SAMPLE Midterm Exam – 1 Hour 20 Minutes

Name:

Grade (out of 100pts):

<table>
<thead>
<tr>
<th>Problem Set</th>
<th>Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. (18 points)</td>
<td></td>
</tr>
<tr>
<td>2. (20 points)</td>
<td></td>
</tr>
<tr>
<td>3. (25 points)</td>
<td></td>
</tr>
<tr>
<td>4. (20 points)</td>
<td></td>
</tr>
<tr>
<td>5. (17 points)</td>
<td></td>
</tr>
</tbody>
</table>

Instructions:

1. Put your name and student number on your exam now.

2. Give your answer in the provided space. If you need more space, use the back of the pages and make a note of it, BUT it should not be necessary.

3. Show all your work! Partial credit is possible for a wrong answer but only if you show the intermediate steps in obtaining an answer.

4. The number of points for each part of each question is given in parentheses. There is a total of 100 points.

5. Good Luck!
Problem Set 1: General Concepts (18 points)

1. (4 pts) What is a data model?

2. (4 pts) What is the difference between a relational table, a file of records, and a spreadsheet?

3. (5 pts) Why is referential integrity not necessary in the ER model, although it is important in the relational model?

4. (5 pts) Define outer union and explain its usage.
1. In the following relation schemas all three attributes form the primary key:

   \[ R = (A,B,C) \quad S = (D,E,F) \]

Let relations \( r(R) \) and \( s(S) \) be given. Give an expression in SQL that is equivalent to each of the three queries below:

(a) (9 pts) \( \sigma_{B=13}(r \times s) \)

(b) (11 pts) \( \sigma_{B=13}(\pi_{B,D,F}(r \bowtie \sigma_{A=D \land C=D}(s))) \)
Problem Set 3: SQL (25 points)

1. (25 pts) Given the populated relations \( r, s \) below, what would the following SQL queries return:

\[
\begin{array}{ccc}
\text{A} & \text{B} \\
0 & 1 \\
1 & 1 \\
3 & 1 \\
4 & 1 \\
2 & 2 \\
3 & 2 \\
4 & 2 \\
1 & 4 \\
\end{array}
\quad
\begin{array}{ccc}
\text{B} & \text{C} \\
1 & 0 \\
1 & 0 \\
1 & 2 \\
1 & 3 \\
2 & 1 \\
3 & 4 \\
4 & 2 \\
4 & 3 \\
\end{array}
\]

a. (7 pts)
select A from r where not A = any (select C from s)

b. (9 pts)
select avg(A) as A_AVG, r.B as B
from (r natural join s)
where s.C > 0
group by r.B
having count(*) > 3;

c. (9 pts)
select avg(A), r.B
from r, s
where r.B = s.B and C > 20
group by r.B
having avg(A) >= (select avg(C) from s);
1. (20 points) Consider the following schema: (SSN stands for social security number and is the key of EMPLOYEE. SUPERSSN is the social security number of a person's supervisor. DNUMBER is the key of DEPARTMENT and MGRSSN is the social security number of the department manager. Both attributes of DEPTLOCATIONS form a key.)

\[
\text{EMPLOYEE (SSN, FName, LName, Address, City, Salary, SuperSSN, DNO)} \\
\text{DEPARTMENT (DNUMBER, DName, MGRSSN)} \\
\text{DEPTLOCATIONS (DNUMBER, DLocation)}
\]

Translate to SQL the two queries below:

Query 1: List the full names and the departments of all employees whose salary is between 50000 and 75000 and who live and work in Phg.

Query 2: List the full names of all employees who are directly supervised by their department manager.
• Consider the following set F of functional dependencies for relational schema R = (A, B, C, D, E):

  1. A → BC
  2. CD → E
  3. B → D
  4. E → A

Compute B+. 