

CS 1656 – Introduction to Data Science

Instructor: Alexandros Labrinidis / Teaching Assistant: Tahereh Arabghalizi

Additional credits: Evangelos Karageorgos, Zuha Agha, Phuong Pham

Lab 8: Advanced SQL

Provided below is the student database schema for queries in this lab.

- Students (sid: int, firstName: str, lastName: str, yearStarted: int)
 - Majors (sid: int, major: str)
 - Note: a student may have more than one major.
 - Grades (sid: int, cid: int, credits: int, grade: int)
 - Note: sid, cid=foreign keys, grades: A=4, B=3, C=2, D=1, F=0.
 - Courses (cid: int, number: int, professor: str, major: str, year: int, semester: str)
 - Note: cid is unique across semesters.
 - semester is either Summer, Fall, or Spring.
 - two course offerings are the same if they have same number + major.
-

Q1. Provide the SQL query that will print the first and last name of all sophomore students (i.e., yearStarted = 2017).

```
SELECT firstName, lastName
FROM students
WHERE yearStarted = 2017;
```

Q2. Provide the SQL query that will print the first and last name of all students who have either a CS or a COE major.

```
SELECT firstName, lastName
FROM students NATURAL JOIN majors
WHERE major = 'COE' OR major = 'CS';
```

Q3. Provide the SQL query that will generate the number of students who have ASTRO as their major.

CS 1656 – Introduction to Data Science

Instructor: Alexandros Labrinidis / Teaching Assistant: Tahereh Arabghalizi

Additional credits: Evangelos Karageorgos, Zuha Agha, Phuong Pham

```
SELECT COUNT(major) AS NumberOfStudentsWhoMajorInASTRO
FROM majors
WHERE major = 'ASTRO';
```

Q4. Provide the SQL query that will generate the first name, last name, yearStarted, and the total number of credits for every student. You should not consider courses with a 0 grade, since these correspond to failed courses.

```
SELECT s.firstName, s.lastName, s.yearStarted, SUM(g.credits) AS 'total credits'
FROM students AS s, grades AS g
WHERE g.sid = s.sid AND g.grade > 0
GROUP BY g.sid;
```

Q5. Provide the SQL query that will generate the professor name and how many courses that professor has taught, for every professor.

```
SELECT professor, COUNT(cid) AS courses_taught
FROM courses
GROUP BY professor;
```

Q6. Provide the SQL query that will show all courses and the distribution of grades (i.e., how many 4s, how many 3s, how many 2s, how many 1s, and how many 0s) for each course.

```
SELECT cid, grade, COUNT(*) AS n
FROM GRADES
GROUP BY CID, GRADE
```

Q7. Modify the query from Q6 to only consider A grades.

```
SELECT cid, grade, COUNT(*) AS n
FROM GRADES
```

CS 1656 – Introduction to Data Science

Instructor: Alexandros Labrinidis / Teaching Assistant: Tahereh Arabghalizi

Additional credits: Evangelos Karageorgos, Zuha Agha, Phuong Pham

WHERE GRADE=4

GROUP BY CID, GRADE