

CS443 - Lab 3

Question 1:

Suppose that our database has the following table.

Person

| Con ID | Con Name | Con Pop | Con Size | State code | State Name | State Rgn | State Size | State Pop | Cty Code | Cty Name | Cty Size | Per SSN | Per Name | Per Age | Per DofB | Per Add |
|--------|----------|---------|----------|------------|------------|-----------|------------|-----------|----------|----------|----------|---------|----------|---------|----------|---------|
|--------|----------|---------|----------|------------|------------|-----------|------------|-----------|----------|----------|----------|---------|----------|---------|----------|---------|

Field Explanation:

Con: Stands for Country
Pop: Stands for population
Rgn: Stands for region (like west, east, central, etc.)
Cty: Stands for City
Per: Stands for Person
DofB: Stands for date of birth
Add: Stands for Address

It is assumed that

- Every country in the world has a different country ID
- Every city in the world has a different city code
- Every state in the world has a different state code, and
- Every person in the world has a different SSN
- Every person in the world has only one citizenship and has only one address

Other Assumptions:

- There is no village, county, area, etc. A country consists of several states and each state has several cities

1) Based on the above assumptions, what do you choose to be the primary key of Person table? Why?

- In the person table, using the assumption that every person in the world has a different SSN, we will use the Person's SSN to be primary key of the Person table because we know that there will not be any duplication in this field. The other fields would most likely have multiple duplicates, such as DofB and Age.

2) Explain the anomalies exist in the Person table. Choose only one example of insert anomaly, one example of delete anomaly and one example of update anomaly. Note that update does not mean adding or deleting records. It only refers to modifications of values in some rows of the table.

- The Person table has multiple anomalies.
- An example of the insert anomaly would be trying to add information about a new country, state, or city. In order to add new information about these other fields, it would require forcing a new person from that specific country, state, or city to input that new information. If we were to have a city with no one from it, then the person SSN would be null, and we cannot have a null as a primary key.

- An example of the deletion anomaly would be removing a person from the table, but if they are the only person within the table from that country, state, or city, then any information that exist of that country, state, or city would be deleted as well.
- An example of the modification anomaly would be adding the information of a person who was just born in a country, state, or city that already exists in the table. Since the population has just been increased, we would have to modify the population field for all the other rows that contain the information of people that live in same country, state, or city.

3) Normalize the table; create as many as tables necessary such that all new tables are in third normal form. All the transitive and derived dependencies must be removed..

Person(SSN, Name, Age, DofB, Add)
 City(Code, Name, Size)
 State(Code, name, Rgn, Size, Pop)
 Country(ID, Name, Pop, Size)

These are the strong entities, but they do have some dependencies in Person, for example, we can remove Age because if we know their date of birth, we can calculate the persons age.

The city table, there are no dependencies.

In the state table, we can derive the size of the state from all the cities' sizes, we don't need to know the state size.

In the country table, we can derive the size from all the cities, just as we did for the states. This is not the only redundancy here. We can derive the population of the country by adding the population of all the states that are in the country.

Update Tables:
 Person (SSN, Name, DofB, Add)
 City (Code, Name, Size)
 State (Code, Name, Rgn, Pop)
 Country (ID, Name)

4) Draw your ERD based on fully normalized table (Reverse Engineering).

