



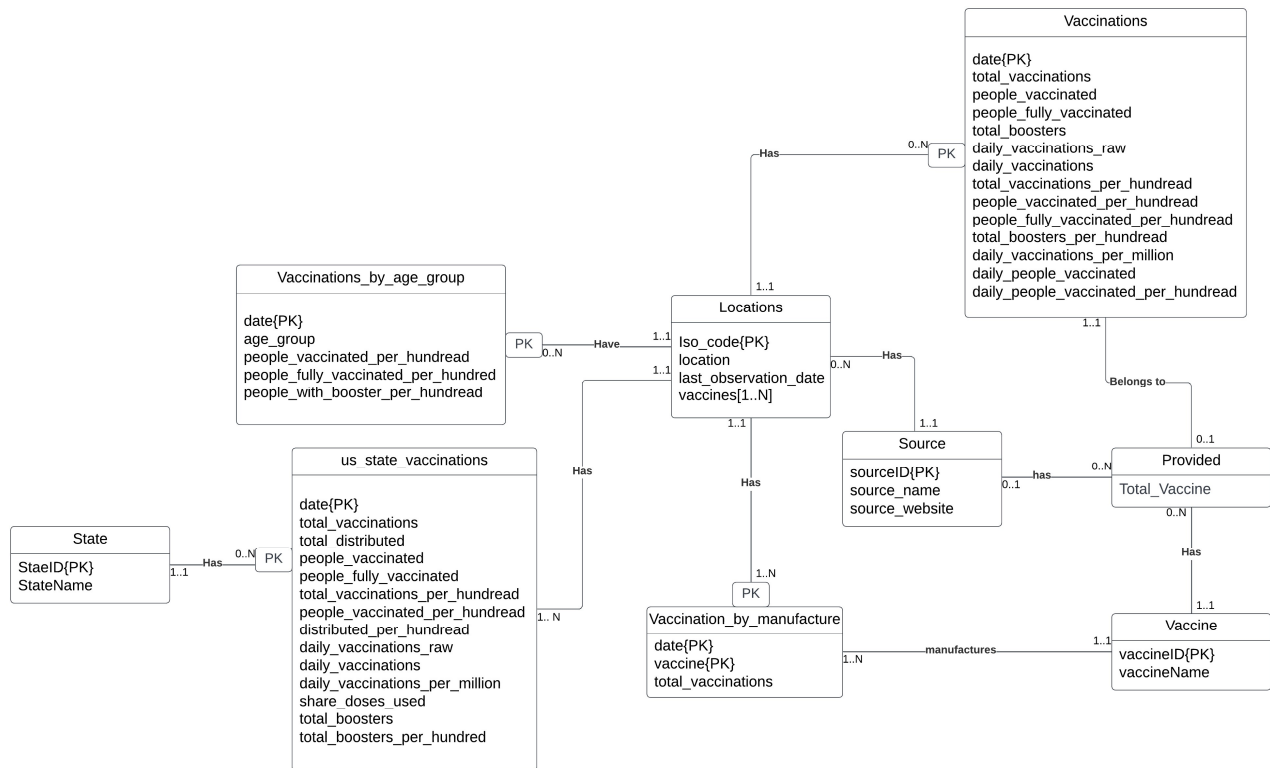
DATABASE CONCEPTS ASSIGNMENT-3

Name: [REDACTED]

StudentID: [REDACTED]

Part B: Designing the Database

- **Entity Relationship Modelling**



Assumptions:

- At one location there can be more than one vaccination occur and vaccination record belongs to only one location.
- Vaccinations record may have more than one vaccine type and some vaccine may be newly introduced and not used by any vaccinations record.
- One vaccine may be manufactured by many manufacturers but a vaccine is not co-manufactured.
- A manufacturer has one location but one location has more than one manufacturer.
- Us_state_vaccinations record belong to one state and one state may have more than one vaccination record.
- One source belong to more than one location and one location have only one source. Some source might not belong to any location.
- Vaccination by age group belongs to one location and one location may have more than one vaccination by age group record.

- **Mapping an ER Model to a Relational Database Schema**

Step-1: Strong Entities

Convert all the entities having primary key directly into relation with all the attributes as it is.

Resultant Schema:

Location(iso_code, location, last_observation_date, vaccines[0...N])

Source(sourceID, source_name, source_website)

State(stateID, stateName)

Vaccine(vaccineID, vaccineName)

Step-2: Weak Entities

Resultant Schema:

Vaccination_by_age_group(date, iso_code*, age_group, people_vaccinated_per_hundread, people_fully_vaccinated_per_hundread, people_with_booster_per_hundread)

Us_state_vaccinations(date, stateID *, total_vaccinations, total_distributed, people_vaccinated, people_fully_vaccinated, total_vaccinations_per_hundread, people_vaccinated_per_hundread, distributed_per_hundread, daily_vaccinations_raw, daily_vaccinations, daily_vaccinations_per_million, share_doses_used, total_boosters, total_boosters_per_hundred)

Vaccination_by_manufacture(date, vaccine, iso_code*, total_vaccinations)

Vaccinations(date, iso_code*, total_vaccinations, people_vaccinated, people_fully_vaccinated, total_boosters, daily_vaccinations_raw, daily_vaccinations, total_vaccinations_per_hundread, people_vaccinated_per_hundread, people_fully_vaccinated_per_hundread, total_boosters_per_hundread, daily_vaccinations_per_million, daily_people_vaccinated, daily_people_vaccinated_per_hundread)

Provided(date*, iso_code*, Total_vaccine)

Step-3: One-to-one Relationships

No actions required.

Step-4: One-to-many Relationships

- One side is source and many side is location so copy primary key from source to location.
- One side is location and many side is Us_state_vaccinations so copy primary key from location to Us_state_vaccinations.
- One side is vaccine and many side is vaccination by manufacture so copy primary key from vaccine to vaccination by manufacture.

Resultant Schema:

Location(iso_code, location, last_observation_date, **sourceID***, vaccines[0...N])

Source(sourceID, source_name, source_website)

State(stateID, stateName)

Vaccine(vaccineID, vaccineName)

Vaccination_by_age_group(date, iso_code*, age_group, people_vaccinated_per_hundred, people_fully_vaccinated_per_hundred, people_with_booster_per_hundred)

Us_state_vaccinations(date, stateID*, **iso_code***, total_vaccinations, total_distributed, people_vaccinated, people_fully_vaccinated, total_vaccinations_per_hundred, people_vaccinated_per_hundred, distributed_per_hundred, daily_vaccinations_raw, daily_vaccinations, daily_vaccinations_per_million, share_doses_used, total_boosters, total_boosters_per_hundred)

Vaccination_by_manufacture(date, vaccine, iso_code*, **vaccineID***, total_vaccinations)

Vaccinations(date, iso_code*, total_vaccinations, people_vaccinated, people_fully_vaccinated, total_boosters, daily_vaccinations_raw, daily_vaccinations, total_vaccinations_per_hundred, people_vaccinated_per_hundred, people_fully_vaccinated_per_hundred, total_boosters_per_hundred, daily_vaccinations_per_million, daily_people_vaccinated, daily_people_vaccinated_per_hundred)

Provided(date*, iso_code*, Total_vaccine, vaccineID*, sourceID*)

Step-5: Many-to-many Relationships

No actions required.

Step-6: Multi-valued Attributes

- Here Location has multivalued attribute vaccine, so add new relation location_vaccine with primary key as a composition of owner relation's primary key and multivalued attribute value itself.

Location_vaccine(iso_code*,vaccineID*)

Resultant Schema:

Location(iso_code, location, last_observation_date, sourceID*)

Source(sourceID, source_name, source_website)

State(stateID, stateName)

Vaccine(vaccineID, vaccineName)

Vaccination_by_age_group(date, iso_code*, age_group, people_vaccinated_per_hundred, people_fully_vaccinated_per_hundred, people_with_booster_per_hundred)

Us_state_vaccinations(date, stateID *, iso_code*, total_vaccinations, total_distributed, people_vaccinated, people_fully_vaccinated, total_vaccinations_per_hundred, people_vaccinated_per_hundred, distributed_per_hundred, daily_vaccinations_raw, daily_vaccinations, daily_vaccinations_per_million, share_doses_used, total_boosters, total_boosters_per_hundred)

Vaccination_by_manufacture(date, vaccine, iso_code*, vaccineID*, total_vaccinations)

Vaccinations(date, iso_code*, total_vaccinations, people_vaccinated, people_fully_vaccinated, total_boosters, daily_vaccinations_raw, daily_vaccinations, total_vaccinations_per_hundred, people_vaccinated_per_hundred, people_fully_vaccinated_per_hundred,

total_boosters_per_hundread, daily_vaccinations_per_million,
daily_people_vaccinated, daily_people_vaccinated_per_hundread)

Provided(vaccineID*, date*, iso_code*, Total_vaccine, sourceID*)

Location_vaccine(iso_code*, vaccineID*)

Step-7: Map higher-degree relationships

No actions required.

Final Schema:

Location(iso_code, location, last_observation_date, sourceID*)

Source(sourceID, source_name, source_website)

State(stateID, stateName)

Vaccine(vaccineID, vaccineName)

Vaccination_by_age_group(date, iso_code*, age_group,
people_vaccinated_per_hundread, people_fully_vaccinated_per_hundread,
people_with_booster_per_hundread)

Us_state_vaccinations(date, stateID *, iso_code*, total_vaccinations,
total_distributed, people_vaccinated, people_fully_vaccinated,
total_vaccinations_per_hundread, people_vaccinated_per_hundread,
distributed_per_hundread, daily_vaccinations_raw, daily_vaccinations,
daily_vaccinations_per_million, share_doses_used, total_boosters,
total_boosters_per_hundred)

Vaccination_by_manufacture(date, vaccine, iso_code*, vaccineID*,
total_vaccinations)

Vaccinations(date, iso_code*, total_vaccinations, people_vaccinated,
people_fully_vaccinated, total_boosters, daily_vaccinations_raw,
daily_vaccinations, total_vaccinations_per_hundread,
people_vaccinated_per_hundread, people_fully_vaccinated_per_hundread,
total_boosters_per_hundread, daily_vaccinations_per_million,
daily_people_vaccinated, daily_people_vaccinated_per_hundread)

Provided(vaccineID*, date*, iso_code*, Total_vaccine, sourceID*)

Location_vaccine(iso_code*,vaccineID*)

- **Functional Dependencies & Normalization**

- Location**

- iso_code -> location, last_observation_date, sourceID

- Source**

- sourceID -> source_name, source_website

- State**

- stateID -> stateName

- Vaccine**

- vaccineID -> vaccineName

- Vaccination_by_age_group**

- date, iso_code, age_group -> people_vaccinated_per_hundred, people_fully_vaccinated_per_hundred, people_with_booster_per_hundred

- Us_state_vaccinations**

- date, stateID, iso_code -> total_vaccinations, total_distributed, people_vaccinated, people_fully_vaccinated, total_vaccinations_per_hundred, people_vaccinated_per_hundred, distributed_per_hundred, daily_vaccinations_raw, daily_vaccinations, daily_vaccinations_per_million, share_doses_used, total_boosters, total_boosters_per_hundred

- Vaccination_by_manufacture**

- date, iso_code, vaccineID -> vaccine, total_vaccinations

- Vaccinations**

- date, iso_code -> total_vaccinations, people_vaccinated, people_fully_vaccinated, total_boosters, daily_vaccinations_raw, daily_vaccinations, total_vaccinations_per_hundred,

people_vaccinated_per_hundread, people_fully_vaccinated_per_hundread,
total_boosters_per_hundread, daily_vaccinations_per_million,
daily_people_vaccinated, daily_people_vaccinated_per_hundread

Provided

- vaccineID, date, iso_code, sourceID -> Total_vaccine

Normalization:

- Looking at all the FD's the relations are in 3NF so no need to decompose the relations.

Final Schema:

Location(iso_code, location, last_observation_date, sourceID*)

Source(sourceID, source_name, source_website)

State(stateID, stateName)

Vaccine(vaccineID, vaccineName)

Vaccination_by_age_group(date, iso_code*, age_group,
people_vaccinated_per_hundread, people_fully_vaccinated_per_hundread,
people_with_booster_per_hundread)

Us_state_vaccinations(date, stateID *, iso_code*, total_vaccinations,
total_distributed, people_vaccinated, people_fully_vaccinated,
total_vaccinations_per_hundread, people_vaccinated_per_hundread,
distributed_per_hundread, daily_vaccinations_raw, daily_vaccinations,
daily_vaccinations_per_million, share_doses_used, total_boosters,
total_boosters_per_hundred)

Vaccination_by_manufacture(date, vaccine, iso_code*, vaccineID*,
total_vaccinations)

Vaccinations(date, iso_code*, total_vaccinations, people_vaccinated,
people_fully_vaccinated, total_boosters, daily_vaccinations_raw,
daily_vaccinations, total_vaccinations_per_hundread,
people_vaccinated_per_hundread, people_fully_vaccinated_per_hundread,
total_boosters_per_hundread, daily_vaccinations_per_million,
daily_people_vaccinated, daily_people_vaccinated_per_hundread)

Provided(vaccineID*, date*, iso_code*, Total_vaccine)

Location_vaccine(iso_code*,vaccineID*)