DATABASE CONCEPTS

FINAL ASSIGNMENT

Name:Pratham Radhakrishna

ID:3997064

I) ER - Diagram:

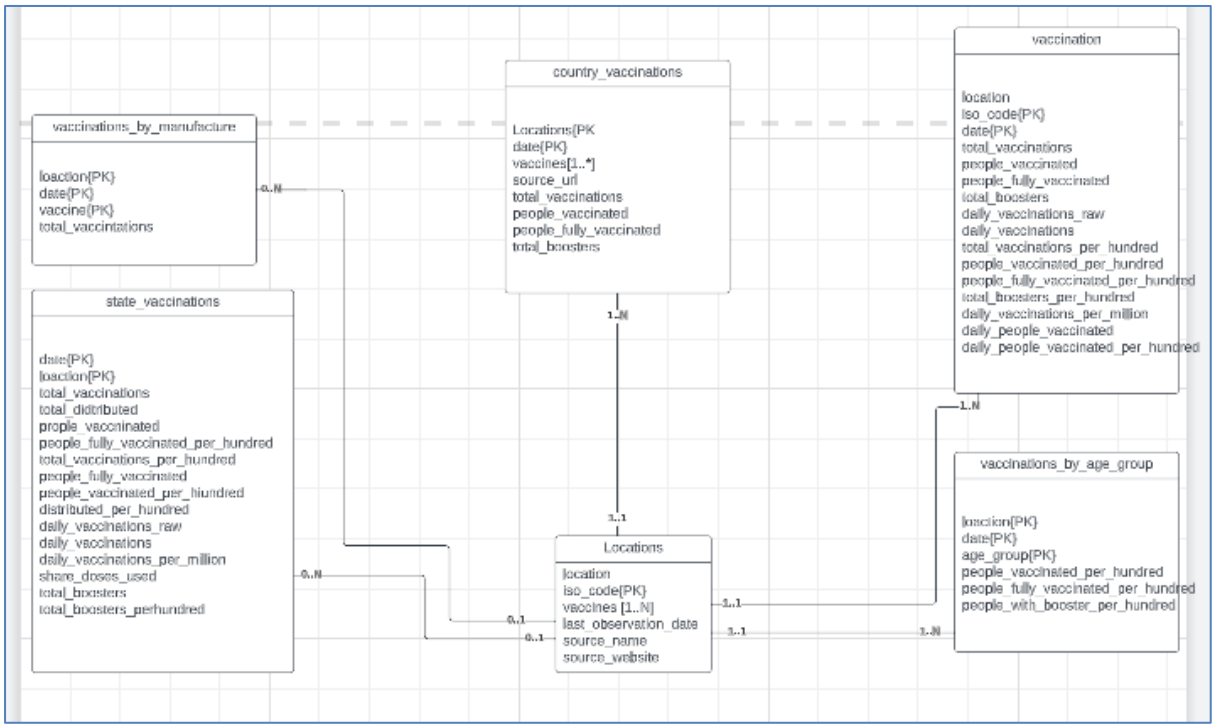
● Australia, England, New Zealand, and the United States csvs are combined into a single table called country_vaccinations .

● Us_state_vaccinations csv is ----→ state_vaccinations table where location specifies the country and state specifies which state in the country.

Assumption:

a) Each country does not have a manufacturer.

b) Every country will have vaccination.



SCHEMA

1.Strong entities:

Locations(location, <u>iso_code</u>, vaccines, last_observation_date, source_name, source_website)

Country_vaccinations(location, date, vaccines, source_url, total_vaccinations, people_vaccinated, people_fully_vaccinated, total_boosters)

Vaccinations_by_age_group(location, date, age_group, people_vaccinated_per_hundred, people_fully_vaccinated_per_hundred, people_with_booster_per_hundred)

State_vaccinations(date, location, state, total_vaccinations,total_distributed,people_vaccinated, people_fully_vaccinated_per_hundred,total_vaccinations_per_hundred, people_fully_vaccinated, 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_manufacturer(location, date, vaccine, total_vaccinations)

Vaccination(location, date, 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)

2) Map weak entities: Null

3) Map 1:1 relations: Null

4) Map 1:N relations:

Locations(location, iso_code, vaccines, last_observation_date, source_name, source_website)

Country_vaccinations(location, date, vaccines, source_url, total_vaccinations, people_vaccinated, people_fully_vaccinated, total_boosters, iso_code*)

Vaccinations_by_age_group(location, date, age_group, people_vaccinated_per_hundred, people_fully_vaccinated_per_hundred, people_with_booster_per_hundred, iso_code*)

State_vaccinations(date, location, state, total_vaccinations,total_distributed,people_vaccinated, people_fully_vaccinated_per_hundred,total_vaccinations_per_hundred, people_fully_vaccinated, 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, iso_code*)

Vaccination_by_manufacturer(location, date, vaccine, total_vaccinations, iso_code*)

Vaccination(location, date, 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, iso_code*)

5) Map M:N relation: Null

6) Map Multi- valued attributes:

Country_vaccinations(location, date, source_url, total_vaccinations, people_vaccinated, people_fully_vaccinated, total_boosters, iso_code*)

Country_vaccines_list(vaccines, (location, date)*)

Locations(location, iso_code, last_observation_date, source_name, source_website)

locations_vaccines_list(vaccines, iso_code*)

7) Map higher-degree relationships: Null

Final Scheme:

Country_vaccinations(location, date, source_url, total_vaccinations, people_vaccinated, people_fully_vaccinated, total_boosters, iso_code*)

Country_vaccines_list(vaccines, (location, date)*)

Locations(location, iso_code, last_observation_date, source_name, source_website)

locations_vaccines_list(vaccines, iso_code*)

Vaccinations_by_age_group(location, date, age_group, people_vaccinated_per_hundred, people_fully_vaccinated_per_hundred, people_with_booster_per_hundred, iso_code*)

State_vaccinations(date, location, state, total_vaccinations,total_distributed,people_vaccinated, people_fully_vaccinated_per_hundred,total_vaccinations_per_hundred, people_fully_vaccinated, 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, iso_code*)

Vaccination_by_manufacturer(location, date, vaccine, total_vaccinations, iso_code*)

Vaccination(location, date, 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, iso_code*)

Schema Normalization:

a) **Locations Locations**(location, iso_code, vaccines, last_observation_date, source_name, source_website)
   Dependencies:
   Location ←→ iso_code
   iso_code → vaccines, last_observation_date, source_name, source_website

   3NF:

Locations(location, iso_code*)
Location_details(iso_code, vaccines, last_observation_date, source_name, source_website)
Locations_vaccines_list(vaccines, iso_code*)

b) Country_vaccinations
Dependencies:
(location, date) → source_url, total_vaccinations, people_vaccinated, people_fully_vaccinated, total_boosters, iso_code
Location←→ iso_code
3NF:
Country_vaccinations((location, date)*, iso_code*)
Country_vaccination_Details(location, date, source_url, total_vaccinations, people_vaccinated, people_fully_vaccinated, total_boosters)
Country_vaccines_list(vaccines, (location, date)*)

c) Vaccinations_by_age_group
Vaccinations_by_age_group(location, date, age_group, people_vaccinated_per_hundred, people_fully_vaccinated_per_hundred, people_with_booster_per_hundred, iso_code*)
Dependencies:
(location, date) → age_group, people_vaccinated_per_hundred, people_fully_vaccinated_per_hundred, people_with_booster_per_hundred, iso_code
Location←→ iso_code
3NF:
Vaccinations_by_age_group(location, date, iso_code*)
Vaccinations_by_age_group_details(location, date, age_group, people_vaccinated_per_hundred, people_fully_vaccinated_per_hundred, people_with_booster_per_hundred)

d) State_vaccinations
State_vaccinations(date, location, state, total_vaccinations,total_distributed,people_vaccinated, people_fully_vaccinated_per_hundred,total_vaccinations_per_hundred, people_fully_vaccinated, 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, iso_code*)

Dependencies:
date, location, state → total_vaccinations,total_distributed,people_vaccinated, people_fully_vaccinated_per_hundred,total_vaccinations_per_hundred, people_fully_vaccinated, 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

State_vaccinations_details(date, location, state, total_vaccinations,total_distributed,people_vaccinated, people_fully_vaccinated_per_hundred,total_vaccinations_per_hundred,

people_fully_vaccinated, 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)

state_vaccinations(date, location, state, iso_code*)

e) **Vaccination_by_manufacturer**(location, date, vaccine, total_vaccinations, iso_code*)
Dependencies:
location, date, vaccine → total_vaccinations
Location → iso_code
3NF:
Vaccination_by_manufacturer_details(location, date, vaccine, total_vaccinations)
Vaccination_by_manufacturer(location, date, iso_code*)

f) Vaccination
Vaccination(location, date, 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, iso_code*)

Dependencies:
Location, date → total_vaccinations, people_vaccinated, people_fully_vaccinated, total_boosters,daily_vaccinations_raw, daily_vaccinations, daily_people_vaccinated
Location ←→ iso_code
3NF:
Vaccination(location, date, iso_code*)
Vaccination_details(location, date, 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)

Comment:
The relations between Column and Column_per_hundred/ per_million are unreliable as the population is not constant and will lead to foreign key constraint failure.

FInal Schema:
Vaccination(location, date, iso_code*)

Vaccination_details(location, date, 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)

Vaccination_by_manufacturer_details(location, date, vaccine, total_vaccinations)

Vaccination_by_manufacturer(location, date, iso_code*)

State_vaccinations_details(date, location, state,
total_vaccinations,total_distributed,people_vaccinated,
people_fully_vaccinated_per_hundred,total_vaccinations_per_hundred,
people_fully_vaccinated, 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)

state_vaccinations(date, location, state, iso_code*)

Vaccinations_by_age_group(location, date, iso_code*)

Vaccinations_by_age_group_details(location, date, age_group,
people_vaccinated_per_hundred, people_fully_vaccinated_per_hundred,
people_with_booster_per_hundred)

Country_vaccinations((location, date)*, iso_code*)

Country_vaccination_Details(location, date, source_url, total_vaccinations,
people_vaccinated, people_fully_vaccinated, total_boosters)

Country_vaccines_list(vaccines, (location, date)*)

Locations(location, iso_code*)

Location_details(iso_code, vaccines, last_observation_date, source_name, source_website)

Locations_vaccines_list(vaccines, iso_code*)