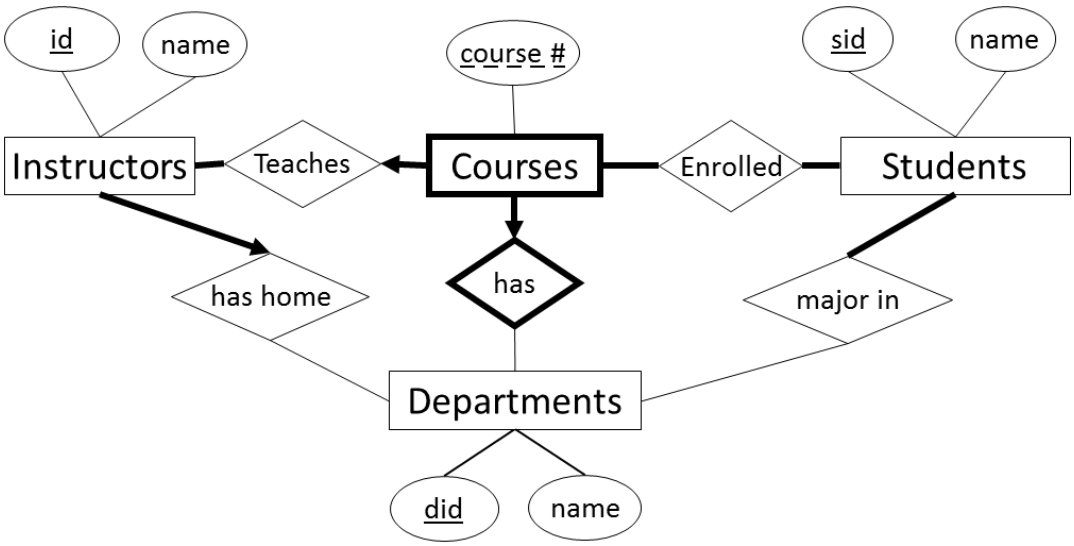
CS4380/7380 Spring 2016 Project 2
Relational Algebra and Structured Query Language

Due: Tuesday, 3/1/2016 by 9:30am

Perform all of the instructions below on your MySQL database on Azure. The project will be graded in two parts. The requirements for each part are described below.



Load the provided SQL file (project2.sql) into MySQL on your Azure account (hint: Log on to MySQL and create a new database, like "project2". Then quit MySQL and use the command-line command `mysql -h localhost -u root -p project2 < project2.sql` to load project2.sql into the newly created project2 database).

Explore the database to see how the above ERD was represented in tables. There may be some minor differences between the ERD and the database. Show the available tables (e.g. "show tables;"), use the Describe command (e.g. "describe department;") and use simple queries (e.g. "select * from student;") to understand the data that is in the database and how you can use it to answer the following problems.

Part 1. Relational Algebra

Write the **Relational Algebra** statements to denote the criteria described below. Only the operations discussed in the lecture videos are allowed (e.g. there are no "count" operations or the like). There is not necessarily one correct answer to these problems, and any reasonable solution will be accepted as long as it satisfies the criteria.

Submit your answers in a **PDF** file on Canvas (additional submission instructions are at the bottom of this document). Relational Algebra statements must be typed. **Points will be subtracted for handwritten answers.** If you do not have a method in mind to type the relational algebra statements, we recommend using Microsoft Word's Equation Editor (click INSERT > Equation (far right)).

Write the relational algebra for the following statements:

1. Find the names (first and last) of students who are enrolled in a class taught by an instructor with first name "Philip".

2. Find the id and full name of instructors who teach at least two courses.

3. Find the did and name of departments that have no students majoring in it.

4. Find the sid and full name of students who are enrolled all of the courses in department #200847.

5. Find the id and full name of instructors who teach both a 4000-level course and a 7000-level course.

6. Find the id and name of instructors in whose class there is at least one student majoring in the "Informatics Institute" department.

Part 2. Structured Query Language

Write the **SQL queries** to find the results for the problems below. There is not necessarily one correct answer to these problems, and any reasonable solution will be accepted as long as it satisfies the criteria. Test your queries on the provided dataset. Note that we will use your queries on a larger data set that may have different IDs for many of the tuples to test the robustness of your submission. Thus, it is important to write an effective query that thoroughly solves the problem, not just to get the correct results for the example dataset.

Submit your answers in an SQL file on Canvas (additional submission instructions are at the bottom of this document). We have provided an example SQL file to demonstrate the file format, called "ExampleQueryFile.sql". Please test your SQL files by using a command like "`mysql -h localhost -u root –p project2 < ExampleQueryFile.sql`" (note that this is very similar to the command you use to load data into your database). If the file has no errors and executed the queries correctly, it should show the results of the queries to the command line. The output for running the above command on the ExampleQueryFile.sql file is below:

```
mcsgx2@cs4380-mcsgx2:~/project2$ mysql -h localhost -u root -p project2 < ExampleQueryFile.sql
Enter password:
instructor
10088
sid     fname   lname
5000022 Wyatt   Johnson
5000038 Logan   Johnson
```

Write the SQL queries for the following statements:

1. All of the relational algebra statements from Part 1.

2. Find the did and name of department that is home to the most instructors.

3. Find the names of students who are enrolled in exactly two courses.

4. Find the instructor(s) who teaches the most courses.

5. Display the department name and course number, along with the average, max, and minimum age of students for each course that belongs to the "Computer Science" department. Sort the results by course number from lowest to highest.

On the Canvas site, navigate to Modules > Assignments > Project 2. Click "Submit Assignment". There you will upload two files:

PAWPRINT_project2.pdf	PDF file containing the RA answers to Part 1
PAWPRINT_project2.sql	SQL file containing the SQL query answers to Part 2