

SQL Practice Exercise

This assignment will use a database containing data about a university. The schema of the database is provided below (keys are in bold, field types are omitted).

- student(**sid**, sname, sex, age, year, gpa)
- dept(**dname**, numphds)
- prof(**pname**, dname)
- course(**cno**, cname, **dname**)
- major(**dname**, **sid**)
- section(**dname**, **cno**, **sectno**, pname)
- enroll(**sid**, grade, **dname**, **cno**, **sectno**)

You can create the tables in a database of your choice using the data.sql file provided.

Questions

Write SQL queries that answer the questions below (one query per question) and run them on the Database System using its SQL interpreter. The query answers must not contain duplicates, but you should use the SQL keyword **distinct** only when necessary.

1. Print the names of professors who work in departments that have fewer than 50 PhD students.
2. Print the name(s) of student(s) with the lowest gpa.
3. For each Computer Sciences class, print the cno, sectno, and the average gpa of the students enrolled in the class.
4. Print the course names, course numbers and section numbers of all classes with less than six students enrolled in them.
5. Print the name(s) and sid(s) of the student(s) enrolled in the most classes.
6. Print the names of departments that have one or more majors who are under 18 years old.
7. Print the names and majors of students who are taking one of the College Geometry courses. (Hint: You'll need to use the "like" predicate and the string matching character in your query.)
8. For those departments that have no majors taking a College Geometry course, print the department name and the number of PhD students in the department.
9. Print the names of students who are taking both a Computer Sciences course and a Mathematics course.
10. Print the age difference between the oldest and youngest Computer Sciences major(s).
11. For each department that has one or more majors with a GPA under 1.0, print the name of the department and the average GPA of its majors.
12. Print the ids, names, and GPAs of the students who are currently taking all of the Civil Engineering courses.

FAQ (more recent questions on the top)

Questions 9/22

Q: For all queries having 'less than' requirements, do we take 'strictly less than' or 'less than equal to', as for query #4 and #6.

A: Strictly less than.

Q: Can we assume that a student is only enrolled in one section for a given dname and cno? This makes sense but I need that assumption for how I'm counting how many courses a student takes in Query 12. I want to just be able to count how many tuples a student has in the enroll table with dname='Civil Engineering'

A: No, you can't assume that. Try to think of the query in terms of division in relational algebra and how the division is translated into SQL. Be careful when using count in what you are counting. Feel welcome to talk to me tomorrow during the office hours or after class tomorrow.

Q: By class do you mean section or course....basically is the condition (print out stuff for every course with less than 6 students) or (print out stuff for every section with less than 6 students).

A: For every section.

Q: 12. Print the ids, names, and GPAs of the students who are currently taking all of the Civil Engineering courses. Can a student be enrolled in the same class twice?

A: Given that we do not know any other constraints, a student could potentially be enrolled in the same class in different sections.

Q: Should a student who has taken only Civil Engineering courses be in the output of Query 12?

A: Not necessarily. Only if the student is taking all Civil Engineering courses, the student should be in the output of Query 12. (The student could be enrolled in other courses.)

Q: For q7, we were asked to output "the names and majors of students who have a major and are taking one of the College Geometry courses." For the case when a student double-majors, do we print out the same name twice with different majors? Or just once with one of his majors?

A: Twice with different majors.

Q: I'm trying to do query #5 and Access doesn't like the following syntax

```
SELECT MAX(cnt)  
FROM (SELECT COUNT(*) ...
```

Any suggestions/hints?

A: Access does not support nested queries in the from clause. Try to write the query in a different way.

Q: I have a question regarding query #3: do you want the average gpa of the students enrolled for each computer science class, or for each computer science class' section?

A: For each section. The text of the query has been modified to make this clear.

Q: Is each student supposed to have at least one major? And some can have more than one? Is each student supposed to only have one major?

A: Students can have more than one major and are not required to have a major. The query text has been slightly modified to only output students who have majors.