Exercise 1: Relational Model

1. Consider the relational database of next relational schema with 3 relations. What are the best possible primary keys in each relation?

   employ(person_name, street, city)

   works(person_name, company_name, salary)

   company(company_name, city)

2. Consider the foreign key constraint from the dept name attribute of instructor to the department relation. Give examples of inserts and deletes to these relations, which can cause a violation of the foreign key constraint.

3. Consider the time slot relation. Given that a particular time slot can meet more than once in a week, explain why day and start time are part of the primary key of this relation, while end time is not.

4. In the instance of instructor shown in next figure, no two instructors have the same name. From this, can we conclude that name can be used as a superkey (or primary key) of instructor?

<table>
<thead>
<tr>
<th>id</th>
<th>name</th>
<th>dept_name</th>
<th>salary</th>
</tr>
</thead>
<tbody>
<tr>
<td>10101</td>
<td>Srinivasan</td>
<td>Comp. Sci.</td>
<td>65000</td>
</tr>
<tr>
<td>12121</td>
<td>Wu</td>
<td>Finance</td>
<td>90000</td>
</tr>
<tr>
<td>15151</td>
<td>Mozart</td>
<td>Music</td>
<td>40000</td>
</tr>
<tr>
<td>22222</td>
<td>Einstein</td>
<td>Physics</td>
<td>95000</td>
</tr>
<tr>
<td>32343</td>
<td>El Said</td>
<td>History</td>
<td>60000</td>
</tr>
<tr>
<td>33456</td>
<td>Gold</td>
<td>Physics</td>
<td>87000</td>
</tr>
<tr>
<td>45565</td>
<td>Katz</td>
<td>Comp. Sci.</td>
<td>75000</td>
</tr>
<tr>
<td>58583</td>
<td>Caliñieri</td>
<td>History</td>
<td>62000</td>
</tr>
<tr>
<td>75543</td>
<td>Singh</td>
<td>Finance</td>
<td>80000</td>
</tr>
<tr>
<td>76766</td>
<td>Crick</td>
<td>Biology</td>
<td>72000</td>
</tr>
<tr>
<td>83821</td>
<td>Brandt</td>
<td>Comp. Sci.</td>
<td>92000</td>
</tr>
<tr>
<td>98345</td>
<td>Kim</td>
<td>Elec. Eng.</td>
<td>89000</td>
</tr>
</tbody>
</table>

5. What is the result of first performing the cross product of student and advisor, and then performing a selection operation on the result with the predicate s_id = ID? (Using the symbolic notation of relational algebra, this query can be written as $\sigma_{s\_id = ID}(\text{student} \times \text{advisor}$.)
6. Consider the following expressions, which use the result of a relational algebra operation as the input to another operation. For each expression, explain in words what the expression does.

   a. $\sigma_{\text{year} \geq 2009}(\text{takes}) \bowtie \text{student}$
   b. $\sigma_{\text{year} \geq 2009}(\text{takes} \bowtie \text{student})$
   c. $\Pi_{\text{id}, \text{name}, \text{course}_\text{id}}(\text{student} \bowtie \text{takes})$

7. Consider the relational database of the following schema with 3 relations.

   $\text{employee (person\_name, street, city)}$
   $\text{works (person\_name, company\_name, salary)}$
   $\text{company (company\_name, city)}$

Give an expression in the relational algebra to express each of the following queries:

a. Find the names of all employees who live in city “Miami”.

b. Find the names of all employees whose salary is greater than $100,000.

c. Find the names of all employees who live in “Miami” and whose salary is greater than $100,000.

8. Consider the following schema for a bank database.

   $\text{branch (branch\_name, branch\_city, assets)}$
   $\text{customer (customer\_name, customer\_street, customer\_city)}$
   $\text{loan (loan\_number, branch\_name, amount)}$
   $\text{borrower (customer\_name, loan\_number)}$
   $\text{account (account\_number, branch\_name, balance)}$
   $\text{depositor (customer\_name, account\_number)}$

Give an expression in the relational algebra for each of the following queries:

a. Find the names of all branches located in “Chicago”.

b. Find the names of all borrowers who have a loan in branch “Downtown”. 
Exercise 2A: Simple SQL

1. Write the following queries in SQL, using the university schema.

a. Find the titles of courses in the Comp. Sci. department that have 3 credits.

b. Find the IDs of all students who were taught by an instructor named Dale; make sure there are no duplicates in the result.

c. Find the highest salary of any instructor.

d. Find all instructors earning the highest salary (there may be more than one with the same salary).

e. Find the enrollment of each section that was offered in Fall 2009 (i.e., find, for each pair of course_id and section_id, the number of students that take it in year 2009 and in Fall semester). How can you take into account the case where a section does not have any students taking it then?

f. Find the maximum enrollment, across all sections, in Fall 2009.

g. Find the sections (course_id, section_id) that had the maximum enrollment in Fall 2009.

2. Suppose you are given a relation grade_points (grade, points), which provides a conversion from letter grades in the takes relation to numeric scores; for example an “A” grade could be specified to correspond to 4 points, an “A−” to 3.7 points, a “B+” to 3.3 points, a “B” to 3 points, and so on. The grade points earned by a student for a course offering (section) is defined as the number of credits for the course multiplied by the numeric points for the grade that the student received.

Given the above relation (download and run the file grade_points.sql), and the university schema, write each of the following queries in SQL. You can assume for simplicity that no takes tuple has the null value for grade.

a. Find the total grade-points earned by the student with ID 1000, across all courses taken by the student.

b. Find the grade-point average (GPA) for the above student, that is, the total grade-points divided by the total credits for the associated courses.

c. Find the ID and the grade-point average of every student.
3. Suppose that we have a relation marks(ID, score) and we wish to assign grades to students based on the score as follows:

- grade F if score < 40,
- grade C if 40 ≤ score < 60,
- grade B if 60 ≤ score < 80, and
- grade A if 80 ≤ score.

Create the relation marks and insert some sample values. Write SQL queries to do the following:

a. Display the grade for each student, based on the marks relation.

b. Find the number of students with each grade.

4. The SQL like operator is case sensitive, but the lower() function on strings can be used to perform case insensitive matching. To show how, write a query that finds departments whose names contain the string "sci" as a substring, regardless of the case.

5. Consider the following insurance database schema, where the primary keys are underlined.

```
person (driver_id, name, address)
car (license, model, year)
accident (report_number, date, location)
owns (driver_id, license)
participated (driver_id, car, report_number, damage_amount)
```

Construct the following SQL queries for this relational database.

a. Find the total number of people who owned cars that were involved in accidents in 2009.

6. Consider the following bank database schema, where the primary keys are underlined.

```
branch(branch_name, branch_city, assets)
customer (customer_name, customer_street, customer_city)
loan (loan_number, branch_name, amount)
borrower (customer_name, loan_number)
account (account_number, branch_name, balance)
depositor (customer_name, account_number)
```

Construct the following SQL queries for this relational database.
a. Find all customers of the bank who have an account but not a loan.

b. Find the names of all customers who live on the same street and in the same city as “Smith”.

c. Find the names of all branches with customers who have an account in the bank and who live in city called “Harrison”.

7. Consider the following employee database schema, where the primary keys are underlined.

\[ \begin{align*}
  employee (employee\_name, street, city) \\
  works (employee\_name, company\_name, salary) \\
  company (company\_name, city) \\
  manages (employee\_name, manager\_name)
\end{align*} \]

Give an expression in SQL for each of the following queries.

a. Find the names and cities of residence of all employees who work for First Bank Corporation.

b. Find the names, street address, and cities of residence of all employees who work for First Bank Corporation and earn more than $10,000.

c. Find all employees in the database who do not work for First Bank Corporation.

d. Find all employees in the database who earn more than each employee of Small Bank Corporation.

e. Assume that the companies may be located in several cities. Find all companies located in every city in which Small Bank Corporation is located.

f. Find the company that has the most employees.

g. Find those companies whose employees earn a higher salary, on average, than the average salary at First Bank Corporation.
Exercise 2B: Simple SQL

1. Write the following inserts, deletes or updates in SQL, using the university schema.

   a. Increase the salary of each instructor in the ‘Comp. Sci.’ department by 10%.

   b. Try to insert every student whose tot_cred attribute is greater than 100 as an instructor in the same department, with a salary of $10,000. Explain what happens. What about having salary $30,000?

2. Consider the following insurance database schema, where the primary keys are underlined.

   \[
   \text{person (driver_id, name, address)} \\
   \text{car (license, model, year)} \\
   \text{accident (report_number, date, location)} \\
   \text{owns (driver_id, license)} \\
   \text{participated (driver_id, car, report_number, damage_amount)}
   \]

   Construct the following SQL queries for this relational database.

   a. Add a new accident to the database; assume any values for required attributes.

   b. Delete the Mazda belonging to “John Smith”.

3. Consider the following relational database schema.

   \[
   \text{employee (employee_name, street, city)} \\
   \text{works (employee_name, company_name, salary)} \\
   \text{company (company_name, city)} \\
   \text{manages (employee_name, manager_name)}
   \]

   Give an expression in SQL for each of the following queries.

   a. Modify the database so that Jones now lives in Newtown.

   b. Give all managers of First Bank Corporation a 10-percent raise unless the salary becomes greater than $100,000; in such cases, give only a 3-percent raise.
Exercise 3: Intermediate SQL

1. Write the following queries in SQL:

a. Display a list of all instructors, showing their ID, name, and the number of sections that they have taught. Make sure to show the number of sections as 0 for instructors who have not taught any section. Your query should use an outerjoin, and should not use scalar subqueries.

b. Write the same query as above, but using a scalar subquery, without outerjoin.

c. Display the list of all course sections offered in Spring 2010, along with the names of the instructors teaching the section. If a section has more than one instructor, it should appear as many times in the result as it has instructors. If it does not have any instructor, it should still appear in the result with the instructor name set to “−”. [Hint: use function coalesce(name, ‘-’)];

d. Display the list of all departments, with the total number of instructors in each department, without using scalar subqueries. Make sure to correctly handle departments with no instructors.

2. Show how to define the view student_grades (ID, GPA) giving the grade-point average of each student, based on the following query (used in Exercise 2A); recall that we used a relation grade_points(grade, points) to get the numeric points associated with a letter grade.

3. Complete the SQL DDL definition of the university database of the following schema to include the relations student, takes, advisor, and prereq.

4. Consider the relational database with the following schema. Give an SQL DDL definition of this database. Identify referential-integrity constraints that should hold, and include them in the DDL definition.

   employee (employee_name, street, city)
   works (employee_name, company_name, salary)
   company (company_name, city)
   manages (employee_name, manager_name)

5. For the manages relation of the previous exercise, add a foreign-key clause to require that every manager also be an employee. Explain exactly what happens when a tuple in the relation employee is deleted.
Exercise 4: ER Modeling

1. Construct an E-R diagram for a car insurance company whose customers own one or more cars each. Each car has associated with it zero to any number of recorded accidents. Each insurance policy covers one or more cars, and has one or more premium payments associated with it. Each payment is for a particular period of time, and has an associated due date, and the date when the payment was received.

2. Consider a database used to record the marks that students get in different exams of different course offerings (sections). Construct an E-R diagram that models exams as entities, and uses a ternary relationship, for the database.

3. Develop the E-R diagram for an online bookstore. Books are identified by ISBN codes, written by authors and sold at a given price. Books are published by publishers. The online bookstore has several warehouses with a given stock-level for each title, whereas each book may be present (at different quantities) in different warehouses. Customers create accounts (assume that user_id is their email) and in each session they make a shopping basket.

4. Consider in the following figure, the alternative representation of a ternary relationship (a) using binary relationships (b or c).

   ![Diagram](image)

   a. Show a simple instance of E, A, B, C, RA, RB, and RC that cannot correspond to any instance of A, B, C, and R.
b. Modify the E-R diagram of Figure 7.27b to introduce constraints that will guarantee that any instance of E, A, B, C, RA, RB, and RC that satisfies the constraints will correspond to an instance of A, B, C, and R.

5. A weak entity set can always be made into a strong entity set by adding to its attributes the primary-key attributes of its identifying entity set. Outline what sort of redundancy will result if we do so.

6. The advisor relationship is one-to-many (see figure). Suppose the advisor relationship were one-to-one. What extra constraints are required on the relation advisor to ensure that the one-to-one cardinality constraint is enforced?

7. Consider a many-to-one relationship R between entity sets A and B. Suppose the relation created from R is combined with the relation created from A. In SQL, attributes participating in a foreign key constraint can be null. Explain how a constraint on total participation of A in R can be enforced using not null constraints in SQL.

8. The following figure shows a lattice structure of generalization and specialization (attributes not shown). For entity sets A, B, and C, explain how attributes are inherited from the higher-level entity sets X and Y. Discuss how to handle a case where an attribute of X has the same name as some attribute of Y.

9. Design a database for a world-wide package delivery company (e.g., DHL or FedEX). The database must be able to keep track of customers (who ship items) and customers (who receive items); some customers may do both. Each package must be identifiable and trackable, so the database must be able to store the location of the package and its history of locations. Locations include trucks, planes, airports, and warehouses. Your design should include an E-R diagram, a set of relational schemas, and a list of constraints, including primary-key and foreign-key constraints.
10. Design a database for an airline. The database must keep track of customers and their reservations, flights and their status, seat assignments on individual flights, and the schedule and routing of future flights. Your design should include an E-R diagram, a set of relational schemas, and a list of constraints, including primary-key and foreign-key constraints.