CS411 Project1 Stage2

Team028-TBD

qilong2, bingjun3, jiayu9, hz75

1 ER Diagram

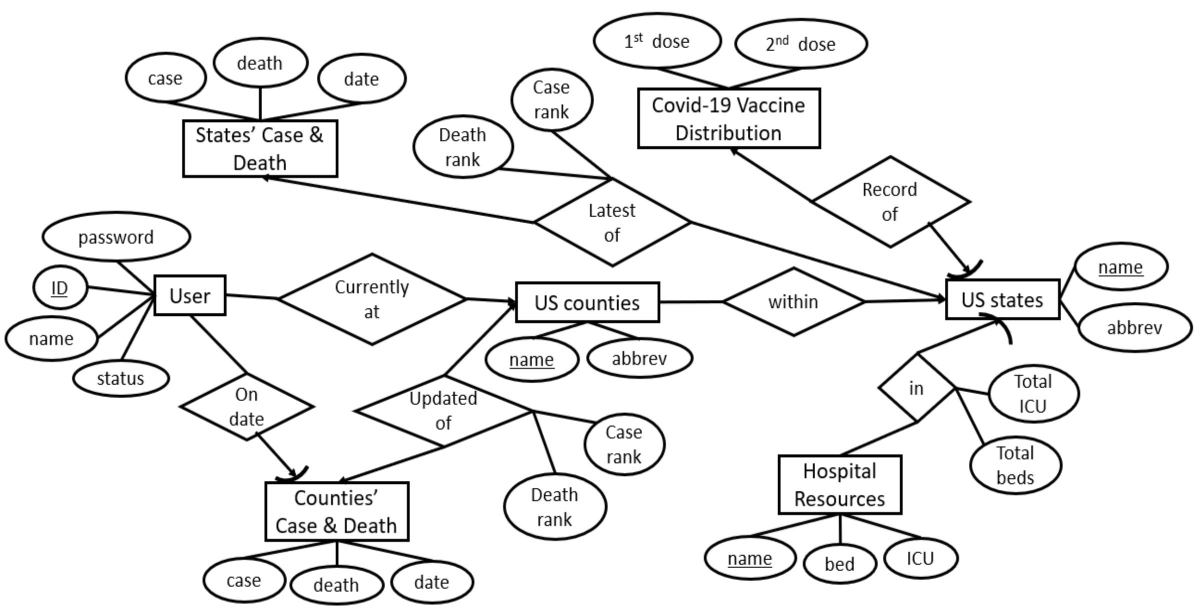

Figure 1: ER diagram.

2 Describe ER Diagram

2.1 Overview

Currently, 7 entities in total are taken into construction:

- States: name, abbrev.
- Counties: (state, name), abbrev.
- User: id, name, password, status, state, county.
- StateCaseDeath: name, case, death, date.
- CountyCaseDeath: (state, name), case, death, date.
- Vaccine: state, firDose, secDose.
- Hospital: name, state, bed, ICU.

Among them, Entity User is regarding user login information.
7 relationship are taken into construction:

- User-Counties: connection between user and counties, with foreign key county referencing name in county entity.
  Cardinality: number of users.
- Counties-States: connection between county and state, county has foreign key state referencing name in state entity.
  Cardinality: number of counties.
- States-StateCaseDeath: The record of number of cases and deaths belong to every state, StateCaseDeath has foreign key name referencing name in state entity.
  Cardinality: number of states.
- Counties-CountyCaseDeath: The record of number of cases and deaths belong to every county, CountyCaseDeath has foreign key (state, name) referencing (state, name) in county entity.
  Cardinality: number of counties.
- States-Vaccine: The record of number of vaccines distributed to every state, Vaccine has foreign key state referencing name in state entity.
  Cardinality: number of states.
- States-Hospital: The record of hospital resources in every state, Vaccine has foreign key state referencing name in state entity.
  Cardinality: number of states.

- User-CountyCaseDeath: the user status will update relevant county case and death record, user has foreign key county referencing county in CountyCaseDeath entity. Cardinality: number of users.

2.2 Assumptions

For each entity, we made these assumptions:

- States: includes all states of US with unique name.
- Counties: includes all counties of US with unique (state, county name).
- User: each user has a unique ID.
- StateCaseDeath: only store the total amount to the latest updated date.
- CountyCaseDeath: only store the total amount to the latest updated date.
- Vaccine: only store the total amount to the latest updated date.
- Hospital: only store the total amount to the latest updated date.

For each relationship, we made these assumptions:

- User-Counties: (many-1) each user can only be currently at one county.
- Counties-States: (many-1) each county can only be within one state.
- States-StateCaseDeath: (1-1) exactly one state matches exactly one row of data recording latest number of cases and deaths.
- Counties-CountyCaseDeath: (1-1) exactly one county matches exactly one row of data recording lastly updated number of cases and deaths.
- States-Vaccine: (1-1) exactly one state matches exactly one row of data recording number of vaccines distributed to that state, state with no vaccine data available is acceptable.
- States-Hospital: (1-1) exactly one state matches exactly one row of data recording hospital resources in that state, state with no hospital resources data available is acceptable.

- User-CountyCaseDeath: (many-1) each user can only contributed to data of one county in a day.

3 Normalize Database

- States: name→abbrev.
- Counties: (state, name)→abbrev.
- User: ID→password, ID→name, ID→status, name→password, name→ID, ID→state, ID→county.
- StateCaseDeath: name→case, name→death, name→date.
- CountyCaseDeath: (state, name)→case, (state, name)→death, (state, name)→date.
- Vaccine: state→firDose, state→secDose.
- Hospital: name→state, name→bed, name→ICU.
- CountyRank: (state, name)→deathRank, (state, name)→CaseRank.
- StateRank: state→deathRank, state→CaseRank.
- StateHospital: state→totalICU, state→totalBeds.

3.1 Choice of Method

Method: 3NF
Reason: BCNF may cause the problem of losing important functional dependencies. For example, in our user entity, besides ID→name, ID→password, also name→password, name→ID should be preserved. BCNF may lose such functional dependency, causing trouble in our log in functionality. Therefore, 3NF was selected.

3.2 Process

Firstly, find the functional dependencies in each entity:

- States: name→abbrev.
- Counties: (state, name)→abbrev.
- User: ID→password, ID→name, ID→status, name→password, name→ID, ID→county, ID→state.
- StateCaseDeath: name→case, name→death, name→date.
- CountyCaseDeath: (state, name)→case, (state, name)→death, (state, name)→date.
- Vaccine: state→firDose, state→secDose.
- Hospital: name→state, name→bed, name→ICU.
- CountyRank: (state, name)→deathRank, (state, name)→caseRank.
- StateRank: state→deathRank, state→caseRank.
- StateRank: state→deathRank, state→caseRank.
- StateHospital: state→totalICU, state→totalBeds.

Secondly, apply 3NF on each entity:

3.2.1 States

a according to name→abbrev, let name=A, abbrev=B, then A→B, A+ = (A,B), B+ = B

b name = A is the primary key

Therefore, States satisfies 3NF rule.

3.2.2 Counties

a according to (state, name)$\rightarrow$abbrev, let state=A, name=B, abbrev=C, then (A,B)$\rightarrow$C, A+ = A, B+ = B, C+ = C, (A,B)+ = C

b (state, name) = (A,B) is the primary key

Therefore, Counties satisfies 3NF rule.

3.2.3 User

a according to ID$\rightarrow$password, ID$\rightarrow$name, ID$\rightarrow$status, name$\rightarrow$password, name$\rightarrow$ID, mame$\rightarrow$county, let ID=A, name=B, password=C, status=D, state=E, county=F, then A$\rightarrow$B, A$\rightarrow$C, A$\rightarrow$D, B$\rightarrow$C, B$\rightarrow$A, A+ = (A, B, C, D, E, F), B+ = (A, B, C), C+ = C, D+ = D, E+ = E, F+ = F.

b Since A and B are mutually dependent, the User entity should be segregated into two: A$\rightarrow$B and A$\rightarrow$(C,D), We should rewrite User into User1: ID$\rightarrow$name, User2: ID$\rightarrow$status, ID$\rightarrow$password, ID$\rightarrow$state, ID$\rightarrow$county

3.2.4 StateCaseDeath

a) minimum basis = {name$\rightarrow$case, name$\rightarrow$death, name$\rightarrow$date}
b) relationship: X(name, case, death, date)
c) candidate key: name+ = name, case, death, date.
Therefore, StateCaseDeath satisfies 3NF rule.

3.2.5 CountryCaseDeath

a) minimum basis = {(state, name)$\rightarrow$case, (state, name)$\rightarrow$death, (state, name)$\rightarrow$date}
b) relationship: X((state, name), case, death, date)
c) candidate key: (state, name)+ = (state, name), case, death, date.
Therefore, CountryCaseDeath satisfies 3NF rule.

3.2.6 Vaccine

a) minimum basis = {state$\rightarrow$firDose, state$\rightarrow$secDose}
b) relationship: X(state, firDose, secDose)

c) candidate key: state+ = firDose, secDose
Therefore, Vaccine satisfies 3NF rule.

3.2.7 Hospital

a minimum basis = {name→state, name→bed, name→ICU}

b relationship: X(name, state, bed, ICU)

c candidate key: $name^{+} = name, state, bed, ICU$
Therefore, Hospital satisfy 3NF rule.

3.2.8 CountyRank

a minimum basis = {(state, name)→deathRank, (state, name)→caseRank}

b relationship: X(state, name, deathRank, caseRank)

c candidate key: $(state, name)^{+} = state, name, deathRank, caseRank$
Therefore, CountyRank satisfy 3NF rule.

3.2.9 StateRank

a minimum basis = {state→deathRank, state→caseRank}

b relationship: X(state, deathRank, caseRank)

c candidate key: $state^{+} = state, deathRank, caseRank$
Therefore, StateRank satisfy 3NF rule.

3.2.10 StateHospital

a) minimum basis = {state→totalICU, state→totalBed}
b) relationship: X(state, totalICU, totalBed)
c) candidate key: state+ = state, totalICU, totalBed
Therefore, StateHospital satisfies 3NF rule.

3.3 Logical Design Relational Schema

1. States(name: VARCHAR(20)[PK], abbrev: VARCHAR(3))
2. Counties((state: VARCHAR(20), name: VARCHAR(20))[PK], abbrev: VARCHAR(3))
3. User1(ID: VARCHAR(20)[PK], name: VARCHAR(20))
4. User2(ID: VARCHAR(20)[PK], password: VARCHAR(20), status: VARCHAR(10), state: VARCHAR(20)[FK to Counties.state], county:VARCHAR(20)[FK to Counties.name])
5. StateCaseDeath(name: VARCHAR(20)[PK][FK to States.name], case: INT, death: INT, date: DATE)
6. CountyCaseDeath((state: VARCHAR(20), name: VARCHAR(20))[PK][FK to (Counties.state, Counties.name)], case: INT, death: INT, date: DATE)
7. Vaccine(state: VARCHAR(20) [PK], firDose: INT, secDose: INT)
8. Hospital(name: VARCHAR(20) [PK], state: VARCHAR(20) [FK to States.name], bed: INT, ICU: INT)
9. CountyRank((state: VARCHAR(20), name: VARCHAR(20)) [PK] [FK to (Counties.state, Counties.name)], deathRank: INT, caseRank: INT)
10. StateRank(state: VARCHAR(20) [PK] [FK to States.name], deathRank: INT, caseRank: INT)
11. StateHospital(state: VARCHAR(20) [PK] [FK to States.name], totalICU: INT, totalBed: INT)