**Homework Assignment 1**

Student ID: 21600635

Name: Bomoon Jung

**1. (5 pt.) What are the major disadvantages of keeping organizational information in a file-processing system?**

- Data redundancy and inconsistency, difficulty in accessing data, data isolation, hard to sustain integrity, concurrency problems, security problems, atomicity problems

**2. (1 pt. per blank) Fill in the blanks.**

(a) Underlying the structure of a database is the (**data – model**): a collection of conceptual tools for describing data, data relationships, data semantics, and consistency constraints.

(b) In the relational model, data are represented in the form of (**table**). A table has multiple (**columns**); Each column has a unique (**name**). Each (**row**) of the table represents one piece of information.

(c) The collection of information stored in the database at a particular moment is called a/an (**instance**). The overall design of the database is called the (**schema**).

(d) A database system provides a/an (**DDL**) to specify the database schema and a/an (**DML**) to express database queries and updates. In practice, these are not two separate languages; instead, they simply form parts of a single language.

(e) A/An (**query**) is a statement requesting the retrieval of information.

(f) A/An (**transaction** **manager**) ensures that the database remains in a consistent state despite system failures, and that concurrent transaction executions proceed without conflicts.

(g) A/An (**file manager**) manages the allocation of space on disk storage and the data structures used to represent information stored on disk.

(h) A/An (**buffer manager**) is responsible for fetching data from disk storage into main memory, and deciding what data to cache in main memory.

(i) The (**null value**) is a special value that signifies the value is unknown or does not exist

**3. (5 pt.; Exercise problem 1.7) List four significant differences between a file-processing system and a DBMS.**

a. Data redundancy and inconsistency: As opposed to DMBS, file-processing system tries to manage by storing unnecessary redundant information in different files.

b. Difficulty in accessing data: Thus, when we get new task, manipulating prior programs take a lot of time.

c. Data isolation: As opposed to DMBS, data of file-processing system is scattered and stored in each file in different formats.

d. Security problems: Information consistency must be guaranteed, but file-processing systems cannot.

**4. (3 pt. each; Exercise problems 2.6, 2.8, and 2.14) Consider the employee database of Figure 2.17 (page 60). Give an expression in the relational algebra to express each of the following queries:**

a. Find the name of each employee who lives in city “Miami”.

b. Find the name of each employee whose salary is greater than $100,000.

c. Find the name of each employee who lives in “Miami” and whose salary is greater than $100,000.

d. Find the ID and name of each employee who does not work for “BigBank”.

e. Find the ID and name of each employee who earns at least as much as every employee in the database.

f. Find the ID, name, and city of residence of each employee who works for “BigBank”.

g. Find the ID, name, street address, and city of residence of each employee who works for “BigBank” and earns more than $10,000.

h. Find the ID and name of each employee in this database who lives in the same city as the company for which s/he works.

**5. (3 pt. each; Exercise problem 2.18) Write the following queries in relational algebra, using the university schema (Figure 2.8, page 46).**

a. Find the ID and name of each instructor in the Physics department.

b. Find the ID and name of each instructor in a department located in the building “Watson”.

c. Find the ID and name of each student who has taken at least one course in the “Comp. Sci.” department.

d. Find the ID and name of each student who has taken at least one course section in the year 2018.

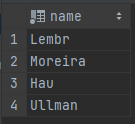
e. Find the ID and name of each student who has not taken any course section in the year 2018.

**Warming up with MySQL**: Launch and access the MySQL databases distributed with the class virtual machine. Below uses the **“university”** database (NOT university\_small), which shares the same schemas with the database used in the lectures but contains a larger set of data records collected within a different period of time.

**6. (5 pt. each) Find the answers to the following questions and provide the SQL queries showing how you find them. All queries should be complete to obtain the listed answers solely by themselves.**

a. List all instructor names in the Accounting department.

Answer:



SQL Query to obtain your answer:

SELECT name FROM instructor  
WHERE dept\_name = 'Accounting';

b. How many students are in the Statistics department?

Answer:



SQL Query to obtain your answer:

SELECT COUNT(ID) as count FROM student  
WHERE dept\_name = 'Statistics';

c. How many unique student names are in the Astronomy department?

Answer:

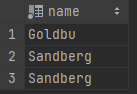


SQL Query to obtain your answer:

SELECT COUNT(DISTINCT name) as count FROM student  
WHERE dept\_name = 'Astronomy';

d. Find all students who have “db” as a substring in their name.

Answer:



SQL Query to obtain your answer:

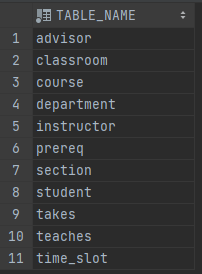
SELECT name FROM student  
WHERE name LIKE '%db%';

**7. (3 pt.) List the names of all tables that the “university” database has.**

Answer:

SELECT table\_name FROM information\_schema.TABLES  
WHERE table\_schema = 'university';

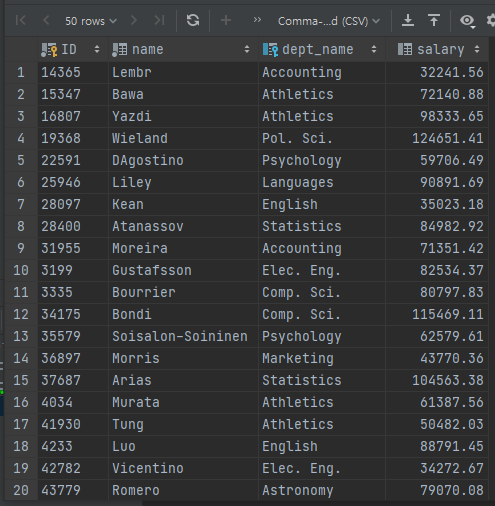
SQL Query to obtain your answer:



**8. (4 pt.) Execute and explain the differences among the results of the following queries.**

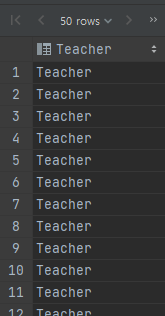
(i) SELECT \* FROM instructor;

All attributes and tuples(50) of Instructor are displayed.



(ii) SELECT 'Teacher' FROM instructor;

Teachers are displayed as many as the number of tuples(50) of the Instructor, and the attribute name is also Teacher.



(iii) SELECT 'Teacher';

Attribute is ‘Teacher’ and only one tuple appears.



(iv) SELECT \*, 'Teacher' FROM instructor;

All attributes and tuples(50) of the Instructor are printed, and Teachers are printed as many as the number of tuples at the end, and the table whose attribute name is also Teacher is printed.

