

How does the Redundant Information comes in a Poor Design?

What problems do come due to Redundant data?

Compare the two relations EMP, DEPT with EMP-DEPT. In EMP-DEPT, the attribute values pertaining to a particular department (Dnumber, Dname, Dmgr-SSN) are repeated for every employee who works for that department. ~~In~~ While in case of DEPT the department no, name, and mgr-SSN comes only once and is not repeated.

EMP-DEPT

Redundancy

EMP_DEPT

Ename	Ssn	Bdate	Address	Dnumber	Dname
Smith, John B.	123456789	1965-01-09	731 Fondren, Houston, TX	5	Research
Wong, Franklin T.	333445555	1955-12-08	638 Voss, Houston, TX	5	Research
Zelaya, Alicia J.	999887777	1968-07-19	3321 Castle, Spring, TX	4	Administration
Wallace, Jennifer S.	987654321	1941-06-20	291 Berry, Bellaire, TX	4	Administration
Narayan, Ramesh K.	666884444	1962-09-15	975 FireOak, Humble, TX	5	Research
English, Joyce A.	453453453	1972-07-31	5631 Rice, Houston, TX	5	Research
Jabbar, Ahmad V.	987987987	1969-03-29	980 Dallas, Houston, TX	4	Administration
Borg, James E.	888665555	1937-11-10	450 Stone, Houston, TX	1	Headquarters

Redundancy

Redundancy

EMP_PROJ

Ssn	Pnumber	Hours	Ename	Pname	Plocation
123456789	1	32.5	Smith, John B.	ProductX	Bellaire
123456789	2	7.5	Smith, John B.	ProductY	Sugarland
666884444	3	40.0	Narayan, Ramesh K.	ProductZ	Houston
453453453	1	20.0	English, Joyce A.	ProductX	Bellaire
453453453	2	20.0	English, Joyce A.	ProductY	Sugarland
333445555	2	10.0	Wong, Franklin T.	ProductY	Sugarland
333445555	3	10.0	Wong, Franklin T.	ProductZ	Houston
333445555	10	10.0	Wong, Franklin T.	Computerization	Stafford
333445555	20	10.0	Wong, Franklin T.	Reorganization	Houston
999887777	30	30.0	Zelaya, Alicia J.	Newbenefits	Stafford
999887777	10	10.0	Zelaya, Alicia J.	Computerization	Stafford
987987987	10	35.0	Jabbar, Ahmad V.	Computerization	Stafford
987987987	30	5.0	Jabbar, Ahmad V.	Newbenefits	Stafford
987654321	30	20.0	Wallace, Jennifer S.	Newbenefits	Stafford
987654321	20	15.0	Wallace, Jennifer S.	Reorganization	Houston
888665555	20	Null	Borg, James E.	Reorganization	Houston

Figure 14.4

Sample states for EMP_DEPT and EMP_PROJ resulting from applying NATURAL JOIN to the relations in Figure 14.2. These may be stored as base relations for performance reasons.

In case of EMP only the D number is repeated which is a foreign key and it defines the Employee.

Also in case of EMP-Proj you can see the data in Page (16) that the project number, name and location is repeated. So, ~~the~~ the repetition occurs if the design is poor as illustrated by the data.

The redundant data creates/results in problems/anomalies that are detailed here below:

1. INSERTION Anomaly

i) To insert a new employee tuple into EMP-Dept, we must include either the attribute values for the department that the employee works for, or NULLs (if the employee is not allocated a department or is not working for a department). For example, to insert a new tuple for an employee who works in department number 5, we must enter all the attributes values of the department repeatedly again and again consistently.

and an error may come due to data entry. Imagine, ~~the~~ one department(s) has 100000 employee and we have to enter dept-name, mgr-ssn and other things 10000 times. With the operator may create or inadvertently may enter wrong data. So the data may become inconsistent.

- ii) It is difficult to enter a new department that has no employees as yet. While in case of Dept alone the department may be created without any employee which is to start is logically correct.

This also violates entity integrity for EMP-DEPT because SSN is its primary key.

2. DELETION Anomaly

If we delete an employee who is the last employee of the department, department data will also be lost. ~~and~~ If we have separate EMP and DEPT

Relations, such a problem will not happen.

Modification / change / update Anomaly

In EMP-dept, if we change the values of one of the attributes of a particular department - say manager of department 5 - we must update the tuples of all employees who work in that department, the database will become inconsistent. If we do not update some records/tuples the same department will be shown to have different values for manager in different employee, which would be wrong.