

Assignment 1 Solution

Name: Yanjun Chen, PSU ID: yfc5289

Part I. DBMS Overview

1. Include a DBMS in application stack:

- advantages: Efficient access; data integrity and security.
- disadvantages: applying DBMS will increase the overall complexity and requirement of the resources.

Manually implement the necessary data access features:

- advantages: Fewer resources (compare with DBMS); Flexibility.
- disadvantages: Less functionalities (compare with DBMS), which makes it less scalable.

2. It means the database maintains its structure after any data transactions. In another words, the database must still meets all defined rules of this database after any transactions.

3. This is an example of physical data independence.

Because for a user's persepective, they will saw another kind of layout of the data they want to see. Instead of how the data is structured in the database, such as an array or search tree, in a physical place.

4. The best type I choose will be VARCHAR.

Because first of all, the ZIP codes will not contain any dots inside. So it cannot be FLOAT. Secondly, as the ZIP code could start with number 0. It cannot be represented in INT.

By the rules of US ZIP codes, we can either use the 5-digit ZIP code such as 16801 or use the "ZIP + 4" version for more accuracy, which will also increase the length. For the case of US ZIP codes, as we didn't set any specific rules of which version we are going to use, I think VARCHAR could be a better choice.

Part II. Relational Model

1. - ID: INT

- FNAME: VARCHAR(20)
- LNAME: VARCHAR(20)
- MNAME: VARCHAR(20)
- SSN: CHAR(9)
- BDAY: VARCHAR(10)
- DEPT: VARCHAR (4)
- COLLEGE: VARCHAR(20)

Notes: all domains are considered only by the data in the given table. And the relation name / table name would be "faculty".

2. - {SSN}
- {ID, SSN}
- {ID, FNAME, LNAME}
- {FNAME, SSN}
- {LNAME, SSN}

NOTES: all sets are only suitable for the given table above.

3. - {ID, LNAME}
- {FNAME, LNAME}
- {LNAME, SSN}

4. The primary key I choose for this question would be {FNAME, LNAME, SSN}.

Because according to the given table, the major and most unique difference between the records is their SSN. They could have same other attributes such as ID, MNAME, DEPT and COLLEGE. But their SSN are unique. So I also think any kinds of 3-attribute sets with SSN would be fine.

5. SQL statement:

```
INSERT INTO faculty (ID, FNAME, LNAME, MNAME, SSN, BDAY, DEPT, COLLEGE)
VALUES (15, 'John', 'Doe', NULL, '987364830', NULL, 'ECE', 'ENG');
```

6. SQL statement:

```
DELETE FROM faculty
WHERE COLLEGE = 'SCI';
```

7. SQL statement:

```
UPDATE faculty
SET BDAY = '12/11/2001'
WHERE FNAME = 'John' AND LNAME = 'Doe';
```

8. SQL statement:

```
ALTER TABLE faculty
DROP BDAY;
```

9. The deletion of the (ENG, CS) record in the table "Departments" will not delete the rows in the "faculty" table. Instead, values for column "COLLEGE" and "DEPT" in the "faculty" table for rows which COLLEGE = 'ENG' and DEPT = 'CS' will be set to 'NULL'.