 3.38), “nested subqueries” (slides 3.41 – 3.51), “derived relations” (slide 3.52), “with” clause (slide 3.53). On the other hand, the following are NOT considered as a single SQL such as “**views**” (slides 4.12 – 4.16), and a sequence of SQL queries (excluding nested SQL queries).

# 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.pptx, 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. Note that NULL value exists, and an answer is considered correct only if it addresses all possible cases.



 **Question** 1: Consider the 4 relations, namely, student, takes, section, and course in the relational database schema （above）(the slide 3.13 in ch3.pptx). In the database schema design, there are foreign keys defined: from takes to student, from takes to section, and from section to course. Suppose that you are requested to find student IDs and the course information (e.g., courseid and course title) they have taken in 2021.

1. Do you have to follow the foreign keys in the database design to answer this question using all the 4 relations? Give you answer with justification.
2. Show an SQL query to answer this question using natural joins.
3. Show an SQL query to answer this question using *θ*-joins.

Note: we have discussed *θ*-join (slides 3.14-3.15 ch3.pptx) and natural join (slides 3.16-3.20 ch3.pptx).

 **Question** 2: Consider the following query: “For all courses, show its courseid if one of its section was once taught by a SEEM department instructor and was once taught by a CSE department instructor.”

1. Answer this query using SQL in which you must have set operations.
2. Answer this query using SQL in which you must only use joins and selections excluding set operations, nested subquery, or derived relation.
3. Answer this query using SQL in which you must use nested subquery without correlation variables.
4. Answer this query using SQL in which you must use nested subquery with correlation variables.
5. Answer this query using SQL in which you must use derived relations.

 **Question** 3: Consider a query to find the instructor names who have the highest salary.

1. Give your answer using aggregate function.
2. There is an incorrect answer to this question.

select name, salary from instructor order by name desc

Explain why this SQL is considered as incorrect.

 **Question** 4: We discussed assertion in Chapter 4. Write an assertion to ensure a student cannot take more than 30 credits per semester. Note that the credits of a course is specified in the credits attribute in the course relation.

 **Question** 5: Suppose that the grade attribute in the takes relation are not letter grades (e.g., A, A-, ...), but are numeric scores in the range of 0-100. Answer the following questions using SQL.

1. For all students, show their student ID, name, and their total numeric scores.
2. For all students, show their student ID, name, and average numeric scores.
3. For every department, show the department name, and the maximum average numeric score.
4. Show the department name with the lowest average numeric score over the answer of (3).

 **Question** 6: Answer the following questions using SQL statements.

1. Show the full department names that contain ’Eng.’ twice, assuming that you do not know the full name.
2. Show candidate classrooms with building name, room no, and capacity for a section that has 100 students, which are not used for any other sections on Thursday in term Fall, 2022. Note that the day in the relation of timeslot is week day.
3. Find student names and the instructor names if the students have taken any course taught by their advisor. Note that the advisor relation records that an instructor (iid) is the advisor of a student (sid).
4. For each student who has not taken any section taught by his/her advisor, output his/her ID, assuming that each student has exactly one advisor.
5. To help students to plan their study plan over 4 years, you are required to show all prerequisites (directly and indirectly) for all courses offered by SEEM department.

1. [↑](#footnote-ref-1)