1. This is a CLOSED book examination. Personal copies of the textbook and class notes cannot be used. However, each student may bring one 8.5x11 inch sheet of summary notes. This sheet cannot be shared with other students. Laptop computers and calculators are not allowed.

2. There are 4 questions. Use a blue book to answer these questions. Please make sure your name, student id, email address, course id, semester, and year are on the cover of the blue book.

3. All questions about SQL refer to SQL2 or SQL3 standard.

4. Clearly state any assumptions that you make regarding possible ambiguities.

**Question 1. (4 pts)** Consider the following relation instances...

<table>
<thead>
<tr>
<th>id</th>
<th>name</th>
<th>age</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Alice</td>
<td>25</td>
</tr>
<tr>
<td>2</td>
<td>Bob</td>
<td>23</td>
</tr>
<tr>
<td>3</td>
<td>Carla</td>
<td>33</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>id</th>
<th>name</th>
<th>age</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>Dustin</td>
<td>53</td>
</tr>
<tr>
<td>2</td>
<td>Bob</td>
<td>23</td>
</tr>
<tr>
<td>5</td>
<td>Emily</td>
<td>53</td>
</tr>
<tr>
<td>3</td>
<td>Carla</td>
<td>33</td>
</tr>
</tbody>
</table>

Table 1: S1  
Table 2: S2

Write the result **CARDINALITY** of the following relational expressions involving the above instances of S1 and S2...

Q1a. S1 \( \cup \) S2 (Union)
Q1b. S1 \( \cap \) S2 (Intersection)
Q1c. S1 \( \times \) S2 (Cross Product)
Q1d. S1 - S2 (Difference)
Q1e. Selection on S2 where age < 30
Q1f. Projection of S2 on age
Q1g. Projection of S2 on id, age
Q1h. Join of S1 and S2 on S1.id = S2.id

**Question 2. (12 pts)** Consider the following schema...

```sql
CREATE TABLE T (C INT PRIMARY KEY, D INT);
CREATE TABLE S (B INT PRIMARY KEY, C INT REFERENCES T(C) ON DELETE SET NULL);
CREATE TABLE R (A INT PRIMARY KEY, B INT REFERENCES S(B) ON DELETE CASCADE);
```

Q2a. (4 pts)
Suppose the current contents of R, S, and T are as follows...

R(A,B) has \{ (10,10), (20,20) \}
S(B,C) has \{ (10,10) and (20,10) \}
T(C,D) has \{ (10,10) and (20,10) \}

After executing the command: DELETE FROM S;

What tuples will each R and S contain? (For example: no tuples in R, \{ (10,null),(null,null) \} in S)

Q2b. (4 pts)
Suppose the tables R, S, and T contain r, s, and t tuples respectively. Let n be the number of tuples in the result of the equi-join query: SELECT * FROM R,S,T WHERE R.B = S.B AND S.C=T.C;

What is the most restrictive (smallest) yet correct upper bond on n?

(A) t (B) r (C) \( \min(r,s,t) \) (D) \( r + s + t \) (E) \( r \times s \times t \) where * is integer multiplication

Q2c. (4 pts)
Suppose the tables R, S, and T contain r, s, and t tuples respectively. Let n be the number of tuples in the result of the outer join query: (R OUTER JOIN S ON R.B = S.B) OUTER JOIN T ON S.C = T.C

What is the most restrictive (smallest) yet correct upper bond on n?

(A) t (B) r (C) \( \min(r,s,t) \) (D) \( r + s + t \) (E) \( r \times s \times t \) where * is integer multiplication
Question 3. Consider the following schema...

Student(`snum, sname, major, level, age`)  
Class(`name, meets_at, room, fid`)  
Enrolled(`snum, cname`)  
Faculty(`fid, fname, deptid`)

A student is uniquely identified by a student number `snum`, and information about their name, major, grade level, and age is stored. Classes are uniquely identified by their `name`, and their meeting time, room number, and faculty instructor are stored. Enrolled lists which students (`snum`) are enrolled in which classes (`cname`). Faculty are uniquely identified by `fid`, and information is kept about their name and department.

Write SQL queries to answer the following...

Q3a. (15 pts) Find the names of all Seniors (level = SR) who are enrolled in a class taught by J. Doe.

Q3b. (15 pts) For each level, print the level and the average age of all students for that level.

Q3c. (15 pts) For levels other than FR, list average age of students for that level. Result includes 2 columns, namely, level and average age.

Q3d. (15 pts) List following information about faculty members who have taught classes only in room R128 (i.e. they never taught in any room other than R128). Result includes three columns, namely, fid, fname and number of distinct classes taught.

Question 4. Birdland Records has asked you to design a database to store information about musicians who perform on their albums. After listening to their needs you end up with the following tables (Primary keys are underlined, Foreign keys are in italics. Note that a foreign key shares the name with the referenced primary key)

Musician(`mssn, name, address, phone_num`)  
Instrument(`instr_id, instr_name, musical_key`)  
Album(`album_title, album_author, mssn`)  
Song(`song_id, album_title`)  
Performed-By(`songid, mssn`)  
Played-By(`instr_id, mssn`)  
Live-Performance(`l-p_id, location, time`)  
Participates-In(`l-p_id, mssn`)

- Each musician at Birdland has a unique SSN, a name, an address, and a phone number.
- Each instrument used at Birdland has a unique ID, name and a musical key.
- Each album recorded at Birdland has a title and an author.
- Musicians also participate in several live performances. Each performance may be attended by several musicians.
- Each musician may play several instruments, and a given instrument may be played by several musicians.
- Some albums have a musician act as its producer. A musician may produce any number of albums.
- Each album has a number of songs on it, a song can only appear only in one album.
- Each song is performed by one or more musicians, and a musician may perform a number of songs.

Q4a. (15 pts) Draw an entity relationship diagram based on the above relationship schema and information. For each entity, specify identifying attributes. For each relationship specify the maximum cardinality. Briefly justify the correctness of the solution by transferring it to relational schema using the conversion rules discussed in the textbook.

Q4b. (5 pts) Identify cycles with two or more entities in your ERD. Do these cycles indicate redundancy (i.e. Can one infer one of the relationships using other relationships in the cycle)? Briefly justify your answer using the available description of the data.
Question 5. The following Facebook logical data model is reverse engineered by a contributor from archive.org (modified for this class)\(^1\). In this model, tables are represented as classes, and edges between tables are foreign key constraints. The primary key of each table is underlined, and the foreign keys are followed by a “*”. The attributes of each table are listed as properties of classes, and the relationships with one arrow represent 1:1 relationship, and double-arrow represents 1:N relationship, and the name of each relationship is shown along the edges.

Q5a. (5 pts) Critique this logical data model based on what you have learned in this course so far.

Q5b. (5 pts) Create a conceptual data model using ERD in Crow’s foot notation.